

# On the Relation Between the Gaussian Orthogonal Ensemble and Reflections, or a Self-Adjoint Version of the Marcus–Pisier Inequality

Roy Wagner

*Abstract.* We prove a self-adjoint analogue of the Marcus–Pisier inequality, comparing the expected value of convex functionals on random reflection matrices and on elements of the Gaussian orthogonal (or unitary) ensemble.

## 1 Introduction

The starting point for this paper is the Marcus–Pisier inequality from [MP]

$$\left( E_G \sup_{X \in M} |\operatorname{tr}(GX)|^p \right)^{1/p} \leq C \left( E_U \sup_{X \in M} |\operatorname{tr}(UX)|^p \right)^{1/p},$$

where  $G$  is a matrix whose entries are independent  $N(0, 1/n)$  Gaussian random variables,  $U$  a random unitary matrix,  $M$  any collection of  $n$  by  $n$  matrices, and  $C$  a constant independent of the dimension and of  $1 \leq p \leq \infty$ . Note that expressions of the form “ $\sup_{X \in M} |\operatorname{tr}(AX)|$ ” include all norms on the matrix  $A$ . The reverse inequality also holds, but here the constant diverges as  $p$  approaches  $\infty$ . According to [S2], where the proof of both directions is reproduced, the reverse inequality is due to Davis and Garling (unpublished).

Motivated by considerations arising from free probability and statistical physics on the one hand, and from the geometry of Banach spaces on the other, this paper will study a symmetric version of this equivalence:  $G$  is replaced by a random self-adjoint Gaussian matrix  $\hat{G}$  (the so called GOE or GUE), and  $U$  is replaced by a random self-adjoint unitary matrix  $R$  (a random reflection). If we restrict our attention to reflections with zero trace, which we will denote  $R_0$ , we obtain the following equivalence

$$\begin{aligned} \left( \frac{8}{3\pi} - o(1) \right) E_{R_0} \left( \sup_{X \in M} |\operatorname{tr}(R_0 X)| \right) &\leq E_{\hat{G}} \left( \sup_{X \in M} |\operatorname{tr}(\hat{G} X)| \right) \\ &\leq (2 + o(1)) E_{R_0} \left( \sup_{X \in M} |\operatorname{tr}(R_0 X)| \right) \\ &\quad + \frac{2}{\sqrt{\pi n}} \sup_{X \in M} |\operatorname{tr}(X)|, \end{aligned}$$

---

Received by the editors April 1, 2004; revised February 19, 2005.  
 AMS subject classification: Primary: 15A52; secondary: 46B09, 46L54.  
 ©Canadian Mathematical Society 2006.

where the  $o$  is relative to increasing dimension but uniform in all other parameters. The extra term on the right is necessary, as can be checked by substituting  $M = \{I\}$ . However, this extra term cannot exceed the Gaussian term, and is usually considerably smaller. Note that the existence of zero-trace reflections implies that the dimension is even. The full statement of the result in the next section covers reflections with non-zero trace and odd dimension as well.

The motivation, as mentioned, arises from two different fields. First, as the relationship between Gaussian and Rademacher random variables proved a useful tool in the study of probability in Banach spaces, one should not overlook its counterpart in statistical physics and free probability, relating the GOE/GUE to reflections, as well as their free probability analogues (even though, admittedly, the level of resolution of the inequality presented here is finer than what free probability currently requires).

The other motivation comes from the geometry of Banach spaces. Given two spaces,  $A$  and  $B$ , the infimum of  $\|T\|_{A \rightarrow B} \|T^{-1}\|_{B \rightarrow A}$  over all linear transformations  $T$  is called the *Banach–Mazur distance*. This parameter measures how similar the spaces  $A$  and  $B$  are, when choice of basis or coordinates is ignored. One way to bound this infimum from above is to cleverly define a Euclidean structure relative to  $A$  and  $B$ , and then estimate the expected value of  $\|U\|_{A \rightarrow B} \|U^{-1}\|_{B \rightarrow A}$ , where  $U$  is a random rotation (orthogonal matrix). Applying the Cauchy–Schwarz inequality and the fact that  $\|U\|_{A \rightarrow B}$  can be written as  $\sup_{X \in M} |\text{tr}(UX)|$  for some  $M$ , we are left with the task of estimating  $(E_U \sup_{X \in M} |\text{tr}(UX)|^2)^{1/2}$ . This is where the Marcus–Pisier inequality comes in handy. This method has been applied in [DMT], [BG], [BM], and is reviewed in [T].

Now if we obtain, as we do in this paper, an estimate for reflections rather than rotations, we can apply this method with reflections rather than rotations, and get an interesting bonus: whatever distance estimate is obtained by a reflection operator, it is already applicable to the initial relative position of  $A$  and  $B$  (namely, the distance estimate is already obtained by the identity operator), provided we restrict our attention to either of two orthogonally complementing subspaces—the eigenspaces of the reflection.

The section below clarifies some points regarding the setting, and then states and proves our inequalities. We continue to comment briefly on the relation between the symmetric and non-symmetric cases. The last section contains two auxiliary Lemmas. We will use complex terminology, but the same proofs hold for the real case as well.

## 2 Statements and Proofs

In this section  $M$  is any collection of  $n$  by  $n$  matrices. We consider the following  $n$  by  $n$  matrix-valued random variables:  $\hat{G}$  is an element of the GUE ( $\hat{G}_{ij} = a_{ij} + ib_{ij}$ , where  $a_{ij}$  and  $b_{ij}$  are independent  $N(0, 1/n)$  real Gaussians for  $i < j$ ,  $\hat{G}_{ij} = a_{ji} - ib_{ji}$  for  $i > j$ , and  $\hat{G}_{ii}$  are independent real  $N(0, 2/n)$  Gaussians),  $U$  is a random unitary matrix, and  $R_t$  is a random reflection (self-adjoint unitary matrix) with trace  $t$ . The measure according to which we draw our random reflections is induced by the adjoint action of  $U(n)$  on reflections with trace  $t$ . Explicitly, a random reflection with trace  $t$

is the product

$$U \begin{pmatrix} I_{\frac{n+t}{2}} & 0 \\ 0 & -I_{\frac{n-t}{2}} \end{pmatrix} U^*,$$

where  $U$  is a random unitary matrix. This choice coincides with drawing a reflection via a random choice of eigenspaces according to the Haar measure on the Grassmannian.

Our notation of variables suppresses the dimension  $n$ , but wherever there is a  $o$  notation, it refers to increasing  $n$ . Note that our results are asymptotically correct for general  $M$ , but that, as the proofs show, information on  $M$  can induce substantial improvements.

In order not to interrupt the proof, we will consider here some properties of the eigenvalues of  $\hat{G}$ . The quantities which are of interest to us are  $\lambda_1$ , the expected largest eigenvalue, the expected norm of  $\hat{G}$ , which equals the expectation of  $\max\{\lambda_1(\hat{G}), -\lambda_n(\hat{G})\}$ , and  $\bar{\lambda}_k$ , the expected value of the average of the largest  $k$  eigenvalues of  $\hat{G}$ . We recall that  $\lambda_1$  and the expected norm converge to 2 as the dimension increases (see [BY] and [G]), and that they are strictly smaller than 2 for any dimension (proved in [DS] Theorem 2.11, and [S1] Appendix C). We also use the fact that

$$\lim_{\frac{k}{n} \rightarrow c} \bar{\lambda}_k = \frac{\int_{2-4c}^2 x \sqrt{4-x^2} dx}{\int_{2-4c}^2 \sqrt{4-x^2} dx},$$

in particular  $\bar{\lambda}_{n/2}$  converges to  $\frac{8}{3\pi}$ . This can be proved either by expressing this average via Hermite polynomials and applying their asymptotic formulas, or alternatively, by combining Arnold's strengthening [A] of Wigner's semicircle law with a concentration inequality for the least eigenvalue. For further details refer to the surveys [DS], [HT] and [PL]. If exact constants are disregarded, then all we really use is that  $\bar{\lambda}_k$  are uniformly bounded for all  $n$  and  $1 \leq k \leq n/2$ .

**Theorem 1** For every  $1 \leq p \leq \infty$  and every integer  $0 \leq t \leq n - 1$  with the same parity as the dimension  $n$  we have:

$$C(t + 1) \left( E_{\hat{G}} \sup_{X \in M} |\text{tr}(\hat{G}X)|^p \right)^{1/p} \geq \left( E_{R_t} \sup_{X \in M} |\text{tr}(R_t X)|^p \right)^{1/p},$$

where  $C$  is a universal constant.

Note that for the purpose of the inequality in the theorem,  $R_t$  and  $R_{-t}$  are equivalent, so we need only consider positive  $t$ . To place the theorem in the free-probability context, note that the proof still holds if we replace  $G$  and  $R_t$  by sequences of such matrices, and the collection  $M$  by a collection of sequences (yielding norms on matrix sequences).

**Proof** The first manipulation consists of inserting unitary matrices into our expression,

$$\left( E_{\hat{G}} \sup_{X \in M} |\text{tr}(\hat{G}X)|^p \right)^{1/p} = \left( E_{\hat{G}} \sup_{X \in M} |\text{tr}(U \hat{G} U^* X)|^p \right)^{1/p},$$

relying on the fact that GUE is invariant under unitary conjugation. We may now integrate over the unitary group and get that the above expression equals

$$(E_U E_{\hat{G}} \sup_{X \in M} |\text{tr}(U \hat{G} U^* X)|^p)^{1/p}.$$

Next, we decompose the self-adjoint  $\hat{G}$  into a product  $U_{\hat{G}} D(\hat{G}) U_{\hat{G}}^*$ , where the  $U_{\hat{G}}$ 's are unitary, and  $D(\hat{G})$  is real and diagonal ( $D(\hat{G})$  may not be unique, because there is no constraint over the order of the entries on the diagonal; at this point any measurable selection will do). Changing the order of integration allows to absorb the  $U_{\hat{G}}$ 's into the  $U$ 's, so we get

$$(E_{\hat{G}} E_U \sup_{X \in M} |\text{tr}(U D(\hat{G}) U^* X)|^p)^{1/p},$$

which leads us to the inequality

$$(1) \quad (E_{\hat{G}} \sup_{X \in M} |\text{tr}(\hat{G} X)|^p)^{1/p} \geq (E_U \sup_{X \in M} |\text{tr}(U (E_{\hat{G}} D(\hat{G})) U^* X)|^p)^{1/p}.$$

In order to analyse the expression  $E_{\hat{G}} D(\hat{G})$ , recall that  $D(\hat{G})$  can consist of any ordering of the eigenvalues on the diagonal. We may replace  $D(\hat{G})$ , then, by the average of the  $\frac{n+t}{2}! \frac{n-t}{2}!$  orderings which restrict the largest  $\frac{n+t}{2}$  entries to the top of the diagonal, and the smallest  $\frac{n-t}{2}$  entries to the bottom. This average results in a matrix of the form

$$\begin{pmatrix} \lambda_{\frac{n+t}{2}}(\hat{G}) I_{\frac{n+t}{2}} & 0 \\ 0 & \mu_{\frac{n-t}{2}}(\hat{G}) I_{\frac{n-t}{2}} \end{pmatrix},$$

where  $\lambda_{\frac{n+t}{2}}(\hat{G})$  is the average of the largest  $\frac{n+t}{2}$  eigenvalues of  $\hat{G}$ , and  $\mu_{\frac{n-t}{2}}(\hat{G})$  is the average of the smallest  $\frac{n-t}{2}$  eigenvalues. Since  $\hat{G}$  and  $-\hat{G}$  have the same distribution, the expected trace must be 0. Therefore, applying  $E_{\hat{G}}$  to our replacement for  $D(\hat{G})$ , we are left with

$$\begin{pmatrix} \bar{\lambda}_{\frac{n+t}{2}} I_{\frac{n+t}{2}} & 0 \\ 0 & -\frac{n+t}{n-t} \bar{\lambda}_{\frac{n+t}{2}} I_{\frac{n-t}{2}} \end{pmatrix} = \frac{n}{n-t} \bar{\lambda}_{\frac{n+t}{2}} \begin{pmatrix} I_{\frac{n+t}{2}} & 0 \\ 0 & -I_{\frac{n-t}{2}} \end{pmatrix} - \frac{t}{n-t} \bar{\lambda}_{\frac{n+t}{2}} I_n,$$

where  $\bar{\lambda}_k$  is the expected value of the average of the largest  $k$  eigenvalues of  $\hat{G}$ .

Plugging this into (1) we get

$$(E_{\hat{G}} \sup_{X \in M} |\text{tr}(\hat{G} X)|^p)^{1/p} \geq \frac{n}{n-t} \bar{\lambda}_{\frac{n+t}{2}} (E_{R_t} \sup_{X \in M} |\text{tr}(R_t X)|^p)^{1/p} - \frac{t}{n-t} \bar{\lambda}_{\frac{n+t}{2}} \sup_{X \in M} |\text{tr}(X)|.$$

As Lemma 4 below shows, we can replace the scalar term by a Gaussian term. Carrying it to the left-hand side yields

$$\left( \frac{n-t}{n \bar{\lambda}_{\frac{n+t}{2}}} + \frac{\sqrt{\pi} t}{2} \right) (E_{\hat{G}} \sup_{X \in M} |\text{tr}(\hat{G} X)|^p)^{1/p} \geq (E_{R_t} \sup_{X \in M} |\text{tr}(R_t X)|^p)^{1/p}.$$

Recalling the estimates for  $\bar{\lambda}_k$  formulated before the statement of the theorem concludes the proof both as stated and for the case  $t = 0$  presented in the introduction. ■

**Theorem 2** *The following inequalities hold:*

(1) *For even dimension  $n$ , and  $t = 0$*

$$E_{\hat{G}}\left(\sup_{X \in M} |\text{tr}(\hat{G}X)|\right) \leq (2 + o(1)) E_{R_0}\left(\sup_{X \in M} |\text{tr}(R_0X)|\right) + \frac{2}{\sqrt{\pi n}} \sup_{X \in M} |\text{tr}(X)|.$$

(2) *For every dimension  $n$  and every  $0 < t \leq n - 2$  with the same parity as  $n$*

$$E_{\hat{G}}\left(\sup_{X \in M} |\text{tr}(\hat{G}X)|\right) \leq \frac{Cn}{n - t} E_{R_t}\left(\sup_{X \in M} |\text{tr}(R_tX)|\right),$$

where  $C$  is an absolute constant.

**Proof** As in the previous proof, the first step is to insert unitary matrices:

$$E_{\hat{G}}\left(\sup_{X \in M} |\text{tr}(\hat{G}X)|\right) = E_U E_{\hat{G}}\left(\sup_{X \in M} |\text{tr}(U\hat{G}U^*X)|\right).$$

Next, let's add and subtract a scalar matrix, so that  $\hat{G}$  is split into constant times a normalised  $t$ -trace reflection plus a scalar matrix. Note first that the absolute value around the trace in the inequality allows us to replace  $\hat{G}$  by  $-\hat{G}$ , and therefore assume that  $\text{tr}(\hat{G}) < 0$ . Now set  $c(\hat{G})$  to be the positive real number for which

$$\frac{\text{tr}(\hat{G} + c(\hat{G})I_n)}{\|\hat{G} + c(\hat{G})I_n\|_{\ell_2 \rightarrow \ell_2}} = t.$$

Such  $c$  exists due to continuity and to the fact that, when  $c$  converges to  $\infty$ , the expression tends to  $n$ . We get

$$E_{\hat{G}}\left(\sup_{X \in M} |\text{tr}(\hat{G}X)|\right) \leq E_{\hat{G}}(|c(\hat{G})|) \sup_{X \in M} |\text{tr}(X)| + E_{\hat{G}} E_U \left( \|\hat{G} + c(\hat{G})I_n\| \sup_{X \in M} \left| \text{tr} \left( U \frac{\hat{G} + c(\hat{G})I_n}{\|\hat{G} + c(\hat{G})I_n\|} U^* X \right) \right| \right).$$

As stated by Lemma 5, a Hermitian matrix with trace  $t$ , whose  $\ell_2 \rightarrow \ell_2$  operator norm equals 1, is a convex combination of reflections with trace  $t$  (provided  $t$  and the dimension have the same parity). Supplanting such matrices for  $\frac{\hat{G} + c(\hat{G})I_n}{\|\hat{G} + c(\hat{G})I_n\|_{\ell_2 \rightarrow \ell_2}}$  yields

$$E_{\hat{G}}\left(\sup_{X \in M} |\text{tr}(\hat{G}X)|\right) \leq E_{\hat{G}}(|c(\hat{G})|) \sup_{X \in M} |\text{tr}(X)| + E_{\hat{G}}\left(\|\hat{G} + c(\hat{G})I_n\|_{\ell_2 \rightarrow \ell_2}\right) E_{R_t}\left(\sup_{X \in M} |\text{tr}(R_tX)|\right).$$

We are left with the task of estimating the expected value of  $c(\hat{G})$  and of the norm. Recall that  $c(\hat{G}) > 0$ . Using the triangle inequality for the operator norm, the definition of  $c$  yields

$$\text{tr}(\hat{G}) + nc(\hat{G}) \leq t(\|\hat{G}\|_{\ell_2 \rightarrow \ell_2} + c(\hat{G})),$$

and subsequently

$$c(\hat{G}) \leq \frac{t}{n-t} \|\hat{G}\|_{\ell_2 \rightarrow \ell_2} + \frac{|\text{tr}(\hat{G})|}{n-t}.$$

As recalled before the statement of the first theorem, the expected norm of  $\hat{G}$  is bounded by 2. Since, in addition,  $\text{tr}(\hat{G})$  is just  $\sqrt{2}$  times a normalised Gaussian, we get

$$E_{\hat{G}}(|c(\hat{G})|) \leq \frac{2t + \sqrt{4/\pi}}{n-t},$$

and therefore

$$E_{\hat{G}}(\|\hat{G} + c(\hat{G})I_n\|_{\ell_2 \rightarrow \ell_2}) \leq \frac{2n + \sqrt{4/\pi}}{n-t}.$$

Putting the estimates together we get

$$\begin{aligned} E_{\hat{G}}\left(\sup_{X \in M} |\text{tr}(\hat{G}X)|\right) &\leq \left(\frac{2n + \sqrt{4/\pi}}{n-t}\right) E_{R_t}\left(\sup_{X \in M} |\text{tr}(R_t X)|\right) \\ &\quad + \left(\frac{2t + \sqrt{4/\pi}}{n-t}\right) \sup_{X \in M} |\text{tr}(X)|. \end{aligned}$$

This already establishes the case  $t = 0$ . All that is left is to absorb the scalar term into the symmetric term for  $t > 0$ . This is easily accomplished by the inequality:

$$E_U |\text{tr}(UR_t U^* X)| \geq |\text{tr}(E_U (UR_t U^*) X)| = \frac{t}{n} |\text{tr}(X)|. \quad \blacksquare$$

**Remark 3** The relation between the self-adjoint and non self-adjoint cases is easy to express.

First, if we add a random Gaussian matrix to its adjoint we get  $\sqrt{2}$  times a member of GUE. Therefore,

$$\left(E_{\hat{G}} \sup_{X \in M} |\text{tr}(\hat{G}X)|^p\right)^{1/p} \leq \sqrt{2} \left(E_G \sup_{X \in M} |\text{tr}(GX)|^p\right)^{1/p}.$$

For the opposite direction, note that if we decompose  $G$  into self-adjoint and anti-self-adjoint parts, then the self-adjoint part is  $1/\sqrt{2}$  times a member of GUE. Therefore

$$\begin{aligned} \sqrt{2} \left(E_G \sup_{X \in M} |\text{tr}(GX)|^p\right)^{1/p} &\leq \left(E_{\hat{G}} \sup_{X \in M} |\text{tr}(\hat{G}X)|^p\right)^{1/p} \\ &\quad + \left(E_{\check{G}} \sup_{X \in M} |\text{tr}(\check{G}X)|^p\right)^{1/p}, \end{aligned}$$

where  $\check{G}$  stands for the anti-self-adjoint part. The anti-self-adjoint part is necessary, as can be verified by considering an  $M$  composed strictly of anti-self-adjoint matrices.

### 3 Lemmas

**Lemma 4** For any matrix collection  $M$  we have

$$\sup_{x \in M} |\operatorname{tr}(X)| \leq \frac{\sqrt{\pi n}}{2} \left( E_{\hat{G}} \sup_{X \in M} |\operatorname{tr}(\hat{G}X)|^p \right)^{1/p}.$$

**Proof** Take  $X \in M$  such that  $|\operatorname{tr}(X)|$  is maximal. Since  $\operatorname{tr}(\hat{G}X)$  is a scalar Gaussian with variance at least  $\frac{2}{n} \sum_i X_{ii}^2 \geq \frac{2}{n^2} (\operatorname{tr}(X))^2$ , we get  $E_{\hat{G}} |\operatorname{tr}(\hat{G}X)| \geq \frac{2}{\sqrt{\pi n}} |\operatorname{tr}(X)|$ . Adding the supremum and  $p$ -th moment may only increase the left hand side. ■

**Lemma 5** Let  $t$  be a non-negative integer of the same parity as the dimension  $n$ . Every Hermitian matrix  $T$ , such that  $\operatorname{tr}(T) = t$  and  $\|T\|_{\ell_2 \rightarrow \ell_2} = 1$ , is a convex combination of reflections with trace  $t$ .

**Proof** Consider first  $t = 0$  and even  $n$ . The extreme points of the set of matrices under hand is the set of zero-trace reflections. Indeed,  $T$  decomposes into  $UDU^*$ , with  $U$  unitary and  $D$  real diagonal, such that the diagonal entries form a zero-sum sequence bounded by  $\pm 1$ . It is easy to verify that the extreme points of the space of such sequences are zero-sum  $\pm 1$  sequences. Substituting such sequences into the diagonal matrix yields zero-trace reflections.

For general  $t$ , reduce as above to sequences bounded by  $\pm 1$  and summing to  $t$ . Appending to the sequences  $t$  entries equal to  $-1$ , we reduce the problem to the  $t = 0$  case, and are done. ■

**Acknowledgements** I would like to thank S. Szarek for useful comments, and G. Pisier for helpful clarifications.

### References

- [A] L. Arnold, *On the asymptotic distribution of the eigenvalues of random matrices*. J. Math. Anal. Appl. **20**(1967), 262–268.
- [BG] Y. Benyamini and Y. Gordon, *Random factorizations of operators between Banach spaces*. J. Anal. Math. **39**(1981), 45–74.
- [BM] J. Bourgain and V. Milman, *Distances between normed spaces, their subspaces and quotient spaces*. Integral Equations Operator Theory **9**(1986), 31–46.
- [BY] Z. D. Bai and Y. K. Yin, *Necessary and sufficient conditions for the almost sure convergence of the largest eigenvalue of a Wigner matrix*. Ann. Probab. **16**(1988), 1729–1741.
- [DMT] W. Davis, V. Milman and N. Tomczak-Jaegermann, *The distance between certain  $n$ -dimensional spaces*. Israel J. Math. **39**(1981), 1–15.
- [DS] K. R. Davidson and S. J. Szarek, *Local operator theory, random matrices and Banach spaces*. In: Handbook of the geometry of Banach spaces, volume I, (eds., W. B. Johnson and J. Lindenstrauss), North-Holland, 2001, 317–366.
- [G] S. Geman, *A limit theorem for the norm of random matrices*. Ann. Probab. **8**(1980), 252–261.
- [HT] U. Haagerup and S. Thorbjørnsen, *Random matrices with complex Gaussian entires*. Exposition. Math. **21**(2003), 293–337.
- [MP] M. B. Marcus and G. Pisier, *Random Fourier Series with Applications to Harmonic Analysis*. Center for Statistics and Probability, Northwestern University **44**, 1980.
- [PL] L. Pastur and A. Lejay, *Matrices aléatoires: statistique asymptotique des valeurs propres*. In: Séminaire de probabilités 23, Lecture Notes in Math. **1801**, Springer, 2003, 135–164.

- [S1] S. J. Szarek, *The volume of separable states is super-doubly-exponentially small*. preprint, <http://arxiv.org/abs/quant-ph/0310061>.
- [S2] ———, *Condition numbers of random matrices*. *J. Complexity* (2) 7(1991), 131–149.
- [T] N. Tomczak-Jaegermann, *Banach–Mazur Distance and finite-dimensional Operator Ideals*. Pitman Monographs and Surveys in Pure and Applied Math. **38**, 1989.

*Computer Science Department  
Academic College of Tel-Aviv-Yaffo  
Israel  
e-mail: rwagner@mta.ac.il*