TSUKUBA J. MATH.
Vol. 25 No. 2 (2001), 443–464

DERIVATION OF WIGNER'S SEMI-CIRCLE LAW FOR A CLASS OF MATRIX ENSEMBLES VIA BROWNIAN MOTION

By

Tsuyoshi HIRATSUKA* and Nariyuki MINAMI†

Abstract. Introducing a fictitious time evolution in a random matrix model having Gaussian entries, we prove that the empirical distribution of the scaled eigenvalues of the random matrix converges in probability to the Wigner's semi-circle law.

1. Introduction

The semicircle law as a limiting distribution of eigenvalues of large random matrices is well known since Wigner's work in 1950's. ([11]. See the introduction of [9] for a nice historical account.) As a theorem in probability theory, the following beautiful result was obtained as early as in 1970's by L. Arnold and R. Wegmann. Namely let $\{X_{kl}; k, l \geq 1\}$ be a family of real random variables defined on a common probability space $(\Omega, \mathscr{F}, P)$. Assume that $\{X_{kl}; l \geq k \geq 1\}$ are independent, among which $\{X_{kk}; k \geq 1\}$ are identically distributed with distribution function G, and $\{X_{kl}; l > k \geq 1\}$ are also identically distributed with distribution function H which has finite positive variance $v = \sigma^2$. Suppose further that $X_{kl} = X_{lk}$, so that $Q^{(n)} = (X_{kl}/\sqrt{n})_{k,l=1}^n$ is a real symmetric $n \times n$ random matrix. Then it was proved ([2], [1], [10]) that with probability one, the empirical distribution

$$\mu^{(n)}(dx) = \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j^{(n)}}(dx)$$

of the eigenvalues

$$\lambda_1^{(n)} \leq \cdots \leq \lambda_n^{(n)}$$

* Institute of Mathematics, University of Tsukuba, E-mail: tsuyoshi@math.tsukuba.ac.jp
† Institute of Mathematics, University of Tsukuba, E-mail: minami@sakura.cc.tsukuba.ac.jp
Received July 19, 2000.
Revised January 22, 2001.

of $Q^{(n)}$ converges weakly to the semicircle law with "radius" $2\sqrt{v}$ defined by

$$\mu_w^v(dx) = \frac{1}{2\pi v} 1_{[-\sqrt{4v},\sqrt{4v}]}(x)\sqrt{4v - x^2}\,dx. \tag{1}$$

Since then, many different approaches have been made to the semicircle law, revealing its various different aspects. Among these works, Chan [3], Rogers and Shi [8] considered fictitious time evolution of a Gaussian matrix ensemble, to obtain another proof of the convergence (in probability) to the semicircle law. (The idea of introducing a fictitious time evolution goes back to Dyson [4], and further discussed by McKean [6].) Namely, suppose each X_{kl} above not only is a mean zero Gaussian random variable, but also is the Ornstein-Uhlenbeck process $\{X_{kl}(t)\}_{t\geq 0}$ leaving its Gaussian distribution invariant. If we let $Q^{(n)}(t) = (X_{kl}(t)/\sqrt{n})_{k,l=1}^n$, then we have a matrix valued stationary Gaussian process, and the empirical distribution

$$\mu_t^{(n)} = \mu_t^{(n)}(dx) = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j^{(n)}(t)}(dx)$$

of the eigenvalues

$$\lambda_1^{(n)}(t) \leq \cdots \leq \lambda_n^{(n)}(t)$$

of $Q^{(n)}(t)$ also forms a stationary stochastic process $\{\mu_t^{(n)}\}$ taking values in the space $\mathscr{M}_1(\boldsymbol{R})$ of all probability measures on $\boldsymbol{R}$. It is easily seen that it has continuous sample paths, if $\mathscr{M}_1(\boldsymbol{R})$ is equipped with the topology of weak convergence. Then Chan, Rogers and Shi proved that as $n \to \infty$, the sequence of processes $\{\mu_t^{(n)}\}$, $n = 1, 2, \ldots$ converges in distribution to the trivial deterministic process $\{\mu_t\}$ such that $\mu_t = \mu_w^v$ for all $t \geq 0$. If we look at one-dimensional distributions of these processes, we obtain the convergence in probability of $\mu^{(n)} = \mu_0^{(n)}$ to the semicircle law μ_w^v.

Now the purpose of the present paper is twofold: First, we shall simplify the above summerized work by Chan, Rogers and Shi by directly deriving the stochastic differential equation satisfied by the empirical measure process $\{\mu_t^{(n)}\}_t$. In fact, Chan et al. (and also Dyson and McKean) investigated the stochastic differential equation satisfied by the eigenvalues $\lambda_j^{(n)}(t)$ of $Q^{(n)}(t)$, but because of the singularity of its coefficients, special effort was needed to prove the absense of collision between the eigenvalues. But it should be noted that the empirical measure process $\{\mu_t^{(n)}\}_t$, which is the object of our study, is well defined without caring the possible degeneracy of eigenvalues. On the other hand, the stochastic differential equation governing $\{\mu_t^{(n)}\}_t$ can be derived by simply applying Itô's formula to the trace of the resolvent of $Q^{(n)}(t)$. Secondly, we treat a wider class of Gaussian matrix ensembles so that it includes the so called GOE, GUE and GSE (see Mehta [7] for definitions), thus showing again the universal nature of Wigner's semicircle law.

Let us now turn to the precise formulation of our result. For $s = 0, 1, 2, 3$ and $1 \leq k \leq l$, let $X_{k,l}^{(s)}$ are independent real random variables obeying the Gaussian distribution of mean zero and variance $(\sigma_{kl}^{(s)})^2$. Let us further define $\{X_{k,l}^{(s)}(t)\}_{t \geq 0}$ to be the solution of the following stochastic differential equation:

$$(2) \qquad \begin{cases} dX_{k,l}^{(s)}(t) = -\dfrac{1}{2} X_{k,l}^{(s)}(t)\, dt + \sigma_{kl}^{(s)}\, dB_{k,l}^{(s)}(t) \\ X_{k,l}^{(s)}(0) = X_{k,l}^{(s)}, \end{cases}$$

where $B_{k,l}^{(s)}$'s are mutually independent 1-dimensional standard Brownian motions, which are also independent of $X_{k,l}^{(s)}$'s. Then each process $\{X_{k,l}^{(s)}(t)\}_{t \geq 0}$ is the stationary Ornstein-Uhlenbeck process which has the normal distribution $N(0, (\sigma_{kl}^{(s)})^2)$ as its invariant distribution. If we set, for $k > l$,

$$X_{k,l}^{(s)} := \begin{cases} X_{l,k}^{(0)} & \text{if } s = 0 \\ -X_{l,k}^{(s)} & \text{if } s = 1, 2, 3 \end{cases}$$

and

$$B_{k,l}^{(s)} := \begin{cases} B_{l,k}^{(s)} & \text{if } s = 0 \\ -B_{l,k}^{(s)} & \text{if } s = 1, 2, 3, \end{cases}$$

then we automatically have

$$X_{k,l}^{(s)}(t) = \begin{cases} X_{l,k}^{(0)}(t) & \text{if } s = 0 \\ -X_{l,k}^{(s)}(t) & \text{if } s = 1, 2, 3, \end{cases}$$

so that the $n \times n$ real matrix

$$X_n^{(s)}(t) := (X_{j,k}^{(s)}(t))_{1 \leq j,k \leq n}$$

is symmetric for $s = 0$ and skew-symmteric for $s = 1, 2, 3$.

Finally let

$$Q^{(n)}(t) := \frac{1}{\sqrt{n}} \sum_{s=0}^{3} X_n^{(s)}(t) e_s,$$

where e_s are 2×2 matrices defined by

$$e_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad e_1 := \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix}$$

$$e_2 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad e_3 := \begin{pmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}.$$

Here we consider e_s as if they were scalar, so that its product with any $n \times n$ matrix A actually means

$$Ae_s = \begin{pmatrix} a_{11}e_s & \dots & a_{1n}e_s \\ \vdots & & \vdots \\ a_{n1}e_s & \dots & a_{nn}e_s \end{pmatrix},$$

which is a $2n \times 2n$ matrix, each $a_{jk}e_s$ being a 2×2 block. With this definition, $Q^{(n)}(t)$ is a stationary stochastic process whose values are $2n \times 2n$ Hermitian, self-dual matrices in the sense that

$$(I_n e_2)Q^{(n)}(t)(I_n e_2)^T = Q^{(n)}(t),$$

where I_n is the $n \times n$ identity matrix and ^{T} denotes the transpose.

Now let

$$\lambda_1^{(n)}(t) \le \cdots \le \lambda_{2n}^{(n)}(t)$$

be the eigenvalues of $Q^{(n)}(t)$ and define their empirical distribution by

$$\mu_t^{(n)}(dx) = \frac{1}{2n}\sum_{j=1}^{2n} \delta_{\lambda_j^{(n)}(t)}(dx).$$

It is clear that $\{\mu_t^{(n)}\}_{t\ge 0}$ is a stationary stochastic process taking values in $\mathcal{M}_1(\boldsymbol{R})$. Since $Q^{(n)}(t)$ is continuous in t, $\mu_t^{(n)}$ is also continuous in t as we have suggested. This may be most easily seen by considering the Stieltjes transform of the measure $\mu_t^{(n)}(dx)$, which turns out to be equal to the trace of the resolvent of $Q^{(n)}(t)$. Namely if for z in the complex upper half-plane $\boldsymbol{H}$ we set

$$R_z^{(n)}(t) = (Q^{(n)}(t) - zI_{2n})^{-1},$$

then we have

$$\int_{\boldsymbol{R}} \frac{1}{x-z}\mu_t^{(n)}(dx) = \frac{1}{2n}\operatorname{Tr}(R_z^{(n)}(t)).$$

The right hand side is continuously dependent on the entries of the Hermitian matrix $Q^{(n)}(t)$, and hence is continuous in t. On the other hand, continuity in t of the left hand side implies that of $\mu_t^{(n)}$, because a sequence of probability measures $\{\nu_n\}_n$ converges weakly to a probability measure ν if and only if

$$\int \frac{\nu_n(dx)}{x-z} \longrightarrow \int \frac{\nu(dx)}{x-z}$$

for each $z \in \boldsymbol{H}$.

In order to state our result, let $\mathscr{P}_n$ be the probability distribution of the process $\{\mu_t^{(n)}\}$ induced on the space of all continuous functions from $[0, \infty)$ to $\mathscr{M}_1(\boldsymbol{R})$

$$\mathscr{C}_{\mathscr{M}} := C([0, \infty); \mathscr{M}_1(\boldsymbol{R})).$$

$\mathscr{C}_{\mathscr{M}}$ is equipped with the topology of uniform convergence on compacta. Let also δ_v be the probability measure concentrated on the constant path $\mu(t) \equiv \mu_w^v$. Now we can state our main result.

THEOREM. *Suppose $\sigma_{kl}^{(s)}$'s satisfy the following conditions:*
1. $\sup_{k \geq 1} \sigma_{kk}^{(0)} < \infty$;
2. $v = \sum_{s=0}^{3} (\sigma_{jk}^{(s)})^2$ *does not depend on j, k $(j < k)$.*

Then as $n \to \infty$, $\mathscr{P}_n$ converges weakly to δ_v on $\mathscr{C}_{\mathscr{M}}$.

Looking at the distribution of $\mu_0^{(n)}$, we obtain the following

COROLLARY. *Set $\mu^{(n)} = \mu_0^{(n)}$, namely let $\mu^{(n)}$ be the empirical distribution of the eigenvalues of $Q^{(n)} = \sum_{s=0}^{3} X_n^{(s)} e_s/\sqrt{n}$. Then as a sequence of $\mathscr{M}_1(\boldsymbol{R})$-valued random variables, $\{\mu^{(n)}\}$ converges in probability to the semicircle law μ_w^v. That is, if $\rho(\cdot, \cdot)$ is a metric on $\mathscr{M}_1(\boldsymbol{R})$ which generates the topology of weak convergence; then for any $\varepsilon > 0$,*

$$P(\rho(\mu^{(n)}, \mu_w^v) > \varepsilon) \to 0 \quad (n \to \infty).$$

Before closing this section, let us discuss some special cases. See Mehta [7] for details.

(i) Suppose $\sigma_{jk}^{(s)} = \sqrt{v}/2$ for all $j > k$ and $s = 0, 1, 2, 3$, with a positive constant v. In this case, every $X_{jk}^{(s)}$ $(j < k)$ are identically distributed according to $N(0, v/4)$. If we further suppose $\sigma_{kk}^{(s)} = \sqrt{v/2}$, then the resulting matrix ensemble coincides with the so called Gaussian symplectic ensemble (GSE).

(ii) Suppose $\sigma_{jk}^{(0)} = \sigma_{jk}^{(1)} = \sqrt{v/2}$ $(j < k)$ and $\sigma_{jk}^{(2)} = \sigma_{jk}^{(3)} = 0$. In this case, $Q^{(n)} = Q^{(n)}(0)$ is unitarily equivalent to

$$\frac{1}{\sqrt{n}} \left(\begin{array}{c|c} X_n^{(0)} + \sqrt{-1} X_n^{(1)} & O \\ \hline O & X_n^{(0)} + \sqrt{-1} X_n^{(1)} \end{array} \right),$$

so that each of the eigenvalues $\lambda_j^{(n)}$ of $Q^{(n)}$ is double. Hence if $\xi_j^{(n)}$ $(j = 1, \dots, n)$ are the eigenvalues of $(X_n^{(0)} + \sqrt{-1} X_n^{(1)})/\sqrt{n}$, we have

$$\mu^{(n)} = \frac{1}{n} \sum_{j=1}^{n} \delta_{\xi_j^{(n)}}.$$

Now if we further suppose $\sigma_{kk}^{(0)} = \sqrt{v}$, then the random Hermitian matrix $X_n^{(0)} + \sqrt{-1} X_n^{(1)}$ forms the so called Gaussian unitary ensemble (GUE).

(iii) Finally suppose $\sigma_{jk}^{(0)} = \sqrt{v}$ $(j < k)$ and $\sigma_{jk}^{(s)} = 0$ $(s = 1, 2, 3)$. Then $Q^{(n)}$ is unitarily equivalent to

$$\frac{1}{\sqrt{n}} \left(\begin{array}{c|c} X_n^{(0)} & O \\ \hline O & X_n^{(0)} \end{array} \right),$$

and again as in (ii), the empirical distribution of the eigenvalues of $Q^{(n)}$ is equal to the empirical distribution of the eigenvalues of the $n \times n$ matrix $X_n^{(0)}/\sqrt{n}$. If $\sigma_{kk}^{(0)} = \sqrt{2v}$, then the random real symmetric matrix $X_n^{(0)}$ forms the so called Gaussian orthogonal ensemble (GOE).

2. Proof of the Main Theorem

2.1. A Tightness Criterion

As is usual in the proof of limit theorems for stochastic processes, we shall first show the tightness of the sequence $\{\mathscr{P}_n\}_{n \geq 1}$ of probability measures on $\mathscr{C}_{\mathscr{M}}$, and then show the uniqueness of its limit point. For this purpose, we prepare a tightness criterion of a fairly general nature, which we have formulated after an analogous proposition in [5].

Let $f_0 \geq 0$ be a continuous real valued function on $\boldsymbol{R}$ which tends to infinity as $|x| \to \infty$. Let further $\{f_j\}_{j \geq 1}$ be a sequence of complex-valued bounded continuous functions on $\boldsymbol{R}$. We assume that the sequence $\{f_j\}_{j \geq 1}$ determines a probability measure on $\boldsymbol{R}$ in the sense that if $\mu, \nu \in \mathscr{M}_1(\boldsymbol{R})$, and if $\langle \mu, f_j \rangle = \langle \nu, f_j \rangle$ for every $j \geq 1$, then one has $\mu = \nu$. Finally let $\mathscr{C}_C$ be the space of all continuous complex-valued paths.

PROPOSITION 1. *Suppose that for $n \geq 1$ and for $\mathscr{P}_n$-almost every $\mu \in \mathscr{C}_{\mathscr{M}}$, $\langle \mu_t, f_0 \rangle = \int_{\boldsymbol{R}} f_0(x)\mu_t(dx)$ is finite and is continuous in t. Let P_n^j, $j \geq 0$, $n \geq 1$, be the image measure of $\mathscr{P}_n$ induced on $\mathscr{C}_C$ by the mapping $\mu. \mapsto \langle \mu., f_j \rangle$. If for each $j \geq 0$, the sequence $\{P_n^j\}_{n\geq 1}$ of probability measures is tight on $\mathscr{C}_C$, then $\{\mathscr{P}_n\}_n$ is tight on $\mathscr{C}_{\mathscr{M}}$.*

The proof of this proposition will be given in the Appendix.

Let us apply this criterion to $\{\mathscr{P}_n\}$ in our Theorem. If we let $f_0(x) = x^2$ and $f_j(x) = 1/(x - z_j)$ with a sequence $\{z_j\}$ which is dense in $\boldsymbol{H}$, then these functions satisfy the requirements stated just before Proposition 1. Since $\mathscr{P}_n$ is the distribution in $\mathscr{C}_{\mathscr{M}}$ of the process $\{\mu_t^{(n)}\}$, P_n^0 is that of the process

$$\langle \mu_t^{(n)}, f_0 \rangle = \frac{1}{2n}\sum_{j=1}^{2n}(\lambda_j^{(n)}(t))^2 = \frac{1}{2n}\operatorname{Tr}(Q^{(n)}(t))^2,$$

which is obviously finite and continuous in t. Hence the first assumption of the Proposition 1 is satisfied.

In order to verify the tightness of the sequence $\{P_n^j\}_{n\geq 1}$ for each $j \geq 0$, we shall derive the stochastic differential equation satisfied by the process $\{\langle \mu_t^{(n)}, f_j \rangle\}_{t\geq 0}$. But before doing so, let us prove that under the condition 1 of Theorem, one can assume $\sigma_{kk}^{(0)} = 0$ for all $k \geq 1$, which we shall do in order to simplify the calculation. For this purpose, let $\tilde{Q}^{(n)}(t)$ be the same matrix as $Q^{(n)}(t)$ except that all its diagonal entries are set to be zero, and let $\tilde{\mu}_t^{(n)}$ be the empirical measure of the eigenvalues of $\tilde{Q}_n(t)$. Then by the resolvent identity, we get for any $f_z(x) = 1/(x - z)$ with $z \in \boldsymbol{H}$,

$$\begin{aligned}
&\langle \mu_t^{(n)}, f_z \rangle - \langle \tilde{\mu}_t^{(n)}, f_z \rangle \\
&= \left| \int_{\boldsymbol{R}} \frac{\mu_t^{(n)}(dx)}{x - z} - \int_{\boldsymbol{R}} \frac{\tilde{\mu}_t^{(n)}(dx)}{x - z} \right| \\
&= \frac{1}{2n}\left|\operatorname{Tr}\{Q^{(n)}(t) - zI_{2n}\}^{-1} - \operatorname{Tr}\{\tilde{Q}^{(n)}(t) - zI_{2n}\}^{-1}\right| \\
&= \frac{1}{2n}\left|\frac{1}{\sqrt{n}}\operatorname{Tr}\{\operatorname{diag}(X_{11}^{(0)}(t), \ldots, X_{nn}^{(0)}(t))e_0 \cdot \{\tilde{Q}^{(n)}(t) - zI_{2n}\}^{-1}\{Q^{(n)}(t) - zI_{2n}\}^{-1}\}\right| \\
&\leq \frac{1}{2n}\left(\max_{1\leq j\leq n}\frac{1}{\sqrt{n}}|X_{jj}^{(0)}(t)|\right) \cdot \sqrt{2n} \cdot \|\{\tilde{Q}^{(n)}(t) - zI_{2n}\}^{-1}\{Q^{(n)}(t) - zI_{2n}\}^{-1}\| \\
&\leq \frac{1}{(\operatorname{Im} z)^2}\frac{1}{\sqrt{n}}\max_{1\leq j\leq n}|X_{jj}^{(0)}(t)|,
\end{aligned}$$

where $\|\cdot\|$ denotes the norm by the inner product $(A, B) = \mathrm{Tr}(AB^*)$ and * means the adjoint. Hence for any $T > 0$ and $\varepsilon > 0$,

$$
\begin{aligned}
P&\left(\sup_{0\le t\le T}|\langle\mu_t^{(n)}, f_z\rangle - \langle\tilde{\mu}_t^{(n)}, f_z\rangle| > \varepsilon\right) \\
&\le P\left(\frac{1}{(\mathrm{Im}\, z)^2}\frac{1}{\sqrt{n}}\max_{1\le j\le n}\sup_{0\le t\le T}|X_{jj}^{(0)}(t)| > \varepsilon\right) \\
&\le \sum_{j=1}^{n} P\left(\sup_{0\le t\le T}|X_{jj}^{(0)}(t)| > (\mathrm{Im}\, z)^2\sqrt{n}\varepsilon\right).
\end{aligned}
$$

In order to estimate the probability on the right hand side, let $\{Y(t)\}$ be the stationary Ornstein-Uhlenbeck process with the standard normal distribution as its stationary measure. Then the process $\{X_{jj}^{(0)}(t)\}$ has the same distribution as $\{\sigma_{jj}^{(0)}Y(t)\}$. On the other hand, $\{Y(t)\}$ is equivalent in distribution to the process $\{e^{-t}B(e^{2t})\}$, $\{B(t)\}$ being the standard Brownian motion. Hence letting $C = \sup_{j\ge 1}\sigma_{jj}^{(0)} < \infty$, we obtain the following bound:

$$
\begin{aligned}
\sum_{j=1}^{n} &P\left(\sup_{0\le t\le T}|X_{jj}^{(0)}(t)| > (\mathrm{Im}\, z)^2\sqrt{n}\varepsilon\right) \\
&\le nP\left(\sup_{0\le t\le T}|Y(t)| \ge \frac{(\mathrm{Im}\, z)^2}{C}\varepsilon\sqrt{n}\right) \\
&\le nP\left(\sup_{0\le t\le e^{2T}}|B(t)| \ge \frac{(\mathrm{Im}\, z)^2}{C}\varepsilon\sqrt{n}\right),
\end{aligned}
$$

which tends to zero as $n \to \infty$. From these considerations, it is now clear that the two sequences of processes $\{\mu_t^{(n)}\}$ and $\{\tilde{\mu}_t^{(n)}\}$ have the same limiting distribution (if any) in $\mathscr{C}_{\mathscr{M}}$.

2.2. Stochastic Differential Equations for $\langle\mu_t^{(n)}, f_0\rangle$ and the Tightness of $\{P_n^0\}_{n\ge 1}$

As was already noted, we have

$$
\langle\mu_t^{(n)}, f_0\rangle = \frac{1}{2n}\mathrm{Tr}(Q^{(n)}(t))^2 = \frac{1}{2n^2}\mathrm{Tr}(X_n(t))^2.
$$

Since $X_n(t) = \sum_{s=0}^{3} X_n^{(s)}(t)e_s$ is Hermitian and self-dual, $X_n^{(0)}(t)$ is real symmetric and $X_n^{(s)}(t)$ is real skew-symmetric for $s = 1, 2, 3$. Since we are assuming $X_{jj}^{(0)}(t) = 0$, it is not difficult to see that

$$\begin{aligned}\operatorname{Tr}(X_n(t)^2) &= \operatorname{Tr}\left\{\left(\sum_{s=0}^{3} X_n^{(s)}(t)e_s\right)^2\right\}\\ &= 2\operatorname{Tr}\left\{X_n^{(0)}(t)^2 - \sum_{s=1}^{3} X_n^{(s)}(t)^2\right\}\\ &= 4\sum_{s=0}^{3}\sum_{1\le j<k\le n} X_{jk}^{(s)}(t)^2.\end{aligned}$$

On the other hand, each $X_{jk}^{(s)}(t)$ satisfies the stochastic differential equation (2). Hence applying Itô's formula, we can proceed as follows:

$$\begin{aligned}d\langle \mu_t^{(n)}, f_0\rangle &= \frac{4}{2n^2}\sum_{s=0}^{3}\sum_{1\le j<k\le n}\{2X_{jk}^{(s)}\, dX_{jk}^{(s)} + (dX_{jk}^{(s)})^2\}\\ &= \frac{2}{n^2}\sum_{s=0}^{3}\sum_{1\le j<k\le n}\left[2X_{jk}^{(s)}\left\{-\frac{1}{2}X_{jk}^{(s)}\, dt + \sigma_{jk}^{(s)}\, dB_{jk}^{(s)}\right\} + (\sigma_{jk}^{(s)})^2\, dt\right]\\ &= -\frac{2}{n^2}\sum_{s=0}^{3}\sum_{1\le j<k\le n}(X_{jk}^{(s)})^2\, dt + \frac{2}{n^2}\sum_{1\le j<k\le n}\left(\sum_{s=0}^{3}(\sigma_{jk}^{(s)})^2\right) dt + dM_n(t)\\ &= -\frac{1}{2n^2}\operatorname{Tr} X_n(t)^2\, dt + \frac{2}{n^2}\frac{n(n-1)}{2}v\, dt + dM_n(t)\\ &= -\langle \mu_t^{(n)}, f_0\rangle\, dt + \left(1-\frac{1}{n}\right)v\, dt + dM_n(t),\end{aligned}$$

where we have used the second condition of Theorem and have set

$$\begin{aligned}M_n(t) &= \frac{4}{n^2}\int_0^t\sum_{s=0}^{3}\sum_{1\le j<k\le n}\sigma_{jk}^{(s)}X_{jk}^{(s)}\, dB_{jk}^{(s)}\\ &=: \frac{4}{n^2}\sum_{s=0}^{3}\sum_{1\le j<k\le n} M_{n,jk}^{(s)}(t).\end{aligned}$$

Having obtained the stochastic differential equation governing $\langle \mu_t^{(n)}, f_0\rangle$, let us now prove the tightness of the sequence of processes $\{\langle \mu_t^{(n)}, f_0\rangle\}_{n\ge 1}$.

Actually much stronger assertion holds:

PROPOSITION 2. *With probability one, we have* $\lim_{n\to\infty}\langle \mu_t^{(n)}, f_0\rangle = v$ *uniformly in* $t\in[0,T]$ *for any* $T>0$.

PROOF. For notational simplicity, set $Z_n(t) = \langle \mu_t^{(n)}, f_0 \rangle$. Then $Z_n(t)$ satisfies

$$Z_n(t) = Z_n(0) + \int_0^t \left\{ \left(1 - \frac{1}{n}\right) v - Z_n(s) \right\} ds + M_n(t).$$

By Doob's martingale inequality, the orthogonality of martingales $M_{n,jk}^{(s)}(t)$, $s = 0,1,2,3$, $j < k$, and $X_{jk}^{(s)} \sim N(0, (\sigma_{jk}^{(s)})^2)$, one shows for each $T > 0$ and $a > 0$,

$$\begin{aligned} P\left(\sup_{0 \le t \le T} M_n(t)^2 > a \right) &\le \frac{1}{a} E[M_n(T)^2] \\ &= \frac{1}{a} \frac{16}{n^4} \sum_{s=0}^{3} \sum_{1 \le j < k \le n} E\left[\int_0^T (\sigma_{jk}^{(s)} X_{jk}^{(s)}(t))^2 \, dt \right] \\ &= \frac{1}{a} \frac{16}{n^4} \sum_{1 \le j < k \le n} \sum_{s=0}^{3} (\sigma_{jk}^{(s)})^4 T \\ &\le 16 \frac{v^2 T}{a n^4} \frac{n(n-1)}{2} = \mathcal{O}\left(\frac{1}{a n^2} \right). \end{aligned}$$

Letting $a = n^{-\alpha}$ with $0 < \alpha < 1$, we have

$$\sum_n P\left(\sup_{0 \le t \le T} |M_n(t)|^2 > n^{-\alpha} \right) \le \text{const.} \sum_n \frac{1}{n^{2-\alpha}} < \infty,$$

and hence by the Borel-Cantelli's lemma, we get for any $T > 0$,

$$\sup_{0 \le t \le T} |M_n(t)|^2 = \mathcal{O}(n^{-\alpha}) \quad (n \to \infty)$$

with probability one.

Next we prove that $Z_n(0) \to v$ almost surely. For this purpose we note that

$$\begin{aligned} E[Z_n(0)] &= \frac{2}{n^2} \sum_{1 \le j < k \le n} \sum_{s=0}^{3} E[X_{jk}^{(s)}(0)^2] \\ &= \frac{2}{n^2} \sum_{1 \le j < k \le n} \sum_{s=0}^{3} (\sigma_{jk}^{(s)})^2 \\ &= \left(1 - \frac{1}{n}\right) v \to v, \end{aligned}$$

and that

$$
\begin{aligned}
E[(Z_n(0) - E[Z_n(0)])^2] &= \frac{4}{n^4} \sum_{1 \le j < k \le n} \sum_{s=0}^{3} E[(X_{jk}^{(s)}(0)^2 - (\sigma_{jk}^{(s)})^2)^2] \\
&= \frac{12}{n^4} \sum_{1 \le j < k \le n} \sum_{s=0}^{3} (\sigma_{jk}^{(s)})^4 \\
&\le \frac{6v^2}{n^2}.
\end{aligned}
$$

Since the right hand side is summable in n, we see that $Z_n(0) - E[Z_n(0)] \to 0$ almost surely.

Now let us fix an ω from our basic probability space for which the above two assertions on $M_n(t)$ and $Z_n(0)$ hold. Since $Z_n(t) \ge 0$, we see from the equation for $Z_n(t)$,

$$0 \le Z_n(t) \le Z_n(0) + vt + M_n(t).$$

Hence $Z_n(t)$, $n = 1, 2, \dots$, are uniformly bounded on each finite interval $[0, T]$. If we denote this bound by C_T, then the same equation shows that

$$|Z_n(t) - Z_n(s)| \le (2v + C_T)|t - s| + |M_n(t) - M_n(s)|$$

for $s, t \in [0, T]$. Since the sequence of functions $\{M_n(\cdot)\}$ are equi-continuous on $[0, T]$, we see from this inequality that $\{Z_n(\cdot)\}_n$ is also equi-continuous on $[0, T]$. By the Ascoli-Arzelà's theorem, any subsequence of $\{Z_n\}$ contains a further subsequence which converges uniformly to some $z(\cdot)$ on each finite interval $[0, T]$. This limit $z(\cdot)$ satisfies the equation

$$z(t) = v + \int_0^t (v - z(s))\, ds,$$

and hence we conclude $z(t) \equiv v$. □

2.3. The Stochastic Differential Equation for $\langle \mu_t^{(n)}, f_z \rangle$ and the Tightness of $\{P_n^j\}_{n \ge 1}$

In order to obtain the desired stochastic differential equation, we shall apply Itô's formula to

$$\langle \mu_t^{(n)}, f_z \rangle = \frac{1}{2n} \operatorname{Tr}(Q^{(n)}(t) - zI_{2n})^{-1} =: \frac{1}{2n} \operatorname{Tr} R(t),$$

where we set $R(t) = R_z^{(n)}(t) = (Q^{(n)}(t) - zI_{2n})^{-1}$. But for this purpose, we need some formula for the derivatives of the right hand side with respect to the entries of $Q^{(n)}(t)$. Namely let $Q = \sum_{s=0}^{3} Q^{(s)} e_s$ be a $2n \times 2n$ self-dual Hermitian matrix so that $Q^{(0)}$ is real symmetric, and $Q^{(s)}$ $(s = 1, 2, 3)$ are real skew-symmetric. If we let $R = (Q - zI_{2n})^{-1} = \sum_{s=0}^{3} R^{(s)} e_s$ for a $z \in \boldsymbol{H}$, then since R is also self-dual, $R^{(0)}$ is symmetric and $R^{(s)}$ $(s = 1,2,3)$ are skew-symmetric. Now it is easy to see for $k \neq l$, $1 \le k, l \le n$, that

$$\frac{\partial R}{\partial Q_{k,l}^{(0)}} = -R\{(E^{kl} + E^{lk})e_0\}R;$$

$$\frac{\partial R}{\partial Q_{k,l}^{(s)}} = -R\{(E^{kl} - E^{lk})e_s\}R, \quad s = 1,2,3,$$

and

$$\frac{\partial^2 R}{\partial (Q_{k,l}^{(0)})^2} = 2R\{(E^{kl} + E^{lk})e_0\}R\{(E^{kl} + E^{lk})e_0\}R;$$

$$\frac{\partial^2 R}{\partial (Q_{k,l}^{(s)})^2} = 2R\{(E^{kl} - E^{lk})e_s\}R\{(E^{kl} - E^{lk})e_s\}R, \quad s = 1,2,3,$$

where we have defined the matrix E^{kl} as $(E^{kl})_{ij} = \delta_{ik}\delta_{jl}$ for $i,j,k,l = 1,2,\ldots,n$.

Now remembering that $X_{kk}^{(s)}(t) = 0$ for all s and k, and that $X_{kl}^{(s)} = \pm X_{lk}^{(s)}$ according to $s = 0$ or $s \neq 0$, we apply Itô's formula to $\mathrm{Tr}\, R(t)$, to obtain

$$\begin{aligned}
& d\langle \mu_t^{(n)}, f_z \rangle \\
&= \frac{1}{2n} d\{\mathrm{Tr}\, R(t)\} \\
&= \frac{1}{2n} \sum_{s=0}^{3} \sum_{1 \le k < l \le n} \mathrm{Tr}\left\{\frac{\partial R(t)}{\partial Q_{k,l}^{(s)}}\right\} \frac{1}{\sqrt{n}}\, dX_{kl}^{(s)}(t) \\
&\quad + \frac{1}{4n} \sum_{s=0}^{3} \sum_{1 \le k < l \le n} \mathrm{Tr}\left\{\frac{\partial^2 R(t)}{\partial (Q_{k,l}^{(s)})^2}\right\} \frac{1}{n} (dX_{kl}^{(s)}(t))^2 \\
&= -\frac{1}{2n\sqrt{n}} \sum_{1 \le k < l \le n} \mathrm{Tr}[R(t)\{(E^{kl} + E^{lk})e_0\}R(t)]\left\{-\frac{1}{2} X_{kl}^{(0)}(t)\, dt + \sigma_{kl}^{(0)}\, dB_{kl}^{(0)}(t)\right\} \\
&\quad - \frac{1}{2n\sqrt{n}} \sum_{1 \le k < l \le n} \sum_{s=1}^{3} \mathrm{Tr}[R(t)\{(E^{kl} - E^{lk})e_s\}R(t)]\left\{-\frac{1}{2} X_{kl}^{(s)}(t)\, dt + \sigma_{kl}^{(s)}\, dB_{kl}^{(s)}(t)\right\} \\
&\quad + \frac{2}{4n^2} \sum_{1 \le k < l \le n} \mathrm{Tr}[R(t)\{(E^{kl} + E^{lk})e_0\}R(t)\{(E^{kl} + E^{lk})e_0\}R(t)](\sigma_{kl}^{(0)})^2\, dt \\
&\quad + \frac{2}{4n^2} \sum_{1 \le k < l \le n} \sum_{s=1}^{3} \mathrm{Tr}[R(t)\{(E^{kl} - E^{lk})e_s\}R(t)\{(E^{kl} - E^{lk})e_s\}R(t)](\sigma_{kl}^{(s)})^2\, dt
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4n\sqrt{n}} \sum_{s=0}^{3} \sum_{k,l=1}^{n} \operatorname{Tr}\{R(t)(X_{kl}^{(s)}(t)E^{kl}e_s)R(t)\}\, dt \\
&\quad - \frac{1}{2n\sqrt{n}} \sum_{s=0}^{3} \sum_{k,l=1}^{n} \operatorname{Tr}\{R(t)(\sigma_{kl}^{(s)}E^{kl}e_s\, dB_{kl}^{(s)})R(t)\} \\
&\quad + \frac{1}{2n^2} \sum_{s=0}^{3} \sum_{k,l=1}^{n} (\sigma_{kl}^{(s)})^2 \operatorname{Tr}\{R(t)(E^{kl}e_s)R(t)(E^{kl}e_s)R(t)\}\, dt \\
&\quad + \frac{1}{2n^2} \sum_{k,l=1}^{n} (\sigma_{kl}^{(0)})^2 \operatorname{Tr}\{R(t)(E^{kl}e_0)R(t)(E^{lk}e_0)R(t)\}\, dt \\
&\quad - \frac{1}{2n^2} \sum_{s=0}^{3} \sum_{k,l=1}^{n} (\sigma_{kl}^{(s)})^2 \operatorname{Tr}\{R(t)(E^{kl}e_s)R(t)(E^{lk}e_s)R(t)\}\, dt \\
&=: a_1^{(n)}(t)\, dt - dM_n(t) + a_2^{(n)}(t)\, dt + a_3^{(n)}(t)\, dt + a_4^{(n)}(t)\, dt.
\end{aligned}
$$

Let us examine $a_i^{(n)}(t)$, $i = 1, 2, 3, 4$, and $M_n(t)$ separately.

To begin with, we can rewrite the expression for $a_1^{(n)}(t)$ as follows:

$$
\begin{aligned}
a_1^{(n)}(t) &= \frac{1}{4n\sqrt{n}} \sum_{s=0}^{3} \operatorname{Tr}\{R(t)(X_n^{(s)}(t)e_s)R(t)\} \\
&= \frac{1}{4n} \operatorname{Tr}\{R(t)Q^{(n)}(t)R(t)\} \\
&= \frac{1}{4n} \{\operatorname{Tr} R(t) + z \operatorname{Tr}(R(t))^2\} \\
&= \frac{1}{2}\langle \mu_t^{(n)}, f_z\rangle + \frac{z}{2}\left\langle \mu_t^{(n)}, \frac{\partial}{\partial z} f_z \right\rangle.
\end{aligned}
$$

In treating $a_2^{(n)}(t), a_3^{(n)}(t)$ and $a_4^{(n)}(t)$, we need the following formula: for general complex $2n \times 2n$ matrices $A = \sum_{t=0}^{3} A^{(t)}e_t$ and $B = \sum_{t=0}^{3} B^{(t)}e_t$, one has

$$
\operatorname{Tr}\{(E^{\alpha\beta}e_s)A(E^{\gamma\delta}e_s)B\} = 2\left(A_{\beta\gamma}^{(s)}B_{\delta\alpha}^{(s)} - \sum_{t \neq s} A_{\beta\gamma}^{(t)}B_{\delta\alpha}^{(t)} \right), \quad s = 0, 1, 2, 3.
$$

We apply this formula to $A = R(t) = \sum_{s=0}^{3} R^{(s)}e_s$ and $B = (R(t))^2 = \sum_{s=0}^{3} (R^2)^{(s)}e_s$. If we note that $R_{kk}^{(s)} = (R^2)_{kk}^{(s)} = 0$ holds for $s = 1, 2, 3$ and $1 \leq k \leq n$, we get

$$\begin{aligned}
a_3^{(n)}(t) &= \frac{1}{2n^2}\sum_{k,l=1}^{n}(\sigma_{kl}^{(0)})^2\ \mathrm{Tr}\{(E^{kl}e_0)R(t)(E^{lk}e_0)(R(t))^2\} \\
&= \frac{2}{2n^2}\sum_{k,l=1}^{n}(\sigma_{kl}^{(0)})^2\left\{R_{ll}^{(0)}(R^2)_{kk}^{(0)} - \sum_{t\neq 0}R_{ll}^{(t)}(R^2)_{kk}^{(t)}\right\} \\
&= \frac{1}{n^2}\sum_{k,l=1}^{n}(\sigma_{kl}^{(0)})^2 R_{ll}^{(0)}(R^2)_{kk}^{(0)},
\end{aligned}$$

and

$$\begin{aligned}
a_4^{(n)}(t) &= \frac{-1}{2n^2}\sum_{s=1}^{3}\sum_{k,l=1}^{n}(\sigma_{kl}^{(s)})^2\ \mathrm{Tr}\{(E^{kl}e_s)R(t)(E^{lk}e_s)(R(t))^2\} \\
&= \frac{-2}{2n^2}\sum_{s=1}^{3}\sum_{k,l=1}^{n}(\sigma_{kl}^{(s)})^2\left\{R_{ll}^{(s)}(R^2)_{kk}^{(s)} - \sum_{t\neq s}R_{ll}^{(t)}(R^2)_{kk}^{(t)}\right\} \\
&= \frac{1}{n^2}\sum_{k,l=1}^{n}\left\{\sum_{s=1}^{3}(\sigma_{kl}^{(s)})^2\right\}R_{ll}^{(0)}(R^2)_{kk}^{(0)}.
\end{aligned}$$

Hence from the second condition of our Theorem,

$$\begin{aligned}
a_3^{(n)}(t) + a_4^{(n)}(t) &= \frac{v}{n^2}\left\{\sum_{k,l=1}^{n}R_{ll}^{(0)}(R^2)_{kk}^{(0)} - \sum_{k=1}^{n}R_{kk}^{(0)}(R^2)_{kk}^{(0)}\right\} \\
&= \frac{v}{n^2}\left(\frac{1}{2}\ \mathrm{Tr}\ R(t)\right)\left(\frac{1}{2}\ \mathrm{Tr}\ R(t)^2\right) - \frac{v}{n^2}\sum_{k=1}^{n}R_{kk}^{(0)}(R^2)_{kk}^{(0)} \\
&= v\langle \mu_t^{(n)}, f_z\rangle\left\langle \mu_t^{(n)}, \frac{\partial}{\partial z}f_z\right\rangle - a_5^{(n)}(t),
\end{aligned}$$

where we have set

$$a_5^{(n)}(t) =: \frac{v}{n^2}\sum_{k=1}^{n}R_{kk}^{(0)}(R^2)_{kk}^{(0)}.$$

Collecting these terms, we obtain

$$\begin{aligned}
(3)\quad d\langle \mu_t^{(n)}, f_z\rangle &= \left\{\frac{1}{2}\langle \mu_t^{(n)}, f_z\rangle + \frac{z}{2}\left\langle \mu_t^{(n)}, \frac{\partial}{\partial z}f_z\right\rangle + v\langle \mu_t^{(n)}, f_z\rangle\left\langle \mu_t^{(n)}, \frac{\partial}{\partial z}f_z\right\rangle\right\} dt \\
&\quad + (a_2^{(n)}(t) - a_5^{(n)}(t))\, dt + dM_n(t).
\end{aligned}$$

Now let P_n^z be the probability distribution of the process $\{\langle \mu_t^{(n)}, f_z \rangle\}_t$ induced on $\mathscr{C}_C$. Since the sequence of the probability distribution of $\langle \mu_0^{(n)}, f_z \rangle$ is tight, and since

$$\left| \frac{1}{2} \langle \mu_t^{(n)}, f_z \rangle + \frac{z}{2} \left\langle \mu_t^{(n)}, \frac{\partial}{\partial z} f_z \right\rangle + \frac{v}{2} \langle \mu_t^{(n)}, f_z \rangle \left\langle \mu_t^{(n)}, \frac{\partial}{\partial z} f_z \right\rangle \right|$$
$$\leq \frac{1}{2} \frac{1}{\operatorname{Im} z} + \frac{|z|}{2} \frac{1}{(\operatorname{Im} z)^2} + \frac{v}{2} \frac{1}{\operatorname{Im} z} \frac{1}{(\operatorname{Im} z)^2}$$

is bounded in n and t, the tightness of the sequence $\{P_n^z\}_n$ of probability measures on $\mathscr{C}_C$ is an immediate consequence of the following proposition.

PROPOSITION 3. *With probability one, we have*

$$\lim_{n \to \infty} a_2^{(n)}(t) = \lim_{n \to \infty} a_5^{(n)}(t) = 0$$

uniformly on $[0, \infty)$, *and*

$$\lim_{n \to \infty} M_n(t) = 0$$

uniformly on each interval $[0, T]$, $T > 0$.

PROOF. For a general complex $2n \times 2n$ matrix $A = \sum_{s=0}^{3} A^{(s)} e_s$, $A^{(s)} = (A_{jk}^{(s)})_{j,k=1}^n$, we have

$$A^* = A^{(0)*} e_0 - \sum_{s=1}^{3} A^{(s)*} e_s$$

and hence

$$\operatorname{Tr}(A^* A) = 2 \operatorname{Tr}\left(\sum_{s=0}^{3} A^{(s)*} A^{(s)} \right) = 2 \sum_{s=0}^{3} \sum_{j,k=1}^{n} |A_{jk}^{(s)}|^2.$$

If we apply this to $R = (Q - zI_{2n})^{-1}$ or R^2 with Q self-dual Hermitian and $\operatorname{Im} z > 0$, then

$$\sum_{s=0}^{3} \sum_{j,k=1}^{n} |R_{jk}^{(s)}|^2 = \frac{1}{2} \operatorname{Tr}(R^* R) \leq \frac{n}{(\operatorname{Im} z)^2}$$

and

$$\sum_{s=0}^{3}\sum_{j,k=1}^{n}|(R^2)_{jk}^{(s)}|^2=\frac{1}{2}\operatorname{Tr}((R^*)^2R^2)\le\frac{n}{(\operatorname{Im}z)^4}.$$

We now estimate, for a fixed $z\in\boldsymbol{H}$, $a_5^{(n)}(t)$ and $a_2^{(n)}(t)$ as follows:

$$\begin{aligned}
|a_5^{(n)}(t)| &\le \frac{v}{n^2}\sqrt{\sum_{k=1}^{n}|R_{kk}^{(0)}(t)|^2}\sqrt{\sum_{k=1}^{n}|(R^2)_{kk}^{(0)}|^2}\\
&\le \frac{v}{n^2}\sqrt{\frac{1}{2}\operatorname{Tr}\{R(t)^*R(t)\}}\sqrt{\frac{1}{2}\operatorname{Tr}\{(R(t)^2)^*(R(t)^2)\}}\\
&\le \frac{v}{n^2}\sqrt{\frac{n}{(\operatorname{Im}z)^2}}\sqrt{\frac{n}{(\operatorname{Im}z)^4}};
\end{aligned}$$

$$\begin{aligned}
|a_2^{(n)}(t)| &= \frac{1}{2n^2}\left|\sum_{s=0}^{3}\sum_{j,k=1}^{n}(\sigma_{jk}^{(s)})^2\operatorname{Tr}\{(E^{kl}e_s)R(t)(E^{kl}e_s)R(t)^2\}\right|\\
&= \frac{2}{2n^2}\left|\sum_{s=0}^{3}\sum_{j,k=1}^{n}(\sigma_{jk}^{(s)})^2\left\{R_{kj}^{(s)}(R^2)_{jk}^{(s)}-\sum_{t\ne s}R_{kj}^{(t)}(R^2)_{jk}^{(t)}\right\}\right|\\
&\le \frac{1}{n^2}\sum_{s=0}^{3}\sum_{j,k=1}^{n}(\sigma_{jk}^{(s)})^2\sum_{t=0}^{3}|R_{kj}^{(t)}(R^2)_{jk}^{(t)}|\\
&\le \frac{v}{n^2}\sum_{t=0}^{3}\sum_{j,k=1}^{n}|R_{kj}^{(t)}(R^2)_{jk}^{(t)}|\\
&\le \frac{v}{n^2}\sqrt{\frac{1}{2}\operatorname{Tr}\{R(t)^*R(t)\}}\sqrt{\frac{1}{2}\operatorname{Tr}\{(R(t)^2)^*(R(t)^2)\}}\\
&\le \frac{v}{n^2}\sqrt{\frac{n}{(\operatorname{Im}z)^2}}\sqrt{\frac{n}{(\operatorname{Im}z)^4}}.
\end{aligned}$$

Hence we arrive at

$$\sup_{t\ge 0}|a_2^{(n)}(t)|=\mathcal{O}(n^{-1});\quad \sup_{t\ge 0}|a_5^{(n)}(t)|=\mathcal{O}(n^{-1})$$

as $n\to\infty$.

In order to treat $M_n(t)$, we introduce the matrix valued Brownian motion $\tilde{B}(t)=\sum_{s=0}^{3}\tilde{B}^{(s)}(t)e_s$ with $\tilde{B}_{jk}^{(s)}(t)=\sigma_{jk}^{(s)}B_{jk}^{(s)}$. Then we get

$$
\begin{aligned}
dM_n(t) &= \frac{1}{2n\sqrt{n}} \operatorname{Tr}\{R(t)^2\, d\tilde{B}(t)\} \\
&= \frac{1}{n\sqrt{n}} \operatorname{Tr}\left\{ (R(t)^2)^{(0)}\, d\tilde{B}^{(0)}(t) - \sum_{s=1}^{3} (R(t)^2)^{(s)}\, d\tilde{B}^{(s)}(t) \right\} \\
&= \frac{2}{n\sqrt{n}} \sum_{s=0}^{3} \sum_{1 \le j < k \le n} (R(t)^2)^{(s)}_{jk} \sigma^{(s)}_{jk}\, dB^{(s)}_{jk},
\end{aligned}
$$

so that

$$
\begin{aligned}
E[|M_n(t)|^2] &= \frac{4}{n^3} \sum_{s=0}^{3} \sum_{1 \le j < k \le n} (\sigma^{(s)}_{jk})^2 \int_0^T |(R(t)^2)^{(s)}_{jk}|^2\, dt \\
&\le \frac{2v}{n^3} \int_0^T \frac{1}{2} \operatorname{Tr}\{(R(t)^*)^2 (R(t))^2\}\, dt \\
&\le \frac{2vT}{n^2 (\operatorname{Im} z)^4}.
\end{aligned}
$$

Hence by the martingale inequality,

$$
P\left(\sup_{0 \le t \le T} |M_n(t)|^2 > a \right) \le \frac{1}{a} E[|M_n(T)|^2] = \mathcal{O}\left(\frac{1}{an^2} \right).
$$

The rest is the same as in the proof of Proposition 2. □

2.4 Identification of the Limiting Process. Completion of the Proof of Theorem

The results of the previous two subsections show that the sequence $\{\mathscr{P}_n\}_n$ consisting of the probability distributions of the empirical measure process $\{\mu_t^{(n)}\}$ satisfies the conditions of Proposition 1, and hence is tight. To prove the weak convergence $\mathscr{P}_n \to \delta_v$ on $\mathscr{C}_{\mathscr{M}}$, let $\mathscr{P}_\infty$ be any weak limit of $\{\mathscr{P}_n\}$ along a subsequence $\{n'\}$. If we define, for each $T > 0$ and $z \in \boldsymbol{H}$,

$$
\begin{aligned}
\Phi_{T,z}(\mu.) = 1 \wedge \int_0^T \Bigg| &\langle \mu_t, f_z \rangle - \langle \mu_0, f_z \rangle \\
&- \int_0^t \left\{ \frac{1}{2} \langle \mu_s, f_z \rangle + \frac{z}{2} \left\langle \mu_s, \frac{\partial}{\partial z} f_z \right\rangle + v \langle \mu_s, f_z \rangle \left\langle \mu_s, \frac{\partial}{\partial z} f_z \right\rangle \right\} ds \Bigg| \, dt,
\end{aligned}
$$

then $\Phi_{T,z}(\cdot)$ is a bounded continuous functional on $\mathscr{C}_{\mathscr{M}}$.

Hence by the stochastic differential equation for $\langle \mu_t^{(n)}, f_z \rangle$ and Proposition 3,

$$\begin{aligned} E^{\mathscr{P}_\infty}[\Phi_{T,z}] &= \lim_{n'\to\infty} E^{\mathscr{P}_{n'}}[\Phi_{T,z}] \\ &= \lim_{n'\to\infty} E[\Phi_{T,z}(\mu^{(n')})] \\ &= \lim_{n'\to\infty} E\left[1 \wedge \int_0^T \left| M^{(n')}(t) + \int_0^t (a_2^{(n')}(s) - a_5^{(n')}(s))\, ds \right| dt \right] \\ &= 0. \end{aligned}$$

This shows that if we let $M(t,z;\mu) := \langle \mu_t, f_z \rangle$, then for $\mathscr{P}_\infty$-almost all $\mu \in \mathscr{C}_{\mathscr{M}}$, $M(t,z;\mu)$ satisfies the partial differential equation

$$\frac{\partial M}{\partial t} = \left(vM + \frac{1}{2} z \right) \frac{\partial M}{\partial z} + \frac{1}{2} M. \tag{4}$$

It was shown by Rogers and Shi [8] that for any $\mu_0 \in \mathscr{M}_1(\boldsymbol{R})$, the unique solution of this partial differential equation under the initial condition $M(0,z) = \langle \mu_0, f_z \rangle$ converges to

$$\frac{1}{2v}\left(-z + \sqrt{z^2 - 4v}\right) = \int_{\boldsymbol{R}} \frac{1}{x - z} \mu_w^v(dx)$$

as $t \to \infty$. Since the process $\{\mu_t\}$ is stationary under the probability measure $\mathscr{P}_\infty$, this shows that the process $\{M(t,z;\mu)\}$ is the trivial non-random process which equals to $\int (x-z)^{-1} \mu_w^v(dx)$ independently of $t \geq 0$. Since $z \in \boldsymbol{H}$ is arbitrary, we see $\mathscr{P}_\infty = \delta_v$. This completes the proof of our Theorem.

2.5 Almost Sure Convergence in the Path Space

Suppose $\sigma_{jk}^{(0)} = \sqrt{v}$ for $(j < k)$, $\sigma_{kk}^{(0)} = \sqrt{2v}$, and $\sigma_{jk}^{(s)} = 0$ for $s = 1,2,3$. Then as was noted in the introduction, $\mu_t^{(n)}$ is equal to the empirical distribution of the eigenvalues of $X_n^{(0)}(t)/\sqrt{n}$. On the other hand, $X_n^{(0)}(t)/\sqrt{n}$ satisfies the conditions of the theorem of Arnold and Wegmann which we quoted at the beginning of this paper. Hence we have the weak convergence $\lim_{n\to\infty} \mu_0^{(n)} = \mu_w^v$ with probability one. If we *use* this fact, then we can prove the following assertion, which is stronger than our main theorem.

PROPOSITION 4. *With probability one, we have the weak convergence* $\lim_{n\to\infty} \mu_t^{(n)} = \mu_w^v$ *uniformly in* $t \in [0,T]$, *for any* $T > 0$.

PROOF. With probability one, we can pick and fix an $\omega \in \Omega$ for which the following conditions hold:

(i) $\mu_0^{(n)}(\omega)(\cdot) \to \mu_w^v$;

(ii) the conclusion of Proposition 2 is valid;

(iii) the conclusion of Proposition 3 is valid for the choice of $z = z_j$, $j \geq 1$, where $\{z_j\}$ is supposed to be dense in $\boldsymbol{H}$.

Now let δ_n^ω be the probability measure on $\mathscr{C}_{\mathscr{M}}$ which is concentrated on the single path $\{\mu_t^{(n)}(\omega)\}_t$. Then (ii) means in particular that the image measures of δ_n^ω induced on $\mathscr{C}_C$ by the mapping $\mu. \mapsto \langle \mu., f_0 \rangle$, where $f_0(x) = x^2$, is tight. On the other hand, from (i), (ii) and the stochastic differential equation (4), we see that the sequence of functions $\langle \mu_t^{(n)}, f_{z_j} \rangle$, where $f_z(x) = 1/(x - z)$, is uniformly bounded and equi-continuous on each interval $[0, T]$, $T > 0$. Hence by Ascoli-Arzelà's theorem, the sequence of image measures of δ_n^ω under the mapping $\mu. \mapsto \langle \mu., f_{z_j} \rangle$ is tight. Hence by Proposition 1, the sequence $\{\delta_n^{(\omega)}\}_n$ of probability measures on $\mathscr{C}_{\mathscr{M}}$ is tight, or equivalently the sequence $\{\mu.^{(n)}\}_n$ of functions is relatively compact in $\mathscr{C}_{\mathscr{M}}$. Let $\mu.^{(\infty)}$ be any of its limit along a subsequence $\{n'\}$. Letting $n = n' \to \infty$ in the equation (4), we see that $M^{(\infty)}(t, z) := \langle \mu_t^{(\infty)}, f_z \rangle$ satisfies the partial differential equation (4) for $z = z_j$. Since $\{z_j\}$ is dense in $\boldsymbol{H}$, this is true for all $z \in \boldsymbol{H}$. But $\mu_0^{(\infty)} = \mu_w^v$ and $M(t, z) \equiv \langle \mu_w^v, f_z \rangle$ is a solution of (4), we must have $M^{(\infty)}(t, z) = \langle \mu_w^v, f_z \rangle$ for all $z \in \boldsymbol{H}$ and $t \geq 0$. Hence $\mu_t^{(\infty)} = \mu_w^v$ for all $t \geq 0$, completing the proof of Proposition 4. □

A Proof of the Tightness Criterion

In this appendix, we give a proof of Proposition 1.

By the tightness of the sequence $\{P_n^0\}_{n=1}^\infty$, we can choose, for each $\varepsilon > 0$, a compact subset K_0 of $\mathscr{C}_{\boldsymbol{R}}$ such that

$$\inf_{n \geq 1} P_n^0(K_0) \geq 1 - \frac{\varepsilon}{2}.$$

Then we have

$$A_T := \sup_{c(\cdot) \in K_0} \sup_{t \in [0, T]} |c(t)| < \infty$$

for any $T > 0$. On the other hand,

$$M^T := \{\mu \in \mathscr{M}_1(\boldsymbol{R}); \langle \mu, f_0 \rangle \leq A_T\}$$

is a compact subset of $\mathscr{M}_1(\boldsymbol{R})$ for any $T > 0$, and

$$\inf_{n \geq 1} \mathscr{P}_n(\mu_t \in M^T, \text{ for any } t \in [0, T] \text{ and } T \geq 0) \geq \inf_{n \geq 1} \mathscr{P}_n(K_0) \geq 1 - \varepsilon/2.$$

Next, for every $j \geq 1$, we can choose a compact subset K_j of $\mathscr{C}_C$ such that

$$\inf_{n\geq 1} P_n^j(K_j) \geq 1 - \frac{\varepsilon}{2^{j+1}}$$

because of the tightness of the sequence $\{P_n^j\}_{n=1}^{\infty}$. Let $\mathscr{K}$ be the Borel subset of $\mathscr{C}_{\mathscr{M}}$ defined by

$$\mathscr{K} := \bigcap_{T>0} \{\mu(\cdot) \in \mathscr{C}_{\mathscr{M}}; \mu_t \in M^T, t \in [0, T]\} \cap \bigcap_{j\geq 1} \{\mu(\cdot) \in \mathscr{C}_{\mathscr{M}}; \langle \mu(\cdot), f_j \rangle \in K_j\}.$$

Then it is easily seen that

$$\inf_{n\geq 1} \mathscr{P}_n(\mathscr{K}) \geq 1 - \varepsilon.$$

Now to finish the proof, we need to verify that the set $\mathscr{K}$ is compact in $\mathscr{C}_{\mathscr{M}}$.

For this purpose, let $\{\nu^{(n)}(\cdot)\}_{n=1}^{\infty}$ be any sequence in $\mathscr{K}$. Since, for every $j \geq 1$, the sequence $\{\langle \nu^{(n)}(\cdot), f_j \rangle\}_{n=1}^{\infty}$ is contained in K_j which is compact in $\mathscr{C}_C$, we can choose a subsequence $\{\nu^{(n_l)}(\cdot)\}_{l=1}^{\infty}$ of $\{\nu^{(n)}(\cdot)\}_{n=1}^{\infty}$ such that, for every j, there exists a $c^j(\cdot) \in K_j$ which satisfies

$$\langle \nu^{(n_l)}(\cdot), f_j \rangle \to c^j(\cdot) \quad (l \to \infty).$$

On the other hand, for every $t \geq 0$, $\{\nu^{(n)}(t)\}_{n=1}^{\infty}$ is a sequence in the compact set $M^t \subset \mathscr{M}_1(\boldsymbol{R})$, so we can choose a subsequence $\{n_{l_m(t)}\}_{m=1}^{\infty}$ (which depends on t) of $\{n_l\}_{l=1}^{\infty}$ and $\nu(t) \in \mathscr{M}_1(\boldsymbol{R})$ such that

$$\nu^{(n_{l_m})}(t) \to \nu(t).$$

Since we have

$$\langle \nu(t), f_j \rangle = \lim_{m\to\infty} \langle \nu^{(n_{l_m})}(t), f_j \rangle = c^j(t),$$

for all $j \geq 1$, according to the assumption for $\{f_j\}_j$, this $\nu(t)$ is uniquely determined by $\{n_l\}_{l=1}^{\infty}$ and does not depend on the choice of its subsequence. Hence for every $t \geq 0$, we have

$$\lim_{l\to\infty} \nu^{(n_l)}(t) = \nu(t)$$

and

$$\langle \nu(t), f_j \rangle = c^j(t).$$

Next let us prove that $\nu(t)$ is a continuous function of $t \geq 0$. In fact, let $t \geq 0$ and let $\{t_k\}_{k=1}^{\infty}$ be a sequence in $[0, \infty)$ which converges to t. Then there exists a $T \in (0, \infty)$ such that $t \in [0, T]$ and $t_k \in [0, T]$ for every $k \geq 1$. Since $\nu(t_k) \in M^T$

and M^T is compact in $\mathscr{M}_1(\boldsymbol{R})$, we can choose a subsequence $\{t_{k_m}\}_{m=1}^\infty$ and $\tilde{\nu}$ such that

$$\lim_{m\to\infty} \nu(t_{k_m}) = \tilde{\nu} \quad \text{in } \mathscr{M}_1(\boldsymbol{R}).$$

Hence, according to the continuity of $c^j(\cdot)$, we have, for every $j \geq 1$,

$$\langle \tilde{\nu}, f_j \rangle = \lim_{m\to\infty} \langle \nu(t_{k_m}), f_j \rangle = \lim_{m\to\infty} c^j(t_{k_m}) = c^j(t) = \langle \nu(t), f_j \rangle.$$

Since $\{f_j\}_{j=1}^\infty$ determines the probability measure uniquely,

$$\tilde{\nu} = \nu(t).$$

That is, the limit

$$\lim_{m\to\infty} \nu(t_{k_m}) = \nu(t)$$

is independent of the choices of the subsequence $\{k_m\}_{m=1}^\infty$, and one has

$$\lim_{k\to\infty} \nu(t_k) = \nu(t).$$

Thus we have proved that $\{\nu^{(n_l)}(t)\}_{l=1}^\infty$ converges to $\nu(t)$ for each $t \geq 0$.

Let us finally prove that this convergence is uniform in $t \in [0, T]$ for any $T > 0$. For this purpose, let ρ be a metric on $\mathscr{M}_1(\boldsymbol{R})$ which generates the topology of the weak convergence of probability measures. If the convergence were not uniform on some interval $[0, T]$, then we would have

$$\limsup_{l\to\infty} \sup_{t\in[0,T]} \rho(\nu^{(n_l)}(t), \nu(t)) > 0.$$

Then there would exist a $\delta > 0$, a subsequence $\{n_{l_m}\}_{m=1}^\infty$ of the sequence $\{n_l\}_{l=1}^\infty$, and a sequence $\{t_{n_{l_m}}\}_{m=1}^\infty$ in $[0, T]$ such that

$$\rho(\nu^{(n_{l_m})}(t_{n_{l_m}}), \nu(t_{n_{l_m}})) \geq \delta.$$

Furthermore, there exists a subsequence $\{n'_m\}_{m=1}^\infty$ of the subsequence $\{n_{l_m}\}$, $\tau \in [0, T]$, and $\mu^* \in M^T$ such that

$$\lim_{m\to\infty} t_{n'_m} = \tau \quad \text{and} \quad \lim_{m\to\infty} \nu^{(n'_m)}(t_{n'_m}) = \mu^*.$$

Thus letting $m \to \infty$, we have

$$\rho(\mu^*, \nu(\tau)) \geq \delta$$

This means that,

$$|\langle \mu^*, f_j \rangle - \langle \nu(\tau), f_j \rangle| > 0$$

for some $j \geq 1$. Denote by η the left hand side of this inequality. Then there exists an $m_0 \geq 1$ such that, for every $m \geq m_0$,

$$|\langle \nu^{(n'_m)}(t_{n'_m}), f_j \rangle - \langle \nu(t_{n'_m}), f_j \rangle| > \frac{\eta}{2}.$$

This implies

$$\limsup_{m \to \infty} \sup_{t \in [0,T]} |\langle \nu^{(n'_m)}(t), f_j \rangle - \langle \nu(t), f_j \rangle| \geq \frac{\eta}{2},$$

which contradicts the uniform convergence on $[0, T]$ of $\{\langle \nu^{(n_l)}(t), f_j \rangle\}_{l=1}^{\infty}$ to $c_j(t) = \langle \nu(t), f_j \rangle$.

References

[1] L. Arnold. On Wigner's semicircle law for the eigenvalues of random matrices. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, **19**:191–198, 1971.

[2] L. Arnold. Deterministic version of Wigner's semicircle law for the distribution of matrix eigenvalues. Linear Algebra and Appl., **13**(3):185–199, 1976.

[3] T. Chan. The Wigner semi-circle law and eigenvalues of matrix-valued diffusions. Probab. Theory Related Fields, **93**(2):249–272, 1992.

[4] F. J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. J. Math. Phys., **3**:1191–1198, 1962.

[5] J. Gärtner. On the McKean-Vlasov limit for interacting diffusions. Math. Nachr., **137**:197–248, 1988.

[6] H. P. McKean, Jr. Stochastic integrals. Academic Press, New York, 1969. Probability and Mathematical Statistics, No. 5.

[7] M. L. Mehta. Random matrices. Academic Press Inc., Boston, MA, second edition, 1991.

[8] L. C. G. Rogers and Z. Shi. Interacting Brownian particles and the Wigner law. Probab. Theory Related Fields, **95**(4):555–570, 1993.

[9] H. F. Trotter. Eigenvalue distributions of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdock, and Szegö. Adv. in Math., **54**(1):67–82, 1984.

[10] R. Wegmann. Bemerkungen zum Halbkreissatz für die Eigenwerte hermitescher Matrizen. Arch. Math. (Basel), **27**(5):535–538, 1976.

[11] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. Ann. of Math. (2), **67**:325–327, 1958.