DIAGRAMMATIC APPROACH TO FLUCTUATIONS IN THE WISHART ENSEMBLE

Jerzy Jurkiewicz[a,b] Grzegorz Łukaszewski[b]
Maciej A. Nowak[a,b,c]

[a]Mark Kac Center for Complex Systems Research, Jagellonian University
Reymonta 4, 30-059 Kraków, Poland
[b]Institute of Physics, Jagellonian University
Reymonta 4, 30-059 Kraków, Poland
[c] GSI, Planckstr. 1, 64291 Darmstadt, Germany

(Received December 3, 2007)

Using diagrammatic techniques, we calculate two-point Green's function for complex Wishart ensemble.

PACS numbers: 02.10.Yn

1. Introduction

Nowadays, information is easily being stored in form of huge matrices, and random matrix theory, in particular the spectral analysis, provides a powerful tool in the study of multivariate analysis of several array processing applications, ranging from telecommunication [1], through genomics [2] to finances and economy [3]. A typical problem is the analysis of covariance matrices, *i.e.* ensembles of the Wishart type [4]. Standard inference methodologies linking estimators to "true" covariance matrices are based on relation between corresponding spectral densities of both ensembles [5]. Very recently, a new conceptual idea was proposed [6], enhancing the inference methodology by exploiting the properties of the fluctuations of eigenvalues. This progress was possible due to new results for two-point Green's function for Wishart ensemble. First, Bai and Silverstein [7] obtained this result using two-dimensional complex integral representation. Then, Speicher and collaborators [8] obtained more general result, introducing so-called freeness of the second kind, a notable extension of the free random variable concept [9, 10]. Despite the final result for correlator for the Wishart ensemble is simple, mathematical tools leading to this result are quite involved and may not be well known among the practitioners. The main aim of this note is to provide a simple re-derivation of the two-point Green's function for the Wishart ensemble, using standard tools of Random Matrix Theory. In the first chapter we introduce diagrammatic technique and we re-derive known results for one-point and two-point Green's functions for Gaussian Unitary Ensemble. In the second chapter, we parallel this construction in the case of Wishart ensemble, providing final formulae. We also point the connection of the final result to universality properties, noticed some time ago by one of the authors [11].

2. One and two-point Green's functions

We are interested in one- and two-point spectral densities, in the limit when the dimension N of the matrices tends to infinity. Corresponding distributions are defined as

$$\rho(\lambda) = \frac{1}{N}\left\langle \sum_{i=1}^{N} \delta(\lambda-\lambda_i) \right\rangle, \tag{1}$$

$$\rho(\lambda,\lambda') = \frac{1}{N^2}\left\langle \sum_{i,j}^{N} \delta(\lambda-\lambda_i)\delta(\lambda'-\lambda_j) \right\rangle_c, \tag{2}$$

where $\langle AB\rangle_c = \langle AB\rangle - \langle A\rangle\langle B\rangle$ for any A, B and averaging $\langle \ldots \rangle$ is performed over the ensemble of matrices $N\times N$ matrices H drawn with probability measure

$$P(H)dH = e^{-N\mathrm{Tr}V(H)}dH \tag{3}$$

with some potential $V(H)$ defining the type of randomness considered.

It is usually convenient to define so-called Green's functions, and then use them to extract corresponding spectral densities.

We define one-point Green's function as

$$G(z) = \frac{1}{N}\left\langle \mathrm{Tr}\,\frac{1}{z-H} \right\rangle \tag{4}$$

and two-point Green's function as

$$G(z,w) = \frac{1}{N^2}\left\langle \mathrm{Tr}\,\frac{1}{z-H}\mathrm{Tr}\,\frac{1}{w-H} \right\rangle_c . \tag{5}$$

Corresponding spectral distributions are easy to reconstruct from the discontinuities of the Green's functions, using the relation

$$\lim_{\epsilon\to 0}\frac{1}{\lambda \pm i\epsilon} = \boldsymbol{P}\frac{1}{\lambda} \mp i\pi\delta(\lambda)\,. \tag{6}$$

Indeed,

$$-\frac{1}{2\pi i}(G(+)-G(-)) = \frac{1}{N}\langle \mathrm{Tr}\,\delta(\lambda - H)\rangle = \frac{1}{N}\left\langle \sum_{i=1}^{N}\delta(\lambda-\lambda_i)\right\rangle = \rho(\lambda) \quad (7)$$

and similarly

$$\rho_c(\lambda,\lambda^{'}) = -\frac{1}{4\pi^2}(G(+,+)-G(+,-)+G(-,-)-G(-,+))\,, \qquad (8)$$

where we introduced the shorthand notation

$$\pm \equiv \lim_{\epsilon\to 0} z|_{z=\lambda\pm i\epsilon} \qquad (9)$$

for generic complex z.

In order to calculate one and two-point Green's functions for Wishart ensemble in the large N limit, we apply diagrammatic method, following general prescription proposed in [13]. We would like to stress, that despite being apparently a perturbative method, the structure of the planar graphs that contribute allow for a resummation of the whole perturbative series and give the full *exact* result in the planar limit.

3. Diagrammatic technique — GUE ensemble

In this section we recall well-known results for one-point and two-point Green's functions for Gaussian Unitary Ensemble, following diagrammatic technique [13]. In this way we introduce the necessary notation which will allow us later to repeat easily similar calculation in the case of the Wishart ensemble.

A starting point of our analysis is the expression allowing for the reconstruction of the Green's function from all the moments $\langle \mathrm{Tr} H^n\rangle$,

$$\begin{aligned} G(z) \;&=\; \frac{1}{N}\left\langle \mathrm{Tr}\,\frac{1}{z-H}\right\rangle = \frac{1}{N}\sum_n \frac{1}{z^{n+1}}\langle \mathrm{Tr} H^n\rangle \\ &=\; \frac{1}{N}\left\langle \mathrm{Tr}\left[\frac{1}{z}+\frac{1}{z}H\frac{1}{z}+\frac{1}{z}H\frac{1}{z}H\frac{1}{z}+\cdots\right]\right\rangle . \end{aligned} \qquad (10)$$

We exploit the diagrammatic method to evaluate efficiently the sum of the moments on the right hand side. Note, that in general this series expansion is convergent only in a neighborhood of $z=\infty$. So the results of the diagrammatic calculation apply directly only there. For hermitian random matrices this however does not pose a problem, since the eigenvalues lie only on some intervals on the real axis, the Green's functions are *holomorphic functions* of one or more complex variables except some cuts, and we can reconstruct the Green's function everywhere by analytical continuation.

For illustration, let us consider well-known case of a random hermitian ensemble with Gaussian distribution. We introduce a generating function with a matrix source J:

$$Z(J) = \int dH e^{-\frac{N}{2}\mathrm{Tr}H^2 + \mathrm{Tr}H\,J}\,. \tag{11}$$

All the moments follow from $Z(J)$,

$$\langle \mathrm{Tr}H^n \rangle = \frac{1}{Z(0)} \mathrm{Tr}\left(\frac{\partial}{\partial J}\right)^n Z(J)\Big|_{J=0} \tag{12}$$

and are straightforward to calculate, since in our case the partition function reads $Z(J) = \exp \frac{1}{2N}\mathrm{Tr}J^2$.

However, instead of calculating all the moments individually alike by the Wick theorem, we could draw them using "Feynman rules" derived from our generating function, and perform a resummation of *all* relevant graphs. The "propagator" reads

$$\langle H^a_b H^c_d \rangle = \frac{1}{Z(0)} \frac{\partial^2 Z(J)}{\partial J^b_a \partial J^d_c}\Big|_{J=0} = \frac{1}{N}\delta^c_b \delta^a_d\,. \tag{13}$$

The $1/z$ in (10) is represented by a horizontal straight line. We depict the "Feynman" rules in Fig. 1.

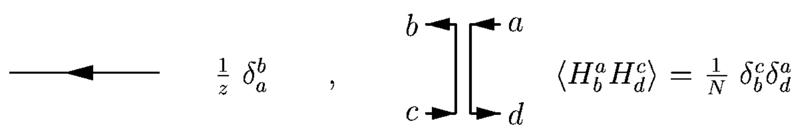

Fig. 1. Large N "Feynman rules" for GUE.

The diagrammatic expansion of Green's function in the large N limit is visualized in Fig. 2. Each "propagator" brings a factor of $1/N$, and each loop a factor of N, therefore only planar graphs survive the large N limit. Introducing the self-energy Σ comprising the sum of all one-particle irreducible

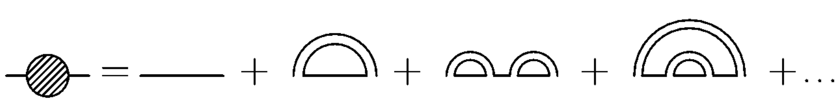

Fig. 2. Diagrammatic expansion of Green's function (4) for Gaussian ensemble. The 2nd and 4th graphs are "rainbow" graphs contributing to the self-energy Σ.

graphs (rainbow-like), the Green's function reads

$$G(z) = \frac{1}{z} + \frac{1}{z}\Sigma(z)\frac{1}{z} + \frac{1}{z}\Sigma(z)\frac{1}{z}\Sigma(z)\frac{1}{z} + \ldots = \frac{1}{z - \Sigma(z)}\,. \tag{14}$$

This equation is represented diagrammatically as a geometric series (*cf.* figure 3). In the large N limit the equation for the self energy Σ, follows from resumming the rainbow-like diagrams of Fig. 2. The resulting equation

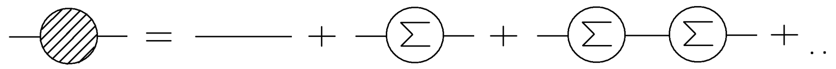

Fig. 3. Green's function expressed in terms of self-energy Σ.

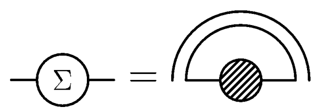

Fig. 4. Schwinger–Dyson equation for GUE.

("Schwinger–Dyson" equation of Fig. 4) encodes pictorially the structure of these graphs and reads

$$\Sigma = \frac{1}{N}\mathrm{Tr}G\mathbf{1}_N = G\,. \tag{15}$$

Equations (14) and (15) give immediately $G(z-G)=1$, so the normalizable solution for the Green's function reads

$$G(z) = \frac{1}{2}\left(z - \sqrt{z^2-4}\right) \tag{16}$$

which, via the discontinuity (cut) leads to Wigner's semicircle for the distribution of the eigenvalues for hermitian random matrices

$$\rho(\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2}\,. \tag{17}$$

We have put a priori arbitrary variance equal to 1, one can easily recover more general result via rescaling $\lambda \to 2\lambda/a$, so the spectrum is localized on interval $[-a,a]$ and normalized density reads $2\sqrt{a^2-\lambda^2}/\pi a^2$.

Same technique we can apply in the case of the two-point Green's function. We express Green's function (5) as

$$G(z,w) = \partial_z\partial_w \left\langle \frac{1}{N^2}\mathrm{Tr}\log(z-H)\mathrm{Tr}\log(w-H)\right\rangle_c\,, \tag{18}$$

and we expand the logarithms, getting

$$G(z,w) = \partial_z \partial_w \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{z^n w^k} \left\langle \frac{1}{Nn} \mathrm{Tr} H^n \frac{1}{Nk} \mathrm{Tr} H^k \right\rangle_c . \tag{19}$$

Let us first consider "diagonal" contributions coming from $n = k$. The corresponding graph can be represented as a wheel diagram, when outer rim corresponds *e.g.* to $1/z$ lines, and inner rim to $1/w$, and rungs are just double lines depicted in Fig. 1. Since $\langle \mathrm{Tr} H^n \mathrm{Tr} H^n \rangle_c = n$ for Gaussian variables, the diagram splits into n disconnected sectors. Resummation of diagonal contributions gives therefore

$$N^2 G(z,w)|_{\mathrm{diag}} = -\partial_z \partial_w \log \left(1 - \frac{1}{zw}\right). \tag{20}$$

The non-diagonal contributions can in general contribute in two-ways. First, they can modify bare line $1/z(1/w)$ by arbitrary insertions of self-energies Σ, corresponding to replacement of the $1/z$ by $G(z)$ ($1/w$ by $G(w)$). Second, they can in general modify the elementary rung, by replacing the double line by some general, two-particle irreducible kernel $\Gamma(z,w)$, as represented graphically in figure 5. Interference between these two contributions is forbidden by the large N (planarity) argument. Therefore in general, the mandala-like diagram yields the result [13]

$$N^2 G(z,w) = -\partial_z \partial_w \log(1 - G(z)G(w)\Gamma(z,w)) . \tag{21}$$

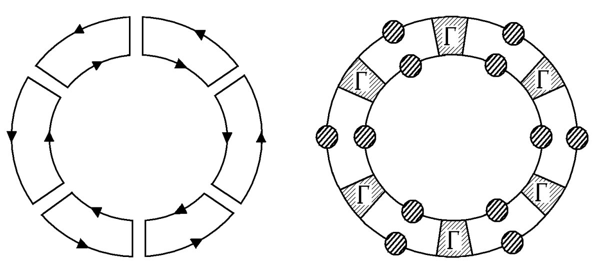

Fig. 5. Two-point correlator for GUE ensemble: sample diagonal contribution for $n = 6$ (left), "dressed" general contribution for $n = 6$ (right).

In the case of Gaussian Unitary Ensemble, the two-point irreducible kernel $\Gamma_{ab,cd} = 1/N \delta_{cb} \delta_{ad}$, since potential $V(H)$ is only quadratic and does not involve "interaction" terms H^k, for $k > 2$. Therefore the final result for two-point correlator in case of the GUE reads

$$G_{\mathrm{GUE}}(z,w) = \frac{1}{N^2} \partial_z \partial_w \log(1 - G(z)G(w)) \tag{22}$$

This simple result hides deep universality properties. First, let us note that two-point Green's function is rephrased in terms of one-point Green's functions only, a point observed for GUE already by [14]. This feature is the consequence of deep universality, noticed first by [11]. Let us rewrite this result using an explicit form of the Green's function, *i.e.* Wigner semicircle spanned on the interval $(-a, a)$. Then

$$G(z,w) = \frac{1}{4(z-w)^2}\left(\frac{z^2+w^2-2a^2}{\sqrt{(z^2-a^2)(w^2-a^2)}}\right) - \frac{1}{4\sqrt{(z^2-a^2)(w^2-a^2)}}\,. \tag{23}$$

The power of universality argument stems from the fact, that this result is valid for any symmetric potential $V(H)$, provided the spectrum is localized on one-cut. In other words, the dependence of the detailed form of the potential comes solely via the endpoints $\pm a$.

Another way of rephrasing this universality is to rewrite above equation using "Schwinger–Dyson" equation $G(z)(z - G(z)) = 1$ Since

$$\frac{1}{G(z)} - \frac{1}{G(w)} = \frac{G(w)-G(z)}{G(z)G(w)} \tag{24}$$

above relation allows to write Green's function as

$$N^2 G(z,w) = -\partial_z \partial_w \log \frac{G(w)-G(z)}{z-w} G(z)G(w)\,. \tag{25}$$

So universality also means [15], that two-point correlator, modulo irrelevant factorized terms depends solely on the universal kernel of the form $\log \frac{G(w)-G(z)}{z-w}$.

Explicit differentiation leads to another form

$$G(z,w) = \frac{G'(z)G'(w)}{(G(z)-G(w))^2} - \left(\frac{1}{z-w}\right)^2. \tag{26}$$

This form has also an interesting interpretation — could be viewed as a consequence of the freeness of the second kind, introduced recently by Speicher and collaborators [8]. Note that from geometric point of view this kernel has a form of Bergmann kernel, suggesting further links of universality to differential geometry [16].

4. Complex Wishart ensemble

We parallel now diagrammatic construction for GUE in case of the complex Wishart N by N ensemble where $H = V^\dagger V$, where V is N by M matrix. We can assume $m = M/N$ larger than 1, and we will work in the limit when both N, M tend to infinity keeping the ratio constant. Note also that the case of M by M ensemble $H' = VV^\dagger$ is basically equivalent to the case considered - due to cyclic properties of the trace all the moments of H and H' are identical, and the only spectral difference comes from the fact that H' ensemble has precisely $(M-N)$ trivial algebraic zero eigenvalues comparing to ensemble H. Since now matrices V come always in pairs, we denote basic building block $\left\langle V_k^a V_l^{*b} \right\rangle = \frac{1}{N}\delta^{ab}\delta_{kl}$, where $a, b = 1, \ldots N$ and $k, l = 1, \ldots M$. Corresponding graph is presented in Fig. 6. To distinguish lines carrying N entries from lines carrying M entries we introduced "dashed propagators" in the case of the second type of lines. All other features are similar comparing to Gaussian case. As in the previous case, Green's function is expressed via one-particle irreducible self energy $\Sigma(z)$ as a geometric series (Fig. 3), *i.e.*

$$G(z) = \frac{1}{z - \Sigma(z)}\,. \tag{27}$$

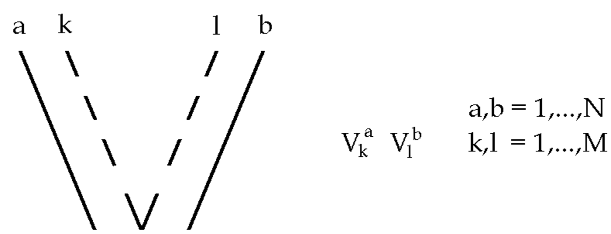

Fig. 6. "Feynman" rule for Wishart propagator.

Self-energy comes from resummation of the rainbow diagrams, represented in Fig. 7. So we read

$$\Sigma(z) \equiv mF(z)\,, \tag{28}$$

where m comes from the dashed loops inside the rainbows. Finally, equation for $F(z)$ is represented diagrammatically in figure 8, therefore depicted iteration written in algebraic way yields

$$F(z) = 1 + F(z)G(z)\,. \tag{29}$$

From (27,28,29) we get quadratic equation

$$zG^2(z) + (m - z - 1)G(z) - 1 = 0 \tag{30}$$

so normalizable solution reads

$$G(z) = \frac{1}{2z}\left(z + 1 - m - \sqrt{(m-1-z)^2 - 4z}\right) \tag{31}$$

and spectral distribution reads

$$\rho(\lambda) = \frac{1}{2\pi}\frac{\sqrt{(\lambda-\lambda_-)(\lambda_+-\lambda)}}{\lambda}\,, \tag{32}$$

where $\lambda_\pm = (\sqrt{m}\pm 1)^2$. This is indeed Pastur–Marchenko result [12].

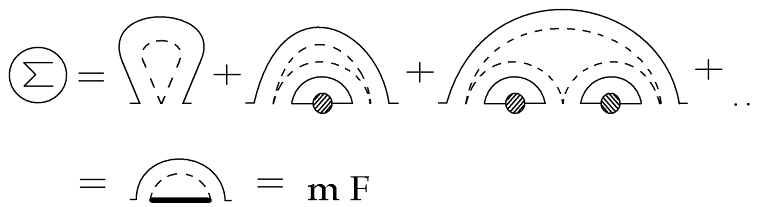

Fig. 7. Sample graphs contributing to self-energy in the case of Wishart ensemble.

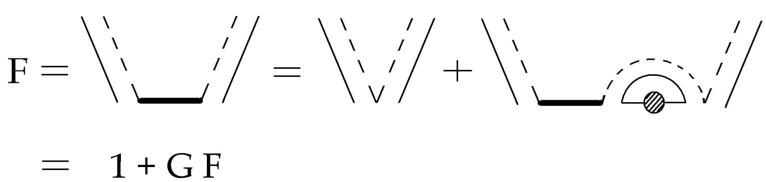

Fig. 8. Schwinger–Dyson equation for Wishart ensemble.

The case of two-point correlator is equally simple. We can use general formula (5), we have only to see what is the two-point irreducible part Γ. From the diagrammatic representation of the rung presented in Fig. 9 we read $\Gamma(z,w) = mF(z)F(w)$, so two-point correlator for Wishart reads

$$G(z,w) = -\partial_z\partial_w \log\left(1 - m\frac{G(z)}{1-G(z)}\frac{G(w)}{1-G(w)}\right), \tag{33}$$

where we used that $F = 1/(1-G)$. Sample diagonal and "dressed" diagrams are depicted in Fig. 10. Using Schwinger–Dyson equation for $G(z)$

$$\frac{1}{G(z)} = z - \frac{m}{1-G(z)} \tag{34}$$

and similar for $G(w)$, one can easily realize (by subtracting formulae (34) for z and w variables), that the argument of the logarithm is given (modulo irrelevant factorized terms) by the universal kernel, and final formula is identical functionally to the considered previously case of the GUE, and two-point correlator for the Wishart has the form

$$G(z,w) = \frac{G'(z)G'(w)}{(G(z)-G(w))^2} - \left(\frac{1}{z-w}\right)^2. \tag{35}$$

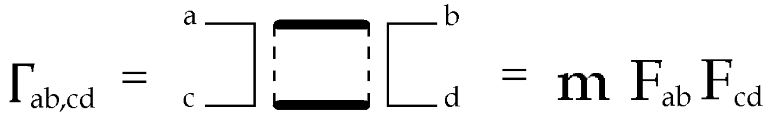

Fig. 9. Two-point irreducible kernel in case of Wishart ensemble.

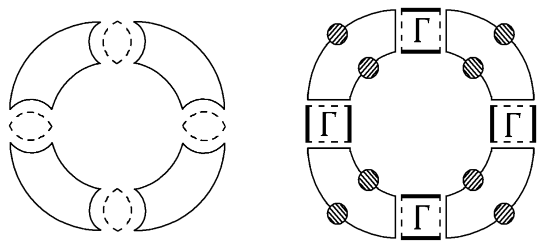

Fig. 10. Sample (left) and general (right) "mandala" graph for two-point correlator for Wishart ensemble.

It is interesting that this explicit result for two-point Green's function for Wishart ensemble was obtained in mathematical literature only very recently [7]. Recently, it was also derived as a consequence of the freeness of the second kind [8]. We would like to point, that this result can be also obtained as a trivial (holomorphic) reduction of the case of more involved two-point correlators for non-hermitian ensembles considered in [18]. Again, the universal form is another consequence of the AJM universality. This observation allows to generalize the analysis of fluctuations to the case when Wishart ensemble is correlated, *i.e.* we consider Gaussian measure of the type [17]. As long as we consider one-cut solutions, the dependence of correlations enters solely via the endpoints, as a consequence of AJM universality. One can check this point via direct, although more painful method, *i.e.* repeating the above calculation in the case of modified diagrams corresponding to correlated Wishart distributions.

5. Conclusions

We have presented a fast derivation for the two-point correlations for the Wishart ensemble. As a next step we are planning to check, to what extent the knowledge of two-point correlations improves cleaning signals from noise for large sets of data gathered for real complex systems.

This work was partially supported by the Marie Curie TOK Grant MTKD-CT-2004-517186 "Correlations in Complex Systems"(COCOS).

REFERENCES

[1] A. Tulino, S. Verdu, *Random Matrix Theory and Wireless Communications*, Now Publishers, ISBN: 193301900X, 2004.

[2] J. Schaefer, K. Strimmer, *Stat. Appl. Gen. Mol. Biol.* **4**, 32 (2005).

[3] Proceedings of Conference on Application of Random Matrices to Economy and Other Complex Systems, *Acta Phys. Pol. B* **36**, 2603 (2005).

[4] J. Wishart, *Biometrica* **20A**, 32 (1928).

[5] L.R. Haff, *Ann. Statist.* **8**, 586 (1980).

[6] N.R. Rao, J. Mingo, R. Speicher, A. Edelman, *Adv. Math.* **209**, 212 (2007).

[7] Z.D. Bai, J. Silverstein, *Ann. Prob.* **32**, 533 (2004).

[8] B. Collins, J.A. Mingo, P. Śniady, R. Speicher, *Documenta Mathematica* **12**, 1 (2007).

[9] D.V. Voiculescu, K.J. Dykema, A. Nica, *CRM Monograph Series, Am. Math. Soc.*, Vol. 1 Providence 1992.

[10] R. Speicher, *Math. Ann.* **298**, 611 (1994).

[11] J. Ambjørn, J. Jurkiewicz, Y. Makeenko, *Phys. Lett.* **B251**, 517 (1990).

[12] V.A. Marchenko, L.A. Pastur, *Math. USSR-Sb* **1**, 457 (1967).

[13] E. Brézin, A. Zee, *Nucl. Phys.* **B453**, 531 (1995).

[14] J.B. French, P.A. Mello, A. Pandey, *Ann. Phys. (N.Y.)* **113**, 277 (1978); J.J.M. Verbaarschot, H.A. Weidenmuller, M. Zirnbauer, *Ann. Phys. (N.Y.)* **153**, 367 (1984).

[15] E. Brézin, A. Zee, *Nucl. Phys.* **B402**, 613 (1993).

[16] L. Chekhov, B. Eynard, *J. High Energy Phys.* **0603**, 4 (2006).

[17] Z. Burda, A. Goerlich, A. Jarosz, J. Jurkiewicz, *Physica A* **343**, 295 (2004).

[18] R.A. Janik, M.A. Nowak, G. Papp, I. Zahed, *Nucl. Phys.* **B501**, 603 (1997).