Leopoldt's Conjecture for Abelian and non-Abelian cases

Eric Stansifer

Master's Thesis, 2012 December 5

Advised by Prof. Massimo Bertolini

ALGANT Erasmus Mundus

Universiteit Leiden

Universitá degli Studi di Milano

Contents

1 Introduction

The Dirichlet L-series $L(s, \chi)$ is a generalization of the Riemann-zeta function,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

given by

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

for a Dirichlet character $\chi : \mathbb{Z} \to \mathbb{C}$ (that is, $\chi(ab) = \chi(a)\chi(b)$ and for some positive n, $\chi^n = 1$ except where χ is zero). In both cases, the series converges for complex numbers s with real part greater than 1, but can be extended meromorphically to the whole complex plane, with a simple pole at $s = 1$ in $\zeta(s) = L(s, 1)$.

The behavior of these functions near $s = 1$ is of central importance, and a classical result of number theory gives that

$$L(1, \chi) \neq 0$$

for $\chi \neq 1$. A corollary of this is the famed **Dirichlet's Theorem**, that for relatively prime integers a and n, there are infinitely many prime numbers congruent to a modulo n. The value of $L(1, \chi)$ also encodes information related to the *regulator* of a number field K, and $L(1, \chi) \neq 0$ implies that the regulator is nonzero, which gives certain linear independence results on the units of the number field K.

Given the importance of the Dirichlet L-series, Heinrich-Wolfgang Leopoldt defined a p-adic analog $L_p(s, \chi)$ to the Dirichlet L-series and raised the question in 1962 [7, p. 57] of whether one still has $L_p(1, \chi) \neq 0$. Similarly, the value of $L_p(1, \chi)$ is related to the *p-adic regulator* of a number field K, and one has $L_p(1, \chi) \neq 0$ if and only if the p-adic regulator is nonzero.

In this document we will concern ourselves with an equivalent form of the question that Leopoldt raised, namely whether a collection of linear independent units in a number field K will also be $\mathbb{Z}_p$-linearly independent in the p-adic completions of K. **Leopoldt's Conjecture** states that this always holds true: that linearly independent units in K will always be $\mathbb{Z}_p$-linearly independent in the p-adic completions.

This conjecture was settled in the affirmative for abelian extensions K of $\mathbb{Q}$ by Armand Brumer in 1967 [4], by completing a proof of James Ax in 1965 [2] that the p-adic regulator of such K is nonzero. While the abelian situation is the only case considered by Leopoldt, as that is the only situation in which the value of $L_p(1, \chi)$ is relevant, the equivalent forms of Leopoldt's Conjecture are meaningful for any number field K of $\mathbb{Q}$ and it is natural to ask whether Leopoldt's Conjecture holds in these situations.

So far no general proof for Leopoldt's Conjecture has been found. We can measure the degree of failure of Leopoldt's Conjecture for a particular number field with the *Leopoldt defect*, and several results have been found that give bounds on the Leopoldt defect. One substantial result was given by D.W. Masser [8] and Michel Waldschmidt [10] in 1981.

This thesis then follows in two main movements. In our first part, sections 3 and 4, we prove that Leopoldt's Conjecture is true for abelian extensions of $\mathbb{Q}$ (Theorem 22). We continue this proof in section 5 with a variation for non-abelian galois extensions, which gives a weak bound on the Leopoldt defect (Theorem 29). In our second part, sections 6 and 7, we prove the bound on the Leopoldt defect found by Masser and Waldschmidt (Theorem 48).

1.1 Outline

Part of our motivation is to avoid dependencies on deep external results (except that in section 6 we will need some facts of commutative algebra which we quickly review in section 6.1), so we begin in section 2 with introducing the conventions that we will use in the rest of the paper, as well as proving a number of general results that will be needed.

The field $\mathbb{C}_p$, which is the p-adic equivalent of the complex numbers $\mathbb{C}$, will play a key role throughout this paper, so in sections 2.1, 2.2, and 2.3 we develop the basic notions of p-adic analysis, which is closely analogous to ordinary analysis on $\mathbb{C}$, and in particular investigate the p-adic exponential and logarithmic functions. This culminates in section 2.5 with a p-adic form of Schwarz's Lemma, which is nicer to state than its analog for $\mathbb{C}$.

In section 2.4 we state Leopoldt's Conjecture and define the Leopoldt defect.

Several times in this paper we will be interested in the growth behavior of functions $f(x)$ for arbitrarily large $x \in \mathbb{R}$; in particular, we can show that certain *algebraic* numbers are zero if they are the values of functions that get arbitrarily small, using Theorem 12 (section 2.7) to give a lower bound on the size of a non-zero algebraic number. The standard notation for analyzing asymptotic behavior is big-Oh notation, but to avoid problems with that notation we introduce a variation on big-Oh notation in section 2.6.

We begin our proof of Leopoldt's Conjecture in the abelian case with the proof of Brumer's Theorem (Theorem 19), which is the goal of section 3. This theorem is a p-adic variation of *Baker's Theorem*, proven by Baker in 1966 [3] (for which he received the Fields Medal), which states that for algebraic numbers $\alpha_1, \ldots, \alpha_n$, if the $\log \alpha_i$ are $\mathbb{Q}$-linearly independent, then they are $\overline{\mathbb{Q}}$-linearly independent. Brumer proved the analogous result with p-adic logarithms. While his proof closely follows the original by Baker, Brumer was able to make a number of simplifications as he was not trying to establish a computable lower bound on $\left|\beta_1 \log_p \alpha_1 + \cdots + \beta_n \log_p \alpha_n\right|$, but only to demonstrate that it is nonzero. Our proof presented here follows again

the general form of Brumer's proof very closely. Our Lemma 15 is a stronger form of Brumer's Lemma 1, which he states without proof (by analogy to Baker's Lemma 2); we found that by strengthening Lemma 15, certain later steps for proving Brumer's Theorem could be skipped altogether.

Given Brumer's Theorem, in section 4 we prove Leopoldt's Conjecture for abelian extensions K of $\mathbb{Q}$. The key step of this proof is Lemma 26, which is based on a technique by James Ax [2, Lemma, p. 2]. An alternative approach, also using Brumer's Theorem, is due to Washington [11, Theorem. 5.25]. In both proofs the essential step requires representation theory of abelian groups. As Ax proves Leopoldt's Conjecture in the form of the p-adic regulator being nonzero, which is not the form that we stated the conjecture, we need Lemmas 23 and 24 to prove that these two forms of the conjecture are equivalent (we will need these again for Theorem 48). The direction of the equivalence that we require is trivial, but the converse can be found in [11, Theorem. 5.31].

In [2, p.587], Ax alludes to extending his work to non-abelian cases, writing "[This method] would also give information even when the Galois group G of $K/\mathbb{Q}$ is not abelian.". As we have isolated the requirement that G is abelian to a single step of Lemma 26, we follow his advice in section 5 to give a Lemma 28, a trivial variation of Lemma 26 that assumes no knowledge of the irreducible representations of G of dimension greater than 1. This immediately gives us Theorem 29, a weak bound on the Leopoldt defect for galois extensions $K/\mathbb{Q}$.

In the second half of this paper we pursue a theorem of Waldschmidt, which uses a deep result of Masser to give a bound on the Leopoldt defect (which is in fact never weaker than the bound we found in Theorem 29).

In section 6 and section 7.2 we very closely follow the work of Masser in [8] with the goal to prove Theorem 39. We start in section 6.1 by stating a number of general results of commutative algebra, which can be found in [1] or [12]: namely, the essential results of primary decompositions and their behavior in localizations; Krull's height theorem; and the Hilbert polynomial. We also need a few facts related to the leading coefficient of the Hilbert polynomial proven in [5]. We then continue in the remainder of section 6 and in section 7.2 closely mimicking the technique of Masser.

This results in Masser's Theorem (Theorem 39), which we will use in section 7. Given a polynomial of exponential functions with zeros of a particular form, this theorem shows a relationship between the number of such zeros, the total degree of the polynomial, and the distribution of the zeros. In section 7.3, we construct (following the approach of Waldschmidt [10]) such a polynomial with many zeros and low degree, proving that the coefficient $\chi(X, Y)$ is small, which encodes the fact that the zeros are not highly concentrated in low-dimensional subspaces. The construction of the polynomial closely resembles the similar construction we performed in section 3.1; both use the pigeon-hole principle in a very similar way to show that the desired degree of the polynomial is high enough to find one that is

small in many places, and uses Theorem 12 to show that, as the values are algebraic numbers, they are so small that they must be zero.

We put together these facts in Theorem 47, loosely following the approach of Waldschmidt, to show that the dimension of the $\mathbb{C}_p$-space spanned by these zeros is at least half the rank of the zeros. Then in Theorem 48, we show that these zeros can be chosen to be the locations of the p-adic logarithms of the units of a number field K. While this does not show that these zeros are $\mathbb{C}_p$-linearly independent, we do find that the dimension of the $\mathbb{C}_p$-space spanned by them is at least half of their $\mathbb{Z}$-rank, which gives us our bound on the Leopoldt defect.

2 Notation and definitions

2.1 Construction of $\mathbb{C}_p$

Throughout this document we will fix a prime p. We denote by $\mathbb{Q}_p$ the field of p-adic numbers (i.e. the completion of $\mathbb{Q}$ under the p-adic valuation) and by $\overline{\mathbb{Q}_p}$ the algebraic closure of $\mathbb{Q}_p$. The valuation on $\mathbb{Q}_p$ can be uniquely extended to $\overline{\mathbb{Q}_p}$, and we denote by $\mathbb{C}_p$ the completion of $\overline{\mathbb{Q}_p}$ under that valuation. Then we naturally obtain a valuation $z \mapsto |z|$ on $\mathbb{C}_p$, whose normalization we fix by

$$|p| = \frac{1}{p}.$$

The logarithm of the p-adic valuation gives the p-adic order $v : \mathbb{C}_p \to \mathbb{Q}$ with $v(p) = 1$.

The field $\mathbb{Q}$ also carries the ordinary absolute value, which can be non-uniquely extended to all of $\mathbb{C}_p$. Let $z \mapsto |z|_\infty$ be the ordinary absolute value on $\mathbb{C}_p$. The field $\mathbb{C}_p$ under $|\cdot|_\infty$ is isomorphic to the field of complex numbers as a field-with-valuation, therefore let $\mathbb{C} = \mathbb{C}_p$. We will use the symbol $\mathbb{C}$ when we wish to emphasize our interest in the absolute value $|\cdot|_\infty$ as opposed to the valuation $|\cdot|$.

Let $\overline{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$, and let $\mathbb{A}$ be the integral closure of $\mathbb{Z}$ in $\overline{\mathbb{Q}}$, i.e., the ring of algebraic integers. We say a *number field* is a finite extension of $\mathbb{Q}$ in $\mathbb{C}$, and a *local field* is a finite extension of $\mathbb{Q}_p$ in $\mathbb{C}_p$. Then for every number field K and local field L we have inclusions

$$\mathbb{Q} \subset K \subset \overline{\mathbb{Q}} \subset \mathbb{C}$$
$$\mathbb{Q}_p \subset L \subset \overline{\mathbb{Q}_p} \subset \mathbb{C}_p.$$

For a number field K, let Θ_K be the set of all embeddings of K into $\mathbb{C}_p$; then $|\Theta_K| = [K : \mathbb{Q}]$. We write Θ for Θ_K when K is clear. Every $\theta \in \Theta$ induces a valuation on K that extends the p-adic valuation on $\mathbb{Q}$, so it must induce the $\mathfrak{p}$-adic valuation for some prime $\mathfrak{p}$ of K lying above p. Conversely, every $\mathfrak{p}$-adic valuation on K is induced by some $\theta \in \Theta$. For each prime $\mathfrak{p}$ of K lying above p, fix a choice of $\theta_\mathfrak{p} \in \Theta$ which induces the $\mathfrak{p}$-adic valuation on K. Let $K_\mathfrak{p}$ be the completion of $\theta_\mathfrak{p}(K)$ in $\mathbb{C}_p$, and let K_p be the compositum of the $K_\mathfrak{p}$ for all $\mathfrak{p}$ above p. In particular, $K_\mathfrak{p}$ and K_p are local fields.

Let $\mathcal{U} = \{z \in \mathbb{C}_p \mid |z| = 1\}$. For reals $0 < r \leq 1$, define

$$\mathcal{U}_r = \{z \in \mathbb{C}_p \mid |z - 1| < r\} \subset \mathcal{U}.$$

Each $u \in \mathcal{U}$ can be written uniquely as $u = \zeta \cdot u_1$ where $\zeta^n = 1$ for some n relatively prime to p and $u_1 \in \mathcal{U}_1$.

For a number field K let $\mathcal{O}_K = K \cap \mathbb{A}$ be the integral closure of $\mathbb{Z}$ in K; for a local field L let $\mathcal{O}_L = \{z \in L \mid |z| \leq 1\}$ be the integral closure of $\mathbb{Z}_p$ in L.

2.2 p-adic analytic functions

We give $\mathbb{C}_p^n$ the ultrametric norm; for $z = (z_i) \in \mathbb{C}_p^n$, let

$$|z| = \max_i |z_i| .$$

Then for real $r > 0$, let

$$B_n^-(r) = \left\{ z \in \mathbb{C}_p^n \;\middle|\; |z| < r \right\},$$
$$B_n(r) = \left\{ z \in \mathbb{C}_p^n \;\middle|\; |z| \leq r \right\}.$$

We will use multi-index notation in this document; for $\kappa \in \mathbb{Z}^n$, we write

$$z^\kappa = z_1^{\kappa_1} \cdots z_n^{\kappa_n},$$
$$\|\kappa\| = \kappa_1 + \cdots + \kappa_n,$$
$$\kappa! = \kappa_1! \cdots \kappa_n!,$$

where $z \in R^n$ for some ring R (generally $\mathbb{C}_p$).

We give $\mathbb{Z}^n$ the usual partial order (coordinate-wise comparison), and when indexing over κ we shall always implicitly mean only $\kappa \in \mathbb{Z}^n$ such that $\kappa \geq 0$ (i.e., each coordinate non-negative). Furthermore, for nonnegative real numbers x we write $\mathbb{Z}^n(x)$ to mean

$$\mathbb{Z}^n(x) = \{ a \in \mathbb{Z}^n \mid 0 \leq a_i \leq x \text{ for each } i \} .$$

Definition 1. A function $f : B_n(r) \to \mathbb{C}_p$ is *analytic* (on $B_n(r)$) if there exist $f_\kappa \in \mathbb{C}_p$ such that

$$\lim_{\|\kappa\| \to \infty} \left(|f_\kappa| \, r^{\|\kappa\|} \right) = 0$$

and $f(z) = \sum_{\kappa \geq 0} f_\kappa z^\kappa$ for all $z \in B_n(r)$. We say $f : B_n^-(r) \to \mathbb{C}_p$ is *analytic* (on $B_n^-(r)$) if it is analytic on $B_n(r')$ for all $0 < r' < r$.

If f is analytic on $B_n(r_0)$ then the coefficients f_κ of its power series are uniquely defined (and do not depend on r_0). Let the *radius of convergence* R of f be the supremum of the r such that $\lim_{\|\kappa\| \to \infty} \left(|f_\kappa| \, r^{\|\kappa\|} \right) = 0$; then we can extend f via its power series to all of $B_n^-(R)$. We will not need to deal with the case of the radius of convergence being infinite.

Now if $f = \sum f_\kappa z^\kappa$ is analytic on $B_n(r)$, we write

$$|f|_r = \sup_\kappa \left(|f_\kappa| \, r^{\|\kappa\|} \right) .$$

If L is a complete subfield of $\mathbb{C}_p$ containing all of the f_κ, then the restriction of f to $B_n(r) \cap L^n$ takes values in L. In this case we say that f is *L-analytic (on $B_n(r)$)*; similarly for $B_n^-(r)$.

As a first application, we define $\exp_p$ and $\log_p$ via their power series. Let

$$\varepsilon = p^{-1/(p-1)},$$

the radius of convergence of $\exp_p$. Now define

$$\log_p(1-z) = -\sum_{k=1}^{\infty} \frac{z^k}{k},$$
$$\exp_p z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

The former is $\mathbb{Q}_p$-analytic on $B_1^-(1)$, and the latter on $B_1^-(\varepsilon)$. In fact, $z \mapsto \log_p(1-z)$ and $z \mapsto 1 - \exp_p(z)$ are two-sided inverses from $B_1^-(\varepsilon)$ to $B_1^-(\varepsilon)$. Furthermore, the definition of $\log_p(z)$ can be extended uniquely to all of $\mathbb{C}_p^\times$ via $\log_p p = 0$ and $\log_p \zeta = 0$ for all $\zeta^n = 1$ with n relatively prime to p. These properties can be found in Propositions 5.4, 5.5, and 5.7 of [11].

Finally, we define a differential operator Δ for f analytic on $B_1(r)$, via

$$\Delta f = \sum_{k \geq 0} (k+1) f_{k+1} z^k.$$

Then Δf is analytic on $B_1(r)$, and satisfies the expected properties. For integers $m \geq 0$, let Δ^m be the iterated differential operator. One sees that f has a zero at $z_0 \in B_1(r)$ of multiplicity at least k if and only if $(\Delta^i f)(z_0) = 0$ for all $0 \leq i < k$.

2.3 p-adic exponentiation

We will define the notation u^z for $u \in \mathcal{U}_\varepsilon$ and $z \in B_1^-(r_u)$ for some $r_u > 1$ as an extension of the ordinary definition for $z \in \mathbb{Z}$. Before we can give this definition, we will need to investigate the convergence properties of u^k for $u \in \mathcal{U}_1, k \in \mathbb{Z}$.

Lemma 2. For $u \in \mathcal{U}_1$, we have

$$\left|\log_p u\right| \leq \frac{1}{-e \cdot \log|u-1|}.$$

If further $u \in \mathcal{U}_\varepsilon$, then $\left|\log_p u\right| = |u-1|$.

Proof. Write $u = 1 + z$, with $|z| < 1$. We have

$$\begin{aligned} v(\log_p u) &\geq \min_{a \geq 0} \left(v\left(\frac{z^{p^a}}{p^a} \right) \right) \\ &= \min_{a \geq 0} \left(p^a v(z) - a \right) \end{aligned}$$

Regarding $p^a v(z) - a$ as a function of $a \in \mathbb{R}$, we find it has a unique global minimum $a_0 \in \mathbb{R}$, given by $p^{a_0} = 1/(v(z)\log p)$ or $a_0 = -\log(v(z)\log p)/\log p$. Thus

$$v(\log_p u) \geq p^{a_0} v(z) - a_0 = \frac{1}{\log p} + \frac{\log(v(z)\log p)}{\log p} = \frac{\log(-e \cdot \log|z|)}{\log p},$$

which gives us the desired result. If additionally $u \in \mathcal{U}_\varepsilon$, then we find that $a_0 < 1$, so it suffices to consider only $a = 0$ and $a = 1$. Comparing shows that the minimum for $p^a v(z) - a$ is achieved at $a = 0$, so $v(\log_p u) = v(z)$. □

Lemma 3. For $u \in \mathcal{U}_1$ and nonnegative $k \in \mathbb{Z}$, we have

$$\left|u^k - 1\right| \leq |k| \cdot \frac{1}{-e \cdot \log|u-1|}.$$

If further $u \in \mathcal{U}_\varepsilon$, then $\left|u^k - 1\right| = |k(u-1)|$.

Proof. Write $u = 1 + z$, with $|z| < 1$. For nonnegative integers m, n, we have

$$\begin{aligned} v\left(u^{mp^n} - 1\right) &= v\left((1+z)^{mp^n} - 1\right) \\ &= v\left(mp^n z + \binom{mp^n}{2} z^2 + \cdots + z^{mp^n}\right) \\ &\geq \min_{a \geq 0} v\left(p^{n-a} z^{p^a}\right) \\ &= n + \min_{a \geq 0}(p^a v(z) - a). \end{aligned}$$

(This calculation required use of the p-adic valuation of binomial coefficients.) Proceeding as in the previous lemma, we eventually find that for nonnegative $k \in \mathbb{Z}$,

$$\left|u^{\pm k} - 1\right| = \left|u^k - 1\right| \leq |k| \cdot \frac{1}{-e \cdot \log|u-1|}.$$

If $u \in \mathcal{U}_\varepsilon$, we continue as in the previous lemma. In general, for $k \in \mathbb{Z}$ with $|k|$ sufficiently small, we have $\left|u^k - 1\right| = \left|k \log_p u\right|$, which can be derived in other ways (e.g. [11, Lemma 5.5]). □

Corollary 4. For every positive real $r, r' < 1$, we have $(\mathcal{U}_{r'})^{p^k} \subset \mathcal{U}_r$ for sufficiently large k. In particular, for every $u \in \mathcal{U}_1$, we have $u^{p^k} \in \mathcal{U}_r$ for sufficiently large k.

For $u \in \mathcal{U}_\varepsilon$, let $r_u = \varepsilon / \left|\log_p u\right| = \frac{\varepsilon}{|u-1|} > 1$. Then the function

$$\exp_p(z \log_p u)$$

is analytic on $z \in B_1^-(r_u)$. Now if $k \in B_1^-(r_u) \cap \mathbb{Z}$, we have

$$\exp_p(k \log_p u) = \exp_p(\log_p(u^k)) = u^k$$

by Lemmas 2 and 3. Thus for $u \in \mathcal{U}_\varepsilon$ and $z \in B_1^-(r_u)$, we define

$$u^z = \exp_p(z \log_p u).$$

One sees that $\log_p(u^z) = z \log_p u$ where defined, by taking $\exp_p$ (which is injective) of both sides.

While the above definition is sufficient for our purposes, we can alternatively define u^z for all $u \in \mathcal{U}_1$ and $z \in \mathbb{Z}_p$ as a limit of some u^{k_i} for a Cauchy sequence $k_i \in \mathbb{Z}$ converging to z. Then Corollary 4 shows us that this sequence u^{k_i} is also a Cauchy sequence and has a unique limit u^z. However we will not find it necessary to do this.

2.4 Leopoldt's Conjecture

For a number field K, let $E_K = \mathcal{O}_K^\times$ be the units of K, and let

$$E_{K,\varepsilon} = \{u \in E_K \mid \theta_\mathfrak{p}(u) \in \mathcal{U}_\varepsilon \text{ for every } \mathfrak{p}\}.$$

$E_{K,\varepsilon}$ is a subgroup of E_K, and we claim that they have the same $\mathbb{Z}$-rank. Certainly $\theta_\mathfrak{p}(E_K) \subset \mathcal{U}$. We know that E_K is finitely generated, so there exists n such that $(E_K)^n$ is torsion-free. Then $\theta_\mathfrak{p}((E_K)^n)$ is torsion-free, so it contains no roots of unity, and in fact $\theta_\mathfrak{p}((E_K)^n) \subset \mathcal{U}_1$. Then by Corollary 4, as $(E_K)^n$ is finitely generated, for sufficiently large k one has $\theta_\mathfrak{p}((E_K)^{np^k}) \subset \mathcal{U}_\varepsilon$. Then

$$(E_K)^{np^k} \subset E_{K,\varepsilon} \subset E_K,$$

and as E_K and $(E_K)^{np^k}$ have the same $\mathbb{Z}$-rank, so does $E_{K,\varepsilon}$.

Now define

$$\mathfrak{K} = K \otimes_\mathbb{Q} \mathbb{Q}_p = \prod_{\mathfrak{p} \mid p} K_\mathfrak{p},$$

and let $\theta_p = \prod \theta_\mathfrak{p}$ be the diagonal embedding $K \to \mathfrak{K}$. Then $\theta_p(E_{K,\varepsilon}) \subset \prod U_\varepsilon$. We can give the product on the right a structure as a $\mathbb{Z}_p$-module, as u^z is defined for $u \in U_\varepsilon, z \in \mathbb{Z}_p \subset B_1(1)$. Now let $\overline{E_{K,\varepsilon}}$ be the $\mathbb{Z}_p$-module generated by $\theta_p(E_{K,\varepsilon})$; equivalently, it is the closure under the topology from $\mathfrak{K}$.

Let $r_K = \operatorname{rank}_\mathbb{Z}(E_K) = \operatorname{rank}_\mathbb{Z}(E_{K,\varepsilon})$, and let $r_{K,p} = \operatorname{rank}_{\mathbb{Z}_p}(\overline{E_{K,\varepsilon}})$. A $\mathbb{Z}$-basis for $E_{K,\varepsilon}$ will generate $E_{K,\varepsilon}$ over $\mathbb{Z}_p$, so certainly $r_K \geq r_{K,p}$. However, it is not clear whether these generators will be $\mathbb{Z}_p$-linearly independent or not, and thus whether or not $r_K = r_{K,p}$.

It was conjectured by Leopoldt in [7] that for all number fields K, we in fact have equality.

Leopoldt's Conjecture. For all number fields K,

$$r_{K,p} = r_K.$$

For any number field K, we define the *Leopoldt defect* as

$$r_K - r_{K,p}.$$

2.5 p-adic Schwarz's Lemma

For an analytic function f, we will need to be able to relate $|f|_r$ to values of $f(z)$ for $z \in B_n(r)$. One direction is easy.

Lemma 5. Suppose f is analytic on $B_n(r)$. Then for all $z \in B_n(r)$, $|f(z)| \leq |f|_r$.

Proof. Let $m = |f|_r$. We need to show that $f(z) \in B_1(m)$ for $z \in B_n(r)$. As $B_1(m)$ is closed under addition and topologically closed, it is therefore complete, and closed under infinite (convergent) sums. Then as each $f_\kappa z^\kappa$ lies in $B_1(m)$, so does their sum $f(z)$, which gives us the result. □

We investigate the other direction, and find a p-adic analog to the maximum modulus principle. Before the proof, we need a quick computation.

Lemma 6. For nonzero $\alpha, \beta \in \mathbb{C}_p$, we have $|\alpha - \beta| = \max(|\alpha|, |\beta|)$ if and only if $\frac{\alpha}{\beta} \notin \mathcal{U}_1$.

Proof. If $|\alpha| \neq |\beta|$, the result is immediate. If $|\alpha| = |\beta|$, we divide through by β and see that the claim becomes the definition of $\mathcal{U}_1$. □

Theorem 7. Suppose f is analytic on $B_1(r)$, and suppose that there exists $z_0 \in \mathbb{C}_p$ such that $|z_0| = r$. Then there exists $z_1 \in \mathbb{C}_p$ with $|z_1| = r$ such that $|f(z_1)| = |f|_r$.

Proof. We can assume that f is nonzero. Write $f(z) = \sum_{i \geq 0} f_i z^i$, and let I be the set of indices $i \geq 0$ such that $|f_i| r^i = |f|_r$, which is nonempty and finite due to the convergence properties of the f_i. Let

$$g(z) = \sum_{i \in I} f_i z^i$$

and $h = f - g$, so that $|g|_r = |f|_r$ and $|h|_r < |f|_r$.

Let n be the degree of g, and write

$$g(z) = f_n \prod_{j=1}^{n} (z - \alpha_i)$$

for some $\alpha_i \in \mathbb{C}_p$. Let $\omega \in \mathcal{U}$ be a primitive k-th root of unity, with k relatively prime to p and larger than n. Let $z_0 \in \mathbb{C}_p$ be any element with $|z_0| = r$. Each $\alpha_i \mathcal{U}_1$ can contain at most one of the $z_0, z_0\omega, \ldots, z_0\omega^{k-1}$ (as $\omega^a \in \mathcal{U}_1$ implies k divides a), so one of them (call it z_1) does not lie in any of the $\alpha_i \mathcal{U}_1$. Then $|z_1| = r$ and by Lemma 6, $|z_1 - \alpha_i| = \max(r, |\alpha_i|)$ for all i. Then we have

$$|g(z_1)| = |f_n| \prod_{j=1}^{n} |z - \alpha_i| = |f_n| \prod_{j=1}^{n} \max(r, |\alpha_i|) \geq |f_n| r^n = |g|_r = |f|_r,$$

but $|g(z_1)| \leq |g|_r$ so $|g(z_1)| = |f|_r$. As $|h(z_1)| \leq |h|_r < |f|_r$, we find that

$$|f(z_1)| = |g(z_1) + h(z_1)| = |f|_r$$

as desired. □

As $\left|\mathbb{C}_p^\times\right|$ is dense in $(0, \infty)$ (in fact $v(\mathbb{C}_p^\times) = \mathbb{Q}$), we get the following corollary.

Corollary 8. Suppose f is analytic on $B_1(R)$. Then

$$|f|_R = \sup_{|z| \leq R} |f(z)| .$$

Indeed for any $0 < r < R$,

$$|f|_R = \sup_{r < |z| \leq R} |f(z)| .$$

The following is an immediate corollary of the preceding in the case that $n = 1$. For $n > 1$, we can instead prove the result by computing directly from the definition of $|\cdot|_r$; we will omit these computations.

Theorem 9. If f and g are analytic on $B_n(r)$, then

$$|f + g|_r \leq \max(|f|_r , |g|_r)$$
$$|fg|_r \leq |f|_r \cdot |g|_r .$$

From the p-adic maximum modulus principle, we can derive a p-adic form of Schwarz's Lemma.

Theorem 10. Suppose f is analytic on $B_1(R)$, and that for some $0 < r < R$, f has at least n roots (counting multiplicities) in $B_1(r)$. Then

$$|f|_r \leq \left(\frac{r}{R}\right)^n |f|_R .$$

Proof. Write

$$f(z) = (z - \alpha_1) \cdots (z - \alpha_n) g(z)$$

for some $\alpha_i \in B_1(r)$ and $g(z)$ analytic on $B_1(R)$. Now for $z \in \mathbb{C}_p$ with $r < |z| \leq R$, we have $|f(z)| = |z|^n |g(z)|$, so by Corollary 8 we have $|f|_R = R^n |g|_R$.

Now for all $z \in B_1(r)$, we have

$$|f(z)| \leq r^n |g(z)| \leq r^n |g|_R = \left(\frac{r}{R}\right)^n |f|_R ,$$

so taking the supremum over all $z \in B_1(r)$ and applying Corollary 8 gives us the claim. □

2.6 Alternative asymptotic notation

We will frequently find ourself interested in the asymptotic growth behavior of functions of real numbers; typically one uses big-Oh notation to capture these semantics, writing $f(x) = O(g(x))$ to describe the growth of a function $f(x)$ in terms of the growth of a (presumably simpler) function $g(x)$. This notation allows us to neglect constants and lower order terms which are not of interest to us.

However, big-Oh notation changes the semantics of the $=$ symbol (e.g. under the Knuth standard for big-Oh notation, $=$ is no longer symmetric, so that we can have $A = B$ but not $B = A$ [6, p. 108]), and makes it incompatible with standard algebraic manipulations. To avoid this problem, we will not use big-Oh notation in this document, and instead use standard tools from mathematics to capture the same concept in a manner that is compatible with algebraic manipulations.

Let $\mathcal{F}$ be the set of all functions $f : D_f \to \mathbb{R}$ such that $D_f \subset \mathbb{R}$ is not bounded from above. (In most cases, either $D_f = \mathbb{Z}$ or $D_f = \mathbb{R}$.) We will define a partial order $\leq_O$ on $\mathcal{F}$.

Definition 11. We write $f \leq_O g$ if there exist $x_0, c \in \mathbb{R}$ with $c > 0$ such that $D_f \cap [x_0, \infty) = D_g \cap [x_0, \infty)$ and $f(x) \leq cg(x)$ for all $x \in D_f \cap [x_0, \infty)$.

We may write $f(x) \leq_{O,x} g(x)$ if we need to make the choice of variable x explicit. Having defined $\leq_O$, we immediately get

$f \geq_O g$	if $g \leq_O f$
$f =_O g$	if $f \leq_O g$ and $g \leq_O f$
$f <_O g$	if $f \leq_O g$ but not $f =_O g$
$f >_O g$	if $g <_O f$.

Of course this notation does not obviate the need to be careful, especially when working with signed quantities, as for example $x^3 =_O \frac{1}{2}x^3$ but $x^3 - \frac{1}{2}x^3 \neq_O 0$.

2.7 Bounds on nonzero algebraic elements

For a number field K, let P_K be the set of all valuations (i.e., "finite primes") normalized in the same way as $x \mapsto |x|$. In particular, $x \mapsto |x|$ is an element of P_K, and the product rule holds. Then as a consequence of the product rule [9, Theorem 8.8] we get the following result.

Theorem 12. Suppose K is a number field, and let $\{\sigma_i\}$ be all of the embeddings $\sigma_i : K \to \mathbb{C}$. Suppose $x \in K^\times$ and $D \in \mathbb{Z}$ is nonzero such that $Dx \in \mathbb{A}$ is an algebraic integer. Then

$$|x| \geq \frac{1}{(D \max_i |\sigma_i x|_\infty)^{[K:\mathbb{Q}]}}.$$

Proof. Let $y = Dx \in \mathbb{A} \cap K^\times$. Then the product rule states that

$$\prod_i |\sigma_i y|_\infty \cdot \prod_{abs \in P_K} abs(y) = 1,$$

but as $y \in \mathbb{A}$, we have $abs(y) \leq 1$ for all $abs \in P_K$. Then dividing out all of the $abs \in P_K$ except for $x \mapsto |x|$, we find

$$\prod_i |\sigma_i y|_\infty \cdot |y| \leq 1$$

so

$$|x| \geq |Dx| \geq \left(\prod |\sigma_i Dy|\right)^{-1} \geq \left(\max_i |D\sigma_i y|^{[K:\mathbb{Q}]}\right)^{-1},$$

which gives us the result. □

At times the logarithmic form is more convenient to work with:

$$\log |x| \geq -[K:\mathbb{Q}] \log(D \max_i |\sigma_i x|_\infty).$$

2.8 Vandermonde determinant

A *Vandermonde matrix* is a special form of matrix which contains iterated powers of some elements $x_1, \ldots, x_n$. A classical result is that if the determinant of a Vandermonde matrix is zero, then some two of the x_i are equal. We prove an equivalent form of that result here.

Theorem 13. Let F be a field and $x_1, \ldots, x_n \in F^\times$. Suppose there exist $a_1, \ldots, a_n \in F$, not all zero, such that

$$\sum_{i=1}^n a_i x_i^j = 0$$

for $1 \leq j \leq n$. Then for some $i \neq j$ we have $x_i = x_j$.

Proof. We prove this by induction on n, the case $n = 1$ being immediate.

The (x_i^j) give us a linear function $F^n \to F^n$, with $(a_1, \ldots, a_n)$ a nonzero element in the kernel. Therefore we have a nonzero linear relation $(b_1, \ldots, b_n)$ on the image,

$$\sum_{j=1}^n b_j x_i^j = 0$$

for all $1 \leq i \leq n$. Let

$$\begin{aligned} c_0 &= b_1 + \cdots + b_n x_1^{n-1} = 0, \\ c_1 &= b_2 + \cdots + b_n x_1^{n-2}, \\ c_2 &= b_3 + \cdots + b_n x_1^{n-3}, \\ &\cdots \\ c_{n-2} &= b_{n-1} + b_n x_1, \\ c_{n-1} &= b_n. \end{aligned}$$

(We know $c_0 = 0$ as $c_0 x_1 = 0$ but $x_1 \neq 0$.) We claim that the $(c_1, \ldots, c_{n-1})$ gives us a relation on the $(x_i, \ldots, x_i^{n-1})$ for all i such that $x_i \neq x_1$. We compute

$$\begin{aligned} &(x_i - x_1)(c_0 + c_1 x_i + \cdots + c_{n-1} x_i^{n-1}) \\ &= c_{n-1} x_i^n + (c_{n-2} - x_1 c_{n-1}) x_i^{n-1} + \cdots + (c_1 - x_1 c_2) x_i^2 + (c_0 - x_1 c_1) x_i \\ &= b_n x_i^n + b_{n-1} x_i^{n-1} + \cdots + b_1 x_i = 0. \end{aligned}$$

As the b_i are not all zero, therefore the c_i are not all zero. Now if $x_i = x_1$ for some $i > 1$, we are done; in not, then the $(c_1, \ldots, c_{n-1})$ are a relation on the $(x_i, \ldots, x_i^{n-1})$ for all $i > 1$, as we have shown, so by induction we have the result. $\square$

3 Brumer's Theorem on $\overline{\mathbb{Q}}$-relations of logarithms

3.1 Construction of a function with many zeros

We will use the pigeon-hole principle to find a large degree polynomial with many zeros. The key idea behind our use of the pigeon-hole principle is captured by the following lemma.

Lemma 14. Suppose m, n, A are positive integers with $n > 2m$, and $a_{i,j}, 1 \leq i \leq m, 1 \leq j \leq n$ are integers with $|a_{i,j}|_\infty \leq A$. Then there exist integers $b_1, \ldots, b_n$, not all zero, with $|b_j|_\infty \leq B = An + 2$ such that

$$\sum_{j=1}^{n} a_{i,j} b_j = 0$$

for all $1 \leq i \leq m$.

Proof. Let $f_i : \mathbb{Z}^n \to \mathbb{Z}$ be the linear transformation sending $b \in \mathbb{Z}^n$ to $\sum_j a_{i,j} b_j$. There are at least $(B-1)^n$ different $b \in \mathbb{Z}^n$ with $|b_j|_\infty < \frac{B}{2}$; for each such b, we find that $|f_i(b)|_\infty < \frac{1}{2}ABn$. Thus there are fewer than $(ABn+1)^m$ possible distinct values of $(f_1(b), \ldots, f_m(b))$. However $(B-1)^2 = ABn + 1$, so

$$(B-1)^n > (B-1)^{2m} = (ABn+1)^m$$

and by the pigeon-hole principle we have two different tuples b, b' with the same image under f. Then the difference $b - b'$ gives the desired result. □

In the statement of Brumer's Theorem, we are given $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{Q}}$ with some kind of relation on their logarithms. We will be interested in polynomials in the α_i^z for some z, so we will have to make sure that α_i^z is well-defined. For now, we assume that $\alpha_1, \ldots, \alpha_n \in \mathcal{U}_\varepsilon \cap \overline{\mathbb{Q}}$, and that

$$\log_p \alpha_n = \beta_1 \log_p \alpha_1 + \cdots + \beta_{n-1} \log_p \alpha_{n-1}$$

for some $\beta_1, \ldots, \beta_{n-1} \in \mathbb{A}$; we will justify these restrictions at the very end. Let $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_{n-1})$, a number field, and let $d = [K : \mathbb{Q}]$. For any $\lambda \in \mathbb{Z}^n$, let $\mu \in \mathbb{Z}^{n-1}$ be given by

$$\mu_i = \lambda_i + \lambda_n \beta_i.$$

As $\alpha_i \in \mathcal{U}_\varepsilon$, we have $r_{\alpha_i} > 1$ (see section 2.3), and let $r_\alpha = (1/2)(1+\min(r_{\alpha_1}, \ldots, r_{\alpha_n}))$. Then $1 < r_\alpha < r_{\alpha_i}$ and $(\alpha^\lambda)^z$ is analytic on $B_1(r_\alpha)$ for all $\lambda \in \mathbb{Z}^n$.

Now given a collection of coefficients $a : \mathbb{Z}^n \to \mathbb{Z}$, only finitely many nonzero, and given $s \in \mathbb{Z}^{n-1}, s \geq 0$, we define a function

$$Q_a(s, z) = \sum_\lambda a(\lambda)(\alpha^\lambda)^z \mu^s$$

which is analytic on $z \in B_1(r_\alpha)$.

Given the above definitions, the following lemma allows us to construct integer coefficients $a(\lambda)$ such that many of the $Q_a(s,z)$ have many zeros.

Lemma 15. Suppose $\alpha_1, \dots, \alpha_n \in \mathcal{U}_\varepsilon \cap \overline{\mathbb{Q}}$ and that

$$\log_p \alpha_n = \beta_1 \log_p \alpha_1 + \cdots + \beta_{n-1} \log_p \alpha_{n-1}$$

for some $\beta_1, \dots, \beta_{n-1} \in \mathbb{A}$. Define K, d, μ, and $Q_a(s,z)$ as above. For real numbers x, let $L_x = x^{2-1/(2n)}$.

Then for all sufficiently large x, there exist integers $a_x(\lambda)$, for $\lambda \in \mathbb{Z}^n(L_x)$, not all $a_x(\lambda)$ zero, with $|a_x(\lambda)|_\infty < e^{x^3}$, and such that

$$Q_{a_x}(s,l) = 0$$

for all $s \in \mathbb{Z}^{n-1}(x^2)$ and $1 \le l \le x^{1+1/(4n)}$.

Proof. We will be interested in finding linear combinations of

$$(\alpha^\lambda)^l \mu^s$$

which are zero, using Lemma 14. However, the lemma is only applicable for finding linear combinations of integers, whereas the $(\alpha^\lambda)^l \mu^s$ are elements from the field K. To remedy this, let $f_1, \dots, f_d$ be an integral basis for K; that is, let $\mathbb{A} \cap K = \mathbb{Z}f_1 + \cdots + \mathbb{Z}f_d$. For $z \in \mathbb{A} \cap K$, we write

$$z = C_1(z) f_1 + \cdots C_d(z) f_d.$$

Now we will apply Lemma 14 to find integers $a_x(\lambda)$, not all zero, such that

$$\sum_{\lambda \in \mathbb{Z}^n(L_x)} a_x(\lambda) C_i((\alpha^\lambda)^l \mu^s) = 0$$

for all $1 \le i \le d$, $1 \le l \le x^{1+1/(4n)}$ and $s \in \mathbb{Z}^{n-1}(x^2)$. The number of equations to be satisfied is

$$\begin{aligned} m_x &= d \cdot \lfloor x^{1+1/(4n)} \rfloor \cdot (\lfloor x^2 + 1 \rfloor)^{n-1} \\ &=_{O,x} x^{1+1/(4n)} \cdot x^{2(n-1)} \cdot 1 = x^{2n-1+1/(4n)} \end{aligned}$$

whereas the number of variables $a_x(\lambda)$ is

$$n_x =_O L_x^n = x^{2n-(1/2)},$$

so $n_x >_O m_x$, and for sufficiently large x, $n_x > 2m_x$ and the lemma is applicable. (We delay finding a bound A_x on the $C_i((\alpha^\lambda)^l \mu^s)$ momentarily.)

Thus the lemma gives us integers $a_x(\lambda)$ such that $\sum a_x(\lambda)C_i((\alpha^\lambda)^l\mu^s) = 0$. Therefore we have

$$\begin{aligned} Q_{a_x}(s,l) &= \sum_{\lambda\in\mathbb{Z}^n(L_x)} a_x(\lambda)(\alpha^\lambda)^l\mu^s \\ &= \sum_{i=1}^{d}\sum_{\lambda\in\mathbb{Z}^n(L_x)} a_x(\lambda)C_i((\alpha^\lambda)^l\mu^s)f_i \\ &= \sum_{i=1}^{d} 0\cdot f_i = 0. \end{aligned}$$

All that remains is to establish the stated bound on the $a_x(\lambda)$. For $z\in\mathbb{A}\cap K$, let $C(z) = \max_i |C_i(z)|_\infty$. Let H be a maximum of the $C(\alpha_i), C(\beta_i)$, and $d\cdot C(f_if_j)$ for the appropriate ranges of i and j. We find that for $z_1,\dots,z_m\in\mathbb{A}\cap K$,

$$C(z_1\cdots z_m) \le H^{m-1}\prod_j C(z_j).$$

Now we have sufficient information to bound the $C_i((\alpha^\lambda)^l\mu^s)$. Fix $s\in\mathbb{Z}^{n-1}(x^2)$; also fix integer l with $1\le l\le x^{1+1/(4n)}$. Then we have

$$\begin{aligned} (\alpha^\lambda)^l\mu^s &= \alpha_1^{l\lambda_1}\cdots\alpha_n^{l\lambda_n}\prod_{i=1}^{n-1}\sum_{j=0}^{s_i}\binom{s_i}{j}(\beta_i\lambda_n)^j\lambda_i^{s_i-j} \\ C((\alpha^\lambda)^l\mu^s) &\le H^{2lL_x+2nx^2}\prod_{i=1}^{n-1}\sum_{j=0}^{s_i}\binom{s_i}{j}\lambda_n^j\lambda_i^{s_i-j} \\ &\le H^{2lL_x+2nx^2}\prod_{i=1}^{n-1}L_x^{s_i}\sum_{j=0}^{s_i}\binom{s_i}{j} \\ &\le H^{2lL_x+2nx^2}(2L_x)^{x^2(n-1)} \\ \log C((\alpha^\lambda)^l\mu^s) &\le (2lL_x+2nx^2)\log H + x^2(n-1)\log(2L_x) <_O x^3, \end{aligned}$$

where we have used that $lL_x \le_O x^{1+1/(4n)}x^{2-1/(2n)} = x^{3-1/(4n)} <_O x^3$.

We have $\log A_x <_O x^3$. Now $N_x =_O x^{2n-(1/2)}$, so

$$\log B = \log(A_xN_x+2) =_O \log A_x + \log N_x <_O x^3,$$

and for sufficiently large x we have $|a_x(\lambda)|_\infty < e^{x^3}$. □

3.2 Finding more zeros

Our definition of $Q_a(s,z)$ was chosen so that we would have a nice relationship with the derivatives (with respect to z). For $m \geq 0$, we find

$$\begin{aligned}(\Delta^m Q_a(s,\cdot))(z) &= \sum_\lambda a(\lambda)(\alpha^\lambda)^z \mu^s (\lambda_1 \log_p \alpha_1 + \cdots + \lambda_n \log_p \alpha_n)^m \\ &= \sum_\lambda a(\lambda)(\alpha^\lambda)^z \mu^s (\mu_1 \log_p \alpha_1 + \cdots + \mu_{n-1} \log_p \alpha_{n-1})^m \\ &= \sum_{t \in \mathbb{Z}^{n-1}, \|t\|=m} \sum_\lambda a(\lambda)(\alpha^\lambda)^z \mu^s \frac{m!}{t!} \mu^t (\log_p \alpha)^t \\ &= \sum_{t \in \mathbb{Z}^{n-1}, \|t\|=m} \frac{m!}{t!} (\log_p \alpha)^t Q_a(s+t,z).\end{aligned}$$

Then, for small s, the a_x we constructed in the previous lemma gives us $Q_{a_x}(s,z)$ with zeros of high multiplicity (roughly of the order x^2). Then Schwarz's Lemma shows us that $Q_{a_x}(s,z)$ must be very small for $z \in B_1(1)$, which allows us to find even more zeros (though of smaller order). In each iteration, we find $x^{1/(4n)}$ as many zeros as before, but of half the order; for large x this allows us to find very many zeros.

Lemma 16. Given the same conditions as in Lemma 15, let the $a_x(\lambda)$ be the coefficients constructed by the lemma. For integers $j \geq 0$, let

$$\begin{aligned}R_j &= \lfloor x^{1+(j+1)/(4n)} \rfloor, \\ S_j &= \lfloor x^2/2^j \rfloor.\end{aligned}$$

Then for every integer $j \geq 0$, for sufficiently large x we have

$$Q_{a_x}(s,l) = 0$$

for all $1 \leq l \leq R_j$ and $s \in \mathbb{Z}^{n-1}(S_j)$.

Proof. We induct on j, the case $j = 0$ being given by Lemma 15. Now suppose the claim is proven for $j = k$ and we wish to prove it for $j = k+1$.

Let $g(z) = Q_{a_x}(s,z)$ for some fixed s with $\|s\| \leq S_{k+1}$; we need to prove that $g(l) = 0$ for all $1 \leq l \leq R_{k+1}$. We will show that, as g has a large number of zeros in $B_1(1)$, $|g(l)|$ must be so small that by Theorem 12 it is zero. First we count how many zeros g has in $B_1(1)$. By the induction hypothesis, we know that for every $0 \leq m \leq S_{k+1}$ and $1 \leq l \leq R_k$, we have

$$(\Delta^m g)(l) = \sum_{\|t\|=m} \frac{m!}{t!} (\log_p \alpha)^t Q_{a_x}(s+t,l) = 0,$$

as $s+t \in \mathbb{Z}^{n-1}(2S_{k+1}) \subset \mathbb{Z}^{n-1}(S_k)$. In particular, g has at at least $R_k S_{k+1}$ zeros (counting multiplicity) in $B_1(1)$. Then by Theorem 10, we have

$$|g|_1 \leq r_\alpha^{-R_k S_{k+1}} |g|_{r_\alpha}.$$

Recalling that $\sup_{|z|<\varepsilon} \left|\exp_p(z)\right| = 1$, we have (by Corollary 8)

$$|g|_{r_\alpha} = \sup_{z \in B_1(r_\alpha)} |g(z)| \leq \sup \left| \sum_\lambda a_x(\lambda)(\alpha^\lambda)^z \mu^s \right| \leq \max_{1 \leq i \leq n-1} \left|\log_p \alpha_i\right|^{\|s\|}.$$

Thus

$$\begin{aligned} \frac{1}{|g|_1} &\geq r_\alpha^{R_k S_{k+1}} \max_i \left|\log_p \alpha_i\right|^{-\|s\|}, \\ -\log |g|_1 &\geq (R_k S_{k+1}) \log r_\alpha - x^2(n-1) \max_i \left|\log_p \alpha_i\right| \\ &=_O x^{1+(k+1)/(4n)} \cdot x^2 = x^{3+(k+1)/(4n)}. \end{aligned}$$

Now, assume towards a contradiction that $g(l)$ is nonzero for some $1 \leq l \leq R_{k+1}$. Let $M \in \mathbb{Z}$ be so large that for all i

$$|\sigma\alpha_i|_\infty < M \qquad \text{and} \qquad |\sigma\beta_i|_\infty < M$$

for all embeddings $\sigma : K \to \mathbb{C}$, and $M\alpha_i \in \mathbb{A}$. Then for all embeddings $\sigma : K \to \mathbb{C}$, we have

$$\begin{aligned} |\sigma g(l)|_\infty = |\sigma Q_{a_x}(s,l)|_\infty &\leq L_x^n e^{x^3} M^{nlL_x} ((M+1)L_x)^{\|s\|} \\ \log |\sigma Q_{a_x}(s,l)|_\infty &\leq n \log L_x + x^3 + nlL_x \log M + \|s\| \log((M+1)L_x), \\ &=_O \log x + x^3 + x^{1+(k+2)/(4n)} x^{2-1/(2n)} + x^2(n-1)\log x \\ &=_O x^3 + x^{3+k/(4n)} =_O x^{3+k/(4n)} \end{aligned}$$

Now $M^{nlL_x} g(l) \in \mathbb{A}$ and is nonzero, so by Theorem 12 we have

$$\begin{aligned} -\log |g(l)| &\leq d(\log M^{nlL_x} + \log \max_\sigma |\sigma g(l)|_\infty) \\ &\leq_O lL_x + x^{3+k/(4n)} \\ &\leq_O x^{1+(k+2)/(4n)} x^{2-1/(2n)} + x^{3+k/(4n)} \\ &=_O x^{3+k/(4n)}. \end{aligned}$$

Combining with our other bound, we find

$$x^{3+(k+1)/(4n)} \leq_O -\log |g|_1 \leq -\log |g(l)| \leq_O x^{3+k/(4n)},$$

which is impossible for sufficiently large x. Thus we have that $g(l) = 0$ for all $1 \leq l \leq R_{k+1}$, which gives the lemma. □

As we have found sufficiently many zeros, we are no longer interested in the order of the zeros and thus do not need $Q_a(s,z)$ for $s \neq 0$. Now for any collection $a(\lambda)$ of coefficients, finitely many nonzero, f_a be defined by

$$f_a(z) = Q_a(0,z) = \sum_\lambda a(\lambda)(\alpha^\lambda)^z \mu^0 = \sum_\lambda a(\lambda)(\alpha^\lambda)^z.$$

Then as an immediate corollary of Lemma 16 we get

Corollary 17. Given the same conditions as in Lemma 15, let the $a_x(\lambda)$ be the coefficients constructed by the lemma. For every real $t \geq 0$, for sufficiently large x we have

$$f_{a_x}(l) = 0$$

for all $1 \leq l \leq x^t$.

3.3 Brumer's Theorem

The function constructed by Corollary 17 allows us to impose so many relations on the powers of the α_i that the Vandermonde theorem gives us a $\mathbb{Q}$-relation on their logarithms.

Lemma 18. If $\alpha_1, \ldots, \alpha_n$ are in $\mathcal{U}_\varepsilon \cap \overline{\mathbb{Q}}$ and

$$\log_p \alpha_n = \beta_1 \log_p \alpha_1 + \cdots + \beta_{n-1} \log_p \alpha_{n-1}$$

for some $\beta_i \in \mathbb{A}$, not all zero, then there exist $b_i \in \mathbb{Q}$, not all zero, such that

$$b_1 \log_p \alpha_1 + \cdots + b_{n-1} \log_p \alpha_{n-1} + b_n \log_p \alpha_n = 0$$

Proof. We can apply Corollary 17 to the α_i with $t = 2n$. Then

$$(L_x + 1)^n =_O x^{2n-1/2} <_O x^t,$$

so $(L_x+1)^n < x^t$ for sufficiently large x. Then for all $1 \leq l \leq \lfloor L_x + 1 \rfloor^n$, we have

$$\sum_{\lambda \in \mathbb{Z}^n(L_x)} a_x(\lambda) \left(\alpha^\lambda\right)^l = 0.$$

Now by Theorem 13, $\alpha^\lambda = \alpha^{\lambda'}$ for some $\lambda \neq \lambda'$. That is, taking logarithms we have

$$(\lambda_1 - \lambda_1') \log_p \alpha_1 + \cdots + (\lambda_n - \lambda_n') \log_p \alpha_n = 0,$$

as desired. □

All that remains to prove Brumer's Theorem is to put the α_i in the correct form to use the preceding results.

Theorem 19. If $\alpha_1, \ldots, \alpha_n$ are in $\overline{\mathbb{Q}}^\times$ and the $\log_p(\alpha_i)$ are $\mathbb{Q}$-linearly independent, then they are also $\overline{\mathbb{Q}}$-linearly independent.

Proof. Suppose that

$$\beta_1 \log_p \alpha_1 + \cdots + \beta_n \log_p \alpha_n = 0$$

with $\beta_i \in \overline{\mathbb{Q}}$, not all zero. We can suppose that β_n is nonzero.

Write $\alpha_i = w_i \alpha_i'$ where $\log_p w_i = 0$ and $\alpha_i' \in \mathcal{U}_1 \cap \overline{\mathbb{Q}}$ (choose w_i to be the appropriate power of p times the appropriate root of unity). Then we get

$$\log_p \alpha_n' = \beta_1' \log_p \alpha_1' + \cdots + \beta_{n-1}' \log_p \alpha_{n-1}'$$

where $\beta_i' = -\beta_i/\beta_n$. Choose a positive integer C_1 such that $(\alpha_i')^{C_1} \in \mathcal{U}_\varepsilon$ for all i (which exists by Corollary 4), and a positive integer C_2 such that $C_2\beta_i' \in \mathbb{A}$ for $1 \leq i \leq n-1$. Then we have a relation

$$\log_p(\alpha_n')^{C_1 C_2} = (C_2\beta_1') \log_p(\alpha_1')^{C_1} + \cdots + (C_2\beta_{n-1}') \log_p(\alpha_{n-1}')^{C_1},$$

so by Lemma 18 there exist $b_i \in \mathbb{Q}$, not all zero, such that

$$b_1 \log_p(\alpha_1')^{C_1} + \cdots + b_{n-1} \log_p(\alpha_{n-1}')^{C_1} + b_n \log_p(\alpha_n')^{C_1 C_2} = 0,$$

so therefore

$$b_1 \log_p \alpha_1 + \cdots + b_{n-1} \log_p \alpha_{n-1} + C_2 b_n \log_p \alpha_n = 0,$$

showing that the $\log_p \alpha_i$ are $\mathbb{Q}$-linearly dependent, as desired. □

4 Leopoldt's Conjecture for abelian extensions

Throughout this section, fix a number field K. Let $\Upsilon = (E_{K,\varepsilon})^{\Theta}$, that is, tuples of elements of $E_{K,\varepsilon}$ indexed by $\Theta = \Theta_K$. In the case that $K/\mathbb{Q}$ is Galois, each embedding $\theta \in \Theta$ sends K to itself, so $\Theta = G = \mathrm{Gal}(K/\mathbb{Q})$. Then we say that $\upsilon \in \Upsilon$ is *compatible with the galois action* if there exists $u \in E_{K,\varepsilon}$ such that $\upsilon_\theta = \theta u$. In particular, $\mathrm{Gal}(K/\mathbb{Q})$ acts on υ by permuting its coordinates.

For $\upsilon \in \Upsilon$ and a subring A of $\mathbb{C}_p$ we define $\psi_\upsilon : \mathbb{Z}_p^\Theta \to \overline{E_{K,\varepsilon}}$ and $\phi_{A,\upsilon} : A^\Theta \to \mathbb{C}_p^\Theta$ by

$$(\psi_\upsilon(\alpha))_{\mathfrak{p}} = \prod_{\theta\in\Theta} (\theta_{\mathfrak{p}}\upsilon_\theta)^{\alpha_\theta},$$
$$(\phi_{A,\upsilon}(\alpha))_\delta = \sum_{\theta\in\Theta} \alpha_\theta(\log_p(\delta\upsilon_\theta)).$$

(We see that $\phi_{\mathbb{Z}_p,\upsilon}$ is related to ψ_υ by taking the logarithm and indexing over all of Θ and not just the $\theta_{\mathfrak{p}}$.)

When $K/\mathbb{Q}$ is Galois, we say that a unit $u \in E_K$ is a *Minkowski unit* if $\langle gu\rangle_{g\in G}$ has finite index in E_K. The existence of Minkowski units is one of the essential steps in proving Dirichlet's unit theorem. It can be proven directly, or by working backwards from Dirichlet's unit theorem; neither would be instructive here, so we give the claim without proof. (See [11, Lemma 5.27] for proof.)

Theorem 20. If $K/\mathbb{Q}$ is Galois, there exists a Minkowski unit $u \in E_K$.

We get an easy corollary.

Corollary 21. If $K/\mathbb{Q}$ is Galois, there exists a Minkowski unit u with $u \in E_{K,\varepsilon}$.

Proof. Given u by Theorem 20, u^k suffices when k is so large that $E_K^k \subset E_{K,\varepsilon}$. □

We wish to prove that Leopoldt's Conjecture holds for K when $K/\mathbb{Q}$ is abelian.

Theorem 22. If $K/\mathbb{Q}$ is abelian, then

$$r_{K,p} \geq r_K,$$

so Leopoldt's Conjecture holds for K.

Proof. Let $u \in E_K$ be a Minkowski unit, and let $\upsilon \in \Upsilon$ be given by $\upsilon_\theta = \theta u$. Then we see that υ is compatible with the galois action and that $\langle \upsilon_\theta\rangle_{\theta\in\Theta} = \langle gu\rangle_{g\in G}$ has finite index in E_K. Then by Lemmas 23, 24, 26, and 27, we have

$$\begin{aligned}
\mathrm{rank}_{\mathbb{Z}_p}(\ker\psi_\upsilon) &\leq \mathrm{rank}_{\mathbb{Z}_p}(\ker\phi_{\mathbb{Z}_p,\upsilon}) \\
&\leq \dim_{\mathbb{C}_p}(\ker\phi_{\mathbb{C}_p,\upsilon}) \\
&= \dim_{\overline{\mathbb{Q}}}(\ker\phi_{\overline{\mathbb{Q}},\upsilon}) \\
&\leq [K:\mathbb{Q}] - r_K,
\end{aligned}$$

so we conclude that

$$\operatorname{rank}_{\mathbb{Z}_p}(\overline{E_{K,\varepsilon}}) \geq \operatorname{rank}_{\mathbb{Z}_p}(\operatorname{im}\psi_\upsilon) = [K:\mathbb{Q}] - \operatorname{rank}_{\mathbb{Z}_p}(\ker\psi_\upsilon) \geq r_K.$$

□

Our proof divides the problem into four steps, which we prove in the following four lemmas. The first two, taken together, amount to the equivalence of Leopoldt's conjecture as stated here with the non-vanishing of the p-adic regulator; we skip past the question of the vanishing of the p-adic regulator to avoid the difficulties in defining it. The third step uses that G is abelian in a key way. The fourth step amounts to an invocation of Theorem 19 by Brumer.

Later in this document we will give a bound, due to Waldschmidt, of the Leopoldt defect $r_K - r_{K,p}$; as Waldschmidt gave his original argument in terms of the p-adic regulator, we will need Lemmas 23 and 24 again later.

Lemma 23. For $\upsilon \in \Upsilon$ we have

$$\ker\psi_\upsilon \subset \ker\phi_{\mathbb{Z}_p,\upsilon}.$$

Proof. Suppose $\alpha \in \ker\psi_\upsilon$, so that for all $\mathfrak{p}$ above p we have

$$\prod_{\theta\in\Theta}(\theta_{\mathfrak{p}}\upsilon_\theta)^{\alpha_\theta} = 1.$$

We claim that the same equation holds with $\theta_{\mathfrak{p}}$ replaced by any $\delta \in \Theta$. For any such δ, there exists $\mathfrak{p}$ above p such that $|\delta(z)| = |\theta_{\mathfrak{p}}(z)|$ for all $z \in K$. Thus for any $\delta \in \Theta$, we have for $a \in \mathbb{Z}^\Theta$

$$\left|\prod_{\theta\in\Theta}(\delta\upsilon_\theta)^{a_\theta} - 1\right| = \left|\delta\left(\prod_{\theta\in\Theta}\upsilon_\theta^{a_\theta} - 1\right)\right| = \left|\theta_{\mathfrak{p}}\left(\prod_{\theta\in\Theta}\upsilon_\theta^{a_\theta} - 1\right)\right| = \left|\prod_{\theta\in\Theta}(\theta_{\mathfrak{p}}\upsilon_\theta)^{a_\theta} - 1\right|,$$

so by letting the $a \in \mathbb{Z}^\Theta$ approach $\alpha \in \mathbb{Z}_p^\Theta$, we get

$$\prod_{\theta\in\Theta}(\delta\upsilon_\theta)^{\alpha_\theta} = 1.$$

Now taking logarithms we get $\alpha \in \ker\phi_{\mathbb{Z}_p,\upsilon}$ as desired.

(The converse can be shown similarly with some care with roots of unity.) □

Lemma 24. For any subfield L of $\mathbb{C}_p$ containing $\mathbb{Q}_p$ and for $\upsilon \in \Upsilon$ we have

$$\operatorname{rank}_{\mathbb{Z}_p}\ker\phi_{\mathbb{Z}_p,\upsilon} \leq \dim_L\ker\phi_{L,\upsilon}$$

Proof. We will show that

$$L \otimes_{\mathbb{Z}_p} \ker \phi_{\mathbb{Z}_p,\upsilon} \subset \ker \phi_{L,\upsilon},$$

from which the claim follows. An element of the left side can be written as $\sum_i \ell_i \alpha_i$ for some $\ell_i \in L$ and $\alpha_i \in \ker \phi_{\mathbb{Z}_p,\upsilon} \subset \ker \phi_{L,\upsilon}$. But $\phi_{L,\upsilon}$ is L-linear, so $\sum_i \ell_i \alpha_i$ is also in the kernel.

In fact the converse also holds. First apply $\mathbb{C}_p \otimes_L -$ to get to $\ker \phi_{\mathbb{C}_p,\upsilon}$. The image of $\phi_{K_p,\upsilon}$ lies in K_p^Θ (instead of $\mathbb{C}_p^\Theta$); this requires that $\log_p$ is $\mathbb{Q}_p$-analytic, so $\log_p(K_p^\times) \subset K_p$. Thus we can get from $\ker \phi_{\mathbb{C}_p,\upsilon}$ to $\mathbb{C}_p \otimes_{K_p} \ker \phi_{K_p,\upsilon}$ by choosing an K_p-basis for K_p^Θ extending an K_p-basis for $\ker \phi_{K_p,\upsilon}$ and applying $\mathbb{C}_p \otimes_{K_p} -$. We can get from $\ker \phi_{K_p,\upsilon}$ to $K_p \otimes_{\mathbb{Q}_p} \ker \phi_{\mathbb{Q}_p,u}$ via a trick with taking the trace (see [11, Theorem 5.31] for details), and can get from $\ker \phi_{\mathbb{Q}_p,\upsilon}$ to $\mathbb{Q}_p \otimes_{\mathbb{Z}_p} \ker \phi_{\mathbb{Z}_p,\upsilon}$ by clearing denominators. □

If $K/\mathbb{Q}$ is Galois, then $G = \Theta$, and for a subring A of $\mathbb{C}_p$ we can naturally identify $A^\Theta = A^G$ with $A[G]$. The group ring $A[G]$ carries a G-action, so this identification induces a G-action on A^Θ, namely $(h \cdot \alpha)_g \mapsto \alpha_{h^{-1}g}$. For some $\upsilon \in \Upsilon$, let $I_A = \ker \phi_{A,\upsilon} \subset A[G]$. This is certainly an A-submodule of $A[G]$, and we wish to know if it is an $A[G]$-submodule, i.e., an ideal.

Lemma 25. If $K/\mathbb{Q}$ is Galois, A a subring of $\mathbb{C}_p$, and $\upsilon \in \Upsilon$ is compatible with the galois action, then I_A is an ideal of $A[G]$.

Proof. One must show that $hI_A \subset I_A$ for $h \in G$. Suppose $\alpha \in I_A$ and $h \in G$; we need to show that $\phi_{A,\upsilon}(h \cdot \alpha) = 0$, i.e., that for all $\delta \in \Theta$, we have

$$\sum_{\theta \in \Theta} (h \cdot \alpha)_\theta (\log_p(\delta \upsilon_\theta)) = 0.$$

As υ is compatible with the galois action, we have

$$\upsilon_{gh} = ghu = g\upsilon_h$$

for some $u \in E_{K,\varepsilon}$. We find that

$$\begin{aligned}
\sum_{\theta \in \Theta} (h \cdot \alpha)_\theta (\log_p(\delta \upsilon_\theta)) &= \sum_{\theta \in \Theta} \alpha_{h^{-1}\theta} (\log_p(\delta \upsilon_\theta)) \\
&= \sum_{\theta \in \Theta} \alpha_\theta (\log_p(\delta \upsilon_{h\theta})) \\
&= \sum_{\theta \in \Theta} \alpha_\theta (\log_p(\delta h \upsilon_\theta)) = 0,
\end{aligned}$$

as $\delta h \in \Theta$. □

The following lemma is the heart of the proof of Leopoldt's Theorem for abelian extensions. It uses an argument of representation theory which depends upong G being abelian in a key way.

Lemma 26. If $K/\mathbb{Q}$ is abelian and $\upsilon \in \Upsilon$ is compatible with the galois action, then

$$\dim_{\mathbb{C}_p}(\ker \phi_{\mathbb{C}_p,\upsilon}) = \dim_{\overline{\mathbb{Q}}}(\ker \phi_{\overline{\mathbb{Q}},\upsilon}).$$

Proof. Suppose that L is a subfield of $\mathbb{C}_p$ containing $\chi(g)$ for all $g \in G$ and $\chi \in \hat{G}$, and identify $\chi \in \hat{G}$ with $\sum_g \chi(g)g \in L[G]$. For $g \in G$ and $\chi \in \hat{G}$, we have

$$g \cdot \chi = \chi(g^{-1})\chi$$

so that $L \cdot \chi$ is G-invariant. As G is abelian (this is the only place we require G to be abelian), $|\hat{G}| = |G| = \dim_L L[G]$, so

$$L[G] = \bigoplus_{\chi \in \hat{G}} L \cdot \chi$$

decomposes $L[G]$ into 1-dimensional irreducible $L[G]$-modules. Then as I_L is an ideal of $L[G]$ by Lemma 25, we have

$$I_L = \bigoplus_{\chi \in H_L} L \cdot \chi$$

for some subset H_L of $\hat{G}$ possibly depending on L. Now we consider the cases $L = \overline{\mathbb{Q}}$ and $L = \mathbb{C}_p$. We find that

$$\overline{\mathbb{Q}}[G] \cap I_{\mathbb{C}_p} = I_{\overline{\mathbb{Q}}} = \bigoplus_{\chi \in H_{\overline{\mathbb{Q}}}} \overline{\mathbb{Q}} \cdot \chi.$$

However $\chi \in \overline{\mathbb{Q}}[G] \cap I_{\mathbb{C}_p}$ if and only if $\chi \in I_{\mathbb{C}_p}$, so $H_{\overline{\mathbb{Q}}} = H_{\mathbb{C}_p}$. Then

$$\dim_{\mathbb{C}_p}(I_{\mathbb{C}_p}) = \left|H_{\mathbb{C}_p}\right| = \left|H_{\overline{\mathbb{Q}}}\right| = \dim_{\overline{\mathbb{Q}}}(I_{\overline{\mathbb{Q}}})$$

as desired. □

This last lemma translates Brumer's Theorem into a form compatible with the notation we've been using for our proof of Leopoldt's Conjecture for abelian extensions.

Lemma 27. Suppose $\upsilon \in \Upsilon$ is such that $\langle \upsilon_\theta \rangle_{\theta \in \Theta}$ has finite index in $E_{K,\varepsilon}$. Then

$$\dim_{\overline{\mathbb{Q}}}(\ker \phi_{\overline{\mathbb{Q}},\upsilon}) \leq [K : \mathbb{Q}] - r_K.$$

Proof. Fix some $\theta_0 \in \Theta$ (say the identity function).

There exists a subset $\Delta \subset \Theta$ with $|\Delta| = \operatorname{rank}_{\mathbb{Z}}(E_{K,\varepsilon}) = r_K$ such that the υ_δ for $\delta \in \Delta$ are $\mathbb{Z}$-linearly independent in $E_{K,\varepsilon}$; then the $\log_p(\theta_0 \upsilon_\delta) \in \mathbb{C}_p$ are $\mathbb{Z}$-linearly independent, and thus $\mathbb{Q}$-linearly independent, and thus by Theorem 19 also $\overline{\mathbb{Q}}$-linearly independent.

Now for each $\delta \in \Delta$ we pick $x_\delta \in (\overline{\mathbb{Q}})^\Theta$ via $(x_\delta)_\theta = 1$ when $\theta = \delta$ and $(x_\delta)_\theta = 0$ otherwise, and claim that the $\phi_{\overline{\mathbb{Q}},\upsilon}(x_\delta) \in \mathbb{C}_p^\Theta$ for $\delta \in \Delta$ are $\overline{\mathbb{Q}}$-linearly independent. In fact, it suffices to look at their θ_0-coordinates; we have

$$(\phi_{\overline{\mathbb{Q}},\upsilon}(x_\delta))_{\theta_0} = \sum_{\theta \in \Theta} (x_\delta)_\theta \log_p(\theta_0 \upsilon_\theta) = \log_p(\theta_0 \upsilon_\delta).$$

Thus $\dim_{\overline{\mathbb{Q}}}(\operatorname{im} \phi_{\overline{\mathbb{Q}},\upsilon}) \geq r_K$, and subtracting from $\dim_{\overline{\mathbb{Q}}}(\overline{\mathbb{Q}}^\Theta) = |G|$ gives us the result.

In fact we have equality; we can construct one relation in $\ker \phi_{\overline{\mathbb{Q}},\upsilon}$ via $\prod_{\theta \in \Theta} \theta(x) = 1$ for $x \in E_{K,\varepsilon}$ and one additional relation for every pair of complex embeddings $\theta, \overline{\theta} \in \Theta$. This is a total of $[K : \mathbb{Q}] - r_K$ relations by Dirichlet's unit theorem. $\square$

5 Leopoldt defect for non-abelian extensions

The assumption that G is abelian was needed so that $|\hat{G}| = |G|$; the argument given above fails in the presence of representations for G with dimension more than one. With additional knowledge about the representations of G we can extend our result.

Let $[G, G]$ be the commutator subgroup of G and $G_a = G/[G, G]$ be the abelianization of G. Then $\hat{G} \cong \hat{G}_a$.

Lemma 28. If $K/\mathbb{Q}$ is Galois and $\upsilon \in \Upsilon$ is compatible with the galois action, then

$$\dim_{\mathbb{C}_p}(\ker \phi_{\mathbb{C}_p,\upsilon}) \leq \dim_{\overline{\mathbb{Q}}}(\ker \phi_{\overline{\mathbb{Q}},\upsilon}) + (|G| - |G_a|).$$

Proof. For a subfield L of $\mathbb{C}_p$ containing the $\chi(g)$ for $\chi \in \hat{G}$ and $g \in G$, we can write $L[G]$ as a sum of $L[G]$-modules

$$L[G] = \left(\bigoplus_{\chi \in \hat{G}} L \cdot \chi \right) \oplus X_L$$

where $X_L \subset L[G]$ is some $L[G]$-module (which encodes all the information about the irreducible representations of G of dimension greater than one). Now I_L is an ideal of $L[G]$ (Lemma 25), so

$$I_L = \left(\bigoplus_{\chi \in H_L} L \cdot \chi \right) \oplus Y_L$$

for some subset $H_L \subset \hat{G}$, possibly depending on L, and some submodule $Y_L \subset X_L$. As before, $\chi \in \overline{\mathbb{Q}}[G] \cap I_{\mathbb{C}_p}$ if and only if $\chi \in I_{\mathbb{C}_p}$, so $H_{\overline{\mathbb{Q}}} = H_{\mathbb{C}_p}$. Then we have

$$\begin{aligned} \dim_{\mathbb{C}_p}(\ker \phi_{\mathbb{C}_p,u}) - \dim_{\overline{\mathbb{Q}}}(\ker \phi_{\overline{\mathbb{Q}},u}) &= \dim_{\mathbb{C}_p} Y_{\mathbb{C}_p} - \dim_{\overline{\mathbb{Q}}} Y_{\overline{\mathbb{Q}}} \\ &\leq \dim_{\mathbb{C}_p} X_{\mathbb{C}_p} = |G| - |\hat{G}| = |G| - |G_a|\,, \end{aligned}$$

which yields the claim. □

Then, proceeding exactly as in Theorem 22 with Lemma 28 replacing Lemma 26, we get

Theorem 29. If $K/\mathbb{Q}$ is Galois, then

$$r_{K,p} \geq r_K - (|G| - |G_a|).$$

This gives us an upperbound on the Leopoldt defect $r_K - r_{K,p}$ (which, using the converses of Lemmas 23, 24, 27, exactly equals $\dim_{\mathbb{C}_p} Y_{\mathbb{C}_p} - \dim_{\overline{\mathbb{Q}}} Y_{\overline{\mathbb{Q}}}$ when $\langle \upsilon_g \rangle_{g \in G}$ has finite index in E_K). In the case that $K/\mathbb{Q}$ is totally complex (but not abelian), or that $[G, G] = G$, this upper bound is trivial, but in all other cases this is a non-trivial upper bound. In no cases, however, is this bound superior to that found in Theorem 48.

6 Masser's Theorem

6.1 General results from commutative algebra

We here recall a number of definitions and fundamental results from the theory of commutative algebra.

A *primary decomposition* of an ideal $\mathfrak{a}$ in the ring R is a finite collection of primary ideals whose intersection is $\mathfrak{a}$. A *minimal primary decomposition* is a primary decomposition where no member of the decomposition contains the intersection of the others, and no two members have the same radical. Every proper ideal of the ring $k[X_1, \ldots, X_n]$ (k a field) has a minimal primary decomposition [1, Theorem 7.13].

The radical of a primary ideal is always prime, and the radicals of the primary ideals in a minimal primary decomposition of $\mathfrak{a}$ are said to be *associated* to $\mathfrak{a}$, and do not depend on the choice of minimal primary decomposition [1, Theorem 4.5].

The *height* of a prime ideal $\mathfrak{p}$ of a ring R is h if there is a chain of proper inclusions of prime ideals

$$\mathfrak{p}_0 \subset \cdots \subset \mathfrak{p}_{h-1} \subset \mathfrak{p}_h = \mathfrak{p}$$

and there is no longer such chain. The *height* of an ideal $\mathfrak{a}$, written $\mathrm{ht}(\mathfrak{a})$, is the least height of the primes associated to $\mathfrak{a}$. An ideal $\mathfrak{a}$ is *unmixed* if all its associated primes have the same height. Krull's height theorem states that if R is Noetherian, then the ideal $(x_1, \ldots, x_r)$ for any $x_i \in R$ has height at most r [1, Corollary 11.16].

Let S be a multiplicative set in R. If $\mathfrak{a}$ is an ideal of R with minimal primary decomposition $\mathfrak{q}_i$, then removing those $\mathfrak{q}_i$ which intersect S gives a minimal primary decomposition for $(S^{-1}\mathfrak{a}) \cap R$ (A-M, Proposition 4.9, page 54). It follows that the associated primes of $(S^{-1}\mathfrak{a}) \cap R$ are those of $\mathfrak{a}$ that do not intersect S. In particular, if $\mathfrak{p}$ is a prime ideal of R, then $(S^{-1}\mathfrak{p}) \cap R = \mathfrak{p}$ if $\mathfrak{p} \cap S = \varnothing$, and $(S^{-1}\mathfrak{p}) \cap R = R$ otherwise.

A *graded ring* R is a ring that can be written

$$R = R_0 \oplus R_1 \oplus \cdots$$

as abelian groups, where $R_m R_n \subset R_{m+n}$. The elements of some R_d are called the *homogeneous elements* of R (of degree d).

A *graded R-module* M is an R-module that can be written

$$M = M_0 \oplus M_1 \oplus \cdots$$

as abelian groups, where $R_m M_n \subset M_{m+n}$. A *graded R-submodule* N of M is an R-submodule $N = \oplus N_i$ for subgroups $N_i \subset M_i$. Then a *homogeneous ideal* of R is a graded R-submodule of R (as a graded R-module over itself in the obvious way).

We can endow $k[X_0, \ldots, X_n]$, k a field, with the structure of a graded ring in the usual way, where the d-part consists of the homogeneous polynomials of degree

d. Then for a finitely generated module M, we define the *Hilbert function* of M at n as $\dim_k M_n$.

Theorem 30. There exists a unique polynomial $H_M(n)$ (the *Hilbert polynomial*), with coefficients in $\mathbb{Q}$, such that $\dim_k M_n = H_M(n)$ for sufficiently large n.

Proof. [1, Corollary 11.2] □

If $H_M(n) = a_r n^r + \cdots + a_1 n + a_0$, write $d(M) = r$ and $\ell(M) = a_r r!$. We see that $\ell(M)$ is always an integer (see proof of [1, Corollary 11.2]).

Now let R be the ordinary ring $k[X_1, \ldots, X_n]$ for some field k, and let $R^h = R[Z]$, which we give the structure of a graded ring as above. For any $f \in R$, we define its *homogenization* f^h in R^h by sending X_i to X_i/Z and multiplying through by the least power of Z needed; then f^h is homogeneous. If $\mathfrak{a}$ is an ideal of R, let $\mathfrak{a}^h$ be the ideal of R^h generated by the f^h for f in $\mathfrak{a}$; then $\mathfrak{a}^h$ is a homogeneous ideal, and so $R^h/\mathfrak{a}^h$ is a graded R^h-module. Then we define

$$d(\mathfrak{a}) = d(R^h/\mathfrak{a}^h),$$
$$\ell(\mathfrak{a}) = \ell(R^h/\mathfrak{a}^h),$$

which we call the *dimension* and *degree* of $\mathfrak{a}$, when not ambiguous.

We get a few basic properties: $d(\mathfrak{a}) + \mathrm{ht}(\mathfrak{a}) = n$, so in particular $d(0) = n$, and also $\ell(0) = 1$. If $f \in R$ is nonconstant, then the degree $\ell((f))$ of the ideal (f) equals the total degree of f as a polynomial. If $\mathfrak{a} \subset \mathfrak{b}$ and $\mathrm{ht}(\mathfrak{a}) = \mathrm{ht}(\mathfrak{b})$, then also $d(\mathfrak{a}) = d(\mathfrak{b})$, and together with $H_{R^h/\mathfrak{a}^h} \geq_O H_{R^h/\mathfrak{b}^h}$ this implies that $\ell(\mathfrak{a}) \geq \ell(\mathfrak{b})$. If $\mathfrak{a}$ and $\mathfrak{b}$ are both unmixed of height r, then so is $\mathfrak{a} \cap \mathfrak{b}$, and $\ell(\mathfrak{a} \cap \mathfrak{b}) = \ell(\mathfrak{a}) + \ell(\mathfrak{b})$. In particular, for an unmixed ideal $\mathfrak{a}$, the number of primes associated to $\mathfrak{a}$ is at most $\ell(\mathfrak{a})$. Furthermore for homogeneous ideals $\mathfrak{a}$ and $\mathfrak{b}$ of R^h, $\ell(R^h/(\mathfrak{a} + \mathfrak{b})) \leq \ell(R^h/\mathfrak{a}) \cdot \ell(R^h/\mathfrak{b})$. These results are found in [5], pages 161-165.

6.2 Preliminary lemmas

We first prove a few general results of commutative algebra which we will be needing in the next subsection. Throughout this subsection, we will take $R = k[X_1, \ldots, X_n]$ for some field k of characteristic 0.

Lemma 31. If $\mathfrak{a}$ is unmixed of height r, and $f \in R$ is not in any of the primes associated to $\mathfrak{a}$, then either $\mathfrak{a} + (f) = R$ or $\mathfrak{a} + (f)$ has height $r + 1$. In the latter case, $\ell(\mathfrak{a} + (f)) \leq \ell(\mathfrak{a}) \cdot \deg f$.

Proof. Let $\mathfrak{a} = \mathfrak{q}_1 \cap \cdots \cap \mathfrak{q}_m$ be a minimal primary decomposition of $\mathfrak{a}$, and $\mathfrak{p}_i$ the radical of $\mathfrak{q}_i$.

If $\mathfrak{p}_i + (f) = R$ for all i, then $(1 - f) \in \mathfrak{p}_i$ for each $\mathfrak{p}_i$, so there is some $(1 - f)^a \in \mathfrak{q}_1 \cap \cdots \cap \mathfrak{q}_m = \mathfrak{a}$. Together with $f \in \mathfrak{a} + (f)$ this gives $1 \in \mathfrak{a} + (f)$, so $\mathfrak{a} + (f) = R$.

Otherwise we have $\mathfrak{p}_i + (f) \neq R$ for some prime $\mathfrak{p}_i$ associated to $\mathfrak{a}$; call that prime $\mathfrak{p}$.

The image of $\mathfrak{p} + (f)$ in $R/\mathfrak{p}$ is principal, so by Krull's height theorem has height 1 (as $R/\mathfrak{p}$ is an integral domain and the image is nonzero, the height cannot be 0). Consider some prime associated to (f) in $R/\mathfrak{p}$ of height 1, and let its preimage in R be $\mathfrak{p}'$. Then $\mathfrak{p}'$ is prime and contains $\mathfrak{p} + (f)$. In particular $\mathfrak{p}'$ strictly contains $\mathfrak{p}$ and there can be no primes properly between them (as $\mathfrak{p}'$ has height 1 in $R/\mathfrak{p}$), so $\mathrm{ht}(\mathfrak{p}') = \mathrm{ht}(\mathfrak{p}) + 1 = \mathrm{ht}(\mathfrak{a}) + 1 = r + 1$. (We have used that $\mathfrak{a}$ is unmixed of height r.)

Now $\mathfrak{a} + (f) \subset \mathfrak{p} + (f) \subset \mathfrak{p}'$, so $r \leq \mathrm{ht}(\mathfrak{a} + (f)) \leq r + 1$.

If $\mathrm{ht}(\mathfrak{a} + (f)) = r$, then $\mathfrak{a} + (f)$ has some associated prime $\mathfrak{p}''$ of height r. Then $\mathfrak{p}'' \supset \mathfrak{a}$ and $\mathfrak{a}$ has height r, so $\mathfrak{p}''$ is a minimal (and thus associated) prime of $\mathfrak{a}$. But $f \in \mathfrak{p}''$, which is impossible, so we must instead have $\mathrm{ht}(\mathfrak{a} + (f)) = r + 1$.

This proves the first part of the lemma. It remains to be shown the claim about the degree of $\mathfrak{a} + (f)$.

Now $\mathfrak{a}^h$ is unmixed of height r, with associated primes $\mathfrak{p}_1^h, \ldots, \mathfrak{p}_m^h$, and f^h does not lie in any of these primes. Then by the above we find that $\mathfrak{a}^h + (f^h)$ is also of height $r + 1$, and thus $d(R^h/(\mathfrak{a}^h + (f^h))) = d(R^h/(\mathfrak{a} + (f))^h)$. Now $\mathfrak{a}^h + (f^h) \subset (\mathfrak{a} + (f))^h$, so we find

$$\begin{aligned}\ell(\mathfrak{a} + (f)) = \ell(R^h/(\mathfrak{a} + (f))^h) &\leq \ell(R^h/(\mathfrak{a}^h + (f^h))) \\ &\leq \ell(R^h/\mathfrak{a}^h) \cdot \ell(R^h/(f^h)) = \ell(\mathfrak{a}) \cdot \deg f.\end{aligned}$$

□

Lemma 32. Suppose $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ are prime ideals of R and $f_1, \ldots, f_r \in R$ are such that $f_i \notin \mathfrak{p}_i$. Then there exists a $\mathbb{Z}$-linear combination of the f_i that does not lie in any of the $\mathfrak{p}_i$.

Proof. We prove this by induction on r, the case $r = 1$ being immediate. Suppose that g is a linear combination of the $f_1, \ldots, f_{r-1}$ which does not lie in any of the $\mathfrak{p}_1, \ldots, \mathfrak{p}_{r-1}$. If g does not lie in $\mathfrak{p}_r$, then we are done, so suppose otherwise. By the pigeon-hole principle, there exist integers $0 \leq a < b \leq 2^{r-1}$ such that $f_r + ag$ and $f_r + bg$ lie in the same subset of the $\mathfrak{p}_1, \ldots, \mathfrak{p}_{r-1}$. If $f_r + ag \in \mathfrak{p}_i$ then $f_r + bg \in \mathfrak{p}_i$ so $(b - a)g \in \mathfrak{p}_i$, impossible (recall that the characteristic of k is 0), so $f_r + ag \notin \mathfrak{p}_1, \ldots, \mathfrak{p}_{r-1}$. Furthermore since $f_r \notin \mathfrak{p}_r$ and $g \in \mathfrak{p}_r$, we have $f_r + ag \notin \mathfrak{p}_r$, so $f_r + ag$ suffices. □

Lemma 33. Let H be a subgroup of the abelian group $\mathbb{Z}^h$ of rank $h' < h$, and x a nonnegative real number. Then the image of $\mathbb{Z}^h(x) \subset \mathbb{Z}^h$ in $\mathbb{Z}^h/H$ has more than $x^{h-h'}$ elements.

Proof. We prove the result by induction on h. If $h = 0$ the result is vacuous, so take $h > 0$.

Let $\pi : \mathbb{Z}^h \to \mathbb{Z}^{h-1}$ be the projection onto the first $h-1$ coordinates. If $\ker \pi \cap H$ is nonempty, then $\pi(H)$ has rank $h' - 1$, so we can apply the lemma to $\pi(H)$ as a subset of $\mathbb{Z}^{h-1}$ to find that $\mathbb{Z}^{h-1}(x)$ has more than $x^{(h-1)-(h'-1)} = x^{h-h'}$ distinct images in $\mathbb{Z}^{h-1}/\pi(H)$, which gives us the result. Now suppose that $\ker \pi \cap H$ is empty.

Now $\pi(H)$ has rank at most h', and if $h' < h-1$ we can apply the lemma to $\pi(H)$ as a subset of $\mathbb{Z}^{h-1}$ to find that $\mathbb{Z}^{h-1}(x)$ has at least $x^{h-h'-1}$ distinct images in $\mathbb{Z}^{h-1}/\pi(H)$; if $h' = h-1$, we get the same result, as $x^{h-h'-1} = 1$. Choose $A \subset \mathbb{Z}^{h-1}(x)$ so that the images of A in $\mathbb{Z}^{h-1}/\pi(H)$ are distinct and $|A| \geq x^{h-h'-1}$. We claim that the elements of $\pi^{-1}(A) \cap \mathbb{Z}^h(x)$ (of which there are more than $x^{h-h'}$) all have distinct images in $\mathbb{Z}^h/H$, from which the result follows. Take distinct $a, b \in \pi^{-1}(A) \cap \mathbb{Z}^h(x)$. If $\pi(a) = \pi(b)$, then $a - b \in \ker \pi$ and therefore not in H, as desired. If $\pi(a) \neq \pi(b)$, then $\pi(a-b) = \pi(a) - \pi(b)$ does not lie in $\pi(H)$, so $a - b$ is not in H, as desired. □

6.3 Central lemma

As in the previous subsection, we take $R = k[X_1, \ldots, X_n]$ for some field k of characteristic 0. For each n-tuple $(\alpha_1, \ldots, \alpha_n)$ of nonzero elements of k, we define the k-algebra automorphism $\gamma(\alpha_1, \ldots, \alpha_n)$ of R by

$$\gamma(\alpha_1, \ldots, \alpha_n)(X_i) = \alpha_i X_i.$$

We will call k-algebra automorphisms of this form *scaling*. Any two scaling automorphisms commute with each other. Given h such scaling automorphisms $\gamma_1, \ldots, \gamma_h$, let Γ be the abelian group generated by the $\gamma_1, \ldots, \gamma_h$, and for nonnegative reals x let

$$\Gamma(x) = \left\{ \gamma^a \;\middle|\; a \in \mathbb{Z}^h(x) \right\}.$$

For an element f of R, write Γf for the set of all γf for $\gamma \in \Gamma$, and write $\Gamma(x) f$ for the set of all γf for $\gamma \in \Gamma(x)$.

Let $\mathfrak{m} = (X_1 - 1, \ldots, X_n - 1)$, a maximal ideal of R. Let $\mathfrak{M}$ be the union of the $\gamma\mathfrak{m}$ for all $\gamma \in \Gamma$. For some subset T of R, let S_T be the complement of T in R, so that $S_\mathfrak{m}$ and $S_\mathfrak{M}$ are both multiplicative sets in R. Write

$$R_\mathfrak{m} = S_\mathfrak{m}^{-1} R,$$
$$R_\mathfrak{M} = S_\mathfrak{M}^{-1} R.$$

Now for any ideal $\mathfrak{a}$ of R, let

$$\mathfrak{a}^* = (S_\mathfrak{M}^{-1}\mathfrak{a}) \cap R,$$

i.e. the contraction of the extension of $\mathfrak{a}$. Thus for a prime ideal $\mathfrak{p}$ of R, we have $\mathfrak{p}^* = \mathfrak{p}$ if and only if $\mathfrak{p} \subset \gamma\mathfrak{m}$ for some $\gamma \in \Gamma$.

For any ideal $\mathfrak{a}$ of R, we can define its *stabilizer* $\mathcal{S}(\mathfrak{a}) \subset \mathbb{Z}^h$ under the action from the γ_i via

$$\mathcal{S}(\mathfrak{a}) = \left\{ a \in \mathbb{Z}^h \,\middle|\, \gamma^a \mathfrak{a} = \mathfrak{a} \right\}.$$

Then for any $1 \leq r \leq n$, let h_r be the maximal $\mathbb{Z}$-rank of the stabilizers $\mathcal{S}(\mathfrak{p})$ as $\mathfrak{p} \subset \mathfrak{m}$ varies over all prime ideals of height exactly r contained in $\mathfrak{m}$. Certainly $h_r \leq h$; if $h_r < h$, we further define $\eta_r = r/(h - h_r)$.

We are now able to state and prove our central lemma.

Lemma 34. Let $R = k[X_1, \ldots, X_n]$ for a field k of characteristic zero, fix some scaling automorphisms $\gamma_1, \ldots, \gamma_h$, and define $\Gamma, \Gamma(x), \mathfrak{m}, \mathfrak{M}, h_r$ and η_r as above. Suppose that each $h_r < h$ for $1 \leq r \leq n$. Fix nonzero $f \in \mathfrak{m}$ of total degree D, and let

$$N_r = D^{\eta_1} + \cdots + D^{\eta_{r-1}}$$

for $1 \leq r \leq n+1$.

For each $1 \leq r \leq n+1$, there exists a polynomial $f_r \in \Gamma(N_r)f$ of total degree at most D that satisfy the following. Let $\mathfrak{a}_r = (f_1, \ldots, f_r)$. Then either $\mathfrak{a}_r^* = R$ or $\mathfrak{a}_r^*$ is unmixed of height r and degree at most D^r.

Proof. We can suppose $n > 0$. For each $1 \leq r \leq n+1$, we will construct a polynomial $f_r \in \Gamma(N_r)f$ with the desired properties, assuming that we have already constructed polynomials $f_1, \ldots, f_{r-1}$.

For now let us assume that $r > 1$, and we need to construct f_r. We will join the case $r = 1$ near the end.

If $\mathfrak{a}_{r-1}^* = R$, then we can pick $f_r = f$ and get $\mathfrak{a}_r^* = R$, as desired. Assume instead that $\mathfrak{a}_{r-1}^*$ is unmixed of height $r-1$ and degree at most D^{r-1}.

We consider for now some prime $\mathfrak{p}$ associated to $\mathfrak{a}_{r-1}^*$. As $\mathfrak{a}_{r-1}^*$ is unmixed of height $r-1$, thus the height of $\mathfrak{p}$ is also $r-1$. Furthermore $\mathfrak{p} \subset \mathfrak{M}$ (as the primes associated to $\mathfrak{a}_{r-1}^*$ are those associated to $\mathfrak{a}_{r-1}$ that don't intersect $S_{\mathfrak{M}}$), so $\mathfrak{p} \subset g_0\mathfrak{m}$ for some $g_0 \in \Gamma$. As $g_0^{-1}\mathfrak{p} \subset m$, by the definition of h_{r-1} the stabilizer $\mathcal{S}(g_0^{-1}\mathfrak{p}) \subset \mathbb{Z}^h$ has rank at most h_{r-1}. However $\mathcal{S}(\mathfrak{p}) = \mathcal{S}(g_0^{-1}\mathfrak{p})$, as the γ_i all commute. Let $\pi : \mathbb{Z}^h \to \mathbb{Z}^h/\mathcal{S}(\mathfrak{p})$ be the quotient map. Then for $b, b' \in \mathbb{Z}^h$, we have $\pi(b) = \pi(b')$ if and only if $b - b' \in \mathcal{S}(\mathfrak{p})$, if and only if $\gamma^{b-b'}\mathfrak{p} = \mathfrak{p}$, if and only if $\gamma^{-b}\mathfrak{p} = \gamma^{-b'}\mathfrak{p}$.

Now applying Lemma 33 to the subgroup $\mathcal{S}(\mathfrak{p})$ of $\mathbb{Z}^h$, we find that the set $B = \mathbb{Z}^h(D^{\eta_{r-1}})$ has more than

$$D^{\eta_{r-1}(h-h_{r-1})} = D^{r-1}$$

distinct images in $\mathbb{Z}^h/\mathcal{S}(\mathfrak{p})$. Thus there are more than D^{r-1} distinct primes of the form $\gamma^{-b}\mathfrak{p}$ for $b \in B$. As the number of primes associated to $\mathfrak{a}_{r-1}^*$ is at most its degree $\ell(\mathfrak{a}_{r-1}^*) \leq D^{r-1}$, therefore there exists $b \in B$ such that $\gamma^{-b}\mathfrak{p}$ is not associated to $\mathfrak{a}_{r-1}^*$. Equivalently, $\mathfrak{p}$ is not associated to $\gamma^b\mathfrak{a}_{r-1}^*$. Now $\mathfrak{p}$ and $\gamma^b\mathfrak{a}_{r-1}^*$ are both unmixed of height $r-1$, so $\gamma^b\mathfrak{a}_{r-1}^* \not\subset \mathfrak{p}$.

We claim that furthermore $\gamma^b\mathfrak{a}_{r-1} \not\subset \mathfrak{p}$. Take $g \in \mathfrak{a}_{r-1}^*$ such that $\gamma^b g \notin \mathfrak{p}$, and write $gg' \in \mathfrak{a}_{r-1}$ for some $g' \in S_{\mathfrak{M}}$. Now $g' \notin \gamma^{-b}\mathfrak{p} \subset \mathfrak{M}$ and $\gamma^b g \notin \mathfrak{p}$, so $\gamma^b(gg') \notin \mathfrak{p}$,

and thus $\gamma^b \mathfrak{a}_{r-1} \not\subset \mathfrak{p}$. As $\mathfrak{a}_{r-1} = (f_1, \ldots, f_{r-1})$, therefore there exists i such that $\gamma^b f_i \notin \mathfrak{p}$.

Now for every prime $\mathfrak{p}$ associated to $\mathfrak{a}_{r-1}^*$, we found some $b \in B$ and i such that $\gamma^b f_i \notin \mathfrak{p}$. Then Lemma 32 shows that there exists a linear combination f_r of these $\gamma^b f_i$ such that f_r does not lie in any of the $\mathfrak{p}$ associated to $\mathfrak{a}_{r-1}^*$.

As each of these $f_1, \ldots, f_{r-1}$ is in $\Gamma(N_{r-1})f$, and $N_r = N_{r-1} + D^{\eta_{r-1}}$, we find that $f_r \in \Gamma(N_r)f$. Certainly the total degree of f_r is at most D. It only remains to show that either $\mathfrak{a}_r^* = R$ or $\mathfrak{a}_r^*$ is unmixed of height r and degree at most D^r.

Let $\mathfrak{b} = \mathfrak{a}_{r-1}^* + (f_r)$. As f_r does not lie in any of the primes associated to $\mathfrak{a}_{r-1}^*$, from Lemma 31 we know that either $\mathfrak{b} = R$ or $\mathfrak{b}$ has height r and degree at most D^r. As $\mathfrak{b} \subset \mathfrak{a}_r^*$, if $\mathfrak{b} = R$ then we are done, so suppose otherwise.

Thus far for any $2 \leq r \leq n+1$ we have constructed an ideal $\mathfrak{b}$ such that $\mathrm{ht}(\mathfrak{b}) = r, \ell(\mathfrak{b}) \leq D^r$, and $\mathfrak{b} \subset \mathfrak{a}_r^*$. If $r = 1$, we can choose $f_1 = f$ (which has total degree D) and $\mathfrak{b} = (f_1)$ and the same three properties hold (the first follows from Krull's height theorem). Now assuming only those three properties, we finish the proof for $1 \leq r \leq n+1$.

Let $\mathfrak{p}$ be a prime associated to $\mathfrak{a}_r^*$; we wish to show that the height of $\mathfrak{p}$ is r. Certainly $\mathrm{ht}(\mathfrak{p}) \geq \mathrm{ht}(\mathfrak{a}_r^*) \geq r$. As before $\mathfrak{p} \subset \mathfrak{M}$, so we can write $\mathfrak{p} \subset g_0\mathfrak{m}$ for some $g_0 \in \Gamma$. Let $\mathfrak{m}' = g_0\mathfrak{m}$. We will work in the localization $S_{\mathfrak{m}'}^{-1}R$; this is a local ring, whose Krull dimension is equal to $\mathrm{ht}(\mathfrak{m}') = n$. Now $\mathfrak{m}'$ is generated by n elements, so $S_{\mathfrak{m}'}^{-1}R$ is a regular local ring. Then

$$\mathrm{ht}(S_{\mathfrak{m}'}^{-1}\mathfrak{a}_r) = \mathrm{ht}((S_{\mathfrak{m}'}^{-1}\mathfrak{a}_r) \cap R) \geq \mathrm{ht}(\mathfrak{a}_r^*) \geq \mathrm{ht}(\mathfrak{b}) = r,$$

but $S_{\mathfrak{m}'}^{-1}\mathfrak{a}_r$ is generated by the r elements $f_1, \ldots, f_r$, so Krull's height theorem shows that its height is at most r, and thus exactly r. Then being in a regular local ring, it must be unmixed. As $S_{\mathfrak{m}'}^{-1}\mathfrak{p}$ is a prime associated to $S_{\mathfrak{m}'}^{-1}\mathfrak{a}_r$, and the latter is unmixed of height r, we have

$$\mathrm{ht}(\mathfrak{p}) = \mathrm{ht}((S_{\mathfrak{m}'})^{-1}\mathfrak{p} \cap R) = \mathrm{ht}(S_{\mathfrak{m}'}^{-1}\mathfrak{p}) = r,$$

so every prime associated to $\mathfrak{a}_r^*$ is of height r. In particular $\mathfrak{a}_r^*$ is unmixed of height r. Now $\mathfrak{a}_r^* \supset \mathfrak{b}$ both have the same height, so $\ell(\mathfrak{a}_r^*) \leq \ell(\mathfrak{b}) \leq D^r$ as desired. □

6.4 Upper bound on the number of zeros of a polynomial

Assuming $\Gamma f \subset \mathfrak{m}$, the central lemma of the previous section allows us to construct ideals in R of arbitrarily high height, which is impossible. Of course, we cannot have $\Gamma f \subset \mathfrak{m}$ anyhow, because that would imply f has an infinite number of zeros. The usefulness of the central lemma is that it allows us to explicitly an ideal in R of some height depending on how many zeros f has. Thus, because we know an upperbound on the height of an ideal in R, we get an upperbound on the number of zeros that f has of a particular form, depending on the total degree of f.

Theorem 35. Let $R = k[X_1, \ldots, X_n]$ for a field k of characteristic zero, fix some scaling automorphisms $\gamma_1 \ldots, \gamma_h$, and define $\Gamma(x), \mathfrak{m}, h_r$ and η_r as above. Suppose that $h_r < h$ for $1 \leq r \leq n$. Suppose f is a nonzero polynomial of total degree D, and let

$$N = D^{\eta_1} + \cdots + D^{\eta_n}.$$

Then $\Gamma(N)f \not\subset \mathfrak{m}$, i.e., there exists $b \in \mathbb{Z}^h(N)$ such that $\gamma^b f \notin \mathfrak{m}$.

Proof. Construct polynomials $f_1, \ldots, f_{n+1} \in \Gamma(N)f$ as in Lemma 34. Then either $\mathfrak{a}^*_{n+1} = R$, or $\mathfrak{a}_{n+1}$ has height $n+1$. The latter is impossible, as R has Krull dimension n, so every proper ideal has height at most n. Thus $\mathfrak{a}^*_{n+1} = R$, so $\mathfrak{a}_{n+1} \not\subset \mathfrak{M}$, and in particular we must have $f_i \notin \mathfrak{m} \subset \mathfrak{M}$ for some i. As $f_i \in \Gamma(N)f$, that gives us the result. □

As a useful application of Theorem 35, we get Theorem 39, which we delay until the next section to take advantage of the notation introduced there.

7 Leopoldt defect for non-abelian extensions, II

7.1 Generalized Dirichlet exponents μ and χ

Throughout this section, we will be working with finitely generated free abelian groups which lie inside of a vector space which carries the structure of an inner product. We capture the essential notions with a definition, so named for its superficial similarity to lattices.

Definition 36. A *lettuce* (or F-lettuce) is a finitely generated subgroup X of a vector space V over a field F of charecteristic 0, together with a fixed choice of F-basis for V. We write $X \subset V$ to specify the F-space V, and write FX for the F-span of X.

Given a lettuce $X \subset V$ with F-basis $a_1, \dots, a_n$, we define an inner product $\langle \cdot, \cdot \rangle : V \times V \to F$ via

$$\langle \kappa_1 a_1 + \cdots + \kappa_n a_n, \lambda_1 a_1 + \cdots + \lambda_n a_n \rangle = \kappa_1 \lambda_1 + \cdots + \kappa_n \lambda_n.$$

When we work with multiple lettuces (a salad, perhaps), we implicity take them to be within the same larger F-space V and with the same choice of basis on V, so that the inner product is consistently defined across them.

Definition 37. A *sublettuce* X' of X is a lettuce of the form $W \cap X$ for some F-space $W \subset FX$.

Equivalently, X' is a sublettuce of X if and only if $X' = FX' \cap X$. Any sublettuce of X is a direct summand of X as a free abelian group, but the converse need not be true. In fact, this distinction is exactly why we find it convenient to define the notion of a lettuce as distinct from the notion of a finitely generated subgroup of a vector space.

For a lettuce X we write $r(X)$ for $\operatorname{rank}_{\mathbb{Z}}(X)$. For any nonzero lettuce X, we define

$$\mu(X) = \min \frac{r(X) - r(X')}{\dim_F FX - \dim_F FX'},$$

where the minimum is taken over all proper sublettuces $X' \subset X$. Certainly $\mu(X) \leq r(X)/\dim_F FX$. We say that a nonzero lettuce X is *maximal* if $\mu(X) = r(X)/\dim_F FX$.

For any nonzero lettuces X, Y, let

$$\chi(X, Y) = \min \frac{r(X) - r(X')}{r(Y')},$$

where the minimum is taken over all $X' \subset X, Y' \subset Y$ with Y' nonzero and $\langle X', Y' \rangle = 0$. We have $\chi(X, Y) \leq r(X)/r(Y)$.

For any F-subspace $W \subset V$, write $W^\perp = \{\alpha \in V \mid \langle W, \alpha \rangle = 0\}$, and define $\pi_W : V \to W$ as the projection of $V = W \oplus W^\perp$ onto its first component. For a lettuce X, we write $X^\perp = (FX)^\perp$.

7.2 Bound on χ for polynomials with many zeros

We continue from section 6 to show that there is a relationship between the number of zeros a polynomial has of a specific form, the total degree of the polynomial, and the χ-coefficient that describes the distribution of the zeros. Before we complete the main proof we isolate a short lemma.

Lemma 38. Suppose F is a subfield of $\mathbb{C}_p$, and X, Y are lettuces with $Y = \mathbb{Z}y_1 + \cdots + \mathbb{Z}y_d$, with $d = r(Y)$. Suppose there is a nonzero polynomial Q with coefficients in F such that

$$Q(\exp_p \langle y_1, x\rangle, \ldots, \exp_p \langle y_s, x\rangle) = 0$$

for all $x \in X$. Then $X^\perp \cap Y \neq 0$.

Proof. Say that Q has degree at most C in any variable. Write

$$\begin{aligned} f(x) &= Q(\exp_p \langle y_1, x\rangle, \ldots, \exp_p \langle y_s, x\rangle) \\ &= \sum_{\lambda \in \mathbb{Z}^h(C)} a(\lambda)(\exp_p \langle y_1, x\rangle)^{\lambda_1} \cdots (\exp_p \langle y_s, x\rangle)^{\lambda_s} \end{aligned}$$

for some coefficients $a(\lambda) \in F$, not all zero. We know $f(x) = 0$ for all $x \in X$. For some fixed value of x, let $\alpha_i = \exp_p \langle y_i, x\rangle$, so that we get

$$0 = f(nx) = \sum_{\lambda \in \mathbb{Z}^h(C)} a(\lambda)(\alpha^\lambda)^n$$

for all integers n. Then by Theorem 13, there exists $\lambda, \lambda' \in \mathbb{Z}^h(C)$ such that $\lambda \neq \lambda'$ and $\alpha^\lambda = \alpha^{\lambda'}$. Let $\nu = \lambda - \lambda' \neq 0$, so that $\alpha^\nu = 1$, and thus

$$\exp_p \langle \nu_1 y_1 + \cdots + \nu_s y_s, x\rangle = 1$$

so $\langle \nu_1 y_1 + \cdots + \nu_s y_s, x\rangle = 0$.

Thus we have shown that FX is contained in the union of the $(Fy)^\perp$, as y runs over the nonzero $\nu_1 y_1 + \cdots + \nu_s y_s \in Y$, with $|\nu_i|_\infty \leq C$. As this is only finitely many values of y, and F is infinite, we find that one of the vector spaces $(Fy)^\perp$ must contain FX. That is, there exists y such that $FX \subset (Fy)^\perp$, or equivalently, $y \in X^\perp$. □

Theorem 39. Suppose F is a subfield of $\mathbb{C}_p$ and that X and Y are lettuces with $FX = FY$. Suppose $X = \mathbb{Z}x_1 + \cdots + \mathbb{Z}x_\ell$ and $Y = \mathbb{Z}y_1 + \cdots + \mathbb{Z}y_d$ where $\ell = r(X)$ and $d = r(Y)$. Suppose that P is a nonzero polynomial with integer coefficients such that

$$P(\exp_p \langle y_1, x\rangle, \ldots, \exp_p \langle y_d, x\rangle) = 0$$

for all $x = \lambda_1 x_1 + \cdots + \lambda_\ell x_\ell$ with $\lambda \in \mathbb{Z}^\ell(N)$ for some integer N. Then the total degree of P is at least

$$(N/d)^{\chi(X,Y)}.$$

Proof. For $1 \le j \le d$, let

$$\ell_j = \max_{Y' \subset Y, r(Y') \ge j} r((Y')^\perp \cap X).$$

Notice that

$$\chi(X,Y) = \min_{1 \le j \le n} \frac{\ell - \ell_j}{j}.$$

If $\ell_j = \ell$ for some j, then $\chi(X,Y) = 0$ and we are done immediately, so suppose otherwise that $\ell_j < \ell$ for each j.

We wish to use Theorem 35: choose $R = F[W_1, \dots, W_d]$, and choose scaling automorphisms

$$\gamma_i = \gamma(\exp_p \langle y_1, x_i \rangle, \dots, \exp_p \langle y_d, x_i \rangle)$$

for $1 \le i \le \ell$. For $\lambda \in \mathbb{Z}^\ell$, let $x = \lambda_1 x_1 + \cdots + \lambda_\ell x_\ell$, and we have

$$\gamma^\lambda = \gamma(\exp_p \langle y_1, x \rangle, \dots, \exp_p \langle y_d, x \rangle).$$

Define $\Gamma(x), \mathfrak{m}, h_r$, and η_r as before.

Now consider some prime $\mathfrak{p} \subset \mathfrak{m}$ of height $1 \le j \le d$, which has a stabilizer $\mathcal{S}(\mathfrak{p}) \subset \mathbb{Z}^\ell$. Corresponding to this stabilizer we get a subgroup $X_0 \subset X$ (not necessarily sublettuce)

$$X_0 = \left\{ \lambda_1 x_1 + \cdots \lambda_\ell x_\ell \,\middle|\, \gamma^\lambda \mathfrak{p} = \mathfrak{p} \right\}.$$

Let X' be the sublettuce of X generated by X_0, that is, $X' = (FX_0) \cap X$, and in particular $FX' = FX_0$. We have $r(X') \ge \operatorname{rank}_{\mathbb{Z}}(X_0) = \operatorname{rank}_{\mathbb{Z}}(\mathcal{S}(\mathfrak{p}))$.

Let $Y' = (X')^\perp \cap Y = X_0^\perp \cap Y$, and then $Y = Y' \oplus Y''$. We aim to show that $r(Y') \ge j$, or equivalently $r(Y'') \le d - j$. Let $s = r(Y'')$, and choose a $\mathbb{Z}$-basis $u_1, \dots, u_s$ for Y''. Write

$$u_i = \nu_{i,1} y_1 + \cdots + \nu_{i,d} y_d;$$

the $\nu_i \in \mathbb{Z}^d$ are $\mathbb{Z}$-linearly independent. Let $M_i = W^{\nu_i} = W_1^{\nu_{i,1}} \cdots W_d^{\nu_{i,d}} \in R$. As $\mathfrak{p}$ has height j, the the transcendence degree of the ring $R/\mathfrak{p}$ over the field F is $d - j$.

Suppose towards a contradiction that $s > d - j$. Then the $M_1, \dots, M_s$ are algebraically dependent in the ring $R/\mathfrak{p}$ (as $\dim R = d$ and $\operatorname{ht}(\mathfrak{p}) = j$), so there exists a nonzero polynomial Q with coefficients in F such that

$$Q(M_1, \dots, M_s) \in \mathfrak{p}.$$

For every $\lambda \in \mathcal{S}(\mathfrak{p})$, we have $\gamma^\lambda \mathfrak{p} = \mathfrak{p}$, so also

$$Q(\gamma^\lambda M_1, \dots, \gamma^\lambda M_s) = \gamma^\lambda(Q(M_1, \dots, M_s)) \in \gamma^\lambda \mathfrak{p} = \mathfrak{p} \subset \mathfrak{m}.$$

Let $x = \lambda_1 x_1 + \cdots + \lambda_\ell x_\ell$. Evaluating the left term at $W_1 = \cdots = W_d = 1$, we find that

$$\gamma^\lambda M_i \mapsto \exp_p \langle \nu_{i,1} y_1 + \cdots + \nu_{i,d} y_d, \lambda_1 x_1 + \cdots + \lambda_\ell x_\ell \rangle = \exp_p \langle u_i, x \rangle ,$$
$$Q(\gamma^\lambda M_1, \ldots, \gamma^\lambda M_s) \mapsto Q(\exp_p \langle u_1, x \rangle , \ldots, \exp_p \langle u_s, x \rangle).$$

However $Q(\gamma^\lambda M_1, \ldots, \gamma^\lambda M_s) \in \mathfrak{m}$, so we find that

$$Q(\exp_p \langle u_1, x \rangle , \ldots, \exp_p \langle u_s, x \rangle) = 0.$$

Now as λ varies over $\mathcal{S}(\mathfrak{p})$, then x varies over X_0, so this holds for all $x \in X_0$.

Thus we can apply Lemma 38 to find that there exists nonzero y in the sublettuce

$$X_0^\perp \cap Y'' = (X')^\perp \cap Y'' = Y' \cap Y'' = 0,$$

which is a contradiction. Thus we must have $r(Y'') = s \leq d - j$, so $r(Y') \geq j$.

Recall $Y' \subset (X')^\perp$, so $\langle X', Y' \rangle = 0$, and in particular $X' \subset (Y')^\perp \cap X$, so

$$\operatorname{rank}_{\mathbb{Z}}(\mathcal{S}(\mathfrak{p})) \leq r(X') \leq r((Y')^\perp \cap X) \leq \max_{U \subset Y, r(U) \geq j} r(U^\perp \cap X) = \ell_j.$$

This holds for all primes $\mathfrak{p} \subset \mathfrak{m}$ of height j. Taking the maximum on the left side over all such primes $\mathfrak{p}$, we find that $h_j \leq \ell_j$. As we have assumed $\ell_j < \ell$, in particular we have $h_j < \ell$, and we can apply Theorem 35.

Now for $g \in \Gamma(N)$, we have $g = \gamma^\lambda$ for some $\lambda \in \mathbb{Z}^\ell(N)$, and thus

$$\begin{aligned}(gP)(1, \ldots, 1) &= (\gamma^\lambda P)(1, \ldots, 1) \\ &= P(\exp_p \langle y_1, x \rangle , \ldots, \exp_p \langle y_d, x \rangle) = 0,\end{aligned}$$

so therefore $(gP) \in \mathfrak{m} = (W_1 - 1, \ldots, W_d - 1)$ and $\Gamma(N)P \subset \mathfrak{m}$. The theorem tells us therefore that if D is the total degree of P, then

$$D^{\eta_1} + \cdots + D^{\eta_d} > N.$$

Now

$$\eta_j = \frac{j}{\ell - h_j} \leq \frac{j}{\ell - \ell_j} \leq \frac{1}{\chi(X, Y)},$$

so

$$N < dD^{1/\chi(X,Y)},$$

and thus $D > (N/d)^{\chi(X,Y)}$ as desired. □

7.3 Construction of a polynomial with many zeros

Having found a bound on χ in terms of the total degree of a polynomial and the number of zeros it has of a particular form, it remains to construct such a polynomial to apply the theorem to. We aim to construct such a polynomial, with many zeros but low degree. First we need another application of the pigeon-hole principle, this time a variation of Lemma 14.

Lemma 40. Let L be a local field. Suppose m, n are positive integers and $a_{i,j} \in L$ are nonzero, with $1 \le i \le m, 1 \le j \le n$. Suppose $0 < A < 1$. Then there exist integers $b_j, 1 \le j \le n$, not all zero, with $0 \le b_j \le B = A^{-[L:\mathbb{Q}_p](m/n)}$ such that

$$\left| \sum_{j=1}^{n} a_{i,j} b_j \right| \le A \max_j |a_{i,j}|$$

for all $1 \le i \le m$.

Proof. If we divide all the $a_{i,j}$ by the $a_{i,j'}, 1 \le j' \le n$, which maximizes $|a_{i,j'}|$, we may assume that the $\max_j |a_{i,j}| = 1$, and in particular that $a_{i,j} \in \mathcal{O}_L$ for all i, j.

Let $I = \{z \in L \mid |z| \le A\}$; this is an ideal of $\mathcal{O}_L$, generated by some power of p. As $(\mathcal{O}_L : p) = p^{[L:\mathbb{Q}_p]}$, we get

$$(\mathcal{O}_L : I) \le A^{-[L:\mathbb{Q}_p]}.$$

Now we view $a_{i,j}$ as elements of the ring $\mathcal{O}_L/I$. Let $f_i : \mathbb{Z}^n \to \mathcal{O}_L/I$ be the linear transformation that sends $b \in \mathbb{Z}^n$ to $\sum_j a_{i,j} b_j$. There are more than B^n different tuples of integers $b \in \mathbb{Z}^n(B)$, but there are only

$$|(\mathcal{O}_L/I)^m| \le A^{-[L:\mathbb{Q}_p]m} = B^n$$

possible distinct values for the $(f_1(b), \ldots, f_m(b))$. By the pigeon-hole principle, we have two different tuples b, b' with the same image. Then the image of $b - b'$ is zero, so that $b - b'$ gives the desired result. □

Lemma 41. Suppose f is analytic on $B_n(R)$. Suppose $0 < r < R$ and T is a positive integer. Then

$$|f|_r \le \max\left(|f|_R (r/R)^T, \max_{\|\tau\| < T} \left(|f_\tau| \, r^{\|\tau\|} \right) \right).$$

Proof. We have

$$\begin{aligned} \sup_{\|\tau\| \ge T} \left(|f_\tau| \, r^{\|\tau\|} \right) &\le r^T \sup_{\|\tau\| \ge T} \left(|f_\tau| \, R^{\|\tau\| - T} \right) \\ &\le (r/R)^T \sup_{\tau} \left(|f_\tau| \, R^{\|\tau\|} \right) = (r/R)^T |f|_R , \end{aligned}$$

which together with the definition of $|f|_r$ gives us the result. □

Theorem 42. Let L be a local field. Suppose $f_1, \ldots, f_d$ are L-analytic on $B_n(R)$, with $d > n$. Suppose that $|f_i|_R \leq 1$ for all i. For positive reals x, define

$$D_x = x^{n/(d-n)}(\log x)^{n+2},$$
$$U_x = x^{d/(d-n)}(\log x)^{d+2}.$$

Suppose $0 < r < R$. Then for sufficiently large x, there exists a nonzero polynomial $P_x \in \mathbb{Z}[X_1, \ldots, X_d]$ such that

$$\begin{aligned} \deg_{X_i} P_x &\leq D_x, \\ H(P_x) &\leq \exp(D_x x), \\ |P_x(f_1, \ldots, f_d)|_r &< \exp(-U_x), \end{aligned}$$

where $H(\cdot)$ is the *height* of a polynomial, i.e. the maximum of the absolute values of its coefficients.

Proof. Let $g_\lambda = f^\lambda$ for $\lambda \in \mathbb{Z}^d, \lambda \geq 0$. Furthermore write

$$g_\lambda(z) = \sum_{\tau \geq 0} g_{\lambda,\tau} z^\tau$$

for $z \in B_n(r)$. As the f_i are L-analytic, so are the g_λ, and $g_{\lambda,\tau} \in L$.

Let $T_x = \lceil 2U_x/(\log(R/r)) \rceil$. We apply Lemma 40 to the $g_{\lambda,\tau}$, for $\lambda \in \mathbb{Z}^d(D_x)$ and $\tau \in \mathbb{Z}^n$, with $0 \leq \tau$ and $\|\tau\| < T_x$ to get integer coefficients $a(\lambda)$, not all zero, such that

$$\left| \sum_\lambda g_{\lambda,\tau} a(\lambda) \right| \leq \exp(-U_x) \max_\lambda |g_{\lambda,\tau}|$$

for each τ. We also get a bound on the $|a(\lambda)|_\infty$ which we will compute momentarily.

Let $P_x = \sum_{\lambda \in \mathbb{Z}^d(D_x)} a(\lambda) X^\lambda$, a nonzero polynomial in $\mathbb{Z}[X_1, \ldots, X_d]$ with $\deg_{X_i} P_x \leq D_x$. Let

$$f = P_x(f_1, \ldots, f_d) = \sum_\lambda a(\lambda) g_\lambda$$

Now for each $\|\tau\| < T_x$, we have

$$\begin{aligned} |f_\tau|\, r^{\|\tau\|} &= \left| \sum_\lambda a(\lambda) g_{\lambda,\tau} \right| r^{\|\tau\|} \\ &\leq \exp(-U_x) \max_\lambda (|g_{\lambda,\tau}|\, r^{\|\tau\|}) \\ &\leq \exp(-U_x) \max_\lambda |g_\lambda|_r \\ &\leq \exp(-U_x). \end{aligned}$$

We also have,

$$|f|_R \le \sum_\lambda |g_\lambda|_R \le (D_x+1)^d,$$
$$|f|_R (r/R)^{T_x} \le \exp(-U_x).$$

Then by Lemma 41, we get $|f|_r \le \exp(-U_x)$. It only remains to compute the bound on the height of P_x.

We have

$$D_x^d = x^{dn/(d-n)}(\log x)^{dn+2d}$$
$$\binom{T_x+n-1}{n} =_O T_x^n =_O x^{dn/(d-n)}(\log x)^{dn+2n}$$
$$D_x^d \cdot \binom{T_x+n}{n}^{-1} =_O (\log x)^{2(d-n)} >_O (\log x)^{d-n}.$$

In our application of Lemma 40, we had to satisfy $\binom{T_x+n-1}{n}$ equations with at least D_x^d variables (as there are $\binom{T_x+n-1}{n}$ values of $\tau \in \mathbb{Z}^n$ with $0 \le \tau$ and $\|\tau\| < T_x$). For sufficiently large x, their ratio exceeds $[L:\mathbb{Q}_p](\log x)^{d-n}$, so we get from the lemma

$$H(P_x) = \max_\lambda |a(\lambda)|_\infty \le \exp(-U_x)^{-[L:\mathbb{Q}_p]/((\log x)^{d-n}[L:\mathbb{Q}_p])}$$
$$= \exp(U_x(\log x)^{n-d}) = \exp(D_x x)$$

as desired. □

7.4 Bound on the Leopoldt defect

Now we combine the polynomial we constructed in Theorem 42 with Theorem 39 to get an essential upperbound on χ.

Theorem 43. Suppose F is a local field. Suppose that X, Y are nonzero lettuces with $FX = FY$ and $\exp_p \langle X, Y\rangle \subset \overline{\mathbb{Q}}$, and that $r(Y) > \dim_F FY$. Then

$$\chi(X,Y)(\mu(Y)-1) \le 1.$$

Proof. Let $d = r(Y)$ and $n = \dim_F FY$, and write $Y = \mathbb{Z}y_1 + \cdots + \mathbb{Z}y_d$. Write $X = \mathbb{Z}x_1 + \cdots + \mathbb{Z}x_\ell$ where $\ell = r(X)$. Let $\alpha_{i,j} = \exp_p \langle x_i, y_j\rangle \in \overline{\mathbb{Q}}$, and let $K = \mathbb{Q}(\alpha_{1,1}, \ldots, \alpha_{\ell,d})$, a number field.

Let $r = \max_\ell |x_\ell|$, so that $X \subset B_n(r)$. Now each $|\langle x_i, y_j\rangle| < \varepsilon$ (as otherwise $\exp_p \langle X, Y\rangle$ would not even make sense), so there exists $R > r$ such that each $|\langle B_n(R), y_j\rangle| < \varepsilon$.

Thus we define $f_1, \ldots, f_d$, F-analytic functions on $B_n(R)$, by

$$f_j : z \mapsto \exp_p \langle z, y_j\rangle .$$

These functions satisfy $|f_j|_R \leq 1$ for each j. Then for sufficiently large x, Theorem 42 gives us a nonzero polynomial $P_x \in \mathbb{Z}[X_1, \ldots, X_d]$ such that $f = P_x(f_1, \ldots, f_d)$ satisfies certain properties. We will show that, for sufficiently large x, we have

$$f(\lambda_1 x_1 + \cdots + \lambda_\ell x_\ell) = 0$$

for all $\lambda \in \mathbb{Z}^\ell(x)$.

Write the polynomial P_x as $P_x = \sum_\mu a(\mu) X^\mu$, where $a(\mu)$ are the integer coefficients. Fix $\lambda \in \mathbb{Z}^\ell(x)$ and let $z = \lambda_1 x_1 + \cdots + \lambda_\ell x_\ell$. Then

$$f(z) = \sum_\mu a(\mu) \prod_{i,j} \alpha_{i,j}^{\lambda_i \mu_j}.$$

Recall the definitions of D_x and U_x from Theorem 42. We know that

$$|f(z)| \leq |f|_r < \exp(-U_x),$$

and $\deg_{X_i} P_x \leq D_x$, and $|a(\mu)|_\infty \leq \exp(D_x x)$. We wish to apply Theorem 12. Let $M \in \mathbb{Z}$ be so large that $M\alpha_{i,j} \in \mathbb{A}$ for each i, j, and

$$|\sigma \alpha_{i,j}|_\infty < M$$

for each i, j and every embedding $\sigma : K \to \mathbb{C}$. Thus $M^{d\ell x D_x} f(z) \in \mathbb{A}$, and $|\sigma f(z)|_\infty \leq D_x^d \exp(D_x x) M^{d\ell x D_x}$ for each embedding $\sigma : K \to \mathbb{C}$. If $f(z)$ is nonzero, then by Theorem 12 we can conclude that

$$\log |f(z)| \geq -[K : \mathbb{Q}] \log(M^{d\ell x D_x} D_x^d \exp(D_x x) M^{d\ell x D_x}).$$

Combining with $|f(z)| < \exp(-U_x)$, we find

$$\begin{aligned} U_x &< [K : \mathbb{Q}] \log(M^{d\ell x D_x} D_x^d \exp(D_x x) M^{d\ell x D_x}) \\ &=_O (d\ell x D_x) \log M + d \log D_x + D_x x \\ &=_O D_x x. \end{aligned}$$

However $U_x = (D_x x)(\log x)^{2(d-n)} >_O D_x x$, and for sufficiently large x (not depending on λ) the inequality $U_x \leq_O D_x x$ is impossible. Thus we must have $f(z) = 0$.

Consequentially, we can apply Theorem 39 to P_x and find that

$$(x/\ell)^{\chi(X,Y)} \leq \sum_i \deg_{X_i} P_x \leq dD_x =_O x^{n/(d-n)} (\log x)^{n+2}.$$

Thus we must have $\chi(X, Y) \leq n/(d-n)$. Now $\mu(Y) \leq d/n$, so $\mu(Y) - 1 \leq (d-n)/n$, and $\chi(X, Y)(\mu(Y) - 1) \leq 1$, as desired. □

Now to get to our main result, we will need three simple lemmas that only use the definitions of μ and χ directly.

Lemma 44. Let X be a nonzero lettuce. There exists nonzero $X' \subset X$ such that $\mu(X') \geq \frac{r(X)}{\dim_F FX}$ and X' is maximal.

Proof. We induct on $r(X)$. If X is maximal then we are done. Otherwise, there exists a proper sublettuce X' of X such that

$$\frac{r(X) - r(X')}{\dim_F FX - \dim_F FX'} = \mu(X) < \frac{r(X)}{\dim_F FX}.$$

(Clearly X' is nonzero.) Rearranging, we find

$$\frac{r(X')}{\dim_F FX'} > \frac{r(X)}{\dim_F FX},$$

so by applying the lemma to X' we get an $X'' \subset X'$ with the desired properties. □

Lemma 45. If $X \subset V$ is a nonzero lettuce and $\alpha : V \to V$ a linear transformation, then $\mu(\alpha X) \geq \mu(X)$.

Proof. Let $Y = \alpha X$. Choose nonzero $Y' \subset Y$ such that $\mu(Y) = \frac{r(Y)-r(Y')}{\dim_F FY - \dim_F FY'}$. Let $X' = \alpha^{-1}(Y') \cap X$. (To see that X' is a sublettuce of X, we observe that $\alpha^{-1}(Y') \cap X = \alpha^{-1}(FY') \cap X$, for if $x \in X$ with $\alpha x \in FY'$, then $\alpha x \in FY' \cap Y = Y'$ as Y' is a sublettuce of Y. Now $\alpha^{-1}(FY') \cap X$ is evidently a sublettuce of X.)

As $r(X') = r(Y') + r(\ker \alpha \cap X)$ and $r(X) = r(Y) + r(\ker \alpha \cap X)$, we get $r(X) - r(X') = r(Y) - r(Y')$.

Now as $\dim_F FX' \leq \dim_F FY' + (\dim_F FX - \dim_F FY)$, combining with the above we find

$$\mu(X) \leq \frac{r(X) - r(X')}{\dim_F FX - \dim_F FX'} \leq \frac{r(Y) - r(Y')}{\dim_F FY - \dim_F FY'} = \mu(Y)$$

as desired. □

Lemma 46. Let X, Y be nonzero lettuces with $FX = FY$. If Y is maximal, then

$$\mu(Y)\chi(X, Y) \geq \mu(X).$$

Proof. The proof amounts to computations on the definitions of μ and χ. Let $n = \dim_F FX$. Choose $X' \subset X$ and $Y' \subset Y$ such that

$$\chi(X, Y) = \frac{r(X) - r(X')}{r(Y')},$$

with Y' nonzero and $\langle X', Y' \rangle = 0$. In particular, $FX' \cap FY' = 0$ and $\dim_F FX' + \dim_F FY' \leq n$.

As Y is maximal, we find that

$$\frac{r(Y)}{n} = \mu(Y) \leq \frac{r(Y) - r(Y')}{n - \dim_F FY'}$$

so

$$\frac{r(Y')n}{r(Y)} \leq \dim_F FY' \leq n - \dim_F FX' \leq \frac{r(X) - r(X')}{\mu(X)},$$

the three inequalities coming respectively from the maximality of Y, the condition $\langle X', Y' \rangle = 0$, and the definition of $\mu(X)$. Rearranging gives the claim. □

The following result brings together all of the preceding preliminary work. The most important hypothesis of the theorem is that $\chi(Y, X) \geq 1$, which encodes the fact that the regulator of the number field K (that will appear in the next theorem) is zero.

Theorem 47. If F is a local field and X, Y are nonzero lettuces with $FX = FY$, $\exp_p \langle X, Y \rangle \subset \overline{\mathbb{Q}}$, and $\chi(Y, X) \geq 1$, then $\dim_F FX \geq r(X)/2$.

Proof. By Lemma 44, we can choose nonzero $X_1 \subset X$ such that $\mu(X_1) \geq \frac{r(X)}{\dim_F FX}$, so it suffices to show that $\mu(X_1) \leq 2$. Let $V = FX_1$.

Now let $Y_1 = \pi_V(Y)$, that is, the projection of Y to V. We certainly have $V = FX_1 = FY_1$. As $X_1 \subset V$ we have $\left\langle X_1, V^\perp \cap Y \right\rangle = 0$, so by the definition of $\chi(Y, X)$,

$$1 \leq \chi(Y, X) \leq \frac{r(Y) - r(V^\perp \cap Y)}{r(X_1)}.$$

Then we have $r(Y) \geq r(V^\perp \cap Y) + r(X_1)$. As $V^\perp \cap Y$ is the kernel of the map $\pi_V : Y \to Y_1$, we get $r(Y) = r(V^\perp \cap Y) + r(Y_1)$, so $r(Y_1) \geq r(X_1)$.

By Lemma 44, we can find nonzero $Y_2 \subset Y_1$ such that Y_2 is maximal with

$$\mu(Y_2) \geq \frac{r(Y_1)}{\dim_F V} \geq \frac{r(X_1)}{\dim_F V} \geq \mu(X_1).$$

Let $W = FY_2$ and $X_2 = \pi_W(X_1)$. Then $W = FX_2 = FY_2$. By Lemma 45 we have $\mu(X_2) \geq \mu(X_1)$. Furthermore, we have

$$\langle X_2, Y_2 \rangle = \langle X_1, Y_2 \rangle \subset \langle X_1, Y_1 \rangle = \langle X_1, Y \rangle \subset \langle X, Y \rangle$$

so we may apply Lemma 46 and Theorem 43 to X_2, Y_2 and find that

$$\begin{aligned} \mu(X_2) \leq & \mu(Y_2)\chi(X_2, Y_2), \\ & \chi(X_2, Y_2)(\mu(Y_2) - 1) \leq 1. \end{aligned}$$

Combining all of the above gives

$$\mu(X_1) \le \mu(X_2) \le \mu(Y_2)\chi(X_2, Y_2) \le \frac{\mu(Y_2)}{\mu(Y_2) - 1} = \frac{1}{1 - (1/\mu(Y_2))} \le \frac{1}{1 - (1/\mu(X_1))}$$

so

$$\mu(X_1)(1 - (1/\mu(X_1))) \le 1$$

which gives $\mu(X_1) \le 2$ as desired.

(To perform the above manipulations we assumed that $\mu(X_1) > 1$; if this is not true we get $\mu(X_1) \le 1 < 2$ immediately. Also to apply Theorem 43 we require that $r(Y_2) > \dim_F FY_2$, which is implied by $\mu(Y_2) \ge \mu(X_1) > 1$.) □

Finally we conclude with a bound on the Leopoldt defect. The following result shows that the Leopoldt defect $r_K - r_{K,p}$ is at most $\frac{1}{2}r_K$.

Theorem 48. For all number fields K,

$$r_{K,p} \ge \frac{1}{2} r_K.$$

Proof. Let $F = K_p$.

Again, let $\Upsilon = (E_{K,\varepsilon})^\Theta$, the set of tuples of elements from $E_{K,\varepsilon}$ indexed by $\Theta = \Theta_K$. Let $\Delta \subset \Theta$ be any subset of size r_K, and choose $\upsilon \in \Upsilon$ so that $\operatorname{rank}_\mathbb{Z} \langle \upsilon_\delta \rangle_{\delta \in \Delta} = r_K$. (The coordinates υ_θ for $\theta \in \Theta \setminus \Delta$ can be chosen arbitrarily from $E_{K,\varepsilon}$.) For $\delta \in \Delta$, we define $x_\delta \in \mathbb{C}^\Theta$ and $x'_\delta \in F^\Theta$ by

$$(x_\delta)_\theta = \log(\theta \upsilon_\delta),$$
$$(x'_\delta)_\theta = \log_p(\theta \upsilon_\delta).$$

Now we define lettuces

$$X = \sum_{\delta \in \Delta} \mathbb{Z} x_\delta \subset \mathbb{C}^\Theta, \qquad Y = \mathbb{Z}^\Theta \subset \mathbb{C}^\Theta,$$
$$X_p = \sum_{\delta \in \Delta} \mathbb{Z} x'_\delta \subset F^\Theta, \qquad Y_p = \mathbb{Z}^\Theta \subset F^\Theta.$$

The x_δ and x'_δ are each $\mathbb{Z}$-linearly independent (otherwise we'd have a relation on the υ_δ in $E_{K,\varepsilon}$ for $\delta \in \Delta$), so $r(X) = r(X_p) = r_K$ and $r(Y) = r(Y_p) = [K : \mathbb{Q}]$. The x_δ are $\mathbb{C}$-linearly independent (equivalently, the regulator of K is nonzero) so $\dim_\mathbb{C} \mathbb{C}X = r_K$. We wish to investigate $\dim_F FX_p$.

We define group isomorphisms $\rho : X_p \to X$ and $\sigma : Y_p \to Y$ by sending the corresponding bases to each other; that is, $\rho(x'_\delta) = x_\delta$ and $\sigma((b_\theta)_{\theta \in \Theta}) = (b_\theta)_{\theta \in \Theta}$. As these are group isomorphisms, they send subgroups to subgroups, and furthermore send direct summands (as free abelian groups) to direct summands.

In general, group isomorphisms between lettuces need not send sublettuces to sublettuces. However, as the bases of X and Y are $\mathbb{C}$-linearly independent, any direct summand of X or Y is necessarily a sublettuce of X or Y, respectively. Any sublettuce of X_p or Y_p is a direct summand, and as ρ and σ sends direct summands to direct summands, therefore ρ and σ also send sublettuces to sublettuces.

Let $W = FX_p$, and $Z_p = \pi_W Y_p$, so that $FZ_p = FX_p$.

We now claim that if $X'_p \subset X_p$ and $Z'_p \subset Z_p$ are sublettuces with $\langle X'_p, Z'_p \rangle = 0$, then $r(X'_p) + r(Z'_p) \le r(Z_p)$. Let $Y'_p \subset Y_p$ be the preimage of Z'_p under the map $\pi_W : Y_p \to Z_p$. Y'_p is a sublettuce as all direct summands of $Y_p = \mathbb{Z}^\Theta$ are sublettuces (in fact, any preimage of a sublettuce under a linear transformation is also a sublettuce, as shown in Lemma 45). We have

$$r(Y'_p) = r(Z'_p) + r(\ker \pi_W \cap Y_p) = r(Z'_p) + (r(Y_p) - r(Z_p)),$$

so it suffices to show that $r(X'_p) + r(Y'_p) \le r(Y_p) = [K : \mathbb{Q}]$. Now let $X' = \rho(X'_p) \subset X$ and $Y' = \sigma(Y'_p) \subset Y$. Certainly $r(X') = r(X'_p)$ and $r(Y') = r(Y'_p)$, so we need to show that $r(X') + r(Y') \le [K : \mathbb{Q}]$. Once we show that $\langle X', Y' \rangle = 0$, we will have

$$r(X') + r(Y') = \dim_{\mathbb{C}} \mathbb{C}X' + \dim_C \mathbb{C}Y' = \dim_{\mathbb{C}}(\mathbb{C}X' + \mathbb{C}Y') \le \dim_{\mathbb{C}} \mathbb{C}^\Theta = [K : \mathbb{Q}]$$

as $\langle X', Y' \rangle = 0$ implies that $\mathbb{C}X' \cap \mathbb{C}Y' = 0$. Thus, to demonstrate our claim, it remains to be shown that $\langle X', Y' \rangle = 0$. This is the only step of our proof that requires use of the properties of the υ_δ.

Suppose we have $x \in X'$ and $y \in Y'$; write $x = \rho(x_p)$ and $y = \sigma(y_p)$ with $x_p \in X'_p$ and $y_p \in Y'_p$. Let

$$x_p = \sum_{\delta \in \Delta} a_\delta x'_\delta = \left(\sum_{\delta \in \Delta} a_\delta \log_p(\theta \upsilon_\delta) \right)_{\theta \in \Theta}, \qquad y_p = (b_\theta)_{\theta \in \Theta},$$

which gives us

$$x = \left(\sum_{\delta \in \Delta} a_\delta \log(\theta \upsilon_\delta) \right)_{\theta \in \Theta}, \qquad y = (b_\theta)_{\theta \in \Theta}.$$

Then $\langle x_p, y_p \rangle \in \langle X'_p, Y'_p \rangle = \langle X'_p, Z'_p \rangle = 0$ so

$$\sum_{\delta \in \Delta, \theta \in \Theta} a_\delta b_\theta \log_p(\theta \upsilon_\delta) = 0.$$

As $\upsilon_\delta \in E_{K,\varepsilon}$, therefore $\theta_{\mathfrak{p}}(\upsilon_\delta) \in \mathcal{U}_\varepsilon$ and $|\theta_{\mathfrak{p}}(\upsilon_\delta - 1)| < \varepsilon$ for every $\mathfrak{p}$, so in particular $|\theta(\upsilon_\delta - 1)| < \varepsilon$ and $\theta(\upsilon_\delta) \in \mathcal{U}_\varepsilon$ for every $\theta \in \Theta$.

In particular, $\prod_{\delta\in\Delta,\theta\in\Theta}(\theta u_\delta)^{a_\delta b_\theta}$ is in $\mathcal{U}_\varepsilon \subset \mathcal{U}_1$, and being in the kernel of $\log_p$, equals a k-th root of unity for some positive integer k (in fact a power of p). Then we have

$$\prod_{\delta\in\Delta,\theta\in\Theta}(\theta u_\delta)^{ka_\delta b_\theta} = 1,$$

so, taking the log, we get

$$\sum_{\delta\in\Delta,\theta\in\Theta} ka_\delta b_\theta \log(\theta u_\delta) = 0,$$

so $\langle x, y\rangle = 0$. Therefore $\langle X', Y'\rangle = 0$, which completes the claim that if $X'_p \subset X_p$ and $Z'_p \subset Z_p$ satisfy $\langle X'_p, Z'_p\rangle = 0$, then $r(X'_p) + r(Z'_p) \le r(Z_p)$. Furthermore, if X'_p is nonzero, then

$$\frac{r(Z_p) - r(Z'_p)}{r(X'_p)} \ge 1.$$

Since $\chi(Z_p, X_p)$ equals the expression on the left for some such X'_p, Z'_p, therefore $\chi(Z_p, X_p) \ge 1$.

Now it is evident that

$$\exp_p \langle X_p, Z_p\rangle = \exp_p \langle X_p, Y_p\rangle \subset \overline{\mathbb{Q}}$$

as the $\upsilon_\delta \in E_{K,\varepsilon} \subset K$ are algebraic over $\mathbb{Q}$. Then we can apply Theorem 47 to X_p, Z_p and find that

$$\dim_F FX_p \ge r(X_p)/2 = r_K/2.$$

Now recall the functions $\psi_\upsilon : \mathbb{Z}_p^\Theta \to \overline{E_{K,\varepsilon}}$ and $\phi_{A,\upsilon} : A^\Theta \to \mathbb{C}_p^\Theta$ from section 4, the latter defined by

$$(\phi_{A,\upsilon}(\alpha))_\delta = \sum_{\theta\in\Theta} \alpha_\theta(\log_p(\delta\upsilon_\theta)),$$

with A a subring of $\mathbb{C}_p$. We wish to bound $\dim_F \ker \phi_{F,\upsilon}$. Certainly

$$X_p \subset \operatorname{im} \phi_{F,\upsilon}$$

so, knowing that $\dim_F F^\Theta = |\Theta| = [K : \mathbb{Q}]$, we get

$$\dim_F \ker \phi_{F,\upsilon} \le [K : \mathbb{Q}] - \dim_F FX_p \le [K : \mathbb{Q}] - (r_K/2).$$

Now by Lemmas 23 and 24 we get

$$\operatorname{rank}_{\mathbb{Z}_p} \ker \psi_\upsilon \le \operatorname{rank}_{\mathbb{Z}_p} \ker \phi_{\mathbb{Z}_p,\upsilon} \le \dim_F \ker \phi_{F,\upsilon} \le [K : \mathbb{Q}] - (r_K/2).$$

Now subtracting from $\operatorname{rank}_{\mathbb{Z}_p} \mathbb{Z}_p^\Theta = [K : \mathbb{Q}]$ we get

$$r_{K,p} = \operatorname{rank}_{\mathbb{Z}_p} \overline{E_{K,\varepsilon}} \ge \operatorname{rank}_{\mathbb{Z}_p} \operatorname{im} \psi_\upsilon \ge r_K/2,$$

as desired. □

References

[1] M.F. Atiyah, I.G. MacDonald, *Introduction to Commutative Algebra*, Addison-Wesley Publishing Company, 1969.

[2] James Ax, *On the units of an algebraic number field*, Illinois J. Math., vol. 9 (1965), 584-589.

[3] A. Baker, *Linear forms in the logarithms of algebraic numbers*, Mathematika, vol. 13 (1966), 204-216.

[4] A. Brumer, *On the units of algebraic number fields*, Mathematika, vol. 14 (1967), 121-124.

[5] W. Gröbner, *Moderne Algebraisc Geometrie, die idealtheoretischen Grundlagen*, Springer-Verlag, 1949.

[6] D.E. Knuth, *The Art of Computer Programming I*, 3rd ed., Addison-Wesley, 1997.

[7] H.W. Leopoldt, *Zur Arithmetic in Abelschen Zahlkörpern*, J. Reine und Angew. Math., vol. 209 (1962), 54 - 71.

[8] D.W. Masser, *On polynomials and exponential polynomials in several complex variables*, Invent. math., vol. 63 (1981), 81-95.

[9] J.S. Milne, *Algebraic Number Theory (v3.02)*, Available at `www.jmilne.org/math/`, 2009.

[10] M. Waldschmidt, *Transcendence et exponentielles en plusiers variables*, Invent. math., vol. 63 (1981), 97-128.

[11] L.C. Washington, *Introduction to Cyclotomic Fields*, Springer-Verlag, 1982.

[12] O. Zariski, P. Samuel, *Commutative Algebra I, II*, Springer-Verlag, 1968.

v1: 2012 December 11