

ASYMPTOTIC DISTRIBUTION OF EIGENVALUES OF RANDOM
MATRICES AND CHARACTERIZATION OF THE GAUSSIAN
DISTRIBUTION BY ROTATIONAL INVARIANCE

by

William H. Olson

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

Dr. Boyd Harshbarger

Dr. V. R. R. Uppuluri

Dr. Raymond H. Myers

Dr. Donald A. Gardiner

Dr. Chris P. Tsokos

Dr. Richard G. Krutchkoff

May 1970

Blacksburg, Virginia

ACKNOWLEDGEMENTS

The author wishes to express his indebtedness and appreciation to the following:

Dr. V. R. R. Uppuluri for his guidance and constant encouragement throughout this project; and for the many hours of enlightening and delightful discussion on matters both related to the project and not.

Dr. Boyd Harshbarger for his interest, encouragement, and help throughout the period of the author's graduate studies.

Dr. Marvin A. Kastenbaum for his interest, encouragement, and assistance.

Dr. Donald A. Gardiner for his interest, encouragement, and assistance.

The National Institute of Environmental Health Sciences for partial support of this investigation through NIH training grant numbers 5 T01 ES00033-03, 5 T01 ES00033-04, 5 T01 ES00033-05.

The Oak Ridge Associated Universities for fellowship support during the period of dissertation research.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
1.1 Statistical Theory of Spectra	1
1.2 Outline of Contents	3
II. NOTATION, DEFINITIONS AND PRELIMINARIES	4
2.1 Random Matrices	4
2.2 Continuity of Ordered Eigenvalues	4
2.3 Modes of Convergence	6
2.4 Empirical Distribution Function	7
2.5 Some Lemmas	7
III. COMBINATORIAL ARGUMENTS	10
IV. RANDOM SIGN ENSEMBLE AND WIGNER'S CONJECTURE	21
4.1 Wigner's 1955 Paper	21
4.2 Wigner's 1958 Conjecture	27
V. THE RESULTS OF GRENANDER AND ARNOLD	29
5.1 Convergence in Probability	29
5.2 Convergence Almost Surely	34
VI. ON WIGNER'S CONJECTURE	35
6.1 Theorem	35
6.2 Comments	43
VII. ROTATIONAL INVARIANCE	45
7.1 Characterization of the Distribution of a Random Matrix	45

TABLE OF CONTENTS

Chapter	Page
VII. 7.2 Gaussian Orthogonal Ensemble	51
VIII. BIBLIOGRAPHY	53
IX. VITA	54

LIST OF TABLES

Table	Page
I. Enumerations	20

I. INTRODUCTION

The impetus for this paper comes mainly from work done in recent years by a number of physicists on a statistical theory of spectra. The book by M. L. Mehta (1967) and the collection of reprints edited by C. E. Porter (1965) are excellent references for this work. The discussion in section 1.1 is an attempt to present a rationale for such investigations. Our interpretation of linear operators as used in quantum mechanics is based largely on the book by T. F. Jordan (1969).

1.1 Statistical Theory of Spectra

In quantum mechanics knowledge of the value of measurable quantities of a system is expressed in terms of probabilities. A state of the system specifies these probabilities. Measurable quantities are represented by self-adjoint linear operators on a separable Hilbert space. The only possible values of the measurable quantities are those in the spectrum of the self-adjoint operator which represents the measurable quantity.

Experience indicates that energy is represented by the Hamiltonian operator. We are interested in the point spectrum of the Hamiltonian, which is its set of eigenvalues. The eigenvalues, E , of the Hamiltonian operator, H , which are real since H is self-adjoint, are those values of energy for which some state of the system specifies a probability of one that the energy is exactly equal to E . This is expressed in the Schrodinger time-independent equation,

$$H\psi = E\psi ,$$

where ψ is an eigenvector associated with E .

In ordinary statistical mechanics renunciation of exact knowledge of the state of a system is made and only properties of averages are considered. An exact knowledge of the system itself is assumed known; it is the impossibility in practice of observing the state of the system in all its detail that leads to the consideration of properties of averages.

An analogous situation exists with respect to the Hamiltonian operator. It is possible to choose an orthonormal basis for the separable Hilbert space in such a way that the matrix representation of the Hamiltonian with respect to this basis is in a form with blocks (finite dimensional square matrices) along the diagonal and zeros elsewhere. Each block corresponds uniquely to each set of values of a certain set of parameters. These parameters are variables which may be used to describe certain aspects of the system, whatever state it may be in. We are interested in the eigenvalues of the very large blocks. There are two difficulties. First, we do not know the Hamiltonian and, second, even if we did, it would be far too complicated to attempt to solve it. These difficulties lead to a renunciation of an exact knowledge of the system itself (i.e., of the Hamiltonian); one considers an ensemble of possible Hamiltonians, subsets of which are assigned weights according to some assumed relative frequency of occurrence. In other words the elements of the matrices are random variables. It is desirable, due to our ignorance of the system, that the statistical properties of the eigenvalues be independent of as many of the properties of the distributions of the elements of the matrices as possible. At best the elements of these matrices are random variables whose distributions are restricted only by the general symmetry properties we might impose on the ensemble of operators.

For our study we impose certain conditions (viz., time reversal symmetry and rotational symmetry, or time reversal symmetry with no rotational symmetry but integral spin) which lead to the consideration of real symmetric matrices. In the last chapter a further restriction is imposed on the ensemble of operators and the resulting distributions of the matrix elements is derived.

1.2 Outline of Contents

There are three basic parts. Chapter III contains the combinatorial arguments which are essential for the proofs of the theorems in the second part, Chapters IV, V, and VI. These chapters all deal with the asymptotic distribution of the empirical distribution function of the eigenvalues of a symmetric random matrix from the points of view of weakening the conditions placed on the distribution of the elements of the matrix and of strengthening the mode of convergence of the empirical distribution functions. The last part, Chapter VII, is on the Gaussian orthogonal ensemble, a model with which the physicists have worked extensively.

II. NOTATION, DEFINITIONS AND PRELIMINARIES

2.1 Random Matrices

The definition of random matrix is established in this section.

Let (Ω, \mathcal{F}, P) denote a probability space, i.e., Ω is a nonempty abstract set, \mathcal{F} is a σ -algebra of subsets of Ω , and P is a probability measure on \mathcal{F} ; and let (R_n, β_n) be the measurable space where R_n is n -dimensional Euclidean space and β_n is the Borel σ -algebra of subsets of R_n .

A mapping $X : \Omega \rightarrow R_n$ is called a random vector if $\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$ for all $B \in \beta_n$. When $n = 1$, X is called a random variable.

A mapping $A : \Omega \times R_n \rightarrow R_n$ is called a random operator if $A(\omega)[x]$ is for every $x \in R_n$ a random vector. A random operator A is said to be linear if $A(\omega) [\alpha x_1 + \beta x_2] = \alpha A(\omega) [x_1] + \beta A(\omega) [x_2]$ for every $\omega \in \Omega$, $x_1, x_2 \in R_n$ and $\alpha, \beta \in R_1$.

A linear random operator defined by the $n \times n$ matrix

$$A = (a_{ij})_{i,j=1}^n$$

where the a_{ij} are random variables is called a random matrix. Thus, a random matrix is a linear random operator on $\Omega \times R_n$ to R_n .

Throughout the paper all random quantities will be assumed to be defined on some fixed probability space (Ω, \mathcal{F}, P) .

2.2 Continuity of Ordered Eigenvalues

It is established in this section that ordered eigenvalues of symmetric random matrices are indeed random variables.

The following lemma is needed. Denote by $\lambda_i(A)$, $i = 1, \dots, n$, the

eigenvalues of any $(n \times n)$ matrix A .

Lemma 1: Let A be an $(n \times n)$ matrix and suppose $\epsilon > 0$ is given. Then there exists a $\delta > 0$ such that for any matrix $D = (d_{ij})_{i,j=1}^n$ such that $\sum_{i,j=1}^n |d_{ij}| < \delta$, there exists a permutation σ of $\{1, 2, \dots, n\}$ for which

$$|\lambda_i(A) - \lambda_{\sigma(i)}(A+D)| < \epsilon, \quad i = 1, \dots, n.$$

A proof of this lemma may be found in Ostrowski (1960).

Denote by $\lambda_1(A) \leq \lambda_2(A) \leq \dots \leq \lambda_n(A)$ the ordered eigenvalues of any $(n \times n)$ Hermitian matrix A .

Corollary: $\lambda_1, \lambda_2, \dots, \lambda_n$ are continuous functions of the elements of $A = (a_{ij})$.

Proof: The proof is by contradiction. Let $A = (a_{ij})$ be given. By the above lemma it is known that for a given $\epsilon > 0$ there exists a $\delta > 0$ such that for any $A' = (a'_{ij})$ such that $\sum_{i,j=1}^n |a_{ij} - a'_{ij}| < \delta$ one has for a suitable permutation, σ , of $\{1, 2, \dots, n\}$,

$$|\lambda_i(A) - \lambda_{\sigma(i)}(A')| < \epsilon$$

for $i = 1, \dots, n$. Assume $|\lambda_i(A) - \lambda_i(A')| > \epsilon$; to be definite assume $\lambda_i(A') > \lambda_i(A)$. Then $\lambda_1(A) \leq \dots \leq \lambda_i(A) < \lambda_i(A') \leq \dots \leq \lambda_n(A')$. With each $\lambda_j(A)$, $j = 1, \dots, i$, is associated $\lambda_{\sigma(j)}(A')$ such that $|\lambda_j(A) - \lambda_{\sigma(j)}(A')| < \epsilon$. But only $\lambda_1(A'), \dots, \lambda_{i-1}(A')$ are available for this purpose since $\lambda_i(A') - \lambda_i(A) > \epsilon$ and hence $\lambda_i(A') - \lambda_j(A) > \epsilon$, $j = 1, \dots, i$. Thus, one must conclude $|\lambda_i(A) - \lambda_i(A')| < \epsilon$, $i = 1, \dots, n$. This completes the proof.

Let $A = (a_{ij})_{i,j=1}^n$ be a random matrix such that $a_{ij} = a_{ji}$ a.s. (referred to as symmetric random matrix), and denote by $\lambda_1(A) \leq \lambda_2(A) \leq \dots \leq \lambda_n(A)$ its ordered eigenvalues. Then by the above corollary the eigenvalues $\lambda_i(A)$ are random variables since they are continuous functions of random variables.

2.3 Modes of Convergence

Three types of convergence of a sequence of random variables are considered in this paper: convergence in law, convergence in probability, and convergence almost surely. Let S be a random variable and let $\{X_n\}_{n=1}^{\infty}$ be an infinite sequence of random variables; let F_n and F denote the distribution functions of X_n and X respectively.

The sequence $\{X_n\}_{n=1}^{\infty}$ is said to converge in law to X as $n \rightarrow \infty$, written $X_n \xrightarrow{\mathcal{L}} X$ as $n \rightarrow \infty$, if $F_n(x) \rightarrow F(x)$ as $n \rightarrow \infty$ at all points of continuity of F .

The sequence $\{X_n\}_{n=1}^{\infty}$ is said to converge in probability to X as $n \rightarrow \infty$, written $X_n \xrightarrow{P} X$ as $n \rightarrow \infty$, if for any given $\epsilon > 0$,

$$P(|X_n - X| > \epsilon) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

The sequence $\{X_n\}_{n=1}^{\infty}$ is said to converge a.s. to X as $n \rightarrow \infty$, written $X_n \rightarrow X$ a.s. as $n \rightarrow \infty$, if

$$P(\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1$$

The following implication structure exists among these modes of convergence: $X_n \rightarrow X$ a.s. as $n \rightarrow \infty$ implies $X_n \xrightarrow{P} X$ as $n \rightarrow \infty$ implies $X_n \xrightarrow{L} X$ as $n \rightarrow \infty$.

2.4 Empirical Distribution Function

Let $\{X_1, X_2, \dots, X_n\}$ be a set of random variables. For any $B \in \mathcal{J}$, let I_B denote the indicator function of B , i.e.,

$$I_B(\omega) = \begin{cases} 1 & \text{if } \omega \in B \\ 0 & \text{if } \omega \notin B. \end{cases}$$

The empirical distribution function of $\{X_1, X_2, \dots, X_n\}$ is a mapping $F_n : R_1 \times \Omega \rightarrow [0, 1]$ defined by

$$F_n(x)(\omega) = \frac{1}{n} \sum_{i=1}^n I_{[X_i \in (-\infty, x)]}(\omega).$$

Let $A_n = (a_{ij})_{i,j=1}^n$ be a random matrix such that $a_{ij} = a_{ji}$ a.s. Let $\lambda_1(A_n) \leq \lambda_2(A_n) \leq \dots \leq \lambda_n(A_n)$ denote the a.s. real ordered random eigenvalues of A_n . Denote by W_n the empirical distribution function of $\{\lambda_1(A_n), \lambda_2(A_n), \dots, \lambda_n(A_n)\}$. A basic question examined in this paper is (for any $x \in R_1$): how does $W_n(x)$ behave as $n \rightarrow \infty$?

2.5 Some Lemmas

In this section are listed some lemmas which will be used below.

Given a random variable X and a sequence of random variables $\{X_n\}_{n=1}^{\infty}$, let F_n and F denote the distribution functions of X_n and X respectively. Furthermore, let

$$m^{(k)} = \int_R x^k dF(x)$$

define the k th moment of the distribution function F , if it exists.

If F possesses subscripts, we affix the same subscripts to its moments.

Lemma 2: If, for $k \geq k_0$ arbitrary but fixed, the sequence $m_n^{(k)} \rightarrow m^{(k)}$ finite, then these sequences converge for every value of k , and if the sequence $\{m^{(k)}\}_{k=1}^{\infty}$ uniquely determines F , then $F_n(x) \rightarrow F(x)$ as $n \rightarrow \infty$ at all points of continuity of F .

A proof of this lemma may be found in Loeve (1963).

For any infinite sequence of sets, $\{A_n\}_{n=1}^{\infty}$, $A_n \in \mathcal{J}$, define

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$$

Lemma 3: $X_n \rightarrow 0$ a.s. as $n \rightarrow \infty$ if and only if for all $\epsilon > 0$

$$P(\limsup_{n \rightarrow \infty} \{\omega : |X_n(\omega)| > \epsilon\}) = 0$$

A proof of this lemma may be found in Chung (1968).

Lemma 4: (Borel-Cantelli) If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\limsup_{n \rightarrow \infty} A_n) = 0$.

A proof of this may be found in Loeve (1963).

Let W_n be the empirical distribution function of $\{X_1, X_2, \dots, X_n\}$. Let W be a distribution function which is uniquely determined by its sequence of moments, $\{X_k\}_{k=1}^{\infty}$. Let us define

$$\begin{aligned} M_{k,n}(\omega) &= \int_{\mathbb{R}} x^k dW_n(x)(\omega) \\ &= \frac{1}{n} \sum_{i=1}^n X_i^k(\omega) \end{aligned}$$

Lemma 5: If $M_{k,n} \xrightarrow{P} \alpha_k$ as $n \rightarrow \infty$ for all $k = 1, 2, \dots$, then $W_n(x) \xrightarrow{P} W(x)$ as $n \rightarrow \infty$, at all points of continuity of W .

Proof: The following result is used to establish the lemma: $X_n \xrightarrow{P} X$ as $n \rightarrow \infty$ if and only if every subsequence $\{X_{n_i}\}$ contains a subsequence which converges a.s. to X . Let $\{n_i\}$ be any subsequence of the positive integers. Then

$$\int_{\mathbb{R}} x^k dW_{n_i}(x) \xrightarrow{P} \int_{\mathbb{R}} x^k dW(x)$$

for all $k = 1, 2, \dots$. By the diagonal procedure it is possible to select a subsequence $\{n'_i\}$ of $\{n_i\}$ such that

$$\int_{\mathbb{R}} x^k dW_{n'_i}(x) \rightarrow \int_{\mathbb{R}} x^k dW(x) \quad \text{a.s.}$$

for all $k = 1, 2, \dots$. Then by Lemma 2

$$W_{n'_i}(x) \rightarrow W(x) \quad \text{a.s.}$$

The above quoted result then gives $W_n(x) \xrightarrow{P} W(x)$ as $n \rightarrow \infty$ at all points of continuity of W . This completes the proof.

Lemma 6: (Cramer's Theorem) Let X and Y be independent random variables. If $X + Y$ is distributed normally, then X and Y are each distributed normally.

A proof of this may be found in Loeve (1963).

III. COMBINATORIAL ARGUMENTS

The following combinatorial lemmas are of central importance in the proofs of the limit theorems to follow. They are slight extensions of results given by Wigner (1955).

Denote by $A_{k,n}$, $k \geq 1$, $n \geq 1$, the class of all finite sequences $f : \{1,2,\dots,k+1\} \rightarrow \{1,2,\dots,n\}$. Any ordered pair of positive integers, (i,j) , will be called a step. The step (j,i) will be called the reverse step of (i,j) . With each $f \in A_{k,n}$ is associated a sequence $g_f : \{1,2,\dots,k\} \rightarrow \{\text{all steps}\}$ defined as follows: $g_f(v) = (f(v), f(v+1))$, $1 \leq v \leq k$. The sequence g_f will be called the sequence of steps associated with f . The cardinality of any set A will be denoted by $\#A$. Let

$$D_f = \{f(i), 2 \leq i \leq k+1 : f(i) \notin \{f(1), \dots, f(i-1)\}\},$$

and let $d_f = \#D_f + 1$. By definition, f has b different members if and only if $d_f = b$. Let $\#(i,j)_f$ denote

$$\#\{g_f(v) = (f(v), f(v+1)), 1 \leq v \leq k : f(v) = i, f(v+1) = j\}.$$

For $1 \leq v \leq k$, $g_f(v) = (f(v), f(v+1))$ is called a free step if and only if $f(v+1) \notin \{f(1), \dots, f(v)\}$ and a repetitive step if and only if $f(v+1) \in \{f(1), \dots, f(v)\}$. Let

$$F_f = \{g_f(v), 1 \leq v \leq k : g_f(v) \text{ is free}\},$$

$$R_f = \{g_f(v), 1 \leq v \leq k : g_f(v) \text{ is repetitive}\}.$$

It is immediate that

$$(1) \quad \#F_f + \#R_f = k$$

$$\begin{aligned} \text{and } \#F_f &= \#\{g_f(v), 1 \leq v \leq k : g_f(v) \text{ is free}\} \\ &= \#\{f(v+1), 1 \leq v \leq k : f(v+1) \notin \{f(1), \dots, f(v)\}\} \\ &= \#\{f(v), 2 \leq v \leq k+1 : f(v) \notin \{f(1), \dots, f(v-1)\}\} \end{aligned}$$

$$(2) \quad \#F_f = \#D_f = d_f - 1 .$$

Lemma 7: Let $f \in A_{k,n}$ be such that if $(i,j) \in \{g_f(1), g_f(2), \dots, g_f(k)\}$, then $\#(i,j)_f + \#(j,i)_f \geq 2$. Then $d_f \leq \lfloor \frac{k}{2} \rfloor + 1$.

Proof: Let $f \in A_{k,n}$ satisfy the conditions of the lemma. If $f(i) \in \{f(v), 2 \leq v \leq k+1 : f(v) \notin \{f(1), \dots, f(v-1)\}\}$ then $(f(i-1), f(i))$ is a free step. The condition of the lemma implies at least one step among $g_f(i), \dots, g_f(k)$ must equal $(f(i-1), f(i))$ or $(f(i), f(i-1))$ (no step among $g_f(1), \dots, g_f(i-2)$ equals $(f(i-1), f(i))$ or $(f(i), f(i-1))$ since $f(i) \notin \{f(i), \dots, f(i-1)\}$). Any such occurrence, say $(f(l-1), f(l))$, must be repetitive since $f(l) \in \{f(1), \dots, f(l-1)\}$. Hence with each free step is associated a repetitive step which is equal to the free step or its reverse. This implies $\#F_f \leq \#R_f$, since all free steps are different. This, with (1), implies $\#F_f \leq \lfloor \frac{k}{2} \rfloor$. Hence, by (2), $d_f - 1 = \#F_f \leq \lfloor \frac{k}{2} \rfloor$, or $d_f \leq \lfloor \frac{k}{2} \rfloor + 1$. This completes the proof of Lemma 7.

Lemma 8: Let $f \in A_{k,n}$ be such that:

- (i) if $(i,j) \in \{g_f(1), g_f(2), \dots, g_f(k)\}$, then $\#(i,j)_f + \#(j,i)_f \geq 2$;
- (ii) $f(l) = f(l+1)$ for some $l, 1 \leq l \leq k$.

Then $d_f \leq \lfloor \frac{k}{2} \rfloor$.

Proof: If f is constant one is through. Assume f is not constant. If $f(\ell) = f(\ell+1)$ for some ℓ , $1 \leq \ell \leq k$, a new sequence of steps may be formed from $g_f(1), g_f(2), \dots, g_f(k)$ by omitting all those steps equal to $(f(\ell), f(\ell+1))$ (there will be two or more such steps, by condition (i)). The sequence of steps thus formed is associated with a sequence $h : \{1, 2, \dots, i\} \rightarrow \{1, 2, \dots, n\}$, $2 \leq i \leq k-1$, (a lower bound of 2 since f is not constant) which satisfies condition (i) and which is such that $d_h = d_f$. Lemma 1 then gives $d_f = d_h \leq \lfloor \frac{i-1}{2} \rfloor + 1 \leq \lfloor \frac{k-2}{2} \rfloor + 1 = \lfloor \frac{k}{2} \rfloor$. This completes the proof of Lemma 8.

Lemma 9: Let k be even, say $k = 2v$. Let $f \in A_{2v, n}$ be such that:

- (i) if $(i, j) \in \{g_f(1), g_f(2), \dots, g_f(2v)\}$, then $\#(i, j)_f + \#(j, i)_f \geq 2$;
- (ii) $f(1) = f(2v+1)$;
- (iii) $d_f = v + 1$.

If $(i, j) \in \{g_f(1), g_f(2), \dots, g_f(2v)\}$, then $\#(i, j)_f = 1$, $\#(j, i)_f = 1$.

Proof: Let $f \in A_{2v, n}$ satisfy conditions (i), (ii), and (iii) of the lemma.

(For $n \geq v + 1$ such an f is easily constructed. For example, let $f(1) = 1$, $f(2) = 2$, $f(v) = v$, $f(v+1) = v + 1$, $f(v+2) = v, \dots, f(2v) = 2$, $f(2v+1) = 1$.)

Lemma 7 shows $d_f \leq v + 1$. By Lemma 8 one must have $f(\ell) \neq f(\ell+1)$,

$1 \leq \ell \leq 2v$. Equation (2) holds:

$$\#F_f = \#D_f = d_f - 1 = v.$$

Consider the first step $g_f(1) = (f(1), f(2))$. If $f(1) \notin \{f(3), \dots, f(2v-1)\}$, then by condition (i) the last step must be the reverse of the first since $f(1) \notin \{f(2), \dots, f(2v)\}$. On the other hand, if $f(1) \in \{f(3),$

..., $f(2\nu-1)$ the following argument applies. Let l , $3 \leq l \leq 2\nu-1$ be the least integer such that $f(l) = f(1)$. Assume $f(l-1) \neq f(2)$. Condition (i) implies the repetitive step $g_f(l-1) = (f(l-1), f(1))$ must be matched by at least one further occurrence among $g_f(l), \dots, g_f(2\nu)$ of a step equal to $(f(l-1), f(1))$ or $(f(1), f(l-1))$, these occurrences being repetitive steps, since no free step equals $(f(l-1), f(1))$ or $(f(1), f(l-1))$; which is so because: (1) the first step does not since $f(l-1) \neq f(2)$; (2) no step among $g_f(2), \dots, g_f(l-2)$ involves an $f(1)$; and (3) any further free step among $g_f(l), \dots, g_f(2\nu)$, say $(f(i-1), f(i))$, could not have $f(i) = f(1)$ or $f(i) = f(l-1)$ because in either case $f(i) \in \{f(1), \dots, f(i-1)\}$. For each free step there is an occurrence in the sequence of steps of a repetitive step equal to the free step itself or its reverse, by condition (i). Since there are ν different free steps one must have at least 2ν steps in the sequence equaling these or their reverses. This is apart from the 2 or more repetitive steps equaling $(f(l-1), f(1))$ or $(f(1), f(l-1))$, since no free step equals either. Altogether one would need at least $2\nu + 2$ steps; but only 2ν are available. Hence one must have $f(l-1) = f(2)$. Thus the reverse of the first step occurs.

Now define a sequence $h : \{1, 2, \dots, 2\nu + 1\} \rightarrow \{1, 2, \dots, n\}$ as follows:

$$\begin{aligned} h(1) &= f(2), h(2) = f(3), \dots, h(i) = f(i+1), \dots, \\ h(2\nu) &= f(2\nu+1) = f(1), h(2\nu+1) = f(2) . \end{aligned}$$

Associated with h is the sequence of steps $g_h(1) = (f(2), f(3))$, $g_h(2) = (f(3), f(4)), \dots, g_h(2\nu-1) = (f(2\nu), f(1))$, $g_h(2\nu) = (f(1), f(2))$. It is immediate that h satisfies conditions (i), (ii), and (iii) of the lemma. The above argument shows that the reverse of $g_h(1) = (f(2), f(3))$

occurs among $g_h(2) = (f(3), f(4)), \dots, g_h(2v) = (f(1), f(2))$. Continuing in the same manner one concludes if $(i, j) \in \{g_f(1), g_f(2), \dots, g_f(2v)\}$, then $(j, i) \in \{g_f(1), g_f(2), \dots, g_f(2v)\}$. Since there are v different free steps and $2v$ steps altogether, one must have $\#(i, j)_f = 1$, $\#(j, i)_f = 1$ for each $(i, j) \in \{g_f(1), g_f(2), \dots, g_f(2v)\}$. This completes the proof of Lemma 9.

Lemma 10: Let k be odd, say $k = 2v + 1$. Let $f \in A_{2v+1, n}$ be such that:

- (i) if $(i, j) \in \{g_f(1), g_f(2), \dots, g_f(2v+1)\}$, then $\#(i, j)_f + \#(j, i)_f \geq 2$;
- (ii) $f(1) = f(2v+2)$.

Then $d_f \leq v = \lfloor \frac{k}{2} \rfloor$.

Proof: Let $f \in A_{2v+1, n}$ satisfy conditions (i) and (ii) of the lemma.

Lemma 7 shows $d_f \leq v + 1$. Assume $d_f = v + 1$. Then by Lemma 8, $f(l) \neq f(l+1)$, $1 \leq l \leq 2v + 1$. There are $\#F_f = \#D_f = d_f - 1 = v$ different free steps. For each free step there is a repetitive step equal to the free step itself or its reverse. This occupies $2v$ of the $2v + 1$ steps associated with f . By condition (i) the remaining step must equal one of the free steps or its reverse. In other words, $\#(i, j)_f + \#(j, i)_f = 2$ for all $(i, j) \in \{g_f(1), \dots, g_f(2v+1)\}$ except one, say (k, l) , for which $\#(k, l)_f + \#(l, k)_f = 3$.

All possibilities are now considered. First consider the case $\#(k, l)_f = 3$. (The case $\#(l, k)_f = 3$ is the same.) With f is associated a sequence of steps $g_f(1), \dots, g_f(r), \dots, g_f(s), \dots, g_f(t), \dots, g_f(2v+1)$, where $g_f(r) = g_f(s) = g_f(t) = (k, l)$ and $s - r \geq 2$, $t - s \geq 2$. Let $g_f^*(i)$ denote the reverse of $g_f(i)$. From the sequence of steps $g_f(1), g_f(2),$

..., $g_f(2\nu+1)$ forma sequence of steps associated with a sequence

$h : \{1, 2, \dots, 2\nu-1\} \rightarrow \{1, 2, \dots, n\}$ in the following manner:

$$g_h(1) = g_f(t+1) = (\ell, f(t+2)) = (h(1), h(2))$$

$$g_h(2) = g_f(t+2) = (f(t+2), f(t+3)) = (h(2), h(3))$$

⋮

$$g_h((2\nu+1)-t) = g_f(2\nu+1) = (f(2\nu+1), f(2\nu+2))$$

$$= (h((2\nu+1)-t), h((2\nu+1)-(t-1)))$$

$$g_h((2\nu+1)-(t-1)) = g_f(1) = (f(1), f(2))$$

$$= (h((2\nu+1)-(t-1)), h((2\nu+1)-(t-2)))$$

⋮

$$g_h((2\nu+1)-(t-r+1)) = g_f(r-1) = (f(r-1), k)$$

$$= (h((2\nu+1)-(t-r+1)), h((2\nu+1)-(t-r)))$$

$$g_h((2\nu+1)-(t-r)) = g_f^*(s-1) = (k, f(s-1))$$

$$= (h((2\nu+1)-(t-r)), h((2\nu+1)-(t-r-1)))$$

$$g_h((2\nu+1)-(t-r-1)) = g_f^*(s-2) = (f(s-1), f(s-2))$$

$$= (h((2\nu+1)-(t-r-1)), h((2\nu+1)-(t-r-2)))$$

⋮

$$g_h((2\nu+1)-(t-s+2)) = g_f^*(r+1) = (f(r+2), \ell)$$

$$= (h((2\nu+1)-(t-s+2)), h((2\nu+1)-(t-s+1)))$$

$$g_h((2\nu+1)-(t-s+1)) = g_f(s+1) = (\ell, f(s+2))$$

$$= (h((2\nu+1)-(t-s+1)), h((2\nu+1)-(t-s)))$$

⋮

$$g_h((2\nu+1)-4) = g_f(t-2) = (f(t-2), f(t-1)) = (h((2\nu+1)-4), h((2\nu+1)-3))$$

$$g_h((2\nu+1)-3) = g_f(t-1) = (f(t-1), k) = (h((2\nu+1)-3), h(2\nu+1)-2))$$

The steps associated with h are $g_f(1), \dots, g_f(r-1), g_f^*(r+1), \dots, g_f^*(s-1), g_f(s+1), \dots, g_f(t-1), g_f(t+1), \dots, g_f(2v+1)$, in other words, the same as those associated with f except all steps equaling (k, l) have been dropped and some of the steps associated with f have been reversed. It is easily seen that if $(i, j) \in \{g_h(1), \dots, g_h(2v-2)\}$, then $\#(i, j)_h + \#(j, i)_h = 2$, and $d_h = d_f = v + 1$. But Lemma 1 shows $d_h \leq v$. One must conclude, by contradiction, that $d_f \leq v$.

The other possibility is $\#(k, l)_f = 2, \#(l, k)_f = 1$ (or, what is the same thing, $\#(k, l)_f = 1, \#(l, k)_f = 2$) for which an argument similar to the above may be given. The details are not given here. This completes the proof of Lemma 10.

Wigner's Combinatorial Theorem

Let $B_{2v, n}$ be the set of all $f \in A_{2v, n}$ such that:

- (i) if $(i, j) \in \{g_f(1), g_f(2), \dots, g_f(2v)\}$, then $\#(i, j)_f + \#(j, i)_f \geq 2$;
- (ii) $f(1) = f(2v+1)$;
- (iii) $d_f = v + 1$.

Then

$$\#B_{2v, n} = \frac{(2v)!}{v!(v+1)!} n^{v+1} + o(n^{v+1}).$$

Proof: Let $f \in B_{2v, n}$. By Lemma 8, $f(l) \neq f(l+1)$, $1 \leq l \leq 2v$. By Lemma 9, if $(i, j) \in \{g_f(1), \dots, g_f(2v)\}$, then $\#(i, j)_f = 1, \#(j, i)_f = 1$. A sequence $t : \{1, 2, \dots, m\} \rightarrow \{\text{integers}\}$ is called a type sequence if and only if $t(l) \geq 0$, $1 \leq l \leq m$, $t(1) = 1$, $t(m) = 0$, and $t(l+1) - t(l) = \pm 1$, $1 \leq l \leq m-1$. For each $f \in B_{2v, n}$ define the type sequence $t_f : \{1, 2, \dots, 2v\} \rightarrow \{\text{integers}\}$ as follows:

$$t_f(\ell) = \#\{g_f(i), 1 < i < \ell : g_f(i) \text{ is free}\} - \#\{g_f(i), 1 < i < \ell : g_f(i) \text{ is repetitive}\}.$$

For a given type sequence $t : \{1, 2, \dots, 2v\} \rightarrow \{\text{integers}\}$ one has

$$\#\{f \in B_{2v, n} : t_f(\ell) = t(\ell), 1 \leq \ell \leq 2v\} = n(n-1)\dots(n-v).$$

This is so because: (1) there are n choices for $f(1)$; (2) for $2 \leq i \leq 2v$, if $t(i) - t(i-1) = 1$, then $(f(i-1), f(i))$ is a free step and $f(i)$ may be any number which has not been used yet; and (3) for $2 \leq i \leq 2v$, if $t(i) - t(i-1) = -1$, then $(f(i-1), f(i))$ is a repetitive step and must be the reverse of the step which originally led to $f(i-1)$ (Lemma 9) and hence $f(i)$ is completely determined. Let S_v denote the number of type sequences with domain $\{1, 2, \dots, 2v\}$. Then

$$\#B_{2v, n} = S_v n(n-1)\dots(n-v) = S_v n^{v+1} + o(n^{v+1}).$$

To find S_v one argues as follows. The number of type sequences t such that $t(i) > 0$, $1 \leq i \leq 2v-1$, $t(2v) = 0$, i.e., no 0 before the last value, will be denoted by S'_v . From such sequences one can obtain a type sequence with domain $\{1, 2, \dots, 2v-2\}$ by omitting $t(1), t(2v)$ and subtracting 1 from each $t(i)$, $2 \leq i \leq 2v-1$. Hence

$$S'_v = S_{v-1} \quad (S'_1 = S_0 = 1).$$

Given a type sequence with domain $\{1, 2, \dots, 2v\}$ let $2k$ be the smallest integer such that $t(2k) = 0$ for the first time. Then $t_1 : \{1, 2, \dots, 2k\} \rightarrow \{\text{integers}\}$ forms a 0 free type sequence while $t_2 : \{2k+1, \dots, 2v\} \rightarrow \{\text{integers}\}$ forms an arbitrary type sequence. Hence

$$(3) \quad S_v = \sum_{k=1}^v S_k' S_{v-k} = \sum_{k=1}^v S_{k-1} S_{v-k}, \quad v = 1, 2, \dots$$

These recursive equations permit the successive calculation of the S_v .

Formally one can obtain a closed formula for them by writing

$$t(x) = \sum_{v=0}^{\infty} t_v x^v.$$

The recursive formula (3) then gives

$$t(x) = 1 + xt^2(x).$$

The 1 on the right hand side is necessary because (3) is not valid for $v = 0$. It follows that

$$t(x) = (1 \pm (1-4x)^{\frac{1}{2}})/2x.$$

Actually, the lower sign has to be taken. It gives

$$S_v = \frac{1}{2} \binom{\frac{1}{2}}{v+1} (-4)^{v+1} = \frac{(2v)!}{v!(v+1)!}, \quad v = 1, 2, \dots$$

And finally,

$$\#B_{2v,n} = \frac{(2v)!}{v!(v+1)!} n^{v+1} + o(n^{v+1}).$$

This completes the proof of the theorem.

Let $C_{2k,n}$ denote the set of all $f \in A_{2k,n}$ such that:

- (i) $f(1) = f(2k+1)$;
- (ii) $f(i) \neq f(i+1)$, $1 \leq i \leq 2k$;

(iii) if $(i, j) \in \{g_f(1), g_f(2), \dots, g_f(2^v)\}$,

then $\#(i, j)_f + \#(j, i)_f$ is even.

Let $C_{2k, n}^j$ denote the set of all $f \in C_{2k, n}$ such that $d_f = j$. By Lemma 7, if $f \in C_{2k, n}$, then $d_f \leq k + 1$. Thus $C_{2k, n} = \bigcup_{j=1}^{k+1} C_{2k, n}^j$ and

$$\#C_{2k, n} = \sum_{j=1}^{k+1} \#C_{2k, n}^j.$$

Lemma 11: $\#C_{2k, n}^j = \binom{n}{j} (\#C_{2k, j}^j)$

Proof: The relation \sim determined by $f \sim f^*$ iff $f \in C_{2k, n}^j$, $f^* \in C_{2k, n}^j$ and $\{f(1), f(2), \dots, f(2k+1)\} = \{f^*(1), f^*(2), \dots, f^*(2k+1)\}$ is an equivalence relation. $C_{2k, n}^j$ is split into $\binom{n}{j}$ equivalence classes by \sim , each containing $\#C_{2k, j}^j$ members. Hence $\#C_{2k, n}^j = \binom{n}{j} (\#C_{2k, j}^j)$. This completes the proof of the lemma.

Using Lemma 11, one has

$$\#C_{2k, n} = \sum_{j=1}^{k+1} \binom{n}{j} (\#C_{2k, j}^j).$$

Thus $\#C_{2k, n}$ is determined for all n by $\#C_{2k, 1}^1, \#C_{2k, 2}^2, \dots, \#C_{2k, k+1}^{k+1}$. An unsuccessful attempt to determine these numbers in a closed form was made. In an attempt to solve the problem the enumerations found in Table 1 were made on a computer. It will be pointed out in the next chapter in what context these numbers may be of interest.

$k, n \backslash j$	2	3	4	5	6	7
3,2	1					
3,3	2	20				
3,4	3	60	30			
3,5	4	120	120			
3,6	5	200	300			
4,2	1					
4,3	2	84				
4,4	3	252	390			
4,5	4	504	1,560	336		
4,6	5	840	3,900	1,680		
5,2	1					
5,3	2	340				
5,4	3	1,020	3,840			
5,5	4	2,040	15,360	8,544		
5,6	5	3,400	38,400	42,720	5,040	
6,2	1					
6,3	2	1,364				
6,4	3	4,092	34,980			
6,5	4	8,184	139,920	153,600		
6,6	5	13,640	349,800	768,000	214,080	
6,7	6	20,460	699,600	2,304,000	1,284,480	95,040

Table I. Enumerations. The numbers in the body of the table are

$$\frac{1}{n}(\#c_{2k,n}^j) = \frac{1}{n} \binom{n}{j} (\#c_{2k,j}^j) .$$

IV. RANDOM SIGN ENSEMBLE AND WIGNER'S CONJECTURE

4.1 Wigner's 1955 Paper

In 1955, Wigner proved the result discussed below.

Let $A_n = (a_{ij})_{i,j=1}^n$ be a random matrix such that:

- (i) $a_{ij} = a_{ji}$ a.s.;
- (ii) $\{a_{ij}, i \leq j\}$ is independent;
- (iii) $P(a_{ij} = \sigma) = \frac{1}{2}, i \neq j,$
 $P(a_{ij} = -\sigma) = \frac{1}{2}, i \neq j,$
 $P(a_{ii} = 0) = 1 .$

Let B_n denote the normalized matrix

$$B_n = \frac{1}{2\sigma\sqrt{n}} A_n .$$

Denote by $\lambda_1(B_n) \leq \lambda_2(B_n) \leq \dots \leq \lambda_n(B_n)$ the ordered random eigenvalues of B_n and by $W_n(x)$ the empirical distribution function of $\{\lambda_1(B_n), \lambda_2(B_n), \dots, \lambda_n(B_n)\}$, i.e.,

$$W_n(x)(\omega) = \frac{1}{n} \sum_{i=1}^n I_{[\lambda_i(B_n) \in (-\infty, x)]}(\omega) .$$

Then one has the following theorem.

Theorem 1 (Wigner, 1955): $\lim_{n \rightarrow \infty} E(W_n(x)) = W(x)$ for all $x \in R_1$, where W is the absolutely continuous distribution function with semi-circle density

$$w(x) = \begin{cases} \frac{2}{\pi} (1-x^2)^{\frac{1}{2}}, & |x| \leq 1 \\ 0, & |x| > 1 . \end{cases}$$

Proof: The distribution function W is uniquely determined by its moment sequence since

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{\gamma_k}{k!} (it)^k &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(k+1)!} \left(\frac{t}{2}\right)^{2k} = \frac{2}{t} J_1(t) \\ &= \frac{2}{\pi} \int_{-1}^1 e^{itx} (1-x^2)^{\frac{1}{2}} dx \end{aligned}$$

where J_1 denotes the Bessel function of order 1 of the first kind and

$$\gamma_k = \int x^k dW(x) = \begin{cases} 0 & , \text{ for } k \text{ odd} \\ \frac{k!}{2^k \left(\frac{k}{2}\right)! \left(\frac{k}{2}+1\right)!} & , \text{ for } k \text{ even} \end{cases}$$

It is immediate that $EW_n(x)$ is a distribution function in k . Thus, if it can be established that

$$\int x^k dEW_n(x) \rightarrow \gamma_k \text{ as } n \rightarrow \infty$$

for all $k = 1, 2, \dots$, then Lemma 2 will yield the desired result.

Consider the set, T , of all ordered $(n-1)n/2$ -tuples of the numbers $+\sigma$ and $-\sigma$. For each $(i_{12}, \dots, i_{1n}, i_{23}, \dots, i_{2n}, \dots, i_{n-1,n}) \in T$ define $D(i_{12}, \dots, i_{n-1,n}) = \{\omega \in \Omega : a_{11}(\omega) = 0, a_{12}(\omega) = i_{12}, \dots, a_{1n}(\omega) = i_{1n}, a_{22}(\omega) = 0, a_{23}(\omega) = i_{23}, \dots, a_{2n}(\omega) = i_{2n}, \dots, a_{n-1,n}, a_{nn}(\omega) = 0\}$. Then using assumptions (i), (ii), and (iii), we have

$$P(D(i_{12}, \dots, i_{n-1,n})) = \frac{1}{2^{(n-1)n/2}},$$

for all points in T . One has $W_n(x) = \frac{N_n(x)}{n}$, where $N_n(x)$ equals the number of eigenvalues of B_n less than x . $W_n(x)$ is constant on each D_i , $i \in T$;

denote these values by $W_n(x)_i = \frac{N_n(x)_i}{n}$, $i \in T$. Then, since

$\Omega = (\bigcup_{i \in T} D_i) \cup N$, where $(\bigcup_{i \in T} D_i) \cap N = \emptyset$ and $P(N) = 0$,

$$\begin{aligned} EW_n(x) &= \int_{\bigcup_{i \in T} D_i} W_n(x) dP \\ &= \sum_{i \in T} \int_{D_i} W_n(x) dP \\ &= \sum_{i \in T} \frac{N_n(x)_i}{n^2(n-1)n/2}. \end{aligned}$$

This shows that $EW_n(x)$ is a discrete distribution function with jumps of length $\frac{1}{n^2(n-1)n/2}$ or multiples thereof at the eigenvalues of the $2^{(n-1)n/2}$ possible (i.e., occurrence with positive probability) values of B_n . Each $i \in T$ represents one of the possible $2^{(n-1)n/2}$ values of B_n on Ω ; denote these values by $B_n(i)$, $i \in T$. Let $e_1, e_2, \dots, e_{n^2(n-1)n/2}$ be the set of all eigenvalues of all possible values of B_n . Then we have

$$\begin{aligned} \int x^k dEW_n(x) &= \frac{1}{n^2(n-1)n/2} \sum_{j=1}^{n^2(n-1)n/2} e_j^k \\ &= \frac{1}{n^2(n-1)n/2} \sum_{i \in T} \text{trace} (B_n^k(i)) \\ &= \frac{1}{2^{(n-1)n/2} (2\sigma)_n^{k/2+1}} \sum_{i \in T} \sum_{j_1=1}^n \sum_{j_k=1}^n \prod_{\ell=1}^k a_{j_\ell j_{\ell+1}}(i) \end{aligned}$$

where $j_{k+1} = j_1$ and $a_{jk}(i)$ equals the value a_{jk} assumes on D_i . Interchanging the order of summation and denoting by $A_{k,n}$ the class of all sequences $f : \{1, 2, \dots, k+1\} \rightarrow \{1, 2, \dots, n\}$, one has

$$\int x^k dEW_n(x) = \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in B_{k,n}} \frac{1}{2^{(n-1)n/2}} \sum_{i \in T} \prod_{l=1}^k a_{f(l)f(l+1)}(i)$$

where $B_{k,n} = \{f \in A_{k,n} : f(1) = f(k+1)\}$, or

$$\begin{aligned} \int x^k dEW_n(x) &= \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in B_{k,n}} \sum_{i \in T} \int_{D_i} \prod_{l=1}^k a_{f(l)f(l+1)} dP \\ &= \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in B_{k,n}} \int_{\left(\bigcup_{i \in T} D_i\right) \cup N = \Omega} \prod_{l=1}^k a_{f(l)f(l+1)} dP \\ &= \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in B_{k,n}} E \prod_{l=1}^k a_{f(l)f(l+1)}. \end{aligned}$$

Since $a_{ii} = 0$ a.s. this becomes

$$\int x^k dEW_n(x) = \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in C_{k,n}} E \prod_{l=1}^k a_{f(l)f(l+1)}$$

where

$$C_{k,n} = \{f \in B_{k,n} : f(l) \neq f(l+1), \quad l = 1, \dots, k\}.$$

Two cases are now considered, k odd and k even. Note that all the random variables a_{ij} are symmetric about 0, so that all odd moments vanish. Let $k = 2v + 1$. For $f \in C_{2v+1,n}$ one has

$\prod_{\ell=1}^{2\nu+1} a_{f(\ell)f(\ell+1)} = a_{f(i)f(i+1)}^m \prod a_{jk}$ a.s., for some i , where m is odd and the product $\prod a_{jk}$ involves no $a_{f(i)f(i+1)}$ or $a_{f(i+1)f(i)}$. Then, by independence and symmetry,

$$E \prod_{\ell=1}^{2\nu+1} a_{f(\ell)f(\ell+1)} = E a_{f(i)f(i+1)}^m E \prod a_{jk} = 0.$$

Thus,

$$\int_{\mathbb{R}} x^{2\nu+1} dEW_n(x) = 0 = \gamma_{2\nu+1}.$$

Now assume $k = 2\nu$. One need only consider those $f \in C_{2\nu,n}$ for which if $(i,j) \in \{g_f(1), g_f(2), \dots, g_f(2\nu)\}$, then $\#(i,j)_f + f(j,i)_f$ is even, for otherwise the argument of the odd case applies and the term vanishes.

Thus,

$$\int_{\mathbb{R}} x^{2\nu} dEW_n(x) = \frac{1}{(2\sigma)^{2\nu} n^{\nu+1}} \sum_{f \in D_{2\nu,n}} E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)}$$

where $D_{2\nu,n} = \{f \in C_{2\nu,n} : \text{if } (i,j) \in \{g_f(1), g_f(2), \dots, g_f(2\nu)\}, \text{ then } \#(i,j)_f + \#(j,i)_f \text{ is even}\}$. For $f \in D_{2\nu,n}$ one has, by Lemma 9,

$$E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)} = \sigma^{2\nu}.$$

Thus,

$$\int_{\mathbb{R}} x^{2\nu} dEW_n(x) = \frac{\#D_{2\nu,n}}{2^{2\nu} n^{\nu+1}}.$$

By Lemma 7, $f \in D_{2v,n}$ is such that $d_f \leq v + 1$. Thus,

$$\#D_{2v,n} = \sum_{j=1}^{v+1} \#D_{2v,n}^j$$

where $D_{2v,n}^j = \{f \in D_{2v,n} : d_f = j\}$. As with Lemma 11, one has

$$\#D_{2v,n}^j = \binom{n}{j} (\#D_{2v,j}^j).$$

Let

$$f_1(v,n) = \frac{\#D_{2v,n}^{v+1}}{2^{2v} n^{v+1}},$$

$$f_2(v,n) = \frac{\sum_{j=1}^v \binom{n}{j} (\#D_{2v,j}^j)}{2^{2v} n^{v+1}}.$$

Then

$$\int_{\mathbb{R}} x^{2v} dEW_n(x) = f_1(v,n) + f_2(v,n).$$

Since $\frac{\binom{n}{j}}{n^{v+1}} \rightarrow 0$ as $n \rightarrow \infty$ for $j = 1, \dots, v$, one has $f_2(v,n) \rightarrow 0$ as $n \rightarrow \infty$.

By Wigner's Combinatorial Theorem

$$\#D_{2v,n}^{v+1} = \frac{(2v)!}{v!(v+1)!} n^{v+1} + o(n^{v+1}).$$

Thus,

$$f_1(v,n) = \frac{(2v)!}{2^{2v} v!(v+1)!} + \frac{o(n^{v+1})}{2^{2v} n^{v+1}}$$

and

$$f_1(v, n) \rightarrow \frac{(2v)!}{2^{2v} v!(v+1)!}$$

as $n \rightarrow \infty$. Thus,

$$\int_{\mathbb{R}} x^{2v} dEW_n(x) \rightarrow \frac{(2v)!}{2^{2v} v!(v+1)!} = \gamma_{2v}$$

as $n \rightarrow \infty$. This completes the proof of Theorem 1.

Note the equation,

$$\int_{\mathbb{R}} x^{2v} dEW_n(x) = \frac{\#D_{2v, n}}{2^{2v} n^{v+1}},$$

derived during the course of the proof. A knowledge of $\#D_{2v, n}$ would give the sequence of moments of the distribution function $EW_n(x)$. As mentioned at the end of the Chapter III, these numbers are determined for all n by a knowledge of only $\#D_{2v, 1}^1, \#D_{2v, 2}^2, \dots, \#D_{2v, v+1}^{v+1}$.

4.2 Wigner's 1958 Conjecture

In 1958 Wigner (1958) conjectured the following result. Let

$A_n = (a_{ij})_{i, j=1}^n$ be a random matrix such that:

- (i) $a_{ij} = a_{ji}$ a.s;
- (ii) $\{a_{ij}, i \leq j\}$ is independent;
- (iii) the distribution function of each a_{ij} is absolutely continuous with density p_{ij} ;
- (iv) each a_{ij} is symmetric;
- (v) $E a_{ij}^2 = \sigma^2$ for all $1 \leq i, j \leq n$;

(vi) $E |a_{ij}|^k < C_k$ for all $1 \leq i, j \leq n$,

where C_k is independent of n .

Let $B_n = \frac{1}{2\sigma\sqrt{n}} A_n$ and denote by $W_n(x)$ the empirical distribution function of $\{\lambda_1(B_n), \lambda_2(B_n), \dots, \lambda_n(B_n)\}$ where $\lambda_1(B_n) \leq \lambda_2(B_n) \leq \dots \leq \lambda_n(B_n)$ are the ordered random eigenvalues of B_n . Under the above conditions one has

Wigner's Conjecture: $EW_n(x) \rightarrow W(x)$ as $n \rightarrow \infty$ for all $x \in R_1$, where $W(x)$ is the absolutely continuous distribution function with semi-circle density

$$w(x) = \begin{cases} \frac{2}{\pi} (1-x^2)^{\frac{1}{2}}, & |x| \leq 1 \\ 0 & , |x| > 1 . \end{cases}$$

In the next chapter we shall discuss work of Grenander (1963) who sketched a proof of convergence in probability of the empirical distribution functions to the semi-circle law. We shall also discuss the work of Arnold in this connection.

V. THE RESULTS OF GREANDER AND ARNOLD

5.1 Convergence in Probability

Grenander (1963) sketches a proof leading to the result given in this section.

Let $A_n = (a_{ij})_{i,j=1}^n$ be a random matrix such that:

- (i) $a_{ij} = a_{ji}$ a.s.;
- (ii) $\{a_{ij}, i \leq j\}$ is independent;
- (iii) a_{ij} is symmetric;
- (iv) $E a_{ij}^2 = \sigma^2$;
- (v) $|E a_{ij}^k| \leq C_k$, $k = 1, 2, \dots$, where C_k is independent of n .

Let $B_n = \frac{1}{2\sigma\sqrt{n}} A_n$ and denote by $W_n(x)$ the empirical distribution function of $\{\lambda_1(B_n), \lambda_2(B_n), \dots, \lambda_n(B_n)\}$, where $\lambda_1(B_n) \leq \lambda_2(B_n) \leq \dots \leq \lambda_n(B_n)$ are the ordered random eigenvalues of B_n .

Theorem 2 (Grenander (1963)): $W_n(x) \xrightarrow{P} W(x)$ as $n \rightarrow \infty$ for all $x \in R_1$, where W is the absolutely continuous distribution function with semi-circle density

$$w(x) = \begin{cases} \frac{2}{\pi} (1-x^2)^{\frac{1}{2}}, & |x| \leq 1 \\ 0, & |x| > 1. \end{cases}$$

Proof: Let

$$M_{k,n} = \int_{R_1} \lambda^k dW_n(x).$$

By Lemma 6 it will be sufficient to prove $M_{k,n} \rightarrow \nu_k = \int x^k dW(x)$ as $n \rightarrow \infty$

for all $k = 1, 2, \dots$. This will be achieved in two steps. First it will be shown that $E M_{k,n} \rightarrow \gamma_k$ as $n \rightarrow \infty$ as for all $k = 1, 2, \dots$. Then it will be shown that $E(M_{k,n} - EM_{k,n})^2 \rightarrow 0$ as $n \rightarrow \infty$ for all $k = 1, 2, \dots$. Tchebichev's inequality,

$$P(|M_{k,n} - EM_{k,n}| > \epsilon) \leq \frac{E(M_{k,n} - EM_{k,n})^2}{\epsilon^2},$$

then gives $M_{k,n} - EM_{k,n} \xrightarrow{P} 0$ as $n \rightarrow \infty$ for all $k = 1, 2, \dots$. This and $EM_{k,n} \rightarrow \gamma_k$ imply $M_{k,n} \xrightarrow{P} \gamma_k$ as $n \rightarrow \infty$ for all $k = 1, 2, \dots$.

It is now shown that $EM_{k,n} \rightarrow \gamma_k$ as $n \rightarrow \infty$. Let $A_{k,n}$ denote the class of all sequences $f : \{1, 2, \dots, k+1\} \rightarrow \{1, 2, \dots, n\}$. Then

$$\begin{aligned} EM_{k,n} &= E \int \lambda^k dW_n(x) = E \frac{1}{n} \sum_{i=1}^n \lambda_i(B_n) \\ &= E \frac{1}{n} \text{trace } B_n^k \\ &= E \frac{1}{(2\sigma)^k n^{\frac{k}{2} + 1}} \sum_{f \in B_{k,n}} \prod_{l=1}^k a_{f(l)f(l+1)} \\ &= \frac{1}{(2\sigma)^k n^{\frac{k}{2} + 1}} \sum_{f \in B_{k,n}} E \prod_{l=1}^k a_{f(l)f(l+1)}, \end{aligned}$$

where $B_{k,n} = \{f \in A_{k,n} : f(1) = f(k+1)\}$. As in Theorem 1, two cases are considered, k odd and k even. For exactly the reasons given in the proof of Theorem 1 one concludes immediately that for $k = 2\nu + 1$

$$EM_{2\nu+1,n} = 0 = \gamma_{2\nu+1}.$$

Now let $k = 2v$. If $f \in B_{2v,n}$ is such that there exists an $(f(i), f(i+1))$ such that $\#(f(i), f(i+1))_f + \#(f(i+1), f(i))_f = 1$, then by independence and symmetry, $E \prod_{l=1}^{2v} a_{f(l)f(l+1)} = E a_{f(i)f(i+1)} E \prod_{\substack{l=1 \\ l \neq i}}^{2v} a_{f(l)f(l+1)} = 0$. Thus

$$EM_{2v,n} = \frac{1}{(2\sigma)^{2v} n^{v+1}} \sum_{f \in C_{2v,n}} E \prod_{l=1}^{2v} a_{f(l)f(l+1)},$$

where $C_{2v,n} = \{f \in B_{2v,n} : \text{if } (i,j) \in \{g_f(1), g_f(2), \dots, g_f(2v)\}, \text{ then } \#(i,j)_f + \#(j,i)_f \geq 2\}$. By Lemma 7, $d_f \leq v+1$ for all $f \in C_{2v,n}$.

Let

$$f_1(v,n) = \frac{1}{(2\sigma)^{2v} n^{v+1}} \sum_{f \in C_{2v,n}^{v+1}} E \prod_{l=1}^{2v} a_{f(l)f(l+1)},$$

$$f_2(v,n) = \frac{1}{(2\sigma)^{2v} n^{v+1}} \sum_{j=1}^v \sum_{f \in C_{2v,n}^j} E \prod_{l=1}^{2v} a_{f(l)f(l+1)},$$

where $C_{2v,n}^j = \{f \in C_{2v,n} : d_f = j\}$. Then

$$EM_{2v,n} = f_1(v,n) + f_2(v,n).$$

By assumption (v) one has

$$|f_2(v,n)| \leq \frac{D_{2v}}{(2\sigma)^{2v} n^{v+1}} \sum_{j=1}^v \#C_{2v,n}^j$$

for some constant $D_{2v} < \infty$, or

$$|f_2(v,n)| \leq \frac{D_{2v} \sum_{j=1}^v \binom{n}{j} (\#C_{2v,n}^j)}{(2\sigma)^{2v} n^{v+1}}.$$

Since $\frac{\binom{n}{j}}{n^{v+1}} \rightarrow 0$ as $n \rightarrow \infty$ for $j = 1, \dots, v$, one has $f_2(v, n) \rightarrow 0$ as $n \rightarrow \infty$.

Let $f \in C_{2v, n}^{v+1}$. Then, using Lemma 9, one has

$$E \prod_{\ell=1}^{2v} a_{f(\ell)} f(\ell+1) = \sigma^{2v}.$$

Hence

$$f_1(v, n) = \frac{\#C_{2v, n}^{v+1}}{2^{2v} n^{v+1}}.$$

By Wigner's Combinatorial Theorem

$$f_1(v, n) = \frac{(2v)!}{2^{2v} v!(v+1)!} + \frac{o(n^{v+1})}{2^{2v} n^{v+1}}.$$

Hence

$$f_1(v, n) \rightarrow \frac{(2v)!}{2^{2v} v!(v+1)!} = \gamma_{2v}$$

as $n \rightarrow \infty$. Thus $EM_{k, n} \rightarrow \gamma_k$ as $n \rightarrow \infty$ for all $k = 1, 2, \dots$.

It will now be shown that $E(M_{k, n} - EM_{k, n})^2 \rightarrow 0$ as $n \rightarrow \infty$ for all $k = 1, 2, \dots$. For $f \in A_{2k+1, n}$ let $E(f)$ denote $E \prod_{i=1}^k a_{f(i)} f(i+1) \prod_{j=k+2}^{2k+1} a_{f(j)} f(j+1) - E \prod_{i=1}^k a_{f(i)} f(i+1) E \prod_{j=k+2}^{2k+1} a_{f(j)} f(j+1)$. One has after some manipulation

$$E(M_{k, n} - EM_{k, n})^2 = \frac{1}{(2\sigma)^{2k} n^{k+2}} \sum_{f \in B_{2k+1, n}} [E(f)]$$

where $B_{2k+1,n} = \{f \in A_{2k+1,n} : (i)f(1) = f(k+1); (ii) f(k+2) = f(2k+2);$
 (iii) if $(i,j) \in \{g_f(1), \dots, g_f(k), g_f(k+2), \dots, g_f(2k+1)\}$, then $\#(i,j)_f$
 $+ \#(j,i)_f \geq 2\}$. Condition (iii) follows from the independence and sym-
 metry conditions. It allows one to conclude, by arguments exactly as
 those of the proof of Lemma 7, that $d_f \leq k+1$ for all $f \in B_{2k+1,n}$. By
 assumption (v),

$$|E(f)| \leq \delta_k < \infty$$

where δ_k is independent of n . Thus

$$E(M_{k,n} - EM_{k,n})^2 \leq \frac{\delta_k (\#B_{2k+1,n})}{(2\sigma)^{2k} n^{k+2}}.$$

Now

$$\begin{aligned} \#B_{2k+1,n} &= \sum_{j=1}^{k+1} \#B_{2k+1,n}^j \\ &= \sum_{j=1}^{k+1} \binom{n}{j} (\#B_{2k+1,j}^j), \end{aligned}$$

where $B_{2k+1,n}^j = \{f \in B_{2k+1,n} : d_f = j\}$. Thus

$$E(M_{k,n} - EM_{k,n})^2 \leq \frac{\delta_k \sum_{j=1}^{k+1} \binom{n}{j} (\#B_{2k+1,n}^j)}{(2\sigma)^{2k} n^{k+2}}.$$

Since $\frac{\binom{n}{j}}{n^{k+2}} \rightarrow 0$ as $n \rightarrow \infty$ for $j = 1, 2, \dots, k+1$, one has

$$E(M_{k,n} - EM_{k,n})^2 \rightarrow 0$$

as $n \rightarrow \infty$. This completes the proof of Theorem 2.

5.2 Convergence Almost Surely

Arnold (1967) sketches a proof leading to the result given in this section.

Let $A_n = (a_{ij})_{i,j=1}^n$ be a random matrix such that:

- (i) $a_{ij} = a_{ji}$ a.s.;
- (ii) $\{a_{ij}, i \leq j\}$ is independent;
- (iii) the $a_{ij}, i \neq j$ are identically distributed with distribution function F , and the a_{ii} are identically distributed with distribution function G ;
- (iv) $E a_{ij} = \int x dF = 0, i \neq j$;
- (v) $E a_{ij}^2 = \int x^2 dF = \sigma^2, i \neq j$;
- (vi) (a) $E a_{ii}^2 = \int x^2 dG < \infty$,
 $E a_{ij}^4 = \int x^4 dF < \infty$;
- (b) $E a_{ii}^4 = \int x^4 dG < \infty$,
 $E a_{ij}^6 = \int x^6 dF < \infty$.

Let $B_n = \frac{1}{2\sigma\sqrt{n}} A_n$ and denote by $W_n(x)$ the empirical distribution function of $\{\lambda_1(B_n), \lambda_2(B_n), \dots, \lambda_n(B_n)\}$ where $\lambda_1(B_n) \leq \lambda_2(B_n) \leq \dots \leq \lambda_n(B_n)$ are the ordered random eigenvalues of B_n . Arnold (1967) then gives the

Theorem 3: Under conditions (i)-(v) and (vi)(a), $W_n(x) \xrightarrow{P} W(x)$ as $n \rightarrow \infty$ for all $x \in \mathbb{R}$, and under conditions (i)-(v) and (vi)(b), $W_n(x) \rightarrow W(x)$ a.s. as $n \rightarrow \infty$ for all $x \in \mathbb{R}_1$, where W is an absolutely continuous distribution function with semi-circle density

$$w(x) = \begin{cases} \frac{2}{\pi} (1-x^2)^{\frac{1}{2}}, & |x| \leq 1 \\ 0, & |x| > 1. \end{cases}$$

VI. ON WIGNER'S CONJECTURE (1958)

6.1 Theorem

We shall prove in this section the following theorem.

Let $A_n = (a_{ij})_{i,j=1}^n$ be a random matrix such that:

- (i) $a_{ij} = a_{ji}$ a.s.;
- (ii) $\{a_{ij}, i \leq j\}$ is independent;
- (iii) $Ea_{ij} = 0, 1 \leq i, j \leq n$;
- (iv) $Ea_{ij}^2 = \sigma^2, 1 \leq i \neq j \leq n$;
- (v) $E|a_{ij}|^k \leq M_k, 1 \leq i, j \leq n$, where M_k is not independent on n .

Let $B_n = \frac{1}{2\sigma\sqrt{n}} A_n$ and denote by $W_n(x)$ the empirical distribution function of $\{\lambda_1(B_n), \lambda_2(B_n), \dots, \lambda_n(B_n)\}$, where $\lambda_1(B_n) \leq \lambda_2(B_n) \leq \dots \leq \lambda_n(B_n)$ are the ordered random eigenvalues of B_n .

Theorem 4: $W_n(x) \rightarrow W(x)$ a.s. as $n \rightarrow \infty$ for all $x \in R_1$, where W is an absolutely continuous distribution function with semi-circle density

$$w(x) = \begin{cases} \frac{2}{\pi} (1-x^2)^{\frac{1}{2}}, & |x| \leq 1 \\ 0, & |x| > 1. \end{cases}$$

Proof: It is to be proved that

$$P(\lim_{n \rightarrow \infty} W_n(x) = W(x)) = 1.$$

Let

$$M_{k,n} = \int_{-\infty}^{\infty} x^k dW_n(x)$$

and

$$\gamma_k = \int_{-\infty}^{\infty} x^k dW(x) \\ = \begin{cases} 0 & , \text{ odd } k \\ \frac{k!}{2^k \left(\frac{k}{2}\right)! \left(\frac{k}{2}+1\right)!} & , \text{ even } k \end{cases} .$$

By Lemma 2, it will be sufficient to prove

$$P(\lim_{n \rightarrow \infty} M_{k,n} = \gamma_k, \quad k \geq 1) = 1 .$$

This will be true if

$$P(\lim_{n \rightarrow \infty} M_{k,n} = \gamma_k) = 1$$

for all $k > 1$. By the triangle inequality

$$|M_{k,n} - \gamma_k| \leq |M_{k,n} - EM_{k,n}| + |EM_{k,n} - \gamma_k| ,$$

and it will be sufficient to prove:

- (1) $\lim_{n \rightarrow \infty} EM_{k,n} = \gamma_k, \quad k \geq 1;$
- (2) $P(\lim_{n \rightarrow \infty} (M_{k,n} - EM_{k,n}) = 0) = 1, \quad k \geq 1 .$

For (2) it will be sufficient to prove

$$\sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty, \quad k \geq 1 .$$

This is seen as follows. The statement

$$P(\lim_{n \rightarrow \infty} (M_{k,n} - EM_{k,n}) = 0) = 1$$

is equivalent to

$$P(\limsup_{n \rightarrow \infty} \{\omega : |M_{k,n}(\omega) - EM_{k,n}| > \epsilon\}) = 0$$

for every $\epsilon > 0$, by Lemma 3. Let $A_n = \{\omega : |M_{k,n}(\omega) - EM_{k,n}| > \epsilon\}$. It is to be shown that $P(\limsup_{n \rightarrow \infty} A_n) = 0$. Tchebichev's inequality gives

$$P(|M_{k,n} - EM_{k,n}| > \epsilon) \leq \frac{E(M_{k,n} - EM_{k,n})^2}{\epsilon^2}.$$

This and $\sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty$ implies $\sum_{n=1}^{\infty} P(A_n) < \infty$. Then Lemma 4 (Borel-Cantelli) gives $P(\limsup_{n \rightarrow \infty} A_n) = 0$. Altogether, then, it will be sufficient to prove:

- (i) $\lim_{n \rightarrow \infty} EM_{k,n} = \gamma_k, \quad k \geq 1;$
(ii) $\sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty, \quad k \geq 1.$

The proof of (i) follows.

One has, letting $A_{k,n}$ denote the class of all sequence $f: \{1, 2, \dots, k+1\} \rightarrow \{1, 2, \dots, n\}$,

$$\begin{aligned} EM_{k,n} &= E \int \lambda^k dW_n(x) = E \frac{1}{n} \sum_{i=1}^n \lambda_i^k(B_n) \\ &= E \frac{1}{n} \text{trace } B_n^k \\ &= E \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in B_{k,n}} \prod_{\ell=1}^k a_{f(\ell)f(\ell+1)} \\ &= \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in B_{k,n}} E \prod_{\ell=1}^k a_{f(\ell)f(\ell+1)}, \end{aligned}$$

where $B_{k,n} = \{f \in A_{k,n} : f(1) = f(k+1)\}$. Let $f \in B_{k,n}$ be such that there exists $(f(i), f(i+1)) \in \{g_f(1), g_f(2), \dots, g_f(k)\}$ such that $\#(f(i), f(i+1))_f + \#(f(i+1), f(i))_f = 1$. Then, by the independence and zero mean assumptions, $E \prod_{l=1}^k a_{f(l)f(l+1)} = E a_{f(i)f(i+1)} E \prod_{\substack{l=1 \\ l \neq i}}^k a_{f(l)f(l+1)} = 0$. Thus one has

$$E N_{k,n} = \frac{1}{(2\sigma)^k n^{\frac{k}{2}+1}} \sum_{f \in C_{k,n}} E \prod_{l=1}^k a_{f(l)f(l+1)}$$

where $C_{k,n} = \{f \in B_{k,n} : \text{if } (i,j) \in \{g_f(1), g_f(2), \dots, g_f(k)\}, \text{ then } \#(i,j)_f + \#(j,i)_f \geq 2\}$. Two cases are now considered, k odd and k even. Let $k = 2\nu + 1$. By Lemma 7, $d_f \leq \nu + 1$ for all $f \in C_{2\nu+1,n}$. Thus

$$E M_{2\nu+1,n} = \frac{1}{(2\sigma)^{2\nu+1} n^{\nu+\frac{3}{2}}} \sum_{j=1}^{\nu+1} \sum_{f \in C_{2\nu+1,n}^j} E \prod_{l=1}^{2\nu+1} a_{f(l)f(l+1)}$$

where $C_{2\nu+1,n}^j = \{f \in C_{2\nu+1,n} : d_f = j\}$. By assumption (v),

$$\left| E \prod_{l=1}^{2\nu+1} a_{f(l)f(l+1)} \right| \leq D_\nu < \infty,$$

for some constant D_ν . Hence

$$\begin{aligned} |E M_{2\nu+1,n}| &\leq \frac{D_\nu \sum_{j=1}^{\nu+1} \#C_{2\nu+1,n}^j}{(2\sigma)^{2\nu+1} n^{(\nu+1)+\frac{1}{2}}} \\ &= \frac{D_\nu \sum_{j=1}^{\nu+1} \binom{n}{j} (\#C_{2\nu+1,n}^j)}{(2\sigma)^{2\nu+1} n^{(\nu+1)+\frac{1}{2}}}. \end{aligned}$$

Since $\frac{\binom{n}{j}}{n^{(v+1)+\frac{1}{2}}} \rightarrow 0$ as $n \rightarrow \infty$ for $j = 1, 2, \dots, v+1$, one has $EM_{2v+1, n} \rightarrow 0$ as $n \rightarrow \infty$. Now consider $k = 2v$. By Lemma 7, $d_f \leq v+1$ for all $f \in C_{2v, n}$. Let

$$f_1(v, n) = \frac{1}{(2\sigma)^{2v} n^{v+1}} \sum_{f \in C_{2v, n}^{v+1}} E \prod_{l=1}^{2v} a_{f(l)f(l+1)},$$

$$f_2(v, n) = \frac{1}{(2\sigma)^{2v} n^{v+1}} \sum_{f \in \bigcup_{j=1}^v C_{2v, n}^j} E \prod_{l=1}^{2v} a_{f(l)f(l+1)},$$

where $C_{2v, n}^j = \{f \in C_{2v, n} : d_f = j\}$. Note that

$$EM_{2v, n} = f_1(v, n) + f_2(v, n).$$

By assumption (v),

$$\left| E \prod_{l=1}^{2v} a_{f(l)f(l+1)} \right| \leq \delta_v < \infty,$$

for some constant δ_v . Thus

$$\begin{aligned} |f_2(v, n)| &\leq \frac{\delta_v \sum_{j=1}^v \#C_{2v, n}^j}{(2\sigma)^{2v} n^{v+1}} \\ &= \frac{\delta_v \sum_{j=1}^v \binom{n}{j} (\#C_{2v, j}^j)}{(2\sigma)^{2v} n^{v+1}}. \end{aligned}$$

Since $\frac{\binom{n}{j}}{n^{v+1}} \rightarrow 0$ as $n \rightarrow \infty$ for $j = 1, \dots, v$, one has $f_2(v, n) \rightarrow 0$ as $n \rightarrow \infty$.

By Lemma 9,

$$E \prod_{l=1}^{2\nu} a_{f(l)f(l+1)} = \sigma^{2\nu}$$

for all $f \in C_{2\nu, n}^{\nu+1}$. Thus

$$f_1(\nu, n) = \frac{\sigma^{2\nu} (\#C_{2\nu, n}^{\nu+1})}{(2\sigma)^{2\nu} n^{\nu+1}}.$$

By Wigner's Combinatorial Theorem,

$$f_1(\nu, n) = \frac{(2\nu)!}{2^{2\nu} \nu! (\nu+1)!} + \frac{o(n^{\nu+1})}{(2\sigma)^{2\nu} n^{\nu+1}}.$$

Thus

$$f_1(\nu, n) \rightarrow \frac{(2\nu)!}{2^{2\nu} \nu! (\nu+1)!} = \gamma_{2\nu}$$

as $n \rightarrow \infty$. Altogether

$$EM_{2\nu, n} \rightarrow \frac{(2\nu)!}{2^{2\nu} \nu! (\nu+1)!} = \gamma_{2\nu}$$

as $n \rightarrow \infty$. This completes the proof of (i). The proof of (ii) follows.

Consider $E(M_{k, n} - EM_{k, n})^2$. For $f \in A_{2k+1, n}$ let $E(f)$ denote

$$E \prod_{i=1}^k a_{f(i)f(i+1)} \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)} - E \prod_{i=1}^k a_{f(i)f(i+1)} E \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)}.$$

One has

$$E(M_{k, n} - EM_{k, n})^2 = \frac{1}{(2\sigma)^{2k} n^{k+2}} \sum_{f \in B_{2k+1, n}} E(f)$$

where $B_{2k+1,n} = \{f \in A_{2k+1,n} : \text{(i) } f(1) = f(k+1); \text{(ii) } f(k+2) = f(2k+2);$
 (iii) $\{g_f(1), g_f(2), \dots, g_f(k)\} \cap \{g_f(k+2), \dots, g_f(2k+1), g_f^*(k+2), \dots,$
 $g_f^*(2k+1)\} = \emptyset$, where $g_f^*(l)$ denotes the reverse of $g_f(l)$; (iv) $E(f) = 0\}$.

Reasons for conditions (i) and (ii) are obvious. If condition (iii) is not met by f , the term in the summation corresponding to f will be zero, by the independence assumption. Condition (iv) is trivial. It will now

be shown that if $f \in B_{2k+1,n}$ then $d_f \leq k$. Using condition (iii), suppose, for the sake of definiteness, that $g_f(s) = g_f(t)$ for some s ,

$1 \leq s \leq k$, and some t , $k+2 \leq t \leq 2k+1$. (The only other case to consider

is when $g_f(s) = g_f^*(t)$ for some s , $1 \leq s \leq k$, and some t , $k+2 \leq t \leq 2k+1$,

for which the following argument also applies.) Define a new sequence

$h \in A_{2k+1,n}$ as follows: $h(1) = f(s), h(2) = f(s+1), \dots, h(k-s+1) = f(k),$

$h(k-s+2) = f(1), h(k-s+3) = f(2), \dots, h(k) = f(s-1), h(k+1) = f(s), h(k+2)$

$= f(t+1), h(k+3) = f(t+2), \dots, h(2k-t+2) = f(2k+1), h(2k-t+3) = f(k+3), \dots,$

$h(2k+1) = f(t-1), h(2k+2) = f(t)$. It is immediate that $d_h = d_f$. The

sequence of steps associated with h is

$$g_h(1) = (f(s), f(s+1)), \dots, g_h(k-s+1) = (f(k), f(1)),$$

$$g_h(k-s+2) = (f(1), f(2)), \dots, g_h(k) = (f(s-1), f(s)),$$

$$g_h(k+1) = (f(s), f(t+1)) = (f(s), f(s+1)), \dots,$$

$$g_h(2k-t+2) = (f(2k+1), f(k+2)), g_h(2k-t+3)$$

$$= (f(k+2), f(k+3)), \dots, g_h(2k+1) = (f(t-1), f(t)).$$

It is true that:

$$\text{(i) } h(1) = h(2k+2);$$

$$\text{(ii) if } (i, j) \in \{g_h(1), g_h(2), \dots, g_h(2k+1)\},$$

$$\text{then } \#(i, j)_h + \#(j, i)_f \geq 2.$$

(i) is immediate. To see (ii) one proceeds as follows. If (i,j) equals $g_h(k+1) = (f(t), f(t+1))$ or $g_h^*(k+1) = (f(t+1), f(t))$, then $\#(i,j)_h + \#(j,i)_h \geq 2$ since $g_h(1) = g_h(k+1)$. On the other hand if (i,j) equals any other step among $g_h(1), \dots, g_h(k), g_h(k+2), \dots, g_h(2k+1)$, and $\#(i,j) + \#(j,i)_h = 1$, then the independence assumption implies

$$\begin{aligned} & E\left(\prod_{i=1}^k a_{f(i)f(i+1)} \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)}\right) - E\left(\prod_{i=1}^k a_{f(i)f(i+1)} \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)}\right) \\ & = E\left(\prod_{i=1}^k a_{h(i)h(i+1)} \prod_{j=k+2}^{2k+1} a_{h(j)h(j+1)}\right) - E\left(\prod_{i=1}^k a_{h(i)h(i+1)} \prod_{j=k+2}^{2k+1} a_{h(j)h(j+1)}\right) = 0 \end{aligned}$$

contrary to the assumption that for $f \in B_{2k+1,n}$ this term is nonzero.

Hence h must satisfy condition (ii) above. Lemma 10 then applies, giving

$d_f = d_h \leq k$. Now consider

$$E(M_{k,n} - EM_{k,n})^2 = \frac{1}{(2\sigma)^{2k} n^{k+2}} \sum_{j=1}^k \sum_{f \in B_{2k+1,n}^j} E(f)$$

where $B_{2k+1,n}^j = \{f \in B_{2k+1,n} : d_f = j\}$. By assumption (v), one has

$$|E(f)| \leq G_v < \infty$$

for some constant G_v . Thus

$$\begin{aligned} |E(M_{k,n} - EM_{k,n})^2| & \leq \frac{G_v \sum_{j=1}^k (\#B_{2k+1,n}^j)}{(2\sigma)^{2k} n^{k+2}} \\ & = \frac{G_v \sum_{j=1}^k \binom{n}{j} (\#B_{2k+1,j}^j)}{(2\sigma)^{2k} n^{k+2}}. \end{aligned}$$

Since

$$\sum_{n=1}^{\infty} \frac{\binom{n}{j}}{n^{k+2}} < \infty$$

for $j = 1, 2, \dots, k$, one has, by the comparison test for series

$$\sum_{n=1}^{\infty} |E(M_{k,n} - EM_{k,n})^2| < \infty$$

which implies

$$\sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty$$

which was to be proved. This completes the proof of Theorem 4.

6.2 Comments

A little reflection will reveal that the assumption of zero means for the diagonal elements is not necessary. For, in proving $EM_{k,n} \rightarrow \gamma_k$ as $n \rightarrow \infty$ it was established that the only sequences of interest were those $f \in A_{k,n}$ for which (i) $f(1) = f(k+1)$ and (ii) if $(i, j) \in \{g_f(1), \dots, g_f(k)\}$, then $\#(i, j)_f + (j, i)_f \geq 2$. Condition (ii) alone implies $d_f \leq \lfloor \frac{k}{2} \rfloor + 1$. If one assumes, however, that (iia) if $(i, j) \in \{g_f(1), \dots, g_f(k)\}$, where $i \neq j$, then $\#(i, j)_f + \#(j, i)_f \geq 2$, then conditions (i) and (iia) together imply $d_f \leq \lfloor \frac{k}{2} \rfloor + 1$. For odd k , say $k = 2\nu + 1$, one has $EM_{2\nu+1, n} \rightarrow 0 = \gamma_{2\nu+1}$ as $n \rightarrow \infty$ exactly as before. For even k , say $k = 2\nu$, the only sequences of interest are those f such that $d_f = \nu + 1$. If $d_f = \nu + 1$, then $f(i) = f(i+1)$ for some i is not possible, since under conditions (i) and (iia) arguments similar to those of the proof of Lemma 8 would give $d_f \leq \nu$. Thus Wigner's Combinatorial Theorem holds under condi-

tions (i), (iia), and (iii) $d_F = \nu + 1$; application of this theorem then gives $EM_{2\nu, n} \rightarrow \gamma_{2\nu}$ as $n \rightarrow \infty$ exactly as before. Note that use of the property of zero expectation of diagonal elements has been eliminated by substitution of condition (iia) for condition (ii). Similar arguments also hold for the proof that $\sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty$.

That the off-diagonal elements all have second moments equal to σ^2 is not necessary. An examination of the proof shows that it is sufficient to assume that the ratio of the number of elements of the matrix having the same second moment to the total number of elements of the matrix approach 1 as the dimension becomes arbitrarily large.

It should be noted that Wigner's conjecture of 1958 is a special case of Theorem 4. The result indicated by Grenander is a special case of Theorem 4. Wigner's conjecture is not a special case of the theorem indicated by Arnold, for Arnold assumes the diagonal random variables are identically distributed and the off-diagonal random variables are identically distributed. Arnold does drop the requirements, given by Wigner, of symmetric random variables and the existence of higher order moments. Theorem 4 is not only more general than the result conjectured by Wigner in the sense that it deals with almost sure convergence, but it also drops Wigner's requirement of symmetric random variables.

VII. ROTATIONAL INVARIANCE

7.1 Characterization of the Distribution of a Random Matrix

In quantum mechanics, under certain symmetry conditions, energy is represented by a real symmetric matrix X . If for a first observer energy is represented by X , then for a second observer with a rotated coordinate system energy is represented by OXO' , where O is the orthogonal matrix relating the axes of the observers. Descriptions based on X and OXO' are completely equivalent physically (Wigner, 1959). Thus if a statistical hypothesis is made on X then it is natural to make the same statistical hypothesis on OXO' . The following makes this precise and characterizes the possible statistical hypotheses.

Let $\{x_{ij}\}_{i \leq j}$, $i, j = 1, 2, \dots, n$ be an independent set of random variables on a probability space (Ω, \mathcal{F}, P) . Let

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{bmatrix}$$

where $x_{ij} = x_{ji}$ a.s., and let

$$Y_0 = (y_{ij}^0) = OXO' ,$$

where O is any orthogonal matrix. Let $x = (x_{11}, \dots, x_{1n}, x_{22}, \dots, x_{2n}, \dots, x_{nn})$ and $y_0 = (y_{11}^0, \dots, y_{1n}^0, y_{22}^0, \dots, y_{2n}^0, \dots, y_{nn}^0)$. Let β_n denote the Borel σ -algebra of subsets of the n -dimensional Euclidean space R_n . We do not consider the case where $x = 0$ a.s.

This theorem seems to have been first proved in this context under more restrictive conditions than those given here by Porter and Rosenzweig (1960).

Theorem 5: $P(a \in B) = P(y_0 \in B)$

for all $B \in \beta_{n(n+1)/2}$ and all orthogonal O if and only if x_{ii} is normal with mean μ and variance $2a^2$ and x_{ij} , $i < j$, is normal with 0 and variance a^2 , for some constants μ and $a^2 > 0$.

Proof: Assume the x_{ij} are normal as stated in the theorem and let $O = (O_{ij})$ by any orthogonal matrix. Then

$$y_{ij}^0 = \sum_{k=1}^n O_{ik} O_{jk} x_{kk} + \sum_{k < m} (O_{ik} O_{jm} O_{im} O_{jk}) x_{km} \text{ a.s.}$$

is normal since it is a linear combination of independent normal random variables. Elementary calculations show:

$$E(y_{ij}^0) = \delta_{ij} \mu$$

$$\text{Var}(y_{ij}^0) = a^2 + \delta_{ij} a^2$$

$$\text{Cov}(y_{ij}^0, y_{kl}^0) = 0$$

for all $i \leq j$, $k \leq l$, where $i \neq k$ or $j \neq l$ and δ_{ij} is the Kronecker symbol. Thus y_{ij}^0 has the same distribution as x_{ij} and $\{y_{ij}^0\}_{i \leq j}$ is independent since covariance between normal random variables equaling zero implies their independence. Thus x and y_0 have the same multivariate normal distribution.

Now assume $P(x \in B) = P(y_0 \in B)$ for all $B \in \beta_{n(n+1)/2}$ and all orthogonal O . This implies $P(x_{ij} \in I) = P(y_{ij}^0 \in I)$ for all $I \in \beta$ and all orthogonal O . By considering orthogonal matrices formed from the identity matrix by permuting rows and columns in any manner one concludes, using the relation $P(x_{ij} \in I) = P(y_{ij}^0 \in I)$, the x_{ii} , $i = 1, \dots, n$, have the same distribution, and the x_{ij} , $i < j$, have the same distribution. Now consider the special orthogonal matrix

$$O = \begin{bmatrix} O_1 & O_2 \\ O_2' & I_{n-2} \end{bmatrix}$$

where

$$O_1 = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix},$$

I_{n-2} is the $(n-2) \times (n-2)$ identity matrix and O_2 is the $2 \times (n-2)$ matrix with elements equal to zero. The transformation

$$Y_0 = (y_{ij}^0) = OXO'$$

gives rise to the following equations:

$$\begin{aligned} y_{11}^0 &= x_{11} \cos^2 \theta + x_{12} 2 \sin \theta \cos \theta + x_{22} \sin^2 \theta \\ y_{12}^0 &= x_{11} (-\sin \theta \cos \theta) + x_{22} (\cos^2 \theta - \sin^2 \theta) + x_{22} \sin \theta \cos \theta \\ y_{22}^0 &= x_{11} \sin^2 \theta + x_{12} (-2 \sin \theta \cos \theta) + x_{22} \cos^2 \theta \\ y_{1j}^0 &= x_{1j} \cos \theta + x_{2j} \sin \theta, \quad 3 \leq j \leq n \\ y_{2j}^0 &= x_{1j} (-\sin \theta) + x_{2j} \cos \theta, \quad 3 \leq j \leq n \\ y_{ij}^0 &= x_{ij}, \quad i \leq j, \quad i \neq 1, 2. \end{aligned}$$

Consider

$$y_{2j}^0 = x_{1j} (-\sin \theta) + x_{2j} \cos \theta .$$

Using this equation for $\theta = 0, \pi/2$ (say) and the relation $P(x_{ij} \in I) = P(y_{ij}^0 \in I)$, one concludes the off-diagonal elements are symmetric about zero mean. Let φ_d and φ_0 denote the characteristic functions of the diagonal and off-diagonal elements, respectively. Then φ_0 is real valued and since $\varphi_0(t) = \varphi_0(-t), \varphi_0(0) = 1$, one need determine φ_0 only for $t > 0$. The equation

$$y_{1j}^0 = x_{1j} \cos \theta + x_{2j} \sin \theta$$

gives rise to the functional equation

$$\varphi_0(t) = \varphi_0(t \cos \theta) \varphi_0(t \sin \theta)$$

by the independence of x_{1j} and x_{2j} . Letting $\alpha = \cos \theta, \beta = \sin \theta$, one has

$$\varphi_0(t) = \varphi_0(\alpha t) \varphi_0(\beta t)$$

where $\alpha^2 + \beta^2 = 1$.

For $t > 0$ define

$$f(t) = \varphi_0(\sqrt{t}) .$$

Then

$$\begin{aligned} f(t^2) &= \varphi_0(t) = \varphi_0(\alpha t) \varphi_0(\beta t) \\ &= f(\alpha^2 t^2) f(\beta^2 t^2) . \end{aligned}$$

Letting $u = \alpha^2 t^2$, $v = \beta^2 t^2$, $u + v = t^2$, one has

$$f(u + v) = f(u) f(v), \quad \text{for all } u, v > 0 .$$

Every solution of $f(u + v) = f(u) f(v)$, $u, v > 0$, is either everywhere or nowhere zero. For if there were a z_0 with $f(z_0) = 0$, then it would follow that

$$f(t) = f[(t - z_0) + z_0] = f(t - z_0) f(z_0) = 0$$

for all $t \geq z_0$. If there were a $t_0 \in (0, z_0)$ such that $f(t_0) \neq 0$, then there would be an n such that $nt_0 \geq z_0$. Thus $0 = f(nt_0) = [f(t_0)]^n \neq 0$, which is impossible. Hence $f(t) = 0$ for all $t > 0$ if $f(t) = 0$ for some $t > 0$. But since $\varphi_0(t) \rightarrow 1$ as $t \rightarrow 0$, one has $f(t) \rightarrow 1$ as $t \rightarrow 0$ and thus f must be nowhere zero. Let $u = v = \frac{t}{2}$; then

$$f(u + v) = f(t) = f^2\left(\frac{t}{2}\right) > 0 .$$

Thus $f(t) > 0$ for all $t > 0$. Then

$$\ln f(u + v) = \ln f(u) + \ln f(v) .$$

Let $g = \ln f$. Then for $u_i > 0$, $i = 1, \dots, n$,

$$g(u_1 + u_2 + \dots + u_n) = g(u_1) + g(u_2) + \dots + g(u_n) .$$

Let $u_k = u > 0$, $k = 1, \dots, n$; then

$$g(nu) = n g(u) .$$

If $u = \frac{m}{n} t$,

$$nu = mt$$

and

$$g(nu) = g(mt)$$

or

$$n g(u) = m g(t)$$

or

$$g\left(\frac{m}{n} t\right) = \frac{m}{n} g(t) .$$

Let $t = 1$; then

$$g\left(\frac{m}{n}\right) = \frac{m}{n} g(1)$$

or

$$g(\lambda) = \lambda g(1)$$

for all rational $\lambda > 0$. For any real $x > 0$ there exists a sequence $\{\lambda_n\}$ of rationals $\lambda_n > 0$ such that $\lambda_n \rightarrow x$ as $n \rightarrow \infty$. By the continuity of g ,

$$\begin{aligned} g(x) &= g(\lim_{n \rightarrow \infty} \lambda_n) = \lim_{n \rightarrow \infty} g(\lambda_n) \\ &= \lim_{n \rightarrow \infty} \lambda_n g(1) \\ &= x g(1) \end{aligned}$$

for all $x > 0$.

Thus

$$g(x) = cx, \quad c = g(1)$$

i.e.,

$$\ln f(x) = cx.$$

Since $|f(x)| < 1$ for at least one $x > 0$ (the distributions are not degenerate at 0 by assumption), one has $c < 0$. Let $-c = \frac{a^2}{2}$. Then

$$f(x) = e^{-\frac{a^2}{2}x}$$

and thus

$$\varphi_0(t) = f(t^2) = e^{-\frac{t^2}{2}a^2}$$

for all $t \in \mathbb{R}_1$. This is the characteristic function of a normal distribution with mean 0 and variance a^2 . The equation for y_{12}^0 with $\theta = \frac{\pi}{4}$ is

$$y_{12}^0 = \left(-\frac{1}{2}x_{11}\right) + \left(\frac{1}{2}x_{22}\right).$$

Thus the sum of two independent random variables is normally distributed and hence each component is normally distributed by Lemma 6. Denote by μ and σ^2 the mean and variance of the diagonal elements. The above equation gives

$$a^2 = \text{var}(y_{12}^0) = \frac{1}{4}\sigma^2 + \frac{1}{4}\sigma^2 = \frac{1}{2}\sigma^2$$

or $\sigma^2 = 2a^2$. This completes the proof of the theorem.

7.2 Gaussian Orthogonal Ensemble

If one assumes that X is a random matrix such that: (i) X is symmetric; (ii) the set of diagonal and superdiagonal elements of X form an independent set of random variables; and (iii) the distribution of X is

invariant under orthogonal similarity transforms, then Theorem 5 allows one to say the elements of X are normally distributed as indicated in the theorem. The physicists call this model the Gaussian orthogonal ensemble. Theorem 4 gives the semi-circle law as the almost sure limit of the empirical distribution function of the eigenvalues of the normalized random matrix $X/2a\sqrt{n}$.

Mehta and Gaudin (1961) examine the asymptotic eigenvalue distribution for the Gaussian orthogonal ensemble. Using standard methods of multivariate analysis (see Theorems 13.3.4 and 13.3.5 in Anderson (1958)) they derive the joint density of the eigenvalues. Due to exchangeability of the unordered eigenvalues the marginal densities of the same number of variables are the same (see Loeve (1963) for the concept of exchangeability). An expression for the marginal density of a single random eigenvalue is given and indications of why this tends asymptotically to the semi-circle density are outlined by Mehta and Gaudin.

VIII. BIBLIOGRAPHY

1. Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York.
2. Arnold, L. (1967), "On the Asymptotic Distribution of the Eigenvalues of Random Matrices," J. Math. Anal. Appl. 20, 262.
3. Chung, K. L. (1968), A Course in Probability Theory, Harcourt, Brace and World, New York.
4. Grenander, U. (1963), Probabilities on Algebraic Structures, John Wiley and Sons, New York.
5. Jordan, T. F. (1969), Linear Operators for Quantum Mechanics, John Wiley and Sons, Inc., New York.
6. Loeve, M. (1963), Probability Theory, Third Edition, D. Van Nostrand Company, Inc., Princeton, New Jersey.
7. Mehta, M. L. and Gaudin, M. (1961), "Addendum and Erratum," Nucl. Phys. 22, 340.
8. Mehta, M. L. (1967), Random Matrices, Academic Press, New York.
9. Ostrowski, A. M. (1960), Solution of Equations and Systems of Equations, Academic Press, New York.
10. Porter, C. E., editor (1965), Statistical Theories of Spectra: Fluctuations, Academic Press, New York.
11. Rosenzweig, N. and Porter, C. E. (1960), "Statistical Properties of Atomic and Nuclear Spectra," Suomalaisen Tiedeakatemia Toimituksia (Ann. Acad. Sci. Fennicae) AVI (Reprinted in Porter, C. E. (1965)).
12. Wigner, E. P. (1955), "Characteristic Vectors of Bordered Matrices with Infinite Dimensions," Ann. Math. 62, 548.
13. Wigner, E. P. (1958), "On the Distribution of the Roots of Certain Symmetric Matrices," Ann. Math. 67, 325.
14. Wigner, E. P. (1959), Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra, Academic Press, New York.

**The vita has been removed from
the scanned document**

ASYMPTOTIC DISTRIBUTION OF EIGENVALUES OF RANDOM
MATRICES AND CHARACTERIZATION OF THE GAUSSIAN
DISTRIBUTION BY ROTATIONAL INVARIANCE

by

William H. Olson

Abstract

The study falls in the area of random equations; in particular properties of random matrices have been studied. The dissertation makes precise some statistical theories of spectra developed in recent years by a number of physicists. Two basic results have been achieved.

The first result is a characterization of the distribution of a symmetric random matrix. Assuming independence of the diagonal and super-diagonal random variables of the symmetric random matrix the following theorem is proved: the distribution of the matrix is invariant under orthogonal similarity transforms if and only if the diagonal random variables are normally distributed with mean μ and variance $2a^2$, and the off-diagonal elements are normally distributed with mean 0 and variance a^2 , for some constants μ and $a^2 > 0$. The proof is achieved by solving a functional equation in characteristic functions. This seems to have been first proved in this context by Porter and Rosenzweig (Ann. Acad. Sci. Fennicae. AVI, No. 44, 1960) by a different method and under more restrictive conditions than those given here.

The second result deals with the asymptotic distribution of eigenvalues of a symmetric random matrix as the dimension approaches infinity. Let A_n be an appropriately normalized $n \times n$ symmetric random matrix and

let $W_n(x)$ denote the empirical distribution function of the eigenvalues of A_n . Under suitable conditions on the random variables of the matrix it is proved that $W_n(x) \rightarrow W(x)$ as $n \rightarrow \infty$, where W is the absolutely continuous distribution function with a semi-circle density,

$$w(x) = \begin{cases} \frac{2}{\pi} (1-x^2)^{\frac{1}{2}}, & |x| \leq 1 \\ 0 & , \quad |x| > 1 . \end{cases}$$

The proof is achieved by an intricate combinatorial analysis in conjunction with the method of moments. This result extends a conjecture made by E. P. Wigner ("On the Distribution of the Roots of Certain Symmetric Matrices," Ann. Math. 67, 1958, 325).