RANDOM FIELDS AND THE ENUMERATIVE GEOMETRY OF LINES ON REAL AND COMPLEX HYPERSURFACES

SAUGATA BASU, ANTONIO LERARIO, ERIK LUNDBERG, CHRIS PETERSON

Abstract. We introduce a probabilistic framework for the study of real and complex enumerative geometry of lines on hypersurfaces. This can be considered as a further step in the original Shub-Smale program of studying the real zeros of random polynomial systems. Our technique is general, and it also applies, for example, to the case of the enumerative geometry of flats on complete intersections.

We derive a formula expressing the average number E_n of real lines on a random hypersurface of degree $2n-3$ in $\mathbb{R}\mathrm{P}^n$ in terms of the expected modulus of the determinant of a special random matrix. In the case $n = 3$ we prove that the average number of real lines on a random cubic surface in $\mathbb{R}\mathrm{P}^3$ equals:

$$E_3 = 6\sqrt{2} - 3.$$

This technique can also be applied to express the number C_n of complex lines on a *generic* hypersurface of degree $2n-3$ in $\mathbb{C}\mathrm{P}^n$ in terms of the expectation of the square of the modulus of the determinant of a random Hermitian matrix. As a special case, we recover the classical statement $C_3 = 27$.

We determine, at the logarithmic scale, the asymptotic of the quantity E_n, by relating it to C_n (whose asymptotic has been recently computed in [19]). Specifically we prove that:

$$\lim_{n\to\infty} \frac{\log E_n}{\log C_n} = \frac{1}{2}.$$

Finally we show that this approach can be used to compute the number $R_n = (2n-3)!!$ of real lines, counted with their *intrinsic signs* (as defined in [28]), on a generic real hypersurface of degree $2n-3$ in $\mathbb{R}\mathrm{P}^n$.

1. Introduction

1.1. **Overview.** A classical problem in enumerative geometry is the count of the number of linear spaces satisfying some geometric conditions (e.g. the number of lines on a generic cubic surface). These problems are usually approached with cohomological techniques, which turn out to be very powerful over the complex numbers but which give almost no information over the reals. In this paper we introduce a novel, more analytical approach to these questions. This comes after adopting a probabilistic point of view – the main idea is the replacement of the word *generic* with *random*. Of course over the complex numbers this gives the same answer, but it also allows to compute other quantities especially meaningful over the reals, where the generic number of solutions is not defined (e.g. the signed count or the average count).

Our contribution with this paper is thus in two different directions. On one hand we present new results in the emerging field of *random real algebraic geometry*, with the investigation of the real average count. A program on random real algebraic geometry started in the 1990s with pioneer works of A. Edelman, E. Kostlan, M. Shub and S. Smale [10, 34, 11, 23, 36, 35] on random polynomial system solving, with a particular emphasis on asymptotic studies. Our result can be

2010 *Mathematics Subject Classification.* 14N15, 14Pxx, 14Nxx, 60G60.

considered as a further step in the original Shub-Smale program: if the "complexity of Bezout's theorem" papers dealt with the basic problem of counting points in the intersection of random hypersurfaces in projective space, here we count solutions to more advanced intersection theory problems. The results presented here are the first results on sections of a homogeneous bundle on a more general Grassmannian (beyond projective space).

On the other hand our methods provide a new way for performing the classical computations over the complex numbers and the signed computation over the reals (see Section 1.4), essentially reducing them to computing the average of some random determinants. Although the technique that we introduce is very general, as it will become clear throughout the paper, we concentrate on the problem of enumeration of lines on projective hypersurfaces.

1.2. **Lines on a random cubic.** One of the most classical statements from enumerative geometry asserts that there are 27 lines lying on a generic cubic surface in $\mathbb{C}\mathrm{P}^3$: Cayley proved this result in the nineteenth century [7]. Today many proofs exist, some more geometric (using the isomorphism between the blowup of $\mathbb{C}\mathrm{P}^2$ at six generic points and the generic cubic in $\mathbb{C}\mathrm{P}^3$), others more topological (exploiting properties of Chern classes), see for example [9, 12]. (A new proof, based on a probabilistic argument, will be given in this paper, see Corollary 8 below.)

In this paper we will mostly be interested in similar problems over the reals, where the situation is more complicated. Schläfli [31] showed that if the cubic surface is *real* and smooth then the number of real lines lying on it is either 27, 15, 7, or 3. In the blow-up point of view, these correspond respectively to the following four cases for the blow-up points: all of them are real (27); 4 are real and the other two are complex conjugate (15); 2 are real and there are two pairs of complex conjugate (7); there are 3 pairs of complex conjugate (3). Reading the number of real lines from the coefficients of the polynomial f defining the cubic is a difficult problem. It is interesting to ask for a probabilistic treatment:

(1) *"How many real lines are expected to lie on a random real cubic surface in* $\mathbb{R}\mathrm{P}^3$*?"*.

To make this question rigorous we should clarify what we mean by "random". Here we will sample f from the so-called *Kostlan* ensemble. We endow the space $\mathbb{R}[x_0, x_1, x_2, x_3]_{(3)}$ of real homogeneous polynomials of degree 3 with a probability distribution by defining a random polynomial f as a linear combination:

$$f = \sum_{|\alpha|=3} \xi_\alpha x_0^{\alpha_0} x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}, \quad \alpha = (\alpha_0, \ldots, \alpha_3),$$

where $|\alpha| = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3$ denotes the length of the multi-index α, and the coefficients ξ_α are real, independent, centered Gaussian variables with variances

$$\sigma_\alpha^2 = \binom{3}{\alpha} := \frac{3!}{\alpha_0!\alpha_1!\alpha_2!\alpha_3!}.$$

(A similar definition is considered for homogeneous polynomials of degree d in several variables, see Section 2.1 below.) This probability distribution is $O(4)$-invariant, that is, it is invariant under an orthogonal change of variables (so there are no preferred points or directions in $\mathbb{R}\mathrm{P}^3$). Moreover it is the only (up to a multiplicative constant) $O(4)$-invariant Gaussian probability distribution on $\mathbb{R}[x_0, x_1, x_2, x_3]_{(3)}$ for which f can be defined as a linear combination of monomials with independent Gaussian coefficients. This distribution is also natural from the point of view of algebraic geometry, as it can be equivalently[1] obtained by sampling a random polynomial

[1] "Equivalently" for our purposes, as we will only be interested in properties that depend on the zero set of a polynomial.

uniformly from the (projectivization of the) space of real polynomials with the metric induced by the inclusion in the space of complex polynomials with the Fubini-Study metric, see Section 2.1 below for more details.

Let $Gr(2,4)$ denote the Grassmannian of two dimensional subspaces of $\mathbb{R}^4$ (or equivalently, the Grassmannian of lines in $\mathbb{R}\mathrm{P}^3$). Denoting by $\tau_{2,4}$ the tautological bundle of $Gr(2,4)$ we see that the homogeneous cubic polynomial f defines, by restriction, a section σ_f of $\mathrm{sym}^3(\tau_{2,4}^*)$. A line ℓ lies on the cubic surface defined by $\{f=0\}$ if and only if $\sigma_f(\ell)=0$. In this way, when f is random, σ_f is a random section of $\mathrm{sym}^3(\tau_{2,4}^*)$ and the section σ_f can be considered as a generalization of a *random field*. In this case, the random field takes values in a vector bundle but on a trivializing chart it reduces to the standard construction (see [1]). There is a powerful technique that allows one to compute the average number of zeros of a random field, the so called *Kac-Rice formula* [1]. The adaptation of this technique to random sections of vector bundles (we discuss it in the proofs below) yields the existence of a function $\rho : Gr(2,4) \to \mathbb{R}$ (the *Kac-Rice density*) such that the answer to question (1) can be written as:

$$E_3 = \int_{Gr(2,4)} \rho\, \omega_{Gr(2,4)}$$

(here $\omega_{Gr(2,4)}$ is the volume density of $Gr(2,4)$ with respect to the metric induced by the Plücker embedding). In our case, by invariance of the problem under the action of the orthogonal group, ρ is a constant function. However, the computation of this constant is still delicate and is one of the main results of this paper. To be precise, we show in Theorem 5 that its value equals:

$$\rho = \frac{6\sqrt{2}-3}{2\pi^2}.$$

Since the volume of $Gr(2,4)$ is equal to $2\pi^2$, we are able to conclude that $E_3 = 6\sqrt{2}-3$.

Notice that over the complex numbers whatever reasonable (i.e. absolutely continuous with respect to Lebesgue measure) probability distribution we consider on the space of the coefficients of the defining polynomial, a random complex cubic surface will be generic in the sense of algebraic geometry with probability one (this is essentially due to the fact that complex discriminants have real codimension two). For this reason our technique, which produces an *expected* answer over the reals, gives the *generic* answer when adapted to the complex setting.

If we take the coefficients of f to be *complex* Gaussians (with the same variances as for the real case) we get a *complex Kostlan* polynomial. Equivalently we can sample from the projectivization of the space of complex polynomials with the natural Fubini-Study metric (see Section 2.1 below). In this way σ_f is a random holomorphic section of $\mathrm{sym}^3(\tau_{2,4}^*)$ where $\tau_{2,4}$ is now the tautological bundle of $Gr_{\mathbb{C}}(2,4)$ (the Grassmannian of complex two dimensional subspaces of $\mathbb{C}^4$). Here again one can use the Kac-Rice approach, interpreting 4 complex variables as 8 real variables and the number C_3 of complex lines on a generic cubic can be rewritten as an integral (with respect to the volume density induced by the complex Plücker embedding) over $Gr_{\mathbb{C}}(2,4)$ of a function $\rho_{\mathbb{C}}$. The model is invariant under the action of $U(4)$ so that $\rho_{\mathbb{C}}$ is constant. The evaluation of this constant is now much easier than its real analogue (see Corollary 8), and it seems more combinatorial in nature. We compute the value to be:

$$\rho_{\mathbb{C}} \equiv \frac{324}{\pi^4}.$$

The volume of $Gr_{\mathbb{C}}(2,4)$ is equal to $\frac{\pi^4}{12}$ and this implies $C_3 = 27$.

Going back to the real case, Segre [32] divided the real lines lying on a cubic surface into *hyperbolic* lines and *elliptic* lines, and showed that the difference between the number, h, of hyperbolic lines and the number, e, of elliptic lines always satisfy the relation $h - e = 3$. In this direction we note that the technique below can also be used to compute this difference; this will correspond to a signed Kac-Rice density and its computation amounts to removing the modulus in the expectation of the determinant of the matrix in Theorem 2. Here we know that the signed count is an invariant number R_3 of the generic cubic [28] and consequently that the average answer is the generic one; the explicit computation $R_3 = \frac{3}{2}\mathbb{E}\det \hat{J}_3 = 3$ recovers this number.

Remark 1. Allcock, Carlson, and Toledo [3] have studied the moduli space of real cubic surfaces from the point of view of hyperbolic geometry. They compute the orbifold Euler characteristic (which is proportional to the hyperbolic volume) of each component of the moduli space. One could take the weighted average of the number of real lines, weighted by the volume of the corresponding component, as an "average count". This yields the number $\frac{239}{37}$, see [3, Table 1.2]. However this way of counting does not have a natural probabilistic interpretation and generalization to $n > 3$.

Remark 2. A Monte Carlo experiment with 1500 random trials drawn from the Kostlan ensemble yielded 868 cubics with 3 lines, 501 with 7 lines, 127 with 15 lines, and 4 with 27 lines. The average number of lines obtained from these 1500 samples is 5.416. The approximate value of $6\sqrt{2} - 3$ is 5.485. These values are quite close. Each of the 4 cases, for the number of lines, occur with positive probability. It would be interesting to obtain exact formulas expressing the probability of a random trial to yield each of the 4 individual cases.

1.3. **Lines on hypersurfaces.** The same scheme can be applied to the problem of enumeration of lines on hypersurfaces of degree $2n-3$ in $\mathbb{RP}^n$ and $\mathbb{CP}^n$ (if the degree is larger than $2n-3$ then a random hypersurface is too "curved" and will contain no lines while if the degree is smaller than $2n-3$ then lines will appear in families). Theorem 2 and its complex analogue Theorem 7 give a general recipe for the computation of the average number E_n of real lines on a real Kostlan hypersurface and the number C_n of lines on a generic complex hypersurface. Zagier [19] showed that C_n can be computed as the coefficient of x^{n-1} in the polynomial

$$p_n(x) = (1-x)\prod_{j=0}^{2n-3}(2n-3-j+jx)$$

and found that the asymptotic behavior for C_n as $n \to \infty$ is given by:

$$C_n \sim \sqrt{\frac{27}{\pi}}(2n-3)^{2n-\frac{7}{2}}(1+O(n^{-1})). \tag{2}$$

One could notice some similarities between the expression that we derive for C_n in Theorem 7 and the analogous expression for E_n from Theorem 2. In Theorem 12 we will make this comparison more rigorous, by proving that the two quantities are related by:

$$\lim_{n\to\infty}\frac{\log E_n}{\log C_n} = \frac{1}{2}. \tag{3}$$

This is roughly saying that the average number of real lines on a random real hypersurface is the "square root" of the number of complex lines (this is true up to an exponential factor; note that the numbers E_n and C_n are super-exponential).

After this paper was written, it was brought to our attention in a private communication that Peter Bürgisser conjectured the square root law (as stated in Theorem 12) in a talk given at a conference in 2008, at the Bernoulli Center in Lausanne, where he also gave numerical evidence with regards to the value of E_3.

Remarkably, Okonek and Teleman [28] and Finashin and Kharlamov [14] have shown that over the reals it is still possible to associate a numerical invariant R_n to a generic hypersurface, counting the number of real lines with canonically assigned *signs*. Even though the bundles $\mathrm{sym}^{2n-3}(\tau^*_{2,n+1})$ and $TGr(2,n+1)$ might not be orientable themselves, Okonek and Teleman [28] introduce a notion of *relative orientation* between them that allows one to define an Euler number up to a sign. In this case it turns out there is a canonical relative orientation, and each zero of a generic section of $\mathrm{sym}^{2n-3}(\tau^*_{2,n+1})$ comes with an *intrinsic sign* (see Section 2.4 below). The relative Euler number has the usual property that the sum of the zeroes of a generic section counted with signs is invariant; for the problem of lines on hypersurfaces it equals[2]:

$$R_n = (2n-3)!!.$$

In particular note that $E_n \geq (2n-3)!!$. Using the random approach, in Proposition 3 we give an alternative proof of $R_n = (2n-3)!!$.

1.4. Generic, signed, and average counts.

We can summarize the above discussion by stating the following inequalities, that offer a broad point of view on the problem of lines on hypersurfaces (here $b > 1$ is a universal constant whose existence is proved in Proposition 11):

$$R_n \leq E_n \leq \sqrt{n} b^n C_n^{1/2}.$$

These inequalities compare three quantities: the generic, the average and the signed count. More generally, given $k, d, n \in \mathbb{N}_0$ such that

$$\binom{k+d}{d} = (k+1)(n-k), \tag{4}$$

one can consider the problem of determining the number of projective subspace of dimension k lying on a hypersurface H_d of degree d in $\mathbb{R}\mathrm{P}^n$. (Equation (4) says that the rank of $\mathrm{sym}^d(\tau^*_{k+1,n+1})$ equals the dimension of $Gr(k+1,n+1)$; this, combined with [8, Theorem 2.1], ensures that a generic section of $\mathrm{sym}^d(\tau^*_{k+1,n+1})$ vanishes at isolated points).

For this class of problems, the number of complex projective k-spaces on the *generic* $H_d \subset \mathbb{C}\mathrm{P}^n$ can be computed (nontrivially!) using the splitting principle. It is interesting to observe that the method that we introduce here for the problem of lines can be easily extended to the more general case (in fact the first part of the proof of Theorem 7 is completely general and works for all Grassmannians), and yields a new way for performing the computation in the generic case.

The *signed* count, however, in general is not always well-defined and even if defined it might give *zero*. As observed in [15] and [28] this count is well defined[3] as soon as the following identity of first Stiefel-Whitney classes holds:

$$w_1(Gr(k+1,n+1)) = w_1(\mathrm{sym}^d(\tau^*_{k+1,n+1})).$$

(This identity asserts that the two bundles are relatively orientable.) The simplest case is of course when $(k,d,n) = (0,d,1)$, which corresponds to counting zeroes of a univariate polynomial. In this case, since $w_1(\mathbb{R}\mathrm{P}^1) = 0$, and $w_1(\mathrm{sym}^d(\tau^*_{1,d})) = \frac{1}{2}(1+(-1)^{d+1}) \bmod 2$, the signed count in the above sense is well defined only when d is even – but this count gives zero. We observe that, as for the complex case, our method allows in principle to compute the Euler class of

[2] Here and below in the paper we use the notation $m!!$ (the "double-factorial") to denote the product of all positive integers that have the same parity as n; e.g. when $m = 2n-3$ is odd, $(2n-3)!!$ denotes the product of all odd numbers smaller than or equal to $2n-3$.

[3] In fact recently Kharlamov and Finashin have identified a class of problems for which this type of signed count is defined [15], including some classical Schubert problems.

the corresponding bundle, essentially reducing it to the evaluation of the expectation of the determinant (without the modulus!) of a random matrix, as done in Proposition 3.

On the other hand for this class of problems the *average* count is always well defined, but its evaluation is much more difficult, because of the presence of the modulus in the random determinant of Theorem 2. In this direction, the second author of the current paper together with P. Bürgisser have recently studied the average count in the case of special Schubert problems [6], relating it to the volume of a convex body in the tangent space to the Grassmannian.

Remark 3. An interesting question on the subject is whether an enumerative problem is *fully real*, meaning that there are real configurations of the constraints for which all a priori complex solutions are real. Some fundamental results on this aspect of real Schubert calculus are contained in papers of Sottile [39, 37, 38, 40], Eremenko and Gabrielov [13], Mukhin, Tarasov and Varchenko [25], Vakil [43] (this list is by no means complete). In particular it is proved in [43] that Schubert calculus is fully real. We do not know if the problem of counting lines on hypersurfaces is fully real, and Theorem 12, combined with Theorem 13, indicates that "typically" this problem is far from full reality.

Acknowledgements. The question (1) motivating this paper was posed to the second author by F. Sottile. This paper originated during the stay of the authors at SISSA (Trieste), supported by Foundation Compositio Mathematica. We are very grateful to the anonymous referee whose careful reading and valuable insights improved the exposition dramatically.

2. The real case

2.1. **Real Kostlan polynomials.** Let $E = \mathbb{R}^{n+1}$. Then, $\mathrm{sym}^d(E^*)$ is identified with the space, $\mathbb{R}[x_0, x_1, \ldots, x_n]_{(d)}$, of homogeneous polynomials of degree d in $n+1$ variables. A Gaussian ensemble can be specified by choosing a scalar product on $\mathrm{sym}^d(E^*)$ in which case f is sampled according to the law:

$$\text{Probability}(f \in A) = \frac{1}{v_{n,d}} \int_A e^{-\frac{\|f\|^2}{2}} df,$$

where $v_{n,d}$ is a normalizing constant that makes the integrand into a probability density function, and df is the volume form on $\mathrm{sym}^d(E^*)$ induced by the chosen scalar product. The ensemble we will consider, known as the *Kostlan ensemble*, results from choosing as a scalar product the Bombieri product[4], defined as:

$$\langle f, g \rangle_B = \frac{1}{d!\pi^{n+1}} \int_{\mathbb{C}^{n+1}} f(z)\overline{g(z)} e^{-\|z\|^2} dz.$$

The following expression relates the Bombieri norm of $f = \sum_{|\alpha|=d} f_\alpha z^\alpha$ with its coefficients in the monomial basis (see [26, Equation (10)]):

$$\|f\|_B = \left(\sum_{|\alpha|=d} |f_\alpha|^2 \binom{d}{\alpha}^{-1} \right)^{\frac{1}{2}}.$$

Here we are using multi-index notation $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_n)$, with $|\alpha| := \alpha_0 + \alpha_1 + \cdots + \alpha_n$, $z^\alpha := z_0^{\alpha_0} \cdots z_n^{\alpha_n}$, and $\binom{d}{\alpha} := \frac{d!}{\alpha_0! \cdots \alpha_n!}$. The Bombieri product defined above can be described in

[4] This inner product has also been referred to as the "Fischer product", especially in the field of holomorphic PDE (e.g., see [21, Ch. 15, 18]) after H.S. Shapiro made a detailed study [33] reviving methods from E. Fischer's 1917 paper [16]. The names of H. Weyl, V. Bargmann, and V. A. Fock have also been attached to this inner product.

a more algebro-geometric language as follows. The standard Hermitian inner product on $\mathbb{C}^{n+1}$ induces a Hermitian inner product, $(\cdot,\cdot)_d$ on the line bundles $\mathcal{O}_{\mathbb{C}\mathrm{P}^n}(d)$, and this in turn induces a Hermitian inner product on the finite-dimensional space of holomorphic sections, $H^0(\mathcal{O}_{\mathbb{C}\mathrm{P}^n}(d))$, by

$$\langle f, g\rangle = \int_{\mathbb{C}\mathrm{P}^n} (f(z), g(z))_d \, dz,$$

where the integration is with respect to the volume form on $\mathbb{C}\mathrm{P}^n$ induced by the standard Fubini-Study metric on $\mathbb{C}\mathrm{P}^n$. The restriction of this last Hermitian inner product to real sections of $\mathcal{O}_{\mathbb{C}\mathrm{P}^n}(d)$ agrees up to normalization by a constant with the Bombieri product (after identifying $\mathbb{R}[x_0, x_1, \dots, x_n]_{(d)}$ with the space of real holomorphic sections of $\mathcal{O}_{\mathbb{C}\mathrm{P}^n}(d)$). Thus, amongst all possible $O(n+1)$-invariant measures on $\mathbb{R}[x_0, x_1, \dots, x_n]_{(d)}$, the Kostlan measure seems to be the most natural one from the point of view of algebraic geometry.

An equivalent, and often more practical, approach to Gaussian ensembles is to build f as a linear combination, using independent Gaussian coefficients, of the vectors in an orthonormal basis for the associated scalar product. The Kostlan ensemble has the distinguished property of being the unique Gaussian ensemble that is invariant under any orthogonal changes of variables while simultaneously having the monomials as an orthogonal basis. We can sample f by placing independent Gaussians ξ_α in front of the monomial basis (here again α is a multi-index):

$$f(x) = \sum_{|\alpha|=d} \xi_\alpha x^\alpha, \quad \xi_\alpha \sim N\left(0, \binom{d}{\alpha}\right).$$

Note that there are other orthogonally-invariant models, but their explicit description requires special functions (Gegenbauer polynomials — in order to build a basis of spherical harmonics), see [22, 17].

2.2. **The Kac-Rice formula.** A technical tool that we will use in the paper is the Kac-Rice formula, which allows to compute the expectation of the number of real zeroes of a real Gaussian map. There are various (more or less equivalent) formulations of this result, see for instance [1, 4]; here we quote the version from [4].

Theorem 1 (Theorem 6.3 from [4]). *Let $U \subset \mathbb{R}^N$ be an open set and $Z : U \to \mathbb{R}^N$ be a random map such that:*

(i) *Z is Gaussian;*
(ii) *Z is almost surely of class C^1;*
(iii) *for every $t \in U$ the random variable $Z(t)$ has a nondegenerate distribution;*
(iv) *the probability that Z has a degenerate zero in U is zero.*

Then, denoting by $p_{Z(t)}$ the density function of $Z(t)$, for every Borel subset $B \subset U$ we have:

$$\mathbb{E}\#\left(\{Z=0\}\cap B\right) = \int_B \mathbb{E}\left\{|\det(JZ(t))| \,\middle|\, Z(t)=0\right\} p_{Z(t)}(0)dt, \tag{5}$$

where $JZ(t)$ denotes the Jacobian matrix of $Z(t)$.

Remark 4. When U has the additional structure of a Riemannian manifold, we can use the volume form ω_U induced by the Riemannian metric to write (5) in a more pleasant way. Arguing exactly as in the proof of [1, Theorem 12.1.1], we have:

$$\mathbb{E}\#\left(\{Z=0\}\cap B\right) = \int_U \mathbb{E}\left\{|\det(\hat{J}Z(t))| \,\middle|\, Z(t)=0\right\} p_{Z(t)}(0)\cdot\omega_U(t), \tag{6}$$

where now $\hat{J}Z(t)$ denotes the matrix of the derivatives of the components of Z with respect to an orthonormal frame field.

2.3. **A general construction.** Consider a random vector $v = (v_1, \ldots, v_{2n-3})$ in $\mathbb{R}^{2n-3}$ whose entries are independent Gaussian variables distributed as:

$$v_j \sim N\left(0, \binom{2n-4}{j-1}\right) \qquad j = 1, \ldots, 2n-3.$$

Let now $v^{(1)}, \ldots, v^{(n-1)}$ be independent random vectors all distributed as v. We define the random $(2n-2) \times (2n-2)$ matrix $\hat{J}_n$ as:

$$\hat{J}_n = \begin{bmatrix} v_1^{(1)} & 0 & \ldots & v_1^{(n-1)} & 0 \\ v_2^{(1)} & v_1^{(1)} & & v_2^{(n-1)} & v_1^{(n-1)} \\ \vdots & v_2^{(1)} & & \vdots & v_2^{(n-1)} \\ v_{2n-3}^{(1)} & \vdots & & v_{2n-3}^{(n-1)} & \vdots \\ 0 & v_{2n-3}^{(1)} & \ldots & 0 & v_{2n-3}^{(n-1)} \end{bmatrix}.$$

Theorem 2. *The average number E_n of real lines on a random Kostlan hypersurface of degree $2n-3$ in $\mathbb{R}\mathrm{P}^n$ is*

$$E_n = \left(\frac{(2n-3)^{n-1}}{\Gamma(n)} \prod_{k=0}^{2n-3} \binom{2n-3}{k}^{-1/2}\right) \mathbb{E}|\det \hat{J}_n|.$$

Proof. We will start with a general construction that works for the Grassmannian $Gr(k,m)$ and only at the end of the proof will we specialize to the case $k = 2, m = n+1$.

Let $Gr^+(k,m)$ denote the Grassmannian of *oriented* k-planes in $\mathbb{R}^m$. As a Riemannian manifold $Gr^+(k,m)$ is identified with its image under the spherical Plücker embedding. The image is equal to the set of simple, norm-one vectors in $\Lambda^k(\mathbb{R}^m)$; moreover the Riemannian metric induced on $Gr^+(k,m)$ from this embedding is the same as the metric induced by declaring the quotient map $SO(m) \to Gr^+(k,m)$ to be a Riemannian submersion (see [24] for more details). We denote by g this metric and by $\omega_{Gr(k,m)}$ the associated volume density.

Let also $\tau_{k,m}^*$ denote the dual of the tautological bundle (on $Gr^+(k,m)$) and consider its symmetric d-th power $\mathrm{sym}^d(\tau_{k,m}^*)$. Note that a homogeneous polynomial $f \in \mathbb{R}[x_1, \ldots, x_m]_{(d)}$ defines a real holomorphic section σ_f of the bundle $\mathrm{sym}^d(\tau_{k,m}^*)$, simply by considering the restriction $\sigma_f(w) = f|_w$ (here w denotes the variable in $Gr^+(k,m)$). Clearly σ_f also defines a section of $\tau_{k,m}^*$ as a bundle over $Gr(k,m)$ (the standard Grassmannian) and a k-space $w \in Gr(k,m)$ is contained in the zero set of f if and only if the section σ_f vanishes at w.

Since $Gr^+(k,m)$ is a Riemannian double covering of $Gr(k,m)$, when the rank of $\mathrm{sym}^d(\tau_{k,m}^*)$ equals the dimension of $Gr^+(k,m)$, generically σ_f has finitely many zeroes; in this case the number of zeros of σ_f as a section of the bundle $\mathrm{sym}^d(\tau_{k,m}^*)$ on $Gr^+(k,m)$ is twice the number of zeros this section has as a section of $\mathrm{sym}^d(\tau_{k,m}^*)$ on $Gr(k,m)$, and consequently:

$$2 \cdot \mathbb{E}\#\{w \in Gr(k,m) \,|\, \sigma_f(w) = 0\} = \mathbb{E}\#\{w \in Gr^+(k,m) \,|\, \sigma_f(w) = 0\}. \tag{7}$$

We will now divide the rest of the proof into two parts:

(a) first prove that there exists a constant $\rho > 0$ such that:

$$\mathbb{E}\#\{w \in Gr(k,m) \,|\, \sigma_f(w) = 0\} = \rho \cdot |Gr(k,m)|, \tag{8}$$

where $|Gr(k,m)|$ denotes the volume of $Gr(k,m)$ with respect to the metric g;

(b) then we identify this expectation with the constant E_n from the claim.

(a) Proof of (8).
For $i = 1, \dots, k$ and $j = k+1, \dots, m$ consider the matrix $E_{ij} \in \mathfrak{so}(m)$ which consists of all zeros except for having a 1 in position (i, j) and a -1 in position (j, i). Then $e^{tE_{ij}} \in O(m)$ is the clockwise rotation matrix of an angle t in the plane, $\operatorname{span}\{e_i, e_j\}$, spanned by the i^{th} and j^{th} standard basis vectors. Let now $t = (t_{ij}) \in \mathbb{R}^{k(m-k)}$ and consider the map $R : \mathbb{R}^{k(m-k)} \times \operatorname{span}\{e_1, \dots, e_k\} \to \mathbb{R}^m$ defined by:

$$R(t, y) = \left(e^{\sum_{ij} t_{ij} E_{ij}}\right) \cdot y = R_t \cdot y.$$

The map $\psi : \mathbb{R}^{k(m-k)} \to Gr^+(k, m)$ defined by:

$$\psi(t) = R_t e_1 \wedge \cdots \wedge R_t e_k$$

is a local parametrization of $Gr^+(k, m)$ near $e_1 \wedge \cdots \wedge e_k = \psi(0)$. In fact the map ψ is the Riemannian exponential map centered at $e_1 \wedge \cdots \wedge e_k$ (see [24]). As a consequence, the map $\varphi : U \to \mathbb{R}^{k(m-k)}$ (the inverse of ψ) gives a coordinate chart on a neighborhood U of $e_1 \wedge \cdots \wedge e_k$. (This chart has the property of being an isometry at the origin, and this will be useful for computations.)

By construction, $R(\varphi(w), \cdot)^* f$ defines for every $w \in Gr^+(k, m)$ a polynomial in $\mathbb{R}[y_1, \dots, y_k]_{(d)}$. As a consequence, the map

$$q : \operatorname{sym}^d(\tau_{k,m}^*)|_U \to U \times \mathbb{R}[y_1, \dots, y_k]_{(d)}$$

defined by:

$$f \mapsto (w, R(\varphi(w), \cdot)^* f), \quad f \in \operatorname{sym}^d(\tau_{k,m}^*)|_w$$

gives a trivialization for $\operatorname{sym}^d(\tau_{k,m}^*)$ over U. To summarize, we have a commutative diagram (the trivialization of the bundle):

$$\begin{tikzcd}
\operatorname{sym}^d(\tau_{k,m}^*)|_U \arrow[rr, "q"] \arrow[dr, "\pi"'] & & U \times \mathbb{R}[y_1, \dots, y_k]_{(d)} \arrow[dl, "p_1"] \\
& U &
\end{tikzcd}$$

Observation: Since $Gr^+(k, m)$ is compact and connected, the map ψ (the Riemannian exponential map) is surjective and we can choose U such that $Gr^+(k, m) \backslash U$ has measure zero (in this way, integrating a continuous function over U with respect to the volume density of $Gr^+(k, m)$ gives the same result as integrating the continuous function over all of $Gr^+(k, m)$).

Let now $f \in \mathbb{R}[x_1, \dots, x_m]_{(d)}$ be a random Kostlan polynomial. Using f we can consider the random map:

$$\tilde{\sigma}_f : U \longrightarrow \mathbb{R}[y_1, \dots, y_k]_{(d)} \simeq \mathbb{R}^{\binom{k+d-1}{d}},$$

which is defined by $q(\sigma_f(w)) = (w, \tilde{\sigma}_f(w))$ (in other words, $\tilde{\sigma}_f$ is nothing but σ_f in the trivialization given by q).

Assume now that $k(m-k) = \binom{k+d-1}{d}$, so that $\operatorname{rank}\left(\operatorname{sym}^d(\tau_{k,m}^*)\right) = \dim Gr(k, m)$. In this case denoting by $N = \dim(\mathbb{R}[y_1, \dots, y_k]_{(d)}) = \dim(Gr(k, m))$, we have that

$$Z = \tilde{\sigma}_f \circ \psi : \psi^{-1}(U) \to \mathbb{R}^N$$

is a random Gaussian map between two open sets in $\mathbb{R}^N$. Then we can use the Kac-Rice formula from Theorem 1 to count the average number of zeros of Z (and consequently of $\tilde{\sigma}_f$):

$$\begin{aligned}\mathbb{E}\#\{\tilde{\sigma}_f=0\} &= \mathbb{E}\#\{Z=0\} \qquad \text{(since } \psi|_U \text{ is a diffeomorphism)}\\ &= \int_{\psi^{-1}(U)} \mathbb{E}\left\{|\det JZ(t)|\,\big|\,Z(t)=0\right\}\cdot p_{Z(t)}(0)dt=(*). \end{aligned} \tag{9}$$

Let us verify that the hypotheses from Theorem 1 are satisfied: (i) the fact that $Z(t)$ is Gaussian is clearly true; (ii) the fact that $Z(t)$ is almost surely of class C^1 is also clearly true (in fact it is C^∞); (iii) the fact that it has a nondegenerate distribution follows form the fact that, in the monomial basis, the random vector $Z(t)$ is just the list of coefficients of a Kostlan polynomial in k homogeneous variables (the restriction of a Kostlan polynomial to a subspace is still a Kostlan polynomial, see below), and the Kostlan distribution is nondegenerate; finally, for property (iv), we observe that [8, Theorem 2.1] guarantees that the section σ_f is transversal to the zero section of $\mathrm{sym}^d(\tau_{k,m}^*)$ for the generic f; because the set of polynomials f whose corresponding section σ_f *is not* transversal to the zero section of $\mathrm{sym}^d(\tau_{k,m}^*)$ is a semialgebraic set, [8, Theorem 2.1] implies that this semialgebraic set has codimension one, hence its complement has probability one (the Kostlan measure on the space of polynomials is absolutely continuous with respect to the Lebesgue measure).

Let us now endow $\psi^{-1}(U)$ with the pull-back metric ψ^*g. Then, following Remark 4, we can present (9) in a more geometric way using (6):

$$\begin{aligned}(*) &= \int_{\psi^{-1}(U)} \mathbb{E}\left\{|\det \hat{J}Z(t)|\,\big|\,Z(t)=0\right\} p_{Z(t)}(0)\cdot\left(\psi^*\omega_{Gr^+(k,m)}\right)(t)\\ &= \int_U \mathbb{E}\left\{|\det J(w)|\,\big|\,\tilde{\sigma}_f(w)=0\right\} p(0;w)\cdot\omega_{Gr^+(k,m)}(w),\end{aligned} \tag{10}$$

where we have denoted by $p(0;w)$ the density at zero of the random vector $\tilde{\sigma}_f(w)$ and

$$J(w)=(\nabla_i(\tilde{\sigma}_f)_j(w))$$

is the matrix of the derivatives at w of the components of $\tilde{\sigma}_f$ with respect to an orthonormal frame field.

Because the polynomial f is $O(m)$-invariant, the quantity

$$\rho(w)=\mathbb{E}\left\{|\det J(w)|\,\big|\,\tilde{\sigma}_f(w)=0\right\} p(0;w)$$

does not depend on $w\in U$ and $\rho(w)\equiv\rho$ is constant. Using also the above observation that we can restrict the integration to be over U since it has full measure in $Gr^+(k,m)$, we have

$$\mathbb{E}\#\{\sigma_f=0\}=\mathbb{E}\#\{\tilde{\sigma}_f=0\}\underset{(10)}{=}\int_U \rho(w)\cdot\omega_{Gr^+(k,m)}\underset{\rho(w)\equiv\rho}{=}|Gr^+(k,m)|\cdot\rho,$$

where $|Gr^+(k,m)|$ denotes the volume of $Gr^+(k,m)$. Together with (7), this proves (8).

(b) Computation of the constant ρ.
We proceed now with the computation of the constant ρ. Let $w_0=e_1\wedge\cdots\wedge e_k$, we compute $\rho=\rho(w_0)$. It is easy to compute the density at zero of $\tilde{\sigma}_f(w_0)$. In fact, in the trivialization q we have:

$$\tilde{\sigma}_f|_{w_0}=\sum_{|\alpha|=d}\xi_{\alpha_1,\ldots,\alpha_k,0,\ldots,0}\,y_1^{\alpha_1}\cdots y_k^{\alpha_k}.$$

Hence the coefficients of $\tilde{\sigma}_f|_{w_0}$ are independent random variables and are still Kostlan distributed (in fact, the restriction of a Kostlan polynomial to a subspace is still a Kostlan polynomial). As a consequence:

$$p(0;w_0)=\prod_{\alpha_1+\cdots+\alpha_k=d}\left(\frac{\alpha_1!\cdots\alpha_k!}{d!2\pi}\right)^{1/2}.$$

For the computation of $\mathbb{E}\left\{|\det J(w_0)|\,\middle|\,\tilde{\sigma}_f(w_0)=0\right\}$ we proceed as follows. For $i=1,\ldots,k$ and $j=k+1,\ldots,m$, consider the curve $\gamma_{ij}:(-\epsilon,\epsilon)\to Gr^+(k,m)$:

$$\gamma_{ij}(s)=e^{sE_{ij}}e_1\wedge\cdots\wedge e^{sE_{ij}}e_k.$$

For the curve γ_{ij}, note that:

$$\gamma_{ij}(0)=w_0\quad\text{and}\quad\dot{\gamma}_{ij}(0)=\psi_*(\partial_{t_{ij}}).$$

By construction, the frame field

$$\mathcal{F}(w)=\{\psi_*(\partial_{t_{ij}})(w)\,|\,i=1,\ldots,k,j=k+1,\ldots,m\}$$

at $w=w_0$ is an orthonormal basis for $T_{w_0}Gr^+(k,m)$. Even though the frame field $\mathcal{F}$ is not orthonormal in a full neighborhood of w_0, we can still use it for computing $J(w_0)$, since the value at a point of the derivative of a function with respect to a vector field only depends on the value of the field at the given point (and not on its extension).

By definition of the trivialization q we have:

$$\tilde{\sigma}_f(\gamma_{ij}(s))(y_1,\ldots,y_k)=f\left(e^{sE_{ij}}y\right),\quad y=(y_1,\ldots,y_k,0,\ldots,0).$$

Note that:

$$e^{sE_{ij}}y=(y_1,y_2,\ldots,y_{i-1},\underbrace{y_i\cos s}_{i\text{-th entry}},y_{i+1},\ldots,y_k,0,\ldots,0,\underbrace{y_i\sin s}_{j\text{-th entry}},0,\ldots,0).$$

In particular:

$$\begin{aligned}\tilde{\sigma}_f(\gamma_{ij}(s))(y)&=\sum_{|\alpha|=d}\xi_\alpha y_1^{\alpha_1}\cdots y_{i-1}^{\alpha_{i-1}}(y_i\cos s)^{\alpha_i}(y_i\sin s)^{\alpha_j}y_{i+1}^{\alpha_{i+1}}\cdots y_k^{\alpha_k}\\&=\sum_{|\alpha|=d}\xi_\alpha(\cos s)^{\alpha_i}(\sin s)^{\alpha_j}y_1^{\alpha_1}\cdots y_{i-1}^{\alpha_{i-1}}y_i^{\alpha_i+\alpha_j}y_{i+1}^{\alpha_{i+1}}\cdots y_k^{\alpha_k}.\end{aligned}$$

Taking the derivative of each coefficient with respect to s and evaluating at $s=0$ we get:

$$\begin{aligned}\frac{d}{ds}\left(\tilde{\sigma}_f(\gamma_{ij}(s))(y)\right)\big|_{s=0}&=\sum_{\alpha_1+\cdots+\alpha_k+1=d}\xi_\alpha y_1^{\alpha_1}\cdots y_{i-1}^{\alpha_{i-1}}y_i^{\alpha_i+1}y_{i+1}^{\alpha_{i+1}}\cdots y_k^{\alpha_k}.\\&=\sum_{|\beta|=d,\beta_i\geq1}q_\beta y_1^{\beta_1}\cdots y_k^{\beta_k}\end{aligned}$$

from which we deduce that the coefficient q_β is distributed as:

$$q_\beta=\xi_{\beta_1,\ldots,\beta_{i-1},\beta_i-1,\beta_{i+1},\ldots,\beta_k,0,\ldots,0,1,\ldots,0}\quad\text{(there is a 1 in position } j).$$

From this we immediately see that $|\det J(w_0)|$ and $\tilde{\sigma}_f(w_0)$ are independent random variables (because of the "1 in position j" above) and:

$$\mathbb{E}\left\{|\det J(w_0)|\,\middle|\,\tilde{\sigma}_f(w_0)=0\right\}=\mathbb{E}\left\{|\det J(w_0)|\right\}.$$

The matrix $J(w_0)$ is in general quite complicated, but when we specialize to the case $k=2$ it becomes simpler (this is due to the fact that there is a natural way to order the monomials of a one-variable polynomial). In fact, working in the above basis $\{\partial_{1,3}, \partial_{2,3}, \dots, \partial_{1,m}, \partial_{2,m}\}$ for $T_{w_0}G^+(2,m)$ and $\{y_1^d, y_1^{d-1}y_2, \dots, y_1y_2^{d-1}, y_2^d\}$ for $\mathbb{R}[y_1,y_2]_{(d)}$, we see that:

$$\partial_{1,j}\tilde{\sigma}_f = \sum_{|\alpha|=d,\alpha_j=1} \xi_\alpha y_1^{1+\alpha_1}y_2^{\alpha_2} \quad \text{and} \quad \partial_{2,j}\tilde{\sigma}_f = \sum_{|\alpha|=d,\alpha_j=1} \xi_\alpha y_1^{\alpha_1}y_2^{1+\alpha_2}.$$

For example, in the case $m = 4$ the matrix $J(w_0)$ is:

$$J(w_0) = \begin{bmatrix} \xi_{2010} & 0 & \xi_{2001} & 0 \\ \xi_{1110} & \xi_{2010} & \xi_{1101} & \xi_{2001} \\ \xi_{0210} & \xi_{1110} & \xi_{0201} & \xi_{1101} \\ 0 & \xi_{0210} & 0 & \xi_{0201} \end{bmatrix},$$

In the case $k = 2, m = n+1$, for the Grassmannian $Gr(2, n+1)$ of lines in $\mathbb{R}\mathrm{P}^n$, the matrix $J(w_0)$ has the same shape as $\hat{J}_n$; moreover collecting $\sqrt{2n-3}$ from each row of $J(w_0)$ we get a matrix that is distributed as $\hat{J}_n$:

$$\mathbb{E}\left\{|\det J(w_0)|\right\} = (2n-3)^{\frac{2n-2}{2}}\mathbb{E}|\det \hat{J}_n|.$$

Thus the average number, E_n, of real lines on a random Kostlan hypersurface of degree $2n-3$ in $\mathbb{R}\mathrm{P}^n$ is:

$$\begin{aligned} E_n &= |Gr(2,n+1)| \underbrace{p(0;w_0)\ (2n-3)^{\frac{2n-2}{2}}\ \mathbb{E}|\det \hat{J}_n|}_{\rho} \\ &= \frac{\pi^{n-\frac{1}{2}}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n+1}{2}\right)} \prod_{k=0}^{2n-3}\left(\frac{k!\cdot(2n-3-k)!}{(2n-3)!2\pi}\right)^{1/2} (2n-3)^{n-1}\ \mathbb{E}|\det \hat{J}_n| \\ &= \frac{(2n-3)^{n-1}}{\Gamma(n)} \prod_{k=0}^{2n-3}\binom{2n-3}{k}^{-1/2} \mathbb{E}|\det \hat{J}_n|, \end{aligned}$$

where for the volume of the Grassmannian we have used the formula (see Remark 5 below)

$$|Gr(2,n+1)| = \frac{\pi^{n-\frac{1}{2}}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}, \tag{11}$$

and in the subsequent step above we also used the duplication formula for the Γ function:

$$\Gamma(z)\cdot\Gamma(z+1/2) = 2^{1-2z}\sqrt{\pi}\Gamma(2z).$$

This concludes the proof. □

Remark 5 (Volume of Grassmannians). The formula (11) follows from $|Gr(k,m)| = \frac{|O(m)|}{|O(k)||O(m-k)|}$ and the formula for the volume of the Orthogonal group (see [20, Section 3.12]):

$$|O(k)| = \frac{2^k\pi^{\frac{k^2+k}{4}}}{\Gamma(k/2)\Gamma((k-1)/2)\cdots\Gamma(1/2)}.$$

Note that an analogous formula holds for the complex Grassmannian:

$$|Gr_{\mathbb{C}}(k,m)| = \frac{|U(m)|}{|U(k)||U(m-k)|} \quad \text{and} \quad |U(k)| = \frac{2^k\pi^{\frac{k^2+k}{2}}}{\prod_{i=1}^{k-1} i!}.$$

2.4. **Intrinsic signs and lower bound.** Generalizing Segre's observation $h-e=3$, Okonek and Teleman [28] have shown that to each real line ℓ on a generic real hypersurface H of degree $2n-3$ in $\mathbb{R}\mathrm{P}^n$ it is possible to canonically associate a sign $\epsilon(\ell)$ which satisfies the property that $\sum_{\ell\subset H}\epsilon(\ell)$ does not depend on H, and that this number equals $(2n-3)!!$. The crucial observation is that there exists a line bundle $L\to Gr(2,n+1)$ such that [28, Proposition 12]:

$$K=\det(\mathrm{sym}^{2n-3}(\tau_{2,n+1}^*))\otimes\det(TGr(2,n+1))=L\otimes L.$$

Hence K has a trivialization and $\mathrm{sym}^{2n-3}(\tau_{2,n+1}^*)$ and $TGr(2,n+1)$ are said to be *relatively oriented* (even if $\mathrm{sym}^{2n-3}(\tau_{2,n+1}^*)$ and the Grassmannian themselves might not be orientable). Relative orientation is all one needs to introduce an Euler number, which has the usual properties and is well defined up to a sign [28, Lemma 5]. The fact that $K=L\otimes L$ allows in this specific case to canonically choose the relative orientation and the sign $\epsilon(\ell)$ for $\ell\subset H$ is defined in an intrinsic way. Using this notation Okonek and Teleman [28] and Kharlamov and Finashin [14] have independently proved that:

$$R_n\doteq\left|\sum_{\ell\subset H}\epsilon(\ell)\right|=(2n-3)!!.$$

We now show that using the Kac-Rice formula, and the knowledge that the modulus of a signed count is invariant, one can recover the value of R_n.

Proposition 3. *Using the above notation, we have:*

$$R_n=(2n-3)!!.$$

Proof. Let us fix a trivialization of $\mathrm{sym}^{2n-3}(\tau_{2,n+1}^*)$ on an open dense set $U\subset Gr(2,n+1)$ as in the proof of Theorem 2. The lines on $H=\{f=0\}$ contained in U are the zeroes of:

$$\tilde\sigma_f:U\to\mathbb{R}^{2n-2}.$$

By [28, Lemma 5] we know that:

$$\sum_{\tilde\sigma_f(w)=0}\mathrm{sign}\det(J\tilde\sigma_f)(w)=\epsilon\cdot\sum_{\ell\subset H,\ell\in U}\epsilon(\ell)\tag{12}$$

and that the choice of the sign $\epsilon=\pm1$ is determined once the trivialization is fixed. On the other hand, for a random map $Z:\mathbb{R}^{2n-2}\to\mathbb{R}^{2n-2}$ (sufficiently smooth, e.g. satisfying the hypothesis of [1, Theorem 12.1.1]), we have:

$$\mathbb{E}\sum_{Z(w)=0}\mathrm{sign}\det(JZ)(t)=\int_{\mathbb{R}^{2n-2}}\mathbb{E}\{\det(JZ)(t)\,|\,Z(t)=0\}\cdot p_{Z(t)}(0)dt.$$

This follows from [27, Lemma 3.5], as the function $\mathrm{sign}(\det(\cdot)):\mathbb{R}^{(2n-2)\times(2n-2)}\to\mathbb{R}$ is admissible (in the terminology of [27]). Applying this to the case $Z=\tilde\sigma_f\circ\psi$ and arguing exactly as in the proof of Theorem 2 we get:

$$\mathbb{E}\sum_{\tilde\sigma_f(w)=0}\mathrm{sign}\det(J\tilde\sigma_f)(w)=\left(\frac{(2n-3)^{n-1}}{\Gamma(n)}\prod_{k=0}^{2n-3}\binom{2n-3}{k}^{-1/2}\right)\mathbb{E}\det\hat J_n\tag{13}$$

Let us first evaluate the factor $\mathbb{E}\det\hat J_n$. To simplify notations, let us denote by $x_{i,j}$ the random variable $v_j^{(i)}$ in the matrix $\hat J_n$ (we will also use this notation later in Remark 6). After taking expectation, by independence, the only monomials in the expansion of $\det\hat J_n$ that give a nonzero contribution are those with all squared variables (for example $x_{1,1}^2x_{2,3}^2\cdots x_{n-1,2n-3}^2$). These monomials all have the form:

$$x_{i_1,1}^2x_{i_2,3}^2\cdots x_{i_{n-1},2n-3}^2, \quad \{i_1,\ldots,i_{n-1}\}=\{1,\ldots,n-1\}.$$

There are $(n-1)!$ many such monomials and because every second subscript is shifted by two, in the expansion of $\det\hat{J}_n$ they all appear with the same sign. Moreover the product of all the variances of the variables in each of these monomials equal:

$$\prod_{j \text{ odd}}\binom{2n-4}{j-1}=\prod_{k=1}^{n-1}\binom{2n-4}{2k-2}.$$

From this we obtain:

$$\mathbb{E}\det\hat{J}_n=(n-1)!\prod_{k=1}^{n-1}\binom{2n-4}{2k-2}.$$

We look now at the term:

$$\begin{aligned}\prod_{k=0}^{2n-3}\binom{2n-3}{k}^{-1/2} &= \prod_{k=0}^{2n-3}\left(\frac{k!(2n-3-k)!}{(2n-3)!}\right)^{1/2}\\ &= \frac{1}{(2n-3)!^{n-1}}\left(\prod_{k=0}^{2n-3}k!\right)^{1/2}\left(\prod_{k=0}^{2n-3}(2n-3-k)!\right)^{1/2}\\ &= \frac{1}{(2n-3)!^{n-1}}\prod_{k=0}^{2n-3}k!.\end{aligned}$$

Collecting all this together, we obtain:

$$\begin{aligned}\left(\frac{(2n-3)^{n-1}}{\Gamma(n)}\prod_{k=0}^{2n-3}\binom{2n-3}{k}^{-1/2}\right)\mathbb{E}\det\hat{J}_n &= \frac{\left(\prod_{k=0}^{2n-3}k!\right)\left(\prod_{k=1}^{n-1}\binom{2n-4}{2k-2}\right)}{(2n-4)!^{n-1}}\\ &= \left(\prod_{k=0}^{2n-3}k!\right)\left(\prod_{k=1}^{n-1}\frac{1}{(2k-2)!(2n-2k-2)!}\right)\\ &= \left(\prod_{j=1}^{n-1}(2j-1)!\right)\left(\prod_{k=1}^{n-1}\frac{1}{(2n-2k-2)!}\right)\\ &= \left(\prod_{j=1}^{n-1}(2j-1)!\right)\left(\prod_{j=1}^{n-1}\frac{1}{(2j-2)!}\right)\\ &= \prod_{j=1}^{n-1}\frac{(2j-1)!}{(2j-2)!}\\ &= \prod_{j=1}^{n-1}(2j-1)=(2n-3)!!\end{aligned}$$

Combining the last equation with (13) and (12) concludes the proof. □

Let us now denote by ρ_n the number:

$$\rho_n = \frac{(2n-3)^{n-1}}{\Gamma(n)} \prod_{k=0}^{2n-3} \binom{2n-3}{k}^{-1/2}.$$

Then the last computation in the proof of Proposition 3 shows that $\rho_n \cdot |\mathbb{E} \det J_n| = (2n-3)!!$. As a corollary, since:

$$E_n = \rho_n \cdot \mathbb{E}|\det J_n| \geq \rho_n \cdot |\mathbb{E} \det J_n|,$$

we derive the following lower bound for E_n. Note however that the proof of $\rho_n \cdot |\mathbb{E} \det J_n| \geq (2n-3)!!$ does not require [28, Corollary 17], hence this lower bound does not depend on the knowledge that the signed count is invariant.

Corollary 4. *The following inequality holds:*

$$E_n \geq (2n-3)!!$$

3. The average number of real lines on a random cubic

Theorem 5. *The average number of real lines on a random cubic surface in* $\mathbb{R}\mathrm{P}^3$ *is* $6\sqrt{2}-3$.

Proof. Applying Theorem 2 for the special case $n=3$ we get:

$$E_3 = \frac{3}{2}\mathbb{E}|\det \hat{J}_3|. \tag{14}$$

where:

$$\hat{J}_3 = \begin{bmatrix} a & 0 & d & 0 \\ \sqrt{2}b & a & \sqrt{2}e & d \\ c & \sqrt{2}b & f & \sqrt{2}e \\ 0 & c & 0 & f \end{bmatrix},$$

with a, b, c, d, e, f independent standard normal random variables.

In order to compute $\mathbb{E}|\det \hat{J}_3|$, we first observe that:

$$\det \hat{J}_3 = (af-cd)^2 - 2(bf-ce)(ae-bd),$$

which will lead us to work in terms of the random variables:

$$\begin{aligned} x &= (bf-ce), \\ y &= (af-cd), \\ z &= (ae-bd), \end{aligned}$$

in order to compute the expectation of $|\det \hat{J}_3| = |2xz-y^2|$. We use the method of characteristic functions (Fourier analysis) in order to compute the joint density $\rho(x,y,z)$ of x, y, z.

By the Fourier inversion formula, we have:

$$\rho(x,y,z) = \frac{1}{(2\pi)^3} \lim_{R\to\infty} \int_{|(t_1,t_2,t_2)|<R} e^{-i(t_1x+t_2y+t_3z)}\ \hat{\rho}(t_1,t_2,t_3)\ dt_1dt_2dt_3, \tag{15}$$

where

$$\hat{\rho}(t_1,t_2,t_3) = \mathbb{E}e^{i(t_1(bf-ce)+t_2(ae-bd)+t_3(af-cd))}$$

is the characteristic function (Fourier transform) of ρ. In (15) the reason we are taking a limit as $R \to \infty$ of an integral over the ball of radius R (rather than directly integrating over $\mathbb{R}^3$ as in the more familiar Fourier inversion formula) is because it will turn out (see (16) below) that the characteristic function $\hat{\rho}$ is not in the space $L^1(\mathbb{R}^3)$ of integrable functions. However, $\hat{\rho}$ is in the space $L^2(\mathbb{R}^3)$, so one can justify (15) using the fact that the Fourier transform extends to $L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ as an isometry (once an appropriate normalizing constant is introduced) – see Plancherel's Theorem in [29]. In particular, the extended Fourier transform and its inverse are continuous. Since the sequence of truncations converges in $L^2(\mathbb{R}^3)$, this justifies (15). Of course the Fourier transform and its inverse can be further extended to include more exotic classes of functions (and distributions), but its extension to L^2 suits our current purposes.

We notice that the expression:

$$t_1(bf - ce) + t_2(ae - bd) + t_3(af - cd) = (a, b, c, d, e, f)^T Q(a, b, c, d, e, f)$$

is a symmetric quadratic form in the Gaussian vector:

$$(a, b, c, d, e, f),$$

where the matrix for the quadratic form is given by:

$$Q = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 & t_2 & t_3 \\ 0 & 0 & 0 & -t_2 & 0 & t_1 \\ 0 & 0 & 0 & -t_3 & -t_1 & 0 \\ 0 & -t_2 & -t_3 & 0 & 0 & 0 \\ t_2 & 0 & -t_1 & 0 & 0 & 0 \\ t_3 & t_1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Using [30, Thm. 2.1] while taking $t = 1$ (and treating t_1, t_2, t_3 as parameters), we have:

$$\mathbb{E} e^{i(t_1(bf-ce)+t_2(ae-bd)+t_3(af-cd))} = \frac{1}{\sqrt{\det(\mathbb{1} - 2iQ)}},$$

which can be simplified to give

$$\hat{\rho}(t_1, t_2, t_3) = \frac{1}{1 + t_1^2 + t_2^2 + t_3^2}. \tag{16}$$

We substitute this into (15) to compute the density ρ next. Note that in the first step of the computation below, we switch to spherical coordinates with the direction of the vector (x, y, z) treated as the zenith, ϕ being the polar angle measured from the zenith, and θ being the azimuthal angle.

$$\begin{aligned}
\rho(x,y,z) &= \frac{1}{(2\pi)^3} \lim_{R\to\infty} \int_{|(t_1,t_2,t_2)|<R} \frac{e^{-i(t_1x+t_2y+t_3z)}}{1+t_1^2+t_2^2+t_3^2}\, dt_1dt_2dt_3 \\
&= \frac{1}{(2\pi)^3} \lim_{R\to\infty} \int_0^{2\pi}\int_0^{\pi}\int_0^{R} \frac{e^{-i|(x,y,z)|r\cos\phi}}{1+r^2} r^2\sin\phi\, drd\phi d\theta \\
&= \frac{1}{(2\pi)^2} \lim_{R\to\infty} \int_0^{\pi}\int_0^{R} \frac{e^{-i|(x,y,z)|r\cos\phi}}{1+r^2} r^2\sin\phi\, drd\phi \\
&= \frac{1}{(2\pi)^2} \lim_{R\to\infty} \int_0^{R} \frac{r^2}{1+r^2} \frac{e^{-i|(x,y,z)|r\cos\phi}}{i|(x,y,z)|r}\Bigg|_{\phi=0}^{\phi=\pi} dr \\
&= \frac{1}{2\pi^2}\frac{1}{|(x,y,z)|} \lim_{R\to\infty} \Im \int_0^{R} \frac{r}{1+r^2} e^{i|(x,y,z)|r}\, dr \\
&= \frac{1}{4\pi^2}\frac{1}{|(x,y,z)|} \lim_{R\to\infty} \Im \int_{-R}^{R} \frac{r}{1+r^2} e^{i|(x,y,z)|r}\, dr \\
&= \frac{1}{4\pi^2}\frac{1}{|(x,y,z)|} \Im \oint_{C_R} \frac{r}{1+r^2} e^{i|(x,y,z)|r}\, dr \\
&= \frac{1}{4\pi}\frac{e^{-|(x,y,z)|}}{|(x,y,z)|},
\end{aligned}$$

where $\Im$ denotes the imaginary part, and we used residues [2, Ch. 5] to evaluate the final integral. Let us give some details explaining this step. First one must modify the path of integration to get a closed contour C_R. We do this by including a semicircle of radius R (the upper half of a circle about the origin). We can freely add this semicircle since it does not affect the limit; indeed, the contribution along the semicircle can be estimated as, say, $O(1/\sqrt{R})$ by separately considering the part of the semicircle that is within distance $\sqrt{R}$ to the real axis, where we use that the integrand is $O(1/R)$, and the remaining part of the semicircle, where we can use the exponential decay of the integrand with respect to the imaginary direction. Once we include the semicircle (thus arriving at the contour integral over C_R in the penultimate line among the displayed equations above), we drop the limit notation because the resulting contour integral is independent of R as long as R is large enough that C_R surrounds $r = i$. We can then compute

$$\oint_{C_R} \frac{r}{1+r^2} e^{i|(x,y,z)|r}\, dr = 2\pi i \cdot \left(\frac{1}{2} e^{-|(x,y,z)|}\right),$$

since the integrand has one simple pole inside this contour at $r = i$, and the residue is easily computed by factoring out $\frac{1}{r-i}$ from the integrand and evaluating the remaining factor at $r = i$.

Having computed $\rho(x,y,z)$, the expectation can now be expressed as:

$$\begin{aligned}
\mathbb{E}|\det \hat{J}_3| &= \frac{1}{4\pi}\int_{\mathbb{R}^3} |2xz-y^2| \frac{e^{-|(x,y,z)|}}{|(x,y,z)|}\, dxdydz \\
&= \frac{1}{8\pi}\int_{\mathbb{R}^3} |a_1a_3-a_2^2| \frac{e^{-\sqrt{\frac{a_1^2}{2}+a_2+\frac{a_3^2}{2}}}}{\sqrt{\frac{a_1^2}{2}+a_2+\frac{a_3^2}{2}}}\, da_1da_2da_3
\end{aligned}$$

where we have made the change of variables $a_1 = \sqrt{2}x, a_2 = y, a_3 = \sqrt{2}z$. Let us view a_1, a_2, a_3 as the entries of a symmetric matrix:

$$A = \begin{bmatrix} a_1 & a_2 \\ a_2 & a_3 \end{bmatrix}.$$

Consider the coordinates given by the eigenvalues λ_1, λ_2 of A along with the angle $\alpha \in [0, \pi/2)$ associated with the orthogonal transformation that diagonalizes A:

$$M = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

The Jacobian determinant for changing coordinates to $(\lambda_1, \lambda_2, \alpha)$ is $|\lambda_1 - \lambda_2|$. We recognize $\frac{a_1^2}{2} + a_2 + \frac{a_3^2}{2} = \frac{\lambda_1^2+\lambda_2^2}{2}$ as half the Frobenius norm of A, and $a_1a_3 - a_2^2 = \lambda_1\lambda_2$ as the determinant of A. With respect to these coordinates, we have:

$$\begin{aligned}
\mathbb{E}|\det \hat{J}_3| &= \frac{\sqrt{2}}{8\pi}\int_0^{\pi/2}\int_{\mathbb{R}^2} |\lambda_1\lambda_2| \frac{e^{-\sqrt{\frac{\lambda_1^2+\lambda_2^2}{2}}}}{\sqrt{\lambda_1^2+\lambda_2^2}} |\lambda_1 - \lambda_2| \; d\lambda_1 d\lambda_2 \; d\alpha \\
&= \frac{\sqrt{2}}{16}\int_{\mathbb{R}^2} |\lambda_1\lambda_2| \frac{e^{-\sqrt{\frac{\lambda_1^2+\lambda_2^2}{2}}}}{\sqrt{\lambda_1^2+\lambda_2^2}} |\lambda_1 - \lambda_2| \; d\lambda_1 d\lambda_2 \\
&= \frac{\sqrt{2}}{16}\int_0^{2\pi}\int_0^{\infty} r^3 |\cos\theta\sin\theta(\cos\theta - \sin\theta)| e^{-r/\sqrt{2}} \; dr \; d\theta, \\
&= \frac{3\sqrt{2}}{2}\int_0^{2\pi} |\cos\theta\sin\theta(\cos\theta - \sin\theta)| \; d\theta, \\
&= \frac{3\sqrt{2}}{2}\left(\frac{8-2\sqrt{2}}{3}\right) = 4\sqrt{2} - 2
\end{aligned}$$

where we have utilized polar coordinates $\lambda_1 = r\cos\theta, \lambda_2 = r\sin\theta$. Substituting the obtained number into (14) gives the desired result $E_3 = 6\sqrt{2} - 3$.

□

4. The complex case

4.1. **Complex Kostlan polynomials.** We consider now the space $\mathbb{C}[x_0, x_1, \dots, x_n]_{(d)}$ of homogeneous polynomials of degree d in $n+1$ variables. A complex Kostlan polynomial is obtained by replacing real Gaussian variables with complex Gaussian variables in the definition from section 2.1. Specifically we take:

$$f(x) = \sum_{|\alpha|=d} \xi_\alpha x^\alpha, \quad \xi_\alpha \sim N_{\mathbb{C}}\left(0, \binom{d}{\alpha}\right),$$

where as before the ξ_α are independent. The resulting probability distribution on the space $\mathbb{C}[x_0, x_1, \dots, x_n]_{(d)}$ is invariant by the action of $U(n+1)$ by change of variables (see [22, 5]).

4.2. **A general construction.** We will need the following elementary Lemma.

Lemma 6. *Consider the $2m \times 2m$ real matrix $\tilde{A}$ defined by:*

$$\tilde{A} = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix} \quad \text{where} \quad A_{ij} = \begin{bmatrix} a_{ij} & b_{ij} \\ -b_{ij} & a_{ij} \end{bmatrix}$$

and the $m \times m$ complex matrix $A^{\mathbb{C}}$ defined by:

$$A^{\mathbb{C}} = \begin{bmatrix} a_{11} + ib_{11} & \cdots & a_{1m} + ib_{1m} \\ \vdots & & \vdots \\ a_{m1} + ib_{m1} & \cdots & a_{mm} + ib_{mm} \end{bmatrix}.$$

Then:

$$\det(\tilde{A}) = \det(A^{\mathbb{C}})\, \overline{\det(A^{\mathbb{C}})} \quad \text{and} \quad \det(\tilde{A}) \geq 0.$$

Proof. Let $P \in GL(2m, \mathbb{R})$ be the permutation matrix corresponding to the permutation:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & m-1 & m & m+1 & m+2 & \cdots & 2m-1 & 2m \\ 1 & 3 & \cdots & 2m-3 & 2m-1 & 2 & 4 & \cdots & 2m-2 & 2m \end{pmatrix}.$$

Then the matrix $P^{-1}\tilde{A}P$ is of the form:

$$\begin{bmatrix} M & N \\ -N & M \end{bmatrix} \quad \text{where} \quad M = (a_{jl}) \quad \text{and} \quad N = (b_{jl}).$$

Note that $A^{\mathbb{C}} = M + iN$ and that:

$$\begin{bmatrix} \mathbb{1} & 0 \\ -i\mathbb{1} & \mathbb{1} \end{bmatrix} \cdot \begin{bmatrix} M & N \\ -N & M \end{bmatrix} \cdot \begin{bmatrix} \mathbb{1} & 0 \\ i\mathbb{1} & \mathbb{1} \end{bmatrix} = \begin{bmatrix} M + iN & N \\ 0 & M - iN \end{bmatrix},$$

which implies $\det(P^{-1}\tilde{A}P) = \det(M + iN)\det(M - iN)$. Consequently:

$$\begin{aligned} \det(\tilde{A}) &= \det(P^{-1}\tilde{A}P) \\ &= \det(M + iN)\det(M - iN) \\ &= \det(A^{\mathbb{C}})\, \overline{\det(A^{\mathbb{C}})}. \end{aligned}$$

□

Consider now a random vector $w = (w_1, \ldots, w_{2n-3})$ in $\mathbb{C}^{2n-3}$ whose entries are independent Gaussian variables distributed as:

$$w_j \sim N_{\mathbb{C}}\left(0, \binom{2n-4}{j-1}\right) \qquad j = 1, \ldots, 2n-3.$$

In other words:

$$w_j \sim \sqrt{\frac{1}{2}\binom{2n-4}{j-1}}(\xi_1 + i\xi_2)$$

where ξ_1, ξ_2 are two standard independent Gaussians.

Let $w^{(1)}, \ldots, w^{(n-1)}$ be independent random vectors all distributed as w. We define the random $(2n-2) \times (2n-2)$ matrix $\hat{J}_n^{\mathbb{C}}$ as:

$$\hat{J}_n^{\mathbb{C}} = \begin{bmatrix} w_1^{(1)} & 0 & \ldots & w_1^{(n-1)} & 0 \\ w_2^{(1)} & w_1^{(1)} & & w_2^{(n-1)} & w_1^{(n-1)} \\ \vdots & w_2^{(1)} & & \vdots & w_2^{(n-1)} \\ w_{2n-3}^{(1)} & \vdots & & w_{2n-3}^{(n-1)} & \vdots \\ 0 & w_{2n-3}^{(1)} & \ldots & 0 & w_{2n-3}^{(n-1)} \end{bmatrix}$$

Theorem 7. *The number C_n of lines on a generic hypersurface of degree $2n-3$ in $\mathbb{C}\mathrm{P}^n$ is:*

$$C_n = \left(\frac{(2n-3)^{2n-2}}{\Gamma(n)\Gamma(n+1)} \prod_{k=0}^{2n-3} \binom{2n-3}{k}^{-1} \right) \mathbb{E} |\det \hat{J}^{\mathbb{C}}|^2$$

Proof. (sketch) The proof proceeds similarly to the proof of Theorem 2; again we use a general approach and then specialize to the case of the Grassmannian of lines in $\mathbb{C}\mathrm{P}^n$.

We view the complex Grassmannian $Gr_{\mathbb{C}}(k,m)$ as a real manifold of dimension $2k(m-k)$ and the space $\mathbb{C}[z_1, \ldots, z_m]_{(d)}$ of homogeneous polynomials of degree d as a real vector space of dimension $2\binom{m+d-1}{d}$:

$$f(z_1, \ldots, z_m) = \sum_{|\alpha|=d} (a_\alpha + i b_\alpha) z_1^{\alpha_1} \cdots z_m^{\alpha_m}, \quad a_\alpha, b_\alpha \in \mathbb{R}.$$

We put a Riemannian structure on $Gr_{\mathbb{C}}(k,m)$ by declaring the quotient map $U(m) \to Gr_{\mathbb{C}}(k,m)$ to be a Riemannian submersion.

Consider the bundle $\mathrm{sym}^d(\tau_{k,m}^*)$ on $Gr_{\mathbb{C}}(k,m)$; a polynomial $f \in \mathbb{C}[z_1, \ldots, z_m]_{(d)}$ defines a section σ_f of this bundle and, in the case $k(m-k) = \binom{k+d-1}{d}$, the number of zeros of σ_f coincides with the number of k-planes on $\{f = 0\}$. We build a random section of $\mathrm{sym}^d(\tau_{k,m}^*)$ by taking f to be a *complex* Kostlan polynomial:

$$f(z) = \sum_{|\alpha|=d} \xi_\alpha z_1^{\alpha_1} \cdots z_m^{\alpha_m}$$

where the ξ_α are independent and distributed as:

$$\xi_\alpha = \sqrt{\frac{1}{2}\binom{d}{\alpha}} (\xi_1 + i \xi_2) \tag{17}$$

and ξ_1, ξ_2 are standard, independent Gaussians. In this way the resulting probability distribution on the space of polynomials is invariant by the action of the unitary group $U(m)$.

Note that in the case $k = 2, m = n+1$, with probability one the number of zeros of σ_f equals the number of k-planes on a generic hypersurface of degree d in $\mathbb{C}\mathrm{P}^n$.

To trivialize the bundle we proceed as in the proof of Theorem 2, using now complex variables. The invariance of f under the action of the unitary group allows to reduce the computation for the Kac-Rice density at a point (which we again assume is $w_0 = \mathrm{span}\{e_1, \ldots, e_k\}$). We obtain:

$$\mathbb{E}\#\{w \in Gr_{\mathbb{C}}(k,m) \,|\, \sigma_f(w) = 0\} = |Gr_{\mathbb{C}}(k,m)| \rho_{\mathbb{C}}$$

where now:

$$\rho_{\mathbb{C}} = \mathbb{E} \left\{ |\det \tilde{J}(w_0)| \,\big|\, \tilde{\sigma}_f(w_0) = 0 \right\} p^{\mathbb{C}}(0; w_0),$$

with $p^{\mathbb{C}}(0; w_0)$ the density at zero of the vector of the real coefficients of $\tilde{\sigma}_f|_{w_0} \in \mathbb{C}[z_1, \dots, z_k]_{(d)}$, and the matrix

$$\tilde{J}(w_0) = (\nabla_i (\tilde{\sigma}_f)_j(w_0))$$

the $2k(m-k) \times 2k(m-k)$ matrix of the derivatives of the coordinates of $\tilde{\sigma}_f$ with respect to an orthonormal frame field at w_0.

Note that $\tilde{\sigma}_f(w_0) = f|_{w_0}$ is the random polynomial:

$$\tilde{\sigma}_f(w_0)(z_1, \dots, z_k) = \sum_{|\alpha|=d} \xi_\alpha z_1^{\alpha_1} \cdots z_k^{\alpha_k}$$

from which we immediately see that in the case $k = 2, m = n+1$, for the Grassmannian $Gr_{\mathbb{C}}(2, n+1)$ of lines in $\mathbb{C}\mathrm{P}^n$ we have:

$$p^{\mathbb{C}}(0, w_0) = \prod_{k=0}^{2n-3} \frac{1}{\pi} \binom{2n-3}{k}^{-1}.$$

For the computation of $\tilde{J}(w_0)$ we use the orthonormal basis of $T_{w_0} Gr_{\mathbb{C}}(k, m)$ given by derivatives at zero of the curves:

$$\gamma_{lj}^1, \gamma_{lj}^2 : (-\epsilon, \epsilon) \to Gr_{\mathbb{C}}(2, m)$$

defined for $l = 1, \dots, k$ and $j = k+1, \dots, m$ by:

$$\gamma_{lj}^1(s) = e^{sE_{lj}} e_1 \wedge \cdots \wedge e^{sE_{lj}} e_k \quad \text{and} \quad \gamma_{lj}^2(s) = e^{isE_{lj}} e_1 \wedge \cdots \wedge e^{isE_{lj}} e_k.$$

It is immediate to verify that:

$$\begin{aligned} \frac{d}{ds}\left(\tilde{\sigma}_f(\gamma_{kj}^1(s))(z)\right) &= \sum_{\alpha_1 + \cdots + \alpha_k + 1 = d} \xi_\alpha z_1^{\alpha_1} \cdots z_i^{\alpha_i+1} \cdots z_k^{\alpha_k} \\ &= \sum_{|\beta|=d, \beta_i \geq 1} q_\beta^{(1)} z_1^{\beta_1} \cdots z_k^{\beta_k} \end{aligned}$$

and that:

$$\begin{aligned} \frac{d}{ds}\left(\tilde{\sigma}_f(\gamma_{kj}^2(s))(z)\right) &= \sum_{\alpha_1 + \cdots + \alpha_k + 1 = d} i\xi_\alpha z_1^{\alpha_1} \cdots z_i^{\alpha_i+1} \cdots z_k^{\alpha_k} \\ &= \sum_{|\beta|=d, \beta_1 \geq 1} q_\beta^{(2)} z_1^{\beta_1} \cdots z_k^{\beta_k} \end{aligned}$$

From this we see that the coefficients $q_\beta^{(1)}$ and $q_\beta^{(2)}$ are distributed as:

$$q_\beta^{(1)} = \xi_{\beta_1, \dots, \beta_{i-1}, \beta_i - 1, \beta_{i+1}, \dots, \beta_k, 0, \dots, 0, 1, \dots, 0} \quad \text{(there is a 1 in position } j) \tag{18}$$

and:

$$q_\beta^{(2)} = i \cdot \xi_{\beta_1, \dots, \beta_{i-1}, \beta_i - 1, \beta_{i+1}, \dots, \beta_k, 0, \dots, 0, 1, \dots, 0} \quad \text{(there is a 1 in position } j). \tag{19}$$

In particular, we deduce from (18) and (19) that $\det \tilde{J}(w_0)$ and $\tilde{\sigma}_f(w_0)$ are independent, and consequently:

$$\mathbb{E}\left\{ |\det \tilde{J}(w_0)| \,\middle|\, \tilde{\sigma}_f(w_0) = 0 \right\} = \mathbb{E}|\det \tilde{J}(w_0)|.$$

Recalling the definition (17), and specializing to the case $k = 2, m = n+1$ of the Grassmannian of lines, we see again from (18) and (19) that the matrix $\tilde{J}(w_0)$ has the same shape as the matrix

$\tilde{A}$ from Lemma 6. Collecting a factor of $\sqrt{d} = \sqrt{2n-3}$ from each row and using Lemma 6 we obtain:

$$\mathbb{E}|\det \tilde{J}(w_0)| = (2n-3)^{2n-2}\mathbb{E}\left(\det \hat{J}_n^{\mathbb{C}}\ \overline{\det \hat{J}_n^{\mathbb{C}}}\right).$$

Putting all the pieces together, and using the formula $|Gr_{\mathbb{C}}(2,n+1)| = \frac{\pi^{2n-2}}{\Gamma(n)\Gamma(n+1)}$ (see Remark 5), we get:

$$\begin{aligned} C_n &= |Gr_{\mathbb{C}}(2,n+1)|\ p^{\mathbb{C}}(0,w_0)\ \mathbb{E}|\det \tilde{J}(w_0)| \\ &= \frac{\pi^{2n-2}}{\Gamma(n)\Gamma(n+1)} \prod_{k=0}^{2n-3} \frac{1}{\pi}\binom{2n-3}{k}^{-1} (2n-3)^{2n-2}\mathbb{E}\left(\det \hat{J}_n^{\mathbb{C}}\ \overline{\det \hat{J}_n^{\mathbb{C}}}\right) \\ &= \left(\frac{(2n-3)^{2n-2}}{\Gamma(n)\Gamma(n+1)} \prod_{k=0}^{2n-3}\binom{2n-3}{k}^{-1}\right)\mathbb{E}|\det \hat{J}_n^{\mathbb{C}}|^2. \end{aligned}$$

□

Remark 6 (Real versus complex Gaussians). Consider the matrix:

$$A_n(x) = \begin{bmatrix} x_{1,1} & 0 & \cdots & x_{n-1,1} & 0 \\ \vdots & x_{1,1} & & \vdots & x_{n-1,1} \\ \binom{2n-4}{j-1}^{1/2} x_{1,j} & \vdots & & \binom{2n-4}{j-1}^{1/2} x_{n-1,j} & \vdots \\ \vdots & \binom{2n-4}{j-1}^{1/2} x_{1,j} & & \vdots & \binom{2n-4}{j-1}^{1/2} x_{n-1,j} \\ x_{1,2n-3} & \vdots & & x_{n-1,2n-3} & \vdots \\ 0 & x_{1,2n-3} & \cdots & 0 & x_{n-1,2n-3} \end{bmatrix}.$$

The determinant $P_n(x)$ of $A_n(x)$ is a homogeneous polynomial of degree $D = 2n-2$ in $N = (n-1)(2n-3)$ many variables and by construction we have:

$$\mathbb{E}|\det \hat{J}_n| = \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} |P_n(x)|\ e^{-\frac{1}{2}\|x\|^2} dx$$

and:

$$\mathbb{E}|\det \hat{J}_n^{\mathbb{C}}|^2 = \frac{1}{\pi^N} \int_{\mathbb{C}^N} P_n(z)\overline{P_n(z)}\ e^{-\|z\|^2} dz. \tag{20}$$

Recall that, given a homogeneous polynomial $P(z) = \sum_{|\alpha|=D} P_\alpha z_1^{\alpha_1}\cdots z_N^{\alpha_N}$ of degree D in N variables, we have denoted by $\|P\|_B$ its Bombieri norm:

$$\|P\|_B = \left(\sum_{|\alpha|=D} |P_\alpha|^2 \frac{\alpha_1!\cdots\alpha_N!}{D!}\right)^{\frac{1}{2}}. \tag{21}$$

Then, it is possible to rewrite (20) as:

$$\mathbb{E}|\det \hat{J}_n^{\mathbb{C}}|^2 = (2n-2)!\ \|P_n\|_B^2. \tag{22}$$

4.3. The 27 lines on a complex cubic.

Corollary 8. *There are* 27 *lines on a generic cubic in* $\mathbb{C}\mathrm{P}^3$.

Proof. This is the case $n = 3$ in the previous theorem, which gives:

$$C_3 = \frac{3}{4}\mathbb{E}|\det \hat{J}_3^{\mathbb{C}}|^2$$

For the computation of $\mathbb{E}|\det \hat{J}_3^{\mathbb{C}}|^2$ we use (22). We have the following expression for $P_3(x) = \det A_3(x)$:

$$x_{13}^2x_{21}^2 - 2x_{12}x_{13}x_{21}x_{22} + 2x_{11}x_{13}x_{22}^2 + 2x_{12}^2x_{21}x_{23} - 2x_{11}x_{13}x_{21}x_{23} - 2x_{11}x_{12}x_{22}x_{23} + x_{11}^2x_{23}^2.$$

Recalling (21) we can immediately compute the Bombieri norm of P_3:

$$\|P_3\|_B^2 = \frac{36}{4!}.$$

From this we get $\mathbb{E}|\det \hat{J}_3^{\mathbb{C}}|^2 = 4!\|P_3\|_B^2 = 36$, and consequently $C_3 = 27$. □

5. Asymptotics

The main purpose of this section is to prove Theorem 12, which gives the asymptotic (3) of E_n in the logarithmic scale, as discussed above in Section 1.3 (the "square root law"). This will follow from a combination of the lower bound given in Corollary 4 and the upper bound that we will prove in Proposition 11.

Remark 7. Our proof of the upper bound uses the Leibniz expansion of the determinant and requires some delicate combinatorics. Before stepping into this, let us briefly recall how the Leibniz expansion easily leads to an estimate for the absolute moment of a random determinant in the simpler setting of a random real Ginibre $n \times n$ matrix, or more generally a random matrix M_n with i.i.d. entries m_{ij} each of mean zero and variance one:

$$\mathbb{E}|\det M_n| \le \sqrt{\mathbb{E}(\det M_n)^2} \le \sqrt{\mathbb{E}\left(\sum_\sigma P_\sigma\right)^2} = \sqrt{\sum_{\sigma,\tau}\mathbb{E}(P_\sigma P_\tau)} = \sqrt{n!}, \tag{23}$$

where σ and τ range over permutations of $\{1, 2, ..., n\}$, $P_\sigma = \mathrm{sgn}(\sigma)\prod_{i=1}^{n} m_{i\sigma(i)}$, and the last equality follows from independence of m_{ij} and the fact that they are each of mean zero and variance one (so that only the diagonal terms in the expansion of $\left(\sum_\sigma P_\sigma\right)^2$ survive after taking expectation).

This simple estimate (23) was observed by Turán [42]. It turns out to be off by only a factor of $\sqrt{n}$ from the true asymptotic behavior of $\mathbb{E}|\det M_n|$ [18] (in particular, it is sharp at the logarithmic scale).

The combinatorial difficulties in adapting this method to our setting are due to the repetition of variables in the matrix $A_n(x)$. A similar complication is faced in studying the moments of the determinant of a random Wigner matrix (where there are pairs of repeating variables in off-diagonal entries since the matrix is symmetric); the Leibniz expansion approach has been adapted to that setting in [41].

5.1. **The upper bound.** The strategy of our proof can be described as follows. In order to simplify notations, let us absorb the variances in the variables of the matrix $A_n(x)$ and consider it as the matrix $B_n(u)$ with entries the random variables $u_{i,j} = \binom{2n-4}{j-1} x_{i,j}$, so that $\det A_n(x) = \det B_n(u)$.

$$B_n(u) = \left[\begin{array}{ccccc} u_{1,1} & 0 & \cdots & u_{n-1,1} & 0 \\ \vdots & u_{1,1} & & \vdots & u_{n-1,1} \\ u_{1,j} & \vdots & & u_{n-1,j} & \vdots \\ \vdots & u_{1,j} & & \vdots & u_{n-1,j} \\ u_{1,2n-3} & \vdots & & u_{n-1,2n-3} & \vdots \\ 0 & u_{1,2n-3} & \cdots & 0 & u_{n-1,2n-3} \end{array} \right]. \tag{24}$$

We will use both the double-index notation $u_{i,j}$ as well as single-index notation u_k, $k = 1, 2, .., N$ for the $N = (n-1)(2n-3)$ many variables appearing in $B_n(u)$.

Given a permutation $\sigma \in S_{2n-2}$ we consider the product:

$$\prod_{i=1}^{2n-2} B_n(u)_{\sigma(i),i} = u_1^{\alpha_1} \cdots u_N^{\alpha_N}.$$

We will call $u_1^{\alpha_1} \cdots u_N^{\alpha_N}$ the monomial generated by the permutation σ. We will denote by I_1 the set of all the multi-indices of all possible monomials generated by permutations in S_{2n-2}. In this way we can write:

$$Q_n(u) = \det B_n(u) = \sum_{\alpha \in I_1} Q_\alpha u_1^{\alpha_1} \cdots u_N^{\alpha_N}.$$

We first prove that for each permutation $\pi \in S_{2n-2}$, the monomial α_π generated by π occurs with a non-zero integral coefficient in the polynomial $Q_n(u)$. In other words there are no cancellations occuring in the Leibniz expansion of the determinant of $B_n(u)$: such cancellations are a priori possible since the same monomial can be generated by several permutations, which is evident (24) from the structure of the matrix. The proof of the above statement follows immediately from Lemma 9 proved below. We then prove (Lemma 10) that in the expansion of the polynomial $Q_n(u)^2$, each "cross-term" $Q_\alpha Q_\beta u^\alpha u^\beta$ that appears with a non-zero coefficient can be "charged" to some square term $Q_\gamma^2 u^{2\gamma}$. Of course, many cross-terms might be charged to the same square-term, but the number of different pairs (α, β) such that $\alpha + \beta = 2\gamma$, is bounded by some number at most exponential in n. Together, these two lemmas reduce the problem of bounding $\mathbb{E}(Q_n(u))^2$ (up to a loss of an exponential factor) to the problem of bounding

$$\mathbb{E} \sum_\gamma Q_\gamma^2 u_\gamma^2,$$

and the latter can be bounded using linearity of expectation in terms of the Bombieri norm of the polynomial $P_n(x) = \det A_n(x)$ (Proposition 11).

We now prove the necessary preliminary results required to carry through the argument sketched above.

Lemma 9. *For each $\alpha \in I_1$, $Q_\alpha \neq 0$ and $|Q_\alpha| > 1$.*

Proof of Lemma. For a given multi-index α suppose that $\sigma = \sigma(1)\sigma(2)\cdots\sigma(2n-2)$ and $\tau = \tau(1)\tau(2)\cdots\tau(2n-2)$ are two permutations (each given in one-line notation) that are associated with the monomial u^α in the Leibniz expansion of the determinant of $B_n(u)$. In other words, the corresponding terms in the Leibniz expansion are $\operatorname{sgn}(\sigma)u^\alpha$ and $\operatorname{sgn}(\tau)u^\alpha$, respectively. It suffices to show, for arbitrary such σ and τ, that we have $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau)$, or equivalently that $\theta = \tau^{-1}\sigma$ is even.

Note that in the matrix $B_n(u)$ each variable appears exactly twice and in positions that are separated by exactly one increment in the row and the column values. The assumption that σ and τ generate the same monomial u^α then implies the following claim.

Claim 1. *Suppose $\sigma(j) \neq \tau(j)$. Then if j is odd, $\sigma(j+1) = \tau(j) + 1$ and $\tau(j+1) = \sigma(j) + 1$, and if j is even, $\sigma(j-1) = \tau(j) - 1$ and $\tau(j-1) = \sigma(j) - 1$.*

We will also need the next claim.

Claim 2. *Suppose $\tau(j) = \sigma(k)$ for $j \neq k$. Then the parities of k and j are the same; moreover if j is odd $\sigma(j+1) = \tau(j) + 1$ and $\tau(k+1) = \sigma(k) + 1$, if j is even $\sigma(j-1) = \tau(j) - 1$ and $\tau(k-1) = \sigma(k) - 1$.*

We prove Claim 2 in the case that j is odd; the case j is even is similar and is omitted. The variable in the matrix $B_n(u)$ in position $(\tau(j), j)$ must be selected by σ as well. Since $\sigma(j) \neq \tau(j)$ and j is odd, the only option is that $\sigma(j+1) = \tau(j)+1$. We will see that the case k is even leads to a contradiction. Since $\tau(k) \neq \sigma(k)$ and k is even, in order for the variable in position $(\sigma(k), k)$ to be selected by τ we must have $\tau(k-1) = \sigma(k)-1$ (by Claim 1). This implies that $\sigma(k-1) \neq \tau(k-1)$, so that there is some $\ell \neq k-1$ such that $\sigma(\ell) = \tau(k-1)$. In order for the variable in position $(\sigma(\ell), \ell)$ to be selected by τ we must have $\tau(\ell - 1) = \sigma(\ell) - 1 = \tau(k-1) - 1$ (note that the alternative option $\tau(\ell + 1) = \sigma(\ell) + 1$ is prevented since $\sigma(\ell) + 1 = \tau(k-1) + 1 = \sigma(k) = \tau(j)$). Iterating this argument $\tau(k-1) - 2$ more steps, and recalling equation (24), we reach the first row of the matrix $B_n(u)$ where we are forced to select an unpaired variable contradicting the assumption that σ and τ generate the same monomial. This shows that k must be odd as well and, by the same reasoning as above, $\tau(k+1) = \sigma(k) + 1$ as stated in the claim.

Let us now go back to the permutation $\theta = \tau^{-1}\sigma$. Claim 2 implies that θ preserves parity, so it can be written as the product of two permutations $\theta = \theta_{\text{even}} \cdot \theta_{\text{odd}}$, where θ_{even} (respectively, θ_{odd}) is in the symmetric group $S_{n-1,\text{even}}$ (respectively, $S_{n-1,\text{odd}}$) on the set of *even* (respectively, *odd*) numbers belonging to $\{1, \dots, 2n-2\}$. We identify $S_{n-1,\text{even}}$ (respectively, $S_{n-1,\text{odd}}$) with the subgroup of S_{2n-2} of permutations which fixes each odd (respectively, even) number in $\{1, \dots, 2n-2\}$. Claim 2 can now be rewritten as:

$$\theta_{\text{odd}}(2k-1) = 2j - 1 \iff \theta_{\text{even}}(2j) = 2k. \tag{25}$$

Note that the bijection:

$$\{2, 4, \dots, 2n-2\} \to \{1, 3, \dots, 2n-3\}, \quad 2k \mapsto 2k - 1,$$

induces an isomorphism $\psi : S_{n-1,\text{odd}} \to S_{n-1,\text{even}}$. Equation (25) shows that $\psi(\theta_{\text{odd}}) = \theta_{\text{even}}^{-1}$. In particular, since the sign of θ_{even} and $\theta_{\text{even}}^{-1}$ are the same it follows that θ is even. □

We will also need the following lemma.

Lemma 10. *Let I_2 be the set of all multi-indices $\gamma = (\gamma_1, \dots, \gamma_N)$ such that there exist $\alpha, \beta \in I_1$ with $\alpha_i + \beta_i = 2\gamma_i$ for all $i = 1, \dots, N$. Then $I_2 \subseteq I_1$.*

Proof. Let σ and τ be two permutations that are associated with the monomials u^α and u^β, respectively, in the Leibniz expansion of $\det B_n(u)$. It suffices to construct a third permutation ω that is associated to the monomial u^γ. In order for $u^{\alpha+\beta}$ to be the square of a monomial, we must have that $\alpha_i + \beta_i$ is either $0, 2$, or 4. If $\alpha_i + \beta_i = 0$ then $\alpha_i = \beta_i = 0$. If $\alpha_i + \beta_i = 4$ then $\alpha_i = 2$ and $\beta_i = 2$. If $\alpha_i + \beta_i = 2$ then there are three possibilities: we can have $\alpha_i = 2$ and $\beta_i = 0$ or $\alpha_i = 1$ and $\beta_i = 1$ or $\alpha_i = 0$ and $\beta_i = 2$. In terms of σ and τ, for each pair of columns of $B_n(u)$ with column numbers $2k-1$ and $2k$, only the following three cases can occur.

Case 1. We have $\sigma(2k-1) = \tau(2k-1)$, which implies $\sigma(2k) = \tau(2k)$.
Case 2. We have $\sigma(2k-1) \neq \tau(2k-1)$ (which implies $\sigma(2k) \neq \tau(2k)$) and $\tau(2k) = \sigma(2k-1)+1$ (which implies $\sigma(2k) = \tau(2k-1)+1$).
Case 3. We have $\sigma(2k-1) \neq \tau(2k-1)$ (which implies $\sigma(2k) \neq \tau(2k)$), and $\tau(2k) = \tau(2k-1)+1$ and $\sigma(2k) = \sigma(2k-1)+1$.

Now we define ω, basing for our choice of $\omega(2k-1)$ and $\omega(2k)$ (for $1 \leq k \leq n-1$) in terms of the three cases enumerated above. In each of Cases 1 and 2, we simply take $\omega(2k-1) = \sigma(2k-1)$ and $\omega(2k) = \sigma(2k)$. In the remaining Case 3, we take $\omega(2k-1) = \sigma(2k-1)$ and $\omega(2k) = \tau(2k)$.

In order to see that ω is indeed a permutation, it suffices to see that it is one-to-one. Let $T \subseteq \{1, 2, ..., 2n-2\}$ denote the set of even integers $2k$ for which the pair of columns (c_{2k-1}, c_{2k}) with column numbers $2k-1$ and $2k$ falls into Case 3, and let $S = \{1, 2, ..., 2n-2\} \setminus T$. We have $\omega|_S = \sigma|_S$, and $\omega|_T = \tau|_T$, hence ω is injective on these subsets and one only needs to check that if $i \in S$ and $j \in T$ then $\omega(i) \neq \omega(j)$. We will need the following claim (note that we will only need the part of the claim pertaining to Case 3, but the statement concerns both Cases 2 and 3 as this is used in the inductive step of the proof).

Claim 3. *Assume a pair of columns with column numbers $2k-1$ and $2k$ falls into Case 2 (respectively Case 3). Then there exists another pair of columns that falls into Case 2 (respectively Case 3) with column numbers $2j-1$ and $2j$ with $j \neq k$ such that $\sigma(2k-1) = \tau(2j-1)$ (which then forces $\sigma(2j) = \tau(2k)$ in Case 2 and $\sigma(2k) = \tau(2j)$ in Case 3).*

We note that this claim can be verified pictorially from the equation (24) of the matrix $B_n(u)$ by "chasing" the row-column structure. However, we will proceed formally. We prove the claim simultaneously for Case 2 and Case 3 by induction on $q = \min\{\sigma(2k-1), \tau(2k-1)\}$. For the base of the induction, let $\min\{\sigma(2k-1), \tau(2k-1)\} = 1$. Since the roles of σ and τ are symmetric, we can assume without loss of generality that $\sigma(2k-1) = 1$. Consider ℓ such that $\tau(\ell) = 1$. Observe that ℓ must be odd, so we can write $\ell = 2j-1$, and the variable in position $(\ell+1, 2) = (2j, 2)$ must be selected by either σ or τ. Namely, we have $\sigma(2j) = 2$ in Case 2 and $\tau(2j) = 2$ in Case 3. i.e., if the pair of columns $p = (c_{2k-1}, c_{2k})$ falls into Case 2 (respectively Case 3), this forces the pair of columns (c_{2j-1}, c_{2j}) to fall into Case 2 (respectively Case 3).

Assume now that the Claim holds for $\min\{\sigma(2k-1), \tau(2k-1)\} \leq q$. For the inductive step, suppose the pair (c_{2k-1}, c_{2k}) falls into Case 2 or Case 3 with $\min\{\sigma(2k-1), \tau(2k-1)\} = q+1$, and again using the symmetry of the roles played by σ and τ we assume without loss of generality that $\sigma(2k-1) = q+1$. Let us assume that the pair is in Case 2 (a similar argument addresses Case 3). Pick ℓ such that $\tau(\ell) = \sigma(2k-1) = q+1$. First, notice that ℓ must be odd. In fact, if ℓ were even then the pair $p' = (c_{\ell-1}, c_\ell)$ would fall into either Case 2 or 3 and would give (by the inductive hypothesis) another pair p'' which is in the same Case as p'; this contradicts the injectivity of σ. Since ℓ is odd we write $\ell = 2j-1$. The variable in position $(2j, q+2)$ must be selected by either σ or τ, and we see that it must be σ, i.e., we have $\sigma(2j) = \tau(2j-1)+1$. Otherwise, if $\tau(2j) = \tau(2j-1)+1$ this would contradict the injectivity of τ. A similar argument applies when the pair (c_{2k-1}, c_{2k}) is in Case 3. This completes the inductive step and proves the claim.

Let us now return to verifying the injectivity of ω. Fixing an arbitrary $2k \in T$ and $i \in S$, let us prove that $\omega(i) \neq \omega(2k)$. We can apply Claim 3 to the pair of columns (c_{2k-1}, c_{2k}) (exchanging the roles of σ and τ) to produce j such that (c_{2j-1}, c_{2j}) falls into Case 3 and $\tau(2k-1) = \sigma(2j-1)$ and $\tau(2k) = \sigma(2j)$. We have $2j \in T$ which implies $i \neq 2j$, and hence $\sigma(2j) \neq \sigma(i)$. This implies $\omega(2k) = \tau(2k) = \sigma(2j) \neq \sigma(i) = \omega(i)$, showing that ω is injective.

Considering the variables that are selected by ω it is simple to check that it is associated with γ as desired. □

We will now prove the claimed upper bound for E_n.

Proposition 11. *Using the notation of Theorem 2 and Theorem 7, there exists $b > 1$ such that:*

$$E_n \leq \sqrt{n} b^n C_n^{1/2}.$$

Proof. As a first step, we claim that there exists $b > 1$ such that:

$$\mathbb{E}|\det \hat{J}_n^{\mathbb{R}}| \leq b^n \left(\mathbb{E}|\det \hat{J}_n^{\mathbb{C}}|^2\right)^{1/2}. \tag{26}$$

We want to apply the Cauchy-Schwarz inequality, and estimate the quantity:

$$\mathbb{E}|\det \hat{J}_n^{\mathbb{R}}| = \mathbb{E}|\det B_n(u)| \leq \left(\mathbb{E}\left(\det B_n(u)\right)^2\right)^{1/2} = \left(\mathbb{E}Q_n(u)^2\right)^{1/2}.$$

We write now:

$$Q_n(u)^2 = \sum_{\gamma \in I_1} Q_\gamma^2 u_1^{2\gamma_1} \cdots u_N^{2\gamma_N} + \sum_{\alpha \neq \beta} Q_\alpha Q_\beta u_1^{\alpha_1+\beta_1} \cdots u_N^{\alpha_N+\beta_N}, \tag{27}$$

where I_1 is the index set defined just before Lemma 9. Observe that, after taking expectation and using independence, in the second sum in (27) only terms such that $\alpha_i + \beta_i$ is even for all $i = 1, \ldots, N$ give a nonzero contribution:

$$\mathbb{E}Q_n(u)^2 = \sum_{\gamma \in I_1} Q_\gamma^2 \mathbb{E} u_1^{2\gamma_1} \cdots u_N^{2\gamma_N} + \sum_{\alpha \neq \beta \text{ and } \alpha+\beta \text{ ``even''}} Q_\alpha Q_\beta \mathbb{E} u_1^{\alpha_1+\beta_1} \cdots u_N^{\alpha_N+\beta_N}. \tag{28}$$

We now rewrite the double sum on the right as:

$$\sum_{\alpha \neq \beta \text{ and } \alpha+\beta \text{ even}} Q_\alpha Q_\beta \mathbb{E} u_1^{\alpha_1+\beta_1} \cdots u_N^{\alpha_N+\beta_N} = \sum_{\gamma \in I_2} \left(\sum_{\alpha+\beta=2\gamma} Q_\alpha Q_\beta\right) u_1^{2\gamma_1} \cdots u_N^{2\gamma_N}$$

where I_2 is the index set defined in Lemma 10.

Note that there exists $b_1 > 1$ such that $|Q_\gamma| \leq b_1^n$ for every $\gamma \in I_1$. In fact, it is easy to see that there are at most 2^{n-1} many ways a given monomial can appear as one of the summands in the Leibniz expansion of $\det B_n(u)$. Moreover there exists $b_2 > 1$ such that, for every fixed γ, the cardinality of the set of pairs (α, β) such that $\alpha + \beta = 2\gamma$ is bounded by b_2^n. In fact, given γ_i there are at most six possible pairs for (α_i, β_i) such that $\alpha_i + \beta_i = 2\gamma_i$ (namely $(0,0)$, $(0,2)$, $(1,1)$, $(1,2)$, $(2,0)$ and $(2,2)$). In our case each monomial $u_1^{2\gamma_1} \cdots u_N^{2\gamma_N}$ can have at most $2n-2$ many variables with nonzero exponents, hence combinatorially we have at most $6^{2n-2} \leq b_2^n$ many pairs (α, β) with $\alpha + \beta = 2\gamma$, as claimed. As a consequence we can bound:

$$\sum_{\gamma \in I_2} \left(\sum_{\alpha+\beta=2\gamma} Q_\alpha Q_\beta\right) u_1^{2\gamma_1} \cdots u_N^{2\gamma_N} \leq (b_1^2 b_2)^n \sum_{\gamma \in I_2} u_1^{2\gamma_1} \cdots u_N^{2\gamma_N}. \tag{29}$$

We now use the fact that in the expansion of $\det B_n(u)$ no coefficient Q_γ is zero for $\gamma \in I_1$ (by Lemma 9; moreover $|Q_\gamma| \geq 1$ and as a consequence we can write:

$$\sum_{\gamma \in I_2} u_1^{2\gamma_1} \cdots u_N^{2\gamma_N} \leq \sum_{\gamma \in I_1} u_1^{2\gamma_1} \cdots u_N^{2\gamma_N} \leq \sum_{\gamma \in I_1} Q_\gamma^2 u_1^{2\gamma_1} \cdots u_N^{2\gamma_N}.$$

In the first inequality we have used the fact that $I_2 \subseteq I_1$ (Lemma 10). Combining this with (28) and (29) we get:

$$\mathbb{E}Q_n(u)^2 \leq b_3^n \sum_{\gamma\in I_1} Q_\gamma^2 \mathbb{E}u_1^{2\gamma_1}\cdots u_N^{2\gamma_N}. \tag{30}$$

We now switch back to the variables $x_1, \dots, x_N$ (which are standard independent Gaussians). Recalling the definition of $P_n(x) = \det A_n(x) = \det B_n(u) = Q_n(u)$, we have:

$$\sum_{\gamma\in I_1} Q_\gamma^2 \mathbb{E}u_1^{2\gamma_1}\cdots u_N^{2\gamma_N} = \sum_{\gamma\in I_1} P_\gamma^2 \mathbb{E}x_1^{2\gamma_1}\cdots x_N^{2\gamma_N}.$$

We now look at $\mathbb{E}x_1^{2\gamma_1}\cdots x_N^{2\gamma_N}$ for $\gamma\in I_1$. Using independence:

$$\mathbb{E}x_1^{2\gamma_1}\cdots x_N^{2\gamma_N} = \prod_{i=1}^N \mathbb{E}x_i^{2\gamma_i} = \prod_{\{i\,|\,\gamma_i\neq 0\}} \mathbb{E}x_i^{2\gamma_i} \leq b_4^n \prod_{\{i\,|\,\gamma_i\neq 0\}} \gamma_i! = b_4^n \prod_{i=1}^N \gamma_i!.$$

For the inequality in the line above we have used the fact that only $2n-1$ many variables appear with a nonzero power, and for those variables we have $\gamma_i \leq 2$ which implies that the moment $\mathbb{E}x_i^{2\gamma_i} = (2\gamma_i - 1)!! \leq 2\gamma_i!$ (so we can take say $b_4 = 4$). In particular, continuing from (30) we get:

$$\begin{aligned}\mathbb{E}P_n(x)^2 = \mathbb{E}Q_n(u)^2 &\leq b_3^n \sum_{\gamma\in I_1} P_\gamma^2 \mathbb{E}x_1^{2\gamma_1}\cdots x_N^{2\gamma_N} \\ &\leq b_5^n \sum_{\gamma\in I_1} P_\gamma^2 \gamma_1!\cdots\gamma_N! \\ &= b_5^n (2n-2)! \sum_\gamma P_\gamma^2 \frac{\gamma_1!\cdots\gamma_N!}{(2n-2)!} \\ &= b_5^n (2n-2)! \|P_n\|_B^2.\end{aligned}$$

Recalling (22), we have $(2n-2)!\ \|P_n\|_B^2 = \mathbb{E}|\det \hat{J}_n^{\mathbb{C}}|^2$, which finally implies (26):

$$\mathbb{E}|\det \hat{J}_n^{\mathbb{R}}| \leq b^n \left(\mathbb{E}|\det \hat{J}_n^{\mathbb{C}}|^2\right)^{1/2}.$$

We can finally use Theorem 2, and estimate the quantity E_n as:

$$\begin{aligned}E_n &= \left(\frac{(2n-3)^{n-1}}{\Gamma(n)} \prod_{k=0}^{2n-3} \binom{2n-3}{k}^{-1/2}\right) \mathbb{E}|\det \hat{J}_n^{\mathbb{R}}| \\ &= \sqrt{n}\left(\frac{(2n-3)^{2n-2}}{\Gamma(n)\Gamma(n+1)} \prod_{k=0}^{2n-3} \binom{2n-3}{k}^{-1}\right)^{1/2} \mathbb{E}|\det \hat{J}_n^{\mathbb{R}}| \\ &\leq \sqrt{n}\left(\frac{(2n-3)^{2n-2}}{\Gamma(n)\Gamma(n+1)} \prod_{k=0}^{2n-3} \binom{2n-3}{k}^{-1}\right)^{1/2} b^n \left(\mathbb{E}|\det \hat{J}_n^{\mathbb{C}}|^2\right)^{1/2} \\ &\leq \sqrt{n} b^n C_n^{1/2}.\end{aligned}$$

This proves the statement in the proposition.

□

5.2. **The square root law.**

Theorem 12. *Using the notation of Theorem 2 and Theorem 7, we have:*

$$\lim_{n\to\infty} \frac{\log E_n}{\log C_n} = \frac{1}{2}.$$

Proof. Applying Corollary 4 and Proposition 11 we get:

$$(2n-3)!! \leq E_n \leq b^n \sqrt{n} C_n^{1/2}.$$

We note that $\log(2n-3)!! = n\log(n) + O(n)$ and that, by (2), $\log C_n = 2n\log(n) + O(n)$. As a consequence:

$$n\log(n) + O(n) = \log(2n-3)!! \leq \log E_n \leq \frac{1}{2}\log C_n + n\log b + O(\log n)$$

and, dividing by $\log C_n$:

$$\frac{n\log(n) + O(n)}{2n\log(n) + O(n)} \leq \frac{\log E_n}{\log C_n} \leq \frac{1}{2} + \frac{n\log b + O(\log n)}{2n\log(n) + O(n)}.$$

Taking the limit $n \to \infty$ yields the result. □

Note that the above proof entails an error estimate

$$\log E_n = \frac{\log C_n}{2} + O(n).$$

Moreoever, the fact that the lower bound in the above proof is deterministic leads to a basic concentration of measure estimate. Namely, let X_n be the number of real lines on a random hypersurface (i.e., X_n is the random variable whose expectation defines E_n). Then, first applying Markov's inequality to the non-negative random variable $\log X_n - \log(2n-3)!!$ and then applying Jensen's inequality, we have:

$$\begin{aligned}\mathbb{P}\{\log X_n - \log(2n-3)!! > t\} &\leq \frac{\mathbb{E}\log X_n - \log(2n-3)!!}{t} \\ &\leq \frac{\log \mathbb{E}X_n - \log(2n-3)!!}{t} \\ &= \frac{O(n)}{t}.\end{aligned}$$

This implies that, with high probability, the random variable $\log X_n$ is within $o\left(\frac{\log C_n}{2}\right)$ of $\frac{\log C_n}{2}$. More precisely, choosing, say, $t = t_n = n\sqrt{\log n}$ in the above estimate, we have the following concentration of measure result.

Theorem 13. *Let X_n be the number of real lines on a random Kostlan hypersurface of degree $2n-3$ in $n+1$ variables. Then*

$$\mathbb{P}\left\{\left|\log X_n - \frac{\log C_n}{2}\right| \geq n\sqrt{\log n}\right\} = O\left(\frac{1}{\sqrt{\log n}}\right), \quad \textit{as } n \to \infty.$$

References

[1] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York, 2007.

[2] Lars V. Ahlfors. *Complex analysis*. McGraw-Hill Book Co., New York, third edition, 1978. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.

[3] Daniel Allcock, James A. Carlson, and Domingo Toledo. Hyperbolic geometry and moduli of real cubic surfaces. *Ann. Sci. Éc. Norm. Supér. (4)*, 43(1):69–115, 2010.

[4] Jean-Marc Azais and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, Inc., Hoboken, NJ, 2009.

[5] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.

[6] Peter Bürgisser and Antonio Lerario. Probabilistic schubert calculus. *J. Reine Angew. Math. (Crelle's Journal)*, Published Online, doi:10.1515/crelle-2018-0009, 2018.

[7] Prof. Cayley. A Memoir on Cubic Surfaces. *Philosophical Transactions of the Royal Society of London Series I*, 159:231–326, 1869.

[8] Olivier Debarre and Laurent Manivel. Sur la variété des espaces linéaires contenus dans une intersection complète. *Math. Ann.*, 312(3):549–574, 1998.

[9] I. V. Dolgachev. *Classical algebraic geometry*. Cambridge University Press, Cambridge, 2012. A modern view.

[10] Alan Edelman and Eric Kostlan. How many zeros of a random polynomial are real? *Bull. Amer. Math. Soc. (N.S.)*, 32(1):1–37, 1995.

[11] Alan Edelman, Eric Kostlan, and Michael Shub. How many eigenvalues of a random matrix are real? *J. Amer. Math. Soc.*, 7(1):247–267, 1994.

[12] D. Eisenbud and J. Harris. 3264 and all that: A second course in algebraic geometry. *Cambridge University PRess*, 2016.

[13] A. Eremenko and A. Gabrielov. Rational functions with real critical points and the B. and M. Shapiro conjecture in real enumerative geometry. *Ann. of Math. (2)*, 155(1):105–129, 2002.

[14] S. Finashin and V. Kharlamov. Abundance of real lines on real projective hypersurfaces. *Int. Math. Res. Not. IMRN*, (16):3639–3646, 2013.

[15] S. Finashin and V. Kharlamov. Abundance of 3-planes on real projective hypersurfaces. *Arnold Math. J.*, 1(2):171–199, 2015.

[16] E. Fischer. Über die Differentiationsprozesse der Algebra. *J. für Math.*, 148:1–17, 1917.

[17] Y. V. Fyodorov, A. Lerario, and E. Lundberg. On the number of connected components of random algebraic hypersurfaces. *J. Geom. Phys.*, 95:1–20, 2015.

[18] N. R. Goodman. Statistical analysis based on a certain multivariate complex Gaussian distribution. (An introduction). *Ann. Math. Statist.*, 34:152–177, 1963.

[19] D. B. Grünberg and P. Moree. Sequences of enumerative geometry: congruences and asymptotics. *Experiment. Math.*, 17(4):409–426, 2008. With an appendix by Don Zagier.

[20] Ralph Howard. The kinematic formula in Riemannian homogeneous spaces. *Mem. Amer. Math. Soc.*, 106(509):vi+69, 1993.

[21] Dmitry Khavinson and Erik Lundberg. *Linear holomorphic partial differential equations and classical potential theory*, volume 232 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2018.

[22] E. Kostlan. On the distribution of roots of random polynomials. In *From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990)*, pages 419–431. Springer, New York, 1993.

[23] Eric Kostlan. On the expected number of real roots of a system of random polynomial equations. In *Foundations of computational mathematics (Hong Kong, 2000)*, pages 149–188. World Sci. Publ., River Edge, NJ, 2002.

[24] S. E. Kozlov. Geometry of real Grassmannian manifolds. I, II, III. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 246(Geom. i Topol. 2):84–107, 108–129, 197–198, 1997.

[25] Evgeny Mukhin, Vitaly Tarasov, and Alexander Varchenko. The B. and M. Shapiro conjecture in real algebraic geometry and the Bethe ansatz. *Ann. of Math. (2)*, 170(2):863–881, 2009.

[26] D. J. Newman and H. S. Shapiro. Certain Hilbert spaces of entire functions. *Bull. Amer. Math. Soc.*, 72:971–977, 1966.

[27] Liviu I. Nicolaescu. A stochastic Gauss-Bonnet-Chern formula. *Probab. Theory Related Fields*, 165(1-2):235–265, 2016.

[28] Ch. Okonek and A. Teleman. Intrinsic signs and lower bounds in real algebraic geometry. *J. Reine Angew. Math.*, 688:219–241, 2014.

[29] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[30] I. Scarowsky. Quadratic forms in normal variables. *Thesis (M.Sc.)–McGill University*, 1973.

[31] Dr. Schläfli. On the distribution of surfaces of the third order into species, in reference to the absence or presence of singular points, and the reality of their lines. *Philosophical Transactions of the Royal Society of London*, 153:193–241, 1863.

[32] B. Segre. *The Non-singular Cubic Surfaces*. Oxford University Press, Oxford, 1942.

[33] H. S. Shapiro. An algebraic theorem of E. Fischer, and the holomorphic Goursat problem. *Bull. London Math. Soc.*, 21(6):513–537, 1989.

[34] M. Shub and S. Smale. Complexity of Bezout's theorem. II. Volumes and probabilities. In *Computational algebraic geometry (Nice, 1992)*, volume 109 of *Progr. Math.*, pages 267–285. Birkhäuser Boston, Boston, MA, 1993.

[35] Michael Shub and Steve Smale. Complexity of Bézout's theorem. I. Geometric aspects. *J. Amer. Math. Soc.*, 6(2):459–501, 1993.

[36] Michael Shub and Steve Smale. Complexity of Bezout's theorem. III. Condition number and packing. *J. Complexity*, 9(1):4–14, 1993. Festschrift for Joseph F. Traub, Part I.

[37] F. Sottile. Enumerative real algebraic geometry. In *Algorithmic and quantitative real algebraic geometry (Piscataway, NJ, 2001)*, volume 60 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 139–179. Amer. Math. Soc., Providence, RI, 2003.

[38] F. Sottile. *Real solutions to equations from geometry*, volume 57 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2011.

[39] Frank Sottile. Enumerative geometry for the real Grassmannian of lines in projective space. *Duke Math. J.*, 87(1):59–85, 1997.

[40] Frank Sottile. Pieri's formula via explicit rational equivalence. *Canad. J. Math.*, 49(6):1281–1298, 1997.

[41] Terence Tao and Van Vu. A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.*, 231(1):74–101, 2012.

[42] P. Turán. On a problem in the theory of determinants. *Acta Math. Sinica*, 5:411–423, 1955.

[43] Ravi Vakil. Schubert induction. *Ann. of Math. (2)*, 164(2):489–512, 2006.