Two short pieces around the Wigner problem

Jean-Philippe Bouchaud and Marc Potters

Capital Fund Management, 23–25, rue de l’Université, 75007 Paris, France

E-mail: jean-philippe.bouchaud@cfm.fr

Received 6 September 2018, revised 12 November 2018
Accepted for publication 19 November 2018
Published 10 December 2018

Abstract
We revisit the classic Wigner semi-circle from two different angles. One consists in studying the Stieltjes transform directly on the real axis, which does not converge to a fixed value but follows a Cauchy distribution that depends on the local eigenvalue density. This result was recently proven by Aizenman and Warzel for a wide class of eigenvalue distributions. We shed new light onto their result using a Coulomb gas method. The second angle is to derive a Langevin equation for the full (matrix) resolvent, extending Dyson’s Brownian motion framework. The full matrix structure of this equation allows one to recover known results on the overlaps between the eigenvectors of a fixed matrix and its noisy counterpart, in particular in the ‘spiked’ case, for which we formulate a new conjecture.

Keywords: random matrix theory, Dyson Brownian motion, resolvent

(Some figures may appear in colour only in the online journal)

1. Introduction

Wigner’s semi-circle law is certainly the most famous in random matrix theory. There are many different ways to obtain it, each of them shedding a different light on the result. In the present paper written for this special issue on random matrix theory, we revisit once again the Wigner problem. In the first part of this paper, we show that the semi-circle can be obtained by studying the (power-law) tail of the normalized trace of the resolvent $g(x)$ on the real axis, rather than ‘just above’ the real axis in the complex plane. On the support on the eigenvalue density, $g(x)$ does not converge to a constant value (the limiting Stieltjes transform $g_0(x)$ is ill-defined for such $x$) but rather converges in probability to a Cauchy law, a result obtained by Fyodorov and collaborators [1–3] for the GOE and the GUE, and later proven by Aizenman and Warzel [4] for a very large class of point processes. We re-derive this result

\[ \text{Note that this result had in fact previously appeared in [26].} \]
using a Coulomb gas method, which amounts to study the response of such a gas to a singular perturbation. The second part of our paper follows Dyson’s Brownian motion framework to derive a Langevin equation for the full (matrix) resolvent. This Langevin equation becomes deterministic in the large \( N \) limit; its trace leads to the well known Burgers equation the solution of which again produces Wigner’s semi-circle. But the full matrix structure allows one to characterize the evolution of the eigenvectors as well, and recover known results on the overlaps between the eigenvectors of a fixed matrix \( \mathbf{C} \) and its noisy counterpart \( \mathbf{C} + \mathbf{W} \), where \( \mathbf{W} \) is the Wigner–Dyson Brownian random matrix. The case of an isolated (spike) eigenvalue is also discussed within the same framework.

2. The resolvent of a Wigner matrix on the real axis is Cauchy distributed

2.1. Schur elimination and the Cauchy fixed point

A standard way to approach the distribution of eigenvalues of random matrices is to write a recursion relation for the elements of the resolvent matrix \( \mathbf{G} \), defined as

\[
\mathbf{G}(z) = (z \mathbf{I} - \mathbf{W})^{-1},
\]

where \( z \) is in the complex plane, but outside the real axis to avoid the poles of \( \mathbf{G} \) (i.e. the eigenvalues of \( \mathbf{W} \)). The standard Schur relation then allows one to relate the elements of the resolvent matrix for a problem of size \( N \) and the same problem with one row and one column added to the matrix \( \mathbf{W} \). Denoting conventionally be ‘0’ the index of the added row and column, one readily finds:

\[
\frac{1}{G_{00}^{(N+1)}} = z - W_{00} - \sum_{i,j=1}^{N} W_{0i} G_{ij}^{(N)} W_{j0}.
\]

(2.2)

For Wigner random matrices with independent entries of order \( N^{-1/2} \), one can further argue that the contributions of terms with \( i \neq j \) in the above formula are negligible in the large \( N \) limit, leading to:

\[
\frac{1}{G_{00}^{(N+1)}} \approx z - \sum_{i=1}^{N} W_{0i}^2 G_{ii}^{(N)}.
\]

(2.3)

A crucial point, which makes this formula useful, is that the new matrix elements \( W_{0i} \) and the resolvent elements \( G_{ii}^{(N)} \) are completely independent.

As recalled above, the standard route is to study the above iteration for \( z \) outside the real axis, in which case the diagonal elements of \( \mathbf{G} \) converge, for large \( N \), to the normalized trace (or Stieltjes transform)

\[
g_0(z) := \lim_{N \to \infty} \frac{1}{N} \text{Tr} \mathbf{G}(z),
\]

(2.4)

where \( g_0(z) \), is the solution of

\[
\frac{1}{g(z)} = z - \sigma^2 g(z),
\]

(2.5)

where \( \sigma^2 := NV[W_0] \). One then recovers the classic result, from which the semi-circle law ensues:
\[ g_0(z) = \frac{1}{2\pi^2} \left[ z \pm \sqrt{z^2 - 4\sigma^2} \right] \rightarrow \rho(x) = \frac{1}{\pi} \lim_{\epsilon \to 0} \text{Im} \ g_0(x - i\epsilon) = \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - x^2}. \]  

(2.6)

For \( \varepsilon = x \) real and within Wigner’s band \([-2\sigma, 2\sigma]\), \( g(x) \) cannot be well defined, because by definition \( x \) is then always very close to a pole of \( G \). An idea, proposed in the context of Lévy matrices in [5], is to turn this predicament on its head and actually exploit the divergence of \( g(x) \) when \( x \) is equal to any eigenvalue of \( W \). In a hand-waving manner, the probability that the difference \( d_i = |x - \lambda_i| \) between \( x \) and a given eigenvalue \( \lambda_i \), is very small is:

\[ P[d_i < \varepsilon/N] = 2\varepsilon\rho(x), \]  

(2.7)

where \( \rho(x) \) is the normalized density of eigenvalues around \( x \). But as \( \varepsilon \to 0 \), the resolvent becomes dominated by a unique contribution—that of the \( \lambda_i \) term. In other words, \( g(x) \approx \pm (Nd_i)^{-1} \), and therefore

\[ P[|g| > \varepsilon^{-1}] = P[d_i < \varepsilon/N] = 2\varepsilon\rho(x). \]  

(2.8)

Hence, the tail of the distribution of non-self averaging Stieltjes transform \( g \) should decay precisely as \( \rho(x)/g^2 \). Studying this tail allows one to extract the eigenvalue density \( \rho(x) \) while working directly on the real axis. This was the strategy used to [5] to obtain the eigenvalue density of Lévy matrices (see below), a result later revisited in [6] and rigorously established by Ben Arous and Guionnet in [7]. Here, we want to revisit this issue in the standard Wigner case, in order to shed light on an admittedly weird strategy that Ben Arous and Guionnet ‘unfortunately (could not) make sense of’ [7].

The first remark is that for a rotationally invariant problem, the distribution of a randomly chosen diagonal element of the resolvent (say \( G_{00} \)) is the same as the distribution \( P(g) \) of its normalized trace. Therefore, equation (2.3) can be interpreted as giving the evolution of \( P(g) \) itself, i.e.:

\[ P^{(N+1)}(g) = \int_{-\infty}^{+\infty} dg' P^{(N)}(g') \delta \left( g - \frac{1}{x - \sigma^2 g^2} \right), \]  

(2.9)

where we have used the fact that for large \( N \), \( \sum_{i=0}^{N-1} W_{0i} G_{ii}^{(N)} \sim \sigma^2 g^{(N)} \). Now, this functional iteration admits the following Cauchy distribution as a fixed point [8]:

\[ P^\infty(g) = \frac{\rho(x)}{(g - \frac{1}{x - \sigma^2})^2 + \pi^2 \rho(x)}. \]  

(2.10)

This simple result, that the resolvent of a Wigner matrix on the real axis is a Cauchy variable, is illustrated in figure 1 and calls for several comments. First, one finds that \( P^\infty(g) \) indeed behaves as \( \rho(x)/g^2 \) for large \( g \), as expected. Second, it would have been entirely natural to find a Cauchy distribution for \( g \) had the eigenvalues been independent. Indeed, since \( g \) is then the sum of \( N \) random variables (i.e. the \( 1/d_i \)'s) distributed with an inverse square power, the generalized CLT predicts that the resulting sum is Cauchy distributed. In the present case, however, the eigenvalues are strongly correlated—in fact the spectrum is so rigid that the approximation \( \lambda_i - \lambda_j \approx (i - j)/(N\rho(x)) \) holds locally to a good approximation. In this case, the Cauchy distribution for \( g \) was actually derived by Fyodorov and collaborators for the GOE and GUE ensembles, using rather specific random matrix theory techniques [1–3]. It was recently proven by Aizenman and Warzel [4] that the Cauchy distribution is in fact super-universal and holds not only for all Coulomb gas models, for arbitrary values of \( \beta \), but in
fact for a much wider class of point processes on the real axis. The gist of the argument of Aizenman and Warzel is summarized in appendix A. In the next subsection, we want to give a physicist’s approach to the problem for Coulomb gas models, where we recover the super-universal Cauchy distribution, that is also valid for an arbitrary confining potential. We believe that the direct calculation, using a saddle point method, is quite interesting in its own right and could be used to obtained some refined results at finite $N$.

2.2. From Coulomb to Cauchy

Let us consider the resolvent of a $\beta$-ensemble matrix $W$ on the real axis:

$$g(x) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{x - \lambda_i}$$

(2.11)

where $\lambda_i$ are the eigenvalues of $W$. The joint distribution of the $\lambda_i$'s is well known to be:

$$P(\{\lambda_i\}) = Z \prod_{i<j} |\lambda_i - \lambda_j|^{\beta} \exp \left[ -\frac{N\beta}{2} \sum_j V(\lambda_j) \right].$$

(2.12)

The case $\beta = 0$ corresponds to independent (Poisson) random eigenvalues for which, as mentioned above, the Cauchy result is a trivial consequence of the generalized CLT. The case $\beta = \infty$ corresponds to a perfectly periodic ‘crystal’ of eigenvalues, for which an explicit calculation is also possible, see appendix B.

The potential $V(x)$ can be any confining potential. To simplify the discussion we start by considering the GxE potential

$$V(x) = \frac{x^2}{2}$$

and introduce later an arbitrary potential.\footnote{Note that we set $\sigma^2 = 1$, so that the spectrum of $W$ is $[-2, 2]$.}

Figure 1. Numerical simulation of the law of $g(x)$ for a GOE matrix. For a fixed $N = 5000$ GOE matrix, we have sampled 4000 times the distribution of $g(x + u/\sqrt{N})$, with $x = 1$ and $u$ distributed uniformly between $-1$ and $1$. Plotted is the left and right sample cumulative probability with the analytical result for the corresponding Cauchy distribution.
Let us fix an arbitrary value of $x$ within the spectrum $[-2, 2]$ and study the characteristic function of the distribution of $g$:

$$
\hat{P}(k) = \int \prod \, d\lambda_i \mathcal{P}(\{\lambda_i\}) e^{\frac{i}{\beta} \sum_{i=1}^{N} \frac{1}{x - \lambda_i}}. 
$$

(2.13)

Introducing a density field $\rho(\lambda)$ and neglecting the entropy term, one finds (see [9] for a detailed account):

$$
\hat{P}(k) = Z \int \mathcal{D}\rho \mathcal{N}^{-\frac{N}{2}} \sum_{\lambda' \lambda''} \rho(\lambda') \rho(\lambda'') \log |\lambda' - \lambda''| - \frac{1}{4} \sum_{\lambda'} \rho(\lambda') \lambda'^2 + i k \sum_{\lambda'} \frac{\lambda' - x}{(x - \lambda')^2}.
$$

(2.14)

The saddle point equation on $\rho$ then reads:

$$
\int d\lambda' \rho(\lambda') \log |\lambda - \lambda'| - \frac{1}{4} \lambda^2 + i \hat{k} \frac{1}{x - \lambda} + K = 0
$$

(2.15)

where $K$ is the Lagrange multiplier insuring that $\rho(x)$ is normalized, and $\hat{k} = N^{-2} k / \beta$. Taking the derivative of this equation w.r.t. $\lambda$ yields:

$$
\int d\lambda' \rho(\lambda') \frac{1}{\lambda - \lambda'} = \frac{1}{2} \lambda - i \hat{k} \frac{1}{(x - \lambda)^2}.
$$

(2.16)

Note that since typically the closest eigenvalue is at a distance $1/N$ from $x$, the last term is $O(1)$ there, and cannot be neglected even if $\hat{k} \sim N^{-2}$.

For $k = 0$, the solution of equation (2.16) is the familiar Wigner distribution:

$$
\rho_0(\lambda) = \frac{1}{2\pi} \sqrt{4 - \lambda^2}.
$$

(2.17)

Since the equation for $\rho$ is linear, the solution for $k \neq 0$ simply reads:

$$
\rho(\lambda) = \rho_0(\lambda) + \delta \rho(\lambda); \quad \delta \rho(\lambda) = -i \hat{k} \delta(\lambda - x).
$$

(2.18)

This corresponds to a shift of the unperturbed eigenvalues $\lambda_0$ by a singular quantity $\delta \lambda = -i \hat{k} \delta(\lambda - x) / \rho_0(x)$. So, although this solution is valid in the continuum limit, we cannot use it directly for finite $N$. We need to ‘zoom’ into the Dirac delta function to resolve the shift on atomic distances. Nevertheless, we will later make use of the above result to fix the asymptotic behaviour of the Stieltjes transform of the saddle-point solution:

$$
\delta \bar{g}(z) := \int d\lambda \frac{\delta \rho(\lambda)}{z - \lambda} \approx -i \hat{k} \frac{1}{(x - z)^2}.
$$

(2.19)

when $|x - z|$ is large enough.

In order to make progress, we write the discrete analogue of the saddle-point equation (2.16) as

$$
\frac{1}{N} \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} - \frac{1}{2} \lambda_i^2 = -i \hat{k} \frac{1}{(x - \lambda_i)^2}.
$$

(2.20)

Multiplying both sides by $(N(z - \lambda_i))^{-1}$ and summing over $i$ leads, after a few standard manipulations, to an equation for the Stieltjes transform of the saddle-point solution, defined as:
\[ \tilde{g}(z) := \frac{1}{N} \sum_{i} \frac{1}{z - \lambda_i}, \]  
which reads (neglecting a \(1/N\) contribution):
\[ \tilde{g}^2(z) - V'(z)\tilde{g}(z) + P(z) = -2i\tilde{\kappa}\partial_z \left[ \frac{\tilde{g}(z) - \tilde{g}(x)}{z - x} \right], \]  
where we have introduced a more general confining potential \(V(z)\) and associated function \(P(z)\). For the standard case, \(V(z) = z^2/2\) and \(P(z) = 1\).

Now, the strange thing that happens is that when \(z\) is the vicinity of \(x\), the right hand side remains of order unity in the limit \(k \to 0\). More precisely, the eigenvalue density perturbation turns out to take the following scaling form:
\[ \delta \rho(\lambda) = i\zeta F \left( \frac{x - \lambda}{\sqrt{|k|}} \right), \quad (\tilde{k} \to 0), \]  
where \(\zeta = \text{sign}(k)\) and \(F(u)\) is a certain odd function of \(u\). This means that the extra (imaginary) ‘charge’ located at \(\lambda = x\) substantially affects the Coulomb gas density in a neighbourhood of size \(\sqrt{|k|} \sim N^{-1}\) around \(x\), but has a negligible influence at larger distances.

From our scaling ansatz one obtains, after setting \(\lambda = x - \sqrt{|k|}u\) and \(z = x - \sqrt{|k|}y\):
\[ \delta \tilde{g}(z) = \int d\lambda \frac{\delta \rho(\lambda)}{z - x + \lambda} \approx i\zeta \int du \frac{F(u)}{u - y} + O(\tilde{k}). \]  
Inserting such a scaling form into the right hand side of equation (2.22) yields a result independent of \(\tilde{k}\) in the scaling regime:
\[ -i\tilde{\kappa}\partial_z \left[ \frac{\delta \tilde{g}(z) - \delta \tilde{g}(x)}{z - x} \right] = \partial_y \left[ \frac{\Gamma(0) - \Gamma(y)}{y} \right] + O(\tilde{k}), \]  
where
\[ \Gamma(y) := \int du \frac{F(u)}{u - y}. \]  
Now, we re-write equation (2.22) in terms of \(\tilde{g}(z) = \tilde{g}_0(z) + \delta \tilde{g}(z)\), where \(\tilde{g}_0(z)\) is the Stieltjes transform of the \(k = 0\) saddle-point, solution of:
\[ \tilde{g}_0^2(z) - V'(z)\tilde{g}_0(z) + P(z) = 0 \quad \Rightarrow \tilde{g}_0(z) = \frac{1}{2} \left[ V'(z) \pm \sqrt{V'^2(z) - 4P(z)} \right], \]  
corresponding to an unperturbed eigenvalue density
\[ \rho_0(\lambda) = \frac{1}{2\pi} \sqrt{4P(\lambda) - V'^2(\lambda)}. \]  
Neglecting terms of order \(\tilde{k}\), the equation for \(\delta \tilde{g}(z)\) then reads:
\[ P(z) := \sum (V(z) - V'(\lambda_i))/(z - \lambda_i), \text{ it is a polynomial of degree } n - 1 \text{ when } V(z) \text{ is a polynomial of degree } n, \text{ see e.g. } [14]. \]
\[
\delta \bar{g}^2(z) + (2\bar{g}_0(z) - V'(z)) \delta \bar{g}(z) = 2\partial_y \left[ \frac{\Gamma(0) - \Gamma(y)}{y} \right]
\]

(2.29)

or, for \( z = x - \sqrt{\hat{k}}y + i\zeta \) and \( x \) inside the eigenvalue spectrum,

\[
-\frac{1}{2} \Gamma^2(y) + \pi \rho_0(x) \Gamma(y) = \partial_y \left[ \frac{\Gamma(0) - \Gamma(y)}{y} \right].
\]

(2.30)

One can immediately deduce from this ODE that the asymptotic behaviour of \( \Gamma(y) \) is:

\[
\Gamma(y) \sim_{|y| \rightarrow \infty} - \frac{\Gamma(0)}{\pi \rho_0(x)} y^{-2} + O(y^{-4}),
\]

(2.31)

or

\[
\delta \bar{g}(z) \sim_{|z-x| \gg \sqrt{\hat{k}}} - i\hat{k} \frac{\Gamma(0)}{\pi \rho_0(x)} (x-z)^{-2}.
\]

(2.32)

Identifying with the asymptotic result equation (2.19) garnered from the continuum approximation, we get an equation fixing the value of \( \Gamma(0) \):

\[
\Gamma(0) = \pi \rho_0(x).
\]

(2.33)

Note that this result is super-universal, in the sense that it does not depend on \( \beta \), nor on the shape of the confining potential \( V(\lambda) \). Noting that equation (2.30) is of the Ricatti type and that \( \Gamma(y) \) is a regular, even function of \( y \), one finds that the solution can be expressed in terms of the modified Bessel function of the first type as:

\[
\frac{\Gamma(y)}{\Gamma(0)} = 1 - \frac{\Psi'(v)}{\Psi(v)}; \quad \text{with} \quad \Psi(v) = v^{3/4} I_{-\frac{3}{4}}(v) \quad \text{and} \quad v := \Gamma(0) y^2 / 4,
\]

(2.34)

which is plotted in figure 2 together with the corresponding density perturbation \( F(u) \). The continuum limit completely disregards the non-monotonic nature of the function, while only retaining the \( -y^{-2} \) behaviour for large arguments.

Endowed with these results, let us go back to our initial problem, which was to estimate the characteristic function of the distribution of \( g(x) \), equation (2.13). Using the fact that \( \rho_0(\lambda) \) is a saddle point for \( k = 0 \), one finds:

\[
\log \tilde{P}(k) \approx i k \int d\lambda \rho_0(\lambda) + \delta \rho(\lambda) \frac{\lambda}{x - \lambda}.
\]

(2.35)

The first contribution is simply the usual real part of the Stieltjes function \( \bar{g}_0(z) \) for \( z = x \), given by \( V'(x)/2 \) when \( x \) is inside the spectrum. The second term is imaginary and precisely given, to lowest order in \( \hat{k} \), by \( i\zeta \Gamma(0) \). Therefore, for large \( N \) (and thus small \( \hat{k} \)) one finally obtains:

\[
\log \tilde{P}(k) = i k V'(x) - |\hat{k}| \pi \rho(x),
\]

(2.36)

which corresponds precisely the Cauchy distribution obtained in the previous section, that decays for large \( g \) as \( \rho(x)/g^2 \), as expected from general arguments. Our detailed description of the saddle-point solution may allow one to characterize finite \( N \) and/or large deviation effects (see e.g. [11]).
2.3. The Lévy case

Most of the arguments of section 2.1 only rely on the fact that the elements of \( W \) are independent, identically distributed random random variables with a finite second moment, but not necessarily Gaussian. If the second moment is finite, we expect that all the above results will generalize, since they only rely on local properties of the spectrum (see [4, 10]). Let us remind the reader how such arguments must be adapted to the case where the matrix elements of \( W \) have a diverging second moment, or more precisely when:

\[
P(W_{0}) \sim_{W_{0} \to \pm \infty} \frac{1}{N|W_{0}|^{1+\mu}},
\]

with \( \mu < 2 \). Then it is well known that the sum \( \sum_{i} W_{0}^{2} G_{ii} \) converges towards a Lévy stable distribution \( L_{\mu/2}^{c,\beta} \) of index \( \mu/2 \), and scale and asymmetry parameters respectively given by [12]:

\[
C := \frac{1}{N} \sum_{i} |G_{ii}|^{\mu/2}; \quad \beta := \frac{1}{NC} \sum_{i} \text{sign}(G_{ii}) |G_{ii}|^{\mu/2}.
\]

Since \( S := x - 1/G_{00} \) is distributed according to \( L_{\mu/2}^{c,\beta}(S) \), one deduces from equation (2.2) that \( G_{00} \) itself is distributed according to:

\[
P(G_{00}) = \frac{1}{G_{00}^{2}} L_{\mu/2}^{c,\beta}(x - 1/G_{00}).
\]

Now, assuming that for large \( N \) \( G_{00}^{(N+1)} \) has the same distribution as the \( G_{00}^{(N)} \) allows one to find the following self-consistent relations for \( C \) and \( \beta \) for a given value of \( x \) [5–7]:

\[\text{Figure 2.}\] Plot of the scaling function \( \Gamma(y) \) for \( y \geq 0 \) and \( \Gamma(0) = \pi \rho_{0}(x) = 1 \). Note the asymptotic behaviours \( \Gamma(y) \approx 1 - y^{2}/2 \) for small \( y \) and \( \approx -y^{-2} \) for large \( y \). Inset: corresponding density perturbation scaling function \( F(u) \), defined in equation (2.23). Note that \( F(-u) = -F(u) \).
Finally, the distribution of eigenvalues $\rho_L(x)$ of Lévy matrices is obtained, as above, as the coefficient of the $G^{-2}$ tail of $P(G)$, i.e. from equation (2.39):

$$\rho_L(x) = L_{\mu/2}^{C/\beta}(x).$$  

(2.41)

As shown in [7], this result coincides (rather miraculously) with the one obtained using the standard route, i.e. working in the complex plane, with $z = x - i\epsilon$. However, note that the distribution of a single diagonal element (say $G_{00}$) is no longer a Cauchy distribution (see equation (2.39)), although it shares the same power-law tail. The difference with the Wigner case lies in the fact that the Lévy ensemble is not rotationally invariant, and is fact characterized by strong correlations between eigenvectors and eigenvalues. In this case, there is no reason to expect that the distribution of the normalized trace of $G$ and of its diagonal elements is the same. Still, one knows from Aizenman and Warzel that the super-universal Cauchy distribution also holds for the Stieltjes transform (or normalized trace) of Lévy matrices.

3. A Dyson Brownian motion for the resolvent

Since the seminal paper of Dyson in 1962 [13], it is well known that the spectrum of Gaussian random matrices can be described in terms of the (fictitious) motion of $N$ interacting ‘particles’ representing the position of the eigenvalues. More precisely, let us introduce a fictitious time $t$ and define a symmetric random matrix $M(t)$ as:

$$M(t) = C + W(t)$$  

(3.1)

where the $W_{ij}(t)$, $i \leq j$ are independent and identically distributed real Brownian motions, of variance $\sigma^2t/N$ for $i \neq j$ and $2\sigma^2t/N$ for $i = j$. As is well known, the dynamics of the eigenvalues of $M(t)$ is then characterized by a stochastic differential equation (SDE), known as Dyson’s Brownian motion:

$$d\lambda_i(t) = \sqrt{2\sigma^2/N}db_i(t) + \frac{1}{N} \sum_{j \neq i}^{N} \lambda_j(t) - \lambda_i(t),$$  

(3.2)

$$\lambda_i(0) = \mu_i,$$

for $i = 1, \ldots, N$, and where the $b_i(t)$ are independent real Brownian motions. The initial conditions $\lambda_i(t = 0)$ are given by the eigenvalues of $C$, $\mu_1 \geq \mu_2 \geq \ldots \mu_N$.

Here, we present an approach that considers directly the time evolution of the full resolvent matrix $G(z,t)$, which we have not seen in the literature before its publication in our review paper [14]. To that end, we define

$$G(z,t) := (zI - M(t))^{-1}.$$  

(3.3)

Using Itô formula and the fact that $dM_{kl} = dW_{kl}$, one has

$$dG_{ij}(z,t) = \sum_{k,l=1}^{N} \frac{\partial G_{ij}}{\partial M_{kl}} dW_{kl} + \frac{1}{2} \sum_{k,l,m,n=1}^{N} \frac{\partial^2 G_{ij}}{\partial M_{kl} \partial M_{mn}} d[W_{kl}W_{mn}],$$  

(3.4)
where we have treated $M_{kl}$ and $M_{lk}$ as independent variables following 100% correlated Brownian motions. Next, we compute the derivatives:

$$\frac{\partial G_{ij}}{\partial M_{kl}} = \frac{1}{2} [G_{ik}G_{jl} + G_{jk}G_{il}], \quad (3.5)$$

from which we deduce the second derivatives

$$\frac{\partial^2 G_{ij}}{\partial M_{kl} \partial M_{mn}} = \frac{1}{4} \left[ (G_{im}G_{kn} + G_{im}G_{kn}) G_{jl} + \ldots \right], \quad (3.6)$$

where we have not written the other 6 $G G G$ products. Now, using the properties of the Brownian noise $W$, the quadratic co-variation reads

$$d[W_{kl}W_{mn}] = \sigma^2 dt N \left( \delta_{km} \delta_{ln} + \delta_{kn} \delta_{lm} \right), \quad (3.7)$$

so that we get from (3.4) and taking into account symmetries:

$$d[G_{ij}] = \left[ \frac{1}{N} \sum_{k,l=1}^{N} G_{ik}G_{jl} dW_{kl} + \frac{\sigma^2}{N} \sum_{k,j=1}^{N} \left( G_{ik}G_{lj} + G_{ik}G_{lj}G_{jl} \right) dt \right]. \quad (3.8)$$

If we now take the average over with respect to the Brownian motion $W_{kl}$, we find the following evolution for the average resolvent:

$$\partial_t E[G(z,t)] = \sigma^2 g(z,t) E[G^2(z,t)] + \frac{1}{N} E[G^3(z,t)]. \quad (3.9)$$

Now, one can notice that:

$$G^2(z,t) = -\partial_z G(z,t); \quad G^3(z,t) = \frac{1}{2} \partial_z^2 G(z,t), \quad (3.10)$$

which hold even before averaging. By sending $N \to \infty$, we obtain the following matrix PDE for the resolvent:

$$\partial_t E[G(z,t)] = -\sigma^2 g(z,t) \partial_z E[G(z,t)], \quad \text{with} \quad E[G(z,0)] = GC(z). \quad (3.11)$$

Note that this equation is linear in $G(z,t)$ once the Stieltjes transform $g(z,t)$ is known. Taking the trace of equation (3.11) immediately leads to a Burgers equation for $g(z,t)$ itself [15, 16]:

$$\partial_t g(z,t) = -\sigma^2 g(z,t) \partial_z g(z,t), \quad \text{with} \quad g(z,0) = gc(z) := \frac{1}{N} \text{Tr}(zI - C)^{-1}. \quad (3.12)$$

Its solution can be found using the method of characteristics and reads:

$$g(z,t) = gc(Z(z,t)), \quad Z(z,t) := z - \sigma^2 t g(z,t). \quad (3.13)$$

It is plain to see that when $C = 0$, one has $gc(Z) = 1/Z$, leading to the familiar second degree equation for $g(z) := g(z,1)$:

$$g(z)(z - \sigma^2 g(z)) = 1, \quad (3.14)$$

identical to equation (2.5). More generally, equation (3.13) is identical to the well known free addition law for R-transforms [14, 18].

$^5$ See e.g. [18].
More interesting is the solution of equation (3.11) for the full resolvent, that simply reads [17, 19]:

$$G(z, t) = G_C(Z(z, t)), \quad \text{ (3.15)}$$

as can be checked by inserting this ansatz in equation (3.11), and making use of equation (3.13). One can use this result to extract the mean squared overlap between the eigenvectors $u_i(t)$ of the perturbed matrix $M$ and the unperturbed eigenvectors $u_i(0) = v_j$ of $C$. Indeed, let us consider the following projection $\langle v_j G(z, t) v_j \rangle$ with $z = \lambda_i - i\varepsilon$. In the large $N$ limit, this quantity converges to

$$\langle v_j, G(z, t) v_j \rangle \sim \frac{1}{N! \infty} \int \Phi(\lambda, \mu) \rho_M(\lambda) \left( \lambda_i - \lambda - i\varepsilon \right) d\lambda,\quad \text{ (3.16)}$$

where $\varepsilon \gg N^{-1}$ and $\Phi(\lambda, \mu)$ is the smoothed squared overlap between the eigenvector of $C$ associated with eigenvalue $\mu_j$ and eigenvectors of $M$ around eigenvalue $\lambda$, averaged over a small interval of width $\varepsilon$. Therefore, one gets

$$\Phi(\lambda, \mu) = \frac{1}{\pi \rho_M(\lambda)} \lim_{\varepsilon \rightarrow 0^+} \text{Im} \left( \langle v_j, G(\lambda_i - i\varepsilon, t) v_j \rangle \right).\quad \text{ (3.17)}$$

Using equation (3.15) with $\langle v_j, G_C(Z) v_j \rangle = (Z - \mu_j)^{-1}$, one finally obtains, for $t = 1$:

$$N\text{E}[|u_i, v_j|^2] = \frac{\sigma^2}{(\lambda_i - \mu_j - \sigma^2 \text{Re} g(\lambda_i))^2 + \sigma^4 \pi^2 \rho^2(\lambda_i)}.\quad \text{ (3.18)}$$

This result was first obtained in [16, 19], and is the counterpart of the Ledoit and Pêché result [20] in the context of multiplicative models (see [14, 21] for more details). Note that the square overlap is of order $N^{-1}$ as soon as $\sigma > 0$. For a given $\lambda_i$, the overlap has a Lorentzian shape as a function of $\mu_j$, that peaks at $\lambda_i - \sigma^2 \text{Re} g(\lambda_i)$.

Note that equation (3.13) obviously generalizes to any intermediate time $t_0$ as

$$g(z, t) = g(Z(z, t; t_0), t_0), \quad Z(z, t; t_0) := z - \sigma^2 (t - t_0) g(z, t), \quad \text{ (3.19)}$$

and correspondingly

$$G(z, t) = G(Z(z, t; t_0), t_0).\quad \text{ (3.20)}$$

This enables one to compute the eigenvector overlaps between any two times $t_0$ and $t$.

Another interesting case is when the initial matrix $C$ is of rank one, with a single non zero eigenvalue $\mu_1$ and eigenvector $v_1$, and all the other $N - 1$ ones are zero, as above. If one carefully keeps terms of order $1/N$, the whole formalism allows one to keep track of the isolated eigenvalue for $t > 0$, and the corresponding overlap $\Phi_1 := |u_1, v_1|^2$. One readily finds that the isolated eigenvalues persists up to $t = t^* = (\mu_1/\sigma)^2$, and is located at $[22]$: 

$$\lambda_1(t) = \mu_1 + \frac{\sigma^2 t}{\mu_1},\quad \text{ (3.21)}$$

before colliding with the edge of the Wigner spectrum $2\sigma \sqrt{t}$ precisely at $t = t^*$ and disappearing altogether in the Wigner sea for $t > t^*$. This is the famous BBP transition [23]. The overlap $\Phi_1$ is contained in the corresponding pole of $G(z = \lambda_1, t)$, as given by equation (3.13) and is found to be given by:

$$\Phi_1(t) = 1 - \frac{t}{t^*}, \quad \text{ (3.22)}$$
which goes to zero at $t = t^*$, as it should be [17, 24]. The way $\Phi_1(t^*)$ behaves for finite $N$ is apparently not known. A natural conjecture is that, for $t \approx t^*$ and $N \gg 1$, the following scaling result holds:

$$
\Phi_1(t, N) = N^{-1/3} \varphi \left( N^{1/3} \frac{t^* - t}{t^*} \right),
$$

(3.23)

where $\varphi(u \gg 1) = u$ and $\varphi(u = 0) = \varphi_0$ a positive constant.

Finally, it is interesting to write the dynamical equation for $g(z, t)$ keeping terms of order $1/N$, since some non zero noise survives in that limit. One finds the following Langevin equation:

$$
\partial_t g(z, t) = -\sigma^2 g(z, t) \partial_z g(z, t) + \frac{1}{2N} \partial_{zz} g(z, t) + \frac{\sigma}{N} \xi(z, t),
$$

(3.24)

where $\xi$ is a white (Langevin) complex noise, such that:

$$
\mathbb{E}(\xi(z, t) \xi(z, t')) = -2\partial_{zz} g(z, t) \delta(t - t').
$$

(3.25)

It would be interesting to see if one can extract some useful information from this formalism. In particular, is it possible to recover the Cauchy distribution discussed in the first part of this paper as a stationary distribution of the above Langevin equation?

We thank R Allez, G Biroli, J Bun, Y Fyodorov, A Guionnet, V Hakim and D Shlyakhtenko for very fruitful exchanges on these topics.

**Appendix A. The Aizenman–Warzel approach**

The paper of Aizenman and Warzel [4] is not easy to penetrate (at least for us). Here we give a simplified version of their approach, that allows one to understand the super-universal nature of the Cauchy law. Assume that instead of fixing the (real) value of $x$ at which we want to compute the distribution of Stieltjes transform $\tilde{g}(x)$ over the considered ensemble of random matrices, one rather fixes the position of the eigenvalues $\lambda_i$ in a typical configuration, such that the local density is $\rho(x)$ and the typical distance between consecutive eigenvalues is $(N\rho(x))^{-1}$. We now want to study the distribution of $g(x + \eta u)$, with $u \sim O(1)$ a random variable with an arbitrary distribution $R(u)$ and $\eta$ a small parameter such that $N^{-1} \ll \eta \ll 1$.

The characteristic function of the distribution of $g$ is then:

$$
\hat{P}(k) = \int du R(u) e^{ik(x + \eta u)}.
$$

(A.1)

One of the lemma of [4] is that the choice of $R(u)$ is arbitrary provided the range of $\eta u$ covers many eigenvalues, i.e. $N^{-1} \ll \eta$. They choose

$$
R(u) = \frac{1}{\pi} \frac{1}{1 + u^2},
$$

(A.2)
i.e. a Cauchy distribution—but unrelated to the final (Cauchy) result we are looking for. Now, one should note that the simple pole structure of \( g(x) \) implies that it maps the upper complex plane \( \mathbb{C}^+ \) into the lower complex plane \( \mathbb{C}^- \). Hence, \( \exp[ik g(x + u)] \) is bounded when \( u \) is in the lower (resp. upper) complex plane when \( k > 0 \) (resp. \( k < 0 \)). For \( k > 0 \), one can therefore calculate equation (A.1) using a contour integration in the lower complex plane, enclosing the pole of \( R(u) \) at \( u = -i \). The result is:

\[
\hat{P}(k) = e^{ikg(x-iu)}.
\]  
(A.3)

Similarly, for \( k < 0 \), the pole is at \( u = +i \) and:

\[
\hat{P}(k) = e^{ikg(x+iu)}.
\]  
(A.4)

For large \( N \), \( g(x + i\eta) \) converges to the limiting Stieltjes distribution \( g_0(x + i\eta) \) provided that \( \eta \gg N^{-1} \). Taking the limit \( \eta \to 0 \) in this sense leads to the characteristic function of the Cauchy distribution, identical to equation (2.36):

\[
\hat{P}(k) = e^{ikg(x) - |\pi \rho(x)|},
\]  
(A.5)

where we have used the standard result:

\[
\lim_{\eta \to 0^+} g_0(x - i\eta) = g_\rho(x) - i\pi \rho(x).
\]  
(A.6)

That the process of averaging over the matrix ensemble or over the position of \( x \) leads to the same result is not unexpected, but not totally trivial either.

**Appendix B. The case of a periodic array of eigenvalues**

Assume that the eigenvalues are locally equally spaced. Near a certain \( x \), eigenvalues are spaced by \( \Delta = 1/(N\rho(x)) \), \( \rho(x) \) is assumed to be constant for \( L = \sqrt{N} \) eigenvalues above and below \( \lambda_m \), defined as the closest eigenvalue to \( x \).

Let \( x - \lambda_m = \Delta u \), so \( u \) is uniform in \([-1/2, 1/2]\). For the \( L \) first eigenvalues larger than \( \lambda_m \), we have \( x - \lambda_m + k = \Delta(u + k) \). For the \( L \) eigenvalues immediately below \( \lambda_m \) we have \( x - \lambda_m - k = \Delta(u - k) \). We split \( g(x) \) into two parts, one near \( x \) one far from \( x \):

\[
g(x) = \frac{1}{N} \left[ \sum_{k \notin [m-L,m+L]} \frac{1}{x - \lambda_k} + \frac{1}{\Delta u} + \sum_{k=1}^{L} \left( \frac{1}{\Delta(u-k)} + \frac{1}{\Delta(u+k)} \right) \right].
\]  
(B.1)

The first sum can be replaced by a principal part integral, using \( \rho(x) = 1/(N\Delta) \) and grouping the last two terms we get

\[
g(x) = g_\rho(x) + \rho(x) \left( \frac{1}{u} + 2u \sum_{k=1}^{L} \frac{1}{u^2 - k^2} \right),
\]  
(B.2)

where \( g_\rho(x) \) is again the real part of the limiting Stieltjes transform. The sum on the right is convergent, we can replace \( L = \sqrt{N} \to \infty \). Mathematica says this sum is

\[
\frac{1}{2u} \left( \pi \cot(\pi u) - \frac{1}{u} \right).
\]  
(B.3)

---

6 One can easily check that taking other forms that allow one to use residues, such as \( R(u) = (2/\sqrt{3}) \pi (1 + u^2)^{-3} \), leads to the same final result.
We can rewrite equation (B.2) as

\[
\frac{1}{\pi \rho(x)} (g(x) - g_R(x)) = \cot(\pi u)
\]

(B.4)

\[
u = \pi^{-1} \cot^{-1} \left( \frac{g(x) - g_R(x)}{\pi \rho(x)} \right)
\]

(B.5)

which is equivalent to saying that \( g(x) \) is distributed according to the Cauchy law centered at \( g_R(x) \) and of width \( \pi \rho(x) \), as for the Poisson case, or any other value of \( \beta \) for Coulomb gas models.

ORCID iDs

Jean-Philippe Bouchaud https://orcid.org/0000-0002-0427-6456

References

[1] Fyodorov Y V and Sommers H J 1997 Statistics of resonance poles, phase shifts and time delays in quantum chaotic scattering: random matrix approach for systems with broken time-reversal invariance J. Math. Phys. 38 1918–81
[2] Fyodorov Y V and Savin D V 2004 Statistics of impedance, local density of states, and reflection in quantum chaotic systems with absorption J. Exp. Theor. Phys. Lett. 80 725–9
[3] Fyodorov Y V and Williams I 2007 Replica symmetry breaking condition exposed by random matrix calculation of landscape complexity J. Stat. Phys. 129 1081–116
[4] Aizenman M and Warzel S 2015 On the ubiquity of the Cauchy distribution in spectral problems Probab. Theory Relat. Fields 163 61–87
[5] Cizeau P and Bouchaud J P 1994 Theory of Lévy matrices Phys. Rev. E 50 1810
[6] Burda Z, Jurkiewicz J, Nowak M A, Papp G and Zahed I 2007 Free random Lévy and Wigner–Lévy matrices Phys. Rev. E 75 051126
[7] Ben Arous G and Guionnet A 2008 The spectrum of heavy tailed random matrices Commun. Math. Phys. 278 715–51
[8] Griniasty M and Hakim V 1994 Correlations and dynamics in ensembles of maps: simple models Phys. Rev. E 49 2661
[9] Dean D S and Majumdar S N 2008 Extreme value statistics of eigenvalues of Gaussian random matrices Phys. Rev. E 77 041108
[10] Erdos L, Schlein B and Yau H T 2010 Wegner estimate and level repulsion for Wigner random matrices Int. Math. Res. Not. 2010 436–79
[11] Grabisch A and Texier C 2016 Distribution of spectral linear statistics on random matrices beyond the large deviation function: Wigner time delay in multichannel disordered wires J. Phys. A: Math. Theor. 49 465002
[12] Bouchaud J P and Georges A 1990 Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications Phys. Rep. 195 127–293
[13] Dyson F J 1962 A Brownian-motion model for the eigenvalues of a random matrix J. Math. Phys. 3 1191–8
[14] Bun J, Bouchaud J P and Potters M 2017 Cleaning large correlation matrices: tools from random matrix theory Phys. Rep. 666 1–109
[15] Rogers L C G and Shi Z 1993 Interacting Brownian particles and the Wigner law Probab. Theory Relat. Fields 95 555–70
[16] Allez R and Bouchaud J P 2014 Eigenvector dynamics under free addition Random Matrices: Theory Appl. 3 1450010
[17] Allez R, Bun J and Bouchaud J P 2014 The eigenvectors of gaussian matrices with an external source (arXiv:1412.7108)
[18] Tulino A M and Verdú S 2004 Random matrix theory and wireless communications *Found. Trends Commun. Inf. Theory* 1 1–182
[19] Shlyakhtenko D 1996 Random Gaussian band matrices and freeness with amalgamation *Int. Math. Res. Not.* 1996 1013–25
[20] Ledoit O and Pécé S 2011 Eigenvectors of some large sample covariance matrix ensembles *Probab. Theory Relat. Fields* 151 233–64
[21] Bun J, Allez R, Bouchaud J P and Potters M 2016 Rotational invariant estimator for general noisy matrices *IEEE Trans. Inf. Theory* 62 7475–90
[22] Féral D and Pécé S 2007 The largest eigenvalue of rank one deformation of large Wigner matrices *Commun. Math. Phys.* 272 185–228
[23] Baik J, Ben Arous G and Pécé S 2005 Phase transition of the largest eigenvalue for non-null complex sample covariance matrices *Ann. Probab.* 33 1643–97
[24] Biroli G, Bouchaud J P and Potters M 2007 On the top eigenvalue of heavy-tailed random matrices *Europhys. Lett.* 78 10001
[25] Bun J, Bouchaud J P and Potters M in preparation
[26] Mello P A 1955 *Mesoscopic Quantum Physics (Les Houches Summer School)* ed E Akkermans *et al* (Amsterdam: Elsevier) p 435 (Session LXI)