ASYMPTOTIC VARIATION OF L FUNCTIONS OF ONE-VARIABLE EXPONENTIAL SUMS

HUI JUNE ZHU

Abstract. Fix an integer $d \geq 3$. Let $\mathbb{A}^d$ be the dimension-d affine space over the algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$, identified with the coefficient space of degree-d monic polynomials $f(x)$ in one variable x. For any $f(x)$ in $\mathbb{A}^d(\overline{\mathbb{Q}})$, let $\mathbb{Q}(f)$ be the field generated by coefficients of f in $\overline{\mathbb{Q}}$. For each prime p coprime to d, pick an embedding from $\overline{\mathbb{Q}}$ to $\overline{\mathbb{Q}}_p$, once and for all. Let $\mathcal{P}$ be the place in $\mathbb{Q}(f)$ lying over p specified by the embedding of $\overline{\mathbb{Q}}$ in $\overline{\mathbb{Q}}_p$ (with residue field $\mathbb{F}_q$ say). Suppose $f \in \mathbb{A}^d(\overline{\mathbb{Q}} \cap \overline{\mathbb{Z}}_p)$, let $\mathrm{NP}(f(x) \bmod \mathcal{P})$ denote the q-adic Newton polygon of the L function $L(f(x) \bmod \mathcal{P}; T)$ of exponential sums of $f \bmod \mathcal{P}$. We prove that there is a Zariski dense open subset $\mathcal{U}$ defined over $\mathbb{Q}$ in $\mathbb{A}^d$ such that for every geometric point $f(x)$ in $\mathcal{U}(\overline{\mathbb{Q}})$ and p large enough (depending only on f) one has $\mathrm{NP}(f \bmod \mathcal{P}) = \mathrm{GNP}(\mathbb{A}^d; \mathbb{F}_p)$ and

$$\lim_{p\to\infty} \mathrm{NP}(f(x) \bmod \mathcal{P}) = \mathrm{HP}(\mathbb{A}^d),$$

where $\mathrm{GNP}(\mathbb{A}^d; \mathbb{F}_p)$ and $\mathrm{HP}(\mathbb{A}^d)$ are the generic Newton polygon and the Hodge polygon, respectively (see [23]).

1. Introduction

In this paper we fix an integer $d \geq 3$. Let $\overline{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$. Let $\mathbb{A}^d$ be the affine variety of dimension d over $\overline{\mathbb{Q}}$, identified with the coefficient space of degree-d monic polynomials $f(x)$ in one variable x. For any $f \in \mathbb{A}^d(\overline{\mathbb{Q}})$ let $\mathbb{Q}(f)$ denote the field generated by coefficients of f over $\mathbb{Q}$. Let p be any prime coprime to d. Let $\overline{\mathbb{Q}}_p$ be the algebraic closure of $\mathbb{Q}_p$ and let $\overline{\mathbb{Z}}_p$ be its ring of integers. For each p we pick an embedding from $\overline{\mathbb{Q}}$ into $\overline{\mathbb{Q}}_p$. Let $\mathcal{P}$ be the place in $\mathbb{Q}(f)$ lying over p specified by the embedding. Henceforth we tacitly understand that such embeddings are already picked once and for all. Suppose the residue field at $\mathcal{P}$ is $\mathbb{F}_q$ for $q = p^a$ for some a. Let $\mathrm{ord}_p(\cdot)$ denote the p-adic valuation in an extension of $\mathbb{Q}_p$ with $\mathrm{ord}_p(p) = 1$; let $\mathrm{ord}_q(\cdot)$ be the q-adic valuation, i.e., $\mathrm{ord}_q(\cdot) := \frac{1}{a}\mathrm{ord}_p(\cdot)$. Let $E(x) = \exp(\sum_{j=0}^{\infty} \frac{x^{p^j}}{p^j})$ be the Artin-Hasse p-adic 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. We fix this p-th root of unity for the entire paper and denote it by ζ_p. Note that $\mathbb{Z}_p[\gamma] = \mathbb{Z}_p[\zeta_p]$.

Date: Sept. 13, 2003.

1991 *Mathematics Subject Classification.* 11,14.

Key words and phrases. Newton polygons; Hodge polygons; exponential sums; L functions; Zeta functions; p-adic variation.

Let $f \in (\overline{\mathbb{Z}}_p \cap \overline{\mathbb{Q}})[x]$ be a degree-d monic polynomial. For every positive integer ℓ let

$$S_\ell(f \bmod \mathcal{P}) := \sum_{x\in\mathbb{F}_{q^\ell}} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_{q^\ell}/\mathbb{F}_p}(f(x) \bmod \mathcal{P})}. \tag{1}$$

Then the L function of the exponential sum of f over $\mathbb{F}_q$ is defined by

$$L(f \bmod \mathcal{P}; T) := \exp(\sum_{\ell=1}^{\infty} S_\ell(f \bmod \mathcal{P})\frac{T^\ell}{\ell}). \tag{2}$$

It is well known that $L(f \bmod \mathcal{P}; T)$ is a polynomial in $1+T\mathbb{Z}[\zeta_p][T]$ of degree $d-1$ (e.g., see remarks in the Introduction of [23]). So we may write,

$$L(f \bmod \mathcal{P}; T) \;=\; 1 + b_1(f)T + b_2(f)T^2 + \ldots + b_{d-1}(f)T^{d-1} \tag{3}$$

for some $b_n(f) \in \mathbb{Z}[\zeta_p][T]$. It is easy to see that $L(f \bmod \mathcal{P}; T)$ becomes independent of the choice of embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$ if p is large enough.

For any polynomial $\sum_{n=0}^{k} c_n T^n$ over $\overline{\mathbb{Q}}$ let $\mathrm{NP}_q(\sum_{n=0}^{k} c_n T^n)$ denote its q-adic Newton polygon, i.e., the lower convex hull in $\mathbb{R}^2$ of the points $(n, \mathrm{ord}_q(c_n))$ with $0 \le n \le k$. Now let $\mathrm{NP}(f \bmod \mathcal{P}) := \mathrm{NP}_q(L(f \bmod \mathcal{P}; T))$. We shall note below that $\mathrm{NP}(f \bmod \mathcal{P})$ is independent of the choice of embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. By the Dieudonné-Manin classification (see [13]), the Newton polygon of an abelian variety over a finite field is determined by certain 'formal isogeny types'. Hence the Newton polygon of a smooth projective curve over a finite field $\mathbb{F}_q$, same as that of its Jacobian variety, is independent of the choice of $\mathbb{F}_q$. Let $\mathrm{NP}(X_f \bmod \mathcal{P})$ be the Newton polygon of the Artin-Schreier curve X_f given by affine equation $y^p - y = f \bmod \mathcal{P}$. One knows that $\mathrm{NP}(f \bmod \mathcal{P}) = \mathrm{NP}(X_f \bmod \mathcal{P})/(p-1)$ where the latter Newton polygon is shrunk by a factor of $p-1$ horizontally and vertically (see [23, Introduction]), all these above imply that $\mathrm{NP}(f \bmod \mathcal{P})$ is independent of the choice of embedding of $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$.

The *Hodge polygon* $\mathrm{HP}(\mathbb{A}^d)$ of $\mathbb{A}^d$ is the lower convex hull in $\mathbb{R}^2$ of the points $(n, \frac{n(n+1)}{2d})$ with $0 \le n \le d-1$. It is known that $\mathrm{HP}(\mathbb{A}^d)$ is a lower bound of $\mathrm{NP}(f \bmod \mathcal{P})$ (see [20, Propositions 2.2 and 2.3]) and that for every $f \in \mathbb{A}^d(\overline{\mathbb{Q}} \cap \overline{\mathbb{Z}}_p)$ one has $\mathrm{NP}(f \bmod \mathcal{P}) = \mathrm{HP}(\mathbb{A}^d)$ if and only if $p \equiv 1 \bmod d$ (see [2, (3.11)]). Our Hodge polygon, which inherits that from [20, 21], is defined combinatorially (so we shall refer to it as Wan's Hodge polygon in the remark below). We shall compare it with classical Hodge polygons in the literature.

Remark 1.1. (i) Wan's Hodge polygon does not generally acquire a geometric meaning, and we do not know of one for the one-variable exponential sum case studied in this paper. Nevertheless there is a well-known case in which it does. If f is an n-variable Laurent polynomial over a finite field, Wan's Hodge polygon of the exponential sum of f is defined in [21, Section 1]. It is known that if $f = 0$ defines a toric variety (denoted by X) then Wan's Hodge polygon of the exponential sum of f does coincide with a variant of classical Hodge polygon which is defined by 'Hodge numbers' of certain subgroup of the cohomology $H_c^{n-1}(X, \mathbb{C})$ with compact support over the complex $\mathbb{C}$ (see [1, Section 5] for details and proofs). This explains Wan's terminology of 'Hodge polygon'.

(ii) Let X be a smooth projective scheme of finite type over a finite field $\mathbb{F}_q$ such that the Hodge cohomology groups $H^j(\hat{X}, \Omega^i_{\hat{X}/W(\mathbb{F}_q)})$ are free $W(\mathbb{F}_q)$-modules of finite ranks $h^{i,j}(\hat{X})$, where $W(\mathbb{F}_q)$ is the ring of Witt vectors over $\mathbb{F}_q$ and $\hat{X}/W(\mathbb{F}_q)$ is a lift of $X/\mathbb{F}_q$. Recall that the (m-dimensional) classical Hodge polygon of X (à la Katz-Mazur) consists of line segments of slope i of horizontal length $h^i := h^{i,m-i}(\hat{X})$ for all $i = 0, \ldots, m$. For example, the (1-dimensional) classical Hodge polygon of a curve X over a finite field of genus g consists of a slope-0 segment of length $h^0 = g$ and a slope-1 one of length $h^1 = g$. It is known that this classical Hodge polygon of X is a lower bound of its Newton polygon (see [8, 14]). However, for Artin-Schreier curves this lower bound is not sharp and there is a sharp lower bound, that is precisely Wan's Hodge polygon (after blown up by a factor of $p-1$ horizontally and vertically). In summary, Wan's Hodge polygon is an analog of Katz-Mazur's classical Hodge polygon as a lower bound to Newton polygons.

(iii) From a geometrical point of view, it has long been a myth why the Newton polygons of Artin-Schreier curves X_f are much higher than the classical Hodge polygon (see [19] for example). It is discovered recently that for the exponential sum of a one-variable *rational* function f there is a generalized Wan's Hodge polygon that determined by the orders of poles of f. (This was conjectured by Poonen and Adolphson-Sperber independently, proved by [24, Theorem 1.1].) Our Theorem 1.3 below is asserting that if f is a polynomial then Wan's Hodge polygon is asymptotically (for p large) a best lower bound! We anticipate that a natural generalization of this sort should hold true also for rational functions.

Remark 1.2. The choice of the primitive p-th root of unity ζ_p does not affect the Newton polygon. In fact, if ζ_p is replaced by ζ_p^i in (1), then every coefficient b_n in the L function in (3) will be changed by replacing every ζ_p in its expression by ζ_p^i. The p-adic valuation of b_n is invariant under Galois conjugation.

In this paper we prove Theorems 1.3 and 3.3. Part of Theorem 1.3 was formulated as a conjecture by Daqing Wan communicated to me in 2001 and it is now a one-dimensional case of a new conjecture collected in [21, Section 1.4]. The case $d = 3$ of 1.3 is proved by [18, (3.14)] using Dwork's method. Theorem 1.3 also yields a complete answer to a question (in one-variable case) proposed by Katz on page 151 [10, Chapter 5.1]. The first slope case is proved recently by [16] by a slope estimate technique essentially following Katz [9]. A weaker version of this theorem, which restricts to $f \in \mathcal{U}(\mathbb{Q})$, is proved in [23] recently. Recall from [23, Section 5] that $\mathrm{GNP}(\mathbb{A}^d; \mathbb{F}_p) := \inf_{\overline{f} \in \mathbb{A}^d(\mathbb{F}_p)} \mathrm{NP}(\overline{f})$ if exists.

Theorem 1.3. *There is a Zariski dense open subset $\mathcal{U}$ defined over $\mathbb{Q}$ in $\mathbb{A}^d$ such that if $f \in \mathcal{U}(\overline{\mathbb{Q}})$ and if $\mathcal{P}$ is a prime ideal in the ring of integers of $\mathbb{Q}(f)$ lying over p, we have for p large enough (depending only on f),* $\mathrm{NP}(f \bmod \mathcal{P}) = \mathrm{GNP}(\mathbb{A}^d; \mathbb{F}_p)$. *In particular, for every $f \in \mathcal{U}(\overline{\mathbb{Q}})$ one has* $\lim_{p \to \infty} \mathrm{NP}(f \bmod \mathcal{P}) = \mathrm{HP}(\mathbb{A}^d)$.

In the proof of Theorem 1.3 (Section 5) we give an explicit formula for the asymptotic generic Newton polygon $\mathrm{GNP}(\mathbb{A}^d; \mathbb{F}_p)$ for p large enough, which depends only on d and the residue class of $p \bmod d$. We consider Theorem 3.3 as a major technical breakthrough of the present paper. Since it is more involved we postpone its discussion to Section 3. Our theorem has the following application in approximating slopes of Artin-Schreier curves (see [23, Corollary 1.3] for a proof).

Corollary 1.4. *There exists a Zariski dense open subset $\mathcal{U}$ defined over $\mathbb{Q}$ in $\mathbb{A}^d$ such that if $f \in \mathcal{U}(\overline{\mathbb{Q}})$ and $\mathcal{P}$ is any prime ideal in the ring of integers of $\mathbb{Q}(f)$ lying* *over p, we have*

$$\lim_{p\to\infty}\frac{\mathrm{NP}(X_f \bmod \mathcal{P})}{p-1}=\mathrm{HP}(\mathbb{A}^d).$$

The paper is organized as follows. Section 2 carries on our exposition of Dwork p-adic analysis from [23, Section 2], with emphasis on the semilinear theory, that is, the spot light is at exponential sums of finite fields that are *not* prime fields (in contrast to that in [23, Section 2]). Section 3 contains the main technical theorem (in Theorem 3.3) of this paper, in which we prove that for a large class of matrix F representing τ^{-1}-linear Frobenius map over a p-adic ring $\mathcal{O}_a$ the Newton polygon of its characteristic polynomial coincides with that of the matrix F_a representing linear Frobenius map. Section 4 defines appropriate Zariski dense open subset $\mathcal{W}_r$ in $\mathbb{A}^{d-1}$ for each residue class $r \bmod d$. Proofs for Theorems 1.3 lie in Section 5. In the same section we also prove that a certain stronger version of Theorem 1.3 is false, which answers a question of Daqing Wan.

Acknowledgments. I am deeply indebt to Hanfeng Li and Daqing Wan for enlightening suggestions and warm encouragements. I also thank Alan Adolphson and the referee for comments on an earlier version. This research was partially supported by a grant of Bjorn Poonen from the David and Lucile Packard Foundation and the University of California at Berkeley.

2. Dwork p-adic theory in a nutshell

The present section is in a sequel to [23, Section 2] yet it is self-contained for the convenience of the reader. We formulates the Dwork trace formula following [4, 5, 17, 18] at various stages without further notice. Our trace formula of this article concerns a Frobenius action on a finite dimensional quotient space, while that of [23] is considering a Frobenius action on an infinite dimensional vector space.

Let $\mathbb{Q}_{p^a}$ denote the unramified extension of $\mathbb{Q}_p$ of degree a. Let $\Omega_1=\mathbb{Q}_p(\zeta_p)$ and let $\Omega_a=\mathbb{Q}_{p^a}(\zeta_p)$. So Ω_a is the unramified extension of Ω_1 of degree a. Let $\mathcal{O}_1=\mathbb{Z}_p[\zeta_p]$ and $\mathcal{O}_a$ be the rings of integers in Ω_1 and Ω_a, respectively. Let τ be the lift of Frobenius endomorphism $c\mapsto c^p$ of $\mathbb{F}_q$ to Ω_a which fixes Ω_1.

Fix $f(x)=x^d+\sum_{i=1}^{d-1}a_ix^i\in\overline{\mathbb{Q}}[x]$. Suppose $\overline{f}(x)=x^d+\sum_{i=1}^{d-1}\overline{a}_ix^i\in\mathbb{F}_q[x]$ is a reduction of $f(x)$. Let $\hat{f}(x)=x^d+\sum_{i=1}^{d-1}\hat{a}_ix^i$, where $\hat{a}_i$ is the Teichmüller lifting of $\overline{a}_i$, that is, $\hat{a}_i$ lies in $\mathbb{Z}_{p^a}$ such that $\hat{a}_i\equiv\overline{a}_i \bmod p$ and $\hat{a}_i^q=\hat{a}_i$. We shall write $\vec{\overline{a}}:=(\overline{a}_1,\ldots,\overline{a}_{d-1})$ and $\vec{\hat{a}}:=(\hat{a}_1,\ldots,\hat{a}_{d-1})$.

Let $\theta(x)=E(\gamma x)$ where $E(\cdot)$ is the p-adic Artin-Hasse exponential function and γ is a root of $\log E(x)$ with $\mathrm{ord}_p\gamma=\frac{1}{p-1}$ (as defined in Section 1). We may write $\theta(x)=\sum_{m=0}^{\infty}\lambda_mx^m$ for $\lambda_m\in\mathcal{O}_1$. In fact, $\lambda_m=\frac{\gamma^m}{m!}$ and $\mathrm{ord}_p\lambda_m=\frac{m}{p-1}$ for $0\le m\le p-1$, and $\mathrm{ord}_p\lambda_m\ge\frac{m}{p-1}$ for $m\ge p$. Let $\vec{m}=(m_1,\ldots,m_{d-1})\in\mathbb{Z}_{\ge 0}^{d-1}$ and let $\vec{A}=(A_1,\ldots,A_{d-1})$ be variables. Denote $\vec{A}^{\vec{m}}:=A_1^{m_1}\cdots A_{d-1}^{m_{d-1}}$. Define for any $n\ge 0$ a polynomial in $\Omega_1[\vec{A}]$ below

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

Let $G(X):=\prod_{i=1}^d\theta(\hat{a}_iX^i)$. So $G(X)\in\mathcal{O}_a[[X]]$ and its expansion is precisely $G(X)=\sum_{n=0}^{\infty}G_n(\vec{\hat{a}})X^n\in\mathcal{O}_a[[X]]$.

Let K be a p-adic field over Ω_a. For any $c>0$ and $b\in\mathbb{R}$ let $\mathcal{L}_K(c,b)$ be the set of power series $\sum_{n=0}^{\infty}B_nX^n\in\Omega_a[[X]]$ with $B_n\in K$ and $\mathrm{ord}_pB_n\geq cn+b$. Let $\mathcal{L}_K(c)=\bigcup_{b\in\mathbb{R}}\mathcal{L}(c,b)$. For example, one may check $G(X)\in\mathcal{L}_K(\frac{1}{d(p-1)})$. Note that $\mathcal{L}_K(c)$ is a infinite dimensional vector space over K.

Consider the composition $\alpha:=\tau^{-1}\cdot\psi\cdot G(X)$ on $\mathcal{L}_K(c)$, where ψ is the Dwork ψ-operator on $\mathcal{L}_K(c)$ defined by $\psi(\sum_{n=0}^{\infty}B_nX^n)=\sum_{n=0}^{\infty}B_{np}X^n$, and $G(X)$ denotes the multiplication map by $G(X)$. One observes that α is τ^{-1}-linear Ω_a-endomorphism of $\mathcal{L}_{\Omega_a}(\frac{p}{d(p-1)})$. Write $\mathcal{L}$ and $\mathcal{L}^1$ for $\mathcal{L}_{\Omega_a}(\frac{p}{d(p-1)})$ and its subspace with no constant terms, respectively. For $\ell\geq 0$ let $\gamma_\ell:=\sum_{j=0}^{\ell}\frac{\gamma^{p^j}}{p^j}$. Let

$$R(X):=\sum_{\ell=0}^{\infty}\gamma_\ell\hat{f}^{\tau^\ell}(X^{p^\ell})=\sum_{\ell=0}^{\infty}\gamma_\ell\sum_{i=1}^{d}\hat{a}_i^{p^\ell}X^{ip^\ell}.$$

Let ∇ be a differential operator on $\mathcal{L}$ defined formally by

$$\nabla:=\exp(-R(X))\cdot X\frac{\partial}{\partial X}\cdot\exp(R(X)).$$

For any $\sum_{n=0}^{\infty}B_nX^n\in\mathcal{L}$, we have

$$\nabla(\sum_{n=1}^{\infty}B_nX^n)=\sum_{n=0}^{\infty}nB_nX^n+(X\frac{\partial R(X)}{\partial X})(\sum_{n=0}^{\infty}B_nX^n).$$

Clearly $\nabla(\mathcal{L})\subseteq\mathcal{L}^1$ and we define $\mathcal{M}:=\mathcal{L}^1/\nabla(\mathcal{L})$. Then $\mathcal{M}$ has the induced τ^{-1}-linear endomorphism α. (See [18, page 279] for more details.)

Let $\vec{e}$ denote the set of images of $\{X,X^2,\ldots,X^{d-1}\}$ in the quotient space $\mathcal{M}$. Then $\vec{e}$ form a basis for $\mathcal{M}$ over Ω_a, and $\dim_{\Omega_a}\mathcal{M}=d-1$. Let F be the matrix representation of α on $\mathcal{M}$ with respect to the basis $\vec{e}$. Let $G^{[a]}(X):=\prod_{j=0}^{a-1}G^{\tau^j}(X^{p^j})$. Let $\alpha_a:=\psi^a\cdot G^{[a]}(X)$, which is a (linear!) endomorphism of $\mathcal{M}$ over Ω_a. The case $a=1$ is thoroughly studied in [23]. Let F_a be the matrix representation of α_a on $\mathcal{M}$ with respect to this monomial basis $\vec{e}$. The map α on $\mathcal{M}$ is given by $\alpha\vec{e}=\vec{e}F$. Since $\alpha_a=\alpha^a$ and α is τ^{-1}-linear we see easily that $F_a=FF^{\tau^{-1}}\cdots F^{\tau^{-(a-1)}}$.

For any positive integer n let $\mathbb{M}_n(\cdot)$ denote the set of all n by n matrix over some ring. Let $\mathbb{I}_n$ denote the n by n identity matrix. By Dwork trace formula (see [6, Theorem 2.2] or [18, Section 2 and in particular (2.35)] [4, discussions in Section 2] for details) for any prime $\mathcal{P}$ over p of degree a we have

$$L(f \bmod \mathcal{P};T)=\det(\mathbb{I}_{d-1}-T\alpha_a|\mathcal{M})=\det(\mathbb{I}_{d-1}-TF_a). \tag{4}$$

Remark 2.1. One observes that the computation of the above L function is reduced to the process of diagonalization (or triangularization) of the matrix F_a. Write $\mathbb{Q}_{p^\infty}$ for the fraction field of $W(\overline{\mathbb{F}}_p)$. Even though the Dieudonné-Manin classification [13] asserts that it is plausible over $\mathbb{Q}_{p^\infty}(\zeta_p)$, it is far more than a small business in practice. One of our hardest tasks broils down to proving a stronger version of the Dieudonné-Manin classification holds in the sense that our matrix can be diagonalized (or triangularized) over the base field Ω_a. This is accomplished in Section 3. Our other challenging tasks include finding the Zariski open subset set defined over $\mathbb{Q}$, which is done in Section 4.

Example 2.2. Below we give a simple example only to demonstrate the essential difficulty and new effects amounted in Wan's conjecture when passing from $\mathbb{F}_p$ to

$\mathbb{F}_{p^a}$. Let $p \equiv -1 \bmod 4$ and $f(x) = x^4 + cx$ in $\mathbb{F}_{p^a}$. For p large enough we can compute and get

$$F = \begin{pmatrix} \frac{\gamma^{p-1}}{(p-1)!}\hat{c}^{p-1} & \frac{\gamma^{p-2}}{(p-2)!}\hat{c}^{p-2} & \frac{\gamma^{p-3}}{(p-3)!}\hat{c}^{p-3} \\ \frac{\gamma^{2p-1}}{(2p-1)!}\hat{c}^{2p-1} & \frac{\gamma^{2p-2}}{(2p-2)!}\hat{c}^{2p-2} & \frac{\gamma^{2p-3}}{(2p-3)!}\hat{c}^{2p-3} \\ \frac{\gamma^{3p-1}}{(3p-1)!}\hat{c}^{3p-1} & \frac{\gamma^{3p-2}}{(3p-2)!}\hat{c}^{3p-2} & \frac{\gamma^{3p-3}}{(3p-3)!}\hat{c}^{3p-3} \end{pmatrix}$$

in $\mathbb{M}_3(\Omega_a)$, where $\hat{c}$ is the Teichmüller lifting of c in Ω_a. Then our Frobenius matrix is $F_a = FF^{\tau^{-1}} \cdots F^{\tau^{-(a-1)}}$. If $a = 1$, the diagonalization process is linear algebra, that is, one needs $C^{-1}FC$ diagonal for some C over Ω_1. For $a > 1$, the process is τ^{-1}-linear, that is, one needs $C^{-\tau}FC$ diagonal for some C over Ω_a. The semilinear algebra involved is highly nontrivial (see Proposition 3.1). The reader who is interested in complete numerical analysis of lower degree cases are referred to two new papers [6] and [7].

3. The two Newton polygons

This section is technical and a key technical ingredient in our argument is a version of p-adic Banach fix point theorem.

Let m be a positive integer. For any m by m matrix M in with coefficients in Ω_a (i.e., $M \in \mathbb{M}_m(\Omega_a)$) and any $1 \leq n \leq m$ let $M^{[n]}$ denote the submatrix of M consisting of its first n rows and columns. Let $M_a := MM^{\tau^{-1}} \cdots M^{\tau^{-(a-1)}}$. We observe that if $\mathrm{ord}_p M_{ij} \to \mathrm{ord}_p M_{i1}$ for every j, $\mathrm{ord}_p M_{i,1} - \mathrm{ord}_p M_{i-1,1} > \xi$ for some constant $\xi > 0$ for every i, and $\mathrm{ord}_p \det M^{[n]} \to \sum_{i=1}^{n} \mathrm{ord}_p M_{i1}$ for every $1 \leq n \leq m$, then $\mathrm{NP}_q(\det(\mathbb{I}_m - TM_a)) = \mathrm{NP}_p(\det(\mathbb{I}_m - TM))$. This observation is highly nontrivial, so we will prove it in 3.1 below. Let

$$\begin{aligned} \delta(M) &:= (p-1) \min_{1\leq i\leq m-1} (\min_{1\leq j\leq m} \mathrm{ord}_p M_{i+1,j} - \max_{1\leq j\leq m} \mathrm{ord}_p M_{ij}); \\ \eta(M) &:= (p-1) \max_{1\leq n\leq m-1} (\mathrm{ord}_p \det M^{[n]} - \sum_{i=1}^{n} \min_{1\leq j\leq n+1} \mathrm{ord}_p M_{ij}). \end{aligned}$$

It is easy to observe that η and δ are nonnegative integers.

Write $\det(\mathbb{I}_m - TM) = 1 + c'_1 T + \cdots + c'_m T^m \in \Omega_a[T]$, and $\det(\mathbb{I}_m - TM_a) = 1 + b'_1 T + \cdots + b'_m T^m$.

Proposition 3.1. *Let M be in $\mathbb{M}_m(\mathcal{O}_a)$ (recall $\mathcal{O}_a$ is the ring of integers in Ω_a) such that $\delta(M) > m\,\eta(M)$. Then $\mathrm{ord}_p c'_n = \mathrm{ord}_p \det M^{[n]}$. There exists a unique upper triangular matrix C in $\mathbb{M}_m(\Omega_a)$ with all 1's on its diagonal and with $\mathrm{ord}_p C_{ij} \geq -\frac{\eta}{p-1}$ such that $M' := C^{-\tau}MC$ in $\mathbb{M}_m(\Omega_a)$ is lower triangular. Set $\det M^{[0]} := 1$. For any $1 \leq n \leq m$ one has*

$$\mathrm{ord}_p M'_{nn} = \mathrm{ord}_p \det M^{[n]} - \mathrm{ord}_p \det M^{[n-1]}. \quad (5)$$

Moreover, M' has strictly increasing p-adic orders down its diagonal.

To be useful to the reader, we make some remarks on what leads us to the formulation of the hypothesis in 3.1 and 3.3: Choose a different basis $\vec{e}_w := \{(\gamma^{1/d}X)^i\}_{1\leq i\leq d-1}$ for $\mathcal{L}_{\Omega_a(\gamma^{1/d})}(\frac{p}{d(p-1)})$ over $\Omega_a(\gamma^{1/d})$. Let F^w be the matrix for α under this basis. Then one notes that $\mathrm{ord}_p F^w_{ij} \geq i/d + r_{ij}/d(p-1)$ where r_{ij} is the least nonnegative residue of $-(pi-j) \bmod d$. As $p \to \infty$ this lower bound of $\mathrm{ord}_p F^w_{ij}$ converges to i/d for every i and j.

Proof. 1) Let C be an upper triangular matrix with (i,j)-th entry denoted by indeterminant C_{ij} and with all 1's on its diagonal. For any $1 \leq j \leq m$ and all $i = 1, \ldots, j-1$ set

$$(C^{-\tau}MC)_{ij} \quad = \quad 0. \tag{6}$$

Write $D := C^{-\tau}$. Then D is also upper triangular with all 1's on its diagonal. So we have

$$(C^{-\tau}MC)_{ij} = \sum_{k=i}^{m}\sum_{\ell=1}^{j} D_{ik}M_{k\ell}C_{\ell j} = \sum_{\ell=1}^{j-1} M_{i\ell}C_{\ell j} + M_{ij} + \sum_{k=i+1}^{m}\sum_{\ell=1}^{j} M_{k\ell}(D_{ik}C_{\ell j}) = 0.$$

Then one verifies that (6) is equivalent to

$$M^{[j-1]}\begin{pmatrix} C_{1j} \\ \vdots \\ C_{j-1,j} \end{pmatrix} + \begin{pmatrix} M_{1j} \\ \vdots \\ M_{j-1,j} \end{pmatrix} + \begin{pmatrix} \sum_{k=2}^{m}\sum_{\ell=1}^{j} M_{k\ell}(D_{1k}C_{\ell j}) \\ \vdots \\ \sum_{k=j}^{m}\sum_{\ell=1}^{j} M_{k\ell}(D_{j-1,k}C_{\ell j}) \end{pmatrix} = 0. \tag{7}$$

Now we introduce some notations. For $1 \leq i, j \leq m$, let $M_{(i,j)}$ denote the submatrix of M with its i-th row and the j-th column removed. Let M^* denote the *adjoint matrix* of M, that is, the matrix whose (i,j)-th entry is equal to $(-1)^{i+j}\det M_{(j,i)}$. From linear algebra we have that $MM^* = M^*M = \det M$. Consider $\vec{C}$-monomials, i.e., consider all C_{ij}'s and C^{τ}_{ij}'s as variables where $1 \leq i < j \leq m$. Our hypothesis on M implies that $\det M^{[n]} \neq 0$ for every $1 \leq n \leq m$. Thus we may multiply $(M^{[j-1]})^{-1}$ on the left-hand-side of (7) and get for all $1 \leq i < j$

$$C_{ij} \quad = \quad w_i(\vec{C}) + v_i \tag{8}$$

where $w_i(\vec{C}) = -\frac{(M^{[j-1]})^*}{\det(M^{[j-1]})}\sum_{k=i+1}^{m}\sum_{\ell=1}^{j} M_{k\ell}(D_{ik}C_{\ell j})$ and $v_i = -\frac{(M^{[j-1]})^*}{\det(M^{[j-1]})}M_{ij}$. It is easy to see that

$$\begin{aligned} \mathrm{ord}_p(M^{[j-1]})^*M_{ij} \quad &\geq \quad \min_{1\leq i\leq j-1} \mathrm{ord}_p \sum_{\ell=1}^{j-1}(-1)^{i+\ell}(\det M^{[j-1]}_{(\ell,i)})M_{\ell j} \\ &\geq \quad \sum_{\ell=1}^{j-1}\min_{1\leq k\leq j}(\mathrm{ord}_p M_{\ell k}). \end{aligned} \tag{9}$$

Thus

$$\mathrm{ord}_p(v_i) \geq \sum_{\ell=1}^{j-1}\min_{1\leq k\leq j}(\mathrm{ord}_p M_{\ell k}) - \mathrm{ord}_p \det M^{[j-1]} \geq -\frac{\eta}{p-1}. \tag{10}$$

Then the p-adic valuation of coefficients of any $\vec{C}$-monomial in $\vec{w}(\vec{C})$ is $\geq \frac{\delta-\eta}{p-1}$ by comparing to $\vec{v}$.

Now change variables by setting $X_{ij} := \gamma^{\eta}C_{ij}$ for all $1 \leq i < j \leq m$. Write $z_i(\vec{X})$ for $\gamma^{\eta}w_i(\vec{C})$ as polynomials in X_{ij} and X^{τ}_{ij}, one has $X_{ij} = z_i(\vec{X}) + v_i\gamma^{\eta}$. We claim that the right-hand-side of (8) for all $j = 1, \ldots, m$ together defines a contraction map with regard to X_{ij}'s on $\mathcal{O}_a^{\frac{m(m-1)}{2}}$. It suffices to show that $z_i(\vec{X})$ has all coefficients of p-adic valuation positive. If this is the case, then the Banach fixed point theorem applies and one has integral solutions X_{ij} and consequently $C, M' \in \mathbb{M}_m(\Omega_a)$ with $\mathrm{ord}_p C_{ij} \geq -\frac{\eta}{p-1}$.

Note that $D = C^{-\tau} = (C^*)^\tau$. For any $1 \le t < k \le m$, it is an exercise to show that $D_{tk} = (-1)^{t+k}\det(C_{(k,t)})^\tau$ is a degree $k-t$ polynomial in $\vec{C}$. Thus $w_i(\vec{C})$ is of degree $\le m$ in $\vec{C}$ Now it is another elementary exercise to show that coefficients of $\vec{X}$ in $z_i(\vec{X})$ has p-adic order $\ge \frac{\delta - m\eta}{p-1}$, which is positive by our hypothesis upon M.

2) From now on we assume C is as chosen above. It is an exercise to see for all $1 \le j \le m$ one has

$$M'_{jj} = (DMC)_{jj} = \sum_{1 \le i < j \le m} M_{ji}C_{ij} + M_{jj} + \sum_{1 \le i \le j < k \le m} M_{ki}D_{jk}C_{ij}.$$

Also note that

$$\det M^{[j]} = \det M^{[j-1]}(M_{jj} + \sum_{i=1}^{j-1} M_{ji}v_i),$$

so one has

$$M'_{jj} = \frac{\det M^{[j]}}{\det M^{[j-1]}} + \sum_{i=1}^{j-1} M_{ji}w_i + \sum_{1 \le i \le j < k \le m} M_{ki}D_{jk}C_{ij}. \tag{11}$$

By some simple computations, one finds every term exact the first one on the right-hand-side of (11) has p-adic valuation $> \mathrm{ord}_p \det M^{[j]} - \mathrm{ord}_p \det M^{[j-1]}$. Applying the isoscele principle, one concludes that $\mathrm{ord}_p(M'_{jj}) = \det M^{[j]} - \mathrm{ord}_p \det M^{[j-1]}$. This proves (5).

3) By a similar argument as above, one can show that there exists an upper triangular matrix C' with all 1's on the diagonal such that $M'' := {C'}^{-1}MC'$ is lower triangular and

$$\mathrm{ord}_p M''_{jj} \quad = \quad \mathrm{ord}_p \det M^{[j]} - \mathrm{ord}_p \det M^{[j-1]}$$

for $1 \le j \le m$. It follows easily that $\mathrm{ord}_p c'_n = \sum_{\ell=1}^n \mathrm{ord}_p M''_{\ell\ell} = \mathrm{ord}_p \det M^{[n]}$.

4) Finally we shall omit the proof of the last statement. The basic idea is using (5) to reduce to show $\mathrm{ord}_p \det M^{[n+1]} + \mathrm{ord}_p \det M^{[n-1]} > 2\,\mathrm{ord}_p \det M^{[n]}$. $\square$

Remark 3.2. For the purpose of computing L function according to Dwork's trace formula (4), the relation between $\mathrm{NP}_q(\det(\mathbb{I}_{d-1} - TF_a))$ and $\mathrm{NP}_p(\det(\mathbb{I}_{d-1} - TF))$ has been explored in the literature (see for example [20]) since the latter is much more straightforward to compute. However, as passing from F to F_a, the only thing we knew previously is that their corresponding Newton polygons have the same lower bounds (i.e., the Hodge polygon) and these Newton polygons are not generally equal. See some discussion including a good example in [8, Section 1.3] and a study of ordinary case in [20, Theorem 2.4]. Little is known besides these, yet the passage of Newton polygon data from F to F_a is the bottleneck in sharp slope estimations generally. In the theorem below we formulate an explicit criterion under which the two aforementioned Newton polygons coincide.

Theorem 3.3. *Let M be in $\mathbb{M}_m(\mathcal{O}_a)$ such that $\delta(M) > m\,\eta(M)$. Then $\mathrm{NP}_q(\det(\mathbb{I}_m - TM_a)) = \mathrm{NP}_p(\det(\mathbb{I}_m - TM))$, and they are equal to the lower convex hull in $\mathbb{R}^2$ of the points $(n, \mathrm{ord}_p \det M^{[n]})$ for $0 \le n \le m$.*

Proof. Let $C, M' \in \mathbb{M}_{d-1}(\Omega_a)$ be as in 3.1, then $C^{\tau^a} = C$ and ${M'}^{\tau^a} = M'$. So

$$\begin{aligned} C^{-\tau} M_a C^{\tau} &= C^{-\tau^{a+1}}(M^{\tau^a} M^{\tau^{a-1}} \cdots M^{\tau})C^{\tau} \\ &= (C^{-\tau}MC)^{\tau^a}(C^{-\tau}MC)^{\tau^{a-1}} \cdots (C^{-\tau}MC)^{\tau} \\ &= {M'}^{\tau^a}{M'}^{\tau^{a-1}} \cdots {M'}^{\tau} \\ &= (M')_a. \end{aligned}$$

By 3.1, one knows that $(M')_a$ is lower triangular with

$$\mathrm{ord}_p((M')_a)_{jj} = \sum_{\ell=0}^{a-1} \mathrm{ord}_p(M'_{jj})^{\tau^\ell} = a(\mathrm{ord}_p M'_{jj}) = a(\mathrm{ord}_p \det M^{[j]} - \mathrm{ord}_p \det M^{[j-1]}).$$

Since

$$\det(\mathbb{I}_m - TM_a) = \det(\mathbb{I}_m - TC^{-\tau}M_aC^{\tau}) = \det(\mathbb{I}_m - T(M')_a),$$

and that $\mathrm{ord}_p M'_{jj}$ being strictly increasing according to j, one has

$$\begin{aligned} \mathrm{ord}_q b'_n &= \frac{1}{a}\sum_{j=1}^{n} \mathrm{ord}_p((M')_a)_{jj} = \frac{1}{a}\sum_{j=1}^{n} a(\mathrm{ord}_p \det M^{[j]} - \mathrm{ord}_p \det M^{[j-1]}) \\ &= \mathrm{ord}_p(\det M^{[n]}). \end{aligned}$$

Thus $\mathrm{ord}_q b'_n = \mathrm{ord}_p \det M^{[n]} = \mathrm{ord}_q c'_n$ for every n. This finishes the proof. $\square$

4. Zariski dense open subset $\mathcal{W}_r$ in $\mathbb{A}^{d-1}$

We shall use an auxiliary $d-1$ by $d-1$ matrix $F^{\dagger}$ defined by $F^{\dagger}_{ij}(\vec{\hat{a}}) := G^{\tau^{-1}}_{pi-j}(\vec{\hat{a}})$ for every $\vec{\hat{a}}$. We outline our approach as below: (1) we find a Zariski dense open subset $\mathcal{X}_r$ of f's in which $\mathrm{ord}_p F_{ij} = \mathrm{ord}_p F^{\dagger}_{ij} = \frac{\lceil \frac{pi-j}{d} \rceil}{p-1}$; (2) we find a Zariski dense open subset $\mathcal{W}_r$ of f's in which $\mathrm{ord}_p(\det F^{[n]}) = \mathrm{ord}_p(\det(F^{\dagger})^{[n]}) = \frac{n(n+1)}{2d} + \epsilon_n$, both for all $p \equiv r \bmod d$ and p large enough. Basically we are looking for sufficient condition on p and f such that the Frobenius matrix F satisfies the hypothesis of 3.1. This is the key observation prepared for the proof in Section 5.

We adopt the same notation as that in [23, Section 3]. For convenience of the reader, we give complete definitions for all statements of our theorems. Let r be a positive integer with $1 \le r \le d-1$ and $\gcd(r,d) = 1$ for the rest of the section. For any $1 \le i, j \le d-1$, let r_{ij} (resp. r'_{ij}) be the least nonnegative residue of $-(ri-j) \bmod d$ (resp. $ri - j \bmod d$). Let $\delta_{ij} = 0$ for $j < r'_{i1}+1$ and let $\delta_{ij} = 1$ for $j \ge r'_{i1}+1$.

Let $1 \le n \le d-1$. Let $\vec{v} := (v_1, \ldots, v_n) \in \mathbb{Z}^n_{\ge 0}$ let $|\vec{v}| := \sum_{\ell=1}^{n} v_\ell$ and $\vec{v}! := v_1! \cdots v_n!$. For any $0 \le t \le n$ let S^t_n denote the subset of the symmetric group S_n consisting of all σ such that $\sum_{i=1}^{n} r_{i,\sigma(i)} = \min_{\sigma' \in S_n} \sum_{i=1}^{n} r_{i,\sigma'(i)} + dt$. For any $1 \le i, j \le n$ and $0 \le s \le n$ define a subset of $\mathbb{Z}^{d-1}_{\ge 0}$ by

$$\mathcal{M}^s_{ij} := \{\vec{m} = (m_1, \ldots, m_{d-1}) \in \mathbb{Z}^{d-1}_{\ge 0} | \sum_{\ell=1}^{d-1} \ell m_{d-\ell} = r_{ij} + ds\}.$$

Then let

$$H^s_{ij}(\vec{A}) := \sum_{\vec{m} \in \mathcal{M}^s_{ij}} \frac{(\frac{r_{i1}-1}{d} + d - 1)(\frac{r_{i1}-1}{d} + d - 2) \cdots (\frac{r_{i1}-1}{d} - \delta_{ij} + s + 1 - |\vec{m}|)}{\vec{m}!} \vec{A}^{\vec{m}}.$$

Clearly H_{ij}^s lies in $\mathbb{Q}[\vec{A}]=\mathbb{Q}[A_1,\ldots,A_{d-1}]$. For any $0\le t\le n$ let

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

Let t_n be the least nonnegative integer t such that $f_n^t\neq 0$. Let

$$\Psi_{d,r}(\vec{A}) := \prod_{0\le j-1\le i\le d-1} H_{ij}^0(\vec{A}), \qquad \Phi_{d,r}(\vec{A}) := \prod_{1\le n\le d-1} f_n^{t_n}(\vec{A}).$$

Let $\mathcal{X}_r$ and $\mathcal{Y}_r$ be the subset of $\mathbb{A}^{d-1}$ consisting of all $f(x)=x^d+a_{d-1}x^{d-1}+\cdots+a_1x$ with $\Psi_{d,r}|_{\vec{A}=\vec{a}}\neq 0$ and $\Phi_{d,r}|_{\vec{A}=\vec{a}}\neq 0$, respectively. Let $\mathcal{W}_r:=\mathcal{X}_r\cap\mathcal{Y}_r$.

For any $b\in\mathbb{Z}$ let $\gamma^{>b}$ denote a term with $\mathrm{ord}_p(\cdot)>\frac{b}{p-1}$. We define $\gamma^{\ge b}$ similarly. For any m by m matrix M and $n\le m$ let $M^{[n]}$ denote the truncated submatrix of M consisting of its first n rows and columns. Let

$$\epsilon_n \quad := \quad \frac{\min_{\sigma\in S_n}\sum_{\ell=1}^n r_{\ell,\sigma(\ell)}+dt_n}{d(p-1)}. \tag{12}$$

Lemma 4.1. *1) Let $\vec{a}\in\mathcal{X}_r(\overline{\mathbb{Q}})$. There exists $N>0$ such that for $p>N$ we have for any $0\le j-1\le i\le d-1$ that*

$$\mathrm{ord}_p F_{ij}^{\dagger}(\hat{\vec{a}}) \quad = \quad \frac{\lceil\frac{pi-j}{d}\rceil}{p-1}. \tag{13}$$

2) Let $\vec{a}\in\mathcal{Y}_r(\overline{\mathbb{Q}})$. There exists $N>0$ such that for $p>N$ we have for every $1\le n\le d-1$ that

$$\mathrm{ord}_p \det(F^{\dagger})^{[n]}(\hat{\vec{a}}) \quad = \quad \frac{n(n+1)}{2d}+\epsilon_n. \tag{14}$$

Proof. Since $F_{ij}^{\dagger}=G_{pi-j}^{\tau^{-1}}$, it suffices to prove our assertion for G_{pi-j}. Let notation be as in [23, 4.2 and 4.3]. By [23, 4.2], for $p\ge(d^2+1)(d-1)$ we have

$$K_{ij}^0(\vec{A}) \quad = \quad u_nH_{ij}^0(\vec{A})+\gamma^{\ge p-1} \tag{15}$$

for some p-adic unit u_n in $\mathbb{Z}_p$. Thus by [23, 4.3] one has

$$G_{pi-j}(\vec{A})=u_nH_{ij}^0(\vec{A})\gamma^{\lceil\frac{pi-j}{d}\rceil}+\gamma^{>\lceil\frac{pi-j}{d}\rceil}.$$

By the hypothesis $\vec{a}\in\mathcal{X}_r(\overline{\mathbb{Q}})$, we have $H_{ij}^0(\vec{a})\neq 0$. So for p large enough one gets

$$\mathrm{ord}_p(G_{pi-j}(\hat{\vec{a}}))=\frac{\lceil\frac{pi-j}{d}\rceil}{p-1}.$$

This proves 1). Part 2) follows immediately from [23, 4.3]. $\square$

Lemma 4.2. *Let $\vec{a}\in\mathcal{X}_r(\overline{\mathbb{Q}})$. For $0\le j-1\le i\le d-1$ and for p large enough one has*

$$\mathrm{ord}_p F_{ij}(\hat{\vec{a}}) \quad = \quad \frac{\lceil\frac{pi-j}{d}\rceil}{p-1}.$$

Proof. The auxiliary matrix $F^{\dagger}$ is p-adically close to F in the following sense. For any $\vec{a}\in\mathbb{A}^{d-1}(\overline{\mathbb{Q}})$ and $1\le i,j\le d-1$ we have

$$\mathrm{ord}_p(F_{ij}(\hat{\vec{a}})-F_{ij}^{\dagger}(\hat{\vec{a}})) \quad \ge \quad \frac{pi-j}{d(p-1)}+\frac{p}{d(p-1)}. \tag{16}$$

(See [6, Lemma 3.2] for a complete proof or follow the proof of Theorem 3.10 in [18].)

By (13) and (16) there exists $N > 0$ such that for all $p > N$ we have

$$\mathrm{ord}_p(F_{ij}(\vec{\hat{a}}) - F_{ij}^{\dagger}(\vec{\hat{a}})) \quad > \quad \mathrm{ord}_p F_{ij}^{\dagger}(\vec{\hat{a}}).$$

By the isosceles triangle principle, we have $\mathrm{ord}_p F_{ij}(\vec{\hat{a}}) = \mathrm{ord}_p F_{ij}^{\dagger}(\vec{\hat{a}})$, hence our assertion follows from (13). $\square$

Proposition 4.3. *The subset $\mathcal{W}_r$ is Zariski dense open in $\mathbb{A}^{d-1}$ defined over $\mathbb{Q}$. For $\vec{a} \in \mathcal{W}_r(\overline{\mathbb{Q}})$ and p large enough, one has for all $1 \leq n \leq d-1$ and $0 \leq j-1 \leq i \leq d-1$ that*

$$\mathrm{ord}_p F_{ij}(\vec{\hat{a}}) = \frac{\lceil \frac{pi-j}{d} \rceil}{p-1}, \quad \mathrm{ord}_p \det F^{[n]}(\vec{\hat{a}}) = \frac{n(n+1)}{2d} + \epsilon_n.$$

Proof. By [23, Section 3] one knows that $\Psi_{d,r} \neq 0$ and $\Phi_{d,r} \neq 0$. Thus the first assertion follows. The first equality is precisely proved in Lemma 4.2 above. We shall focus on the second equality for the rest of our proof. Write Δ for the set $\{1, \ldots, n\}$. By definition,

$$\begin{aligned}
\det F^{[n]} &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{\ell=1} ((F_{\ell,\sigma(\ell)} - F_{\ell,\sigma(\ell)}^{\dagger}) + F_{\ell,\sigma(\ell)}^{\dagger}) \\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{\ell=1}^{n} F_{\ell,\sigma(\ell)}^{\dagger} \\
&\quad + \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \sum_{\Delta_1 \subsetneq \Delta} (\prod_{\ell \in \Delta_1} (F_{\ell,\sigma(\ell)} - F_{\ell,\sigma(\ell)}^{\dagger}) \prod_{\ell' \in \Delta - \Delta_1} F_{\ell',\sigma(\ell')}^{\dagger}).
\end{aligned}$$

By (16) and (13) (since $\vec{a} \in \mathcal{X}_r(\overline{\mathbb{Q}})$), for p large enough we have

$$\begin{aligned}
& \mathrm{ord}_p(\det F^{[n]} - \det(F^{\dagger})^{[n]}) \\
\geq \quad & \min_{\sigma \in S_n, \Delta_1 \subsetneq \Delta} (\sum_{\ell \in \Delta_1} \mathrm{ord}_p(F_{\ell,\sigma(\ell)} - F_{\ell,\sigma(\ell)}^{\dagger}) + \sum_{\ell' \in \Delta - \Delta_1} \mathrm{ord}_p F_{\ell',\sigma(\ell')}^{\dagger}) \\
\geq \quad & \min_{\sigma \in S_n, \Delta_1 \subsetneq \Delta} (\sum_{\ell \in \Delta_1} (\frac{p\ell - \sigma(\ell)}{d(p-1)} + \frac{p}{d(p-1)}) + \sum_{\ell' \in \Delta - \Delta_1} \frac{p\ell' - \sigma(\ell')}{d(p-1)}) \\
\geq \quad & \frac{n(n+1)}{2d} + \frac{p}{d(p-1)}.
\end{aligned}$$

Since $\vec{a} \in \mathcal{Y}_r(\overline{\mathbb{Q}})$ and since ϵ_n goes to 0 as p approaches ∞, for p large enough this is strictly greater than $\mathrm{ord}_p \det(F^{\dagger})^{[n]}$ by (14). By the isosceles principle, we concludes our assertion. $\square$

5. The asymptotic generic Newton polygon and slope filtration

Let notations be as in previous sections. In particular, recall the matrices F and F_a represent the τ^{-1}-linear and linear Frobenius maps α and α_a, respectively.

Let $\mathcal{W} := \bigcap_r \mathcal{W}_r$ where r ranges in all $1 \leq r \leq d-1$ with $\gcd(r,d) = 1$. Let $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0 \in \mathbb{A}^d$. Let $\mathcal{U}$ be the pre-image of $\mathcal{W}$ in $\mathbb{A}^d$ under the projection map $\iota : \mathbb{A}^d \to \mathbb{A}^{d-1}$ by $\iota(f) = \vec{a}$ with $\vec{a} = (a_1, \ldots, a_{d-1})$. Let

$\mathcal{P}$ be any degree-a prime ideal in $\mathbb{Q}(f)$ over p. By (4) one has

$$\begin{aligned} L(x^d+\cdots+a_1x \bmod \mathcal{P};T) &= \det(\mathbb{I}_{d-1}-F_a(\vec{\hat{a}})\cdot T)\\ &= 1+b_1(\vec{a})T+\cdots+b_{d-1}(\vec{a})T^{d-1}\in\mathbb{Z}[\zeta_p][T]. \end{aligned}$$

One observes (see [23, Section 5]) that $\mathrm{ord}_q(b_n(f))=\mathrm{ord}_q(b_n(\vec{a}))$.

Proof of Theorem 1.3. By [23, Theorem 5.1], for p large enough one has $\mathrm{GNP}(\mathbb{A}^d;\mathbb{F}_p)$ equal to the the lower convex hull of points

$$(n,\frac{n(n+1)}{2d}+\epsilon_n)\quad\text{for } 0\le n\le d-1, \tag{17}$$

each of which is a vertex (recall ϵ_n from (12)). By 4.3, one sees that $\mathcal{W}$ is Zariski dense open in $\mathbb{A}^{d-1}$ and so is $\mathcal{U}$ in $\mathbb{A}^d$ by its definition. According to the discussion preceding the proof, it then suffices to prove our statements for $\vec{a}\in\mathcal{W}(\overline{\mathbb{Q}})$. Suppose $\vec{a}\in\mathcal{W}(\overline{\mathbb{Q}})$. From 4.3 it is not hard to verify that as p increases $\delta(F)$ is unbounded while $\eta(F)$ is bounded. Thus for p large enough, $F(\vec{\hat{a}})\in\mathbb{M}_{d-1}(\Omega_a)$ clearly satisfies the hypothesis of 3.3. So, by 3.3, one has $\mathrm{ord}_q b_n(\vec{a})=\mathrm{ord}_p(\det F^{[n]})$. Then by 4.3 one sees that $\mathrm{NP}(f \bmod \mathcal{P})$ is equal to the aforementioned convex hull given in (17). Finally, one notes that ϵ_n goes to 0 as p goes to infinity, this proves the theorem. □

Remark 5.1. We remark on another consequence of our results in 3.1 and 3.3 from a different viewpoint. It can be shown, by a symmetric argument as that in 3.1, that the two Frobenius matrices F and F_a for the exponential sums of f are diagonalizable over the base field Ω_a provided $f\in\mathcal{U}(\overline{\mathbb{Q}})$ and p is large enough. This implies that for $f\in\mathcal{U}(\overline{\mathbb{Q}})$ and p large enough, the F-crystal $\mathcal{M}$ (arisen from one-variable exponential sum in Section 2) has a slope filtration over Ω_a. More precisely, it is isogenous over Ω_a to the direct sum of rank-one F-crystals of slopes $\frac{1}{d}+\epsilon_1$, $\frac{2}{d}+(\epsilon_2-\epsilon_1)$, $\cdots$, and $\frac{d-1}{d}+(\epsilon_{d-1}-\epsilon_{d-2})$ respectively (necessarily in a strictly increasing order). This provides an improvement, in the case of one-variable exponential sums, to the Dieudonné-Manin classification which asserts that the F-crystal has a slope filtration over $\mathbb{Q}_{p^\infty}(\zeta_p)$ (see the classic of Manin [13, Chapter II], or see [11, Theorem 5.6] and [8]).

In the proposition below we show that a certain *stronger version* of Theorem 1.3 is false. This answers a question of Daqing Wan, proposed to me via email.

Proposition 5.2. *There does not exist any Zariski dense open subset $\mathcal{U}$ over $\overline{\mathbb{Q}}$ of $\mathbb{A}^d$ such that the following is satisfied:*

For any strictly increasing sequence $\{p_i\}_{i\ge1}$ of primes, and for any sequence $\{f_i(x)\}_{i\ge1}\in\mathcal{U}(\overline{\mathbb{Q}})$, where $\mathcal{P}_i$ is a prime idea of $\overline{\mathbb{Q}}(f_i)$ lying over p_i, one has

$$\lim_{i\to\infty}\mathrm{NP}(f_i \bmod \mathcal{P}_i)=\mathrm{HP}(\mathbb{A}^d).$$

Proof. Suppose there is such a Zariski dense open subset $\mathcal{U}$ defined over $\overline{\mathbb{Q}}$ in $\mathbb{A}^d$. Then $\mathcal{U}$ contains the complementary set of zeros of $h(\vec{t})$ for some nonzero polynomial $h(\vec{t})\in\overline{\mathbb{Q}}[\vec{t}]$ with $\vec{t}=(t_0,\ldots,t_{d-1})$ as the variable. Since $h(\vec{t})\ne0$ we also have $h(p\vec{t})\ne0$ for any prime p. We will construct a contradiction. Choose primes $p_i\equiv-1 \bmod d$ such that the sequence $\{p_i\}_{i\ge1}$ is strictly increasing. Let $f_i(x)=x^d+p_ic_{i,d-1}x^{d-1}+\cdots+p_ic_{i,1}x+p_ic_{i,0}$ where $\vec{c_i}=(c_{i,0},\ldots,c_{i,d-1})\in\overline{\mathbb{Q}}^d$ satisfies $h(p_i\vec{c_i})\ne0$. This exists because $h(p_i\vec{t})\ne0$. We observe easily that $f_i\in\mathcal{U}(\overline{\mathbb{Q}})$ and

$f_i(x) \equiv x^d \bmod \mathcal{P}_i$ for every $\mathcal{P}_i$ over p_i in $\mathbb{Q}(\vec{c_i})$. The latter congruence implies that $\mathrm{NP}(f_i \bmod \mathcal{P}_i) = \mathrm{NP}(x^d \bmod p_i)$. It is well-known that for $p_i \equiv -1 \bmod d$ the $\mathrm{NP}(x^d \bmod p_i)$ is a straight line of slope $1/2$ (see [23, Section 6]). Apparently this limit is not equal to the Hodge polygon. This proves the proposition. □

References

[1] Alan Adolphson; Steven Sperber: On the zeta function of a complete intersection. *Ann. Sci. École Norm. Sup.* **29** (1996), no. 3, 287–328.

[2] Alan Adolphson; Steven Sperber: Newton polyhedra and the degree of the L-function associated to an exponential sum. *Invent. Math.* **88** (1987), 555–569.

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

[4] Bernard Dwork: On the zeta function of a hypersurface. *Publication Math. IHES,* (1962), 5–68.

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

[6] Shaofang 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] Shaofang Hong: Newton polygons for L functions of exponential sums of polynomials of degree six over finite fields. *J. Number Theory* **97** (2002), 368–396.

[8] Nicholas Katz: Slope filtration of F-crystals. *Astérisque* **63** (1979), 113–164. Société Mathématique de France.

[9] Nicholas Katz: Crystalline cohomology, Dieudonné modules, and Jacobi sums, *Automorphic forms, Representation theory and Arithmetic.* Tata Institute of Fundamental Research, Bombay, 1979. 165–246.

[10] Nicholas Katz: Sommes exponentielles, *Astérisque* **79**, Société mathématique de France, 1980.

[11] Kiran Kedlaya: A p-adic local monodromy theorem. `arXiv.math.AG/0110124`, To appear in Annals of Math.

[12] Neal Koblitz: p-adic numbers, p-adic analysis, and Zeta-functions, (Second edition), *Graduate Texts in Mathematics* **58**. Springer-Verlag, 1984.

[13] Yuri Manin: The theory of commutative formal groups over fields of finite characteristic, *Russian Math. Surveys* **18** (1963), 1–83.

[14] Barry Mazur: Frobenius and the Hodge filtration, *Bull. of the American Math. Soc.*, **78** (1972), 653–667.

[15] Jasper Scholten; Hui June Zhu: The first slope case of Wan's conjecture. *Finite fields Appl.* **8** (2002), 414–419.

[16] Jasper Scholten; Hui June Zhu: Slope estimates of Artin-Schreier curves, *Compositio Math.* **137** (2003), 275–292.

[17] Steven Sperber: Congruence properties of the hyper-Kloosterman sum. *Compositio Math.* **40** (1980), 3–33.

[18] Steven Sperber: On the p-adic theory of exponential sums, *American J. Math.,* **109** (1986), 255–296.

[19] Gerard van der Geer; Marcel van der Vlugt: On the existence of supersingular curves of given genus, *J. Reine Angew. Math.,* **458** (1995), 53–61.

[20] Daqing Wan: Newton polygons of zeta functions and L functions. *Annals of Mathematics* **137** (1993), 249–293.

[21] Daqing Wan: Variation of p-adic Newton polygons of L functions for exponential sums. Preprint available at `http://www.math.uci.edu/dwan/Overview.html`.

[22] Lawrance C. Washington: An introduction to cyclotomic fields. Second edition. Graduate textbook in mathematics, vol. **83**. Springer, 1997.

[23] Hui June Zhu: p-adic variation of L functions of one variable exponential sums. I. *American J. Math.* **125** (2003), 669-690.

[24] Hui June Zhu: L-functions of exponential sums over one dimensional affinoids: Newton over Hodge. `http://arXiv.org/abs/math.NT/0302085`.

Department of mathematics and statistics, McMaster University, Hamilton, ON L8S 4K1, CANADA.

E-mail address: zhu@cal.berkeley.edu