

# *Algebra & Number Theory*

Volume 2

2008

No. 8

**The number of  $2 \times 2$  integer matrices having a  
prescribed integer eigenvalue**

Greg Martin and Erick Wong



**mathematical sciences publishers**

# The number of $2 \times 2$ integer matrices having a prescribed integer eigenvalue

Greg Martin and Erick Wong

What is the probability that an integer matrix chosen at random has a particular integer as an eigenvalue, or an integer eigenvalue at all? For a random real matrix, what is the probability of there being a real eigenvalue in a particular interval? This paper solves these questions for  $2 \times 2$  matrices, after specifying the probability distribution suitably.

## 1. Introduction

Random matrices arise in many mathematical contexts, and it is natural to ask about the properties that such matrices satisfy. If we choose a matrix with integer entries at random, for example, we would like to know the probability that it has a particular integer as an eigenvalue, or an integer eigenvalue at all. Similarly, if we choose a matrix with real entries at random, we would like to know the probability that it has a real eigenvalue in a particular interval. Certainly the answer depends on the probability distribution from which the matrix entries are drawn.

In this paper, we are primarily concerned with uniform distribution, so for both integer-valued and real-valued cases we must restrict the entries to a bounded interval. In [Martin and Wong 2008], we showed that random  $n \times n$  matrices of integers almost never have integer eigenvalues. An explicit calculation by Hetzel, Liew and Morrison [Hetzel et al. 2007] shows that a  $2 \times 2$  matrix with entries independently chosen uniformly from  $[-1, 1]$  has real eigenvalues with probability  $\frac{49}{72}$ . This calculation gives hope that our more precise questions about eigenvalues of a particular size might be accessible in the  $2 \times 2$  setting. The purpose of this paper is to resolve these questions, once we make them suitably precise.

For an integer  $k \geq 1$ , let  $\mathcal{M}_2(k)$  denote the uniform probability space of  $2 \times 2$  matrices of integers with absolute value at most  $k$ . Note that

$$|\mathcal{M}_2(k)| = (2k + 1)^4 = 16k^4 + O(k^3).$$

---

*MSC2000:* primary 15A36, 15A52; secondary 11C20, 15A18, 60C05.

*Keywords:* random matrix, eigenvalue, integer eigenvalue, integer matrix, distribution of eigenvalues.

We will obtain exact asymptotics for the number of matrices in  $\mathcal{M}_2(k)$  having integer eigenvalues and, more precisely, for the number of matrices with a given integer eigenvalue  $\lambda$ .

For any integer  $\lambda$ , define

$$\mathcal{M}_2^\lambda(k) = \{M \in \mathcal{M}_2(k) : \lambda \text{ is an eigenvalue of } M\},$$

and let

$$\mathcal{M}_2^{\mathbb{Z}}(k) = \bigcup_{\lambda \in \mathbb{Z}} \mathcal{M}_2^\lambda(k).$$

**Theorem 1.** *Define the function  $V : [-2, 2] \rightarrow \mathbb{R}$  by  $V(-\delta) = V(\delta)$  and*

$$V(\delta) = \begin{cases} 4 - 2\delta - \delta^2 + \delta^2 \log(1 + \delta) - 2(1 - \delta) \log(1 - \delta) & \text{if } 0 \leq \delta < 1, \\ 1 + \log 2 & \text{if } \delta = 1, \\ 4 - 2\delta - \delta^2 + \delta^2 \log(\delta + 1) + 2(\delta - 1) \log(\delta - 1) & \text{if } 1 < \delta \leq \sqrt{2}, \\ \delta^2 - 2\delta - (\delta^2 - 2\delta + 2) \log(\delta - 1) & \text{if } \sqrt{2} < \delta \leq 2 \end{cases} \quad (1)$$

(where  $\log$  is the natural logarithm). Then for any integer  $\lambda$  between  $-2k$  and  $2k$ ,

$$|\mathcal{M}_2^\lambda(k)| = \frac{24V(\lambda/k)}{\pi^2} k^2 \log k + O(k^2), \quad (2)$$

where the implied constant is absolute. On the other hand, if  $|\lambda| > 2k$  then  $\mathcal{M}_2^\lambda(k)$  is empty.

We remark that the function  $V(\delta)$  is continuous and, with the exception of the points of infinite slope at  $\delta = \pm 1$ , differentiable everywhere (even at  $\delta = \pm 2$ , if we imagine that  $V(\delta)$  is defined to be 0 when  $|\delta| > 2$ ). Notice that (2) is technically not an asymptotic formula when  $\lambda$  is extremely close to  $\pm 2k$ , because then the value of  $V(\lambda/k)$  can have order of magnitude  $1/\log k$  or smaller, making the “main term” no bigger than the error term. However, (2) is truly an asymptotic formula for  $|\lambda| < 2k - \psi(k)k/(\log k)^{1/3}$ , where  $\psi(k)$  is any function tending to infinity (the exponent  $\frac{1}{3}$  arises because  $V(\delta)$  approaches 0 cubically as  $\delta$  tends to 2 from below).

By summing the formula (2) over all possible values of  $\lambda$ , we obtain an asymptotic formula for  $|\mathcal{M}_2^{\mathbb{Z}}(k)|$ . We defer the details of the proof to Section 3.

**Corollary 2.** *Let*

$$C = \frac{7\sqrt{2} + 4 + 3 \log(\sqrt{2} + 1)}{3\pi^2} \approx 0.55873957.$$

The probability that a randomly chosen matrix in  $\mathcal{M}_2(k)$  has integer eigenvalues is asymptotically  $C(\log k)/k$ . More precisely,

$$|\mathcal{M}_2^{\mathbb{Z}}(k)| = 16Ck^3 \log k + O(k^3).$$

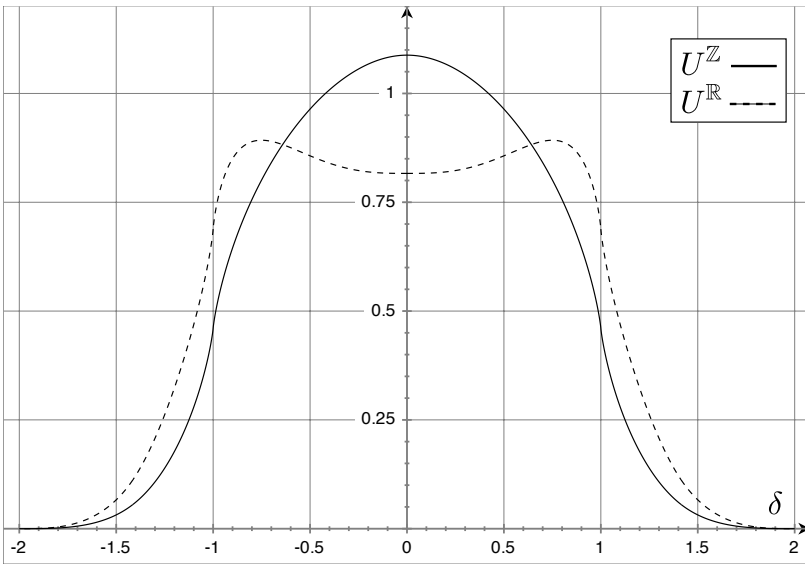
If  $M \in \mathcal{M}_2(k)$  has eigenvalue  $\lambda$ , then the scaled matrix  $k^{-1}M$  has eigenvalue  $\lambda/k$ , which is the argument of  $V$  that appears on the right-hand side of (2). Thus one interpretation of Theorem 1 is that for large  $k$ , the rational eigenvalues of  $k^{-1}M$  tend to be distributed like the function  $V$ .

Note that the entries of  $k^{-1}M$  are sampled uniformly from a discrete, evenly-spaced subset of  $[-1, 1]$ . As  $k \rightarrow \infty$  this probability distribution converges in law to the uniform distribution on the interval  $[-1, 1]$ . Let  $\mathcal{M}_2([-1, 1])$  denote the probability space of all  $2 \times 2$  matrices whose entries are independent random variables drawn from this distribution. One might ask whether the distribution given by Theorem 1 is just a discrete approximation to the distribution of eigenvalues in  $\mathcal{M}_2([-1, 1])$ ; the answer, perhaps surprisingly, is no. The next theorem provides this latter distribution.

**Theorem 3.** *Define  $W(\delta)$  to be the density function for real eigenvalues of matrices in  $\mathcal{M}_2([-1, 1])$ : if  $M$  is a randomly chosen matrix from  $\mathcal{M}_2([-1, 1])$ , then the expected number of eigenvalues of  $M$  in the interval  $[s, t]$  is  $\int_s^t W(\delta) d\delta$ . Then  $W(-\delta) = W(\delta)$  and*

$$W(\delta) = \begin{cases} \frac{80 + 20\delta + 90\delta^2 + 52\delta^3 - 107\delta^4}{144(1+\delta)} - \frac{1}{4}\delta(1-\delta^2)\log(1+\delta) \\ \qquad - \frac{1}{12}(5 - 7\delta + 8\delta^2)(1-\delta)\log(1-\delta) & \text{if } 0 \leq \delta \leq 1, \\ \frac{\delta(20 + 10\delta - 12\delta^2 - 3\delta^3)}{16(1+\delta)} + \frac{1}{4}\delta(\delta^2 - 1)\log(\delta + 1) \\ \qquad + \frac{1}{4}(3\delta - 1)(\delta - 1)\log(\delta - 1) & \text{if } 1 \leq \delta \leq \sqrt{2}, \\ \frac{\delta(\delta - 2)(2 - 6\delta + 3\delta^2)}{16(\delta - 1)} - \frac{1}{4}(\delta - 1)^3\log(\delta - 1) & \text{if } \sqrt{2} \leq \delta \leq 2, \\ 0 & \text{if } \delta \geq 2. \end{cases}$$

As in the case of  $V(\delta)$ , the function  $W(\delta)$  is continuous and differentiable everywhere, with the exception of the points of infinite slope at  $\delta = \pm 1$ . (The value  $W(1) = \frac{15}{32}$  makes the function continuous there, although the value of a density function at a single point is irrelevant to the probability distribution.) It also shares the same cubic decay as  $\delta$  tends to 2 from below. However, there are obvious qualitative differences between the functions  $V$  and  $W$ . In Figure 1 we plot  $V$  and  $W$  on the same axes, normalized so that the area under each is 2 (these normalized versions are denoted  $U^{\mathbb{Z}}$  and  $U^{\mathbb{R}}$  in our earlier paper [Martin and Wong 2008]). In the case of  $\mathcal{M}_2(k)$ , this normalization corresponds to conditioning on having integer eigenvalues, that is, scaling by the probability  $C(\log k)/k$  from Corollary 2. For  $\mathcal{M}_2([-1, 1])$  we are conditioning on having real eigenvalues, which occurs with probability  $\frac{49}{72}$  (this can be obtained by integrating  $W(\delta)$ , analogously to the proof of Corollary 2; the computation in [Hetzl et al. 2007] is more direct).



**Figure 1.** Graph of  $U^{\mathbb{Z}}$  ( $V$  normalized) versus  $U^{\mathbb{R}}$  ( $W$  normalized).

Note that the distribution  $W(\delta)$  is bimodal, having its maxima at

$$\delta \approx \pm 0.75030751.$$

Thus, a random matrix in  $\mathcal{M}_2([-1, 1])$  is more likely to have an eigenvalue of magnitude near  $\frac{3}{4}$  than one of magnitude near 0. We expect this would still hold if we were to condition on matrices in  $\mathcal{M}_2([-1, 1])$  having rational eigenvalues, since any matrix with real eigenvalues is a small perturbation from one with rational eigenvalues. That this is not true for  $V(\delta)$  shows that the eigenvalue distribution of  $\mathcal{M}_2(k)$  is not purely the result of magnitude considerations but also encodes some of the arithmetic structure of the integers up to  $k$ .

We remark that [Theorem 1](#) can also be obtained from a powerful result of Katznelson [[1993](#)]. Let  $\mathcal{B}$  be a convex body containing the origin in  $\mathbb{R}^4$ , and embed the set of  $2 \times 2$  integer matrices

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

as lattice points  $(a, b, c, d) \in \mathbb{Z}^4$ . Then [[Katznelson 1993](#), Theorem 1] gives an asymptotic formula for the number of singular integer matrices inside the dilate  $k \cdot \mathcal{B}$ . Taking

$$\mathcal{B} = [-1, 1]^4$$

then yields an asymptotic formula for  $|M_2^0(k)|$ , and more generally one can obtain  $|M_2^\lambda(k)|$  by adding and subtracting appropriate shifts of  $\mathcal{B}$ . The asymptotic formula in [[Katznelson 1993](#)] is defined in terms of an unusual singular measure

supported on the Zariski-closed subset of  $\mathbb{R}^4$  corresponding to singular matrices. The explicit computation of this measure is roughly analogous to our case-by-case considerations in Section 4, modulo the significant complications of carrying error terms. Our techniques are more elementary, but Katznelson’s results apply in theory to matrices of any size, whereas our methods become unwieldy even for  $3 \times 3$  matrices.

In the case of  $n \times n$  matrices with entries independently chosen from a Gaussian distribution, a great deal more is known. Edelman [1997] has computed the exact distributions of the real and complex eigenvalues for any  $n$ , as well as the number of real eigenvalues (for instance, the probability of having all  $n$  eigenvalues real is precisely  $2^{-n(n-1)/4}$ ). As  $n \rightarrow \infty$ , the real eigenvalues, suitably rescaled by a factor of  $1/\sqrt{n}$ , converge to the uniform distribution on  $[-1, 1]$ . Similarly, the complex eigenvalues converge to the “circular law” predicted by Girko [1984], namely the uniform distribution on the unit disk centered at the origin. Very recently, Tao and Vu [2008] have shown that the circular law is universal: one can replace the Gaussian distribution by an arbitrary distribution with mean 0 and variance 1. Similar results have been established for random symmetric matrices with entries independently chosen from a Gaussian distribution (Wigner law) or from other distributions.

Those who are interested in the connections between analytic number theory and random matrix theory might wonder whether those connections are related to the present paper. The matrices in that context, however, are selected from classical matrix groups, such as the group of  $n \times n$  Hermitian matrices, randomly according to the Haar measures on the groups. The relationship to our results is therefore minimal.

## 2. Preliminaries about matrices

We begin with some elementary observations about  $2 \times 2$  matrices that will simplify our computations. The first lemma explains why the functions  $V$  and  $W$  are supported only on  $[-2, 2]$ .

**Lemma 4.** *Any eigenvalue of a matrix in  $\mathcal{M}_2(k)$  is bounded in absolute value by  $2k$ . Any eigenvalue of a matrix in  $\mathcal{M}_2([-1, 1])$  is bounded in absolute value by 2.*

*Proof.* We invoke Gershgorin’s circle theorem [1931], a standard result in spectral theory: let  $M = (m_{ij})$  be an  $n \times n$  matrix, and let  $D(z, r)$  denote the disk of radius  $r$  around the complex number  $z$ . Then Gershgorin’s theorem says that all of the eigenvalues of  $M$  must lie in the union of the disks

$$D\left(m_{11}, \sum_{\substack{1 \leq j \leq n \\ j \neq 1}} |m_{1j}|\right), \quad D\left(m_{22}, \sum_{\substack{1 \leq j \leq n \\ j \neq 2}} |m_{2j}|\right), \quad \dots, \quad D\left(m_{nn}, \sum_{\substack{1 \leq j \leq n \\ j \neq n}} |m_{nj}|\right).$$

In particular, if all of the entries of  $M$  are bounded in absolute value by  $B$ , then all the eigenvalues are bounded in absolute value by  $nB$ .  $\square$

The key to the precise enumeration of  $\mathcal{M}_2^\lambda(k)$  is the simple structure of singular integer matrices:

**Lemma 5.** *For any singular matrix  $M \in \mathcal{M}_2(k)$ , either at least two entries of  $M$  equal zero, or else there exist nonzero integers  $a, b, c, d$  with  $(a, b) = 1$  such that*

$$M = \begin{pmatrix} ac & bc \\ ad & bd \end{pmatrix}. \quad (3)$$

*This representation of  $M$  is unique up to replacing each of  $a, b, c, d$  by its negative.*

*Proof.* If one of the entries of  $M$  equals zero, then a second one must equal zero as well for the determinant to vanish. Otherwise, given

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

with none of the  $m_{ij}$  equal to zero, define  $c = (m_{11}, m_{12})$ , and set  $a = m_{11}/c$  and  $b = m_{12}/c$ , so that  $(a, b) = 1$ . Since  $M$  is singular, the second row of  $m$  must be a multiple of the first row, that is, there exists a real number  $d$  such that  $m_{21} = ad$  and  $m_{22} = bd$ . And since  $a$  and  $b$  are relatively prime,  $d$  must in fact be an integer.

This argument shows that every such matrix has one such representation. If

$$M = \begin{pmatrix} a'c' & b'c' \\ a'd' & b'd' \end{pmatrix}$$

is another such representation, then  $(a', b') = 1$  implies  $(a'c', b'c') = |c'|$ , which shows that  $|c'| = c$ ; the equalities  $|a'| = |a|$ ,  $|b'| = |b|$ , and  $|d'| = |d|$  follow quickly.  $\square$

For a  $2 \times 2$  matrix  $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ , we define

$$\text{disc } M = (\text{tr } M)^2 - 4 \det M = (a - d)^2 + 4bc.$$

It is easily seen that this is the discriminant of the characteristic polynomial of  $M$ . We record the following elementary facts, which will be useful in the proofs of [Lemma 7](#) and [Proposition 13](#).

**Lemma 6.** *Let  $M$  be a  $2 \times 2$  matrix with real entries*

- (a)  *$M$  has repeated eigenvalues if and only if  $\text{disc } M = 0$ .*
- (b)  *$M$  has real eigenvalues if and only if  $\text{disc } M \geq 0$ .*
- (c)  *$\det M < 0$  if and only if  $M$  has two real eigenvalues of opposite sign.*
- (d) *If  $\det M > 0$  and  $\text{disc } M \geq 0$ , then the eigenvalues of  $M$  have the same sign as  $\text{tr } M$ .*

*Proof.* Let  $\lambda_1, \lambda_2$  denote the eigenvalues of  $M$ , so

$$\operatorname{tr} M = \lambda_1 + \lambda_2, \quad \det M = \lambda_1 \lambda_2 \quad \text{and} \quad \operatorname{disc} M = (\lambda_1 - \lambda_2)^2,$$

each of which is real. Parts (a), (b) and (d) follow immediately from these observations, and part (c) from the fact that  $\lambda_2 = \overline{\lambda_1}$  if  $\lambda_1, \lambda_2$  are complex.  $\square$

The next lemma gives a bound for the probability of a matrix having repeated eigenvalues. It is natural to expect this probability to converge to 0 as  $k$  increases, and indeed such a result was obtained in [Hetzel et al. 2007] for matrices of arbitrary size. We give a simple proof of a stronger bound for the  $2 \times 2$  case, as well as an analogous qualitative statement for real matrices which will be helpful in the proof of Theorem 3.

**Lemma 7.** *The number of matrices in  $\mathcal{M}_2(k)$  with a repeated eigenvalue, for every  $\varepsilon > 0$ , is  $\ll_\varepsilon k^{2+\varepsilon}$ . The probability that a random matrix in  $\mathcal{M}_2([-1, 1])$  has a repeated eigenvalue or is singular is 0.*

*Proof.* By Lemma 6 (a), the  $2 \times 2$  matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  has a double eigenvalue if and only if  $4bc = -(a - d)^2$ . For matrices in  $\mathcal{M}_2([-1, 1])$  this is easily seen to be a zero-probability event, as is the event that  $\det M = ad - bc = 0$ .

For matrices in  $\mathcal{M}_2(k)$ , we enumerate how many can satisfy  $4bc = -(a - d)^2$ . If  $a = d$  then there are  $4k + 1$  choices for  $b, c \in \{-k, \dots, k\}$ ; otherwise there are at most  $2\tau((a - d)^2/4)$  choices if  $a \equiv d \pmod{2}$  and no choices otherwise. (Here  $\tau(n)$  is the number-of-divisors function; the factor of 2 comes from the fact that  $b$  and  $c$  can be positive or negative, while the “at most” is due to the fact that not all factorizations of  $-(a - d)^2$  result in two factors not exceeding  $k$ .) Therefore the number of matrices in  $\mathcal{M}_2(k)$  with a repeated eigenvalue is at most

$$\sum_{|a| \leq k} (4k + 1) + 2 \sum_{|a| \leq k} \sum_{\substack{|d| \leq k, d \neq a \\ a \equiv d \pmod{2}}} \tau\left(\frac{(a-d)^2}{4}\right) \ll_\varepsilon k^{2+\varepsilon},$$

where the inequality follows from  $(a - d)^2/4 \leq k^2$  and the well-known fact that  $\tau(n) \ll_\varepsilon n^\varepsilon$  for any  $\varepsilon > 0$  — see for instance [Montgomery and Vaughan 2007, p. 56].  $\square$

### 3. Enumeration theorems for integer eigenvalues

Let  $\mu(n)$  be the Möbius function, characterized by the identity

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases} \tag{4}$$



The well-known Dirichlet series identity  $1/\zeta(s) = \sum_{n=1}^{\infty} \mu(n)n^{-s}$  is valid for  $\text{Re } s > 1$  (see [Montgomery and Vaughan 2007, Corollary 1.10], for example). In particular,

$$\sum \frac{\mu(n)}{n^2} = \frac{6}{\pi^2},$$

and we can estimate the tail of this series (using  $|\mu(n)| \leq 1$ ) to obtain the quantitative estimate

$$\sum_{d \leq k} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2} + O\left(\frac{1}{k}\right). \tag{5}$$

**Lemma 8.** For nonzero integers  $a, b$  and parameters  $k, \lambda$ , define the function

$$N_{k,\lambda}(a, b) = \#\{(c, d) \in \mathbb{Z}^2, c \neq 0, d \neq 0 : |ac + \lambda|, |bc|, |ad|, |bd + \lambda| \leq k\}. \tag{6}$$

Then

$$|\mathcal{M}_2^\lambda(k)| = 4 \sum_{d \leq k} \mu(d) \sum_{1 \leq \alpha < \beta \leq k/d} N_{k,\lambda}(d\alpha, d\beta) + O(k^2),$$

where the implied constant is independent of  $\lambda$  and  $k$ .

*Proof.* Fix an integer  $0 \leq \lambda \leq 2k$ , and let  $M \in \mathcal{M}_2^\lambda(k)$ , so that  $M - \lambda I$  is singular. By Lemma 5, either at least two entries of  $M - \lambda I$  equal zero, or else  $M - \lambda I$  has exactly two representations of the form (3). In the former case, there are  $2k + 1$  choices for each of the two potentially nonzero entries, hence  $O(k^2)$  such matrices in total (even taking into account the several different choices of which two entries are nonzero). In the latter case, there are exactly two corresponding quadruples  $a, b, c, d$  of integers as in Lemma 5. Taking into account that each entry of  $M$  must be at most  $k$  in absolute value, we deduce that

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= \frac{1}{2} \#\{(a, b, c, d) \in \mathbb{Z}^4 : a, b, c, d \neq 0, (a, b) = 1, \\ &\quad |ac + \lambda|, |bc|, |ad|, |bd + \lambda| \leq k\} + O(k^2) \\ &= \frac{1}{2} \sum_{\substack{1 \leq |a|, |b| \leq k \\ (a,b)=1}} N_{k,\lambda}(a, b) + O(k^2). \end{aligned}$$

Because of the symmetries  $N_{k,\lambda}(a, b) = N_{k,\lambda}(-a, b) = N_{k,\lambda}(a, -b) = N_{k,\lambda}(-a, -b)$ , we have

$$|\mathcal{M}_2^\lambda(k)| = 2 \sum_{\substack{1 \leq a, b \leq k \\ (a,b)=1}} N_{k,\lambda}(a, b) + O(k^2).$$

The only term in the sum where  $a = b$  is the term  $a = b = 1$ , and for all other terms we can invoke the additional symmetry  $N_{k,\lambda}(a, b) = N_{k,\lambda}(b, a)$ , seen to be valid by switching the roles of  $c$  and  $d$  in the definition (6) of  $N_{k,\lambda}(a, b)$ . We obtain

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= 4 \sum_{\substack{1 \leq a < b \leq k \\ (a,b)=1}} N_{k,\lambda}(a, b) + 2N_{k,\lambda}(1, 1) + O(k^2) \\ &= 4 \sum_{\substack{1 \leq a < b \leq k \\ (a,b)=1}} N_{k,\lambda}(a, b) + O(k^2), \end{aligned}$$

where the last step used the fact that  $N_{k,\lambda}(1, 1) \leq \#\{(c, d) \in \mathbb{Z}^2 : |c|, |d| \leq k\} \ll k^2$ .

Using the characteristic property of the Möbius function (4), we can write the last expression as

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= 4 \sum_{1 \leq a < b \leq k} N_{k,\lambda}(a, b) \sum_{d|(a,b)} \mu(d) + O(k^2) \\ &= 4 \sum_{d \leq k} \mu(d) \sum_{\substack{1 \leq a < b \leq k \\ d|a, d|b}} N_{k,\lambda}(a, b) + O(k^2) \\ &= 4 \sum_{d \leq k} \mu(d) \sum_{1 \leq \alpha < \beta \leq k/d} N_{k,\lambda}(d\alpha, d\beta) + O(k^2). \quad \square \end{aligned}$$

**Lemma 9.** *Let  $k$  and  $\lambda$  be integers with  $0 \leq \lambda \leq 2k$ , and let  $x$  and  $y$  be integers with  $1 \leq x \leq y \leq k$ . Then*

$$N_{k,\lambda}(x, y) = k^2 C\left(\frac{\lambda}{k}; x, y\right) D\left(\frac{\lambda}{k}; x, y\right) + O\left(\frac{k}{y}\right),$$

where

$$C(\delta; x, y) = \max\left\{0, \min\left\{\frac{1-\delta}{x} + \frac{1}{y}, \frac{2}{y}\right\}\right\}, \quad D(\delta; x, y) = \min\left\{\frac{1-\delta}{y} + \frac{1}{x}, \frac{2}{y}\right\}. \quad (7)$$

*Proof.* We have

$$\begin{aligned} N_{k,\lambda}(x, y) &= \#\{(c, d) \in \mathbb{Z}^2, c \neq 0, d \neq 0 : |xc + \lambda|, |yc|, |xd|, |yd + \lambda| \leq k\} \\ &= \#\{c \in \mathbb{Z}, c \neq 0 : -k \leq xc + \lambda \leq k, -k \leq yc \leq k\} \\ &\quad \times \#\{d \in \mathbb{Z}, d \neq 0 : -k \leq xd \leq k, -k \leq yd + \lambda \leq k\}. \end{aligned}$$

Since  $x$  and  $y$  are positive, we can rewrite this product as

$$\begin{aligned} N_{k,\lambda}(x, y) &= \#\left\{c \in \mathbb{Z}, c \neq 0 : \frac{-k-\lambda}{x} \leq c \leq \frac{k-\lambda}{x}, -\frac{k}{y} \leq c \leq \frac{k}{y}\right\} \\ &\quad \times \#\left\{d \in \mathbb{Z}, d \neq 0 : -\frac{k}{x} \leq d \leq \frac{k}{x}, \frac{-k-\lambda}{y} \leq d \leq \frac{k-\lambda}{y}\right\} \\ &= \#\left\{c \in \mathbb{Z}, c \neq 0 : -\frac{k}{y} \leq c \leq \min\left\{\frac{k-\lambda}{x}, \frac{k}{y}\right\}\right\} \\ &\quad \times \#\left\{d \in \mathbb{Z}, d \neq 0 : \max\left\{-\frac{k}{x}, \frac{-k-\lambda}{y}\right\} \leq d \leq \frac{k-\lambda}{y}\right\}, \quad (8) \end{aligned}$$

where we have used  $\lambda \geq 0$  and  $x \leq y$  to simplify the inequalities slightly. The first factor on the right-hand side of (8) is

$$\min\left\{\frac{k-\lambda}{x}, \frac{k}{y}\right\} - \left(-\frac{k}{y}\right) + O(1) = k \min\left\{\frac{1-\lambda/k}{x} + \frac{1}{y}, \frac{2}{y}\right\} + O(1)$$

if this expression is positive, and 0 otherwise; it is thus precisely  $kC(\lambda/k; x, y) + O(1)$ . Similarly, the second factor on the right-hand side of (8) is

$$\frac{k-\lambda}{y} - \max\left\{-\frac{k}{x}, \frac{-k-\lambda}{y}\right\} + O(1) = k \min\left\{\frac{1-\lambda/k}{y} + \frac{1}{x}, \frac{2}{y}\right\} + O(1)$$

(note that this expression is always positive under the hypotheses of the lemma), which is simply  $kD(\lambda/k; x, y) + O(1)$ . Multiplying these two factors yields

$$N_{k,\lambda}(x, y) = k^2 C\left(\frac{\lambda}{k}; x, y\right) D\left(\frac{\lambda}{k}; x, y\right) + k \cdot O\left(C\left(\frac{\lambda}{k}; x, y\right) + D\left(\frac{\lambda}{k}; x, y\right)\right) + O(1).$$

The lemma follows upon noting that both  $C(\lambda/k; x, y)$  and  $D(\lambda/k; x, y)$  are  $\ll 1/y$  by definition, so the second summand becomes simply  $O(k/y)$ , and the  $O(1)$  term may be subsumed into  $O(k/y)$  since  $y \leq k$ . □

We have already used the trivial estimate

$$\sum_{L \leq \alpha < U} 1 = (U - L) + O(1),$$

provided  $0 < L < U$ . We will also use, without further comment, the estimates

$$\sum_{L \leq \alpha < U} \frac{1}{\alpha} = \log \frac{U}{L} + O\left(\frac{1}{L}\right),$$

with its particular case

$$\sum_{1 \leq \alpha < U} \frac{1}{\alpha} = \log U + O(1),$$

and

$$\sum_{L \leq \alpha < U} \frac{1}{\alpha^2} = \frac{1}{L} - \frac{1}{U} + O\left(\frac{1}{L^2}\right).$$

These estimates (also valid for  $0 < L < U$ ) follow readily from comparison to the integrals  $\int_L^U dx/x$  and  $\int_L^U dx/x^2$ .

Most of the technical work in proving [Theorem 1](#) lies in establishing an estimate on a sum of the form  $\sum_{1 \leq \alpha < \beta} C(\delta; \alpha, \beta) D(\delta; \alpha, \beta)$  for a fixed  $\beta$ . The following proposition provides an asymptotic formula for this sum; we defer the proof until the next section. Assuming this proposition, though, we can complete the proof of [Theorem 1](#), as well as [Corollary 2](#).

**Proposition 10.** *Let  $\beta \geq 1$  and  $0 \leq \delta \leq 2$  be real numbers, and let  $C$  and  $D$  be the functions defined in (7). Then*

$$\sum_{1 \leq \alpha < \beta} C(\delta; \alpha, \beta) D(\delta; \alpha, \beta) = \frac{V(\delta)}{\beta} + O\left(\frac{1 + \log \beta}{\beta^2}\right),$$

where  $V(\delta)$  was defined in (1).

*Proof of Theorem 1 assuming Proposition 10.* The functions  $C$  and  $D$  defined in (7) are homogeneous of degree  $-1$  in the variables  $x$  and  $y$ , so that Lemma 9 implies

$$N_{k,\lambda}(d\alpha, d\beta) = \frac{k^2}{d^2} C\left(\frac{\lambda}{k}; \alpha, \beta\right) D\left(\frac{\lambda}{k}; \alpha, \beta\right) + O\left(\frac{k}{d\beta}\right).$$

Inserting this formula into the conclusion of Lemma 8 yields

$$|\mathcal{M}_2^\lambda(k)| = 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \sum_{1 \leq \alpha < \beta \leq k/d} C\left(\frac{\lambda}{k}; \alpha, \beta\right) D\left(\frac{\lambda}{k}; \alpha, \beta\right) + O\left(\sum_{d \leq k} \sum_{1 \leq \alpha < \beta \leq k/d} \frac{k}{d\beta}\right) + O(k^2).$$

We bound the first error term by summing over  $1 \leq \alpha < \beta$  to obtain

$$\sum_{d \leq k} \sum_{1 \leq \alpha < \beta \leq k/d} \frac{k}{d\beta} \leq \sum_{d \leq k} \sum_{1 < \beta \leq k/d} \frac{k}{d} < \sum_{d \leq k} \frac{k^2}{d^2} \ll k^2,$$

so that we have the estimate

$$|\mathcal{M}_2^\lambda(k)| = 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \sum_{1 \leq \alpha < \beta \leq k/d} C\left(\frac{\lambda}{k}; \alpha, \beta\right) D\left(\frac{\lambda}{k}; \alpha, \beta\right) + O(k^2).$$

We now apply Proposition 10 to obtain

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \sum_{1 \leq \beta \leq k/d} \left( \frac{V(\lambda/k)}{\beta} + O\left(\frac{1 + \log \beta}{\beta^2}\right) \right) + O(k^2) \\ &= 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \left( V\left(\frac{\lambda}{k}\right) \left( \log \frac{k}{d} + O(1) \right) + O(1) \right) + O(k^2) \\ &= 4k^2 \left( V\left(\frac{\lambda}{k}\right) \log k \sum_{d \leq k} \frac{\mu(d)}{d^2} + O\left(\sum_{d \leq k} \frac{\log d}{d^2}\right) \right) + O(k^2) \\ &= 4k^2 \left( V\left(\frac{\lambda}{k}\right) \log k \left( \frac{6}{\pi^2} + O\left(\frac{1}{k}\right) \right) + O(1) \right) + O(k^2) \\ &= \frac{24}{\pi^2} V\left(\frac{\lambda}{k}\right) k^2 \log k + O(k^2), \end{aligned}$$

where we have used (5) and the convergence of  $\sum 1/n^2$  and  $\sum (\log n)/n^2$  (so the partial sums are uniformly bounded).  $\square$

*Proof of Corollary 2 from Theorem 1.* Note that for any  $M \in \mathcal{M}_2(k)$ , if one eigenvalue is an integer then both of them are (since the trace of  $M$  is an integer). Thus if we add up the cardinalities of all of the  $\mathcal{M}_2^\lambda(k)$ , we get twice the cardinality of  $\mathcal{M}_2^{\mathbb{Z}}(k)$ , except that matrices with repeated eigenvalues only get counted once. However, the number of such matrices is  $\ll_\varepsilon k^{2+\varepsilon}$  by Lemma 7. Therefore

$$\begin{aligned} 2|\mathcal{M}_2^{\mathbb{Z}}(k)| &= \sum_{\lambda \in \mathbb{Z}} |\mathcal{M}_2^\lambda(k)| + O_\varepsilon(k^{2+\varepsilon}) \\ &= \sum_{-2k \leq \lambda \leq 2k} \left( \frac{24}{\pi^2} V\left(\frac{\lambda}{k}\right) k^2 \log k + O(k^2) \right) + O_\varepsilon(k^{2+\varepsilon}) \\ &= \frac{24}{\pi^2} k^3 \log k \sum_{-2k \leq \lambda \leq 2k} \frac{V\left(\frac{\lambda}{k}\right)}{k} + O(k^3). \end{aligned}$$

The sum is a Riemann sum of a function of bounded variation, so this becomes

$$2|\mathcal{M}_2^{\mathbb{Z}}(k)| = \frac{24}{\pi^2} k^3 \log k \left( \int_{-2}^2 V(\delta) d\delta + O\left(\frac{1}{k}\right) \right) + O(k^3).$$

The corollary then follows from the straightforward computation of the integral

$$\int_{-2}^2 V(\delta) d\delta = \frac{4}{9}(7\sqrt{2} + 4 + 3 \log(\sqrt{2} + 1)),$$

noting that  $\log(\sqrt{2} - 1) = -\log(\sqrt{2} + 1)$ . □

### 4. Proving the key proposition

It remains to prove Proposition 10. Recalling that the functions  $C$  and  $D$  defined in (7) are formed by combinations of minima and maxima, we need to separate our arguments into several cases depending on the range of  $\delta$ . The following lemma addresses a sum that occurs in two of these cases ( $0 < \delta < 1$  and  $1 < \delta < \sqrt{2}$ ). Note that because of the presence of terms like  $\log(\delta - 1)$  in the formula for  $V(\delta)$ , we need to exercise some caution near  $\delta = 1$ .

**Lemma 11.** *Let  $\beta \geq 1$  and  $0 \leq \delta \leq \sqrt{2}$  be real numbers, with  $\delta \neq 1$ . Then*

$$\sum'_\alpha \left( \frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} = \frac{2}{\beta} \left( \frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) + O\left( \frac{1+\log \beta}{\beta^2} \right),$$

where  $\sum'_\alpha$  is the sum subject to the condition  $\max\{1, |1-\delta|\beta\} \leq \alpha < \beta/(1+\delta)$ .

*Proof.* Suppose first that  $|1-\delta|\beta \geq 1$ . Then the sum in question is

$$\begin{aligned}
 & \frac{2(1-\delta)}{\beta} \sum_{\substack{\alpha \geq |1-\delta|\beta \\ \alpha < \beta/(1+\delta)}} \frac{1}{\alpha} + \frac{2}{\beta^2} \sum_{\substack{\alpha \geq |1-\delta|\beta \\ \alpha < \beta/(1+\delta)}} 1 \\
 &= \frac{2(1-\delta)}{\beta} \left( \log \frac{\beta}{(1+\delta)|1-\delta|\beta} + O\left(\frac{1}{|1-\delta|\beta}\right) \right) + \frac{2}{\beta^2} \left( \frac{\beta}{1+\delta} - |1-\delta|\beta + O(1) \right) \\
 &= \frac{2}{\beta} \left( \frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) + O\left(\frac{1}{\beta^2}\right),
 \end{aligned}$$

which establishes the lemma in this case. On the other hand, if  $|1-\delta|\beta < 1$  then the sum in question is

$$\begin{aligned}
 & \frac{2(1-\delta)}{\beta} \sum_{1 \leq \alpha < \beta/(1+\delta)} \frac{1}{\alpha} + \frac{2}{\beta^2} \sum_{1 \leq \alpha < \beta/(1+\delta)} 1 \\
 &= \frac{2(1-\delta)}{\beta} \left( \log \frac{\beta}{1+\delta} + O(1) \right) + \frac{2}{\beta^2} \left( \frac{\beta}{1+\delta} + O(1) \right) \\
 &= \frac{2}{\beta} \left( \frac{1}{1+\delta} - (1-\delta) \log(1+\delta) \right) + O\left(\frac{1}{\beta^2} + \frac{|1-\delta| \log \beta}{\beta}\right).
 \end{aligned}$$

We subtract  $2(|1-\delta| + (1-\delta) \log |1-\delta|)/\beta$  from the main term and compensate in the error term to obtain

$$\begin{aligned}
 & \frac{2(1-\delta)}{\beta} \sum_{1 \leq \alpha < \beta/(1+\delta)} \frac{1}{\alpha} + \frac{2}{\beta^2} \sum_{1 \leq \alpha < \beta/(1+\delta)} 1 \\
 &= \frac{2}{\beta} \left( \frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) \\
 &\quad + O\left(\frac{1}{\beta^2} + \frac{|1-\delta| \log \beta}{\beta} + \frac{|1-\delta| + |1-\delta| \log |1-\delta|^{-1}}{\beta}\right) \\
 &= \frac{2}{\beta} \left( \frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) + O\left(\frac{1 + \log \beta}{\beta^2} + \frac{|1-\delta| \log |1-\delta|^{-1}}{\beta}\right),
 \end{aligned}$$

since we are working with the assumption that  $|1-\delta| < 1/\beta$ . Because the function  $t \log t^{-1}$  is increasing on the interval  $(0, 1/e)$  and bounded on the interval  $(0, 1]$ , we have  $|1-\delta| \log |1-\delta|^{-1} < (1/\beta) \log \beta$  if  $\beta > e$  and  $|1-\delta| \log |1-\delta|^{-1} \ll 1 \ll \frac{1}{\beta}$  if  $1 \leq \beta \leq e$ . In either case, the last error term can be simplified to

$$O\left(\frac{1 + \log \beta}{\beta^2}\right),$$

which establishes the lemma in second case.  $\square$

*Proof of Proposition 10.* We consider separately the four cases corresponding to the different parts of the definition (1) of  $V(\delta)$ . To lighten the expressions, we use the  $\sum'_\alpha$  notation from the statement of Lemma 11 and omit the dependence on  $\delta$ ,  $\alpha$ , and  $\beta$  from the notation for  $C$  and  $D$ .

**Case 1:  $0 \leq \delta < 1$ .** In this case we have  $0 < 1 - \delta < \frac{1}{1+\delta} \leq 1$  and

$$C = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq (1 - \delta)\beta, \\ \frac{1-\delta}{\alpha} + \frac{1}{\beta}, & \text{if } (1 - \delta)\beta \leq \alpha, \end{cases} \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{1+\delta}, \\ \frac{1-\delta}{\beta} + \frac{1}{\alpha}, & \text{if } \frac{\beta}{1+\delta} \leq \alpha. \end{cases}$$

Therefore,

$$\begin{aligned} & \sum_{1 \leq \alpha < \beta} CD \\ &= \sum_{1 \leq \alpha < (1-\delta)\beta} \frac{2}{\beta} \cdot \frac{2}{\beta} + \sum'_\alpha \left( \frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} + \sum_{\beta/(1+\delta) \leq \alpha < \beta} \left( \frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \left( \frac{1-\delta}{\beta} + \frac{1}{\alpha} \right). \end{aligned} \tag{9}$$

(The first sum might be empty, but this does not invalidate the argument that follows.) The first sum is simply

$$\frac{4}{\beta^2} \sum_{1 \leq \alpha < (1-\delta)\beta} 1 = \frac{4}{\beta^2} ((1 - \delta)\beta + O(1)) = \frac{4(1-\delta)}{\beta} + O\left(\frac{1}{\beta^2}\right).$$

By [Lemma 11](#), the second sum is

$$\sum'_\alpha \left( \frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} = \frac{2}{\beta} \left( \frac{1}{1+\delta} + \delta - 1 - (1-\delta) \log(1-\delta^2) \right) + O\left(\frac{1+\log \beta}{\beta^2}\right),$$

while the third sum is

$$\begin{aligned} & \sum_{\beta/(1+\delta) \leq \alpha < \beta} \left( \frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \left( \frac{1-\delta}{\beta} + \frac{1}{\alpha} \right) \\ &= (1 - \delta) \sum_{\beta/(1+\delta) \leq \alpha < \beta} \frac{1}{\alpha^2} + \frac{\delta^2 - 2\delta + 2}{\beta} \sum_{\beta/(1+\delta) \leq \alpha < \beta} \frac{1}{\alpha} + \frac{1-\delta}{\beta^2} \sum_{\beta/(1+\delta) \leq \alpha < \beta} 1 \\ &= (1 - \delta) \left( \frac{1+\delta}{\beta} - \frac{1}{\beta} + O\left(\frac{1}{\beta^2}\right) \right) \\ &\quad + \frac{\delta^2 - 2\delta + 2}{\beta} \left( \log \frac{\beta}{(1+\delta)^{-1}\beta} + O\left(\frac{1}{\beta}\right) \right) + \frac{1-\delta}{\beta^2} \left( \beta - \frac{\beta}{1+\delta} + O(1) \right) \\ &= \frac{1}{\beta} \left( 2 - \delta^2 - \frac{2}{1+\delta} + (\delta^2 - 2\delta + 2) \log(1 + \delta) \right) + O\left(\frac{1}{\beta^2}\right). \end{aligned} \tag{10}$$

This case of the proposition then follows from (9) by noting that

$$\begin{aligned} & 4 - 4\delta + \frac{2}{1+\delta} + 2\delta - 2 - 2(1-\delta) \log(1-\delta^2) + 2 - \delta^2 - \frac{2}{1+\delta} + (\delta^2 - 2\delta + 2) \log(1+\delta) \\ &= 4 - 2\delta - \delta^2 + \delta^2 \log(1 + \delta) - 2(1 - \delta) \log(1 - \delta). \end{aligned}$$

**Case 2:  $\delta = 1$ .** In this case we have

$$C = \frac{1}{\beta}, \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{2}, \\ \frac{1}{\alpha} & \text{if } \frac{\beta}{2} \leq \alpha. \end{cases}$$

Therefore,

$$\begin{aligned} \sum_{1 \leq \alpha < \beta} CD &= \sum_{1 \leq \alpha < \beta/2} \frac{1}{\beta} \cdot \frac{2}{\beta} + \sum_{\beta/2 \leq \alpha < \beta} \frac{1}{\beta} \cdot \frac{1}{\alpha} \\ &= \frac{2}{\beta^2} \left( \frac{\beta}{2} + O(1) \right) + \frac{1}{\beta} \left( \log \frac{\beta}{\beta/2} + O\left(\frac{1}{\beta}\right) \right) = \frac{1 + \log 2}{\beta} + O\left(\frac{1}{\beta^2}\right), \end{aligned}$$

as desired.

**Case 3:  $1 < \delta \leq \sqrt{2}$ .** In this case we have

$$C = \begin{cases} 0 & \text{if } \alpha \leq (\delta - 1)\beta, \\ \frac{1 - \delta}{\alpha} + \frac{1}{\beta} & \text{if } (\delta - 1)\beta \leq \alpha, \end{cases} \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{\delta + 1}, \\ \frac{1 - \delta}{\beta} + \frac{1}{\alpha} & \text{if } \frac{\beta}{\delta + 1} \leq \alpha. \end{cases}$$

Therefore,

$$\sum_{1 \leq \alpha < \beta} CD = \sum'_\alpha \left( \frac{1 - \delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} + \sum_{\beta/(\delta + 1) \leq \alpha < \beta} \left( \frac{1 - \delta}{\alpha} + \frac{1}{\beta} \right) \left( \frac{1 - \delta}{\beta} + \frac{1}{\alpha} \right). \quad (11)$$

(We note that  $(\delta - 1)\beta \leq \beta/(\delta + 1)$  for  $\delta$  between 1 and  $\sqrt{2}$ . For very small  $\beta$  we might have  $1 > \beta/(\delta + 1)$ , in which case the first sum is empty, but that does not invalidate the argument that follows.) By [Lemma 11](#), the first sum is

$$\sum'_\alpha \left( \frac{1 - \delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} = \frac{2}{\beta} \left( \frac{1}{1 + \delta} + 1 - \delta - (1 - \delta) \log(\delta^2 - 1) \right) + O\left(\frac{1 + \log \beta}{\beta^2}\right),$$

while the second sum has already been evaluated in (10) above. This case of the proposition then follows from (11) by noting that

$$\begin{aligned} \frac{2}{1 + \delta} + 2 - 2\delta - 2(1 - \delta) \log(\delta^2 - 1) + 2 - \delta^2 - \frac{2}{\delta + 1} + (\delta^2 - 2\delta + 2) \log(\delta + 1) \\ = 4 - 2\delta - \delta^2 + \delta^2 \log(\delta + 1) + 2(\delta - 1) \log(\delta - 1). \end{aligned}$$

**Case 4:  $\sqrt{2} < \delta \leq 2$ .** Just as in Case 3, we have

$$C = \begin{cases} 0 & \text{if } \alpha \leq (\delta - 1)\beta, \\ \frac{1 - \delta}{\alpha} + \frac{1}{\beta} & \text{if } (\delta - 1)\beta \leq \alpha, \end{cases} \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{\delta + 1}, \\ \frac{1 - \delta}{\beta} + \frac{1}{\alpha} & \text{if } \frac{\beta}{\delta + 1} \leq \alpha. \end{cases}$$



However, the inequality  $(\delta - 1)\beta \leq \alpha$  automatically implies that  $\frac{\beta}{\delta+1} \leq \alpha$  when  $\delta \geq \sqrt{2}$ . Therefore,

$$\sum_{1 \leq \alpha < \beta} CD = \sum_{(\delta-1)\beta \leq \alpha < \beta} \left( \frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \left( \frac{1-\delta}{\beta} + \frac{1}{\alpha} \right).$$

(In this case we will not need to use  $\max\{1, (\delta - 1)\beta\} \leq \alpha$ , which is the more precise lower bound.) This yields

$$\begin{aligned} & \sum_{1 \leq \alpha < \beta} CD \\ &= (1 - \delta) \sum_{(\delta-1)\beta \leq \alpha < \beta} \frac{1}{\alpha^2} + \frac{\delta^2 - 2\delta + 2}{\beta} \sum_{(\delta-1)\beta \leq \alpha < \beta} \frac{1}{\alpha} + \frac{1-\delta}{\beta^2} \sum_{(\delta-1)\beta \leq \alpha < \beta} 1 \\ &= (1 - \delta) \left( \frac{1}{(\delta-1)\beta} - \frac{1}{\beta} + O\left(\frac{1}{(\delta-1)^2\beta^2}\right) \right) \\ &\quad + \frac{\delta^2 - 2\delta + 2}{\beta} \left( \log \frac{\beta}{(\delta-1)\beta} + O\left(\frac{1}{(\delta-1)\beta}\right) \right) + \frac{1-\delta}{\beta^2} (\beta - (\delta-1)\beta + O(1)) \\ &= \frac{1}{\beta} \left( \delta^2 - 2\delta - (\delta^2 - 2\delta + 2) \log(\delta - 1) \right) + O\left(\frac{1}{\beta^2}\right), \end{aligned}$$

the error terms having been simplified since  $\delta - 1$  is bounded away from 0. □

### 5. Distribution of real eigenvalues

In proving [Theorem 3](#), it will be convenient to define the odd function

$$G(z) = \int_0^z -\log |t| dt = z(1 - \log |z|), \tag{12}$$

whose relevance is demonstrated by the following lemma.

**Lemma 12.** *If  $B$  and  $C$  are independent random variables uniformly distributed on  $[-1, 1]$ , the product  $BC$  is a random variable whose distribution function is*

$$F_{BC}(z) = \Pr(BC < z) = \frac{1}{2}(1 + G(z))$$

for  $z \in [-1, 1]$ .

Of course, for  $z < -1$  we have  $F_{BC}(z) = 0$ , and likewise  $F_{BC}(z) = 1$  for  $z > 1$ .

*Proof.* Note that  $|B|$  and  $|C|$  are uniformly distributed on  $[0, 1]$ . For  $0 \leq z \leq 1$ , we easily check that

$$\Pr(|BC| < z) = \int_0^1 \min\left\{1, \frac{z}{s}\right\} ds = G(z).$$

Thus  $|BC|$  is distributed on  $[0, 1]$  with density  $f_{|BC|}(z) = -\log z$ , and by symmetry  $BC$  has density  $f_{BC}(z) = -\frac{1}{2} \log |z|$  on  $[-1, 1]$ . The lemma follows upon computing

$$F_{BC}(z) = \int_{-1}^z f_{BC}(s) ds. \quad \square$$

It will also be helpful to define the following functions, which are symmetric in  $x$  and  $y$ :

$$v_1(x, y) = \frac{1}{2} + \frac{1}{2}G(xy) + G\left(\frac{(x-y)^2}{4}\right), \quad (13)$$

$$v_2(x, y) = \frac{1}{2} - \frac{1}{2}G(xy), \quad (14)$$

$$v(x, y) = \begin{cases} v_1(x, y), & \text{if } xy < 1 \text{ and } x + y < 0, \\ v_2(x, y), & \text{if } xy < 1 \text{ and } x + y > 0, \\ 1 + G\left(\frac{(x-y)^2}{4}\right), & \text{if } xy > 1 \text{ and } x + y < 0, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

To prove [Theorem 3](#), we first consider the distribution function

$$F_W(\delta) = \int_{t < \delta} W(t) dt$$

associated to the density  $W(\delta)$ . For a random matrix  $M$  in  $\mathcal{M}_2([-1, 1])$  and a real number  $\delta$ , we will derive an expression for the expected number of real eigenvalues of  $M$  falling below  $\delta$ , then differentiate it to obtain  $W(\delta)$ .

It is clear that  $W(-\delta) = W(\delta)$  since the set  $\mathcal{M}_2([-1, 1])$  is closed under negation, so it suffices to compute  $W(\delta)$  for  $\delta \in [0, 2]$ . It turns out that our calculations for  $F_W$  will be somewhat simplified by considering  $F_W(-\delta)$  rather than  $F_W(\delta)$ .

**Proposition 13.** *We have*

$$F_W(-\delta) = \frac{1}{4} \int_{-1+\delta}^{1+\delta} \int_{-1+\delta}^{1+\delta} v(x, y) dx dy$$

for all  $0 \leq \delta \leq 2$ , where  $v$  is defined in (15).

*Proof.* Denote the entries of  $M$  by  $A, B, C, D$ , which by assumption are independent random variables uniformly distributed in  $[-1, 1]$ . Let  $\delta$  be fixed in the range  $[0, 2]$ , and consider the shifted matrix  $M' = M + \delta I$ , which we write as

$$M' = \begin{pmatrix} X & B \\ C & Y \end{pmatrix},$$

where  $X, Y$  range independently and uniformly in  $[-1 + \delta, 1 + \delta]$  and  $B, C$  are as before. Clearly the eigenvalues of  $M$  less than  $-\delta$  correspond to the negative (real) eigenvalues of  $M'$ . By [Lemma 7](#), we are free to exclude the null set where

$M'$  is singular or has repeated eigenvalues. Outside of this null set,  $M'$  has exactly one negative eigenvalue if and only if

$$\det M' = XY - BC < 0,$$

by [Lemma 6 \(c\)](#). Likewise by [Lemma 6 \(d\)](#),  $M'$  has exactly two negative eigenvalues if and only if

$$XY - BC > 0, \quad X + Y < 0 \quad \text{and} \quad \text{disc } M' = (X - Y)^2 + 4BC > 0.$$

We thus have:

$$F_W(-\delta) = \Pr(BC > XY) + 2\Pr\left(X + Y < 0 \text{ and } -\frac{(X - Y)^2}{4} < BC < XY\right).$$

We may express this probability as the average value

$$F_W(-\delta) = \frac{1}{4} \int_{-1+\delta}^{1+\delta} \int_{-1+\delta}^{1+\delta} \rho(x, y) \, dx \, dy,$$

where for fixed  $x$  and  $y$ ,

$$\begin{aligned} \rho(x, y) &= \Pr(BC > xy) + 2\Pr\left(x + y < 0 \text{ and } -\frac{(x - y)^2}{4} < BC < xy\right) \\ &= \Pr(BC > xy) + 2\Pr\left(-\frac{(x - y)^2}{4} < BC < xy\right) \mathbf{1}\{x + y < 0\} \end{aligned} \tag{16}$$

(here  $\mathbf{1}\{\cdot\}$  denotes the indicator function of the indicated relation). To complete the proof it suffices to show that  $\rho$  equals the function  $\nu$  defined in [\(15\)](#).

The probabilities appearing in [\(16\)](#) are effectively given by [Lemma 12](#). However, there is some case-checking involved in applying this lemma, since the value of, say,

$$\Pr(BC > xy) = 1 - F_{BC}(xy)$$

depends on whether  $xy < -1$ ,  $-1 \leq xy \leq 1$ , or  $xy > 1$ . We make some observations to reduce the number of cases we need to examine.

Note that  $(x - y)^2/4$  is bounded between 0 and 1 for any  $x, y \in [-1 + \delta, 1 + \delta]$ , so that  $-\frac{(x - y)^2}{4}$  always lies in the interval  $[-1, 1]$  prescribed by [Lemma 12](#). From the identity

$$(x + y)^2 - (x - y)^2 = 4xy$$

we see also that  $xy \geq -(x - y)^2/4$ . Thus  $xy$  is never lower than  $-1$ , and we need only consider whether  $xy > 1$  (in which case  $F_{BC}(xy) = 1$ ). We therefore have

$$\begin{aligned} \Pr(BC > xy) &= 1 - F_{BC}(xy) = \frac{1}{2}(1 - G(xy)) \mathbf{1}\{xy < 1\}, \\ 2\Pr\left(-\frac{(x - y)^2}{4} < BC < xy\right) &= 2F_{BC}(xy) - 2F_{BC}\left(-\frac{(x - y)^2}{4}\right) \\ &= \mathbf{1}\{xy > 1\} + G(xy) \mathbf{1}\{xy < 1\} + G\left(\frac{(x - y)^2}{4}\right). \end{aligned}$$

Inserting these two evaluations into the formula (16), we obtain

$$\rho(x, y) = \frac{1}{2}(1 - G(xy))\mathbf{1}\{xy < 1\} + \left(\mathbf{1}\{xy > 1\} + G(xy)\mathbf{1}\{xy < 1\} + G\left(\frac{(x-y)^2}{4}\right)\right)\mathbf{1}\{x+y < 0\}.$$

It can be verified that this last expression is indeed equal to the right-hand side of the definition (15) of  $v$ . □

Since

$$W(\delta) = W(-\delta) = -\frac{d}{d\delta}F_W(-\delta),$$

to finish the proof of [Theorem 3](#) it therefore suffices to prove that  $-\frac{d}{d\delta}F_W(-\delta)$  equals the formula given in [Theorem 3](#).

### 6. The derivative of the distribution

[Proposition 13](#) expresses  $F_W(\delta)$  as an integral, of a function  $v$  that is independent of  $\delta$ , over the square

$$S_\delta = [-1 + \delta, 1 + \delta]^2.$$

Since the region  $S_\delta$  varies continuously with  $\delta$ , we can compute the derivative

$$-\frac{d}{d\delta}F_W(-\delta)$$

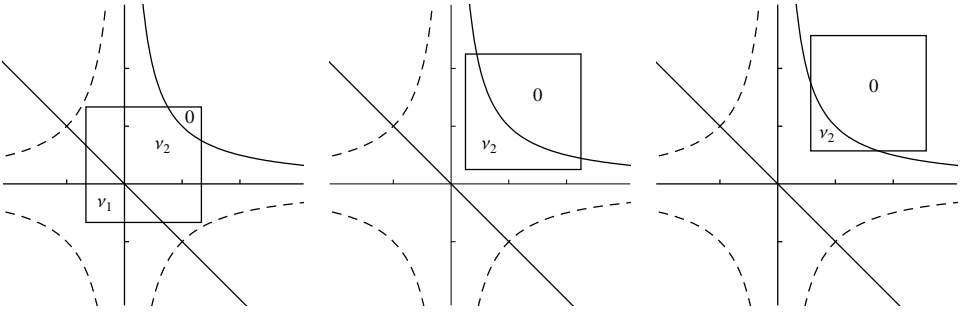
by an appropriate line integral around the boundary of  $S_\delta$ . Indeed, by the fundamental theorem of calculus, we have

$$\begin{aligned} -\frac{d}{d\delta}F_W(-\delta) &= -\frac{1}{4}\frac{d}{d\delta}\left(\int_{-1+\delta}^{1+\delta}\int_{-1+\delta}^{1+\delta}v(x, y)dx dy\right) \\ &= -\frac{1}{4}\left(\int_{-1+\delta}^{1+\delta}v(1+\delta, y)dy - \int_{-1+\delta}^{1+\delta}v(-1+\delta, y)dy \right. \\ &\quad \left. + \int_{-1+\delta}^{1+\delta}v(x, 1+\delta)dx - \int_{-1+\delta}^{1+\delta}v(x, -1+\delta)dx\right) \\ &= \frac{1}{2}\int_{-1+\delta}^{1+\delta}v(x, -1+\delta)dx - \frac{1}{2}\int_{-1+\delta}^{1+\delta}v(x, 1+\delta)dx, \end{aligned} \tag{17}$$

where we have used the symmetry  $v(x, y) = v(y, x)$  to reduce the integral to just the top and bottom edges of  $S_\delta$  (where  $y = 1 + \delta$  and  $y = -1 + \delta$ , respectively).

The evaluation of (17) divides into three cases depending on the behavior of the indicator functions  $\mathbf{1}\{x+y < 0\}$  and  $\mathbf{1}\{xy < 1\}$  on the boundary of  $S_\delta$  ([Figure 2](#)).

**Case 1:  $0 \leq \delta \leq 1$ .** For this range of  $\delta$ , the line  $x + y = 0$  intersects the bottom edge of  $S_\delta$  at  $x = 1 - \delta$ , while the hyperbola  $xy = 1$  intersects the top edge at



**Figure 2.** The three cases:  $0 \leq \delta \leq 1$ ,  $1 \leq \delta \leq \sqrt{2}$  and  $\sqrt{2} \leq \delta \leq 2$ .

$x = (1 + \delta)^{-1}$ . Thus by the definition of  $\nu$ , (17) becomes

$$-\frac{d}{d\delta} F_W(-\delta) = \frac{1}{2} \left( \int_{-1+\delta}^{1-\delta} \nu_1(x, -1 + \delta) dx + \int_{1-\delta}^{1+\delta} \nu_2(x, -1 + \delta) dx - \int_{-1+\delta}^{(1+\delta)^{-1}} \nu_2(x, 1 + \delta) dx \right).$$

The following elementary antiderivatives, which are readily obtained by substitution and integration by parts, follow for any fixed nonzero real number  $y$  from the definitions (12), (13), and (14) of  $G$ ,  $\nu_1$ , and  $\nu_2$ :

$$\begin{aligned} \int \nu_1(x, y) dx &= \frac{1}{2}x + \frac{1}{8}x^2y(3 - 2 \log |xy|) + \frac{1}{36}(x-y)^3 \left( 5 - 6 \log \left| \frac{x-y}{2} \right| \right), \\ \int \nu_2(x, y) dx &= \frac{1}{2}x - \frac{1}{8}x^2y(3 - 2 \log |xy|). \end{aligned} \tag{18}$$

Therefore in this case,

$$\begin{aligned} -\frac{d}{d\delta} F_W(-\delta) &= \frac{1}{2} \left( \left( \frac{1}{2}x + \frac{1}{8}x^2(-1 + \delta)(3 - 2 \log |x(-1 + \delta)|) + \frac{1}{36}(x + 1 - \delta)^3 \left( 5 - 6 \log \left| \frac{x+1-\delta}{2} \right| \right) \right) \Big|_{x=-1+\delta}^{1-\delta} \right. \\ &\quad + \left( \frac{1}{2}x - \frac{1}{8}x^2(-1 + \delta)(3 - 2 \log |x(-1 + \delta)|) \right) \Big|_{x=1-\delta}^{1+\delta} \\ &\quad \left. - \left( \frac{1}{2}x - \frac{1}{8}x^2(1 + \delta)(3 - 2 \log |x(1 + \delta)|) \right) \Big|_{x=-1+\delta}^{(1+\delta)^{-1}} \right) \\ &= \frac{80 + 20\delta + 90\delta^2 + 52\delta^3 - 107\delta^4}{144(1 + \delta)} \\ &\quad - \frac{(5 - 7\delta + 8\delta^2)(1 - \delta)}{12} \log(1 - \delta) - \frac{\delta(1 - \delta^2)}{4} \log(1 + \delta) \end{aligned}$$

(after some algebraic simplification), which verifies the first case of [Theorem 3](#). (Note that the integrands really are continuous, despite terms that look like  $\log 0$ , because the function  $G$  is continuous at 0; hence evaluating the integrals by antiderivatives is valid.)

**Case 2:**  $1 \leq \delta \leq \sqrt{2}$ . Now, the line  $x + y = 0$  does not intersect  $S_\delta$ , while the hyperbola  $xy = 1$  intersects the top edge at  $x = (1 + \delta)^{-1}$ . Thus by the definition of  $\nu$  and the antiderivative (18) of  $\nu_2$ , (17) becomes

$$\begin{aligned} -\frac{d}{d\delta} F_W(-\delta) &= \frac{1}{2} \left( \int_{-1+\delta}^{1+\delta} \nu_2(x, -1+\delta) dx - \int_{-1+\delta}^{(1+\delta)^{-1}} \nu_2(x, 1+\delta) dx \right) \\ &= \frac{1}{2} \left( \left( \frac{1}{2}x - \frac{1}{8}x^2(-1+\delta)(3 - 2 \log |x(-1+\delta)|) \right) \Big|_{x=-1+\delta}^{1+\delta} \right. \\ &\quad \left. - \left( \frac{1}{2}x - \frac{1}{8}x^2(1+\delta)(3 - 2 \log |x(1+\delta)|) \right) \Big|_{x=-1+\delta}^{(1+\delta)^{-1}} \right) \\ &= \frac{\delta(20 + 10\delta - 12\delta^2 - 3\delta^3)}{16(1+\delta)} + \frac{(3\delta - 1)(\delta - 1)}{4} \log(\delta - 1) \\ &\quad + \frac{\delta(\delta^2 - 1)}{4} \log(\delta + 1), \end{aligned}$$

which verifies the second case of [Theorem 3](#).

**Case 3:**  $\sqrt{2} < \delta \leq 2$ . As before, the line  $x + y = 0$  does not intersect  $S_\delta$ , while the hyperbola  $xy = 1$  now intersects the bottom edge at  $x = (\delta - 1)^{-1}$ . Thus by the definition of  $\nu$  and the antiderivative (18) of  $\nu_2$ , (17) becomes

$$\begin{aligned} -\frac{d}{d\delta} F_W(-\delta) &= \frac{1}{2} \int_{-1+\delta}^{(-1+\delta)^{-1}} \nu_2(x, -1+\delta) dx \\ &= \frac{1}{2} \left( \frac{1}{2}x - \frac{1}{8}x^2(1+\delta)(3 - 2 \log |x(1+\delta)|) \right) \Big|_{x=-1+\delta}^{(-1+\delta)^{-1}} \\ &= \frac{\delta(\delta - 2)(2 - 6\delta + 3\delta^2)}{16(\delta - 1)} - \frac{(\delta - 1)^3}{4} \log(\delta - 1), \end{aligned}$$

which verifies the third case of [Theorem 3](#).

Since the last case of [Theorem 3](#) is a consequence of [Lemma 4](#), the proof of the theorem is complete.

**Remark.** One could also use the same method to extract the individual distributions of the greater and lesser eigenvalues of  $M$ : for instance, eliminating the factor of 2 from (16) would yield an expression for the distribution of just the lesser eigenvalue of  $M$ .

## References

- [Edelman 1997] A. Edelman, “The probability that a random real Gaussian matrix has  $k$  real eigenvalues, related distributions, and the circular law”, *J. Multivariate Anal.* **60**:2 (1997), 203–232. [MR 98b:15025](#) [Zbl 0886.15024](#)
- [Gershgorin 1931] S. Gershgorin, “Über die Abgrenzung der Eigenwerte einer Matrix”, *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk* **7** (1931), 749–754.
- [Girko 1984] V. L. Girko, “The circular law”, *Teor. Veroyatnost. i Primenen.* **29**:4 (1984), 669–679. [MR 87c:15042](#) [Zbl 0565.60034](#)
- [Hetzel et al. 2007] A. J. Hetzel, J. S. Liew, and K. E. Morrison, “The probability that a matrix of integers is diagonalizable”, *Amer. Math. Monthly* **114**:6 (2007), 491–499. [MR 2008d:15054](#) [Zbl 1140.15018](#)
- [Katznelson 1993] Y. R. Katznelson, “Singular matrices and a uniform bound for congruence groups of  $SL_n(\mathbf{Z})$ ”, *Duke Math. J.* **69**:1 (1993), 121–136. [MR 94g:11083](#) [Zbl 0785.11050](#)
- [Martin and Wong 2008] G. Martin and E. B. Wong, “Almost all integer matrices have no integer eigenvalues”, preprint, 2008. To appear in *Amer. Math. Monthly*. [arXiv 0712.3060](#)
- [Montgomery and Vaughan 2007] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics **97**, Cambridge University Press, Cambridge, 2007. [MR 2378655](#) [Zbl 1142.11001](#)
- [Tao and Vu 2008] T. Tao and V. Vu, “Random matrices: Universality of ESDs and the circular law”, preprint, 2008. [arXiv 0807.4898](#)

Communicated by Andrew Granville

Received 2008-05-28

Revised 2008-08-13

Accepted 2008-10-16

[gerg@math.ubc.ca](mailto:gerg@math.ubc.ca)

*Department of Mathematics, University of British Columbia,  
Room 121, 1984 Mathematics Road,  
Vancouver, BC V6T 1Z2, Canada  
[www.math.ubc.ca/~gerg](http://www.math.ubc.ca/~gerg)*

[erick@math.ubc.ca](mailto:erick@math.ubc.ca)

*Department of Mathematics, University of British Columbia,  
Room 121, 1984 Mathematics Road,  
Vancouver, BC V6T 1Z2, Canada  
[www.math.ubc.ca/~erick](http://www.math.ubc.ca/~erick)*