THE DISTRIBUTION OF POINTS ON CURVES OVER FINITE FIELDS IN SOME SMALL RECTANGLES

KIT-HO MAK

ABSTRACT. Let $p$ be a prime. We study the distribution of points on a class of curves $C$ over $\mathbb{F}_p$ inside very small rectangles $B$ for which the Weil bound fails to give nontrivial information. In particular, we show that the distribution of points on $C$ over some long rectangles is Gaussian.

1. Introduction and statements of results

Let $p$ be a large prime, and let $C \subseteq \mathbb{A}^2 := \mathbb{A}^2(\mathbb{F}_p)$ be an absolutely irreducible affine plane curve over $\mathbb{F}_p$ of degree $d > 1$. We identify the affine plane with the set of points with integer coordinates in the square $[0, p - 1]^2$. For a rectangle $B = I \times J \subseteq [0, p - 1]^2$, we define $N_B(C)$ to be the number of (rational) points on $C$ inside $B$. When $B = [0, p - 1]^2$, we will write $N(C) = N_{[0, p - 1]^2}(C)$ for the number of points on $C$. It is widely believed that the points on $C$ are uniformly distributed in the plane. That is,

(1) \[ N_B(C) \sim N(C) \cdot \frac{\text{vol}(B)}{p^2}. \]

In fact, using some standard techniques involving exponential sums, one can show that the classical Weil bound \cite{11} together with the Bombieri estimate \cite{1} imply

(2) \[ N_B(C) = N(C) \cdot \frac{\text{vol}(B)}{p^2} + O(d^2 \sqrt{p} \log^2 p), \]

where the implied constant is absolute. If $f$ and $g$ are two functions of $p$, we write

(3) \[ f = \Omega(g) \]

to denote the function $f/g$ tends to infinity as $p$ tends to infinity. In other words, \cite{3} is equivalent to $g = o(f)$. The main term of \cite{2} dominates the error term when

(4) \[ \text{vol}(B) \gg p^{3/2} \log^{2+\epsilon} p. \]

In those cases \cite{1} holds. A natural and intriguing question that arises is whether \cite{1} continues to hold for smaller boxes $B$. However, very few is known for the number of points $N_B(C)$ in a small $B$. Indeed, given a particular small $B$ that do not satisfy \cite{3}, we do not even know if $B$ contains a point or not.

One way to study $N_B(C)$ for small $B$ is to consider results on average. For instance, Chan \cite{2} considered the number of points on average on the modular hyperbola $xy \equiv c$ modulo an odd number $q$, and showed that almost all (here “almost all” means with probability one) boxes satisfying

\[ \text{vol}(B) \gg O(q^{1/4+\epsilon}) \]
have the expected number of points. Recently, Zaharescu and the author [11] generalizes the result of Chan to all curves over \( \overline{\mathbb{F}}_p \).

Another result of similar sort with only one moving side for \( C \) being the modular hyperbola was obtained by Gonek, Krishnaswami and Sondhi [4]. In our language, they showed that if \( B = (x, x + H) \times \mathcal{J} \) with \( H \) very small and \( \mathcal{J} \) of size comparable to \( p \), then the numbers of points inside \( B \) exhibit a Gaussian distribution when we move the box \( B \) horizontally. A Gaussian distribution is also obtained by Zaharescu and the author [8] in a similar situation. More precisely, we show that under some natural conditions, for \( C, B \) as above, and if at least one of the character \( \chi, \psi \) is nontrivial, the projections of the values of the hybrid exponential sum

\[
S = \sum_{P_i \in C \cap B} \chi(g(P_i)) \psi(f(P_i))
\]

to any straight lines passing through the origin exhibit a Gaussian distribution when we move \( B \) horizontally. We note that when \( C \) is the affine line, a two-dimensional distribution of \( S \) is obtained by Lamzouri [5].

The aim of this paper is continue the study of \( N_B(C) \) for small rectangle \( B \). In particular, we show that for a large class of curves \( C \), the distribution of \( N_B(C) \) for the \( B \) above is Gaussian. Our first step is to study the patterns of points on curves, which is crucial for our study of \( N_B(C) \) and may be of independent interest.

The study of patterns was first introduced by Cobeli, Gonek and Zaharescu [3], where they get results for the distribution of patterns of multiplicative inverses modulo \( p \). We generalize their definition of patterns to curves by viewing the patterns in [3] as patterns on the two coordinates for the curve \( xy = 1 \). For any positive integer \( s \), let \( a = (a_1, \ldots, a_s), b = (b_1, \ldots, b_s) \) be two vectors so that all \( a_i \)'s are coprime to \( p \), and all \( a_i^{-1}b_1, \ldots, a_i^{-1}b_s \) are distinct modulo \( p \). Define an \((a, b)\)-pattern to be an \( s \)-tuple of points \( (P_1, \ldots, P_s) \), where each \( P_i \) is of the form \((a_i x + b_i, y_i)\) for some \( x \). As in [3], we may further restrict all the \( y_i \) to lie in a specific interval \( \mathcal{J} \) as we see fit.

In the case of the modular hyperbola in [3], if \( \mathcal{J} = [0, p-1] \) the number of patterns is just \( p - 1 \), since each \( x \) corresponds to exactly one \( y \) on the curve. However, for a general curve \( C \) and any vector \( a, b \), we do not know a priori that even one pattern exists, since the two coordinates will not in general correspond bijectively. Nevertheless, we are able to estimate the number of patterns for a large class of curves. Let \( P(I, \mathcal{J}) := P_{a, b}(C; I, \mathcal{J}) \) be the number of patterns with \( x \in I \) and all \( y \)-coordinates lie in \( \mathcal{J} \), then we have the following.

**Theorem 1.** Let \( C \) be a plane curve given by the equation \( f(x, y) = 0 \). Let

\[
\pi : C \to \mathbb{A}^1, \ (x, y) \mapsto x
\]

be the projection of \( C \) to the first coordinates over \( \overline{\mathbb{F}}_p \). Suppose there is an \( x \in \overline{\mathbb{F}}_p \) so that \( \pi \) ramifies completely, and let \( a = (a_1, \ldots, a_s), b = (b_1, \ldots, b_s) \) be two vectors so that all \( a_i \)'s are coprime to \( p \), and all \( a_i^{-1}b_1, \ldots, a_i^{-1}b_s \) are distinct modulo \( p \), then

\[
P(I, \mathcal{J}) = |I| \left( \left\lfloor \frac{|\mathcal{J}|}{p} \right\rfloor \right)^s + O(d^{2s} \sqrt{p} \log^{s+1} p).
\]

In the case \( I = [0, p-1] \), the error term can be slightly improved.

\[
P([0, p-1], \mathcal{J}) = p \left( \left\lfloor \frac{|\mathcal{J}|}{p} \right\rfloor \right)^s + O(d^{2s} \sqrt{p} \log^s p).
\]
Note that our estimation for the number of patterns is independent of \(a\) and \(b\).

We are now ready for the study of distribution of \(N_B(C)\) for small \(B\). We fix an interval \(J \subseteq [0, p - 1]\), and let \(N = |J|\). For any \(H > 0\) (which may depends on \(p\)), let \(B_x = (x, x + H] \times J\). From now on, we will assume the following condition.

\[
\text{(7) For any given } x, \text{ there is at most one } y \text{ so that } (x, y) \in C \cap J.
\]

This is the same condition we imposed in [8] when Zaharescu and the author study the distribution of hybrid exponential sums over curves.

Define

\[
M_k(H) = \sum_{x=0}^{p-1} \left( N_{B_x}(C) - \frac{HN}{p} \right)^k
\]

to be the \(k\)-th moment of the number of points in \(C \cap B_x\) about its mean. We also define \(\mu_k(H, P)\) to be the \(k\)-th moment of a binomial random variable \(X\) with parameter \(H\) and \(P\), i.e.

\[
\mu_k(H, P) := E((X - HP)^k) = \sum_{h=1}^{H} \binom{H}{h} P^h (1 - P)^{H-h} (h - HP)^k.
\]

We estimate the moment \(M_k(H)\) using the binomial model with parameter \(H\) and \(N/p\).

**Theorem 2.** Fix a positive integer \(k\). Let \(C\) be a curve satisfying the assumptions in Theorem 1 and the additional condition (7). Set \(B_x, H, N\) as above, we have

\[M_k(H) = p \nu_k(N/p) + O_k(d^2H^k \sqrt{p \log k}).\]

For a fixed \(k\), it is well-known (see Montgomery and Vaughan [10], Lemma 11) that

\[\mu_k(H, P) \ll (HP)^{k/2} + HP\]

uniformly for \(0 \leq P \leq 1\) and \(H = 1, 2, 3, \ldots\). Therefore, Theorem 2 immediately implies the following.

**Corollary 3.** Assumptions as in Theorem 2. For any fixed \(k\), we have

\[M_k(H) \ll_k p(HN/p)^{k/2} + HN/p + d^2H^k \sqrt{p \log k}.\]

**Remark 1.1.** For the case of curves, [8, Theorem 2] gives an upper bound for the second moment of \(N_B(C)\) when \(B\) is allowed to move freely on the plane. That theorem implies \(M_2(H) \ll p \mu_2(H, N/p)\). Since \(\mu_2(H, P) = HP(1 - P)\), Theorem 2 shows that [8, Theorem 2] has the correct main term, and therefore is best possible for the case of curves (with suitable \(H\) and \(N\)).

Let

\[\nu_k = \begin{cases} 1 \cdot 3 \cdot \ldots \cdot (k - 1) & \text{if } k \text{ even,} \\ 0 & \text{if } k \text{ odd,} \end{cases}\]

then (see [10, Lemma 10])

\[\mu_k(H, P) = (\nu_k + o(1))(HP(1 - P))^{k/2}\]

as \(HP(1 - P)\) tends to infinity. From this and Theorem 2 we obtain the following.
Corollary 4. For any fixed \( k \), if \( H = o\left(\frac{d^{3/2k}}{d^2 \log p}\right) \) and \((HN/p)(1 - N/p) \to \infty\) as \( p \) tends to infinity, then

\[ M_k(H) = p(\nu_k + o(1)) \left( \frac{HN}{p} \left( 1 - \frac{N}{p} \right) \right)^{k/2}. \]

In particular, when \( N \sim cp \), \( 0 < c < 1 \) and \( \log H/\log p \to 0 \) as \( p \) tends to infinity, the distribution of \( N_{B_x}(C) \) tends to a Gaussian distribution with mean \( HN/p \) and variance \((HN/p)(1 - N/p)\).

Remark 1.2. If condition (7) does not hold, we may still have Gaussian distribution for the \( N_{B_x}(C) \). For example, if \( C \) is a hyperelliptic curve, and choose \( J \) to be the interval \((-\alpha p, \alpha p]\) for some \( 0 < \alpha < 1/2 \), then generically one \( x \)-coordinate on the curve corresponds to two \( y \)-coordinates. From Corollary 4, we have Gaussian distribution for \( J_1 = [0, \alpha p] \), and also for \( J_2 = [-\alpha p, 0] \), with the same mean and variance. After combining the two of them we will have Gaussian distribution for the whole interval \( J \).

2. Preliminary lemmas

In this section we collect all the preliminary lemmas that will be used in the subsequent sections. The first lemma is the Weil bound for space curves. For a proof, see [9, Theorem 2.1].

Lemma 2.1. Let \( C \) be an absolutely irreducible curve in the affine \( r \)-space \( \mathbb{A}^r_{\mathbb{F}_p} \) of degree \( d > 1 \), which is not contained in any hyperplane. Let \( B = I_1 \times \ldots \times I_r \) be a box, then

\[ N_B(C) = p \cdot \frac{\text{vol}(B)}{p^r} + O(d^2 \sqrt{p} \log^4 p), \]

where \( t \) is the number of intervals \( I_i \) that are not the full interval \([0, p-1]\).

The next lemma states that if we translate a set in \( \mathbb{F}_p \) a small number of times, it will always reach a new element. This lemma allows us to show later that some curves are absolutely irreducible.

Lemma 2.2. Let \( r \geq 2 \), \( x_1, \ldots, x_r \in \mathbb{F}_p \) be \( r \) distinct elements. Suppose \( \mathcal{M} \) is a nonempty finite subset of the algebraic closure \( \overline{\mathbb{F}}_p \) with \( 4|\mathcal{M}| < p^{7/5} \). Then there exists a \( j \in \{1, \ldots, r\} \) such that the translate \( \mathcal{M} + x_j \) is not contained in \( \cup_{i \neq j}(\mathcal{M} + x_i) \).

Proof. Suppose \((x_1, \ldots, x_r, \mathcal{M})\) provides a counterexample to the statement of the lemma. Then it is clear that for any nonzero \( t \in \mathbb{F}_p \), the tuple \((tx_1, \ldots, tx_r, t\mathcal{M})\) is another counterexample.

By Minkowski’s theorem on lattice points in a convex symmetric body, there exists a nonzero integer \( t \) such that

\[
\begin{aligned}
\|t\| &\leq p - 1 \\
\left\|\frac{tx_1}{p}\right\| &\leq (p - 1)^{-\frac{1}{5}} \\
&\vdots \\
\left\|\frac{tx_r}{p}\right\| &\leq (p - 1)^{-\frac{1}{5}}.
\end{aligned}
\]
Thus there are integers $y_j$ such that

$$
\begin{aligned}
|y_j| &\leq p(p-1)^{-\frac{1}{p}} \\
y_j &\equiv tx_j \pmod{p}
\end{aligned}
$$

(10)

for any $j \in \{1, \ldots, r\}$, and $(y_1, \ldots, y_r, t\mathcal{M})$ provides a counterexample. Now let $j_0$ be such that $|y_{j_0}| = \max_{1 \leq j \leq r} |y_j|$. Choose $\alpha \in t\mathcal{M}$ and consider the set $\mathcal{M} = t\mathcal{M} \cap (\alpha + \mathbb{F}_p)$. Then $(y_1, \ldots, y_r, \mathcal{M})$ will also be a counterexample.

Note that $\alpha + \mathbb{F}_p$ can be written as a union of at most $|\mathcal{M}|$ intervals (i.e. subsets of $\mathbb{F}_p$ consisting of consecutive integers or its translate in $\mathbb{F}_p$) whose endpoints are in $\mathcal{M}$. Let $\{a + \alpha, a + a + 1, \ldots, a + b\}$ be the longest of these intervals. Then

$$
|b - a| \geq \frac{p}{|\mathcal{M}|} \geq \frac{p}{|\mathcal{M}|}.
$$

By this, (10) and the hypothesis $4|\mathcal{M}| < p^\frac{4}{3}$, we have

$$
|b - a| > 4p^{1-\frac{4}{3}} > 2|y_{j_0}|.
$$

Now if $y_{j_0} > 0$, then $\alpha + a + y_{j_0}$ belongs to $\mathcal{M} + y_{j_0}$ but does not belong to $\cup_{i \neq j_0} (\mathcal{M} + y_i)$, while if $y_{j_0} < 0$, then $\alpha + b + y_{j_0}$ belongs to $\mathcal{M} + y_{j_0}$ but does not belong to $\cup_{i \neq j_0} (\mathcal{M} + y_i)$. This contradicts the fact that $(y_1, \ldots, y_r, \mathcal{M})$ is a counterexample, and completes our proof. $\square$

Recall that the Stirling number of second kind, $S(r, t)$, is by definition the number of partition of a set of cardinality $r$ into exactly $t$ nonempty subsets. The proof of the following lemma can be found in [10].

**Lemma 2.3.** Let $\mu_k(H, P)$ be defined by (9), then

$$
\mu_k(H, P) = \sum_{r=0}^{k} \binom{k}{r} (-HP)^{k-r} \left( \sum_{t=0}^{r} \binom{H}{t} S(r, t)! P^t \right).
$$

3. Patterns of curves: Proof of Theorem 1

Let $C$ be a plane curve given by the equation $f(x, y) = 0$ and two vectors $a = (a_1, \ldots, a_s), b = (b_1, \ldots, b_s)$ so that $p \nmid a_i$ and $a_1^{-1}b_1, \ldots, a_s^{-1}b_s$ are all distinct modulo $p$, we define the $x$-shifted curve $C_{a, b}$ to be the space curve in the affine $(s+1)$-space with variables $x, y_1, \ldots, y_r$ and equations

$$
f(a_i x + b_i, y_i) = 0, \forall 1 \leq i \leq s.
$$

(11)

It is not difficult to see that $C_{a, b}$ is indeed a curve, and its degree is less than or equal to $d^s$. Note that similar constructions appeared in [7, 8, 9].

It is clear from the defining equations (11) that a point on $C_{a, b}$ corresponds to an $(a, b)$-pattern on $C$, i.e.

$$
P_{a, b}(C, \mathcal{I}, \mathcal{J}) = N_B(C_{a, b}),
$$

where $B = \mathcal{I} \times (\mathcal{J})^s$. We want to show that $C_{a, b}$ is absolutely irreducible. Currently we are not able to prove this for all curves $C$, but we are able to show the irreducibility for the class of curves so that the projection $\pi$ defined by (6) has a completely ramified point.

**Proposition 3.1.** If $C$ satisfies the assumptions in Theorem 1, then $C_{a, b}$ is absolutely irreducible.
Proof. For $1 \leq j \leq s$ we define $C_j$ to be the curve given by the first $j$ equations in (11), i.e.

$$f(a_i x + b_i, y_i) = 0, \ \forall 1 \leq i \leq j.$$  

We have a chain of coverings of curves,

$$C_{a,b} = C_s \to C_{s-1} \to \ldots \to C_1 \cong C,$$

where each arrow represent a projection $\pi$.

Let $S \subseteq \mathbb{P}_p$ be the set of completely ramified points for the map $\pi : C \to A^1$. Since all the $x_i = b_i a_i^{-1}$ are distinct, we can apply Lemma 2.2 with $x_i = b_i a_i^{-1}$ to conclude that there are new completely ramified points in each level of the above chain of coverings. Since $C$ is absolutely irreducible, this shows that $C_{a,b}$ is also absolutely irreducible.

We are now ready to prove Theorem 1. By Proposition 2.1 if $C$ satisfies the assumptions in the theorem, then $C_{a,b}$ is absolutely irreducible in $A^{s+1}$. Theorem 1 now follows easily from Lemma 2.4.

4. Estimation of $M_k(H)$: Proof of Theorem 2

Using the binomial theorem to expand the right hand side of (8), we obtain

$$M_k(H) = \sum_{x=0}^{p-1} \sum_{r=0}^{k} \binom{k}{r} N_{B_x}(C)^r \left( -\frac{HN}{p} \right)^{k-r} \sum_{x=0}^{p-1} N_{B_x}(C)^r.$$

Here we make the convention that if $r = 0$, $N_{B_x}(C)^r = 1$ even when $N_{B_x}(C) = 0$. Define

$$S_r(H) = \sum_{x=0}^{p-1} N_{B_x}(C)^r.$$

Clearly $S_0(H) = p$ (by our convention). For $r \geq 1$, we have

$$S_r(H) = \sum_{x=0}^{p-1} \sum_{(x_1, y_1) \in C \cap B_x} \cdots \sum_{(x_r, y_r) \in C \cap B_x} 1 = \sum_{x=0}^{p-1} \sum_{(x, y_1) \in C, y_i \in J, \{x_1, \ldots, x_r\} \subseteq (x, x+H)} 1.$$

For each $1 \leq i \leq r$, let $x_i = x + a_i$, and let $A$ be the set of distinct $a_i$’s. Set $|A| = t$. We have $A \subseteq \{1, 2, \ldots, H\}$. From the definition of the Stirling number of second kind, we see that for any given $A$, the number of sets with $\{x_1, \ldots, x_r\} = A$ is $S(r, t)!$. Grouping the terms in (12) according to different values of $t$, we obtain

$$S_r(H) = \sum_{t=1}^{r} S(r, t)! \sum_{|A|=t, A \subseteq [1, H]} \sum_{x=0}^{p-1} \sum_{(x + b_i, y_i) \in C, 1 \leq i \leq r} 1.$$

By condition (7), the inner sum

$$\sum_{x=0}^{p-1} \sum_{(x + b_i, y_i) \in C, 1 \leq i \leq r} 1.$$
is the number of \((a, b)\)-pattern of \(C\) with \(a = (1, 1, \ldots, 1)\), \(b\) is any \(t\)-tuple ordering of the set \(A\), and all \(y\) coordinates lie in \(J\). By Theorem [1] this sum is
\[
\sum_{x=0}^{p-1} \sum_{(x+b,y) \in C, 1 \leq i \leq r} 1 = p \cdot \frac{N^t}{p^t} + O(d^{2t} \sqrt{p} \log^t p).
\]

Put this into (13) yields
\[
S_r(H) = \sum_{t=1}^{r} S(r, t)! \sum_{|A|=t, A \subseteq [1.H]} \left( p \cdot \frac{N^t}{p^t} + O(d^{2t} \sqrt{p} \log^t p) \right)
= p \sum_{t=1}^{r} S(r, t)! \left( \frac{H}{t} \right) \left( \frac{N}{p} \right)^t + O \left( \sum_{t=1}^{r} S(r, t)! \left( \frac{H}{t} \right) d^{2t} \sqrt{p} \log^t p \right).
\]

Therefore,
\[
M_k(H) = p \sum_{r=0}^{k} \binom{k}{r} \left( -\frac{HN}{p} \right)^{k-r} \sum_{t=1}^{r} S(r, t)! \left( \frac{H}{t} \right) \left( \frac{N}{p} \right)^t + O_k(d^{2k} H^k \sqrt{p} \log^k p).
\]

We can insert the terms with \(t = 0\) without altering the sum since \(S(r, 0) = 0\) for any \(r \geq 1\) (and for \(r = 0\) the inner sum is understood to be zero), thus we may apply Lemma [2] with \(P = N/p\) to conclude that
\[
M_k(H) = p^\nu(H, N/p) + O_k(d^{2k} H^k \sqrt{p} \log^k p).
\]

This completes the proof of Theorem [2].

References

[1] E. Bombieri, On exponential sums in finite fields, Amer. J. Math. 88 (1966), no. 1, 71–105.
[2] T. H. Chan, An almost all result on \(q_1q_2 \equiv c \pmod{q}\), Monatsh. Math. 162 (2011), no. 1, 29–39.
[3] C. I. Cobeli, S. M. Gonek, and A. Zaharescu, The distribution of patterns of inverses modulo a prime, J. Number Theory 101 (2003), no. 2, 209–222.
[4] S. M. Gonek, G. S. Krishnaswami, and V. L. Sondhi, The distribution of inverses modulo a prime in short intervals, Acta Arith. 102 (2002), no. 4, 315–322.
[5] Y. Lamzouri, The distribution of short character sums, arXiv:1108.6072 [math.NT].
[6] K.-H. Mak and A. Zaharescu, The distribution of rational points and polynomial maps on an affine variety over a finite field on average, arXiv:1301.1359 [math.NT].
[7] , On the distribution of the number of points on a family of curves over finite fields, arXiv:1110.4693 [math.NT].
[8] , The distribution of values of short hybrid exponential sums on curves over finite fields, Math. Res. Lett. 18 (2011), no. 1, 155–174.
[9] , Poisson type phenomena for points on hyperelliptic curves modulo \(p\), Funct. Approx. Comment. Math. 47 (2012), no. 1, 65–78.
[10] H. L. Montgomery and R. C. Vaughan, On the distribution of reduced residues, Ann. of Math. (2) 123 (1986), no. 2, 311–333.
[11] André Weil, Sur les courbes algébriques et les variétés qui s’en déduisent, Actualités Sci. Ind., no. 1041 = Publ. Inst. Math. Univ. Strasbourg 7 (1945), Hermann et Cie., Paris, 1948.