p-ADIC VARIATION OF L FUNCTIONS OF ONE VARIABLE EXPONENTIAL SUMS, I

By Hui June Zhu

Abstract. For a polynomial $f(x)$ in $(\mathbb{Z}_p \cap \mathbb{Q})[x]$ of degree $d \geq 3$ let $L(f \otimes \mathbb{F}_p; T)$ be the L function of the exponential sum of f mod p. Let $\mathrm{NP}\,(f \otimes \mathbb{F}_p)$ denote the Newton polygon of $L(f \otimes \mathbb{F}_p; T)$. Let $\mathrm{HP}\,(\mathbb{A}^d)$ denote the Hodge polygon of $\mathbb{A}^d$, which is the lower convex hull in $\mathbb{R}^2$ of the points $(n, \frac{n(n+1)}{2d})$ for $0 \leq n \leq d-1$. Let $\mathbb{A}^d$ be the space of degree-d monic polynomials parameterized by their coefficients. Let $\mathrm{GNP}\,(\mathbb{A}^d; \mathbb{F}_p) := \inf_{\overline{f} \in \mathbb{A}^d(\mathbb{F}_p)} \mathrm{NP}\,(\overline{f})$ be the lowest Newton polygon over $\mathbb{F}_p$ if exists. We prove that for p large enough $\mathrm{GNP}\,(\mathbb{A}^d; \mathbb{F}_p)$ exists and we give an explicit formula for it. We also prove that there is a Zariski dense open subset $\mathcal{U}$ defined over $\mathbb{Q}$ in $\mathbb{A}^d$ such that for $f \in \mathcal{U}(\mathbb{Q})$ and for p large enough we have $\mathrm{NP}\,(f \otimes \mathbb{F}_p) = \mathrm{GNP}\,(\mathbb{A}^d; \mathbb{F}_p)$; furthermore, as p goes to infinity their limit exists and is equal to $\mathrm{HP}\,(\mathbb{A}^d)$. Finally we prove analogous results for the space of polynomials $f(x) = x^d + ax$ with one parameter. In particular, for any nonzero $a \in \mathbb{Q}$ we show that $\lim_{p \to \infty} \mathrm{NP}\,((x^d + ax) \otimes \mathbb{F}_p) = \mathrm{HP}\,(\mathbb{A}^d)$.

1. Introduction. In this paper d is an integer ≥ 3. Let $\mathbb{A}^d$ be the d-dimensional affine space identified with the space of degree-d monic polynomials parameterized by their coefficients. We always assume that p is a prime coprime to d. Let $\overline{\mathbb{Q}}_p$ and $\overline{\mathbb{Z}}_p$ be the algebraic closure of $\mathbb{Q}_p$ and its ring of integers respectively. Let $f(x)$ a polynomial of one variable in $\mathbb{A}^d(\mathbb{Z}_p \cap \mathbb{Q})$. Let $E(x) = \exp\,(\sum_{j=0}^{\infty} \frac{x^{p^j}}{p^j})$ be the Artin-Hasse exponential function. Let γ be a root of $\log\,(E(x))$ in $\overline{\mathbb{Q}}_p$ with $\mathrm{ord}_p\,\gamma = \frac{1}{p-1}$. Then $E(\gamma)$ is a primitive p-th root of unity. Denote it by ζ_p. It is observed that $\mathbb{Z}_p[\gamma] = \mathbb{Z}_p[\zeta_p]$. For every $\ell \geq 1$, recall the exponential sums of the reduction $f \otimes \mathbb{F}_p$ of f modulo p

$$S_\ell(f \otimes \mathbb{F}_p) := \sum_{x \in \mathbb{F}_{p^\ell}} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_{p^\ell}/\mathbb{F}_p}(f(x) \otimes \mathbb{F}_p)}.$$

The L-function of the exponential sum of $f \otimes \mathbb{F}_p$ is defined by

$$L(f \otimes \mathbb{F}_p; T) := \exp\left(\sum_{\ell=1}^{\infty} S_\ell(f \otimes \mathbb{F}_p) \frac{T^\ell}{\ell}\right). \tag{1}$$

Manuscript received December 3, 2001; revised August 2, 2002.

Research supported in part by U. C. Berkeley, and a grant of Bjorn Poonen from the David and Lucile Packard Foundation.

American Journal of Mathematics 125 (2003), 669–690.

It is well known (or simply using the Weil Conjecture for curves combined with (3) below) that

$$L(f\otimes\mathbb{F}_p;T)=1+b_1T+b_2T^2+\cdots+b_{d-1}T^{d-1}\in\mathbb{Z}[\zeta_p][T]. \tag{2}$$

Let $\mathrm{ord}_p(\cdot)$ denote the unique extension of the (additive) p-adic valuation in $\mathbb{Q}_p$ to $\overline{\mathbb{Q}}_p$. We also denote by $\mathrm{ord}_p(\cdot)$ the p-adic valuation of the *content of a power series over* $\overline{\mathbb{Z}}_p$ (see [10, pages 209 and 181] for its standard definition). Define the *Newton polygon* of the L-function of $f\otimes\mathbb{F}_p$, denoted by $\mathrm{NP}(f\otimes\mathbb{F}_p)$, as the lower convex hull of the points $(n,\mathrm{ord}_p\, b_n)$ in $\mathbb{R}^2$ for $0\le n\le d-1$, where we set $b_0=1$. The *Hodge polygon* of f, denoted by $\mathrm{HP}(\mathbb{A}^d)$, is the lower convex hull in $\mathbb{R}^2$ of the points $(n,\frac{n(n+1)}{2d})$ for $0\le n\le d-1$. It is known that $\mathrm{HP}(\mathbb{A}^d)$ is a lower bound of $\mathrm{NP}(f\otimes\mathbb{F}_p)$ (see [17, Propositions 2.2 or 2.3]) and that if $p\equiv 1 \bmod d$ then $\mathrm{NP}(f\otimes\mathbb{F}_p)=\mathrm{HP}(\mathbb{A}^d)$ for every $f\in\mathbb{A}^d(\mathbb{Q})$ (see [16, (3.11)]).

The main results of this paper are Theorems 1.1, 5.1 and 6.2. Theorem 1.1 was a conjecture of Daqing Wan, proposed in the following form in the number theory seminar at Berkeley in the fall of 2000 (see also [18, Section 2.5] for developments related to this topic). This theorem follows from Theorem 5.1.

THEOREM 1.1. *There is a Zariski dense open subset $\mathcal{U}$ defined over $\mathbb{Q}$ in $\mathbb{A}^d$ such that for all $f(x)\in\mathcal{U}(\mathbb{Q})$ we have*

$$\lim_{p\to\infty}\mathrm{NP}(f\otimes\mathbb{F}_p)=\mathrm{HP}(\mathbb{A}^d).$$

Remark 1.2. The case $d=3$ follows from [16, (3.14)], and the case $d=4$ is discussed in [6, Corollary 4.7]. The first slope case was proved recently by an elementary method in [12] (see also [13]). Results concerning Wan's conjecture "over $\overline{\mathbb{Q}}$" are forthcoming in [22].

Theorem 1.1 yields an answer toward questions (in one variable case) proposed by Katz which asked how the Newton polygon of the L function of exponential sums varies with the prime p (see Katz's questions and Sperber's example on page 151 of [7, Chapter 5.1]). For more developments in these directions see [14], [15] and [2] and their bibliographies.

Let X_f: $y^p-y=f(x)\otimes\mathbb{F}_p$ be an Artin-Schreier curve over $\mathbb{F}_p$. The Newton polygon of X_f, denoted by $\mathrm{NP}(X_f\otimes\mathbb{F}_p)$, is the p-adic Newton polygon of the numerator of the Zeta function $\mathrm{Zeta}(X_f\otimes\mathbb{F}_p;T)$ of X_f over $\mathbb{F}_p$. It is well known that (see, for example, [3, Section VI, (93)])

$$\mathrm{Zeta}(X_f\otimes\mathbb{F}_p;T)=\frac{\mathrm{N}_{\mathbb{Q}(\zeta_p)/\mathbb{Q}}(L(f\otimes\mathbb{F}_p;T))}{(1-T)(1-pT)}, \tag{3}$$

where the norm $\mathrm{N}_{\mathbb{Q}(\zeta_p)/\mathbb{Q}}$ being interpreted as the product of the conjugates of $L(f\otimes\mathbb{F}_p;T)$ in $\mathbb{Q}(\zeta_p)$ over $\mathbb{Q}$, the automorphism acting trivially on the variable T. Thus $\mathrm{NP}(f\otimes\mathbb{F}_p)$ is precisely equal to $\mathrm{NP}(X_f\otimes\mathbb{F}_p)$ shrunk by a factor of $\frac{1}{p-1}$

horizontally and vertically, which is denoted by $\frac{\mathrm{NP}\,(X_f\otimes\mathbb{F}_p)}{p-1}$. From these remarks the following corollary is obvious.

COROLLARY 1.3. *There exists a Zariski dense open subset $\mathcal{U}$ defined over $\mathbb{Q}$ in $\mathbb{A}^d$ such that for every $f\in\mathcal{U}(\mathbb{Q})$ we have*

$$\lim_{p\to\infty}\frac{\mathrm{NP}\,(X_f\otimes\mathbb{F}_p)}{p-1}=\mathrm{HP}\,(\mathbb{A}^d).$$

Remark 1.4. (a) The behavior of $\mathrm{NP}\,(x^d\otimes\mathbb{F}_p)$ is well understood. For the reader's convenience we describe them briefly below. Let σ be a permutation in the symmetric group S_{d-1} such that for every $1\leq n\leq d-1$ we let $\sigma(n)$ be the least positive residue of $pn \bmod d$. Write σ as a product of disjoint cycles (including 1-cycles). Let σ_i be a ℓ_i-cycle in σ. Let $\lambda_i:=\sum n/(d\ell_i)$ where the sum ranges over all n in the standard representation of the ℓ_i-cycle σ_i. Arrange σ_i in such an order that $\lambda_1\leq\lambda_2\leq\cdots$. For every σ_i in σ let the pair (λ_i,ℓ_i) of rational numbers represent the line segment of (horizontal) length ℓ_i and of slope λ_i. The *joint of line segments* (λ_i,ℓ_i) is the lower convex hull consisting of line segment (λ_i,ℓ_i)'s connected at their end-points. The eigenvalues of $L(x^d\otimes\mathbb{F}_p;T)$ are Gauss sums (see [8, chapter III]). By the Stickelberger theorem (see [20, Chapter 6]), the p-adic Newton polygon of $L(x^d\otimes\mathbb{F}_p;T)$ and hence $\mathrm{NP}\,(x^d\otimes\mathbb{F}_p)$ is the joint of (λ_i,ℓ_i)'s. (I thank Kiran Kedlaya for discussions here.)

(b) For every $d\geq 3$ the Newton polygon $\mathrm{NP}\,(x^d\otimes\mathbb{F}_p)$ does not have a limit as p approaches ∞. Indeed, it is clear from the above that for $p\equiv 1 \bmod d$ the Newton polygon is equal to the Hodge polygon while for $p\equiv -1 \bmod d$ the Newton polygon is a straight line of slope $1/2$.

This paper is organized as follows. In Section 2 notations and terminologies are introduced. Using Dwork's p-adic analysis, we define the *Fredholm polygon* of $f(x)$ over $\mathbb{F}_p$ and show that it is equal to $\mathrm{NP}\,(f\otimes\mathbb{F}_p)$. Section 3 is a key step in the proof; it constructs an nth *generic polynomial*, denoted by f_n^{tn}, proves that they are nonzero and hence defines some Zariski dense open subset $\mathcal{V}_n$ in $\mathbb{A}^{d-1}$. It is recommended that the reader skip Section 3 at first and continue with Section 4, where we immediately apply Dwork's p-adic theory to determine the Fredholm polygon. In Section 5 we prove that in some Zariski dense open subset the Fredholm polygon and Newton polygon coincide if p is large enough. We prove Theorem 5.1 there. Finally in Section 6 we study the families $f(x)=x^d+ax$ and prove Theorem 6.2 there.

Acknowledgments. It is my great pleasure to thank Alan Adolphson and Steven Sperber who exposed me to Dwork theory during the Dwork trimester in Italy (2001). I thank Hanfeng Li, Daqing Wan, and the referees for careful reading and very helpful comments on earlier versions. Most of all I thank Bjorn Poonen for generously sharing ideas and answering questions.

2. Dwork p-adic theory. The fundamental material in our exposition follows [3, Sections II and III] (see also [5] [4] and [1]). Recall that p is a prime number coprime to d. Let $\overline{f}(x) = x^d + \sum_{i=1}^{d-1} \overline{a}_i x^i \in \mathbb{F}_p[x]$. Let $f(x) = x^d + \sum_{i=1}^{d-1} a_i x^i \in (\mathbb{Z}_p \cap \mathbb{Q})[x]$ and $a_d = 1$ such that reduction of $f(x)$ at p is equal to $\overline{f}(x)$. For any $a_0 \in \mathbb{Z}_p \bigcap \mathbb{Q}$, by a simple computation with (1), one easily concludes that $L((f + a_0) \otimes \mathbb{F}_p; T) = L(f \otimes \mathbb{F}_p; \zeta_p^{a_0} T)$. Thus we have

$$\mathrm{NP}\,((f + a_0) \otimes \mathbb{F}_p) = \mathrm{NP}\,(f \otimes \mathbb{F}_p). \tag{4}$$

Write $\vec{\hat{a}} = (\hat{a}_1, \ldots, \hat{a}_{d-1})$ where $\hat{a}_i$ is the Teichmüller lifting of $\overline{a}_i$, that is, $\hat{a}_i \equiv a_i \bmod p$ and $\hat{a}_i^p = \hat{a}_i$. Let $\vec{a} := (a_1, \ldots, a_{d-1}) \in (\mathbb{Z}_p \cap \mathbb{Q})^{d-1}$. Let $\theta(x) = E(\gamma x)$, where $E(\cdot)$ and γ are as defined in Section 1. Then we may write $\theta(x) = \sum_{m=0}^{\infty} \lambda_m x^m$ for $\lambda_m \in \mathbb{Z}_p[\zeta_p]$. Note the following properties,

$$\mathrm{ord}_p\, \lambda_m \geq \frac{m}{p-1}; \tag{5}$$

for $0 \leq m \leq p-1$ we have,

$$\lambda_m = \frac{\gamma^m}{m!} \text{ and } \mathrm{ord}_p\, \lambda_m = \frac{m}{p-1}. \tag{6}$$

Let $\vec{A} = (A_1, \ldots, A_{d-1})$ be a vector of variables and $\vec{m} = (m_1, \ldots, m_{d-1})$. Write $\vec{A}^{\vec{m}}$ for the monomial $A_1^{m_1} \cdots A_{d-1}^{m_{d-1}}$. Let $G_n(\vec{A}) = 0$ for $n < 0$. For every integer $n \geq 0$ let

$$G_n(\vec{A}) := \sum_{\substack{m_\ell \geq 0 \\ \sum_{\ell=1}^{d} \ell m_\ell = n}} \lambda_{m_1} \cdots \lambda_{m_d} \vec{A}^{\vec{m}}. \tag{7}$$

Clearly we observe that $G_n(\vec{A}) \in \mathbb{Z}_p[\zeta_p][\vec{A}]$, that is, $G_n(\vec{A})$ is a polynomial in variable $\vec{A}$ and with coefficients in $\mathbb{Z}_p[\zeta_p]$. For all integers $m_1, \ldots, m_d \geq 0$ such that $\sum_{\ell=1}^{d} \ell m_\ell = n$, we have $d(m_1 + \cdots + m_d) \geq \sum_{\ell=1}^{d} \ell m_\ell = n$ and so $\min\,(m_1 + \cdots + m_d) = \lceil \frac{n}{d} \rceil$. Therefore by (7) we have

$$\mathrm{ord}_p\, G_n(\vec{A}) \geq \frac{\min\,(m_1 + \cdots + m_d)}{p-1} \geq \frac{\lceil \frac{n}{d} \rceil}{p-1} \geq \frac{n}{d(p-1)}. \tag{8}$$

Let $G(X) := \prod_{i=1}^{d} \theta(\hat{a}_i X^i) \in \mathbb{Z}_p[\zeta_p][[X]]$. We have

$$G(X) = \left(\sum_{m_1=0}^{\infty} \lambda_{m_1} \hat{a}_1^{m_1} X^{m_1} \right) \cdots \left(\sum_{m_d=0}^{\infty} \lambda_{m_d} \hat{a}_d^{m_d} X^{d m_d} \right) = \sum_{n=0}^{\infty} G_n(\vec{\hat{a}}) X^n.$$

Let $C_0(\vec{A}) = 1$, and for every $n \geq 1$ let

$$C_n(\vec{A}) := \sum_{1\leq u_1<u_2<\cdots<u_n} \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} G_{pu_i - u_{\sigma(i)}}(\vec{A}), \tag{9}$$

where $\operatorname{sgn}(\sigma)$ is the signature of the permutation σ in the nth symmetric group S_n. It can be verified that this definition makes sense and that $C_n(\vec{A}) \in \mathbb{Z}_p[\zeta_p][[\vec{A}]]$.

LEMMA 2.1. *Let p be a prime coprime to d. For $f(x) = x^d + \sum_{i=1}^{d-1} a_i x^i \in (\mathbb{Z}_p \cap \mathbb{Q})[x]$, write $\vec{a} = (a_1, \ldots, a_{d-1})$. Let $\vec{\hat{a}} = (\hat{a}_1, \ldots, \hat{a}_{d-1})$ be Teichmüller lifting of $\vec{\overline{a}} = (\overline{a}_1, \ldots, \overline{a}_{d-1})$. Then*

$$\begin{aligned} L(f\otimes\mathbb{F}_p;T) &= 1 + b_1(\vec{a})T + \cdots + b_{d-1}(\vec{a})T^{d-1} \\ &= \frac{1+\sum_{n=1}^{\infty}(-1)^n C_n(\vec{\hat{a}})T^n}{(1-pT)(1+\sum_{n=1}^{\infty}(-1)^n C_n(\vec{\hat{a}})p^nT^n)}, \end{aligned} \tag{10}$$

where $b_1(\vec{a}), \ldots, b_{d-1}(\vec{a}) \in \mathbb{Z}[\zeta_p]$.

Proof. The first equality is a rephrasing of (2). For every positive integer ℓ let

$$S_\ell^*(f\otimes\mathbb{F}_p) := \sum_{x\in\mathbb{F}_{p^\ell}^*} \zeta_p^{\operatorname{Tr}_{\mathbb{F}_{p^\ell}/\mathbb{F}_p}(f(x)\otimes\mathbb{F}_p)}.$$

Let

$$L^*(f\otimes\mathbb{F}_p;T) := \exp\left(\sum_{\ell=1}^{\infty} S_\ell^*(f\otimes\mathbb{F}_p)\frac{T^\ell}{\ell}\right).$$

Note that $S_\ell^*(f\otimes\mathbb{F}_p) = S_\ell(f\otimes\mathbb{F}_p) - 1$ so

$$\begin{aligned} L^*(f\otimes\mathbb{F}_p;T) &= \exp\left(\sum_{\ell=1}^{\infty}(S_\ell(f\otimes\mathbb{F}_p)-1)\right) \\ &= (1-T)\exp\left(\sum_{\ell=1}^{\infty} S_\ell(f\otimes\mathbb{F}_p)\frac{T^\ell}{\ell}\right) \\ &= (1-T)L(f\otimes\mathbb{F}_p;T). \end{aligned} \tag{11}$$

For any $c > 0$ and $b \in \mathbb{R}$ let $\mathcal{L}(c,b)$ be the set of power series defined by

$$\mathcal{L}(c,b) := \left\{ \sum_{n=0}^{\infty} A_n X^n \mid A_n \in \mathbb{Q}_p(\zeta_p), \operatorname{ord}_p A_n \geq \frac{cn}{d} + b \right\}.$$

Let $\mathcal{L}(c) := \bigcup_{b\in\mathbb{R}} \mathcal{L}(c,b)$. From (8) we have $G(X) = \sum_{n=0}^{\infty} G_n(\vec{a})X^n$ lie in $\mathcal{L}(1/(p-1))$. For any $\sum B_nX^n$ in $\mathcal{L}(c)$, let ψ be the Hecke operator from $\mathcal{L}(c)$ to $\mathcal{L}(cp)$ given by $\psi(\sum B_nX^n) = \sum B_{pn}X^n$. Let $\alpha_1 := \psi \cdot G(X)$ be the endomorphism of $\mathcal{L}(p/(p-1))$ defined by the composition of the multiplication map by $G(X)$ then ψ, namely,

$$\alpha_1\left(\sum_{i=0}^{\infty} B_iX^i\right) = \sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} G_{pi-j}(\vec{a})B_j\right)X^i.$$

Choose the standard monomial basis $\{1, x, x^2, \ldots\}$ for the p-adic space $\mathcal{L}(p/(p-1))$. Then the $\mathbb{Q}_p(\zeta_p)$-endomorphism α_1 of $\mathcal{L}(p/(p-1))$ has a matrix representation by $\{G_{pi-j}(\vec{a})\}_{i,j\geq 0}$. We denote this matrix by F_1. By the Dwork trace formula (see [3, Section III]) we have

$$L^*(f\otimes\mathbb{F}_p;T) = \frac{\det(1-F_1T)}{\det(1-F_1pT)}.$$

For the first row (i.e., $i=0$) of F_1, we have $G_{pi-j}(\vec{a}) = 0$ for all $j \geq 1$ and $G_0(\vec{a}) = 1$. By (9) we have

$$\det(1-F_1T) = (1-T)\det(1-\{G_{pi-j}(\vec{a})T\}_{i,j\geq 1}) = (1-T)\sum_{n=0}^{\infty}(-1)^nC_n(\vec{a})T^n.$$

Therefore, by (11) we have

$$(1-T)L(f\otimes\mathbb{F}_p;T) = L^*(f\otimes\mathbb{F}_p;T) = \frac{(1-T)\sum_{n=0}^{\infty}(-1)^nC_n(\vec{a})T^n}{(1-pT)\sum_{n=0}^{\infty}(-1)^nC_n(\vec{a})p^nT^n}.$$

By simplification of the above formula, our assertion follows. □

Proposition 2.2. *Let the* Fredholm polygon *of* $f\otimes\mathbb{F}_p$, *denoted by* $\mathrm{FP}(f\otimes\mathbb{F}_p)$, *be the lower convex hull of points* $(n, \mathrm{ord}_p\, C_n(\vec{a}))$ *in* $\mathbb{R}^2$ *for* $0\leq n\leq d-1$. *Then*

$$\mathrm{NP}(f\otimes\mathbb{F}_p) = \mathrm{FP}(f\otimes\mathbb{F}_p).$$

Proof. By (10) we have

$$L(f\otimes\mathbb{F}_p;T)(1-pT)(1-C_1pT+C_2p^2T^2-\cdots) = 1-C_1T+C_2T^2-\cdots.$$

The (p-adic) Newton polygon of $1-C_1T+C_2T^2-\cdots$ has only positive slopes (see [3, III]), so the Newton polygon of $1-C_1pT+C_2p^2T^2-\cdots$ has every slope >1. On the other hand, the Newton polygon of $L(f\otimes\mathbb{F}_p;T)$ is symmetric in the sense that for every slope segment α there is a slope segment $1-\alpha$ of the same horizontal length. This property is derived from the same fact for Newton polygons of Zeta functions of abelian varieties and hence of Artin-Schreier curves (see, for example, [11, Introduction]). Thus the slopes of $\mathrm{NP}(f\otimes\mathbb{F}_p)$ are positive and <1. Note that the power series $1-C_1T+C_2T^2-\cdots$ is entire (see [9, page 121] for a proof), so are the three factors on the left-hand side. By the p-adic Weierstrass preparation theorem (see [9, IV.4 Theorem 14]), $\mathrm{NP}(f\otimes\mathbb{F}_p)$ coincides with the p-adic Newton polygon of $1-C_1T+\cdots+(-1)^{d-1}C_{d-1}T^{d-1}$. □

We remark that it is *not* generally true that $L(f\otimes\mathbb{F}_p;T)=1-C_1(\vec{a})T+\cdots+(-1)^{d-1}C_{d-1}(\vec{a})T^{d-1}$.

3. Generic polynomials and Zariski dense subsets.

The following notations and conventions are adopted for the remainder of this section. Given a polynomial as a sum (or several sums) of polynomials, its *formal expansion* means the formal summation of its monomials (so one does not do *"arithmetic,"* e.g, cancellations, among its terms). For any $\vec{m}=(m_1,\ldots,m_{d-1})\in\mathbb{Z}_{\geq 0}^{d-1}$ let $|\vec{m}|=\sum_{\ell=1}^{d-1}m_\ell$ and $\vec{m}!=m_1!\cdots m_{d-1}!$. Fix an integer r with $1\leq r\leq d-1$ and $\gcd(d,r)=1$. Let $1\leq n\leq d-1$.

3.1. The residue matrix $\mathbf{r}_n$

Let $1\leq i,j\leq d-1$. Let r_{ij} be the least nonnegative residue of $-(ri-j)\bmod d$. That is, $r_{ij}:=d\left\lceil\frac{ri-j}{d}\right\rceil-(ri-j)$. Let r'_{ij} be the least nonnegative residue of $ri-j\bmod d$.

Lemma 3.1. *Let $\mathbf{r}_n$ be the matrix $\mathbf{r}_n:=\{r_{ij}\}_{1\leq i,j\leq n}$. Then $0\leq r_{ij}\leq d-1$ and there are no two identical entries in any row (column). In $\mathbf{r}_{d-1}$ for every $1\leq i\leq d-1$ one has $r_{ij}=0$ if and only if $j=r'_{i1}+1$.*

Proof. By definition, r_{ij} is the least nonnegative residue of $-(ri-j)\bmod d$ so we have $0\leq r_{ij}\leq d-1$. We prove for rows. The argument for columns is almost identical. Suppose we have $r_{ij}=r_{ij'}$ then $ri-j\equiv ri-j'\bmod d$ by definition. Then $j\equiv j'\bmod d$. Since $1\leq j,j'\leq n\leq d-1$ we have $j=j'$. So there are no identical entries in any row of $\mathbf{r}_n$. Note that r is coprime to d so for every $1\leq i\leq d-1$ there is a unique $1\leq j\leq d-1$ (more precisely $j=r'_{i1}+1$) such that $ri\equiv j\bmod d$. This is equivalent to $r_{ij}=0$ by definition. This proves the last assertion. □

Let A_d be an auxiliary variable. Define a homogeneous auxiliary polynomial $\dot{D}_n:=\det(\{A_{d-r_{ij}}\}_{1\leq i,j\leq n})$ of degree n in $\mathbb{Q}[A_1,\ldots,A_d]$. Note that

$$\dot{D}_n=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{i=1}^{n}A_{d-r_{i,\sigma(i)}}=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{k=0}^{d-1}A_{d-k}^{\#\{1\leq i\leq n|r_{i,\sigma(i)}=k\}}. \tag{12}$$

Lemma 3.2. *There is a unique highest-lexicographic-order-monomial in the formal expansion of $\dot{D}_n$ in $\mathbb{Q}[A_1,\ldots,A_d]$.*

Proof. It is a combinatorial problem and we shall give an intuitive proof. We shall do so by verifying the correctness of the following algorithm which can really be used to obtain the desired highest-lexicographic-order-monomial.

Fix a residue matrix $\mathbf{r}_n$. Let σ be a permutation in S_n awaiting to be defined. For every entry in $\mathbf{r}_n$ with $r_{i_0,j_0}=0$, assign $\sigma(i_0):=j_0$ and cross off the i_0-row and the j_0-column; Let ℓ_0 be the number of all such entries. For every leftover entry in $\mathbf{r}_n$ with $r_{i_1,j_1}=1$, assign $\sigma(i_1):=j_1$ and cross off the i_1-row and the j_1-column; Let ℓ_1 be the number of all such entries. Continue this process until all entries are crossed off.

It is straightforward to verify that this algorithm uniquely defines a permutation σ by the first statement in Lemma 3.1. Moreover, σ yields the highest-lexicographic-order-monomial. Indeed, from (12) one notes that ℓ_0 is the highest-A_d-exponent in the formal expansion of $\dot{D}_n$; and ℓ_1 is the highest-A_{d-1}-exponent in a monomial containing $A_d^{\ell_0}$; and so on. Thus σ yields the (unique) highest-lexicographic-order-monomial $A_d^{\ell_0}A_{d-1}^{\ell_1}\cdots A_1^{\ell_{d-1}}$ in the formal expansion of $\dot{D}_n$. □

LEMMA 3.3. *Let M be the (unique) highest-lexicographic-order-monomial of formal expansion of $\dot{D}_n$ derived in Lemma 3.2. Then $M|_{A_d=1}$ is the (unique) highest-lexicographic-order-monomial of lowest degree in the formal expansion of $\dot{D}_n|_{A_d=1}$.*

Proof. It is clear that the evaluation map $\dot{D}_n\longrightarrow \dot{D}_n|_{A_d=1}$ (on the formal expansions) yields a bijective map sending the set of highest-A_d-exponents monomials in the formal expansion of $\dot{D}_n$ to the set of lowest-degree-monomials in the formal expansion of $\dot{D}_n|_{A_d=1}$. Applying the same argument for the rest of the variables inductively, we conclude our assertion immediately. □

3.2. The nth generic polynomial f_n^{tn} For any $0\le s\le n$ one obtains a nonempty subset in $\mathbb{Z}_{\ge 0}^{d-1}$

$$\mathcal{M}_{ij}^s := \left\{\vec{m}=(m_1,m_2,\dots,m_{d-1})\in\mathbb{Z}_{\ge0}^{d-1} \mid \sum_{\ell=1}^{d-1}\ell m_{d-\ell}=r_{ij}+ds\right\}.$$

Recall $\vec{A}:=(A_1,\dots,A_{d-1})$, and $\vec{A}^{\vec{m}}:=A_1^{m_1}\cdots A_{d-1}^{m_{d-1}}$. For $1\le i,j\le d-1$ let

$$\delta_{ij}:=\begin{cases}0 & \text{for } j<r'_{i1}+1\\ 1 & \text{for } j\ge r'_{i1}+1.\end{cases} \tag{13}$$

For $0\le s\le n$ and $1\le i,j\le n$ define an auxiliary polynomial

$$H_{ij}^s(\vec{A}):=\sum_{\vec{m}\in\mathcal{M}_{ij}^s} h_{\vec{m},i,j}^s\vec{A}^{\vec{m}} \tag{14}$$

where

$$h^s_{\vec{m},i,j} := \frac{(\frac{r_{i1}-1}{d}+n)(\frac{r_{i1}-1}{d}+n-1)\cdots(\frac{r_{i1}-1}{d}-\delta_{ij}+s+1-|\vec{m}|)}{\vec{m}!}.$$

LEMMA 3.4. *Let $0 \le s \le n$ and $1 \le i,j \le n$.*

(a) *The polynomial $H^s_{ij}(\vec{A})$ in $\mathbb{Q}[\vec{A}]$ is nonzero and is supported on every $\vec{m} \in \mathcal{M}^s_{ij}$. The degree of its monomials ranges from $s + \left\lceil \frac{r_{ij}+s}{d-1} \right\rceil$ up to $r_{ij} + ds$, where the maximal degree is attained at exactly one monomial $A_{d-1}^{r_{ij}+ds}$ while the minimal degree is attained at one or more monomials.*

(b) *The polynomial $H^s_{ij}(\vec{A})$ has a constant term if and only if $s = r_{ij} = 0$; it has a linear term if and only if $s = 0$ and $r_{ij} \ne 0$, in which case this linear monomial is exactly $A_{d-r_{ij}}$.*

Proof. (a) Since $\gcd(r,d) = 1$ we have $-(ir-1) \not\equiv 1 \bmod d$. Hence $\frac{r_{i1}-1}{d} \notin \mathbb{Z}$ and $h^s_{\vec{m},i,j} \ne 0$. Now it remains to show

$$\max_{\vec{m}\in\mathcal{M}^s_{ij}} |\vec{m}| = r_{ij} + ds, \qquad \min_{\vec{m}\in\mathcal{M}^s_{ij}} |\vec{m}| = s + \left\lceil \frac{r_{ij}+s}{d-1} \right\rceil.$$

For $\vec{m} \in \mathcal{M}^s_{ij}$ we have $|\vec{m}| \le \sum_{\ell=1}^{d-1} \ell m_{d-\ell} = r_{ij} + ds$ and the equality holds precisely for $m_1 = \cdots = m_{d-2} = 0$ and $m_{d-1} = r_{ij} + ds$. For $\vec{m} \in \mathcal{M}^s_{ij}$ one has clearly $(d-1)|\vec{m}| \ge \sum_{\ell=1}^{d-1} \ell m_{d-\ell} = r_{ij} + ds$. So

$$|\vec{m}| \ge \left\lceil \frac{r_{ij}+ds}{d-1} \right\rceil = s + \left\lceil \frac{r_{ij}+s}{d-1} \right\rceil.$$

It is easy to see that there are $m_1, \ldots, m_{d-1} \ge 0$ satisfying

$$\sum_{\ell=1}^{d-1} m_\ell = \left\lceil \frac{r_{ij}+ds}{d-1} \right\rceil \quad \text{and} \quad \sum_{\ell=1}^{d-1} (d-\ell) m_\ell = r_{ij} + ds. \tag{15}$$

For example, let κ be the least nonnegative residue of $-(r_{ij} + ds) \bmod (d-1)$ then let $m_1 = \left\lceil \frac{r_{ij}+ds}{d-1} \right\rceil - \kappa$, $m_2 = \kappa$ and let the rest $m_\ell = 0$. This says that there are $\vec{m} \in \mathcal{M}^s_{ij}$ with $|\vec{m}| = \left\lceil \frac{r_{ij}+ds}{d-1} \right\rceil$.

(b) Suppose $H^s_{ij}(\vec{A})$ has a linear term, then by part (a) we have $s + \left\lceil \frac{r_{ij}+s}{d-1} \right\rceil = 1$, which implies $s = 0$ and $r_{ij} \ne 0$. In this case the only solution to (15) is $m_{d-r_{ij}} = 1$ and $m_\ell = 0$ for all $\ell \ne r_{ij}$. So the linear monomial is $A_{d-r_{ij}}$. In the same vein we obtain the assertion about the constant term. □

For $1 \leq n \leq d-1$ and $0 \leq t \leq c_n$, let

$$c_n := \frac{1}{d}\left(\max_{\sigma\in S_n}\sum_{i=1}^{n} r_{i,\sigma(i)} - \min_{\sigma\in S_n}\sum_{i=1}^{n} r_{i,\sigma(i)}\right); \tag{16}$$

$$S_n^t := \left\{\sigma \in S_n \mid \sum_{i=1}^{n} r_{i,\sigma(i)} = \min_{\sigma\in S_n}\sum_{i=1}^{n} r_{i,\sigma(i)} + dt\right\}; \tag{17}$$

$$f_n^t(\vec{A}) := \sum_{\substack{s_0+s_1+\cdots+s_n=t\\ s_0,\ldots,s_n\geq 0}} \sum_{\sigma\in S_n^{s_0}} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} H_{i,\sigma(i)}^{s_i}(\vec{A}). \tag{18}$$

Note that $c_n \leq n$. The polynomial $f_n^t(\vec{A}) \in \mathbb{Q}[\vec{A}]$ will play a central role in this paper.

LEMMA 3.5. (Key-Lemma) *Let* $1 \leq n \leq d-1$. *Then there exists t with* $0 \leq t \leq c_n$ *such that the polynomial* $f_n^t(\vec{A}) \neq 0$. *Let t_n be the least such t. Let $\mathcal{V}_n$ be the complement in $\mathbb{A}^{d-1}$ of the variety defined by* $f_n^{t_n} = 0$. *Then $\mathcal{V}_n$ is a Zariski dense open subset defined over $\mathbb{Q}$ of $\mathbb{A}^{d-1}$.*

Proof. It suffices to prove the first assertion. We first show that among the lowest-degree-terms in the formal expansion of $\sum_{t=0}^{c_n} f_n^t$ there is a unique highest-lexicographic-order-monomial. This suffices because the polynomial f_n^t (for some t) whose formal expansion contains this unique monomial has to be nonzero.

Partition the summands of the formal expansion of $\sum_{t=0}^{c_n} f_n^t$ into two parts:

$$\sum_{t=0}^{c_n} f_n^t = \sum_{t=0}^{c_n}\sum{}' \sum_{\sigma\in S_n^{s_0}} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} H_{i,\sigma(i)}^{s_i}(\vec{A}) + \sum_{t=0}^{c_n}\sum{}'' \sum_{\sigma\in S_n^{s_0}} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} H_{i,\sigma(i)}^{s_i}(\vec{A})$$

where $\sum'$ ranges over the set of all $s_0,\ldots,s_n \geq 0$ with $s_1 = \cdots = s_n = 0$ and $s_0 = t$ while $\sum''$ ranges over the set of all $s_0,\ldots,s_n \geq 0$ with $s_0 + \cdots + s_n = t$ and $s_\ell \geq 1$ for some $\ell = 1,\ldots,n$. Denote them by $H'(\vec{A})$ and $H''(\vec{A})$, respectively. Note that $S_n = \bigcup_{t=0}^{c_n} S_n^t$, by which we find

$$H' = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} H_{i,\sigma(i)}^{0}(\vec{A}).$$

Let μ, μ' and μ'' denote the lowest degrees in the formal expansions of $\sum_{t=0}^{c_n} f_n^t$, H' and H'', respectively. By Lemma 3.4(a), we have $\mu'' = \sum_{i=1}^{n}\left(s_i + \left\lceil \frac{r_{i,\sigma'(i)}+s_i}{d-1}\right\rceil\right)$

for some $\sigma' \in S_n$. By the definition of H'' we have $s_i \geq 1$ for some $1 \leq i \leq n$. Thus $\sum_{i=1}^{n} \left\lceil \frac{r_{i,\sigma'(i)}}{d-1} \right\rceil < \mu''$. On the other hand, we have

$$\mu \leq \mu' = \min_{\sigma \in S_n} \sum_{i=1}^{n} \left\lceil \frac{r_{i,\sigma(i)}}{d-1} \right\rceil \leq \sum_{i=1}^{n} \left\lceil \frac{r_{i,\sigma'(i)}}{d-1} \right\rceil .$$

Combining these above, we have $\mu \leq \mu' < \mu''$. Hence $\mu = \mu' < \mu''$ and it follows that all degree-μ monomials in the formal expansion of $\sum_{t=0}^{c_n} f_n^t$ lie in the formal expansion of H'.

Recall from Lemma 3.4(b) that for every i the lowest-degree-monomial of $H_{i,\sigma(i)}^0$ is 1 or $A_{d-r_{i,\sigma(i)}}$ depending on $r_{i,\sigma(i)} = 0$ or not, respectively. Then the set of degree-μ monomials of the formal expansion of H' is equal to the set of degree-μ monomials in $\dot{D}_n|_{A_d=1}$ by a perusal of the definition of $\dot{D}_n$ in (12). This finishes the proof by Lemma 3.3. □

4. Fredholm polygons. Let notations be as in previous sections. This section will study the shape of Fredholm polygons of $f \in \mathbb{A}^{d-1}$. We do this by considering the p-adic valuation of the content of $G_{pi-j}(\vec{A}) \in \mathbb{Z}_p[\zeta_p][\vec{A}]$ and that of the $C_n(\vec{A}) \in \mathbb{Z}_p[\zeta_p][[\vec{A}]]$. We shall consider $G_{pi-j}(\vec{A})$ as formal expressions in $\mathbb{Z}_p[\vec{A}][\gamma]$.

Throughout this section we adopt the following convention. Fix an integer r with $1 \leq r \leq d-1$ and $\gcd(r,d) = 1$. Let p be a prime that $p \equiv r \bmod d$. Let $\vec{a} = (a_1, \ldots, a_{d-1}) \in (\mathbb{Q} \cap \mathbb{Z}_p)^{d-1}$. Let n be an integer with $1 \leq n \leq d-1$. For any rational number R let $\gamma^{>R}$ denote the terms in $\mathbb{Q}_p(\zeta_p)[[\vec{A}]]$ whose coefficients have p-adic valuation $> R/(p-1)$. We also use it to denote algebraic numbers in $\mathbb{Q}_p(\zeta_p)$ with p-adic valuation $> R/(p-1)$ and this should not cause any confusion. We define $\gamma^{\geq R}$ analogously. Let

$$M_n := \min_{\sigma \in S_n} \sum_{i=1}^{n} \left\lceil \frac{pi - \sigma(i)}{d} \right\rceil .$$

LEMMA 4.1. *For any $s \geq 0$ we have*

$$c_n = \max_{\sigma \in S_n} \sum_{i=1}^{n} \left\lceil \frac{pi - \sigma(i)}{d} \right\rceil - M_n \leq n; \tag{19}$$

$$M_n = \frac{n(n+1)(p-1)}{2d} + \frac{1}{d} \min_{\sigma \in S_n} \sum_{i=1}^{n} r_{i,\sigma(i)}; \tag{20}$$

$$S_n^s = \left\{ \sigma \in S_n \mid \sum_{i=1}^{n} \left\lceil \frac{pi - \sigma(i)}{d} \right\rceil = M_n + s \right\} . \tag{21}$$

Proof. Suppose $\sigma_1, \sigma_2 \in S_n$ are minimizer and maximizer of $\sum_{i=1}^n \left\lceil \frac{pi-\sigma(i)}{d} \right\rceil$, respectively. Note that $\left\lceil \frac{Ppi-j}{d} \right\rceil = \frac{pi-j+r_{ij}}{d}$ thus

$$\max_{\sigma\in S_n} \sum_{i=1}^n \left\lceil \frac{pi-\sigma(i)}{d} \right\rceil - M_n = \frac{1}{d}\left(\max_{\sigma\in S_n} \sum_{i=1}^n r_{i,\sigma(i)} - \min_{\sigma\in S_n} \sum_{i=1}^n r_{i,\sigma(i)} \right) = c_n.$$

For any i, since $1 \leq \sigma(i) \leq d-1$, we have

$$\left\lceil \frac{pi-\sigma_2(i)}{d} \right\rceil \leq \left\lceil \frac{pi-\sigma_1(i)}{d} \right\rceil + 1.$$

Taking sum both sides and get

$$\max_{\sigma\in S_n} \sum_{i=1}^n \left\lceil \frac{pi-\sigma(i)}{d} \right\rceil \leq M_n + n.$$

This proves (19). Since

$$\sum_{i=1}^n \left\lceil \frac{pi-\sigma(i)}{d} \right\rceil = \frac{n(n+1)(p-1)}{2d} + \frac{1}{d}\sum_{i=1}^n r_{i,\sigma(i)},$$

we see that (20) and (21) follows. □

For $0 \leq s, t \leq c_n$, and $i, j \geq 1$ let

$$K_{ij}^s(\vec{A}) := \sum_{\vec{m}\in\mathcal{M}_{ij}^s} \frac{\vec{A}^{\vec{m}}}{\vec{m}!\left(\left\lceil \frac{pi-j}{d} \right\rceil + s - |\vec{m}|\right)!}. \tag{22}$$

$$f_{n,p}^t(\vec{A}) := \sum_{\substack{s_0+\cdots+s_n=t \\ s_0,\ldots,s_n\geq 0}} \sum_{\sigma\in S_n^{s_0}} \mathrm{sgn}\,(\sigma) \prod_{i=1}^n K_{i,\sigma(i)}^{s_i}(\vec{A}). \tag{23}$$

For $p \geq d^2$, one notes that $H_{ij}^s(\vec{A}), K_{ij}^s(\vec{A}) \in \mathbb{Z}_p[\vec{A}]$, hence $f_n^t(\vec{A})$, $f_{n,p}^t(\vec{A}) \in \mathbb{Z}_p[\vec{A}]$. But $f_n^t(\vec{A})$ evaluates at $\vec{A} = \vec{a}$ while $f_{n,p}^t(\vec{A})$ at $\vec{A} = \vec{\hat{a}}$.

LEMMA 4.2. *Let $p \geq (d^2+1)(d-1)$. Then $f_n^t(\vec{A}) \equiv u_n f_{n,p}^t(\vec{A}) \bmod p$ for some p-adic unit u_n, where the reduction is taken at coefficients. Moreover, $f_n^t(\vec{a}) \equiv u_n f_{n,p}^t(\vec{\hat{a}}) \bmod p$.*

Proof. Since $p \geq d^2 - 1$ we always have $1 \leq \left\lceil \frac{pi-1}{d} \right\rceil + n \leq p-1$. Then

$$u_n := \prod_{i=1}^n \left(\left\lceil \frac{pi-1}{d} \right\rceil + n\right)!$$

is a p-adic unit in $\mathbb{Z}_p$.

Recall δij defined in (13). It is an elementary exercise to get

$$\begin{aligned}\left\lceil \frac{pi-1}{d} \right\rceil &= \frac{pi+r_{i1}-1}{d} \equiv \frac{r_{i1}-1}{d} \bmod p\\ \left\lceil \frac{pi-j}{d} \right\rceil &= \left\lceil \frac{pi-1}{d} \right\rceil - \delta_{ij} \equiv \frac{r_{i1}-1}{d} - \delta_{ij} \bmod p.\end{aligned}$$

Then we have

$$\begin{aligned} H_{ij}^{s}(\vec{A}) &\equiv \sum_{\vec{m}\in\mathcal{M}_{ij}^{s}} \frac{\left(\left\lceil \frac{pi-1}{d}\right\rceil + n\right)\left(\left\lceil \frac{pi-1}{d}\right\rceil + n - 1\right)\cdots\left(\left\lceil \frac{pi-j}{d}\right\rceil + s + 1 - |\vec{m}|\right)}{\vec{m}!}\vec{A}^{\vec{m}}\\ &\equiv \sum_{\vec{m}\in\mathcal{M}_{ij}^{s}} \frac{\left(\left\lceil \frac{pi-1}{d}\right\rceil + n\right)!}{\vec{m}!\left(\left\lceil \frac{pi-j}{d}\right\rceil + s - |\vec{m}|\right)!}\vec{A}^{\vec{m}}\\ &\equiv \left(\left\lceil \frac{pi-1}{d}\right\rceil + n\right)!\ K_{ij}^{s}(\vec{A}) \bmod p.\end{aligned}$$

Our first assertion follows easily. The second assertion follows from the fact that $\vec{a} \equiv \vec{\hat{a}} \bmod p$. □

PROPOSITION 4.3. *Let $p \geq (d^2+1)(d-1)$. For any $1 \leq i,j \leq n$ we have*

$$G_{pi-j}(\vec{A}) = \sum_{s=0}^{c_n} \gamma^{\left\lceil \frac{pi-j}{d}\right\rceil + s} K_{ij}^{s}(\vec{A}) + \gamma^{>\left\lceil \frac{pi-j}{d}\right\rceil + c_n}. \tag{24}$$

$$\det\{G_{pi-j}(\vec{A})\}_{1\leq i,j\leq n} = \sum_{t=0}^{c_n} \gamma^{M_n+t} f_{n,p}^{t}(\vec{A}) + \gamma^{>M_n+c_n}. \tag{25}$$

Proof. For $0 \leq s \leq c_n$, $m_\ell \geq 0$ and $m_1+\cdots+m_d = \left\lceil \frac{pi-j}{d}\right\rceil + s$, since $p \geq d^2-1$, we have $m_\ell \leq \left\lceil \frac{pi-j}{d}\right\rceil + c_n \leq p-1$. And by (6) and (7), we have

$$\begin{aligned} G_{pi-j}(\vec{A}) &= \sum_{\substack{m_1+\cdots+m_d \leq \left\lceil \frac{pi-j}{d}\right\rceil + c_n \\ \sum_{\ell=1}^{d} \ell m_\ell = pi-j}} \lambda_{m_1}\cdots\lambda_{m_d}\vec{A}^{\vec{m}} + \gamma^{>\left\lceil \frac{pi-j}{d}\right\rceil + c_n}\\ &= \sum_{s=0}^{c_n}\sum \frac{\gamma^{m_1+\cdots+m_d}\vec{A}^{\vec{m}}}{m_1!\cdots m_d!} + \gamma^{>\left\lceil \frac{pi-j}{d}\right\rceil + c_n},\end{aligned}$$

where the last sum ranges over all $m_\ell \geq 0$ such that $m_1+\cdots+m_d = \left\lceil \frac{pi-j}{d}\right\rceil + s$ and $\sum_{\ell=1}^{d} \ell m_\ell = pi-j$. It is easy to see that this is a subset of $\mathcal{M}_{ij}^{s}$. Conversely, if $\vec{m} \in \mathcal{M}_{ij}^s$ then

$$\sum_{\ell=1}^{d-1} m_\ell \leq \sum_{\ell=1}^{d-1} \ell m_{d-\ell} = r_{ij} + ds \leq d-1+ds \leq \left\lceil \frac{pi-j}{d} \right\rceil + s,$$

since $p \geq (d^2+1)(d-1)$. Set $m_d = \left\lceil \frac{pi-j}{d} \right\rceil + s - \sum_{\ell=1}^{d-1} m_\ell$, then $m_1 + \cdots + m_d = \left\lceil \frac{pi-j}{d} \right\rceil + s$ and $\sum_{\ell=1}^{d} \ell m_\ell = pi - j$ where $m_\ell \geq 0$. Thus we have

$$G_{pi-j}(\vec{A}) = \sum_{s=0}^{c_n} \gamma^{\left\lceil \frac{pi-j}{d} \right\rceil + s} \sum_{\vec{m} \in \mathcal{M}_{ij}^s} \frac{\vec{A}^{\vec{m}}}{\vec{m}! \left(\left\lceil \frac{pi-j}{d} \right\rceil + s - |\vec{m}| \right)!} + \gamma^{> \left\lceil \frac{pi-j}{d} \right\rceil + c_n}.$$

To prove (25) we have

$$\begin{aligned}
\det\{G_{pi-j}(\vec{A})\}_{1\leq i,j\leq n} &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} G_{pi-\sigma(i)}(\vec{A}) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \sum_{s_i=0}^{c_n} \left(\gamma^{\left\lceil \frac{pi-\sigma(i)}{d} \right\rceil + s_i} K_{i,\sigma(i)}^{s_i}(\vec{A}) + \gamma^{> \left\lceil \frac{pi-\sigma(i)}{d} \right\rceil + c_n} \right) \\
&= \sum_{s_0=0}^{c_n} \sum_{\sigma \in S_n^{s_0}} \operatorname{sgn}(\sigma) \sum_{\ell=0}^{c_n - s_0} \gamma^{M_n + s_0 + \ell} \\
&\quad \times \sum_{s_1+\cdots+s_n=\ell} \prod_{i=1}^{n} K_{i,\sigma(i)}^{s_i}(\vec{A}) + \gamma^{>M_n+c_n} \\
&= \sum_{t=0}^{c_n} \gamma^{M_n+t} \left(\sum_{s_0+\cdots+s_n=t} \sum_{\sigma \in S_n^{s_0}} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} K_{i,\sigma(i)}^{s_i}(\vec{A}) \right) + \gamma^{>M_n+c_n},
\end{aligned}$$

where the second equality follows from (24) and the third from Lemma 4.1. □

LEMMA 4.4. *Let $p \geq (d^2+1)(d-1)$. Then $\operatorname{ord}_p C_n(\hat{\vec{a}}) \geq \frac{M_n+t_n}{p-1}$ for all $\vec{a} \in (\mathbb{Z}_p \cap \mathbb{Q})^{d-1}$, and the equality holds if and only if $\bar{\vec{a}} \in \mathcal{V}_n(\mathbb{F}_p)$.*

Proof. First we show that

$$C_n(\vec{A}) = \sum_{t=0}^{c_n} \gamma^{M_n+t} f_{n,p}^t(\vec{A}) + \gamma^{>M_n+c_n}. \tag{26}$$

By (9) and (25) it suffices to show that if there is a t with $u_t > n$ then

$$\min_{\sigma \in S_n} \operatorname{ord}_p \prod_{t=1}^{n} G_{pu_t - u_{\sigma(t)}}(\vec{A}) > \frac{M_n + c_n}{p-1}. \tag{27}$$

Since $\sum_{t=1}^{n} u_t > \sum_{i=1}^{n} i$, we have

$$\begin{aligned}
(28)\qquad \frac{1}{p-1}\sum_{t=1}^{n}\left\lceil\frac{pu_t - u_{\sigma(t)}}{d}\right\rceil &\geq \frac{1}{p-1}\sum_{t=1}^{n}\frac{pu_t - u_{\sigma(t)}}{d}\\
&= \frac{1}{d}\sum_{t=1}^{n} u_t\\
&\geq \frac{1}{d}\sum_{i=1}^{n} i + \frac{1}{d}\\
&= \frac{1}{p-1}\sum_{i=1}^{n}\frac{pi-\delta(i)}{d} + \frac{1}{d}
\end{aligned}$$

for any $\delta \in S_n$. For $p \geq (d^2+1)(d-1) > d^2 - d + 1$ we have

$$\begin{aligned}
(29)\qquad \frac{1}{p-1}\sum_{i=1}^{n}\left\lceil\frac{pi-\delta(i)}{d}\right\rceil &\leq \frac{1}{p-1}\sum_{i=1}^{n}\frac{pi-\delta(i)}{d} + \frac{n}{p-1}\\
&\leq \frac{1}{p-1}\sum_{i=1}^{n}\frac{pi-\delta(i)}{d} + \frac{d-1}{p-1}\\
&< \frac{1}{p-1}\sum_{i=1}^{n}\frac{pi-\delta(i)}{d} + \frac{1}{d}
\end{aligned}$$

for any $\delta \in S_n$. Therefore,

$$\begin{aligned}
\min_{\sigma\in S_n} \mathrm{ord}_p \prod_{t=1}^{n} G_{pu_t - u_{\sigma(t)}}(\vec{A}) &\geq \frac{1}{p-1}\min_{\sigma\in S_n}\sum_{t=1}^{n}\left\lceil\frac{pu_t - u_{\sigma(t)}}{d}\right\rceil\\
&> \frac{1}{p-1}\max_{\delta\in S_n}\sum_{i=1}^{n}\left\lceil\frac{pi-\delta(i)}{d}\right\rceil = \frac{M_n + c_n}{p-1},
\end{aligned}$$

where the first inequality is due to (8), the second inequality by (28) and (29), and the last by (19).

Let $0 \leq t < t_n$. We have $f_n^t(\vec{A}) = 0$ and hence by Lemma 4.2 we have $f_{n,p}^t(\vec{a}) \equiv 0 \bmod p$. So

$$\mathrm{ord}_p\,(\gamma^{M_n+t} f_{n,p}^t(\vec{a})) \geq \frac{M_n+t}{p-1} + 1 > \frac{M_n + c_n}{p-1}.$$

Therefore, for all $\vec{a} \in (\mathbb{Z}_p \cap \mathbb{Q})^{d-1}$ by (26) we have

$$C_n(\vec{a}) = f_{n,p}^{t_n}(\vec{a})\gamma^{M_n+t_n} + \gamma^{>M_n+t_n}.$$

So

$$\operatorname{ord}_p C_n(\vec{\hat{a}}) \geq \frac{M_n + t_n}{p-1}$$

and the equality holds if and only if $f^{t_n}_{n,p}(\vec{\hat{a}}) \equiv f^{t_n}_n(\vec{a}) \not\equiv 0 \bmod p$. This proves the lemma. □

5. Generic Newton polygons. Let the *generic Newton polygon* of $\mathbb{A}^d$ over $\mathbb{F}_p$ be the lowest Newton polygon over all $\bar{f} \in \mathbb{A}^d(\mathbb{F}_p)$, that is,

$$\operatorname{GNP}(\mathbb{A}^d; \mathbb{F}_p) := \inf_{\bar{f} \in \mathbb{A}^d(\mathbb{F}_p)} \operatorname{NP}(\bar{f}).$$

Note that it is equal to $\inf_{f \in \mathbb{A}^d(\mathbb{Z}_p \cap \mathbb{Q})} \operatorname{NP}(f \otimes \mathbb{F}_p)$ and one does not know *a priori* whether this infimum exists. Note that Wan has shown that the generic Newton polygon over $\overline{\mathbb{F}}_p$ defined by $\operatorname{GNP}(\mathbb{A}^d; \overline{\mathbb{F}}_p) := \inf_{\bar{f} \in \mathbb{A}^d(\overline{\mathbb{F}}_p)} \operatorname{NP}(\bar{f})$ exists by the Grothendieck specialization theorem (see [19, Section 1.1]). In the theorem below we show that $\operatorname{GNP}(\mathbb{A}^d; \mathbb{F}_p)$ exists for p large enough. One may ask if it is true that $\operatorname{GNP}(\mathbb{A}^d; \mathbb{F}_p) = \operatorname{GNP}(\mathbb{A}^d; \overline{\mathbb{F}}_p)$ for p large enough.

We shall proceed to prove Theorem 5.1 below by first introducing some notations. Let $\epsilon_0 = 0$ and for $1 \leq n \leq d-1$ let

$$\epsilon_n := \frac{\min_{\sigma \in S_n} \sum_{i=1}^{n} r_{i,\sigma(i)} + d t_n}{d(p-1)}, \tag{30}$$

where r_{ij} and t_n are defined in Section 3.1 and Lemma 3.5, respectively. One observes easily

$$\frac{M_n + t_n}{p-1} = \frac{n(n+1)}{2d} + \epsilon_n. \tag{31}$$

Note that $0 \leq r_{ij} \leq d-1$ for all $1 \leq i,j \leq d-1$, and $t_n \leq c_n \leq n \leq d-1$ by (19), so we have $\epsilon_n \leq \frac{n(2d-1)}{d(p-1)}$. Thus ϵ_n goes to 0 as p approaches ∞.

For every integer r with $1 \leq r \leq d-1$ and $\gcd(r,d) = 1$, let $\mathcal{W}_r := \bigcap_{n=1}^{d-1} \mathcal{V}_n$ (recall from Key-Lemma 3.5 that $\mathcal{V}_n$ consists of all $f \in \mathbb{A}^{d-1}$ whose coefficients $\vec{a}$ satisfy $f^{t_n}_n(\vec{a}) \neq 0$.) Let $\mathcal{W} := \bigcap_{\substack{1 \leq r \leq d-1 \\ \gcd(r,d)=1}} \mathcal{W}_r$. Consider the natural projection map $\iota\colon \mathbb{A}^d \to \mathbb{A}^{d-1}$ by $\iota(f) = \vec{a} = (a_1, \ldots, a_{d-1})$ for every $f = x^d + a_{d-1}x^{d-1} + \cdots + a_0 \in \mathbb{A}^d$. Let $\mathcal{U} := \iota^{-1}(\mathcal{W})$. For every residue class r denote by $f^{t_n}_{n,r}$ the $f^{t_n}_n$ in Lemma 3.5, then $\mathcal{U}$ consists of all $f \in \mathbb{A}^d$ whose coefficients satisfy $\prod_r \prod_{n=1}^{d-1} f^{t_n}_{n,r}(\vec{a}) \neq 0$ where r ranges over all $1 \leq r \leq d-1$ coprime to d. Since $\prod_r \prod_{n=1}^{d-1} f^{t_n}_{n,r}$ is a nonzero polynomial over $\mathbb{Q}$ by Lemma 3.5, one concludes that $\mathcal{U}$ is Zariski dense open in $\mathbb{A}^d$ over $\mathbb{Q}$. One notes that, even though $\mathcal{U}(\mathbb{F}_p)$ is not necessarily nonempty, it is nonempty when p is large enough.

THEOREM 5.1. *Let notations be as above.*

(a) *For p large enough (depending only on d)* $\mathrm{GNP}(\mathbb{A}^d;\mathbb{F}_p)$ *exists and is equal to the lower convex hull of points* $(n, \frac{n(n+1)}{2d}+\epsilon_n)$ *for* $0 \le n \le d-1$*, each of which is a vertex.*

(b) *Fix* $f \in \mathbb{A}^d(\mathbb{Q})$*. For p large enough (depending only on d and f) we have*

$$\mathrm{NP}(f \otimes \mathbb{F}_p) \ge \mathrm{GNP}(\mathbb{A}^d;\mathbb{F}_p)$$

where the equality holds for all p large enough if and only if $f \in \mathcal{U}(\mathbb{Q})$*. Here $\ge$ means "lies above".*

(c) *For* $f \in \mathcal{U}(\mathbb{Q})$ *we have*

$$\lim_{p\to\infty} \mathrm{NP}(f \otimes \mathbb{F}_p) = \mathrm{HP}(\mathbb{A}^d).$$

Proof. (a) Because of (4), we consider $f(x) \in \mathbb{A}^d(\mathbb{Z}_p \cap \mathbb{Q})$ with no constant term, that is, $f(x) = x^d + \sum_{i=1}^{d-1} a_i x^i$. By Lemma 4.4, for p large enough we have

$$\mathrm{ord}_p\, C_n(\vec{a}) \ge \frac{n(n+1)}{2d} + \epsilon_n$$

for all $0 \le n \le d-1$ and the equality holds if and only if $\vec{\bar{a}} \in \mathcal{W}(\mathbb{F}_p)$. On the other hand, by the remarks preceding the theorem, ϵ_n approaches 0. Thus for p large enough the lower convex hull of points $(n, \frac{n(n+1)}{2d}+\epsilon_n)$ with $0 \le n \le d-1$ passes all these points as vertices. By Proposition 2.2, for p large enough,

$$\mathrm{ord}_p\, b_n(\vec{a}) \ge \frac{n(n+1)}{2d} + \epsilon_n \tag{32}$$

and the equality holds if and only if $\vec{\bar{a}} \in \mathcal{W}(\mathbb{F}_p)$. Now (a) clearly follows.

(b) Now let $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0 \in \mathbb{A}^d(\mathbb{Q})$. Note that (32) says for p large enough,

$$\mathrm{NP}(f \otimes \mathbb{F}_p) = \mathrm{NP}((x^d + a_{d-1}x^{d-1} + \cdots + a_1 x) \otimes \mathbb{F}_p) \ge \mathrm{GNP}(\mathbb{A}^d;\mathbb{F}_p)$$

where the equality holds if and only if $(\bar{a}_1, \ldots, \bar{a}_{d-1}) \in \mathcal{W}(\mathbb{F}_p)$. Note that a rational number N is nonzero if and only if N is not divisible by all primes large enough. Thus for p large enough the above equality holds if and only if $(a_1, \ldots, a_{d-1}) \in \mathcal{W}(\mathbb{Q})$, that is, $f \in \mathcal{U}(\mathbb{Q})$. This proves (b). Note that (c) follows from (a) and (b). □

Remark 5.2. (1) Let $d \ge 3$. Let *generic polynomial* $F_d := \prod_r \prod_{n=1}^{d-1} f_{n,r}^{t_n}$ where r ranges over $1 \le r \le d-1$ coprime to d. From the theorem above, the set of polynomials $f(x) = x^d + \cdots + a_1 x + a_0 \in \mathbb{A}^d(\mathbb{Q})$ with $\mathrm{NP}(f \otimes \mathbb{F}_p) = \mathrm{GNP}(\mathbb{A}^d)$ corresponds precisely to the set of $(a_0, \ldots, a_{d-1}) \in \mathbb{Q}^d$ with $F_d|_{\vec{A}=(a_1,\ldots,a_{d-1})} \ne 0$.

(2) In practice, for any $d \geq 3$ one may compute the polynomial $P_d := \prod_r \prod_{n=1}^{\lceil \frac{d-1}{2} \rceil} f_n^{t_n}$ in $\mathbb{Q}[\vec{A}]$ where r ranges over all $2 \leq r \leq d-1$ with $\gcd(r,d) = 1$. (Remark: One notes that the $r = 1$ case is explained in remarks before Theorem 1.1. One also notes that $\mathrm{NP}(f \otimes \mathbb{F}_p)$ is symmetric in the sense that every slope α segment comes with a slope $1-\alpha$ segment with the same length.) Then every $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0 \in \mathbb{A}^d(\mathbb{Q})$ with $P_d|_{\vec{A}=(a_1,\ldots,a_{d-1})} \neq 0$ satisfies $\lim_{p\to\infty} \mathrm{NP}(f \otimes \mathbb{F}_p) = \mathrm{HP}(\mathbb{A}^d)$.

6. Generic Newton polygon for $x^d + ax$. Recall $d \geq 3$. In this section we consider the Newton polygon of the L function of exponential sums of $f(x) = x^d + ax$ over $\mathbb{Q}$. This family has drawn some attention recently (see [21] for some progress). When $a = 0$ see Remark 1.4(b). Let $\mathbb{A}^d(1)$ denote the space of all such $f(x)$ with parameter a. Let $\mathrm{GNP}(\mathbb{A}^d(1); \mathbb{F}_p)$ be the corresponding analog of $\mathrm{GNP}(\mathbb{A}^d; \mathbb{F}_p)$.

Let r be $1 \leq r \leq d-1$ coprime to d. Recall that r'_{ij} is the least nonnegative residue of $ri-j \bmod d$. That is, $r'_{ij} = ri-j-d\left\lfloor \frac{ri-j}{d} \right\rfloor$. Let $M'_n := \min_{\sigma\in S_n} \sum_{i=1}^n r'_{i,\sigma(i)}$. Let S'_n be the subset of $\sigma \in S_n$ with $\sum_{i=1}^n r'_{i,\sigma(i)} = M'_n$. Let $\epsilon'_0 := 0$; for $n \geq 1$ and for $p \equiv r \bmod d$ let

$$\epsilon'_n := \frac{(d-1)M'_n}{d(p-1)}.$$

LEMMA 6.1. *Let $1 \leq n \leq d-1$ and $p \equiv r \bmod d$. The following statements are equivalent:*

(1) $\sigma \in S'_n$;

(2) $\sigma(i) \leq r'_{i1} + 1$ *for all* $1 \leq i \leq n$;

(3) $r'_{i,\sigma(i)} = r'_{i1} - \sigma(i) + 1$ *for all* $1 \leq i \leq n$;

(4) $\left\lfloor \frac{pi-1}{d} \right\rfloor = \left\lfloor \frac{pi-\sigma(i)}{d} \right\rfloor$ *for all* $1 \leq i \leq n$.

Proof. Define $\delta'_{ij} := 0$ if $j \leq r'_{i1}+1$ and $\delta'_{ij} := 1$ if $j > r'_{i1}+1$. From Lemma 3.1 one notes that $r'_{11}+1, \ldots, r'_{n1}+1$ are n distinct integers in the interval $[1, d-1]$. So there exists $\sigma \in S_n$ such that $\sigma(i) \leq r'_{i1} + 1$ for every $1 \leq i \leq n$, that is, $\delta'_{i,\sigma(i)} = 0$ for every $1 \leq i \leq n$. Thus $\min_{\sigma\in S_n} \sum_{i=1}^n \delta'_{i,\sigma(i)} = 0$ and it is achieved if and only if (2) holds.

By recalling Lemma 3.1, it is straightforward to see that

$$r'_{ij} = r'_{i1} - j + 1 + \delta'_{ij}(d-1).$$

Thus for any $\sigma \in S_n$,

$$\sum_{i=1}^n r'_{i,\sigma(i)} = \sum_{i=1}^n (r'_{i1}-\sigma(i)+1)+(d-1)\sum_{i=1}^n \delta'_{i,\sigma(i)} = \sum_{i=1}^n r'_{i1}-\frac{n(n-1)}{2}+(d-1)\sum_{i=1}^n \delta'_{i,\sigma(i)}.$$

One notes that (1) holds if and only if $\sum_{i=1}^{n} r'_{i,\sigma(i)}$ achieves its minimum, and if and only if $\sum_{i=1}^{n} \delta'_{i,\sigma(i)} = 0$ by the previous paragraph. Thus (1), (2) and (3) are equivalent to each other. Since $r'_{ij} = pi-j-d\lfloor\frac{pi-j}{d}\rfloor$, it is easy to see (3) and (4) are equivalent to each other. This proves the lemma. □

By the lemma above, $M'_n = \sum_{i=1}^{n}(r'_{i1} - \sigma(i) + 1)$. So one gets an explicit formula

$$\epsilon'_n = \frac{(d-1)(\sum_{i=1}^{n} r'_{i1} - \frac{n(n-1)}{2})}{d(p-1)}.$$

Note that ϵ'_n converges to 0 as p approaches ∞.

THEOREM 6.2. (a) *For p large enough (depending only on d)* GNP $(\mathbb{A}^d(1); \mathbb{F}_p)$ *exists and is equal to the lower convex hull of points* $(n, \frac{n(n+1)}{2d} + \epsilon'_n)$ *for* $0 \le n \le d-1$, *each of which is a vertex.*

(b) *Fix* $f = x^d + ax \in \mathbb{A}^d(\mathbb{Q})$. *For p large enough (depending only on d and a) we have*

$$\mathrm{NP}\,(f \otimes \mathbb{F}_p) \ge \mathrm{GNP}\,(\mathbb{A}^d(1); \mathbb{F}_p),$$

where the equality holds for all p large enough if and only if $a \ne 0$. Here $\ge$ means "lies above."

(c) *For any* $a \ne 0$ *we have*

$$\lim_{p\to\infty} \mathrm{NP}\,((x^d + ax) \otimes \mathbb{F}_p) = \mathrm{HP}\,(\mathbb{A}^d).$$

LEMMA 6.3. *Let $p \equiv r$* mod *d. Let $a \in \mathbb{Q} \cap \mathbb{Z}_p$ and let $\hat{a}$ be the Teichmüller lifting of a* mod *p. Let $p \ge d$. For any* $1 \le i, j \le d-1$ *we have*

$$G_{pi-j} = \gamma^{r'_{ij}+\lfloor\frac{pi-j}{d}\rfloor}\hat{a}^{r'_{ij}}\frac{1}{r'_{ij}!\,\left\lfloor\frac{pi-j}{d}\right\rfloor!} + \gamma^{>r'_{ij}+\lfloor\frac{pi-j}{d}\rfloor}.$$

Proof. Note that $G_{pi-j} = \sum \lambda_{m_1}\lambda_{m_d}\hat{a}^{m_1}$ where the sum ranges in $m_1 + dm_d = pi-j$ with $m_1, m_d \ge 0$. But in this range of m_1 and m_d, one notices that the minimum of $m_1 + m_d$ is achieved precisely at $m_1 = r'_{ij}$ and $m_d = \lfloor\frac{pi-j}{d}\rfloor$, that is, $\min(m_1 + m_d) = r'_{ij} + \lfloor\frac{pi-j}{d}\rfloor$. The rest of the proof is analogous to Proposition 4.3. □

LEMMA 6.4. *Let $p \equiv r$* mod *d and $p \ge (d-1)^3 + 2$. Let* $1 \le n \le d-1$. *Then we have*

$$C_n = \gamma^{(p-1)(\frac{n(n+1)}{2d}+\epsilon'_n)}\hat{a}^{M'_n}f'_{n,p} + \gamma^{>(p-1)(\frac{n(n+1)}{2d}+\epsilon'_n)},$$

where

$$f'_{n,p} := \sum_{\sigma\in S'_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \frac{1}{r'_{i,\sigma(i)}!\left\lfloor \frac{pi-\sigma(i)}{d}\right\rfloor !}.$$

Proof. The proof is analogous to Lemma 4.4, so we will only give an outline. First one shows that for $1 \leq n \leq d-1$ one has

$$C_n = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} G_{pi-\sigma(i)} + \gamma^{\geq (p-1)(\frac{n(n+1)}{2d}+\frac{1}{d})}. \tag{33}$$

Since $r'_{ij} = pi - j - d\left\lfloor \frac{pi-j}{d}\right\rfloor$, we have

$$r'_{ij} + \left\lfloor \frac{pi-j}{d}\right\rfloor = \frac{pi-j}{d} + \frac{d-1}{d} r'_{ij}.$$

Thus

$$\begin{aligned}\min_{\sigma\in S_n} \sum_{i=1}^{n} \left(r'_{i,\sigma(i)} + \left\lfloor \frac{pi-\sigma(i)}{d}\right\rfloor\right) &= \frac{(p-1)n(n+1)}{2d} + \frac{(d-1)M'_n}{d}\\ &= (p-1)\left(\frac{n(n+1)}{2d} + \epsilon'_n\right).\end{aligned}$$

Consequently the minimum is achieved precisely at all $\sigma \in S'_n$. Note that $p \geq (d-1)^3 + 2$ implies that $(p-1)(\frac{n(n+1)}{2d} + \frac{1}{d}) > (p-1)(\frac{n(n+1)}{2d} + \epsilon'_n)$. By (33) and Lemma 6.3 we have

$$\begin{aligned}C_n &= \gamma^{(p-1)(\frac{n(n+1)}{2d}+\epsilon'_n)} \hat{a}^{M'_n} \sum_{\sigma\in S'_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \frac{1}{r'_{i,\sigma(i)}!\left\lfloor \frac{pi-\sigma(i)}{d}\right\rfloor !}\\ &\quad + \gamma^{>(p-1)(\frac{n(n+1)}{2d}+\epsilon'_n)}.\end{aligned}$$

The lemma follows. □

LEMMA 6.5. *Let notation and hypothesis be as in Lemma 6.4. Then*

$$\operatorname{ord}_p C_n \geq \frac{n(n+1)}{2d} + \epsilon'_n$$

and the equality holds if and only if $a \not\equiv 0 \bmod p$.

Proof. Let

$$u_n := \prod_{i=1}^{n} r'_{i1}!\left(\left\lfloor \frac{pi-1}{d}\right\rfloor !\right).$$

By Lemma 6.1, one sees that

$$u_n f'_{n,p} = \sum_{\sigma \in S'_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \frac{r'_{i1}!}{r'_{i,\sigma(i)}!}.$$

By Lemma 6.1, we have

$$\begin{aligned} u_n f'_{n,p} &= \sum_{\sigma \in S'_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (r'_{i1}(r'_{i1}-1)\cdots(r'_{i1}-(\sigma(i)-2)) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (r'_{i1}(r'_{i1}-1)\cdots(r'_{i1}-(\sigma(i)-2)), \end{aligned}$$

where we set $(r'_{i1}(r'_{i1}-1)\cdots(r'_{i1}-(\sigma(i)-2)) := 1$ if $\sigma(i)=1$. One observes that this is equal to the determinant of a matrix M shown as below

$$M = \begin{bmatrix} 1 & r'_{11} & r'_{11}(r'_{11}-1) & \cdots \\ 1 & r'_{21} & r'_{21}(r'_{21}-1) & \cdots \\ & & \vdots & \\ 1 & r'_{n1} & r'_{n1}(r'_{n1}-1) & \cdots \end{bmatrix}.$$

Under natural column transformation M becomes a Vandermonde matrix, we get

$$u_n f'_{n,p} = \det M = \det \begin{bmatrix} 1 & r'_{11} & (r'_{11})^2 & \cdots \\ 1 & r'_{21} & (r'_{21})^2 & \cdots \\ & & \vdots & \\ 1 & r'_{n1} & (r'_{n1})^2 & \cdots \end{bmatrix} = \prod_{1 \le i < k \le n} (r'_{k1} - r'_{i1}).$$

As in Lemma 3.1, one notes that $r'_{i1} \neq r'_{k1}$ for any $i < k$. One also notes that u_n is a p-adic unit. Therefore, $f'_{n,p} \not\equiv 0 \bmod p$ for all p. □

Proof of Theorem 6.2. Theorem 6.2 follows from Lemma 6.5, using the same arguments as in the proof of Theorem 5.1. □

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCMASTER UNIVERSITY, HAMILTON,
ON L8S 4K1, CANADA
E-mail: zhu@cal.berkeley.edu

REFERENCES

[1] A. Adolphson and S. Sperber, Newton polyhedra and the degree of the L-function associated to an exponential sum, *Invent. Math.* **88** (1987), 555–569.

[2] ———, Exponential sums and Newton polyhedra: Cohomology and estimates, *Ann. of Math.* **130** (1989), 367–406.

[3] E. Bombieri, On exponential sums in finite fields, *Amer. J. Math.* **88** (1966), 71–105.

[4] B. Dwork, On the zeta function of a hypersurface, *Inst. Hautes Études Sci. Publ. Math.* **12** (1962), 5–68.

[5] ———, On the zeta function of a hypersurface. II. *Ann. of Math.* **80** (1964), 227–299.

[6] S. Hong, Newton polygons of L functions associated with exponential sums of polynomials of degree four over finite fields, *Finite Fields Appl.* **7** (2001), 205–237.

[7] N. M. Katz, Sommes exponentielles, *Astérisque* **79** (1980).

[8] N. Koblitz, *p-adic Analysis: A Short Course on Recent Work*, *London Math. Soc. Lecture Note Ser.*, vol. 46, Cambridge University Press, 1980.

[9] ———, *p-adic Numbers, p-adic Analysis, and Zeta-functions*, 2nd edition, *Grad. Texts in Math.*, vol. 58, Springer-Verlag, New York, 1984.

[10] S. Lang, *Algebra*, 3rd edition, Addison-Wesley, Reading, MA, 1993.

[11] K.-Z. Li and F. Oort, *Moduli of Supersingular Abelian Varieties*, *Lecture Notes in Math.*, vol. 1680, Springer-Verlag, Berlin, 1998.

[12] J. Scholten and H. J. Zhu, The first slope case of Wan's conjecture, *Finite Fields Appl.* **8** (2002), 414–419.

[13] ———, Slope estimates of Artin-Schreier curves, *Compositio Math.* (to appear), math.AG/0105005.

[14] S. Sperber, Congruence properties of the hyperkloosterman sum, *Compositio Math.* **40** (1980), 3–33.

[15] ———, Newton polygons for general hyperkloosterman sums, *p-adic Cohomology*, *Astérisque* **119–120** (1984), 267–330.

[16] ———, On the p-adic theory of exponential sums, *Amer. J. Math.* **109** (1986), 255–296.

[17] D. Wan, New polygons of zeta functions and L functions, *Ann. of Math.* **137** (1993), 249–293.

[18] ———, An introduction to the theory of Newton polygons for L-functions of exponential sums, preprint available at http://www.math.uci.edu/dwan/Overview.html.

[19] ———, Variation of p-adic Newton polygons of L functions for exponential sums, preprint available at http://www.math.uci.edu/dwan/Overview.html.

[20] L. Washington, *Introduction to Cyclotomic Fields*, 2nd edition, *Grad. Texts in Math.*, vol. 83, Springer-Verlag, Berlin, 1997.

[21] R. Yang, Newton polygons of L-functions of polynomials of the form $x^d + \lambda x$ (to appear).

[22] H. J. Zhu, p-adic variation of L functions of one variable exponential sums, II, math.AG/0206284.