5-9-2016

# Generalized p-adic Gauss Sums

Sandi Xhumari
*University of Connecticut - Storrs*, sandi.xhumari@uconn.edu

# Generalized $p$-adic Gauss Sums

Sandi Xhumari,

University of Connecticut, 2016

## ABSTRACT

In 2005 Blache studied certain generalized Gauss sums and established an analogue for them of Stickelberger's congruence for classical Gauss sums over finite fields. We improve on Blache's work in two ways: (i) simplify Blache's proof and give a second proof that works for a larger family of generalized Gauss sums, and (ii) give a $p$-adic lifting of Stickelberger's congruence for the larger family of generalized Gauss sums that is partial progress towards a version of the Gross–Koblitz formula for these sums. In addition, we study this larger family of generalized Gauss sums, prove a formula for them which simplifies computations, prove Stickelberger-type congruences for power series representations of these sums, make a conjecture for their degree over $\mathbb{Q}_p$ and prove cases of it. We conclude by making a family of conjectures for generalized Gross–Koblitz formulas regarding generalized Gauss sums and power series representations of them.

# Generalized $p$-adic Gauss Sums

Sandi Xhumari

M.S. University of Connecticut

B.S. Grand Valley State University

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2016

## APPROVAL PAGE

Doctor of Philosophy Dissertation

# Generalized $p$-adic Gauss Sums

Presented by

Sandi Xhumari, B.S. Math., M.S. Math.

Major Advisor _____
Keith Conrad

Associate Advisor _____
Álvaro Lozano-Robledo

Associate Advisor _____
Kyu-Hwan Lee

University of Connecticut

2016

# ACKNOWLEDGMENTS

It is a pleasure to finish writing my thesis by acknowledging everyone who has contributed to it. First, I am not sure how to begin thanking my advisor, Keith Conrad, who has gone well beyond his duty with his expert guidance, teachings and feedback. Keith tirelessly exchanged hundreds of emails with me. We met frequently for extended hours especially when I had a presentation or special event coming up. His concise advice is priceless and always amazingly full of knowledge. Writing this thesis would have been impossible without Keith's immense help. The quality of my mathematical writing and attention to detail has improved so much because of him. I am extremely lucky to have had the pleasure to learn from the best and hope that one day I will be able to help and guide my students just as Keith did with me.

My gratitude also goes to Liang Xiao, who helped me understand areas of mathematics I was unfamiliar with, gave me useful suggestions to improve my work and helped me complete one of my main results. I especially appreciate both him and Álvaro Lozano-Robledo for their helpful comments and feedback on my thesis. In addition, I would like to thank Francesco Baldassarri and Kiran Kedlaya among others for helping me understand some of the literature I was struggling with.

I am indebted to all the staff and professors at UConn for helping me painlessly navigate through graduate school. In particular, Monique Roy has always been there to lend a hand, cheer me up and provide a warm and friendly atmosphere to me and everyone else who needed it.

Of course, I would not be here without the unconditional love and support of my family. My parents have put my well-being above theirs and have always let me do what I desire in life. I also want to thank my sister for her encouragement. There are no words to express my gratitude for all of them.

Most of all, I would like to thank my dear and lovely wife Tze-Chun for always supporting and encouraging me to write my thesis as well as for taking care of me. Through her love, advice and attention to detail, she has played a crucial role in my success.

# Contents

## Table of Notation

| Notation | Definition |
|---|---|
| $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ | integers, rationals, reals, complexes |
| $p$, $q$ | prime number, $p^f$ for $f \geq 1$ |
| $\mathbb{Z}_p$, $\mathbb{Q}_p$, $\mathbb{C}_p$ | $p$-adic integers, $p$-adic numbers, $p$-adic complex numbers |
| $\mathfrak{m}_p$ | maximal ideal in the ring of integers of $\mathbb{C}_p$ |
| $\zeta_k$ | root of unity of order $k$ |
| $\mu_k$, $\mu'_k$ | groups of roots of unity of order dividing $k$ and of order exactly $k$ |
| $\mathbb{F}_p$, $\mathbb{F}_q$ | the finite fields of order $p$ and $q$ |
| $\mathbb{Z}_q$, $\mathbb{Q}_q$ | $\mathbb{Z}_p[\zeta_{q-1}]$, $\mathbb{Q}_p(\zeta_{q-1})$ |
| $K$ | finite extension of $\mathbb{Q}_q(\zeta_{p^l})$ |
| $\lvert K^\times\rvert_p = \lvert K^\times\rvert$ | set of $p$-adic absolute values of elements in $K^\times$ |
| $\mathcal{L}(r)$ | $\{\sum_{k=0}^{\infty} c_k X^k \in K[[X]] : \lvert c_k\rvert_p r^k \to 0\}$ |
| Tr | trace map $\mathbb{Q}_q \to \mathbb{Q}_p$, $\mathcal{L}(r) \to \mathcal{L}(r)$ or $\mathcal{L}(r)/D_{0,l} \to \mathcal{L}(r)/D_{0,l}\mathcal{L}(r)$ |
| $\omega$ | Teichmuller character $\mathbb{F}_q^\times \to \mathbb{C}_p^\times$ or lifting $\mathbb{F}_p[[X]] \to \mathbb{Z}_p[X]$ |
| $\chi$, $\psi$ | multiplicative and additive characters |
| $\psi_l$, $\psi_{l,v}$ | continuous additive characters $\mathbb{Z}_q \to \mathbb{C}_p^\times$ of order $p^l$ |
| $G(\chi, \psi)$ | Gauss sum associated to $\chi$ and $\psi$ |
| $G(\chi)$ | $-\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} \in \mathbb{C}$, where $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ |
| $G(a)$ | $-\sum_{x\in\mathbb{F}_q^\times} \omega(x)^{-a}\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} \in \mathbb{C}_p$ |
| $G_l(\chi)$ | $-\sum_{t\in\mu_{q-1}} \chi(t)\zeta_{p^l}^{\mathrm{Tr}(t)} \in \mathbb{C}$, where $\chi\colon \mu_{q-1} \to \mathbb{C}^\times$ |
| $G_l(a)$ | $-\sum_{t\in\mu_{q-1}} t^{-a}\zeta_{p^l}^{\mathrm{Tr}(t)} \in \mathbb{C}_p$ |
| $G_{l,v}(\chi)$ | $-\sum_{t\in\mu_{q-1}} \chi(t)\zeta_{p^l}^{\mathrm{Tr}(tv)} \in \mathbb{C}$, where $\chi\colon \mu_{q-1} \to \mathbb{C}^\times$ |
| $G_{l,v}(a)$ | $-\sum_{t\in\mu_{q-1}} t^{-a}\zeta_{p^l}^{\mathrm{Tr}(tv)} \in \mathbb{C}_p$ |
| $r_l$ | $(1/p)^{1/(p^{l-1}(p-1))}$ |
| $L_n(X)$, $L(X)$ | $X + X^p/p + X^{p^2}/p^2 + \cdots + X^{p^n}/p^n$, $L_\infty(X) = X + X^p/p + X^{p^2}/p^2 + \cdots$ |
| $z$, $\pi$ | $\zeta_p - 1 \in \mathbb{C}_p$, root of $L_1(X)$ in $\mathbb{C}_p$ closest to $z$ |
| $z_l$, $\pi_l$ | $\zeta_{p^l} - 1 \in \mathbb{C}_p$, root of $L(X)$ in $\mathbb{C}_p$ closest to $z_l$ |
| $\pi_{n,l}$ | root of $L_n(X)$ in $\mathbb{C}_p$ closest to $z_l$ |
| $\mathrm{AH}_n(X)$, $\mathrm{AH}(X)$ | $\exp(L_n(X))$, $\mathrm{AH}_\infty(X) = \exp(L(X))$ |
| $\Psi_n(T, X)$ | $\mathrm{AH}_n(TX)\,\mathrm{AH}_n(T^pX)\,\mathrm{AH}_n(T^{p^2}X)\cdots\mathrm{AH}_n(T^{p^{f-1}}X)$ |
| $\Psi(T, X)$ | $\Psi_\infty(T, X) = \mathrm{AH}(TX)\,\mathrm{AH}(T^pX)\,\mathrm{AH}(T^{p^2}X)\cdots\mathrm{AH}(T^{p^{f-1}}X)$ |
| $G(a, X)$, $G_v(a, X)$ | $-\sum_{t\in\mu_{q-1}} t^{-a}\,\mathrm{AH}(X)^{\mathrm{Tr}(t)}$, $-\sum_{t\in\mu_{q-1}} t^{-a}\,\mathrm{AH}(X)^{\mathrm{Tr}(tv)}$ |
| $G_{n,v}(a, X)$ | $-\sum_{t\in\mu_{q-1}} t^{-a}\,\mathrm{AH}_n(X)^{\mathrm{Tr}(tv)}$ |
| $\theta_{n,l}(X)$ | $\mathrm{AH}_n(X\pi_{n,l})$ |
| $\widehat{\theta_{n,l}}(X)$ | $\theta_{n,l}(X)\theta_{n,l}(X^p)\theta_{n,l}(X^{p^2})\cdots = \exp\left(\sum_{i=0}^{n-1} L_i(\pi_{n,l})X^{p^i}\right)$ |
| $a$, $a_i$ | integer $0 \leq a < q-1$, base $p$ expansion $a = a_0 + a_1p + \cdots + a_{f-1}p^{f-1}$ |
| $S(a)$ | sum of base $p$ digits of $a$, i.e. $a_0 + a_1 + \cdots + a_{f-1}$ |
| $a^{(i)}$ | $a_i + a_{i+1}p + \cdots + a_{f-1}p^{f-1-i} + a_0p^{f-i} + \cdots + a_{i-1}p^{f-1}$ |
| $R(f)$ | radius of convergence of series $f = f(X)$ |
| $D(f)$ | disc of convergence of series $f = f(X)$ |
| $n \gg l$ | $n \geq l \geq 1$ such that $p + p^2 + \cdots + p^{n-l+1} > n$, i.e, $r_l < R(\mathrm{AH}_n)$ |

# Chapter 1

# Gauss sums over finite fields

## 1.1 Definition over the complex numbers

In this section, we define Gauss sums over $\mathbb{C}$ and look at some examples. Let $p$ be a prime, $f \geq 1$ an integer, $q = p^f$ and denote by $\mathbb{F}_q$ the finite field of $q$ elements.

**Definition 1.1.1.** For a multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ and an additive character $\psi\colon \mathbb{F}_q \to \mathbb{C}^\times$, the *Gauss sum* associated to $\chi$ and $\psi$ is

$$G(\chi, \psi) = -\sum_{x \in \mathbb{F}_q^\times} \chi(x)\psi(x).$$

The minus sign in front of the sum makes later formulas look neater (see Theorem 1.2.2). The image of $\chi$ is in the $(q-1)^{\text{th}}$ roots of unity in $\mathbb{C}$, whereas the image of $\psi$ is in the $p^{\text{th}}$ roots of unity in $\mathbb{C}$. Fixing a nontrivial $p^{\text{th}}$ root of unity $\zeta_p \in \mathbb{C}$, there is a unique $y \in \mathbb{F}_q$ such that $\psi(x) = \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(xy)}$ for all $x \in \mathbb{F}_q$. When $y = 1$, we call $\psi(x) = \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)}$ the *basic additive character* $\mathbb{F}_q \to \mathbb{C}^\times$.

**Definition 1.1.2.** Fix a root of unity $\zeta_p \in \mathbb{C}$ of order $p$. For a multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$, the *basic Gauss sum* associated to $\chi$ is

$$G(\chi) = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{x+x^p+x^{p^2}+\cdots+x^{p^{f-1}}}.$$

Note that $G(\chi)$ is a special case of $G(\chi,\psi)$, where $\psi$ is the basic additive character. We now show that the basic additive character is essentially the only additive character $\mathbb{F}_q \to \mathbb{C}^\times$ we need to worry about for the Gauss sums $G(\chi,\psi)$.

**Lemma 1.1.3.** *If $\psi : \mathbb{F}_q \to C^\times$ is a nontrivial additive character, then*

$$G(\chi,\psi) = \overline{\chi(y)}G(\chi),$$

*where $\overline{\chi(y)} = 1/\chi(y) = \chi(y^{-1})$.*

*Proof.* Let $\psi : \mathbb{F}_q \to C^\times$ be a nontrivial additive character, or equivalently $\psi(x) = \zeta_p^{\mathrm{Tr}(xy)}$ for $y \neq 0$. By the change of variables $x \mapsto x/y$ we get

$$G(\chi,\psi) = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(xy)} = -\sum_{x\in\mathbb{F}_q^\times} \chi(x/y)\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} = \overline{\chi(y)}G(\chi),$$

where $\overline{\chi(y)} = 1/\chi(y) = \chi(y^{-1})$. □

Since $G(\chi,\psi)$ differs from $G(\chi)$ by the simple scaling factor $\overline{\chi(y)}$, it is enough to focus on the basic Gauss sums $G(\chi)$.

**Example 1.1.4.** Let $p = q = 5$ and $i \in \mathbb{C}$ be a primitive fourth root of unity. Set $\mathbb{F}_5 = \mathbb{Z}[i]/(2-i)$ so that $i \equiv 2 \bmod (2-i)$. We let $\eta\colon \mathbb{F}_5^\times \to \mathbb{C}^\times$ be defined by setting

$\eta(x)$ to be the unique fourth root of unity in $\mathbb{Z}[i]$ such that $\eta(x) \equiv x \bmod (2-i)$. Since every multiplicative character of $\mathbb{F}_5^\times$ is a power of $\eta$, we have the following table.

| $\chi$ | $G(\chi)$ |
|---|---|
| $\mathbb{1}$ | $-\zeta_5 - \zeta_5^2 - \zeta_5^3 - \zeta_5^4 = 1$ |
| $\eta$ | $-\zeta_5 - i\zeta_5^2 + i\zeta_5^3 + \zeta_5^4$ |
| $\eta^2$ | $-\zeta_5 + \zeta_5^2 + \zeta_5^3 - \zeta_5^4$ |
| $\eta^3$ | $-\zeta_5 + i\zeta_5^2 - i\zeta_5^3 + \zeta_5^4$ |

## 1.2 Properties of Gauss sums over $\mathbb{C}$

In this section, we discuss the Archimedean absolute value of $G(\chi)$, the Hasse-Davenport relation, $L$-functions of Gauss sums, and the Riemann Hypothesis associated to them.

### 1.2.1 Archimedean absolute value and Hasse-Davenport relation

When $\chi = \mathbb{1}$ is trivial,

$$G(\mathbb{1}) = -\sum_{x \in \mathbb{F}_q^\times} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} - 1 + 1 = -\sum_{x \in \mathbb{F}_q} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} + 1 = 1, \tag{1.2.1}$$

since the sum of all the values of a non-trivial additive character $\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)}$ over its full domain $\mathbb{F}_q$ is 0. For non-trivial $\chi$, while the number $G(\chi)$ can vary, its absolute value does not.

**Theorem 1.2.1.** *For non-trivial multiplicative characters* $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$*, we have*

$$|G(\chi)| = \sqrt{q}.$$

*Proof.* See [11, p. 4]. □

In particular, this implies that $G(\chi) \neq 0$, which will be used in Section 1.4.2 to get the degree over $\mathbb{Q}_p$ of the $p$-adic-valued version of $G(\chi)$.

The Hasse-Davenport relation connects Gauss sums on $\mathbb{F}_q$ and extensions of $\mathbb{F}_q$. For a multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ and integer $n \geq 1$, $\chi^{(n)} := \chi \circ \mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}$ is a multiplicative character $\mathbb{F}_{q^n}^\times \to \mathbb{C}^\times$. Hence we have

$$G(\chi^{(n)}) = -\sum_{x \in \mathbb{F}_{q^n}^\times} \chi^{(n)}(x)\zeta_p^{\mathrm{Tr}_{\mathbb{F}_{q^n}/\mathbb{F}_p}(x)},$$

where the sum is over $\mathbb{F}_{q^n}^\times$ instead of $\mathbb{F}_q^\times$.

**Theorem 1.2.2** (Hasse-Davenport)**.** *For any multiplicative character* $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$*, we have* $G(\chi^{(n)}) = G(\chi)^n$ *for all* $n \geq 1$*.*

*Proof.* See [9, Theorem 1, p. 162], noting Gauss sums there are not defined with an overall minus sign. □

### 1.2.2 $L$-functions and the Riemann Hypothesis

In this section, we derive a formula for a certain $L$-function in terms of $G(\chi)$ and then verify the Riemann Hypothesis for that $L$-function.

**Definition 1.2.3.** Fix a root of unity $\zeta_p \in \mathbb{C}$ of order $p$ and a multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$. Let $\eta\colon (\mathbb{F}_q[X]/(X^2))^\times \to \mathbb{C}^\times$ be a multiplicative character defined by $\eta(a(1+bX) \bmod X^2) = \overline{\chi(a)}\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(b)}$. Extend $\eta$ to 0 mod $X^2$ by $\eta(0 \bmod X^2) = 0$ and lift $\eta$ to $\mathbb{F}_q[X]$ by declaring $\eta(g) = \eta(g \bmod X^2)$. This function $\eta$ on $\mathbb{F}_q[X]$ is totally multiplicative. The $L$-function associated to this data is

$$L(T) = \sum_{\text{monic } g} \eta(g) T^{\deg(g)} = \sum_{k=0}^{\infty} \left( \sum_{\deg(g)=k} \eta(g) \right) T^k,$$

where the sum is over all monic $g(X) \in \mathbb{F}_q[X]$.

Plugging in $T = q^{-s} = 1/q^s$ we get

$$L(q^{-s}) = \sum_{\text{monic } g} \frac{\eta(g)}{\mathrm{N}(g)^s} = \sum_{k=0}^{\infty} \left( \sum_{\deg(g)=k} \eta(g) \right) \frac{1}{q^{ks}},$$

where $\mathrm{N}(g) = |\mathbb{F}_q[X]/g| = q^{\deg(g)}$. As we can see, $L(q^{-s})$ is analogous to the $L$-function of a Dirichlet character.

By [6, Theorem 2.1], the power series $L(T)$ is a linear polynomial. The only non-zero monic element of degree zero in $\mathbb{F}_q[X]$ is $g(X) = 1$. Letting $g_c(X) = X + c =$

$c(1 + X/c)$ denote all the degree one monic polynomials in $\mathbb{F}_q[X]$, we get

$$\begin{aligned} L(T) &= 1 + \sum_{c \in \mathbb{F}_q^\times} \eta(g_c) T \\ &= 1 + \sum_{c \in \mathbb{F}_q^\times} \overline{\chi(c)} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(1/c)} T \\ &= 1 + \sum_{c \in \mathbb{F}_q^\times} \overline{\chi(1/c)} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(c)} T \\ &= 1 - G(\chi) T. \end{aligned}$$

In particular, when $\chi = \mathbb{1}$ is trivial, (1.2.1) implies $G(\mathbb{1}) = 1$ and so $L(T) = 1 - T$. The corresponding $L$-function is then $L(q^{-s}) = 1 - G(\chi)/q^s$, whose roots have real part equal to $1/2$ if and only if $|G(\chi)| = \sqrt{q}$. Therefore the Riemann Hypothesis for $L(q^{-s})$ is equivalent to Theorem 1.2.1.

Keeping in mind the power series expansion for $\log(1 - T)$, we get

$$L(T) = \exp\left(\log(1 - G(\chi)T)\right) = \exp\left(\sum_{n=1}^{\infty} -\frac{(G(\chi)T)^n}{n}\right).$$

Theorem 1.2.2 implies

$$L(T) = \exp\left(\sum_{n=1}^{\infty} -G(\chi^{(n)})\frac{T^n}{n}\right). \tag{1.2.2}$$

This exponential form of $L(T)$ is equivalent to its definition through Theorem 1.2.2. In Section 2.2, we use this form to generalize the definition of $L(T)$.

## 1.3 Definition over the $p$-adic numbers

We now replace the complex-valued additive and multiplicative characters $\chi$ and $\psi$ in the Gauss sums $G(\chi, \psi)$ with $p$-adic-valued characters in order to formulate Stickelberger's congruence and compute the non-Archimedean absolute value of Gauss sums. Let $\mathbb{C}_p$ be the completion of the algebraic closure of $\mathbb{Q}_p$. For any $n \geq 1$, we will denote by $\zeta_n$ a root of unity of order $n$ in $\mathbb{C}_p$. In addition, we let $\mu_n = \{\zeta_n^k \in \mathbb{C}_p : 0 \leq k < n\}$ denote the group of $n^{\text{th}}$ roots of unity in $\mathbb{C}_p$. Note that $G(\chi, \psi)$ is a finite sum of roots of unity in $\mathbb{C}$ and thus it is an algebraic integer. Therefore the $p$-adic analogue of $G(\chi, \psi)$ we work with in this chapter is an embedding of $G(\chi, \psi)$ in $\mathbb{C}_p$.

Let $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}_p^\times$ and $\psi\colon \mathbb{F}_q \to \mathbb{C}_p^\times$ be $p$-adic-valued multiplicative and additive characters, respectively. Just as for complex-valued additive characters, the image of $\psi$ is in the $p^{\text{th}}$ roots of unity $\mu_p \subset \mathbb{C}_p$. Fixing a nontrivial $p^{\text{th}}$ root of unity $\zeta_p \in \mathbb{C}_p$, there is a unique $y \in \mathbb{F}_q$ such that $\psi(x) = \zeta_p^{\text{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(xy)}$ for all $x \in \mathbb{F}_q$. When $y = 1$, we call $\psi(x) = \zeta_p^{\text{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)}$ the *basic p-adic additive character* $\mathbb{F}_q \to \mathbb{C}_p^\times$. The computation in the proof of Lemma 1.1.3 applies to $p$-adic-valued Gauss sums and shows that we only need to consider Gauss sums involving the basic $p$-adic additive character.

Set $\mathbb{Q}_q = \mathbb{Q}_p(\zeta_{q-1})$, the unramified extension of $\mathbb{Q}_p$ with residue field of order $q$, and $\mathbb{Z}_q = \mathbb{Z}_p[\zeta_{q-1}]$, the ring of integers of $\mathbb{Q}_q$. Then $\mathbb{F}_q$ can be realized as the residue field $\mathbb{Z}_q/p\mathbb{Z}_q$. With this notation, every $p$-adic multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}_p^\times$ has values in the group of $(q-1)^{\text{th}}$ roots of unity $\mu_{q-1} \subset \mathbb{Z}_q^\times$ and the character group of $\mathbb{F}_q^\times$ with values in $\mathbb{C}_p$ is a cyclic group of order $q-1$, since $\mathbb{F}_q^\times$ is a cyclic group of order $q-1$. There's a preferred choice for a generator of the character group of $\mathbb{F}_q^\times$ with values in $\mathbb{C}_p$, which we define next.

**Definition 1.3.1.** The *Teichmüller character* $\omega\colon \mathbb{F}_q^\times \to \mu_{q-1}$ is the $p$-adic multi-

plicative character where $\omega(x)$ is the unique $(q-1)^{\text{th}}$ root of unity in $\mathbb{Z}_q$ such that $\omega(x) \equiv x \bmod p\mathbb{Z}_q$.

Note that $\mathbb{F}_q^\times = \mathbb{F}_p\left(\overline{\zeta}_{q-1}\right)^\times = \left\{\overline{\zeta}_{q-1}^n : 0 \leq n < q-1\right\}$, where $\overline{\zeta}_{q-1} \in \mathbb{Z}_q/p\mathbb{Z}_q = \mathbb{F}_q$ is the reduction of $\zeta_{q-1} \in \mathbb{Z}_q$ modulo $p\mathbb{Z}_q$. Hence, we explicitly have $\omega\left(\overline{\zeta}_{q-1}^n\right) = \zeta_{q-1}^n$. This implies that $\omega$ has order $q-1$ and therefore every multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}_p^\times$ is a power of $\omega$.

**Remark 1.3.2.** When $p = q = 5$, the Teichmüller character $\omega$ is analogous to $\eta$ in Example 1.1.4. However, in the construction of $\mathbb{F}_5$ there, the ideal $(2-i)\mathbb{Z}[i]$ could have been replaced by $(2+i)\mathbb{Z}[i]$, which makes the definition of $\eta$ depend on the ideal over $5\mathbb{Z}$ we mod out by. In the 5-adic case, on the other hand, the prime ideal $5\mathbb{Z}_5$ is the unique ideal we can mod out $\mathbb{Z}_5$ by to construct $\mathbb{F}_5$. Thus, the Teichmüller character $\omega$ is canonical in this sense.

**Definition 1.3.3.** Fix $\zeta_p \in \mathbb{C}_p$ a root of unity of order $p$. For $0 \leq a < q-1$, the *basic p-adic Gauss sum* associated to $a$ is

$$G(a) = -\sum_{x \in \mathbb{F}_q^\times} \omega(x)^{-a} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)}.$$

The signs in front of the sum as well as in the exponent of $\omega(x)^{-a}$ make later formulas look nicer (see Remark 1.4.6). As was noted above, every $p$-adic multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}_p^\times$ is a power of $\omega$, i.e. of the form $\chi = \omega^{-a}$ for some unique $0 \leq a < q-1$. Compare the following example with Example 1.1.4.

**Example 1.3.4.** Let $p = q = 5$ and $i \in \mathbb{Z}_5$ be the primitive fourth root of unity such that $i \equiv 2 \bmod 5\mathbb{Z}_5$. We have the following table.

| $a$ | $G(a)$ |
|---|---|
| 0 | $-\zeta_5 - \zeta_5^2 - \zeta_5^3 - \zeta_5^4 = 1$ |
| 1 | $-\zeta_5 + i\zeta_5^2 - i\zeta_5^3 + \zeta_5^4$ |
| 2 | $-\zeta_5 + \zeta_5^2 + \zeta_5^3 - \zeta_5^4$ |
| 3 | $-\zeta_5 - i\zeta_5^2 + i\zeta_5^3 + \zeta_5^4$ |

## 1.4 Properties of $p$-adic Gauss sums

In this section, we find the degree of basic $p$-adic Gauss sums over $\mathbb{Q}_p$, state Stickelberger's congruence, and use it to find $|G(a)|_p$. Then we recall how Stickelberger's congruence lifts to an equality, the Gross–Koblitz formula. After sketching a proof of the Gross–Koblitz formula, we state Baldassarri's generalization.

### 1.4.1 Degree over $\mathbb{Q}_p$

From Definition 1.3.3 and the fact that the Teichmüller character takes values in $\mu_{q-1}$, we have $G(a) \in \mathbb{Z}_p[\zeta_{q-1}, \zeta_p] = \mathbb{Z}_q[\zeta_p]$.

**Theorem 1.4.1.** *The basic $p$-adic Gauss sums $G(a)$ lie in $\mathbb{Z}_p[\zeta_p]$ for all $0 \leq a < q-1$. More precisely, $\mathbb{Q}_p(G(a))/\mathbb{Q}_p$ is the unique extension of degree $(p-1)/(a, p-1)$ that lies between $\mathbb{Q}_p(\zeta_p)$ and $\mathbb{Q}_p$.*

*Proof.* Recall that $q = p^f$, $\mathbb{Q}_q(\zeta_p) = \mathbb{Q}_p(\zeta_{q-1}, \zeta_p)$ has degree $f(p-1)$ over $\mathbb{Q}_p$, and the Galois group $\mathrm{Gal}(\mathbb{Q}_q(\zeta_p)/\mathbb{Q}_p) \cong \mathbb{Z}/f\mathbb{Z} \times \mathbb{F}_p^\times$ is generated by the automorphisms $\phi_p, \sigma_y \colon \mathbb{Q}_q(\zeta_p) \to \mathbb{Q}_q(\zeta_p)$ for $y \in \mathbb{F}_p^\times$, which are determined respectively by their values

on $\zeta_{q-1}$ and $\zeta_p$:

$$\phi_p(\zeta_{q-1}) = \zeta_{q-1}^p, \;\; \phi_p(\zeta_p) = \zeta_p \quad \text{and} \quad \sigma_y(\zeta_{q-1}) = \zeta_{q-1}, \;\; \sigma_y(\zeta_p) = \zeta_p^y.$$

The automorphism $\phi_p$ is called the Frobenius automorphism. Set $\overline{\phi}_p\colon \mathbb{F}_q \to \mathbb{F}_q$ to be the corresponding Frobenius automorphism of the residue field $\mathbb{F}_q = \mathbb{Z}_q[\zeta_p]/(\zeta_p - 1)$ over $\mathbb{F}_p$. Applying $\phi_p$ to $G(a)$ and using the facts: (1) $\omega(x) \in \mu_{q-1}$, (2) $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x) = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x^p)$, (3) $\overline{\phi}_p\colon \mathbb{F}_q^\times \to \mathbb{F}_q^\times$ is a group isomorphism, we get

$$\phi_p(G(a)) \overset{(1)}{=} -\sum_{x\in\mathbb{F}_q^\times} \omega(x)^{p(-a)} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} \overset{(2)}{=} -\sum_{x\in\mathbb{F}_q^\times} \omega(x^p)^{-a} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x^p)} \overset{(3)}{=} G(a).$$

This implies that $G(a)$ lies in $\mathbb{Q}_p(\zeta_p)$, the fixed field of $\mathrm{Gal}(\mathbb{Q}_q(\zeta_p)/\mathbb{Q}_p(\zeta_p)) = \langle \phi_p \rangle$. Hence $G(a)$ lies in $\mathbb{Z}_p[\zeta_p]$, as claimed. Thus $\mathbb{Q}_p(G(a))/\mathbb{Q}_p$ is totally ramified since $\mathbb{Q}_p(\zeta_p)/\mathbb{Q}_p$ is.

Now following the same steps as in the proof of Lemma 1.1.3, for $y \in \mathbb{F}_p^\times$ we have

$$\sigma_y(G(a)) = -\sum_{x\in\mathbb{F}_q^\times} \omega(x)^{-a} \zeta_p^{y\,\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} = -\sum_{x\in\mathbb{F}_q^\times} \omega(x)^{-a} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(xy)} = \omega(y)^a G(a).$$

The number $G(a)$ is an algebraic integer, and we can view it in $\mathbb{C}$. Theorem 1.2.1 implies $|G(\chi)| = \sqrt{q} \neq 0$ and so $G(a) \neq 0$ for $a \neq 0$, while (1.2.1) implies $G(0) = 1 \neq 0$. Hence, $\sigma_y(G(a)) = G(a)$ if and only if $\omega(y)^a = 1$, which is equivalent to $y^a = 1$ since $\omega\colon \mathbb{F}_q^\times \to \mu_{q-1}$ is an injective homomorphism. The set $\{y \in \mathbb{F}_p^\times : y^a = 1\}$ is a

subgroup of order $(a, p-1)$ in the cyclic group $\mathbb{F}_p^\times$. Galois theory implies

$$[\mathbb{Q}_p(G(a)) : \mathbb{Q}_p] = \frac{p-1}{\left|\{y \in \mathbb{F}_p^\times : y^a = 1\}\right|} = \frac{p-1}{(a, p-1)}.$$

Since the Galois group of $\mathbb{Q}_p(\zeta_p)/\mathbb{Q}_p$ is cyclic, there is a unique intermediate extension of degree $(p-1)/(a, p-1)$ over $\mathbb{Q}_p$. Hence, $\mathbb{Q}_p(G(a))$ is the unique extension of degree $(p-1)/(a, p-1)$ over $\mathbb{Q}_p$ that lies in $\mathbb{Q}_p(\zeta_p)$. □

### 1.4.2 Computing $p$-adic Gauss sums

For computations of $G(a)$ we can use Sage, but at this time it supports only unramified or totally ramified $p$-adic fields, not composite extensions of $\mathbb{Q}_p$. Hence, to construct $G(a)$ and run computations, I found a formula for $G(a)$ as an element of $\mathbb{Z}_p[\zeta_p]$ instead of as an element of the composite extension $\mathbb{Z}_q[\zeta_p]$ of $\mathbb{Z}_p$, which is how it arises from its definition.

In order to state such a formula for $G(a)$ we use some additional notation. For any polynomial $g(X)$, denote by $d(g)$ its degree and by $s(g)$ the coefficient of $X^{d(g)-1}$ (if $d(g) = 0$ then we define $s(g) = 0$). By Hensel's lemma, there is a lifting $\omega$ of any monic irreducible polynomial $g(X) \in \mathbb{F}_p[x]$ dividing $X^{q-1} - 1$ in $\mathbb{F}_p[X]$ to a monic irreducible polynomial $\omega(g)(X) \in \mathbb{Z}_p[X]$ dividing $X^{q-1} - 1$ in $\mathbb{Z}_p[X]$ such that $\omega(g)(X) \equiv g(X) \bmod p\mathbb{Z}_p[X]$ (i.e., the coefficients of $\omega(g)$ reduce to the coefficients of $g \bmod p$). In particular, $d(g) = d(\omega(g))$ and $s(g) \equiv s(\omega(g)) \bmod p\mathbb{Z}_p$. We call this lifting $\omega$ because the Teichmüller character (Definition 1.3.1) $\omega\colon \mathbb{F}_q^\times \to \mu_{q-1}$ sends each root of $g(X)$ to a root of $\omega(g)(X)$: if $g(X) = \prod_{k=1}^{d(g)} (X - \alpha_k)$ where $\alpha_k \in \mathbb{F}_q^\times$, then $\omega(g)(X) = \prod_{k=1}^{d(g)} (X - \omega(\alpha_k))$.

**Example 1.4.2.** Let $q = 8$. In $\mathbb{F}_2[X]$, we have

$$X^7 - 1 = (X - 1)(X^3 + X + 1)(X^3 + X^2 + 1),$$

where each factor is monic irreducible. The lifting $\omega$, sends

$$\begin{aligned}
\omega(X - 1) &= X - 1, \\
\omega(X^3 + X + 1) &= X^3 + (2 + 2^2 + 2^5 + \cdots)X^2 + (1 + 2^2 + 2^5 + \cdots)X - 1, \\
\omega(X^3 + X^2 + 1) &= X^3 + (1 + 2 + 2^3 + \cdots)X^2 + (2 + 2^3 + 2^4 + \cdots)X - 1,
\end{aligned}$$

which are the irreducible factors of $X^7 - 1$ in $\mathbb{Z}_2[X]$.

**Theorem 1.4.3.** *Fix $0 \leq a < q - 1$ and $\zeta_p \in \mathbb{C}_p$. If $g_a$ and $\omega(g)_a$ respectively denote the minimal polynomials of $x^{-a}$ over $\mathbb{F}_p$ and $\omega(x)^{-a}$ over $\mathbb{Q}_p$ for any root $x \in \mathbb{F}_q^\times$ of $g(X)$, then*

$$G(a) = \sum_{g(X)|(X^{q-1}-1)} \frac{d(g)}{d(g_a)} s(\omega(g)_a)\zeta_p^{-fs(g)/d(g)}, \tag{1.4.1}$$

*where the sum is over all the monic irreducible factors $g(X)$ of $X^{q-1} - 1$ in $\mathbb{F}_p[X]$.*

The key point in this theorem is that each term in the sum in (1.4.1) lies in $\mathbb{Z}_p[\zeta_p]$.

*Proof.* For any monic irreducible polynomial $g(X)$ that is a factor of $X^{q-1} - 1$ in $\mathbb{F}_p[X]$, $d(g)$ necessarily divides $f$ (recall $q = p^f$) and the trace $\mathbb{F}_q \to \mathbb{F}_p$ of any root of $g(X)$ is equal to $-\frac{f}{d(g)}s(g)$. Hence, collecting together in $G(a)$ all $x \in \mathbb{F}_q^\times$ with a common minimal polynomial over $\mathbb{F}_p$, we have

$$G(a) = -\sum_{g(X)|(X^{q-1}-1)} \left( \sum_{x \text{ root of } g(X)} \omega(x)^{-a} \right) \zeta_p^{-fs(g)/d(g)}, \tag{1.4.2}$$

where the outer sum is over all the irreducible monic factors $g(X)$ of $X^{q-1}-1$ in $\mathbb{F}_p[X]$ and the inner sum is over all the roots $x \in \mathbb{F}_q$ of $g(X)$.

In addition, if $x$ is a root of $g(X)$ in $\mathbb{F}_q$, then the full set of roots of $g(X)$ is $\{x, x^p, x^{p^2}, \ldots, x^{p^{d-1}}\}$, where $d = d(g)$, and therefore

$$\sum_{x \text{ root of } g(X)} \omega(x)^{-a} = \omega(x)^{-a} + \omega(x^p)^{-a} + \omega(x^{p^2})^{-a} + \cdots + \omega(x^{p^{d-1}})^{-a}. \tag{1.4.3}$$

Denote by $S(g,a)$ the sum in (1.4.3). It is in $\mathbb{Z}_q$ and invariant under $\mathrm{Gal}(\mathbb{Q}_q/\mathbb{Q}_p)$, so $S(g,a) \in \mathbb{Z}_p$.

We now want an explicit formula for $S(g,a)$. It looks like a trace and indeed we can realize it as one. For $x \in \mathbb{F}_q^\times$ with minimal polynomial $g(X)$ over $\mathbb{Q}_p$, the minimal polynomial of $\omega(x)^{-a}$ over $\mathbb{Q}_p$ is a monic irreducible factor, say $\omega(g)_a(X) \in \mathbb{Z}_p[X]$, of $X^{q-1}-1$. In addition, since $\mathbb{Q}_p(\omega(x)^{-a}) \subseteq \mathbb{Q}_p(\omega(x))$, the degree $d(g_a) = d(\omega(g)_a)$ divides $d(g) = d(\omega(g))$. It follows that $S(g,a) = \mathrm{Tr}_{\mathbb{Q}_{p^{d(g)}}/\mathbb{Q}_p}(\omega(x)^{-a}) = -\frac{d(g)}{d(g_a)} s(\omega(g)_a)$. Plugging this formula into (1.4.2), we get

$$G(a) = \sum_{g(X) | (X^{q-1}-1)} \frac{d(g)}{d(g_a)} s(\omega(g)_a) \zeta_p^{-fs(g)/d(g)},$$

which is what we wanted and can be used to encode $G(a)$ in Sage. □

### 1.4.3 Stickelberger's congruence

From the previous section we know $G(a) \in \mathbb{Z}_p[\zeta_p]$. Every element of $\mathbb{Z}_p[\zeta_p]$ can be expanded in powers of $z := \zeta_p - 1$, the standard uniformizer of $\mathbb{Z}_p[\zeta_p]$. Stickelberger's congruence gives the leading term of this expansion for $G(a)$ in terms of the base $p$

digits of $a$.

**Theorem 1.4.4** (**Stickelberger's congruence**). *For every $0 \leq a < q-1$, we have*

$$G(a) \equiv \frac{z^{a_0+a_1+\cdots+a_{f-1}}}{a_0!a_1!\cdots a_{f-1}!} \bmod z^{a_0+a_1+\cdots+a_{f-1}+1},$$

*where $z = \zeta_p - 1$ and $a = a_0 + a_1p + a_2p^2 + \cdots + a_{f-1}p^{f-1}$ with $0 \leq a_i \leq p-1$ for $0 \leq i \leq f-1$.*

*Proof.* See [11, Chapter 1, Theorem 2.1], where the congruence is proved in the number field $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$. Through $p$-adic completion, we get the above result. □

**Example 1.4.5.** Substituting $\zeta_5 = 1 + z$ in Example 1.3.4, we have the following table.

| $a$ | $G(a)$ |
|---|---|
| 0 | $1 \equiv \frac{z^0}{0!} \bmod z$ |
| 1 | $z + 2z^2 + 2z^3 + z^4 + \cdots \equiv \frac{z^1}{1!} \bmod z^2$ |
| 2 | $3z^2 + 2z^3 + 4z^4 + z^6 + \cdots \equiv \frac{z^2}{2!} \bmod z^3$ |
| 3 | $z^3 + z^4 + 3z^5 + 2z^7 + \cdots \equiv \frac{z^3}{3!} \bmod z^4$ |

**Remark 1.4.6.** If we defined the basic $p$-adic Gauss sum $G(a)$ without the minus sign in front of the sum in Definition 1.3.3, then in Stickelberger's congruence we would need an extra minus sign on the right side. More importantly, if we had used $a$ in place of $-a$ as the power of $\omega(x)$ for the multiplicative character in the definition of $G(a)$, then we would have to write $p-1-a_i$ in place of $a_i$ everywhere in Theorem 1.4.4.

**Corollary 1.4.7.** *With notation as in Theorem* 1.4.4*, for every* $0 \leq a < q-1$ *we have*

$$|G(a)|_p = |z|_p^{a_0+a_1+a_2+\cdots+a_{f-1}} = \left(\frac{1}{p}\right)^{\frac{a_0+a_1+a_2+\cdots+a_{f-1}}{p-1}}.$$

*Proof.* The modulus in Theorem 1.4.4 has larger exponent than the power of $z$ on the right side, and $a_0!a_1!\cdots a_{f-1}! \in \mathbb{Z}_p^\times$. Also $|z|_p = |\zeta_p - 1|_p = (1/p)^{1/(p-1)}$. □

We can improve the exponent in the modulus of Stickelberger's congruence if we use another uniformizer of $\mathbb{Z}_p[\zeta_p]$ instead of $z = \zeta_p - 1$. Let $\pi$ be the unique solution of $X^{p-1} = -p$ closest to $z$ in $\mathbb{C}_p$. This is a uniformizer of $\mathbb{Z}_p[\zeta_p]$ by Corollary B.8 where $\pi = \pi_{1,1}$.

**Theorem 1.4.8** (**Stickelberger's congruence**)**.** *For every* $0 \leq a < q-1$*, we have* $G(a) \in \pi^{a_0+a_1+\cdots+a_{f-1}}\mathbb{Z}_p^\times$ *and*

$$G(a) \equiv \frac{\pi^{a_0+a_1+\cdots+a_{f-1}}}{a_0!a_1!\cdots a_{f-1}!} \bmod \pi^{a_0+a_1+\cdots+a_{f-1}+p-1},$$

*where* $a = a_0 + a_1p + a_2p^2 + \cdots + a_{f-1}p^{f-1}$ *with* $0 \leq a_i \leq p-1$ *for* $0 \leq i \leq f-1$.

*Proof.* Setting $l = n = v = 1$, Corollary 3.4.6 implies $G(a) \in \pi^{a_0+a_1+\cdots+a_{f-1}}\mathbb{Z}_p^\times$ and Sections 3.3.2 and 3.3.3 give two proofs of the above congruence. Alternatively, using Galois theory to show $G(a)/\pi^{a_0+a_1+\cdots+a_{f-1}} \in \mathbb{Q}_p$ and replacing $z_l$ by $\pi_l$ in Theorem 1.4.4, we get another proof of Theorem 1.4.8. □

Note the higher power of $\pi$ in the modulus of Theorem 1.4.8 compared to Theorem 1.4.4. It is not generally true that $G(a)$ is a $p$-adic integer multiple of $z^{a_0+a_1+\cdots+a_{f-1}}$ so $G(a) \in \pi^{a_0+a_1+\cdots+a_{f-1}}\mathbb{Z}_p^\times$ is truly a special feature of the uniformizer $\pi$.

In order to compute the $\pi$-expansion of $G(a)$ in Sage, we use a formula for $\zeta_p$ in terms of $\pi$. From [11, Chapter 14, Section 2 and 3], $\zeta_p = \mathrm{AH}_1(\pi)$, where $\mathrm{AH}_1(X) =$

$\exp(X + X^p/p) \in \mathbb{Q}_p[[X]]$ converges for $|x|_p < (1/p)^{1/(p-1)}p^{(p-1)/p^2}$, i.e., $\mathrm{AH}_1(\pi)$ converges since $|\pi|_p = (1/p)^{1/(p-1)} < (1/p)^{1/(p-1)}p^{(p-1)/p^2}$. Hence we may simply write $\zeta_p$ as $\mathrm{AH}_1(\pi)$ in (1.4.1) to get the $\pi$-expansion of $G(a) \in \mathbb{Z}_p[\zeta_p] = \mathbb{Z}_p[\pi]$.

**Example 1.4.9.** Fix $p = q = 5$, $\zeta_5 \in \mathbb{C}_5$ and let $\pi$ be the root of $X^4 = -5$ in $\mathbb{C}_p$ closest to $\zeta_5 - 1$. Hence $\pi^4 = -5$, $f = 1$ and $a = a_0$. Using Sage and substituting $\zeta_5 = \mathrm{AH}_1(\pi)$ in Example 1.3.4, where $\mathrm{AH}_1(X) = \exp(X + X^5/5)$, we get the following table. Compare the exponents in the moduli with those in Example 1.4.5: they are larger.

| $a$ | $G(a)$ |
|---|---|
| 0 | $1 \equiv \frac{\pi^0}{0!} \bmod \pi^4$ |
| 1 | $\pi + \pi^5 + \pi^9 + 3\pi^{17} + \cdots \equiv \frac{\pi^1}{1!} \bmod \pi^5$ |
| 2 | $3\pi^2 + 2\pi^6 + 3\pi^{10} + 2\pi^{14} + \cdots \equiv \frac{\pi^2}{2!} \bmod \pi^6$ |
| 3 | $\pi^3 + 4\pi^7 + \pi^{11} + \pi^{15} + \cdots \equiv \frac{\pi^3}{3!} \bmod \pi^7$ |

Note that $G(a)/\pi^a$ is a power series in $\pi^4 = -5$ and thus in $G(a) \in \pi^a\mathbb{Z}_5^\times$.

### 1.4.4 The Gross–Koblitz Formula

Besides being used to get a higher modulus in Stickelberger's congruence, the uniformizer $\pi$ of $\mathbb{Q}_p(\zeta_p)$ allows us to lift Stickelberger's congruence to an equality, which is the Gross–Koblitz formula. In addition, Theorem 1.4.8 says that $G(a)/\pi^{a_0+a_1+\cdots+a_{f-1}}$ is a $p$-adic integer unit and the Gross–Koblitz formula below explicitly describes this unit as an $f$-product of the $p$-adic Gamma function $\Gamma_p$ evaluated at certain rational numbers depending on $a$.

**Definition 1.4.10.** For any positive integer $n$ the *p-adic Gamma function* is given by

$$\Gamma_p(n+1) = (-1)^{n+1} \prod_{\substack{k=1 \\ p \nmid k}}^{n} k.$$

We extend $\Gamma_p \colon \mathbb{Z}_p \to \mathbb{Z}_p^\times$ by continuity with respect to the $p$-adic absolute value.

**Theorem 1.4.11** (**Gross–Koblitz formula**). *For $0 \leq a < q-1$ we have*

$$G(a) = \pi^{a_0 + a_1 + \cdots + a_{f-1}} \prod_{i=0}^{f-1} \Gamma_p\left(\frac{a^{(i)}}{q-1}\right),$$

*where $a^{(i)}$ is $a$ with its digits in base $p$ cyclically permuted $i$ positions: writing the base $p$ expansion of $a = a_0 + a_1 p + a_2 p^2 + \cdots + a_{f-1} p^{f-1}$ as $[a_0 a_1 \ldots a_{f-1}]_p$, set $a^{(0)} = a = [a_0 a_1 \ldots a_{f-1}]_p$, $a^{(1)} = [a_1 a_2 \ldots a_{f-1} a_0]_p$, $a^{(2)} = [a_2 a_3 \ldots a_{f-1} a_0 a_1]_p$, and so on.*

*Proof.* See [8] for the original proof for $p \neq 2$, [16] for an elementary proof for all primes $p$ or [5] and [11, Chapter 15] for the proof we will partially generalize below. □

**Example 1.4.12.** The Gross–Koblitz formula lifts the congruences in Example 1.4.9 to equality as follows.

| $a$ | $G(a)$ |
|---|---|
| 0 | $\Gamma_5\left(\frac{0}{5-1}\right) = \Gamma_5(0) = 1$ |
| 1 | $\pi\Gamma_5\left(\frac{1}{5-1}\right) = \pi\Gamma_5\left(\frac{1}{4}\right)$ |
| 2 | $\pi^2\Gamma_5\left(\frac{2}{5-1}\right) = \pi^2\Gamma_5\left(\frac{2}{4}\right)$ |
| 3 | $\pi^3\Gamma_5\left(\frac{3}{5-1}\right) = \pi^3\Gamma_5\left(\frac{3}{4}\right)$ |

In general, if $x \equiv y \bmod p\mathbb{Z}_p$ then $\Gamma_p(x) \equiv \Gamma_p(y) \bmod p\mathbb{Z}_p$, so

$$\Gamma_p\left(\frac{a^{(i)}}{q-1}\right) \equiv \Gamma_p\left(-a_i\right) \equiv \frac{1}{a_i!} \bmod p\mathbb{Z}_p,$$

where the second congruence relies on $a_i$ lying between 0 and $p-1$. This explains where the factorials of the digits in Stickelberger's congruence come from. In addition, since $\Gamma_p\left(\mathbb{Z}_p\right) \subseteq \mathbb{Z}_p^\times$, the Gross–Koblitz formula implies $G(a)/\pi^{a_0+a_1+\cdots+a_{f-1}}$ is a unit in $\mathbb{Z}_p$, not just in $\mathbb{Z}_p[\zeta_p]$. This ratio lying in $\mathbb{Z}_p$ is also clear from Galois theory.

## Sketch of a proof of the Gross–Koblitz formula

We present a brief sketch of the proof of the Gross–Koblitz formula from [11, Chapter 15]. It has three main steps.

**Step 1:** Show $G(a) = (1-q)\operatorname{Tr}(\alpha) = \operatorname{Tr}(\overline{\alpha})$ for a completely continuous linear map $\alpha$ on an infinite-dimensional $K$-Banach space that induces a map $\overline{\alpha}$ on a 1-*dimensional* quotient space. Moreover, $\overline{\alpha}$ decomposes as a composition of $f$ $K$-linear maps $\overline{\alpha_i}$, where $q = p^f$.

**Step 2:** Show $G(a) = \prod_{i=0}^{f-1} \mu_i$ for certain $\mu_i \in K$ associated to the maps $\overline{\alpha_i}$.

**Step 3:** Show $\mu_i = \pi^{a_i}\Gamma_p\left(\frac{a^{(i)}}{q-1}\right)$ for all $0 \leq i \leq f-1$.

We will focus on Step 1 and Step 2, which are already interesting: how can the sum $G(a)$ be turned into a trace and then a product?

Fix a finite extension $K$ of $\mathbb{Q}_q(\zeta_p)$. For $r \in |K^\times|_p := \{|x|_p : x \in K^\times\}$, let

$$\mathcal{L}(r) \;=\; \left\{ \sum_{k\geq 0} c_k X^k \in K[[X]] : |c_k|_p r^k \to 0 \right\}.$$

This is a $K$-Banach space with respect to the norm $\left|\left| \sum_{k\geq 0} c_k X^k \right|\right| = \max_{k\geq 0} \left\{ |c_k|_p r^k \right\}$ since $K$ contains an element $c$ such that $|c|_p = 1/r$, which implies that $\mathcal{L}(r)$ has an orthonormal basis $\{1, cX, c^2X^2, \dots, \}$. In particular, for reasons that will be clear in Remark 4.4.8 with $n = l = 1$, we are interested in the spaces $\mathcal{L}(r)$ with

$$1 < r < p^{(p-1)/p} \tag{1.4.4}$$

such that $r \in |K^\times|_p$.

**Remark 1.4.13.** In [11, Chapter 15], Lang defines spaces $L(\delta)$ where $\delta > 0$. For rational $\delta$ such that $K$ has an element of valuation $\delta$, $L(\delta) = \mathcal{L}(r)$ with $r = p^\delta > 1$ in our notation.

Since $\mathcal{L}(r)$ has an orthonormal basis, the trace of the completely continuous linear operator $\alpha$ on $\mathcal{L}(r)$ is well-defined as the sum of the diagonal entries in its matrix representation with respect to any orthonormal basis of $\mathcal{L}(r)$. We denote this trace by $\mathrm{Tr}(\alpha)$.

**Key fact 1**: For $r \in |K^\times|_p$ such that $1 < r < p^{(p-1)/p}$, there is a completely continuous linear operator $\alpha\colon \mathcal{L}(r) \to \mathcal{L}(r)$ such that

$$G(a) = (1 - q)\,\mathrm{Tr}(\alpha). \tag{1.4.5}$$

We may explicitly set $K := \mathbb{Q}_q(\zeta_p)$ and $r = p^{1/(p-1)}$ for $p$ odd; or $K := \mathbb{Q}_q(\zeta_{p^3}) = \mathbb{Q}_q(\zeta_8)$ and $r = p^{1/(p(p-1))} = 2^{1/4}$ for $p = 2$. In either case, inequality (1.4.4) is satisfied and $K$ is the smallest extension of $\mathbb{Q}_q(\zeta_p)$ that contains an element of absolute value $r = p^{1/(p-1)}$ when $p$ is odd and $r = 2^{1/4}$ when $p = 2$.

Having $G(a)$ as the trace on an infinite–dimensional $K$-vector space is not that helpful. To reduce to a finite–dimensional space, we quotient by the image of certain differential operators. Fix $0 \leq a < q-1$. For $i \geq 0$, define the differential operators

$$D_i = X\frac{d}{dX} + \frac{a^{(i)}}{q-1} + \pi X, \tag{1.4.6}$$

where the constants $a^{(i)}/(q-1)$ in the operators $D_i$ appear in the Gross–Koblitz formula. Note that $D_i$ are $f$-periodic: $D_i = D_{i+f}$ for all $i \geq 0$. While every $D_i$ as an operator $K[[X]] \to K[[X]]$ is a bijection for $a \neq 0$, as an operator $\mathcal{L}(r) \to \mathcal{L}(r)$ it is injective but not surjective for any $a$ since by [11, Lemma 13, p. 335] the unique solution to the differential equation $D_i(y) = 1$ in $K[[X]]$ does not lie in $\mathcal{L}(r)$ for any $r > 1$.

**Key fact 2**: For all $r > 1$ in $|K^\times|_p$ and $i \geq 0$, $\operatorname{coker}(D_i) = \mathcal{L}(r)/D_i\mathcal{L}(r)$ is 1-dimensional over $K$ and we can take $\{\bar{1}\}$ to be a basis. Moreover we have the following commutative relation

$$\alpha \circ D_0 = D_0 \circ q\alpha.$$

By this relation, there is an induced linear map $\overline{\alpha}\colon \mathcal{L}(r)/D_0\mathcal{L}(r) \to \mathcal{L}(r)/D_0\mathcal{L}(r)$ and we have commutativity of the diagram below.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{L}(r) & \xrightarrow{D_0} & \mathcal{L}(r) & \longrightarrow & \mathcal{L}(r)/D_0\mathcal{L}(r) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle q\alpha} & & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \overline{\alpha}} & & \\ 0 & \longrightarrow & \mathcal{L}(r) & \xrightarrow{D_0} & \mathcal{L}(r) & \longrightarrow & \mathcal{L}(r)/D_0\mathcal{L}(r) & \longrightarrow & 0 \end{array}$$

By [11, p. 358] this implies $\mathrm{Tr}(q\alpha) - \mathrm{Tr}(\alpha) + \mathrm{Tr}(\overline{\alpha}) = 0$, so

$$(q-1)\,\mathrm{Tr}(\alpha) + \mathrm{Tr}(\overline{\alpha}) = 0.$$

Therefore by Key fact 1 the Gauss sum $G(a)$ is a trace:

$$G(a) = \mathrm{Tr}(\overline{\alpha}). \tag{1.4.7}$$

By Key fact 2, $\overline{\alpha}$ is an operator on a 1-dimensional $K$-vector space $\mathcal{L}(r)/D_0\mathcal{L}(r)$ so equation (1.4.7) can be rewritten as

$$\overline{\alpha}(\bar{1}) = G(a) \cdot \bar{1}, \tag{1.4.8}$$

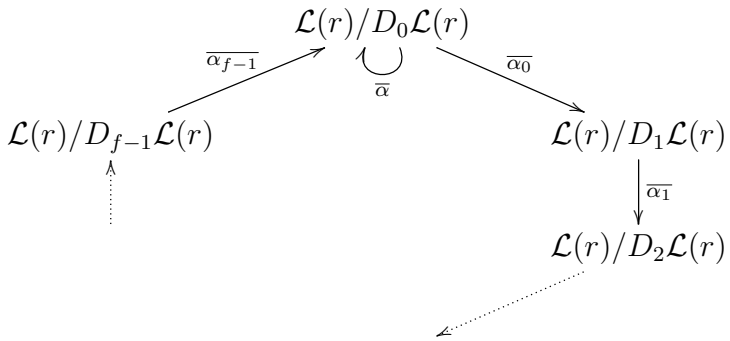
showing $G(a)$ is an eigenvalue of $\overline{\alpha}$. We will now factor the operator $\overline{\alpha}$ in order to finish Step 1.

**Key fact 3**: For each $i \geq 0$ there is a completely continuous linear operator $\alpha_i \colon \mathcal{L}(r) \to \mathcal{L}(r)$ such that

$$\alpha = \alpha_{f-1} \circ \alpha_{f-2} \circ \cdots \circ \alpha_0 \text{ and } \alpha_i \circ D_i = D_{i+1} \circ p\alpha_i.$$

Thus we get well-defined linear maps $\overline{\alpha_i} \colon \mathcal{L}(r)/D_i\mathcal{L}(r) \to \mathcal{L}(r)/D_{i+1}\mathcal{L}(r)$ between

1-dimensional $K$-vector spaces such that $\overline{\alpha} = \overline{\alpha_{f-1}} \circ \cdots \circ \overline{\alpha_1} \circ \overline{\alpha_0}$, or equivalently the diagram below commutes.




This completes Step 1. Since the quotient spaces in the diagram above each have $\overline{1}$ as a basis, for each $i \geq 0$ there are numbers $\mu_i$ in $K$ defined by

$$\overline{\alpha_i}(\overline{1}) = \mu_i \cdot \overline{1}, \tag{1.4.9}$$

which are not eigenvalues of $\overline{\alpha_i}$ since $\overline{1}$ lies in different spaces on each side.

By applying (1.4.9) repeatedly we obtain

$$\overline{\alpha}(\overline{1}) = \overline{\alpha_{f-1}} \circ \overline{\alpha_{f-2}} \circ \cdots \circ \overline{\alpha_0}(\overline{1}) = \prod_{i=0}^{f-1} \mu_i \cdot \overline{1}$$

where $\overline{1}$ on both sides is in $\mathcal{L}(r)/D_0\mathcal{L}(r)$. This shows by (1.4.8) that

$$G(a) = \prod_{i=0}^{f-1} \mu_i.$$

This completes Step 2 in the proof of the Gross–Koblitz formula.

We collect the key facts used above in the following theorem, whose generalization

in Chapter 2 is one of our main results.

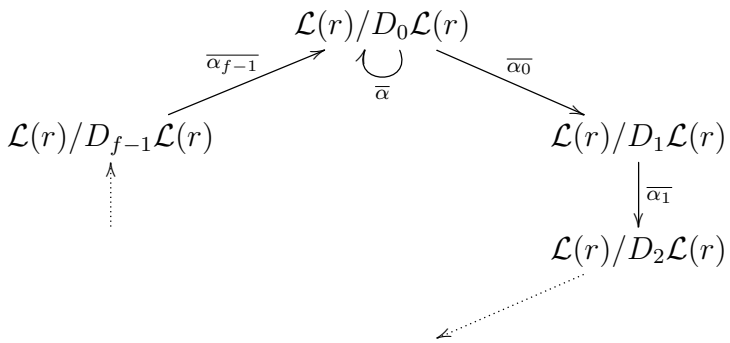
**Theorem 1.4.14** (Dwork-Lang)**.** *Fix* $0 \leq a < q-1$ *and* $\zeta_p \in \mathbb{C}_p$*. For* $i \geq 0$ *and* $r \in |K^\times|_p$ *such that* $1 < r \leq p^{(p-1)/p}$*, there are differential operators* $D_i \colon \mathcal{L}(r) \to \mathcal{L}(r)$ *defined by* (1.4.6) *and completely continuous* $K$*-linear maps* $\alpha \colon \mathcal{L}(r) \to \mathcal{L}(r)$ *and* $\alpha_i \colon \mathcal{L}(r) \to \mathcal{L}(r)$ *such that* (1) *each quotient space* $\mathcal{L}(r)/D_i\mathcal{L}(r)$ *is* 1*-dimensional over* $K$ *and* (2) *the induced maps* $\overline{\alpha} \colon \mathcal{L}(r)/D_0\mathcal{L}(r) \to \mathcal{L}(r)/D_0\mathcal{L}(r)$ *and* $\overline{\alpha_i} \colon \mathcal{L}(r)/D_i\mathcal{L}(r) \to \mathcal{L}(r)/D_{i+1}\mathcal{L}(r)$ *satisfy*

$$G(a) = (1-q)\operatorname{Tr}(\alpha) = \operatorname{Tr}(\overline{\alpha})$$

*and*

$$\overline{\alpha} = \overline{\alpha_{f-1}} \circ \overline{\alpha_{f-2}} \circ \cdots \circ \overline{\alpha_1} \circ \overline{\alpha_0},$$

*i.e., the following diagram commutes.*




*Proof.* See [11, Chapter 15] or Chapter 4 using $n = l = 1$. □

The $p$-adic analytic representation $\zeta_p = \mathrm{AH}_1(\pi)$ where $\mathrm{AH}_1(X) = \exp(X + X^p/p)$ is the first of a family of such formulas. For $n \geq 1$ or $n = \infty$, there is a unique root of

$$L_n(X) = X + X^p/p + X^{p^2}/p^2 + \cdots + X^{p^n}/p^n$$

in $\mathbb{C}_p$ with absolute value $(1/p)^{1/(p-1)}$ that is closest to $\zeta_p - 1$. Denote this root by $\pi_n$ (not to be confused with $\pi_l$ in Appendix A and other Chapters). In this notation, the uniformizer $\pi$ from before is $\pi_1$. It turns out that $\pi_n$ is also a uniformizer of $\mathbb{Z}_p[\zeta_p] = \mathbb{Z}_p[\pi_n]$ and by Theorem B.7 we have $\zeta_p = \mathrm{AH}_n(\pi_n)$, where $\mathrm{AH}_n(X) = \exp(L_n(X)) \in \mathbb{Q}_p[[X]]$ (in Appendix B $\pi_n$ is denoted by $\pi_{n,1}$). Baldassarri [3] used the uniformizers $\pi_n$ and defined higher $p$-adic Gamma functions $\Gamma_{n,p}(x)$ associated to $\pi_n$ to prove the following generalization of the Gross–Koblitz formula. We omit the explicit definition of $\Gamma_{n,p}(x)$ in [3] since it requires too much setup and is unnecessary for our discussion here.

**Theorem 1.4.15** (Baldassarri)**.** *With notation as in the Gross–Koblitz formula* (*Theorem* 1.4.11)*, for* $n \geq 1$ *and* $0 \leq a < q-1$ *we have*

$$G(a) = \pi_n^{a_0+a_1+\cdots+a_{f-1}} \prod_{i=0}^{f-1} \Gamma_{n,p}\left(\frac{a^{(i)}}{q-1}\right).$$

*Proof.* See [3]. □

Baldassarri showed this for $p \neq 2$, but with a little more care his argument works for $p = 2$ as well. It is also worth noting that we may replace $\pi$ by $\pi_n$ everywhere in Theorem 1.4.8, which would lift to an equality given by Theorem 1.4.15.

# Chapter 2

# Generalized Gauss sums

## 2.1 Definition over the complex numbers

Set $\mathbb{F}_q = \mathbb{Z}_q/(p)$, which is a canonical residue field of order $q$. We will generalize the Gauss sums $G(\chi, \psi)$ by using continuous additive characters $\psi\colon \mathbb{Z}_q \to \mathbb{C}^\times$. The key point is that the domain of $\psi$ is not $\mathbb{F}_q$ as before, so the role of $\zeta_p \in \mathbb{C}$ in the additive character of Gauss sums can be replaced by roots of unity of $p$-power order in $\mathbb{C}$. Recall from the previous chapter that the basic Gauss sum associated to a multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ (and a choice of $\zeta_p$) is

$$G(\chi) = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)} = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{x+x^p+x^{p^2}+\cdots+x^{p^{f-1}}}.$$

If we want to replace $\zeta_p$ by higher $p$-power order roots of unity, the exponent $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)$ has to be modified since its values only make sense mod $p$. We will use a $p$-adic trace. Let Tr be the trace map $\mathrm{Tr}_{\mathbb{Q}_q/\mathbb{Q}_p}\colon \mathbb{Q}_q \to \mathbb{Q}_p$ and recall $\omega$, the Teichmüller character

(Definition 1.3.1). Since $\mathrm{Gal}(\mathbb{Q}_q/\mathbb{Q}_p) \cong \mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ by reduction mod $p$, we have $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x) \equiv \mathrm{Tr}(\omega(x)) \bmod p$ for $x \in \mathbb{F}_q$. Thus

$$G(\chi) = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{\mathrm{Tr}(\omega(x))} = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_p^{\omega(x)+\omega(x)^p+\omega(x)^{p^2}+\cdots+\omega(x)^{p^{f-1}}}.$$

Let's see how this change in the trace map that we use modifies Example 1.1.4.

**Example 2.1.1.** With notation as in Example 1.1.4, we have the following table.

| $\chi$ | $G(\chi)$ |
|---|---|
| $\mathbb{1}$ | $-\zeta_5-\zeta_5^2-\zeta_5^3-\zeta_5^4 = -\zeta_5^{\omega(1)}-\zeta_5^{\omega(2)}-\zeta_5^{\omega(3)}-\zeta_5^{\omega(4)}$ |
| $\eta$ | $-\zeta_5-i\zeta_5^2+i\zeta_5^3+\zeta_5^4 = -\zeta_5^{\omega(1)}-i\zeta_5^{\omega(2)}+i\zeta_5^{\omega(3)}+\zeta_5^{\omega(4)}$ |
| $\eta^2$ | $-\zeta_5+\zeta_5^2+\zeta_5^3-\zeta_5^4 = -\zeta_5^{\omega(1)}+\zeta_5^{\omega(2)}+\zeta_5^{\omega(3)}-\zeta_5^{\omega(4)}$ |
| $\eta^3$ | $-\zeta_5+i\zeta_5^2-i\zeta_5^3+\zeta_5^4 = -\zeta_5^{\omega(1)}+i\zeta_5^{\omega(2)}-i\zeta_5^{\omega(3)}+\zeta_5^{\omega(4)}$ |

Since the exponent on $\zeta_p$ is now in $\mathbb{Z}_p$, rather than $\mathbb{F}_p$, we can replace $\zeta_p$ with roots of unity of higher $p$-power order. For $l \geq 1$, $\zeta_{p^l} \in \mathbb{C}$ and a multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$, set

$$G_l(\chi) = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_{p^l}^{\mathrm{Tr}(\omega(x))} = -\sum_{x\in\mathbb{F}_q^\times} \chi(x)\zeta_{p^l}^{\omega(x)+\omega(x)^p+\omega(x)^{p^2}+\cdots+\omega(x)^{p^{f-1}}}.$$

When $l = 1$ the sum $G_1(\chi)$ is the same as the basic Gauss sum $G(\chi)$. In the sum $G_l(\chi)$, the factor $\zeta_{p^l}^{\mathrm{Tr}(\omega(x))}$ is not an additive character $\mathbb{F}_q \to \mathbb{C}^\times$. However, substituting $t = \omega(x)$, the function $t \mapsto \zeta_{p^l}^{\mathrm{Tr}(t)}$ is a continuous additive character $\mathbb{Z}_q \to \mathbb{C}^\times$ of order $p^l$. Since $\omega\colon \mathbb{F}_q^\times \to \mu_{q-1}$ is an isomorphism with inverse the reduction mod $p$ map, substituting $t = \omega(x)$ in $G_l(\chi)$ results in replacing summation over $x \in \mathbb{F}_q^\times$ by

summation over $t \in \mu_{q-1} \subset \mathbb{Z}_q$. Hence

$$G_l(\chi) = -\sum_{t \in \mu_{q-1}} \chi(\bar{t})\zeta_{p^l}^{\mathrm{Tr}(t)} = -\sum_{t \in \mu_{q-1}} \chi(\bar{t})\zeta_{p^l}^{t+t^p+\cdots+t^{p^{f-1}}},$$

where the sum is over $p$-adic roots of unity but the values in the sum are complex. It is now natural to replace the domain of the multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ by $\mu_{q-1}$ and record this as the definition of $G_l(\chi)$.

**Definition 2.1.2.** Fix an integer $l \geq 1$ and a root of unity $\zeta_{p^l} \in \mathbb{C}$ of order $p^l$. For a multiplicative character $\chi\colon \mu_{q-1} \to \mathbb{C}^\times$, the *generalized basic Gauss sum* of level $l$ associated to $\chi$ is

$$G_l(\chi) = -\sum_{t \in \mu_{q-1}} \chi(t)\zeta_{p^l}^{\mathrm{Tr}(t)} = -\sum_{t \in \mu_{q-1}} \chi(t)\zeta_{p^l}^{t+t^p+\cdots+t^{p^{f-1}}}.$$

Every continuous additive character $\mathbb{Z}_q \to \mathbb{C}^\times$ has finite $p$-power order since the only subgroup arbitrarily close to 1 in $\mathbb{C}^\times$ is $\{1\}$. The continuous additive characters $\mathbb{Z}_q \to \mathbb{C}^\times$ of order dividing $p^l$ are in one-to-one correspondence with the additive characters $\mathbb{Z}_q/p^l \to \mathbb{C}^\times$. By a counting argument they are of the form $\psi_{l,v}(t) = \zeta_{p^l}^{\mathrm{Tr}(tv)}$ for some $v \in \mathbb{Z}_q$ that is uniquely determined in $\mathbb{Z}_q/p^l$, and $\psi_{l,v}$ has order exactly $p^l$ if and only if $v \in \mathbb{Z}_q^\times$. When $v = 1$, we define the *basic continuous additive character* $\psi_l(t) = \psi_{l,1}(t) = \zeta_{p^l}^{\mathrm{Tr}(t)}$.

**Definition 2.1.3.** Fix an integer $l \geq 1$ and a root of unity $\zeta_{p^l} \in \mathbb{C}$ of order $p^l$. For $v \in \mathbb{Z}_q^\times$ and $\chi\colon \mu_{q-1} \to \mathbb{C}^\times$ a multiplicative character, the *generalized Gauss sum* of

level $l$ associated to $\chi$ and $v$ is

$$G_{l,v}(\chi) = -\sum_{t \in \mu_{q-1}} \chi(t)\zeta_{p^l}^{\mathrm{Tr}(tv)}.$$

This definition includes the previous Gauss sums over $\mathbb{C}$: $G_l(\chi) = G_{l,1}(\chi)$ and $G(\chi) = G_{1,1}(\chi)$.

**Remark 2.1.4.** For a multiplicative character $\chi\colon \mu_{q-1} \to \mathbb{C}^\times$ and a continuous additive character $\psi\colon \mathbb{Z}_q \to \mathbb{C}^\times$, the Gauss sum

$$G(\chi, \psi) = -\sum_{t \in \mu_{q-1}} \chi(t)\psi(t)$$

generalizes of Definition 1.1.1: $G_l(\chi) = G(\chi, \psi_l)$ and $G_{l,v}(\chi) = G(\chi, \psi_{l,v})$.

When $v \in \mu_{q-1}$, using the change of variables $t \mapsto t/v$ as in the proof of Lemma 1.1.3 we get

$$G_{l,v}(\chi) = -\sum_{t \in \mu_{q-1}} \chi(t/v)\zeta_{p^l}^{\mathrm{Tr}(tv)} = \overline{\chi(v)}G_l(\chi).$$

Thus we may restrict to studying $G_{l,v}(\chi)$ with $v \in 1 + p\mathbb{Z}_q$. Note that for general $v \in \mathbb{Z}_q^\times$ this change of variables is invalid since $t/v$ may not be in $\mu_{q-1}$ for $t \in \mu_{q-1}$. Hence, unlike in the case of Gauss sums $G(\chi)$, generalized Gauss sums $G_{l,v}(\chi)$ as $v$ runs over $\mathbb{Z}_q^\times$ do not reduce to a simple multiple of $G_l(\chi)$. See Theorem 3.1.1 for a decomposition of the $p$-adic analogue of $G_{l,v}(\chi)$ in terms of the $p$-adic analogue of $G_k(\chi)$ for $1 \leq k \leq l$. Since all the quantities involved are algebraic, the complex analogue of Theorem 3.1.1 holds as well.

## 2.2 Properties of generalized Gauss sums over $\mathbb{C}$

In this section we use $L$-functions to show that a generalization of the Hasse-Davenport relation (Theorem 1.2.2) to $G_l(\chi)$ fails and also note through examples that $|G_l(\chi)| \neq \sqrt{q}$ when $l > 1$. Recall from Chapter 1 that the basic Gauss sums $G(\chi) = G_1(\chi)$ satisfy both of these properties.

Generalizing $L(T)$ as in equation 1.2.2 we have the following definition.

**Definition 2.2.1.** Fix an integer $l \geq 1$ and a choice of multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$. The $L$-function of level $l$ associated $\chi$ is

$$L_l(T) = \exp\left(\sum_{n=1}^{\infty} -G_l(\chi^{(n)})\frac{T^n}{n}\right) = 1 - G_l(\chi)T + \cdots .$$

Adolphson, Sperber and Liu showed that $L_l(T)$ is in fact a polynomial of degree $p^{l-1}$ (see [1, 2, 12, 13]). Hence for $l > 1$, if a generalized Hasse-Davenport relation were true, i.e. $G_l(\chi^{(n)}) = G_l(\chi)^n$ for all $n \geq 1$, then by the same argument as in Section 1.2.2, we get that $L_l(T)$ is a polynomial of degree 1. This would contradict the fact that $L_l(T)$ is a polynomial of degree $p^{l-1} > 1$ when $l > 1$.

By [17, Theorem 3], $L_l(T)$ is a polynomial of degree at most $p^{l-1}$ and has nonzero reciprocal roots in $\mathbb{C}$ of absolute value $\sqrt{q}$. Thus, since $G_l(\chi)$ is the linear coefficient of $L_l(T)$ up to a sign and $L_l(T)$ has constant term 1, we have

$$|G_l(\chi)| = \left|\sum_{\rho \text{ root of } L_l(T)} \frac{1}{\rho}\right| \leq p^{l-1}\sqrt{q},$$

where the sum is over all the complex roots $\rho$ of $L_l(T)$. When $l = 1$ and $\chi$ is nontrivial, the above inequality turns into an equality, i.e., $|G_l(\chi)| = \sqrt{q}$ (Theorem

1.2.1). For $l > 1$ generally we have $|G_l(\chi)| \neq \sqrt{q}$, which can be seen computationally: for $q = 3, 5$, $l = 2$ and all $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$, we have $|G_2(\chi)| \neq \sqrt{q}$.

## 2.3 Definition over the $p$-adic numbers

To define the generalized $p$-adic Gauss sums, set $\mathbb{F}_q = \mathbb{Z}_q/(p)$ as before. The $p$-adic Gauss sums will use multiplicative and additive characters taking values in $\mathbb{C}_p$. Recall the basic $p$-adic Gauss sum definition

$$G(a) = -\sum_{x \in \mathbb{F}_q^\times} \omega(x)^{-a} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)}.$$

Just like for its complex version, we replace $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)$ by the $p$-adic trace $\mathrm{Tr}(\omega(x))$ first and then substitute $t = \omega(x)$. This gives

$$G(a) = -\sum_{x \in \mathbb{F}_q^\times} \omega(x)^{-a} \zeta_p^{\mathrm{Tr}(\omega(x))} = -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_p^{\mathrm{Tr}(t)}.$$

Since $\mathrm{Tr}(t) \in \mathbb{Z}_p$, we may now replace $\zeta_p$ by any $p$-power order root of unity in $\mathbb{C}_p$.

**Definition 2.3.1.** Fix an integer $l \geq 1$ and $\zeta_{p^l} \in \mathbb{C}_p$. For $0 \leq a < q-1$, the *generalized basic p-adic Gauss sum* of level $l$ associated to $a$ is

$$G_l(a) = -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{\mathrm{Tr}(t)} = -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{t + t^p + t^{p^2} + \cdots + t^{p^{f-1}}}.$$

Since the multiplicative group $\mu_{q-1} = \{t \in \mathbb{Z}_q^\times : t^{q-1} = 1\}$ is isomorphic to $\mathbb{F}_q^\times$ by reduction mod $p$, with inverse the Teichmüller map $\omega$, when $l = 1$, the generalized basic $p$-adic Gauss sum $G_1(a)$ is the same as the basic $p$-adic Gauss sum $G(a)$.

**Example 2.3.2.** For $p = q = 5$, $l = 2$ and $i := \omega(2) = 2 + 5 + 2 \cdot 5^2 + \cdots \in \mathbb{Z}_5$, we have the following table.

| $a$ | $G_2(a)$ |
|---|---|
| 0 | $-\zeta_{25} - \zeta_{25}^{i} - \zeta_{25}^{-i} - \zeta_{25}^{-1} = -\zeta_{25} - \zeta_{25}^{7} - \zeta_{25}^{18} - \zeta_{25}^{24}$ |
| 1 | $-\zeta_{25} + i\zeta_{25}^{i} - i\zeta_{25}^{-i} + \zeta_{25}^{-1} = -\zeta_{25} + i\zeta_{25}^{7} - i\zeta_{25}^{18} + \zeta_{25}^{24}$ |
| 2 | $-\zeta_{25} + \zeta_{25}^{i} + \zeta_{25}^{-i} - \zeta_{25}^{-1} = -\zeta_{25} + \zeta_{25}^{7} + \zeta_{25}^{18} - \zeta_{25}^{24}$ |
| 3 | $-\zeta_{25} - i\zeta_{25}^{i} + i\zeta_{25}^{-i} + \zeta_{25}^{-1} = -\zeta_{25} - i\zeta_{25}^{7} + i\zeta_{25}^{18} + \zeta_{25}^{24}$ |

In the definition of $G_l(a)$, the factor $t^{-a}$ is a $p$-adic multiplicative character $\mu_{q-1} \to \mathbb{C}_p^{\times}$ and $\zeta_{p^l}^{\mathrm{Tr}(t)}$ is a continuous $p$-adic additive character $\mathbb{Z}_q \to \mathbb{C}_p^{\times}$ of order $p^l$ restricted to $\mu_{q-1}$.

**Definition 2.3.3.** For a $p$-adic multiplicative character $\chi\colon \mu_{q-1} \to \mathbb{C}_p^{\times}$ and a continuous $p$-adic additive character $\psi\colon \mathbb{Z}_q \to \mathbb{C}_p^{\times}$, the *generalized p-adic Gauss sum* associated to $\chi$ and $\psi$ is

$$G(\chi, \psi) = -\sum_{t \in \mu_{q-1}} \chi(t)\psi(t).$$

Every $p$-adic multiplicative character $\mu_{q-1} \to \mathbb{C}_p^{\times}$ is of the form $\chi_a(t) = t^{-a}$ for a unique integer $0 \leq a < q - 1$ and every finite order continuous $p$-adic additive character $\mathbb{Z}_q \to \mathbb{C}_p^{\times}$ has $p$-power order by continuity. Unlike the continuous additive characters of $\mathbb{Z}_q$ valued in $\mathbb{C}^{\times}$, continuous $p$-adic additive character $\mathbb{Z}_q \to \mathbb{C}_p^{\times}$ do not necessarily have finite order since there is infinitely many subgroups of $\mathbb{C}_p^{\times}$ arbitrarily close to 1. The continuous $p$-adic additive character $\mathbb{Z}_q \to \mathbb{C}_p^{\times}$ of order dividing $p^l$ are in one-to-one correspondence with the additive characters $\mathbb{Z}_q/p^l \to \mathbb{C}_p^{\times}$ and by a counting argument they are of the form $\psi_{l,v}(t) = \zeta_{p^l}^{\mathrm{Tr}(vt)}$ for some $v \in \mathbb{Z}_q$ that is

uniquely determined in $\mathbb{Z}_q/p^l$. Additionally, $\psi_{l,v}$ has order exactly $p^l$ if and only if $v \in \mathbb{Z}_q^\times$. When $v = 1$, we define the *basic p-adic additive character* $\psi_l = \psi_{l,1}$ and note that $\psi_l(t) = \zeta_{p^l}^{\mathrm{Tr}(t)}$ and $G_l(a) = G(\chi_a, \psi_l)$.

**Definition 2.3.4.** Fix an integer $l \geq 1$ and $\zeta_{p^l} \in \mathbb{C}_p$. For $0 \leq a < q-1$, the *generalized p-adic Gauss sum* of level $l$ associated to $a$ and $v$ is

$$G_{l,v}(a) = -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{\mathrm{Tr}(tv)}.$$

This definition includes the previous Gauss sums over $\mathbb{C}_p$: $G_l(a) = G_{l,1}(a)$ and $G(a) = G_{1,1}(a)$. When $v \in \mu_{q-1}$, by the change of variables $t \mapsto t/v$ as in the proof of Lemma 1.1.3 we get

$$G_{l,v}(a) = -\sum_{t \in \mu_{q-1}} (t/v)^{-a} \zeta_{p^l}^{\mathrm{Tr}(t)} = v^a G_l(a), \tag{2.3.1}$$

where $v^a = \overline{\chi_a(v)}$. However, for general $v \in \mathbb{Z}_q^\times$, this change of variables is invalid since $t/v$ may not be in $\mu_{q-1}$ for $t \in \mu_{q-1}$. Hence, like for the complex-valued generalized Gauss sums $G_{l,v}(\chi)$, the generalized $p$-adic Gauss sums $G_{l,v}(a)$ do not reduce to a simple multiple of $G_l(a)$. If $v \in \mathbb{Z}_q^\times$ and we set $v_0 = \omega(v)$, so $v_0 \in \mu_{q-1}$ and $v \equiv v_0 \bmod p\mathbb{Z}_q$, then by a similar argument we get

$$G_{l,v}(a) = v_0^a G_{l,v/v_0}(a). \tag{2.3.2}$$

Thus it suffices to understand the Gauss sums $G_{l,v}(a)$ for $v \in 1 + p\mathbb{Z}_q$.

## 2.4 Properties of generalized $p$-adic Gauss sums

In this section, we narrow down the degree of generalized $p$-adic Gauss sums over $\mathbb{Q}_p$, state the analogue of Stickelberger's congruence for $G_l(a)$ (proved by Blache [4]) and $G_{l,v}(a)$ (one of our main results), and use it to find $|G_{l,v}(a)|_p$. Then we state a partial analogue of the Gross–Koblitz formula, which will be gradually proved throughout Chapter 4.

### 2.4.1 Degree over $\mathbb{Q}_p$

From Definition 2.3.1, we have $G_l(a) \in \mathbb{Z}_p[\zeta_{q-1}, \zeta_{p^l}] = \mathbb{Z}_q[\zeta_{p^l}]$.

**Theorem 2.4.1.** *The Gauss sums $G_l(a)$ lie in $\mathbb{Z}_p[\zeta_{p^l}]$ for all $0 \leq a < q-1$. More precisely, $\mathbb{Q}_p(G_l(a))/\mathbb{Q}_p$ is an extension of degree dividing $p^{l-1}(p-1)/(a, p-1)$ that lies between $\mathbb{Q}_p(\zeta_{p^l})$ and $\mathbb{Q}_p$.*

*Proof.* We follow an analogous approach to the proof of Theorem 1.4.1. Recall that $q = p^f$, $\mathbb{Q}_q(\zeta_{p^l}) = \mathbb{Q}_p(\zeta_{q-1}, \zeta_{p^l})$ has degree $fp^{l-1}(p-1)$ over $\mathbb{Q}_p$, and the Galois group $\mathrm{Gal}(\mathbb{Q}_q(\zeta_{p^l})/\mathbb{Q}_p) \cong \mathbb{Z}/f\mathbb{Z} \times \left(\mathbb{Z}_p/p^l\right)^\times \cong \mathbb{Z}/f\mathbb{Z} \times \mathbb{Z}_p^\times/(1+p^l\mathbb{Z}_p)$ is generated by the automorphisms $\phi_p, \sigma_c \colon \mathbb{Q}_q(\zeta_{p^l}) \to \mathbb{Q}_q(\zeta_{p^l})$ for $c \in \mathbb{Z}_p^\times$ that matters only modulo $p^l$, which are determined respectively by $\phi_p(\zeta_{q-1}) = \zeta_{q-1}^p$, $\phi_p(\zeta_{p^l}) = \zeta_{p^l}$ and $\sigma_c(\zeta_{p^l}) = \zeta_{p^l}^c$, $\sigma_c(\zeta_{q-1}) = \zeta_{q-1}$. The automorphism $\phi_p$ is the Frobenius automorphism as in the proof of Theorem 1.4.1. Applying $\phi_p$ to $G(a)$ and using the facts: (1) $\phi_p(t) = t^p$ for any $t \in \mu_{q-1}$, (2) $\mathrm{Tr}(t) = \mathrm{Tr}(t^p)$, (3) $\phi_p \colon \mu_{q-1} \to \mu_{q-1}$ is a group isomorphism, we get

$$\phi_p(G_l(a)) \overset{(1)}{=} -\sum_{t \in \mu_{q-1}} (t^p)^{-a} \zeta_{p^l}^{\mathrm{Tr}(t)} \overset{(2)}{=} -\sum_{t \in \mu_{q-1}} (t^p)^{-a} \zeta_{p^l}^{\mathrm{Tr}(t^p)} \overset{(3)}{=} G_l(a).$$

This implies that $G_l(a)$ lies in $\mathbb{Q}_p(\zeta_{p^l})$, the fixed field of $\mathrm{Gal}(\mathbb{Q}_q(\zeta_{p^l})/\mathbb{Q}_p(\zeta_{p^l})) = \langle \phi_p \rangle$. Hence $G(a)$ lies in $\mathbb{Z}_p[\zeta_{p^l}]$, as claimed. Thus $\mathbb{Q}_p(G(a))/\mathbb{Q}_p$ is totally ramified, since $\mathbb{Q}_p(\zeta_{p^l})/\mathbb{Q}_p$ is.

Since $\mathbb{Z}_p^\times \cong \mu_{p-1} \times (1 + p\mathbb{Z}_p)$, we get $\left(\mathbb{Z}_p/p^l\right)^\times \cong \mu_{p-1} \times (1 + p\mathbb{Z}_p) \,/\, \left(1 + p^l\mathbb{Z}_p\right)$, where for $l \geq 2$, $(1 + p\mathbb{Z}_p) \,/\, \left(1 + p^l\mathbb{Z}_p\right) \cong \mathbb{Z}_p/p^{l-1}$ when $p$ is odd and $(1 + 2\mathbb{Z}_2)/(1 + 2^l\mathbb{Z}_2) \cong \mathbb{Z}_2/2 \times \mathbb{Z}_2/2^{l-2}$ when $p = 2$. For $c \in \mathbb{Z}_p^\times$, we have

$$\sigma_c(G_l(a)) = -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{\mathrm{Tr}(tc)} = G_{l,c}(a).$$

We would like to know for what $c \in \mathbb{Z}_p^\times$, we have $G_{l,c}(a) = G_l(a)$. Following the same steps as in (2.3.1), for $c \in \mu_{p-1} \subset \mathbb{Z}_p^\times$ we have

$$G_{l,c}(a) = c^a G_l(a).$$

From Corollary 2.4.7 (independent of this result), we get $|G_l(a)|_p \neq 0$ and so $G_l(a) \neq 0$. (Note that in the proof of Theorem 1.4.1 we used the Archimedean absolute value to get $G(a) \neq 0$, but in the generalized case, we only know the upper bound $|G_l(\chi)| \leq p^{l-1}\sqrt{q}$ and so we are forced to use the $p$-adic result.) Hence $\sigma_c(G_l(a)) = G_l(a)$ for $c \in \mu_{p-1}$ if and only if $c^a = 1$. The set $\{c \in \mu_{p-1} : c^a = 1\}$ is a subgroup of order $(a, p-1)$ in the cyclic group $\mu_{p-1}$. Galois theory implies

$$[\mathbb{Q}_p(G(a)) : \mathbb{Q}_p] \text{ divides } \frac{p^{l-1}(p-1)}{|\{c \in \mu_{p-1} : c^a = 1\}|} = \frac{p^{l-1}(p-1)}{(a, p-1)}.$$

Hence, $\mathbb{Q}_p(G(a))$ is an extension of degree dividing $p^{l-1}(p-1)/(a, p-1)$ over $\mathbb{Q}_p$ that lies between $\mathbb{Q}_p(\zeta_{p^l})$ and $\mathbb{Q}_p$. $\square$

Computations suggest more is true: for odd $p$, $a \neq 0$ and $c \in \mathbb{Z}_q^\times$ (that matter only modulo $p^l$), we believe $G_{l,c}(a) = G_l(a)$ if and only if $c^{(a,p-1)} = 1$.

**Conjecture 2.4.2.** *For odd $p$ and $a \neq 0$, $\mathbb{Q}_p(G_l(a))/\mathbb{Q}_p$ is the unique extension of degree $p^{l-1}(p-1)/(a,p-1)$ that lies between $\mathbb{Q}_p(\zeta_{p^l})$ and $\mathbb{Q}_p$.*

**Remark 2.4.3.** For $v \in \mathbb{Z}_p^\times \subseteq \mathbb{Z}_q^\times$, we may replace $G_l(a)$ by $G_{l,v}(a)$ in Theorem 2.4.1, Conjecture 2.4.2 and any later formula regarding $G_l(a)$ since that amounts to replacing $\zeta_{p^l}$ in $G_l(a)$ by another root of unity $\zeta_{p^l}^v$ of order $p^l$. However, for $v \in \mathbb{Z}_q^\times - \mathbb{Z}_p^\times$, $G_{l,v}(a)$ does not generally lie in $\mathbb{Z}_p[\zeta_{p^l}]$.

### 2.4.2 Computing generalized $p$-adic Gauss sums

We work parallel to Section 1.4.2 in order to give a formula for $G_l(a)$ as an element of $\mathbb{Z}_p[\zeta_{p^l}]$ instead of as an element of the composite extension $\mathbb{Z}_q[\zeta_{p^l}]$ of $\mathbb{Z}_p$, which is how it arises from its definition. To this end, we borrow the notation of Section 1.4.2. Since $G_l(a)$ is written as a sum over $t \in \mu_{q-1} \subset \mathbb{Z}_q$, an important difference is that now we work entirely over the $p$-adics rather than over both $p$-adic and finite fields. Thus, we will not need the lifting $\omega$, which simplifies the notation in the following theorem. Recall that for any polynomial $g(X)$, we denote by $d(g)$ its degree and by $s(g)$ the coefficient of $X^{d(g)-1}$ (if $d(g) = 0$ then we define $s(g) = 0$).

**Theorem 2.4.4.** *Fix $0 \leq a < q-1$, $l \geq 1$ and $\zeta_{p^l} \in \mathbb{C}_p$. We have,*

$$G_l(a) = \sum_{g(X)|(X^{q-1}-1)} \frac{d(g)}{d(g_a)} s(g_a) \zeta_{p^l}^{-fs(g)/d(g)},$$

*where the sum is over all the monic irreducible factors $g(X)$ of $X^{q-1}-1$ in $\mathbb{Z}_p[X]$, $g_a$ denotes the minimal polynomials of $t^{-a}$ over $\mathbb{Q}_p$ for any root $t \in \mu_{q-1}$ of $g(X)$.*

The main differences compared to Theorem 1.4.1 are: (1) the sum in Theorem 2.4.4 runs over $g(X)$ in $\mathbb{Z}_p[X]$ rather than in $\mathbb{F}_p[X]$, (2) $\zeta_p$ has been replaced by $\zeta_{p^l}$ and (3) $s(g)$ in the exponent of $\zeta_{p^l}$ is a $p$-adic integer rather than an element of $\mathbb{F}_p$, and thus makes sense modulo $p^l$.

*Proof.* For any monic irreducible polynomial $g(X)$ that is a factor of $X^{q-1}-1$ in $\mathbb{Z}_p[X]$, we know that $d(g)$ necessarily divides $f$ (recall $q = p^f$) and the trace $\mathrm{Tr}\colon \mathbb{Q}_q \to \mathbb{Q}_p$ of any root of $g(X)$ is equal to $-\frac{f}{d(g)}s(g)$. Hence, collecting together in $G(a)$ all $t \in \mu_{q-1}$ with a common minimal polynomial over $\mathbb{Q}_p$, we have

$$G_l(a) = -\sum_{g(X)|(X^{q-1}-1)} \left( \sum_{t \text{ root of } g(X)} t^{-a} \right) \zeta_{p^l}^{-fs(g)/d(g)}, \tag{2.4.1}$$

where the outer sum is over all the irreducible monic factors $g(X)$ of $X^{q-1}-1$ in $\mathbb{Z}_p[X]$ and the inner sum is over all the roots $t \in \mu_{q-1}$ of $g(X)$.

In addition, if $t$ is a root of $g(X)$, then it lies in $\mu_{q-1}$ and the full set of roots of $g(X)$ is $\{t, t^p, t^{p^2}, \ldots, t^{p^{d-1}}\}$, where $d = d(g)$. It follows that

$$\sum_{t \text{ root of } g(X)} t^{-a} = t^{-a} + (t^p)^{-a} + (t^{p^2})^{-a} + \cdots + (t^{p^{d-1}})^{-a}. \tag{2.4.2}$$

Denote by $S(g,a)$ the sum in (2.4.2). It is in $\mathbb{Z}_q$ and invariant under $\mathrm{Gal}(\mathbb{Q}_q/\mathbb{Q}_p)$, so $S(g,a) \in \mathbb{Z}_p$. For $t \in \mu_{q-1}$ with minimal polynomial $g(X)$ over $\mathbb{Q}_p$, the minimal polynomial of $t^{-a}$ over $\mathbb{Q}_p$ is a monic irreducible factor $g_a(X) \in \mathbb{Z}_p[X]$ of $X^{q-1}-1$. In addition, since $\mathbb{Q}_p(t^{-a}) \subseteq \mathbb{Q}_p(t)$, the degree $d(g_a)$ divides $d(g)$. It follows that

$S(g,a) = \mathrm{Tr}_{\mathbb{Q}_{p^{d(g)}}/\mathbb{Q}_p}(t^{-a}) = -\frac{d(g)}{d(g_a)}s(g_a)$. Plugging this formula into (2.4.1), we get

$$G_l(a) = \sum_{g(X)|(X^{q-1}-1)} \frac{d(g)}{d(g_a)} s(g_a)\zeta_{p^l}^{-fs(g)/d(g)},$$

which is what we wanted and can be used to encode $G_l(a)$ in Sage. □

By Remark 2.4.3, $G_{l,v}(a)$ does not generally lie in $\mathbb{Z}_p[\zeta_{p^l}]$ for $v \in \mathbb{Z}_q^\times - \mathbb{Z}_p^\times$. Therefore, we can't hope for a formula that would allow us to encode it in Sage at the present time.

### 2.4.3 Stickelberger's congruence

From the previous section, we know $G_l(a) \in \mathbb{Z}_p[\zeta_{p^l}]$. Every element of $\mathbb{Z}_p[\zeta_{p^l}]$ can be expanded in powers of $z_l := \zeta_{p^l} - 1$, the standard uniformizer of $\mathbb{Z}_p[\zeta_{p^l}]$. Blache's analogue of Stickelberger's congruence below gives the leading term of this expansion for $G_l(a)$ in terms of the base $p$ digits of $a$.

**Theorem 2.4.5** (Blache)**.** *For every $0 \le a < q-1$ and $l \ge 1$, we have*

$$G_l(a) \equiv \frac{z_l^{a_0+a_1+\cdots+a_{f-1}}}{a_0!a_1!\cdots a_{f-1}!} \bmod z_l^{a_0+a_1+\cdots+a_{f-1}+1},$$

*where $a = a_0 + a_1p + a_2p^2 + \cdots + a_{f-1}p^{f-1}$ with $0 \le a_i \le p-1$ for $0 \le i \le f-1$.*

*Proof.* See [4]. We also give a more elementary proof in Section 3.2 that avoids Witt vectors altogether. □

**Example 2.4.6.** Substituting $\zeta_{25} = 1 + z_2$ in Example 2.3.2, we have the following table.

| $a$ | $G_2(a)$ |
|---|---|
| 0 | $1+4z_2^4+2z_2^5+3z_2^6 \equiv \frac{z_2^0}{0!} \bmod z_2$ |
| 1 | $z_2+2z_2^2+2z_2^3+z_2^4+\cdots \equiv \frac{z_2^1}{1!} \bmod z_2^2$ |
| 2 | $3z_2^2+2z_2^3+4z_2^4+z_2^7+\cdots \equiv \frac{z_2^2}{2!} \bmod z_2^3$ |
| 3 | $z_2^3+z_2^4+3z_2^5+2z_2^8+\cdots \equiv \frac{z_2^3}{3!} \bmod z_2^4$ |

**Corollary 2.4.7.** *With notation as in Theorem* 2.4.5*, for every* $0 \le a < q-1$ *we have*

$$|G_l(a)|_p = |z_l|_p^{a_0+a_1+a_2+\cdots+a_{f-1}} = \left(\frac{1}{p}\right)^{\frac{a_0+a_1+a_2+\cdots+a_{f-1}}{p^{l-1}(p-1)}}.$$

*Proof.* The modulus in Theorem 2.4.5 has larger exponent than the power of $z_l$ on the right side, and $a_0!a_1!\cdots a_{f-1}! \in \mathbb{Z}_p^\times$. Also $|z_l|_p = |\zeta_{p^l}-1|_p = (1/p)^{1/(p^{l-1}(p-1))}$. □

We are able to show an analogue of Stickelberger's congruence for the generalized $p$-adic Gauss sums $G_{l,v}(a)$.

**Theorem 2.4.8.** *For every* $0 \le a < q-1$*,* $l \ge 1$ *and* $v \in 1+p\mathbb{Z}_p$*, we have*

$$G_{l,v}(a) \equiv \frac{z_l^{a_0+a_1+\cdots+a_{f-1}}}{a_0!a_1!\cdots a_{f-1}!} \bmod z_l^{a_0+a_1+\cdots+a_{f-1}+1},$$

*where* $a = a_0+a_1p+a_2p^2+\cdots+a_{f-1}p^{f-1}$ *with* $0 \le a_i \le p-1$ *for* $0 \le i \le f-1$.

*Proof.* See Section 3.1. □

Theorem 2.4.5 is a special case of the above theorem for $v = 1$.

**Remark 2.4.9.** By (2.3.2) and Theorem 2.4.8, if $v \in \mathbb{Z}_q^\times$, and $v_0 = \omega(v)$, then

$$G_{l,v}(a) = v_0^a G_{l,v/v_0}(a) \equiv v_0^a \frac{z_l^{a_0+a_1+\cdots+a_{f-1}}}{a_0!a_1!\cdots a_{f-1}!} \bmod z_l^{a_0+a_1+\cdots+a_{f-1}+1}.$$

**Corollary 2.4.10.** *With notation as in Theorem* 2.4.8*, for every* $0 \leq a < q-1$*,* $v \in 1+p\mathbb{Z}_q$*,* $l \geq 1$ *and* $n \gg l$*, we have*

$$|G_{l,v}(a)|_p = |z_l|_p^{a_0+a_1+a_2+\cdots+a_{f-1}} = \left(\frac{1}{p}\right)^{\frac{a_0+a_1+a_2+\cdots+a_{f-1}}{p^{l-1}(p-1)}}.$$

*Proof.* The modulus in Theorem 2.4.8 has larger exponent than the power of $z_l$ on the right side, and $a_0!a_1!\cdots a_{f-1}! \in \mathbb{Z}_p^\times$. Also $|z_l|_p = |\zeta_{p^l}-1|_p = (1/p)^{1/(p^{l-1}(p-1))}$. □

In fact, we can improve the exponent in the modulus of the congruences in both Theorem 2.4.5 and Theorem 2.4.8 if we use different uniformizers of $\mathbb{Z}_p[\zeta_{p^l}]$ instead of $z_l = \zeta_{p^l}-1$. Following the notation and results of Appendix B, for $n \geq 0$ or $n = \infty$ let

$$L_n(X) = X + \frac{X^p}{p} + \frac{X^{p^2}}{p^2} + \cdots + \frac{X^{p^n}}{p^n}.$$

For each choice of $\zeta_{p^l}$ and integer $n$ big enough compared to $l$, a uniformizer of $\mathbb{Z}_q[\zeta_{p^l}]$ can be singled out using a root of $L_n(X)$: for $l \geq 1$ and $n \geq l$ such that

$$p + p^2 + \cdots + p^{n-l+1} > n, \tag{2.4.3}$$

there is a unique root $\pi_{n,l} \in \mathbb{C}_p$ of $L_n(X)$ with $p$-adic valuation $1/(p^{l-1}(p-1))$ that is closest to $z_l = \zeta_{p^l}-1$. This $\pi_{n,l}$ is a uniformizer of $\mathbb{Z}_q[\zeta_{p^l}]$ by Corollary B.8. When $l = 1$, the inequality (2.4.3) holds for any $n \geq 1$, and when $n = \infty$ the condition (2.4.3) is dropped. Blache actually states Theorem 2.4.5 in [4] using the uniformizer $\pi_{\infty,l}$ instead of $z_l$.

We will write $n \gg l$ when $n \geq l \geq 1$ and (2.4.3) holds.

**Theorem 2.4.11.** *For every $0 \leq a < q-1$, $v \in 1+p\mathbb{Z}_q$, $l \geq 1$ and $n \gg l$, we have*

$$G_{l,v}(a) \equiv \frac{\pi_{n,l}^{a_0+a_1+\cdots+a_{f-1}}}{a_0!a_1!\cdots a_{f-1}!} \bmod \pi_{n,l}^{a_0+a_1+\cdots+a_{f-1}+p-1},$$

*where $a = a_0 + a_1p + a_2p^2 + \cdots + a_{f-1}p^{f-1}$ with $0 \leq a_i \leq p-1$ for $0 \leq i \leq f-1$.*

*Proof.* See two proofs in Chapter 3: Sections 3.3.2 and 3.3.3. □

Note the higher power of $\pi_{n,l}$ in the modulus of this theorem compared to Theorem 2.4.5.

In order to compute the $\pi_{n,l}$-expansion of $G_l(a)$ in Sage, we need a way to express $\zeta_{p^l}$ in terms of $\pi_{n,l}$. Theorem B.7 implies $\zeta_{p^l} = \mathrm{AH}_n(\pi_{n,l})$, where $\mathrm{AH}_n(X) = \exp(L_n(X)) = \exp(L_n(X)) \in \mathbb{Q}_p[[X]]$. Hence, we may simply substitute $\zeta_{p^l} = \mathrm{AH}_n(\pi_{n,l})$ in (1.4.1) and get the $\pi_{n,l}$-expansion of $G_l(a) \in \mathbb{Z}_p[\zeta_{p^l}] = \mathbb{Z}_p[\pi_{n,l}]$.

**Example 2.4.12.** Fix $p = q = 5$, $n = l = 2$, $v = 1$ and $\zeta_{25} \in \mathbb{C}_5$. Using Sage and substituting $\zeta_{25} = \mathrm{AH}_2(\pi_{2,2})$ in Example 2.3.2, where $\mathrm{AH}_2(X) = \exp(X + X^5/5 + X^{25}/25)$, we get the following table. Compare the exponents in the moduli with those in Example 2.4.6.

| $a$ | $G_{2,1}(a)$ |
|---|---|
| 0 | $1 + 4\pi_{2,2}^4 + \pi_{2,2}^8 + 4\pi_{2,2}^{12} + \cdots \equiv \frac{\pi^0}{0!} \bmod \pi_{2,2}^4$ |
| 1 | $\pi_{2,2} + 3\pi_{2,2}^{17} + 3\pi_{2,2}^{25} + 3\pi_{2,2}^{29} + \cdots \equiv \frac{\pi_{2,2}^1}{1!} \bmod \pi_{2,2}^5$ |
| 2 | $3\pi^2 + \pi_{2,2}^6 + \pi_{2,2}^{18} + 2\pi_{2,2}^{22} + \cdots \equiv \frac{\pi_{2,2}^2}{2!} \bmod \pi_{2,2}^6$ |
| 3 | $\pi_{2,2}^3 + 2\pi_{2,2}^7 + \pi_{2,2}^{11} + \pi_{2,2}^{15} + \cdots \equiv \frac{\pi_{2,2}^3}{3!} \bmod \pi_{2,2}^7$ |

In the above example, it looks like $G_{2,1}(a) \in \pi_{2,2}^a\mathbb{Z}_p[\pi_{2,2}^4]$ and from Example 1.4.9, we have $G(a) \in \pi^a\mathbb{Z}_p^\times$. These are not coincidences and follow from one of the proofs of Theorem 2.4.11. See Corollary 3.4.6 for the precise statement.

### 2.4.4 A partial analogue of the Gross–Koblitz formula

Besides being used to get a higher modulus in our analogue of Stickelberger's congruence, we wonder if the uniformizers $\pi_{n,l}$ of $\mathbb{Q}_p(\zeta_{p^l})$ allow us to lift Theorem 2.4.11 to an equality for $l \geq 2$ and $n \gg l$. For $n = l = 1$, we have $\pi_{1,1} = \pi$ and the Gross–Koblitz formula lifts Stickelberger's congruence (or equivalently Theorem 2.4.11) to an equality. For $l = 1$ and $n \geq 1$ or $n = \infty$, we have $\pi_{n,1} = \pi_n$ and Baldassarri's generalization of the Gross–Koblitz formula (Theorem 1.4.15) lifts Theorem 2.4.11 to an equality.

Let $K$ be an extension of $\mathbb{Q}_q(\zeta_{p^l})$ containing an element of absolute value $r$ such that

$$1 < r < p^{(p+p^2+p^3+\cdots+p^{n-l+1}-n)/p^n}. \tag{2.4.4}$$

When $n = l = 1$, inequality (2.4.4) turns into (1.4.4). The following theorem is a generalization of Theorem 1.4.14. We borrow the notation from Section 1.4.4.

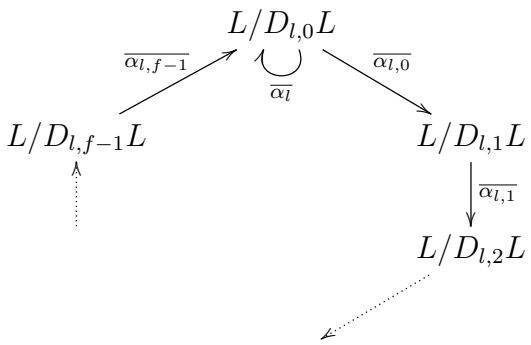
**Theorem 2.4.13.** *Fix $0 \leq a < q-1$, $l \geq 1$, $n \gg l$, $\zeta_{p^l} \in \mathbb{C}_p$ and $v \in 1+p\mathbb{Z}_q$ in standard form mod $p^l$. For $i \geq 0$ and $r \in |K^\times|_p$ such that (2.4.4) holds, there are differential operators $D_{l,i}\colon \mathcal{L}(r) \to \mathcal{L}(r)$ and completely continuous $K$-linear maps $\alpha_l\colon \mathcal{L}(r) \to \mathcal{L}(r)$ and $\alpha_{l,i}\colon \mathcal{L}(r) \to \mathcal{L}(r)$ such that (1) each quotient space $\mathcal{L}(r)/D\mathcal{L}(r)$ is $p^{l-1}$-dimensional over $K$ and (2) the induced maps $\overline{\alpha_l}\colon \mathcal{L}(r)/D_{l,i}\mathcal{L}(r) \to \mathcal{L}(r)/D_{l,i}\mathcal{L}(r)$ and $\overline{\alpha_{l,i}}\colon \mathcal{L}(r)/D_{l,i}\mathcal{L}(r) \to \mathcal{L}(r)/D_{l,i+1}\mathcal{L}(r)$ satisfy*

$$G_{l,v}(a) = (1-q)\operatorname{Tr}(\alpha_l) = \operatorname{Tr}(\overline{\alpha_l})$$

*and*

$$\overline{\alpha_l} = \overline{\alpha_{l,f-1}} \circ \overline{\alpha_{l,f-2}} \circ \cdots \circ \overline{\alpha_{l,1}} \circ \overline{\alpha_{l,0}},$$

*i.e., the following diagram commutes.*




*Proof.* See all of Chapter 4. Each section shows different parts of this theorem and the maps $\alpha_l$, $\alpha_{l,i}$ and differential operators $D_{l,i}$ are explicitly defined. □

# Chapter 3

# Proofs of generalized Stickelberger's congruences

In this chapter we present several proofs of the generalized Stickelberger's congruences discussed in Section 1.4.3 and Section 2.4.11.

## 3.1 Using Blache's analogue of Stickelberger's congruence

We use Theorem 2.4.5 to prove Theorem 2.4.8. From (2.3.2) we know that we may reduce to $v \in 1+p\mathbb{Z}_q$. In this section we express $G_{l,v}(a)$ in terms of $G_i(k)$ for $1 \leq i \leq l$, $0 \leq k < q-1$ and any $v \in \mathbb{Z}_q^\times$. Since for our purposes $v \in \mathbb{Z}_q^\times$ only matters mod $p^l$, using Teichmüller expansions there is no loss in only considering $v \in \mathbb{Z}_q$ such that $v = v_0 + v_1 p + \cdots + v_{l-1}p^{l-1}$ where $v_i \in \mu_{q-1} \cup \{0\}$ for $0 \leq i \leq l-1$. We will refer to such $v$ as "*in standard form mod* $p^l$."

**Theorem 3.1.1.** *For $l \geq 1$, $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$ and $0 \leq a < q-1$ we have*

$$G_{l,v}(a) = \frac{1}{(1-q)^{l-1}} \sum_{i_0+i_1+\cdots+i_{l-1} \equiv a \bmod q-1} G_{l,v_0}(i_0) G_{l-1,v_1}(i_1) \cdots G_{1,v_{l-1}}(i_{l-1}),$$

*where the sum is over all integers $0 \leq i_j < q-1$ for any $0 \leq j \leq l-1$ and in the Gauss sums $G_{i,v_k}(i_k)$ use compatible $\zeta_{p^i}$: $\zeta_{p^i}^p = \zeta_{p^{i-1}}$ for all $i \geq 1$.*

Recall that $G_{l,v}(a)$ depends also on the choice of $\zeta_{p^l}$ even though this is not clear in the notation. Thus we have to relate choices of $\zeta_{p^i}$ in $G_i(k)$ to the fixed $\zeta_{p^l}$ in $G_{l,v}(a)$.

**Remark 3.1.2.** We make two observations:

(1) For any $l \geq 1$ and $0 \leq a < q-1$: if $v \in \mu_{q-1}$, $G_{l,v}(a) = v^a G_l(a)$ and if $v = 0$,

$$G_{l,v}(a) = G_{l,0}(a) = \begin{cases} 1-q & \text{if } a = 0, \\ 0 & \text{if } 0 < a < q-1. \end{cases}$$

Thus we may write $G_{l,v}(a)$ in terms of the generalized Gauss sums $G_i(k)$ considered by Blache where $1 \leq i \leq l$ and $0 \leq k < q-1$ .

(2) The above theorem holds for complex-valued Gauss sums $G_{l,v}(\chi)$ as well since all the operations involved are algebraic.

*Proof.* Fix $l \geq 1$, $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$, $0 \leq a < q-1$ and $\zeta_{p^l} \in \mathbb{C}_p$. Define a compatible list $\{\zeta_{p^i}\}_{0 \leq i \leq l}$ of $p^{\text{th}}$ power roots of unity such that $\zeta_{p^i}^p = \zeta_{p^{i-1}}$ for all $1 \leq i \leq l$. It's easier to simplify the right hand side. To be able to fit computations, we set up some notation. Let $\underline{k} = (k_0, k_1, k_2, \ldots, k_{l-1})$, $[k] = \{0, 1, 2, \ldots, k\}$, $S(\underline{k}) =$

$k_0 + k_1 + \cdots + k_{l-1}$ and RHS be the right hand side of the formula in Theorem 3.1.1. Unfolding the definition of $G_{l,v}(a)$ and the notation, we get

$$\begin{aligned}
\text{RHS} &= \frac{(-1)^l}{(1-q)^{l-1}} \sum_{S(\underline{i}) \equiv a \bmod q-1} \left( \sum_{\underline{t} \in \mu_{q-1}^l} t_0^{-i_0} t_1^{-i_1} \cdots t_{l-1}^{-i_{l-1}} \zeta_{p^l}^{\mathrm{Tr}(t_0 v_0) + \cdots + \mathrm{Tr}(t_{l-1} v_{l-1}) p^{l-1}} \right) \\
&= -\frac{1}{(q-1)^{l-1}} \sum_{\underline{i} \in [q-2]^l} \left( \sum_{\underline{t} \in \mu_{q-1}^l} t_0^{i_1 + i_2 + \cdots + i_{l-1} - a} t_1^{-i_1} \cdots t_{l-1}^{-i_{l-1}} \zeta_{p^l}^{\mathrm{Tr}(t_0 v_0 + \cdots + t_{l-1} v_{l-1} p^{l-1})} \right) \\
&= -\frac{1}{(q-1)^{l-1}} \sum_{\underline{t} \in \mu_{q-1}^l} t_0^{-a} \zeta_{p^l}^{\mathrm{Tr}(t_0 v_0 + \cdots + t_{l-1} v_{l-1} p^{l-1})} \left( \sum_{\underline{i} \in [q-2]^l} (t_0/t_1)^{i_1} \cdots (t_0/t_{l-1})^{i_{l-1}} \right).
\end{aligned}$$

The inner sum $S(t_0, t_1, \ldots, t_{l-1})$ factors as follows.

$$S(t_0, t_1, \ldots, t_{l-1}) = \left( \sum_{i_1=0}^{q-2} (t_0/t_1)^{i_1} \right) \left( \sum_{i_2=0}^{q-2} (t_0/t_2)^{i_2} \right) \cdots \left( \sum_{i_{l-1}=0}^{q-2} (t_0/t_{l-1})^{i_{l-1}} \right).$$

Each of the sums in parenthesis is a geometric series. For any $1 \leq j \leq l-1$, we know that $t_j \in \mu_{q-1}$. Hence, for any $1 \leq j \leq l-1$ we have

$$\sum_{i_j=0}^{q-2} (t_0/t_j)^{i_j} = \begin{cases} q-1 & \text{if } \ t_0 = t_j \\ 0 & \text{if } \ t_0 \neq t_j \end{cases}.$$

Therefore, we have

$$S(t_0, t_1, \ldots, t_{l-1}) = \begin{cases} (q-1)^{l-1} & \text{if } \ t_0 = t_1 = t_2 = \cdots = t_{l-1} \\ 0 & \text{if } \ t_0 \neq t_j \text{ for some } 1 \leq j \leq l-1 \end{cases}.$$

Plugging $S(t_0, t_1, \ldots, t_{l-1})$ back into RHS and setting $t = t_0$, we get

$$\begin{aligned}
\text{RHS} &= -\frac{1}{(q-1)^{l-1}} \sum_{\underline{t} \in \mu_{q-1}^l} t_0^{-a} \zeta_{p^l}^{\mathrm{Tr}(t_0 v_0 + t_1 v_1 p + \cdots + t_{l-1} v_{l-1} p^{l-1})} S(t_0, t_1, \ldots, t_{l-1}) \\
&= -\frac{1}{(q-1)^{l-1}} \sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{\mathrm{Tr}(t v_0 + t v_1 p + \cdots + t v_{l-1} p^{l-1})} (q-1)^{l-1} \\
&= -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{\mathrm{Tr}(tv)} \\
&= G_{l,v}(a).
\end{aligned}$$

□

**Example 3.1.3.** Let $p = q = 5$, $a = 3$, $l = 2$, $\zeta_{25} \in \mathbb{C}_5$, $i = \omega(2)$ and $v = i - 5$ in standard form mod 25. Then we have

$$\begin{aligned}
G_{2,i-5}(3) &= \frac{1}{4} \left( G_{2,i}(0) G_{1,-1}(3) + G_{2,i}(1) G_{1,-1}(2) + G_{2,i}(2) G_{1,-1}(1) + G_{2,i}(3) G_{1,-1}(0) \right) \\
&= \frac{1}{4} \left( G_2(0) G(3) - i G_2(1) G(2) - G_2(2) G(1) + i G_2(3) G(0) \right).
\end{aligned}$$

Theorem 2.4.8 follows if we show that the summand with the biggest $p$-adic absolute value in Theorem 3.1.1 is the one where $i_0 = a$ and $i_j = 0$ for all $1 \leq j \leq l-1$. This looks deceptively easy, but we need a non-trivial combinatorial result to show it.

Let $[m] = \{0, 1, 2, \ldots, m\}$ and $a = a_0 + a_1 p + \cdots + a_{f-1} p^{f-1}$ be the base $p$ expansion of $a \in [q-2]$, where we recall $q = p^f$. We denote by $S(a) = a_0 + a_1 + \cdots + a_{f-1}$ the sum of the base $p$ digits of $a$, $\underline{n} = (n_0, n_1, \ldots, n_{f-1}) \in \mathbb{Z}^f$ an $f$-tuple of integers to which we associate the sum of its components $S(\underline{n}) = n_0 + n_1 + \cdots + n_{f-1}$ and the base $p$ expansion $B(\underline{n}) = n_0 + n_1 p + \cdots + n_{f-1} p^{f-1}$. Note that with this notation, we

have $S(\underline{a}) = S(a)$ and $B(\underline{a}) = a$, but generally $B(\underline{n})$ is not the base $p$ expansion of an integer since the components $n_j$ need not be in $[p-1] = \{0, 1, 2 \ldots, p-1\}$. We define $\underline{n} \geq 0$ to mean that $n_i \geq 0$ for all $0 \leq i \leq f-1$. The question we are considering is: for fixed non-negative integers $a < q-1$ and $S$, what are all possible $f$-tuples $\underline{n} \geq 0$ such that $B(\underline{n}) \equiv a \bmod p^f - 1$ and $S(\underline{n}) = S$?

**Lemma 3.1.4.** *Let $q = p^f$, $0 \leq a < q-1$ and $S$ be non-negative integers. If $S$ is* not *of the form $S = S(a) + m(p-1)$ for some non-negative integer $m$, then there are no possible $f$-tuples $\underline{n} \geq 0$ such that $B(\underline{n}) \equiv a \bmod q-1$ and $S(\underline{n}) = S$. On the other hand, if $S = S(a) + m(p-1)$ for some non-negative integer $m$, then all the possible $f$-tuples $\underline{n} \geq 0$ such that $B(\underline{n}) \equiv a \bmod q-1$ and $S(\underline{n}) = S$ are given by*

$$\underline{n} = \underline{a} + m_0(-1, 0, 0, \ldots, p) + m_1(p, -1, 0, 0, \ldots, 0) + \cdots + m_{f-1}(0, 0, \ldots, 0, p, -1),$$

*for all possible non-negative integers $m_i$ with $m = m_0 + m_1 + \cdots + m_{f-1}$. In particular, when $S = S(a)$, $m = 0$ and so there is only one $f$-tuple $\underline{n} = \underline{a}$ with the desired properties.*

*Proof.* First, notice that by further reducing $B(\underline{n}) \equiv a \bmod q-1$ modulo $p-1$ (using $p \equiv 1 \bmod p-1$), we get $S(\underline{n}) \equiv S(a) \bmod p-1$. Hence, $S = S(\underline{n}) = S(a) + m(p-1)$ for some integer $m$. So, only such values of $S$ are allowed to begin with, because otherwise there is no $f$-tuple satisfying the requirements. In addition, we claim that $m \geq 0$. Since $0 \leq a < q-1$, $B(\underline{n}) \equiv a \bmod q-1$ and $B(\underline{n}) \geq 0$, there is a non-negative integer $m_0$ such that $B(\underline{n}) = a + m_0(q-1)$. Adding $m_0$ to both sides and unfolding the notation, we get

$$n_0 + m_0 + n_1 p + \cdots + n_{f-1} p^{f-1} = a_0 + a_1 p + \cdots + a_{f-1} p^{f-1} + m_0 p^f. \qquad (3.1.1)$$

Reducing modulo $p$, we have $n_0+m_0 \equiv a_0 \bmod p$. Since $n_0, m_0 \geq 0$ and $0 \leq a_0 \leq p-1$, there exists a non-negative integer $m_1$ such that $n_0 + m_0 = a_0 + m_1 p$. Plugging this into 3.1.1, canceling $a_0$ and dividing by $p$, we have

$$n_1 + m_1 + n_2 p + \cdots + n_{f-1} p^{f-2} = a_1 + a_2 p + \cdots + a_{f-1} p^{f-2} + m_0 p^{f-1}.$$

Again, reducing modulo $p$, we have $n_1 + m_1 \equiv a_1 \bmod p$. Since $n_1, m_1 \geq 0$ and $0 \leq a_1 \leq p-1$, there exists a non-negative integer $m_2$ such that $n_1 + m_1 = a_1 + m_2 p$. Continuing this process, we have non-negative integers $m_0, m_1, m_2, \ldots, m_{f-1}$ such that $n_i + m_i = a_i + m_{i+1} p$ for $0 \leq i \leq f-2$ and $n_{f-1} + m_{f-1} = a_{f-1} + m_0 p$. Summing up all these equalities, we get

$$(n_0 + n_1 + \cdots + n_{f-1}) + (m_0 + m_1 + m_2 + \cdots + m_{f-1}) = S(a) + (m_0 + m_1 + \cdots + m_{f-1})p.$$

Therefore $S = S(\underline{n}) = S(a) + (m_0 + m_1 + \cdots + m_{f-1})(p-1)$ and so $m = m_0 + m_1 + \cdots + m_{f-1} \geq 0$ as claimed. In addition, $S = S(a)$ gives $m = 0$ which implies that $m_i = 0$ for $0 \leq i \leq f-1$ since all $m_i$ are non-negative. So we get that $n_i = a_i$ for $1 \leq i \leq f-1$. Hence there is only one $f$-tuple, namely $\underline{n} = (a_0, a_1, \ldots, a_{f-1})$ that answers our question for $S = S(a)$. In general, for $S = S(a) + m(p-1)$, we know that $\underline{n} + \underline{m} = \underline{a} + p(m_1, m_2, \ldots, m_{f-1}, m_0)$. Rewriting this equation, we also have

$$\underline{n} = \underline{a} + m_0(-1, 0, 0, \ldots, p) + m_1(p, -1, 0, 0, \ldots, 0) + \cdots + m_{f-1}(0, 0, \ldots, 0, p, -1),$$

which concludes our proof. □

Using Theorem 3.1.1, Lemma 3.1.4 and Blache's analogue of Stickelberger's con-

gruence for $G_l(a)$ (Theorem 2.4.5), we present the first proof of a Stickelbeger-type congruence.

*Proof of Theorem* 2.4.8. Let $0 \leq a < q-1$, $l \geq 1$ and $v \in 1+p\mathbb{Z}_p$ be in standard form mod $p^l$. Theorem 3.1.1 implies

$$G_{l,v}(a) = \frac{1}{(1-q)^{l-1}} \sum_{i_0+i_1+\cdots+i_{l-1} \equiv a \bmod q-1} G_l(i_0)G_{l-1,v_1}(i_1)\cdots G_{1,v_{l-1}}(i_{l-1}), \quad (3.1.2)$$

where the sum is over all integers $0 \leq i_j < q-1$ for any $0 \leq j \leq l-1$ and in the Gauss sums $G_{i,v_k}(i_k)$ use compatible $\zeta_{p^i}$: $\zeta_{p^i}^p = \zeta_{p^{i-1}}$ for all $i \geq 1$. To prove the theorem we now show that Corollary 2.4.7 implies the unique summand with the biggest $p$-adic absolute value in Theorem 3.1.1 is the one where $i_0 = a$ and $i_j = 0$ for all $1 \leq j \leq l-1$: by (2.3.1) and Theorem 2.4.5, we have

$$\begin{aligned} G_l(a)G_{l-1,v_1}(0)\cdots G_{1,v_{l-1}}(0) &= G_l(a)G_{l-1}(0)\cdots G_1(0) \\ &\equiv \frac{z_l^{a_0+a_1+\cdots+a_{f-1}}}{a_0!a_1!\cdots a_{f-1}!} \bmod z_l^{a_0+a_1+\cdots+a_{f-1}+1}. \end{aligned}$$

The theorem follows by the non-Archimedean property of the $p$-adic absolute value provided all the other terms in the sum (3.1.2) have smaller absolute value. By Corollary 2.4.7 we have

$$|G_l(i_0)G_{l-1,v_1}(i_1)\cdots G_{1,v_{l-1}}(i_{l-1})|_p = \left(\frac{1}{p}\right)^{\frac{S(i_0)+S(i_1)p+\cdots+S(i_{l-1})p^{l-1}}{p^{l-1}(p-1)}}.$$

Let $\tilde{B}(\underline{i}) = S(i_0)+S(i_1)p+\cdots+S(i_{l-1})p^{l-1}$. Our task is to find all $(i_0, i_1, \ldots, i_{l-1}) \in [q-1]^l$ that minimize $\tilde{B}(\underline{i})$. To do this, we use Lemma 3.1.4. Set $i_{j,k} \in [p-1]$ for all $k \in [f-1]$ to be the base $p$ digits of $i_j = i_{j,0}+i_{j,1}p+\cdots+i_{j,f-1}p^{f-1}$ for any $j \in [l-1]$.

We want $B(\underline{i}) = i_0 + i_1 + \cdots + i_{l-1}$ in Lemma 3.1.4. Thus, we let the components of $\underline{n}$ be $n_k = i_{0,k} + i_{1,k} + \cdots + i_{0,k}$ for all $k \in [f-1]$ and note that $S(\underline{n}) = \sum_{j\in[l-1],k\in[f-1]} i_{j,k} = S(i_0) + S(i_1) + \cdots + S(i_{l-1})$. Lemma 3.1.4 says that $S(\underline{n}) = S(a) + m(p-1)$ for some $m \geq 0$. Hence, $\tilde{B}(\underline{i}) = S(\underline{n}) + S(i_1)(p-1) + S(i_2)(p^2-1) + \cdots + S(i_{l-1})(p^{l-1}-1) = S(a) + (p-1)M$ where $M = m + S(\underline{i_1}) + S(i_2)(1+p)\cdots + S(\underline{i_{l-1}})(1+p+\cdots+p^{l-2}) \geq 0$. Thus, $\tilde{B}(\underline{i})$ is minimized precisely when $\tilde{B}(\underline{i}) = S(a)$. This happens only when $m = 0$, $i_j = 0$ for all $1 \leq j \leq l-1$ and $i_0 = a$, as claimed. □

**Remark 3.1.5.** If we knew Theorem 2.4.11 for $v = 1$, the above proof shows Theorem 2.4.11 for any $v \in 1 + p\mathbb{Z}_q$ since for all $\underline{i}$ as above, we have $\tilde{B}(\underline{i}) = S(a) + (p-1)M$ for some $M \geq 0$ and $M = 0$ if and only if $\underline{i} = (a, 0, 0, \ldots, 0)$.

## 3.2 A new proof of Theorem 2.4.5

Blache's proof of Theorem 2.4.5 relies on a $p$-adic series representation of the additive character $t \mapsto \zeta_{p^l}^{\mathrm{Tr}(t)}$. We present a more elementary proof.

*Proof of Theorem* 2.4.5. Let $z_l = \zeta_{p^l} - 1$. Then

$$\begin{aligned} G_l(a) &= -\sum_{t\in\mu_{q-1}} t^{-a}(1+z_l)^{\mathrm{Tr}(t)} \\ &= -\sum_{t\in\mu_{q-1}} t^{-a} \sum_{k\geq 0} \binom{\mathrm{Tr}(t)}{k} z_l^k \\ &= -\sum_{k\geq 0} \left( \sum_{t\in\mu_{q-1}} t^{-a} \binom{\mathrm{Tr}(t)}{k} \right) z_l^k. \end{aligned}$$

Let $H_{k,i}$ be the coefficients of $T(T-1)(T-2)\cdots(T-(k-1)) = H_{k,k}T^k + H_{k,k-1}T^{k-1} +$

$\cdots + H_{k,1}T + H_{k,0}$ and $C_{a,k} = \sum_{t\in\mu_{q-1}} t^{-a}\binom{\mathrm{Tr}(t)}{k}$. Then

$$\begin{aligned}
C_{a,k} &= \sum_{t\in\mu_{q-1}} t^{-a}\binom{\mathrm{Tr}(t)}{k} \\
&= \frac{1}{k!}\sum_{t\in\mu_{q-1}}\sum_{i=0}^{k} H_{k,i}\,\mathrm{Tr}(t)^i t^{-a} \\
&= \frac{1}{k!}\sum_{i=0}^{k} H_{k,i}\sum_{t\in\mu_{q-1}} (t + t^p + \cdots + t^{p^{f-1}})^i t^{-a} \\
&= \frac{1}{k!}\sum_{i=0}^{k} H_{k,i}\sum_{t\in\mu_{q-1}}\sum_{n_0+n_1+\cdots+n_{f-1}=i}\binom{i}{n_0,n_1,\ldots,n_{f-1}} t^{n_0+n_1p+\cdots+n_{f-1}p^{f-1}-a} \\
&= \frac{1}{k!}\sum_{i=0}^{k} H_{k,i}\sum_{n_0+n_1+\cdots+n_{f-1}=i}\binom{i}{n_0,n_1,\ldots,n_{f-1}}\sum_{t\in\mu_{q-1}} t^{n_0+n_1p+\cdots+n_{f-1}p^{f-1}-a} \\
&= \frac{q-1}{k!}\sum_{\substack{\underline{n}\geq 0, S(\underline{n})\leq k \\ B(\underline{n})\equiv a \bmod q-1}} H_{k,S(\underline{n})}\binom{S(\underline{n})}{n_0,n_1,\ldots,n_{f-1}}.
\end{aligned}$$

Therefore, the first non-zero $C_{k,n}$ in the sum $G(\chi,\psi_l)$ is the one that minimizes $S(\underline{n})$ such that $B(\underline{n}) \equiv a \bmod q-1$ and $\underline{n} \geq 0$. Lemma 3.1.4 implies that this happens only when $\underline{n} = \underline{a}$ and $k = S(a)$. In that case, we get

$$C_{a,S(a)} = \frac{q-1}{S(a)!}H_{S(a),S(a)}\binom{S(a)}{a_0,a_1,\ldots,a_{f-1}} = \frac{q-1}{a_0!a_1!\cdots a_{f-1}!},$$

and so the corollary follows. □

**Remark 3.2.1.** Note in particular the following:

1. Liang Xiao noted that replacing $z_l$ by a variable $X$ everywhere in the above proof, we get a Stickelbeger-type congruence for power series. This was our inspiration for the following proofs of Theorem 2.4.11.

2. Following the strategy in the above proof, we get another proof of Theorem 2.4.8. To do this it is convenient to use $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$ and expand $\mathrm{Tr}(tv) = \mathrm{Tr}(tv_0) + \mathrm{Tr}(tv_1)p + \cdots + \mathrm{Tr}(tv_{l-1})p^{l-1}$ using $\mathrm{Tr}(tv_i) = tv_i + (tv_i)^p + \cdots + (tv_i)^{p^{f-1}}$ for all $t \in \mu_{q-1}$ and $0 \le i \le l-1$. Since this other proof of Theorem 2.4.8 has the same main ideas as the above proof, we omit the details.

## 3.3 Stickelberger-type congruences through $\mathrm{AH}(X)$

A first step in the proof of the generalized Stickelberger's congruence with the higher exponent in the modulus (Theorem 2.4.11) is to represent the additive character $\psi_l \colon \mathbb{Z}_q \to \mathbb{C}_p^\times$ restricted to $\mu_q$ as a power series evaluated at $t \in \mu_{q-1}$. Blache [4] used this technique to prove Theorem 2.4.5, but did not realize that with a few more observations, his proof leads to a higher exponent in the modulus for $p \neq 2$. We first prove Theorem 2.4.11 for $n = \infty$, i.e. for the Gauss sums $G_l(a)$ with respect to the uniformizers $\pi_l = \pi_{\infty,l}$.

### 3.3.1 Power series representations through $\mathrm{AH}(X)$

We start by recalling some notations and results. As in Appendix A, let $\mathrm{AH}(X) = e^{L(X)} = \sum_{i\ge 0} A_i X^i = 1 + X + \cdots$, where $L(X) = \sum_{i\ge 0} X^{p^i}/p^i$. Both series converge on the open unit disc in $\mathbb{C}_p$ and $\mathrm{AH}(X)$ is the well-known Artin–Hasse series, which is in $\mathbb{Z}_p[[X]]$ (see Lemma A.2 and the discussion after it). However the numerical equality $\mathrm{AH}(x) = e^{L(x)}$ only holds in general for $|x|_p < (1/p)^{1/(p-1)}$. In fact $\mathrm{AH}(\pi) \neq e^{L(\pi)} = 1$ for any non-zero $p$-adic root $\pi$ of $L(X)$. By Theorem A.5, for any primitive $p^l$-th root

of unity $\zeta_{p^l}$ in $\mathbb{C}_p$, where $l \geq 1$, there is a unique root $\pi_l$ of $L(X)$ in $\mathbb{C}_p$ such that $\zeta_{p^l} = \mathrm{AH}(\pi_l)$. Therefore we can use the Artin–Hasse series to represent $p$-th power roots of unity, which will help us get a handle on the Gauss sums $G_l(a)$. With this in mind we write

$$G_l(a) = -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{t+t^p+\cdots+t^{p^{f-1}}} = -\sum_{t \in \mu_{q-1}} t^{-a}\, \mathrm{AH}(\pi_l)^{t+t^p+\cdots+t^{p^{f-1}}}.$$

Since it is easier to deal with power series, for any $z \in \mathbb{Z}_q$ we have

$$\mathrm{AH}(X)^{\mathrm{Tr}(z)} = (1 + (\mathrm{AH}(X) - 1))^{\mathrm{Tr}(z)} := \sum_{k \geq 0} \binom{\mathrm{Tr}(z)}{k} (\mathrm{AH}(X) - 1)^k.$$

**Definition 3.3.1.** For $0 \leq a < q-1$, the *Artin–Hasse Gauss series* associated to $a$ is

$$G(a, X) = -\sum_{t \in \mu_{q-1}} t^{-a}\, \mathrm{AH}(X)^{t+t^p+\cdots+t^{p^{f-1}}}.$$

Note that in particular $G(a, \pi_l) = G_l(a)$ for any $l \geq 1$. Hence $G(a, X)$ is a power series representation of $G_l(a)$ for all levels $l \geq 1$. The goal now is to understand the power series $G(a, X)$.

We will rewrite the power series $\mathrm{AH}(X)^{\mathrm{Tr}(t)}$ for $t \in \mu_{q-1}$ so that we can more easily compute its coefficients and the coefficients of the Gauss sum power series $G(a, X)$.

For $t \in \mu_{q-1}$, we have

$$\begin{aligned}
\mathrm{AH}(X)^{\mathrm{Tr}(t)} &= e^{\mathrm{Tr}(t)L(X)} \\
&= \exp(\mathrm{Tr}(t)X + \mathrm{Tr}(t)X^p/p + \mathrm{Tr}(t)X^{p^2}/p^2 + \cdots) \\
&= \exp(\mathrm{Tr}(t)X + \mathrm{Tr}(t^p)X^p/p + \mathrm{Tr}(t^{p^2})X^{p^2}/p^2 + \cdots) \\
&= \exp((t + t^p + \cdots + t^{p^{f-1}})X + (t^p + t^{p^2} + \cdots + t^{p^f})X^p/p + \cdots) \\
&= \exp(L(tX) + L(t^pX) + \cdots + L(t^{p^{f-1}}X)).
\end{aligned}$$

By definition of $\mathrm{AH}(X)$ we get

$$\mathrm{AH}(X)^{\mathrm{Tr}(t)} = \mathrm{AH}(tX)\,\mathrm{AH}(t^pX)\cdots\mathrm{AH}(t^{p^{f-1}}X), \tag{3.3.1}$$

even though the above calculation of power series breaks down numerically with $\pi_l$ in place of $X$ since $L(\pi_l) = 0$. This is why it is convenient to work with power series.

Note that for any $x \in \mathfrak{m}_p$ and $t \in \mu_{q-1}$, $\mathrm{AH}(x)^{\mathrm{Tr}(t)} = \mathrm{AH}(tx)\,\mathrm{AH}(t^px)\cdots\mathrm{AH}(t^{p^{f-1}}x)$ and in particular

$$\psi_l(t) = \zeta_{p^l}^{\mathrm{Tr}(t)} = \mathrm{AH}(\pi_l)^{\mathrm{Tr}(t)} = \mathrm{AH}(t\pi_l)\,\mathrm{AH}(t^p\pi_l)\cdots\mathrm{AH}(t^{p^{f-1}}\pi_l).$$

Replacing $t$ by the indeterminate $T$ in (3.3.1), we make the following definition.

**Definition 3.3.2.** The *Artin–Hasse additive character series* associated to $q = p^f$ is

$$\Psi(T, X) = \mathrm{AH}(TX)\,\mathrm{AH}(T^pX)\cdots\mathrm{AH}(T^{p^{f-1}}X) \in \mathbb{Z}_p[[T, X]].$$

If $\mathrm{Tr}(T) := T + T^p + \cdots + T^{p^{f-1}}$, note that

$$\Psi(T,X) \neq \mathrm{AH}(X)^{\mathrm{Tr}(T)} := \sum_{k\geq 0} \binom{\mathrm{Tr}(T)}{k} (\mathrm{AH}(X)-1)^k.$$

However, for $t \in \mu_{q-1}$, we do get $\Psi(t,X) = \mathrm{AH}(X)^{\mathrm{Tr}(t)}$ and so

$$G(a,X) = -\sum_{t\in\mu_{q-1}} t^{-a}\Psi(t,X).$$

### 3.3.2 Stickelberger-type congruences for $G(a,X)$ and $G_{l,v}(a)$

To get a Stickelberger-type congruence for $G(a,X)$, it is convenient to introduce some more notation. Let $f \geq 1$, $\underline{n} = (n_0, n_1, \ldots, n_{f-1}) \in \mathbb{Z}^f$, $B(\underline{n}) = n_0 + n_1 p + \cdots + n_{f-1}p^{f-1} \in \mathbb{Z}$ and $S(\underline{n}) = n_0 + n_1 + \cdots + n_{f-1}$. The notation $\underline{n} \geq 0$ means that $n_i \geq 0$ for all $0 \leq i \leq f-1$. Recalling $\mathrm{AH}(X) = \sum_{k\geq 0} A_k X^k$, we have

$$\begin{aligned}
\Psi(T,X) &= \left(\sum_{n_0\geq 0} A_{n_0} T^{n_0} X^{n_0}\right)\left(\sum_{n_1\geq 0} A_{n_1} T^{n_1 p} X^{n_1}\right)\cdots\left(\sum_{n_{f-1}\geq 0} A_{n_{f-1}} T^{n_{f-1}p^{f-1}} X^{n_{f-1}}\right) \\
&= \sum_{n_0,n_1,\ldots,n_{f-1}\geq 0} A_{n_0} A_{n_1}\cdots A_{n_{f-1}} T^{n_0+n_1 p+\cdots+n_{f-1}p^{f-1}} X^{n_0+n_1+\cdots n_{f-1}} \\
&= \sum_{\underline{n}\geq 0} A_{\underline{n}} T^{B(\underline{n})} X^{S(\underline{n})}.
\end{aligned}$$

where $A_{\underline{n}} = A_{n_0} A_{n_1}\cdots A_{n_{f-1}}$. Therefore for $t \in \mu_{q-1}$

$$\Psi(t,X) = \sum_{\underline{n}\geq 0} A_{\underline{n}} t^{B(\underline{n})} X^{S(\underline{n})}$$

and so

$$\begin{aligned} G(a,X) &= -\sum_{t\in\mu_{q-1}} t^{-a}\Psi(t,X) \\ &= -\sum_{t\in\mu_{q-1}} \left(\sum_{\underline{n}\geq 0} A_{\underline{n}} X^{S(\underline{n})} t^{B(\underline{n})-a}\right) \\ &= -\sum_{\underline{n}\geq 0} A_{\underline{n}} X^{S(\underline{n})} \left(\sum_{t\in\mu_{q-1}} t^{B(\underline{n})-a}\right) \\ &= (1-q) \sum_{\underline{n}\geq 0, B(\underline{n})\equiv a \bmod q-1} A_{\underline{n}} X^{S(\underline{n})}, \end{aligned}$$

The last equality follows from the geometric series identity: for any $k \in \mathbb{Z}$

$$\sum_{t\in\mu_{q-1}} t^k = \sum_{i=0}^{q-2} \zeta_{q-1}^{ik} = \begin{cases} q-1 & \text{if } k \equiv 0 \bmod q-1 \\ 0 & \text{if } k \not\equiv 0 \bmod q-1 \end{cases}.$$

We can use this computation to get the following theorem.

**Theorem 3.3.3.** *Let $q = p^f$, $0 \leq a < q-1$, $a = a_0 + a_1 p + \cdots + a_{f-1}p^{f-1}$ be the $p$-adic expansion of $a$ and $S(a) = a_0 + a_1 + \cdots + a_{f-1}$. Then*

$$G(a,X) \equiv (1-q)\frac{X^{S(a)}}{a_0!\, a_1! \cdots a_{f-1}!} \bmod X^{S(a)+p-1}\mathbb{Z}_p[[X^{p-1}]]$$

*and $G(a,X)/X^{S(a)} \in \mathbb{Z}_p[[X^{p-1}]]^\times$.*

*Proof.* With the notation of the theorem and the previous observations, we have

$$G(a,X) = (1-q) \sum_{\underline{n}\geq 0, B(\underline{n})\equiv a \bmod q-1} A_{\underline{n}} X^{S(\underline{n})},$$

where $A_{\underline{n}} \in \mathbb{Z}_p$. Lemma 3.1.4 implies that $S(\underline{n}) = S(a) + m(p-1)$ for some integer $m \geq 0$ and $S(\underline{n})$ is minimized when $m = 0$. Thus $S(\underline{n}) = S(a)$ and $\underline{n} = \underline{a}$. Since $0 \leq a_i \leq p-1$ for all $0 \leq i \leq f-1$ and the coefficients of the Artin–Hasse power series match the coefficients of the exponential power series up to degree $p-1$, we get $A_{a_i} = \frac{1}{a_i!}$. In addition, since $\mathrm{AH}(X) \in 1 + X\mathbb{Z}_p[[X]]$, we have $A_i \in \mathbb{Z}_p$ for all $i \geq 0$. Hence

$$G(a, X) \equiv (1-q)\frac{X^{S(a)}}{a_0!\, a_1! \cdots a_{f-1}!} \bmod X^{S(a)+p-1}\mathbb{Z}_p[[X^{p-1}]]$$

and $G(a, X)/X^{S(a)} \in \mathbb{Z}_p[[X^{p-1}]]^\times$, as wanted. □

We now use the power series congruence of $G(a, X)$ to prove Theorem 2.4.11.

*Proof of Theorem* 2.4.11. Recall that $G(a, \pi_l) = G_l(a)$, where $\pi_l = \pi_{\infty,l}$ is a root of $L(X) = L_\infty(X)$ of absolute value $r_l$. Hence we can simply plug in $X = \pi_l$ into Theorem 3.3.3 and get the desired result when $v = 1$ and $n = \infty$ due to the coefficients being $p$-adic integers. To show Theorem 2.4.11 for $n = \infty$ and all $v \in 1 + p\mathbb{Z}_q$, it is enough to recall Remark 3.1.5. Hence we have

$$G_{l,v}(a) \equiv \frac{\pi_l^{S(a)}}{a_0! a_1! \cdots a_{f-1}!} \bmod \pi_l^{S(a)+p-1}.$$

For $n \gg l$ such that $n \neq \infty$, we have $|\pi_{n,l} - \pi_l| = R(n,l) < r$, where $\pi_l = \pi_{\infty,l}$ by Theorem B.10. Thus $|\pi_l/\pi_{n,l} - 1| \leq r_l^{p-1}$ and so $|\pi_l^{S(a)}/\pi_{n,l}^{S(a)} - 1| \leq r_l^{p-1}$. Hence $|\pi_l^{S(a)} - \pi_{n,l}^{S(a)}| \leq r_l^{S(a)+p-1}$, which proves Theorem 2.4.11 for all finite $n$ such that $n \gg l$. □

**Remark 3.3.4.** In Theorem 3.3.3, we may plug in any $x \in \mathbb{C}_p$ such that $|x| < 1$ to

get

$$G(a,x) = -\sum_{t\in\mu_{q-1}} t^{-a}\,\mathrm{AH}(x)^{t+t^p+\cdots+t^{p^{f-1}}} =\equiv (1-q)\frac{x^{S(a)}}{a_0!\,a_1!\cdots a_{f-1}!} \bmod x^{S(a)+p-1}\mathbb{Z}_p[[x^{p-1}]]$$

and $G(a,x)/x^{S(a)} \in \mathbb{Z}_p[[x^{p-1}]]^\times$. For $z \in \mathbb{Z}_q$, the map $z \mapsto \mathrm{AH}(x)^{\mathrm{Tr}(z)}$ is a continuous additive character $\mathbb{Z}_q \to \mathbb{C}_p^\times$ for any $|x| < 1$. When $x = \pi_l$ is a root of $L(X)$ of absolute value $r_l = (1/p)^{1/(p^{l-1}(p-1))}$, the map $z \mapsto \mathrm{AH}(\pi_l)^{\mathrm{Tr}(z)} = \zeta_{p^l}^{\mathrm{Tr}(z)}$ is a continuous additive character of order $p^l$. For $x \in \mathbb{C}_p$ not a root of $L(X)$ and $|x| < 1$, the map $z \mapsto \mathrm{AH}(x)^{\mathrm{Tr}(z)}$ is a continuous additive character of infinite order.

### 3.3.3 Stickelberger-type congruence for $G_v(a,X)$ and $G_{l,v}(a)$

Since $G_{l,v}(a) \equiv \pi_l^{S(a)}/(a_0!a_1!\cdots a_{f-1}!) \bmod \pi_l^{S(a)+p-1}$ for any $v \in 1+p\mathbb{Z}_q$, it is natural to wonder if this congruence really comes from a corresponding congruence of power series as it happens when $v = 1$ by Theorem 3.3.3. We now follow the same strategy to represent $G_{l,v}(a)$ as a power series. Note that

$$\begin{aligned}\psi_{l,v}(t) &= \zeta_{p^l}^{\mathrm{Tr}(tv_0)+\mathrm{Tr}(tv_1)p+\cdots+\mathrm{Tr}(tv_{l-1})p^{l-1}} \\ &= \zeta_{p^l}^{\mathrm{Tr}(tv_0)}\zeta_{p^l}^{\mathrm{Tr}(tv_1)p}\cdots\zeta_{p^l}^{\mathrm{Tr}(tv_{l-1})p^{l-1}}.\end{aligned}$$

Hence we can decompose $\psi_{l,v}$ in terms of the basic additive characters $\psi_l$ for $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$ as

$$\psi_{l,v}(t) = \psi_l(tv_0)\psi_l(tv_1)^p\cdots\psi_l(tv_{l-1})^{p^{l-1}}. \tag{3.3.2}$$

Keeping in mind (3.3.1) and the above computation, we make the following defi-

nition.

**Definition 3.3.5.** For $0 \leq a < q-1$ and $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$, the *Artin–Hasse Gauss series* associated to $a$ and $v$ is

$$G_v(a,X) = -\sum_{t\in\mu_{q-1}} t^{-a}\Psi(tv_0,X)\Psi(tv_1,X)^p\cdots\Psi(tv_{l-1},X)^{p^{l-1}},$$

where we can replace $\Psi(tv_0,X)\Psi(tv_1,X)^p\cdots\Psi(tv_{l-1},X)^{p^{l-1}}$ by $\mathrm{AH}(X)^{\mathrm{Tr}(tv)}$ when needed.

Note that in particular $G_v(a,\pi_l) = G_{l,v}(a)$ for all $l \geq 1$. Hence $G_v(a,X)$ is a power series representation of $G_{l,v}(a)$ for all levels $l \geq 1$.

We now prove an analogue of Theorem 3.3.3 for the more general series $G_v(a,X)$ and then use it to get another proof of Theorem 2.4.11.

**Theorem 3.3.6.** *Let $q = p^f$, $0 \leq a < q-1$, $a = a_0 + a_1 p + \cdots + a_{f-1}p^{f-1}$ be the $p$-adic expansion of $a$ and $S(a) = a_0 + a_1 + \cdots + a_{f-1}$. For $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$ such that $\omega(\overline{v}) = v_0$, we have*

$$G_v(a,X) \equiv (1-q)v_0^a \frac{X^{S(a)}}{a_0!\,a_1!\cdots a_{f-1}!} \bmod X^{S(a)+p-1}\mathbb{Z}_p[v][[X^{p-1}]]$$

*and $G_v(a,X)/X^{S(a)} \in \mathbb{Z}_p[v][[X^{p-1}]]^\times$.*

*Proof.* We need a way to get rid of the powers of $p$ of the additive character series $\Psi(tv_0,X)\Psi(tv_1,X)^p\cdots\Psi(tv_{l-1},X)^{p^{l-1}}$ in $G_v(a,X)$. For any $i \geq 0$ we set

$$\mathrm{AH}(X)^{p^i} = \sum_{k\geq 0} A_k^{(i)} X^k,$$

where $A_k^{(i)} \in \mathbb{Z}_p$ since $\mathrm{AH}(X) \in \mathbb{Z}_p[[X]]$. Following a similar argument as in the beginning of Section **??** with $A_k^{(i)}$ in place of $A_k$ and using similar notation, for all $0 \le i \le l-1$ we have

$$\begin{aligned}\Psi(T,X)^{p^i} &= \mathrm{AH}(TX)^{p^i}\,\mathrm{AH}(T^pX)^{p^i}\cdots\mathrm{AH}(T^{p^{f-1}}X)^{p^i}\\ &= \sum_{\underline{n_i}\ge 0} A^{(i)}_{\underline{n_i}} T^{B(\underline{n_i})} X^{S(\underline{n_i})},\end{aligned}$$

where $\underline{n_i} = (n_{i,0}, n_{i,1}, \ldots, n_{i,f-1}) \in \mathbb{Z}^f$, $A^{(i)}_{\underline{n_i}} = A^{(i)}_{n_{i,0}} A^{(i)}_{n_{i,1}} \cdots A^{(i)}_{n_{i,f-1}}$, $B(\underline{n_i}) = n_{i,0} + n_{i,1}p + \cdots + n_{i,f-1}p^{f-1}$ and $S(\underline{n_i}) = n_{i,0} + n_{i,1} + \cdots + n_{i,f-1}$. Therefore, for $t, v_i \in \mu_{q-1}$ we have

$$\Psi(tv_i, X)^{p^i} = \sum_{\underline{n_i}\ge 0} A^{(i)}_{\underline{n_i}} (tv_i)^{B(\underline{n_i})} X^{S(\underline{n_i})},$$

and so

$$\mathrm{AH}(X)^{\mathrm{Tr}(tv)} = \sum_{\underline{\underline{n}}\ge 0} A_{\underline{\underline{n}}} (tv_0)^{B(\underline{n_0})} (tv_1)^{B(\underline{n_1})} \cdots (tv_{l-1})^{B(\underline{n_{l-1}})} X^{S(\underline{\underline{n}})},$$

where $\underline{\underline{n}} = (\underline{n_0}, \underline{n_1}, \ldots, \underline{n_{l-1}}) \in (\mathbb{Z}^f)^l$, $A_{\underline{\underline{n}}} = A^{(0)}_{\underline{n_0}} A^{(1)}_{\underline{n_1}} \cdots A^{(l-1)}_{\underline{n_{l-1}}}$ and $S(\underline{\underline{n}}) = S(\underline{n_0}) + S(\underline{n_1}) + \cdots + S(\underline{n_{l-1}}) = \sum_{i,j} n_{i,j}$. Thus

$$\begin{aligned}G_v(a,X) &= -\sum_{t\in\mu_{q-1}} t^{-a}\,\mathrm{AH}(X)^{\mathrm{Tr}(tv)}\\ &= -\sum_{t\in\mu_{q-1}} \left( \sum_{\underline{\underline{n}}\ge 0} A^{(i)}_{\underline{\underline{n}}} (v_0)^{B(\underline{n_0})} (v_1)^{B(\underline{n_1})} \cdots (v_{l-1})^{B(\underline{n_{l-1}})} X^{S(\underline{\underline{n}})} t^{B(\underline{\underline{n}})-a} \right)\\ &= -\sum_{\underline{\underline{n}}\ge 0} A_{\underline{\underline{n}}} (v_0)^{B(\underline{n_0})} (v_1)^{B(\underline{n_1})} \cdots (v_{l-1})^{B(\underline{n_{l-1}})} X^{S(\underline{\underline{n}})} \left( \sum_{t\in\mu_{q-1}} t^{B(\underline{\underline{n}})-a} \right)\\ &= (1-q) \sum_{\underline{\underline{n}}\ge 0, B(\underline{\underline{n}})\equiv a \bmod q-1} A_{\underline{\underline{n}}} (v_0)^{B(\underline{n_0})} (v_1)^{B(\underline{n_1})} \cdots (v_{l-1})^{B(\underline{n_{l-1}})} X^{S(\underline{\underline{n}})},\end{aligned}$$

where $B(\underline{\underline{n}}) = B(\underline{n_0}) + B(\underline{n_1}) + \cdots + B(\underline{n_{l-1}})$ and the last equality follows from the geometric series identity: for any $k \in \mathbb{Z}$

$$\sum_{t \in \mu_{q-1}} t^k = \sum_{i=0}^{q-2} \zeta_{q-1}^{ik} = \begin{cases} q-1 & \text{if } k \equiv 0 \bmod q-1 \\ 0 & \text{if } k \not\equiv 0 \bmod q-1 \end{cases}.$$

With the previous observations, we have

$$G_v(a, X) = (1-q) \sum_{\underline{\underline{n}} \geq 0, B(\underline{\underline{n}}) \equiv a \bmod q-1} A_{\underline{\underline{n}}} (v_0)^{B(\underline{n_0})} (v_1)^{B(\underline{n_1})} \cdots (v_{l-1})^{B(\underline{n_{l-1}})} X^{S(\underline{\underline{n}})},$$

where $A_{\underline{\underline{n}}} \in \mathbb{Z}_p$. Lemma 3.1.4 implies that $S(\underline{\underline{n}}) = S(a) + m(p-1)$ for some integer $m \geq 0$ and $S(\underline{\underline{n}})$ is minimized when $m = 0$. Thus $S(\underline{\underline{n}}) = S(a)$, $\underline{\underline{n}} = (\underline{a}, 0, 0, \ldots, 0)$ and $B(\underline{n_0}) = a$. Since $0 \leq a_i \leq p-1$ for all $0 \leq i \leq f-1$ and the coefficients of the Artin–Hasse power series match the coefficients of the exponential power series up to degree $p-1$, we get $A_{a_i} = \frac{1}{a_i!}$. Note that $A_k^{(0)} = A_k$ and the coefficients $A_k^{(i)}$ for $i \geq 1$ don't play any role besides the fact that they are in $\mathbb{Z}_p$ since $\underline{\underline{n}} = (\underline{a}, 0, 0, \ldots, 0)$. Since $v \in \mathbb{Z}_q^\times$ is in standard form mod $p^l$, we have $\mathbb{Z}_p[v_0, v_1, \cdots, v_{f-1}] = \mathbb{Z}_p[v]$. Thus

$$G_v(a, X) \equiv (1-q) v_0^a \frac{X^{S(a)}}{a_0!\, a_1! \cdots a_{f-1}!} \bmod X^{S(a)+p-1} \mathbb{Z}_p[v][[X^{p-1}]]$$

and $G(a, X)/X^{S(a)} \in \mathbb{Z}_p[v][[X^{p-1}]]^\times$, as wanted. □

As a result we have another proof of Theorem 2.4.11.

*Proof of Theorem* 2.4.11. Recall that $G_v(a, \pi_l) = G_{l,v}(a)$, where $\pi_l = \pi_{\infty,l}$ is a root of $L(X) = L_\infty(X)$ of absolute value $r_l$. Hence we can simply plug in $X = \pi_l$ into Theorem 3.3.6 and get the desired result for all $v \in 1 + p\mathbb{Z}_q$ and $n = \infty$ due to the

coefficients being $p$-adic integers. Hence we have

$$G_{l,v}(a) \equiv \frac{\pi_l^{S(a)}}{a_0!a_1!\cdots a_{f-1}!} \bmod \pi_l^{S(a)+p-1}.$$

For $n \gg l$ such that $n \neq \infty$, we have $|\pi_{n,l} - \pi_l| = R(n,l) < r$, where $\pi_l = \pi_{\infty,l}$ by Theorem B.10. Thus $|\pi_l/\pi_{n,l} - 1| \leq r_l^{p-1}$ and so $|\pi_l^{S(a)}/\pi_{n,l}^{S(a)} - 1| \leq r_l^{p-1}$. Hence $|\pi_l^{S(a)} - \pi_{n,l}^{S(a)}| \leq r_l^{S(a)+p-1}$, which proves Theorem 2.4.11 for all finite $n$ such that $n \gg l$. □

**Remark 3.3.7.** The statements in Remark 3.3.4 generalize if we replace $G(a,X)$ and $\mathrm{Tr}(z)$ by $G_v(a,X)$ and $\mathrm{Tr}(zv)$ respectively for any $v \in 1 + p\mathbb{Z}_q$. However, the collection of continuous additive characters of the form $z \mapsto \mathrm{AH}(x)^{\mathrm{Tr}(zv)}$ does not cover all possible continuous additive characters $\mathbb{Z}_q \to \mathbb{C}_p^\times$.

## 3.4 Stickelberger-type congruences through $\mathrm{AH}_n(X)$

We follow the layout of Section 3.3 to represent the additive characters $\psi_{l,v} \colon \mathbb{Z}_q \to \mathbb{C}_p^\times$ restricted to $\mu_{q-1}$ as power series evaluated at $t \in \mu_{q-1}$ for $n \geq 1$. We did this before in the case $n = \infty$, but now we do it for all $n \gg l$. Hence we will get infinitely many power series representations of $\psi_{l,v}$ indexed by $n$.

### 3.4.1 Power series representations through $\mathrm{AH}_n(X)$

We start by recalling some notations and results. As in Appendix B, let $\mathrm{AH}_n(X) = e^{L_n(X)} = \sum_{i\geq 0} A_{n,i}X^i = 1 + X + \cdots$, where $L_n(X) = \sum_{i=0}^{n} X^{p^i}/p^i$. The truncated Artin–Hasse series $\mathrm{AH}_n(X)$ have disc of convergence $D(\mathrm{AH}_n)$ given by Lemma B.3. By

Theorem B.7, for $n \gg l$ and any primitive $p^l$-th root of unity $\zeta_{p^l}$ in $\mathbb{C}_p$, there is a unique root $\pi_{n,l}$ of $L_n(X)$ in $\mathbb{C}_p$ such that $\zeta_{p^l} = \mathrm{AH}_n(\pi_{n,l})$. It follows that $\pi_{n,l}$ is the unique root of $L_n(X)$ closest to $\zeta_{p^l} - 1$. Therefore we can use the truncated Artin–Hasse series to represent $p$-th power roots of unity, which will help us get a handle on the Gauss sums $G_{l,v}(a)$. With this in mind, for $n \gg l$ and $v \in \mathbb{Z}_q^\times$ we get

$$G_{l,v}(a) = -\sum_{t \in \mu_{q-1}} t^{-a} \zeta_{p^l}^{\mathrm{Tr}(tv)} = -\sum_{t \in \mu_{q-1}} t^{-a}\, \mathrm{AH}_n(\pi_{n,l})^{\mathrm{Tr}(tv)}.$$

**Definition 3.4.1.** For any $0 \leq a < q-1$, $v \in \mathbb{Z}_q$ and $n \geq 1$ or $n = \infty$ the *Artin–Hasse Gauss series* associated to $a$, $n$ and $v$ is

$$G_{n,v}(a, X) = -\sum_{t \in \mu_{q-1}} t^{-a}\, \mathrm{AH}_n(X)^{\mathrm{Tr}(tv)}.$$

Fix $n \geq 1$. Note that in particular for all $l \geq 1$ such that $n \gg l$, if $\pi_{n,l}$ is the root of $L_n(X)$ closest to $\zeta_{p^l} - 1$, then $G_{n,v}(a, \pi_{n,l}) = G_{l,v}(a)$. So the goal now is to understand the power series $G_{n,v}(a, X)$.

We will rewrite the additive character power series $\mathrm{AH}_n(X)^{\mathrm{Tr}(tv)}$ so that we can more easily compute its coefficients and the coefficients of the power series $G_{n,v}(a, X)$.

For $t \in \mu_{q-1}$, by the properties of the trace map we have

$$\begin{aligned}
\mathrm{AH}_n(X)^{\mathrm{Tr}(t)} &= e^{\mathrm{Tr}(t)L_n(X)} \\
&= \exp(\mathrm{Tr}(t)X + \mathrm{Tr}(t)X^p/p + \cdots + \mathrm{Tr}(t)X^{p^n}/p^n) \\
&= \exp(\mathrm{Tr}(t)X + \mathrm{Tr}(t^p)X^p/p + \cdots + \mathrm{Tr}(t^{p^n})X^{p^n}/p^n) \\
&= \exp((t + t^p + \cdots + t^{p^{f-1}})X + \cdots + (t^{p^n} + t^{p^{n+1}} + \cdots + t^{p^{n+f-1}})X^{p^n}/p^n) \\
&= \exp(L_n(tX) + L_n(t^pX) + \cdots + L_n(t^{p^{f-1}}X)) \\
&= \mathrm{AH}_n(tX)\,\mathrm{AH}_n(t^pX)\cdots\mathrm{AH}_n(t^{p^{f-1}}X),
\end{aligned}$$

even though the calculation above breaks down numerically with $\pi_{n,l}$ in place of $X$ since $L_n(\pi_{n,l}) = 0$. However, we can plug in any $x \in D(\mathrm{AH}_n)$ into the power series equation

$$\mathrm{AH}_n(X)^{\mathrm{Tr}(t)} = \mathrm{AH}_n(tX)\,\mathrm{AH}_n(t^pX)\cdots\mathrm{AH}_n(t^{p^{f-1}}X), \tag{3.4.1}$$

so we have the numerical identity

$$\psi_l(t) = \zeta_{p^l}^{\mathrm{Tr}(t)} = \mathrm{AH}_n(\pi_{n,l})^{\mathrm{Tr}(t)} = \mathrm{AH}_n(t\pi_{n,l})\,\mathrm{AH}_n(t^p\pi_{n,l})\cdots\mathrm{AH}_n(t^{p^{f-1}}\pi_{n,l}).$$

Replacing $t$ by the indeterminate $T$ in (3.4.1), we make the following definition.

**Definition 3.4.2.** The *Artin–Hasse additive character series* associated to $n$ and $q = p^f$ is

$$\Psi_n(T, X) = \mathrm{AH}_n(TX)\,\mathrm{AH}_n(T^pX)\cdots\mathrm{AH}_n(T^{p^{f-1}}X).$$

When $n = \infty$ we get $\Psi_\infty(T, X) = \Psi(T, X)$. Fix $n \geq 1$. From the previous computation for $t \in \mu_{q-1} \cup \{0\}$ and all $l \geq 1$ such that such that $n \gg l$, we have $\psi_l(t) = \Psi_n(t, \pi_{n,l})$. Thus $\Psi_n$ is the power series representation we were looking for

that generalizes $\Psi$ in the previous section. Moreover, we can use $\Psi_n$ to represent the more general additive characters $\psi_{l,v}$ through (3.3.2) as

$$\psi_{l,v}(t) = \Psi_n(tv_0, \pi_{n,l})\Psi_n(tv_1, \pi_{n,l})^p \cdots \Psi_n(tv_{l-1}, \pi_{n,l})^{p^{l-1}}. \tag{3.4.2}$$

### 3.4.2 Generalized Stickelberger-type congruence for $G_{n,v}(X)$

To get a generalization of Theorem 3.3.6, we first need to get some estimates on the absolute value of the coefficients of $G_{n,v}(X)$.

**Lemma 3.4.3.** *Let $K$ be any extension of $\mathbb{Q}_p$ and for each $0 < r \leq 1$ set*

$$\mathcal{B}(r) = \left\{ \sum_{k\geq 0} c_k X^k \in K[[X]] : |c_k|_p r^k \leq 1 \text{ for all } k \geq 0 \right\}.$$

*If $g(X), h(X) \in \mathcal{B}(r)$, then $g(X)h(X) \in \mathcal{B}(r)$ and $g(X) + h(X) \in \mathcal{B}(r)$.*

*Proof.* This follows by applying the strong triangle inequality to estimate the coefficients of $g(X)h(X)$ and $g(X) + h(X)$. □

**Corollary 3.4.4.** *With the notation of the above lemma, for any $0 \leq a \leq q-1$, $v \in \mathbb{Z}_q^\times$, $n \geq 1$ or $n = \infty$, we have*

$$\mathrm{AH}_n(X), \mathrm{AH}_n(X)^{\mathrm{Tr}(tv)}, \Psi_n(1, X), G_{n,v}(a, X) \in \mathcal{B}(R(\mathrm{AH}_n)),$$

*where $R(\mathrm{AH}_n)$ is the radius of convergence of $\mathrm{AH}_n(X)$ and $K = \mathbb{Q}_p$ or $K = \mathbb{Q}_p(v)$ respectively.*

*Proof.* By Lemma B.3 we have $\mathrm{AH}_n(X) \in \mathcal{B}(R(\mathrm{AH}_n))$. The rest follow by Lemma 3.4.3 since they are finite sums of products of $\mathrm{AH}_n(X)$. □

**Theorem 3.4.5.** *Let $q = p^f$, $0 \leq a < q-1$, $a = a_0 + a_1 p + \cdots + a_{f-1}p^{f-1}$ be the $p$-adic expansion of $a$ and $S(a) = a_0 + a_1 + \cdots + a_{f-1}$. For $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$ such that $\omega(\overline{v}) = v_0$ and $n \gg l$, we have*

$$G_{n,v}(a,X) \equiv (1-q)v_0^a \frac{X^{S(a)}}{a_0!\, a_1! \cdots a_{f-1}!} \bmod X^{S(a)+p-1}\mathbb{Q}_p(v)[[X^{p-1}]]$$

*and $G_{n,v}(a,X) = \sum_{k \geq S(a)} g_k X^k \in \mathbb{Q}_p(v)[[X^{p-1}]]$ with*

$$|g_k|_p R(\mathrm{AH}_n)^k \leq 1 \text{ for all } k \geq S(a).$$

*Proof.* We only need to replace $\mathrm{AH}(X)$ by $\mathrm{AH}_n(X)$ in the proof of Theorem 3.3.6 and then apply Lemma 3.4.4. □

Note that when $n = \infty$, $R(\mathrm{AH}_\infty) = 1$ and the condition on the coefficients of $G_{n,v}(a,X)$ says that they lie in $\mathbb{Z}_p[v]$. Hence this theorem generalizes Theorem 3.3.6. The coefficients of $G_{n,v}(a,X)$ match those of $G_{\infty,v}(X) = G_v(a,X)$ up to degree $p^{n+1}-1$ and are therefore in $\mathbb{Z}_p[v]$, but beyond this point they are not guaranteed to be in $\mathbb{Z}_p[v]$. When $a = 0$, the above theorem implies Theorem 2.4.11, but for $a > 0$ we would need a better bound on the coefficients of $G_{n,v}(a,X)$. In any case, we know from previous proofs that indeed when we plug in $X = \pi_{n,l}$ we do get Theorem 2.4.11.

### 3.4.3 Degrees of extensions over $\mathbb{Q}_p$

We now deduce some consequences and make conjectures about the degree of $G_{l,v}(a)$ and $G_{l,v}(a)/\pi_{n,l}^{S(a)}$ over $\mathbb{Q}_p(v)$. For simplicity, in this section we denote by $|\cdot|$ the $p$-adic absolute value $|\cdot|_p$. From Theorem 2.4.11, we have $G_{l,v}(a)/\pi_{n,l}^{S(a)} \in \mathbb{Z}_p[v, \pi_{n,l}]^\times$,

but much more is true. Recall that $v \in \mathbb{Z}_q$ is in standard form mod $p^l$ if $v = v_0 + v_1 p + v_2 p^2 + \cdots + v_{l-1} p^{l-1}$ with $v_i \in \mu_{q-1} \cup \{0\}$. It follows that for such $v$, we have $\mathbb{Z}_p[v] = \mathbb{Z}_p[v_0, v_1, \ldots, v_{l-1}]$.

**Corollary 3.4.6.** *With notation as in Theorem* 2.4.11 *and* $v \in \mathbb{Z}_q^\times$ *in standard form mod* $p^l$*, we have* $G_{l,v}(a)/\pi_{n,l}^{S(a)} \in \mathbb{Z}_p[v, \pi_{n,l}^{p-1}]^\times$*, so* $G_{l,v}(a) \in \mathbb{Z}_p[v, \pi_{n,l}^{(S(a),p-1)}]$.

*Proof.* The first statement is clear for $n = \infty$ and for $n \gg l$ it follows from Theorem 3.4.5 by plugging in $X = \pi_{n,l}$ and Theorem 2.4.11. The second statement follows from the first one upon multiplying by $\pi_{n,l}^{S(a)}$. □

Recall that $\mathbb{Q}_p(\pi_{n,l}) = \mathbb{Q}_p(\zeta_{p^l})$ (Corollary B.8) is an extension of $\mathbb{Q}_p$ of degree $\varphi(p^l) = p^{l-1}(p-1)$. The following lemma gives the degree of $\pi_{n,l}^m$ over $\mathbb{Q}_p$ for any positive integer $m$.

**Lemma 3.4.7.** *Let* $p$ *be an odd prime,* $m$ *a positive integer and* $d = (m, p-1)$*. Then* $\mathbb{Q}_p(\pi_{n,l}^m) = \mathbb{Q}_p(\pi_{n,l}^d)$ *and* $[\mathbb{Q}_p(\pi_{n,l}) : \mathbb{Q}_p(\pi_{n,l}^d)] = d$.

*Proof.* Since $d$ divides $m$, it is clear that $\mathbb{Q}_p(\pi_{n,l}^m) \subseteq \mathbb{Q}_p(\pi_{n,l}^d)$. To show the equality and prove the lemma, it is enough to show that $[\mathbb{Q}_p(\pi_{n,l}) : \mathbb{Q}_p(\pi_{n,l}^m)] = d$. We will do this by finding the subgroup of $\mathrm{Gal}(\mathbb{Q}_p(\pi_{n,l})/\mathbb{Q}_p)$ fixing $\mathbb{Q}_p(\pi_{n,l}^m)$. First recall $\mathbb{Z}_p^\times/(1 + p^l\mathbb{Z}_p) \cong \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^l})/\mathbb{Q}_p) = \mathrm{Gal}(\mathbb{Q}_p(\pi_{n,l})/\mathbb{Q}_p)$ by the map $c \to \sigma_c$ where $\sigma_c$ is the automorphism defined by $\sigma_c(\zeta_{p^l}) = \zeta_{p^l}^c$. Since $\mathbb{Z}_p^\times/(1 + p^l\mathbb{Z}_p) \cong \mu_{p-1} \times (1 + p\mathbb{Z}_p)/(1+p^l\mathbb{Z}_p) \cong \mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/p^{l-1}\mathbb{Z}$ and these two groups $\mathbb{Z}/(p-1)\mathbb{Z}$ and $\mathbb{Z}/p^{l-1}\mathbb{Z}$ have relatively prime orders, it is enough to see how $c \in \mu_{p-1}$ and $c \in 1 + p^k\mathbb{Z}_p^\times$ act on $\pi_{n,l}^m$ for $1 \le k \le l-1$.

Let $c \in \mu_{p-1}$, i.e. $c \in \mathbb{Z}_p^\times$ with $c^{p-1} = 1$. Then we know that $\sigma_c(\pi_{n,l}) = c\pi_{n,l}$ and so $\sigma_c(\pi_{n,l}^m) = \pi_{n,l}^m$ is equivalent to $c^m\pi_{n,l}^m = \pi_{n,l}^m$ or $c^m = 1$. Hence the only $c \in \mu_{p-1}$ that

fix $\pi_{n,l}^m$ are the ones that satisfy $c^d = 1$.

Now let $c \in 1 + p^k\mathbb{Z}_p^\times$ for some $1 \leq k \leq l-1$. There exists $z \in \mathbb{Z}_p^\times$ such that $c = 1 + p^k z$. From the $\mathrm{AH}_n$ isometry, we get

$$\begin{aligned}
|\sigma_c(\pi_{n,l}) - \pi_l| &= |\,\mathrm{AH}_n(\sigma_c(\pi_{n,l})) - \mathrm{AH}(\pi_{n,l})| \\
&= |\zeta_{p^l}^c - \zeta_{p^l}| \\
&= |\zeta_{p^l}^{p^k z} - 1| \\
&= |\zeta_{p^{l-k}}^z - 1| \\
&= r_l^{p^k}.
\end{aligned}$$

Let $u \in \mathbb{Z}_p[\pi_{n,l}]^\times$ such that $\sigma_c(\pi_{n,l}) = u\pi_{n,l}$. Then from the above computation we get

$$|u-1| = |\sigma_c(\pi_{n,l})/\pi_{n,l} - 1| = \frac{|\sigma_c(\pi_{n,l}) - \pi_{n,l}|}{|\pi_{n,l}|} = \frac{r_l^{p^k}}{r_l} = r_l^{p^k-1}.$$

Then $\sigma_c(\pi_{n,l}^m) = \pi_{n,l}^m$ is equivalent to $u^m\pi_{n,l}^m = \pi_{n,l}^m$ or $u^m = 1$. Since by the above computation $u \in 1 + \pi_{n,l}^{p^k-1}\mathbb{Z}_p^\times$, we know that $u$ could only be a $p^{\text{th}}$ power root of unity. So $|u-1| = r_t$ for some $1 \leq t \leq l$. This means that $r_l^{p^k-1} = r_t$, or equivalently

$$\left(\frac{1}{p}\right)^{\frac{p^k-1}{p^{l-1}(p-1)}} = \left(\frac{1}{p}\right)^{\frac{1}{p^{t-1}(p-1)}}.$$

Thus $p^k - 1 = p^{l-t}$ or equivalently $p^{l-t}(p^{k-l+t} - 1) = 1$. Hence $l = t$, $p = 2$ and $k = 1$. So for odd primes $p$, we get that $u^m \neq 1$ for any $m$ which implies that the only $c \in \mathbb{Z}_p^\times$ that fix $\pi_{n,l}^m$ are the one that satisfy $c^d = 1$. Hence $[\mathbb{Q}_p(\pi_{n,l}) : \mathbb{Q}_p(\pi_{n,l}^m)] = d$, as wanted. □

By this lemma, we get that $[\mathbb{Q}_p(\pi_{n,l}) : \mathbb{Q}_p(\pi_{n,l}^{S(a)})] = [\mathbb{Q}_p(\pi_{n,l}) : \mathbb{Q}_p(\pi_{n,l}^{(S(a),p-1)})] =$

$(S(a), p-1)$ and $[\mathbb{Q}_p(\pi_{n,l}) : \mathbb{Q}_p(\pi_{n,l}^{p-1})] = p-1$. In addition, $\pi_{n,l}^{(S(a),p-1)}, \pi_{n,l}^{p-1}$ and $\pi_{n,l}$ are respective uniformizers of $\mathbb{Q}_p(\pi_{n,l}^{(S(a),p-1)})$, $\mathbb{Q}_p(\pi_{n,l}^{p-1})$ and $\mathbb{Q}_p(\pi_{n,l})$ and so $\mathbb{Z}_p[\pi_{n,l}^{(S(a),p-1)}]$, $\mathbb{Z}_p[\pi_{n,l}^{p-1}]$ and $\mathbb{Z}_p[\pi_{n,l}]$ are respective rings of integers of $\mathbb{Q}_p(\pi_{n,l}^{(S(a),p-1)})$, $\mathbb{Q}_p(\pi_{n,l}^{p-1})$ and $\mathbb{Q}_p(\pi_{n,l})$.

We now recall Conjecture 2.4.2 and add to it based on Corollary 3.4.6 and Lemma 3.4.7. Keep in mind $(a, p-1) = (S(a), p-1)$, $v \in \mathbb{Z}_q^\times$ only matters mod $p^l$ and $\mathbb{Q}_p(v)$ is the same as the field generated by the Teichmüller digits of $v$ in its $p$-expansion.

**Conjecture 3.4.8.** *For odd $p$, $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$ and $a \neq 0$, we have $\mathbb{Q}_p(G_{l,v}(a))/\mathbb{Q}_p(v)$ and $\mathbb{Q}_p(G_{l,v}(a)/\pi_{n,l}^{S(a)})/\mathbb{Q}_p(v)$ are the unique extensions of degree $p^{l-1}(p-1)/(a, p-1)$ and $p^{l-1}$ that lie between $\mathbb{Q}_p(v)(\zeta_{p^l})$ and $\mathbb{Q}_p(v)$. In other words,*

$$\mathbb{Q}_p(G_{l,v}(a)) = \mathbb{Q}_p(v)(\pi^{(S(a),p-1)}) \text{ and } \mathbb{Q}_p(G_{l,v}(a)/\pi_{n,l}^{S(a)}) = \mathbb{Q}_p(v)(\pi^{p-1}).$$

The containments

$$\mathbb{Q}_p(G_{l,v}(a)) \subseteq \mathbb{Q}_p(v)(\pi^{(S(a),p-1)}) \text{ and } \mathbb{Q}_p(G_{l,v}(a)/\pi_{n,l}^{S(a)}) \subseteq \mathbb{Q}_p(v)(\pi^{(S(a),p-1)})$$

as well as the degrees of the extensions $\mathbb{Q}_p(v)(\pi^{(S(a),p-1)})$ and $\mathbb{Q}_p(v)(\pi^{p-1})$ over $\mathbb{Q}_p(v)$ follow directly from Corollary 3.4.6 and Lemma 3.4.7. The question then becomes: are these containments in fact equalities?

First, notice that $pa \equiv a^{(f-1)} \bmod q-1$ where $a^{(f-1)} = a_{f-1} + a_0 p + a_1 p^2 + \cdots + a_{f-2} p^{f-1}$ and thus $G_{l,v}(pa) = G_{l,v}(a)$. Hence in our analysis we may assume $a$ is not divisible by $p$. To support Conjecture 3.4.8 for $v = 1$, we use the following strategy based on Galois theory. Recall the notation form the proof of Lemma 3.4.7. Since $\mathbb{Z}_p^\times \cong \mu_{p-1} \times (1 + p\mathbb{Z}_p)$ and the values of $c$ only matter modulo $p^l$, then we

want show that the subgroup of $(\mathbb{Z}_p/p^l)^\times$ which contains all the values of $c$ such that $\sigma_c(G_l(a)) = G_l(a)$ lies entirely in $\mu_{p-1}$. This follows if we can show that for any $0 \leq a < q-1$ and all $c$ of the form $c = 1 + p^k z$ for some $1 \leq k \leq l-1$ and $z \in \mathbb{Z}_p^\times$, $|\sigma_c(G_l(a)) - G_l(a)| \neq 0$. Similarly, the same values of $c$ need to be checked to show that $|\sigma_c(G_l(a)/\pi_{n,l}^{S(a)}) - G_l(a)/\pi_{n,l}^{S(a)}| \neq 0$.

To get started, we show the following lemma.

**Lemma 3.4.9.** *Let $x$ be an algebraic integer in an extension of $\mathbb{Q}_p$ with maximal ideal $\mathfrak{m}$ such that $x \equiv 1 \bmod \mathfrak{m}$. Then for any positive integer $n = p^k n'$ with $n'$ not divisible by $p$ we have*

$$|x^n - 1| = \begin{cases} |x-1| & \text{if } k = 0 \\ |p^k||x-1| & \text{if } k \geq 1 \text{ and } |x-1| < r_1 \\ |p^{k-l}||x-1|^{p^l} & \text{if } r_l < |x-1| < r_{l+1} \text{ for some } 1 \leq l < k \\ |x-1|^{p^k} & \text{if } k \geq 1 \text{ and } |x-1| > r_k. \end{cases},$$

*where $r_l = \left(\frac{1}{p}\right)^{\frac{1}{p^{l-1}(p-1)}}$. At the critical values $|x-1| = r_l$, the above equality turns into $\leq$.*

*Proof.* Let $x = 1 + z$ with $z \in \mathfrak{m}$. Using the binomial theorem we have

$$x^n - 1 = \sum_{i=1}^{n} \binom{n}{i} z^i.$$

Note that the assumption $x \equiv 1 \bmod \mathfrak{m}$ means $|z| = |x-1| < 1$ and that $r_{l+1}^p = r_l$ for any $l \geq 1$.

First, suppose that $n$ is not divisible by $p$. Then for $i \geq 2$, $|\binom{n}{i} z^i| \leq |z^2| < |z| =$

$|nz| = |\binom{n}{1}z|$. So the first term in the above sum has a bigger absolute value than the others. By the non-Archimedean property of the $p$-adic absolute value we have $|x^n - 1| = |z| = |x - 1|$.

Now we consider $n$ being a $p$-th power $n = p^k$ for some $k \geq 1$ and use induction on $k$. Assume first that $n = p$. Since for $0 < i < p$, $\binom{p}{i}$ is divisible by $p$ exactly once we have $|\binom{p}{i}z^i| < |pz|$ for all $1 < i < p$. By the non-Archimedean property of the $p$-adic absolute value

$$|x^p - 1| \leq \max_{1 \leq i \leq p} \left\{ \left| \binom{p}{i} z^i \right| \right\} = \max\{|pz|, |z|^p\}$$

and equality holds if $|z|^p \neq |pz|$, i.e. when $|z| = |x - 1| \neq r_1 = \left(\frac{1}{p}\right)^{\frac{1}{p-1}}$. This is the reason we avoid this case in the statement of the Lemma.[1] From this, one can easily see that

$$|x^p - 1| = \begin{cases} |p||x - 1| & \text{if } |x - 1| < r_1 \\ |x - 1|^p & \text{if } |x - 1| > r_1 \end{cases}. \tag{3.4.3}$$

So if we replace $x$ with $x^p$ in 3.4.3 we get

$$|x^{p^2} - 1| = \begin{cases} |p||x^p - 1| & \text{if } |x^p - 1| < r_1 \\ |x^p - 1|^p & \text{if } |x^p - 1| > r_1 \end{cases}.$$

By (3.4.3) we have: $|x - 1| < r_1$ implies $|x^p - 1| = |p||x - 1| < r_1$; $r_1 < |x - 1| < r_2$ implies $|x^p - 1| = |x - 1|^p < r_1$; $|x - 1| > r_2$ implies $|x^p - 1| = |x - 1|^p > r_1$. Hence

[1] When $x$ is a root of unity of order $p$, for example, $|x - 1| = r_1 \neq 0$, but $|x^p - 1| = 0$.

we get

$$|x^{p^2}-1| = \begin{cases} |p^2||x-1| & \text{if } |x-1| < r_1 \\ |p||x-1|^p & \text{if } r_1 < |x-1| < r_2 \\ |x-1|^{p^2} & \text{if } |x-1| > r_2 \end{cases}.$$

Now let $n = p^k$ for $k \geq 3$ and assume that the lemma is true for $n = p^{k-1}$. Then by replacing $x$ with $x^p$ and (3.4.3) we have

$$\begin{aligned} |x^{p^k}-1| &= \begin{cases} |p^{k-1}||x^p-1| & \text{if } |x^p-1| < r_1 \\ |p^{k-1-l}||x^p-1|^{p^l} & \text{if } r_l < |x^p-1| < r_{l+1} \text{ for some } 1 \leq l < k-1 \\ |x^p-1|^{p^{k-1}} & \text{if } |x^p-1| > r_{k-1} \end{cases} \\ &= \begin{cases} |p^k||x-1| & \text{if } |x-1| < r_1 \\ |p^{k-l}||x-1|^{p^l} & \text{if } r_l < |x-1| < r_{l+1} \text{ for some } 1 \leq l < k \\ |x^p-1|^{p^{k-1}} & \text{if } |x-1| > r_k \end{cases}. \end{aligned}$$

Thus the lemma follows by induction for $n$ any power of $p$.

Finally if $n = p^k n'$ with $n'$ not divisible by $p$ we have $|x^{p^n}-1| = |(x^{n'})^{p^k}-1|$ and thus

$$|x^{p^n}-1| = \begin{cases} |p^k||x^{n'}-1| & \text{if } |x^{n'}-1| < r_1 \\ |p^{k-l}||x^{n'}-1|^{p^l} & \text{if } r_l < |x^{n'}-1| < r_{l+1} \text{ for some } 1 \leq l < k \\ |x^{n'}-1|^{p^k} & \text{if } |x^{n'}-1| > r_k \end{cases}.$$

Now the lemma follows from $|x^{n'}-1| = |x-1|$. □

We now show two lemmas that support our Conjecture 3.4.8.

**Lemma 3.4.10.** *For $p$ odd and $S(a)$ divisible by $p$, we have*

$$\mathbb{Q}_p(G_l(a)) = \mathbb{Q}_p(\pi_{n,l}^{(a,p-1)}).$$

*Proof.* Let $b > 0$ determine the next term in the $\pi_{n,l}$ expansion of $G_l(a)$:

$$G_l(a) = B_{a,S(a)}\pi_{n,l}^{S(a)} + B_{a,S(a)+b}\pi_{n,l}^{S(a)+b} + \cdots$$

Then, applying $\sigma_c$ for $c = 1 + p^k$, we get

$$\sigma_c(G_l(a)) = B_{a,S(a)}{\pi'_{n,l}}^{S(a)} + B_{a,S(a)+b}{\pi'_{n,l}}^{S(a)+b} + \cdots$$

where $\pi'_{n,l} = \sigma_c(\pi_{n,l})$. Using the properties of the Artin–Hasse exponential as in the proof of Lemma 3.4.7, we let $u = \pi'_{n,l}/\pi_{n,l}$ and get $|u-1| = |\pi'_{n,l}-\pi_{n,l}|/|\pi_{n,l}| = r_{l-k}/r_l$. Then by factoring the corresponding power of $\pi_{n,l}$ in each summand, we get

$$\sigma_c(G_l(a)) - G_l(a) = B_{a,S(a)}\pi_{n,l}^{S(a)}(u^{S(a)} - 1) + B_{a,S(a)+b}\pi_{n,l}^{S(a)+b}(u^{S(a)+b} - 1) + \cdots .$$

By Lemma 3.4.9 for $S(a)$ not divisible by $p$, we have $|u^{S(a)}-1| = |u-1| \geq |u^{S(a)+b}-1|$ no-matter what $b > 0$ is since $|u^{S(a)+b} - 1|$ always has a factor of $|u - 1| < 1$. In addition, $b > 0$ and $|B_{a,S(a)}| = 1$ imply that

$$|B_{a,S(a)}\pi_{n,l}^{S(a)}(u^{S(a)} - 1)| > |B_{a,S(a)+b}\pi_{n,l}^{S(a)+b}(u^{S(a)+b} - 1)|$$

and so by the strong triangle inequality of the $p$-adic absolute value, we get

$$|\sigma_c(G_a) - G_a| = |B_{a,S(a)}\pi_{n,l}^{S(a)}(u^{S(a)} - 1)| \neq 0$$

as wanted. □

In addition we have the following lemma.

**Lemma 3.4.11.** *For odd $p$ and $c \in 1 + p\mathbb{Z}_p$ we have*

$$|\sigma_c(G_{l,v}(a)) - G_{l,v}(a)| \neq 0 \text{ or } \left|\sigma_c(G_{l,v}(a)/\pi_{n,l}^{S(a)}) - G_{l,v}(a)/\pi_{n,l}^{S(a)}\right| \neq 0.$$

*Proof.* Let $c \in 1 + p\mathbb{Z}_p$ and suppose

$$|\sigma_c(G_{l,v}(a)) - G_{l,v}(a)| = 0 \text{ and } \left|\sigma_c(G_{l,v}(a)/\pi_{n,l}^{S(a)}) - G_{n,l,v}(a)/\pi_{n,l}^{S(a)}\right| = 0.$$

Let $u = \sigma_c(\pi_{n,l})/\pi_{n,l}$ and so by a similar reasoning as before $|u - 1| < 1$. Hence we get

$$\left|\sigma_c(G_{l,v}(a)/\pi_{n,l}^{S(a)}) - G_{l,v}(a)/\pi_{n,l}^{S(a)}\right| = \left|\frac{\sigma_c(G_{l,v}(a)) - u^{S(a)}G_{l,v}(a)}{\sigma_c(\pi_{n,l})^{S(a)}}\right| = 0.$$

This implies that $(u^{S(a)} - 1)G_{l,v}(a) = 0$ and thus $u$ is a root of unity. Lemma 3.4.7 implies that this is not the case for $p$ odd. □

In the future, if we can show that for odd $p$ and $c \in 1 + p\mathbb{Z}_p$ we have both

$$|\sigma_c(G_{l,v}(a)) - G_{l,v}(a)| \neq 0 \text{ and } \left|\sigma_c(G_{l,v}(a)/\pi_{n,l}^{S(a)}) - G_{l,v}(a)/\pi_{n,l}^{S(a)}\right| \neq 0,$$

then Conjecture 3.4.8 follows.

# Chapter 4

# Proof of a partial generalization of Gross–Koblitz formula

In this chapter we will focus on proving Theorem 2.4.13. We denote the $p$-adic absolute value by $|\cdot|$ since we will not use the Archimedean absolute value. It is useful to keep in mind the sketch of the proof of Step 1 of the Gross–Koblitz formula in Section 1.4.4.

## 4.1 Trace formula on $\mathcal{L}(r)$

Fix a finite extension $K$ of $\mathbb{Q}_q(\zeta_p)$. For $r \in |K^\times|$, let

$$\mathcal{L}(r) \;=\; \left\{ \sum_{n\geq 0} g_n X^n \in K[[X]] : |g_n| r^n \to 0 \right\},$$

as in Section 1.4.4. This is a $K$-Banach space with respect to the norm $\left|\left|\sum_{n\geq 0} g_n X^n\right|\right| = \max_{n\geq 0}\{|g_n|r^n\}$ since $K$ contains an element $c$ such that $|c| = 1/r$, which implies that $\mathcal{L}(r)$ has an orthonormal basis $\{1, cX, c^2X^2, \ldots, \}$. This point is particularly important since it implies that $\mathcal{L}(r)$ is an infinite–dimensional vector space over $K$ where traces and determinants of completely continuous linear maps of $\mathcal{L}(r)$ are natural generalizations of their finite–dimensional counterparts. In general, completely continuous linear maps between Banach spaces are limits of continuous linear maps with finite–dimensional image. See [11, Section 5, p. 348] for more details on completely continuous linear maps and why their traces and determinants are well-defined. Also, note that for $r = 1$, $\mathcal{L}(1)$ is the Tate algebra over $K$ and for general $r \in |K^\times|$, $\mathcal{L}(r)$ is a scaled version of it.

The goal of this section is to represent the Gauss sum $G_{l,v}(a)$ as the trace of a completely continuous linear map $\alpha_l \colon \mathcal{L}(r) \to \mathcal{L}(r)$ for any $r$ in an fixed interval $(1, b)$ for some $b > 1$ depending on $\alpha_l$. Following [11, p. 353], the composition of a completely continuous linear map with a continuous linear map in any order is completely continuous. We will construct $\alpha_l$ as the composition of three linear maps among which one is completely continuous and thus $\alpha_l$ is completely continuous as well. From the definition of $\mathcal{L}(r)$, it is easy to see that $r_1 < r_2$ implies $\mathcal{L}(r_2) \subset \mathcal{L}(r_1)$. In other words, this means that a power series converging on a disc around the origin also converges in a smaller disc around the origin. The inclusion map $\mathcal{L}(r_2) \to \mathcal{L}(r_1)$ is a completely continuous linear map. Note that if $r_1 = r_2 = r$ the identity map $\mathcal{L}(r) \to \mathcal{L}(r)$ is not completely continuous. If that were the case then every continuous map would be completely continuous since you can compose it with the identity map. For any $g \in \mathcal{L}(r)$, the multiplication by $g$ map $g \colon \mathcal{L}(r) \to \mathcal{L}(r)$ (denoted

by the same symbol) is a continuous linear map. Dwork introduced the linear map $\Psi_p\colon K((X)) \to K((X))$ induced by

$$\Psi_p(X^n) = \begin{cases} X^{n/p} & \text{if } \ p|n \\ 0 & \text{if } \ p \not| n \end{cases}$$

and for $t \in \mu_{q-1}$ we have $\Psi_p(t) = t^{1/p}$, where $t^{1/p}$ is the inverse image of the Frobenius automorphism of $t$. We denote by $\Psi_q$ the composition of $\Psi_p$ by itself $f$ times, where $q = p^f$. The restriction $\Psi_p\colon \mathcal{L}(r) \to \mathcal{L}(r^p)$ is a continuous linear map too.

**Theorem 4.1.1.** *Let $r \in |K^\times|$ such that $r \geq 1$ and $g(X) \in \mathcal{L}(r)$. For $0 \leq a < q-1$, the composition map $\Psi_q \circ X^{-a}g(X)\colon L(r^q) \to L(r^q)$ given by*

$$\mathcal{L}(r^q) \to \mathcal{L}(r) \xrightarrow{X^{-a}g(X)} X^{-a}\mathcal{L}(r) \xrightarrow{\Psi_q} \mathcal{L}(r^q)$$

*satisfies*

$$(q-1)\operatorname{Tr}(\Psi_q \circ X^{-a}g(X)) = \sum_{t \in \mu_{q-1}} t^{-a}g(t)$$

*and is completely continuous for $r > 1$.*

The above trace formula holds for $r = 1$ as well even though $\Psi_q \circ X^{-a}g(X)\colon L(1) \to L(1)$ is not completely continuous. Condition $r \geq 1$ is necessary so that $r^q \geq r$ and the inclusion map $\mathcal{L}(r^q) \to \mathcal{L}(r)$ is well–defined. The map $\Psi_q : X^{-a}\mathcal{L}(r) \to \mathcal{L}(r)$ is continuous and well-defined since $a < q-1$ and thus $\Psi_q(X^{-a}g(X)) \in K[[X]]$ for any power series $g(X) \in K[[X]]$.

*Proof.* See [11, Theorem 3.1, p. 341] for the case $r = 1$, where $\mathcal{L}(1)$ is the Tate algebra. The proof generalizes for all $r \in |K^\times|$ such that $r \geq 1$. □

The right hand side of the formula in Theorem 4.1.1 becomes $G_{l,v}(a)$ if we were to find a power series $g(X)$ such that $g(t) = \psi_{l,v}(t)$ for all $t \in \mu_{q-1}$. Here is where we need the power series representation of the additive characters (3.4.2): for $t \in \mu_{q-1}$

$$\psi_l(t) = \Psi_n(t, \pi_{n,l}) \text{ and } \psi_{l,v}(t) = \Psi_n(tv_0, \pi_{n,l})\Psi_n(tv_1, \pi_{n,l})^p \cdots \Psi_n(tv_{l-1}, \pi_{n,l})^{p^{l-1}},$$

where

$$\Psi_n(T, X) = \mathrm{AH}_n(TX)\,\mathrm{AH}_n(T^pX)\cdots\mathrm{AH}_n(T^{p^{f-1}}X).$$

Previously we plugged in $T = t$ and considered $\Psi_n(t, X)$ as a power series in $X$, whereas in this chapter we set $X = \pi_{n,l}$ and consider $\Psi_n(T, \pi_{n,l})$ as a power series in $T$. Since we are used to working with power series in $X$, we use $X$ in place of $T$ and let

$$\theta_{n,l}(X) := \mathrm{AH}_n(X\pi_{n,l}),$$

so that

$$\Psi_n(X, \pi_{n,l}) = \theta_{n,l}(X)\theta_{n,l}(X^p)\cdots\theta_{n,l}(X^{p^{f-1}}).$$

Recall that $|\pi_{n,l}| = r_l = (1/p)^{1/(p^{l-1}(p-1))}$. Therefore by Lemma B.3 and (B.3), the series $\theta_{n,l}(X)$ converges precisely on the open disc of radius

$$R(\theta_{n,l}) = \frac{R(\mathrm{AH}_n)}{r_l} = p^{(p+p^2+\cdots+p^{n-l+1}-n)/(p^{n+1})} > 1. \tag{4.1.1}$$

In particular $R(\theta_{\infty,l}) = 1/r_l = p^{1/(p^{l-1}(p-1))}$. Thus $\Psi_n(X, \pi_{n,l})$ converges precisely on the open disc of radius

$$R(\Psi_n(X, \pi_{n,l})) = R(\theta_{n,l})^{1/p^{f-1}} = p^{\frac{p+p^2+\cdots+p^{n-l+1}-n}{p^n q}}.$$

When $n = \infty$, we get $R(\Psi(X, \pi_{\infty,l})) = p^{1/(p^{l-2}q(p-1))}$.

**Definition 4.1.2.** For $n \gg l$ and $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$, we set

$$\Theta_{n,l,v}(X) = \Psi_n(Xv_0, \pi_{n,l})\Psi_n(Xv_1, \pi_{n,l})^p \cdots \Psi_n(Xv_{l-1}, \pi_{n,l})^{p^{l-1}}.$$

We need to find a suitable $r \geq 1$ so that the power series representation $\Theta_{n,l,v}(X)$ of $\psi_{l,v}(t)$ lies in $\mathcal{L}(r)$ and thus we can apply Theorem 4.1.1. We already know the disc of convergence of $\Psi_n(X, \pi_{n,l})$ from the above computations. Since $R(\Psi_n(X, \pi_{n,l})) > 1$, raising $\Psi_n(X, \pi_{n,l})$ to powers of $p$ increases the radius of convergence, so we know that $R(\Theta_{n,l,v}(X)) \geq R(\Psi_n(X, \pi_{n,l}))$. When $v = 1$, $\Theta_{n,l,1}(X) = \Psi_n(X, \pi_{n,l})$ and so we cannot do any better than $r < R(\Psi_n(X, \pi_{n,l}))$ for all $v \in \mathbb{Z}_q^\times$ simultaneously.

**Definition 4.1.3.** Let

$$1 \leq r < R(\Psi_n(X, \pi_{n,l}))^q = R(\theta_{n,l})^p = p^{(p+p^2+\cdots+p^{n-l+1}-n)/p^n}$$

and define $\alpha_l \colon \mathcal{L}(r) \to \mathcal{L}(r)$ as the composition

$$\mathcal{L}(r) \to \mathcal{L}(r^{1/q}) \xrightarrow{X^{-a}\Theta_{n,l,v}(X)} X^{-a}\mathcal{L}(r^{1/q}) \xrightarrow{\Psi_q} \mathcal{L}(r),$$

i.e., $\alpha_l = \Psi_q \circ X^{-a}\Theta_{n,l,v}$.

When $n = \infty$ we have $1 \leq r < p^{1/(p^{l-2}(p-1))}$. For $n = l = v = 1$, we get condition (1.4.4) in Key fact 1, except for $r \neq 1$. We need this additional constraint later on so that the quotient spaces $\mathcal{L}(r)/D_{l,i}\mathcal{L}(r)$ are finite–dimensional, where $D_{l,i}$ are certain differential operators matching $D_i$ in Section 1.4.4 when $n = l = v = 1$. Even in the case $n = l = v = 1$, one can show in special cases that $\mathcal{L}(r)/D_i\mathcal{L}(r)$ is infinite–

dimensional over $K$ by direct computations as in [11, proof of Lemma 1.1, p. 333]. For sufficiently big $n$, we have $1 < p^{1/(p^{l-1}(p-1))} < p^{(p+p^2+\cdots+p^{n-l+1}-n)/p^n}$ and thus we may take $K = \mathbb{Q}_q(\zeta_{p^l})$. Otherwise for any $n \gg l$, we set $K$ to be the smallest extension of $\mathbb{Q}_q(\zeta_{p^l})$ that contains an element $c$ such that $1 < |c| < p^{(p+p^2+\cdots+p^{n-l+1}-n)/p^n}$.

Therefore by Theorem 4.1.1, we have

$$G_{l,v}(a) = (1-q)\operatorname{Tr}(\alpha_l). \tag{4.1.2}$$

## 4.2 Differential operator and trace on $\mathcal{L}(r)/D_{l,0}\mathcal{L}(r)$

Our goal now is to define a differential operator $D_{l,0}\colon \mathcal{L}(r) \to \mathcal{L}(r)$ that induces a map $\overline{\alpha_l}\colon \mathcal{L}(r)/D_{l,0}\mathcal{L}(r) \to \mathcal{L}(r)/D_{l,0}\mathcal{L}(r)$ such that $G_{l,v}(a) = \operatorname{Tr}(\overline{\alpha_l})$. This is useful since we will show that the space $\mathcal{L}(r)/D_{l,0}\mathcal{L}(r)$ is finite–dimensional over $K$ as opposed to $\mathcal{L}(r)$. In order to define the differential operator $D_{l,0}$ generalizing $D_0$ in Key fact 2, we use the commutativity property that $\alpha_l$ and $D_{l,0}$ need to satisfy:

$$\alpha_l \circ D_{l,0} = D_{l,0} \circ q\alpha_l.$$

For this reason, it is useful to rewrite $\Theta_{n,l,v}(X)$ in $\alpha_l = \Psi_q \circ X^{-a}\Theta_{n,l,v}$ using the following power series.

**Definition 4.2.1.** For $n \gg l$, we set

$$\widehat{\theta_{n,l}}(X) := \prod_{i=0}^{\infty} \theta_{n,l}\left(X^{p^i}\right).$$

This is useful since $\theta_{n,l}(X)$ can be expressed as $\widehat{\theta_{n,l}}(X)/\widehat{\theta_{n,l}}(X^p)$ and

$$\Psi_n(X, \pi_{n,l}) = \prod_{i=0}^{f-1} \theta_{n,l}\left(X^{p^i}\right) = \widehat{\theta_{n,l}}(X)/\widehat{\theta_{n,l}}(X^q).$$

Hence we get

$$\Theta_{n,l,v}(X) = \frac{\widehat{\theta_{n,l}}(Xv_0)\widehat{\theta_{n,l}}(Xv_1)^p \cdots \widehat{\theta_{n,l}}(Xv_{l-1})^{p^{l-1}}}{\widehat{\theta_{n,l}}(X^q v_0)\widehat{\theta_{n,l}}(X^q v_1)^p \cdots \widehat{\theta_{n,l}}(X^q v_{l-1})^{p^{l-1}}}. \tag{4.2.1}$$

The operator $\psi_q$ used in the definition of $\alpha_l$ has the following useful properties: for any $g(X), h(X) \in K((X))$, we have

$$\psi_q(g(X^q)h(X)) = g(X)\psi_q(h(X)) \text{ and } \psi_q \circ X\frac{\mathrm{d}}{\mathrm{d}X} = qX\frac{\mathrm{d}}{\mathrm{d}X} \circ \psi_q \tag{4.2.2}$$

By the first property in (4.2.2) and $-a = \frac{-a(q-1)}{q-1} = \frac{a}{q-1} - \frac{qa}{q-1}$, we get

$$\alpha_l = \frac{1}{X^{a/(q-1)} \prod_{i=0}^{l-1} \widehat{\theta_{n,l}}(Xv_i)^{p^i}} \circ \psi_q \circ X^{a/(q-1)} \prod_{i=0}^{l-1} \widehat{\theta_{n,l}}(Xv_i)^{p^i},$$

which holds as an operator on $K[[X]](X^{a/(q-1)})$. We let

$$D_{l,0} = \frac{1}{X^{a/(q-1)} \prod_{i=0}^{l-1} \widehat{\theta_{n,l}}(Xv_i)^{p^i}} \circ X\frac{\mathrm{d}}{\mathrm{d}X} \circ X^{a/(q-1)} \prod_{i=0}^{l-1} \widehat{\theta_{n,l}}(Xv_i)^{p^i} \tag{4.2.3}$$

so that the commutativity relation $\alpha_l \circ D_{l,0} = D_{l,0} \circ q\alpha_l$ is satisfied by the second property in (4.2.2). By (4.2.3), we have $D_{l,0}$ is an operator on $K[[X]](X^{a/(q-1)})$ so it is not clear that $D_{l,0}$ is a well-defined differential operator on $\mathcal{L}(r)$. However, we now rewrite $\widehat{\theta_{n,l}}(X)$ as the exponential of a polynomial for finite $n$ (power series when

$n = \infty$), which will lead to a simpler formula for $D_{l,0}$. This formula will make it clear that $D_{l,0}$ is a differential operator $\mathcal{L}(r) \to \mathcal{L}(r)$ and reduces to the operator $D_0$ introduced in Chapter 1 when $l = n = v = 1$.

We begin by rewriting $\widehat{\theta_{n,l}}(X)$ as follows:

$$\begin{aligned}
\widehat{\theta_{n,l}}(X) &= \prod_{i=0}^{\infty} \mathrm{AH}_n(\pi_{n,l} X^{p^i}) \\
&= \prod_{i=0}^{\infty} \exp(L_n(\pi_{n,l} X^{p^i})) \\
&= \exp\left( \sum_{i=0}^{\infty} L_n(\pi_{n,l} X^{p^i}) \right) \\
&= \exp\left( \sum_{i=0}^{\infty} \sum_{m=0}^{n} \frac{\pi_{n,l}^{p^m} X^{p^{i+m}}}{p^m} \right) \\
&= \exp\left( \sum_{k=0}^{n-1} L_k(\pi_{n,l}) X^{p^k} \right),
\end{aligned}$$

where the last equality follows from the fact that for $k \geq n$, the coefficient of $X^{p^k}$ in the exponential is $L_n(\pi_{n,l}) = 0$. Hence, for finite $n$, $\widehat{\theta_{n,l}}(X)$ is the exponential of a polynomial $f(X) := \sum_{k=0}^{n-1} L_k(\pi_{n,l}) X^{p^k}$. Then for any $g(X) \in K[[X]]$, by the product rule we get

$$D_{l,0}(g(X)) = \frac{a}{q-1} g(X) + \sum_{i=0}^{l-1} p^i X f'(X v_i) g(X) + X \frac{\mathrm{d}}{\mathrm{d}X} g(X).$$

Hence we have

$$D_{l,0} = X \frac{\mathrm{d}}{\mathrm{d}X} + \frac{a}{q-1} + \sum_{i=0}^{l-1} p^i X f'(X v_i),$$

which is obviously a well–defined operator on $\mathcal{L}(r)$ for finite $n$. When $n = \infty$ we will show at the end of this section that $f'(X) \in \mathcal{L}(r)$ is only true for $r < p^{1/(p^{l-1}(p-1))}$.

For $v = l = n = 1$, we get $D_{1,0} = X\frac{\mathrm{d}}{\mathrm{d}X} + \frac{a}{q-1} + \pi X = D_0$, which we recognize from Section 1.4.4.

**Remark 4.2.2.** We motivated the definition of $D_{l,0}$ using the desired commutativity relation $\alpha_l \circ D_{l,0} = D_{l,0} \circ q\alpha_l$. However, to define $D_{l,0}$ we were originally inspired by Baldassarri's paper [3], which defines the same operator for $l = 1$ and any $n \geq 1$ or $n = \infty$.

The relation $\alpha_l \circ D_{l,0} = D_{l,0} \circ q\alpha_l$ induces the linear map $\overline{\alpha_l}\colon \mathcal{L}(r)/D_0\mathcal{L}(r) \to \mathcal{L}(r)/D_0\mathcal{L}(r)$ and we have commutativity of the diagram below.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{L}(r) & \xrightarrow{D_{l,0}} & \mathcal{L}(r) & \longrightarrow & \mathcal{L}(r)/D_{l,0}\mathcal{L}(r) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle q\alpha_l} & & \downarrow{\scriptstyle \alpha_l} & & \downarrow{\scriptstyle \overline{\alpha_l}} & & \\ 0 & \longrightarrow & \mathcal{L}(r) & \xrightarrow{D_{l,0}} & \mathcal{L}(r) & \longrightarrow & \mathcal{L}(r)/D_{l,0}\mathcal{L}(r) & \longrightarrow & 0 \end{array}$$

By [11, p. 358] this implies $\mathrm{Tr}(q\alpha_l) - \mathrm{Tr}(\alpha_l) + \mathrm{Tr}(\overline{\alpha_l}) = 0$, so

$$(q-1)\,\mathrm{Tr}(\alpha_l) + \mathrm{Tr}(\overline{\alpha_l}) = 0.$$

Therefore by (4.1.2) the Gauss sum $G_{l,v}(a)$ is a trace:

$$G_{l,v}(a) = \mathrm{Tr}(\overline{\alpha_l}). \tag{4.2.4}$$

Later on we will show that the space $\mathcal{L}(r)/D_{l,0}\mathcal{L}(r)$ has $K$-dimension $p^{l-1}$ and thus (4.2.4) gives us $G_{l,v}(a)$ as a trace on a finite–dimensional vector space over $K$. This is an improvement compared to (4.1.2), which realizes $G_{l,v}(a)$ as a trace on an infinite–dimensional vector space over $K$.

## 4.3 Decomposing $\alpha_l$ and $\overline{\alpha_l}$ as compositions

Now we generalize the decomposition of $\alpha$ in terms of $\alpha_i$ in Chapter 1.

**Definition 4.3.1.** Let $1 \leq r < p^{(p+p^2+\cdots+p^{n-l+1}-n)/p^n}$. For all $0 \leq i \leq f-1$ define $\alpha_{l,i} \colon \mathcal{L}(r) \to \mathcal{L}(r)$ as the composition

$$\mathcal{L}(r) \to \mathcal{L}(r^{1/p}) \xrightarrow{X^{-a_i}\prod_{j=0}^{l-1}\theta_{n,l}(Xv_j)^{p^j}} X^{-a_i}\mathcal{L}(r^{1/p}) \xrightarrow{\Psi_p} \mathcal{L}(r),$$

i.e., $\alpha_{l,i} = \Psi_p \circ X^{-a_i}\prod_{j=0}^{l-1}\theta_{n,l}(Xv_j)^{p^j}$, where $a_i$ is the $i^{\text{th}}$ base $p$ digit of $a$.

Since $a = a_0 + a_1 p + \cdots + a_{f-1}p^{f-1}$ and

$$\Theta_{n,l,v}(X) = \prod_{i=0}^{f-1}\prod_{j=0}^{l-1}\theta_{n,l}((Xv_j)^{p^i})^{p^j},$$

by (4.2.2) we have

$$\alpha_l = \alpha_{l,f-1} \circ \alpha_{l,f-2} \circ \cdots \circ \alpha_{l,1} \circ \alpha_{l,0}. \tag{4.3.1}$$

In order to define the differential operators $D_{l,i}$ that generalize $D_i$, we rewrite $\alpha_{l,i}$ so that their definition becomes evident in light of the desired commutativity relation $\alpha_{l,i} \circ D_{l,i} = D_{l,i+1} \circ p\alpha_{l,i}$. As we did for $\alpha_l$ we use the function $\widehat{\theta_{n,l}}(X)$. Since $-a_i = \frac{a^{(i)}}{q-1} - \frac{pa^{(i+1)}}{q-1}$, where $a^{(i)} = a_i + a_{i+1}p + \cdots + a_{i-1}p^{f-1}$, we have

$$\alpha_{l,i} = \frac{1}{X^{a^{(i+1)}/(q-1)}\prod_{j=0}^{l-1}\widehat{\theta_{n,l}}(Xv_j)^{p^j}} \circ \psi_p \circ X^{a^{(i)}/(q-1)}\prod_{j=0}^{l-1}\widehat{\theta_{n,l}}(Xv_j)^{p^j}.$$

For $i \geq 0$ we let

$$D_{l,i} = \frac{1}{X^{a^{(i)}/(q-1)} \prod_{j=0}^{l-1} \widehat{\theta_{n,l}}(Xv_j)^{p^j}} \circ X\frac{\mathrm{d}}{\mathrm{d}X} \circ X^{a^{(i)}/(q-1)} \prod_{j=0}^{l-1} \widehat{\theta_{n,l}}(Xv_j)^{p^j} \tag{4.3.2}$$

so that the commutativity relation $\alpha_{l,i} \circ D_{l,i} = D_{l,i+1} \circ p\alpha_{l,i}$ holds by (4.2.2). Since $a^{(0)} = a$, we see that $D_{l,0}$ matches our previous definition. Note that $D_{l,i}$ are $f$-periodic: $D_{l,i} = D_{l,i+f}$ for all $i \geq 0$.

As for $D_{l,0}$, using the product rule, we have

$$D_{l,i} = X\frac{\mathrm{d}}{\mathrm{d}X} + \frac{a^{(i)}}{q-1} + \sum_{j=0}^{l-1} p^j X f'(Xv_j),$$

where $f(X) = \sum_{j=0}^{p^{n-1}} L_j(\pi_{n,l})X^{p^j}$. Hence $D_{l,i}$ is well–defined as an operator on $\mathcal{L}(r)$ for finite $n$. It is also clear that $D_{l,i}$ generalize the differential operators $D_i$:

$$D_{1,i} = D_i = X\frac{\mathrm{d}}{\mathrm{d}X} + \frac{a^{(i)}}{q-1} + \pi X,$$

where $\pi = \pi_{1,1}$.

The relation $\alpha_{l,i} \circ D_{l,i} = D_{l,i+1} \circ p\alpha_{l,i}$ induces the linear map $\overline{\alpha_{l,i}}\colon \mathcal{L}(r)/D_i\mathcal{L}(r) \to \mathcal{L}(r)/D_{i+1}\mathcal{L}(r)$ and we have commutativity of the diagram below.

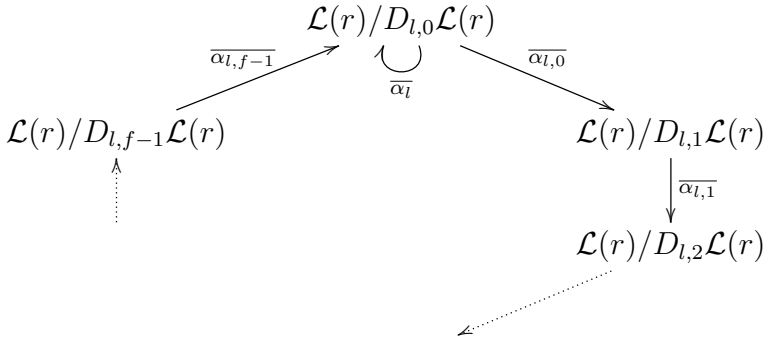
$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathcal{L}(r) & \xrightarrow{D_{l,i}} & \mathcal{L}(r) & \longrightarrow & \mathcal{L}(r)/D_{l,i}\mathcal{L}(r) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle p\alpha_{l,i}} & & \downarrow{\scriptstyle \alpha_{l,i}} & & \downarrow{\scriptstyle \overline{\alpha_{l,i}}} & & \\
0 & \longrightarrow & \mathcal{L}(r) & \xrightarrow{D_{l,i+1}} & \mathcal{L}(r) & \longrightarrow & \mathcal{L}(r)/D_{l,i+1}\mathcal{L}(r) & \longrightarrow & 0
\end{array}$$

Note that in this case $\mathrm{Tr}(\overline{\alpha_{l,i}})$ is not well-defined since $\overline{\alpha_{l,i}}$ is a linear map between

two *different* vector spaces. Nevertheless (4.3.1) reduces to

$$\overline{\alpha_l} = \overline{\alpha_{l,f-1}} \circ \overline{\alpha_{l,f-2}} \circ \cdots \circ \overline{\alpha_{l,1}} \circ \overline{\alpha_{l,0}},$$

i.e., the following diagram commutes.




We still need to show that when $n = \infty$, $D_{l,i}$ is a legitimate differential operator on $\mathcal{L}(r)$. For this purpose we need to show that $f'(X) \in \mathcal{L}(r)$. Hence we prove the following lemma which gives the $p$-adic absolute value of the coefficients of $f(X) = \sum_{j=0}^{p^{n-1}} L_j(\pi_{n,l}) X^{p^j}$.

**Lemma 4.3.2.** *For $n \gg l$ and $i \geq 0$ we have*

$$|L_i(\pi_{n.l})| = \begin{cases} \left| \frac{\pi_{n,l}^{p^i}}{p^i} \right| & \text{if } \ 0 \leq i \leq l-1 \\ \left| \frac{\pi_{n,l}^{p^{i+1}}}{p^{i+1}} \right| & \text{if } \ l \leq i < n \end{cases} = \begin{cases} \left(\frac{1}{p}\right)^{\frac{p^i}{p^{l-1}(p-1)} - i} & \text{if } \ i \leq l-1 \\ \left(\frac{1}{p}\right)^{\frac{p^{i+1}}{p^{l-1}(p-1)} - (i+1)} & \text{if } \ l \leq i < n \end{cases}.$$

*Proof.* For any $k \geq 0$ we have

$$\left| \frac{\pi_{n,l}^{p^k}}{p^k} \right| = \left(\frac{1}{p}\right)^{\frac{p^k}{p^{l-1}(p-1)} - k}.$$

Let $h(k) = \frac{p^k}{p^{l-1}(p-1)} - k$. Note that $h(k) - h(k-1) = \frac{1}{p^{l-k}} - 1 = p^{k-l} - 1$ and so

$$h(k) - h(k-1) = \begin{cases} < 0 & : k < l \\ = 0 & : k = l \\ > 0 & : k > l \end{cases}.$$

This implies that $h(k)$ is decreasing up to $k = l - 1$ and then increasing after $k = l$. Thus for $i < l$, by the strong triangle inequality we have

$$|L_i(\pi_{n,l})| = \left| \frac{\pi_{n,l}^{p^i}}{p^i} \right| = \left( \frac{1}{p} \right)^{\frac{p^i}{p^{l-1}(p-1)} - i},$$

whereas for $l \leq i < n$ we have

$$|L_i(\pi_{n,l})| = |L_i(\pi_{n,l}) - L_n(\pi_{n,l})| = \left| - \sum_{k=i+1}^{n} \frac{\pi_{n,l}^{p^i}}{p^i} \right| = \left| \frac{\pi_{n,l}^{p^{i+1}}}{p^{i+1}} \right| = \left( \frac{1}{p} \right)^{\frac{p^{i+1}}{p^{l-1}(p-1)} - (i+1)}.$$

The lemma follows. □

By Lemma 4.3.2, the radius of convergence of $f(X)$ when $n = \infty$ is

$$R(f(X)) = \liminf_{i \to \infty} \frac{1}{|L_i(\pi_{n,l})|^{1/p^i}} = \liminf_{i \to \infty} p^{1/(p^{l-2}(p-1)) - (i+1)/p^i} = p^{1/(p^{l-2}(p-1))},$$

which is precisely our upper bound for $r$ in (4.4.3) for $n = \infty$. When taking derivatives the radius of convergence does not change and thus $f(X), f'(X) \in \mathcal{L}(r)$ for all $r < p^{1/(p^{l-2}(p-1))}$ when $n = \infty$ and for any $r \geq 0$ when $n \neq \infty$ since in that case $f(X)$ is a polynomial.

This concludes the proof of Theorem 2.4.13, except for showing that the dimension

of $\mathcal{L}(r)/D_{l,i}\mathcal{L}(r)$ is $p^{l-1}$, which is the most challenging part of the proof.

## 4.4 Dimension of $\mathcal{L}(r)/D_{l,i}\mathcal{L}(r)$ for $v = 1$

In this section we find the dimension of $\mathcal{L}(r)/D_{l,i}\mathcal{L}(r)$, i.e., the dimension of the cokernel of $D_{l,i}$ over $K$ in $\mathcal{L}(r)$ for $v = 1$.

We start by finding the disc of convergence of $\widehat{\theta_{n,l}}(X)$ and its reciprocal.

**Lemma 4.4.1.** *For any $n \gg l$, the series $\widehat{\theta_{n,l}}(X)$ and its reciprocal $1/\widehat{\theta_{n,l}}(X)$ converge precisely on the open unit disc around the origin and therefore*

$$R(\widehat{\theta_{n,l}}) = 1.$$

*Proof.* From the infinite product representation of $\widehat{\theta}$ in (4.2.1), we have

$$\widehat{\theta_{n,l}}(X) = \theta_{n,l}(X)\widehat{\theta_{n,l}}(X^p).$$

It is easy to see that $R(\widehat{\theta_{n,l}}) > 0$ from expressing $\widehat{\theta_{n,l}}(X)$ as the exponential of a polynomial (or power series when $n = \infty$). Suppose that $R(\widehat{\theta_{n,l}}) < 1$. Recalling that when you replace $X$ by $X^p$ in a power series the radius of convergence gets closer to 1 (provided it does not have radius 0) and that $R(\theta_{n,l}) > 1$, we get a contradiction with

$$R(\widehat{\theta_{n,l}}) \geq \min\{R(\theta_{n,l}(X)), R(\widehat{\theta_{n,l}}(X^p))\} = R(\widehat{\theta_{n,l}}(X^p)).$$

Hence $R(\widehat{\theta_{n,l}}) \geq 1$. Now suppose that $\theta_{n,l}(X)$ converges on the closed unit disc around the origin and set $q = p^f$. Recall from the proof of Theorem 2.4.11, for any

$t \in \mu_{q-1} \subseteq \mathbb{C}_p$ we have the following representation of the additive character $\psi_l$:

$$\psi_l(t) = \zeta_{p^l}^{\mathrm{Tr}(t)} = \prod_{i=0}^{f-1} \theta_{n,l}(t^{p^i}), \tag{4.4.1}$$

where $\mathrm{Tr} : \mathbb{Z}_q \to \mathbb{Z}_p$ is the trace map. Thus from the infinite product representation of $\widehat{\theta_{n,l}}$, we have

$$\widehat{\theta_{n,l}}(t) = \zeta_{p^l}^{\mathrm{Tr}(t)} \widehat{\theta_{n,l}}(t^q) = \zeta_{p^l}^{\mathrm{Tr}(t)} \widehat{\theta_{n,l}}(t). \tag{4.4.2}$$

This implies that either $\mathrm{Tr}(t) \in p^l\mathbb{Z}_p$ or $\widehat{\theta_{n,l}}(t) = 0$. Note that there are $p^f - p^{f-1}$ elements $\alpha \in \mathbb{F}_q$ such that $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\alpha) = \alpha + \alpha^p + \cdots + \alpha^{p^{f-1}} \neq 0$. Hence, by the Teichmuller lifting there are at least $p^f - p^{f-1}$ elements $t \in \mu_{q-1}$ such that $\mathrm{Tr}(t) \notin p^l\mathbb{Z}_p$ and therefore $\widehat{\theta_{n,l}}(t) = 0$. Letting $f \to \infty$, we find infinitely many roots of $\widehat{\theta_{n,l}}$ in the closed unit disc around the origin, which contradicts the Weierstrass Preparation Theorem. Hence, $\widehat{\theta_{n,l}}$ converges precisely on the open unit disc around the origin as claimed. We could run through the same exact argument with the reciprocal series $1/\widehat{\theta_{n,l}}(X)$. For odd $p$, this also follows by the identity $\widehat{\theta_{n,l}}(-X) = 1/\widehat{\theta_{n,l}}(X)$ so obviously this series has the same radius and disc of convergence as $\widehat{\theta_{n,l}}(X)$. □

Since the differential operators $D_{l,i}$ only differ from each other by the constant $a^{(i)}/(q-1)$, fix $c \in [0,1) \cap \mathbb{Q} \cap \mathbb{Z}_p$ and $n \gg l$. Let

$$D = X\frac{d}{dX} + c + \pi_{n,l}X + L_1(\pi_{n,l})pX^p + L_2(\pi_{n,l})p^2X^{p^2} + \cdots + L_{n-1}(\pi_{n,l})p^{n-1}X^{p^{n-1}}.$$

Recall $f(X) = \sum_{j=0}^{p^{n-1}} L_j(\pi_{n,l})X^{p^j}$. Hence

$$D = X\frac{d}{dX} + c + Xf'(X).$$

For $v = 1$ and $c = a^{(i)}/(q-1)$, the differential operators $D$ and $D_{l,i}$ are the same. Based on the integrating factor method for solving first order differential equations or by a similar computation as for $D_{l,0}$, we may rewrite $D$ as

$$D = \frac{1}{X^c \widehat{\theta_{n,l}}(X)} \circ X \frac{d}{dX} \circ X^c \widehat{\theta_{n,l}}(X),$$

which holds over $K[[X]](X^c)$ (where $\frac{d}{dX} X^c := cX^{c-1}$ as usual). This can be also verified directly as we did for $D_{l,0}$. It is clear that $Dg = 0$ has as solutions the constant multiples of $1/(X^c \widehat{\theta_{n,l}}(X)) = X^{-c} \widehat{\theta_{n,l}}(X)^{-1}$ in the extension $K[[X]](X^c)$ of $K[[X]]$. In particular, it follows that for $c \neq 0$, $D$ is injective on $K[[X]]$. Since $\widehat{\theta_{n,l}}(X) \in 1 + XK[[X]]$ is a unit, the differential operator $D : K[[X]] \to K[[X]]$ is bijective for $c \neq 0$.

Lemma 4.4.1 implies that for $r \geq 1$, $D : \mathcal{L}(r) \to \mathcal{L}(r)$ is injective even for $c = 0$ since $\widehat{\theta_{n,l}}(X)^{-1} \notin \mathcal{L}(r)$. The main question is: for

$$1 < r < R(\theta_{n,l})^p = p^{(p+p^2+\cdots+p^{n-l+1}-n)/p^n}, \tag{4.4.3}$$

what is the dimension $d(r)$ of the cokernel of $D$ (i.e. $\mathcal{L}(r)/D\mathcal{L}(r)$) over $K$ ? To answer this question we will find $d(r)$ for all $r > 0$.

**Lemma 4.4.2.** *The dimension $d(r)$ of $\mathcal{L}(r)/D\mathcal{L}(r)$ over $K$ is the same regardless of $n$ such that $n \gg l$: let $n' > n \gg l$ and*

$$D' = \frac{1}{X^c \widehat{\theta_{n',l}}(X)} \circ X \frac{d}{dX} \circ X^c \widehat{\theta_{n',l}}(X).$$

*Multiplication by $\widehat{\theta_{n,l}}(X)/\widehat{\theta_{n',l}}(X)$ is an isomorphism $\mathcal{L}(r) \to \mathcal{L}(r)$ that induces an*

*isomorphism* $\mathcal{L}(r)/D\mathcal{L}(r) \to \mathcal{L}(r)/D'\mathcal{L}(r)$ *and thus*

$$\dim_K \left(\mathcal{L}(r)/D\mathcal{L}(r)\right) = \dim_K \left(\mathcal{L}(r)/D'\mathcal{L}(r)\right).$$

*Proof.* Since

$$\frac{\widehat{\theta_{n,l}}}{\widehat{\theta_{n',l}}} \circ D = D' \circ \frac{\widehat{\theta_{n,l}}}{\widehat{\theta_{n',l}}}$$

the lemma follows if we can show that the ratio $\widehat{\theta_{n,l}}(X)/\widehat{\theta_{n',l}}(X) \in 1 + XK[[X]]$ is a unit in $\mathcal{L}(r)$. Hence it is enough to show that $R(\widehat{\theta_{n,l}}/\widehat{\theta_{n',l}}) \geq R(\theta_{n,l})^p$.

Since

$$\widehat{\theta_{n,l}}(X)/\widehat{\theta_{n',l}}(X) = \exp\left(\sum_{s=0}^{n-1}(L_s(\pi_{n,l}) - L_s(\pi_{n',l}))X^{p^s} - \sum_{s=n}^{n'-1} L_s(\pi_{n',l})X^{p^s}\right),$$

we set $\delta_s = L_s(\pi_{n',l}) - L_s(\pi_{n,l})$ and get

$$R\left(\widehat{\theta_{n,l}}/\widehat{\theta_{n',l}}\right) \geq \min\left\{\min_{0\leq s<n}\{R\left(\exp\left(\delta_s X^{p^s}\right)\right)\}, \min_{n\leq s<n'}\{R\left(\exp\left(L_s(\pi_{n',l})X^{p^s}\right)\right)\right\}.$$

By Lemma 4.3.2, for $n \leq s < n'$ we have

$$R\left(\exp\left(L_s(\pi_{n',l})X^{p^s}\right)\right) < R(\theta_{s,l})^p \leq R(\theta_{n,l})^p.$$

Let $0 \leq s < l$. By Lemma 4.3.2 we have $|\delta_s| = |(\pi_{n',l}^{p^s} - \pi_{n,l}^{p^s})/p^s|$. Theorem B.10 implies $|\pi_{n',l} - \pi_{n,l}| = r/p^\varepsilon \leq r_l^{p^l} = r_1/p$, where $\varepsilon = p + p^2 + \cdots + p^{n-l+1} - n \geq 1$ by (B.3). Hence $|\pi_{n',l}/\pi_{n,l} - 1| \leq r_1/(p^\varepsilon r_l) \leq r_1$ and by Lemma 3.4.9 we have

$$\left|\left(\frac{\pi_{n',l}}{\pi_{n,l}}\right)^{p^s} - 1\right| \leq \frac{r_1}{p^{s+\varepsilon} r_l}. \tag{4.4.4}$$

Therefore

$$|\delta_s| = |(\pi_{n',l}^{p^s} - \pi_{n,l}^{p^s})/p^s| \leq r_1 r_l^{p^s-1}/p^{\varepsilon}.$$

For $|x| < R(\theta_{n,l})^p = p^{\varepsilon/p^n}$ we have

$$|\delta_s x^{p^s}| < r_1 r_l^{p^s-1} p^{\varepsilon/p^{n-s}-\varepsilon} \leq r_1 r_l^{p^s-1} \leq r_1,$$

and thus for $0 \leq s < l$ we have

$$R\left(\exp\left(\delta_s X^{p^s}\right)\right) \geq R(\theta_{n,l})^p.$$

Finally, let $l \leq s < n$. By Lemma 4.3.2, we have $|\delta_s| = |(\pi_{n',l}^{p^{s+1}} - \pi_{n,l}^{p^{s+1}})/p^{s+1}|$. From (4.4.4) where we replace $s$ by $s+1$, we have

$$|\delta_s| = |(\pi_{n',l}^{p^{s+1}} - \pi_{n,l}^{p^{s+1}})/p^{s+1}| \leq r_1 r_l^{p^{s+1}-1}/p^{\varepsilon}.$$

For $|x| < R(\theta_{n,l})^p = p^{\varepsilon/p^n}$ we have

$$|\delta_s x^{p^s}| < r_1 r_l^{p^{s+1}-1} p^{\varepsilon/p^{n-s}-\varepsilon} \leq r_1 r_l^{p^{s+1}-1} \leq r_1,$$

and thus for $0 \leq s < n$ we have

$$R\left(\exp\left(\delta_s X^{p^s}\right)\right) \geq R(\theta_{n,l})^p,$$

as wanted. □

For $r < 1$, the series $\widehat{\theta_{n,l}}(X) \in \mathcal{L}(r)$ is a unit by Lemma 4.4.1, and using the same argument as before for formal power series in $K[[X]]$, we get that $D : \mathcal{L}(r) \to \mathcal{L}(r)$ is

bijective. Hence, $d(r) = 0$ for $r < 1$.

Let $r \geq 1$. In order to find $d(r)$, we will use a theorem of Robba, which requires finding the radius of convergence of the solution to the differential equation $Dg = 0$ for $g(X)$ a power series centered around a "generic point" $t$ with $|t| = r \geq 1$ and $r \in |K^\times|$ (see [14, p. 201] for the precise definition). First, define a norm on the polynomials $K[t]$ that extends the absolute value of $K$ by

$$\left| \sum_{k \geq 0} a_k t^k \right| = \sup_{k \geq 0} |a_k| r^k.$$

The completion $K_t$ of $K[t]$ with respect to this norm is the space of power series in $K[[t]]$ converging on the closed disc of radius $r$. Let $g \in K_t[[X - t]]$ be a solution to $Dg = 0$.

**Theorem 4.4.3** (Robba–Young)**.** *Let $\mathbb{F}$ be a complete extension of $\mathbb{Q}_p$. Fix $h(X) \in \mathbb{F}[X]$ nonconst. and let $D = X\frac{d}{dX} + h(X)$ be a differential operator. For $r > 0$ let $t$ be a "generic point" with $|t|_p := r$ and set $\rho(r)$ to be the radius of convergence of a nonzero solution to $DY = 0$ in $\widehat{\mathbb{F}(t)}[[X - t]]$; $\rho(r)$ is defined and continuous.*

*(i) (Robba) For $r > 0$ such that $\rho(r) < r$, $d(r) := \dim_{\mathbb{F}}(\mathcal{L}_{\mathbb{F}}(r)/D\mathcal{L}_{\mathbb{F}}(r))$ is finite and $\rho(r)$ is given by*

$$\rho(r) = \frac{C(r)}{r^{d(r)-1}},$$

*where $C(r)$ is a piecewise constant function of $r$.*

*(ii) (Young) If $r > 0$ is big enough so that $|h(t)| > 1$, then*

$$\rho(r) = \left(\frac{1}{p}\right)^{\frac{1}{p-1}} \frac{r}{|h(t)|} < r.$$

*Hence Robba's theorem applies and* $d(r) = \#\{x \in \mathbb{F} : h(x) = 0, |x| \leq r\}$ *with multiplicity.*

*Proof.* See [14, p. 201] and [18, Theorem 3.1, p. 16]. Robba lets $K$ be algebraically closed and complete, but the proof goes through also for $K$ a finite extension of $\mathbb{Q}_q(\zeta_{p^l})$, which is what we need. The space $H_0(r^+)$ in Robba's work is the same as $\mathcal{L}(r)$. The conclusion that $C(r)$ is a piecewise constant function of $r$ follows from Robba's proof. □

Through explicit calculations Lang and Dwork showed that for $l = 1$ the spaces $\mathcal{L}(r)/D_i\mathcal{L}(r) = \mathcal{L}(r)/D_{1,i}\mathcal{L}(r)$ are all 1-dimensional with basis $\bar{1}$. The dimension, but not the basis, can also be calculated from the above theorem of Robba–Young.

For ease of notation, we set $Y = X - t$ and let $f(X) = \sum_{i=0}^{p^{n-1}} L_i(\pi_{n,l})X^{p^i}$ so that

$$\widehat{\theta_{n,l}}(X) = \exp(f(X)) \text{ and } D = X\frac{d}{dX} + c + f'(X).$$

Applying the integrating factor method, a solution to $Dg = 0$ in $K_t[[Y]]$ is

$$g(Y) = \left(1 + \frac{Y}{t}\right)^{-c} \exp(f(t) - f(Y + t)).$$

Since $c \in \mathbb{Z}_p$ and $\binom{c}{k} \in \mathbb{Z}_p$ for all $k \geq 0$, we get

$$R\left(\left(1 + \frac{Y}{t}\right)^c\right) \geq r.$$

When $c = 0$, we have $R\left(\left(1 + \frac{Y}{t}\right)^c\right) = \infty$. When $c \neq 0$ on the other hand, since $\binom{c}{k}$ does not tend to 0 $p$-adically we have $R\left(\left(1 + \frac{Y}{t}\right)^c\right) = r$.

Let $r < 1$ and $c = 0$. Recall that $g_0(Y) := \exp(f(Y)) = 1/\theta_{n,l}(Y)$ converges

precisely on the open unit disc around 0 by Lemma 4.4.1. Hence $|t| = r < R(g_0) = 1$ and $g(Y) = \exp(f(t) - f(Y+t))$ converges for all $|y| < 1$ since $|y+t| < 1 = R(g_0)$ and $g_t(y) = \exp(f(t))/\exp(f(y+t)) = g_0(t)/g_0(y+t)$. If $g(Y)$ converges on the closed unit disc, then $g_0(Y) = g_0(t)/g(Y-t)$ converges on the closed unit disc as well since for $|y| = 1$ we have $|y+t| = 1$ and $g(Y-t)$ converges. This contradicts Lemma 4.4.1 since $g_0(Y)$ converges precisely on the open unit disc. Therefore $\rho(r) = R(g(Y)) = 1$ for all $r < 1$ when $c = 0$. Note that the conclusion of Theorem 4.4.3 implies $d(r) = 1$, which contradicts the fact that $d(r) = 0$ by our previous argument. However, all is well since Theorem 4.4.3 does not apply when $\rho(r) \geq r$.

Now let $r < 1$ and $c \neq 0$. Hence $|t| = r < 1$ and $g(Y) = \left(1 + \frac{Y}{t}\right)^{-c} \exp(f(t) - f(Y+t))$. We know that $R\left(\left(1 + \frac{Y}{t}\right)^{c}\right) = r$ and $R(\exp(f(t) - f(Y+t))) = 1$ by our previous computations and thus $R(g(Y)) = \min\{r, 1\} = r$ for all $r < 1$ when $c \neq 0$. The conclusion of Theorem 4.4.3 implies $d(r) = 0$, which matches our previous computations even though we cannot get this conclusion through Theorem 4.4.3 since the assumption $\rho(r) < r$ is not satisfied.

We now fix $r > 1$ and pull out a factor of $\exp(-f(Y)) = \widehat{\theta_{n,l}}(Y)^{-1}$ from the second factor of $g(Y)$, which by Lemma 4.4.1 has radius of convergence

$$R\left(\widehat{\theta_{n,l}}\right) = 1 \leq r.$$

In order to get the radius of convergence $\rho(r)$ of $g(Y)$, we need to find the radius of convergence of

$$\exp(f(Y) + f(t) - f(Y+t)) = \prod_{i=1}^{p^{n-1}} \prod_{k=1}^{p^i - 1} \exp\left(-L_i(\pi_{n,l}) \binom{p^i}{k} t^{p^i - k} Y^k\right).$$

Our goal now is to find the minimum radius of convergence of each piece of the product decomposition of $g(Y)$ and then make sure that this minimum is unique. We know that

$$\left|\binom{p^i}{k}\right| = \left(\frac{1}{p}\right)^{i-\mathrm{ord}_p(k)}.$$

Since the radius of convergence of the exponential is $r_1 = (1/p)^{1/(p-1)}$, we get

$$\left|L_i(\pi_{n,l})\binom{p^i}{k}t^{p^i-k}Y^k\right| < \left(\frac{1}{p}\right)^{\frac{1}{p-1}},$$

which implies

$$|Y| < \begin{cases} \left(\frac{1}{p}\right)^{\left(\frac{1}{p-1}-\frac{p^i}{p^{l-1}(p-1)}+\mathrm{ord}_p(k)\right)/k}\frac{1}{r^{p^i/k-1}} & \text{if } \ 1 \le i \le l-1 \\ \left(\frac{1}{p}\right)^{\left(\frac{1}{p-1}-\frac{p^{i+1}}{p^{l-1}(p-1)}+\mathrm{ord}_p(k)+1\right)/k}\frac{1}{r^{p^i/k-1}} & \text{if } \ l \le i < n \end{cases}.$$

Let $k = p^s m$ with $m$ not divisible by $p$. If we want the minimum radius above, we only need to focus on the cases where $m = 1$ and $0 \le s \le i-1$. So we are left to consider the minimum $R_{\min}$ of

$$R_{i,s} = \begin{cases} \left(\frac{1}{p}\right)^{\left(\frac{1}{p-1}-\frac{p^i}{p^{l-1}(p-1)}+s\right)/p^s}\frac{1}{r^{p^{i-s}-1}} & \text{if } \ 1 \le i \le l-1 \\ \left(\frac{1}{p}\right)^{\left(\frac{1}{p-1}-\frac{p^{i+1}}{p^{l-1}(p-1)}+s+1\right)/p^s}\frac{1}{r^{p^{i-s}-1}} & \text{if } \ l \le i < n \end{cases}$$

over all $1 \le i \le n-1$ and $0 \le s \le i-1$. For $r > 1$, $R_{\min} < 1 < r$ and so the radius of convergence of $g(Y)$ is $\rho(r) = R_{\min}$ provided that this minimum is unique (and that $R_{\min}$ is not an empty definition as in the first example below). In this case, Robba's theorem says that the dimension $d(r)$ of $\mathcal{L}(r)/D\mathcal{L}(r)$ is $p^{i-s}$ where $i$ and $s$ are the unique numbers that make $R_{\min} = R_{i,s}$. Let's compute some examples.

**Example 4.4.4.** Take $l = n = 1$. Then there's no $R_{i,s}$ to consider and the radius of convergence of $g(Y)$ is 1. Hence, for $r > 1$ we get by Robba's theorem that $d(r) = 1$, which matches Lang's computations in [11, Lemma 1.1, p. 333] (the condition $\delta \geq 1/(p-1)$ may be removed).

**Example 4.4.5.** Take $l = n = 2$ and $p > 2$. Then we need to consider $R_{i,s}$ for $i = 1$ and $s = 0$. Hence, $R_{1,0} = 1/r^{p-1} < 1$ and so the radius of convergence of $g(Y)$ is $1/r^{p-1}$. Hence, for $r > 1$ we get by Theorem 4.4.3 that $d(r) = p$.

**Example 4.4.6.** Take $l = 2$, $n = 3$ and $p = 2$. Then we need to consider $R_{i,s}$ for $i = 1$, $s = 0$ and $i = 2$, $s = 0, 1$ respectively. Hence, $R_{1,0} = 1/r^{p-1} < 1$; $R_{2,0} = p^p/r^{p^2-1}$, $R_{2,1} = p^{p-1}/r^{p-1}$. So the radius of convergence of $g(Y)$ is $1/r^{p-1}$ for $1 < r < p^{1/(p-1)}$ and $p^p/r^{p^2-1}$ for $r > p^{1/(p-1)}$. Hence, by Theorem 4.4.3, we get that for $1 < r < p^{1/(p-1)}$, $d(r) = p$ and for $r > p^{1/(p-1)}$, $d(r) = p^2$.

Continuing in this manner, we fill out the following table. The question marks mean that we do not know the exact answer since there are two radii $R_{i,s}$ equal to each other.

**TABLE 1.** Generic radii of convergence of $g(Y)$ for $r > 1$ and corresponding dimensions $d(r)$.

| $l$ | $n$ | $p$ | $i$ | $s$ | $R_{i,s}$ | $r$ | $R(g)$ | $d(r)$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | any | N/A | N/A | N/A | $r > 1$ | 1 | 1 |
| 2 | 3 | any | 1 | 0 | $1/r^{p-1}$ | $1 < r < p^{1/(p-1)}$ | $1/r^{p-1}$ | $p$ |
| | | | 2 | 0 | $p^p/r^{p^2-1}$ | $r > p^{1/(p-1)}$ | $p^p/r^{p^2-1}$ | $p^2$ |
| | | | | 1 | $p^{(p-1)/p}/r^{p-1}$ | | | |
| 3 | 4 | any | 1 | 0 | $p^{-1/p}/r^{p-1}$ | $1 < r < p^{1/p^2(p-1)}$ | $p^{-1/p}/r^{p-1}$ | $p$ |
| | | | 2 | 0 | $1/r^{p^2-1}$ | $p^{1/p^2(p-1)} < r < p^{1/p(p-1)}$ | $1/r^{p^2-1}$ | $p^2$ |
| | | | | 1 | $p^{-1/p}/r^{p-1}$ | $r > p^{1/p(p-1)}$ | $p^p/r^{p^3-1}$ | $p^3$ |
| | | | 3 | 0 | $p^p/r^{p^3-1}$ | | | |
| | | | | 1 | $p^{(p-1)/p}/r^{p^2-1}$ | | | |
| | | | | 2 | $p^{(p-2)/p^2}/r^{p-1}$ | | | |
| 4 | 5 | any | 1 | 0 | $p^{-(p+1)/p^2}/r^{p-1}$ | $1 < r < p^{1/p^3(p-1)}$ | $p^{-(p+1)/p^2}/r^{p-1}$ | $p$ |
| | | | 2 | 0 | $p^{-1/p}/r^{p^2-1}$ | $p^{1/p^3(p-1)} < r < p^{1/p^2(p-1)}$ | $1/r^{p^3-1}$ | $p^3$ |
| | | | | 1 | $p^{-(p+1)/p^2}/r^{p-1}$ | $r > p^{1/p^2(p-1)}$ | $p^p/r^{p^4-1}$ | $p^4$ |
| | | | 3 | 0 | $1/r^{p^3-1}$ | | | |
| | | | | 1 | $p^{-1/p}/r^{p^2-1}$ | | | |
| | | | | 2 | $p^{-2/p^2}/r^{p-1}$ | | | |
| | | | 4 | 0 | $p^p/r^{p^4-1}$ | | | |
| | | | | 1 | $p^{(p-1)/p}/r^{p^3-1}$ | | | |
| | | | | 2 | $p^{(p-2)/p^2}/r^{p^2-1}$ | | | |
| | | | | 3 | $p^{(p-3)/p^3}/r^{p-1}$ | | | |
| 5 | 7 | any | 1 | 0 | $p^{-(p^2+p+1)/p^3}/r^{p-1}$ | $1 < r < p^{1/p^4(p-1)}$ | $p^{-(p^2+p+1)/p^3}/r^{p-1}$ | $p$ |
| | | | 2 | 0 | $p^{-(p+1)/p^2}/r^{p^2-1}$ | $p^{1/p^4(p-1)} < r < p^{1/p^3(p-1)}$ | $1/r^{p^4-1}$ | $p^4$ |
| | | | | 1 | $p^{-(p^2+p+1)/p^3}/r^{p-1}$ | $r > p^{1/p^3(p-1)}$ | $p^{p^2+p}/r^{p^6-1}$ | $p^6$ |
| | | | 3 | 0 | $p^{-1/p}/r^{p^3-1}$ | | | |
| | | | | 1 | $p^{-(p+1)/p^2}/r^{p^2-1}$ | | | |
| | | | | 2 | $p^{-(2p+1)/p^3}/r^{p-1}$ | | | |
| | | | 4 | 0 | $1/r^{p^4-1}$ | | | |
| | | | | 1 | $p^{-1/p}/r^{p^3-1}$ | | | |
| | | | | 2 | $p^{-2/p^2}/r^{p^2-1}$ | | | |
| | | | | 3 | $p^{-3/p^3}/r^{p-1}$ | | | |
| | | | 5 | 0 | $p^p/r^{p^5-1}$ | | | |
| | | | | 1 | $p^{(p-1)/p}/r^{p^4-1}$ | | | |
| | | | | 2 | $p^{(p-2)/p^2}/r^{p^3-1}$ | | | |
| | | | | 3 | $p^{(p-3)/p^3}/r^{p^2-1}$ | | | |
| | | | | 4 | $p^{(p-4)/p^4}/r^{p-1}$ | | | |
| | | | 6 | 0 | $p^{p^2+p}/r^{p^6-1}$ | | | |
| | | | | 1 | $p^{(p^2+p-1)/p}/r^{p^5-1}$ | | | |
| | | | | 2 | $p^{(p^2+p-2)/p^2}/r^{p^4-1}$ | | | |
| | | | | 3 | $p^{(p^2+p-3)/p^3}/r^{p^3-1}$ | | | |
| | | | | 4 | $p^{(p^2+p-4)/p^4}/r^{p^2-1}$ | | | |
| | | | | 5 | $p^{(p^2+p-5)/p^5}/r^{p-1}$ | | | |

Part of the patterns in the table hold in general. Part (ii) of Theorem 4.4.3 tells us the exact radius of convergence is

$$\rho(r) = (1/p)^{1/(p-1)} r/|c + tf'(t)|$$

provided $|c + tf'(t)| > 1$.

Case 1: $r > p^{1/(p^{l-2}(p-1))}$. Then $|c + tf'(t)| = (1/p)^{p^{n-l+1}/(p-1)-1} r^{p^{n-1}} > 1$ and so $g(Y)$ has radius of convergence $(1/p)^{1/(p-1)-p^{n-l+1}/(p-1)+1}/r^{p^{n-1}-1} < r$. This means that $d(r) = p^{n-1}$ in this case.

Case 2: $p^{1/(p^{l-1}(p-1))} < r < p^{1/(p^{l-2}(p-1))}$. Then $|c + tf'(t)| = (1/p)^{1/(p-1)} r^{p^{l-1}} > 1$ and so $g(Y)$ has radius of convergence $1/r^{p^{l-1}-1} < r$. This means that $d(r) = p^{l-1}$ in this case.

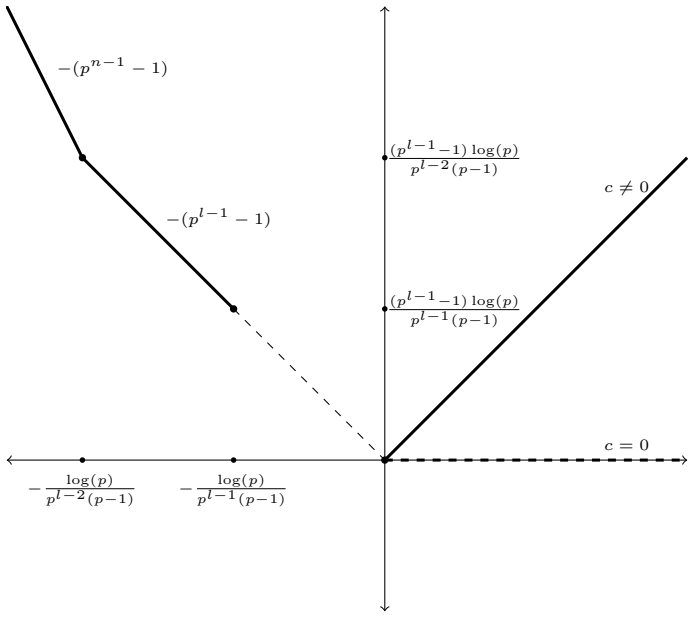
Case 3: $1 < r \leq p^{1/(p^{l-1}(p-1))}$. Then $|c + tf'(t)| = (1/p)^{1/p^{l-1}(p-1)} r < 1$ and so [18, Theorem 3.1, p. 16] only tells us that $g(Y)$ has radius of convergence *at least* $(1/p)^{1/(p-1)} r$. However, we claim that $\rho(r) = 1/r^{p^{l-1}-1} < r$ and $d(r) = p^{l-1}$ as in Case 2. To show this we use the following theorem.

**Theorem 4.4.7** (Kedlaya)**.** *The function $y = -\log(\rho(e^{-x}))$ is piecewise linear, continuous and convex.*

*Proof.* See [10, Theorem 11.3.2]. □

The convexity property of the function $y = -\log(\rho(e^{-x}))$ is most useful to us: For $r = e^{-x} < 1$ we have $x > 0$ and since $\rho(r) = 1$ when $c = 0$ and $\rho(r) = r$ when $c \neq 0$, we

get $y = 0$ and $y = x$ respectively. Doing the same calculation, in Case 1 and 2, we get respectively $y = -(p^{n-1}-1)x-(p^{n-l+1}-p)\log(p)/(p-1)$ for $x < -\log(p)/(p^{l-2}(p-1))$ and $y = -(p^{l-1}-1)x$ for $-\log(p)/(p^{l-2}(p-1)) < x < -\log(p)/(p^{l-1}(p-1))$. Since $y = -\log(\rho(e^{-x}))$ must be convex, we must have $\rho(r) = 1/r^{p^{l-1}-1}$ over the whole interval $1 < r < p^{1/p^{l-2}(p-1)}$ as can be seen dashed in the following graph.




**Remark 4.4.8.** We explain why Theorem 2.4.13 requires $r \in |K^\times|$ such that (2.4.4) holds. First note that for any $n \gg l$ we have

$$1 < r < R(\theta_{n,l})^p = p^{\frac{p+p^2+\cdots+p^{n-l+1}-n}{p^n}} \leq p^{1/p^{l-2}(p-1)}.$$

This falls into Case 3 above, where $d(r) = p^{l-1}$ and thus we have a complete proof of Theorem 2.4.13. The reason we restrict to $r > 1$ in (2.4.4) rather than $r \geq 1$ as needed in Theorem 4.1.1 is Theorem 4.4.3, which only applies when $\rho(r) < r$: when $r = 1$, $\rho(r) = 1$ by continuity, which contradicts the assumption $\rho(r) < r$. In fact

we can construct an explicit example in the classical case $l = n = v = 1$, where the quotient space $\mathcal{L}(1)/D_{0,l}\mathcal{L}(1)$ is infinite–dimensional over $K$. Hence $r > 1$ is a necessary condition. On the other hand, the upper bound in (2.4.4) is needed to make sure $\alpha_l \colon \mathcal{L}(r) \to \mathcal{L}(r)$ is well–defined.

## 4.5 Dimension of $\mathcal{L}(r)/D_{l,i}\mathcal{L}(r)$ for any $v$

For $v \in \mathbb{Z}_q^\times$ in standard form mod $p^l$, we follow the same setup as in the previous section and by part (ii) of Theorem 4.4.3, the exact radius of convergence is

$$\rho(r) = (1/p)^{1/(p-1)} r / \left| c + \sum_{j=0}^{l-1} p^j t f'(t) \right|$$

if $|c + \sum_{j=0}^{l-1} p^j t f'(t)| > 1$.

Analog calculations show that the dimensions are unaffected by $v \in \mathbb{Z}_q^\times$:

Case 1: If $r > p^{1/p^{l-2}(p-1)}$, then $d(r) = p^{n-1}$.

Case 2: If $p^{1/p^{l-1}(p-1)} < r < p^{1/p^{l-2}(p-1)}$, then $d(r) = p^{l-1}$.

Case 3: If $1 < r \leq p^{1/p^{l-1}(p-1)}$, then Theorem 4.4.7 implies $d(r) = p^{l-1}$.

We need $r \in |K^\times|$ such that

$$1 < r < R(\theta_{n,l})^p = p^{\frac{p+p^2+\cdots+p^{n-l+1}-n}{p^n}} \leq p^{1/p^{l-2}(p-1)}.$$

This falls into the range where $d(r) = p^{l-1}$ (Case 2 and 3 above) as claimed. Hence we have $d(r) = p^{l-1}$ for any $v \in \mathbb{Z}_q^\times$ in standard form mod $p^{l-1}$, which completes the proof of Theorem 2.4.13.

# Appendix A

# Roots of unity over $\mathbb{Q}_p$ and the Artin–Hasse series

Roots of unity in $\mathbb{C}$ arise as values of the exponential function at $2\pi i\mathbb{Q}$. The $p$-adic exponential series, which converges on the disc $D_p = \{x \in \mathbb{C}_p : |x| < (1/p)^{1/(p-1)}\}$ and takes values on $1 + D_p$, has no root of unity as a value other than $\exp(0) = 1$ since all other roots of unity lie outside $1 + D_p$. We will see in this appendix that the Artin–Hasse exponential, which converges on a larger domain, the open unit disc in $\mathbb{C}_p$, takes on all $p^{\text{th}}$–power roots of unity among its values. Also, in the same spirit as $\left\{\zeta_{p^l} := e^{\frac{2\pi i}{p^l}}\right\}_{l\geq 0}$ being a compatible family of $p^{\text{th}}$–power roots of unity in $\mathbb{C}$, in the sense that $\zeta_{p^l}^p = \zeta_{p^{l-1}}$ for all $l \geq 1$, we will use the Artin–Hasse exponential to get a compatible family $\{\zeta_{p^l}\}_{l\geq 0}$ of $p^{\text{th}}$ power roots of unity in $\mathbb{C}_p$. Just as the representation of roots of unity in $\mathbb{C}$ through the exponential function is quite useful, we will use the representation of $p^{\text{th}}$-power roots of unity in $\mathbb{C}_p$ through the Artin–Hasse exponential to get the analogue of Stickelberger's theorem for generalized Gauss sums (see Theorem 2.4.11).

We begin with some notation and definitions. For any power series $f(X)$ in $\mathbb{C}_p[[x]]$, denote by $R(f)$ its $p$-adic radius of convergence, $f_k$ its $k^{\text{th}}$ coefficient, and set $|f(X)|_r := \sup_k\{|f_k|r^k\}$ for any $r \geq 0$.

**Definition A.1.** The *Artin–Hasse logarithm series* is

$$L(X) = \sum_{n=0}^{\infty} \frac{X^{p^n}}{p^n} = X + \frac{X^p}{p} + \frac{X^{p^2}}{p^2} + \frac{X^{p^3}}{p^3} + \cdots,$$

and the *Artin–Hasse exponential series* is

$$\mathrm{AH}(X) := \exp(L(X)) = \sum_{n\geq 0} A_n X^n = 1 + X + \cdots.$$

The series $L(X)$ is like $-\log(1-X) = \sum\limits_{n=1}^{\infty} X^n/n$ where we drop all $n$ that are not powers of $p$.

**Lemma A.2** (Dwork)**.** *Let $p$ be a prime number and $F(X) = \sum a_i X^i \in 1+X\mathbb{Q}_p[[X]]$. Then $F(X) \in 1 + X\mathbb{Z}_p[[X]]$ if and only if $\frac{F(X^p)}{F(X)^p} \in 1 + pX\mathbb{Z}_p[[X]]$.*

*Proof.* See [15, p. 392]. □

From Dwork's lemma, it is not hard to see that $\mathrm{AH}(X) \in 1 + X\mathbb{Z}_p[[X]]$ using the following argument. By the properties of the exponential function:

$$\mathrm{AH}(X^p) = \exp\left(X^p + \frac{X^{p^2}}{p} + \frac{X^{p^3}}{p^2} + \cdots\right)$$

and

$$\mathrm{AH}(X)^p = \exp\left(pX + X^p + \frac{X^{p^2}}{p} + \frac{X^{p^3}}{p^2} + \cdots\right).$$

Therefore, we get

$$\frac{\mathrm{AH}(X^p)}{\mathrm{AH}(X)^p} = e^{-pX} = \sum_{n\geq 0} \frac{(-pX)^n}{n!} = 1 - pX + \frac{p^2}{2}X^2 - \frac{p^3}{3!}X^3 + \cdots .$$

The coefficients $\frac{(-p)^n}{n!}$ are in $p\mathbb{Z}_p$ for $n \geq 1$ because $\mathrm{ord}_p(n!) = \frac{n-S(n)}{p-1} \leq \frac{n-1}{p-1} < n$ where $S(n)$ is the sum of the digits of $n$ in base $p$. Hence by Dwork's lemma, we immediately get that the coefficients of the Artin–Hasse series are in $\mathbb{Z}_p$, so the radius of convergence $R(\mathrm{AH})$ of $\mathrm{AH}(X)$ is at least 1. In [15, p. 388] it is shown that in fact $R(\mathrm{AH}(X)) = 1$ and $\mathrm{AH}(X)$ converges precisely on $\mathfrak{m}_p = \{x \in \mathbb{C}_p : |x| < 1\}$, the unit open ball around the origin in $\mathbb{C}_p$. Since $\mathrm{AH}(X) \in 1 + X + X^2\mathbb{Z}_p[[X]]$, we have $|\mathrm{AH}(x) - 1| = |x|$ for all $x \in \mathfrak{m}_p$ and $\mathrm{AH}(\mathfrak{m}_p) \subset 1 + \mathfrak{m}_p$.

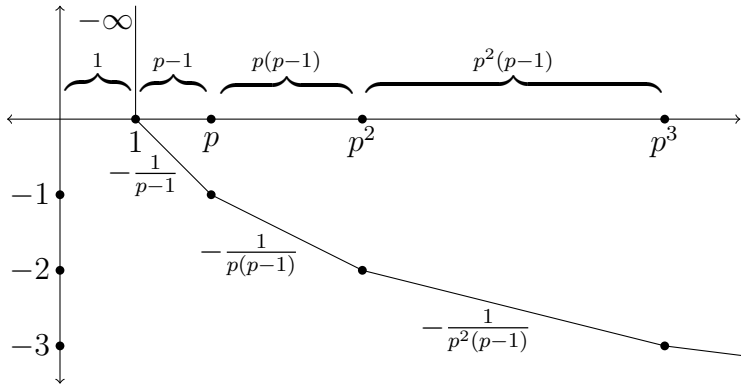
Now let's focus on the power series $L(X)$. It is not hard to show that $L(X)$ converges precisely on $\mathfrak{m}_p$, which strictly contains the disc of convergence $D_p$ of $\exp(X)$. Set $r_0 = 0$, $r := r_1 := (1/p)^{1/(p-1)} < 1$, $r_l := (1/p)^{1/p^{l-1}(p-1)}$, $D_p := \{x \in \mathbb{C}_p : |x| < r = r_1\}$ and $D_{p^l} := \{x \in \mathbb{C}_p : |x| < r_l\}$ for all $l \geq 1$. Note that

$$0 < r_1 < r_2 < \cdots < r_l < \cdots < 1,\ r_l \to 1,$$

and consequently

$$\{0\} \subset D_p \subset D_{p^2} \subset \cdots \subset D_{p^l} \subset \cdots \subset \mathfrak{m}_p.$$

**Theorem A.3.** *In $\mathfrak{m}_p$, we have $L(X)$ has $p^{l-1}(p-1)$ distinct zeros with absolute value $r_l$ for $l \geq 1$ lying in a finite extension of $\mathbb{Q}_p$ and no other zeros besides 0. In particular, the zeros of $L(X)$ other than 0 are algebraic numbers in $\mathfrak{m}_p - D_p$.*




*Proof.* We will use the theory of Newton polygons. For each non-negative integer $l$, $1/p^l$ is the coefficient of $X^{p^l}$ in $L(X)$ and all other coefficients are 0. Also, $\mathrm{ord}_p(1/p^l) = -l$, so the Newton polygon of $L(X)$ has break points $(1, 0)$, $(p, -1)$, $(p^2, -2)$, $(p^3, -3), \ldots, (p^l, -l), \ldots$ in this order. The slopes of the Newton polygon are precisely

$$-\frac{1}{p-1} < -\frac{1}{p(p-1)} < -\frac{1}{p^2(p-1)} < \cdots < -\frac{1}{p^l(p-1)} < \ldots$$

with corresponding horizontal lengths

$$p-1,\ p(p-1),\ p^2(p-1),\ \ldots,\ p^l(p-1),\ \ldots\ .$$

The slopes are in increasing order tending towards 0 with the smallest one corresponding to absolute value $r$. In general, a slope $m$ of horizontal length $n$ tells us that the power series with that Newton polygon has $n$ zeros of $p$-adic absolute value $p^m$, counting multiplicity. Therefore $L(X)$ has precisely one root of multiplicity one at $x = 0$, $p - 1$ roots of absolute value $r_1 = r$, and in general $p^{l-1}(p - 1)$ roots of absolute value $r_l$ for $l \geq 1$ as summarized in the table below.

| Segment | $p^{\text{slope}}$ | Hor. length |
|---|---|---|
| 0 | 0 | 1 |
| 1 | $r_1$ | $p-1$ |
| 2 | $r_2$ | $p(p-1)$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $n$ | $r_n$ | $p^{n-1}(p-1)$ |

In addition, these zeros are simple since $L'(X) \in 1 + X\mathbb{Z}_p[[X]]$, so $|x| < 1 \Rightarrow |L'(x)| = 1$, and thus $L'(x) \neq 0$. □

A naive calculation suggests $\mathrm{AH}(\pi) = 1$ for every zero $\pi$ of $L(X)$:

$$\mathrm{AH}(\pi) = \exp(L(\pi)) = \exp(0) = 1.$$

However, this calculation is incorrect. Indeed, since $|\,\mathrm{AH}(x) - 1| = |x|$ for all $x \in \mathfrak{m}_p$, if $L(x) = 0$ and $x \neq 0$ in $\mathfrak{m}_p$ then $|\,\mathrm{AH}(x) - 1| = |x| \neq 0$, so $\mathrm{AH}(x) \neq 1$.

**Example A.4.** Let $p = 2$. By Theorem A.3, there exists a unique root $\pi$ of $L(X)$ with $|\pi| = r = \frac{1}{2}$. It turns out that $\pi = 2 + 2^3 + 2^5 + \cdots \in \mathbb{Z}_2$. Therefore we have

$$\mathrm{AH}(\pi) = 1 + \pi + O(\pi^2) \equiv 3 \bmod 4.$$

Hence, $\mathrm{AH}(\pi) \neq 1$ since $\mathrm{AH}(\pi) \not\equiv 1 \bmod 4$. What is $\mathrm{AH}(\pi)$ equal to?

It turns out that when $L(\pi) = 0$ and $\pi \neq 0$, $\mathrm{AH}(\pi)$ is a $p^{\text{th}}$–power root of unity other than 1. More precisely, we have the following theorem whose statement is facilitated the following definitions. Let $Z$ denote the set of zeros of $L(X)$ in $\mathfrak{m}_p$ and

$\mu_{p^\infty}$ be the $p^{\text{th}}$–power roots of unity in $\mathbb{C}_p$. Analogously, for $l \geq 0$, let

$$Z_{p^l} := \{\pi \in Z : |\pi| \leq r_l\},\ Z'_{p^l} := \{\pi \in Z : |\pi| = r_l\}$$

and

$$\mu_{p^l} := \{\zeta \in \mu_{p^\infty} : |\zeta - 1| \leq r_l\},\ \mu'_{p^l} := \{\zeta \in \mu_{p^\infty} : |\zeta - 1| = r_l\}.$$

Note that $\mu_{p^l}$ are the $p^l$–th roots of unity in $\mathbb{C}_p$ and $\mu'_{p^l}$ are the roots of unity of order $p^l$ in $\mathbb{C}_p$.

**Theorem A.5.** *For all $l \geq 1$, the maps*

$$\text{AH}\colon \mathfrak{m}_p \to 1 + \mathfrak{m}_p,\ \text{AH}\colon Z \to \mu_{p^\infty},\ \text{AH}\colon Z_{p^l} \to \mu_{p^l} \text{ and } \text{AH}\colon Z'_{p^l} \to \mu'_{p^l}$$

*are bijections. In addition, for any $\pi \in Z'_{p^l}$, $\text{AH}(\pi)$ is the unique $\zeta \in \mu_{p^\infty}$ such that*

$$\left|\zeta - \sum_{i \leq p^{l-1}} A_i \pi^i\right| < r_1 = r,$$

*where $A_i$ are the coefficients of $\text{AH}(X)$.*

To prove $\text{AH}(\pi)$ is a root of unity when $L(\pi) = 0$, we will use the following result that provides conditions under which substitution into a composition of power series is legitimate.

**Theorem A.6.** *Let $f(X)$ and $g(X)$ be power series in $\mathbb{C}_p[[X]]$. For $x \in \mathbb{C}_p$, if*

1) $g(0) = g_0 = 0$,

2) $|x| < R(g)$,

3) $|g_k x^k| < R(f)$ *for all* $k \geq 1$

*then we have* $(f \circ g)(x) = f(g(x))$*, i.e., we are allowed to substitute* $x$ *into the formal power series equation* $(f \circ g)(X) = f(g(X))$*.*

*Proof.* See [15, p. 294]. □

Using notation introduced at the start, for $|x| < R(g)$ the third condition in the theorem says $|g(X)|_{|x|} < R(f)$, and this implies $|g(x)| < R(f)$. In general, $|g(X)|_{|x|} < R(f)$ is a stronger condition than $|g(x)| < R(f)$.

Let's show that the equality $\mathrm{AH}(x) = \exp(L(x))$ holds for all $x \in D_p$ as an application of Theorem A.6. We know that $L(0) = 0$, $R(L) = 1$ and $R(\exp) = r$. Thus the first two conditions are satisfied. Checking the third condition, we have $|L_k x^k| \leq |x^p/p| = |L_1||x|^p$ for $k \geq 1$, and

$$|x^p/p| < pr^p = r = R(\exp).$$

Thus, Theorem A.6 implies $\mathrm{AH}(x) = \exp(L(x))$ when $x \in D_p$. In fact, this is the best we can do since there is an $x$ on the "boundary" of $D_p$ (*i.e.*, $|x| = r$) such that $\mathrm{AH}(x) \neq \exp(L(x))$: there are $p-1$ zeros of $L(X)$ with absolute value $r$ and Theorem A.5 says $\mathrm{AH}(x)$ is a nontrivial $p^{\text{th}}$ root of unity for such $x$, so $\mathrm{AH}(x) \neq 1 = \exp(L(x))$.

Let's see how the hypotheses of Theorem A.6 break down for the equation $\mathrm{AH}(X) = \exp(L(X))$ if $x$ is a zero of $L(X)$ other than 0. Set $|x| = r_l = (1/p)^{1/(p^{l-1}(p-1))}$ for $l$ a positive integer. We have $L(0) = 0$, $R(L) = 1$, and $R(\exp) = r$. Thus the first two conditions of Theorem A.6 are satisfied. However,

$$|L(X)|_{r_l} = |L_l| r_l^{p^l} = |p|^{-l+\frac{p^l}{p^{l-1}(p-1)}} = |p|^{-l+1+\frac{1}{p-1}} = p^l(r/p) \not< r = R(\exp), \qquad \text{(A.1)}$$

so the third condition is not satisfied.

Now we prove Theorem A.5.

*Proof.* Let $l \geq 0$.

Step 1: $\mathrm{AH}(Z) \subseteq \mu_{p^\infty}$ and $\mathrm{AH}(Z'_{p^l}) \subseteq \mu'_{p^l}$.

First, note that $\mathrm{AH}(0) = 1$. Let $\pi_l \in Z - \{0\}$ have $p$-adic absolute value $r_l$(i.e. $\pi_l \in Z'_{p^l}$) for some positive integer $l$. We will show that $\mathrm{AH}(\pi_l)$ is a root of unity with order $p^l$.

In (A.1) we saw that $|L(X)|_{r_l} = p^l(r/p)$. If we could get rid of the $p^l$ on the right side, we'd be left with $r/p$, which is less than $r = R(\exp)$, and we would then be able to make a substitution by Theorem A.6. To remove the $l$, consider the equation

$$\mathrm{AH}(X)^{p^l} = \exp(p^l L(X)). \tag{A.2}$$

We will show that substituting $X = \pi_l$ in this equation is allowed and preserves equality. Then we would get

$$\mathrm{AH}(\pi_l)^{p^l} = \exp(p^l L(\pi_l)) = \exp(0) = 1,$$

so $\mathrm{AH}(\pi_l)$ is a $p^l$-th root of unity. The first two conditions of Theorem A.6 are obviously true for (A.2). Let's check the third condition for (A.2):

$$|p^l L(X)|_{r_l} = |p|^l |L(X)|_{r_l} = \left(\frac{1}{p}\right)^l p^l(r/p) = r/p < r = R(\exp).$$

Hence setting $X = \pi_l$ in (A.2) is valid.

To argue that $\mathrm{AH}(\pi_l)$ has order precisely $p^l$, note that $|\,\mathrm{AH}(\pi_l) - 1| = |\pi_l| = r_l$. Since

we know that in general any $p$-adic root of unity $\zeta$ of order $p^m$ with $m \geq 1$ satisfies $|\zeta - 1| = r_m$, the claim follows.

Step 2: The mapping $\mathrm{AH}\colon \mathfrak{m}_p \to 1 + \mathfrak{m}_p$ is a bijection.

First, this map is well-defined since AH convergences on $\mathfrak{m}_p$. Since the difference $\mathrm{AH}(X) - 1 = X + ... \in X\mathbb{Z}_p[[X]]$ and its linear term has a coefficient in $\mathbb{Z}_p^\times$, the formal power series Inverse Function Theorem says this formal power series has a composition inverse $B(X) = X + .... \in X\mathbb{Z}_p[[X]]$. Now, it is easy to check the conditions of Theorem A.6 are satisfied and therefore the functions $B\colon \mathfrak{m}_p \to \mathfrak{m}_p$ and $\mathrm{AH}(X) - 1\colon \mathfrak{m}_p \to \mathfrak{m}_p$ are inverses of each other: $B(\mathrm{AH}(x) - 1) = x$ and $(\mathrm{AH} - 1)(B(x)) = \mathrm{AH}(B(x)) - 1 = x$ for all $x \in \mathfrak{m}_p$. Thus $\mathrm{AH}\colon \mathfrak{m}_p \to 1 + \mathfrak{m}_p$ is a bijection.

Step 3: The mappings $\mathrm{AH}\colon Z \to \mu_{p^\infty}$, $\mathrm{AH}\colon Z_{p^l} \to \mu_{p^l}$ and $\mathrm{AH}\colon Z'_{p^l} \to \mu'_{p^l}$ are bijections.

By step 1, we know that $\mathrm{AH}(Z'_{p^l}) \subseteq \mu'_{p^l}$ for all $l \geq 1$ and by step 2, the mapping $\mathrm{AH}\colon \mathfrak{m}_p \to 1 + \mathfrak{m}_p$ is a bijection. Thus, its restrictions $\mathrm{AH}\colon Z'_{p^l} \to \mu'_{p^l}$ and $\mathrm{AH}\colon Z_{p^l} \to \mu_{p^l}$ are bijections. In addition, since $Z = \{0\} \cup \bigcup_{l \geq 1} Z'_{p^l}$ and $\mu_{p^\infty} = \{1\} \cup \bigcup_{l \geq 1} \mu'_{p^l}$, we get that $\mathrm{AH}\colon Z \to \mu_{p^\infty}$ is also a bijection.

Step 4: If $\pi_l \in Z - \{0\}$ has absolute value $r_l$ and $\zeta \in \mu_{p^\infty}$ satisfies $|\zeta - \sum_{i \leq p^{l-1}} A_i \pi_l^i| < r$, where the coefficients $A_i$ come from $\mathrm{AH}(X) = \sum_{i \geq 0} A_i X^i$, then $\zeta = \mathrm{AH}(\pi_l)$.

Since the coefficients of $\mathrm{AH}(\pi_l)$ are in $\mathbb{Z}_p$,

$$\left| \mathrm{AH}(\pi_l) - \sum_{i \leq p^{l-1}} A_i \pi_l^i \right| \leq |\pi_l|^{p^{l-1}+1} < |\pi_l|^{p^{l-1}} = r_l^{p^{l-1}} = (1/p)^{1/(p-1)} = r,$$

so $|\,\mathrm{AH}(\pi_l) - \zeta| = |\,\mathrm{AH}(\pi_l) - S + S - \zeta| < r$, where $S = \sum_{i \leq p^{l-1}} A_i \pi_l^i$. Different $p^{\text{th}}$ power roots of unity have distance at least $r$ from each other, so $\zeta = \mathrm{AH}(\pi_l)$. $\square$

In the proof above, we showed that $\mathrm{AH}\colon \mathfrak{m}_p \to 1+\mathfrak{m}_p$ is a bijection. In fact, this map is also an isometry: For any $x, y \in \mathfrak{m}_p$, we have

$$|\,\mathrm{AH}(x) - \mathrm{AH}(y)| = \left|\sum_{i\geq 0} a_i(x^i - y^i)\right| = |x-y|\left|1 + \sum_{i\geq 2}\sum_{j=1}^{i} a_i x^{i-j}y^{j-1}\right| = |x-y|$$

since $|a_i x^{i-j}y^{j-1}| < 1$ for all $i \geq 2$ and $1 \leq j \leq i$.

Inspired by the proof of the above theorem and the fact that $\mathrm{AH}\colon \mathfrak{m}_p \to 1+\mathfrak{m}_p$ is an isometry, we wonder if given a root $\pi$ of $L$, how far away from it do we need to look to encounter another root of $L$. In general, it is reasonable to believe that the roots of $L$ are spread out precisely like the $p^{\text{th}}$ power roots of unity.

**Corollary A.7.** *Let $\pi$ be a root of $L$. Then, there are exactly $p^{l-1}(p-1)$ roots $\pi'$ of $L$ with $|\pi - \pi'| = r_l$ for all $l \geq 1$ and no other roots of $L$. In particular, if $\pi' \neq \pi$, then $|\pi' - \pi| \geq r$. Also, for any collection of $p$ roots of $L$ distinct from $\pi$, at least one of them, say $\pi'$, satisfies $|\pi - \pi'| \geq r_2$.*

*Proof.* Let $\pi$ be a root of $L$. By Theorem A.5, we know that $\mathrm{AH}(\pi) = \zeta$ for some $\zeta \in \mu_{p^\infty}$. There are precisely $p^{l-1}(p-1)$ roots of unity $\zeta'$ with $|\zeta - \zeta'| = |\zeta/\zeta' - 1| = r_l$ for any $l \geq 1$. Now, let $l \geq 1$ and $\zeta'$ as above. Again, by Theorem A.5, we know that $\zeta' = \mathrm{AH}(\pi')$ for a unique $\pi' \in Z_{p^\infty}$. Since $\mathrm{AH}\colon \mathfrak{m}_p \to 1+\mathfrak{m}_p$ is an isometry and $\zeta, \zeta' \in \mu_{p^\infty}$, we have

$$|\pi - \pi'| = |\,\mathrm{AH}(\pi) - \mathrm{AH}(\pi')| = |\zeta - \zeta'| = r_l.$$

The two remaining statements follow by a counting argument. □

It is possible to prove the above corollary without using Theorem A.5. Now we

present a second proof, which is longer, but quite interestingly takes advantage of the theory of Newton polygons. This technique is due to Blache in [4].

*Proof.* When $\pi = 0$, the closest root of $L$ is at a distance $r$ from it by the Newton polygon of $L$ we considered before. Now, let $\pi$ be a non-zero root of $L$. We know that $|\pi| \geq r$. We will prove our theorem by analyzing the Newton polygon of $H(X) := L(X + \pi)$. The constant term of this power series is

$$H_0 = H(0) = L(\pi) = 0.$$

For $k \geq 1$, using the binomial theorem and collecting like terms, we get

$$H_k = \sum_{i \geq 0, p^i \geq k} \frac{\binom{p^i}{k}}{p^i} \pi^{p^i - k}.$$

Since $\operatorname{ord}_p \binom{p^i}{k} = i - \operatorname{ord}_p(k)$ and $\operatorname{ord}_p(\pi^{p^{i+1}-k}) > \operatorname{ord}_p(\pi_l^{p^i-k}) \geq 0$ for all $i$ such that $p^i \geq k$, we have $\operatorname{ord}_p(H_k) \geq -\operatorname{ord}_p(k)$. Furthermore, equality holds if and only if $k$ is a $p^{\text{th}}$ power.

Thus, the Newton polygon of $H(X)$ is identical to the Newton polygon of $L(X)$. Hence, 0 is the unique root of $H$ with absolute value less than $r$. So, if $\pi'$ is a root of $L$ different than $\pi$, then $\alpha = \pi' - \pi$ is a root of $H$ and so $|\pi' - \pi| = r_l$ for some $l \geq 1$. In particular if $|\pi' - \pi| < r$, then $\alpha = \pi' - \pi$ is a root of $H$ with $|\alpha| < r$. But 0 is the unique such root of $H$. Therefore, $\alpha = 0$ and so $\pi' = \pi$. □

The $p^{\text{th}}$ power map sends elements of $\mu_{p^l}$ to elements of $\mu_{p^{l-1}}$ for all $l \geq 1$. How does this map translate to a corresponding map between roots of $L$, i.e. between $Z_{p^l}$ and $Z_{p^{l-1}}$? The following theorem provides an answer.

**Theorem A.8.** *Let $\pi_l$ be a zero of $L$ with $|\pi_l| = r_l$ for some integer $l \geq 1$. Then:*

*(i) There exists a unique root $\pi_{l-1}$ of $L$ with $|\pi_{l-1}| = r_{l-1}$ such that $\mathrm{AH}(\pi_l)^p = \mathrm{AH}(\pi_{l-1})$ and*

*(ii) $\pi_{l-1}$ is the unique root of $L$ such that $|\pi_{l-1} - \pi_l^p| < r_1 = r$. More precisely,*

$$|\pi_{l-1} - \pi_l^p| = r_l/p < 1/p \leq r.$$

*Proof.* To begin with, let $\pi_l$ be a zero of $L$ with $|\pi_l| = r_l$ for some integer $l \geq 2$. Then part $(i)$ of the theorem follows from Theorem A.5 since $\mathrm{AH}(\pi_l)^p$ is a root of unity of order $p^{l-1}$. In other words, there exists a unique root $\pi_{l-1}$ of $L$ with $|\pi_{l-1}| = r_{l-1}$ such that $\mathrm{AH}(\pi_l)^p = \mathrm{AH}(\pi_{l-1})$. Part $(ii)$ tells us how to produce $\pi_{l-1}$ directly from $\pi_l$: take its $p^{\text{th}}$ power and look closeby. We now focus on proving this. From the proof of the fact that $\mathrm{AH}(X)$ has $\mathbb{Z}_p$ coefficients, we have

$$\frac{\mathrm{AH}(X)^p}{\mathrm{AH}(X^p)} = \frac{\exp\left(pX + X^p + \frac{X^{p^2}}{p} + \frac{X^{p^3}}{p^2} + \cdots\right)}{\exp\left(X^p + \frac{X^{p^2}}{p} + \frac{X^{p^3}}{p^2} + \cdots\right)} = \exp(pX) = 1 + pX + \cdots$$

So it easily follows that

$$\mathrm{AH}(X)^p - \mathrm{AH}(X^p) = \mathrm{AH}(X^p)(\exp(pX) - 1).$$

What values can we plug in for $X$ in the above equality of power series?

We know that the Artin–Hasse exponential converges on $\mathfrak{m}_p$ and exp converges on $D_p$. Thus we may plug in any $x \in \mathfrak{m}_p$ such that $|px| < r$ or equivalently $|x| < pr$. However, $pr = p^{1-\frac{1}{p-1}} > 1$ and so we are able to plug in any $x \in \mathfrak{m}_p$, which is quite

convenient. Thus, for any $x \in \mathfrak{m}_p$, we have

$$|\operatorname{AH}(x)^p - \operatorname{AH}(x^p)| = |\operatorname{AH}(x^p)(\exp(px) - 1)| = |\exp(px) - 1| = |px|,$$

where the last equality follows from the fact that $\exp : D_p \to 1 + D_p$ is an isometry. Plugging in $x = \pi_l$, we get

$$|\operatorname{AH}(\pi_l)^p - \operatorname{AH}(\pi_l^p)| = |p\pi_l| = \frac{r_l}{p},$$

which we recognize as part of the statement of the theorem. Since $\operatorname{AH}\colon \mathfrak{m}_p \to 1 + \mathfrak{m}_p$ is an isometry and $\operatorname{AH}(\pi_l)^p = \operatorname{AH}(\pi_{l-1})$, we have

$$|\pi_{l-1} - \pi_l^p| = |\operatorname{AH}(\pi_{l-1}) - \operatorname{AH}(\pi_l^p)| = |\operatorname{AH}(\pi_l)^p - \operatorname{AH}(\pi_l^p)| = \frac{r_l}{p}.$$

As for the uniqueness, suppose $\pi \neq \pi_{l-1}$ is another root of $L$ such that $|\pi - \pi_l^p| < r$. From Corollary A.7, we know that $|\pi - \pi_{l-1}| \geq r$. Since $|\pi_{l-1} - \pi_l^p| < r$, by the strong triangle inequality, we have

$$|\pi - \pi_l^p| = |\pi - \pi_{l-1} + \pi_{l-1} - \pi_l^p| = |\pi - \pi_{l-1}| \geq r,$$

which is a contradiction and so the uniqueness follows.

□

Now that we know how to go down from $\pi_l$ to $\pi_{l-1}$, let's see how to get back to $\pi_l$. In $\mu_{p^\infty}$, we just take a $p^{\text{th}}$ root, for which there are precisely $p$ choices. A similar thing happens in $Z_{p^\infty}$.

**Theorem A.9.** *Let $\pi_{l-1} \in Z_{p^{l-1}}$ for some integer $l \geq 2$.*

*(i) There exist exactly $p$ distinct roots $\{\pi_{l,i}\}_{1\leq i\leq p}$ of $L$, such that*

$$\mathrm{AH}(\pi_{l,i})^p = \mathrm{AH}(\pi_{l-1}) \text{ for all } 1 \leq i \leq p.$$

*Moreover, for all $1 \leq i \leq p$ and $l \geq 2$, $\pi_{l,i} \in Z_{p^l}$.*

*(ii) In addition, if we choose a $p^{th}$ root $\pi_{l-1}^{1/p}$ of $\pi_{l-1}$, then $\{\pi_{l,i}\}_{1\leq i\leq p}$ are the only roots $\pi$ of $L$ such that*

$$|\pi - \pi_{l-1}^{1/p}| < r_2.$$

*In fact, $|\pi_{l,i} - \pi_{l-1}^{1/p}| = (r_l/p)^{1/p} < (1/p)^{1/p} \leq r_2$ for all $l \geq 2$ and $1 \leq i \leq p$.*

*Proof.* To begin with, let $\pi_{l-1}$ be a zero of $L$ with $|\pi_{l-1}| = r_{l-1}$ for some integer $l \geq 2$. Then part $(i)$ of the theorem follows from Theorem A.5 since $\mathrm{AH}(\pi_{l-1})$ is a root of unity of order $p^{l-1}$ and there exist precisely $p$ roots of unity of order $p^l$ whose $p^{\text{th}}$ power is $\mathrm{AH}(\pi_{l-1})$. In other words, there exist exactly $p$ distinct roots $\{\pi_{l,i}\}_{1\leq i\leq p}$ of $L$ such that $\mathrm{AH}(\pi_{l,i})^p = \mathrm{AH}(\pi_{l-1})$. Moreover, again by Theorem A.5, for all $1 \leq i \leq p$ and $l \geq 2$, $\pi_{l,i} \in Z_{p^l}$.

Part $(ii)$ tells us how to produce $\{\pi_{l,i}\}_{1\leq i\leq p}$ directly from $\pi_l$: choose $\pi_{l-1}^{1/p}$ to be any one of its $p^{\text{th}}$ roots and look close by. We now focus on proving this.

For $p$ odd, for all $1 \leq i \leq p$ we get by the binomial theorem

$$(\pi_{l,i} - \pi_{l-1}^{1/p})^p = \pi_{l,i}^p - \pi_{l-1} + \sum_{k=1}^{p-1} \binom{p}{k} \pi_{l,i}^{p-k} \pi_{l-1}^{k/p}.$$

From the previous theorem, we know that $|\pi_{l,i}^p - \pi_{l-1}| = |\pi_{l-1} - \pi_{l,i}^p| = r_l/p$. So, if we can show that all the other summands $\binom{p}{k}\pi_{l,i}^{p-k}\pi_{l-1}^{k/p}$ have absolute value strictly less

than $r_l/p$ for all $1 \leq k \leq p-1$, then by the strong triangle inequality

$$|(\pi_{l,i} - \pi_{l-1}^{1/p})^p| = r_l/p.$$

Taking $p^{\text{th}}$ roots, we get

$$|\pi_{l,i} - \pi_{l-1}^{1/p}| = (r_l/p)^{1/p},$$

as claimed. Now, since $\text{ord}_p \binom{p}{k} = 1$ for all $1 \leq k \leq p-1$, for such $k$, we have

$$\left| \binom{p}{k} \pi_{l,i}^{p-k} \pi_{l-1}^{k/p} \right| = \frac{1}{p} r_l^{p-k} r_{l-1}^{k/p} = \frac{1}{p} r_l^{p-k} r_l^k = \frac{r_l^p}{p} = \frac{r_{l-1}}{p} < \frac{r_l}{p},$$

as wanted. For $p = 2$, for all $1 \leq i \leq 2$, we get

$$(\pi_{l,i} - \pi_{l-1}^{1/2})^2 = \pi_{l,i}^2 - \pi_{l-1} + 2\pi_{l,i} z_{l-1}^{1/2} + 2\pi_{l-1}.$$

From the previous theorem, we know that $|\pi_{l,i}^2 - \pi_{l-1}| = |\pi_{l-1} - \pi_{l,i}^2| = r_l/2$. Also, we have

$$|2\pi_{l,i}\pi_{l-1}^{1/2}| = \frac{1}{2} r_l^2 = \frac{r_{l-1}}{2} \text{ and } |2\pi_{l-1}| = \frac{r_{l-1}}{2},$$

which are both strictly less than $r_l/2$. So by the strong triangle inequality, we have

$$|(\pi_{l,i} - \pi_{l-1}^{1/2})^2| = r_l/2.$$

Taking square roots, we get

$$|\pi_{l,i} - \pi_{l-1}^{1/2}| = (r_l/2)^{1/2},$$

as claimed.

As for the uniqueness, suppose $\pi$ is another root of $L$ such that $|\pi - \pi_{l-1}^{1/p}| < r_2$ and $\pi \neq \pi_{l,i}$ for all $1 \leq i \leq p$. By the last statement of Corollary A.7, we get that there exists $1 \leq i \leq p$ such that $|\pi - \pi_{l,i}| \geq r_2$. From this, the fact that $|\pi_{l,i} - \pi_{l-1}^{1/p}| < r_2$ and the strong triangle inequality, we get

$$|\pi - \pi_{l-1}^{1/p}| = |\pi - \pi_{l,i} + \pi_{l,i} - \pi_{l-1}^{1/p}| = |\pi - \pi_{l,i}| \geq r_2,$$

which contradicts our assumption that $|\pi - \pi_{l-1}^{1/p}| < r_2$.

□

Now, we will use the above two theorems to construct a compatible family of roots of $L$ that maps to a compatible family of $p^{\text{th}}$ power roots of unity. It makes sense to define a sequence of $p^{\text{th}}$ power roots of unity $\{\zeta_{p^l}\}_{l \geq 0}$ to be a compatible family if

(1) $\zeta_1 = 1$,

(2) $\zeta_{p^l}$ is a primitive root of unity of order $p^l$ for all $l \geq 1$,

(3) $\zeta_{p^l}^p = \zeta_{p^{l-1}}$ for all $l \geq 1$.

To define a compatible family of roots of $L$, we first find equivalent formulations of the above natural conditions. Note that condition (2) is equivalent to

(2)' $\zeta_{p^l}$ is a $p^{\text{th}}$ power root of unity such that $|\zeta_{p^l} - 1| = r_l$ for all $l \geq 1$,

whereas (3) is equivalent to

(3)' $|\zeta_{p^{l-1}} - \zeta_{p^l}^p| < r$ for all $l \geq 1$.

(3) $\implies$ (3)$'$ is clear. Whereas, the other direction, suppose (3)$'$ holds and fix $l \geq 1$. Then, since $\zeta_{p^l}$ is a unit, we get

$$|\zeta_{p^{l-1}} - \zeta_{p^l}^p| = \left|\zeta_{p^l}^p \left(\frac{\zeta_{p^{l-1}}}{\zeta_{p^l}^p} - 1\right)\right| = \left|\frac{\zeta_{p^{l-1}}}{\zeta_{p^l}^p} - 1\right| < r$$

However, $\frac{\zeta_{p^{l-1}}}{\zeta_{p^l}^p}$ is itself a $p^{\text{th}}$ power root of unity and so (2)$'$ implies $\zeta_{p^l}^p = \zeta_{p^{l-1}}$. Thus, (3) follows. Motivated by the above equivalent formulations, we make the following definition.

**Definition A.10.** A compatible family of roots of $L$ is a sequence of roots $\{\pi_l\}_{l\geq 0}$ such that

(1) $\pi_0 = 0$,

(2) $L(\pi_l) = 0$ and $|\pi_l| = r_l$ for all $l \geq 1$,

(3) $|\pi_{l-1} - \pi_l^p| < r$ for all $l \geq 1$.

**Theorem A.11.** *$L$ has a compatible family of roots.*

*Proof.* Let $\pi_0 := 0$ and pick $\pi_1$ to be a root of $L$ with $|\pi_1| = r$. Now using Theorem A.9, we choose consecutively for each $l \geq 2$ a root $\pi_l$ of $L$ as required.

□

# Appendix B

# Truncated Artin–Hasse exponential series

At this time, Sage is unable to compute roots of an infinite series (such as $L(X)$ in Appendix A) in $\mathbb{C}_p$. So we are unable to use the results in Appendix A for expressing $p^{\text{th}}$ power roots of unity in terms of roots of $L(X)$. Thus we follow the layout of Appendix A and replace the Artin–Hasse logarithm series with a truncations of it.

**Definition B.1.** For integers $n \geq 0$, the *truncated Artin–Hasse logarithm series* is

$$L_n(X) = \sum_{i=0}^{n} \frac{X^{p^i}}{p^i} = X + \frac{X^p}{p} + \frac{X^{p^2}}{p^2} + \cdots + \frac{X^{p^n}}{p^n}$$

and the *truncated Artin–Hasse exponential series* is

$$\mathrm{AH}_n(X) = \exp(L_n(X)) = \sum_{i=0}^{n} A_{n,i} X^i = 1 + X + \ldots.$$

To make notation easier, we also define $L_\infty(X) = L(X)$ and $\mathrm{AH}_\infty(X) = \mathrm{AH}(X)$.

The series $\mathrm{AH}_n(X)$, unlike the Artin–Hasse exponential series, doesn't have $\mathbb{Z}_p$ coefficients and so we can't plug in any elements of $\mathfrak{m}_p$ that we want. In particular, for $n = 0$, we get $\mathrm{AH}_0(X) = \exp(X)$, which has radius of convergence $r_1 = (1/p)^{1/(p-1)}$. To compute the disc of convergence of $\mathrm{AH}_n(X)$ for any $n \geq 1$, we need the following lemma.

**Lemma B.2.** *For $k \geq 1$,*

$$\frac{1}{p^k}\left(k + \frac{1}{p-1}\right) > \frac{1}{p^{k+1}}\left(k+1+\frac{1}{p-1}\right).$$

*Proof.* Let $k \geq 1$. The following inequalities are all equivalent to each other:

$$\begin{aligned}
\frac{1}{p^k}\left(k + \frac{1}{p-1}\right) &> \frac{1}{p^{k+1}}\left(k+1+\frac{1}{p-1}\right) \\
p(p-1)\left(k + \frac{1}{p-1}\right) &> (p-1)\left(k+1+\frac{1}{p-1}\right) \\
p(p-1)k + p &> (p-1)(k+1) + 1 \\
p(p-1)k + p - 1 &> (p-1)(k+1) \\
pk + 1 &> k + 1 \\
p &> 1.
\end{aligned}$$

Since $p$ is a prime, the last inequality is true and so are all the previous ones. The Lemma follows. □

**Lemma B.3.** *For $n \geq 1$, the truncated Artin–Hasse series $\mathrm{AH}_n(X)$ converges precisely for all $x \in \mathbb{C}_p$ with $|x| < R(\mathrm{AH}_n)$, where*

$$R(\mathrm{AH}_n) = \left(\frac{1}{p}\right)^{\frac{1}{p^{n+1}}\left(n+1+\frac{1}{p-1}\right)} = \left(\frac{1}{p}\right)^{\frac{n(p-1)+p}{p^{n+1}(p-1)}} = \left(\frac{1}{p}\right)^{\frac{1}{p-1}} p^{\frac{p+p^2+\cdots+p^n-n}{p^{n+1}}} > r_1.$$

*In addition, the coefficients of* $\mathrm{AH}_n(X)$ *satisfy*

$$\mathrm{ord}_p(A_{n,i}) \geq -\left\lfloor \frac{i}{p^{n+1}} \right\rfloor \left(n+1+\frac{1}{p-1}\right) + \frac{S\left(\lfloor i/p^{n+1} \rfloor\right)}{p-1}.$$

*Proof.* One can write

$$\mathrm{AH}_n(X) = \mathrm{AH}(X) \prod_{k \geq n+1} \exp(-X^{p^k}/p^k). \tag{B.1}$$

Each of the exponentials $\exp(-X^{p^k}/p^k)$ converges precisely for all $x \in \mathbb{C}_p$ such that $\left|\frac{x^{p^k}}{p^k}\right| < \left(\frac{1}{p}\right)^{\frac{1}{p-1}}$. Thus their radius of convergence are respectively $R_k = \left(\frac{1}{p}\right)^{\frac{1}{p^k}\left(k+\frac{1}{p-1}\right)}$.

By Lemma B.2, these radii $R_k$ are strictly increasing as $k \geq 1$ gets bigger. In other words, for $k \geq 1$, we have

$$\left(\frac{1}{p}\right)^{\frac{1}{p^k}\left(k+\frac{1}{p-1}\right)} < \left(\frac{1}{p}\right)^{\frac{1}{p^{k+1}}\left(k+1+\frac{1}{p-1}\right)}.$$

Hence

$$\min_{k \geq n+1} \left(\frac{1}{p}\right)^{\frac{1}{p^k}\left(k+\frac{1}{p-1}\right)} = \left(\frac{1}{p}\right)^{\frac{1}{p^{n+1}}\left(n+1+\frac{1}{p-1}\right)}.$$

Let $F_{n+1}(X) = \prod_{k \geq n+1} \exp(-X^{p^k}/p^k) = \exp(-\sum_{k \geq n+1} X^{p^k}/p^k)$. We claim that $F_{n+1}(X)$ converges for all $x \in \mathbb{C}_p$ such that $|x| < R_{n+1}$, i.e. $R(F_{n+1}) \geq R_{n+1}$. Let $x \in \mathbb{C}_p$ such that $|x| < R_{n+1}$. Then for $k \geq n+1$, $R_{n+1} \leq R_k$ and so we have $\left|\frac{x^{p^k}}{p^k}\right| < \frac{1}{p-1}$. Thus, by lemma A.6 the claim follows.

Equation (B.1) implies $\mathrm{AH}_n(X)$ converges for all $x \in \mathbb{C}_p$ with $|x| < R_{n+1}$ since $R(\mathrm{AH}) = 1 > R_{n+1}$. We will now show that $\mathrm{AH}_n(X)$ doesn't converge for any other $x \in \mathbb{C}_p$ and thus $R(\mathrm{AH}_n) = R_{n+1}$. Let $x \in \mathbb{C}_p$ have absolute value $R_{n+1}$ and suppose $\mathrm{AH}_n(X)$ converges at $x$. Since $R_{n+1} < R_{n+2}$, $F_{n+2}(X)$ converges at $x$. We also know

that the $\mathrm{AH}(X)$ converges at $x$. Since $R(1/F_{n+2}(X)) \geq R_{n+2}$ by the same argument as for $R(F_{n+1}) \geq R_{n+1}$ in the previous paragraph, the power series equation

$$\exp(-X^{p^{n+1}}/p^{n+1}) = \frac{\mathrm{AH}_n(X)}{\mathrm{AH}(X)F_{n+2}(X)}$$

implies $\exp(-X^{p^{n+1}}/p^{n+1})$ converges at $x$ with $|x| = R_{n+2}$ (recall $R(1/\mathrm{AH}(X)) = 1 > R_{n+2}$ since $1/\mathrm{AH}(X) \in 1 + X\mathbb{Z}_p[[X]]$). This is a contradiction to the disc of convergence of this series $\exp(-X^{p^{n+1}}/p^{n+1})$ being $|x| < R_{n+1}$ as discussed earlier. Therefore the first part of the lemma follows.

Now we show the inequality in the lemma. By direct computation we see that the coefficients of

$$\exp(-X^{p^k}/p^k) = \sum_{i=0}^{\infty} \frac{X^{p^k i}}{p^{ki} i!} = \sum_{m=0}^{\infty} c_{k,m} X^m$$

satisfy

$$\mathrm{ord}_p(c_{k,p^k i}) = -ki - \frac{i - S(i)}{p-1} = -i\left(k + \frac{1}{p-1}\right) + \frac{S(i)}{p-1}$$

for $i \geq 0$ and $c_{k,m} = 0$ for $m$ not divisible by $p^k$. Thus, for all $m \geq 0$ we have

$$\mathrm{ord}_p(c_{k,m}) \geq -\left\lfloor \frac{m}{p^k} \right\rfloor \left(k + \frac{1}{p-1}\right) + \frac{S(\lfloor m/p^k \rfloor)}{p-1}.$$

Denote by $h(k,m)$ the right hand side of this inequality. We claim that $h(k,m)$ is non-decreasing with respect to $k \geq 1$ for all $m \geq 0$. Fix $m \geq 0$. For $k \geq 1$ such that $0 \leq m < p^k$, we have $h(k,m) = 0$ so $h(k+1,m) \geq h(k,m)$. For $k \geq 1$ such that $p^k \leq m < p^{k+1}$, we have $1 \leq \lfloor m/p^k \rfloor < p$. So $S(\lfloor m/p^k \rfloor) = \lfloor m/p^k \rfloor$ and thus $h(k,m) = -\lfloor m/p^k \rfloor k \leq 0 = h(k+1,m)$. Assume $k \geq 1$ and $m > p^{k+1}$. From the

properties of the floor function we have

$$\begin{aligned} h(k+1,m)-h(k,m) &\geq \left(\left\lfloor \frac{m}{p^k}\right\rfloor - \left\lfloor \frac{m}{p^{k+1}}\right\rfloor\right)\left(k+\frac{1}{p-1}\right) - \left\lfloor \frac{m}{p^{k+1}}\right\rfloor - 1 \\ &\geq \left\lfloor \frac{m}{p^{k+1}}\right\rfloor (p-1)\left(k+\frac{1}{p-1}\right) - \left\lfloor \frac{m}{p^{k+1}}\right\rfloor - 1 \\ &\geq \left\lfloor \frac{m}{p^{k+1}}\right\rfloor k(p-1) - 1 \\ &\geq k(p-1) - 1 \\ &\geq 0. \end{aligned}$$

Thus the claim follows, i.e. $h(k,m)$ is non-decreasing with respect to $k \geq 1$ for all $m \geq 0$. Therefore, for $k \geq n+1$ and $m \geq 0$ we get

$$\operatorname{ord}_p(c_{k,m}) \geq h(k,m) \geq h(n+1,m). \tag{B.2}$$

From B.1 we get

$$A_{n,i} = \sum_{j,j_{n+1},j_{n+1},\cdots j_{\lfloor \log_p(i)\rfloor}} A_j c_{n+1,j_{n+1}} c_{n+2,j_{n+2}} \cdots c_{\lfloor \log_p(i)\rfloor, j_{\lfloor \log_p(i)\rfloor}},$$

where $A_j$ are the coefficient of $\mathrm{AH}(X)$ and the sum runs over all $0 \leq j, j_k \leq i$ for any $n+1 \leq k \leq \lfloor \log_p(i)\rfloor$ such that $j + j_{n+1} + j_{n+2} + \cdots + j_{\lfloor \log_p(i)\rfloor} = i$. Note that the sum is finite. Then for $i \geq p^{n+1}$, $\operatorname{ord}_p(A_j) \geq 0$ and (B.2) imply

$$\operatorname{ord}_p(A_{n,i}) \geq \min\left\{ \operatorname{ord}_p(A_j) + \sum_{k=n+1}^{\lfloor \log_p(i)\rfloor} \operatorname{ord}_p(c_{k,j_k}) \right\} \geq h(n+1,i),$$

where the minimum runs over all $0 \leq j, j_k \leq i$ for any $n+1 \leq k \leq \lfloor \log_p(i)\rfloor$ such

that $j + j_{n+1} + j_{n+2} + \cdots + j_{\lfloor \log_p(i) \rfloor} = i$. This is precisely what we had to show. □

From this lemma one can show the following theorem.

**Theorem B.4.** *For every $n \geq 1$ or $n = \infty$, the truncated Artin–Hasse series $\mathrm{AH}_n(X)$ is an isometry in its disc of convergence. In other words, for every $x, y \in \mathbb{C}_p$ such that $|x|, |y| < R(\mathrm{AH}_n)$ we have*

$$|\,\mathrm{AH}_n(x) - \mathrm{AH}_n(y)| = |x - y|.$$

*Proof.* For $n = \infty$, $\mathrm{AH}_\infty(X)$ is the Artin–Hasse exponential series, which was shown to be an isometry in Appendix A. See Keith Conrad's paper [7] for $n \geq 1$. □

Analogously to the Artin–Hasse exponential series, we would like to know the layout of the roots of $L_n(X)$.

**Theorem B.5.** *In $\mathfrak{m}_p$, for any $n \geq 1$ we have $L_n(X)$ has exactly $p^n$ roots: $p^{l-1}(p-1)$ distinct non-trivial zeros with absolute value $r_l$ for $1 \leq l \leq n$ and no other non-trivial zeros. In particular, the zeros of $L_n(X)$ other than 0 are in $\mathfrak{m}_p - D_p$.*

*Proof.* The Newton polygon of $L_n(X)$ is identical to that of $L(X)$ up to vertex $(p^n, -n)$ (see Lemma A.3 ) and then it continues with a straight vertical line. Also, $L_n'(X) \in 1 + X\mathbb{Z}_p[X]$ implies $|L_n'(x)| = 1$ and consequently $L_n'(x) \neq 0$ for all $x \in \mathfrak{m}_p$. These observations conclude the proof. □

We would like to be able to plug in $x \in \mathbb{C}_p$ with $|x| = r_n$ into $\mathrm{AH}_n(X)$ just as we did with $\mathrm{AH}(X)$. So, we want $r_n < R(\mathrm{AH}_n)$, which is equivalent to $p^2 > (n+1)(p-1)+1$ or $p+1 > n+1$. Therefore, this substitution only works for primes $p > n$. In particular, it doesn't work for all primes as it did for $L(X)$.

The natural follow up question to ask is: for what integers $l \leq n$, we may plug in $x \in \mathbb{C}_p$ with $|x| = r_l$ into $\mathrm{AH}_n(X)$? In other words, for what $l \leq n$ is it true that $r_l < R(\mathrm{AH}_n)$. This is equivalent to showing $R(\mathrm{AH}_n)/r_l > 1$. Hence we get

$$\begin{aligned}\frac{R(\mathrm{AH}_n)}{r_l} &= \left(\frac{1}{p}\right)^{\frac{n(p-1)+p}{p^{n+1}(p-1)} - \frac{1}{p^{l-1}(p-1)}} \\ &= \left(\frac{1}{p}\right)^{\frac{n(p-1)+p-p^{n-l+2}}{p^{n+1}(p-1)}} \\ &= p^{\left(p+p^2+\cdots+p^{n-l+1}-n\right)/p^{n+1}}.\end{aligned}$$

Thus, for $n \geq l \geq 1$, we may plug in $x \in \mathbb{C}_p$ with $|x| = r_l$ into $\mathrm{AH}_n(X)$ if and only if

$$p + p^2 + \cdots + p^{n-l+1} > n. \tag{B.3}$$

**Definition B.6.** When we say $n$ is big enough compared to $l$ or $n \gg l$, we mean $n \geq l \geq 1$ and (B.3) is satisfied.

For $n = \infty$, (B.3) is satisfied for any $l \geq 1$ and so fits the above definition. For $\pi_{n,l}$ root of $L_n(X)$ of absolute value $r_l$, is $\mathrm{AH}_n(\pi_{n,l})$ a primitive root of unity of order $p^l$ for all $l$ such that $\pi_{n,l}$ is in within the disk of convergence of $\mathrm{AH}_n(X)$? In other words, is there an analogue to Theorem A.5?

Fix $n \gg l$. We borrow the notation from Theorem A.5. Let $Z$ denote the set of zeros of $L_n(X)$, $Z_{p^l} := \{\pi \in Z : |\pi| \leq r_l\}$ and $Z'_{p^l} := \{\pi \in Z : |\pi| = r_l\}$. We set $D(\mathrm{AH}_n)$ to be the disc of convergence of $\mathrm{AH}_n(X)$.

**Theorem B.7.** *For all $n \gg l$, the maps*

$$\mathrm{AH}_n \colon D(\mathrm{AH}_n) \to 1 + D(\mathrm{AH}_n), \mathrm{AH}_n \colon Z_{p^l} \to \mu_{p^l} \text{ and } \mathrm{AH} \colon Z'_{p^l} \to \mu'_{p^l}$$

*are bijections. In addition,* $\pi_{n,l} \in Z'_{p^l}$, $\mathrm{AH}_n(\pi_{n,l})$ *is the unique* $\zeta \in \mu_{p^\infty}$ *such that*

$$\left| \zeta - \sum_{i \leq p^{l-1}} A_{n,i} \pi_{n,l}^i \right| < r_1 = r,$$

*where* $A_{n,i}$ *are the coefficients of* $\mathrm{AH}_n(X)$.

*Proof.* First, when $n = \infty$, the statements of the theorem hold by Theorem A.5. Hence, fix integers $n \gg l$.

Step 1: $\mathrm{AH}_n(Z_{p^l}) \subseteq \mu_{p^l}$ and $\mathrm{AH}_n(Z'_{p^l}) \subseteq \mu'_{p^l}$.

First, note that $\mathrm{AH}_n(0) = 1$. Now let $\pi_{n,l} \in Z_{p^l} - \{0\}$ have $p$-adic absolute value $r_l$. We will show that $\mathrm{AH}(\pi_{n,l})$ is a root of unity with order $p^l$.

Note that $|L_n(X)|_{r_l} = p^l r_l^{p^l} = p^l(r/p)$. If we could get rid of the $p^l$ on the right side, we'd be left with $r/p$, which is less than $r = R(\exp)$, and we would then be able to make a substitution by Theorem A.6. To remove the $l$, consider the equation

$$\mathrm{AH}_n(X)^{p^l} = \exp(p^l L_n(X)). \tag{B.4}$$

We will show that substituting $X = \pi_{n,l}$ in this equation is allowed and preserves equality. Then we get

$$\mathrm{AH}_n(\pi_{n,l})^{p^l} = \exp(p^l L(\pi_{n,l})) = \exp(0) = 1,$$

so $\mathrm{AH}(\pi_{n,l})$ is a $p^l$-th root of unity. The first two conditions of Theorem A.6 are obviously true for (B.4). Let's check the third condition for (B.4):

$$|p^l L_n(X)|_{r_l} = |p|^l |L_n(X)|_{r_l} = \left(\frac{1}{p}\right)^l p^l(r/p) = r/p < r = R(\exp).$$

Hence setting $X = \pi_{n,l}$ in (B.4) is valid.

Now, to argue that $\mathrm{AH}_n(\pi_{n,l})$ has order precisely $p^l$, we notice that $|\mathrm{AH}_n(\pi_{n,l}) - 1| = |\pi_{n,l}| = r_l$ by Theorem B.4. Since we know that in general any $p$-adic root of unity $\zeta$ of order $p^m$ with $m \geq 1$ satisfies $|\zeta - 1| = r_m$, the claim follows.

Step 2: The mapping $\mathrm{AH}_n \colon D(\mathrm{AH}_n) \to 1 + D(\mathrm{AH}_n)$ is a bijection.

This follows by the fact that $\mathrm{AH}_n(X)$ is an isometry on its domain (Theorem B.4).

Step 3: The mappings $\mathrm{AH}_n \colon Z \to \mu_{p^\infty}$ and $\mathrm{AH}_n \colon Z_{p^l} \to \mu'_{p^l}$ are bijections.

By step 1, we know that $\mathrm{AH}_n(Z_{p^l}) \subseteq \mu_{p^l}$ for all $l \geq 1$ and by step 2, the mapping $\mathrm{AH}_n \colon \mathfrak{m}_p \to 1 + \mathfrak{m}_p$ is a bijection. Thus, its restriction $\mathrm{AH}_n \colon Z_{p^l} \to \mu_{p^l}$ must also be a bijection. In addition, since $Z = \{0\} \cup \bigcup_{l \geq 1} Z_{p^l}$ and $\mu_{p^\infty} = \{1\} \cup \bigcup_{l \geq 1} \mu'_{p^l}$, we get that $\mathrm{AH}_n \colon Z \to \mu_{p^\infty}$ is also a bijection.

Step 4: If $\pi_{n,l} \in Z - \{0\}$ has absolute value $r_l$ and $\zeta \in \mu_{p^\infty}$ satisfies $|\zeta - \sum_{i \leq p^{l-1}} A_{n,i} \pi_{n,l}^i| < r$, where the coefficients $A_{n,i}$ come from $\mathrm{AH}_n(X) = \sum_{i \geq 0} A_{n,i} X^i$, then $\zeta = \mathrm{AH}_n(\pi_{n,l})$.

By the strong triangle inequality, we have

$$\left| \mathrm{AH}_n(\pi_{n,l}) - \sum_{i \leq p^{l-1}} A_{n,i} \pi_{n,l}^i \right| \leq \max_{i > p^{l-1}} \left\{ |A_{n,i} \pi_{n,l}^i| \right\}.$$

Since the coefficients of $\mathrm{AH}_n(X)$ match the coefficients of $\mathrm{AH}(X) \in \mathbb{Z}_p[[X]]$ up to the coefficient of $X^{p^{n+1}-1}$, $A_{n,i} \in \mathbb{Z}_p$ for $p^{l-1} < i < p^{n+1}$ and so we have

$$|A_{n,i} \pi_{n,l}^i| \leq |\pi_l|^{p^{l-1}+1} < |\pi_{n,l}|^{p^{l-1}} = r_l^{p^{l-1}} = (1/p)^{1/(p-1)} = r.$$

For $i \geq p^{n+1}$ we have $S(\lfloor i/p^{n+1} \rfloor) \geq 1$. Thus, since the coefficients of $A_{n,i}$ satisfy the

bound in Lemma B.3, for $i \geq p^{n+1}$ we get

$$|A_{n,i}\pi_{n,l}^i| \leq \max_{i>p^{l-1}} \left\{ \left(\frac{1}{p}\right)^{i\left(\frac{1}{p^{l-1}(p-1)} - \frac{1}{p^{n+1}}\left(n+1+\frac{1}{p-1}\right)\right)+\frac{1}{p-1}} \right\} \overset{(B.3)}{<} r.$$

Hence $|\operatorname{AH}_n(\pi_{n,l}) - S| < r$, where $S = \sum_{i \leq p^{l-1}} A_i \pi_{n,l}^i$, and thus $|\operatorname{AH}_n(\pi_{n,l}) - S + S - \zeta| < r$ by the strong triangle inequality. Different $p^{\text{th}}$ power roots of unity have distance at least $r$ from each other, so $\zeta = \operatorname{AH}_n(\pi_{n,l})$. □

From this theorem and Theorem A.5 we have many ways to represent the same $p^l$-th roots of unity.

**Corollary B.8.** *For any $n \gg l$ we have $\mathbb{Q}_p(\pi_{n,l}) = \mathbb{Q}_p(\zeta_{p^l})$.*

*Proof.* By Theorem B.7 we have $\zeta_{p^l} = \operatorname{AH}_n(\pi_{n,l})$ and so $\mathbb{Q}_p(\zeta_{p^l}) \subseteq \mathbb{Q}_p(\pi_{n,l})$. On the other hand, since $L_n(X)$ has exactly $\varphi(p^l) = p^{l-1}(p-1)$ roots of absolute value $(1/p)^{1/\varphi(p^l)}$ by its Newton polygon, we get that there are at most $p^{l-1}(p-1)$ conjugates of $\pi_{n,l}$. Hence $\mathbb{Q}_p(\pi_{n,l})$ has degree at most $p^{l-1}(p-1)$ over $\mathbb{Q}_p$. However since $\mathbb{Q}_p \subseteq \mathbb{Q}_p(\zeta_{p^l}) \subseteq \mathbb{Q}_p(\pi_{n,l})$ and $\mathbb{Q}_p(\zeta_{p^l})$ has degree at most $p^{l-1}(p-1)$ over $\mathbb{Q}_p$, the corollary follows. □

Let's see how $\pi_{n,l}$ are related to $\pi_l$, the roots of $L(X)$.

**Lemma B.9.** *Let $n \gg l$ and $\pi_{n,l}$ a root of $L_n$ of absolute value $r_l$. Then, there exists a unique root $\pi_l$ of $L$ such that*

$$|\pi_l - \pi_{n,l}| < r.$$

*In addition, $|\pi_l| = r_l$ and $|\pi_l - \pi_{n,l}| = R(n,l)$ where $R(n,l) = r_l^{p^{n+1}-(n+1)p^{l-1}(p-1)} \leq r/p = r_l^{p^l}$. For $n = \infty$, this means $\pi_{\infty,l} = \pi_l$.*

*Proof.* First, for $n = \infty$, the statement is true by Theorem A.3. So let $n \gg l$ be integers and $\pi_{n,l}$ a root of $L_n$ of absolute value $r_l$, i.e. $\pi_{n,l}$ is within the radius of convergence of $\mathrm{AH}_n(X)$. Consider the Newton polygon of $H(X) := L(X + \pi_{n,l})$. Since $L_n(\pi_{n,l}) = 0$, we know that the constant term of this power series has absolute value

$$|H_0| = |L(\pi_{n,l})| = \left| L_n(\pi_{n,l}) + \sum_{i=n+1}^{\infty} \pi_{n,l}^{p^i}/p^i \right| = |\pi_{n,l}^{p^{n+1}}/p^{n+1}| == r_l^{p^{n+1}-(n+1)p^{l-1}(p-1)}$$

So, $\mathrm{ord}_p(H_0) = \frac{p^{n-l+2}-(n+1)(p-1)}{p-1}$. For $k \geq 1$, using the binomial theorem and collecting like terms, we get

$$H_k = \sum_{i \geq 0, p^i \geq k} \frac{\binom{p^i}{k}}{p^i} \pi_{n,l}^{p^i-k}.$$

Since $\mathrm{ord}_p \binom{p^i}{k} = i - \mathrm{ord}_p(k)$ and $\mathrm{ord}_p(\pi_{n,l}^{p^{i+1}-k}) > \mathrm{ord}_p(\pi_{n,l}^{p^i-k}) \geq 0$ for all $i$ such that $p^i \geq k$, we have $\mathrm{ord}_p(H_k) = -\mathrm{ord}_p(k) + \frac{p^{\lceil \log_p(k) \rceil}-k}{p^{l-1}(p-1)} \geq -\mathrm{ord}_p(k)$. Furthermore, equality holds if and only if $k$ is a $p^{\text{th}}$ power.

Thus, the Newton polygon of $H(X)$ is identical to the Newton polygon of $L(X)$ except perhaps the very first line. To see what happens with the first line, the line passing through the points $(1, 0)$ and $(p, -1)$ has equation $y = \frac{-1}{p-1}(x - 1)$. So, its $y$-intercept is $\frac{1}{p-1}$. Thus, if $\mathrm{ord}_p(H_0) > \frac{1}{p-1}$, then the first line starts at $(0, \mathrm{ord}_p(H_0))$ and ends at $(1, 0)$, which completes the Newton polygon $H$. The condition $\mathrm{ord}_p(H_0) > \frac{1}{p-1}$ is equivalent to

$$\frac{p^{n-l+2} - (n+1)(p-1)}{p-1} > \frac{1}{p-1} \iff p^{n-l+2} > (n+1)(p-1) + 1,$$

which is equivalent to inequality (B.3). This is follows from our assumption $n \gg l$.

Hence, $H$ has a unique root of absolute value $R(n,l) := r/p^{p+p^2+\cdots+p^{n-l+1}-n} \leq r/p < r$. So, if $\pi$ is a root of $L$, then $\alpha = \pi - \pi_{n,l}$ is a root of $H$ and so $|\alpha| = |\pi - \pi_{n,l}| = r_l \geq r$ for some $l \geq 1$ or $|\pi - \pi_{n,l}| = R(n,l) < r$. In particular, there exists a unique root $\pi_l$ of $L$ of absolute value $r_l$ such that $|\pi_l - \pi_{n,l}| = R(n,l)$ and all other roots of $L$ are a distance of at least $r$ from $\pi_{n,l}$. □

Hence, if $n \to \infty$, then $R(n,l) \to 0$ and so $\pi_{n,l} \to \pi_l$. This means that as $n$ increases the roots $\pi_{n,l}$ of $L_n(X)$ closest to $\pi_l$ approximate better and better $\pi_l$ and at each step we know the exact error by the above lemma.

We can represent any $p^l$-th root of unity $\zeta$ as $\zeta = \mathrm{AH}_n(\pi_{n,l})$ for $\pi_{n,l}$ a root of absolute value $r_l$ of $L_n(X)$ for some $n$ big enough so that $\mathrm{AH}_n(\pi_{n,l})$ converges or as $\zeta = \mathrm{AH}(\pi_l)$ for $\pi_l$ a root of $L(X)$ of absolute value $r_l$.

**Theorem B.10.** *For any $n \gg l$ and $n' \gg l$ such that $n' > n$, the roots of $L_n(X)$ of absolute value at most $r_l$ are in a one-to-one correspondence with the roots of $L_{n'}(X)$ of absolute value at most $r_l$ given by $\pi_{n,l} \mapsto \pi_{n',l}$, where $\pi_{n',l}$ is the unique root of $L_{n'}$ such that*

$$|\pi_{n',l} - \pi_{n,l}| < r,$$

*i.e. $\pi_{n',l}$ is the unique root of $L_{n'}$ closest to $\pi_{n,l}$. More precisely, $|\pi_{n',l} - \pi_{n,l}| = R(n,l) < r$, where $R(n,l) = r/p^{p+p^2+\cdots+p^{n-l+1}-n} \leq r/p = r_l^{p^l} < r$. Moreover, $\mathrm{AH}_n(\pi_{n,l})$ and $\mathrm{AH}_{n'}(\pi_{n',l})$ are the same root of unity.*

*Proof.* If $n' = \infty$, then the above theorem follows by Lemma B.9 except for the very last statement. For integers $n \gg l$ and $n' \gg l$, fix $\pi_{n,l}$ as in the theorem. By Lemma B.9 there is a unique root $\pi_l$ of $L(X)$ closest to $\pi_{n,l}$ and in addition there's a unique root $\pi_{n',l}$ closest to $\pi_l$, which satisfy $|\pi_l - \pi_{n,l}| = R(n,l) < r$ and

$|\pi_{n',l} - \pi_l| = R(n', l) < r$. Since $R(n, l)$ is decreasing with respect to $n$ for $n \gg l$, we have $R(n', l) < R(n, l)$ and thus

$$|\pi_{n',l} - \pi_{n,l}| = |\pi_{n',l} - \pi_l + \pi_l - \pi_{n,l}| = R(n, l) < r.$$

By a similar argument, any other root of $L_{n'}(X)$ will be at least distance $r$ away from $\pi_{n,l}$. It remains to show that $\mathrm{AH}_n(\pi_{n,l})$ and $\mathrm{AH}_{n'}(\pi_{n',l})$ are the same root of unity. It is enough to do this for $n' = \infty$. Let $n \gg l$, $\zeta_{p^l} = \mathrm{AH}(\pi_l)$ and suppose $\pi_{n,l}$ is the root of $L_n(X)$ closest to $\pi_l$. Then $\mathrm{AH}_n(\pi_{n,l}) = \zeta'_{p^l}$ for some $p^l$-th order root of unity $\zeta'_{p^l}$. We want to show that $\zeta_{p^l} = \zeta'_{p^l}$. It is enough to show that $|\zeta_{p^l} - \zeta'_{p^l}| < r_1 = r$. By the strong triangle inequality we get

$$\begin{aligned} |\zeta_{p^l} - \zeta'_{p^l}| &= |\,\mathrm{AH}(\pi_l) - \mathrm{AH}_n(\pi_{n,l})| \\ &= |\,\mathrm{AH}(\pi_l) - \mathrm{AH}(\pi_{n,l}) + \mathrm{AH}(\pi_{n,l}) - \mathrm{AH}_n(\pi_{n,l})| \\ &\leq \max\{|\,\mathrm{AH}(\pi_l) - \mathrm{AH}(\pi_{n,l})|, |\,\mathrm{AH}(\pi_{n,l}) - \mathrm{AH}_n(\pi_{n,l})|\}. \end{aligned}$$

Using the fact that the Artin–Hasse series is an isometry and lemma B.9 we know that $|\,\mathrm{AH}(\pi_l) - \mathrm{AH}(\pi_{n,l})| = |\pi_l - \pi_{n,l}| = R(n, l) < r$. Thus it remains to show that $|\,\mathrm{AH}(\pi_{n,l}) - \mathrm{AH}_n(\pi_{n,l})| < r$. Recall that for $k \geq 0$, $|A_k| \leq 1$, $|A_{n,k}|$ is bounded by Lemma B.3 and $A_k = A_{n,k}$ for $0 \leq k < p^{n+1}$. Moreover, from the proofs of Step 4 in both Theorem A.5 and Theorem B.7, we get

$$|\,\mathrm{AH}(\pi_{n,l}) - \mathrm{AH}_n(\pi_{n,l})| \leq \max_{i \geq p^{n+1}} \{|A_i \pi_l^i|, |A_{n,i} \pi_{n,l}^i|\} < r,$$

as wanted. □

# Bibliography


[1] A. Adolphson and S. Sperber, *On Twisted Exponential Sums*, Ann. of Math. **290** (1991), 713–726.

[2] ———, *Twisted Exponential Sums and Newton Polyhedra*, J. Reine Angew. Math. **443** (1993), 151–177.

[3] F. Baldassarri, *Higher p-adic Gamma Functions and Dwork Cohomology*, Astérisque **119/120** (1984), 111–127.

[4] R. Blache, *A Stickelberger Theorem for p-adic Gauss Sums*, Acta Arith. **118** (2005), 11–26.

[5] M. Boyarsky, *p-adic Gamma Function and Dwork Cohomology*, Transactions of the American Math. Soc. **257** (1980), 359–369.

[6] K. Conrad, *L-functions for Gauss and Jacobi sums*, (2016), `http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/LfunctionGaussJacobi.pdf`.

[7] ———, *Truncated Artin-Hasse Series and Roots of Unity*, (2016), `http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/AHrootofunity.pdf`.

[8] B. H. Gross and N. I. Koblitz, *Gauss Sums and the p-adic Gamma Function*, Ann. of Math. **109** (1979), 569–581.

[9] K. Ireland and M. Rosem, *A Classical Introduction to Modern Number Theory, 2nd Ed.*, Graduate texts in math., vol. 121, Springer-Verlag, New York, 1992.

[10] K. Kedlaya, *p-adic Differential Equations*, Cambridge Studies in Advanced Mathematics, vol. 125, Cambridge University Press, Cambridge, 2010.

[11] S. Lang, *Cyclotomic Fields I and II Combined 2nd Ed.*, Graduate texts in math., vol. 121, Springer-Verlag, New York, 1990.

[12] C. Liu, *The L-functions of Twisted Witt Extensions*, J. Number Theory **125** (2007), 267–284.

[13] C. Liu and C Niu, *Generic Twisted $T$-adic Exponential Sums of Polynomials*, J. Number Theory **140** (2014), 38–59.

[14] P. Robba, *Index of p-adic Differential Operators*, Astérisque, vol. 119-120, Springer-Verlag, New York, 2000.

[15] A. M. Robert, *A Course in p-adic Analysis*, Graduate texts in math., vol. 198, Springer-Verlag, New York, 2000.

[16] ———, *The Gross-Koblitz Formula Revisited*, Rend. Semin. Mat. Univ. Padova **105** (2001), 157–170.

[17] A. G. Shanbhag, P. V. Kumar, and T. Helleseth, *Upper Bound for a Hybrid Sum Over Galois Rings with Applications to Aperiodic Correlation of Some q-ary Sequences*, IEEE Trans. of Inf. Theory **42** (1996), 250–254.

[18] P. Young, *Radii of p-adic Convergence of Generic Solutions of Homogeneous Linear Differential Equations*, (1988), http://youngp.people.cofc.edu/thesis.pdf.