Simons Symposia

Jennifer S. Balakrishnan
Noam Elkies · Brendan Hassett
Bjorn Poonen · Andrew V. Sutherland
John Voight *Editors*

Arithmetic Geometry, Number Theory, and Computation

Springer

Simons Symposia

Working to foster communication and enable interactions between experts, volumes in the Simons Symposia series bring together leading researchers to demonstrate the powerful connection of ideas, methods, and goals shared by mathematicians, theoretical physicists, and theoretical computer scientists.

Symposia volumes feature a blend of original research papers and comprehensive surveys from international teams of leading researchers in thriving fields of study. This blend of approaches helps to ensure each volume will serve not only as an introduction for graduate students and researchers interested in entering the field, but also as the premier reference for experts working on related problems.

The Simons Foundation at its core exists to support basic, discovery-driven research in mathematics and the basic sciences, undertaken in pursuit of understanding the phenomena of our world without specific applications in mind. The foundation seeks to advance the frontiers of research in mathematics and the basic sciences by creating strong collaborations and fostering cross-pollination of ideas between investigators, leading to unexpected breakthroughs and a deeper understanding of the world around us.

More information about this series at https://link.springer.com/bookseries/15045

Jennifer S. Balakrishnan • Noam Elkies
Brendan Hassett • Bjorn Poonen
Andrew V. Sutherland • John Voight
Editors

Arithmetic Geometry, Number Theory, and Computation

Springer

Editors
Jennifer S. Balakrishnan
Department of Mathematics and Statistics
Boston University
Boston, MA, USA

Noam Elkies
Department of Mathematics
Harvard University
Cambridge, MA, USA

Brendan Hassett
Institute for Computational and
Experimental Research in Mathematics
Brown University
Providence, RI, USA

Bjorn Poonen
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA, USA

Andrew V. Sutherland
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA, USA

John Voight
Department of Mathematics
Dartmouth College
Hanover, NH, USA

ISSN 2365-9564 ISSN 2365-9572 (electronic)
Simons Symposia
ISBN 978-3-030-80913-3 ISBN 978-3-030-80914-0 (eBook)
https://doi.org/10.1007/978-3-030-80914-0

Mathematics Subject Classification: 11G40, 11G10, 11G30, 14H40, 11G05

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

In fall 2017, the collaboration "Arithmetic Geometry, Number Theory, and Computation" was created through awards from the Simons Foundation to our institutions. We are grateful for the support of Jim Simons, Yuri Tschinkel, and other leaders at the Foundation that made this possible.

Our common perspective is that advances in computational techniques accelerate research in arithmetic geometry and number theory, both as a source of data and examples and as an impetus for effective results. The dynamic interplay between experiment, theory, and computation has historically played a pivotal role in the development of number theory. In the eighteenth and nineteenth centuries, Euler and Gauss undertook extensive calculations by hand in the pursuit of data to help formulate and refine conjectures and to produce a source of counterexamples. In the twentieth century, systematic computations of elliptic curves and their L-functions led to the formulation of the Sato–Tate and modularity conjectures, both of which have now been proved, and the conjecture of Birch and Swinnerton-Dyer, which remains open but has been proved in some special cases.

In the twenty-first century, the frontier of research in arithmetic geometry has moved on to curves of higher genus, abelian varieties, and K3 surfaces. Although available computational resources have grown dramatically, the development and implementation of practical algorithms have lagged behind the theory; we seek to correct this imbalance. In contrast to the situation with elliptic curves, in higher dimensions, brute force computation yields very little. To obtain practical algorithms, one must exploit the theoretical infrastructure of modern arithmetic geometry.

We pursue these goals by recruiting graduate students and research scientists with expertise in both computation and fundamental mathematics. Members of our group prove theorems, develop algorithms, explore guiding examples, collect data on arithmetic objects, and support scholarly resources like the *L-functions and Modular Forms Database (LMFDB)* http://www.lmfdb.org.

This volume offers a sampling of our work over the last three years. In addition to the six of us, many authors were directly supported by the collaboration: Eran Assaf, Alex Best, Raymond van Bommel, Edgar Costa, Alex Cowan, Lassina

Dembélé, Maarten Derickx, Sachi Hashimoto, Dohyeong Kim, David Lowry-Duda, Benjamin Matschke, David Roe, and Ciaran Schembri. We are also grateful for the contributions of senior affiliated scientists Andrew Booker, John Cremona, and Kiran Kedlaya.

A number of manuscripts address fundamental questions informed by computation or inspired by the need for effective algorithms. Cowan's contribution "Conjecture: 100% of Elliptic Surfaces over $\mathbb{Q}$ Have Rank Zero" builds on previous work of Nagao and Rosen–Silverman about quantitative measures of growth of ranks of fibers of elliptic surfaces. "Congruent Number Triangles with the Same Hypotenuse" by Lowry-Duda explores coincidences among solutions to the congruent number problem; motivated by observations on Dirichlet series—and encouraged by numerical experiments—the author uncovers some interesting Diophantine geometry. Poonen offers a new proof of Siegel's 1926 theorem on finiteness of integral points in "The S-Integral Points on the Projective Line Minus Three Points via étale Covers and Skolem's Method." His work was inspired by, but is different from, M. Kim's proof via a nonabelian analogue of the Skolem–Chabauty method.

Several manuscripts focus on effective approaches to S-unit equations. "A Robust Implementation for Solving the S-Unit Equation and Several Applications"—by Alvarado, Koutsianas, Malmskog, Rasmussen, Vincent, and West—presents their new algorithms for these problems. Best and Matschke use a reduction to Mordell equations to enumerate "Elliptic Curves with Good Reduction Outside of the First Six Primes."

The development of databases of low genus curves has demanded improved techniques for computing their invariants. "Efficient Computation of BSD Invariants in Genus 2" by van Bommel discusses quantities associated with the Birch and Swinnerton-Dyer conjecture. Picard groups and other invariants of higher-dimensional varieties are also of interest. A p-adic variational approach is developed in "Effective Obstructions to Lifting Tate Classes from Positive Characteristic" by Costa and E. Sertöz.

The work of our collaboration tends to focus attention on examples with unusual properties. Fontaine and Abrashkin showed there are no nonzero abelian varieties over $\mathbb{Q}$ with good reduction everywhere, and we can enumerate elliptic curves with just a few small primes of bad reduction. Mascot, Sijsling, and Voight offer a nice example in this vein in "A Prym Variety with Everywhere Good Reduction over $\mathbb{Q}(\sqrt{61})$." The manuscript "On Rational Bianchi Newforms and Abelian Surfaces with Quaternionic Multiplication"—by Cremona, Dembélé, Pacetti, Schembri, and Voight—presents beautiful modular forms described in the title and the associated very special abelian surfaces.

Modular forms are a major focus of the collaboration, given their historical importance for number theory and across mathematics. Assaf focuses on algorithms for nonstandard levels in "Computing Classical Modular Forms for Arbitrary Congruence Subgroups." We offer a comprehensive survey of computations of modular forms by Best, Bober, Booker, Costa, Cremona, Derickx, Lowry-Duda, Lee, Roe, Sutherland, and Voight. D. Kim lays the groundwork for new algorithms

in his contribution "Linear Dependence Among Hecke Eigenvalues." "A Database of Hilbert Modular Forms" by Donnelly and Voight presents what is available on Hilbert modular forms in the LMFDB.

One of the most important tools for effective computation of rational points on curves is the Chabauty–Coleman method, the prototype for nonabelian extensions developed over the last decade. We are especially interested in implementing these techniques for large numbers of examples. Contributions in this direction include "Curves with Sharp Chabauty–Coleman Bound" by Gajović, "Square Root Time Coleman Integration on Superelliptic Curves" by Best, "Computing Rational Points on Genus 3 Hyperelliptic Curves" by de Frutos Fernández and Hashimoto, and "Chabauty–Coleman Computations on Rank 1 Picard Curves" by Hashimoto and Morrison.

Classification of algebraic varieties over finite fields and their Weil polynomials is an emerging focus of number-theoretic databases. "Restrictions on Weil Polynomials of Jacobians of Hyperelliptic Curves"—by Costa, Donepudi, Fernando, Karemaker, Springer, and West—develops results that emerged from experiments, building on theorems in genus two by Howe, Nart, and Ritzenthaler. Dupuy, Kedlaya, Roe, and Vincent present recent progress on building comprehensive databases in "Isogeny Classes of Abelian Varieties over Finite Fields in the LMFDB."

Finally, we include two manuscripts directly documenting recent innovations to the LMFDB. "Zen and the Art of Database Maintenance" by Costa and Roe describe their efforts to modernize the internal architecture—how objects are represented in the system. This work was crucial to recent expansions to the LMFDB and also formed the foundation for the new online directory of research talks and conferences http://researchseminars.org supported by the collaboration. "Visualizing Modular Forms" by Lowry-Duda is a product of serendipity with the program *Illustrating Mathematics* held at the Institute for Computational and Experimental Research in Mathematics (ICERM) in Fall 2019. The "badges" used to represent modular forms on the LMFDB have been updated based on these ideas. We expect these will lead to new insights linking modular forms to geometric and probabilistic processes.

Acknowledgments Balakrishnan is supported by NSF grant DMS-1702196, the Clare Boothe Luce Professorship (Henry Luce Foundation), Simons Foundation grant #550023, and a Sloan Research Fellowship. Elkies is supported by NSF grant DMS-1502161 and Simons Foundation grant #550031. Hassett is supported by NSF grant DMS-1701659 and Simons Foundation grant #546235. Poonen is supported by NSF grant DMS-1601946 and Simons Foundation grants #402472 and #550033. Sutherland also is supported by Simons Foundation grant #550033. Voight is supported by Simons Foundation grant #550029.

Boston University — Jennifer S. Balakrishnan
Harvard University — Noam Elkies
Brown University/ICERM — Brendan Hassett
MIT — Bjorn Poonen
MIT — Andrew V. Sutherland
Dartmouth College — John Voight
December 28, 2020

Contents

A Robust Implementation for Solving the *S*-Unit Equation and Several Applications

Alejandra Alvarado, Angelos Koutsianas, Beth Malmskog, Christopher Rasmussen, Christelle Vincent, and Mckenzie West

Abstract Let K be a number field, and S a finite set of places in K containing all infinite places. We present an implementation for solving the S-unit equation $x + y = 1$, $x, y \in \mathcal{O}_{K,S}^{\times}$ in the computer algebra package SageMath. This paper outlines the mathematical basis for the implementation. We discuss and reference the results of extensive computations, including exponent bounds for solutions in many fields of small degree for small sets S. As an application, we prove an asymptotic version of Fermat's Last Theorem for totally real cubic number fields with bounded discriminant where 2 is totally ramified. In addition, we use the implementation to find all solutions to some cubic Ramanujan-Nagell equations.

A. Alvarado
Department of Mathematics and Computer Science, Eastern Illinois University, Charleston, IL, USA
e-mail: aalvarado2@eiu.edu

A. Koutsianas
Department of Mathematics, University of British Columbia, Vancouver, BC, Canada
e-mail: akoutsianas@math.ubc.ca

B. Malmskog
Department of Mathematics and Computer Science, Colorado College, Colorado Springs, CO, USA

C. Rasmussen (✉)
Department of Mathematics and Computer Science, Wesleyan University, Middletown, CT, USA
e-mail: crasmussen@wesleyan.edu

C. Vincent
Department of Mathematics and Statistics, University of Vermont, Burlington, VT, USA
e-mail: christelle.vincent@uvm.edu

M. West
Department of Mathematics, University of Wisconsin-Eau Claire, Eau Claire, WI, USA
e-mail: westmr@uwec.edu

J. S. Balakrishnan et al. (eds.), *Arithmetic Geometry, Number Theory, and Computation*, Simons Symposia, https://doi.org/10.1007/978-3-030-80914-0_1

1 Introduction

In 1909, Thue proved there are only finitely many integral solutions to what we now call the Thue equation; i.e, that for any $\mathbb{Q}$-irreducible binary form $F(X, Y)$ of degree at least 3, defined over the integers, there are only finitely many solutions $(x, y) \in \mathbb{Z}^2$ to the equation

$$F(x, y) = c,$$

where c is any non-zero integer [38]. Thue accomplished this by formally factoring F into linear terms of the form $(x - \alpha y)$, where α is algebraic, then bounding the quality of rational approximations of α in terms of the size of x and y. Thus bounds on integer solutions to the Thue equation arose out of the theory of approximating algebraic numbers by rationals. Thue's theorem was generalized by Siegel [33][1] and then Mahler [25]. These generalizations gave rise to a central fact of modern computational number theory: if K is a number field, and S a finite list of places of K including all infinite places, then there are only finitely many solutions (x, y) to the equation

$$x + y = 1, \qquad x, y \in \mathscr{O}_{K,S}^{\times}. \tag{1}$$

Here, $\mathscr{O}_{K,S}^{\times}$ is the unit group of the ring $\mathscr{O}_{K,S}$ of S-integers in K. We refer to (1) as the *S-unit equation*. In this paper, we describe an algorithm to determine the complete set of solutions to the S-unit equation for general K and S. More generally, for fixed $a, b \in \mathscr{O}_{K,S}$, we can see that the equation $ax + by = 1$ will also have only finitely many solutions by expanding the set S to include all primes dividing a and b and searching for solutions to (1). Thus it suffices to solve (1) to address the more general case, and we focus on (1) here (though it should be remarked that this is not the most efficient way to solve $ax + by = 1$).

The work of Gelfond and Schneider, resolving Hilbert's seventh problem in the affirmative (all irrational algebraic powers of algebraic numbers are transcendental once trivial cases are ignored), determined lower bounds on the absolute value of a $\mathbb{Q}$-linear combination of two $\mathbb{Q}$-linearly independent logarithms of algebraic numbers. Alan Baker's 1967 theorem [1] generalized these results to the case of many logarithms. Baker, Wüstholz, and many others continued to improve these bounds. Naturally, one should ask if similar results are available over local fields, and indeed such results began to appear quickly. In 1968, Brumer proved the first analogue of Baker's work for p-adic logarithms [8], followed by many improvements and generalizations, such as the results of Yu [44, 45, 46]. Improvements in both the archimedean and nonarchimedean cases continue to appear, such as in [20, 4, 47, 19].

[1] See also the recent translation [17] by Fuchs.

For any choice of K and S, $\mathscr{O}_{K,S}^{\times}$ is a finitely generated $\mathbb{Z}$-module. Fixing a basis $\rho_1, \ldots, \rho_t$ for the torsion free part, we can express any $x \in \mathscr{O}_{K,S}^{\times}$ as $x = \xi \cdot \prod_{i=1}^{t} \rho_i^{a_i}$ for some root of unity $\xi \in K$ and some $a_i \in \mathbb{Z}$. Building on the lower bounds for linear combinations of logarithms, Győry [18] determined effectively computable bounds for the exponents a_i. This was a great victory for computational number theory, as this provably restricted all solutions to (1) to a finite search space. Unfortunately, the demonstrated bounds were enormous and as a matter of practice, it was computationally infeasible to conduct an exhaustive search for solutions, even in the very simplest cases. Baker and Davenport devised a clever method of reducing the bounds in special cases in [2]. However, in [14], de Weger built on the ideas of Baker-Davenport to develop a powerful general method of algorithmically reducing the bounds to a manageable size, relying on the lattice basis reduction algorithm of Lenstra, Lenstra, and Lovász [24] (henceforth referred to as the "LLL algorithm"). Though it is has not been proven that de Weger's method will always reduce the bounds coming from the results in linear forms of logarithms, this is the rule in practice. In many cases, de Weger's approach provides sufficient improvements that, with careful sieving (or sometimes even with only brute force), the entire search space can be exhausted and complete lists of solutions can be enumerated.

Beyond the improvements provided by LLL-based reduction, many mathematicians have developed further algorithms for efficiently searching below the "LLL bounds" provided by de Weger's work. Two powerful examples are reported in [43] and [37]. Increasingly, the theoretical improvements (assisted by technological improvements) have pushed ambitious and interesting computational problems within reach. For example, Smart determined the entire set of all genus 2 curves over $\mathbb{Q}$ with good reduction away from 2, based in part on solving (1) for a family of number fields unramified away from 2 [35].

We have written a package of Python functions for inclusion in the computer algebra system SageMath [31], which solves the S-unit equation (1) over any number field K and for any finite set S of finite places. As experienced readers may expect, the package is not practical when either $[K : \mathbb{Q}]$ or $|S|$ is too large, although there is no theoretical obstruction. While this package is the independent creation of the authors, it is based in part on the descriptions of algorithms implemented by Smart [34, 35, 36]. Specifically, we follow Smart's development in determining initial large bounds, including the numbering of constants, in [34], with some adjustments and small corrections. In reducing the bounds, we follow [36], again with some adjustments. The sieving step is based on ideas cited by Smart [35] as due to others (as noted in Sect. 6) but has been redeveloped in new notation and style. We include proofs of our versions of results when we made adjustments to versions in the literature. To the authors' knowledge, our package is the first publicly available implementation for solving the S-unit equation over any field other than $\mathbb{Q}$; the present article describes the algorithm and its implementation. The implementation was a highly non-trivial undertaking, involving efforts spreading over more than 7 years on the parts of individuals and the entire team.

We also provide new results facilitated by our implementation. In particular, we first provide a discussion of and link to explicit exponent bounds for solutions of the S-unit equation in all cases (K, S) where $K/\mathbb{Q}$ is ramified only at primes above some subset of $\{2, 3\}$ and

$$[K : \mathbb{Q}] \leq 5, \qquad S \subseteq \{\mathfrak{p} \subseteq \mathscr{O}_K : \mathfrak{p} \mid 6\}.$$

We solve the S-unit equation in the 13 totally real cubic number fields K in which 2 is totally ramified and the absolute discriminant of K, Δ_K, satisfies $|\Delta_K| \leq 2000$, and we use these results to verify that an asymptotic version of Fermat's Last Theorem holds over these fields. Finally, we find all solutions to certain cubic Ramanujan-Nagell equations.

1.1 Overview

The organization of the paper proceeds as follows. We introduce certain notations in Sect. 2. In Sect. 3, we review the relevant work of Baker-Wüstholz and Yu. This is used in Sect. 4 to establish a "pre-LLL" exponent bound for each place in S. In Sect. 5, we explain the process of using LLL to reduce these exponent bounds—the approach is different for archimedean and nonarchimedean places. In Sect. 6, we describe the sieve for further constraining the final search space. We devote Sect. 7 to a discussion of our experimental observations, having now executed our algorithm in several dozen cases. We highlight a special condition (S contains only one finite place) under which a significant improvement in the search space can be obtained. Although narrow in scope, the special condition is sufficiently natural, and the savings sufficiently nontrivial, as to warrant its discussion. Finally, Sect. 8 introduces two applications: an asymptotic version of Fermat's Last Theorem over totally real cubic fields and a solution to a cubic variant of the Ramanujan-Nagell equation.

Acknowledgments We are delighted to recognize the Institute for Computational and Experimental Research in Mathematics for both funding and hosting a 2017 collaboration during which a great deal of this project was completed. Part of this work began at the 2014 workshop SageDays 62, and we would like to thank Anna Haensch and Lola Thompson for organizing that workshop and Microsoft Research and The Beatrice Yormark Fund for Women in Mathematics for funding. Some of the work was supported by the van Vleck fund at Wesleyan University. The authors would like to thank many people for helpful conversations that led to improvements in the code and gave direction to this project, including Bjorn Poonen, Andrew Sutherland, and Norman Danner. We would also like especially to thank David Roe for his contributions to refining and reviewing the code for inclusion in SageMath.

The third author was partially supported in this work by NSA Grant #H98230-16-1-0300. We are very grateful to the anonymous referees for their careful reading of this work and their many helpful comments which have improved the quality of this paper.

2 Notation

2.1 S-Units in Number Fields

Throughout this paper, we let $\bar{\mathbb{Q}}$ denote the algebraic closure of $\mathbb{Q}$ inside $\mathbb{C}$, the field of complex numbers. Unless stated otherwise, we fix the following notation throughout:

K	a number field (assumed to be a subfield of $\bar{\mathbb{Q}}$),
d_K	the absolute degree $[K : \mathbb{Q}]$
w	the number of distinct roots of unity in K
Δ_K	the absolute discriminant of $K/\mathbb{Q}$
$\mathscr{O}_K$	the ring of integers of K
$e_{\mathfrak{p}}$	the ramification index of $\mathfrak{p}$ in $K/\mathbb{Q}$
$f_{\mathfrak{p}}$	the inertial degree of the prime $\mathfrak{p} \subseteq \mathscr{O}_K$ over the rational prime $\mathfrak{p} \cap \mathbb{Z}$
r	the rank of $\mathscr{O}_K^\times$ as a $\mathbb{Z}$-module
S_{fin}	a set $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_s\}$ of s finite places of K
S_∞	the set $\{\mathfrak{p}_{s+1}, \ldots, \mathfrak{p}_{r+s+1}\}$ of all infinite places of K
S	$S_{\text{fin}} \cup S_\infty = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_{r+s+1}\}$
$S_{\mathbb{Q}}$	the set of places of $\mathbb{Q}$ which extend to places of K in S
$\mathscr{O}_{K,S}$	the ring of S-integers in K
$\mathscr{O}_{K,S}^\times$	the group of S-units in K
t	the rank of $\mathscr{O}_{K,S}^\times$ as a $\mathbb{Z}$-module (so $t = r + s$)
ρ_0	a root of unity generating the torsion part of $\mathscr{O}_{K,S}^\times$
$\rho_1, \ldots, \rho_t$	an ordered basis for the torsion-free part of the $\mathbb{Z}$-module $\mathscr{O}_{K,S}^\times$
$\boldsymbol{\rho}$	the ordered list $[\rho_0, \rho_1, \ldots, \rho_t]$

If $f(x) \in \mathbb{Z}[x]$ is a monic and irreducible polynomial, we let K_f denote the number field $\mathbb{Q}(\xi)$, where ξ is a root of $f(x)$. Always, log denotes the principal branch of the complex logarithm function, with argument in $(-\pi, \pi]$.

2.2 Absolute Values and Completions

Each place of K determines an associated absolute value, $|\cdot|_{\mathfrak{p}}$, which we now describe.

Let $|\cdot|$ denote the usual absolute value on $\mathbb{C}$. If $\mathfrak{p}$ is an infinite place, choose $\sigma_{\mathfrak{p}}\colon K \to \mathbb{C}$, an embedding corresponding to $\mathfrak{p}$. The associated absolute value depends on whether $\mathfrak{p}$ is a real or complex (meaning non-real) place of K:

$$|\alpha|_{\mathfrak{p}} := \begin{cases} \left|\sigma_{\mathfrak{p}}(\alpha)\right| & \mathfrak{p} \text{ is real,} \\ \left|\sigma_{\mathfrak{p}}(\alpha)^2\right| & \mathfrak{p} \text{ is complex.} \end{cases}$$

Now suppose $\mathfrak{p}$ is a finite place. View $\mathfrak{p}$ as a prime ideal of $\mathscr{O}_K$, and let p be the characteristic of the residue field $\mathscr{O}_K/\mathfrak{p}$. Let $\operatorname{ord}_{\mathfrak{p}}$ denote the ordinal function for $\mathfrak{p}$. On $\mathscr{O}_K^{\times}$ this is defined by

$$\operatorname{ord}_{\mathfrak{p}}(\beta) = m \quad \text{if } \beta \in \mathfrak{p}^m - \mathfrak{p}^{m+1},$$

and it extends to $K^{\times}$ in the obvious way. We let $|\cdot|_p$ denote the usual absolute value of the p-adic field $\mathbb{Q}_p$. The absolute value associated to $\mathfrak{p}$ on K is

$$|\alpha|_{\mathfrak{p}} := p^{-f_{\mathfrak{p}} \operatorname{ord}_{\mathfrak{p}}(\alpha)}.$$

Let $K_{\mathfrak{p}}$ be the $\mathfrak{p}$-adic completion of K with respect to $|\cdot|_{\mathfrak{p}}$; we also use $|\cdot|_{\mathfrak{p}}$ for the absolute value on $K_{\mathfrak{p}}$.

We fix once and for all an algebraic closure $\bar{\mathbb{Q}}_p$ of $\mathbb{Q}_p$, and let $\mathbb{C}_p$ denote the completion of $\bar{\mathbb{Q}}_p$. We use $|\cdot|_p$ to denote the natural extension of $|\cdot|_p$ to all of $\mathbb{C}_p$. We define ord_p on $\mathbb{C}_p^{\times}$ to satisfy

$$|\alpha|_p = p^{-\operatorname{ord}_p \alpha}, \qquad \alpha \in \mathbb{C}_p^{\times}.$$

As $p\mathscr{O}_K$ may split into several prime ideals, the absolute value $|\cdot|_p$ on $\mathbb{Q}$ may have several inequivalent extensions to K, of which $|\cdot|_{\mathfrak{p}}$ is just one; so we must take care when viewing $K_{\mathfrak{p}}$ as a subfield of $\bar{\mathbb{Q}}_p$.

For any embedding $\vartheta\colon K \to \bar{\mathbb{Q}}_p$, we obtain a subfield of $\bar{\mathbb{Q}}_p$ as the composite $\vartheta(K)\cdot\mathbb{Q}_p$. By the Prolongation Theorem [21, §18.5], there exists a choice of ϑ such that $(K_{\mathfrak{p}}, |\cdot|_{\mathfrak{p}})$ is value-isomorphic to $(\vartheta(K)\mathbb{Q}_p, |\cdot|_p)$. Henceforth, we always use this isomorphism to view $K_{\mathfrak{p}}$ as a subfield of $\bar{\mathbb{Q}}_p$. As the isomorphism respects the valuations, we know ord_p and $\operatorname{ord}_{\mathfrak{p}}$ satisfy

$$\operatorname{ord}_{\mathfrak{p}} \beta = e_{\mathfrak{p}} \operatorname{ord}_p \beta, \qquad \beta \in K_{\mathfrak{p}}. \tag{2}$$

2.3 Height Functions

Suppose $n \geq 1$. We let h denote the standard logarithmic Weil height on $\mathbb{P}^n(K)$. This is defined as follows: for any $\mathbf{x} = (x_0 : \cdots : x_n) \in \mathbb{P}^n(K)$,

$$h(\mathbf{x}) = \frac{1}{d_K} \sum_{\mathfrak{p}} \log \left(\max_j \{|x_j|_{\mathfrak{p}}\}\right),$$

where the sum runs over all places of K. It is a consequence of the product formula [21, Ch. 20, pgs. 326–327] that $h(\mathbf{x})$ is independent of the choice of coordinates for $\mathbf{x}$. For any $\alpha \in K$, set $h(\alpha) = h((1 : \alpha))$. Note that this height is *absolute* in the sense that it is not dependent on which field extension K containing the coordinates of $\mathbf{x}$ is considered.

We introduce a modified version of this height function, used in Sect. 3. Suppose $\alpha_1, \ldots, \alpha_n \in K$, and let $K' = \mathbb{Q}(\alpha_1, \ldots, \alpha_n) \subseteq K$. For any nonzero element $\beta \in K'$, we define the function h' by

$$h'(\beta) = \frac{1}{d_{K'}} \max \left\{d_{K'} \cdot h(\beta), |\log \beta|, 1\right\}.$$

The definition of another height function, $h_{\mathfrak{p}}$, is slightly more technical and will be introduced when needed in Sect. 3.

2.4 p-Adic Logarithms

Inside $\mathbb{C}_p$, consider the open disk

$$\Delta_1 := \{z \in \mathbb{C}_p : |z - 1|_p < 1\}.$$

On Δ_1, we define the p-adic logarithm by the series

$$\log_p z = -\sum_{n \geq 1} \frac{(1 - z)^n}{n}. \tag{3}$$

The series is convergent on Δ_1; moreover, on Δ_1 it satisfies the identity

$$\log_p(xy) = \log_p x + \log_p y. \tag{4}$$

If $|z|_p < p^{-\frac{1}{p-1}}$ we have

$$\operatorname{ord}_p \left(\log_p(1 + z)\right) = \operatorname{ord}_p z. \tag{5}$$

Based on an idea due to Iwasawa, the p-adic logarithm can be extended to any $z \in \mathbb{C}_p$ such that $|z|_p = 1$; this extension continues to satisfy (4) (see [36, II.2.4]).

2.5 Solutions to the S-Unit Equation

We let $A_{K,S}$ denote the additive $\mathbb{Z}$-module $(\mathbb{Z}/w\mathbb{Z}) \times \mathbb{Z}^t$. This is isomorphic to $\mathscr{O}_{K,S}^{\times}$, and the list of generators $\boldsymbol{\rho}$ determines an isomorphism

$$\Phi_{\boldsymbol{\rho}} \colon A_{K,S} \longrightarrow \mathscr{O}_{K,S}^{\times}, \qquad \mathbf{a} := (a_0, a_1, \ldots, a_t) \mapsto \prod_{i=0}^{t} \rho_i^{a_i}.$$

We use the shorthand $\boldsymbol{\rho}^{\mathbf{a}} := \Phi_{\boldsymbol{\rho}}(\mathbf{a})$. For obvious reasons, we call the elements of $A_{K,S}$ *exponent vectors*. Much of our discussion will focus on bounds for the entries of an exponent vector. For $\mathbf{a} \in A_{K,S}$, we use the notation $|\mathbf{a}| \leq B$ to signify

$$\max_{0 < i \leq t} |a_i| \leq B.$$

Within $\mathscr{O}_{K,S}^{\times}$, we wish to determine

$$X_{K,S} := \{\tau \in \mathscr{O}_{K,S}^{\times} : 1 - \tau \in \mathscr{O}_{K,S}^{\times}\}.$$

Solving the S-unit equation is equivalent to determining the set $X_{K,S}$. We let $E_{K,S}$ denote the corresponding subset $\Phi_{\boldsymbol{\rho}}^{-1}(X_{K,S})$ of $A_{K,S}$.

3 The Bounds of Baker-Wüstholz and Yu

Suppose $\tau_1, \tau_2 \in \mathscr{O}_{K,S}^{\times}$ provide a solution to the S-unit equation, so that $\tau_1 + \tau_2 = 1$. With respect to the ordered generating set $\boldsymbol{\rho}$, there are unique vectors $\mathbf{b}_i = (b_{i,0}, \ldots, b_{i,t}) \in A_{K,S}$ such that

$$\tau_i = \boldsymbol{\rho}^{\mathbf{b}_i} = \prod_{j=0}^{t} \rho_j^{b_{i,j}}, \qquad i = 1, 2. \tag{6}$$

The techniques of lattice reduction discussed in Sect. 5 will not produce an absolute bound for $|b_{i,j}|$ on their own; they can only be used to improve a known bound. So in this section, we recall bounds established by Baker-Wüstholz [3] and Kunrui Yu [47]. An excellent treatment of the background material appears in [15].

3.1 Statement of Yu's Bound

Let $\mathfrak{p}$ be a finite place of K, and let p denote the rational prime below $\mathfrak{p}$. We let q be the smallest rational prime distinct from p (so $q = 2$ unless $p = 2$, in which case $q = 3$). Let $\zeta_m := \exp(2\pi i/m)$. We say K *satisfies Yu's auxiliary condition* if any of the following hold:

1. $q = 2$ and $p^{f_\mathfrak{p}} \equiv 1 \bmod 4$,
2. $q = 2$ and $\zeta_4 \in K$,
3. $q = 3$ and $\zeta_3 \in K$.

At the end of this section, we explain how the algorithm finds a bound in cases where K does not satisfy Yu's auxiliary condition.

Theorem 3.1 (Yu, [47, pg. 190]) *Suppose $n \geq 1$ and $\mathfrak{p}$ is a prime of $\mathcal{O}_K$. Suppose K is a number field satisfying Yu's auxiliary condition and $\mu_0, \mu_1, \ldots, \mu_{n-1} \in K^\times$ are chosen which satisfy*

$$\operatorname{ord}_\mathfrak{p} \mu_j = 0, \qquad 0 \leq j \leq n-1. \tag{7}$$

Suppose $b_j \in \mathbb{Z}$ and $\Theta := \prod_{j=0}^{n-1} \mu_j^{b_j} \neq 1$. Finally, suppose B satisfies

$$B \geq \max\{|b_0|, \ldots, |b_{n-1}|, 3\}.$$

Then there exist explicit constants C_1^* and Ω, given below, such that

$$\operatorname{ord}_\mathfrak{p} (\Theta - 1) < C_1^* \Omega \log B.$$

3.2 The Constants Ω and Ω'

We first discuss the constant Ω, and the variant Ω' used in the algorithm. In Theorem 3.1, Ω is roughly a product of the logarithmic heights of the μ_j. More precisely, decompose the set $\{\mu_j\}_j$ into a disjoint union $\mathfrak{a} \cup \mathfrak{b}$, where $\mathfrak{a}$ is a maximal subset of $\{\mu_j\}_j$ which is multiplicatively independent. Such a decomposition need not be unique. Because of the possible dependence among the μ_j, Yu requires a modified height function:

$$h_\mathfrak{p}(\mu) := \max\left\{h(\mu), \frac{f_\mathfrak{p}}{\kappa_1(n+4)d_K}\right\}.$$

(The value κ_1 is explained in the following subsection.) The constant Ω, which depends on n, d_K, $\mathfrak{p}$ as well as the μ_j, is then defined by

$$\Omega(n, d_K, \mathfrak{p}) := \prod_{\mu \in \mathfrak{a}} h(\mu) \cdot \prod_{\mu \in \mathfrak{b}} h_{\mathfrak{p}}(\mu).$$

As shown in [47], one may choose any maximal independent set $\mathfrak{a}$ for the computation of Ω. If optimization of the bound is critical, one may search over all possible $\mathfrak{a}$ and take the smallest possible bound. This observation is moot in our use, however.

Corollary 3.2 *Keeping the hypotheses of the previous theorem, suppose also that* $\mu_1, \ldots, \mu_{n-1}$ *are multiplicatively independent. Set*

$$\Omega'(n, d_K, \mathfrak{p}) := h_{\mathfrak{p}}(\mu_0) \prod_{j=1}^{n-1} h(\mu_j).$$

Then $\operatorname{ord}_{\mathfrak{p}}(\Theta - 1) < C_1^* \Omega' \log B$.

Proof In this case, $\mathfrak{b}$ is unique; either $\mathfrak{b} = \{\mu_0\}$ or $\mathfrak{b} = \varnothing$. In either case, $\Omega \leq \Omega'$ and the result follows immediately. □

In the algorithm, we are always in the situation of the Corollary. Rather than decide the question of independence between μ_0 and the other μ_j, we just use the constant Ω'.

3.3 The Constant C_1^*

The value of $C_1^* := C_1^*(n, d_K, \mathfrak{p})$ is dependent on n, d_K, and $\mathfrak{p}$, as follows. Let $u := \operatorname{ord}_q w$, so that q^u is the q-part of w. Set

$$k_2 := c^{(1)} a^{(1)} \cdot \frac{n^n \cdot (n+1)^{n+1}}{n!},$$

$$k_3 := \frac{p^{f_{\mathfrak{p}}}}{q^u} \left(\frac{d_K}{f_{\mathfrak{p}} \log p} \right)^{n+2} \cdot \log \max\{d_K, e\},$$

$$k_4 := \max \left\{ \log \left(e^4 (n+1) d_K \right), e_{\mathfrak{p}}, f_{\mathfrak{p}} \log p \right\}.$$

Here, e denotes the base of the natural logarithm. The constants $a^{(1)}$, κ_1, and $c^{(1)}$ are given in Tables 1 and 2. Finally,

$$C_1^*(n, d_K, \mathfrak{p}) := (n+1) k_2 k_3 k_4. \tag{8}$$

Table 1 The constants $a^{(1)}$ and κ_1

Case	$a^{(1)}$	κ_1
$p = 2$	32	40
$p = 3$	16	20
$p > 3$ and $e_\mathfrak{p} \geq 2$	16	20
$p > 3$ and $e_\mathfrak{p} = 1$	$\frac{8(p-1)}{p-2}$	10

Table 2 The constant $c^{(1)}$

$p \leq 5$		$p > 5$	
Case	$c^{(1)}$	Case	$c^{(1)}$
$p = 2$	160	$p \equiv 1\ (4)$ and $e_\mathfrak{p} = 1$	1473
$p = 3$ and $d_K = 1$	537	$p \equiv 1\ (4)$ and $e_\mathfrak{p} \geq 2$	1502
$p = 3$ and $d_K \geq 2$	759	$p \equiv 3\ (4)$, $e_\mathfrak{p} = 1$, $d_K = 1$	1288
$p = 5$ and $e_\mathfrak{p} = 1$	1473	$p \equiv 3\ (4)$, $e_\mathfrak{p} = 1$, $d_K \geq 2$	1282
$p = 5$ and $e_\mathfrak{p} \geq 2$	319	$p \equiv 3\ (4)$, $e_\mathfrak{p} \geq 2$	2190

3.4 A Remark About Implementation

For this subsection only, suppose all hypotheses in Theorem 3.1 are satisfied, except K does **not** satisfy Yu's auxiliary condition. Set

$$K' := \begin{cases} K(\zeta_4) & q = 2, \\ K(\zeta_3) & q = 3. \end{cases}$$

Let $\mathfrak{P}$ be a prime of $\mathscr{O}_{K'}$ above $\mathfrak{p}$. Let $e_{\mathfrak{P}|\mathfrak{p}}$ be the ramification index of $\mathfrak{P}$ over $\mathfrak{p}$. Because

$$\operatorname{ord}_\mathfrak{p} \alpha = e_{\mathfrak{P}|\mathfrak{p}} \operatorname{ord}_\mathfrak{P} \alpha, \qquad \alpha \in K^\times,$$

we see $\operatorname{ord}_\mathfrak{P} \mu_j = 0$ for all j. Now Theorem 3.1 applies with K' and $\mathfrak{P}$ in place of K and $\mathfrak{p}$, respectively.

Corollary 3.3 *Under the conditions of this subsection,*

$$\operatorname{ord}_\mathfrak{p} (\Theta - 1) < e_{\mathfrak{P}|\mathfrak{p}} \cdot C_1^*(n, d_{K'}, \mathfrak{P}) \cdot \Omega'(n, d_{K'}, \mathfrak{P}) \cdot \log B.$$

Note that even if $\mathfrak{p}$ splits as $\mathfrak{P}\mathfrak{P}'$ in K', the choice of $\mathfrak{P}$ is irrelevant; both give the exact same bound in the Corollary.

3.5 Bound of Baker-Wüstholz

We now give an effective version of Baker's theorem. (Notations are as in Sects. 2.2 and 2.3.)

Theorem 3.4 (Baker-Wüstholz, [3, pg. 20]) *Let L be a linear form in $t+1$ indeterminates,*

$$L(z_0, \ldots, z_t) = b_0 z_0 + \cdots + b_t z_t, \qquad b_i \in \mathbb{Z}.$$

Let $B = \max\{|b_0|, \ldots, |b_t|\}$, and let $\rho_0, \ldots, \rho_t \in \overline{\mathbb{Q}} - \{0, 1\}$. Let K' be the subfield of $\overline{\mathbb{Q}}$ generated by the ρ_i. If $B > 3$ and

$$\Lambda = L(\log \rho_0, \log \rho_1, \ldots, \log \rho_t) \neq 0,$$

then

$$\log |\Lambda| > -C(t, d_{K'}) \log(B) \prod_{j=0}^{t} h'(\rho_j),$$

where the constant $C(t, d_{K'})$ is defined by

$$C(t, d_{K'}) = 18(t+2)!(t+1)^{(t+2)}(32 d_{K'})^{(t+3)} \log\left(2(t+1)d_{K'}\right).$$

Note that we may be sure $\Lambda \neq 0$ if the set $\{\log \rho_i\}$ is linearly independent over $\mathbb{Q}$.

3.6 Obtaining the Initial Bound

The theorems of Baker-Wüstholz and Yu both provide inequalities of the form "a polynomial function of B is bounded by a polynomial function of $\log(B)$," which in turn guarantee an absolute bound on B. The analysis to determine such a bound explicitly is standard; we will use the following result of Pethő and de Weger for this purpose.

Lemma 3.5 (Pethő and de Weger [29, Lemma 2.2]) *Suppose the real numbers a, b, h satisfy $a \geq 0$, $h \geq 1$, $b > \left(\frac{e^2}{h}\right)^h$, and let $x \in \mathbb{R}$ be the largest solution to the equation*

$$x = a + b(\log x)^h.$$

Then

$$x < 2^h \left(a^{\frac{1}{h}} + b^{\frac{1}{h}} \log\left(h^h b\right)\right)^h.$$

4 Initial Exponent Bounds

4.1 An Upper Bound at the Extremal Place

Suppose (τ_1, τ_2) is a solution to the S-unit equation, with τ_i specified as in (6). We set $B = \max_{i,j} \left|b_{i,j}\right|$, and assume $B \geq \max\{4, w\}$. Relabeling τ_1 and τ_2 if necessary, we assume $B = |b_{1,j}|$ for some $1 \leq j \leq t$. Recall that S contains precisely $t+1$ places, $\mathfrak{p}_1, \ldots, \mathfrak{p}_{t+1}$. We choose the indices $k, \ell \in \{1, 2, \ldots, t+1\}$ so that

$$\left|\log |\tau_1|_{\mathfrak{p}_k}\right| = \max_{\mathfrak{p} \in S} \left|\log |\tau_1|_{\mathfrak{p}}\right|, \qquad |\tau_1|_{\mathfrak{p}_\ell} = \min_{\mathfrak{p} \in S} |\tau_1|_{\mathfrak{p}}.$$

Remark 4.1 In the sequel, we number our constants in an effort to stay consistent with the enumeration given in Smart's paper [34]. There, Smart considers a more general unit equation, and so introduces certain constants $c_4(i)$, $c_6(i)$, $c_7(i)$, ... whose values are trivial in the present application. So while the alert reader may notice gaps in the enumeration of constants, this is intentional. (Adjusting our implementation to the more general setting is not difficult, but we are satisfied to limit the discussion to match the current state of the implementation.)

For any choice of $U := \{\mathfrak{u}_1, \ldots, \mathfrak{u}_t\} \subseteq S$ define the $t \times t$ matrix

$$M = (m_{i,j}), \qquad m_{i,j} = \log |\rho_j|_{\mathfrak{u}_i}.$$

One may always choose U so that M is invertible (see [15, §5.1]), and so we assume this is the case. We have

$$\begin{pmatrix} b_{1,1} \\ b_{1,2} \\ \vdots \\ b_{1,t} \end{pmatrix} = M^{-1} \begin{pmatrix} \log |\tau_1|_{\mathfrak{u}_1} \\ \log |\tau_1|_{\mathfrak{u}_2} \\ \vdots \\ \log |\tau_1|_{\mathfrak{u}_t} \end{pmatrix}.$$

Let $\|M\|$ be the row norm of M^{-1}, i.e. $\|M\| = \max_i \sum_{j=1}^{t} |m_{i,j}|$, and set

$$c_1 := \max\left\{1, \max_{U \subseteq S} \{||M|| : M \text{ is invertible }\}\right\}.$$

Note that this differs slightly from Smart's definition, to ensure that $c_1 \geq 1$. Then $B \leq c_1 \left|\log |\tau_1|_{\mathfrak{p}_k}\right|$. We define

$$c_2 := \frac{1}{c_1} \qquad c_3 := \frac{0.9999999 c_2}{r+s}.$$

By Smart [34, Lemma 2], we have

$$|\tau_1|_{\mathfrak{p}_\ell} \leq e^{-c_3 B}. \tag{9}$$

We now have an upper bound on $|\tau_1|_{\mathfrak{p}_\ell}$ in terms of B. We next establish a lower bound, also involving B, which will force a limit on the size of B. The precise argument depends on whether $\mathfrak{p}_\ell$ is a finite or infinite place. For the purposes of the algorithm, we must compute this bound on B for each possible index $1 \leq \ell \leq t$; we have no choice but to take the largest possible bound, i.e., the larger of the two values K_0 and K_1 determined in the remainder of this section.

4.2 Case I: $\mathfrak{p}_\ell$ Is Finite

If $\mathfrak{p}_\ell$ is finite, then let $\mathfrak{p}_\ell$ also denote the associated prime ideal in $\mathcal{O}_K$. Let p be the prime of $\mathbb{Z}$ lying below $\mathfrak{p}_\ell$, and let e_ℓ and f_ℓ denote the ramification index and inertial degree of $\mathfrak{p}_\ell$ over p, respectively. From (9) we have

$$N_{K/\mathbb{Q}}(\mathfrak{p}_\ell)^{-\operatorname{ord}_{\mathfrak{p}_\ell}(\tau_1)} \leq e^{-c_3 B}. \tag{10}$$

Setting

$$c_5(\ell) := \frac{c_3}{e_\ell \log N_{K/\mathbb{Q}}(\mathfrak{p}_\ell)},$$

the inequality (10) yields

$$\operatorname{ord}_{\mathfrak{p}_\ell} \tau_1 \geq \frac{c_3 B}{\log N_{K/\mathbb{Q}}(\mathfrak{p}_\ell)} = e_\ell c_5(\ell) B > 0, \tag{11}$$

and so $\operatorname{ord}_{\mathfrak{p}_\ell} \tau_2 = 0$. We would like to apply Yu's Theorem to $\Theta = \tau_2$, but unfortunately the generators ρ_i may have nonzero order with respect to $\mathfrak{p}_\ell$. So we now replace the ρ_i with a different set of generators, as in [34, pp. 824–825]. First, set $n_i := \operatorname{ord}_{\mathfrak{p}_\ell} \rho_i$. Necessarily, there exist indices i for which $n_i \neq 0$. Choose i_0 so that

$$|n_{i_0}| = \min\{|n_i| : n_i \neq 0\},$$

and now relabel so that $i_0 = t$. For $1 \leq i \leq t-1$, define

$$\mu_i = \rho_i^{n_t} \rho_t^{-n_i},$$

so that $\operatorname{ord}_{\mathfrak{p}_\ell} \mu_i = 0$. Next, for each i with $1 \leq i \leq t-1$, choose integers d_i, r_i such that

$$0 \leq r_i < |n_t| \qquad \text{and} \qquad b_{2,i} = n_t d_i + r_i.$$

Necessarily, $|d_i| \le B$. Set $N := \sum_{i=1}^{t-1} n_i r_i$.

Lemma 4.1 *We have* $N \equiv 0 \pmod{n_t}$.

Proof Since $\mathrm{ord}_{\mathfrak{p}_\ell}(\tau_2 \rho_0^{-b_{2,0}}) = 0$, we know $\sum_{i=1}^{t} n_i b_{2,i} = 0$. Thus,

$$n_t b_{2,t} = -\sum_{i=1}^{t-1} n_i b_{2,i} = -n_t \sum_{i=1}^{t-1} n_i d_i - \sum_{i=1}^{t-1} n_i r_i,$$

proving the claim. □

Setting $N_0 = \frac{N}{n_t}$ and $\mu_0 := \rho_0^{b_{2,0}} \rho_t^{-N_0} \cdot \prod_{i=1}^{t-1} \rho_i^{r_i}$, we have arranged that

$$\tau_2 = \mu_0 \prod_{i=1}^{t-1} \mu_i^{d_i}, \quad |d_i| \le B, \quad \mathrm{ord}_{\mathfrak{p}_\ell} \mu_i = 0. \tag{12}$$

Since $0 \le b_{2,0} < w$ and $0 \le r_i < |n_t|$, there are only finitely many possible values for μ_0, and this finite set can be determined without any knowledge of B or the $b_{2,i}$. For each μ_0, we may apply Corollary 3.2 or 3.3 as appropriate, and obtain a constant $c_8'(\ell, \mu_0)$ such that

$$\mathrm{ord}_{\mathfrak{p}_\ell} \tau_1 = \mathrm{ord}_{\mathfrak{p}_\ell}(\tau_2 - 1) < c_8'(\ell, \mu_0) \log B.$$

Setting

$$c_8(\ell) := \max\left\{ \frac{e^2}{\log 2}, \max_{\mu_0}\{c_8'(\ell, \mu_0)\} \right\},$$

we may be sure every S-unit solution satisfies

$$\mathrm{ord}_{\mathfrak{p}_\ell} \tau_1 < c_8(\ell) \log B. \tag{13}$$

Combining inequalities (11) and (13), we have

$$B < \frac{c_8(\ell)}{e_\ell c_5(\ell)} \log B.$$

Since $c_1 \ge 1$ and $c_8(\ell) \ge e^2 (\log 2)^{-1}$, it follows that

$$\frac{c_8(\ell)}{e_\ell c_5} \ge e^2.$$

Applying Lemma 3.5 with $a = 0$, $b = c_8(\ell)/e_\ell c_5(\ell)$, and $h = 1$, we may conclude

$$B \le K_0(\ell) := \frac{2c_8(\ell)}{e_\ell c_5(\ell)} \log\left(\frac{c_8(\ell)}{e_\ell c_5}\right).$$

Set

$$K_0 := \max\{K_0(\ell) : \mathfrak{p}_\ell \in S_{\text{fin}}\}.$$

If ℓ corresponds to a finite place, then $B \le K_0$.

In our implementation, the functions `mus` and `possible_mu0s` are used to recover the μ_i for each finite place $\mathfrak{p}_\ell$. The constants $c_8(\ell)$ determined from Yu's Theorem are computed in `Yu_bound`, while the constant K_0, which may be of independent interest, is computed by `K0_func`.

4.3 Case II: $\mathfrak{p}_\ell$ Is Infinite

We now assume $\mathfrak{p}_\ell$ is infinite. As in §2.3, we let $\sigma_{\mathfrak{p}_\ell}$ denote the embedding of K into $\mathbb{C}$ such that

$$|\alpha|_{\mathfrak{p}_\ell} = \left|\sigma_{\mathfrak{p}_\ell}(\alpha)\right|^{\delta(\ell)}, \quad \text{where } \delta(\ell) = \begin{cases} 1 & \mathfrak{p}_\ell \text{ is real,} \\ 2 & \mathfrak{p}_\ell \text{ is complex.} \end{cases}$$

We let $\alpha^{(\ell)}$ denote $\sigma_{\mathfrak{p}_\ell}(\alpha)$ for any $\alpha \in K$, and we define

$$c_{11}(\ell) := \frac{\delta(\ell)\log 4}{c_3}, \qquad c_{13}(\ell) := \frac{c_3}{\delta(\ell)}.$$

The condition (9) can now be expressed as

$$\left|\tau_1^{(\ell)}\right| \le e^{-c_{13}(\ell)B}.$$

The choices of $c_{11}(\ell)$ and $c_{13}(\ell)$ guarantee that

$$B \ge c_{11}(\ell) \quad \Longrightarrow \quad \left|\tau_1^{(\ell)}\right| \le \frac{1}{4}.$$

Set $\Lambda := \log \tau_2^{(\ell)}$. The estimate $|\log z| \le 2|z-1|$ holds for $|z-1| \le \frac{1}{4}$, and so

$$|\Lambda| \le 2\left|\tau_2^{(\ell)} - 1\right| = 2\left|\tau_1^{(\ell)}\right| \le 2e^{-c_{13}(\ell)B}. \tag{14}$$

The next step is to view Λ as a linear form in logarithms and apply the theorem of Baker and Wüstholz. Set $\zeta := \exp\frac{2\pi\sqrt{-1}}{w} \in \mathbb{C}$. Since ρ_0 is a wth root of unity, there exists $0 \le k < w$ such that $(\rho_0^{(\ell)})^{b_{2,0}} = \zeta^k$. By (6), we have

$$\begin{aligned}\Lambda &= \log\left(\left(\rho_0^{(\ell)}\right)^{b_{2,0}} \cdot \prod_{j=1}^{t}\left(\rho_j^{(\ell)}\right)^{b_{2,j}}\right)\\ &= \log\zeta^k + \sum_{j=1}^{t} b_{2,j}\log\rho_j^{(\ell)} + A\cdot 2\pi\sqrt{-1}\\ &= k\log\zeta + \sum_{j=1}^{t} b_{2,j}\log\rho_j^{(\ell)} + Aw\log\zeta\\ &= (Aw+k)\log\zeta + \sum_{j=1}^{t} b_{2,j}\log\rho_j^{(\ell)},\end{aligned} \tag{15}$$

where we have introduced $A \in \mathbb{Z}$ to adjust for the principal branch of the logarithm. Certainly $|A| \le tB$, and so $|Aw+k| \le (t+1)Bw$. Set

$$b'_{2,j} := \begin{cases} Aw+k & j=0\\ b_{2,j} & j>0\end{cases}$$

and $L'(z_0,\dots,z_t) := \sum_{j=0}^{t} b'_{2,j}z_j$. We now have

$$|\Lambda| = \left|L'(\log\zeta, \log\rho_1^{(\ell)},\dots,\log\rho_t^{(\ell)})\right|.$$

Taking $K' = \mathbb{Q}(\rho_0,\dots,\rho_t) \cong \mathbb{Q}(\zeta,\rho_1^{(\ell)},\dots,\rho_t^{(\ell)})$, we define

$$c_{14}(\ell) := C(t, d_{K'})\prod_{j=0}^{t} h'(\rho_j).$$

(Recall that $C(t,d_{K'})$ is defined in Theorem 3.4.) We have $|b'_{2,j}| \le B' := (t+1)Bw$. Applying Theorem 3.4 to Λ, we obtain

$$\log|\Lambda| > -c_{14}(\ell)\cdot\log B' = -c_{14}(\ell)\log\big((t+1)wB\big).$$

Combining this inequality with (14), we obtain

$$2e^{-c_{13}(\ell)B} \ge |\Lambda| \ge e^{-c_{14}(\ell)\log B'}. \tag{16}$$

This yields the inequality

$$B < a(\ell) + b(\ell) \log B,$$

where

$$a(\ell) := \frac{1}{c_{13}(\ell)} \left(\log 2 + c_{14}(\ell) \log\big((t+1)w\big)\right), \quad b(\ell) := \frac{c_{14}(\ell)}{c_{13}(\ell)}.$$

As $c_{13}(\ell) \leq \frac{1}{t}$ and $c_{14}(\ell) \geq 32^3$, we have $a(\ell) \geq 0$ and $b(\ell) \geq e^2$. So by Lemma 3.5, $B < c_{15}(\ell)$ (provided $B \geq c_{11}(\ell)$), where

$$c_{15}(\ell) := 2\big(a(\ell) + b(\ell) \log b(\ell)\big).$$

Thus, setting

$$K_1(\ell) := \max\{c_{11}(\ell), c_{15}(\ell)\},$$
$$K_1 := \max\{K_1(\ell) : \mathfrak{p}_\ell \text{ is infinite}\},$$

we may be sure $B \leq K_1$. In our implementation, the constant K_1 is computed in the function `K1_func`.

Combining all the results of this section, we obtain the following.

Lemma 4.2 *The constant B satisfies $B \leq \max\{4, w, K_0, K_1\}$.*

5 LLL Reduction

In this section we explain how we can reduce the upper bound we have computed in Sect. 4. This is necessary, because in practice the size of the initial bound is extremely large and cannot be used for practical computations. The idea of the method we will present here has its origin in de Weger's thesis [13, 12, 14] where he develops a method based on multi-dimensional approximation lattices of linear forms of p-adic numbers to solve (among many other equations) S-unit equations[2] over $\mathbb{Q}$. These ideas of de Weger have been extended by himself and others to apply over any number field K, and have also been used for the solution of other exponential Diophantine equations [39, 40, 41, 34].

In the reduction step we use the LLL reduction algorithm on lattices generated by integer matrices. So instead of the classical LLL algorithm [24], we use the algorithm in [12]. If $\mathscr{L}$ is a lattice in $\mathbb{R}^n$, let $\mathscr{L}^* = \mathscr{L} - \{\mathbf{0}\}$. For $\mathbf{y} \in \mathbb{R}^n$, we define

[2] It is worth mentioning the recent results of von Känel and Matschke [42], who solve S-unit equations using modularity.

$$\ell(\mathscr{L}, \mathbf{y}) = \begin{cases} \min\limits_{\mathbf{x} \in \mathscr{L}^*} \|\mathbf{x}\|, & \text{if } \mathbf{y} \in \mathscr{L}, \\ \min\limits_{\mathbf{x} \in \mathscr{L}} \|\mathbf{x} - \mathbf{y}\|, & \text{otherwise.} \end{cases}$$

Computing the exact value of $\ell(\mathscr{L}, \mathbf{y})$ is a very challenging problem in general. Instead, the function `minimal_vector` computes a lower bound using standard properties of a reduced basis of a lattice and the LLL algorithm (see [36, Chapter V]). As in the previous section, we follow Smart's notation in [34]. Most of the material we present in this section can also be found in [36, 15].

We preserve the meaning of $\mathfrak{p}_\ell$ from Sect. 4. When $\mathfrak{p}_\ell$ is a finite place, we let p denote the prime of $\mathbb{Z}$ lying below $\mathfrak{p}_\ell$. We continue to assume $B \geq \max\{4, w\}$ in this section.

5.1 Finite Places

Suppose $\mathfrak{p}_\ell$ is a finite place. Set

$$c_{16}(\ell) := 1 + \frac{1}{c_5(\ell)},$$

and suppose that $B \geq c_{16}(\ell)$. Define $\Delta_2 \in K_{\mathfrak{p}_\ell}$ as $\Delta_2 := \log_p \tau_2$. Combining (11), (2), and $B \geq c_{16}(\ell)$, shows that $\operatorname{ord}_p \tau_1 > 1$. Thus $|\tau_1|_p < p^{-\frac{1}{p-1}}$, and by (5),

$$\operatorname{ord}_p \Delta_2 = \operatorname{ord}_p \log_p \tau_2 = \operatorname{ord}_p \log_p (1 - \tau_1) = \operatorname{ord}_p \tau_1 > 1.$$

Let μ_i, d_i be as given in (12), so that we have

$$\Delta_2 = \log_p \tau_2 = \log_p \mu_0 + \sum_{i=1}^{t-1} d_i \log_p \mu_i.$$

Choose $\theta \in K_{\mathfrak{p}_\ell}$ such that $K_{\mathfrak{p}_\ell} = \mathbb{Q}_p(\theta)$, and let $\operatorname{Disc}(\theta)$ denote the discriminant of θ. Set $D_p(\theta) = \operatorname{ord}_p \operatorname{Disc}(\theta)$ and $n = [K_{\mathfrak{p}_\ell} : \mathbb{Q}_p]$, so that $n = e_\ell f_\ell$. Expressing Δ_2 with respect to the power basis, we obtain $\Delta_{2,k} \in \mathbb{Q}_p$ such that $\Delta_2 = \sum_{k=0}^{n-1} \Delta_{2,k} \theta^k$. Further, we may express

$$\Delta_{2,k} = a_{0,k} + \sum_{j=1}^{t-1} d_j a_{j,k}, \qquad a_{j,k} \in \mathbb{Q}_p,\ 0 \leq k \leq n-1. \tag{17}$$

Using an idea due to Evertse [41, p. 257], we have

$$\operatorname{ord}_p \Delta_{2,k} \geq c_5(\ell)B - \frac{D_p(\theta)}{2}.$$

Define

$$c_{17}(\ell) := \min\{\operatorname{ord}_p a_{j,k} : 1 \leq j \leq t-1,\ 0 \leq k \leq n-1\},$$
$$c_{18}(\ell) := c_{17}(\ell) + \frac{D_p(\theta)}{2},$$

and choose $\lambda \in \mathbb{Q}_p$ such that $\operatorname{ord}_p \lambda = c_{17}(\ell)$.

Should there be some index k such that $c_{17}(\ell) > \operatorname{ord}_p(a_{0,k})$, then $\operatorname{ord}_p \Delta_{2,k} = \operatorname{ord}_p a_{0,k} < c_{17}(\ell)$, and consequently

$$B < \frac{c_{18}(\ell)}{c_5(\ell)}.$$

For the remainder, then, we assume

$$c_{17}(\ell) \leq \min\{\operatorname{ord}_p a_{0,k} : 0 \leq k \leq n-1\}.$$

By the choice of λ, $\kappa_{j,k} := a_{j,k}/\lambda$ is a p-adic integer for all j, k, and we may rewrite (17) as

$$\frac{\Delta_{2,k}}{\lambda} = \kappa_{0,k} + \sum_{j=1}^{t-1} d_j \kappa_{j,k}, \quad \text{with} \quad \operatorname{ord}_p\left(\frac{\Delta_{2,k}}{\lambda}\right) \geq c_5(\ell)B - c_{18}(\ell).$$

For any $a \in \mathbb{Z}_p$ and a positive integer z, let $a^{(z)}$ denote the unique integer between 0 and p^z such that $a \equiv a^{(z)} \pmod{p^z}$. For a positive integer u, let $\mathscr{L}$ be the lattice generated by the columns of the matrix

$$\begin{pmatrix} 1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & & 1 & 0 & \cdots & 0 \\ \kappa_{1,0}^{(u)} & \cdots & \kappa_{t-1,0}^{(u)} & p^u & & 0 \\ \vdots & & \vdots & & \ddots & \\ \kappa_{1,n-1}^{(u)} & \cdots & \kappa_{t-1,n-1}^{(u)} & 0 & & p^u \end{pmatrix} \in \mathbb{Z}^{(t+n-1)\times(t+n-1)}.$$

Define

$$\mathbf{y} = \begin{pmatrix} 0 & \cdots & 0 & -\kappa_{0,0}^{(u)} & \cdots & -\kappa_{0,n-1}^{(u)} \end{pmatrix}^{\mathsf{T}} \in \mathbb{Z}^{t+n-1}.$$