Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Probabilities of incidence between lines and a plane curve over finite fields

M. Makhul, J. Schicho, M. Gallet

RICAM-Report 2017-40

PROBABILITIES OF INCIDENCE BETWEEN LINES AND A PLANE CURVE OVER FINITE FIELDS

MEHDI MAKHUL*, JOSEF SCHICHO, AND MATTEO GALLET*,°

Abstract. We study the probability for a random line to intersect a given plane curve, defined over a finite field, in a given number of points. In particular, we focus on the limits of these probabilities under successive finite field extensions. The main tools we use are the Lang-Weil bound for the number of rational points of an algebraic variety and a geometric interpretation of the Galois group of a curve. Supposing absolute irreducibility for the curve, we prove the existence of these limits, and under a mildly stronger condition we provide an explicit formula for them, depending only on the degree of the curve.

1. Introduction

What is the probability that a random line in the (affine or projective) plane intersects a curve of given degree in a given number of points? More precisely, what happens if we consider a finite field with q elements as base field, and then we ask the same question for a field with $q^2, q^3, \dots, q^N$ elements, analyzing how these probabilities behave as N goes to infinity? In this work we investigate this problem by means of algebro-geometric techniques. In recent years, the interplay between combinatorial problems and algebraic techniques has become more and more common, and has been revealing to be extremely fruitful. Here we refer in particular to the area called *combinatorial geometry*, which deals with, among others, distance or intersection relations between finite sets of geometric objects such as points, lines, or circles (seee [Tao14]). In the last decade, algebraic geometry and algebraic topology helped solving several outstanding problems and conjectures in this area. Amongst the most prominent of such problems, one can mention the *distinct distance problem* (see [GK15]), the *Kakeya problem over finite fields* (see [Dvi09] and later improvements in [SS08] and [DKSS13]) and the *Dirac-Motzkin conjecture* (see [GT13]). For a nice survey about these topics, see [Tao14].

A motivation for the problem we investigate in our paper comes from the famous *Sylvester-Gallai theorem*. Sylvester posed it as a question in [Syl93], which was raised again by Erdös in [EBW+43] and later solved by Melchior (see [Mel41]) and Gallai (see [Gal44]). Given a set of points in the affine plane, a line is called *ordinary* if it passes through exactly two of them.

* Supported by the Austrian Science Fund (FWF): W1214-N15, Project DK9.

° Supported by the Austrian Science Fund (FWF): P26607 and P25652.

Theorem (Sylvester-Gallai)**.** *Suppose that P is a finite set of points in the real plane, not all on a line. Then P admits an ordinary line.*

This theorem is clearly false if we work over finite fields, since in this case we can pick P to be the the whole plane.

In the same circle of ideas, in the early 60s, Erdös asked the following question: is it possible for a set of points in the real plane to contain many collinear four-tuples, but to contain no five points on a line? Here by "many" we mean a number which is quadratic in terms of the cardinality of the set of points. In [SS13], Solymosi pointed out that, if we weaken the quadratic condition, then the problem has a affirmative solution. Given a set P of points in the plane, a line is called *k-rich*, if it contains precisely k points of P. For example, a 2-rich line is an ordinary line. Then, Solymosi's theorem reads as:

Theorem (Solymosi)**.** *For any $k \geq 4$, there is a positive integer n_0 such that for $n > n_0$ there exists $P \subset \mathbb{R}^2$ such that there are at least $n^{2-\frac{c}{\sqrt{\log n}}}$ k-rich lines, but no $k+1$-rich lines. Here, $c = 2\log(4k+1)$.*

A recent outstanding result of Green and Tao (see [GT13]) gives an almost complete description of the structure of sets with few ordinary lines in the real plane. In the same paper, the authors also proved the Dirac-Motzkin conjecture and a less known problem, referred in the literature as the *orchard problem*.

Theorem (Dirac-Motzkin conjecture)**.** *There exists a number $n_0 \in \mathbb{N}$ such that the following holds. Let $n \in \mathbb{N}$ be such that $n \geq n_0$. Suppose that P is a finite set of n points in the real plane, not all on a line. Then P spans at least $\frac{n}{2}$ (respectively, $\lfloor \frac{3n}{4} \rfloor$) ordinary lines if n is even (respectively, if n is odd).*

Theorem (Orchard problem)**.** *There exists a number $n_0 \in \mathbb{N}$ such that the following holds. Let $n \in \mathbb{N}$ be such that $n \geq n_0$. Suppose that P is a finite set of n points in the real plane. Then there are no more than $\lfloor \frac{n(n-3)}{6} + 1 \rfloor$ lines that are 3-rich for P.*

Our work is inspired by these results and conjectures. Here we consider an algebraic plane curve C of degree d over a finite field $\mathbb{F}_q$ with q elements, where q is a prime power, namely the set of points in the projective plane $\mathbb{P}^2(\mathbb{F}_q)$ that are zeros of a homogeneous trivariate polynomial of degree d. Given such a curve, we can define the probability for a line in $\mathbb{P}^2(\mathbb{F}_q)$ to intersect it in exactly k points. Notice that here we consider the mere set-theoretic intersection: no multiplicities are taken into account. We can then consider the same kind of probability, keeping the same curve C — namely, the same trivariate polynomial — but changing the base field from $\mathbb{F}_q$ to $\mathbb{F}_{q^2}$, $\mathbb{F}_{q^3}$ and so on. In this way, for every $N \in \mathbb{N}$ we define the numbers $p_k^N(C)$, namely the probability for a line in $\mathbb{P}^2(\mathbb{F}_{q^N})$ to intersect C in exactly k points. If the limit as N goes to infinity of the sequence $\left(p_k^N(C)\right)_{N\in\mathbb{N}}$ exists, we denote this number by $p_k(C)$. This is the first main result of our paper (which can be interpreted as a finite field version of the classical Cauchy-Crofton formula).

Theorem 1.1. *Let C be an absolutely irreducible plane algebraic curve of degree d over $\mathbb{F}_q$, where q is a prime power. Then for all $k \in \{0, \dots, d\}$ the numbers $p_k(C)$ exist, and it holds $p_0(C) + p_1(C) + \cdots + p_d(C) = 1$.*

Here, by *absolutely irreducible* we mean that the curve is irreducible over the algebraic closure of its field of definition. The two main tools we use to prove Theorem 1.1 are the *Lang-Weil bound* for the number of points of an algebraic variety over a finite field (Theorem 2.3) and a geometric description of the so-called *Galois group of a curve* (Definition 2.7).

It is possible to obtain a variant of Theorem 1.1 for varying characteristic. More precisely, instead of proving the existence of limits of probabilities where the base fields have higher and higher cardinality, but fixed characteristic, one shows the existence of limits of probabilities where the base fields have increasing characteristic. However, we do not pursue this line of research in this work.

If we suppose that the curve C has *simple tangency* — namely that there exists a line whose intersection with C consists of simple intersections except for one, which is a double intersection — then we can provide explicit formulas for the probabilities $p_k(C)$.

Theorem 1.2. *Let C be an absolutely irreducible plane algebraic curve of degree d over $\mathbb{F}_q$, where q is a prime power. Suppose that C has simple tangency. Then for every $k \in \{0, \dots, d\}$ we have*

$$p_k(C) = \sum_{s=k}^{d} \frac{(-1)^{k+s}}{s!} \binom{s}{k}.$$

In particular, $p_{d-1}(C) = 0$ and $p_d(C) = 1/d!$.

A consequence of this theorem is that the question by Erdös we already mentioned — asking whether there exists a set of points in the plane containing many collinear four-tuples, but not containing any five points on a line — has a positive answer for the plane over a finite field (see Corollary 5.9).

The rest of the paper is structured as follows. Section 2 introduces some preliminary results, namely the Lang-Weil bound for the number of points of a variety over a finite field (Subsection 2.1) and some known facts about Galois groups of plane curves (Subsection 2.2). To prove Theorem 1.1 and Theorem 1.2, we cast the notions of intersection probabilities for lines and plane curves into a more general framework, namely the one of morphisms between smooth varieties with suitable hypotheses (see Equation $(*)$). In Section 3, we prove, in this more general context, the equivalence between two notions of Galois group, one more "geometric" and one more "algebraic". Section 4 recalls the notion of simple tangency and its consequences in terms of Galois groups. In Section 5, we prove the analogues of Theorems 1.1 and 1.2 in the setting of morphisms, which imply our desired result.

2. Preliminaries

2.1. **Lang-Weil bound.** One of the main tools we use in our work is the so-called *Lang-Weil bound* for the number of points of a variety over a finite field (see [LW54, Theorem 1]). For a nice exposition of this result, see Terence Tao's blog[1]. Let $\mathbb{F}$ be a field and consider an affine algebraic variety V defined over $\mathbb{F}$. This means that V is the set of common zeros in $\mathbb{F}^n$ of finitely many polynomials $P_1, \dots, P_r \in \mathbb{F}[x_1, \dots, x_n]$. To a variety V defined over $\mathbb{F}$ we can then associate the ideal $I(V)$ of all polynomials $P \in \mathbb{F}[x_1, \dots, x_n]$ that vanish at all points of V. For any extension of fields $\mathbb{F} \subset \mathbb{K}$, we denote by $V(\mathbb{K})$ the set of common zeros in $\mathbb{K}^n$ of the polynomials in the ideal $I(V)$, considered now as an ideal in $\mathbb{K}[x_1, \dots, x_n]$. One says that a variety $V \subset \mathbb{F}^n$ is *irreducible* if $I(V)$ is prime in $\mathbb{F}[x_1, \dots, x_n]$. For our considerations we will need a stronger notion of irreducibility, which we introduce in the following definition.

Definition 2.1. We say that an affine variety V over a field $\mathbb{F}$ is *absolutely irreducible* if the ideal $I(V)$ is prime in $\overline{\mathbb{F}}[x_1, \dots, x_n]$, where $\overline{\mathbb{F}}$ is an algebraic closure of $\mathbb{F}$. This is equivalent to the fact that $V(\overline{\mathbb{F}})$ is irreducible.

Definition 2.2. We say that an affine variety $V \subset \mathbb{F}^n$ defined by polynomials $P_1, \dots, P_r$ has *complexity* M if $n, r \leq M$ and $\deg(P_i) \leq M$ for all $i \in \{1, \dots r\}$.

Theorem 2.3 (Lang-Weil bound)**.** *Let V be an absolutely irreducible variety over a finite field $\mathbb{F}$ of complexity at most M. Then*

$$|V(\mathbb{F})| = \left(1 + O_M(|\mathbb{F}|^{-\frac{1}{2}})\right) |\mathbb{F}|^{\dim(V)}.$$

By writing $O_M(|\mathbb{F}|^{-\frac{1}{2}})$ we mean that there exists a nonnegative constant δ_M depending on M, but not on V, such that

$$\left(1 - \delta_M |\mathbb{F}|^{-\frac{1}{2}})\right) |\mathbb{F}|^{\dim(V)} \leq |V(\mathbb{F})| \leq \left(1 + \delta_M |\mathbb{F}|^{-\frac{1}{2}})\right) |\mathbb{F}|^{\dim(V)}.$$

Using an inclusion-exclusion argument, one obtains by induction on the dimension:

Corollary 2.4. *Let V be a variety over a finite field $\mathbb{F}$ of complexity at most M. Then*

$$|V(\mathbb{F})| = \left(c + O_M(|\mathbb{F}|^{-\frac{1}{2}})\right) |\mathbb{F}|^{\dim(V)},$$

where c is the number of irreducible components of $V(\overline{\mathbb{F}})$.

All the considerations and results we stated so far hold also for projective varieties defined over finite fields. By a *projective variety* defined over a field $\mathbb{F}$ we mean the set of common zeros in the projective space $\mathbb{P}^n(\mathbb{F})$ of finitely many *homogeneous* polynomials $P_1, \dots, P_r \in \mathbb{F}[x_0, \dots, x_n]$. From now on, all the varieties we consider are projective, or are open subsets of projective varieties.

[1] https://terrytao.wordpress.com/2012/08/31/the-lang-weil-bound/

2.2. **Galois group of a plane curve.** The aim of this section is to recall a construction (see [Rat87]) that associates a Galois group to a plane algebraic curve. We will see in the following sections that this group determines the irreducibility of certain surfaces; this will be the key to derive a formula for the probabilities we are interested in.

Let q be a prime power, namely $q = p^r$ for some prime number p. We denote by $\mathbb{F}_q$ the finite field with q elements. Let C be an absolutely irreducible algebraic curve in $\mathbb{P}^2(\mathbb{F}_q)$. Define

$$X_1 := \big\{(w, [\ell]) \in C \times \check{\mathbb{P}}^2(\mathbb{F}_q) : w \in \ell\big\} \quad \text{and} \quad X_0 := \check{\mathbb{P}}^2(\mathbb{F}_q).$$

Here $\check{\mathbb{P}}^2(\mathbb{F}_q)$ denotes the *dual* projective plane, namely the projective plane whose points are in bijection with the lines in $\mathbb{P}^2\big(\mathbb{F}_q\big)$. For a line $\ell \subset \mathbb{P}^2(\mathbb{F}_q)$, we write $[\ell]$ for the corresponding point in $\check{\mathbb{P}}^2(\mathbb{F}_q)$. The correspondence is given by

$$\check{\mathbb{P}}^2(\mathbb{F}_q) \ni (a : b : c) \quad \longleftrightarrow \quad \big\{(x : y : z) \in \mathbb{P}^2(\mathbb{F}_q) \,:\, ax + by + cz = 0\big\}.$$

Definition 2.5. Using the notation we have already introduced, we define the map $\pi\colon X_1 \longrightarrow X_0$ to be the projection onto the second component.

Since X_0 is irreducible, we can define its *function field*, denoted $K(X_0)$. This is the field of equivalence classes of morphisms $\varphi\colon U \longrightarrow \mathbb{F}_q$, where U is any (Zariski) open subset of X_0; two morphisms are considered equivalent if they agree on a non-empty open subset. Consider the projection $\rho\colon X_1 \longrightarrow C$ on the first component: its fibers are lines in $\check{\mathbb{P}}^2\big(\mathbb{F}_q\big)$, because the elements in the fiber over a point $w \in C$ correspond to the lines in $\mathbb{P}^2\big(\mathbb{F}_q\big)$ through w. Hence all these fibers are irreducible varieties of the same dimension. This implies that X_1 is irreducible by [Sha13, Chapter 1, Section 6.3, Theorem 1.26]; its function field is denoted $K(X_1)$.

Lemma 2.6 (see [Rat87, Definition 1.3])**.** *The projection $\pi\colon X_1 \longrightarrow X_0$ is a quasi-finite dominant separable morphism of degree d.*

Because of Lemma 2.6, the induced map $\pi^*\colon K(X_0) \longrightarrow K(X_1)$ between fields of rational functions realizes $K(X_1)$ as a finite separable extension of $K(X_0)$ of degree d. By the primitive element theorem, the field $K(X_1)$ is generated over $K(X_0)$ by a single rational function $h \in K(X_1)$ satisfying $P(h) = 0$ for an irreducible monic polynomial P over $K(X_0)$ of degree d.

Definition 2.7 (Galois group, see [Rat87, Definition 1.3])**.** Using the notation just introduced, we define the *Galois group* $\mathrm{Gal}(C)$ of C to be the Galois group of a splitting field of the polynomial P over $K(X_0)$. In other words, $\mathrm{Gal}(C)$ is the Galois group of a Galois closure (see [Row06, Remark 4.77]) of the field extension $K(X_0) \hookrightarrow K(X_1)$. The group $\mathrm{Gal}(C)$ is independent of the choice of h and it can be regarded as a subgroup of the permutation group S_d of the roots of P.

3. Galois theory for maps

In this section we associate a Galois group to a morphism (satisfying certain conditions) between two irreducible smooth varieties. We show (Proposition 3.6) that

this concept admits a geometric counterpart, and we use this characterization in the next section.

Note. All varieties considered in this section are supposed to be defined over an algebraically closed field.

For technical reasons, we develop the theory for a special class of morphisms, namely the one of *étale* maps. They model, in the algebraic setting, the notion of "local isomorphism" for the analytic topology. Recall that, in differential geometry, a smooth map between two smooth manifolds is a *local diffeomorphism* if it induces an isomorphism at the level of tangent spaces. For an affine variety X cut out by polynomials $P_1, \dots, P_r$, one defines the *tangent cone* $C_x(X)$ of X at the origin as the variety defined by the homogeneous parts of minimal degree of each of the polynomials $P_1, \dots, P_r$; the tangent cone at any other point is obtained by translating it to the origin and by applying the previous definition. The tangent cone plays for étale morphisms the role played by the tangent space for local diffeomorphisms:

Definition 3.1 (Étale map)**.** A morphism $f\colon X \longrightarrow Y$ between irreducible varieties is *étale at a point* $x \in X$ if f induces an isomorphism between the tangent cones $C_x(X)$ and $C_{f(x)}(Y)$. A map is called *étale* if it is étale at every point.

Definition 3.2. Let $f\colon X \longrightarrow Y$ be a finite separable dominant étale map of degree d between two irreducible smooth varieties. We define the *Galois scheme* (see [Vak06, Section 3]) of f as

$$\mathrm{GS}(f) := \big\{(x_1, \dots, x_d) \in X^d \,:\, f(x_1) = \cdots = f(x_d),\ x_i \neq x_j \text{ for all } i \neq j\big\}.$$

Notice that the Galois scheme is the fiber product of d copies of the map f minus the big diagonal. Because of this, and since f is a finite map, we have

$$\dim \mathrm{GS}(f) \;=\; \dim X \;=\; \dim Y. \tag{1}$$

There is an induced map $F\colon \mathrm{GS}(f) \longrightarrow Y$, sending $(x_1, \dots, x_d)$ to $f(x_1)$, which is dominant of degree $d!$.

Because of Definition 3.2, in the following we will consider often morphisms between varieties satisfying the following condition:

$$(*) \qquad \begin{array}{c}\text{the morphism is a finite separable dominant étale map of degree } d\\ \text{between smooth absolutely irreducible varieties.}\end{array}$$

We have a natural action of the symmetric group S_d on $\mathrm{GS}(f)$ given by

$$\sigma \cdot (x_1, \dots, x_d) := \big(x_{\sigma(1)}, \dots, x_{\sigma(d)}\big) \quad \text{for evey } \sigma \in S_d.$$

Lemma 3.3. *Let $f\colon X \longrightarrow Y$ be a morphism satisfying $(*)$. Then, for any pair of irreducible components Z and Z' of $\mathrm{GS}(f)$ there exists an element $\tau \in S_d$ such that $\tau \cdot Z = Z'$. Namely, the action of S_d on $\mathrm{GS}(f)$ is transitive on irreducible components.*

Proof. Since fiber products of étale maps are étale (see [Sta17, Tag 03PA, Proposition 53.26.2]), the morphism F is étale. Moreover, by construction F is dominant, and it is finite, since the fiber product of finite morphisms is finite. Hence F is

surjective. Since Y is irreducible and smooth and F is étale, then $\mathrm{GS}(f)$ is smooth (see [Sta17, Tag 03PA, Proposition 53.26.2]) and equidimensional (namely, all of its irreducible components have the same dimension, see [Har77, Corollary 9.6] and [Har77, Theorem 10.2] taking into account that an étale morphism is smooth of relative dimension 0). For any two irreducible components Z and Z', consider the restriction maps

$$F_{|_Z} \colon Z \longrightarrow Y \quad \text{and} \quad F_{|_{Z'}} \colon Z' \longrightarrow Y.$$

These maps are also étale, in fact open immersions are étale, and the composition of étale maps is étale. Moreover, both restrictions are dominant. In fact, since $\mathrm{GS}(f)$ is equidimensional, it follows $\dim Z = \dim \mathrm{GS}(f)$, and since F is finite, we get $\dim F(Z) = \dim(Z) = \dim(Y)$ because of Equation (1).

Since f is finite and étale, then for every $y \in Y$ the fiber $f^{-1}(y)$ is constituted of d distinct points (see [GW10, Equation 12.6.2][2]). Hence, for all $y \in Y$ we have $|F^{-1}(y)| = d!$. Fix $y \in Y$ and write $f^{-1}(y) = \{x_1, \ldots, x_d\}$. Therefore

$$F^{-1}(y) = \left\{ (x_{\sigma(1)}, \ldots, x_{\sigma(d)}) \,:\, \sigma \in S_d \right\},$$

where S_d is the symmetric group. Suppose that $F^{-1}(y)$ intersects nontrivially both Z and Z'. It follows that there exists $a \in Z$ and $a' \in Z'$ and an element $\tau \in S_d$ such that $\tau \cdot a = a'$. Since the action of S_d on $\mathrm{GS}(f)$ is algebraic, then the action of any element $\sigma \in S_d$ determines an automorphism of $\mathrm{GS}(f)$; in particular, such an action sends irreducible components to irreducible components. It follows that $\tau \cdot Z = Z'$. Hence, we are left to show that such an element $y \in Y$ exists. This is the case because of the following argument. Define

$$\widetilde{Y}_Z = \left\{ y \in Y \,:\, F^{-1}(y) \cap Z \neq \emptyset \right\} \quad \text{and} \quad \widetilde{Y}_{Z'} = \left\{ y \in Y \,:\, F^{-1}(y) \cap Z' \neq \emptyset \right\}.$$

These two sets are open, and since the restrictions $F_{|_Z}$ and $F_{|_{Z'}}$ are dominant, they are non-empty. Since Y is irreducible, then $\widetilde{Y}_Z \cap \widetilde{Y}_{Z'}$ is open and non-empty. Any point in $\widetilde{Y}_Z \cap \widetilde{Y}_{Z'}$ satisfies the desired requirement. □

Lemma 3.3 shows that the action of the symmetric group S_d on the Galois scheme is transitive on irreducible components. Because of this, the stabilizers of these components are conjugate subgroups of S_d.

Definition 3.4. We define the *geometric Galois group* $\mathrm{Gal}_g(f)$ of a morphism $f \colon X \longrightarrow Y$ of degree d satisfying (∗) to be the stabilizer of any irreducible component of $\mathrm{GS}(f)$ with respect to the action of the symmetric group S_d. This definition is well-posed up to conjugacy in S_d.

It follows that the number of irreducible components of the Galois scheme $\mathrm{GS}(f)$ coincides with the number of cosets of the geometric Galois group $\mathrm{Gal}_g(f)$ in S_d.

Definition 3.5. Let $f \colon X \longrightarrow Y$ be a morphism satisfying (∗). Since f is dominant, it determines a field extension $K(Y) \hookrightarrow K(X)$. We define the *algebraic Galois*

[2]See also the answer by Sandor Kovács at https://mathoverflow.net/questions/86221/fibre-cardinality-of-an-unramified-morphism.

group $\mathrm{Gal}_a(f)$ of f to be the Galois group of the extension $K(Y) \hookrightarrow E$, where E is a Galois closure (see [Row06, Remark 4.77]) of $K(Y) \hookrightarrow K(X)$.

Proposition 3.6. *For every morphism $f\colon X \longrightarrow Y$ of degree d satisfying (∗), the two groups $\mathrm{Gal}_g(f)$ and $\mathrm{Gal}_a(f)$ are isomorphic.*

Proof. Let Z be any component of $\mathrm{GS}(f)$ and realize $\mathrm{Gal}_g(f)$ as the stabilizer of Z. Since the restriction $F_{|_Z}\colon Z \longrightarrow Y$ is dominant (see Lemma 3.3) we have a field inclusion $K(Y) \hookrightarrow K(Z)$.

Claim. $\mathrm{Gal}_g(f) \cong \mathrm{Gal}\big(K(Z)/K(Y)\big)$.

Proof of the claim. Since $\mathrm{Gal}_g(f)$ is the stabilizer of Z, then for every $\sigma \in \mathrm{Gal}_g(f)$ we have an automorphism $\varphi_\sigma\colon Z \longrightarrow Z$, which induces an automorphism $\psi_\sigma\colon K(Z) \longrightarrow K(Z)$ fixing $K(Y)$. We define a group homomorphism

$$\begin{array}{rccc} \Psi\colon & \mathrm{Gal}_g(f) & \longrightarrow & \mathrm{Gal}\big(K(Z)/K(Y)\big) \\ & \sigma & \mapsto & \psi_\sigma \end{array} \tag{2}$$

Let b be the number of components of $\mathrm{GS}(f)$. Notice that the field extension $K(Y) \hookrightarrow K(Z)$ has degree $d!/b$, since $K(Y) \hookrightarrow K(\mathrm{GS}(f))$ has degree $d!$ and all components of $\mathrm{GS}(f)$ differ by the action of an element of S_d. In particular, $|\,\mathrm{Gal}\big(K(Z)/K(Y)\big)| \leq d!/b$. The group homomorphism Ψ is injective because any automorphism which is the identity on an open subset is the identity everywhere. Hence the set $\Psi\big(\mathrm{Gal}_g(f)\big)$ has cardinality $|\,\mathrm{Gal}_g(f)|$. By what we noticed before, the number $|\,\mathrm{Gal}_g(f)|$ equals $|S_d|/b = d!/b$ because the number of components of $\mathrm{GS}(f)$ equals the number of cosets of $\mathrm{Gal}_g(f)$. Hence, Ψ is an isomorphism and $K(Y) \hookrightarrow K(Z)$ is a Galois extension (see [Cha05, Definition 3.2.5]).

Claim. $K(Z)$ is a Galois closure of $K(Y) \hookrightarrow K(X)$.

Proof of the claim. We know already that $K(Y) \hookrightarrow K(Z)$ is a Galois extension. Moreover, the latter factors via $K(Y) \hookrightarrow K(X)$ by considering the projection $Z \longrightarrow X$ on the first factor. The only thing left to prove is that $K(Y) \hookrightarrow K(Z)$ is minimal among Galois extensions of $K(Y) \hookrightarrow K(X)$. Namely, we have to show that if $E \subset K(Z)$ is a field such that the image of the inclusion $K(X) \hookrightarrow K(Z)$ is contained in E, and $K(Y) \hookrightarrow E$ is Galois, then $E = K(Z)$. Define $G := \mathrm{Gal}\big(K(Z)/K(Y)\big)$. By the Galois correspondence, and recalling that the inclusion $K(X) \hookrightarrow K(Z)$ is given by the projection on the first factor, we have

$$\begin{array}{ccccc} G & \supset & \mathrm{Stab}_G(1) & \supset & \{\mathrm{id}\} \\ \updownarrow & & \updownarrow & & \updownarrow \\ K(Y) & \hookrightarrow & K(X) & \hookrightarrow & K(Z) \end{array} \tag{3}$$

where $\mathrm{Stab}_G(1) = \big\{\sigma \in G \,:\, \sigma(1) = 1\big\}$ is the stabilizer of 1 under the action of G on $\{1, \ldots, d\}$. This action is given by the identification of G with $\mathrm{Gal}_g(f)$ provided by the map Ψ in Equation (2). Under this correspondence, the field E is associated to a subgroup $H \subset G$, which is normal since $K(Y) \hookrightarrow E$ is Galois, and which is contained in $\mathrm{Stab}_G(1)$. To conclude, we prove that $H = \{\mathrm{id}\}$, which implies $E = K(Z)$. We start by showing that G acts transitively on $\{1, \ldots, d\}$. If we denote by $G \cdot 1$ the orbit of 1 under G, then by standard results in Galois theory

and taking into account Equation (3), we have

$$|G \cdot 1| = [G : \mathrm{Stab}_G(1)] = [K(X) : K(Y)] = d.$$

This implies that $G \cdot 1 = \{1, \dots, d\}$, showing that the action is transitive. By normality, $H \subset \sigma\, \mathrm{Stab}_G(1)\, \sigma^{-1}$ for any $\sigma \in G$. Since G is transitive on $\{1, \dots, d\}$, it follows

$$H \subset \mathrm{Stab}_G(1) \cap \mathrm{Stab}_G(2) \cap \cdots \cap \mathrm{Stab}_G(d). \tag{4}$$

The right hand side of Equation (4) equals $\{\mathrm{id}\}$, so the claim is proved.

Summing up, the first claim says that $\mathrm{Gal}_g(f) \cong \mathrm{Gal}\big(K(Z)/K(Y)\big)$, and the second claim implies that the latter group is isomorphic to $\mathrm{Gal}_a(f)$. This concludes the proof of the proposition. □

4. Simple tangency

In this section, we cast the notions defined in Section 2.2 into the framework of Galois schemes of morphisms (Corollary 4.4). After that, we recall the notion of simple tangency for a curve and highlight its consequences on Galois groups.

Definition 4.1. For an absolutely irreducible curve $C \subset \mathbb{P}^2(\mathbb{F}_q)$ of degree d, define $\mathcal{V}_C$ to be the set of points in $X_0(\overline{\mathbb{F}}_q) = \check{\mathbb{P}}^2\big(\overline{\mathbb{F}}_q\big)$ such that the restriction of the map $\pi\colon X_1(\overline{\mathbb{F}}_q) \longrightarrow X_0(\overline{\mathbb{F}}_q)$ from Definition 2.5 to $\mathcal{U}_C := \pi^{-1}\big(\mathcal{V}_C\big)$ is étale.

Remark 4.2. Notice that the set $\mathcal{V}_C$ is open and non-empty. In fact, since the map $\pi\colon X_1(\overline{\mathbb{F}}_q) \longrightarrow X_0(\overline{\mathbb{F}}_q)$ is separable, it is enough to ensure that $\pi\colon \mathcal{U}_C \longrightarrow \mathcal{V}_C$ is flat. Now, the locus in the domain where a map is flat is open (see [Sta17, Tag 0398, Theorem 36.15.1,]), and flat maps are open morphisms (see [Sta17, Tag 01U2, Lemma 28.24.9]), so this shows that $\mathcal{V}_C$ is open. The fact that $\mathcal{V}_C$ is non-empty is ensured by the generic flatness result (see [Sta17, Tag 0529, Proposition 28.26.1]).

Lemma 4.3. *Let C be an absolutely irreducible curve of degree d defined over $\mathbb{F}_q$. Then the restriction to $\mathcal{U}_C := \pi^{-1}\big(\mathcal{V}_C\big)$ of the map $\pi\colon X_1\big(\overline{\mathbb{F}}_q\big) \longrightarrow X_0\big(\overline{\mathbb{F}}_q\big)$ from Definition 2.5 is a finite separable dominant étale morphism between smooth absolutely irreducible varieties, namely it satisfies condition $(*)$.*

Proof. We know from Section 2.2 that both X_0 and X_1 are smooth and absolutely irreducible. Since $\mathcal{V}_C$ and $\mathcal{U}_C$ are open and non-empty, the same is true for them. Moreover, π is a quasi-finite separable dominant morphism between projective varieties (Lemma 2.6) and so it is finite. Hence, the same holds for its restriction $\pi_{|_{\mathcal{U}_C}}$. By Remark 4.2, the map is étale, and this concludes the proof. □

By unravelling the definition, in the light of Lemma 4.3 we obtain:

Corollary 4.4. *For an absolutely irreducible curve C defined over $\mathbb{F}_q$, we have* $\mathrm{Gal}(C) \cong \mathrm{Gal}_a\big(\pi_{|_{\mathcal{U}_C}}\big)$.

The interpretation of the Galois group of a curve provided by Corollary 4.4 allows to use Proposition 3.6 and hence to deduce the irreducibility of the Galois scheme when the Galois group is the full symmetric group.

Definition 4.5 (Simple tangency)**.** Let C be an absolutely irreducible curve of degree d in $\mathbb{P}^2(\overline{\mathbb{F}_q})$. We say that C has *simple tangency* if there exists a line $\ell \subset \mathbb{P}^2(\overline{\mathbb{F}_q})$ intersecting C in $d-1$ smooth points of C such that ℓ intersects C transversely at $d-2$ points and has intersection multiplicity 2 at the remaining point.

Remark 4.6. A general curve $C \subset \mathbb{P}^2(\mathbb{F}_q)$ of degree d has simple tangency. In fact, notice that having simple tangency is open condition, therefore it is enough to exhibit a single example in order to obtain the claim. To do that, consider the curve of equation

$$x^2\, P(x,y) + z\, Q(x,y,z) = 0,$$

where P is a homogeneous polynomial with $d-2$ distinct roots in $\overline{\mathbb{F}}_a$ and Q is a homogeneous polynomial of degree $d-1$.

Proposition 4.7 ([Rat87, Proposition 2.1])**.** *Let $C \subset \mathbb{P}^2(\mathbb{F}_q)$ be an absolutely irreducible plane curve of degree d with simple tangency. Then the Galois group* $\mathrm{Gal}(C)$ *of C is the whole symmetric group S_d.*

Corollary 4.8. *Suppose that C is an absolutely irreducible curve in $\mathbb{P}^2(\mathbb{F}_q)$ of degree d with simple tangency. Then, the Galois group* $\mathrm{Gal}_g\big(\pi_{|\mathcal{U}_C}\big) \cong \mathrm{Gal}_a\big(\pi_{|\mathcal{U}_C}\big)$ *is the full symmetric group, and so the Galois scheme* $\mathrm{GS}\big(\pi_{|\mathcal{U}_C}\big)$ *is irreducible.*

Proof. This follows from Corollary 4.4 and Proposition 3.6. □

5. Probabilities of incidence

In this section we define probabilities of intersection between a random line and a given curve in the projective plane over a finite field (Definition 5.1). We then prove the two main results of our paper, namely Theorems 1.1 and 1.2, by showing that their counterparts for morphisms hold (Theorems 5.5 and 5.7).

Definition 5.1 (Probabilities of intersection)**.** Let q be a prime power and let $C \subset \mathbb{P}^2\big(\mathbb{F}_q\big)$ be an absolutely irreducible curve of degree d defined over $\mathbb{F}_q$. For every $N \in \mathbb{N}$ and for every $k \in \{0,\dots,d\}$, the *k-th probability of intersection* $p_k^N(C)$ of lines with C over $\mathbb{F}_{q^N}$ is

$$p_k^N(C) := \frac{\Big|\big\{\text{lines } \ell \subset \mathbb{P}^2(\mathbb{F}_{q^N}) \,:\, |\ell(\mathbb{F}_{q^N}) \cap C(\mathbb{F}_{q^N})| = k\big\}\Big|}{q^{2N}+q^N+1}.$$

Notice that $q^{2N}+q^N+1$ is the number of lines in $\mathbb{P}^2\big(\mathbb{F}_{q^N}\big)$.

The aim of this paper is to prove that the limit as N goes to infinity of the quantities $p_k^N(C)$ exists for every k, and to give a formula for these limits, provided that some conditions on the curve C are fulfilled.

The following result is a direct consequence of Definitions 5.1 and 2.5.

Lemma 5.2. *Let $C \subset \mathbb{P}^2(\mathbb{F}_q)$ be an absolutely irreducible curve of degree d defined over $\mathbb{F}_q$. For every $k \in \{0,\dots,d\}$ we have*

$$p_k^N(C) = \frac{\Big|\big\{[\ell] \in \check{\mathbb{P}}^2(\mathbb{F}_{q^N}) \,:\, |\pi^{-1}([\ell])(\mathbb{F}_{q^N})| = k\big\}\Big|}{q^{2N}+q^N+1}.$$

Via Lemma 5.3 and Definition 5.4 we reduce the problem of computing intersection probabilities for curves to the analogous problem for morphisms.

Lemma 5.3. *Let $C \subset \mathbb{P}^2(\mathbb{F}_q)$ be an absolutely irreducible curve of degree d defined over $\mathbb{F}_q$. Let $\mathcal{V}_C \subset \check{\mathbb{P}}^2(\overline{\mathbb{F}}_q)$ be as in Definition 4.1. For every $N \in \mathbb{N}$ and for every $k \in \{0, \dots, d\}$, define*

$$\widetilde{p}_k^N(C) := \frac{\left|\left\{[\ell] \in \mathcal{V}_C(\mathbb{F}_{q^N}) \,:\, |\pi^{-1}([\ell])(\mathbb{F}_{q^N})| = k\right\}\right|}{|\mathcal{V}_C(\mathbb{F}_{q^N})|}.$$

Then $\lim_{N\to\infty} p_k^N(C)$ exists if and only if $\lim_{N\to\infty} \widetilde{p}_k^N(C)$ exists, in which case the two numbers coincide.

Proof. It is enough to show that the probability for a point to lie in $\mathbb{P}^2(\mathbb{F}_{q^N}) \setminus \mathcal{V}_C(\mathbb{F}_{q^N})$ goes to zero as N goes to infinity. This is a consequence of the Lang-Weil bound (Theorem 2.3). In fact, since $\mathbb{P}^2(\mathbb{F}_{q^N}) \setminus \mathcal{V}_C(\mathbb{F}_{q^N})$ has dimension at most 1:

$$\frac{\left|\mathbb{P}^2(\mathbb{F}_{q^N}) \setminus \mathcal{V}_C(\mathbb{F}_{q^N})\right|}{q^{2N} + q^N + 1} \sim \frac{a\, q^N}{q^{2N}} \to 0\,,$$

where the constant a is the number of irreducible components of $\mathbb{P}^2(\overline{\mathbb{F}}_q) \setminus \mathcal{V}_C(\overline{\mathbb{F}}_q)$. □

Definition 5.4. Let $f\colon X \longrightarrow Y$ be a morphism of degree d defined over $\mathbb{F}_q$, where q is a prime power, satisfying $(*)$. For every $N \in \mathbb{N}$ and for every $k \in \{0, \dots, d\}$, we define the *k-th preimage probability* $p_k^N(f)$ to be

$$p_k^N(f) := \frac{\left|\left\{y \in Y(\mathbb{F}_{q^N}) \,:\, |f^{-1}(y)(\mathbb{F}_{q^N})| = k\right\}\right|}{|Y(\mathbb{F}_{q^N})|}.$$

Notice that if C is an absolutely irreducible algebraic plane curve of degree d, then for every $N \in \mathbb{N}$ and for every $k \in \{0, \dots, d\}$ we have $\widetilde{p}_k^N(C) = p_k^N\left(\pi_{|_{\mathcal{U}_C}}\right)$. Hence, by Lemma 5.3, in order to show the existence of the limits of k-th probabilities of intersections for a curve, it is enough to show the existence of k-th preimage probabilities for morphisms over $\mathbb{F}_q$ satisfying $(*)$.

Theorem 5.5. *Let $f\colon X \longrightarrow Y$ be a morphism of degree d defined over $\mathbb{F}_q$, where q is a prime power, satisfying $(*)$. Then for every $k \in \{0, \dots, d\}$ the limit as N goes to infinity of the sequence $\left(p_k^N(f)\right)_{N\in\mathbb{N}}$ exists.*

Proof. We generalize the construction of the Galois scheme of the morphism f. For every $k \in \{0, \dots, d\}$, define

$$G_k(f) := \left\{(x_1, \dots, x_k) \in X^k \,:\, f(x_1) = \dots = f(x_k),\ x_i \neq x_j \text{ for all } i \neq j\right\}.$$

In particular $G_d(f) = \mathrm{GS}(f)$. As we showed for the Galois scheme, see Equation (1), for every k the variety $G_k(f)$ has the same dimension of X and Y. There is a natural finite morphism $F_k\colon G_k(f) \longrightarrow Y$, the fiber product of f with itself k times. This map has degree $d(d-1)\cdots(d-k+1)$. The main idea of the proof is to compute,

in two different ways, the expected cardinality $\mu_k^N(f)$ of a fiber $F_k^{-1}(y)$ over $\mathbb{F}_{q^N}$, where y is a random element in Y. On one hand,

$$\mu_k^N(f) = \frac{\left|G_k(f)(\mathbb{F}_{q^N})\right|}{\left|Y(\mathbb{F}_{q^N})\right|}\,.$$

On the other hand, we can express $\mu_k^N(f)$ in terms of the preimage probabilities:

$$\mu_k^N(f) = \sum_{s=k}^{d} s(s-1)\cdots(s-k+1)\, p_s^N(f)\,. \tag{5}$$

In matrix form:

$$\begin{pmatrix} \mu_0^N(f) \\ \vdots \\ \mu_d^N(f) \end{pmatrix} = \begin{pmatrix} 1 & * & \cdots & & \cdots & * \\ 0 & 1 & * & & & \vdots \\ \vdots & & \ddots & & & \vdots \\ 0 & \cdots & 0 & k! & * & * \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & & & \cdots & d! \end{pmatrix} \begin{pmatrix} p_0^N(f) \\ \vdots \\ p_d^N(f) \end{pmatrix}. \tag{6}$$

Since the matrix in Equation (6) has non-zero determinant, we can write

$$p_k^N(f) = \sum_{s=0}^{d} \alpha_{k,s}\, \mu_s^N \tag{7}$$

for some numbers $(\alpha_{k,s})_{k,s}$. Using the Lang-Weil bound on Equation (5), we have

$$\mu_k^N \;\sim\; \frac{\delta_k\, q^{N\cdot\dim G_k(f)}}{q^{N\cdot\dim Y}} \qquad \text{as } N\to\infty\,, \tag{8}$$

where δ_k is the number of irreducible components of $G_k(f)(\overline{\mathbb{F}}_q)$. Since $\dim G_k(f) = \dim Y$, we conclude that the limit in Equation (8) exists, and so by Equation (7) also $\lim_{N\to\infty} p_k^N(f)$ exists. □

Corollary 5.6. *Theorem 1.1 holds. In fact, the map $\pi_{|\mathcal{U}_C}$ satisfies the hypotheses of Theorem 5.5, so the numbers $p_k\big(\pi_{|\mathcal{U}_C}\big)$ exist, and we have already proved that this implies that the limits $p_k(C)$ exist.*

Theorem 5.7. *Let $f\colon X \longrightarrow Y$ be a morphism of degree d defined over $\mathbb{F}_q$, where q is a prime power, satisfying (*). Suppose that $\mathrm{Gal}_g(f) \cong \mathrm{Gal}_a(f)$ is the full symmetric group S_d. Then for every $k \in \{0,\dots,d\}$ we have*

$$p_k(f) = \sum_{s=k}^{d} \frac{(-1)^{k+s}}{s!}\binom{s}{k}.$$

In particular, $p_{d-1}(f) = 0$ and $p_d(f) = 1/d!$.

Proof. Since $\mathrm{Gal}_g(f)$ is the full symmetric group, the Galois scheme $\mathrm{GS}(f)$ is absolutely irreducible. Hence, using the notation of the proof of Theorem 5.5, for all $k \in \{0,\dots,d\}$ we have

$$\lim_{N\to\infty} \mu_k^N(f) = \lim_{N\to\infty} \frac{q^{N\cdot\dim G_k(f)}}{q^{N\cdot\dim Y}} = 1. \tag{9}$$

In fact, every variety $G_k(f)$ is an image (under a projection) of $\mathrm{GS}(f) = G_d(f)$, thus is absolutely irreducible and so Equation (9) follows from Equation (8). Again using the notation as in Theorem 5.5, we get

$$\lim_{N\to\infty} p_k^N(f) = \sum_{s=0}^{d} \alpha_{k,s}. \tag{10}$$

Therefore, the statement is proved once we are able to explicitly compute the coefficients $(\alpha_{k,s})_{k,s}$. Recall that $\alpha_{k,s}$ is the (k,s)-entry of the inverse of the matrix M_d appearing in Equation (6). A direct inspection of the matrices M_d shows that they admit the following structure:

$$M_d = \left(\begin{array}{ccc|c} & & & 1 \\ & M_{d-1} & & \vdots \\ & & & d!/1! \\ \hline 0 & \cdots & 0 & d!/0! \end{array}\right).$$

A direct computation shows that

$$M_d^{-1} = \left(\begin{array}{ccc|c} & & & \frac{(-1)^d}{d!}\cdot\binom{d}{0} \\ & M_{d-1}^{-1} & & \vdots \\ & & & \frac{(-1)}{d!}\cdot\binom{d}{d-1} \\ \hline 0 & \cdots & 0 & \frac{1}{d!}\cdot\binom{d}{d} \end{array}\right).$$

Hence

$$\alpha_{k,s} = \frac{(-1)^{k+s}}{s!}\binom{s}{k} \quad \text{for all } k,s \in \{0,\dots,d\}.$$

It follows from Equation (10) that for all $k \in \{0,\dots,d\}$,

$$p_k(f) = \sum_{s=0}^{d} \frac{(-1)^{k+s}}{s!}\binom{s}{k} = \sum_{s=k}^{d} \frac{(-1)^{k+s}}{s!}\binom{s}{k}$$

and so the statement is proved. □

As a consequence of Proposition 4.7 and Theorem 5.7, we obtain:

Corollary 5.8. *Theorem 1.2 holds.*

As we pointed out in the Introduction, one of the consequences of Theorem 1.2 is that a question raised by Erdös concerning 4- and 5-rich lines has a positive answer over finite fields.

Corollary 5.9. *Let C be an absolutely irreducible plane algebraic curve of degree 4 in the plane $\mathbb{P}^2(\mathbb{F}_q)$. By Theorem 2.3, the curve C has $cq + O(\sqrt{q})$ elements for some $c > 0$, and by Theorem 1.2 it has ϵq^2 4-rich lines for some $\epsilon > 0$, after possibly taking a finite extension of the base field, since $p_4(C) > 0$. Hence, if we take P as the set of points of C, then P spans a quadratic number of 4-rich lines, but no five points of P are collinear.*

References

[Cha05] Antoine Chambert-Loir, *A field guide to algebra.*, Springer, New York, 2005.

[DKSS13] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan, *Extensions to the method of multiplicities, with applications to Kakeya sets and mergers*, SIAM J. Comput. **42** (2013), no. 6, 2305–2328.

[Dvi09] Zeev Dvir, *On the size of Kakeya sets in finite fields*, J. Amer. Math. Soc. **22** (2009), no. 4, 1093–1097.

[EBW+43] Paul Erdös, Richard Bellman, Hubert S. Wall, James Singer, and Victor Thébault, *Problem 4065*, Amer. Math. Monthly **50** (1943), 65–66.

[Gal44] Tibor Gallai, *Solution to problem 4065*, Amer. Math. Monthly **51** (1944), 169–171.

[GK15] Larry Guth and Nets H. Katz, *On the Erdös distinct distances problem in the plane*, Ann. of Math. (2) **181** (2015), no. 1, 155–190.

[GT13] Ben Green and Terence Tao, *On sets defining few ordinary lines*, Discrete Comput. Geom. **50** (2013), no. 2, 409–468.

[GW10] Ulrich Görtz and Torsten Wedhorn, *Algebraic geometry I*, Advanced Lectures in Mathematics, Vieweg + Teubner, 2010.

[Har77] Robin Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, no. 52, Springer-Verlag, New York-Heidelberg, 1977.

[LW54] Serge Lang and André Weil, *Number of points of varieties in finite fields*, Amer. J. Math. **76** (1954), 819–827.

[Mel41] Eberhard Melchior, *Über Vielseite der projektiven Ebene*, Deutsche Math. **5** (1941), 461–475.

[Rat87] Jürgen Rathmann, *The uniform position principle for curves in characteristic p*, Math. Ann. **276** (1987), no. 4, 565–579.

[Row06] Louis Rowen, *Graduate Algebra: Commutative View*, American Mathematical Society, Providence, RI, 2006.

[Sha13] Igor R. Shafarevich, *Basic algebraic geometry 1. Varieties in projective space*, third ed., Springer, Heidelberg, 2013.

[SS08] Shubhangi Saraf and Madhu Sudan, *An improved lower bound on the size of Kakeya sets over finite fields*, Anal. PDE **1** (2008), no. 3, 375–379.

[SS13] József Solymosi and Miloš Stojaković, *Many collinear k-tuples with no $k+1$ collinear points*, Discrete Comput. Geom. **50** (2013), no. 3, 811–820.

[Sta17] The Stacks Project Authors, *Stacks Project*, http://stacks.math.columbia.edu, 2017.

[Syl93] James J. Sylvester, *Mathematical Question 11851*, Educational Times **59** (1893), 385–394.

[Tao14] Terence Tao, *Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory*, EMS Surv. Math. Sci. **1** (2014), no. 1, 1–46.

[Vak06] Ravi Vakil, *Schubert induction*, Ann. of Math. (2) **164** (2006), no. 2, 489–512.

(Mehdi Makhul) Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Linz and Research Institute for Symbolic Computation (RISC), Johannes Kepler University, Linz

(Josef Schicho) Research Institute for Symbolic Computation (RISC), Johannes Kepler University, Linz

E-mail address: {mmakhul,jschicho}@risc.jku.at

(Matteo Gallet) Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Linz

E-mail address: matteo.gallet@ricam.oeaw.ac.at