Abh. Math. Semin. Univ. Hambg. (2023) 93:149–162
https://doi.org/10.1007/s12188-023-00269-7

Lines on *p*-adic and real cubic surfaces

Rida Ait El Manssour[1] · **Yassine El Maazouz**[2] · **Enis Kaya**[3] · **Kemal Rose**[1]

Received: 1 March 2023 / Accepted: 21 August 2023 / Published online: 16 September 2023

Abstract

We study lines on smooth cubic surfaces over the field of p-adic numbers, from a theoretical and computational point of view. Segre showed that the possible counts of such lines are 0, 1, 2, 3, 5, 7, 9, 15 or 27. We show that each of these counts is achieved. Probabilistic aspects are investigated by sampling both p-adic and real cubic surfaces from different distributions and estimating the probability of each count.We link this to recent results on probabilistic enumerative geometry. Some experimental results on the Galois groups attached to p-adic cubic surfaces are also discussed.

Keywords p-adic cubic surfaces · Real cubic surfaces · Probabilistic enumerative geometry

Mathematics Subject Classification 14Q10 · 14N10 · 60B99

1 Introduction

Smooth cubic surfaces are arguably one of the most famous objects in classical algebraic geometry [1–3]. The problem of counting lines on such surfaces also has a long standing history and there is a vast literature on the topic. In the nineteenth century, Cayley [4] and Schläfli [5] proved that a smooth cubic surface over the complex numbers always has 27 lines, while over the real numbers it can have 3, 7, 15 or 27 lines. Segre [6] showed that over

Communicated by Henrik Bachmann.

✉ Enis Kaya
enis.kaya@kuleuven.be

Rida Ait El Manssour
rida.manssour@mis.mpg.de

Yassine El Maazouz
yassine.el-maazouz@berkeley.edu

Kemal Rose
kemal.rose@mis.mpg.de

[1] MPI MIS, Inselstrasse 22, 04103 Leipzig, Germany

[2] Department of Statistics, UC Berkeley, Evans Hall #3860, Berkeley, CA 94720, USA

[3] Department of Mathematics, KU Leuven, Celestijnenlaan 200B, B-3001 Leuven, Belgium

Table 1 Classification of possible line counts for the fields $\mathbb{C}, \mathbb{R}, \mathbb{Q}, \mathbb{F}_q$

Base field	Numbers attained	References
$\mathbb{C}$	27	[4]
$\mathbb{R}$	3, 7, 15, 27	[5]
$\mathbb{Q}$	0, 1, 2, 3, 5, 7, 9, 15, 27	[6]
$\mathbb{F}_q, q > 5$ odd	0, 1, 2, 3, 5, 7, 9, 15, 27	[8]
$\mathbb{F}_{2^n}, n \geq 2$	0, 1, 2, 3, 5, 7, 9, 15, 27	[8]
$\mathbb{F}_5$	0, 1, 2, 3, 5, 7, 9, 15	[8]
$\mathbb{F}_2,\ \mathbb{F}_3$	0, 1, 2, 3, 5, 9, 15	[8, 9]

any field, a smooth cubic surface can only have 0, 1, 2, 3, 5, 7, 9, 15, or 27 lines[1]. Depending on the base field, some of these numbers may not be attained. Since then, the problem of counting lines on cubic surfaces over different fields attracted many mathematicians and some cases have already been classified; see Table 1.

Recently, McKean [7] showed that all of the numbers 0, 1, 2, 3, 5, 7, 9, 15 and 27 occur when the base field is finitely generated with at least 22 elements or a finite transcendental extension of an arbitrary field; see [7, Corollary 1.6].

In this text, our main aim is to get an idea of how many lines are on a cubic surface over $\mathbb{Q}_p$ or $\mathbb{R}$ from a probabilistic point of view. To this end, we conduct numerical experiments when the base field is $\mathbb{Q}_7$ or $\mathbb{R}$. While over the real numbers, a cubic surface can only have either 3, 7, 15 or 27 lines, by adapting the approach in an earlier version of [7][2] to the case we are interested in, we get the following:

Theorem 1.1 *Let* $n \in \{0, 1, 2, 3, 5, 7, 9, 15, 27\}$. *Then there exists a smooth cubic surface over* $\mathbb{Q}_p$ *that contains exactly* n *lines. In other words, all possible line counts occur when the base field is* $\mathbb{Q}_p$.

In Sect. 2, we explain how to explicitly construct (by blowing up the projective plane in 6 suitable points) a smooth cubic surface having any of the line counts mentioned in Theorem 1.1. Sections 3 and 4 focus on probabilistic computations and heuristics for both the p-adic and real case, respectively. For the p-adic case, we sample from the family of smooth cubic surfaces with four different probability measures (see Sect. 3), and compute the probability of seeing each number of lines. Table 3 summarizes the distributions of the number of lines when $p = 7$. For the real case, we consider a one-parameter family $(\mathbb{P}_\lambda)_{0<\lambda<1}$ of Gaussian distributions studied in [10]. The probability distribution of line counts is then a curve in the 3-simplex which is depicted in Fig. 1. The Galois groups attached to smooth cubic surfaces are of special interest. The final section of this paper (Sect. 5) is devoted to experimental results on which Galois groups appear for cubic surfaces defined over $\mathbb{Q}_p$. Our results suggest that Galois groups should be quite small and usually abelian. This motivates an inverse Galois problem for smooth cubic surfaces over $\mathbb{Q}_p$ in line with Elsenhans and Jahnel's work [11] on the inverse Galois problem for smooth cubic surfaces over $\mathbb{Q}$.

Most of the results in this text were found by computation. The codes and data are made available at

https://mathrepo.mis.mpg.de/27pAdicLines/index.html. (1)

[1] In his work, Segre actually points out that his statement fails in characteristic 2, but this is not correct; see [7, Section 3.1].

[2] This was done prior to the appearance of McKean's stronger result [7, Theorem 1.3].

Our computations were carried out using the computer algebra systems Magma [12], Julia [13] and Macaulay2 [14].

2 Lines on smooth *p*-adic cubic surfaces

In the following, we recall a well-known representation of smooth cubic surfaces: the blow-up of the projective plane $\mathbb{P}^2$ at six $\overline{\mathbb{Q}}_p$-rational points in general position is a smooth cubic surface (see, for example, [15] for a detailed treatment). Up to a change of coordinates, the six points lie on the cuspidal cubic curve

$$\mathscr{C} : \theta \longmapsto (1 : \theta : \theta^3)$$

and can be represented as the columns of the matrix

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 & \theta_6 \\ \theta_1^3 & \theta_2^3 & \theta_3^3 & \theta_4^3 & \theta_5^3 & \theta_6^3 \end{bmatrix}. \tag{2}$$

The six points are said to be in *general position* if no three of them lie on a line and not all of them lie on a conic. Equivalently the points are in general position if both the maximal minors of (2), which are given by

$$[ijk] := (\theta_i - \theta_j)(\theta_i - \theta_k)(\theta_j - \theta_k)(\theta_i + \theta_j + \theta_k) \quad \text{for } 1 \leq i < j < k \leq 6,$$

and the polynomial

$$\begin{aligned} &[134][156][235][246] - [135][146][234][256] \\ &:= (\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5 + \theta_6) \prod_{1 \leq i < j \leq 6} (\theta_i - \theta_j), \end{aligned}$$

do not vanish. These polynomials split into linear forms that together determine the hyperplane arrangement of type E_6 (see [16, Section 6] for more details):

$$\begin{array}{ll} \theta_i - \theta_j & \text{for } 1 \leq i < j \leq 6, \\ \theta_i + \theta_j + \theta_k & \text{for } 1 \leq i < j < k \leq 6, \\ \theta_1 + \theta_2 + \cdots + \theta_6. & \end{array} \tag{3}$$

Definition 2.1 Let $F = \prod_{i=1}^{6} (X - \theta_i) \in \mathbb{Q}_p[X]$ be a univariate polynomial of degree 6 whose roots lie in the complement of the hyperplane arrangement (3). We denote by $\mathcal{S}(F)$ the smooth cubic surface that is the blow-up of the projective plane at the six points $\{[1 : \theta_i : \theta_i^3]\}_{1 \leq i \leq 6}$.

The defining equation of $\mathcal{S}(F)$ in $\mathbb{P}^3$ in terms of the θ_i is determined in [15, Equation 4]. The 27 $\overline{\mathbb{Q}}_p$-rational lines on $\mathcal{S}(F)$ are of three distinct types:

(i) $\{E_i : i = 1, \ldots, 6\}$, where E_i is the exceptional divisor of the point $\mathcal{C}(\theta_i)$;
(ii) $\{F_{i,j} : i, j = 1, \ldots, 6,\ i \neq j\}$, where $F_{i,j}$ is the strict transform of the line passing through the points $\mathcal{C}(\theta_i)$ and $\mathcal{C}(\theta_j)$;
(iii) $\{G_i : i = 1, \ldots, 6\}$, where G_i is the strict transform of the unique conic passing through the points $\{\mathcal{C}(\theta_1), \ldots, \mathcal{C}(\theta_6)\} \setminus \{\mathcal{C}(\theta_i)\}$.

In view of constructing smooth cubic surfaces that have a prescribed number of $\mathbb{Q}_p$-rational lines, we investigate the action of the absolute Galois group $G = \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ on

Table 2 A list of surfaces $\mathcal{S}(F)$ that realize all possible line-counts over $\mathbb{Q}_p$

Polynomial F	(ℓ, q)	# of lines on $\mathcal{S}(F)$
$X^6 + pX^5 + p$	$(0, 0)$	0
$(X^4 + p)(X^2 + pX + p)$	$(0, 1)$	1
$X(X^5 + pX^4 + p)$	$(1, 0)$	2
$(X^2 + p)(X^2 + pX + p)(X^2 + p^2X + p)$	$(0, 3)$	3
$(X + 1)(X + 2)(X^4 + p)$	$(2, 0)$	5
$(X + 1)(X + 2)(X^2 + p)(X^2 + pX + p)$	$(2, 2)$	7
$(X + 1)(X + 2)(X + 3)(X^3 + pX^2 + p)$	$(3, 0)$	9
$(X + 1)(X + 2)(X + 3)(X + 4)(X^2 + p)$	$(4, 1)$	15
$X(X + 1)(X + 2)(X + 3)(X + 4)(X + 5)$	$(6, 0)$	27

$\mathcal{S}(F)$, and show that the line-count depends only on the decomposition of F into irreducible factors over $\mathbb{Q}_p$.

Lemma 2.2 *The smooth cubic surface $\mathcal{S}(F)$ is $\mathbb{Q}_p$-rational and contains exactly*

$$2\ell + q + \binom{\ell}{2}$$

$\mathbb{Q}_p$-rational lines, where ℓ and q are the number of linear and quadratic irreducible factors of F in $\mathbb{Q}_p[X]$, respectively.

Proof The absolute Galois group G acts on $\mathcal{S}(F)$ and lines on it by permuting the roots $\theta_1, \dots, \theta_6$. In particular, for every element σ of G, we have

$$\sigma(\mathcal{S}(F)) = \mathcal{S}(\sigma(F)),\ \sigma(E_i) = E_{\sigma(i)},\ \sigma(F_{i,j}) = F_{\sigma(i),\sigma(j)},\ \sigma(G_i) = G_{\sigma(i)},$$

where we define $\sigma(i)$ by enforcing $\sigma(\theta_i) = \theta_{\sigma(i)}$. Since G acts trivially on $F \in \mathbb{Q}_p[X]$, the surface $\mathcal{S}(F)$ is stable under G. Hence it is $\mathbb{Q}_p$-rational. Note that the number of G-stable E_i's is ℓ, the number of G-stable $F_{i,j}$'s is $q + \binom{\ell}{2}$, and the number of G-stable G_i's is ℓ. This finishes the proof. □

Proof of Theorem 1.1 This is a direct consequence of the above lemma. The polynomials in Table 2 have the desired numbers of irreducible linear and quadratic factors. The discriminant, the degree 5 coefficients $\theta_1 + \cdots + \theta_6$ and the expression

$$\prod_{\#\{i,j,k\}=3} \theta_i + \theta_j + \theta_k$$

are symmetric and can be expressed as polynomials in the coefficients of F. The non-vanishing of these quantities for the polynomials in Table 2 was certified using Macaulay2 (see (1)); therefore, the equations in (3) are not satisfied for any of these 9 polynomials. □

Remark 2.3 (1) It is clear that Theorem 1.1 continues to hold if we replace $\mathbb{Q}_p$ with any finite extension of $\mathbb{Q}_p$. In fact, it holds for any local field (one just needs to modify the above proof slightly). But, as we mentioned in the introduction, this is a special case of a theorem proved by McKean, after our work was completed, in the final version of his paper; see [7, Corollary 1.6]. We think it is still valuable to have explicit polynomials that yield the desired surfaces.

(2) It is important to note that, while any cubic surface over $\overline{\mathbb{Q}}_p$ is isomorphic to the blow-up of 6 points in general position in $\overline{\mathbb{Q}}_p\mathbb{P}^2$ that is not the case over $\mathbb{Q}_p$. In other words, there are cubic surfaces defined over $\mathbb{Q}_p$ which do not arise from the blow-up construction as in Sect. 2. The cubic surfaces arising in that way are the cubic surfaces which have a Galois invariant (i.e. $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-invariant) *double-six*. A double-six is a configuration of 12 lines $\{L_1, \dots, L_6\}$ and $\{L'_1, \dots, L'_6\}$ in $\overline{\mathbb{Q}}_p\mathbb{P}^3$ arranged in matrix form

$$\begin{pmatrix} L_1 & L_2 & L_3 & L_4 & L_5 & L_6 \\ L'_1 & L'_2 & L'_3 & L'_4 & L'_5 & L'_6 \end{pmatrix},$$

such that any two lines of these 12 are secant if and only if they are not on the same row or column. With the notation of Definition 2.1, the configuration of lines $\{E_1, \dots, E_6\}$ and $\{G_1, \dots, G_6\}$ is a Galois invariant double-six; see, for example, [1, Remark V.4.9.1].

3 Heuristics for the p-adic numbers

There are plenty of ways one can construct a smooth cubic surface over $\mathbb{Q}_p$. Each leads to a different way of sampling such surfaces and hence a measure on the space of smooth cubic surfaces. For our probabilistic investigations, we focus on four sampling methods and describe both the measures and how the sampling process is implemented in our computations.

3.1 The haar measure

We endow the space of cubics $\mathbb{Q}_p[x_0, x_1, x_2, x_3]_{(3)}$ with its natural measure, defined by choosing the 20 coefficients ξ_α of the degree 3 cubic

$$f = \sum_{|\alpha|=3} \xi_\alpha x^\alpha \tag{4}$$

to be independent and uniformly distributed in $\mathbb{Z}_p$ with respect to the Haar measure on $\mathbb{Z}_p$.

Definition 3.1 The probability measure of the random smooth cubic surface $\mathcal{S}(f)$ defined as the zero set of a random polynomial f as in (4) is called the *Haar measure* on the space of cubic surfaces.

Note that the singular cubics lie on a hypersurface of $\mathbb{Q}_p[x_0, x_1, x_2, x_3]_{(3)}$ and hence with probability 1 the random cubic f in (4) defines a smooth surface. Moreover, the Haar measure in Definition 3.1 is invariant under change of variables, i.e.

$$\text{for any } g \in \mathrm{GL}_4(\mathbb{Z}_p),\ \ f(g \cdot x) \text{ has the same istribution as } f(x).$$

Remark 3.2 We remark that one may make a different choice of basis in (4) and get another measure on the space of polynomials $\mathbb{Q}_p[x_0, x_1, x_2, x_3]_{(3)}$. The reason we chose the monomial basis in (4) is because it is guaranteed to be invariant under change of variables, moreover when p is big enough, namely $p > 3$, the resulting measure is actually the only one that is invariant under the action of $\mathrm{GL}_4(\mathbb{Z}_p)$, see [17, Theorem 1.1].

For our computational experiments, we can determine the coefficients only up to finite precision. So in practice, we sample the variables ξ_α are considered to be random variables with uniform distribution on the set $\{0, 1, \dots, p^{N+1} - 1\}$ for some precision N.

We note that the Haar measure has already been used on the space of cubic surfaces over $\mathbb{Q}_p$ to compute the expected number of lines; see [18]. To the best of our knowledge, the other measures (to be defined next) are not recorded anywhere in the literature.

3.2 The blow-up measure

As we saw in Sect. 2, one way to construct a cubic surface is to blow the plane $\mathbb{P}^2$ at six points in general position. If the six points are chosen randomly, we then get a random cubic surface. More precisely we define a new measure on the space of cubic surfaces as follows:

Definition 3.3 (The blow-up measure) Let $\xi_0, \ldots, \xi_6$ be independent uniformly distributed random variables in $\mathbb{Z}_p$ and let F be the random polynomial $F = \xi_0 + \xi_1 X + \cdots + \xi_6 X^6$. With probability 1, the polynomial F has degree 6, its roots $\theta_1, \ldots, \theta_6$ are simple and lie in the complement of the hyperplane arrangement (3). This defines a measure on the space of degree 6 polynomials. The map $F \longmapsto \mathcal{S}(F)$ determines a measure on the space of smooth cubic surfaces via pushforward which we call the *blow-up measure*.

Again, in practice, sampling a polynomial F is done by choosing a large integer N and sampling the coefficients of F from the set $\{0, 1, \ldots, p^{N+1} - 1\}$ independently with respect to the uniform distribution. The number of lines on the cubic surface $\mathcal{S}(F)$ is determined by factorizing F over the field K and applying Lemma 2.2.

As mentioned in Remark 2.3, not all cubic surfaces over $\mathbb{Q}_p$ arise from a blow-up of 6 points in $\mathbb{P}^2$. Therefore, the blow-up measure only sees cubic surfaces with a Galois-invariant double-six.

3.3 The tropical (generic) measure

The following theorem suggests that the number of lines on a smooth cubic surface $\mathcal{S}(F)$ is tightly linked to the tropicalization of F:

Theorem 3.4 (Theorem 3.5 in [15]) *Fix a prime $p \geq 5$ and a cubic surface $\mathcal{S}(F)$ over $\mathbb{Q}_p$, as in Definition 2.1. If $\mathcal{S}(F)$ is "tropically smooth", then the 27 lines on $\mathcal{S}(F)$ have distinct tropicalizations in the tropical projective space $\mathbb{TP}^3$. In that case, all 27 lines on $\mathcal{S}(F)$ are defined over $\mathbb{Q}_p$.*

When it comes to the combinatorial type of the cubic equation f (i.e., the regular subdivision it induces on the polytope $3\Delta_3$), sampling from the Haar measure is not the best way of sampling polynomials with a diverse combinatorial types (or *tropicalizations*). Under the Haar measure on $\mathbb{Z}_p$, it is quite rare to see elements with a large valuation. In order to remedy this shortcoming, we may sample from $\mathbb{Q}_p[x_0, x_1, x_2, x_3]_{(3)}$ by prescribing the valuation of each coefficient of f. More precisely,

Definition 3.5 (The tropical measure) Let N be a positive integer and let $(\nu_\alpha)_{|\alpha|=3}$ be independent uniform random variables in $\{0, \ldots, N\}$. We then obtain the random cubic equation

$$f = \sum_{|\alpha|=3} p^{\nu_\alpha} x^\alpha.$$

When N is large enough, the cubic surface $\mathcal{S}(f)$ defined by f is almost surely smooth. This defines a measure on the space of smooth cubic surfaces which we call the *tropical measure*.

It is quite natural to adapt this measure in the following way:

Table 3 The distribution of the number of lines for different probabilistic measures over the 7-adic numbers

# of $\mathbb{Q}_7$-Lines	Haar Mes.	Blow-up Mes.	Trop. Mes.	Trop. Gen. Mes.
0	0.43995	0.19446	0.22230	0.25580
1	0.34534	0.12608	0.29004	0.26891
2	0.08686	0.19604	0.02106	0.02083
3	0.08564	0.21602	0.29145	0.26952
5	0.03169	0.12778	0.04620	0.04988
7	0.00467	0.07337	0.06708	0.06911
9	0.00401	0.05099	0.02868	0.03443
15	0.00045	0.01500	0.02582	0.02666
27	0.00001	0.00026	0.00487	0.00406
Average	1.01023	3.00964	2.68398	2.67200

Definition 3.6 (The tropical generic measure) Let N be a positive integer and let $(\nu_\alpha)_{|\alpha|=3}$ be independent uniform random variables in $\{0, \dots, N\}$ and $(c_\alpha)_{|\alpha|=3}$ be independent uniform random variables in $\mathbb{Z}_p^\times$ (with respect to the Haar measure on $\mathbb{Z}_p$). We then obtain the random cubic equation

$$f = \sum_{|\alpha|=3} p^{\nu_\alpha} u_\alpha x^\alpha.$$

The cubic surface $\mathcal{S}(f)$ is then smooth with probability 1. The measure so defined on the space of smooth cubic surfaces is called the *tropical generic measure*.

3.4 Experiments for $p = 7$

For each probability measure μ on the space of cubic surfaces, we can obtain a probability measure $\pi^{(\mu)}$ on the set $\{0, 1, 2, 3, 5, 7, 9, 15, 27\}$ of possible line counts, where $\pi_i^{(\mu)}$ records the probability under μ that a cubic surface contains i lines. In general, given the measure μ, it is quite hard to determine the distribution $(\pi_i^{(\mu)})$, even for the probability distributions defined in Sect. 3. Hence, we sample a large number of cubic surfaces under each measure and use Monte–Carlo estimation to get an idea of how this distribution looks like for the measures we defined above.

We investigated each of the above measures experimentally by sampling 10^5 instances of smooth cubic surfaces and counting the corresponding number of lines. The latter is accomplished by using Gröbner basis techniques, or rather Lemma 2.2 for the blow-up measure. The resulting distributions on the number of lines are depicted in Table 3. The code we used for sampling surfaces and counting lines can be found at (1).

It is important to note the following:

(1) Our result only estimates[3] the distribution $\pi^{(\mu)} = \left(\pi_k^{(\mu)}\right)$ for any chosen measure μ. This approximation relies on the law of large numbers, so our results are random but converge to the correct distribution as the sample size gets larger and larger.

[3] Using a Monte-Carlo method.

(2) The dimension of the space of cubic equations is 20, which would make the computation of integrals[4] quite expensive. One of the advantages of the estimation method we used is that it is not affected by the dimension of the space we sample from.

(3) When dealing with computations over a p-adic field, one has to be careful with precision. Our computation were conducted with an absolute p-adic precision of 300, while a random p-adic number in $\mathbb{Z}_p$ was sampled from the uniform distribution on the set $\{0, \dots, p^8 - 1\}$. All these choices were made so that the computation runs in a reasonable time all the while keeping the results significant.

(4) Finally, it should also be mentioned that our implementations in (1) can be optimized a great deal from a performance point of view. So it is very much possible to go beyond the sample sizes we have used.

3.5 Interpretation of the results

We observe from Table 3 that under the Haar measure the probabilities of the line counts decrease. Actually most surfaces have 0 or 1 $\mathbb{Q}_p$-lines so under this measure, it is quite rare to find a cubic surfaces with a high count of p–adic lines, in particular 27-lines (around $\simeq 10^{-5}$ probability). This explains that the average number of p-adic lines under the Haar measure is almost 1. In fact, the following theorem quantifies the expected number which is $\simeq 1.01749$ for $p = 7$.

Theorem 3.7 ([18, Theorem 3]) *The expected number of p-adic lines on a random uniform p-adic cubic surface in* $\mathbb{P}^3$ *is* $\frac{(p^3-1)(p^2+1)}{p^5-1}$.

We observe that under both tropical measures, the chance of seeing bigger line counts is significantly higher compared to the Haar measure. Notice also that both measures yield more or less the same distribution of line counts. This is probably due to the fact that the number of lines depends heavily on the tropicalization of the cubic equation f, and hence also on the induced triangulation on the newton polytope $3\Delta_3$. This is in light of [15, Conjecture 4.1].

As far as the blow-up measure is concerned, we see again bigger number of lines with higher probability compared to any of the other measures. This can be explained by the fact that this measure only sees the cubic surfaces with a $\mathbb{Q}_p$-rational double-six, which generally have more lines than those that do not. This is in line with the results from [19, Theorem 1] where the probability $\rho_n(r)$ that a degree n polynomial with random independent and uniform coefficients in $\mathbb{Z}_p$ has exactly r roots is given recursively. The particular case of interest to us is the case $n = 6$ and $p = 7$, where we get

$$\begin{aligned}
\rho_6(0) + \rho_6(1) &= \frac{72800100990600581357013564212293034519 29}{99151249001680027034372294700769267020 00} \simeq 0.7343,\\
\rho_6(2) &= \frac{1321420868520256483059804018443386713 77}{6610083266778668468958152980051284468 00} \simeq 0.19991,\\
\rho_6(3) &= \frac{627473746053959019283493792850291928 3}{1239390612521000337929653683759615837 75} \simeq 0.05063,\\
\rho_6(4) &= \frac{327980509096640494280261674503468580 3}{2203361088926222822986050993350428156 00} \simeq 0.01489,\\
\rho_6(5) &= 0,\\
\rho_6(6) &= \frac{379252487878267254806025930638752849}{110168054446311141149302549667521407 8000} \simeq 0.00035.
\end{aligned}$$

[4] The probability that a random surface has a certain number of lines is an integral.

Table 4 Comparison of theoretical and experimental results for the blow-up measure

# of roots of F	0, 1	2	3	4	6
# of lines on $\mathcal{S}(F)$	0, 1, 2, 3	5, 7	9	15	27
Results from [19]	0.73423	0.19991	0.05063	0.01489	0.00035
Estimates from table 3	0.73260	0.20115	0.05099	0.01500	0.00026

Table 4 compares the exact results from [19] to the estimates obtained in our experiments.

4 Heuristics for the real numbers

Let $V = \mathbb{R}[x_0, x_1, x_2, x_3]_{(3)}$ be the real vector space of homogeneous degree three polynomials in $x = (x_0, x_1, x_2, x_3)$. We endow V with the inner product $\langle\cdot,\cdot\rangle$ defined as

$$\langle f_1, f_2\rangle := \frac{1}{4\pi^2}\int_{\mathbb{R}^4} f_1(x) f_2(x) e^{-\frac{\|x\|^2}{2}} dx.$$

The space V is endowed with the action of the orthogonal group $O(4, \mathbb{R})$ by change of variables, i.e.,

$$(g \cdot f)(x) = f(g^{-1}x) \quad \text{for any } f \in V,\ g \in O(4, \mathbb{R}).$$

This makes V a linear representation of $O(4, \mathbb{R})$ which is moreover unitary, i.e., the inner product $\langle\cdot,\cdot\rangle$ is stable under $O(4, \mathbb{R})$:

$$\langle g \cdot f_1,\ g \cdot f_2\rangle = \langle f_1, f_2\rangle \quad \text{for } f_1, f_2 \in V \text{ and } g \in O(4, \mathbb{R}).$$

However, V is not irreducible, so $\langle\cdot,\cdot\rangle$ is the not the unique $O(4, \mathbb{R})$-invariant inner product on V. The representation V splits into two irreducible orthogonal sub-representations as follows:

$$V = \mathscr{H}_3 \oplus \|x\|^2 \cdot \mathcal{H}_1 \tag{5}$$

where $\mathscr{H}_3$ is the subspace of homogeneous harmonic polynomials of degree 3 and $\mathcal{H}_1$ is the space of homogeneous degree 1 polynomials. Let p_1 and p_2 be the orthogonal projections of V onto $\mathcal{H}_3$ and $\|x\|^2\mathcal{H}_1$, respectively.

Definition 4.1 Let λ and μ be positive real numbers. We define the inner product $\langle\cdot,\cdot\rangle_{\lambda,\mu}$ on V as follows:

$$\langle f_1, f_2\rangle_{\lambda,\mu} = \frac{1}{\lambda^2}\langle p_1(f_1), p_1(f_2)\rangle + \frac{1}{\mu^2}\langle p_2(f_1), p_2(f_2)\rangle.$$

We also define the centered Gaussian measure $(\mathbb{P}_{\lambda,\mu})_{\lambda,\mu>0}$ on V associated to $\langle\cdot,\cdot\rangle_{\lambda,\mu}$ i.e. the measure on V whose density $\phi_{\lambda,\mu}$ is proportional to

$$\phi_{\lambda,\mu}(f) \propto \exp\left(\frac{-\langle f, f\rangle_{\lambda,\mu}}{2}\right), \quad \text{for } f \in V.$$

The two parameter family in Definition 4.1 parametrizes all $O(4, \mathbb{R})$-invariant inner products on V and hence all non-degenerate Gaussian measures on V that are $O(4, \mathbb{R})$-invariant are of the form $\mathbb{P}_{\lambda,\mu}$ (see [10, Section 4] for more details).

To sample a random cubic $f \in V$ with respect to the measure $\mathbb{P}_{\lambda,\mu}$, we use the orthonormal basis $\{H_{3,i}\}_{1\leq i\leq 16}$ and $\{H_{1,j}\}_{1\leq j\leq 4}$ respectively of $\mathcal{H}_3$ and $\|x\|^2\mathcal{H}_1$ given in [20, Table 1].

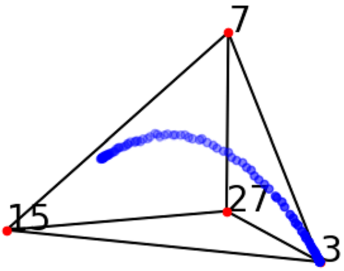

Fig. 1 The curve $(\pi^{(\lambda)})_{\lambda\in(0,1)}$ inside the 3-dimensional probability simplex

Table 5 The distribution of number of real lines under the Kostlan measure

$\pi_3^{(1/3)}$	$\pi_7^{(1/3)}$	$\pi_{15}^{(1/3)}$	$\pi_{27}^{(1/3)}$	Average # of lines
0.570252	0.338973	0.089406	0.001369	5.46162

More precisely, let $(\xi_{3,i})_{1\le i\le 16}$ and $(\xi_{1,j})_{1\le j\le 4}$ be two independent sequences of independent identically distributed standard Gaussian random variables and let us define the random cubic f as follows:

$$f = \lambda \sum_{1\le i\le 16} \xi_{3,i} H_{3,i} \; + \; \mu \sum_{1\le j\le 4} \xi_{1,j} H_{1,j}.$$

Since multiplying with a scalar does not change the zero set of f, we may assume that $\lambda + \mu = 1$. Then we may focus on the 1-parameter family of measures

$$\mathbb{P}_\lambda := \mathbb{P}_{\lambda,1-\lambda}, \quad \lambda \in (0,1).$$

For each $\lambda \in (0,1)$, we have sampled 10^5 cubics under $\mathbb{P}_\lambda$ and estimated the probability distribution $\pi^{(\lambda)} = \left(\pi_3^{(\lambda)}, \pi_7^{(\lambda)}, \pi_{15}^{(\lambda)}, \pi_{27}^{(\lambda)}\right)$ where

$$\pi_k^{(\lambda)} = \mathbb{P}_\lambda\left(\#\{\mathbb{R}\text{-rational lines in } \mathcal{S}(f)\} = k\right) \quad \text{for } k = 3, 7, 15, 27.$$

The result is a curve

$$(0,1) \to \Delta_3, \quad \lambda \mapsto \pi^{(\lambda)} = \left(\pi_3^{(\lambda)}, \pi_7^{(\lambda)}, \pi_{15}^{(\lambda)}, \pi_{27}^{(\lambda)}\right),$$

in the 3-dimensional Δ_3 simplex:

$$\Delta_3 := \left\{(p_1, p_2, p_3, p_4) \in \mathbb{R}^4_{\ge 0} \colon \; p_1 + p_2 + p_3 + p_4 = 1\right\}.$$

The 3-dimensional simplex Δ_3 is contained in the affine hyperplane in $\mathbb{R}^4$ given by $\{x_1+x_2+x_3+x_4 = 1\}$. We pick an orthogonal basis (e_1, e_2, e_3) of the hyperplane $\{x_1+x_2+x_3+x_4 = 0\}$ and write expand the curve $\lambda \mapsto \pi^{(\lambda)}$ as follows

$$\pi^{(\lambda)} = \frac{1}{4}(1,1,1,1) \; + \; y_1(\lambda)e_1 + y_2(\lambda)e_2 + y_3(\lambda)e_3.$$

Figure 1 depicts the curve $\lambda \mapsto (y_1(\lambda), y_2(\lambda), y_3(\lambda))$.

A particular case of interest is the distribution $\pi^{(1/3)}$ obtained when we sample from the Kostlan distribution $\mathbb{P}_{1/3}$. To estimate $\pi^{(1/3)}$, we sampled 10^6 cubics $f \in V$ under $\mathbb{P}_{1/3}$ and counted the number of lines on $\mathcal{S}(f)$. Table 5 summarizes the estimate we obtained for $\pi^{(1/3)}$.

We note that our experiments corroborate the following results of [21] and [20]:

Theorem 4.2 ([21, Theorem 5]) *The average number of real lines on a random cubic surface in* $\mathbb{RP}^3$ *under the Kostlan distribution* $\mathbb{P}_{\frac{1}{3}}$ *is* $6\sqrt{2}-3 \simeq 5.48528$.

Theorem 4.3 ([20, Theorem 1]) *The expected number of real lines on a random real cubic surfaces under the measure* $\mathbb{P}_\lambda$ *is*

$$E_\lambda = \frac{9(8\lambda^2+(1-\lambda)^2)}{2\lambda^2+(1-\lambda)^2}\left(\frac{2\lambda^2}{8\lambda^2+(1-\lambda)^2}-\frac{1}{3}+\frac{2}{3}\sqrt{\frac{8\lambda^2+(1-\lambda)^2}{20\lambda^2+(1-\lambda)^2}}\right).$$

Our computations were carried out using the Julia package HomotopyContinuation.jl [22]. Code and data are available at (1).

4.1 Interpretation of the results

Notice that, as $\lambda \to 0$, the distribution $\pi^{(\lambda)}$ converges to the vertex $\pi^{(0)} = (1, 0, 0, 0)$. So when λ is small, we can only hope to see surfaces with 3 real lines under $\mathbb{P}_\lambda$; see [20, Proposition 2]. As $\lambda \to 1$ we see higher number of lines with bigger probabilities. The curve in Fig. 1 depicts of $\pi^{(\lambda)}$ as λ ranges in $(0, 1)$. We observe that the limit distribution as $\lambda \to 1$ (a random cubic surface under the measure $\mathbb{P}_1$ supported on $\mathcal{H}_3$ is smooth with probability 1) lies in the interior of the 3-dimensional probability simplex but is not easy to determine explicitly.

This poses the following question:

Question 4.4 Is the curve $(\pi^{(\lambda)})_{\lambda\in(0,1)}$ algebraic, and if yes, what are the equations defining it?

5 Galois groups

Let S be a smooth cubic surface defined over a perfect field K, and let L be the field of definition of the 27 lines on S. Then the extension L/K is Galois, and the Galois group $\mathrm{Gal}(L/K)$ is a subgroup of $W(E_6)$, the Weyl group of order $51840 = 2^7 \cdot 3^4 \cdot 5$. A natural question to ask is which subgroups of $W(E_6)$ can be realized in this way. The answer depends, of course, on the base field K. Elsenhans–Jahnel proved the following:

Theorem 5.1 ([11, Theorem 0.1]) *All subgroups of* $W(E_6)$ *can be realized when* $K = \mathbb{Q}$.

Let us now consider the case where K is a non-archimedean local field of characteristic 0; such a field is isomorphic to a finite extension of $\mathbb{Q}_p$ for some prime p. In this case, there are certain constraints on the possible Galois groups that can arise. Let us make this more precise. Let L be a finite Galois extension of K with Galois group G, and consider (the first parts of) the ramification group filtration on G:

$$G \supseteq G_0 \supseteq G_1.$$

Here, G_0 and G_1 are the inertia and ramification groups, respectively. It is well-known that

- the group G_1 is a p-group,
- the quotient G/G_0 must be cyclic, and
- the quotient G_0/G_1 must be cyclic of order coprime to p

Table 6 The 24107 abelian groups that appeared in our sample for $p = 5$

C_2	C_3	C_4	C_5	C_6	C_8	C_9	C_{10}	C_{12}	C_2^2	C_4^2	$C_2 \times C_4$	$C_2 \times C_6$
242	246	2117	2042	5154	2928	2126	2647	3613	475	30	1193	1294

Table 7 The 893 non-abelian groups that appeared in our sample for $p = 5$

D_5	D_6	D_{10}	F_5	S_3	OD_{16}	$C_2 \times F_5$	$C_3 \times S_3$	$C_4 \times S_3$	$C_6 \times S_3$	$C_3 \rtimes C_4$	$C_3 \rtimes C_8$
2	98	2	13	204	7	21	323	4	47	170	2

Table 8 The 24922 abelian groups that appeared in our sample for $p = 7$

C_2	C_3	C_4	C_5	C_6	C_8	C_9	C_{10}	C_{12}	C_2^2	C_3^2	$C_2 \times C_4$	$C_2 \times C_6$	$C_3 \times C_6$
284	286	2263	2149	5513	3001	2303	2835	3931	333	42	787	1043	152

Table 9 The 78 non-abelian groups that appeared in our sample for $p = 7$

D_4	F_5	$C_2 \times F_5$	$C_4 \rtimes C_4$
50	7	3	18

(see, for example, [23, Chapter IV]). In addition to these constraints, we also have the following:

Proposition 5.2 *The Galois group G is solvable.*

Proof This well-known result follows from the fact that every p-group is nilpotent. □

Because of this proposition, the statement in Theorem 5.1 clearly can not be true for K since $W(E_6)$ has non-solvable subgroups. Indeed, there are precisely 19 non-solvable subgroups of $W(E_6)$ up to conjugation, and they can be computed using the following Magma code:

```
R_E6 := RootDatum(“E6”);
Cox_E6 := CoxeterGroup(R_E6);
WE6 := StandardActionGroup(Cox_E6);
print NonsolvableSubgroups(WE6);
```

These observations pose the following question:

Question 5.3 Which subgroups of $W(E_6)$ can arise as Galois groups of lines on smooth cubic surfaces over $\mathbb{Q}_p$? How does this list depend on p?

The results of our experiments suggest that Galois groups should be quite small and they are usually abelian. Tables 6, 7 (resp. Tables 8, 9) summarize the Galois groups[5] obtained for a sample of 25000 surfaces over $\mathbb{Q}_5$ (resp. $\mathbb{Q}_7$) sampled from the Haar measure, and the number of times they occurred. Our code can be found at (1).

[5] In the tables, we use the notation provided by Magma. In particular, F_q is the Frobenius group $\mathbb{F}_q \rtimes \mathbb{F}_q^\times$, and OD_{2^k} is the “other-dihedral” group $C_{2^{k-1}} \rtimes C_2$ with C_2 acting as $2^{k-2} + 1$.

Remark 5.4 Notice that the Galois groups that appeared for $p = 5$ are more complicated. This is expected since the prime 5 divides the order of the Weyl group $W(E_6)$. We have also tried to make the same computation for the other primes with the same property, namely 2 and 3. Determining Galois groups using the Magma's `GaloisGroup()` function, however, was significantly slower in these cases. This means that more interesting groups show up for the primes 2 and 3.

We see that while the generic Galois group for surfaces over $\mathbb{Q}$ is $W(E_6)$, over the p-adics there is no "generic" group but rather a list of "small" groups that can occur with positive probability. Question 5.3 would be interesting to answer in future work. We conclude this section with the following remark.

Remark 5.5 Elsenhans–Jahnel also showed that some Galois groups are possible over any field of odd characteristic as long as a field extention with that group exists; see [24, 25]. This might give some information on what happens over $\mathbb{Q}_p$ by lifting surfaces over $\mathbb{F}_p$ to $\mathbb{Z}_p$.

Acknowledgements We thank Claus Fieker, Tom Fisher, Stevan Gajović, Marta Panizzut, Emre Sertöz and Bernd Sturmfels for valuable discussions. We also thank Avinash Kulkarni, Daniel Loughran and Jean Pierre Serre for their valuable comments on earlier versions of this manuscript. Finally, we thank the referee for suggestions that improved the paper.

References

1. Hartshorne, R.: Algebraic geometry. Graduate texts in mathematics, No. 52. Springer-Verlag, New York-Heidelberg, (1977)
2. Manin, Y.I.: Cubic forms: algebra, geometry, arithmetic. North-Holland Mathematical Library, Vol. 4. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, (1974). Translated from the Russian by M. Hazewinkel
3. Dolgachev, I.V.: Classical algebraic geometry. Cambridge University Press, Cambridge (2012)
4. Cayley, A.: On the triple tangent planes of surfaces of the third order. Camb. Dublin Math. J **4**, 118–138 (1849)
5. Schläfli, L.: An attempt to determine the twenty-seven lines upon a surface of the third order, and to divide such surfaces into species in reference to the reality of the lines upon the surface. Q. J. Math **2**(55), 110 (1858)
6. Segre, B.: Le rette delle superficie cubiche nei corpi commutativi. Boll. Un. Mat. Ital. **3**(4), 223–228 (1949)
7. McKean, S.: Rational lines on smooth cubic surfaces, (2022). Preprint available at arXiv:2101.08217v3
8. Loughran, D., Trepalin, A.: Inverse Galois problem for del Pezzo surfaces over finite fields. Math. Res. Lett. **27**(3), 845–853 (2020)
9. Dickson, L.E.: Projective classification of cubic surfaces modulo 2. Ann. Math. (2), **16**(1-4), 139–157 (1914/15)
10. Kostlan, E.: On the distribution of roots of random polynomials. In: From topology to computation: proceedings of the smalefest (Berkeley, C, 1990), pp. 419–431. Springer, New York, (1993)
11. Elsenhans, A.-S., Jahnel, J.: Moduli spaces and the inverse Galois problem for cubic surfaces. Trans. Am. Math. Soc. **367**(11), 7837–7861 (2015)
12. Bosma, W., Cannon, J., Playoust, C.: The Magma algebra system. I. The user language. J. Symbolic Comput. **24**(3–4), 235–265 (1997)
13. Bezanson, J., Edelman, A., Karpinski, S., Julia, V.B.S.: A fresh approach to numerical computing. SIAM Rev. **59**(1), 65–98 (2017)
14. Grayson, D.R., Stillman, M.E.: Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/
15. Panizzut, M., Sertöz, E.C., Sturmfels, B.: An octanomial model for cubic surfaces. Matematiche (Catania) **75**(2), 517–536 (2020)
16. Ren, Q., Sam, S.V., Sturmfels, B.: Tropicalization of classical moduli spaces. Math. Comput. Sci. **8**(2), 119–145 (2014)

17. Maazouz, Y.E., Lerario, A.: A nonarchimedean version of Kostlan's theorem, (2022). Preprint available at arXiv:2209.13634
18. Manssour, R.A.E., Lerario, A.: Probabilistic enumerative geometry over p-adic numbers: linear spaces on complete intersections. Annales Henri Lebesgue **5**, 1329–1360 (2022)
19. Bhargava, M., Cremona, J., Fisher, T., Gajović, S.: The density of polynomials of degree n over $\mathbb{Z}_p$ having exactly r roots in $\mathbb{Q}_p$. Proc. Lond. Math. Soc, (2021)
20. El Manssour, R.A., Belotti, M., Meroni, C.: Real lines on random cubic surfaces. Arnold Math. J. **7**(4), 541–559 (2021)
21. Basu, S., Lerario, A., Lundberg, E., Peterson, C.: Random fields and the enumerative geometry of lines on real and complex hypersurfaces. Math. Ann. **374**(3), 1773–1810 (2019)
22. Breiding, P., Timme, S.: HomotopyContinuation.jl: A package for homotopy continuation in Julia. In: International congress on mathematical software, pp. 458–465. Springer, (2018)
23. Serre, J.P.: Local fields, volume 67 of Graduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1979. Translated from the French by Marvin Jay Greenberg
24. Elsenhans, A.-S., Jahnel, J.: On plane quartics with a Galois invariant Cayley octad. Eur. J. Math. **5**(4), 1156–1172 (2019)
25. Elsenhans, A.-S., Jahnel, J.: Plane quartics with a Galois-invariant Steiner hexad. Int. J. Number Theory **15**(5), 1075–1109 (2019)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.