Eigenvalue Statistics of the Real Ginibre Ensemble

Peter J. Forrester[1] and Taro Nagao[2]

[1]*Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia*
[2]*Graduate School of Mathematics, Nagoya University, Chikusa-ku, Nagoya 464-8602, Japan*
(Received 3 May 2007; published 3 August 2007)

The real Ginibre ensemble consists of random $N \times N$ matrices formed from independent and identically distributed standard Gaussian entries. By using the method of skew orthogonal polynomials, the general n-point correlations for the real eigenvalues, and for the complex eigenvalues, are given as $n \times n$ Pfaffians with explicit entries. A computationally tractable formula for the cumulative probability density of the largest real eigenvalue is presented. This is relevant to May's stability analysis of biological webs.

Dyson's threefold way [1] is a viewpoint on the foundations of random matrix theory, showing how consideration of time reversal symmetry leads to three classes of ensembles of relevance to quantum mechanics. The three ensembles are catalogued by the classes of unitary matrices which leave the ensemble invariant—orthogonal (time reversal symmetry is an involution), unitary (no time reversal symmetry), and symplectic (time reversal symmetry is an anti-involution). For an ensemble theory of Hermitian matrices, an equivalent characterization is that the matrix elements be real, complex, and real quaternion, respectively.

Both as a concept and as a calculational tool, the threefold way has been highly successful. As a concept, allowing for global symmetries in addition to that of time reversal gives a classification of the former in terms of the ten infinite families of matrix Lie algebras [2]. This classification now provides theoretical underpinning to fundamental phenomena in mesoscopic physics [3], disordered systems [4], and low energy QCD [5], in addition to the study of the statistical properties of quantum spectra for which it was originally intended. A good deal of the success relates to the matrix ensembles of the threefold way and its generalization being exactly solvable—analytic forms are available for all key statistical quantities, allowing for quantitative theoretical predictions.

Soon after the formulation of the threefold way, Ginibre [6] presented as a mathematical extension an analogous theory of non-Hermitian random matrices. The entries are taken to be either real, complex, or real quaternion. Like their Hermitian counterparts, it transpires that such random matrices have physical relevance.

Consider the complex case first. Then the joint eigenvalue probability density function (PDF) is proportional to

$$\prod_{l=1}^{N} e^{-|z_l|^2} \prod_{j<k}^{N} |z_k - z_j|^2, \qquad z_j := r_j e^{i\theta_j}. \quad (1)$$

This can be recognized as the Boltzmann factor for the two-dimensional one-component plasma in a disk at coupling $\beta = 2$, or the absolute value squared of the wave function for free fermions in a plane, subject to a perpendicular magnetic field and confined to the lowest Landau level [7]. In the study of chaotic dissipative quantum systems, the statistical properties of eigenvalues for certain model maps are well described by the corresponding statistical properties implied by this PDF [8].

In the case of real quaternion elements, the eigenvalues come in complex conjugate pairs. The eigenvalue PDF of the eigenvalues in the upper half plane is proportional to

$$\prod_{l=1}^{N} e^{-2|z_l|^2} |z_l - \bar{z}_l|^2 \prod_{j<k}^{N} |z_k - z_j|^2 |z_k - \bar{z}_j|^2 \quad (2)$$

Up to an extra one body factor $\prod_{l=1}^{N} |z_j - \bar{z}_j|$, the eigenvalue PDF of the eigenvalues in the upper half plane is proportional to the Boltzmann factor for the two-dimensional one-component plasma confined to a semidisk at coupling $\beta = 2$, bounded by a dielectric material of dielectric constant $\epsilon = 0$ along the straight edge [9,10].

Both joint PDFs for the complex and real quaternion cases are contained in Ginibre's paper [6]. However, in the case of real elements, it was not until a further 25 or so years later that the joint distribution was computed, first by Lehmann and Sommers [11], then by Edelman [12]. Part of the difficulty is that the joint PDF is not absolutely continuous. Rather, there is a nonzero probability that for N even (odd) there will be an even (odd) number of real eigenvalues for all even (odd) positive integers up to N. The final result is that for k real eigenvalues (k of the same parity as N), the joint PDF is

$$\frac{1}{2^{N(N+1)/4} \prod_{l=1}^{N} \Gamma(l/2)} \frac{2^{(N-k)/2}}{k![(N-k)/2]!} |\Delta(\{\lambda_l\}_{l=1,\ldots,k} \cup \{x_j \pm iy_j\}_{j=1,\ldots,(N-k)/2})| e^{-\sum_{j=1}^{k} \lambda_j^2/2} e^{\sum_{j=1}^{(N-k)/2}(y_j^2 - x_j^2)} \prod_{j=1}^{(N-k)/2} \operatorname{erfc}(\sqrt{2} y_j),$$

$$(3)$$

where $\Delta(\{z_p\}_{p=1,\ldots,m}) := \prod_{j<l}^{m}(z_l - z_j)$. Here $\lambda_l \in (-\infty, \infty)$ while $(x_j, y_j) \in \mathbb{R}^2_+$, $\mathbb{R}^2_+ := \{(x, y) \in \mathbb{R}^2 : y > 0\}$. Integrating (3) over $\{\lambda_l\} \cup \{x_j + iy_j\}$ gives the probability that there are precisely k eigenvalues. The simplest case in this regard is when $k = N$ (i.e., all eigenvalues real), and it is found that the sought probability is equal to $2^{-N(N-1)/4}$ [12]. For $k = 2$ an evaluation in terms of a single definite integral has been given recently in [13], while [14] reduces the calculation for general k down to an expression of the same computational complexity as our Eq. (11) below.

Perhaps the first applied study to draw attention to the eigenvalues of random real matrices was that of May [15], in the context of the stability of large biological webs. The very general setting [16] is to consider an N-dimensional vector $\vec{x}(t)$ with components specified as the solution of the coupled first order system $dx_i(t)/dt = F_i(\vec{x}(t))$, ($i = 1, \ldots, N$) for some nonlinear functions F_i. Assuming an isolated equilibrium solution $\vec{x}^0$, linearization about this point leads to the linear matrix differential equation

$$\frac{d\vec{y}(t)}{dt} = A\vec{y}(t), \qquad (4)$$

where A is an $N \times N$ matrix. The system is stable if all eigenvalues of A have a negative real part. To model the effect of random coupling between components on a stable equilibrium, the matrix A is written $A = -1_N + B$, where B is a dilute matrix (fraction $1 - c$ of its elements zero) with mean zero and variance s^2. The May criterion asserts that stability requires $s\sqrt{Nc} < 1$. Indeed in the case $c = 1$ this is consistent with limit theorems for the spectral radius of random real matrices proved subsequently [17,18]. Neural networks are further examples of complex webs to which such a random matrix based stability analysis is relevant [19–21]. The results obtained below allow the evaluation of the probability of stability in the borderline case of the May stability criterion,

$$s\sqrt{Nc} = 1. \qquad (5)$$

As with the matrix ensembles of Dyson's threefold way, all correlations and a number of key distributions for the complex and real quaternion Ginibre ensembles have been calculated exactly [9,22]. The Fourier transform of the two-point correlation (structure function) is a quantity of key importance to the plasma and fermion interpretation of (1), while the decay of the two-point function along the boundaries indicates general physical principles [nonzero dipole moment of the screening cloud in the case of (1), vanishing dipole moment for (2)]. Further, the distribution function for the spacing between eigenvalues in the bulk can be compared against data obtained from dissipative maps [8], while the density fluctuations in a large disk within the bulk indicate further general physical principles [23].

In contrast to the situation for (1) and (2), the correlations and distributions have not in general been computed

for the real Ginibre ensemble. Exceptions are the density of real eigenvalues [24]

$$\rho^r_{(1)}(x) = \frac{1}{\sqrt{2\pi}}\left[\frac{\Gamma(N - 1, x^2)}{\Gamma(N - 1)}\right.$$
$$\left. + \frac{2^{N/2-3/2}}{\Gamma(N - 1)}|x|^{N-1}e^{-x^2/2}\gamma\left(\frac{N - 1}{2}, \frac{x^2}{2}\right)\right] \quad (6)$$

with $\Gamma(p, x) := \int_x^\infty t^{p-1}e^{-t}dt$, $\gamma(p, x) := \int_0^x t^{p-1}e^{-t}dt$, and the density of complex eigenvalues [12]

$$\rho^c_{(1)}((x, y)) = \sqrt{\frac{2}{\pi}}\frac{\Gamma(N - 1, x^2 + y^2)}{\Gamma(N - 1)}ye^{2y^2}\text{erfc}(\sqrt{2}y). \quad (7)$$

Further, with $Z_{k,(N-k)/2}[u, u]$ denoting the canonical average of $\prod_{l=1}^{k} u(\lambda_l)\prod_{j=1}^{(N-k)/2} u(x_j + iy_j)$ with respect to (3), it has been shown in [25] (taking N even for definiteness) that

$$Z_N[u, u] := \sum_{k=0}^{N} Z_{k,(N-k)/2}[u, u]$$
$$= \frac{2^{-N(N+1)/4}}{\prod_{l=1}^{N}\Gamma(l/2)}\text{Pf}\left[\alpha_{j,k}(u) + \beta_{j,k}(u)\right]_{j,k=1,\ldots,N}, \qquad (8)$$

where, with $p_l(x)$ an arbitrary monic degree l polynomial and $z := x + iy$,

$$\alpha_{j,k}(u) = \int_{-\infty}^{\infty} dx u(x)\int_{-\infty}^{\infty} dy u(y)e^{-(x^2+y^2)/2}$$
$$\times p_{j-1}(x)p_{k-1}(y)$$
$$\times \text{sgn}(y - x) \qquad (9)$$

$$\beta_{j,k}(u) = 2i\int_{\mathbb{R}^2_+} dxdy u(z)e^{y^2-x^2}\text{erfc}(\sqrt{2}y)$$
$$\times [p_{j-1}(z)p_{k-1}(\bar{z}) - p_{k-1}(z)p_{j-1}(\bar{z})]. \qquad (10)$$

It is the purpose of this Letter to report that all the results (6)–(8) can be generalized, thereby fully exhibiting the real Ginibre ensemble as exactly solvable. For convenience it will be assumed throughout that N is even. We first observe that with the second argument u on the left-hand side of (8) replaced by an arbitrary function $v = v(x, y)$ the equality remains valid with u in $\beta_{j,k}(u)$ replaced by v. With $p_{N,2n}$ denoting the probability that $2n$ out of the N eigenvalues are real, it then follows by choosing $v = 1$, $u = \zeta$ that

$$\sum_{n=0}^{N/2}\zeta^n p_{N,2n} = \frac{1}{2^{N(N+1)/4}\prod_{l=1}^{N}\Gamma(l/2)}$$
$$\times \text{Pf}[\zeta\alpha_{j,k}(1) + \beta_{j,k}(1)]_{j,k=1,\ldots,N} \qquad (11)$$

[cf. Eq. (11) of Ref. [14]].

As is well known in random matrix theory [9,22], the correlations of a Pfaffian generating functional (8) are

themselves Pfaffians. However, in general this form involves the inverse of the matrix in (8) with $u = 1$. To make this explicit, one seeks to choose the polynomials $\{p_l(x)\}$ to have the skew orthogonality property

$$\alpha_{2j,2k}(1) + \beta_{2j,2k}(1) = \alpha_{2j-1,2k-1}(1) + \beta_{2j-1,2k-1}(1) = 0, \qquad \alpha_{2j-1,2k}(1) + \beta_{2j-1,2k}(1) = r_{j-1}\delta_{j,k}. \qquad (12)$$

Our key result is that the very simple choice

$$p_{2j}(x) = x^{2j}, \qquad p_{2j+1}(x) = x^{2j+1} - 2jx^{2j-1}, \qquad r_{j-1} = 2\sqrt{2\pi}\Gamma(2j-1) \qquad (13)$$

validates (12). With this established, and $q_{2j}(z) := -p_{2j+1}(z)$, $q_{2j+1}(z) := p_{2j}(z)$, one finds for the correlations between complex eigenvalues

$$\rho_{(n)}^c((x_1, y_1), \ldots, (x_n, y_n)) = \prod_{j=1}^{n}[2ie^{y_j^2 - x_j^2}\mathrm{erfc}\,(\sqrt{2}y_j)]\mathrm{Pf}\begin{bmatrix} S^c(\bar{z}_j, \bar{z}_k) & S^c(\bar{z}_j, z_k) \\ S^c(z_j, \bar{z}_k) & S^c(z_j, z_k) \end{bmatrix}_{j,k=1,\ldots,n}, \qquad (14)$$

where $S^c(w, z) := \sum_{j=1}^{N} p_{j-1}(w)q_{j-1}(z)/r_{[(j-1)/2]}$ and $z_j := x_j + iy_j$. In the case $n = 1$, the Pfaffian equals $S^c(\bar{z}_1, z_1)$ and (7) is reclaimed. In the case $n = 2$ the Pfaffian equals $S^c(\bar{z}_1, z_1)S^c(\bar{z}_2, z_2) + S^c(\bar{z}_1, z_2)S^c(z_1, \bar{z}_2) - S^c(\bar{z}_1, \bar{z}_2)S^c(z_1, z_2)$.

Similarly, the correlations between real eigenvalues are computed as

$$\rho_{(n)}^r(x_1, \ldots, x_n) = \mathrm{Pf}\begin{bmatrix} \mathrm{sgn}\,(x_j - x_k) + I^r(x_j, x_k) & S^r(x_j, x_k) \\ -S^r(x_k, x_j) & D^r(x_j, x_k) \end{bmatrix}_{j,k=1,\ldots,n} \qquad (15)$$

with $S^r(x, y) = \frac{1}{2}\frac{\partial}{\partial y}I^r(x, y)$, $D^r(x, y) = \frac{1}{2}\frac{\partial}{\partial x}S^r(x, y)$, and

$$I^r(x, y) = \sqrt{\frac{2}{\pi}}e^{-x^2/2}\sum_{k=0}^{N/2-1}\frac{x^{2k}}{(2k)!}\int_0^y e^{-u^2/2}u^{2k}du - (x \leftrightarrow y). \qquad (16)$$

In the case $n = 1$ this gives $\rho_{(1)}(x) = S^r(x, x)$ and (6) is reclaimed. In the limit $N \to \infty$ with x, y fixed (16) simplifies to

$$I^r(x, y) = \sqrt{\frac{2}{\pi}}\int_0^{y-x} e^{-u^2/2}du \qquad (17)$$

implying the correlations decay at a Gaussian rate. Integrating (6) over $x \in (-\infty, \infty)$ gives the mean number E_N of real eigenvalues, which is computed [24] to have the large N asymptotic form $\sqrt{2N/\pi}(1 - 3/8N - \ldots)$. The variance V_N of this same number is computed in terms of the two-point correlation according to $V_N = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\rho_{(2)}^{rT}(x, y) + E_N$, $\rho_{(2)}^{rT}(x, y) := \rho_{(2)}^r(x, y) - \rho_{(1)}^r(x)\rho_{(1)}^r(y)$. We read off from (15) that

$$\rho_{(2)}^{rT}(x_1, x_2) = -S^r(x_1, x_2)S^r(x_2, x_1)$$
$$- [\mathrm{sgn}\,(x_1 - x_2) + I^r(x_1, x_2)]D^r(x_1, x_2). \qquad (18)$$

The quantity $\rho_{(2)}^{rT}(x, y)/\rho_{(1)}(x)$ is integrable in y showing that for large N,

$$V_N \sim E_N[1 + \lim_{N\to\infty}[1/\rho_{(1)}(0)]$$
$$\times \int_{-\infty}^{\infty}\rho_{(2)}^{rT}(0, y)dy] = (2 - \sqrt{2})E_N, \qquad (19)$$

with the final equality making use of (15) and (17).

We draw attention to quantitatively similar results which hold for the zeros of the random polynomial $p(z) =$ $\sum_{j=0}^{N}\binom{N}{j}^{1/2}\alpha_j z^j$, where the α_j are independent and identically distributed real Gaussian random variables. This has the interpretation in quantum mechanics as a random superposition of states with spin $N/2$. Moreover, the function $p(e^{i\phi}\cot\theta/2)$ vanishes at the values of (θ_j, ϕ_j) on the sphere corresponding to the stereographic projection of the zeros z_j of $p(z)$, giving the Majorana parametrization [26]. The analogy with the present problem is that the mean number of real zeros is proportional to $\sqrt{N}$, as is the variance, and the correlations decay as Gaussians [27]. A distinction is the lack of a boundary for the eigenvalue support, which is distributed as a Cauchy distribution.

We remark, too, that although not reported on here, the correlations between real and complex eigenvalues can be written as a Pfaffian. Furthermore, we anticipate that the partially symmetric real Ginibre ensemble, introduced in [11], will also yield to the present strategy.

To leading order the support of the eigenvalue densities (6) and (7) is the disk $|z| = \sqrt{N}$, as is consistent with the formula (5) for the boundary of the May stability criterion (here $c = 1$ and $s = 1$; however, the variable s can be reinstated by scaling $z \mapsto z/s$ throughout). Setting $x = \sqrt{N} + X$ and taking $N \to \infty$ in (6) gives for the limiting edge profile

$$\rho_{(1)}^r(X) = \frac{1}{\sqrt{2\pi}}\left[\frac{1}{2}(1 - \mathrm{erf}\,\sqrt{2}X) + \frac{e^{-X^2}}{2\sqrt{2}}(1 + \mathrm{erf}\,X)\right]. \qquad (20)$$

For any fixed angle away from the real axis, as $N \to \infty$ the density of complex eigenvalues near the boundary of support is radially symmetric, and the same as that in the complex Ginibre ensemble. Writing the radius r as $r = \sqrt{N} + R$, for $N \to \infty$ this has the explicit form [28,29]

TABLE I. Tabulation of two probabilities, specified in the text, relating to the probability that all real eigenvalues of an $N \times N$ Gaussian real matrix, entries of mean zero, variance s^2, are less that $s\sqrt{N}$.

N	$\Pr(R_{s\sqrt{N}})$	$\Pr(\tilde{R}_{s\sqrt{N}})$
2	0.81444	0.737 579
4	0.793 864	0.756 706
6	0.784 485	0.762 255
8	0.778 838	0.764 193
10	0.774 963	0.76475
12	0.772 092	0.76469
14	0.769 855	0.76434
16	0.768 048	0.76385

$$\rho^c_{(1)}(R) = \frac{1}{\sqrt{2\pi}}(1 - \mathrm{erf}\, \sqrt{2}R), \qquad (21)$$

and is thus equal to twice the first term in (20).

Suppose now that the variance of the Gaussian entries is reinstated as the variable s^2. Let R_r be the event there are no real eigenvalues, or all real eigenvalues are less than r. By scaling of the eigenvalues, $\Pr(R_{s\sqrt{N}+sr})$ is independent of s, and for $N \to \infty$ it is an order 1 function of r. The latter can be written as an infinite sum over the limiting n-point edge correlations, or equivalently as a Fredholm determinant of the integral operator with kernel given by the edge limit of the general entry in (15). For r large one has $\lim_{N\to\infty} \Pr(R_{s\sqrt{N}+sr}) \sim 1 - \int_r^\infty \rho^r_{(1)}(X)dX$, showing that the corresponding PDF decays as a Gaussian. For general N a practical formula for computing this probability is in terms of the generating functional (8), $\Pr(R_{s\sqrt{N}+sr}) = Z_N[\chi_{\lambda \in (-\infty, s\sqrt{N}+sr)}, 1]$, where $\chi_A = 1$ if A is true, $\chi_A = 0$ otherwise, and with the polynomials in (9) and (10) chosen according to (13). Numerical values of $\Pr(R_{s\sqrt{N}})$ so computed are tabulated in Table I for successive even values of N. The quantity $\Pr(\tilde{R}_{s\sqrt{N}}) := [\Pr(R_{s\sqrt{N}}) - p_{N,0}]/(1 - p_{N,0})$, also listed in Table I, gives the probability that all real eigenvalues are less than $s\sqrt{N}$, given that there is at least one real eigenvalue. In the case $s^2 = 1/N$, this corresponds to the probability that all nonoscillatory solutions of the linear system (4) are stable, given that there is at least one nonoscillatory solution.

With this study, building on the contributions of Lehmann and Sommers [11], Edelman [12], Kanzieper and Akemann [14], and Sinclair [25], the problem began by Ginibre over 40 years ago of calculating the statistical properties of the eigenvalues of non-Hermitian real Gaussian matrices is solved. As a consequence, the distribution of the largest real eigenvalue is presented in a computable form. The largest real eigenvalue determines the stability of nonoscillatory solutions in May's [15] analysis of biological webs.

[1] F. J. Dyson, J. Math. Phys. Sci. **3**, 1199 (1962).
[2] M. R. Zirnbauer, J. Math. Phys. Sci. **37**, 4986 (1996).
[3] C. W. J. Beenakker, Rev. Mod. Phys. **69**, 731 (1997).
[4] K. Efetov, *Supersymmetry in Disorder and Chaos* (Cambridge University Press, Cambridge, England, 1997).
[5] J. J. M. Verbaarschot and T. Wettig, Annu. Rev. Nucl. Part. Sci. **50**, 343 (2000).
[6] J. Ginibre, J. Math. Phys. Sci. **6**, 440 (1965).
[7] P. J. Forrester, Phys. Rep. **301**, 235 (1998).
[8] F. Haake, *Quantum Signatures of Chaos* (Springer, New York, 2000), 2nd ed.
[9] P. J. Forrester, "Log-Gases and Random Matrices," www.ms.unimelb.edu.au/~matpjf/matpjf.html.
[10] E. R. Smith, J. Phys. A **15**, 3861 (1982).
[11] N. Lehmann and H.-J. Sommers, Phys. Rev. Lett. **67**, 941 (1991).
[12] A. Edelman, J. Multivariate Anal. **60**, 203 (1997).
[13] G. Akemann and E. Kanzieper, arXiv:math-ph/0703019v2.
[14] E. Kanzieper and G. Akemann, Phys. Rev. Lett. **95**, 230201 (2005).
[15] R. M. May, Nature (London) **238**, 413 (1972).
[16] R. E. McMurtrie, J. Theor. Biol. **50**, 1 (1975).
[17] Z. D. Bai, Ann. Probab. **25**, 494 (1997).
[18] S. Gemam, Ann. Probab. **14**, 1318 (1986).
[19] D. J. Albers and J. C. Sprott, J. Stat. Phys. **125**, 885 (2006).
[20] J. Feng, V. Jirsa, and M. Ding, Chaos **16**, 015109 (2006).
[21] K. Rajan and L. F. Abbott, Phys. Rev. Lett. **97**, 188104 (2006).
[22] M. L. Mehta, *Random Matrices* (Elsevier, New York, 2004), 3rd ed.
[23] Ph. A. Martin, Rev. Mod. Phys. **60**, 1075 (1988).
[24] A. Edelman, E. Kostlan, and M. Shub, J. Amer. Math. Soc. **7**, 247 (1994).
[25] C. D. Sinclair, arXiv:math-ph/0605006.
[26] J. H. Hannay, J. Phys. A **29**, L101 (1996).
[27] P. Bleher and X. Di, J. Stat. Phys. **88**, 269 (1997).
[28] P. J. Forrester and G. Honner, J. Phys. A **32**, 2961 (1999).
[29] E. Kanzieper, in *Frontiers in Field Theory*, edited by O. Kovras (Nova Science Publishers, New York, 2005), p. 23.