



ON THE DISTRIBUTION OF POINTS OF VALUATION 1 FOR A
POLYNOMIAL IN TWO VARIABLES

Krishnan Rajkumar

School of Engineering, Jawaharlal Nehru University, New Delhi, India
krishnan@jnu.ac.in

Shubham

School of Physical Sciences, Jawaharlal Nehru University, New Delhi, India
shubha76_sps@jnu.ac.in
shubham01nitw@gmail.com

Received: 4/10/25, Revised: 10/30/25, Accepted: 6/3/26, Published: 6/26/26

Abstract

We investigate the variation in the total number of points in a random $p \times p$ square in \mathbb{Z}^2 at which the p -adic valuation of a given polynomial in two variables is greater than 1. We establish that this quantity follows a Poisson distribution as $p \rightarrow \infty$ under a certain conjecture. We also relate this conjecture to certain uniform distribution properties of a vector-valued sequence.

1. Introduction

The p -adic valuation of integer sequences is a fundamental concept in number theory, providing insights into the divisibility properties of integers with respect to a prime number p . The p -adic valuation $\nu_p(n)$ of an integer n is defined as the largest exponent e such that p^e divides n for a prime number p . Historically, the concept of p -adic valuation stems from the work of Hensel [6], who introduced p -adic numbers in 1897. This development allowed mathematicians to study numbers and their divisibility in a local sense, focusing on individual primes. Over time, the theory of p -adic numbers and valuations was expanded by E. Artin, Grothendieck, and others, making it a crucial tool in areas such as Diophantine equations and algebraic geometry.

One of the first and important results in the study of p -adic valuations is Legendre's formula [7] for the p -adic valuation of $n!$. For a prime number p , the formula is given by:

$$\nu_p(n!) = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor,$$

where $\lfloor x \rfloor$ is the greatest integer less than or equal to x . This sums the number of multiples of p, p^2, p^3, \dots less than or equal to n , counting the number of times p divides $n!$. This formula has important applications in number theory and combinatorics, particularly in understanding binomial coefficients and their divisibility by prime numbers.

The work presented by Byrnes et al. [1] studies the 2-adic valuation of $n^2 - a$. It is shown that $\nu_2(n^2 - a)$ has a simple closed form (see [3, 4] for a detailed discussion of the definition of a closed form) when $a \not\equiv 4, 7 \pmod{8}$. For these two remaining cases, the valuation cannot be described in a simple form. However, it can be fully described in a rule-based framework called the *valuation tree*.

These examples form part of a general project, with several notable contributions made by Victor Moll [1, 8] and his coauthors to describe the p -adic valuations $\nu_p(x_n)$, for any given sequence $\{x_n\}_{n \in \mathbb{N}}$ of integers. In the work [10] by the second author, the valuations $\nu_p(f(m, n))$ of a two-dimensional sequence $\{f(m, n)\}_{m, n \in \mathbb{N}}$ of integers are taken up, where $f(x, y) \in \mathbb{Z}[x, y]$ is any polynomial in two variables with integer coefficients. It is fully described by generalizing the notion of the valuation tree.

In this work, we wish to understand the distribution of the p -adic valuations $\nu_p(f(m_1, m_2, \dots, m_k))$ as m_1, m_2, \dots, m_k vary over \mathbb{N} for a polynomial $f \in \mathbb{Z}[x_1, x_2, \dots, x_k]$. Let us first discuss the case $k = 1$ of a polynomial in one variable.

Let $f(x) \in \mathbb{Z}[x]$ have degree d and non-zero discriminant. First, we note that points n where f has non-zero valuation $\nu_p(f(n)) > 0$ follow a periodic pattern with period p , as $\nu_p(f(x)) > 0$ if and only if $\nu_p(f(x + p)) > 0$. Hence, every interval of length p , in particular the intervals $\{kp, kp + 1, \dots, kp + p - 1\}$, will contain the same number of points of non-zero valuation, and this number is the cardinality of the solution set of $f(x) = 0$ over \mathbb{F}_p .

The situation changes when we consider the points of valuation $\nu_p(f(n)) > 1$. In the first place, the pattern is still periodic with period p^2 . However, not all intervals of length p contain the same number of points where $\nu_p(f(n)) > 1$. In particular, at least $p - d$ intervals $\{kp, kp + 1, \dots, kp + p - 1\}$ with $k = 0, 1, \dots, p - 1$ do not contain any points where $\nu_p(f(n)) > 1$. This is because, by Corollary 1, the total number of points where $\nu_p(f(n)) > 1$ in the set $\{0, 1, \dots, p^2 - 1\}$ is at most d . Hence, as $p \rightarrow \infty$, the total number of points n of valuation $\nu_p(f(n)) > 1$ in an interval $\{kp, kp + 1, \dots, kp + p - 1\}$ chosen uniformly at random from $k = 0, 1, \dots, p - 1$ is zero with probability tending to 1.

We now turn to the case $k = 2$ of a polynomial in two variables and answer the following question: what is the distribution of points in \mathbb{N}^2 at which a polynomial $f(x, y) \in \mathbb{Z}[x, y]$ has valuation greater than 1?

For a fixed polynomial $f(x, y) \in \mathbb{Z}[x, y]$, we note that the points (r, s) where f

has non-zero valuation $\nu_p(f(r, s)) > 0$ display a p -periodic pattern in both r and s coordinates. The number of points in any $p \times p$ square is the number of points on the curve $C_f(\mathbb{F}_p) = \{(x, y) \in \mathbb{Z}^2 : \nu_p(f(x, y)) > 0 \text{ and } 0 \leq x, y < p\}$ given by the equation $f(x, y) = 0$ over \mathbb{F}_p . Let $m = |C_f(\mathbb{F}_p)|$ be the number of points on this curve.

Let us observe the patterns that occur for the points in \mathbb{N}^2 at which the valuation of f is exactly 1, for various polynomials f . This is shown in Figure 1. The points missing from $C_f(\mathbb{F}_p)$ in the various $p \times p$ squares seem to exhibit some element of randomness. In order to quantify this, we first study the properties of indicator random variables $Y_{x,y}$ in Section 3. In summary, for each $(x, y) \in C_f(\mathbb{F}_p)$, we define a random variable $Y_{x,y}$ that indicates which translates of (x, y) into a random $p \times p$ square have valuation greater than 1. These random variables are identically distributed Bernoulli random variables with probability $1/p$, though they are not always pairwise independent.

In order to further quantify this variation, we define the following random variable $X : \mathbb{F}_p \times \mathbb{F}_p \rightarrow \mathbb{Z}$:

$$X(k, l) = |\{(x, y) \in \{kp, kp + 1, \dots, kp + p - 1\} \times \{lp, lp + 1, \dots, lp + p - 1\} : \nu_p(f(x, y)) > 1\}|. \tag{1}$$

In other words, X represents the total number of points in a random $p \times p$ square at which the valuation of f is greater than 1. The distribution of X for various primes p is shown in Figure 2.

The distribution of X appears to approach the Poisson distribution for large p . This limiting behavior of X can be explained by the following reasoning under the assumption that $Y_{x,y}$ are mutually independent: by Equation (??) and this assumption, we find that X follows the binomial distribution with parameters m and $1/p$, with their product m/p tending to 1 as $p \rightarrow \infty$ (by Corollary 4). Hence, under this assumption, the distribution of X tends to a Poisson distribution with the parameter $\lambda = 1$. The central result of this work is that this statement holds without the assumption on mutual independence of $Y_{x,y}$, under Conjecture 1. To state this conjecture, we need some notation.

For any k -tuple $\mathbf{P} = ((r_1, s_1), (r_2, s_2), \dots, (r_k, s_k))$ of points on the curve $C_f(\mathbb{F}_p)$, denote the associated matrix $B_k(\mathbf{P})$ with the entries in \mathbb{F}_p as

$$B_k(\mathbf{P}) = \begin{bmatrix} f_x(r_1, s_1) & f_y(r_1, s_1) & f(r_1, s_1)/p \\ f_x(r_2, s_2) & f_y(r_2, s_2) & f(r_2, s_2)/p \\ \vdots & \vdots & \vdots \\ f_x(r_k, s_k) & f_y(r_k, s_k) & f(r_k, s_k)/p \end{bmatrix}. \tag{2}$$

For example, consider $f(x, y) = x^3 + y^2 + xy + 1$ and $p = 7$. Then the curve $C_f(\mathbb{F}_p) = \{(1, 3), (3, 0), (3, 4), (4, 1), (4, 2), (5, 0), (5, 2), (6, 0), (6, 1)\}$. For $k = 3$, and

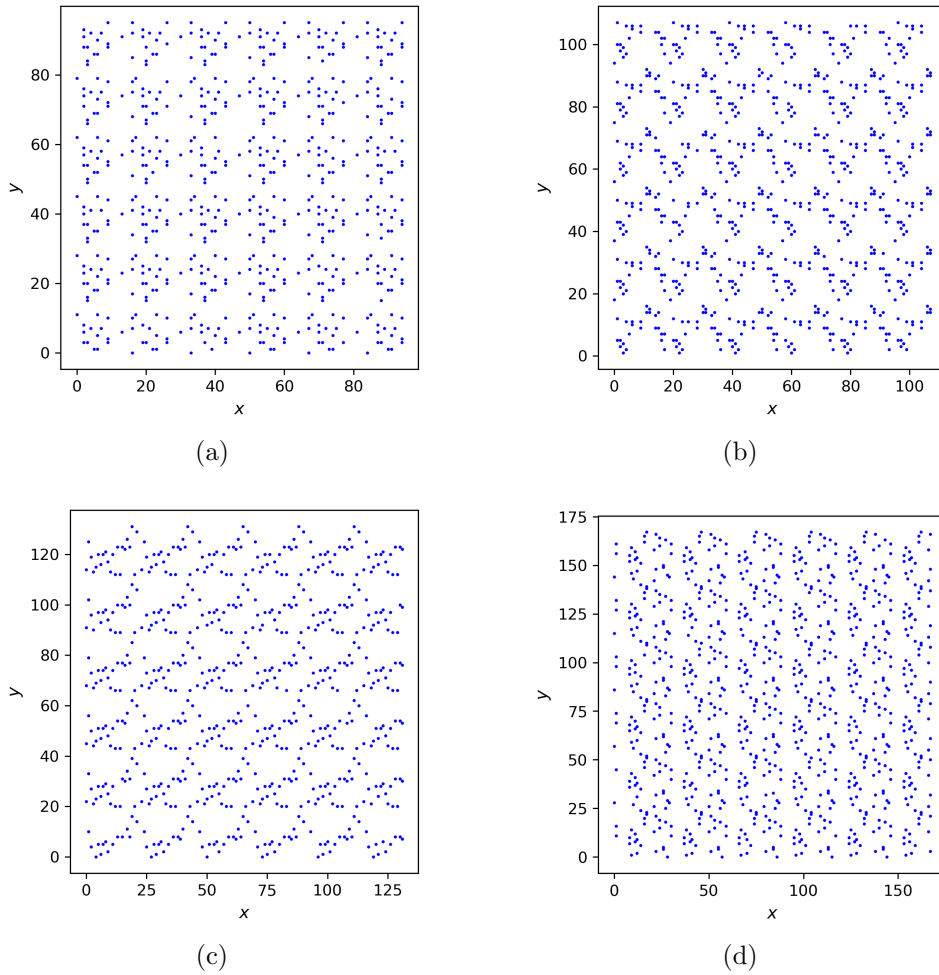
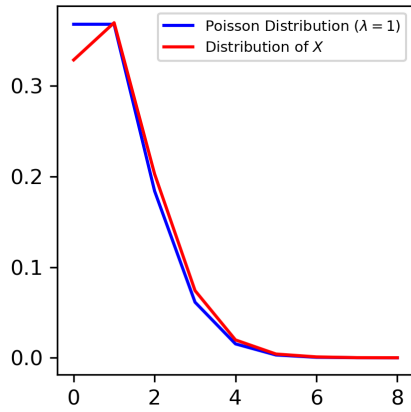
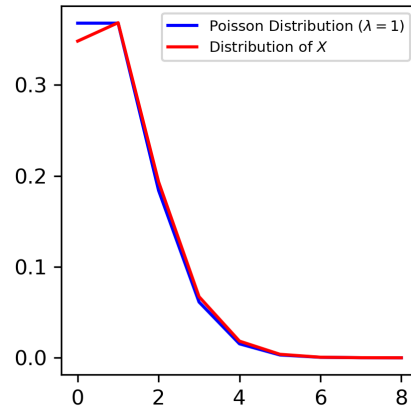


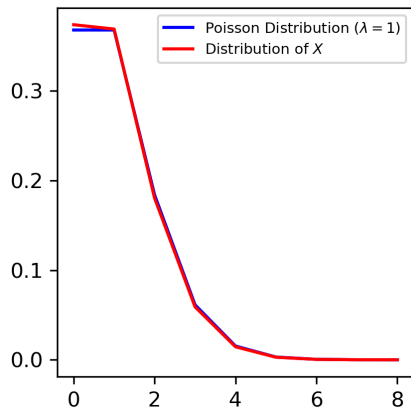
Figure 1: Scatter plots of points (x, y) satisfying $\nu_p(f(x, y)) = 1$ for different polynomials and primes. Subfigures (a)–(d) correspond to $f(x, y) = x^3 + y^3 + x^2y + y + 1$ with $p = 17$, $f(x, y) = x^3 + y^2x + xy + x + y + 1$ with $p = 19$, $f(x, y) = x^3 + xy + x + y + 1$ with $p = 23$, and $f(x, y) = y^2x + xy + x + y + 1$ with $p = 29$, respectively.



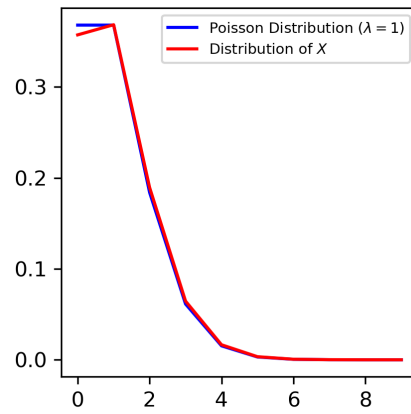
(a)



(b)



(c)



(d)

Figure 2: Comparison of the distribution of X with the Poisson distribution ($\lambda = 1$) for the polynomial $f(x, y) = x^3 + y^2 + xy + 1$. Subfigures (a)-(d) correspond to primes 211, 503, 1013, and 1033, respectively.

$\mathbf{P}_1 = ((1, 3), (3, 0), (3, 4))$, we have

$$B_3(\mathbf{P}_1) = \begin{bmatrix} 6 & 7 & 2 \\ 27 & 3 & 4 \\ 31 & 11 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 2 \\ 6 & 3 & 4 \\ 3 & 4 & 1 \end{bmatrix},$$

in \mathbb{F}_7 , which has rank 3 (its determinant is non-zero in \mathbb{F}_7).

For $\mathbf{P}_2 = ((1, 3), (3, 0), (5, 2))$, we have

$$B_3(\mathbf{P}_2) = \begin{bmatrix} 6 & 7 & 2 \\ 27 & 3 & 4 \\ 77 & 9 & 20 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 2 \\ 6 & 3 & 4 \\ 0 & 2 & 6 \end{bmatrix},$$

over \mathbb{F}_7 , which has rank 2, as the first two columns are linearly independent (any of the 2×2 minors involving the first two columns is non-zero in \mathbb{F}_7) and the third column is a linear combination of the first two columns over \mathbb{F}_7 .

Let $m_k(f, p)$ denote the number of k -tuples $\mathbf{P} \in C_f(\mathbb{F}_p)^k$ that are element-wise distinct, for which $B_k(\mathbf{P})$ has rank 2 and the first two columns of $B_k(\mathbf{P})$ are linearly independent. In the preceding example, $m_3(f, p)$ includes \mathbf{P}_2 and its permutations, but not \mathbf{P}_1 or its permutations.

For $k = 1$, we observe that $m_k(f, p) = 0$ trivially. For $k \geq 2$, we have the following conjecture.

Conjecture 1. For all $k \geq 2$, we have

$$\lim_{p \rightarrow \infty} \frac{m_k(f, p)}{p^2} = 1. \tag{3}$$

Qualitatively, this conjecture implies that, for $k > 2$, the k -tuples counted by $m_k(f, p)$ are very sparse among the element-wise distinct k -tuples in $C_f(\mathbb{F}_p)^k$. In continuation of the preceding example with $f(x, y) = x^3 + y^2 + xy + 1$ and $p = 7$, we calculate $m_3(f, 7) = 90$, which is comparable to $p^2 = 49$, although their ratio is much higher than 1. However, for larger values of p , numerical observations appear to corroborate this conjecture, as seen in Figure 3.

We also show in Section 4 that this conjecture is a natural consequence of certain versions of the uniform distribution of the sequence \mathcal{F}_p in \mathbb{F}_p^3 , where

$$\mathcal{F}_p = \{(f_x(r, s), f_y(r, s), f(r, s)/p) \in \mathbb{F}_p^3 \mid (r, s) \in C_f(\mathbb{F}_p)\}. \tag{4}$$

Now, we are in a position to state the main result of this work.

Theorem 1. *Let $f(x, y) \in \mathbb{Z}[x, y]$ be a polynomial with no solutions to the set of congruences $f(x, y) \equiv 0 \pmod{p}$, $f_x(x, y) \equiv 0 \pmod{p}$, and $f_y(x, y) \equiv 0 \pmod{p}$ for all sufficiently large prime numbers p . Let X be the random variable defined by Equation (1), and assume that Conjecture 1 is satisfied. Then X converges in distribution to a Poisson random variable with parameter $\lambda = 1$ as $p \rightarrow \infty$.*

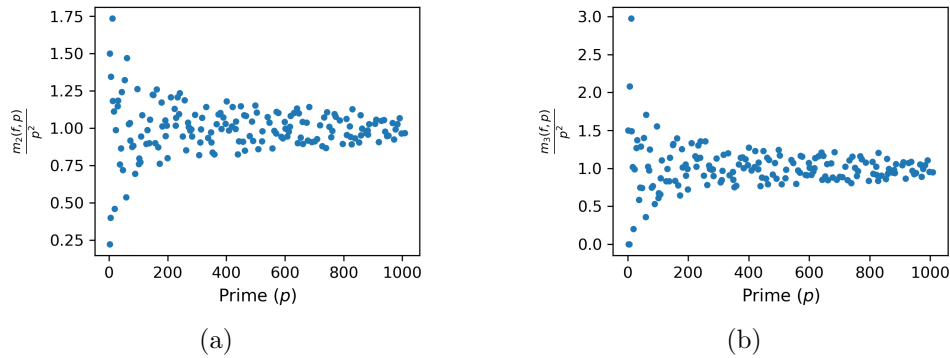


Figure 3: Plots of $m_k(f, p)/p^2$ versus p for the polynomial $f(x, y) = x^3 + y^2 + xy + 1$. Subfigures (a) and (b) correspond to $k = 2$ and $k = 3$, respectively.

We note that the condition on the polynomial $f(x, y)$ mentioned in Theorem 1 is satisfied by the polynomials mentioned in Figures 1, 2, and 3. For instance, the polynomial $f(x, y) = x^3 + y^2 + xy + 1$ does not have any solutions to the set of congruences $f(x, y) \equiv 0 \pmod{p}$, $f_x(x, y) \equiv 0 \pmod{p}$, and $f_y(x, y) \equiv 0 \pmod{p}$ for $p > 431$.

To obtain a sufficient condition on $f(x, y)$ that is independent of p , one can eliminate the variable y from $f(x, y) = 0$ and $f_x(x, y) = 0$, as well as from $f(x, y) = 0$ and $f_y(x, y) = 0$. This would result in two polynomials in x , both of which have roots modulo p at the solutions of the original set of congruences. However, there can be common roots for these two polynomials modulo p if and only if $p \mid R$, where R is the resultant of these two polynomials. Thus, a sufficient condition on $f(x, y)$ for Theorem 1 to hold is that this resultant R is a non-zero integer.

We also remark that the case of polynomials in more than two variables is considerably involved and will be considered in future research.

We now describe the organization of this paper. In Section 2, we prove several elementary results on the number of solutions to certain valuation conditions. In Section 3, we interpret these results as describing the single and joint distributions of the indicator random variables indexed by points in $C_f(\mathbb{F}_p)$. We also derive the implication of Conjecture 1 on the sum of expectations of k -fold products of these random variables. Section 4 contains results on the equidistribution of \mathcal{F}_p and their relation to Conjecture 1. Section 5 contains the proof of Theorem 1.

2. Preliminaries

We start by studying the cardinality of the set $\{x : \nu_p(g(x)) > 1, 0 \leq x < p^2\}$ for a given polynomial $g(x) \in \mathbb{Z}[x]$.

Proposition 1. *Let p be a prime number, let $g(x) \in \mathbb{Z}[x]$, and let $r \in \{0, 1, \dots, p-1\}$ not be a common solution to congruences $g(r) \equiv 0 \pmod{p}$ and $g'(r) \equiv 0 \pmod{p}$. Then the cardinality of the set $\{k \in \mathbb{F}_p : \nu_p(g(r + kp)) > 1\}$ is at most 1.*

Proof. We note that

$$g(r + kp) \equiv g(r) + kpg'(r) \pmod{p^2}. \tag{5}$$

If $g(r) \not\equiv 0 \pmod{p}$, then by Equation (5), there are no solutions to $g(r + kp) \equiv 0 \pmod{p^2}$. On the other hand, if $g(r) \equiv 0 \pmod{p}$ and $g'(r) \not\equiv 0 \pmod{p}$, then Equation (5) provides a unique solution to $g(r + kp) \equiv 0 \pmod{p^2}$, where $k \in \mathbb{F}_p$. □

Corollary 1. *Let $g(x) \in \mathbb{Z}[x]$ have degree d and non-zero discriminant Δ . Then the cardinality of the set $S = \{x : \nu_p(g(x)) > 1, 0 \leq x < p^2\}$ is at most d for sufficiently large prime numbers p .*

Proof. First, we note that for p sufficiently large, we have $\Delta \not\equiv 0 \pmod{p}$, since Δ is a non-zero integer (since it is a polynomial in the coefficients of g). Hence, there are no repeated roots of g in any algebraically closed extension of \mathbb{F}_p . As a result, there are no solutions to $g(r) \equiv 0 \pmod{p}$ and $g'(r) \equiv 0 \pmod{p}$ in \mathbb{F}_p .

Thus, the conditions of Proposition 1 hold for all $r \in \mathbb{F}_p$. Now, from the proof of Proposition 1, it is clear that there are solutions to $g(r + kp) \equiv 0 \pmod{p^2}$ only if r is a root of $g(r) \equiv 0 \pmod{p}$. Hence, there are at most d such roots $r \in \{0, 1, \dots, p-1\}$ to consider, and at most one translate $r + kp$ belongs to the set S . Thus, the cardinality of S is at most d . □

Next, we consider a polynomial $f(x, y) \in \mathbb{Z}[x, y]$ and study the cardinality of the set $\{(x, y) : \nu_p(f(x, y)) = 1, 0 \leq x, y < p^2\}$. We get the following results.

Proposition 2. *Let p be a prime number, let $f(x, y) \in \mathbb{Z}[x, y]$ be a polynomial, and let $(r, s) \in C_f(\mathbb{F}_p)$ be such that at least one of $f_x(r, s)$ or $f_y(r, s) \not\equiv 0 \pmod{p}$. Then the cardinality of the set $\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r + kp, s + lp)) = 1\}$ is $p(p-1)$.*

Proof. Consider the solutions of

$$f(r + kp, s + lp) \equiv 0 \pmod{p^2},$$

for $k, l \in \{0, 1, 2, \dots, p-1\}$. Now, by Taylor series expansion, we have

$$f(r, s) + kpf_x(r, s) + lpf_y(r, s) \equiv 0 \pmod{p^2}. \tag{6}$$

We know that $f(r, s) \equiv 0 \pmod{p}$ which implies there exists an integer α , such that $f(r, s) = \alpha p$. Using Equation (6), we have

$$\alpha + kf_x(r, s) + lf_y(r, s) \equiv 0 \pmod{p}. \tag{7}$$

Assuming, without loss of generality, that $f_x(r, s) \not\equiv 0 \pmod{p}$, we have

$$k \equiv \frac{-\alpha - lf_y(r, s)}{f_x(r, s)} \pmod{p}.$$

So, there are p choices for (k, l) such that $f(r + kp, s + lp) \equiv 0 \pmod{p^2}$. Hence, there are $p^2 - p$ choices for (k, l) such that $f(r + kp, s + lp) \not\equiv 0 \pmod{p^2}$. \square

Corollary 2. *Let m and s be the cardinalities of the sets $C_f(\mathbb{F}_p)$ and S , respectively, for a prime number p , where $S = \{(x, y) : \nu_p(f(x, y)) = 1, 0 \leq x, y < p^2\}$. Then $s = mp(p - 1)$.*

We also derive the following result directly from Equation (7).

Corollary 3. *Let p be a prime number and let $(r, s) \in C_f(\mathbb{F}_p)$ be such that $f_x(r, s) \equiv f_y(r, s) \equiv 0 \pmod{p}$. Then the cardinality of the set $\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r + kp, s + lp)) = 1\}$ is p^2 if $f(r, s) \equiv 0 \pmod{p^2}$, and 0 if $f(r, s) \not\equiv 0 \pmod{p^2}$.*

Next, we examine the mutual influence of points (r_1, s_1) and (r_2, s_2) on the curve $C_f(\mathbb{F}_p)$. Recall the matrix $B_2((r_1, s_1), (r_2, s_2))$ from Equation (2), denoted as A for convenience,

$$A = \begin{pmatrix} f_x(r_1, s_1) & f_y(r_1, s_1) & f(r_1, s_1)/p \\ f_x(r_2, s_2) & f_y(r_2, s_2) & f(r_2, s_2)/p \end{pmatrix}.$$

Let r be the rank of A . In the case where $r = 2$ and the first two columns of A are linearly independent, we establish the following.

Proposition 3. *Let p be a prime number, let $f(x, y) \in \mathbb{Z}[x, y]$ be an integer polynomial, and let (r_1, s_1) and (r_2, s_2) be points on the curve $C_f(\mathbb{F}_p)$. Suppose that the first two columns of A are linearly independent. Then, the cardinality of the set*

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) = 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) = 1\}$$

is $(p - 1)^2$, the cardinality of the set

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) > 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) = 1\}$$

is $p - 1$, and the cardinality of the set

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) > 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) > 1\}$$

is 1.

Proof. For $i, j \in \{0, 1, 2, \dots, p-1\}$, consider the following Taylor series expansion:

$$f(r_1 + ip, s_1 + jp) \equiv f(r_1, s_1) + ipf_x(r_1, s_1) + jpf_y(r_1, s_1) \pmod{p^2}. \quad (8)$$

We know that $f(r_1, s_1) \equiv 0 \pmod{p}$ which implies that there exist integers α and α' such that $f(r_1, s_1) = \alpha p$ and $f(r_1 + ip, s_1 + jp) = \alpha' p$. So, using Equation (8), we have

$$if_x(r_1, s_1) + jf_y(r_1, s_1) \equiv -\alpha + \alpha' \pmod{p}. \quad (9)$$

Similarly, by using $\nu_p(f(r_2, s_2)) = 1$ we have

$$if_x(r_2, s_2) + jf_y(r_2, s_2) \equiv -\beta + \beta' \pmod{p}, \quad (10)$$

for integers β and β' such that $f(r_2, s_2) = \beta p$ and $f(r_2 + ip, s_2 + jp) = \beta' p$. From Equation (9) and Equation (10), we get the following system of two congruences in variables i and j :

$$A \begin{pmatrix} i \\ j \\ 1 \end{pmatrix} \equiv \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} \pmod{p}. \quad (11)$$

Since the first two columns of A are linearly independent, the above system of congruences has a unique solution for each α' and β' .

The set

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) = 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) = 1\}$$

corresponds to $\alpha' \not\equiv 0 \pmod{p}$ and $\beta' \not\equiv 0 \pmod{p}$. Hence, its cardinality is $(p-1)^2$. Similarly, the set

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) > 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) = 1\}$$

corresponds to $\alpha' \equiv 0 \pmod{p}$ and $\beta' \not\equiv 0 \pmod{p}$. So, this set has $p-1$ elements. Finally, the set

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) > 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) > 1\}$$

corresponds to $\alpha' \equiv 0 \pmod{p}$ and $\beta' \equiv 0 \pmod{p}$. So, it has only one element. \square

Next, in the case where the first two columns of A are linearly dependent, we establish the following.

Proposition 4. *Let p be a prime number and let $f(x, y) \in \mathbb{Z}[x, y]$ be a polynomial with no solutions to the set of congruences $f(x, y) \equiv 0 \pmod{p^2}$, $f_x(x, y) \equiv 0 \pmod{p}$, and $f_y(x, y) \equiv 0 \pmod{p}$. Let (r_1, s_1) and (r_2, s_2) be points on the curve $C_f(\mathbb{F}_p)$. Suppose that the first two columns of A are linearly dependent and at least one of them is non-zero. Then, the cardinality of the set*

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) = 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) = 1\}$$

is $p(p - 1)$ if $r = 1$ and $p(p - 2)$ if $r = 2$, the cardinality of the set

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) > 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) = 1\}$$

is 0 if $r = 1$ and p if $r = 2$, and the cardinality of the set

$$\{(k, l) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + kp, s_1 + lp)) > 1 \text{ and } \nu_p(f(r_2 + kp, s_2 + lp)) > 1\}$$

is p if $r = 1$ and 0 if $r = 2$.

Proof. We begin the proof by noting that System (11) can be established just as in the proof of Proposition 3. We also note that $r \neq 0$ by the given assumption on $f(x, y)$. There are three cases to consider.

Case 1: Consider the set

$$\{(i, j) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + ip, s_1 + jp)) = 1 \text{ and } \nu_p(f(r_2 + ip, s_2 + jp)) = 1\}.$$

This corresponds to $\alpha' \not\equiv 0 \pmod{p}$ and $\beta' \not\equiv 0 \pmod{p}$ in System (11). If $r = 1$, then System (11) has solutions if and only if $\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix}$ is a multiple of any non-zero column of A and this has $p - 1$ choices. For each such choice, there are p solutions of System (11), as the first two columns of A are linearly dependent. Hence, System (11) has a total of $p(p - 1)$ solutions in this case.

If $r = 2$, then System (11) has solutions if and only if $\begin{pmatrix} \alpha' - f(r_1, s_1)/p \\ \beta' - f(r_2, s_2)/p \end{pmatrix}$ is a multiple of any of the first two (non-zero) columns of A . This has $p - 2$ choices, since there are two distinct multiples leading to $\alpha' = 0, \beta' = 0$, respectively. Hence, System (11) has $p(p - 2)$ solutions in this case.

Case 2: Consider the set

$$\{(i, j) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + ip, s_1 + jp)) > 1 \text{ and } \nu_p(f(r_2 + ip, s_2 + jp)) = 1\}.$$

This corresponds to $\alpha' \equiv 0 \pmod{p}$ and $\beta' \not\equiv 0 \pmod{p}$ in System (11). In this case, System (11) becomes

$$A \begin{pmatrix} i \\ j \\ 1 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ \beta' \end{pmatrix} \pmod{p}. \tag{12}$$

Note that if $r = 1$, System (12) has solutions only if the first row of A has all zero entries, which is ruled out by the assumption on $f(x, y)$. Hence, System (12) has zero solutions when the rank of A is 1 and p solutions when the rank of A is 2.

Case 3: The set

$$\{(i, j) \in \mathbb{F}_p^2 : \nu_p(f(r_1 + ip, s_1 + jp)) > 1 \text{ and } \nu_p(f(r_2 + ip, s_2 + jp)) > 1\}$$

corresponds to $\alpha' \equiv 0 \pmod{p}$ and $\beta' \equiv 0 \pmod{p}$. In this case, System (11) becomes

$$A \begin{pmatrix} i \\ j \\ 1 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \end{pmatrix} \pmod{p}. \tag{13}$$

Therefore, System (13) has p solutions when $r = 1$, and zero solutions when $r = 2$. □

Next, we state the Hasse–Weil bounds and derive a corollary regarding the growth of $|C_f(\mathbb{F}_p)|$.

Proposition 5 ([11]). *Let C be a curve of genus g , \mathbb{F}_q be the finite field of order q , and N be the number of points of C over \mathbb{F}_q . Then, we have*

$$|N - q - 1| \leq 2g\sqrt{q}.$$

Corollary 4. *Let $m = |C_f(\mathbb{F}_p)|$ be the number of points on the curve $C_f(\mathbb{F}_p)$ given by the solutions of $f(x, y) = 0$ over \mathbb{F}_p . Then, we have $m/p \rightarrow 1$ as $p \rightarrow \infty$.*

3. Indicator Random Variables for Valuation Greater than 1

We shall assume for the remainder of this work that p is a prime number and that f satisfies the condition in Theorem 1, namely that there are no solutions to the set of congruences $f(x, y) \equiv 0 \pmod{p}$, $f_x(x, y) \equiv 0 \pmod{p}$, and $f_y(x, y) \equiv 0 \pmod{p}$ for all sufficiently large prime numbers p .

For each $(x, y) \in C_f(\mathbb{F}_p)$, define a random variable $Y_{x,y} : \mathbb{F}_p \times \mathbb{F}_p \rightarrow \{0, 1\}$ as

$$Y_{x,y}(k, l) = \begin{cases} 1, & \nu_p(f(x + kp, y + lp)) > 1, \\ 0, & \text{otherwise.} \end{cases}$$

In other words, $Y_{x,y}$ is an indicator function on the translates of (x, y) where f has p -adic valuation greater than 1. For all points $(x, y) \in C_f(\mathbb{F}_p)$, at least one of $f_x(x, y)$ or $f_y(x, y)$ is non-zero in \mathbb{F}_p according to the assumption on f . Hence, using Proposition 2, we infer that the probability mass function of $Y_{x,y}$ is

$$P(Y_{x,y} = i) = \begin{cases} 1/p, & i = 1, \\ (p - 1)/p, & i = 0. \end{cases} \tag{14}$$

Throughout, for a random variable X , the symbols $E[X]$ and $\text{Var}(X)$ denote the expectation and variance of X , respectively. Note that $Y_{x,y}$ has the following properties:

- (i) $E[Y_{x,y}] = 1/p$.

- (ii) $\text{Var}(Y_{x,y}) = \frac{p-1}{p^2}$.
- (iii) The r^{th} moment about the origin, denoted by μ'_r , equals $1/p$.
- (iv) By Chebyshev's inequality $P(|Y_{x,y} - 1/p| \geq \epsilon) \leq \frac{p-1}{\epsilon^2 p^2}$, for any $\epsilon > 0$.

For any pair of points (x, y) and $(r, s) \in C_f(\mathbb{F}_p)$ that satisfy the conditions of Proposition 3, the joint probability mass function of $Y_{x,y}, Y_{r,s}$ is

$$P(Y_{x,y} = i, Y_{r,s} = j) = \begin{cases} 1/p^2, & i = j = 1, \\ (p-1)/p^2, & i = 1, j = 0 \text{ or } i = 0, j = 1, \\ (p-1)^2/p^2, & i = j = 0. \end{cases}$$

Clearly, $P(Y_{x,y} = i, Y_{r,s} = j) = P(Y_{x,y} = i)P(Y_{r,s} = j)$ for all i and j , since Equation (14) holds for both points. Hence, $Y_{x,y}$ and $Y_{r,s}$ are independent random variables.

Similarly, in the case $r = 2$, the joint probability mass function for $Y_{x,y}$ and $Y_{r,s}$, where the pair of points satisfies the conditions of Proposition 4, is

$$P(Y_{x,y} = i, Y_{r,s} = j) = \begin{cases} 0, & i = j = 1, \\ 1/p, & i = 1, j = 0 \text{ or } i = 0, j = 1, \\ (p^2 - 2p)/p^2, & i = j = 0. \end{cases}$$

In the other case, $r = 1$, we have

$$P(Y_{x,y} = i, Y_{r,s} = j) = \begin{cases} 1/p, & i = j = 1, \\ 0, & i = 1, j = 0 \text{ or } i = 0, j = 1, \\ (p^2 - p)/p^2, & i = j = 0. \end{cases}$$

Clearly, neither of these cases describes independent random variables. However, the correlation coefficient is $-1/(p-1)$ for the case $r = 2$, and is 1 for the case $r = 1$. Hence, random variables are close to being uncorrelated when $r = 2$, but display a linear relationship when $r = 1$.

For ease of notation, we now fix any order on $C_f(\mathbb{F}_p)$ for the rest of this article and denote by Y_i the random variable corresponding to the i^{th} point.

Recall from Section 1 that $m_k(f, p)$ denotes the number of element-wise distinct k -tuples of points on $C_f(\mathbb{F}_p)$ whose associated matrix $B_k(\mathbf{P})$, given by (2), has rank 2 with the first two columns of $B_k(\mathbf{P})$ being linearly independent. Let $o_k(f, p)$ denote the number of element-wise distinct k -tuples where $B_k(\mathbf{P})$ has rank 1 with at least one of the first two columns non-zero.

Proposition 6. *For $k \geq 1$, we have*

$$\sum_{l_1 \neq l_2 \neq \dots \neq l_k} E[Y_{l_1} Y_{l_2} \dots Y_{l_k}] = m_k(f, p)/p^2 + o_k(f, p)/p. \tag{15}$$

Proof. For any element-wise distinct tuple l_1, l_2, \dots, l_k , it is easy to see that

$$P(Y_{l_1} = Y_{l_2} = \dots = Y_{l_k} = 1) = n/p^2,$$

where n is the number of solutions (i, j) of the system of congruences

$$B_k(\mathbf{P}) \begin{pmatrix} i \\ j \\ 1 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \pmod{p}. \tag{16}$$

In the case where the rank of $B_k(\mathbf{P})$ is 2 and the first two columns are linearly independent, there is a unique solution of System (16) and therefore $n = 1$. In the case where the rank of $B_k(\mathbf{P})$ is 1 with at least one of the first two columns non-zero, there are $n = p$ distinct solutions of System (16). The case where the rank of $B_k(\mathbf{P})$ is 0 is ruled out by the condition on f in Theorem 1.

The remaining cases are those where the rank of $B_k(\mathbf{P})$ is either 3 or 2 with the first two columns linearly dependent, or 1 with the first two columns zero. It can be seen that there are no solutions of System (16) in either of these remaining cases.

Since $E[Y_{l_1} Y_{l_2} \dots Y_{l_k}] = P(Y_{l_1} = Y_{l_2} = \dots = Y_{l_k} = 1)$, Equation (15) follows. \square

Corollary 5. *For $k \geq 1$, assuming that Conjecture 1 holds, we have*

$$\lim_{p \rightarrow \infty} \sum_{l_1 \neq l_2 \neq \dots \neq l_k} E[Y_{l_1} Y_{l_2} \dots Y_{l_k}] = 1.$$

Proof. Recall that $m_1(f, p) = 0$ trivially and $o_1(f, p) = m$. Hence, by Corollary 4, we get $m_1(f, p)/p^2 = 0$ and $o_1(f, p)/p \rightarrow 1$ as $p \rightarrow \infty$. Using this in Equation (15), we get the required result for $k = 1$.

Also, for $k \geq 2$, we note that $o_k(f, p)$ is bounded by a constant depending on f . This is because, for $k = 2$, we can bound $o_k(f, p)$ by the number of solutions of the system of congruences (in four unknowns)

$$\begin{aligned} f(x, y) &\equiv f(r, s) \equiv 0 \pmod{p} \\ f_x(x, y)f_y(r, s) - f_x(r, s)f_y(x, y) &\equiv 0 \pmod{p} \\ f_x(x, y)f(r, s) - f_x(r, s)f(x, y) &\equiv 0 \pmod{p^2}, \end{aligned}$$

and Bezout's theorem gives a bound on this number depending only on f and independent of p . For the case $k > 2$, there are more congruences than there are unknowns, and, in general, there are no solutions. In either case, $o_k(f, p)/p \rightarrow 0$ as $p \rightarrow \infty$. Using this and Conjecture 1 in Equation (15), we get the required result for $k \geq 2$. \square

4. Uniform Distribution Results

We start by describing the idea of uniform distribution of the sequence \mathcal{F}_p (defined in Equation (4)) in \mathbb{F}_p^3 , namely that for every set $E \subseteq \mathbb{F}_p^3$, the proportion of points of \mathcal{F}_p that lie in E is approximated by the proportion of $|E|$ to $|\mathbb{F}_p^3|$ for p sufficiently large, that is,

$$\frac{|\mathcal{F}_p \cap E|}{m} \approx \frac{|E|}{p^3}, \quad p \rightarrow \infty. \tag{17}$$

We shall use several measures to quantify the deviation from uniform distribution, also called discrepancy in the literature (see [5] for an excellent exposition) and define precise versions of uniform distribution.

At the outset, we describe the version of uniform distribution that provably implies Conjecture 1. For that, we give the following definition of *discrepancy*:

$$\mathcal{D}_p(\mathcal{F}_p) = \sup_{E \subseteq \mathbb{F}_p^3} \left| \frac{|\mathcal{F}_p \cap E|}{|E|} \frac{p^3}{m} - 1 \right|. \tag{18}$$

Using this, we formulate the following conjecture, which implies the uniform distribution of \mathcal{F}_p as described in Equation (17).

Conjecture 2. For $p \rightarrow \infty$, we have $\mathcal{D}_p(\mathcal{F}_p) \rightarrow 0$.

We now give the proof of Conjecture 1 assuming this version of uniform distribution.

Proposition 7. *Conjecture 2 implies Conjecture 1.*

Proof. We start by noting that Conjecture 2 can be applied to any independent choice of k vectors from \mathcal{F}_p . More precisely, Conjecture 2 implies

$$\sup_{E \subseteq (\mathbb{F}_p^3)^k} \left| \frac{|\mathcal{F}_p^{k'} \cap E|}{|E|} \frac{p^{3k}}{m(m-1)\dots(m-k+1)} - 1 \right| \rightarrow 0, \tag{19}$$

as $p \rightarrow \infty$ where $\mathcal{F}_p^{k'}$ denotes the element-wise distinct k -tuples of elements of \mathcal{F}_p .

Applying Equation (19) to $E = \{(v_1, v_2, \dots, v_k) \in (\mathbb{F}_p^3)^k \mid \text{the associated matrix has rank 2 with the first two columns linearly independent}\}$ gives

$$\left| \frac{m_k(f, p)}{|E|} \frac{p^{3k}}{m(m-1)\dots(m-k+1)} - 1 \right| \rightarrow 0.$$

Now, $|E| = p^3(p^3 - p)(p^2 - 2)(p^2 - 3)\dots(p^2 - k + 1)$ and $m/p \rightarrow 1$ as $p \rightarrow \infty$ by Corollary 4. Hence, we get (as $p \rightarrow \infty$)

$$\left| \frac{m_k(f, p)}{p^2} - 1 \right| \rightarrow 0,$$

which is the same as Equation (3). □

For comparison with Conjecture 2, we recall the standard version of discrepancy in the 3-dimensional unit cube of the sequence of \mathcal{F}_p scaled down by a factor of p in each component, which we denote by $\Delta_p(\mathcal{F}_p)$:

$$\Delta_p(\mathcal{F}_p) = \sup_{E \subseteq \mathbb{F}_p^3} \left| \frac{|\mathcal{F}_p \cap E|}{m} - \frac{|E|}{p^3} \right|. \tag{20}$$

We now state the following conjecture, which is another version of the uniform distribution as described in Equation (17).

Conjecture 3. For $p \rightarrow \infty$, we have $\Delta_p(\mathcal{F}_p) \rightarrow 0$.

It can be seen that Conjecture 2 implies Conjecture 3, as $\Delta_p(\mathcal{F}_p) \leq \mathcal{D}_p(\mathcal{F}_p)$. It is also easy to see that Conjecture 3 gives $m_{2,2}/p^2 \rightarrow 1$ as $p \rightarrow \infty$. However, it is not yet clear whether this suffices to prove Conjecture 1 in its entirety.

On the other hand, we note the connection of Conjecture 3 with exponential sums by recalling the Erdős–Turan–Koksma inequality.

Proposition 8 (Theorem 1.21, [5]). *There exists a constant $C > 0$ such that, for any positive integer L , the bound*

$$\Delta_p(\mathcal{F}_p) \leq C \left(\frac{1}{L} + \frac{1}{|\mathcal{F}_p|} \sum_{0 < |\mathbf{a}| \leq L} \frac{1}{r(\mathbf{a})} \left| \sum_{\mathbf{v} \in \mathcal{F}_p} e_p(\mathbf{a} \cdot \mathbf{v}) \right| \right),$$

holds, where the sum runs over triples $\mathbf{a} \in \mathbb{Z}^3$ which satisfy the given condition with $|\mathbf{a}| = \max_{i=1}^3 a_i$, $r(\mathbf{a}) = \prod_{i=1}^3 \max(|a_i|, 1)$, and $e_p(x) = \exp(2\pi i x/p)$.

We finally state the following conjecture on exponential sums which appears to be obtainable along the same lines as Bombieri [2].

Conjecture 4. There exists a constant $D > 0$ such that, for any prime p and any $k, l \in \{0, 1, \dots, p-1\}$, the bound

$$\sum_{(x,y) \in C_f(\mathbb{F}_p)} e_{p^2}(f(x+kp, y+lp)) \leq Dp^{\frac{1}{2}},$$

holds.

It is easy to see, using Proposition 8, that Conjecture 4 implies that for any $\epsilon > 0$, there exist a constant $E_\epsilon > 0$ such that $\Delta_p(\mathcal{F}_p) \leq E_\epsilon p^{-\frac{1}{2} + \epsilon}$, which definitely implies Conjecture 3.

Numerical computations appear to support Conjectures 2 and 4 on uniform distribution, though the speed of convergence appears to follow Conjecture 4. More precisely, we compute lower bounds for $\Delta_p(\mathcal{F}_p)$ and $\mathcal{D}_p(\mathcal{F}_p)$ in Table 1 and compare these with the order of decay predicted by Conjecture 4. These entries are obtained by restricting E to cubes of side $p/3$ in Equations (18) and (20).

p	Lower bound for $\Delta_p(\mathcal{F}_p)$	Lower bound for $\mathcal{D}_p(\mathcal{F}_p)$	$p^{-1/2}$
1009	0.01088009538	0.29376257545	0.03148142750
2003	0.01131027426	0.30537740503	0.02234392810
3001	0.00930693824	0.25128733264	0.01825437644
4003	0.00494879436	0.13361744798	0.01580546236
5443	0.00511927708	0.13822048125	0.01355441669
6007	0.00672228668	0.18150174042	0.01290242026
7013	0.00646586788	0.17457843276	0.01194120297
8011	0.00496738448	0.13411938098	0.01117266132
9007	0.00563262285	0.15208081695	0.01053682867
10009	0.00494777953	0.13359004739	0.00999550303

Table 1: Lower bounds on discrepancy for various primes for the polynomial $f(x, y) = x^3 + y^2 + xy + 1$.

5. Proof of Theorem 1

We now present the proof of Theorem 1.

Proof of Theorem 1. We start by noting that the random variable $X(k, l)$ defined in Equation (1), representing the total number of points in $C_f(\mathbb{F}_p)$ with valuation greater than 1 in a random $p \times p$ square, has the following expression:

$$X(k, l) = \sum_{(x,y) \in C_f(\mathbb{F}_p)} Y_{x,y}(k, l) = \sum_{i=1}^m Y_i(k, l).$$

Recall that $m = |C_f(\mathbb{F}_p)|$.

By the properties of the random variables $Y_{x,y}$ discussed in Section 3, we have

$$E[X] = \sum_{i=1}^m E[Y_i] = m/p.$$

In general, for $k \geq 1$

$$E[X^k] = E\left[\left(\sum_{l=1}^m Y_l\right)^k\right] = \sum_{l_1, l_2, \dots, l_k=1}^m E[Y_{l_1} Y_{l_2} \dots Y_{l_k}]. \tag{21}$$

For each term in the last sum in (21), we denote the distinct variables that appear by $Y_{q_1}, Y_{q_2}, \dots, Y_{q_i}$. Since these variables are indicator variables, we have $E[Y_{l_1} Y_{l_2} \dots Y_{l_k}] = E[Y_{q_1} Y_{q_2} \dots Y_{q_i}]$. In addition, the positions where these variables appear form a partition of $\{1, 2, \dots, k\}$ into i parts. So, for each element-wise distinct tuple $(q_1, q_2, \dots, q_i) \in \{1, 2, \dots, m\}^i$, we get $S(k, i)$ many terms in (21), where $S(k, i)$ is the Stirling number of the second kind.

Hence, (21) can be rewritten as

$$E[X^k] = \sum_{i=1}^k S(k, i) \sum_{q_1 \neq q_2 \neq \dots \neq q_i} E[Y_{q_1} Y_{q_2} \dots Y_{q_i}], \tag{22}$$

where $q_i \in \{1, 2, \dots, m\}$.

Thus, from Equation (22) and Corollary 5 (assuming Conjecture 1 as in the theorem), we get the following limit for the k^{th} moment $E[X^k]$:

$$\lim_{p \rightarrow \infty} E[X^k] = \sum_{i=1}^k S(k, i),$$

which is indeed the k^{th} moment of the Poisson distribution with parameter $\lambda = 1$ (see [9]). In other words, the moment generating function of X , $M_X(t)$, converges to

$$\lim_{p \rightarrow \infty} M_X(t) = \exp(e^t - 1),$$

the moment generating function of the Poisson random variable with parameter $\lambda = 1$. Hence, X converges in distribution to a Poisson random variable with parameter $\lambda = 1$ as $p \rightarrow \infty$. □

Acknowledgements. The authors thank the anonymous referee for valuable suggestions and insightful comments that helped greatly improve the exposition.

References

[1] L. Almodovar, A. N. Byrnes, J. Fink, X. Guan, A. Kesarwani, G. Lavigne, L. A. Medina, V. H. Moll, I. Nogues, S. Rajasekaran, E. Rowland, and A. Yuan, A closed-form solution might be given by a tree. Valuations of quadratic polynomials, *Scientia Ser. A Math. Sci.* **29** (2019), 11–28.

- [2] E. Bombieri, On exponential sums in finite fields, *Amer. J. Math.* **88** (1966), 71–105.
- [3] J. M. Borwein and R. Crandall, Closed forms: what they are and why we care, *Notices Amer. Math. Soc.* **60** (2013), 50–65.
- [4] T. Y. Chow, What is a closed-form number?, *Amer. Math. Monthly* **106** (1999), 440–448.
- [5] M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications*, Springer-Verlag, Berlin, 1997.
- [6] F. Q. Gouvêa, *p-Adic Numbers: An Introduction*, 3rd ed., Springer-Verlag, 2020.
- [7] A. M. Legendre, *Théorie des Nombres*, Firmin Didot Frères, Paris, 1830.
- [8] L. A. Medina, V. H. Moll, and E. Rowland, Periodicity in the p -adic valuation of a polynomial, *J. Number Theory* **180** (2017), 139–153.
- [9] J. Riordan, Moment recurrence relations for binomial, Poisson and hypergeometric frequency distributions, *Ann. Math. Statist.* **8** (1937), 103–111.
- [10] Shubham, The 2-adic valuation of the general degree-2 polynomial in 2 variables, *Notes Number Theory Discrete Math.* **29** (2023), 737–751.
- [11] A. Weil, *Sur les Courbes Algébriques et les Variétés qui s'en Déduisent*, Actualités Sci. Ind., no. 1041, Hermann, Paris, 1948.