What is the probability that a random integral quadratic form in n variables has an integral zero?

M. Bhargava, J. E. Cremona, T. A. Fisher, N. G. Jones, and J. P. Keating

28th July 2015

Abstract

We show that the density of quadratic forms in n variables over $\mathbb{Z}_p$ that are isotropic is a rational function of p, where the rational function is independent of p, and we determine this rational function explicitly. When real quadratic forms in n variables are distributed according to the Gaussian Orthogonal Ensemble (GOE) of random matrix theory, we determine explicitly the probability that a random such real quadratic form is isotropic (i.e., indefinite).

As a consequence, for each n, we determine an exact expression for the probability that a random integral quadratic form in n variables is isotropic (i.e., has a nontrivial zero over $\mathbb{Z}$), when these integral quadratic forms are chosen according to the GOE distribution. In particular, we find an exact expression for the probability that a random integral quaternary quadratic form is isotropic; numerically, this probability of isotropy is approximately 98.3%.

1 Introduction

An integral quadratic form Q in n variables is a homogeneous quadratic polynomial

$$Q(x_1, x_2, \ldots, x_n) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j, \tag{1}$$

where all coefficients c_{ij} lie in $\mathbb{Z}$. The quadratic form Q is said to be *isotropic* if it represents 0, i.e., if there exists a nonzero n-tuple $(k_1, \ldots, k_n) \in \mathbb{Z}^n$ such that $Q(k_1, \ldots, k_n) = 0$. We wish to consider the question: what is the probability that a random integral quadratic form in n variables is isotropic?

In this paper, we give a complete answer to this question for all n, when integral quadratic forms in n variables are chosen according to the Gaussian Orthogonal Ensemble (GOE) of random matrix theory [1]. In particular, in the most interesting case $n = 4$, we show that the probability that a random integral quaternary quadratic form is isotropic is given by

$$\left(\frac{1}{2} + \frac{\sqrt{2}}{8} + \frac{1}{\pi}\right) \prod_p \left(1 - \frac{p^3}{4(p+1)^2(p^4+p^3+p^2+p+1)}\right) \approx 98.3\%. \tag{2}$$

More precisely, let D be a piecewise smooth rapidly decaying function on the vector space $\mathbb{R}^{n(n+1)/2}$ of real quadratic forms in n variables (i.e., $D(x)$ and all its partial derivatives are $o(|x|^{-N})$

for all $N > 0$), and assume that $\int_Q D(Q)dQ = 1$; we call such a function D a *nice distribution* on the space of real n-ary quadratic forms. Then we define the probability, with respect to the distribution D, that a random integral n-ary quadratic form Q has a property P by

$$\lim_{X\to\infty} \frac{\sum_{Q \text{ integral, with property } P} D(Q/X)}{\sum_{Q \text{ integral}} D(Q/X)}, \tag{3}$$

if the limit exists. Let ρ_n^D denote the probability with respect to the distribution D that a random integral quadratic form in n variables is isotropic. If $D = \text{GOE}$ is the distribution on the space of $n \times n$ symmetric matrices given by $\frac{1}{\sqrt{2}}(A + A^t)$, where each entry of the matrix A is an identical and independently distributed real Gaussian—i.e., the Gaussian Orthogonal Ensemble—then we use $\rho_n := \rho_n^{\text{GOE}}$ to denote the probability, with respect to the GOE distribution, that a random n-ary quadratic form over $\mathbb{Z}$ is isotropic.

We wish to explicitly determine the probability ρ_n that a random n-ary quadratic form over $\mathbb{Z}$, with respect to the GOE distribution, is isotropic, i.e., has a nontrivial zero over $\mathbb{Z}$. To accomplish this, we first recall the Hasse–Minkowski Theorem, which states that a quadratic form over $\mathbb{Z}$ is isotropic if and only if it is isotropic over $\mathbb{Z}_p$ for all p and over $\mathbb{R}$. For any distribution D as above, let $\rho_n^D(p)$ denote the probability that a random integral quadratic form, with respect to the distribution D, is isotropic over $\mathbb{Z}_p$, and let $\rho_n^D(\infty)$ denote the probability that it is isotropic over $\mathbb{R}$ (i.e., is indefinite). Then it is not hard to show (for the details, see Section 2) that $\rho_n(p) = \rho_n^D(p)$ is independent of D, and is simply given by the probability that a random n-ary quadratic form over $\mathbb{Z}_p$, with respect to the usual additive measure on $\mathbb{Z}_p^{n(n+1)/2}$, is isotropic over $\mathbb{Z}_p$. Moreover, we will also show in Section 2 that the probability $\rho_n^D(\infty)$ that a random *integral* quadratic form is isotropic over $\mathbb{R}$ is equal to the probability that a random *real* quadratic form (with respect to the same distribution D) is indefinite.

For any distribution D as above, it can be proved using the work of Poonen and Voloch [11], together with the Hasse–Minkowski Theorem, that:

Theorem 1.1 *The probability ρ_n^D that a random (with respect to the distribution D) integral quadratic form in n variables is isotropic is given by the product of the local probabilities:*

$$\rho_n^D = \rho_n^D(\infty) \prod_p \rho_n(p). \tag{4}$$

See Section 2 for details. Hence, to determine ρ_n^D, it suffices to determine $\rho_n^D(\infty)$ and $\rho_n(p)$ for all p.

We treat first the probability $\rho_n(p)$ that a random n-ary quadratic form over $\mathbb{Z}_p$ is isotropic. Our main result here is that, for each n, the quantity $\rho_n(p)$ is given by a fixed rational function in p that is independent of p (this even includes the case $p = 2$), and we determine these rational functions explicitly. Specifically, we prove the following theorem:

Theorem 1.2 *Let $\rho_n(p)$ denote the probability that a quadratic form in n variables over $\mathbb{Z}_p$ is isotropic. Then*

$$\rho_1(p) = 0, \quad \rho_2(p) = \frac{1}{2}, \quad \rho_3(p) = 1 - \frac{p}{2(p+1)^2}, \quad \rho_4(p) = 1 - \frac{p^3}{4(p+1)^2(p^4+p^3+p^2+p+1)},$$

and $\rho_n(p) = 1$ for all $n \geq 5$.

Our method of proof for Theorem 1.2 is uniform in n, and relies on establishing certain recursive formulae for densities of local solubility for certain subsets of n-ary quadratic forms defined by their behavior modulo powers of p. In particular, we obtain a new recursive proof of the well-known fact that every n-ary quadratic form over $\mathbb{Q}_p$ is isotropic when $n \geq 5$. See Section 3 for details.

We turn next to the probability $\rho_n(\infty) = \rho_n^{\mathrm{GOE}}(\infty)$ that a real n-ary quadratic form is isotropic. Closed form expressions for $\rho_n(\infty)$ for $n \leq 3$ were first given by Beltran in [5, (7)]; it is known that $1 - \rho_n(\infty)$ decays like $e^{-n^2(\log 3)/4}$ as $n \to \infty$ (see [2] and [3]).

In Section 4, we show how to obtain an exact formula for $\rho_n(\infty)$ for any given n. More precisely, using the de Bruijn identity [4] for calculating certain determinantal integrals, we express $\rho_n(\infty)$ as the Pfaffian of an explicit $n' \times n'$ matrix, where $n' := 2\lceil n/2 \rceil$, whose entries are given in terms of values of the gamma and incomplete beta functions at integers and half-integers. Indeed, let Γ denote the usual gamma function $\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\mathrm{d}x$, and β_t the usual incomplete beta function $\beta_t(i,j) = \int_0^t x^{i-1}(1-x)^{j-1}\mathrm{d}x$. Then we have the following theorem giving expressions for $\rho_n(\infty)$:

Theorem 1.3 *Let $n \geq 1$ be any integer, and define $n' := 2\lceil n/2 \rceil$. When real n-ary quadratic forms are chosen according to the n-dimensional Gaussian Orthogonal Ensemble, the probability of isotropy over $\mathbb{R}$ is given by*

$$\rho_n(\infty) = 1 - \frac{\mathrm{Pf}(A)}{2^{(n-1)(n+4)/4}\prod_{m=1}^{n}\Gamma(\frac{m}{2})}, \tag{5}$$

where A is the $n' \times n'$ skew-symmetric matrix whose (i,j)-entry a_{ij} is given for $i < j$ by

$$a_{ij} = \begin{cases} 2^{i+j-2}\Gamma(\frac{i+j}{2})\left(\beta_{\frac{1}{2}}(\frac{i}{2},\frac{j}{2}) - \beta_{\frac{1}{2}}(\frac{j}{2},\frac{i}{2})\right) & \text{if } i < j \leq n, \\ 2^{i-1}\Gamma(\frac{i}{2}) & \text{if } i < j = n+1. \end{cases} \tag{6}$$

(*Note that the second case in* (6) *arises only when n is odd.*)

Theorem 1.3 allows one to calculate $\rho_n(\infty)$ exactly in closed form for any given n. In particular, it follows from the Pfaffian representation in Theorem 1.3 that $\rho_n(\infty)$ is a polynomial in π^{-1} of degree at most $\lfloor \frac{n+1}{4} \rfloor$ with coefficients in $\mathbb{Q}(\sqrt{2})$ (see Remark 4.1). In Table 1, we give the resulting formulae for $\rho_n(\infty)$ for all $n \leq 8$, and also provide numerical approximations. (For any $n > 8$, we have $\rho_n(\infty) \approx 1$ to more than 10 decimal places!)

Combining Theorems 1.1, 1.2, and 1.3, we finally obtain the following theorem giving the probability ρ_n that a random integral quadratic form in n variables has an integral zero:

Theorem 1.4 *Let D be any nice (i.e., piecewise smooth and rapidly decaying) distribution. Then the probability ρ_n^D that a random integral quadratic form in n variables with respect to the distribution D is isotropic is given by*

$$\rho_n^D = \begin{cases} 0 & \text{if } n \leq 3; \\ \rho_4^D(\infty)\prod_p \left(1 - \dfrac{p^3}{4(p+1)^2(p^4+p^3+p^2+p+1)}\right) & \text{if } n = 4; \\ \rho_n^D(\infty) & \text{if } n \geq 5. \end{cases}$$

If $D = \mathrm{GOE}$ is the GOE *distribution, then the quantities $\rho_n(\infty) = \rho_n^D(\infty)$ are as given in Theorem 1.3.*

n	$\rho_n(\infty) =$	$\rho_n(\infty) \approx$
1	0	0
2	$\sqrt{2}/2$	0.7071067811
3	$1/2 + \sqrt{2}\,\pi^{-1}$	0.9501581580
4	$1/2 + \sqrt{2}/8 + \pi^{-1}$	0.9950865814
5	$3/4 + (2/3 + \sqrt{2}/12)\pi^{-1}$	0.9997197706
6	$3/4 + 7\sqrt{2}/64 + (37/48 - \sqrt{2}/3)\pi^{-1}$	0.9999907596
7	$7/8 + (47/120 + 109\sqrt{2}/480)\pi^{-1} - (32\sqrt{2}/45)\pi^{-2}$	0.9999998239
8	$7/8 + 9\sqrt{2}/256 + (2377/3840 - 53\sqrt{2}/480)\pi^{-1} - (32/45)\pi^{-2}$	0.9999999980

Table 1: Probability $\rho_n(\infty)$ that a random n-ary quadratic form over $\mathbb{R}$ from the GOE distribution is isotropic, for $n \leq 8$.

In particular, when $D = \text{GOE}$, we have $\rho_n = 0$ for $n = 1$, 2, and 3, while for $n = 4$ we obtain the expression (2) for ρ_4. For $n \geq 5$, we have $\rho_n = \rho_n(\infty)$, and so the values of ρ_n are as given by Theorem 1.3. Theorem 1.4 shows that $n = 4$ is in a sense the most interesting case, as all places play a nontrivial role in the final answer.

It is also interesting to compare how the probabilities change if instead of the GOE we use the uniform distribution U on quadratic forms, where each coefficient of the quadratic form is chosen uniformly in the interval $[-1/2, 1/2]$. While the quantities $\rho_n^{\mathrm{U}}(\infty)$ can easily be expressed as explicit definite integrals, it seems unlikely that they can be evaluated in compact and closed form for general n in this case. Using numerical integration, or a Monte Carlo approximation, we can compute $\rho_n^{\mathrm{U}}(\infty) \approx 0, 0.627, 0.901, 0.982, 0.998$, and > 0.999 for $n = 1$, 2, 3, 4, 5, and 6, respectively. It is known (see, e.g., [1, Theorem. 2.3.5]) that $1 - \rho_n^{\mathrm{U}}(\infty)$ decays faster than e^{-cn} for some constant $c > 0$; the actual rate of decay is likely even faster. (Indeed, this follows for Wigner random matrices in general, so applies as well to the GOE case.)

In particular, we have $\rho_4^{\mathrm{U}} = \rho_4^{\mathrm{U}}(\infty) \prod_p \rho_4(p) \approx 97.0\%$, which is slightly smaller than the GOE probability $\rho_4^{\mathrm{GOE}} \approx 98.3\%$. We summarize the values of ρ_n^D, and provide numerical values in the cases of the uniform and GOE distributions, in Table 2.

Let $N_n(X)$ denote the number of integral n-ary quadratic forms that are isotropic over $\mathbb{Q}$ whose coefficients are less than X in absolute value. Since the probabilities of isotropy are equal to 0 for $n \leq 3$, the question arises as to how $N_n(X)$ grows in these cases as $X \to \infty$. For $n = 1$, we have trivially $N(X) = 1$ for any $X > 0$. For $n = 2$, it was shown by Dörge [6] and Kuba [10] that $X^2 \log X \ll N_2(X) \ll X^2 \log X$, and this was recently refined to an exact asymptotic formula, $N_2(X) \sim c_2 X^2 \log X$ for an explicit positive constant c_2, by Dubickas [7]. For $n = 3$, it was shown by Serre [13] that $N_3(X) = O(X^6/\sqrt{\log X})$, who also conjectured that $N_3(X) > EX^6/\sqrt{\log X}$ for some positive constant E; this conjecture was recently resolved by Hooley [9]. In conjunction with these results for $n \leq 3$, the result of Theorem 1.4 determines the rates of growth of $N_n(X)$ for all $n \geq 1$, and indeed proves the existence of main terms in the asymptotics of $N_n(X)$ for all $n \neq 3$. (Whether $N_3(X) \sim c_3 X^6/\sqrt{\log X}$ for some positive constant c_3 remains an open question.)

n	ρ_n^D	ρ_n^{U}	ρ_n
1	0	0	0
2	0	0	0
3	0	0	0
4	$\rho_4^D(\infty)\prod_p\left(1-\frac{p^3}{4(p+1)^2(p^4+p^3+p^2+p+1)}\right)$	$\approx 97.0\%$	$\approx 98.3\%$
5	$\rho_5^D(\infty)$	$\approx 99.8\%$	$> 99.9\%$
≥ 6	$\rho_n^D(\infty)$	$> 99.9\%$	$> 99.9\%$

Table 2: Probability that a random integral quadratic form in n variables is isotropic, for a general distribution D, for the uniform distribution, and for the GOE distribution.

This paper is organized as follows. In Section 2, we prove the product formula in Theorem 1.1. The theorem is known in the case of the uniform distribution U (or indeed any uniform distribution supported on a box) for any $n \geq 4$ by the work of Poonen and Voloch [11], which in turn depends on the Ekedahl sieve [8]. To complete the proof of Theorem 1.1, we first prove directly that both sides of (4) are equal to 0 for $n \leq 3$. For $n \geq 4$, we prove that (4) is true for a general nice distribution D by approximating D by a finite weighted average of uniform box distributions, where the result is already known. The condition that D is rapidly decreasing (as in the case of $D = \mathrm{GOE}$) plays a key role in the proof; indeed, we show how counterexamples to (4) can be constructed when this condition does not hold.

In Section 3, we then prove Theorem 1.2, i.e., we determine for each n the exact p-adic density of n-ary quadratic forms over $\mathbb{Z}_p$ that are isotropic. The outline of the proof is as follows. First, we note that a quadratic form in n variables defined over $\mathbb{Z}_p$ can be anisotropic only if its reduction modulo p has either two conjugate linear factors over $\mathbb{F}_{p^2}$ or a repeated linear factor over $\mathbb{F}_p$. We first compute the probability of each of these cases occurring, which is elementary. We then determine the probabilities of isotropy in each of these two cases by developing certain recursive formulae for these probabilities, in terms of other suitable quantities, which allow us to solve and obtain exact algebraic expressions for these probabilities for each value of n. We note that our general argument shows in particular that quadratic forms in $n \geq 5$ variables over $\mathbb{Q}_p$ are always isotropic, thus yielding a new recursive proof of this well-known fact.

Finally, we prove Theorem 1.3 in Section 4, i.e., we determine for each n the probability that a random real n-ary quadratic form from the GOE distribution is indefinite. We accomplish this by first expressing, as a certain determinantal integral, the probability that an $n \times n$ symmetric matrix from the GOE distribution has all positive eigenvalues. We then show how this determinantal integral can be evaluated using the de Bruijn identity [4], allowing us to obtain an expression for the probability of positive definiteness in terms of the Pfaffian of an explicit skew-symmetric matrix A, as given in Theorem 1.3. We note that the values of these probabilities were known previously for $n \leq 3$ (cf. [5, (7)]).

We end this introduction by remarking that the analogues of Theorems 1.2 and 1.4 also hold over a general local or global field, respectively. Here, we define global densities of quadrics as in [11, §4]; more general densities with respect to "nice distributions" could also be defined in an analogous manner. Indeed, the analogue of Theorem 1.1 holds (with the identical proof), where the product on the right hand side of (4) should be taken over all finite and infinite places of the number field (the densities at the complex places are all equal to 1, since all quadratic forms over $\mathbb{C}$ are isotropic). Theorem 1.2 also holds over any finite extension of $\mathbb{Q}_p$, with the same proof, provided that when making substitutions in the proofs we replace p by a uniformiser, and when computing probabilities we replace p by the order of the residue field.

2 The local product formula: Proof of Theorem 1.1

Let D be any nice (piecewise smooth and rapidly decaying) distribution. Our aim in this section is to prove the following three assertions from the introduction:

(a) $\rho_n^D(p)$ is equal to the probability $\rho_n(p)$ that a random n-ary quadratic form over $\mathbb{Z}_p$, with respect to the usual additive measure on $\mathbb{Z}_p^{n(n+1)/2}$, is isotropic over $\mathbb{Z}_p$;

(b) $\rho_n^D(\infty)$ is equal to the probability that a random n-ary quadratic form over $\mathbb{R}$, with respect to the distribution D, is indefinite; and

(c) $\rho_n^D = \rho_n^D(\infty)\prod_p \rho_n(p)$ (i.e., Theorem 1.1 holds).

Items (a) and (b) are trivial in the case that $D = \mathrm{U}$ is the uniform distribution, or more generally when D is any distribution $\mathrm{U}(\vec{a},\vec{b})$ that is constant on a box $[\vec{a},\vec{b}] := [a_1,b_1]\times\cdots\times[a_{n(n+1)/2},b_{n(n+1)/2}]$ and 0 outside this box; here $\vec{a} = (a_1,\ldots,a_{n(n+1)/2})$ and $\vec{b} = (b_1,\ldots,b_{n(n+1)/2})$ are vectors in $\mathbb{R}^{n(n+1)/2}$ such that $a_i < b_i$ for all i.

Meanwhile, Theorem 1.1 for $n \geq 4$, in the case that D is the uniform distribution U, follows from the work of Poonen and Voloch [11, Theorem 3.6] (which establishes the product formula for the probability that an integral quadratic form with respect to the distribution D is locally soluble), together with the Hasse–Minkowski Theorem (which states that a quadratic form is isotropic if and only if it is locally soluble). In fact, the proof of [11, Theorem 3.6] (which in turn relies on the Ekedahl's sieve [8]) immediately adapts to the case where $D = \mathrm{U}(\vec{a},\vec{b})$ without essential change.

To show that Theorem 1.1 holds also when $D = \mathrm{U}(\vec{a},\vec{b})$ and $n \leq 3$, it suffices to prove that in this case both sides of (4) are equal to 0. To see this, we may use Theorem 1.2, which does not rely on the results of this section, and which states that the probability that a random n-ary quadratic form over $\mathbb{Z}_p$ is isotropic is equal to $\rho_n(p) = 0$, $1/2$, or $1 - p/(2(p+1)^2)$ for $n = 1$, 2, or 3, respectively. This immediately implies that the right hand side of (4) is zero. To see that the left hand side of (4) is zero, we note that if a quadratic form over $\mathbb{Z}$ is isotropic, then it must be isotropic over $\mathbb{Z}_p$ for all p (the easy direction of the Hasse–Minkowski Theorem). By the Chinese Remainder Theorem, the (limsup of the) probability ρ_n^D that a random integral n-ary quadratic form is isotropic with respect to the distribution $D = \mathrm{U}(\vec{a},\vec{b})$ is at most

$$\prod_{p<Y} \rho_n(p)$$

for any $Y > 0$. Letting Y now tend to infinity shows that $\rho_n^D = 0$ for $n = 1$, 2, or 3, i.e., the left hand side of (4) is also zero.

Thus we have established items (a)–(c), for all n, in the case that $D = \mathrm{U}(\vec{a}, \vec{b})$ is a constant distribution supported on a box $[\vec{a}, \vec{b}]$. Clearly (a)–(c) then must hold also for any finite weighted average of such box distributions $\mathrm{U}(\cdot, \cdot)$.

To show that (a)–(c) hold for general nice distributions D, we make use of the following elementary lemma regarding integration of rapidly decaying functions.

Lemma 2.1 *Let f be any piecewise smooth rapidly decaying function on $\mathbb{R}^m$. Then*

$$\int f(y)\mathrm{d}y = \lim_{X\to\infty} \frac{1}{X^m} \sum_{y\in\mathbb{Z}^m} f(y/X). \tag{7}$$

Proof: For any $N > 0$, let $f_N(y)$ be equal to $f(y)$ if $|y| \leq N$, and 0 otherwise. Then f_N is piecewise smooth with bounded support, and so is Riemann integrable. Thus we have

$$\int f_N(y)\mathrm{d}y = \lim_{X\to\infty} \frac{1}{X^m} \sum_{y\in\mathbb{Z}^m} f_N(y/X). \tag{8}$$

Since f is rapidly decreasing, for any $\varepsilon > 0$ we may choose N large enough so that $\int_{|y|>N} |f(y)|\mathrm{d}y < \varepsilon$ *and* $(1/X^m)\sum_{y\in\mathbb{Z}^m,\,|y/X|>N} |f(y/X)| < \varepsilon$ for any $X \geq 1$. For this value of N, the left hand side of (8) is within ε of the left hand side of (7), while for each $X \geq 1$, the expression in the limit on the right hand side of (8) is within ε of the expression in the limit on the right hand side of (7). Since we have equality in (8), we conclude that the left hand side of (7) is within 2ε of both the $\liminf_{X\to\infty}$ and the $\limsup_{X\to\infty}$ of the expression in the limit of the right hand side of (7). Since ε is arbitrarily small, we have proven (7). $\square$

Note that Lemma 2.1 does not necessarily hold if we drop the condition that f is rapidly decaying. For example, if f is the characteristic function of a finite-volume region having a cusp going off to infinity containing a rational line through the origin (and thus infinitely many lattice points on that line), then the left hand side of (7) is finite while the expression in the limit on the right hand side of (7) is infinite for any rational value of X.

Lemma 2.1 implies in particular that

$$\lim_{X\to\infty} \frac{1}{X^{n(n+1)/2}} \sum_{Q \text{ integral}} D(Q/X) = 1 \tag{9}$$

for any nice distribution D.

Now any piecewise smooth rapidly decaying function can be approximated arbitrarily well by a finite linear combination of characteristic functions of boxes. Let D be a nice distribution. For any $\varepsilon > 0$, we may find a nice distribution D_ε that is a finite weighted average of box distributions $\mathrm{U}(\cdot, \cdot)$, such that

$$\int |D(y) - D_\varepsilon(y)|\mathrm{d}y < \varepsilon. \tag{10}$$

By Lemma 2.1, we then have

$$\lim_{X\to\infty} \frac{1}{X^{n(n+1)/2}} \sum_{Q \text{ integral}} |D(Q/X) - D_\varepsilon(Q/X)| < \varepsilon. \tag{11}$$

To show that $\rho_n^D(p) = \rho_n(p)$, we note that

$$\begin{aligned}
\rho_n(p) = \rho_n^{D_\varepsilon}(p) &= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{Z}_p} D_\varepsilon(Q/X)}{\sum_{Q \text{ integral}} D_\varepsilon(Q/X)} &(12)\\
&= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{Z}_p} D_\varepsilon(Q/X)}{X^{n(n+1)/2}} &(13)\\
&= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{Z}_p} D(Q/X) + E(X,\varepsilon)}{X^{n(n+1)/2}} &(14)\\
&= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{Z}_p} D(Q/X) + E(X,\varepsilon)}{\sum_{Q \text{ integral}} D(Q/X)}, &(15)
\end{aligned}$$

where for sufficiently large X we have $|E(X,\varepsilon)| < \varepsilon X^{n(n+1)/2}$ by (11); here the first equality follows because D_ε is a finite weighted average of box distributions $\mathrm{U}(\cdot,\cdot)$, the second equality follows from the definition (3), and the third and fifth equalities follow from (9). Letting ε tend to 0 in (15) now yields $\rho_n(p) = \rho_n^D(p)$, proving item (a) for general nice distributions D.

Analogously, we have

$$\begin{aligned}
\int_{Q \text{ isotropic}/\mathbb{R}} D_\varepsilon(Q)\mathrm{d}Q = \rho_n^{D_\varepsilon}(\infty) &= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{R}} D_\varepsilon(Q/X)}{\sum_{Q \text{ integral}} D_\varepsilon(Q/X)} &(16)\\
&= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{R}} D(Q/X) + E'(X,\varepsilon)}{\sum_{Q \text{ integral}} D(Q/X)}, &(17)
\end{aligned}$$

where again for sufficiently large X we have $|E'(X,\varepsilon)| < \varepsilon X^{n(n+1)/2}$. By (10), the leftmost expression in (16) approaches $\int_{Q \text{ isotropic}/\mathbb{R}} D(Q)\mathrm{d}Q$ as $\varepsilon \to 0$, while expression (17) approaches $\rho_n^D(\infty)$ by definition (3). This thus proves item (b) for general nice distributions. In particular, we have also proven that

$$\lim_{\varepsilon\to 0} \rho_n^{D_\varepsilon}(\infty) = \rho_n^D(\infty). \tag{18}$$

Finally, we have in a similar manner:

$$\begin{aligned}
\rho_n^{D_\varepsilon}(\infty)\prod_p \rho_n(p) = \rho_n^{D_\varepsilon} &= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{Z}} D_\varepsilon(Q/X)}{\sum_{Q \text{ integral}} D_\varepsilon(Q/X)} &(19)\\
&= \lim_{X\to\infty} \frac{\sum_{Q \text{ integral,isotropic}/\mathbb{Z}} D(Q/X) + E''(X,\varepsilon)}{\sum_{Q \text{ integral}} D(Q/X)}, &(20)
\end{aligned}$$

where again for sufficiently large X we have $|E''(X,\varepsilon)| < \varepsilon X^{n(n+1)/2}$. By (18), the leftmost expression in (19) approaches $\rho_n^D(\infty)\prod_p \rho_n(p)$ as $\varepsilon \to 0$, while expression (20) approaches ρ_n^D by definition. We have proven also item (c) for general nice distributions, as desired.

3 The density of n-ary quadratic forms over $\mathbb{Z}_p$ that are isotropic: Proof of Theorem 1.2

3.1 Preliminaries on n-ary quadratic forms over $\mathbb{Z}_p$

Fix a prime p. For any free $\mathbb{Z}_p$-module V of finite rank, there is a unique additive p-adic Haar measure μ_V on V which we always normalize so that $\mu_V(V) = 1$. All densities/probabilities are computed with respect to this measure. In this section, we take $V = V_n$ to be the $n(n+1)/2$-dimensional $\mathbb{Z}_p$-module of n-ary quadratic forms over $\mathbb{Z}_p$. We then work out the density $\rho_n(p)$ (i.e. measure with respect to μ_V) of the set of n-ary quadratic forms over $\mathbb{Z}_p$ that are isotropic.

We start by observing that a primitive n-ary quadratic form over $\mathbb{Z}_p$ can be anisotropic only if, either: (I) the reduction modulo p factors into two conjugate linear factors defined over a quadratic extension of $\mathbb{F}_p$, or (II) the reduction modulo p is a constant times the square of a linear form over $\mathbb{F}_p$. For if the reduced form has rank ≥ 3, then, after setting some variables to zero we obtain a smooth conic. But a conic over a finite field always has a rational point (see, for example, [12, Chapter I, Cor. 2]); this lifts to a $\mathbb{Q}_p$-point by Hensel's Lemma. Note that if $p = 2$, this argument is still valid, provided that we define the rank correctly, i.e. it is not the rank of the corresponding symmetric matrix, but rather the codimension of the singular locus in the ambient projective space.

Let $\xi_1^{(n)}$ and $\xi_2^{(n)}$ be the probabilities of Cases I and II, i.e. the densities of these two types of quadratic forms in V_n. Then

$$\xi_0^{(n)} = 1 - \xi_1^{(n)} - \xi_2^{(n)} - \frac{1}{p^{n(n+1)/2}}$$

is the probability that a form is primitive, but not in Cases I or II. Let $\alpha_1^{(n)}$ (resp. $\alpha_2^{(n)}$) be the probability of isotropy for quadratic forms in Case I (resp. Case II). Then

$$\rho_n(p) = \xi_0^{(n)} + \xi_1^{(n)}\alpha_1^{(n)} + \xi_2^{(n)}\alpha_2^{(n)} + \frac{1}{p^{n(n+1)/2}}\rho_n(p),$$

implying that

$$\rho_n(p) = \frac{p^{n(n+1)/2}}{p^{n(n+1)/2} - 1}\big(\xi_0^{(n)} + \xi_1^{(n)}\alpha_1^{(n)} + \xi_2^{(n)}\alpha_2^{(n)}\big). \tag{21}$$

3.2 Some counting over finite fields

Let $\eta_1^{(n)}$ (resp. $\eta_2^{(n)}$) be the probability that a quadratic form is in Case I (resp. Case II) given the "point condition" that the coefficient of x_1^2 is a unit. Similarly, let $\nu_1^{(n)}$ be the probability that a quadratic form is in Case I given the "line condition" that the binary quadratic form $Q(x_1, x_2, 0, \ldots, 0)$ is irreducible modulo p. Note that it is impossible to be in Case II given the line condition, but we may also define $\nu_2^{(n)} = 0$. Set $\eta_0^{(n)} = 1 - \eta_1^{(n)} - \eta_2^{(n)}$ and $\nu_0^{(n)} = 1 - \nu_1^{(n)} - \nu_2^{(n)} = 1 - \nu_1^{(n)}$. The values of $\xi_j^{(n)}$, $\eta_j^{(n)}$, $\nu_j^{(n)}$, are given by the following easy lemma.

Lemma 3.1 *The probabilities that a random quadratic form over $\mathbb{Z}_p$ is in Case I or Case II are as follows.*

- *Case I (all; relative to point condition; relative to line condition)*

$$\xi_1^{(n)} = \frac{(p^n-1)(p^n-p)}{2(p+1)p^{n(n+1)/2}}; \qquad \eta_1^{(n)} = \frac{p^{n-1}-1}{2p^{n(n-1)/2}}; \qquad \nu_1^{(n)} = \frac{1}{p^{(n-1)(n-2)/2}}.$$

- *Case II (all; relative to point condition; relative to line condition)*

$$\xi_2^{(n)} = \frac{p^n-1}{p^{n(n+1)/2}}; \qquad \eta_2^{(n)} = \frac{1}{p^{n(n-1)/2}}; \qquad \nu_2^{(n)} = 0.$$

Proof: Case I: There are $(p^{2n}-1)/(p^2-1)$ linear forms over $\mathbb{F}_{p^2}$ up to scaling; subtracting the $(p^n-1)/(p-1)$ which are defined over $\mathbb{F}_p$, dividing by 2 to account for conjugate pairs and then multiplying by $p-1$ for scaling gives $\frac{(p^n-1)(p^n-p)}{2(p+1)}$ Case I forms, and hence the value of $\xi_1^{(n)}$.

Similarly, the number of Case I quadratic forms satisfying the point condition is $(p^{2(n-1)} - p^{n-1})(p-1)/2$. Dividing by the probability $1-1/p$ of the point condition holding gives $p^n(p^{n-1}-1)/2$ and hence the value of $\eta_1^{(n)}$.

Lastly, the number of Case I quadratic forms satisfying the line condition is $p^{2n-3}(p-1)^2/2$; dividing by the probability $\xi_1^{(2)}$ of the line condition holding gives p^{2n-1}, and hence the value of $\nu_1^{(n)}$.

Case II is similar and easier: the number of Case II quadratic forms is p^n-1, of which p^n-p^{n-1} satisfy the point condition and none satisfy the line condition; the given formulae follow. $\square$

3.3 Recursive formulae

We now outline our strategy for computing the densities $\rho_n(p)$ using (21), by evaluating $\alpha_j^{(n)}$ for $j = 1, 2$. If a quadratic form is in Case I, then we may make a linear change of variables (using a change of coordinate matrix in $\mathrm{GL}_n(\mathbb{Z}_p)$, which preserves density), transforming it so that its reduction is an irreducible binary form in only two variables. Now isotropy forces the values of those variables, in any primitive vector giving a zero, to be multiples of p; so we may scale those variables by p and divide the form by p. Similarly, if a form is in Case II, then we transform it so that its reduction is the square of a single variable, then scale that variable and divide out. After carrying out this process once, we again divide into cases and repeat the procedure, which leads us back to an earlier situation but with either the line or point conditions, which we need to allow for. All these transformations clearly preserve the property of isotropy.

To make this precise, we introduce some extra notation for the probability of isotropy for quadratic forms which are in Case I or Case II after the initial transformation: let $\beta_1^{(n)}$ (resp. $\beta_2^{(n)}$) be the probability of isotropy given we are in Case I (resp. Case II) after one step when the original quadratic form was in Case I, and similarly $\gamma_1^{(n)}$ (resp. $\gamma_2^{(n)}$) the probability of isotropy given we are in Case I (resp. Case II) after one step when the original quadratic form was in Case II.

Lemma 3.2 .

1. $\alpha_1^{(2)} = 0$, *and for* $n \geq 3$,

$$\alpha_1^{(n)} = \xi_0^{(n-2)} + \xi_1^{(n-2)}\beta_1^{(n)} + \xi_2^{(n-2)}\beta_2^{(n)} + \frac{1}{p^{(n-1)(n-2)/2}}(\nu_0^{(n)} + \nu_1^{(n)}\alpha_1^{(n)} + \nu_2^{(n)}\alpha_2^{(n)}).$$

2. $\alpha_2^{(1)} = 0$, *and for* $n \geq 2$,

$$\alpha_2^{(n)} = \xi_0^{(n-1)} + \xi_1^{(n-1)}\gamma_1^{(n)} + \xi_2^{(n-1)}\gamma_2^{(n)} + \frac{1}{p^{n(n-1)/2}}(\eta_0^{(n)} + \eta_1^{(n)}\alpha_1^{(n)} + \eta_2^{(n)}\alpha_2^{(n)}).$$

Proof: We have $\alpha_1^{(2)} = 0$ since a binary quadratic form that is irreducible over $\mathbb{F}_p$ is anisotropic. Now assume that $n \geq 3$, and (for Case I) $Q(x_1, \ldots, x_n) \pmod p$ has two conjugate linear factors. Without loss of generality, the reduction modulo p is a binary quadratic form in x_1 and x_2. Now any primitive vector giving a zero of Q must have its first two coordinates divisible by p, so replace $Q(x_1, \ldots, x_n)$ by $\frac{1}{p}Q(px_1, px_2, x_3, \ldots, x_n)$. The reduction modulo p is now a quadratic form in $x_3, \ldots, x_n$. If the new Q is identically zero modulo p, then, after dividing it by p, we obtain a new integral form that lands in Cases I and II with probabilities $\nu_1^{(n)}$ and $\nu_2^{(n)}$, respectively, since it satisfies the line condition; otherwise, we divide into cases as before, with the probabilities of being in each case given by $\xi_j^{(n-2)}$.

The result for $\alpha_2^{(n)}$ is proved similarly: without loss of generality the reduction modulo p is a quadratic form in x_1 only, we replace $Q(x_1, \ldots, x_n)$ by $\frac{1}{p}Q(px_1, x_2, \ldots, x_n)$, whose reduction modulo p is a quadratic form in $x_2, \ldots, x_n$. If the new Q is identically zero modulo p, then, after dividing by p, we have an integral form that lands in Cases I and II with probabilities $\eta_1^{(n)}$ and $\eta_2^{(n)}$, respectively, since it satisfies the point condition; otherwise, we divide into cases, with probabilities $\xi_j^{(n-1)}$. □

It remains to compute $\beta_1^{(n)}$ (for $n \geq 4$), $\beta_2^{(n)}$ (for $n \geq 3$), $\gamma_1^{(n)}$ (for $n \geq 3$) and $\gamma_2^{(n)}$ (for $n \geq 2$). Since $\xi_1^{(1)} = 0$, we do not need to compute $\beta_1^{(3)}$ or $\gamma_1^{(2)}$, which are in any case undefined.

Lemma 3.3 .

(i) *If* $n \geq 4$ *then* $\beta_1^{(n)} = \nu_0^{(n-2)} + \nu_1^{(n-2)}\beta_1^{(n)}$*; also,* $\beta_1^{(4)} = 0$.

(ii) *If* $n \geq 3$ *then* $\beta_2^{(n)} = \nu_0^{(n-1)} + \nu_1^{(n-1)}\gamma_1^{(n)}$*; also,* $\beta_2^{(3)} = 0$.

(iii) *If* $n \geq 3$ *then* $\gamma_1^{(n)} = \eta_0^{(n-2)} + \eta_1^{(n-2)}\beta_1^{(n)} + \eta_2^{(n-2)}\beta_2^{(n)}$*; also,* $\gamma_1^{(3)} = 0$.

(iv) *If* $n \geq 2$ *then* $\gamma_2^{(n)} = \eta_0^{(n-1)} + \eta_1^{(n-1)}\gamma_1^{(n)} + \eta_2^{(n-1)}\gamma_2^{(n)}$*; also,* $\gamma_2^{(2)} = 0$.

Proof: In Case I, the initial transformation leads to a quadratic form for which the valuations of the coefficients satisfy[1]

$$
\begin{array}{ccccccc}
\geq 1 & \geq 1 & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\
 & \geq 1 & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\
 & & \geq 0 & \geq 0 & \geq 0 & \dots & \geq 0 \\
 & & & \geq 0 & \geq 0 & \dots & \geq 0 \\
 & & & & \geq 0 & \dots & \geq 0 \\
 & & & & & \ddots & \vdots \\
 & & & & & & \geq 0
\end{array} \tag{22}
$$

and $\beta_1^{(n)}$ (resp. $\beta_2^{(n)}$) are the probabilities of isotropy given that the reduction modulo p of the form in $x_3, x_4, \dots, x_n$ is in Case I (resp. Case II).

Similarly, in Case II the initial transformation leads to

$$
\begin{array}{ccccccc}
= 1 & \geq 1 & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\
 & \geq 0 & \geq 0 & \geq 0 & \geq 0 & \dots & \geq 0 \\
 & & \geq 0 & \geq 0 & \geq 0 & \dots & \geq 0 \\
 & & & \geq 0 & \geq 0 & \dots & \geq 0 \\
 & & & & \geq 0 & \dots & \geq 0 \\
 & & & & & \ddots & \vdots \\
 & & & & & & \geq 0
\end{array} \tag{23}
$$

and $\gamma_1^{(n)}$ (resp. $\gamma_2^{(n)}$) are the probabilities of isotropy given that the reduction modulo p of the form in $x_2, x_3, \dots, x_n$ is in Case I (resp. Case II).

(i) To evaluate $\beta_1^{(n)}$ we may assume, after a second linear change of variables, that we have

$$
\begin{array}{ccccccc}
\geq 1 & \geq 1 & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\
 & \geq 1 & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\
 & & \geq 0 & \geq 0 & \geq 1 & \dots & \geq 1 \\
 & & & \geq 0 & \geq 1 & \dots & \geq 1 \\
 & & & & \geq 1 & \dots & \geq 1 \\
 & & & & & \ddots & \vdots \\
 & & & & & & \geq 1
\end{array}
$$

and that the reductions modulo p of both $\frac{1}{p}Q(x_1, x_2, 0, \dots, 0)$ and $Q(0, 0, x_3, x_4, 0, \dots, 0)$ are irreducible binary quadratic forms. Any zero of Q must satisfy $x_3 \equiv x_4 \equiv 0 \pmod{p}$. This gives a contradiction when $n = 4$, so that $Q(x_1, \dots, x_4)$ is anisotropic, and $\beta_1^{(4)} = 0$. Otherwise, replacing $Q(x_1, \dots, x_n)$ by $\frac{1}{p}Q(x_3, x_4, px_1, px_2, x_5, \dots, x_n)$ brings us back to the situation in (22). Now, however, the line condition holds, so that Cases I and II occur with probabilities $\nu_1^{(n-2)}$ and $\nu_2^{(n-2)} = 0$ instead of $\xi_1^{(n-2)}$ and $\xi_2^{(n-2)}$.

[1] In this and the similar arrays which follow, we put into position (i, j) the known condition on $v(a_{i,j})$, so the top left entry refers to the coefficient of x_1^2, the top right to $x_1 x_n$ and the bottom right to x_n^2.

(ii) To evaluate $\beta_2^{(n)}$, we may assume that the valuations of the coefficients satisfy

$$\begin{array}{cccccc} \geq 1 & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\ & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\ & & =0 & \geq 1 & \dots & \geq 1 \\ & & & \geq 1 & \dots & \geq 1 \\ & & & & \ddots & \vdots \\ & & & & & \geq 1 \end{array}$$

and that the reduction modulo p of $\frac{1}{p}Q(x_1, x_2, 0, \dots, 0)$ is an irreducible binary quadratic form. If $n = 3$ then Q is anisotropic, and $\beta_2^{(3)} = 0$. Otherwise, replacing $Q(x_1, \dots, x_n)$ by $\frac{1}{p}Q(x_2, x_3, px_1, x_4, \dots, x_n)$ brings us back to the situation in (23) but with the line condition, so that Cases I and II occur with probabilities $\nu_1^{(n-1)}$ and $\nu_2^{(n-1)}$ instead of $\xi_1^{(n-1)}$, $\xi_2^{(n-1)}$.

(iii) For $\gamma_1^{(n)}$, we may assume that the valuations of the coefficients satisfy

$$\begin{array}{cccccc} =1 & \geq 1 & \geq 1 & \geq 1 & \dots & \geq 1 \\ & \geq 0 & \geq 0 & \geq 1 & \dots & \geq 1 \\ & & \geq 0 & \geq 1 & \dots & \geq 1 \\ & & & \geq 1 & \dots & \geq 1 \\ & & & & \ddots & \vdots \\ & & & & & \geq 1 \end{array}$$

and the reduction of $Q(0, x_2, x_3, 0, \dots, 0)$ modulo p is irreducible. Any zero of Q now satisfies $x_2 \equiv x_3 \equiv 0 \pmod p$. When $n = 3$ this gives a contradiction, so $Q(x_1, x_2, x_3)$ is anisotropic, and $\gamma_1^{(3)} = 0$. Otherwise, replacing $Q(x_1, \dots, x_n)$ by $\frac{1}{p}Q(x_3, px_1, px_2, x_4, \dots, x_n)$ brings us back to the situation in (22) but with the point condition, so that Cases I and II occur with probabilities $\eta_1^{(n-2)}$ and $\eta_2^{(n-2)}$.

(iv) Lastly, for $\gamma_1^{(n)}$, we may assume that the valuations of the coefficients satisfy

$$\begin{array}{ccccc} =1 & \geq 1 & \geq 1 & \dots & \geq 1 \\ & =0 & \geq 1 & \dots & \geq 1 \\ & & \geq 1 & \dots & \geq 1 \\ & & & \ddots & \vdots \\ & & & & \geq 1. \end{array}$$

If $n = 2$ then $Q(x_1, x_2)$ is anisotropic , and $\gamma_2^{(2)} = 0$. Otherwise, replacing $Q(x_1, \dots, x_n)$ by $\frac{1}{p}Q(x_2, px_1, x_3, \dots, x_n)$ brings us back to the situation in (23) but with the point condition. $\square$

3.4 Conclusion

Using Lemmas 3.1 and 3.3 we can compute $\beta_j^{(n)}$ and $\gamma_j^{(n)}$ for $j = 1, 2$ and all n: we first determine β_1 from Lemma 3.3 (i), then $\beta_2^{(n)}$ and $\gamma_1^{(n)}$ together using Lemma 3.3 (ii,iii), and finally $\gamma_2^{(n)}$ using

Lemma 3.3 (iv). The following table gives the result:

	$\beta_1^{(n)}$	$\beta_2^{(n)}$	$\gamma_1^{(n)}$	$\gamma_2^{(n)}$
$n=2$	$-$	$-$	$-$	0
$n=3$	$-$	0	0	$1/2$
$n=4$	0	$(2p+1)/(2p+2)$	$(p+2)/(2p+2)$	$1-(p/(4(p^2+p+1)))$
$n\geq 5$	1	1	1	1

Now, using Lemma 3.2, we compute $\alpha_1^{(n)}$ and $\alpha_2^{(n)}$:

	$\alpha_1^{(n)}$	$\alpha_2^{(n)}$
$n=2$	0	$1/(2p+2)$
$n=3$	$1/(p+1)$	$(p+2)/(2p+2)$
$n=4$	$1-(p^3/(2(p+1)(p^2+p+1)))$	$1-(p^3/(4(p+1)(p^3+p^2+p+1)))$
$n\geq 5$	1	1

Finally, we compute $\rho_n(p)$ using (21), yielding the values stated in Theorem 1.2.

Note that our proof of Theorem 1.2 also yields a (recursive) algorithm to determine whether a quadratic form over $\mathbb{Q}_p$ is isotropic. Tracing through the algorithm, we see that, for a quadratic form of nonzero discriminant, only finitely many recursive iterations are possible (since we may organize the algorithm so that at each such iteration the discriminant valuation is reduced), i.e., the algorithm always terminates. In particular, when $n\geq 5$, our algorithm always yields a zero for any n-ary quadratic form of nonzero discriminant; hence every nondegenerate quadratic form in $n\geq 5$ variables is isotropic.

4 The density of n-ary quadratic forms over $\mathbb{R}$ that are indefinite: Proof of Theorem 1.3

4.1 Preliminaries on the Gaussian Orthogonal Ensemble (GOE)

We wish to calculate the probability $\rho_n(\infty)$ that a real symmetric matrix M from the n-dimensional GOE has an indefinite spectrum. The distribution of matrix entries in the GOE is invariant under orthogonal transformations. Since real symmetric matrices can be diagonalised by an orthogonal transformation, the GOE measure can be written directly in terms of the eigenvalues $\lambda(M)$, yielding the distribution

$$\mathbb{P}\big(\lambda(M)\in[\lambda+\mathrm{d}\lambda)\big)=\frac{1}{Z_n^{\mathrm{GOE}}}|\Delta(\lambda)|\prod_{i=1}^{n}e^{-\frac{1}{4}\lambda_i^2}\mathrm{d}\lambda_i; \tag{24}$$

here

$$\Delta(\lambda):=\prod_{1\leq i<j\leq n}(\lambda_j-\lambda_i)=\det(\varphi_i(\lambda_j)),$$

where $(\varphi_i(\lambda_j))=(\lambda_j^{i-1})$ is a Vandermonde matrix, and the normalizing factor Z_n^{GOE} is given by

$$Z_n^{\mathrm{GOE}}=n!(2\pi)^{\frac{n}{2}}2^{(n(n-1)/4+n/2)}\prod_{j=1}^{n}\frac{\Gamma(\frac{j}{2})}{\Gamma(\frac{1}{2})}. \tag{25}$$

See, for example, [1, (2.5.4)].

Note that the probability that the matrix M is indefinite is related to the probability p_n^+ that all its eigenvalues are positive by

$$\rho_n(\infty) = 1 - \mathbb{P}(\text{positive definite}) - \mathbb{P}(\text{negative definite}) = 1 - 2p_n^+, \tag{26}$$

where the second equality follows by symmetry. Below we will calculate p_n^+, and hence obtain the value of $\rho_n(\infty)$.

4.2 de Bruijn's identity

We recall a useful result from [4, §4] for calculating determinantal integrals of the type we will need. As a generalisation of an expression for the volume of the space of symmetric unitary matrices, de Bruijn considered integrals of the form:

$$\Omega = \int\cdots\int_{a\le x_1\le\cdots\le x_n\le b} \det_{1\le i,j\le n}(\varphi_i(x_j))\mathrm{d}\mu(x_1)\ldots\mathrm{d}\mu(x_n). \tag{27}$$

Recall that the Pfaffian of a skew-symmetric matrix $A = (a_{ij})$ is given by

$$\mathrm{Pf}(A) = \sum_{\tau} \mathrm{sgn}(\tau) a_{i_1,j_1} a_{i_2,j_2} \cdots a_{i_s,j_s}, \tag{28}$$

where τ ranges over all partitions

$$\tau = \{(i_1, j_1), (i_2, j_2), \ldots (i_s, j_s)\}$$

of $n = 2s$ where $i_k < i_{k+1}$ and $i_k < j_k$. The sign is of the corresponding permutation

$$\tau = \begin{bmatrix} 1 & 2 & 3 & 4 & \cdots & 2s \\ i_1 & j_1 & i_2 & j_2 & \cdots & j_s \end{bmatrix}.$$

The integral (27) may be rewritten as the Pfaffian of either an $n \times n$ skew-symmetric matrix if n is even, or an $(n+1)\times(n+1)$ skew-symmetric matrix if n is odd. More precisely, let $n' := 2\lceil n/2\rceil$; then we have $\Omega = \mathrm{Pf}(A)$, where A is the $n' \times n'$ skew-symmetric matrix whose (i, j)-entry a_{ij} is given for $i < j$ by

$$a_{ij} = \begin{cases} \int_a^b \int_a^b \mathrm{sign}(y-x)\varphi_i(x)\varphi_j(y)\mathrm{d}\mu(x)\mathrm{d}\mu(y) & \text{if } i < j \le n; \\ \int_a^b \varphi_j(x)\mathrm{d}\mu(x) & \text{if } i < j = n+1. \end{cases} \tag{29}$$

The second case occurs only when n is odd. Note that this expression for Ω is valid in any ordered measure space; below, we will use $\mathrm{d}\mu(x) = e^{-x^2/4}\mathrm{d}x$, where $\mathrm{d}x$ is the Lebesgue measure on $\mathbb{R}$.

The Pfaffian form of the integral is found by expanding the determinant and using a signature function to keep track of the signs and the ordering of the x_i. This signature function of n variables can be broken up into a sum of products of two-variable pieces (and a one-variable piece if n is odd) and thus the integral can be factorised into a sum of products of two (and one) dimensional integrals which is recognised as of the form (28) for a matrix with entries (29).

4.3 Calculation of ρ_n^∞

For a matrix M from the GOE, the joint distribution of the eigenvalues $\lambda_1(M) \leq \lambda_2(M) \leq \cdots \leq \lambda_n(M)$ is given by

$$\frac{n!}{Z_n^{\mathrm{GOE}}} \mathbf{1}_{\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n} |\Delta(\lambda)| \prod_{i=1}^n e^{-\frac{1}{4}\lambda_i^2} \mathrm{d}\lambda_i. \tag{30}$$

The ordering in the domain of integration below means that we can replace $|\Delta(\lambda)|$ by $\Delta(\lambda)$. It then follows that p_n^+ is given by the integral

$$\begin{aligned} p_n^+ &= \frac{n!}{Z_n^{\mathrm{GOE}}} \int \cdots \int_{0 \leq \lambda_1 \leq \cdots \leq \lambda_n \leq \infty} \Delta(\lambda) \prod_{i=1}^n e^{-\frac{1}{4}\lambda_i^2} \mathrm{d}\lambda_i \\ &= \frac{n!}{Z_n^{\mathrm{GOE}}} \int \cdots \int_{0 \leq \lambda_1 \leq \cdots \leq \lambda_n \leq \infty} \det(\varphi_i(\lambda_j)) \prod_{i=1}^n e^{-\frac{1}{4}\lambda_i^2} \mathrm{d}\lambda_i \\ &= \frac{n!}{Z_n^{\mathrm{GOE}}} \mathrm{Pf}(A), \end{aligned} \tag{31}$$

where the last equality follows from the result of §4.2. Here, $A = (a_{ij})$, where for $i < j \leq n$ we define

$$\begin{aligned} a_{ij} &= \int_0^\infty \int_0^\infty \mathrm{sign}(y - x) x^{i-1} y^{j-1} e^{-\frac{x^2+y^2}{4}} \mathrm{d}x\mathrm{d}y \\ &= 2^{i+j-2} \Gamma\left(\frac{i+j}{2}\right) \left(\beta_{\frac{1}{2}}\left(\frac{i}{2}, \frac{j}{2}\right) - \beta_{\frac{1}{2}}\left(\frac{j}{2}, \frac{i}{2}\right)\right), \end{aligned} \tag{32}$$

and for n odd we also set $a_{i,n+1} = 2^{i-1}\Gamma(\frac{i}{2})$. Here the gamma and incomplete beta functions are as defined in §1. From the resulting skew-symmetric matrix A, we may evaluate (31) to determine $\rho_n(\infty)$, yielding Theorem 1.3. Explicit values of $\rho_n(\infty)$ are displayed in Table 1 for $n \leq 8$.

Remark 4.1 It is easily shown that the matrix entries a_{ij} in Theorem 1.3 are of the form x or $x\sqrt{\pi}$ for $x \in \mathbb{Q}(\sqrt{2})$, in accordance with whether $i + j$ is even or odd. Let $s = \lceil n/2 \rceil$, so that A is a $2s \times 2s$ matrix. Then after re-ordering the rows and columns we have

$$\mathrm{Pf}(A) = \pm \mathrm{Pf}\begin{pmatrix} A_1 & \sqrt{\pi} A_2 \\ -\sqrt{\pi} A_2^t & A_3 \end{pmatrix} = \pm \pi^{s/2} \mathrm{Pf}\begin{pmatrix} A_1 & A_2 \\ -A_2^t & \pi^{-1} A_3 \end{pmatrix}$$

where A_1, A_2 and A_3 are $s \times s$ matrices with entries in $\mathbb{Q}(\sqrt{2})$. Since $\prod_{m=1}^n \Gamma(m/2) = \pi^{s/2} y$ for some $y \in \mathbb{Q}$, it follows by Theorem 1.3 and the definition of the Pfaffian that $\rho_n(\infty)$ is a polynomial in π^{-1} having coefficients in $\mathbb{Q}(\sqrt{2})$ and degree at most $\lfloor s/2 \rfloor = \lfloor (n+1)/4 \rfloor$.

Acknowledgments

We thank Carlos Beltran, Jonathan Hanke, Peter Sarnak, Jean-Pierre Serre and Terence Tao for helpful conversations. The first author (Bhargava) was supported by a Simons Investigator Grant and NSF grant DMS-1001828; the second (Cremona) and fifth (Keating) were supported by EPSRC Programme Grant EP/K034383/1 LMF: L-Functions and Modular Forms; the fifth (Keating) was also supported by a grant from The Leverhulme Trust, a Royal Society Wolfson Merit Award, a Royal Society Leverhulme Senior Research Fellowship, and by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-10-1-3088.

References

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics **118**, Cambridge University Press, 2009.

[2] G. Ben Arous and A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy, *Probab. Theory Relat. Fields*, **108**, (1997) 517–542.

[3] D. Dean and S. Majumdar, Extreme Value Statistics of Eigenvalues of Gaussian Random Matrices, *Phys. Rev. E* **77**, 041108 (2008) (`http://arxiv.org/abs/0801.1730`).

[4] N. G. de Bruijn, On some multiple integrals involving determinants, *J. Indian Math. Soc.* **19**, 133–151 (1955).

[5] J-P. Dedieu and G. Malajovich, On the number of minima of a random polynomial, *J. Complexity* **24**, 89–108 (2008) (`http://arxiv:math/0702360`).

[6] K. Dörge, Abschätzung der Anzahl der reduziblen Polynome, Math. Ann. **160** (1965), 59-63.

[7] A. Dubickas, On the number of reducible polynomials of bounded naive height, *Manuscripta Mathematica* **144** (2014), 439–456.

[8] T. Ekedahl, An infinite version of the Chinese remainder theorem, *Comment. Math. Univ. St. Paul.* **40**, 53–59 (1991).

[9] C. Hooley, On ternary quadratic forms that represent zero. II. *J. Reine Angew. Math.* **602** (2007), 179-225.

[10] G. Kuba, On the distribution of reducible polynomials, *Math. Slovaca* **59** (2009), 349–356.

[11] B. Poonen and P. Voloch, Random Diophantine equations, *Arithmetic of Higher-Dimensional Algebraic Varieties*, *Progress in Mathematics* **226**, 175–184, Birkhäuser, Boston, MA, 2004.

[12] J. P. Serre, *A Course in Arithmetic*. Springer 1996.

[13] J. P. Serre, Specialisation des élements de $\mathrm{Br}_2(Q(T_1, \ldots, T_n))$, *C. R. Acad. Sci. Paris Sér I Math.* **311** (1990), 397–402.