ASYMPTOTIC DISTRIBUTION OF EIGENVALUES OF RANDOM MATRICES

W. H. OLSON[1]
and
V. R. R. UPPULURI[2]
OAK RIDGE NATIONAL LABORATORY

1. 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 [10] and the collection of reprints edited by C. E. Porter [14] 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 [8].

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 [8]. This is expressed in the Schrödinger time independent equation,

$$H\psi = E\psi, \tag{1.1.1}$$

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 laws governing the system 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.

[1] Research partially supported by the Oak Ridge Associated Universities and the National Institute of Environmental Health Sciences through grant numbers 5 T01 ES 00033-03, 04 and 05.
[2] Research sponsored by the U.S. Atomic Energy Commission under contract with the Union Carbide Corporation.

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 (see [10]). 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, that is, of the Hamiltonian. The basic statistical hypothesis is this: the statistical behavior of energy levels in a simple sequence (a simple sequence is one whose levels all have the same set of values of the parameters mentioned above) is identical with the behavior of the eigenvalues of a random matrix. 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.

1.2. *Outline of contents.* There are three basic parts. Section 3 contains the combinatorial arguments which are essential for the proofs of the theorems in the second part, Sections 4, 5, and 6. These sections 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, Section 7, discusses results of the same type, that is, asymptotic distributions of the empirical distribution function of the eigenvalues of random matrices, for the Gaussian orthogonal ensemble [10], a Toeplitz ensemble [5], and a Wishart ensemble.

2. Notation, definitions and preliminaries

2.1. *Random matrices.* Let $(\Omega, \mathscr{F}, P)$ denote a probability space, that is, Ω is a nonempty abstract set, $\mathscr{F}$ is a σ-algebra of subsets of Ω, and P is a probability measure on $\mathscr{F}$; 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 $X : \Omega \twoheadrightarrow R_n$ is called a *random vector* if $\{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F}$ for all $B \in \beta_n$. When $n = 1$, X is called a *random variable.*

A mapping $A : \Omega \times R_n \twoheadrightarrow R_n$ 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

$$A = (a_{ij})_{i,j=1}^{n}, \tag{2.1.1}$$

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 $(\Omega, \mathscr{F}, P)$.

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 $\lambda_i(A)$, $i = 1, \cdots, n$, the eigenvalues of any $(n \times n)$ matrix A.

LEMMA 2.2.1. *Let A be an $(n \times n)$ matrix and suppose $\varepsilon > 0$ is given. Then there exists a $\delta > 0$ such that for any matrix $D = (d_{ij})_{i,j=1}^{n}$ such that $\Sigma_{i,j=1}^{n} |d_{ij}| < \delta$, and there exists a permutation σ of $\{1, 2, \cdots, n\}$ for which*

$$|\lambda_i(A) - \lambda_{\sigma(i)}(A + D)| < \varepsilon, \qquad i = 1, \cdots, n. \tag{2.2.1}$$

A proof of this lemma may be found in A. M. Ostrowski [13].

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

COROLLARY 2.2.1. *The ordered eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ are continuous functions of the elements of the Hermitian matrix $A = (a_{ij})$.*

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

$$|\lambda_i(A) - \lambda_{\sigma(i)}(A')| < \varepsilon \tag{2.2.2}$$

for $i = 1, \cdots, n$. Assume $|\lambda_i(A) - \lambda_i(A')| > \varepsilon$; to be definite assume $\lambda_i(A') > \lambda_i(A)$. Then $\lambda_1(A) \leqq \cdots \leqq \lambda_i(A) < \lambda_i(A') \leqq \cdots \leqq \lambda_n(A')$. With each $\lambda_j(A)$, $j = 1, \cdots, i$, is associated $\lambda_{\sigma(j)}(A')$ such that $|\lambda_j(A) - \lambda_{\sigma(j)}(A')| < \varepsilon$. But only $\lambda_1(A'), \cdots, \lambda_{i-1}(A')$ are available for this purpose since $\lambda_i(A') - \lambda_i(A) > \varepsilon$, and hence $\lambda_i(A') - \lambda_j(A) > \varepsilon$, $j = 1, \cdots, i$. Thus, one must conclude $|\lambda_i(A) - \lambda_i(A')| < \varepsilon$, $i = 1, \cdots, n$. 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 $\lambda_1(A) \leqq \lambda_2(A) \leqq \cdots \leqq \lambda_n(A)$ its ordered eigenvalues. Then by the above corollary the ordered 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 X 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 $n \to \infty$, written $X_n \xrightarrow{\mathscr{L}} X$ as $n \to \infty$, if $F_n(x) \to F(x)$ as $n \to \infty$ at all points of continuity of F.

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

$$P(|X_n - X| > \varepsilon) \to 0 \qquad \text{as } n \to \infty. \tag{2.3.1}$$

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

$$P(\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1. \tag{2.3.2}$$

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

2.4. *Empirical distribution function.* Let $\{X_1, X_2, \cdots, X_n\}$ be a set of random variables. For any $B \in \mathscr{F}$, let I_B denote the indicator function of B,

$$I_B(\omega) = \begin{cases} 1 & \text{if } \omega \in B, \\ 0 & \text{if } \omega \notin B. \end{cases} \tag{2.4.1}$$

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

$$F_n(x)(\omega) = \frac{1}{n} \sum_{i=1}^{n} I_{[X_i \in (-\infty, x)]}(\omega). \tag{2.4.2}$$

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

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

$$\begin{aligned} \alpha_k &= \int_{R_1} x^k \, dF(x), \\ \alpha_{k,n} &= \int_{R_1} x^k \, dF_n(x) \end{aligned} \tag{2.5.1}$$

define the kth moment of the distribution functions F and F_n, respectively, if they exist.

LEMMA 2.5.1. *If, for all $k \geqq k_0$ arbitrary but fixed, the sequence $\alpha_{k,n} \to \alpha_k$ finite, then these sequences converge for every value of k, and if the sequence $\{\alpha_k\}_{k=1}^{\infty}$ uniquely determines F, then $F_n(x) \to F(x)$ as $n \to \infty$ at all points of continuity of F.*

A proof of this lemma may be found in M. Loève [9].

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

$$\limsup_{n\to\infty} A_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n. \tag{2.5.2}$$

LEMMA 2.5.2. *The sequence* $X_n \to 0$ a.s. *as* $n \to \infty$ *if and only if for all* $\varepsilon > 0$

$$P\big(\limsup_{n\to\infty} \{\omega : |X_n(\omega)| > \varepsilon\}\big) = 0. \tag{2.5.3}$$

A proof of this lemma may be found in K. L. Chung [3].

LEMMA 2.5.3 (Borel-Cantelli). *If* $\Sigma_{n=1}^{\infty} P(A_n) < \infty$, *then* $P(\limsup_{n\to\infty} A_n) = 0$.

A proof of this may be found in Loève [9].

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

$$\begin{aligned} M_{k,n}(\omega) &= \int_R x^k \, dW_n(x)(\omega) \\ &= \frac{1}{n}\sum_{i=1}^{n} X_i^k(\omega). \end{aligned} \tag{2.5.4}$$

LEMMA 2.5.4. *If* $M_{k,n} \xrightarrow{P} \alpha_k$ *as* $n \to \infty$ *for all* $k = 1, 2, \cdots$, *then* $W_n(x) \xrightarrow{P} W(x)$ *as* $n \to \infty$, *at all points of continuity of* W.

PROOF. The following result is used to establish the lemma: $X_n \xrightarrow{P} X$ as $n \to \infty$ 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

$$\int_R x^k \, dW_{n_i}(x) \xrightarrow{P} \int_R x^k \, dW(x) \tag{2.5.5}$$

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

$$\int_R x^k \, dW_{n_i'}(x) \to \int_R x^k \, dW(x) \qquad \text{a.s.} \tag{2.5.6}$$

for all $k = 1, 2, \cdots$. Then by Lemma 2.5.1,

$$W_{n_i'}(x) \to W(x) \qquad \text{a.s.} \tag{2.5.7}$$

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

3. 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 E. P. Wigner [17].

Denote by $A_{k,n}$, $k > 1$, $n \geqq 1$, the class of all finite sequences $f: \{1, 2, \cdots, k+1\} \to \{1, 2, \cdots, n\}$. 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 $g_f : \{1, 2, \cdots, k\} \to \{\text{all steps}\}$ defined as follows: $g_f(\nu) = \big(f(\nu), f(\nu+1)\big)$, $1 \leqq \nu \leqq k$. 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

$$D_f = \{f(i), 2 \leqq i \leqq k + 1 : f(i) \notin \{f(1), \cdots, f(i-1)\}\}, \tag{3.1}$$

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

$$\#\{g_f(\nu) = (f(\nu), f(\nu+1)), 1 \leqq \nu \leqq k : f(\nu) = i, f(\nu+1) = j\}. \tag{3.2}$$

For $1 \leqq \nu \leqq k$, $g_f(\nu) = (f(\nu), f(\nu+1))$ is called a *free step* if and only if $f(\nu+1) \notin \{f(1), \cdots, f(\nu)\}$ and a repetitive step if and only if $f(\nu+1) \in \{f(1), \cdots, f(\nu)\}$. Let

$$\begin{aligned} F_f &= \{g_f(\nu), 1 \leqq \nu \leqq k : g_f(\nu) \text{ is free}\}, \\ R_f &= \{g_f(\nu), 1 \leqq \nu \leqq k : g_f(\nu) \text{ is repetitive}\}. \end{aligned} \tag{3.3}$$

It is immediate that

$$\begin{aligned} \# F_f + \# R_f &= k, \\ \# F_f &= \#\{g_f(\nu), 1 \leqq \nu \leqq k : g_f(\nu) \text{ is free}\} \\ &= \#\{f(\nu+1), 1 \leqq \nu \leqq k : f(\nu+1) \notin \{f(1), \cdots, f(\nu)\}\} \\ &= \#\{F(\nu), 2 \leqq \nu \leqq k+1 : f(\nu) \notin \{f(1), \cdots, f(\nu-1)\}\}, \end{aligned} \tag{3.4}$$

and

$$\# F_f = \# D_f = d_f - 1. \tag{3.5}$$

LEMMA 3.1. *Let $f \in A_{k,n}$ be such that if $(i,j) \in \{g_f(1), g_f(2), \cdots, g_f(k)\}$, then $\#(i,j)_f + \#(j,i)_f \geqq 2$. Then $d_f \leqq [\frac{1}{2}k] + 1$.*

PROOF. Let $f \in A_{k,n}$ satisfy the conditions of the lemma. If $f(i) \in \{f(\nu), 2 \leqq \nu \leqq k + 1 : f(\nu) \notin \{f(1), \cdots, f(\nu - 1)\}\}$, then $(f(i-1), f(i))$ is a free step. The condition of the lemma implies at least one step among $g_f(i), \cdots, g_f(k)$ must equal $(f(i-1), f(i))$ or $(f(i), f(i-1))$ (no step among $g_f(1), \cdots, g_f(i-2)$ equals $(f(i-1), f(i))$ or $(f(i), f(i-1))$ since $f(i) \notin \{f(i), \cdots, f(i-1)\}$). Any such occurrence, say $(f(\ell - 1), f(\ell))$, must be repetitive since $f(\ell) \in \{f(1), \cdots, f(\ell - 1)\}$. Hence, with each free step is associated a repetitive step which is equal to the free step or its reverse. This implies $\# F_f \leqq \# R_f$, since all free steps are different. This, with (3.4), implies $\# F_f \leqq [\frac{1}{2}k]$. Hence, by (3.5), $d_f - 1 = \# F_f \leqq [\frac{1}{2}k]$ or $d_f \leqq [\frac{1}{2}k] + 1$. This completes the proof of Lemma 3.1.

LEMMA 3.2. *Let $f \in A_{k,n}$ be such that:*

(i) *if $(i,j) \in \{g_f(1), g_f(2), \cdots, g_f(k)\}$, then $\#(i,j)_f + \#(j,i)_f \geqq 2$;*

(ii) *$f(\ell) = f(\ell + 1)$ for some ℓ, $1 \leqq \ell \leqq k$.*

Then $d_f \leqq [\frac{1}{2}k]$.

PROOF. If f is constant one is through. Assume f is not constant. If $f(\ell) = f(\ell + 1)$ for some ℓ, $1 \leqq \ell \leqq k$, a new sequence of steps may be formed from $g_f(1), g_f(2), \cdots, g_f(k)$ 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 $h: \{1, 2, \cdots, i\} \to \{1, 2, \cdots, n\}$, $2 \leqq i \leqq k - 1$, (a lower bound of 2 since f is not constant) which satisfies condition (i) and which is such that $d_h = d_f$. Lemma 3.1 then gives $d_f = d_h \leqq [\frac{1}{2}(i - 1)] + 1 \leqq [\frac{1}{2}(k - 2)] + 1 = [\frac{1}{2}k]$. This completes the proof of Lemma 3.2.

LEMMA 3.3. *Let k be even, say $k = 2v$. Let $f \in A_{2v,n}$ be such that:*

(i) *if $(i, j) \in \{g_f(1), g_f(2), \cdots, g_f(2v)\}$, then $\#(i, j)_f + \#(j, i)_f \geqq 2$;*

(ii) $f(1) = f(2v + 1)$;

(iii) $d_f = v + 1$.

If $(i, j) \in \{g_f(1), g_f(2), \cdots, g_f(2v)\}$, 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 $n \geqq v + 1$ such an f is easily constructed. For example, let $f(1) = 1$, $f(2) = 2, f(v) = v, f(v + 1) = v + 1, f(v + 2) = v, \cdots, f(2v) = 2, f(2v + 1) = 1$.) Lemma 3.2 shows $d_f \leqq v + 1$. By Lemma 3.3 one must have $f(v) \neq f(v + 1)$, $1 \leqq \ell \leqq 2v$. Equation (3.5) holds:

$$\#F_f = \#D_f = d_f - 1 = v. \tag{3.6}$$

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

Now define a sequence $h: \{1, 2, \cdots, 2v + 1\} \to \{1, 2, \cdots, n\}$ as follows:

$$\begin{aligned} &h(1) = f(2), \quad h(2) = f(3), \cdots, \quad h(i) = f(i + 1), \cdots, \\ &h(2v) = f(2v + 1) = f(1), \quad h(2v + 1) = f(2). \end{aligned} \tag{3.7}$$

Associated with h is the sequence of steps $g_h(1) = (f(2), f(3))$, $g_h(2) = (f(3), f(4)), \cdots, g_h(2v - 1) = (f(2v), f(1)), g_h(2v) = (f(1), f(2))$. 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 $g_h(2) = (f(3), f(4)), \cdots, g_h(2\nu) = (f(1), f(2))$. Continuing in the same manner, one concludes if $(i, j) \in \{g_f(1), g_f(2), \cdots, g_f(2\nu)\}$, then $(j, i) \in \{g_f(1), g_f(2), \cdots, g_f(2\nu)\}$. Since there are ν different free steps and 2ν steps altogether, one must have $\#(i, j)_f = 1$, $\#(j, i)_f = 1$ for each $(i, j) \in \{g_f(1), g_f(2), \cdots, g_f(2\nu)\}$. This completes the proof of Lemma 3.3.

LEMMA 3.4. *Let k be odd, say $k = 2\nu + 1$. Let $f \in A_{2\nu+1,n}$ be such that:*

(i) *if $(i, j) \in \{g_f(1), g_f(2), \cdots, g_f(2\nu + 1)\}$, then $\#(i, j)_f + \#(j, i)_f \geqq 2$;*

(ii) $f(1) = f(2\nu + 2)$.

Then $d_f \leqq \nu = [\frac{1}{2}k]$.

PROOF. Let $f \in A_{2\nu+1,n}$ satisfy conditions (i) and (ii) of the lemma. Lemma 3.1 shows $d_f \leqq \nu + 1$. Assume $d_f = \nu + 1$. Then by Lemma 3.2, $f(\ell) \neq f(\ell + 1)$, $1 \leqq \ell \leqq 2\nu + 1$. There are $\#F_f = \#D_f = d_f - 1 = \nu$ different free steps. For each free step there is a repetitive step equal to the free step itself or its reverse. This occupies 2ν of the $2\nu + 1$ 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 $(i, j) \in \{g_f(1), \cdots, g_f(2\nu + 1)\}$ except one, say (k, ℓ), for which $\#(k, \ell)_f + \#(\ell, k)_f = 3$.

All possibilities are now considered. First consider the case $\#(k, \ell)_f = 3$. (The case $\#(\ell, k)_f = 3$ is the same.) With f is associated a sequence of steps $g_f(1), \cdots, g_f(r), \cdots, g_f(s), \cdots, g_f(t), \cdots, g_f(2\nu + 1)$, where $g_f(r) = g_f(s) = g_f(t) = (k, \ell)$ and $s - r \geqq 2$, $t - s \geqq 2$. Let $g_f^*(i)$ denote the reverse of $g_f(i)$. From the sequence of steps $g_f(1), g_f(2), \cdots, g_f(2\nu + 1)$ form a sequence of steps associated with a sequence $h : \{1, 2, \cdots, 2\nu - 1\} \to \{1, 2, \cdots, n\}$ in the following manner:

$$
\begin{aligned}
g_h(1) &= g_f(t + 1) = (\ell, f(t + 2)) = (h(1), h(2)) \\
g_h(2) &= g_f(t + 2) = (f(t + 2), f(t + 3)) = (h(2), h(3)) \\
&\ \vdots \\
g_h((2\nu + 1) - t) &= g_f(2\nu + 1) = (f(2\nu + 1), f(2\nu + 2)) \\
&= (h((2\nu + 1) - t), h((2\nu + 1) - (t - 1))) \\
g_h((2\nu + 1) - (t - 1)) &= g_f(1) = (f(1), f(2)) \\
&= (h((2\nu + 1) - (t - 1)), h((2\nu + 1) - (t - 2))) \\
&\ \vdots \\
g_h((2\nu + 1) - (t - r + 1)) &= g_f(r - 1) = (f(r - 1), k) \\
&= (h((2\nu + 1) - (t - r + 1)), h((2\nu + 1) - (t - r))) \\
g_h((2\nu + 1) - (t - r)) &= g_f^*(s - 1) = (k, f(s - 1)) \\
&= (h((2\nu + 1) - (t - r)), h((2\nu + 1) - (t - r - 1))) \\
g_h((2\nu + 1) - (t - r - 1)) &= g_f^*(s - 2) = (f(s - 1), f(s - 2)) \\
&= (h((2\nu + 1) - (t - r - 1)), h((2\nu + 1) - (t - r - 2)))
\end{aligned} \tag{3.8}
$$

$$\vdots$$

$$\begin{aligned}
g_h((2v+1)-(t-s+2)) &= g_f^*(r+1) = (f(r+2), \ell) \\
&= (h((2v+1)-(t-s+2)), h((2v+1)-(t-s+1))) \\
g_h((2v+1)-(t-s+1)) &= g_f(s+1) = (\ell, f(s+2)) \\
&= (h((2v+1)-(t-s+1)), h((2v+1)-(t-s)))
\end{aligned}$$

$$\vdots$$

$$\begin{aligned}
g_h((2v+1)-4) &= g_f(t-2) = (f(t-2), f(t-1)) \\
&= (h((2v+1)-4), h((2v+1)-3)) \\
g_h((2v+1)-3) &= g_f(t-1) = (f(t-1), k) \\
&= (h((2v+1)-3), h((2v+1)-2)).
\end{aligned}$$

The steps associated with h are $g_f(1), \cdots, g_f(r-1), g_f^*(r+1), \cdots, g_f^*(s-1), g_f(s+1), \cdots, g_f(t-1), g_f(t+1), \cdots, g_f(2v+1)$, 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 $(i,j) \in \{g_h(1), \cdots, g_h(2v-2)\}$, then $\#(i,j)_h + \#(j,i)_h = 2$, and $d_h = d_f = v+1$. But Lemma 3.1 shows $d_h \leqq v$. One must conclude, by contradiction, that $d_f \leqq v$.

The other possibility is $\#(k,\ell)_f = 2$, $\#(\ell,k)_f = 1$ (or, what is the same thing, $\#(k,\ell)_f = 1$, $\#(\ell,k)_f = 2$) for which an argument similar to the above may be given. The details are not given here. This completes the proof of Lemma 3.4.

WIGNER'S COMBINATORIAL THEOREM. *Let $B_{2v,n}$ be the set of all $f \in A_{2v,n}$ such that:*

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

Then

$$\#B_{2v,n} = \frac{(2v)!}{v!(v+1)!} n^{v+1} + o(n^{v+1}). \tag{3.9}$$

PROOF. Let $f \in B_{2v,n}$. By Lemma 3.2, $f(\ell) \neq f(\ell+1)$, $1 \leqq \ell \leqq 2v$. By Lemma 3.3, if $(i,j) \in \{g_f(1), \cdots, g_f(2v)\}$, then $\#(i,j)_f = 1$, $\#(j,i)_f = 1$. A sequence $t: \{1, 2, \cdots, m\} \to \{\text{integers}\}$ is called a *type sequence* if and only if $t(\ell) \geqq 0$, $1 \leqq \ell \leqq m$, $t(1) = 1$, $t(m) = 0$, and $t(\ell+1) - t(\ell) = \pm 1$, $1 \leqq \ell \leqq m-1$. For each $f \in B_{2v,n}$ define the type sequence $t_f: \{1, 2, \cdots, 2v\} \to \{\text{integers}\}$ as follows:

$$t_f(\ell) = \#\{g_f(i), 1 < i < \ell : g_f(i) \text{ is free}\} - \#\{g_f(i), 1 < i < \ell : g_f(i) \text{ is repetitive}\}. \tag{3.10}$$

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

$$\#\{f \in B_{2v,n} : t_f(\ell) = t(\ell), 1 \leqq \ell \leqq 2v\} = n(n-1) \cdots (n-v). \tag{3.11}$$

This is so because: (1) there are n choices for $f(1)$; (2) for $2 \leqq i \leqq 2\nu$, 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 $2 \leqq i \leqq 2\nu$, 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 3.3), and hence $f(i)$ is completely determined. Let S_ν denote the number of type sequences with domain $\{1, 2, \cdots, 2\nu\}$. Then

$$\#B_{2\nu,n} = S_\nu n(n-1)\cdots(n-\nu) = S_\nu n^{\nu+1} + o(n^{\nu+1}). \tag{3.12}$$

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

$$S'_\nu = S_{\nu-1}, \qquad S'_1 = S_0 = 1. \tag{3.13}$$

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

$$S_\nu = \sum_{k=1}^{\nu} S'_k S_{\nu-k} = \sum_{k=1}^{\nu} S_{k-1} S_{\nu-k}, \qquad \nu = 1, 2, \cdots. \tag{3.14}$$

These recursive equations permit the successive calculation of the S_ν. Formally, one can obtain a closed formula for them by writing $t(x) = \Sigma_{\nu=0}^{\infty} t_\nu x^\nu$. The recursive formula (3.14) then gives

$$t(x) = 1 + xt^2(x). \tag{3.15}$$

The 1 on the right side is necessary because (3.14) is not valid for $\nu = 0$. It follows that

$$t(x) = \frac{1 \pm (4 - x)^{1/2}}{2x}. \tag{3.16}$$

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

$$S_\nu = \frac{1}{2}\binom{\frac{1}{2}}{\nu + 1}(-4)^{\nu+1} = \frac{(2\nu)!}{\nu!(\nu+1)!}, \qquad \nu = 1, 2, \cdots. \tag{3.17}$$

And finally,

$$\#B_{2\nu,n} = \frac{(2\nu)!}{\nu!(\nu+1)!} n^{\nu+1} + o(n^{\nu+1}). \tag{3.18}$$

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) $f(i) \neq f(i+1)$, $1 \leqq i \leqq 2k$;

(iii) if $(i, j) \in \{g_f(1), g_f(2), \cdots, g_f(2\nu)\}$, then $\#(i, j)_f + \#(j, i)_f$ is even.

Let $C^j_{2k,n}$ denote the set of all $f \in C_{2k,n}$ such that $d_f = j$. By Lemma 3.2, if $f \in C_{2k,n}$, then $d_f \leqq k + 1$. Thus, $C_{2k,n} = \bigcup_{j=1}^{k+1} C^j_{2k,n}$ and

$$\# C_{2k,n} = \sum_{j=1}^{k+1} \# C^j_{2k,n}. \tag{3.19}$$

LEMMA 3.5. *The sets just defined satisfy the relation* $\# C^j_{2k,n} = \binom{n}{j}(\# C^j_{2k,j})$.

PROOF. The relation $\sim$ determined by $f \sim f^*$ if and only if $f \in C^j_{2k,n}$, $f^* \in C^j_{2k,n}$ and $\{f(1), f(2), \cdots, f(2k+1)\} = \{f^*(1), f^*(2), \cdots, f^*(2k+1)\}$ is an equivalence relation. The set $C^j_{2k,n}$ is split into $\binom{n}{j}$ equivalence classes by $\sim$, each containing $\# C^j_{2k,j}$ members. Hence, $\# C^j_{2k,n} = \binom{n}{j}(\# C^j_{2k,j})$. This completes the proof of the lemma.

Using Lemma 3.5, one has

$$\# C_{2k,n} = \sum_{j=1}^{k+1} \binom{n}{j} (\# C^j_{2k,j}). \tag{3.20}$$

Thus, $\# C_{2k,n}$ is determined for all n by $\# C^1_{2k,1}$, $\# C^2_{2k,2}$, $\cdots$, $\# C^{k+1}_{2k,k+1}$. 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 I were made on a computer. It will be pointed out in the next section in what context these numbers may be of interest.

TABLE I

ENUMERATIONS

The numbers in the body of the table are $(1/n)(\# C^j_{3k,n}) = (1/n)(\# C^j_{2k,j})$.

k, n \ 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	768,00	214,080	
6, 7	6	20,460	699,600	2,304,000	1,284,480	95,040

4. Random sign ensemble and Wigner's conjecture

4.1. *Wigner's* 1955 *paper.* In 1955, Wigner [16] 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) $\{a_{ij}, i \leqq j\}$ 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

$$B_n = \frac{1}{2\sigma\sqrt{n}} A_n. \tag{4.1.1}$$

Denote by $\lambda_1(B_n) \leqq \lambda_2(B_n) \leqq \cdots \leqq \lambda_n(B_n)$ the ordered random eigenvalues of B_n and by $W_n(x)$ the empirical distribution function of $\{\lambda_1(B_n), \lambda_2(B_n), \cdots, \lambda_n(B_n)\}$, that is,

$$W_n(x)(\omega) = \frac{1}{n} \sum_{i=1}^n I_{[\lambda_i(B_n) \in (-\infty, x)]}(\omega). \tag{4.1.2}$$

Then one has the following theorem.

THEOREM 4.1.1 (Wigner [16]). *For all* $x \in R_1$, $\lim_{n\to\infty} E(W_n(x)) = W(x)$, *where* W *is the absolutely continuous distribution function with semicircle density*

$$w(x) = \begin{cases} \frac{2}{\Pi}(1 - x^2)^{1/2}, & |x| \leqq 1, \\ 0, & |x| > 1. \end{cases} \tag{4.1.3}$$

PROOF. The distribution function W is uniquely determined by its moment sequence since

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{\gamma_k}{k!}(it)^k &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(k+1)!}\left(\frac{t}{2}\right)^{2k} = \frac{2}{t} J_1(t) \\ &= \frac{2}{\Pi}\int_{-1}^{1} e^{itx}(1 - x^2)^{1/2}\, dx, \end{aligned} \tag{4.1.4}$$

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

$$\gamma_k = \int x^k\, dW(x) = \begin{cases} 0 & \text{for } k \text{ odd}, \\ \dfrac{k!}{2^k(\frac{1}{2}k)!(\frac{1}{2}k + 1)!} & \text{for } k \text{ even}. \end{cases} \tag{4.1.5}$$

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

$$\int x^k\, dEW_n(x) \to \gamma_k \text{ as } n \to \infty \tag{4.1.6}$$

for all $k = 1, 2, \cdots$, then Lemma 2.5.1 will yield the desired result.

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

$$P\left(D_{(i_{12}, \cdots, i_{n-1,n})}\right) = \frac{1}{h}, \tag{4.1.7}$$

for all points in T, where $h = 2^{(n-1)n/2}$ One has $W_n(x) = N_n(x)/n$, where $N_n(x)$ equals the number of eigenvalues of B_n less than x. On each D_i, $i \in T$, $W_n(x)$ is constant; denote these values by $W_n(x)_i = N_n(x)_i/n$, $i \in T$. Then, since $\Omega = \left(\bigcup_{i \in T} D_i\right) \cup N$, where $\left(\bigcup_{i \in T} D_i\right) \cap N = \varphi$ and $P(N) = 0$,

$$\begin{aligned} E W_n(x) &= \int_{\bigcup_{i \in T} D_i} W_n(x)\, dP \\ &= \sum_{i \in T} \int_{D_i} W_n(x)\, dP \\ &= \sum_{i \in T} \frac{N_n(x)_i}{nh}. \end{aligned} \tag{4.1.8}$$

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

$$\begin{aligned} \int x^k\, dE W_n(x) &= \frac{1}{nh} \sum_{j=1}^{nh} e_j^k \\ &= \frac{1}{nh} \sum_{i \in T} \operatorname{tr}\left(B_n^k(i)\right) \\ &= \frac{1}{h(2\sigma)^k n^{1+k/2}} \sum_{i \in T} \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} \prod_{\ell=1}^{k} a_{j_\ell j_{\ell+1}}(i), \end{aligned} \tag{4.1.9}$$

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: $\{1, 2, \cdots, k+1\} \to \{1, 2, \cdots, n\}$, one has

$$\int x^k\, dE W_n(x) = \frac{1}{(2\sigma)^k n^{1+k/2}} \sum_{f \in B_{k,n}} \frac{1}{h} \sum_{i \in T} \prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)}(i) \tag{4.1.10}$$

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

$$(4.1.11)\qquad \int x^k\, dEW_n(x) = \frac{1}{(2\sigma)^k n^{1+k/2}} \sum_{f\in B_{k,n}} \sum_{i\in T} \int_{D_i} \prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)}\, dP$$
$$= \frac{1}{(2\sigma)^k n^{1+k/2}} \sum_{f\in B_{k,n}} E \prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)}.$$

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

$$(4.1.12)\qquad \int x^k\, dEW_n(x) = \frac{1}{(2\sigma)^k n^{1+k/2}} \sum_{f\in C_{k,n}} E \prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)},$$

where

$$(4.1.13)\qquad C_{k,n} = \{f\in B_{k,n} : f(\ell) \neq f(\ell+1), \ell = 1, \cdots, k\}.$$

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 $\prod_{\ell=1}^{2v+1} a_{f(\ell)f(\ell+1)} = a^m_{f(i)f(i+1)} \prod a_{jk}$ 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,

$$(4.1.14)\qquad E \prod_{\ell=1}^{2v+1} a_{f(\ell)f(\ell+1)} = E\, a^m_{f(i)f(i+1)}\, E \prod a_{jk} = 0.$$

Thus,

$$(4.1.15)\qquad \int_R x^{2v+1}\, dEW_n(x) = 0 = \gamma_{2v+1}.$$

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

$$(4.1.16)\qquad \int_R x^{2v}\, dEW_n(x) = \frac{1}{(2\sigma)^{2v} n^{v+1}} \sum_{f\in D_{2v,n}} E \prod_{\ell=1}^{2v} a_{f(\ell)f(\ell+1)}.$$

where $D_{2v,n} = \{f\in C_{2v,n} :$ if $(i, j)\in \{g_f(1), g_f(2), \cdots, g_f(2v)\}$, then $\#(i, j)_f + \#(j, i)_f$ is even$\}$. For $f\in D_{2v,n}$ one has, by Lemma 3.3

$$(4.1.17)\qquad E \prod_{\ell=1}^{2v} a_{f(\ell)f(\ell+1)} = \sigma^{2v}.$$

Thus,

$$(4.1.18)\qquad \int_R x^{2v}\, dEW_n(x) = \frac{\# D_{2v,n}}{2^{2v} n^{v+1}}.$$

By Lemma 3.1, $f\in D_{2v,n}$ is such that $d_f \leqq v + 1$. Thus,

$$(4.1.19)\qquad \# D_{2v,n} = \sum_{j=1}^{v+1} \# D^j_{2v,n},$$

where $D^j_{2\nu,n} = \{f \in D_{2\nu,n} : d_f = j\}$. As with Lemma 3.5, one has

$$\#D^j_{2\nu,n} = \binom{n}{j}(\#D^j_{2\nu,j}). \tag{4.1.20}$$

Let

$$f_1(\nu, n) = \frac{\#D^{\nu+1}_{2\nu,n}}{2^{2\nu}n^{\nu+1}}, \tag{4.1.21}$$

$$f_2(\nu, n) = \frac{\sum_{j=1}^{\nu}\binom{n}{j}(\#D^j_{2\nu,j})}{2^{2\nu}n^{\nu+1}}. \tag{4.1.22}$$

Then

$$\int_R x^{2\nu}\, dEW_n(x) = f_1(\nu, n) + f_2(\nu, n). \tag{4.1.23}$$

Since $\binom{n}{j}/n^{\nu+1} \to 0$ as $n \to \infty$ for $j = 1, \cdots, \nu$, one has $f_2(\nu, n) \to 0$ as $n \to \infty$. By Wigner's combinatorial theorem

$$\#D^{\nu+1}_{2\nu,n} = \frac{(2\nu)!}{\nu!(\nu+1)!}n^{\nu+1} + o(n^{\nu+1}). \tag{4.1.24}$$

Thus,

$$f_1(\nu, n) = \frac{(2\nu)!}{2^{2\nu}\nu!(\nu+1)!} + \frac{o(n^{\nu+1})}{2^{2\nu}n^{\nu+1}} \tag{4.1.25}$$

and

$$f_1(\nu, n) \to \frac{(2\nu)!}{2^{2\nu}\nu!(\nu+1)!} \tag{4.1.26}$$

as $n \to \infty$. Thus,

$$\int_R x^{2\nu}\, dEW_n(x) \to \frac{(2\nu)!}{2^{2\nu}\nu!(\nu+1)!} = \gamma_{2\nu} \tag{4.1.27}$$

as $n \to \infty$. This completes the proof of Theorem 4.1.1.

Note the equation,

$$\int_R x^{2\nu}\, dEW_n(x) = \frac{\#D_{2\nu,n}}{2^{2\nu}n^{\nu+1}}, \tag{4.1.28}$$

derived during the course of the proof. A knowledge of $\#D_{2\nu,n}$ would give the sequence of moments of the distribution function $EW_n(x)$. As mentioned at the end of Section 3, these numbers are determined for all n by a knowledge of only $\#D^1_{2\nu,1}, \#D^2_{2\nu,2}, \cdots, \#D^{\nu+1}_{2\nu,\nu+1}$.

4.2. *Wigner's* 1958 *conjecture.* In 1958 Wigner [18] 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) $\{a_{ij}, i \leqq j\}$ 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 $1 \leqq i, j \leqq n$;

(vi) $E\, |a_{ij}|^k < C_k$ for all $1 \leqq i, j \leqq n$, where C_k is independent of n.

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

WIGNER'S CONJECTURE. *For all* $x \in R_1$, $EW_n(x) \to W(x)$ *as* $n \to \infty$, *where* $W(x)$ *is the absolutely continuous distribution function with semicircle density*

$$w(x) = \begin{cases} \frac{2}{\pi}(1 - x^2)^{1/2}, & |x| \leqq 1, \\ 0, & |x| > 1. \end{cases} \tag{4.2.1}$$

In the next section, we shall discuss work of U. Grenander [7] who sketched a proof of convergence in probability of the empirical distribution functions to the semicircle law. We shall also discuss the work of L. Arnold [2] in this connection.

5. The results of Grenander and Arnold

5.1. *Convergence in probability.* Grenander [7] 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) $\{a_{ij}, i \leqq j\}$ is independent;

(iii) a_{ij} is symmetric;

(iv) $E\, a_{ij}^2 = \sigma^2$;

(v) $|E\, a_{ij}^k| \leqq C_k$, $k = 1, 2, \cdots$, where C_k is independent of n.

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

THEOREM 5.1.1 (Grenander [7]). *For all* $x \in R_1$, $W_n(x) \xrightarrow{P} W(x)$ *as* $n \to \infty$, *where* W *is the absolutely continuous distribution function with semicircle density*

$$w(x) = \begin{cases} \frac{2}{\pi}(1 - x^2)^{1/2}, & |x| \leqq 1, \\ 0, & |x| > 1. \end{cases} \tag{5.1.1}$$

PROOF. Let

$$M_{k,n} = \int_{R_1} \lambda^k \, dW_n(x). \tag{5.1.2}$$

By Lemma 2.5.4, it will be sufficient to prove $M_{k,n} \to \gamma_k = \int x^k \, dW(x)$ as $n \to \infty$ for all $k = 1, 2, \cdots$. This will be achieved in two steps. First it will be shown that $EM_{k,n} \to \gamma_k$ as $n \to \infty$ as for all $k = 1, 2, \cdots$. Then it will be shown that $E(M_{k,n} - EM_{k,n})^2 \to 0$ as $n \to \infty$ for all $k = 1, 2, \cdots$. Chebyshev's inequality,

$$P(|M_{k,n} - EM_{k,n}| > \varepsilon) \leqq \frac{E(M_{k,n} - EM_{k,n})^2}{\varepsilon^2}, \tag{5.1.3}$$

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

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

$$\begin{aligned} EM_{k,n} &= E \int \lambda^k \, dW_n(x) = E \frac{1}{n} \sum_{i=1}^{n} \lambda_i(B_n) \\ &= E \frac{1}{n} \operatorname{tr} B_n^k \\ &= E \frac{1}{(2\sigma)^k n^{1+k/2}} \sum_{f \in B_{k,n}} \prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)} \\ &= \frac{1}{(2\sigma)^k n^{1+k/2}} \sum_{f \in B_{k,n}} E \prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)}, \end{aligned} \tag{5.1.4}$$

where $B_{k,n} = \{f \in A_{k,n} : f(1) = f(k+1)\}$. As in Theorem 4.1.1, two cases are considered, k odd and k even. For exactly the reasons given in the proof of Theorem 4.1.1 one concludes immediately that for $k = 2\nu + 1$

$$EM_{2\nu+1,n} = 0 = \gamma_{2\nu+1}. \tag{5.1.5}$$

Now let $k = 2\nu$. If $f \in B_{2\nu,n}$ 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,

$$E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)} = Ea_{f(i)f(i+1)} E \prod_{\ell \neq i, \ell=1}^{2\nu} a_{f(\ell)f(\ell+1)} = 0. \tag{5.1.6}$$

Thus,

$$EM_{2\nu,n} = \frac{1}{(2\sigma)^{2\nu} n^{\nu+1}} \sum_{f \in C_{2\nu,n}} E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)}, \tag{5.1.7}$$

where $C_{2\nu,n} = \{f \in B_{2\nu,n} : \text{if } (i,j) \in \{g_f(1), g_f(2), \cdots, g_f(2\nu)\}, \text{ then } \#(i,j)_f + \#(j,i)_f \geqq 2\}$. By Lemma 3.1, $d_f \leqq \nu + 1$ for all $f \in C_{2\nu,n}$. Let

$$\begin{aligned} f_1(\nu, n) &= \frac{1}{(2\sigma)^{2\nu} n^{\nu+1}} \sum_{f \in C_{2\nu,n}^{\nu+1}} E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)}, \\ f_2(\nu, n) &= \frac{1}{(2\sigma)^{2\nu} n^{\nu+1}} \sum_{f \in \bigcup_{j=1}^{\nu} C_{2\nu,n}^j} E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)}, \end{aligned} \tag{5.1.8}$$

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

$$EM_{2\nu,n} = f_1(\nu, n) + f_2(\nu, n). \tag{5.1.9}$$

By assumption (v), one has

$$|f_2(\nu, n)| \leqq \frac{D_{2\nu}}{(2\sigma)^{2\nu} n^{\nu+1}} \sum_{j=1}^{\nu} \# C^j_{2\nu,n} \tag{5.1.10}$$

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

$$|f_2(\nu, n)| \leqq \frac{D_{2\nu} \sum_{j=1}^{\nu} \binom{n}{j} (\# C^j_{2\nu,j})}{(2\sigma)^{2\nu} n^{\nu+1}}. \tag{5.1.11}$$

Since $\binom{n}{j}/n^{\nu+1} \to 0$ as $n \to \infty$ for $j = 1, \cdots, \nu$, one has $f_2(\nu, n) \to 0$ as $n \to \infty$. Let $f \in C^{\nu+1}_{2\nu,n}$. Then, using Lemma 3.3, one has

$$E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)} = \sigma^{2\nu}. \tag{5.1.12}$$

Hence,

$$f_1(\nu, n) = \frac{\# C^{\nu+1}_{2\nu,n}}{2^{2\nu} n^{\nu+1}}. \tag{5.1.13}$$

By Wigner's combinatorial theorem,

$$f_1(\nu, n) = \frac{(2\nu)!}{2^{2\nu} \nu! (\nu + 1)!} + \frac{o(n^{\nu+1})}{2^{2\nu} n^{\nu+1}}. \tag{5.1.14}$$

Hence,

$$f_1(\nu, n) \to \frac{(2\nu)!}{2^{2\nu} \nu! (\nu + 1)!} = \gamma_{2\nu} \tag{5.1.15}$$

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

It will now be shown that $E(M_{k,n} - EM_{k,n})^2 \to 0$ as $n \to \infty$ for all $k = 1, 2, \cdots$. For $f \in A_{2k+1,n}$, let

$$E(f) = E \prod_{i=1}^{k} a_{f(i)f(i+1)} \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)} - E \prod_{i=1}^{k} a_{f(i)f(i+1)} E \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)}. \tag{5.1.16}$$

One has after some manipulation

$$E(M_{k,n} - EM_{k,n})^2 = \frac{1}{(2\sigma)^{2k} n^{k+2}} \sum_{f \in B_{2k+1,n}} [E(f)], \tag{5.1.17}$$

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 $(i, j) \in \{g_f(1), \cdots, g_f(k), g_f(k+2), \cdots, g_f(2k+1)\}$, then $\#(i, j)_f + \#(j, i)_f \geqq 2\}$. Condition (iii) follows from the independence and symmetry conditions. It allows one to conclude, by arguments exactly as those of the proof of Lemma 3.1, that $d_f \leqq k + 1$ for all $f \in B_{2k+1,n}$. By assumption (v),

$$|E(f)| \leqq \delta_k < \infty, \tag{5.1.18}$$

where δ_k is independent of n. Thus,

$$E(M_{k,n} - EM_{k,n})^2 \leqq \frac{\delta_k(\#B_{2k+1,n})}{(2\sigma)^{2k} n^{k+2}}. \tag{5.1.19}$$

Now

$$\begin{aligned} \#B_{2k+1,n} &= \sum_{j=1}^{k+1} \#B^j_{2k+1,n} \\ &= \sum_{j=1}^{k+1} \binom{n}{j} (\#B_{2k+1,j}), \end{aligned} \tag{5.1.20}$$

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

$$E(M_{k,n} - EM_{k,n})^2 \leqq \frac{\delta_k \sum_{j=1}^{k+1} \binom{n}{j} (\#B^j_{2k+1,n})}{(2\sigma)^{2k} n^{k+2}}. \tag{5.1.21}$$

Since $\binom{n}{j}/n^{k+2} \to 0$ as $n \to \infty$ for $j = 1, 2, \cdots, k+1$, one has

$$E(M_{k,n} - EM_{k,n})^2 \to 0 \tag{5.1.22}$$

as $n \to \infty$. This completes the proof of Theorem 5.1.1.

5.2. *Convergence almost surely.* Arnold [2] sketches a proof leading to the result given in this section.

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

(i) $a_{ij} = a_{ji}$ a.s.;

(ii) $\{a_{ij}, i \leqq j\}$ 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) $Ea_{ij} = \int x\, dF = 0$, $i \neq j$;

(v) $Ea^2_{ij} = \int x^2\, dF = \sigma^2$, $i \neq j$;

(vi) (a) $Ea^2_{ii} = \int x^2\, dG < \infty$, $Ea^4_{ij} = \int x^4\, dF < \infty$;

(b) $Ea^4_{ii} = \int x^4\, dG < \infty$, $Ea^6_{ij} = \int x^6\, dF < \infty$.

Let $B_n = (2\sigma\sqrt{n})^{-1} A_n$ and denote by $W_n(x)$ the empirical distribution function of $\{\lambda_1(B_n), \lambda_2(B_n), \cdots, \lambda_n(B_n)\}$, where $\lambda_1(B_n) \leqq \cdots \leqq \lambda_n(B_n)$ are the ordered random eigenvalues of B_n. Arnold [2] then gives the following theorem.

THEOREM 5.2.1. *Under conditions* (i) *to* (v) *and* (vi) (a), $W_n(x) \xrightarrow{P} W(x)$ *as* $n \to \infty$ *for all* $x \in R$, *and under conditions* (i) *to* (v) *and* (vi) (*b*), $W_n(x) \to W(x)$ *a.s. as* $n \to \infty$ *for all* $x \in R_1$, *where* W *is an absolutely continuous distribution* *function with semicircle density*

$$w(x) = \begin{cases} \frac{2}{\pi}(1 - x^2)^{1/2}, & |x| \leqq 1, \\ 0, & |x| > 1. \end{cases} \tag{5.2.1}$$

6. On Wigner's 1958 conjecture

6.1. *A limit 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) $\{a_{ij}, i \leqq j\}$ is independent;
(iii) $Ea_{ij} = 0$, $1 \leqq i, j \leqq n$;
(iv) $Ea_{ij}^2 = \sigma^2$, $1 \leqq i \neq j \leqq n$;
(v) $E|a_{ij}|^k \leqq M_k$, $1 \leqq i, j \leqq n$, where M_k is not dependent on n.

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

THEOREM 6.1.1. *For all $x \in R_1$, $W_n(x) \to W(x)$ a.s. as $n \to \infty$, where W is an absolutely continuous distribution function with semicircle density*

$$w(x) = \begin{cases} \frac{2}{\pi}(1 - x^2)^{1/2}, & |x| \leqq 1, \\ 0, & |x| > 1. \end{cases} \tag{6.1.1}$$

PROOF. It is to be proved that

$$P\left(\lim_{n\to\infty} W_n(x) = W(x)\right) = 1. \tag{6.1.2}$$

Let

$$M_{k,n} = \int_{-\infty}^{\infty} x^k \, dW_n(x) \tag{6.1.3}$$

and

$$\gamma_k = \int_{-\infty}^{\infty} x^k \, dW(x) = \begin{cases} 0, & \text{odd } k, \\ \dfrac{k!}{2^k(\frac{1}{2}k)!(\frac{1}{2}k + 1)!}, & \text{even } k. \end{cases} \tag{6.1.4}$$

By Lemma 2.5.1, it will be sufficient to prove

$$P\left(\lim_{n\to\infty} M_{k,n} = \gamma_k, k \geqq 1\right) = 1. \tag{6.1.5}$$

This will be true if

$$P\left(\lim_{n\to\infty} M_{k,n} = \gamma_k\right) = 1 \tag{6.1.6}$$

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

$$|M_{k,n} - \gamma_k| \leqq |M_{k,n} - EM_{k,n}| + |EM_{k,n} - \gamma_k|, \tag{6.1.7}$$

and it will be sufficient to prove:

$$\begin{aligned} &\text{(a)}\ \lim_{n\to\infty} EM_{k,n} = \gamma_k, && k \geqq 1; \\ &\text{(b)}\ P\left(\lim_{n\to\infty} (M_{k,n} - EM_{k,n}) = 0\right) = 1, && k \geqq 1. \end{aligned} \tag{6.1.8}$$

For (b), it will be sufficient to prove

$$\sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty, \qquad k \geqq 1. \tag{6.1.9}$$

This is seen as follows. The statement

$$P\left(\lim_{n\to\infty} (M_{k,n} - EM_{k,n}) = 0\right) = 1 \tag{6.1.10}$$

is equivalent to

$$P\left(\limsup_{n\to\infty} \{\omega : |M_{k,n}(\omega) - EM_{k,n}| > \varepsilon\right) = 0 \tag{6.1.11}$$

for every $\varepsilon > 0$, by Lemma 2.5.2. Let $A_n = \{\omega : |M_{k,n}(\omega) - EM_{k,n}| > \varepsilon\}$. It is to be shown that $P(\limsup_{n\to\infty} A_n) = 0$. Chebyshev's inequality gives

$$P\left(|M_{k,n} - EM_{k,n}| > \varepsilon\right) \leqq \frac{E(M_{k,n} - EM_{k,n})^2}{\varepsilon^2}. \tag{6.1.12}$$

This and $\Sigma_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty$ implies $\Sigma_{n=1}^{\infty} P(A_n) < \infty$. Then Lemma 2.5.3 (Borel–Cantelli) gives $P(\limsup_{n\to\infty} A_n) = 0$. Altogether, then, it will be sufficient to prove:

$$\begin{aligned} &\text{(a)}\ \lim_{n\to\infty} EM_{k,n} = \gamma_k, && k \geqq 1; \\ &\text{(b)}\ \sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty, && k \geqq 1. \end{aligned} \tag{6.1.13}$$

The proof of (6.1.13) (a) follows.

Letting $A_{k,n}$ denote the class of all sequence $f : \{1, 2, \cdots, k + 1\} \to \{1, 2, \cdots, n\}$, one has

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

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 $(f(i), f(i+1))\in\{g_f(1), g_f(2), \cdots, g_f(k)\}$ such that $\#(f(i), f(i+1))_f + \#(f(i+1), f(i))_f = 1$. Then, by the independence and zero mean assumptions,

$$
(6.1.15)\qquad E\prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)} = Ea_{f(i)f(i+1)}E\prod_{\ell\neq i,\ell=1}^{k} a_{f(\ell)f(\ell+1)} = 0.
$$

Thus, one has

$$
(6.1.16)\qquad EM_{k,n} = \frac{1}{(2\sigma)^k n^{1+k/2}}\sum_{f\in C_{k,n}} E\prod_{\ell=1}^{k} a_{f(\ell)f(\ell+1)},
$$

where $C_{k,n} = \{f\in B_{k,n}:$ if $(i,j)\in\{g_f(1), g_f(2), \cdots, g_f(k)\}$, then $\#(i,j)_f + \#(i,j)_f \geqq 2\}$. Two cases are now considered, k odd and k even. Let $k = 2v + 1$. By Lemma 3.1, $d_f \leqq v + 1$ for all $f\in C_{2v+1,n}$. Thus,

$$
(6.1.17)\qquad EM_{2v+1,n} = \frac{1}{(2\sigma)^{2v+1} n^{v+3/2}}\sum_{j=1}^{v+1}\sum_{f} E\prod_{\ell=1}^{2v+1} a_{f(\ell)f(\ell+1)},
$$

where the summation is taken over $f\in C^j_{2v+1,n} = \{f\in C_{2v+1,n} : d_f = j\}$. By assumption (v),

$$
(6.1.18)\qquad \left|E\prod_{\ell=1}^{2v+1} a_{f(\ell)f(\ell+1)}\right| \leqq D_v < \infty,
$$

for some constant D_v. Hence,

$$
\begin{aligned}
(6.1.19)\qquad |EM_{2v+1,n}| &\leqq \frac{D_v\sum_{j=1}^{v+1}\# C^j_{2v+1,n}}{(2\sigma)^{2v+1} n^{(v+1)+1/2}}\\
&= \frac{D_v\sum_{j=1}^{v+1}\binom{n}{j}(\# C^j_{2v+1,j})}{(2\sigma)^{2v+1} n^{(v+1)+1/2}}.
\end{aligned}
$$

Since $\binom{n}{j}/n^{(\nu+1)+1/2} \to 0$ as $n \to \infty$ for $j = 1, 2, \cdots, \nu + 1$, one has $EM_{2\nu+1,n} \to 0$ as $n \to \infty$. Now consider $k = 2\nu$. By Lemma 3.1, $d_f \leqq \nu + 1$ for all $f \in C_{2\nu,n}$. Let

$$f_1(\nu, n) = \frac{1}{(2\sigma)^{2\nu} n^{\nu+1}} \sum_{f \in C_{2\nu,n}^{\nu+1}} E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)}, \tag{6.1.20}$$
$$f_2(\nu, n) = \frac{1}{(2\sigma)^{2\nu} n^{\nu+1}} \sum_{f \in \bigcup_{j=1}^{\nu} C_{2\nu,n}^{j}} E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)},$$

where $C_{2\nu,n}^{j} = \{f \in C_{2\nu,n} : d_f = j\}$. Note that $EM_{2\nu,n} = f_1(\nu, n) + f_2(\nu, n)$. By assumption (v),

$$\left| E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)} \right| \leqq \delta_\nu < \infty, \tag{6.1.21}$$

for some constant δ_ν. Thus,

$$|f_2(\nu, n)| \leqq \frac{\delta_\nu \sum_{j=1}^{\nu} \# C_{2\nu,n}^{j}}{(2\sigma)^{2\nu} n^{\nu+1}} = \frac{\delta_\nu \sum_{j=1}^{\nu} \binom{n}{j} (\# C_{2\nu,j}^{j})}{(2\sigma)^{2\nu} n^{\nu+1}}. \tag{6.1.22}$$

Since $\binom{n}{j}/n^{\nu+1} \to 0$ as $n \to \infty$ for $j = 1, \cdots, \nu$, one has $f_2(\nu, n) \to 0$ as $n \to \infty$. By Lemma 3.3,

$$E \prod_{\ell=1}^{2\nu} a_{f(\ell)f(\ell+1)} = \sigma^{2\nu} \tag{6.1.23}$$

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

$$f_1(\nu, n) = \frac{\sigma^{2\nu}(\# C_{2\nu,n}^{\nu+1})}{(2\sigma)^{2\nu} n^{\nu+1}}. \tag{6.1.24}$$

By Wigner's combinatorial theorem,

$$f_1(\nu, n) = \frac{(2\nu)!}{2^{2\nu}\nu!(\nu+1)!} + \frac{o(n^{\nu+1})}{(2\sigma)^{2\nu} n^{\nu+1}}. \tag{6.1.25}$$

Thus,

$$f_1(\nu, n) \to \frac{(2\nu)!}{2^{2\nu}\nu!(\nu+1)!} = \gamma_{2\nu} \tag{6.1.26}$$

as $n \to \infty$. Altogether

$$EM_{2\nu,n} \to \frac{(2\nu)!}{2^{2\nu}\nu!(\nu+1)!} = \gamma_{2\nu} \tag{6.1.27}$$

as $n \to \infty$. 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

$$E \prod_{i=1}^{k} a_{f(i)f(i+1)} \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)} - E \prod_{i=1}^{k} a_{f(i)f(i+1)} E \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)}. \tag{6.1.28}$$

One has

$$E(M_{k,n} - EM_{k,n})^2 = \frac{1}{(2\sigma)^{2k} n^{k+2}} \sum_{f \in B_{2k+1,n}} E(f), \tag{6.1.29}$$

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) $\{g_f(1), g_f(2), \cdots, g_f(k)\} \bigcap \{g_f(k+2), \cdots, g_f(2k+1), g_f^*(k+2), \cdots, g_f^*(2k+1)\} \neq \emptyset$, where $g_f^*(\ell)$ denotes the reverse of $g_f(\ell)$; (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 $d_f \leqq k$. Using condition (iii), suppose, for the sake of definiteness, that $g_f(s) = g_f(t)$ for some s, $1 \leqq s \leqq k$, and some t, $k + 2 \leqq t \leqq 2k + 1$. (The only other case to consider is when $g_f(s) = g_f^*(t)$ for some s, $1 \leqq s \leqq k$, and some t, $k + 2 \leqq t \leqq 2k + 1$, for which the following argument also applies.) Define a new sequence $h \in A_{2k+1,n}$ as follows: $h(1) = f(s)$, $h(2) = f(s+1)$, $\cdots, h(k-s+1) = f(k), h(k-s+2) = f(1), h(k-s+3) = f(2), \cdots, h(k) = f(s-1)$, $h(k+1) = f(s)$, $h(k+2) = f(t+1)$, $h(k+3) = f(t+2), \cdots$, $h(2k-t+2) = f(2k+1), h(2k-t+4) = f(k+3), \cdots, h(2k+1) = f(t-1)$, $h(2k+2) = f(t)$. It is immediate that $d_h = d_f$. The sequence of steps associated with h is

(6.1.30)

$$\begin{aligned} g_h(1) &= \big(f(s), f(s+1)\big), \cdots, g_h(k-s+1) = \big(f(k), f(1)\big), \\ g_h(k-s+2) &= \big(f(1), f(2)\big), \cdots, g_f(k) = \big(f(s-1), f(s)\big), \\ g_h(k+1) &= \big(f(s), f(t+1)\big) = \big(f(s), f(s+1)\big), \cdots, \\ g_h(2k-t+2) &= \big(f(2k+1), f(k+2)\big), \\ g_h(2k-t+3) &= \big(f(k+2), f(k+3)\big), \cdots, g_h(2k+1) \\ &= \big(f(t-1), f(t)\big). \end{aligned}$$

It is true that:

(i) $h(1) = h(2k+2)$;

(ii) if $(i, j) \in \{g_h(1), g_h(2), \cdots, g_h(2k+1)\}$, then $\#(j, i)_f \geqq 2$.

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

$$E\left(\prod_{i=1}^{k} a_{f(i)f(i+1)} \prod_{j=k+2}^{2k+1} a_{f(j)f(j+1)}\right) - E\left(\prod_{i=1}^{k} a_{f(i)f(i+1)} \mathop{E}_{j=k+2}^{2k+1} a_{f(j)f(j+1)}\right) \tag{6.1.31}$$

$$= E\left(\prod_{i=1}^{k} a_{h(i)h(i+1)} \prod_{j=k+2}^{2k+1} a_{h(j)h(j+1)}\right)$$

$$- E\left(\prod_{i=1}^{k} a_{h(i)h(i+1)} \underset{j=k+2}{\overset{2k+1}{E}} a_{h(j)h(j+1)}\right) = 0,$$

contrary to the assumption that for $f \in B_{2k+1,n}$ this term is nonzero. Hence, h must satisfy condition (ii) above. Lemma 3.4 then applies, giving $d_f = d_h \leqq k$. Now consider

$$E(M_{k,n} - EM_{k,n})^2 = \frac{1}{(2\sigma)^{2k}n^{k+2}} \sum_{j=1}^{k} \sum_{f \in B^j_{2k+1,n}} E(f), \tag{6.1.32}$$

where $B^j_{2k+1,n} = \{f \in B_{2k+1,n} : d_f = j\}$. By assumption (v), one has $|E(f)| \leqq G_\nu < \infty$ for some constant G_ν. Thus,

$$|E(M_{k,n} - EM_{k,n})^2| \leqq \frac{G_\nu \sum_{j=1}^{k} (\# B^j_{2k+1,n})}{(2\sigma)^{2k}n^{k+2}} \tag{6.1.33}$$

$$= \frac{G_\nu \sum_{j=1}^{k} \binom{n}{j} (\# B^j_{2k+1,j})}{(2\sigma)^{2k}n^{k+2}}.$$

Since

$$\sum_{n=1}^{\infty} \frac{\binom{n}{j}}{n^{k+2}} < \infty \tag{6.1.34}$$

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

$$\sum_{n=1}^{\infty} |E(M_{k,n} - EM_{k,n})^2| < \infty, \tag{6.1.35}$$

which implies

$$\sum_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty \tag{6.1.36}$$

which was to be proved. This completes the proof of Theorem 6.1.1.

6.2. *Comments.* A little reflection will reveal that the assumption of zero means for the diagonal elements is not necessary. For, in proving $EM_{k,n} \to \gamma_k$ as $n \to \infty$, 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 $(i, j) \in \{g_f(1), \cdots, g_f(k)\}$, then $\#(i,j)_f + (j,i)_f \geqq 2$. Condition (ii) alone implies $d_f \leqq [\frac{k}{2}] + 1$. If one assumes, however, that (iia) if $(i,j) \in \{g_f(1), \cdots, g_f(k)\}$, *where* $i \neq j$, then $\#(i,j)_f + \#(j,i)_f \geqq 2$, then conditions (i) and (iia) together imply $d_f \leqq [\frac{1}{2}k] + 1$. For odd k, say $k = 2\nu + 1$, one has $EM_{2\nu+1,n} \to 0 = \gamma_{2\nu+1}$ as $n \to \infty$ 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 3.2 would give $d_f \leqq \nu$. Thus, Wigner's combinatorial theorem holds under conditions (i), (iia), and (iii) $d_f = \nu + 1$; application of this theorem then gives $EM_{2\nu,n} \to \gamma_{2\nu}$ as $n \to \infty$ 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 $\Sigma_{n=1}^{\infty} E(M_{k,n} - EM_{k,n})^2 < \infty$.

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 6.1.1. 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 6.1.1 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.

7. Related results

7.1. *The Gaussian orthogonal ensemble.* 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. 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 $\{x_{ij}\}_{i \leqq j}$, $i, j = 1, 2, \cdots, n$ be an independent set of random variables on a probability space $(\Omega, \mathscr{F}, P)$. Let

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \\ x_{n1} & x_{n2} & & x_{nn} \end{bmatrix}, \tag{7.1.1}$$

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

$$Y_0 = (y_{ij}^0) = OXO', \tag{7.1.2}$$

where O is any orthogonal matrix. Let $x = (x_{11}, \cdots, x_{1n}, x_{22}, \cdots, x_{2n}, \cdots, x_{nn})$ and $y_0 = (y^0_{11}, \cdots, y^0_{1n}, y^0_{22}, \cdots, y^0_{2n}, \cdots, y^0_{nn})$. 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 C. E. Porter and N. Rosenzweig [15].

THEOREM 7.1.1. *For all $B \in \beta_{n(n+1)/2}$ and all orthogonal O,*

$$P(a \in B) = P(y_0 \in B) \tag{7.1.3}$$

if and only if x_{ii} is normal with mean μ and variance $2a^2$ and x_{ij}, $i < j$, is normal with O and variance a^2, for some constants μ and $a^2 > 0$.

A proof of this may be found in Olson and Uppuluri [12].

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 7.1.1 allows one to say that the elements of X are normally distributed as indicated in the theorem. The physicists call this model the Gaussian orthogonal ensemble.

For the particular Gaussian orthogonal ensemble $X_{ii} \sim n(0, 1)$ and $X_{ij} \sim n(0, \frac{1}{2})$ the probability density function of the $n \times n$ symmetric random matrix $X = (X_{ij})$ is given by

$$\text{const} \exp\{-\tfrac{1}{2} \operatorname{tr} X^2\}. \tag{7.1.4}$$

By using standard methods of multivariate analysis, one can show that the probability density function of the eigenvalues $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ of X is given by

$$\frac{1}{2^{n/2} n! \prod_{j=1}^{n} \Gamma(\frac{1}{2}j)} \exp\left\{-\tfrac{1}{2}\sum_{i=1}^{n} \varepsilon_i^2\right\} \prod_{i<j} |\varepsilon_i - \varepsilon_j|. \tag{7.1.5}$$

We note from this explicit form of the density function that the eigenvalues $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ in this case are exchangeable (for definition of exchangeability see [6]). M. L. Mehta and M. Gaudin [11] exploit this property by using the technique of integration over alternate variables (see N. G. de Bruijn [4]) to obtain the density function of a single eigenvalue (for the case $n = 2m$) as

$$\sigma_{2m}(\varepsilon) = \sum_{i=0}^{2m-1} \varphi_i^2(\varepsilon) + \sqrt{m}\varphi_{2m-1}(\varepsilon) \int_0^{\varepsilon} \varphi_{2m}(y)\, dy, \tag{7.1.6}$$

where

$$\varphi_j(\varepsilon) = (2^j j! \sqrt{\pi})^{-1/2} e^{\varepsilon^2/2} \left(-\frac{d}{d\varepsilon}\right)^j e^{-\varepsilon^2}. \tag{7.1.7}$$

Then it is claimed by Mehta and Gaudin [11] that $\sigma_{2m}(x)$ is asymptotically equal to $\sigma(x)$, where

$$\sigma(x) = \begin{cases} \frac{1}{\pi}(4m - x^2)^{1/2}, & |x| < (4m)^{1/2}, \\ 0, & \text{otherwise.} \end{cases} \tag{7.1.8}$$

Indications of why this holds are also outlined in an appendix to Mehta's book [10]. For a different approach to the convergence to the semicircle law for a Gaussian orthogonal ensemble one may refer to Wigner [17].

For a normalized Gaussian orthogonal ensemble, Theorem 6.1.1 gives the semicircle law as the almost sure limit of the empirical distribution function of the eigenvalues of the normalized random matrix $X/\sqrt{2n}$. This, however, does not imply the convergence of the corresponding probability density functions mentioned above.

7.2. *A random Toeplitz ensemble.* It is of interest to know whether there exist random ensembles whose empirical distribution functions of their eigenvalues converge to limiting distributions other than Wigner's semicircle distribution. Such an ensemble was recently discussed by V. M. Dubner [5]. He considered the random Toeplitz ensemble described below.

Let $\{Z_k, k = 0, \pm 1, \cdots, \pm 2m\}$ be a set of complex valued random variables such that:

(i) $Z_k = \bar{Z}_{-k}$;

(ii) $Z_k = x_k + iy_k$, where $\{x_k, y_k, k = 0, 1, \cdots, 2m\}$ is an independent set of random variables each of which has a Gaussian distribution with mean 0 and variance σ^2 (except, $y_0 = 0$).

Let $A_{2m+1} = (a_{ij})_{i,j=1}^{2m+1}$, be a random matrix such that:

(i) $a_{ij} = \bar{a}_{ji}$

(ii) For $i < j$, $a_{ij} = Z_{j-i}$, $0 \leqq j - i \leqq [\frac{1}{2}(2m + 1)]$, $a_{ij} = \bar{Z}_{(2m+1)-(j-i)}$, for $[\frac{1}{2}(2m + 1)] + 1 \leqq j - i \leqq 2m$.

For instance, when $m = 2$, we have the 5×5 random matrix

$$\begin{bmatrix} Z_0 & Z_1 & Z_2 & \bar{Z}_2 & \bar{Z}_1 \\ \bar{Z}_1 & Z_0 & Z_1 & Z_2 & \bar{Z}_2 \\ \bar{Z}_2 & \bar{Z}_1 & Z_0 & Z_1 & Z_2 \\ Z_2 & \bar{Z}_2 & \bar{Z}_1 & Z_0 & Z_1 \\ Z_1 & Z_2 & \bar{Z}_2 & \bar{Z}_1 & Z_0 \end{bmatrix} \tag{7.2.1}$$

For this random Toeplitz ensemble, Dubner [5] has indicated that the asymptotic distribution of the sequence of empirical distribution functions of the set of eigenvalues is Gaussian.

7.3. *A Wishart ensemble.* In general, statisticians are interested in the distribution of the eigenvalues of a sample variance-covariance type matrix, in contrast to the physicists' interest in the distribution of the eigenvalues of a random matrix of the most general type. Recently, C. Stein considered the limiting distribution of the expected value of the empirical distribution function of the eigenvalues of random matrices of the variance-covariance type (Stein's result appears in Technical Report No. 42, December 2, 1969, Department of Statistics, Stanford University).

Stein's result may be stated as follows. Let $X = (X_{ij})$ be a $p \times n$ random matrix such that:

(i) $\{X_{ij}, 1 \leqq i \leqq p, 1 \leqq j \leqq n\}$ is an independent set of random variables;
(ii) $EX_{ij} = 0$;
(iii) $EX_{ij}^2 = 1$;
(iv) $E|X_{ij}|^k \leqq C_k < \infty$ for $k = 1, 2, \cdots$.

Let $B = (1/n)XX'$ and denote by $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_p$ the ordered eigenvalues of B. Denote the empirical distribution function of $\lambda_1, \lambda_2, \cdots, \lambda_n$ by $W_{p,n}(x)$ so that

$$W_{p,n}(x) = \frac{1}{p}(\#\lambda_i < x). \tag{7.3.1}$$

THEOREM 7.3.1. *Let F_β be the absolutely continuous distribution function with density*

$$f_\beta(x) = \begin{cases} \dfrac{\beta}{2\pi x}[(x-a)(b-x)]^{1/2} & a \leqq x \leqq b, \\ 0, & \textit{elsewhere}, \end{cases} \tag{7.3.2}$$

where $a = [1 - \beta^{-1/2}]^2$ and $b = [1 + \beta^{-1/2}]^2$, then

$$EW_{p,n}(x) \to F_\beta(x) \tag{7.3.3}$$

as $p \to \infty$, $n \to \infty$ in such a way that $n/p \to \beta > 1$.

It is interesting to note that when $\beta = 1$ there is a relation between this result and Wigner's semicircle distribution. If X is a random variable with a semicircle distribution, then $Y = 4X^2$ has the probability density function

$$g(y) = \begin{cases} \dfrac{1}{2\pi}(4-y)^{1/2}y^{-1/2}, & 0 \leqq y \leqq 4, \\ 0, & \text{elsewhere}. \end{cases} \tag{7.3.4}$$

REFERENCES

[1] T. W. ANDERSON, *An Introduction to Multivariate Statistical Analysis*, New York, Wiley, 1958.

[2] L. ARNOLD, "On the asymptotic distribution of the eigenvalues of random matrices," *J. Math. Anal. Appl.*, Vol. 20 (1967), pp. 262–268.

[3] K. L. CHUNG, *A Course in Probability Theory*, New York, Harcourt, Brace and World, 1968.

[4] N. G. DE BRUIJN, "On some multiple integrals involving determinants," *J. Indian Math. Soc.*, Vol. 19 (1955), pp. 133–151.

[5] V. M. DUBNER, "The distribution of the eigenvalues of random cyclic matrices," *Theor. Probability Appl.*, Vol. 14 (1969), pp. 342–344.

[6] WILLIAM FELLER, *An Introduction to Probability Theory and its Applications*, Vol. 2, New York, Wiley, 1966.
[7] U. GRENANDER, *Probabilities on Algebraic Structures*, New York, Wiley, 1963.
[8] T. F. JORDAN, *Linear Operators for Quantum Mechanics*, New York, Wiley, 1969.
[9] M. LOÈVE, *Probability Theory*, Princeton, Van Nostrand, 1963 (3rd ed.).
[10] M. L. MEHTA, *Random Matrices*, New York, Academic Press, 1967.
[11] M. L. MEHTA and M. GAUDIN, "Addendum and Erratum," *Nuclear Phys.*, Vol. 22 (1961), p. 340.
[12] W. H. OLSON and V. R. R. UPPULURI, "Characterization of the distribution of a random matrix by rotational invariance," *Sankhyā Ser. A*, Vol. 32 (1970), pp. 325–328.
[13] A. M. OSTROWSKI, *Solution of Equations and Systems of Equations*, New York, Academic Press, 1960.
[14] C. E. PORTER, *Statistical Theories of Spectra: Fluctuations*, New York, Academic Press, 1965.
[15] C. E. PORTER and N. ROSENZWEIG, "Statistical properties of atomic and nuclear spectra," *Ann. Acad. Sci. Fennicae A VI*, Vol. 44 (1960). (Reprinted in [14], pp. 235–299.)
[16] E. P. WIGNER, "Characteristic vectors of bordered matrices with infinite dimensions," *Ann. of Math.*, Vol. 62 (1955), pp. 548–564.
[17] ———, "Statistical properties of real symmetric matrices with many dimensions," *Proceedings of the Canadian Mathematical Congress*, Toronto, University of Toronto Press, 1957, pp. 174–184.
[18] ———, "On the distribution of the roots of certain symmetric matrices," *Ann. of Math.*, Vol. 67 (1958), pp. 325–326.