Z. Wahrscheinlichkeitstheorie verw. Geb. 19, 191–198 (1971)

On Wigner's Semicircle Law for the Eigenvalues of Random Matrices

L. ARNOLD

1. Introduction

Wigner asked in [8] for the general conditions of validity for his so-called semicircle law for the distribution of eigenvalues of random matrices which is important in the statistical theory of energy levels of heavy atomic nuclei [6, 7]. We discovered [2] that the semicircle law possesses the following completely deterministic version from which probabilistic applications can be derived relatively easily.

Let $A_n=(a_{ij})$, $1\leqq i, j\leqq n$, be the nth section of an infinite Hermitian matrix, $\{\lambda_k^{(n)}\}_{1\leqq k\leqq n}$ its eigenvalues and $\{u_k^{(n)}\}_{1\leqq k\leqq n}$ the corresponding (orthonormalized column-) eigenvectors. Let $v_n^*=(a_{n1}, a_{n2}, \ldots, a_{n,n-1})$, put

$$X_n(t)=\big(n(n-1)\big)^{-\frac{1}{2}}\sum_{k=1}^{[(n-1)t]} |v_n^* u_k^{(n-1)}|^2, \qquad 0\leqq t\leqq 1 \tag{1}$$

(bookkeeping function for the length of the projections of the new row v_n^* of A_n onto the eigenvectors of the preceding matrix A_{n-1}), let finally

$$F_n(x)=n^{-1}(\text{number of } \lambda_k^{(n)}\leqq x\sqrt{n},\ 1\leqq k\leqq n) \tag{2}$$

(empirical d.f. of the eigenvalues of $A_n/\sqrt{n}$).

Theorem 1. (*Deterministic version of the semicircle law, see* [2].) *Suppose*

(i) $\lim_n$ (*number of* $k\leqq n$ *with* $|a_{kk}|>\sqrt{n})/n=0$,

(ii) $\lim_n X_n(t)=C\,t\ (0<C<\infty, 0\leqq t\leqq 1)$.

Then

$$F_n \Rightarrow W(\cdot, C) \qquad (n\to\infty), \tag{3}$$

where W is absolutely continuous with density (semicircle!)

$$w(x, C)=\begin{cases}(2C\pi)^{-1}(4C-x^2)^{\frac{1}{2}} & \text{for } |x|\leqq 2\sqrt{C},\\ 0 & \text{for } |x|>2\sqrt{C}.\end{cases}$$

Suppose now that the matrix elements a_{ij} are real-valued random variables defined on a fixed probability space $(\Omega, \mathscr{F}, P)$, being independent for $i\geqq j$ and satisfying $a_{ij}=a_{ji}$ a.s. Suppose further that the diagonal elements a_{ii}, $i\geqq 1$, are identically distributed according to the d.f. G, and that the off-diagonal elements a_{ij}, $i>j$, are also identically distributed with d.f. H having variance σ^2.

What we are interested in is the asymptotic behavior of the sequence of stochastic processes defined by (2). We are aiming at a strong (convergence a.s.) and a weak (convergence in probability) form of the semicircle law (3) for (2).

In an earlier paper [1], we proved by a completely different method that the weak form of (3) holds under the conditions $\int x^2\,dG<\infty$, $\int x^4\,dH<\infty$ and $\int x\,dH=0$. If, moreover, $\int x^4\,dG<\infty$ and $\int x^6\,dH<\infty$, then the strong form holds, choosing in both cases $C=\sigma^2$.

In this paper, we are able to eliminate any condition about G and the condition $\int x\,dH=0$ and to reduce the moment restrictions on H by 2. The method used consists in utilizing Theorem 1 by verifying the assumptions (i) and (ii).

2. Stochastic Convergence of $\{X_n(t)\}$

Due to the independence of the a_{ij}'s, the vector v_n^* and the eigenvectors of A_{n-1} are independent, too. This fact will be used without further mentioning.

Lemma 1. *We have*

$$\lim_n X_n(1)=C<\infty \text{ in probability} \Leftrightarrow \int x^2\,dH<\infty.$$

If $\int x\,dH=0$ *and* $\sigma^2<\infty$, *then* $C=\sigma^2$ *and*

$$\lim_n X_n(t)=\sigma^2 t \text{ in probability}, \qquad 0\leqq t\leqq 1.$$

Proof. Since

$$X_n(1)=\big(n(n-1)\big)^{-\frac{1}{2}}\sum_{k=1}^{n-1} a_{kn}^2,$$

the first part of the lemma is essentially the weak law of large numbers (cf. Feller [3], p. 232). For the proof of the second assertion, put

$$\bar{a}_{kn}=\begin{cases} a_{kn} & \text{for } |a_{kn}|\leqq\sqrt{n-1},\\ 0 & \text{for } |a_{kn}|>\sqrt{n-1},\end{cases}$$

$1\leqq k\leqq n-1$. Let $\bar{X}_n(t)$ be the expression obtained from (1) by replacing a_{kn} by $\bar{a}_{kn}$, $1\leqq k\leqq n-1$. Clearly

$$P[|X_n(t)-E\bar{X}_n(t)|>\varepsilon]\leqq P[|\bar{X}_n(t)-E\bar{X}_n(t)|>\varepsilon]+P[X_n(t)\neq\bar{X}_n(t)].$$

We are going to show that the right-hand side tends to 0. Indeed,

$$P[X_n(t)\neq\bar{X}_n(t)]\leqq\sum_{k=1}^{n-1}P[|a_{kn}|>\sqrt{n-1}]=(n-1)\,P[|a_{12}|>\sqrt{n-1}]\to 0,$$

since $\sum P[|a_{12}|>\sqrt{n}]<\infty$, which is the case iff $\sigma^2<\infty$. By Chebyshev's inequality,

$$P[|\bar{X}_n(t)-E\bar{X}_n(t)|>\varepsilon]\leqq\varepsilon^{-2}\big(E\bar{X}_n(t)^2-(E\bar{X}_n(t))^2\big).$$

The proof of the lemma will be completed if we know that $E\bar{X}_n(t)\to\sigma^2 t$ and $E\bar{X}_n(t)^2\to\sigma^4 t^2$.

We have

$$E\overline{X}_n(t)=(n(n-1))^{-\frac{1}{2}}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}E(\bar{a}_{in}\,\bar{a}_{jn})\,E(r_{ij}),$$

where

$$r_{ij}=r_{ij}(t)=\sum_{k=1}^{[(n-1)t]}u_{ik}^{(n-1)}\,u_{jk}^{(n-1)}.$$

Using

$$\sum_{i=1}^{n-1}r_{ii}=[(n-1)\,t],\qquad E\,\bar{a}_{in}\,\bar{a}_{jn}=E\,\bar{a}_{in}^2=\overline{m}_2\quad(i=j),\qquad=(E\,\bar{a}_{in})^2=\overline{m}_1^2\quad(i\neq j),$$

we obtain

$$E\overline{X}_n(t)=(n(n-1))^{-\frac{1}{2}}\left([(n-1)\,t](\overline{m}_2-\overline{m}_1^2)+\overline{m}_1^2\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}E\,r_{ij}\right). \tag{4}$$

By assumption, $\overline{m}_2\to\sigma^2<\infty$ and $\overline{m}_1=o(n^{-\frac{1}{2}})$, the last statement following from

$$\int|x|^k\,dH<\infty,\qquad \int x\,dH=o\Rightarrow n^{(k-1)/2}\int_{|x|\leqq\sqrt{n}}x\,dH=o(1)\qquad(n\to\infty,\,k\geqq 1) \tag{5}$$

(see Arnold [1], p. 265). Putting this into (4) together with the trivial estimate $|\sum\sum E\,r_{ij}|\leqq(n-1)^2$, *we obtain* $E\overline{X}_n(t)\to\sigma^2\,t$.

For estimating

$$E\overline{X}_n(t)^2=(n(n-1))^{-1}\sum_{i_1=1}^{n-1}\cdots\sum_{i_4=1}^{n-1}E(\bar{a}_{i_1n}\,\bar{a}_{i_2n}\,\bar{a}_{i_3n}\,\bar{a}_{i_4n})\,E(r_{i_1i_2}\,r_{i_3i_4})$$

we have to consider seven different cases of index degeneracy. Using again $\overline{m}_1=o(n^{-\frac{1}{2}})$, $\overline{m}_2\to\sigma^2$ and in addition $\overline{m}_3=o(n^{\frac{1}{2}})$ and $\overline{m}_4=o(n)$, which follows from

$$\int|x|^k\,dH<\infty\Rightarrow n^{(k-r)/2}\int_{|x|\leqq\sqrt{n}}|x|^r\,dH=o(1)\qquad(r\geqq k+1) \tag{6}$$

(see Arnold [1], p. 264), we arrive at

$$E\overline{X}_n(t)^2=\sigma^4\,t^2+o(1),$$

q.e.d.

3. Almost Sure Convergence of $\{X_n(t)\}$

A much more delicate truncation technique has to be applied in order to obtain

Lemma 2. *We have*

$$\lim_n X_n(1)=C<\infty\;a.s.\Leftrightarrow\int x^4\,dH<\infty.$$

If $\int x\,dH=0$ *and* $\int x^4\,dH<\infty$, *then* $C=\sigma^2$ *and*

$$\lim_n X_n(t)=\sigma^2\,t\;a.s.,\qquad 0\leqq t\leqq 1.$$

Proof. 1. According to Lemma 1, there is a chance for a.s. convergence of $\{X_n(1)\}$ only if $\sigma^2<\infty$. The prospective limit C must be equal to σ^2. Obviously,

$$X_n(1)\to\sigma^2\text{ a.s.}\Leftrightarrow S_{n-1}=(n-1)^{-1}\sum_{k=1}^{n-1}(a_{kn}^2-\sigma^2)\to 0\text{ a.s.}$$

Since $\{S_n\}$ is a sequence of independent random variables, the Borel-Cantelli lemma yields

$$S_n \to 0 \text{ a.s.} \Leftrightarrow \sum P[|S_n| > \varepsilon] < \infty \qquad \text{for all } \varepsilon > 0.$$

By a theorem of Heyde and Rohatgi [4] (Theorem 2) this is equivalent to

$$\sum n P[|a_{12}^2 - \sigma^2| > n] < \infty \quad \text{and} \quad \int_{|x|<n} (x^2 - \sigma^2)\, dH \to 0.$$

The second condition is fulfilled since $\sigma^2 < \infty$, and the first one is equivalent to $E a_{12}^4 = \int x^4\, dH = m_4 < \infty$, according to the relation

$$E|X|^{(t+1)r} < \infty \Leftrightarrow \sum n^t P[|X| > n^{1/r}] < \infty \tag{7}$$

(see [4], p. 74). This proves the first part of the lemma.

2. Suppose now $m_4 < \infty$, $m_1 = \int x\, dH = 0$, $m_2 = \sigma^2$. We are going to prove that

$$\sum P[|X_n(t) - \sigma^2 t| > \varepsilon] < \infty \qquad \text{for all } \varepsilon > 0,$$

which is sufficient for $X_n(t) \to \sigma^2 t$ a.s. This time, our truncation level for a_{kn}, $1 \leqq k \leqq n-1$, will be

$$\tau_n = (n-1)^{\gamma/2},$$

where the appropriate choice of $\gamma \in (0, 1]$ will result from the proof. By (5) and (6), $\bar{m}_1 = o(n^{-3\gamma/2})$, $\bar{m}_2 \to \sigma^2$, $\bar{m}_3 \to m_3 = \int x^3\, dH$, $\bar{m}_4 \to m_4$, and $\bar{m}_r = o(n^{\gamma(r/2-2)})$ $(r \geqq 5)$, which we have to apply for $r = 5$, 6, 7, and 8. Finally, for the r_{ij}'s introduced in Section 2, we have to take into account that

$$\sum_{\substack{j=1 \\ j \neq i}}^{n-1} |r_{ij}|^p \leqq \tfrac{1}{4} \quad (p \geqq 2), \qquad \left| \sum_{\substack{j=1 \\ j \neq i}}^{n-1} r_{ij} \right| \leqq \sqrt{n}/2,$$

and

$$\sum_{i=1}^{n-1} \sum_{j=1}^{n-1} r_{ij}^2 = [(n-1)\, t].$$

3. Consider the following events (in Ω):

$$A_{1n} = [|a_{kn}| > \sqrt{\varepsilon}\sqrt{n-1}/2 \text{ for at least one } k \leqq n-1],$$
$$A_{2n} = [|a_{kn}| > (n-1)^{\gamma/2} \text{ for at least two } k \leqq n-1],$$
$$A_{3n} = [|\bar{S}_{kn}(t)| > \sqrt{\varepsilon}\sqrt{n}/4 \text{ for at least one } k \leqq n-1],$$

where for $1 \leqq k \leqq n-1$

$$\bar{S}_{kn}(t) = \sum_{\substack{i=1 \\ i \neq k}}^{n-1} \bar{a}_{in}\, r_{ik}(t),$$

$$A_{4n} = [|\bar{X}_n(t) - \sigma^2 t| > \varepsilon/2],$$

$$\Omega_n = A_{1n} \cup A_{2n} \cup A_{3n} \cup A_{4n}$$

and

$$B_n = [|X_n(t) - \sigma^2 t| > \varepsilon].$$

We have

$$B_n \subset \Omega_n,$$

and therefore $P(B_n) \leqq P(\Omega_n) \leqq \sum_{i=1}^{4} P(A_{in})$, whence

$$\sum_n P(B_n) \leqq \sum_{i=1}^{4} \sum_n P(A_{in}).$$

We complete the proof of the lemma by showing that the four series on the right-hand side of the last inequality converge.

4. Convergence of $\sum P(A_{1n})$: We have

$$P(A_{1n}) = (n-1)\, P[|a_{12}| > \sqrt{\varepsilon}\sqrt{n-1}/2],$$

thus

$$\sum_n P(A_{1n}) \leqq \sum_n n P[|a_{1n}| > \sqrt{\varepsilon}\sqrt{n}/2].$$

By (7), the last series is finite iff $m_4 < \infty$.

5. Convergence of $\sum P(A_{2n})$: We have

$$A_{2n} = \bigcup_{i=1}^{n-1} \bigcup_{\substack{j=1 \\ i \neq j}}^{n-1} [|a_{in}| > (n-1)^{\gamma/2} \text{ and } |a_{jn}| > (n-1)^{\gamma/2}],$$

therefore (independence!)

$$P(A_{2n}) \leqq (n-1)^2 (P[|a_{12}| > (n-1)^{\gamma/2}])^2.$$

Since $m_4 < \infty$, $n^{2\gamma} P[|a_{12}| > n^{\gamma/2}] \to 0$, thus

$$\sum (nP[|a_{12}| > n^{\gamma/2}])^2 = \sum o(n^{2-4\gamma}) < \infty,$$

whenever $\gamma > \frac{3}{4}$.

6. Convergence of $\sum P(A_{3n})$: Putting $\varepsilon_1 = \sqrt{\varepsilon}/4$, Chebyshev's inequality yields

$$P(A_{3n}) \leqq \sum_{k=1}^{n-1} P[|\bar{S}_{kn}(t)| > \varepsilon_1 \sqrt{n}] \leqq \varepsilon_1^{-6} n^{-3} \sum_{k=1}^{n-1} E\bar{S}_{kn}^6.$$

In detail,

$$E\bar{S}_{kn}^6 = \sum_{\substack{i_1=1 \\ i_1 \neq k}}^{n-1} \cdots \sum_{\substack{i_6=1 \\ i_6 \neq k}}^{n-1} E(\bar{a}_{i_1 n} \ldots \bar{a}_{i_6 n})\, E(r_{i_1 k} \ldots r_{i_6 k}).$$

A systematic search through possible index degeneracies leads to $E\bar{S}_{kn} = o(n^{\gamma})$, thus

$$\sum P(A_{3n}) \leqq \sum o(1)\, n^{1+\gamma-3}.$$

The last series is finite, if $\gamma < 1$. It turns out that the restrictions

$$\tfrac{3}{4} < \gamma < 1$$

put on γ up to now will also assure the convergence of the remaining series.

7. Convergence of $\sum P(A_{4n})$: Clearly,

$$P(A_{4n}) \leqq (2/\varepsilon)^4 E(\bar{X}_n(t) - \sigma^2 t)^4,$$

furthermore

$$E(\bar{X}_n(t) - \sigma^2 t)^4 \leqq 8 E(\bar{X}_n(t) - E\bar{X}_n(t))^4 + 8(E\bar{X}_n(t) - \sigma^2 t)^4.$$

According to the proof of Lemma 1,

$$E\bar{X}_n(t) - \sigma^2 t = o(n^{-\gamma}),$$

so $\sum (E\bar{X}_n(t) - \sigma^2 t)^4$ certainly converges. After cumbersome, but simple calculations following the lines of the proof of Lemma 1, we obtain

$$\sum E(\bar{X}_n(t) - E\bar{X}_n(t))^4 < \infty,$$

q.e.d.

4. The Semicircle Law

The essential part of Lemma 1 as well as of Lemma 2 requires $\int x\,dH = 0$. The following lemma assures that the limit of $\{F_n\}$ is not perturbed by a non-vanishing expectation of the a_{ij}'s. If F is any d.f.,

$$\hat{F}(z) = \int_{x=-\infty}^{\infty} (x-z)^{-1} dF(x), \qquad \operatorname{Im}(z) > 0,$$

is known as the *Stieltjes transform* of F. F is uniquely determined by $\hat{F}$, and uniform convergence of $\{\hat{F}_n\}$ in compact z sets is equivalent to vague convergence of $\{F_n\}$ (see [2], appendix).

Lemma 3. *Let A be an $n \times n$ Hermitian matrix, E the $n \times n$ matrix having all elements equal to* 1, *$D = \operatorname{diag}(d_1, \ldots, d_n)$ a real diagonal matrix. Denote by F, F_1 and F_2 the empirical d.f. of the eigenvalues of A, $A + aE$ (a real) and $A + D$, respectively. Then*

(i) $|\hat{F}(z) - \hat{F}_1(z)| \leqq (n \operatorname{Im}(z))^{-1}$,

(ii) $|\hat{F}(z) - \hat{F}_2(z)| \leqq (\operatorname{Im}(z))^{-2} \max_{1 \leqq i \leqq n} |d_i|$,

the bounds being independent of A and a.

Proof. (i) We have for $\operatorname{Im}(z) > 0$ $\hat{F}(z) = n^{-1} \operatorname{tr} R(z, A)$ and $\hat{F}_1(z) = n^{-1} \operatorname{tr} R(z, A + aE)$, where $R(z, A) = (A - zI)^{-1}$ denotes the resolvent and $\operatorname{tr} A$ the trace of A. The insertion of

$$(I + aER(z, A))^{-1} = I - (a/(1 + a e' R(z, A) e)) ER(z, A),$$

$$e' = (1, 1, \ldots, 1) \qquad (n \text{ times}),$$

into the second resolvent equation

$$R(z, A + aE) = R(z, A)(I + aER(z, A))^{-1}$$

and passing to traces leads to

$$\hat{F}(z) - \hat{F}_1(z) = n^{-1} \frac{a e' R(z, A)^2 e}{1 + a e' R(z, A) e}.$$

If $f(z)=\int (x-z)^{-1}\,d\mu(x)$, where $\mu(x)=e' S(x)\,e$, S being the spectral matrix of A, we have $e' R(z, A)\,e=f(z)$ and $e' R(z, A)^2\,e=f'(z)$, thus

$$\hat{F}(z)-\hat{F}_1(z)=n^{-1}\frac{af'(z)}{1+af(z)}.$$

Since

$$|1+af(z)|\geqq|a|\,\mathrm{Im}(f(z))=|a|\,\mathrm{Im}(z)\int|x-z|^{-2}\,d\mu(x)$$

and

$$|af'(z)|\leqq|a|\int|x-z|^{-2}\,d\mu(x),$$

we obtain

$$|\hat{F}(z)-\hat{F}_1(z)|\leqq(n\,\mathrm{Im}(z))^{-1}.$$

(ii) Again by the second resolvent equation,

$$R(z, A+D)-R(z, A)=R(z, A+D)\,DR(z, A),$$

therefore

$$\begin{aligned}|\hat{F}(z)-\hat{F}_2(z)|&\leqq n^{-1}\,|\mathrm{tr}\,R(z, A+D)\,DR(z, A)|\\&\leqq\|R(z, A+D)\,DR(z, A)\|\\&\leqq(\mathrm{Im}(z))^{-2}\max_{1\leqq i\leqq n}|d_i|,\end{aligned}$$

q.e.d.

We are now in a position to prove

Theorem 2. *Let F_n be the empirical d.f. of the matrix $A_n/\sqrt{n}$ as defined by* (2), *where A_n is a random matrix satisfying the conditions stated in Section* 1. *Then*

(i) *(Weak semicircle law): If $\sigma^2<\infty$, then*

$$F_n\Rightarrow W(\cdot,\sigma^2)\qquad \textit{in probability},$$

where W is Wigner's semicircle d.f. defined in Theorem 1.

(ii) *(Strong semicircle law): If, moreover, $\int x^4\,dH<\infty$, then*

$$F_n\Rightarrow W(\cdot,\sigma^2)\qquad a.s.$$

Proof. By virtue of Lemma 3, it is no restriction of generality to assume $\int x\,dH=0$. A look at the proof of Theorem 1 shows that it remains true if all limits are interpreted as limits in probability. Condition (i) of Theorem 1 means

$$Z_n=n^{-1}\sum_{k=1}^{n} I_{(|a_{kk}|>\sqrt{n})}\to 0\qquad\text{in probability},$$

which is true since

$$EZ_n=P[|a_{11}|>\sqrt{n}]\to 0.$$

Actually, by the strong law of large numbers (see [5], p. 238), we even have

$$Z_n\to 0\qquad\text{a.s.}$$

Hence, the weak and strong version of the semicircle law follow immediately from Lemma 1 and Lemma 2, resp., q.e.d.

In physical applications of the semicircle law it is sometimes required to determine the eigenvalue distribution of functions of A_n. This can be done by the following

Corollary 1. *Let $f(x)$ be a real-valued measurable function on the real line being continuous $W(\cdot, \sigma^2)$-a.s. Denote by $f(F)$ the image under f of the measure corresponding to the d.f. F. Define the matrix $f(A)$ as usual by $f(A)=\int f(x)\,dS(x)$, $S(x)$ being the spectral matrix of A. Then for the sequences of measures μ_n defined by*

$$\mu_n(M)=n^{-1}(\text{number of eigenvalues of } f(A_n/\sqrt{n}) \text{ belonging to } M)$$

we have

$$\mu_n \Rightarrow f(W(\cdot, \sigma^2)) \qquad \text{in probability or a.s.}$$

whenever

$$F_n \Rightarrow W(\cdot, \sigma^2) \qquad \text{in probability or a.s.}$$

Proof. Clear by observing that $\mu_n = f(F_n)$.

As an example, put $f(x)=x^2$. The asymptotic d.f. of the eigenvalues of A_n^2/n has density

$$g(x)=\begin{cases}(2/\pi\sqrt{x})(1-x)^{\frac{1}{2}} & \text{for } 0<x<1,\\ 0 & \text{for } x<0 \text{ and } x>1,\end{cases}$$

whereas it is sometimes incorrectly assumed that g is a quartercircle (see e.g. [8], p. 7).

We conjecture that the conditions for the strong semicircle law can still be reduced to the finiteness of σ^2. The results essentially carry over to the Hermitian case.

References

1. Arnold, L.: On the asymptotic distribution of the eigenvalues of random matrices. J. math. Analysis Appl. **20**, 262–268 (1967).
2. – Deterministic version of Wigner's semicircle law for the distribution of matrix eigenvalues. (To appear.)
3. Feller, W.: An introduction to probability theory and its applications, Vol. II. New York: John Wiley 1966.
4. Heyde, C. C., Rohatgi, V. K.: A pair of complementary theorems on convergence rates in the law of large numbers. Proc. Cambridge philos. Soc. **63**, 73–82 (1967).
5. Loève, M.: Probability theory, 2nd edn. New York: Van Nostrand 1960.
6. Metha, M. L.: Random matrices. New York: Academic Press 1967.
7. Porter, C. E. (ed.): Statistical theories of spectra: Fluctuations. New York: Academic Press 1965.
8. Wigner, E. P.: Random matrices in physics. SIAM Review **9**, 1–23 (1967).

Dr. L. Arnold
Fachbereich Mathematik der Universität Stuttgart
BRD-7000 Stuttgart 1, Herdweg 51
Germany

(Received May 4, 1970)