NNT/NL : 2022AIXM0580/031ED184

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 18 novembre 2022 par

Leonardo COLÒ

Oriented supersingular elliptic curves and class group actions

Discipline
Mathématiques

Spécialité
Théorie des nombres et cryptographie

École doctorale
ED-184 Mathematiques et Informatique

Laboratoire/Partenaires de recherche
Institut de Mathématiques de Marseille (I2M)
Équipe Arithmétique et Théorie de l'Information (ATI)

Composition du jury

Tanja LANGE Technische Universiteit Eindhoven	Rapporteuse
Felipe VOLOCH University of Canterbury	Rapporteur
Jean-Marc COUVEIGNES Institut de mathématiques de Bordeaux	Président du jury
Samuele ANNI Université d'Aix-Marseille	Examinateur
Luca DE FEO IBM Research Zürich	Examinateur
Elisa LORENZO-GARCÍA Université de Neuchâtel	Examinatrice
Christophe RITZENTHALER Université de Rennes 1	Examinateur
David KOHEL Université d'Aix-Marseille	Directeur de thèse

Affidavit

I, undersigned, Leonardo Colò, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of David Kohel, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the French national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Marseille, 11 September 2022

Liste des publications réalisées dans le cadre du projet de thèse:

1. Orienting supersingular isogeny graphs, with David Kohel, *Journal of Mathematical Cryptology*, **14**, 414–437, 2020. [ArXiv, ePrint, HAL, DOI]

2. On the modular OSIDH protocol, with David Kohel, *preprint*, 2022.

Participation aux conférences et écoles d'été au cours de la période de thèse:

1. Journés Nationales de Calcul Formel 2019 (CIRM, Marseille, France), Feb. 2019.

2. Séminaire CCA - Codage, Cryptologie, Algorithmes (Paris, France), Jun. 2019.

3. Number-Theoretic Methods in Cryptology 2019 (Institut de Mathématiques de Jussieu, Paris, France), Jun. 2019.

4. Séminaire ATI - Arithmétique et Théorie de l'Information (Marseille, France), Sep. 2019.

5. Séminaire de Cryptographie (Rennes, France), Oct. 2019.

6. Séminaire de l'École des mines de Saint-Étienne (Gardanne, France), Dec. 2019.

7. Séminaire de l'équipe IAA de Toulon (Toulon, France), Mar. 2021.

8. AGC^{2}T: Arithmetic, Geometry, Cryptography and Coding Theory (CIRM, Marseille, France), May 2021.

9. Isogeny-based cryptography school (Bristol, UK), Jul. 2021.

Résumé

Cette thèse s'articule autour de divers aspects des courbes elliptiques supersingulières, de leurs anneaux des endomorphismes et des graphes d'isogénies associés. La structure riche de ces graphes en fait un outil important pour aborder plusieurs problèmes en théorie des nombres. Dans ce document, nous étudions différentes facettes et propriétés des graphes d'isogénies supersingulières et nous exploitons leurs nouvelles applications cryptographiques. En particulier, nous décrivons les outils théoriques nécessaires pour orienter les graphes d'isogénies et caractériser leur structure. Nous introduisons une nouvelle catégorie de courbes elliptiques supersingulières orientées, au moyen d'un plongement d'un ordre d'un corps de nombres quadratique imaginaire dans l'anneau d'endomorphismes d'une courbes elliptique supersingulière, et nous obtenons les propriétés du graphe d'isogénie associé. Nous établissons un modèle de calcul qui permet de décrire une action d'un groupe de classes compatible sur ces objets et nous fournissons la description théorique et pratique de la façon dont cette action fonctionne sur les courbes, les chaînes d'isogénies et tout le graphe orienté. Comme application, nous introduisons un protocole d'échange de clés de Diffie et Hellman à base d'isogénies supersingulières orientées (OSIDH), analogue au protocole Diffie et Hellman supersingulier (SIDH) et qui généralise le protocole CSIDH.

En parallèle, nous développons la version modulaire de notre construction. Nous montrons que l'action de groupe peut être effectuée efficacement sur les séquences de points de moduli sur une courbe modulaire uniquement et qu'elle est plus susceptible d'être accéléré en imposant une structure de niveau appropriée. Dans cette direction, nous décrivons une famille de courbes modulaires efficace pour la mise en œuvre de ces idées et nous calculons les propriétés du graphe d'isogénies correspondant, à la fois orienté et non-orienté. Enfin, en utilisant des courbes modulaires avec structure de niveau, nous généralisons l'approche modulaire au protocole OSIDH et nous améliorons sa complexité.

Mots clés: courbes elliptiques supersingulières, anneaux des endomorphismes, graphes d'isogénies, courbes modulaires, actions de groupes de classes, orientations.

Abstract

This thesis revolves around various aspects of supersingular elliptic curves, their endomorphism rings and associated isogeny graphs. The rich structure of these graphs makes them an important implement for approaching several computational problems in number theory. In this document, we study different facets and properties of supersingular isogeny graphs and exploit their potential new cryptographic applications. In particular, we describe the theoretical tools necessary to orient isogeny graphs and characterize their structure. By means of embeddings of quadratic orders in the endomorphism ring of supersingular elliptic curves, we introduce a new category of oriented elliptic curves and derive properties of the associated isogeny graphs. We establish a computational model which permits one to describe a compatible class group action on these objects and provide the theoretical and practical description of how this action works on curves, chains of isogenies and the entire oriented graph. As an application, we introduce an oriented supersingular isogeny Diffie-Hellman protocol (OSIDH), analogous to the supersingular isogeny Diffie-Hellman (SIDH) protocol and generalizing the commutative supersingular isogeny Diffie-Hellman (CSIDH) protocol.

In parallel, we develop an explicit modular version of our construction. We show that the group action can be carried out effectively solely on the sequences of moduli points on a modular curve and is further amenable to speedup by imposing a suitable level structure. In this direction, we describe a suitable family of modular curves and derive the properties of the corresponding covering isogeny graphs, both oriented and non-oriented. Finally, by introducing the use of modular curves of higher level, we expand and improve the complexity of the modular approach to the OSIDH protocol.

Keywords: supersingular elliptic curves, endomorphism rings, isogeny graphs, modular curve, class group actions, orientations.

Acknowledgements

First and foremost, I would like to thank my supervisor, David Kohel. This thesis would not exist without his constant support, help and guidance. I am grateful to him for suggesting interesting research questions, stimulating my curiosity and at the same time giving me all the freedom to explore different topics of interest. My warmest thanks to him for his patience, his time and all his insightful comments and precious advice.

Thanks to Tanja Lange, Felipe Voloch, Samuele Anni, Jean-Marc Couveignes, Luca De Feo, Elisa Lorenzo García and Christophe Ritzenthaler for accepting to be part of my Ph.D. committee and for reading this manuscript and providing valuable comments.

Thanks to the members of the équipe ATI at Marseille University for welcoming me and providing a great working environment. In particular, I am deeply grateful to Samuele Anni for his constant efforts in bringing people together by organizing seminars and working groups.

Thanks to all my colleagues in Marseille for all the great moments shared during these years. A special mention to Elena for being so generous with her time and help when I first arrived and to my office mates Alejandro and Bastien.

Thanks to all the people that I have met along the way at various events, conferences and seminars for great discussions and practical contributions.

Thanks to all my friends, wherever they are, for the moral support despite the distance.

Finally, and most importantly, a huge thank you to my family for the endless love and the constant support in everything I do.

Leonardo Colò
Marseille, October 2022

Contents

List of Figures

List of Tables

List of Algorithms

Introduction

In this thesis we study the theoretical and practical aspects of supersingular isogeny graphs, which constitute the main objects of study; we introduce tools to orient them and exploit the potential of the new augmented graph by developing new cryptographic applications.

Isogeny graphs

Isogeny graphs, whose vertices classify isomorphism classes of elliptic curves and whose edges denote isogenies of a given degree, are combinatorial objects encoding properties of elliptic curves and relations between them. These graphs represent a powerful tool in theoretical and computational number theory due to their remarkable properties. They permit one to study elliptic curves not only by themselves but also as part of a much richer structure where they are put in relation with the others.

Abelian varieties and the isogenies between them have been known and studied since the beginning of the 19th century thanks to the work of Abel and Jacobi. They showed how to associate an abelian variety to any algebraic curve and obtain a group law on their set of points. In case of elliptic curves, this abelian variety — called the Jacobian — is isomorphic to the curve itself and therefore we obtain an additive group directly on its points. Isogenies are maps that preserves this structure. During the last two centuries many important contributions to the theory of abelian varieties and isogenies came from great mathematicians such as Weierstrass, Dedekind, Kronecker, Riemann, Weil, Cartier, Mumford, Serre, Tate and Grothendieck. Although the idea of isogeny graphs was already present, it was not until the development of computer algebra and the consequent rise of interest in effective algorithms and methods that their structure became a central object of investigation.

Isogeny graphs come in two distinct categories: ordinary and supersingular. The latter was first studied by Mestre [Mes] and Pizer [Piz3] who proved important combinatorial and connectivity properties of these graphs. In his thesis, Kohel [Koh1] dealt with the ordinary case; he described the ordinary isogeny graphs and used the information to provide an effective method for computing the endomorphism ring of an ordinary elliptic curve. In the same document, he also introduced graph theoretic algorithms for the supersingular endomorphism ring. Fouquet and Morain [FM] eventually expanded on the work of Kohel and introduced the terminology *volcano* still in use today to describe an ordinary isogeny graph.

At a first glance, the difference between ordinary and supersingular isogeny graphs is that while the former have a well determined shape due to the existence of a commutative action on the set of vertices, the latter normally look less well behaved. However, supersingular graphs are more practical both computationally and in terms of representation on a computer; in fact, supersingular moduli points, endomorphism rings and isogenies can all be defined over $\mathbb{F}_{p^2}$ and the same desirable property holds on modular curves of higher level, see Section 3.3. Further, the lack of a canonical way of providing directions on their vertices and their good mixing properties make supersingular isogeny graphs good candidates for cryptographic applications. One of the contributions of this document is to provide a way of orienting these graphs by recovering a class group action on its vertices. The additional structure blurs the distinction between the category of supersingular and ordinary elliptic curves and their associated isogeny graphs and this permits one to adapt certain CM methods to supersingular curves. For instance, it is worth noting that a similar notion in the equivalent category of maximal quaternion orders can be used to extend p-adic canonical lifts from the ordinary [CH] to the supersingular realm [Bel].

The work of Kohel [Koh1] is an explicit example of how the rich structure of isogeny graphs can provide algorithms and explicit methods to the study of elliptic curves. Other applications include identifying supersingular elliptic curves [Sut4, §3.2], endomorphism ring determination [Koh1; BS1], computing Hilbert class polynomials [Bel+], computing modular polynomials [BLS] and point counting.

Cryptography

Public key encryption schemes are a central idea in modern cryptography. Their introduction resulted in a drastic change of paradigm in the world of information security so to be defined “a revolution in cryptography” [DH]. Contrary to symmetric systems, public key cryptography permits secure communication between parties that have never communicated before making it possible to meet the needs of modern use of the internet. Indeed, all exchanges taking place on the web, from emails to bank transactions to secure navigation, require establishing some sort of shared secret between two (or more) entities over a potentially insecure communication channel in a secure way. Besides that, the continuous and dramatically fast rise of the internet, the ease of the access to it and the rapid emergence of new technologies relying on it poses new problems and demands new cryptographic schemes and protocols such as digital signatures, hash functions, digital identification, multiparty computations etc.

Public key cryptography was first introduced to the public by Whitfield Diffie and Martin Hellman in 1976 [DH]; the idea had been previously conceived by Ellis, Cocks and Williamson at GCHQ, although their work remained classified until 1997. Public key schemes and algorithms are built upon a certain number of theoretical problems that are believed to be computationally hard. Number theory has always been a great source for such problems, the reason being that this branch of mathematics provides optimal examples of trapdoor one-way public functions. These are functions $f : X \to Y$ such that the secret trapdoor permits efficient inversion of the function but the determination of the trapdoor, or more generally of a preimage for the function, is supposed to be computationally difficult, i.e., the computational cost to find x from $y = f(x)$ is assumed to be prohibitively expensive compared to the cost of evaluating $f(x)$.

There are two main classes of hard problems used in classical public key cryptography. The first one consists in the difficulty of factoring integers: RSA, the first public key encryption system proposed in 1977 by Rivest, Shamir and Adleman [RSA] is based on this assumption. The second family includes all instances of the discrete logarithm problem: Diffie-Hellman first key exchange scheme is based on the discrete logarithm problem in finite fields. Building of their idea, ElGamal [ElG] converted it to a public key cryptosystem using the unit group of $\mathbb{Z}/n\mathbb{Z}$; ECDH [Kob; Mill] relies on the discrete logarithm problem taking place in the group of points on an elliptic curve but other less common protocols work in the unit group of number fields or in the set of solution of Pell’s equations.

Public key cryptography relies entirely on computationally hard problems, but the hardness of a mathematical problem depends on the computer model in use. Since their invention, computer science has been studying classical deterministic Turing machines which treats information as sequences of zeroes and ones (classical bits), these are called *classical computers*. However, in 1980, Paul Benioff [Ben] theorized the construction of a quantum version of the Turing machine which, by contrast, elaborate data allowing superposition of different states; we refer to these as *quantum computers*. The importance of this change of paradigm became very clear thanks to the work of Manin [Man1] and Feynman [Fey] who showed that, in certain computations, quantum computers could outperform classical ones. Oddly enough, number theoretical problems represent a field in which quantum computers have shown great relevance. In 1994, Peter Shor [Sho] introduced an algorithm that can solve the integer factorization problem and the discrete logarithm problem on a quantum computer in polynomial time. The work of Shor makes all the cryptographic protocols now in use vulnerable in the case of the advent of a working quantum computer.

The many important steps forward towards the construction of a stable and reliable quantum computer of the last decade have increased the need for new cryptographic protocol capable of withstanding the power of quantum attacks. Despite the fact that a large enough quantum computer is yet to be constructed, the need for durable security of sensitive data and the necessity of being prepared to future advancements in this direction have pushed researchers, government institutions and corporations to search for new cryptographic primitives. In December 2016, the National Institute of Standards and Technology (NIST) launched a selection process for post-quantum secure protocols [NIST]; aim of this call for application was the selection of new standards in cryptographic algorithms. In response, there has been a shift to different mathematical assumptions from the classical ones. To this day most of the new protocols are less efficient, slower or heavier than the ones they are called to replace; however, the growing demand for security have pushed the research in a whole new variety of mathematical areas. The hard problems underlying the security of these schemes can be mainly classified into the following families: lattices (the closest vector problem), codes (the decoding problem - finding the nearest codeword), multivariate polynomials, isogenies and hash functions.

Isogeny based cryptography

Isogeny based hard problems make their appearance in 2006 with the work of Charles, Goren and Lauter [CGL] who proposed a hash function constructed on supersingular isogeny graphs. That same year Rostovtsev and Stolbunov [RS] published an article in which they described a key exchange protocol based on isogenies between ordinary elliptic curves. This had already been described by Couveignes in 1996 but never made public. The construction takes place in a set of elliptic curves on which an action of a certain class group is defined. The great mathematical construction, however, could not supply a certain degree of inefficiency even though at the time there was no sub-exponential time attack against it. In 2010, Childs, Jao and Soukharev [CJS] exploited the fact that this action is commutative to adapt a subexponential-time quantum algorithm by Kuperberg [Kup1] to the isogeny construction of Rostovtsev, Stolbunov and Couveignes. The beautiful ordinary isogeny protocol suffers of an intrinsic contradiction: the mathematical property that makes it so elegant also represent its greater weakness confronted to a quantum computer.

Ordinary elliptic curves behave somehow too well for cryptography. Therefore, De Feo and Jao [DJ] moved their attention to the supersingular realm and successfully described an isogeny-based cryptosystem which takes place on a supersingular isogeny graph; this eliminates the commutativity of the class group action and circumvents the attack of [CJS]. This is known as SIDH, *Supersingular Isogeny Diffie-Hellman*. Since then, isogeny based cryptography has started to take a greater and more prominent role and one of its instantiation, SIKE, made it to the pool of alternates candidates for NIST standardization process; its relative slowness compared to other post-quantum candidates is partially compensated by the small size of the keys. Despite the great recent improvements to the SIDH protocol, it still lacks some flexibility and speed. In particular, commutative group action schemes still seemed useful and amenable of speedups: in 2017 De Feo, Kieffer and Smith [DKS] retrieved the ideas of Couveignes-Rostovtsev-Stolbunov and gave algorithmic improvements for the key exchange protocol; unfortunately, they had to conclude that, in their framework, parameter generation was too hard. However, their work laid the foundation for yet another cryptosystem based on the isogeny graph of a subclass of supersingular elliptic curves by Castryck, Lange, Martindale, Panny and Renes [Cas+] which took the name of CSIDH, *Commutative Supersingular Isogeny Diffie Hellman*. At the time of writing of this document a powerful new attack has been proposed by Castryck and Decru [CD] and a similar approach has been taken by Maino and Martindale [MM] embedding the supersingular isogeny graph in a bigger isogeny graph of supersingular abelian surfaces. This attack breaks all proposed parameters of SIKE. Since then, the attack has been generalized by Robert [Rob]. At present CSIDH and SIDH variants, such as SQISign [De+], resist these attacks

Leaving security aspects aside, among the other candidates for post quantum algorithms, those based on isogeny graphs are arguably the ones with the richest underlying mathematical structure. Lying at the crossroad of purely theoretical research and very practical applications they pose some very interesting questions both to cryptographers and number theorists. The supersingular isogeny path problem takes place in a geometric category of supersingular elliptic curves which has a corresponding algebraic side. In fact, in an equivalent category of left ideals for a quaternion order, there is an analogous path problem. This has been shown to be polynomial time in practice [KLPT]. This is an analogous situation to the one between the discrete logarithms problem (DLP) in finite fields (for which the best known algorithms are subexponential) and the DLP in an additive abelian group (which is trivial). Rather than rendering the supersingular isogeny path problem polynomial time, this result identifies the supersingular endomorphism ring problem [Koh1] as a (potentially hard) computational problem. A solution can be used to pull back the quaternion ideal isogeny path problem to the supersingular isogeny path problem. Nevertheless, the best known solution to this problem remains exponential.

Relevance and contributions

The first objective of this thesis project is to describe an additional structure of the endomorphism ring of supersingular elliptic curves. This comes in the form of an embedding of an imaginary quadratic order which provides a class group action on the set of supersingular curves; the isogeny graphs of curves endowed with this extra piece of information become an infinite volcano. The study of this new category of enhanced supersingular elliptic curves provides information on the entire set of curves and permits to define orientations and directions in an otherwise chaotic graph.

The recovery of a class group action motivates the construction of a computational model in which to frame an equivalence of categories between ideals, lattices, quadratic forms and CM points. This permits one to move from an algebraic to a more geometric point of view depending on the situation. Alongside with this equivalence of categories we provide the basis to endow these objects with level structure.

The tools for this work will be the study of the moduli spaces and invariant theory of curves, the specific analysis of the locus of supersingular points, the study of algorithms for efficient isogenies, and the analysis of equivalent algebraic categories on the quaternion side.

Finally, the project provides the motivation to investigate the possibility of new post-quantum cryptosystem associated to path problems in the new category of *oriented* supersingular curves. However, we stress the fact that this thesis, despite revolving around various aspects of cryptography is not about implementing or optimizing cryptographic algorithms but rather about exploring the mathematical background and providing tools to understand the structure of the objects in use. The main goal is the construction of a general framework in which to approach the study of supersingular-isogeny based constructions.

Structure of the Thesis

The general structure of the thesis is as follows. In the first 4 chapters we build a strong and complete theoretical background necessary to our construction in chapter 5. This consist in a mix of known results, different interpretations, ad-hoc constructions and

Chapter 1 introduces the necessary background theory to proceed in the understanding of the remaining of the thesis. In particular we discuss here general properties of elliptic and modular curves.

Chapter 2 surveys some of the main results about the modular curves $X_0(N)$ both from a geometric point of view and a more algebraic perspective. The Chapter is built to understand how isogenies fit in a modular tower description and aim towards the last section in which we explain how to construct precomputed models for squares of isogenies.

Chapter 3 explores the idea of adding different level structures to supersingular isogeny graphs. Sections 3.1 and 3.2 describe well known constructions while in Section 3.3 we introduce new modular towers motivated by the search for an optimal model to work on in the following chapters.

Chapter 4 investigates the relation between an isogeny and its kernel. We provide a quick survey on different models for elliptic curves with the corresponding modular description and we study how class groups of quadratic imaginary fields act on them. In particular, we specialize the work of Satoh [Sat] to the case of Gaussian and Eisenstein integers.

Chapter 5 represents the core of the thesis. All the previous chapters come together to provide a solid theory of orientations for supersingular elliptic curves. We describe how to impose a structure on an isogeny graph by means of isogeny chains and how to carry out an effective class group action, by means of ladders. We also construct a general computational model for augmented supersingular curves which permits to include additional level structure.

Most of the ideas presented in this Chapter appeared in [CK1].

Chapter 6 constructs a new cryptographic protocol relying on the idea of orientations. This is an application of the model described in previous chapters and despite the fact that its complexity remains exponential to these days, it is very slow and not effective. However, its mathematical interest motivates a generalization including a modular version of it.

This chapter mainly follows [CK1] and [CK2].

Chapter 1

Algebraic curves

1.1 Basic definitions

We start by recalling some basic facts on algebraic varieties. Main references will be [Har], [Sil1], [Sha] and [Gal4].

An affine space $\mathbb{A}^n$ of dimension n over K is a scheme such that $\mathbb{A}^n(L) = L^n$ for every extension L/K.

We define an algebraic set in $\mathbb{A}^n(K)$ as the vanishing locus V of a set of polynomials in $S \subseteq K[x_1, \ldots, x_n]$:

$$X = V(S) = \{P \in \mathbb{A}^n(\overline{K}) \bigm| f(P) = 0 \text{ for all } f \in S\}$$

For any extension L of K, the set of L-rational points of X is

$$X(L) = \{P \in \mathbb{A}^n(L) \mid f(P) = 0 \text{ for all } f \in S\}$$

Reversing our point of view, suppose we have an algebraic subset X of $\mathbb{A}^n(K)$; we define its ideal I as the polynomial ideal vanishing on X:

$$I(X) = \{f \in K[x_1, \ldots, x_n] \bigm| f(P) = 0 \text{ for all } P \in X(\overline{K})\}$$

If X is an algebraic set over K, then $X = V(I(X))$. The affine coordinate ring of an algebraic set X is

$$K[X] = K[x_1, \ldots, x_n]/I(X)$$

Studying the arithmetic of algebraic sets requires additional points, namely those lying at infinity. A projective space $\mathbb{P}^n$ of dimension n over K is a projective scheme such that $\mathbb{P}^n(K) = \left(\mathbb{A}^{n+1}(K) \setminus \{0\}\right) / \sim$ where the equivalence relation is given by $(a_1, \ldots, a_{n+1}) \sim (\lambda a_1, \ldots, \lambda a_{n+1})$ for any $\lambda \in K$. This gives a geometric interpretation of $\mathbb{P}^n(K)$ as the set of lines through the origin in $\mathbb{A}^{n+1}$. Points in $\mathbb{P}^n$ are represented by homogeneous coordinates $(x_0 : \ldots : x_n)$.

In order to adapt the definitions above to the projective realm, we have to keep track of the equivalence relation defining $\mathbb{P}^n$. We therefore have to introduce homogeneous polynomials as those polynomials f such that $f(\lambda x_0, \ldots, \lambda x_n) = \lambda^{\deg f} f(x_0, \ldots, x_n)$. A projective algebraic set is now the vanishing locus of a set of homogeneous polynomials

$$X = \overline{V}(S) = \{P \in \mathbb{P}^n(\overline{K}) \bigm| f(P) = 0 \text{ for all homogeneous polynomials } f \in S\}$$

A homogeneous ideal $I \in \overline{K}[x_0, \ldots, x_n]$ will be an ideal generated by homogeneous polynomials and to any set $X \subseteq \mathbb{P}^n(\overline{K})$ we associate its homogeneous ideal

$$I(X) = \{f \in K[x_0, \ldots, x_n] \mid f \text{ is homogeneous and } f(P) = 0 \text{ for all } P \in X\}$$

The projective coordinate ring of X is $K[x_0, \ldots, x_n]/I(X)$.

We say that X is defined over K if its ideal is generated by polynomials with coefficients in K and we denote its set of K-rational points by $X(K) = X \cap \mathbb{P}^n(K)$.

Definition. An algebraic set X is K-irreducible if it is defined over K and $I(X)$ is a prime ideal.

Definition. A projective K-irreducible algebraic set is called a projective algebraic variety. A curve is a projective variety of dimension 1.

Definition. Let X be a curve over K. This means that $I(X) = (f)$, i.e., its ideal is generated by a single polynomial. Let $P \in X(K)$ be a point on X and define the partial derivatives $\partial f / \partial x_i$ of f as usual.

We say that P is singular if $\partial f / \partial x_i (P) = 0$ for all i and nonsingular (or smooth) otherwise. X is nonsingular if all its points are nonsingular.

1.1.1 Divisors

Let X be a smooth projective curve over an algebraically closed field K.

Definition. A divisor on X is a formal finite sum

$$\sum_{P \in X} n_P\,(P)$$

with integer coefficients $n_P \in \mathbb{Z}$ such that $n_P = 0$ for all but finitely many points $P \in X$. The set of divisors of X is denoted $\mathrm{Div}(X)$ and, in algebraic terms, it is just the free abelian group on the points of X.

Notation. We use parentheses around points in order to distinguish divisors from exact sums of points on algebraic varieties.

Definition. A divisor D is called effective, and we write $D \geq 0$, if $n_P \geq 0$ for all P. The set of effective divisors is denoted $\mathrm{EDiv}(X)$ and it is the free abelian monoid on the points of X.

With this definition we can endow the set of divisors with a partial order: we say that $D_1 \geq D_2$ if $D_1 - D_2 \in \mathrm{EDiv}$ is an effective divisor.

Definition. The degree of a divisor $D = \sum_{P \in X} n_P(P)$ is the integer defined as the sum of all the coefficients n_P:

$$\deg(D) = \sum_{P \in X} n_p$$

The definition above defines a group homomorphism $\deg : \mathrm{Div}(X) \to \mathbb{Z}$ which takes positive values when restricted to EDiv and whose kernel consists of divisors of degree 0; we denote the kernel by

$$\mathrm{Div}^0(X) = \{D \in \mathrm{Div}(X) \mid \deg(D) = 0\}$$

Definition. We define the support of a divisor D to be the set of points $P \in X$ for which $n_P \neq 0$, denoted $\mathrm{supp}(D)$.

Now let f be a non-zero homogeneous polynomial on X. We construct the divisor of f as follows:

$$\mathrm{div}(f) = \sum_{P \in Z(f)} \mathrm{mult}_P(f)\,(P)$$

where $Z(f)$ denotes the zero locus of f on X and mult is the multiplicity of f at P (see [Gat, Ch. 14]).

Another way of constructing divisors is by considering intersection theory: if Y is another curve not containing X then we define

$$X \cdot Y = \sum_{P \in X \cap Y} \mathrm{mult}_P(X, Y)\,(P)$$

where $\mathrm{mult}_P(X, Y)$ is the multiplicity of the intersection.
The resulting divisor is clearly effective; this gives us a way of thinking of effective divisors as divisors encoding information about the intersection between X and another variety.
From the previous definition we can infer the main construction we will need. Let f be a rational map defined locally on X. We define the divisor of f as

$$\mathrm{div}(f) = \sum_{P \in X} \nu_P(f) P$$

having as coefficients the valuation of f at P given by the choice of a local uniformizer.

If $f \in K(X)^\times$ is defined globally, then the associated divisor

$$\operatorname{div}(f) = \sum_{P \in X} \operatorname{ord}_P(f)\,(P)$$

encodes information on the order of vanishing of f at points of X.

Remark. Observe that we have indeed constructed a divisor since every element of $K(X)^\times$ has only finitely many zeros and finitely many poles [NX, Cor. 3.3.2].

Similarly, we can define $\operatorname{div}(\omega)$ for any differential $\omega \in \Omega^1_X$. For any point $P \in X$ we let t be a uniformizer at P, i.e., a function with $\operatorname{ord}_P(t) = 1$; since Ω^1_X is a one dimensional $K(X)$ vector space, ω/dt is a well defined function and we can take $\operatorname{ord}_P(\omega)$ to be $\operatorname{ord}_P(\omega/dt)$. Then

$$\operatorname{div}(\omega) = \sum_{P \in X} \operatorname{ord}_P(\omega)\,(P)$$

Definition. A divisor D on X is said to be principal if it comes from a rational function, i.e., it is of the form $\operatorname{div}(f)$ for some $f \in K(X)^\times$. We denote $\operatorname{Prin}(X)$ the set of principal divisors.

Two divisors D_1 and D_2 are said to be linearly equivalent if their difference is a principal divisor; we write $D_1 \sim D_2$.

Given this equivalence relation we construct the Divisor Class Group (Picard Group) of X as the quotient

$$\operatorname{Pic}(X) = \operatorname{Div}(X)/\operatorname{Prin}(X)$$

For a general variety X, what we have defined is called the Weil divisor class group. The Picard group is more generally defined as the sheaf cohomology group

$$\operatorname{Pic}(X) = \mathrm{H}^1(X, \mathcal{O}_X^*)$$

isomorphic to the group of isomorphism classes of line bundles on X (the group structure being induced by the usual tensor product).

If X is normal, we can define a Cartier divisor to be a divisor D which is locally principal meaning that there exists a cover of X such that, for any open subset in the cover, the restriction of D is principal, i.e. each point $\operatorname{Supp}(D)$ has a neighborhood in which D is principal.

For an irreducible normal variety, the Picard group is isomorphic to the group of Cartier divisors modulo linear equivalence. This is equal to the Weil divisor class group if X is locally factorial and, in particular, if X is smooth, see [Har, §II.6].

Definition. Since Ω^1_X is one dimensional over $K(X)$, then the divisors of all differentials belong to the same equivalence class. We will refer to it as the canonical class. Any divisor coming from a differential is called canonical divisor (generally denoted K_X).

Theorem 1.1. *Let X be a smooth curve over an algebraically closed field K and $f \in K(X)^\times$. Then*

(a) $\operatorname{div}(f) = 0 \iff f \in K^\times$ *is a constant function.*

(b) $\deg(\operatorname{div}(f)) = 0$.

This means that the map $K(X)^\times \xrightarrow{\operatorname{div}} \operatorname{Div}(X)$ has image contained in $\operatorname{Div}^0(X)$ and kernel equal to $K^\times$. Motivated by the previous Theorem, we introduce the following definition.

Definition. Let $D \in \operatorname{Div}(X)$. We associate to it the set of rational functions (Riemann-Roch space)

$$L(D) = \{f \in K(X)^\times \mid \operatorname{div}(f) + D \geq 0\} \cup \{0\}$$

The set $L(D)$ is a vector space over K of finite dimension, which we will denote by $\ell(D)$, [Sil1, Th. II.5.2]. The dimension $\ell(D)$ grows as we allow the functions to have more poles (without imposing restrictions on zeroes); more precisely, $\ell(D_1) \geq \ell(D_2)$ whenever $D_1 \geq D_2$.

Theorem 1.2 (Riemann-Roch)**.** *Let X be a smooth curve over an algebraically closed field K. Let K_X be a canonical divisor and g the genus of the curve. For any divisor $D \in \mathrm{Div}(X)$,*

$$\ell(D) - \ell(K_X - D) = \deg(D) - g + 1$$

Corollary 1.3 (Riemann)**.** *We have the following*

(a) $\ell(K_X) = g$.

(a) $\deg(K_X) = 2g - 2$.

(a) *If* $\deg(D) > 2g - 2$*, then* $\ell(D) = \deg(D) + 1 - g$.

Before moving on describing maps between curves, we devote the remaining part of the section to the study of divisors of rational functions.

Definition. Let $f \in K(X)^\times$ be a rational function. We have already seen that it has a finite number of zeroes and poles. We define $\mathcal{N}(f)$ the set of zeroes of f and $\mathcal{P}(f)$ the set of its poles.

Definition. We introduce the *divisor of zeroes* of f as

$$\mathrm{div}(f)_0 = \sum_{P \in \mathcal{N}(f)} \mathrm{ord}_P(f)\,(P)$$

and, analogously, the *divisor of poles* of f by

$$\mathrm{div}(f)_\infty = \sum_{P \in \mathcal{P}(f)} -\mathrm{ord}_P(f)\,(P)$$

These two definitions will play a central role in the construction of explicit models of modular curves in Chapter 2.
At this stage we limit ourselves to observe that divisors of zeroes and poles of f are intimately related.

Theorem 1.4. *For any $f \in K(X) \setminus K$,* $\deg(\mathrm{div}(f)_0) = [K(X) : K(f)]$.

Proof. See [NX, Th. 3.4.2] □

Corollary 1.5. *For any non-zero $f \in K(X)^\times$,* $\deg(\mathrm{div}(f)) = 0$*, implying that* $\deg(\mathrm{div}(f)_0) = \deg(\mathrm{div}(f)_\infty)$.

Proof. The first part is already in the statement of Theorem 1.1. The second follows from the fact that

$$\mathrm{div}(f) = \mathrm{div}(f)_0 - \mathrm{div}(f)_\infty$$

and both the divisors of zeroes and poles have positive degree by construction. □

1.1.2 Maps on Algebraic Curves

In this section we give a brief look at the theory of maps between curves. The main reference will be [Sil1, §II.2]. We start by stating a fundamental theorem

Theorem 1.6. *Let $\psi : X_1 \to X_2$ be a non-constant rational morphism between projective curves over K. Then:*

(a) *The induced map*

$$\begin{aligned} \psi^* : K(X_2) &\longrightarrow K(X_1) \\ f &\longrightarrow f \circ \psi \end{aligned}$$

is a homomorphism fixing K.

(b) *The field extension $K(X_1)/\psi^*(K(X_2))$ is finite.*

Definition. Let $\psi : X_1 \to X_2$ be a non-constant map of smooth curves. If ψ is constant, we define the degree of ψ to be 0. Otherwise, thanks to the previous theorem, we define

$$\deg(\psi) = [K(X_1) : \psi^* K(X_2)]$$

Definition. Let $\psi : X_1 \to X_2$ be a non-constant map of smooth curves and let P be a point on X_1. The ramification index of ψ at P is defined as

$$e_\psi(P) = \operatorname{ord}_P(t_{\psi(P)})$$

for a choice of a uniformizer $t_{\psi(P)} \in K(X_2)$ at $\psi(P)$.

Proposition 1.7. *With the notation as above, for every point $Q \in X_2$,*

$$\deg \psi = \sum_{P \in \psi^{-1}(Q)} e_\psi(P)$$

Therefore, the degree of the map counts how many points lay over most points of X_2 counted with their "multiplicity" (their ramification index).

We have seen that any rational morphism between two curves induces a map between their function fields. In view of the previous section, a natural question might be how rational maps act on the set of divisors.

For $\psi : X_1 \to X_2$, we construct a map

$$\begin{aligned} \psi^* : \operatorname{Div}(X_2) &\longrightarrow \operatorname{Div}(X_1) \\ (Q) &\longmapsto \sum_{P \in \psi^{-1}(Q)} e_\psi(P)\,(P) \end{aligned}$$

and we extend it $\mathbb{Z}$-linearly.

Proposition 1.8. *Let $\psi : X_1 \to X_2$ be a non-constant map of smooth projective curves.*

(a) *If $D \in \operatorname{Div}(X_2)$, then $\deg(\psi^* D) = \deg(\psi)\deg(D)$.*

(a) *The pullback ψ^* preserves principal divisors; moreover $\psi^*\operatorname{div}(f) = \operatorname{div}(\psi^* f)$ for any rational function f on X_2.*

1.1.3 The Canonical Embedding

Let D_0 be a divisor on X. We define the complete linear system $\mathcal{L}(D_0)$ of D_0 as follows: it is the set of effective divisors which are linearly equivalent to D_0:

$$\mathcal{L}(D_0) = \{D \in \operatorname{EDiv}(X) \mid D \sim D_0\}$$

Since $D \sim D_0$, by definition we know that $D = D_0 + \operatorname{div}(f)$ for some rational function $f \in K(X)^\times$. Hence, we can identify D with the function it comes from, f. In this case the property of being effective becomes the condition $\operatorname{div}(f) \geq D_0$.
We therefore recover the definition of Riemann-Roch space. In some literature, linear systems are indicated as $|D_0|$ and they are called complete linear series.

Remark. Note that $\operatorname{div}(f) = \operatorname{div}(\alpha f)$ for all $\alpha \in K^\times$. For this reason, is more natural to consider the corresponding projective space $|D_0| \simeq \mathbb{P}D_0 = (L(D_0) \setminus \{0\})/K^\times$ (sometimes also called linear series).

Definition. The base locus of a divisor D is the set of all closed points $x \in X$ such that $f(x) = 0$ for all $f \in \mathbb{P}D$. It is a proper closed subset of X. In terms of divisors this condition translates into the property that the point is in the support of all the divisors of the linear system.

We now fix a basis $\{f_0, \ldots, f_{\ell(D)-1}\}$ for $\mathcal{L}(D)$; on the complement of the base locus we can define a morphism

$$\begin{aligned} X &\longrightarrow \mathbb{P}D \\ x &\longrightarrow (f(x)) = (f_0(x) : \ldots : f_{\ell(D)-1}(x)) \end{aligned}$$

Definition. If we consider a canonical divisor K_X, the map constructed above is called canonical embedding.

Remark. If X is a curve of genus 0, then the Riemann-Roch Theorem (see Th. 1.2) gives $\ell(K_X) = 0$, meaning that the canonical linear system is empty. Thus, the canonical map cannot be defined.

Remark. If the genus of X is 1, then the canonical map $X \to \mathbb{P}^0$ collapses X to a single point.

Observe that there is a more general way of constructing these objects for a wider class of curves in terms of global sections of sheaves, see [Har] and [Barb, §1.1].
To any divisor $D \in \mathrm{Div}(X)$ one can associate the line bundle $\mathcal{L}(D)$ defined over any open subset $\mathcal{U}$ of X by the requirement:

$$H^0(\mathcal{U}, \mathcal{L}(D)) = \{f \in K(X)^\times \mid \mathrm{div}(f)|_{\mathcal{U}} + D|_{\mathcal{U}} \geq 0\}$$

i.e., the line bundle whose global sections are locally controlled by D. As for rational functions, the sections of $\mathcal{L}(D)$ can have poles only on the support of D.

Remark. Every line bundle $\mathcal{O}_X$ on X admits a global section s and, therefore, it can be written as $\mathcal{O}_X = \mathcal{L}(D)$ with $D = (s)$ being the divisor corresponding to such a section. This follows from Riemann-Roch Theorem which provides the dimension of the space of global sections.

As before, one can define two divisors to be linearly equivalent if their difference is a principal divisor. Next, if D is a divisor, we denote $|D|$ the set of all effective divisors that are linearly equivalent to D; this set is called the complete linear series of D; by sending global sections $f \in H^0(D)$ to $D + \mathrm{div}(f)$ we get an isomorphism $\mathbb{P}H^0(D) \simeq |D|$ where the right-hand side corresponds to the projectivization of the vector space of global sections of the line bundle $\mathcal{L}(D)$. More generally, to any linear subspace $V \subseteq H^0(D)$ we associate the projective space $\mathbb{P}V$, called linear series.

Definition. If D is a divisor of degree d and $V \subseteq H^0(D)$ is a linear subspace of dimension $r+1$, we define $\mathfrak{g}^r_d$ to be the linear series associated to $\mathbb{P}(V)$.

Remark. We can now mimic the construction for rational function to obtain divisors starting from sections of arbitrary line bundles on X.

For instance, we can consider the cotangent line bundle Ω^1_X of holomorphic differentials on X (called the canonical line bundle). If ω is a section of Ω^1_X, we can consider any open cover $\{\mathcal{U}_i\}_{i\in I}$ of X and a local uniformizer t_i for any i. Then

$$\omega|_{\mathcal{U}_i} = g_i dt_i$$

hence, $\mathrm{div}(\omega) = \sum_{i\in I} \mathrm{div}(g_i)$.

Definition. A linear series $\mathfrak{g}^r_d$ is base-point-free if there is no point contained in the supports of all its divisors.

Any base-point free linear series can now be used to construct a map of X into a projective space.

$$\begin{aligned} \varphi_V : X &\longrightarrow \mathbb{P}(V)^* \\ P &\longrightarrow \{s \in V \mid s(P) = 0\} \end{aligned}$$

where $\mathbb{P}V^*$ is the dual projective space. Finally, we extend the map φ_V to any effective divisor $D = \sum_{i=1}^d P_i$ by

$$\varphi_V(D) = \langle \varphi_V(P_1), \dots, \varphi_V(P_d) \rangle$$

If X has genus $g \geq 2$, then we can choose any canonical divisor K_X and the complete linear series $|K_X|$ gives rise to the canonical map

$$\varphi_K : X \longrightarrow \mathbb{P}^{g-1}$$

Recall. The genus g equals the dimension of Ω^1_X.

Lemma 1.9. *If $g \geq 2$ and X is not hyperelliptic, then the canonical map is injective.*

For genus 0 curves, Ω^1_X has no non-zero global sections implying that the canonical map is not defined. If X is a curve of genus 1 the space of global sections of the cotangent bundle Ω^1_X has dimension 1 and, thus, the canonical map simply reduces X to a point: $X \longrightarrow \mathbb{P}^1$

Finally, for curves of genus $g \geq 2$ we know that the canonical bundle is always globally generated meaning that the canonical map $X \to \mathbb{P}^{g-1}$ is a morphism of algebraic varieties defined everywhere. We distinguish two possibilities: either the canonical map is an embedding or not. The latter occurs if the image is isomorphic to $\mathbb{P}^1$ and the map is generically 2-to-1 from which X is a double cover of a rational normal curve in $\mathbb{P}^{g-1}$. This happens if and only if X is hyperelliptic. The factorization below is obtained following [Har, Prop. IV.5.3] for hyperelliptic (right) and non-hyperelliptic curves (left).

$$\begin{tikzcd} X \arrow[r, hook, "\varphi_K"] & \mathbb{P}^{g-1} \end{tikzcd} \qquad \begin{tikzcd} X \arrow[rr, "\varphi_K"] \arrow[dr] & & \mathbb{P}^{g-1} \\ & \mathbb{P}^1 \arrow[ur, hook] & \end{tikzcd}$$

1.2 Elliptic curves

An elliptic curve E over K is a non singular projective curve of genus 1 together with a distinguished point $O \in E(K)$. From the Riemann-Roch Theorem there are two rational functions $x \in \mathcal{L}(2(O)) \setminus K$ and $y \in \mathcal{L}(3(O)) \setminus \mathcal{L}(2(O))$ such that the triple $(x, y, 1)$ is a basis of $\mathcal{L}(3(O))$ and defines a non-constant rational map to $\mathbb{P}^2(K)$ [Miln, Ch 1]. Since $\ell(6(O)) = 5$, they satisfy a relation in $K[x, y]$

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

called Weierstrass equation. Its projectivization

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

defines the projective model of E in $\mathbb{P}^2$ with O the point at infinity $(0 : 1 : 0)$.

Following [Sil1, §III.1] we define

$$\begin{aligned} &b_2 = a_1^2 + 4a_2 \qquad b_4 = a_1a_3 + 2a_4 \qquad && b_6 = a_2^3 + 4a_6 \\ &b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2 && \text{so that } 4b_8 = b_2b_6 - b_4^2 \end{aligned}$$

and also

$$c_4 = b_2^2 - 24b_4 \qquad c_6 = -b_2^3 + 36b - 2b_4 - 216b_6$$

If K is a field of characteristic different from 2 and 3, then the substitution

$$(x, y) \longmapsto \left(\frac{1}{36}(x - 3b_2), \frac{1}{216}(y - a_1x - a_3) \right)$$

yields the short model $y^2 = x^3 - 27c_4x - 54c_6$.

To an elliptic curve E we usually associate 3 invariants [Sil1, §III.1]. The quantity $\Delta = -b_2^2b_8 - 8b_4^3 27b_6^2 + 9b_2b_4b_6 = (c_4^3 - c_6^2)/12^3$ is called the discriminant of the Weierstrass equation and it classifies singular curves as E is nonsingular if and only if $\Delta \neq 0$. The j invariant of E is $j = c_4^3/\Delta$. It parametrizes isomorphism classes of Elliptic curves over $\overline{K}$. Note that if K is not algebraically closed there exist twists of elliptic curves with the same j invariants that are not isomorphic over K. Finally, the differential

$$\omega_E = \frac{dx}{2y + a_1x + a_3}$$

generates the sheaf of differentials Ω_E of E and it is called the *invariant differential*.

1.2.1 Addition laws

Elliptic curves admit a structure of commutative algebraic group where the identity is the distinguished point O. In order to see this, one could consider the natural map

$$\begin{aligned} X &\longrightarrow \mathrm{Pic}^0(X) \\ P &\longmapsto (P) - (O) \end{aligned}$$

which is injective since $(P) - (O)$ is principal if and only if $P = O$. The Riemann-Roch Theorem permits one to prove that this is also surjective allowing the identification of E with its Jacobian variety.
The group law coming from this isomorphism is equivalent to the geometric *chord and tangent rule* on projective cubic curves introduced by Jacobi.

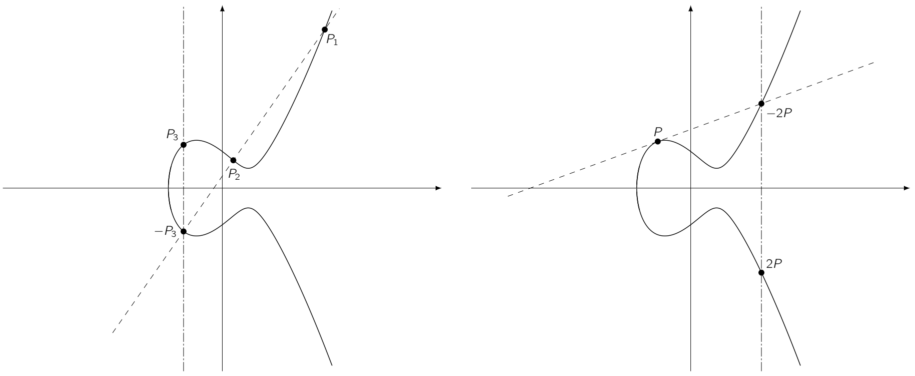


Figure 1.1 – Chord and tangent rule for the Weierstrass cubic

Suppose we have two points P_1 and P_2 on E, the line through them intersects the cubic defining E in a third point $\tilde{P}_3$ by Bezout's Theorem and, if P_1 and P_2, are rational then the third intersection is also rational. The addition law on E is defined by the relation $P_1 + P_2 + \tilde{P}_3 = O$, namely the sum of any three colinear points (counted with their multiplicity) is O. The sum of P_1 and P_2 is therefore the third intersection of E with the line passing through O and $\tilde{P}_3$.
For more details on the chord and tangent rule one may refer to the introduction of [Hus]. The equivalence between the algebraic group law coming from $\mathrm{Pic}^0(E)$ and the geometric interpretation is Proposition III.3.4 of [Sil1]. Explicit formulæ for point addition can be found in [Sil1, p. III.2.3].

Theorem 1.10. *The elliptic curve E admits the structure of a group scheme with O as identity; in particular, the set of K-rational points of $E(K)$ is an abelian group under the addition law defined below.*

(a) *The inverse of $P = (x_0, y_0)$ is $-P = (x_0, -y_0 - a_1x_0 - a_3)$.*

(b) *The sum of $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ is $P_3 = (x_3, y_3)$ such that*

- *If $P_2 = -P_1$ then $P_3 = O$.*
- *Otherwise*

$$\begin{cases} x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2 y_3 = -(\lambda + a_1)x_3 - \nu - a_3 \end{cases}$$

where

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} & \text{if } P_1 \neq P_2 \\ \dfrac{(3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3)} & \text{if } P_1 = P_2 \end{cases} \qquad \text{and} \qquad \nu = y_1 - \lambda x_1$$

1.2.2 Kummer line

The negation map $[-1] : E \to E$ mapping a point P to its inverse is an automorphism of E and the quotient of E modulo $\{[\pm 1]\}$ is a degree two cover $\sigma : E \to \mathbb{P}^1 \simeq E/\{[\pm 1]\}$.

The Kummer line $\mathcal{K}_E$ of E is defined to be the singular projective curve obtained in this way, by quotienting E by the subgroup of $\mathrm{Aut}(E)$ generated by -1. This means that the Kummer line consists of the set of coordinates invariant under the inversion. In particular, for the Weierstrass elliptic curve, the map to $\mathbb{P}^1$ can be defined as the x-map associating to P its x-coordinate.

$$(X : Y : Z) \longmapsto (X : Z) \qquad \text{where } x = X/Z$$

Since we cannot distinguish between a point and its inverse, the group law on E does not induce an addition law on $\mathcal{K}_E$. In particular, there exists no map $\mathcal{K}_E \times \mathcal{K}_E \to \mathcal{K}_E$ mapping $(\sigma(P_1), \sigma(P_2)) \mapsto \sigma(P_1 + P_2)$. The best we could do is to define $(\sigma(P_1), \sigma(P_2)) \mapsto \{\sigma(P_1 - P_2), \sigma(P_1 + P_2)\}$. However, there exist algorithms permitting one to construct $\sigma(P+Q)$ once $\sigma(P_1), \sigma(P_2)$ and $\sigma(P_1 - P_2)$ are known. In fact, one could construct methods to recover any one of $\sigma(P_1), \sigma(P_2), \sigma(P_1 - P_2), \sigma(P+Q)$ when the other three are known. This kind of operation is called differential addition [CS3].

Example. Let E be the elliptic curve $Y^2Z = X^3 + aXZ^2 + bZ^3$ in short Weierstrass form. Let us take $P_1 = (X_1 : Y_1 : Z_1)$ and $P_2 = (X_2 : Y_2 : Z_2)$; we define $P_0 + P_2 = P_1$ and $P_1 + P_2 = P_3$. We note $P_0 = (X_0 : Y_0 : Z_0)$ and $P_3 = (X_3 : Y_3 : Z_3)$.

Then $\sigma(P_1) = (X_1 : Z_1)$, $\sigma(P_2) = (X_2 : Z_2)$, $\sigma(P_0) = (X_0 : Z_0)$ and $\sigma(P_3) = (X_3 : Z_3)$.

```
Algorithm 1. Differential addition on the Kummer curve of a short Weierstrass curve
Input: $\sigma(P_1) = (X_1 : Z_1)$ , $\sigma(P_2) = (X_2 : Z_2)$ , $\sigma(P_0) = (X_0 : Z_0)$
Output: $\sigma(P_3) = (X_3 : Z_3)$ if $P_1 - P_2 \neq O$ or $X_3 = 0,\ Z_3 = 0$ otherwise

1. Compute the addition formula in projective coordinates for $X$
       $X_3 = Z_0\left[(X_1X_2 - aZ_1Z_2)^2 - 4b(Z_1X_2 + X_1Z_2)Z_1Z_2\right]$
2. Compute the addition formula in projective coordinates for $Z$
       $Z_3 = X_0\left(Z_1X_2 - X_1Z_2\right)^2$
3. Return $(X_3 : Z_3)$.
```

Remark. We note that $(E \times E)/\{[\pm 1]\} \to \mathcal{K}_E$ is well defined as a morphism.

In the same way, we can define pseudo doubling, i.e., a strategy to recover $\sigma(2P)$ knowing $\sigma(P)$. We note $P := (X_P : Y_P : Z_P)$ and $2P = (X_{2P} : Y_{2P} : Z_{2P})$.

```
Algorithm 2. Doubling on the Kummer curve of a short Weierstrass curve
Input: $\sigma(P) = (X_P : Z_P)$
Output: $\sigma(2P) = (X_{2P} : Z_{2P})$ if $P \neq O$ or $X_+ = 0,\ Z_+ = 0$ otherwise

1. Compute the doubling formula in projective coordinates for $X$
       $X_{2P} = \left[(X_P^2 - aZ_P^2)^2 - 4b(2X_PZ_P)Z_P^2\right]$
2. Compute the doubling formula in projective coordinates for $Z$
       $Z_{2P} = 4\left(Z_P^4 + aX_Pz_P^3 + bZ_P^4\right)^2$
3. Return $(X_{2P} : Z_{2P})$.
```

Remark. Scalar multiplication is well defined on $\mathcal{K}_E$.

1.2.3 Isogenies

An isogeny $\phi : E_1 \to E_2$ of elliptic curves is a non-constant morphism of curves sending the base point of E_1 to the base point of E_2. Isogenies respect the group structure of the elliptic curves thus inducing homomorphisms. An isogeny induces the usual injection of function fields

$$\phi^* : K(E_2) \longrightarrow K(E_1) \qquad f \mapsto f \circ \phi$$

We define the degree of ϕ as the degree of this extension: $\deg\phi = [K(E_1) : \phi^* K(E_2)]$ and its separable and inseparable degrees accordingly. We also say that ϕ is separable, inseparable or purely inseparable if the associate field extension is.

Proposition 1.11 ([Sil1, Th. III.4.10]). *If $\phi : E_1 \to E_2$ is an isogeny of elliptic curves, then*

(a) *For every point $Q \in E_2$, $\#\phi^{-1}(Q) = \deg_s \phi$. Further, if $P \in E_1$, then $e_\phi(P) = \deg_i \phi$.*

(b) *If ϕ is separable, then the associated field extension is Galois and $\deg\phi = \#\ker\phi$.*

The second statement is actually much stronger than numerical equality; separable isogenies are completely determined (up to isomorphism) by their kernels: for every finite subgroup G of E there exist a unique elliptic curve E/G up to isomorphism and a separable isogeny $\phi_G : E \to E/G$ with kernel G [Sil1, Prop. III.4.12].

We say that two curves E_1 and E_2 are isogenous if there exists an isogeny between them and we indicate by $\mathrm{Hom}(E_1, E_2)$ the set of isogenies between them together with the zero morphism. Note that the group structure on E_2 induces an abelian group structure on $\mathrm{Hom}(E_1, E_2)$. We let $\mathrm{End}(E) = \mathrm{Hom}(E, E)$ denote the set of isogenies from E to itself, called endomorphisms, and $\mathrm{Aut}(E)$ be the subset of invertible elements in $\mathrm{End}(E)$, called automorphisms. The set of endomorphisms has a ring structure given by addition and composition and it is therefore referred to as the *endomorphism ring* of E.

The addition law on an elliptic curve is a morphism of varieties $E \times E \to E$ which induces morphisms

$$\begin{aligned}[m] : E &\longrightarrow E\\ P &\longmapsto [m]P = P + \ldots + P\end{aligned}$$

associating to a point P the sum of P to itself m times. This allows one to define an injective ring homomorphism

$$\mathbb{Z} \longrightarrow \mathrm{End}(E)$$

which induces a $\mathbb{Z}$-module structure on $\mathrm{End}(E)$.

Multiplication-by-m maps are separable isogenies of degree m^2 [Sil1, Cor III.5.4/III.6.4] and their kernels are called the m-torsion subgroups of E: they consist of points of order m and are denoted

$$E[m] = \{P \in E \mid [m]P = O\}$$

Theorem 1.12. *Let $\phi : E_1 \to E_2$ be a non-constant isogeny of degree m. Then there exists a unique isogeny $\hat{\phi} : E_2 \to E_1$, called the dual isogeny, such that*

$$\phi \circ \hat{\phi} = [m] : E_2 \to E_2 \qquad \hat{\phi} \circ \phi = [m] : E_1 \to E_1$$

The dual isogeny permits one to show that the degree map

$$\deg : \mathrm{Hom}(E_1, E_2) \longrightarrow \mathbb{Z}$$

defines a positive definite quadratic form [Sil1, Cor. III.6.3] and the associated bilinear form

$$\Phi : \mathrm{Hom}(E_1, E_2) \times \mathrm{Hom}(E_1, E_2) \longrightarrow \mathbb{Z} \qquad (\phi, \psi) \longmapsto \hat{\phi}\psi + \hat{\psi}\phi$$

makes $\mathrm{Hom}^0(E_1, E_2) = \mathrm{Hom}(E_1, E_2) \otimes \mathbb{Q}$ into a positive definite quadratic space.

Remark. Here we are identifying $\mathbb{Z}$ with its image in $\mathrm{End}(E_1)$.

As a consequence, we obtain a description of the m-torsion subgroup of E.

Lemma 1.13 ([Sil1, Cor. III.6.4]). *Let E be an elliptic curve and m an integer different from 0. If $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) = p$ with $(p, m) = 1$, then*

$$E[m] = \frac{\mathbb{Z}}{m\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$$

If $\mathrm{char}(K) = p$ then one of the following holds

$$E[p^e] = \{O\} \qquad \text{or} \qquad E[p^e] = \frac{\mathbb{Z}}{m\mathbb{Z}}$$

for all $e \geq 1$.

Suppose we have an elliptic curve E and consider a prime ℓ. The ℓ^n torsion group $E[\ell^n]$ has been described above. We construct a compatible system where the maps are given by the multiplication-by-ℓ map

$$E[\ell] \xleftarrow{[\ell]} E[\ell^2] \xleftarrow{[\ell]} E[\ell^3] \xleftarrow{[\ell]} \dots$$

The inverse limit $T_\ell(E) = \varprojlim_n E[\ell^n]$ is called the (ℓ-adic) Tate module attached to E.

Remark. By the lemma above the Tate module $T_\ell(E)$ is isomorphic to $\simeq \mathbb{Z}_\ell^2$ or, in case $\ell = p$, either $\mathbb{Z}_\ell$ or the trivial module.

Notation. It might be sometimes convenient to work with $V_\ell(E) = T_\ell(E) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$.

Suppose now that $P \equiv (x, y) \in E$ is a point; then the absolute Galois group $\mathcal{G}_K = \mathcal{G}a\ell(\overline{K}/K)$ acts on it by $\sigma \cdot P = (\sigma(x), \sigma(y))$. Further, if $P \in E[\ell^n]$ then $\sigma \cdot P \in E[\ell^n]$. We obtain an action of $\mathcal{G}_K$ on the Tate module. This yields the Galois representation

$$\mathcal{G}_K \longrightarrow \mathrm{Aut}\,(T_\ell(E)) \simeq \mathrm{GL}_2(\mathbb{Z}_\ell) \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell)$$

where the isomorphism comes from the choice of a basis of $T_\ell(E)$.

The Tate module is quite useful when it comes to the study of isogenies. In fact, it turns out that there is an injective homomorphism

$$\mathrm{Hom}(E_1, E_2) \otimes \mathbb{Z}_\ell \longrightarrow \mathrm{Hom}(T_\ell(E_1), T_\ell(E_2))$$

and since $\mathrm{Hom}\,(T_\ell(E_1), T_\ell(E_2)) \simeq \mathrm{M}_2(\mathbb{Z}_\ell)$, this implies that $\mathrm{Hom}(E_1, E_2)$ is a free $\mathbb{Z}$-module of rank at most 4.

In the case of endomorphism rings this result is even more precise. A well known Theorem of Deuring states that the endomorphism ring of an elliptic curve E/K is either $\mathbb{Z}$ or an order in an imaginary quadratic field if K has characteristic zero and either an order in an imaginary quadratic field or a maximal order in a quaternion algebra if $\mathrm{char}(K) > 0$.

Definition. Elliptic curves with endomorphism ring isomorphic to a quadratic imaginary order are called *ordinary*. Elliptic curves whose full endomorphism ring is a quaternion order are called *supersingular*.

1.2.4 Elliptic curves over finite fields

The Frobenius endomorphism

Let k be the finite field with q elements, $k = \mathbb{F}_q$; the map $\phi_q : a \mapsto a^q$ is an automorphism of k which represents a profinite generator of the Galois group $\mathcal{G}a\ell(\overline{k}/k)$. We call ϕ_q the Frobenius automorphism of k. Since for any polynomial $f(x_1, \dots, x_k) \in k[x_1, \dots, x_k]$ we have $f(x_1, \dots, x_k)^q = f(x_1^q, \dots, x_k^q)$, we can extend the action of Frobenius to polynomial rings and ideals. One can verify that, in fact, it extends to all varieties over k.

Definition. Let E be an elliptic curve over k. The Frobenius morphism on E, induced by ϕ_q, is an endomorphism

$$\begin{aligned} \pi_q : E &\longrightarrow E \\ (x, y) &\longmapsto (x^q, y^q) \end{aligned}$$

called the Frobenius Endomorphism of E.

The main properties of the Frobenius endomorphism are given by the following proposition.

Proposition 1.14. *Let E be an elliptic curve defined over a finite field k. The Frobenius endomorphism π is a purely inseparable map of degree q.*

This result shows that any isogeny $\psi : E_1 \to E_2$ between elliptic curves over a finite field factors through the Frobenius map as

$$\begin{tikzcd}
E_1 \arrow[r, "\psi"] \arrow[d, "\pi"'] & E_2 \\
E_1^{\pi} \arrow[ur, bend right, "\phi"'] &
\end{tikzcd}$$

where $q = \deg_i(\psi)$, π is the q-th power Frobenius isogeny, and ϕ is separable.

One associate to π its characteristic polynomial

$$\chi(x) = x^2 - tx + q$$

such that $\chi(\phi) = \pi^2 - [t]\pi + [q] = [0]$. The integer t plays an important role in the description of the set of points of E over k.

Cardinality

We describe now an important invariant of isogeny classes of elliptic curves over finite fields: the set of rational points.

Let E be an elliptic curve defined over the finite field k. The set of points fixed by π is exactly $E(k)$. Thus, $E(k) = \ker(\pi - 1)$; since the isogeny $\pi - 1$ is separable (see [Sil1, Cor. III.5.5]), the cardinality of $E(k)$ is $\deg(\pi - 1)$. We obtain the following theorem

Theorem 1.15 (Hasse-Weil)**.** *If E is an elliptic curve defined over a field k with q elements, then*

$$\# E(k) = 1 + q - t \qquad \text{for } |t| \leq 2\sqrt{q}$$

The integer t is called the *trace of Frobenius*. An important result of Tate [Tat] establishes that the trace of Frobenius determines the isogeny class of E:

Theorem 1.16. *Two elliptic curves E_1 and E_2 defined over a finite field k are isogenous over k if and only if $\#E_1(k) = \#E_2(k)$.*

Remark. Equivalently, E_1 and E_2 are isogenous over k if and only if they have the same trace of Frobenius.

If p is a prime and $k = \mathbb{F}_p$, Waterhouse [Wat1] went a bit further showing that the possible values for the trace of Frobenius over $\mathbb{F}_p$, namely the integers in $[-2\sqrt{p}, 2\sqrt{p}]$, are indeed in a one to one correspondence with isogeny classes of elliptic curves over $\mathbb{F}_p$. More in general,

Theorem 1.17 ([Wat1, Th. 4.1])**.** *Let $q = p^e$. There exists an isogeny class of elliptic curves defined over $\mathbb{F}_q$ with trace of Frobenius t if and only if one of the following holds:*

(a) $(p, t) = 1$ *and* $t \leq 2\sqrt{q}$.

(b) *e is even and either*

- **(i)** $t = \pm 2\sqrt{q}$
- **(ii)** $p \not\equiv 1 \mod 3$ *and* $t = \pm\sqrt{q}$;
- **(iii)** $p \not\equiv 1 \mod 4$ *and* $t = 0$.

(c) *e is odd and either*

- **(i)** $t = 0$;
- **(ii)** $p = 2$ *and* $t = \pm\sqrt{2q}$;
- **(iii)** $p = 3$ *and* $t = \pm\sqrt{3q}$.

The endomorphism ring

As we have already stated at the end of section 1.2.3, over a finite field k we have the following classification:

Theorem 1.18. *Let E be an elliptic curve defined over a finite field K. The endomorphism ring* $\mathrm{End}(E)$ *is either an order $\mathcal{O}$ in a quadratic imaginary field K or a maximal order $\mathfrak{O}$ in a quaternion algebra $\mathfrak{A}$.*

We recall that multiplication-by-m maps always provide an embedding $\mathbb{Z} \hookrightarrow \mathrm{End}(E)$ but, over a finite field, the existence over the Frobenius endomorphism satisfying a certain quadratic equation implies that this is not the whole story. However, we know that π commutes with all isogenies and, therefore, it lies in the center of the endomorphism ring.

Lenstra [Len] showed that the the endomorphism ring structure determines the set of rational points over extensions of k.

Theorem 1.19. *Let k be a finite field, E an elliptic curve over k and π the Frobenius endomorphism of E. Further, let κ be a finite field extension of k, of degree $r = [\kappa : k]$.*

(a) *If $\pi \notin \mathbb{Z}$, then* $\mathrm{End}_k(E)$ *has rank 2 over $\mathbb{Z}$ and*

$$E(\kappa) \simeq \frac{\mathrm{End}_k(E)}{(\pi^r - 1)}$$

(b) *If $\pi \in \mathbb{Z}$, then* $\mathrm{End}_k(E)$ *has rank 4 over $\mathbb{Z}$ and*

$$E(\kappa) \simeq \frac{\mathbb{Z}}{\mathbb{Z}(\pi^r - 1)} \oplus \frac{\mathbb{Z}}{\mathbb{Z}(\pi^r - 1)}$$

as abelian groups. Further,

$$E(\kappa) \oplus E(\kappa) \simeq \frac{\mathrm{End}_k(E)}{(\pi^r - 1)}$$

as $\mathrm{End}_k(E)$*-modules.*

This, together with Theorem 1.16, yields the following

Theorem 1.20 (Sato-Tate)**.** *Two elliptic curves E_1 and E_2 defined over a finite field k are isogenous if and only if their endomorphism algebras are isomorphic as $\mathbb{Q}(\pi)$-modules:*

$$\mathrm{End}^0(E_1) \simeq \mathrm{End}^0(E_2) \qquad \text{where } \mathrm{End}^0(E) = \mathrm{End}(E) \otimes \mathbb{Q}$$

Ordinary and supersingular elliptic curves

In Section 1.2.3 we have defined ordinary elliptic curves as those with endomorphism algebra a quadratic imaginary field and supersingular curves as the ones with endomorphism algebra a quaternion algebra. We now want to show how the Frobenius endomorphism provides an equivalent discriminant.

Theorem 1.21 ([Deu])**.** *Let E be an elliptic curve defined over a perfect field k of characteristic p and let π be the Frobenius endomorphism. The following conditions are equivalent.*

(a) *$E[p^e] = 0$ for all $e \geq 1$.*

(b) *The dual $\hat{\pi}$ of the Frobenius endomorphism is purely inseparable.*

(c) *The trace of the Frobenius is divisible by p.*

(d) *The full endomorphism ring* $\mathrm{End}(E)$ *is an order in a quaternion algebra.*

and we say that E is supersingular. If the equivalent conditions above do not hold, then all of the following conditions hold.

(a) *$E[p^e] = \mathbb{Z}/p^r\mathbb{Z}$ for all $e \geq 1$.*

(b) *The dual $\hat{\pi}$ of the Frobenius endomorphism is separable.*

(c) *The trace of the Frobenius is coprime to p.*

(d) *The full endomorphism ring* End(E) *is an order in a quadratic imaginary field.*

And we say that E is ordinary.

Proof. See [Sil1, Th. V.3.1]. □

1.2.5 Complex lattices and elliptic curves

In the literature there are many different definitions of a lattice. We are interested in the following.

Definition. Let V be a vector space of finite dimension n over a field K of characteristic 0. A lattice Λ of V is an additive subgroup $\Lambda \subseteq V$ isomorphic to $\mathbb{Z}^n$ and containing a K-basis of V

Example. Let $d \in \mathbb{Z}$ be a non-square. The field $\mathbb{Q}(\sqrt{d})$ is a quadratic extension of $\mathbb{Q}$ and can be seen as a vector space over the rationals with basis $\{1, \sqrt{d}\}$. Then $\mathbb{Z}[\sqrt{d}]$ is a lattice in $\mathbb{Q}(\sqrt{d})$. More generally, for every element $\alpha \in \mathbb{Q}(\sqrt{d}) \setminus \mathbb{Q}$, $\Lambda_\alpha = \mathbb{Z} + \alpha\mathbb{Z}$ is a lattice.

Example. The previous definition works for a generic number field F over $\mathbb{Q}$. By definition this means that F is a vector space over $\mathbb{Q}$ of dimension $\dim_{\mathbb{Q}} F = [F : \mathbb{Q}]$, the degree of the extension. If $\{\alpha_1, \ldots, \alpha_d\}$ is a $\mathbb{Q}$-basis of F, then $\Lambda_{\alpha_1,\ldots,\alpha_d} = \alpha_1\mathbb{Z} + \ldots + \alpha_d\mathbb{Z}$ is a lattice of F.
It is worth recalling that in case of number fields the two conditions $\Lambda \simeq \mathbb{Z}^n$ suffices to define a lattice as any $\mathbb{Z}$-basis is also $\mathbb{Q}$-linearly independent.

We consider now $V = \mathbb{C}$ as a two-dimensional vector space over $\mathbb{R}$. A lattice of $\mathbb{C}$ is

$$\Lambda_{\omega_1,\omega_2} = \omega_1\mathbb{Z} + \omega_2\mathbb{Z} \qquad \text{such that } \omega_2/\omega_1 \notin \mathbb{R}$$

Definition. The analytic quotient group $\mathbb{C}/\Lambda$ is called a complex torus.

These objects are compact complex manifolds and Riemann showed they are also abelian varieties. In fact, they are elliptic curves.

Definition. Let Λ be a lattice. An elliptic function is a meromorphic function defined on the complex torus. For a meromorphic function $f : \mathbb{C} \to \mathbb{C}$ to be defined on $\mathbb{C}/\Lambda$ we need it to be invariant under translation by points of the lattice: for all $z \in \mathbb{C}$

$$\begin{cases} f(z+\omega_1) = f(z), \\ f(z+\omega_2) = f(z) \end{cases}$$

Elliptic functions are also called doubly periodic functions.
For more details on elliptic functions one can refer to [Sil1, Ch. VI] or [Lan2, Ch. I]

The most important example of elliptic function is represented by the Weierstrass $\wp$ function

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{\omega\in\Lambda \\ \omega\neq 0}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right)$$

by expanding it as a power series and rearranging the terms we obtain

$$\wp(z) = \frac{1}{z^2} + \sum_{k=1}^{+\infty} (2k+1) G_{2k+2}(\Lambda) z^{2k}$$

where

$$G_k(\Lambda) = \sum_{\substack{\omega\in\Lambda \\ \omega\neq 0}} \frac{1}{\omega^k}$$

is the Eisenstein series of weight k for Λ.

Proposition 1.22. *The Eisenstein series $G_k(\Lambda)$ is absolutely convergent for $k > 1$. Note that $G_{2k+1}(\Lambda) = 0$ since $\pm\omega \in \Lambda$.*

The importance of the Weierstrass $\wp$-function is multifold. On the one hand it provides generators for the function field of the complex torus

Theorem 1.23. *The field of elliptic functions for Λ is $\mathbb{C}\left(\wp(z), \wp'(z)\right)$. In other words, every elliptic function on Λ is a rational function of $\wp(z)$ and $\wp'(z)$*

On the other hand, the relation between $\wp$ and $\wp'$ yields an explicit connection between complex tori and the function field of elliptic curves.

Theorem 1.24 ([Sil1, Th. VI.3.5 and VI.3.6])**.** *For all $z \in \mathbb{C} \setminus \Lambda$*

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4(\Lambda)\wp(z) - 140G_6(\Lambda)$$

We set $g_2 = g_2(\Lambda) = 60G_4(\Lambda)$ and $g_3 = g_3(\Lambda) = 140G_6(\Lambda)$. We have an isomorphism

$$\frac{\mathbb{C}[x,y]}{(y^2 - 4x^3 + g_2x + g_3)} \longrightarrow \mathbb{C}\left[\wp, \wp'\right]$$

and the curve $y^2 = 4x^3 - g_2x - g_3$ is an elliptic curve.

We therefore have maps

$$\begin{tikzcd}
\{\text{Lattices of } \mathbb{C}\} \arrow[r] & \{\text{Complex tori}\} \arrow[r] & \{\text{Elliptic curves over } \mathbb{C}\} \\
\Lambda \arrow[r, mapsto] & \mathbb{C}/\Lambda \arrow[r, mapsto] & E_\Lambda : y^2 = x^3 - g_2(\Lambda)x - g_3(\Lambda) \\
 & z \arrow[r, mapsto] & (\wp(z), \wp'(z))
\end{tikzcd}$$

An equivalence of categories

We show now that the correspondences given above extend to an equivalence of categories.

Definition. We say that two complex lattices Λ, Λ' are homothetic if they differ by a scalar multiplication, i.e., $\Lambda = \alpha\Lambda'$ for some $\alpha \in \mathbb{C}^\times$; we write $\Lambda \sim \Lambda'$. Geometrically, this corresponds to being equivalent up to a rotation and a dilation.

Let us consider a lattice $\Lambda = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ with basis $\{\omega_1, \omega_2\}$ such that $\operatorname{Im}(\omega_1/\omega_2) > 0$, a condition usually referred to as positive orientation [Sil2, §1.1] (otherwise we simply switch them). We can normalize this basis

$$\frac{1}{\omega_1}\Lambda = \mathbb{Z} + \frac{\omega_2}{\omega_1}\mathbb{Z} \qquad \text{where } \tau = \frac{\omega_2}{\omega_1}$$

and obtain a map from the Poincaré upper half plane $\mathbb{H} = \{z \in \mathbb{C} \mid \operatorname{Im}(z) > 0\}$ to the set of lattices:

$$\begin{aligned} \mathbb{H} &\longrightarrow \mathcal{L} = \{\text{Lattices in } \mathbb{C}\} \\ \tau &\longrightarrow \Lambda_\tau = \mathbb{Z} + \tau\mathbb{Z} \end{aligned}$$

This gives a cleaner description of complex lattices as now they are represented by a single complex value. Homotheties of lattices are now described by the following theorem

Lemma 1.25 ([Sil2, Lemma 1.2])**.** *Let $\tau_1, \tau_2 \in \mathbb{H}$. Λ_{τ_1} is homothetic to Λ_{τ_2} if and only if*

$$\tau_2 = \gamma \cdot \tau_1 = \frac{a\tau + b}{c\tau + d} \qquad \text{for some } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

This means that the quotient of the upper half plane under the action by $\mathrm{SL}_2(\mathbb{Z})$, is the moduli space of lattices up to homothety

$$\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} \xrightarrow{1:1} \mathcal{L}/\mathbb{C}^\times$$

Homotheties of lattices translate to isomorphisms of corresponding elliptic curves [Sil1, Cor. VI.4.1.1] meaning we have an injection

$$\mathcal{L}/\mathbb{C}^\times \hookrightarrow \mathcal{E}\ell\ell_{\mathbb{C}} = \{\text{Elliptic curves over } \mathbb{C}\}/\simeq$$

and the Uniformization Theorem [Sil2, §I.4] asserts that this is indeed a bijection.

We get one to one correspondences between the following sets

$$\begin{tikzcd}
\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} \arrow[r] & \mathcal{L}/\mathbb{C}^\times \arrow[r] & \mathcal{E}\ell\ell_{\mathbb{C}} \arrow[r] & \mathbb{C} \arrow[lll, bend right, "\text{Uniformization Theorem}"'] \\
\tau \arrow[r, mapsto] & {[\Lambda_\tau]} \arrow[r, mapsto] & {[E_\Lambda]} \arrow[r, mapsto] & j(E_\Lambda)
\end{tikzcd}$$

1.3 Modular curves

The takeaway of last section is that elliptic curves are parametrized by points on the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. In this section we will develop this further and we will study quotients of the upper half plane in a geometric fashion.

1.3.1 The modular group

First of all, we note that $-I \in \mathrm{SL}_2(\mathbb{Z})$ acts trivially on $\mathbb{H}$.

Definition. The modular group, denoted $\Gamma(1)$, is the quotient group

$$\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$$

The importance of this distinction comes from the following

Corollary 1.26. *$\Gamma(1)$ acts faithfully on $\mathbb{H}$.*

The group $\Gamma(1)$ contains two particularly important elements, which are usually denoted

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

acting on $\mathbb{H}$ by

$$S(\tau) = -\frac{1}{\tau} \qquad T(\tau) = \tau + 1$$

Geometrically, S represents inversion on the unit circle followed by reflection with respect to the imaginary axis, while T represents a unit translation to the right.

Note that the elements S and ST have finite order,

$$S^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 = 1 \qquad (ST)^3 = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}^3 = 1$$

Proposition 1.27. *Let $\mathcal{F} \subset \mathbb{H}$ be the subset of $\mathbb{H}$*

$$\mathcal{F} = \left\{ \tau \in \mathbb{H} \mid |\tau| > 1 \text{ and } |Re(\tau)| \leq \frac{1}{2} \right\}$$

(a) *Let $\tau \in \mathbb{H}$. Then there is a $\gamma \in \Gamma(1)$ such that $\gamma\tau \in \mathcal{F}$.*

(b) *Suppose that both τ and $\gamma\tau$ are in $\mathcal{F}$ for some $\gamma \in \Gamma(1), \gamma \neq 1$. Then one of the following is true:*

i. *$Re(\tau) = -\frac{1}{2}$ and $\gamma\tau = \tau + 1$;*
ii. *$Re(\tau) = \frac{1}{2}$ and $\gamma\tau = \tau - 1$;*
iii. *$|\tau| = 1$ and $\gamma\tau = -\frac{1}{\tau}$.*

(c) *Let $\tau \in \mathcal{F}$, and let*

$$\mathrm{Stab}_{\Gamma(1)}(\tau) = \{\gamma \in \Gamma(1) | \gamma\tau = \tau\}$$

the stabilizer of τ. Then

$$\mathrm{Stab}_{\Gamma(1)} = \begin{cases} \{1, S\} & if\ \tau = i \\ \{1, ST, (ST)^2\} & if\ \tau = \rho = e^{2\pi i/3} \\ \{1, TS, (TS)^2\} & if\ \tau = -\overline{\rho} = e^{2\pi i/6} \\ \{1\} & otherwise \end{cases}$$

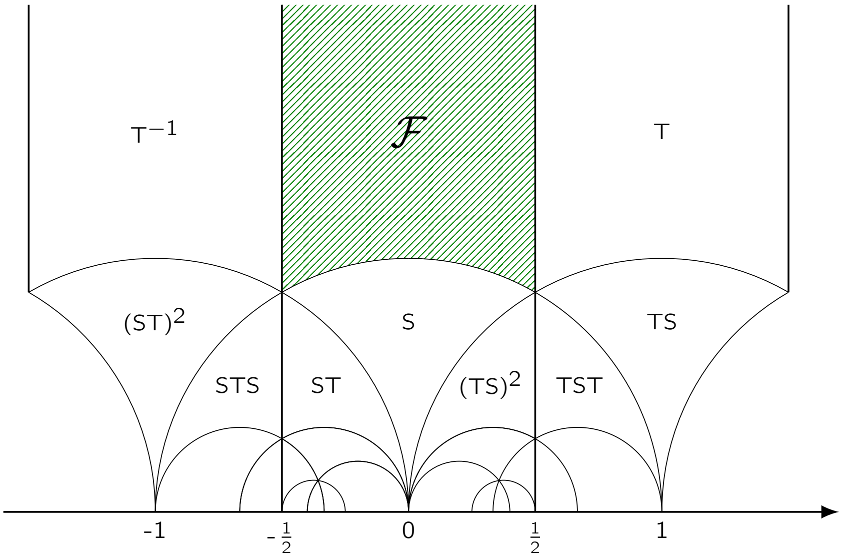


Figure 1.2 – The region $\mathcal{F}$ and some of its images under the action by $\Gamma(1)$.

As a consequence, we obtain the algebraic description of $\Gamma(1)$.

Corollary 1.28. *The modular group $\Gamma(1)$ is generated by the matrices S and T.*

Remark. In fact, $\Gamma(1)$ is the free product of its subgroups $\langle S\rangle$ and $\langle ST\rangle$ of orders 2 and 3.

Corollary 1.29. *The modular group has the presentation:*

$$\Gamma(1) \simeq \langle S, T \mid S^2 = I \ , \ (ST)^3 = I\rangle$$

The curve attached to $\Gamma(1)$

The modular group acts on the upper half plane and its quotient classifies homothety classes of lattices or, equivalently, isomorphism classes of elliptic curves. As a geometric object it is a sphere with one point missing as it corresponds to its fundamental domain with the identification of Proposition 1.27.

Definition. The extended upper-half plane $\mathbb{H}^*$ is the union

$$\mathbb{H}^* = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$$

The points in $\mathbb{P}^1(\mathbb{Q})$ are called cusps.

The action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tau = \frac{a\tau + b}{c\tau + d}$$

extends naturally to an action on $\mathbb{P}^1(\mathbb{Q})$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

Therefore, $\mathrm{SL}_2(\mathbb{Z})$ acts on the extended upper half plane.

Definition. We define $Y(1) = \Gamma(1)\backslash\mathbb{H}$ and $X(1) = \Gamma(1)\backslash\mathbb{H}^*$.

$X(1)$ is now a sphere as we add the cusps (which are all equivalent under $\Gamma(1)$) to fill the missing point of $Y(1)$; it is, in fact, the projective closure of $Y(1)$. Therefore, $X(1)$ is a projective curve. We can endow it with the quotient topology and find that $X(1)$ is a connected compact Hausdorff space, see [Sil2, Prop. I.2.4]. More than that, it has a complex structure, i.e., an open covering $\{U_i\}$ together with compatible homeomorphisms $\psi_i : U_i \to \mathbb{C}$.

Theorem 1.30. *$X(1)$ is a compact Riemann surface of genus 0.*

The details on the topology and the atlas on $X(1)$ can be found in [Sil2, §I.2].

1.3.2 Congruence subgroups

We have seen that the group $\mathrm{SL}_2(\mathbb{Z})$ has a well defined action on the upper-half plane $\mathbb{H}$. Any of its subgroups inherits the same action. We will study a particular class of subgroups of $\mathrm{SL}_2(\mathbb{Z})$.

Definition. We denote $\Gamma(N)$ the kernel of the reduction map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. This is

$$\Gamma(N) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \gamma \equiv I \mod N\} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,\middle|\, a \equiv d \equiv 1,\ b \equiv c \equiv 0 \mod N \right\}$$

As a kernel of a group morphism, it is normal in $\mathrm{SL}_2(\mathbb{Z})$. We call $\Gamma(N)$ the principal congruence subgroup of level N. A congruence subgroup is any subgroup of $\mathrm{SL}_2(\mathbb{Z})$ that contains $\Gamma(N)$ for some N.
Some important examples of congruence subgroups are

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,\middle|\, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \mod N \right\}$$

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,\middle|\, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \mod N \right\}$$

Note that $\Gamma(N)$ is of finite index in $\mathrm{SL}_2(\mathbb{Z})$ and therefore any congruence subgroup also has finite index.

The curve attached to Γ

As we did for the modular curve $X(1)$, we can quotient the extended upper half plane by the action of any congruence subgroup Γ.

Definition. The affine modular curve over $\mathbb{C}$ attached to Γ is an algebraic curve $Y(\Gamma)$ whose complex points are identified with $\Gamma\backslash\mathbb{H}$ with its natural Riemann surface structure.

Definition. The complete modular curve over $\mathbb{C}$ attached to Γ is an algebraic curve $X(\Gamma)$ whose complex points are identified with $\Gamma\backslash\mathbb{H}^*$ equipped with the following Riemann surface structure

(∞) At ∞, $\mathrm{Stab}_\Gamma(\infty) = \left\langle \pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \right\rangle$ where $n \in \mathbb{Z}_{\geq 1}$ is called the width of infinity in $X(\Gamma)$.

(x) At $x = \frac{a}{c} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}(\infty)$ then $A^{-1}\mathrm{Stab}_\Gamma(\frac{a}{c})A = \left\langle \pm \begin{pmatrix} 1 & n_x \\ 0 & 1 \end{pmatrix} \right\rangle$

1.3.3 Simplicial complexes

Modular curves are Riemann surfaces. We provide here a topological discussion expanding on the idea that the tessellation of $\mathbb{H}$ given in Figure 1.2 provides a simplicial decomposition of the Riemann surface $\Gamma\backslash\mathbb{H}^*$; in particular we show how to compute their genus.

Definition. Let X be a compact orientable surface. A simplicial complex on X is a decomposition of X as a union of simplicies of dimension $0, 1, 2$ such that the intersection of two different 2-dimensional simplicies is either empty or a 1-dimensional simplex and, likewise for the intersection of 1-dimensional simplicies.

We set

$$\begin{aligned} \#(\text{2-dimensional simplicies}) &= F \\ \#(\text{1-dimensional simplicies}) &= E \\ \#(\text{0-dimensional simplicies}) &= V \end{aligned}$$

Definition. The quantity $F - E + V$ is independent of the choice of the complex and it is called the Euler characteristic of X

$$\chi(X) = F - E + V$$

Remark. $\chi(X) = 2 - 2g(X)$ where $g(X)$ is the genus of the curve.

Let $f : X_1 \to X_2$ be a branched covering of surface. If $P \in X_1$, the ramification degree at P is the degree of $f_{|\mathcal{U}_P}$ where $\mathcal{U}_P$ is a small punctured neighborhood of P.

Remark. $e_P = 1$ for "almost all" $P \in X_1$.

$$\sum_{P \in f^{-1}(Q)} e_p = \deg(f) = d \qquad \text{for all } Q \in X_2$$

We call $\mathcal{E}$ the set of all the critical points $P \in X_1$ for which $e_P > 1$ and $\mathcal{B} = f(\mathcal{E})$.

Theorem 1.31 (Riemann-Hurwitz)**.** *Let $f : X_1 \to X_2$ be a branched covering of compact orientable surfaces. Then*

$$\chi(X_1) = d \cdot \chi(X_2) - \sum_{P \in X_1} (e_p - 1)$$

Remark. Observe that

$$\sum_{P \in X_1} (e_p - 1) = \sum_{\substack{P \in f^{-1}(Q) \\ Q \in X_2}} (e_p - 1) = \sum_{Q \in \mathcal{B}} (d - \# f^{-1}(Q))$$

Corollary 1.32. *Assume $g(X_2) = 0$ then*

$$g(X_1) = \frac{1}{2}\left(2 + \sum_{Q \in \mathcal{B}} (d - \# f^{-1}(Q)) - 2d\right)$$

Now we consider the map $f : X(\Gamma) \longrightarrow X(1) = \mathrm{SL}_2(\mathbb{Z}) \backslash \mathbb{H}^*$.

Proposition 1.33. *The following properties hold*

i. $d = \#(\Gamma \backslash \mathrm{SL}_2(\mathbb{Z}))$

ii. $B_f \subseteq \{\rho, i, \infty\}$

iii. $f^{-1}(i) = \Gamma \backslash \mathrm{SL}_2(\mathbb{Z}) / \langle S \rangle$

iiii. $f^{-1}(\rho) = \Gamma \backslash \mathrm{SL}_2(\mathbb{Z}) / \langle ST \rangle$

iiiii. $f^{-1}(\infty) = \Gamma \backslash \mathrm{SL}_2(\mathbb{Z}) / \langle T \rangle$

The problem of computing the genus of $X(\Gamma)$ is now reduced to computing

i. $\# \Gamma \backslash \mathrm{SL}_2(\mathbb{Z})$

iii. $\# \Gamma \backslash \mathrm{SL}_2(\mathbb{Z}) / \langle S \rangle$

iiii. $\# \Gamma \backslash \mathrm{SL}_2(\mathbb{Z}) / \langle ST \rangle$

iiiii. $\# \Gamma \backslash \mathrm{SL}_2(\mathbb{Z}) / \langle T \rangle$

1.3.4 Elliptic points and cusps

In this section we recall some notions on the Riemann surfaces structure of modular curves which will enable us to recover an alternative form of the genus formula. We will follow [DS, Ch. 2].

We know that the ramification degree of f at a point P is a local measure encoding the multiplicity with which f takes 0 to 0 as a map in local coordinates at P. It is therefore natural to first describe local coordinates on modular curves. For simplicity, we will do it for the algebraic curve $Y(\Gamma)$ and we will refer to [DS, §2.4] for the approach at the cusps.

The curve $Y(\Gamma)$ inherits the quotient topology from the upper half plane via the projection map $\pi : \mathbb{H} \to Y(\Gamma)$ acting as $\tau \to \Gamma\tau$. At a point $\pi(\tau)$ where $t \in \mathbb{H}$ is fixed only by the identity, we can find a small enough neighborhood $\mathcal{U}$ of τ in $\mathbb{H}$ which is homeomorphic under π to its image $\pi(\mathcal{U})$ in $Y(\Gamma)$. A local inverse $\varphi : \pi(\mathcal{U}) \to \mathcal{U}$ will play the role of local coordinate map.

The same approach does not work at a point $\pi(\tau)$ for which τ has non-trivial stabilizer in $\mathrm{SL}_2(\mathbb{Z})$.

Definition. Let Γ be a congruence subgroup. We define the isotropy group of $\tau \in \mathbb{H}$ as

$$\Gamma_\tau := \mathrm{Stab}_\Gamma(\tau) = \{\gamma \in \Gamma \mid \gamma(\tau) = \tau\}$$

An elliptic point is a point with non-trivial isotropy group. By extension, the image point $\pi(\tau) \in Y(\Gamma)$ is also called elliptic.

To understand $X(\Gamma)$, it will be therefore important to understand the stabilizers Γ_τ of points τ in the upper half plane. We begin by considering the case $\Gamma = \Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$.
Let $\gamma \in \Gamma(1)$ be a non-trivial matrix such that $\gamma(\tau) = \tau$ for some $\tau \in \mathbb{H}$. This means that

$$(a\tau + b)/(c\tau + d) = \tau \;\Rightarrow\; a\tau + b = c\tau^2 + d\tau \;\Rightarrow\; c\tau^2 + (d-a)\tau - b = 0$$

$c = 0$ would imply $a = d = \pm 1$ and $b = 0$ which represents the matrix $\gamma = \pm I$. Since we supposed γ non-trivial we can exclude this case and solve the quadratic equation normally:

$$\tau = \frac{a - d \pm \sqrt{a^2 + d^2 - 2ad - 4cb}}{2c} = \frac{a - d \pm \sqrt{a^2 + d^2 + 2ad - 4}}{2c} = \frac{a - d \pm \sqrt{(a+d)^2 - 4}}{2c}$$

Since $\tau \in \mathbb{H}$ we impose $|a+d| < 2$ which gives $a+d = \pm 1$ or 0. In particular this says that the characteristic polynomials of γ ($P_\gamma(x) = x^2 + (a+d)x + 1$) can only be of the form $x^2 + x + 1$, $x^2 - x + 1$ or $x^2 + 1$. Since $x^2 + 1$ is a factor of $x^4 - 1$ and $x^2 \pm x + 1$ are factors of $x^6 - 1$, either $\gamma^2 = I$ or $\gamma^6 = I$.

Proposition 1.34. *Let $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ Let*

$$R = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \qquad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad W = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

i. *If γ has order 3, then it is conjugate to $R^{\pm 1}$.*

ii. *If γ has order 4, then it is conjugate to $S^{\pm 1}$.*

iii. *If γ has order 6, then it is conjugate to $W^{\pm 1}$.*

and, obviously,

iv. *If γ has order 2, then $\gamma = \pm I$.*

Corollary 1.35. *The elliptic points for $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$ are the points $\Gamma(1)i$ and $\Gamma(1)\rho$ where $\rho = e^{2\pi i/3}$. Thus, the modular curve $Y(1)$ has just two elliptic points.*

Note that the same result comes from the next proposition.

Proposition 1.36. *Suppose $\tau \in \mathbb{H}$. There is a natural identification of $\Gamma(1)_\tau$ with $\mathrm{Aut}(E_\tau)$.*

Corollary 1.37 (of Proposition 1.34)**.** *Let Γ be a congruence subgroup of $\Gamma(1)$. The modular curve $Y(\Gamma)$ has finitely many elliptic points and for each of those the isotropy group is finite cyclic.*

Proof. This is easily deduced from the fact that Γ_τ is a subgroup of $\Gamma(1)_\tau$. □

Thus, to each $\tau \in \mathbb{H}$, we associate a positive integer

$$h_\tau = |\{\pm I\}\Gamma/\{\pm I\}| = \begin{cases} |\Gamma|/2 & \text{If } -I \in \Gamma_\tau \\ |\Gamma| & \text{If } -I \notin \Gamma_\tau \end{cases}$$

counting the number of transformations fixing τ.

Note that $h_\tau > 1$ if and only if τ is elliptic. The integer h_τ is called the period of τ.

Remark. It is straightforward to prove that this is well defined on the modular curve $Y(\Gamma)$.

Now, to put coordinates on $Y(\Gamma)$ about an elliptic point $\pi(\tau)$ we first use the map $\delta_\tau = \begin{pmatrix} 1 & -\tau \\ 1 & -\bar{\tau} \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C})$ to take τ to 0 and its conjugate $\bar{\tau} \to \infty$.

The isotropy group of 0 after the transformation is conjugate to the one of τ and, therefore, it is cyclic of order h_τ. Since this group of fractional linear transformations fixes 0 and ∞, it consists of transformations of the form $z \to cz$ for some complex number c and, since it is cyclic it must be the group of rotations by $2\pi/h_\tau$ about 0.
When we transform a neighborhood $\mathcal{U}$ of τ via the map δ_τ, the Γ-equivalent points are separated by a fixed angle (δ_τ is straightening the neighborhoods of τ). We next wrap the sector around a disc $\mathcal{V}$ via the map $z \to z^{h_\tau}$ and we call this map ρ. Finally, we define $\psi = \rho \circ \delta_\tau$.
By the Open mapping theorem ψ is an open map.

Lemma 1.38. *The projection $\pi : \mathcal{U} \to \pi(\mathcal{U})$ and the map $\psi : \mathcal{U} \to \mathcal{V}$ identify the same points.*

Thus, there is an injection $\varphi : \pi(\mathcal{U}) \to \mathcal{V}$ commuting with π and ψ which will play the role of local coordinates.

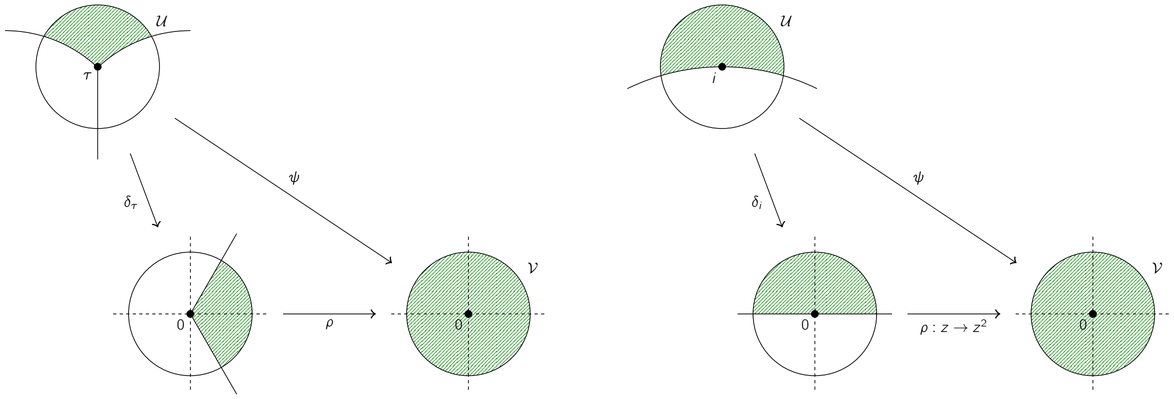


Figure 1.3 – Coordinate chart around an elliptic point and specialization to i.

computing the number of elliptic points of period 2 and 3 for the congruence subgroup $\Gamma_0(N)$.

1.3.5 Genus formula

We will now derive an equivalent formula to compute the genus of $X(\Gamma)$. We suppose again to have a non-constant holomorphic map between modular curves $f : X(\Gamma_1) \to X(\Gamma_2)$. We already know from Proposition 1.33 that

$$\deg(f) = [\{\pm I\}\Gamma_2 : \{\pm I\}\Gamma_1] = \begin{cases} [\Gamma_2 : \Gamma_1]/2 & \text{If } -I \in \Gamma_2 \text{ and } -I \notin \Gamma_1 \\ [\Gamma_2 : \Gamma_1] & \text{Otherwise} \end{cases}$$

Now, from previous sections we have the following commutative diagram where $\mathcal{U}$ is an open neighborhood of a point $\tau \in \mathbb{H}$.

$$\begin{tikzcd}
 & \mathcal{U} \arrow[r, "id"] \arrow[dl, "\delta"'] \arrow[d, "\pi_1"] & \mathcal{U} \arrow[d, "\pi_2"'] \arrow[dr, "\delta"] & \\
\delta(\mathcal{U}) \arrow[dr, "\rho_1"'] & \pi_1(\mathcal{U}) \arrow[r, "f"] \arrow[d, "\varphi_1"] & \pi_2(\mathcal{U}) \arrow[d, "\varphi_2"'] & \delta(\mathcal{U}) \arrow[dl, "\rho_2"] \\
 & \mathcal{V} \arrow[r, "f_{\text{local}}"'] & \mathcal{V} &
\end{tikzcd}$$

$$\pi_i : \mathbb{H} \longrightarrow X(\Gamma_i)$$

$$\rho_i : z \longrightarrow z^{h_i}$$

$$f_{\text{local}} \circ \rho_1 \circ \delta = \rho_2 \circ \delta$$

The local map turns out to be $q \to q^{h_2/h_1}$. Now if Γ_0j, τ is the isotropy group of τ in Γ_j for $j = 1, 2$, then the periods h_j live in $\{1, 2, 3\}$ and since the quotient is integral we have $h_1 = 1$ or $h_1 = h_2$. Therefore the ramification degree is

$$\begin{aligned} e_{\pi_1(\tau)} = h_2/h_1 &= \begin{cases} h_2 & \text{If } \tau \text{ is an elliptic point for } \Gamma_2 \text{ but not for } \Gamma_1 \\ 1 & \text{Otherwise} \end{cases} \\ &= [\{\pm I\}\Gamma_{2,\tau} : \{\pm I\}\Gamma_{1,\tau}] \end{aligned}$$

Viceversa, if $s \in \mathbb{Q} \cup \{\infty\}$ is a cusp, then $\rho_1(z) = e^{2\pi i z/h_1}$ and $\rho_2(z) = e^{2\pi i z/h_2}$ so the local map is $q \to q^{h_1/h_2}$ where $h_j = [\Gamma(1)_\infty : \{\pm 1\}\Gamma_{j,s}]$ and the ramification degree is

$$e_{\pi_1(s)} = h_1/h_1 = [\{\pm 1\}\Gamma_{2,s} : \{\pm 1\}\Gamma_{1,s}]$$

Remark. If Γ_1 is normal in Γ_2 then all points of $X(\Gamma_1)$ lying over a given point of $X(\Gamma_2)$ have the same ramification degree.

Finally, we get back to our map $X(\Gamma) \to X(1)$ for a congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$.

We denote $y_2 = \mathrm{SL}_2(\mathbb{Z})i$, $y_3 = \mathrm{SL}_2(\mathbb{Z})\rho$ and $y_\infty = \mathrm{SL}_2(\mathbb{Z})\infty$ the set of elliptic points of period 2, 3 and the cusps of $X(1)$. In the same way ϵ_2 and ϵ_3 indicate the set of elliptic points of Γ in $f^{-1}(y_2)$ and $f^{-1}(y_3)$ - of period 2 and 3 in $X(\Gamma)$ - and ϵ_∞ the set of cusps. Now,

$$d = \sum_{x \in f^{-1}(y_h)} e_x = h \overbrace{\left(|f^{-1}(y_h)| - \epsilon_h\right)}^{\substack{\text{number of}\\ \text{points that are}\\ \text{not elliptic in } \Gamma}} + 1 \cdot \epsilon_h$$

Thus, if $\nu_h = |f^{-1}(y_h)|$, then $h\nu_h - h\epsilon_h + \epsilon_h = d \implies d - \epsilon_h = h\left(\nu_h - \epsilon_h\right)$, from which

$$\sum_{x \in f^{-1}(y_h)} (e_x - 1) = h\left(|f^{-1}(y_h)| - \epsilon_h\right) + \epsilon_h + |f^{-1}(y_h)| = (h-1)\left(|f^{-1}(y_h)| - \epsilon_h\right) = \frac{h-1}{h}(d - \epsilon_h)$$

And for the cusps

$$\sum_{x \in f^{-1}(y_\infty)} (e_x - 1) = d - |f^{-1}(y_\infty)| = d - \epsilon_\infty$$

Putting all the pieces together, from Theorem 1.31, we get

Theorem 1.39 (Riemann-Hurwitz). *The genus of the modular curve $X(\Gamma)$ is*

$$g\left(X(\Gamma)\right) = 1 + \frac{d}{12} - \frac{\epsilon_2}{4} - \frac{\epsilon_3}{3} - \frac{\epsilon_\infty}{2}$$

Remark. From Corollary 1.32 and the expression for $\nu_h = \# f^{-1}(y_h)$, we obtain the same result

$$g\left(X(\Gamma)\right) = 1 - d + \frac{1}{2}\sum_{Q \in \mathcal{B}} (d - \#f^{-1}(Q)) = 1 - d + \frac{1}{2} \sum_{Q \in y_2 \cup y_3 \cup y_\infty} (d - \#f^{-1}(Q)) =$$

$$= 1 - d + \frac{d - \#f^{-1}(y_2)}{2} + \frac{d - \#f^{-1}(y_3)}{2} + \frac{d - \#f^{-1}(y_\infty)}{2} = 1 + \frac{d}{2} - \frac{\#f^{-1}(y_2)}{2} - \frac{\#f^{-1}(y_3)}{2} - \frac{\#f^{-1}(y_\infty)}{2} =$$

$$= 1 + \frac{d}{2} - \frac{d + \epsilon_2}{4} - \frac{d + 2\epsilon_3}{6} - \frac{\epsilon_\infty}{2} = 1 + \frac{d}{12} - \frac{\epsilon_2}{4} - \frac{\epsilon_3}{3} - \frac{\epsilon_\infty}{2}$$

Chapter 2

Modular towers and squares of isogenies

The goal of this chapter is to present the main properties of the modular curves associated to

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \; c \equiv 0 \bmod N \right\}$$

Since they parametrize elliptic curves together with an isogeny of degree N, we will use information on their models to study isogeny chains and commutative diagrams of isogenies.

We start by studying the action of $\Gamma_0(N)$ on the upper half plane and the resulting quotient; this provides dimension formulæ for the space of elliptic points and cusps of $X_0(N)$. We eventually shift our focus on the function field of these algebraic objects by looking for generators and relations between them. Finally, we explain how to use these parameters in the construction of isogeny structures.

2.1 Signature of the modular curve $X_0(N)$

In this first section we provide a geometric interpretation of the action of $\Gamma_0(N)$ on $\mathbb{H}$. We give all the quantities in the Riemann-Hurwitz formula 1.31 an interpretation in terms of group-theoretic actions and we use it to compute the genus of $X_0(N)$ as a branched covering of $\mathbb{P}^1$.

2.1.1 The projective line $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$

Definition. $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ denotes the projective line over $\mathbb{Z}/N\mathbb{Z}$; it is the quotient of the set

$$\left\{ (a, b) \in (\mathbb{Z}/N\mathbb{Z})^2 \,\middle|\, \gcd(a, b, N) = 1 \right\}$$

by the equivalence relation $\sim$ given by $(a, b) \sim (c, d)$ if and only if there exists $\lambda \in (\mathbb{Z}/N\mathbb{Z})^\times$ such that

$$(c, d) = \lambda (a, b)$$

Remark. Note that this is equivalent to $ad - bc \equiv 0 \bmod N$, [CS2, Lemma 6.1.4].

Lemma 2.1. *Every element of the projective line $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ has a representative $(a : b)$ with $b \mid N$, $b > 0$, and $\gcd(a, b) = 1$, and this representative is unique up to the addition of a multiple of N/b to a (which leaves a coprime to b).*

Proof. See [CS2, Lemma 6.1.7]. □

A list of representatives could thus be obtained by first considering all the divisors b of N; for each of them, we consider all the elements a' in $\mathbb{Z}/(N/b)\mathbb{Z}$ and we look for $a \in \mathbb{Z}/N\mathbb{Z}$ such that $\gcd(a, b) = 1$ and $a \equiv a' \bmod N/b$.

Example. For $N = 12$ the set $\mathbb{P}^1(\mathbb{Z}/12\mathbb{Z})$ consists of

$$\{(0:1), (1:1), (2:1), (3:1), (4:1), (5:1), (6:1), (7:1), (8:1), (9:1), (10:1), (11:1),$$
$$(1:2), (3:2), (5:2), (1:3), (2:3), (7:3), (4:3), (1:4), (5:4), (3:4), (1:6), (1:0)\}$$

As a consequence of Lemma 2.1, we have the following:

$$\#\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}) = N\prod_{p|N}\left(1+\frac{1}{p}\right)$$

2.1.2 The quotient $\Gamma_0(N)\backslash\mathrm{SL}_2(\mathbb{Z})$

From the definition we know that any element $(a : b)$ of the projective line $\#\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ has coprime coordinates: $(a, b) = 1$. Hence, we can always find two integers $c, d \in \mathbb{Z}$ such that $ad - cb = 1$. This gives us another equivalent method of representing elements of $\#\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ as matrices:

$$(v : u) \rightsquigarrow \begin{pmatrix} s & t \\ u & v \end{pmatrix}$$

We focus on the congruence subgroup

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; c \equiv 0 \mod N \right\}$$

Lemma 2.2. *We have the following isomorphism:*

$$\Gamma_0(N)\backslash\mathrm{SL}_2(\mathbb{Z}) \simeq \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$$

Proof. This is Proposition 6.3.22 in [CS2]. See also Proposition 1.43 of [Shi]. □

The action of T

The matrix $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ acts on $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ and corresponds to translations:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^m \cdot (a : b)$$

This shows that T fixes the second term. Let b be a divisor of N; an element of the form $(* : b)$ has the first coordinate differing by an element of $\mathbb{Z}/\left(N/b\right)\mathbb{Z}$. Therefore, we claim that the number of orbits of the action of the matrix T on the subset $\{(* : b)\} \subset \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ is given by

$$\varphi\left(\gcd\left(b, \frac{N}{b}\right)\right)$$

where φ is the Euler function.

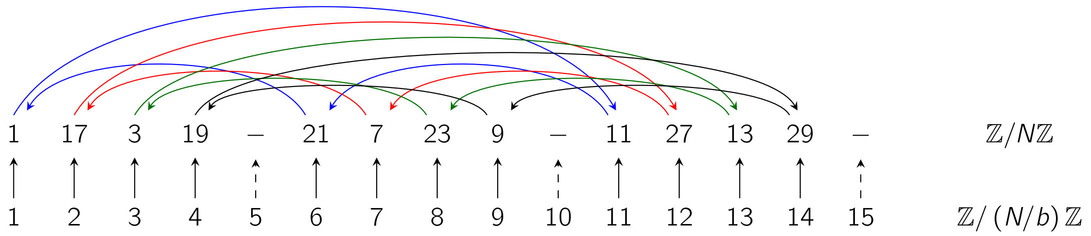
Suppose $(a : b)$ and $(a' : b)$ are in the same orbit. This implies that $a' \equiv a + kb \mod N/b$ for some $k \in \mathbb{Z}$. Hence, modulo changing the sign of k,

$$a' - a = m\frac{N}{b} + kb = \gcd(b, N/b)\left(m\frac{N}{b\gcd(b, N/b)} + k\frac{b}{\gcd(b, N/b)}\right)$$

which clearly says that $a' - a$ is a multiple of $\gcd(b, N/b)$.

The question can be turned into finding the number of elements in $\mathbb{Z}/\left(N/b\right)\mathbb{Z}$ which are less than or equal to $\gcd(b, N/b)$ and lift to an element of $\{(* : b)\} \subset \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$. Now it is clear that if $\alpha \in \mathbb{Z}/\left(N/b\right)\mathbb{Z}$ is not coprime to $\gcd(b, N/b)$ then it is not coprime neither with b nor with N/b and therefore it does not lift. Viceversa if it is coprime to $\gcd(b, N/b)$ then either it is coprime to b and therefore it determines the element $(\alpha : b)$ or it is not coprime to b but $\alpha + N/b$ is. In both cases we get an element of $\{(* : b)\} \subset \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$. This discussion shows that the number of orbits equals the number of elements in $\mathbb{Z}/\left(N/b\right)\mathbb{Z}$ which are less than or equal to $\gcd(b, N/b)$ and coprime to it and this is exactly $\varphi\left(\gcd\left(b, \frac{N}{b}\right)\right)$ as wanted.

Example. Let us look at the situation for $N = 150 = 2 \cdot 3 \cdot 5^2$, $b = 10 = 2 \cdot 5$.



As expected we have $4 = \varphi(5) = \varphi\left(\gcd\left(b, \frac{N}{b}\right)\right)$ orbits.

Thus, we have
$$\#\pi^{-1}(\infty) = \#\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/\langle T\rangle = \sum_{b|N} \varphi\left(\left(b, \frac{N}{b}\right)\right)$$

The action of S

S is the matrix
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
It acts on the left on $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot (a : b) = (b : -a) = (-b : a)$$
It is not difficult to see that S is of order 2 and it satisfies the minimal polynomial $X^2 + 1 = 0$. This implies that all its orbits are of size either 1 or 2. Orbits of order one consists of eigenvalues.

Hence, S has orbits of order one in $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ if and only if its minimal polynomial $X^2 + 1$ splits modulo N. Suppose $N = p_1^{e_1} p_2^{e_2} \dots p_n^{e_n}$, the Chinese Remainder Theorem implies we can solve a polynomial $f(x)$ over $\mathbb{Z}/N\mathbb{Z}$ by first solving it modulo $\mathbb{Z}/p_i^{e_i}\mathbb{Z}$ and then combining the roots together to find a solution modulo N.
$$\frac{\mathbb{Z}}{N\mathbb{Z}} \xrightarrow{\sim} \frac{\mathbb{Z}}{p_1^{e_1}\mathbb{Z}} \times \frac{\mathbb{Z}}{p_2^{e_2}\mathbb{Z}} \times \dots \times \frac{\mathbb{Z}}{p_n^{e_n}\mathbb{Z}}$$
Let us have a look at $P_S(X) = X^2 + 1$ modulo p. We study the roots of this polynomial in $\mathbb{Z}/p\mathbb{Z}$ and we try to see whether they lift or not to solutions in $\mathbb{Z}/p^k\mathbb{Z}$. In particular we study if $P_S(X)$ has solutions in $\mathbb{Z}_p$.

Lemma 2.3. *If $p \neq 2$ and $a \not\equiv 0 \mod p$, the equation $X^2 - a = 0$ has a solution in $\mathbb{Z}_p$ if and only if it has a solution in $\mathbb{F}_p$.*

Proof. This is an immediate consequence of Hensel's Lemma. □

Thus, for an odd prime p, $P_S(X)$ has a solution in $\mathbb{Z}/p^k\mathbb{Z}$ for all k if and only if it has a solution in $\mathbb{Z}/p\mathbb{Z}$. We know that $X^2 - a$ has a solution in $\mathbb{F}_p$ if and only if
$$\left(\frac{a}{p}\right) = 1$$
In particular $X^2 + 1$ has a root in $\mathbb{Z}_p$ (for $p \neq 2$) if and only if
$$\left(\frac{-1}{p}\right) = 1 \iff p \equiv 1 \mod 4$$
In conclusion, supposing $2 \nmid N$, the action of S on $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ has either 0 or 2 orbits of size 1 depending on whether the polynomial $X^2 + 1$ is irreducible or not modulo all the primes diving N.

We can express this condition saying that

$$\#\text{orbits of size } 1 = \prod_{p|N}\left(1+\left(\frac{-1}{p}\right)\right)$$

It remains to study the case $p=2$. Clearly X^2+1 has one solution in $\mathbb{Z}/2\mathbb{Z}$, namely 1, but this does not lift to any solution in $\mathbb{Z}/4\mathbb{Z}$. Also (see [Conr])

Lemma 2.4. *If $u \in \mathbb{Z}_2^\times$, then u is a square in $\mathbb{Q}_2$ if and only if $u \equiv 1 \mod 8\mathbb{Z}_2$.*

Since $-1 \not\equiv 1 \mod 8$, we conclude that $P_S(X)$ has no solutions in $\mathbb{Z}/N\mathbb{Z}$ if $4 \mid N$.

In conclusion,

$$\#\pi^{-1}(i) = \#\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/\langle S\rangle = \frac{\#\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}) - \#\text{orbits of size } 1}{2} + \#\text{orbits of size } 1 =$$
$$= \frac{1}{2}\cdot\begin{cases} N\prod_{p|N}\left(1+\frac{1}{p}\right) & \text{if } 4 \mid N \\ N\prod_{p|N}\left(1+\frac{1}{p}\right) + \prod_{p|N}\left(1+\left(\frac{-1}{p}\right)\right) & \text{if } 4 \nmid N \end{cases}$$

where $(-1/p)$ is 1 if $p \equiv 1 \bmod 4$, -1 if $p \equiv -1 \bmod 4$ and 0 if $p=2$.

Remark. Note that $(-1/p)$ is the classical Legendre symbol for every odd prime. For $p=2$, we use the same notation following [DS] but we define it to be $(-1/2)=0$. This does not correspond neither to the Legendre symbol (which is not defined for $p=2$) nor to the Kronecker symbol for which $(-1|2)=1$.

Example. Let $N=10$; the orbits of S in $\mathbb{P}^1(\mathbb{Z}/10\mathbb{Z})$ are

$$\begin{tikzcd}
(0:1) \arrow[d] & (1:1) \arrow[d] & (2:1) \arrow[d] & (3:1) \arrow[d] & (4:1) \arrow[d] & (5:1) \arrow[d] & (6:1) \arrow[d] & (7:1) \arrow[d] & (8:1) \arrow[d] & (5:2) \arrow[d] \\
(1:0) \arrow[d] & (9:1) \arrow[d] & (9:2) \arrow[d] & (3:1) & (7:2) \arrow[d] & (1:5) \arrow[d] & (3:2) \arrow[d] & (7:1) & (1:2) \arrow[d] & (2:5) \arrow[d] \\
(0:1) & (1:1) & (2:1) & & (4:1) & (5:1) & (6:1) & & (8:1)) & (5:2)
\end{tikzcd}$$

As expected we found

$$\prod_{p|10}\left(1+\left(\frac{-1}{p}\right)\right) = \left(1+\left(\frac{-1}{2}\right)\right)\cdot\left(1+\left(\frac{-1}{5}\right)\right) = 1\cdot 2 = 2$$

orbits of size 1 and

$$\frac{1}{2}\cdot\left(10\prod_{p|10}\left(1+\frac{1}{p}\right) - \prod_{p|10}\left(1+\left(\frac{-1}{p}\right)\right)\right) = \frac{1}{2}\cdot\left(10\cdot\frac{3}{2}\cdot\frac{6}{5}-2\right) = 8$$

orbits of size 2.

The action of ST

The matrix

$$ST = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$$

acts on the left on $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ by sending

$$\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\cdot(a:b) = (b:-a-b) = (b:-(a+b))$$

ST has order 3 and satisfies the minimal polynomial $P_{ST}(X) = X^2+X+1$. As we did for S we study the splitting behavior of this polynomial modulo all the prime divisors of N and we see if they lift to general solution modulo prime powers.

First of all, we have the following lemma:

Lemma 2.5. *The polynomial X^2+X+1 has a solution in $\mathbb{F}_p$ if and only if -3 is a square* mod p.

Proof. Consider the quadratic extension of $\mathbb{Q}(\sqrt{-3})$. The ring of integers of $K=\mathbb{Q}(\sqrt{-3})$ is $\mathcal{O}_K=\mathbb{Z}[\omega]$ where the primitive element $\omega=\left(-1+\sqrt{-3}\right)/2$ satisfies the minimal polynomial $f(X)=X^2+X+1$. Now let $p\in\mathbb{Z}$ be a prime number; a prime $\mathfrak{p}$ ideal of $\mathcal{O}_K$ is a factor of $p\mathcal{O}_K$ if and only if $p\mathcal{O}_K\subseteq\mathfrak{p}$.

Hence there is a bijection between the set of prime factors of $p\mathcal{O}_K$ and the prime ideals of $\mathcal{O}_K$ containing $p\mathcal{O}_K$. But this last set is in bijection with the set of prime ideals of $\mathcal{O}_K/p\mathcal{O}_K$ and now

$$\frac{\mathcal{O}_K}{p\mathcal{O}_K}\simeq\frac{\left(\frac{\mathbb{Z}[X]}{(f(X))}\right)}{\left(\frac{(p)\mathbb{Z}[X]+(f(X))\mathbb{Z}[X]}{(f(X))}\right)}\simeq\frac{\mathbb{Z}[X]}{(p)\mathbb{Z}[X]+(f(X))\,\mathbb{Z}[X]}\simeq\frac{\left(\frac{\mathbb{Z}[X]}{(p)\mathbb{Z}[X]}\right)}{\left(\frac{(p)\mathbb{Z}[X]+(f(X))\mathbb{Z}[X]}{(p)\mathbb{Z}[X]}\right)}\simeq\frac{\mathbb{F}_p[X]}{\left(\overline{f(X)}\right)}$$

In the last isomorphism we use the canonical projection map

$$\begin{aligned}\pi:\mathbb{Z}[X]&\longrightarrow\mathbb{F}_p[X] \qquad\qquad \ker(\pi)=(p)\mathbb{Z}[X]\\ f(X)&\longrightarrow\overline{f(X)}\end{aligned}$$

Then

$$\left\{\begin{matrix}\text{prime ideals of } \mathcal{O}_K\\ \text{containing } (p)\mathcal{O}_K\end{matrix}\right\}\longleftrightarrow\left\{\text{prime ideals of } \frac{\mathbb{F}_p[X]}{\left(\overline{f(X)}\right)}\right\}\longleftrightarrow\left\{\begin{matrix}\text{prime ideals of } \mathbb{F}_p[X]\\ \text{containing } \left(\overline{f(X)}\right)\end{matrix}\right\}$$

and since $\mathbb{F}_p[X]$ is a P.I.D., then this last set is in bijection with the monic irreducible factors of $\overline{f(X)}$.

Thus, the splitting behavior of the polynomial X^2+X+1 modulo p corresponds to the splitting behavior of the ideal $(p)_K$.

To conclude it only remains to observe that the splitting behavior of a prime (p) in any quadratic field L is given (see [Cox, Prop. 5.16])by the Legendre symbol

$$\left(\frac{\Delta_L}{p}\right)\qquad\qquad \Delta_L=\operatorname{disc}(\mathcal{O}_L)$$

In our case $\Delta_K=-3$ which concludes the proof. □

In particular, the number of roots of $P_{ST}(X)$ mod p is given by $1+\left(\frac{-3}{p}\right)$. It remains to observe that a root lifts to a solution modulo p^k.

Lemma 2.6 (Hensel's Lemma)**.** *If $f(X)\in\mathbb{Z}_p[X]$ and $a\in\mathbb{Z}_p$ satisfies*

$$f(a)\equiv 0 \bmod p \qquad f'(a)\not\equiv 0 \bmod p$$

then there is a unique $\alpha\in\mathbb{Z}_p$ such that $f(\alpha)=0$ and $\alpha\equiv a$ mod p.

The minimal polynomial $P_{ST}(X)$ has prime derivative $2X+1$ and $2X+1$ divides X^2+X+1 in $\mathbb{F}_p[X]$ if and only if $p=3$. Indeed,

$$X^2+X+1\equiv(2X+1)(aX+b)\equiv 2aX^2+(2b+a)X+b \mod p$$

implies

$$\begin{cases}b\equiv 1 \mod p\\ 2b+a\equiv 1 \mod p\\ 2a\equiv 1 \mod p\end{cases}\Leftrightarrow\begin{cases}b\equiv 1 \mod p\\ a\equiv -1 \mod p\\ 2a\equiv 1 \mod p\end{cases}\Leftrightarrow\begin{cases}b\equiv 1 \mod p\\ a\equiv -1 \mod p\\ -2\equiv 1 \mod p\end{cases}$$

and the third equation is true if and only if $p=3$.
Hence, all the solution that we find can be lifted to general solutions modulo p^k for $p\neq 3$.

It only remains to study the case $p=3$. $P_{ST}(X)$ mod 3 has a root with multiplicity 2, namely 1. But this does not lift to any solution mod 9.

x	0	1	2	3	4	5	6	7	8
x^2	0	1	4	0	7	7	0	4	1
x^2+x+1	1	3	7	4	3	4	7	3	1

More in general, suppose we have a solution α of X^2+X+1 in $\mathbb{Z}/p^i\mathbb{Z}$; in order to have it lifting to a root in $\mathbb{Z}/p^{i+1}\mathbb{Z}$ we have to solve $\left(\alpha+kp^i\right)^2+\left(\alpha+kp^i\right)+1\equiv 0 \mod p^{i+1}$. This gives

$$0\equiv(\alpha^2+\alpha+1)+(2\alpha+1)kp^i+k^2p^{2i}\equiv(\alpha^2+\alpha+1)+(2\alpha+1)kp^i \mod p^{i+1}$$

If $p\neq 3$, then $2\alpha+1$ is coprime to p and therefore we are left with a linear equation in k:

$$(\alpha^2+\alpha+1)+(2\alpha+1)p^ik\equiv 0 \mod p^{i+1}$$

and since α is a solution modulo p^i, then the latter is equivalent to the following equation

$$1+(2\alpha+1)k\equiv 0 \mod p$$

which has always solution since $2\alpha+1$ is invertible.
If $p=3$ the situation changes since $\alpha=1$ and $2\alpha+1=3=p$. This gives

$$0\equiv(\alpha^2+\alpha+1)+(2\alpha+1)p^ik\equiv(\alpha^2+\alpha+1)+p^{i+1}k\equiv(\alpha^2+\alpha+1) \mod p^{i+1}$$

and we check easily that 1 is not a solution in $\mathbb{Z}/9\mathbb{Z}$.

Remark. More concisely, $f(x)$ and $f(x')$ are clearly coprime if $p\neq 3$ since $p\nmid \operatorname{disc}(f(x))=-3$. Further, the impossibility of lifting a root to $\mathbb{Z}/9\mathbb{Z}$ comes from the fact that 3 ramifies and therefore the root lies in a quadratic extension of $\mathbb{Z}_3$.

In conclusion, the number of orbits of size 1 of ST is given by

$$\begin{cases} 0 & \text{if } 9\mid N\\ \prod_{p\mid N}\left(1+\left(\frac{-3}{p}\right)\right) & \text{if } 9\nmid N\end{cases}$$

where $(-3/p)$ is 1 if $p\equiv 1 \bmod 3$, -1 if $p\equiv -1 \bmod 3$ and 0 if $p=3$.

Remark. As in the previous section, $(-3/p)$ represents the Legendre symbol for any odd prime p. In this situation, however, the case of $p=2$ does fit with the case of $p\equiv -1 \bmod 3$ and also with the Kronecker symbol $(-3|2)=-1$.

All the other orbits are of size 3 (since this is the order of ST) and therefore,

$$\begin{aligned}\#\pi^{-1}(\rho)=\#\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/\langle ST\rangle &= \frac{\#\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})-\#\text{orbits of size 1}}{3}+\#\text{orbits of size 1} =\\ &=\frac{1}{3}\cdot\begin{cases} N\prod_{p\mid N}\left(1+\frac{1}{p}\right) & \text{if } 9\mid N\\ N\prod_{p\mid N}\left(1+\frac{1}{p}\right)+2\prod_{p\mid N}\left(1+\left(\frac{-3}{p}\right)\right) & \text{if } 9\nmid N\end{cases}\end{aligned}$$

Example. We have a look at the case $N=21$. The orbits are

$$\begin{tikzcd}
(0:1) \arrow[d] & (1:1) \arrow[d] & (2:1) \arrow[d] & (3:1) \arrow[d] & (4:1) \arrow[d] & (6:1) \arrow[d] & (7:1) \arrow[d] & (8:1) \arrow[d] & (11:1) \arrow[d] & (14:1) \arrow[d] & (15:1) \arrow[d] & (16:1) \arrow[d] \\
(20:1) \arrow[d] & (10:1) \arrow[d] & (13:3) \arrow[d] & (5:1) \arrow[d] & (4:1) & (2:7) \arrow[d] & (13:1) \arrow[d] & (2:3) \arrow[d] & (5:3) \arrow[d] & (4:3) \arrow[d] & (17:1) \arrow[d] & (16:1) \\
(1:0) \arrow[d] & (19:1) \arrow[d] & (9:1) \arrow[d] & (10:3) \arrow[d] & & (7:3) \arrow[d] & (1:7) \arrow[d] & (12:1) \arrow[d] & (18:1) \arrow[d] & (3:7) \arrow[d] & (1:3) \arrow[d] & \\
(0:1) & (1:1) & (2:1) & (3:1) & & (6:1) & (7:1) & (8:1) & (12:1) & (14:1) & (15:1) &
\end{tikzcd}$$

We have two fixed elements, namely $(4:1)$ and $(16:1)$ (the roots of X^2+X+1), and 10 orbits of size 3.

Observe that 4 comes from the roots $X=4$ in $\mathbb{Z}/7\mathbb{Z}$ and $X=1$ in $\mathbb{Z}/3\mathbb{Z}$ while 16 is the lift of $X=2$ in $\mathbb{Z}/7\mathbb{Z}$ and $X=1$ in $\mathbb{Z}/3\mathbb{Z}$.

2.1.3 Genus of the modular curves $X_0(N)$

Finally, we can assemble all the information that we found in order to compute the genus of our modular curve.

$$g(X_0(N)) = \frac{1}{2}\left(2 + \sum_{Q\in B_\pi}(d_\pi - \#\pi^{-1}(Q)) - 2d_\pi\right) =$$

$$= \frac{1}{2}\left(2 + d_\pi - \#\pi^{-1}(i) - \#\pi^{-1}(\rho) - \#\pi^{-1}(\infty)\right) =$$

$$= \begin{cases} 1 + \frac{N}{12}\prod_{p|N}\left(1+\frac{1}{p}\right) - \frac{1}{2}\sum_{d|N}\varphi\left(\left(d, \frac{N}{d}\right)\right) & \text{if } 36 \mid N \\ 1 + \frac{N}{12}\prod_{p|N}\left(1+\frac{1}{p}\right) - \frac{1}{2}\sum_{d|N}\varphi\left(\left(d, \frac{N}{d}\right)\right) - \frac{1}{3}\prod_{p|N}\left(1+\left(\frac{-3}{p}\right)\right) & \text{if } 4 \mid N,\ 9 \nmid N \\ 1 + \frac{N}{12}\prod_{p|N}\left(1+\frac{1}{p}\right) - \frac{1}{2}\sum_{d|N}\varphi\left(\left(d, \frac{N}{d}\right)\right) - \frac{1}{4}\prod_{p|N}\left(1+\left(\frac{-1}{p}\right)\right) & \text{if } 9 \mid N,\ 4 \nmid N \\ 1 + \frac{N}{12}\prod_{p|N}\left(1+\frac{1}{p}\right) - \frac{1}{2}\sum_{d|N}\varphi\left(\left(d, \frac{N}{d}\right)\right) - \sum_{a\in\{-1,-3\}}\frac{1}{|\mathrm{disc}\,(\mathcal{O}_a)|}\prod_{p|N}\left(1+\left(\frac{a}{p}\right)\right) & \text{if } 4 \nmid N,\ 9 \nmid N \end{cases}$$

$$= 1 + \frac{N}{12}\prod_{p|N}\left(1+\frac{1}{p}\right) - \frac{1}{2}\sum_{d|N}\varphi\left(\left(d, \frac{N}{d}\right)\right) - \begin{cases} 0 & \text{if } 4 \mid N \\ \frac{1}{4}\prod_{p|N}\left(1+\left(\frac{-1}{p}\right)\right) & \text{if } 9 \mid N,\ 4 \nmid N \\ \sum_{a\in\{-1,-3\}}\frac{1}{|\mathrm{disc}\,(\mathcal{O}_a)|}\prod_{p|N}\left(1+\left(\frac{a}{p}\right)\right) & \text{if } 36 \nmid N \end{cases}$$

where we recall that $(-1/2)$ is set to be 0.

In the last equation $\mathcal{O}_a$ indicates the ring of integers of the quadratic field $\mathbb{Q}\left(\sqrt{a}\right)$ which, in our case are the Gaussian and the Eisenstein integers (for $a = -1$ and $a = -3$ respectively).

The last equation (which is a combination of the previous two) has been written in the current form for layout reasons and it is a compact version of the more explicit:

$$1 + \frac{N}{12}\prod_{p|N}\left(1+\frac{1}{p}\right) - \frac{1}{2}\sum_{d|N}\varphi\left(\left(d, \frac{N}{d}\right)\right) - \frac{1}{4}\prod_{p|N}\left(1+\left(\frac{-1}{p}\right)\right) - \frac{1}{3}\prod_{p|N}\left(1+\left(\frac{-3}{p}\right)\right)$$

The genus of the modular curves $X_0(2^n)$

We are interested in the case $N = 2^n$. For the moment we forget about $n = 1$ which gives genus 0 and we suppose $4 \mid N$.

$$g(X_0(2^n)) = 1 + \frac{2^n}{12}\prod_{p|2^n}\left(1+\frac{1}{p}\right) - \frac{1}{2}\sum_{d|2^n}\varphi\left(\left(d, \frac{2^n}{d}\right)\right) - \frac{1}{3}\prod_{p|2^n}\left(1+\left(\frac{-3}{p}\right)\right) =$$

$$= 1 + \frac{2^n}{12}\left(1+\frac{1}{2}\right) - \frac{1}{2}\sum_{k=0}^{n}\varphi\left((2^k, 2^{n-k})\right) - \frac{1}{3}\left(1+\left(\frac{-3}{2}\right)\right) =$$

$$= 1 + \frac{2^n}{12}\cdot\frac{3}{2} - \frac{1}{2}\sum_{k=0}^{n}\varphi\left((2^k, 2^{n-k})\right) =$$

$$= 1 + 2^{n-3} - \begin{cases} \sum_{k=0}^{n/2-1}\varphi\left((2^k, 2^{n-k})\right) + \frac{1}{2}\varphi\left(\left(2^{\frac{n}{2}}, 2^{\frac{n}{2}}\right)\right) & \text{if } n \equiv 0 \mod 2 \\ \sum_{k=0}^{(n-1)/2}\varphi\left((2^k, 2^{n-k})\right) & \text{if } n \equiv 1 \mod 2 \end{cases}$$

where we split the sum into two symmetric parts.

$$=1+2^{n-3}-\begin{cases} \displaystyle\sum_{k=0}^{n/2-1}\varphi\left(2^k\right)+\frac{1}{2}\varphi\left(2^{\frac{n}{2}}\right) & \text{if } n\equiv 0 \mod 2\\ \displaystyle\sum_{k=0}^{(n-1)/2}\varphi\left(2^k\right) & \text{if } n\equiv 1 \mod 2\end{cases}$$

$$=1+2^{n-3}-\begin{cases} 1+\displaystyle\sum_{k=1}^{n/2-1}2^{k-1}+2^{\frac{n}{2}-2} & \text{if } n\equiv 0 \mod 2\\ 1+\displaystyle\sum_{k=1}^{(n-1)/2}2^{k-1} & \text{if } n\equiv 1 \mod 2\end{cases}$$

$$=1+2^{n-3}-\begin{cases} 2^{\frac{n}{2}-1}+2^{\frac{n}{2}-2} & \text{if } n\equiv 0 \mod 2\\ 2^{\frac{n-1}{2}} & \text{if } n\equiv 1 \mod 2\end{cases}$$

In conclusion,

$$g(X_0(2))=0 \qquad \text{and} \qquad g(X_0(2^n))=1+2^{n-3}-\begin{cases} 3\cdot 2^{\frac{n}{2}-2} & \text{if } n\equiv 0 \mod 2\\ 2^{\frac{n-1}{2}} & \text{if } n\equiv 1 \mod 2\end{cases}$$

Example. The following table describes the genus of $X_0(2^n\cdot p)$ for $p=1,3,5,7,11$ and n up to 10.

n	$g(X_0(2^n))$	$g(X_0(2^n\cdot 3))$	$g(X_0(2^n\cdot 5))$	$g(X_0(2^n\cdot 7))$	$g(X_0(2^n\cdot 11))$
0	0	0	0	0	1
1	0	0	0	1	2
2	0	0	1	2	4
3	0	1	3	5	9
4	0	3	7	11	19
5	1	9	17	25	41
6	3	21	37	53	85
7	9	49	81	113	177
8	21	105	169	233	361
9	49	225	353	481	737
10	105	465	721	977	1489

Table 2.1 – Genus of the modular curves $X_0(2^n)$, $X_0(2^n\cdot 3)$, $X_0(2^n\cdot 5)$, $X_0(2^n\cdot 7)$ and $X_0(2^n\cdot 11)$

The genus of the modular curves $X_0(3^n)$

As we did in the previous section we may assume that $9\mid N$. Therefore

$$g(X_0(3^n))=1+\frac{3^n}{12}\prod_{p|3^n}\left(1+\frac{1}{p}\right)-\frac{1}{2}\sum_{d|3^n}\varphi\left(\left(d,\frac{3^n}{d}\right)\right)-\frac{1}{4}\prod_{p|3^n}\left(1+\left(\frac{-1}{p}\right)\right)=$$

$$=1+3^{n-2}-\begin{cases} \displaystyle\sum_{k=0}^{n/2-1}\varphi\left((3^k,3^{n-k})\right)+\frac{1}{2}\varphi\left(\left(3^{\frac{n}{2}},3^{\frac{n}{2}}\right)\right) & \text{if } n\equiv 0 \mod 2\\ \displaystyle\sum_{k=0}^{(n-1)/2}\varphi\left((3^k,3^{n-k})\right) & \text{if } n\equiv 1 \mod 2\end{cases}$$

Once again we obtain a simpler expression

$$
\begin{aligned}
&=1+3^{n-2}-\begin{cases}\displaystyle\sum_{k=0}^{n/2-1}\varphi\left(3^k\right)+\frac{1}{2}\varphi\left(3^{\frac{n}{2}}\right) & \text{if } n\equiv 0 \mod 2\\ \displaystyle\sum_{k=0}^{(n-1)/2}\varphi\left(3^k\right) & \text{if } n\equiv 1 \mod 2\end{cases}\\
&=1+3^{n-2}-\begin{cases}\displaystyle 1+\sum_{k=1}^{n/2-1}2\cdot 3^{k-1}+3^{\frac{n}{2}-1} & \text{if } n\equiv 0 \mod 2\\ \displaystyle 1+\sum_{k=1}^{(n-1)/2}2\cdot 3^{k-1} & \text{if } n\equiv 1 \mod 2\end{cases}\\
&=1+3^{n-2}-\begin{cases}2\cdot 3^{\frac{n}{2}-1} & \text{if } n\equiv 0 \mod 2\\ 3^{\frac{n-1}{2}} & \text{if } n\equiv 1 \mod 2\end{cases}
\end{aligned}
$$

Example. The following table describes the genus of $X_0(3^n\cdot p)$ for $p=1,2,5,7,11$ and n up to 10.

n	$g(X_0(3^n))$	$g(X_0(3^n\cdot 2))$	$g(X_0(3^n\cdot 5))$	$g(X_0(3^n\cdot 7))$	$g(X_0(3^n\cdot 11))$
0	0	0	0	0	1
1	0	0	1	1	3
2	0	0	3	5	9
3	1	4	13	19	31
4	4	16	43	61	97
5	19	64	145	199	307
6	64	208	451	613	937
7	217	676	1405	1891	2863
8	676	2080	4267	5725	8641
9	2107	6400	12961	17335	26083
10	6400	19360	39043	52165	78409

Table 2.2 – Genus of the modular curves $X_0(3^n)$, $X_0(3^n\cdot 2)$, $X_0(3^n\cdot 5)$, $X_0(3^n\cdot 7)$ and $X_0(3^n\cdot 11)$

2.1.4 The fundamental domain of $\Gamma_0(N)$

The aim of this section is to describe the fundamental domain for a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We will mainly follow [Kil, §2.5] and [BD, §3]

Definition. Let Γ be a congruence subgroup. A fundamental domain $\mathbb{F}$ for Γ is a subset of $\mathbb{H}$ such that

(1) every $z\in\mathbb{H}$ is Γ-equivalent to a point of the closure of $\mathcal{F}$;

(2) no two distinct points in the interior of $\mathcal{F}$ are Γ-equivalent.

A well known result describes the fundamental domain for the full modular group $\mathrm{SL}_2(\mathbb{Z})$.

Theorem 2.7. *The set*

$$\mathcal{F}:=\left\{\tau\in\mathbb{H}\,\middle|\,-\frac{1}{2}\leq\mathrm{Re}(\tau)\leq 0,\ |\tau|\geq 1\right\}\cup\left\{\tau\in\mathbb{H}\,\middle|\,0<\mathrm{Re}(\tau)<\frac{1}{2},\ |\tau|>1\right\}\cup\{\infty\}$$

is a fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$.

Proof. See [SchB, §I.5] and [Sil2, §1.1]. □

More in general

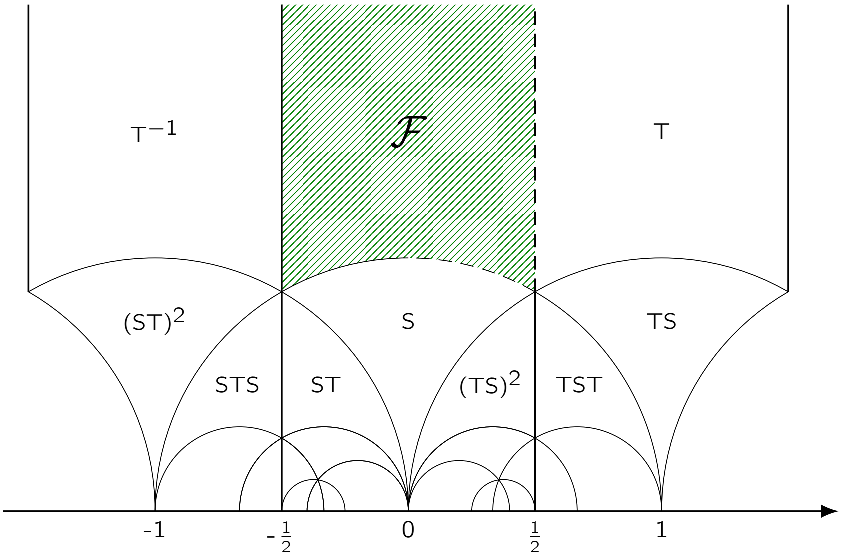


Figure 2.1 – The fundamental region for the full modular group $\mathrm{SL}_2(\mathbb{Z})$.

Proposition 2.8. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, and let R be a set of coset representatives for the quotient $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$. Then the set*

$$\mathcal{F}_\Gamma = \bigcup_{\gamma\in R} \gamma\mathcal{F}$$

has the property **(1)**. *Further, its interior satisfies also property* **(2)**.

For more information about the boundaries (where we lose the faithfulness of the action of $\mathrm{SL}_2(\mathbb{Z})$, one can refer to [SchB, §IV.4].

Proof. Let $z \in \mathbb{H}$. Because of Theorem 2.7 we know that there exists $\tau \in \mathcal{F}$ and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $z = \gamma \cdot \tau$. We also know that $\gamma = \gamma_0^{-1}\gamma'$ with $\gamma_0 \in \Gamma$ and $\gamma' \in R$. Thus,

$$\gamma_0 \cdot z = \gamma_0\gamma\tau = \gamma'\tau \in \mathcal{F}_\Gamma$$

□

Fundamental domain for $\Gamma_0(2)$

We know that

$$\Gamma_0(2)\backslash\mathrm{SL}_2(\mathbb{Z}) \simeq \mathbb{P}^1(\mathbb{Z}/2\mathbb{Z}) = \left\{ \begin{pmatrix}1&0\\0&1\end{pmatrix}, \begin{pmatrix}1&0\\1&1\end{pmatrix}, \begin{pmatrix}0&-1\\1&0\end{pmatrix} \right\}$$

The first matrix is the identity and leaves $\mathcal{F}$ unchanged. The third matrix is clearly the matrix S and therefore it adds to the fundamental domain of $\Gamma_0(2)$ the region labeled S in Figure 2.1. Finally, the second matrix can be decomposed as:

$$\begin{pmatrix}1&0\\1&1\end{pmatrix}\cdot S\cdot T^m\cdot S = \begin{pmatrix}-1&0\\m-1&-1\end{pmatrix} \implies \begin{pmatrix}1&0\\1&1\end{pmatrix}\cdot STS = I\in\mathrm{SL}_2(\mathbb{Z}) \implies \begin{pmatrix}1&0\\1&1\end{pmatrix} = ST^{-1}S$$

We immediately see that $ST^{-1}S$ sends ∞ to $(1:1)$ which is outside the region $|\mathrm{Re}(\tau)| \leq \frac{1}{2}$. Therefore we precompose it with T^{-1} obtaining

$$\left(T^{-1}S\right)^2 = \begin{pmatrix}0&-1\\1&1\end{pmatrix}$$

Remark. This is another representation of the element $(1:1)$ of $\mathbb{P}^1(\mathbb{Z}/2\mathbb{Z})$.

Remark. The matrix T^{-1} belongs to $\Gamma_0(2)$; thus, the regions $ST^{-1}S\cdot\mathcal{F}$ and $\left(T^{-1}S\right)^2\cdot\mathcal{F}$ are connected by a matrix of $\Gamma_0(2)$ and therefore we can choose either of them. In order to have a connected fundamental region we choose $\left(T^{-1}S\right)^2\cdot\mathcal{F}$.

Remark. Observe that $S = S^{-1}$. Hence $\left(T^{-1}S\right)^2 = (ST)^{-2}$. We know recall that ST is of order 3 meaning $(ST)^{-2} = ST$

We conclude that the action associated to the element $(1:1) \in \mathbb{P}^1(\mathbb{Z}/2\mathbb{Z})$ corresponds to the action of the matrix ST.

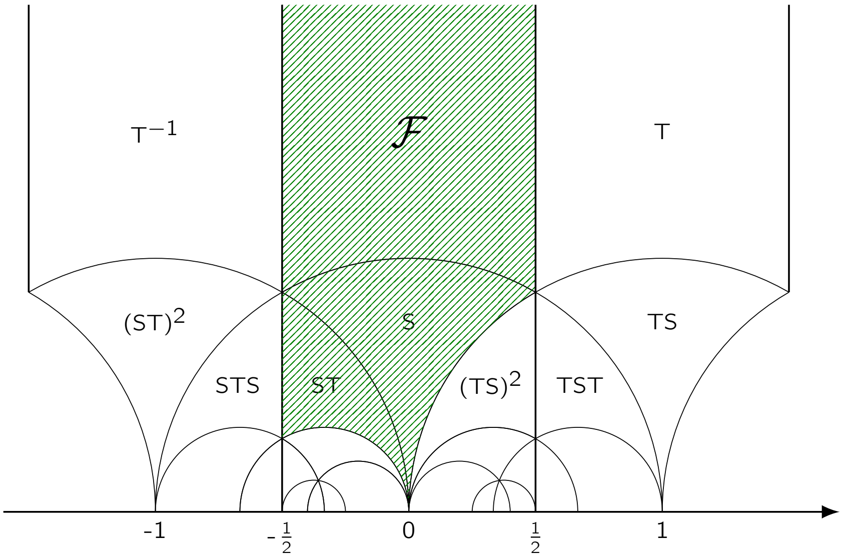


Figure 2.2 – The fundamental region for the congruence subgroup $\Gamma_0(2)$.

We used the following representation of $\mathbb{P}^1(\mathbb{Z}/2\mathbb{Z})$:

$$\Gamma_0(2)\backslash \mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & -1\\ 1 & 1\end{pmatrix}, \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \right\}$$

Fundamental domain for $\Gamma_0(4)$

We repeat the same process for $\Gamma_0(4)$.

$$\Gamma_0(4)\backslash \mathrm{SL}_2(\mathbb{Z}) \simeq \mathbb{P}^1(\mathbb{Z}/4\mathbb{Z}) = \left\{ \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & -1\\ 1 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 3 & 1\end{pmatrix}, \begin{pmatrix}0 & -1\\ 1 & 2\end{pmatrix} \right\}$$

We have already studied the action of the first three matrices in the previous subsection. Let us start with $(2:1)$. We observe that the fourth matrix sends ∞ to $\frac{1}{2}$. But $(2:1) = (-2:1) \in \mathbb{P}^1(\mathbb{Z}/4\mathbb{Z})$ and the matrix $\begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix}$ sends ∞ to $-\frac{1}{2}$. Now, for $a \geq \frac{\sqrt{3}}{2}$ and $-\frac{1}{2} \leq b \leq \frac{1}{2}$ we have

$$\begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix} \cdot \left(\frac{1}{2} + ai\right) = \frac{\frac{1}{2} + ai}{-\frac{2}{2} - 2ai + 1} = \frac{1 + 2ai}{-4ai} = -\frac{(1+2ai)(-i)}{4a} = \frac{-2a + i}{4a} = -\frac{1}{2} + \frac{i}{4a}$$

$$\begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix} \cdot \left(-\frac{1}{2} + ai\right) = \frac{-\frac{1}{2} + ai}{-\frac{-2}{2} - 2ai + 1} = \frac{-1 + 2ai}{4 - 4ai} = \frac{(-1+2ai)(4+4ai)}{16 + 16a^2} = \frac{-2a^2 - 1 + ai}{4(a^2+1)}$$

$$\begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix} \cdot \left(b + \sqrt{1-b^2}i\right) = \frac{b + \sqrt{1-b^2}i}{-2b + 1 - 2\sqrt{1-b^2}i} = \frac{(b + \sqrt{1-b^2}i)(1 - 2b + 2\sqrt{1-b^2}i)}{5 + 8b\sqrt{1-b^2}} =$$
$$= \frac{b - 2 + \sqrt{1-b^2}i}{5 + 8b\sqrt{1-b^2}}$$

Remark. For the last computation we could also use the exponential form of complex numbers keeping in mind that $\rho_1 e^{i\theta_1} + \rho_2 e^{i\theta_2} = \rho e^{i\theta}$ where

$$\rho = \sqrt{\rho_1^2 + \rho_2^2 + 2\rho_1\rho_2\cos(\theta_1 - \theta_2)} \qquad \text{and} \qquad \theta = \arctan\left(\frac{\rho_1 \sin\theta_1 + \rho_2 \sin\theta_2}{\rho_1\cos\theta_1 + \rho_2\cos\theta_2}\right)$$

In particular the matrix associated to $(2:1)$ sends

$$\infty \to (1:-2) \in \mathbb{P}^1(\mathbb{Q}) \qquad \frac{1}{2} + \frac{\sqrt{3}}{2}i \to -\frac{1}{2} + \frac{\sqrt{3}}{6}i \qquad -\frac{1}{2} + \frac{\sqrt{3}}{2}i \to -\frac{5}{14} + \frac{\sqrt{3}}{14}i$$

Remark. The decomposition of the matrix is

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} = ST^2S$$

Now we look at the fifth matrix $\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$. In the same spirit as before we could consider the equivalent representation of $(1:3) = (-1:1) \in \mathbb{P}^1(\mathbb{Z}/4\mathbb{Z})$, i.e., $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ but this sends ∞ to -1. We can therefore change the matrix (not the representation of the projective line element) by writing $\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$ which still lives in $\mathrm{SL}_2(\mathbb{Z})$.

$$\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \cdot \left(\frac{1}{2} + ai\right) = \frac{1}{-\frac{1}{2} - ai + 1} = \frac{2}{1 - 2ai} = \frac{2}{1 + 4a^2}(1 + 2ai) \qquad \frac{1}{2} + \frac{\sqrt{3}}{2}i \to \frac{1}{2} + \frac{\sqrt{3}}{2}i$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \cdot \left(-\frac{1}{2} + ai\right) = \frac{1}{\frac{1}{2} - ai + 1} = \frac{2}{3 - 2ai} = \frac{2}{9 + 4a^2}(3 + 2ai) \qquad -\frac{1}{2} + \frac{\sqrt{3}}{2}i \to \frac{1}{2} + \frac{\sqrt{3}}{6}i$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \cdot \left(b + \sqrt{1 - b^2}i\right) = \frac{1}{-b - \sqrt{1 - b^2}i + 1} = \frac{1 - b + \sqrt{1 - b^2}}{2 - 2b} = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{1 + b}{1 - b}}$$

And for $|b| \leq \frac{1}{2}$ then

$$\frac{\sqrt{3}}{6} = \frac{1}{2}\sqrt{\frac{1/2}{3/2}} \leq \frac{1}{2}\sqrt{\frac{1 + b}{1 - b}} \leq \frac{1}{2}\sqrt{\frac{3/2}{1/2}} = \frac{\sqrt{3}}{2}$$

Remark. We obtain the decomposition

$$\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} = (TS)^2$$

It only remains to study last matrix. Note that $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ is another representative but trying to minimize the first column we consider $\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$. The reason for doing it is to send ∞ to a cusp living in the strip $|\mathrm{Re}(\tau)| \leq \frac{1}{2}$.

$$\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \cdot \left(\frac{1}{2} + ai\right) = \frac{-1}{\frac{1}{2} + ai + 2} = \frac{-2}{5 + 2ai} = \frac{2}{25 + 4a^2}(-5 + 2ai) \qquad \frac{1}{2} + \frac{\sqrt{3}}{2}i \to -\frac{5}{14} + \frac{\sqrt{3}}{14}i$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \cdot \left(-\frac{1}{2} + ai\right) = \frac{-1}{-\frac{1}{2} + ai + 2} = \frac{-2}{3 + 2ai} = \frac{2}{9 + 4a^2}(-3 + 2ai) \qquad -\frac{1}{2} + \frac{\sqrt{3}}{2}i \to -\frac{1}{2} + \frac{\sqrt{3}}{6}i$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \cdot \left(b + \sqrt{1 - b^2}i\right) = \frac{-1}{b + \sqrt{1 - b^2}i + 2} = \frac{-b - 2 + \sqrt{1 - b^2}i}{5 + 4b}$$

and

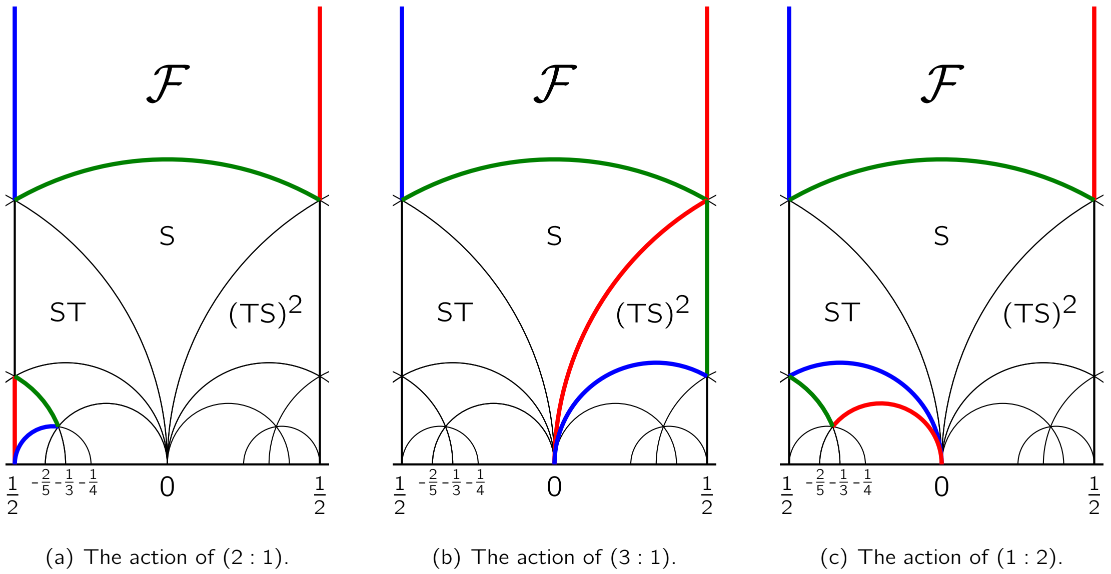
$$\left(\frac{-b-2}{5+4b}+\frac{2}{3}\right)^2+\left(\frac{\sqrt{1-b^2}i}{5+4b}\right)^2=\frac{b^2+4b+4+1-b^2}{(5+4b)^2}+\frac{-4b-8}{15+12b}+\frac{4}{9}=$$
$$=\frac{1}{4b+5}+\frac{-12b-24+20+16b}{45+36b}=\frac{4b-4+9}{45+36b}=\frac{1}{9}$$

Showing that $(2:1)$ sends the bottom boundary of $\mathcal{F}$ on the circle of center $-2/3$ and radius $1/3$.

Remark. We get the decomposition

$$\begin{pmatrix}0 & -1\\ 1 & 2\end{pmatrix}=ST^2$$

We resume the previous computations in the Figure 2.3.



(a) The action of $(2:1)$.

(b) The action of $(3:1)$.

(c) The action of $(1:2)$.

Figure 2.3 – The action of the last three matrices.

We now have all the information to construct the fundamental domain for $\Gamma_0(4)$.

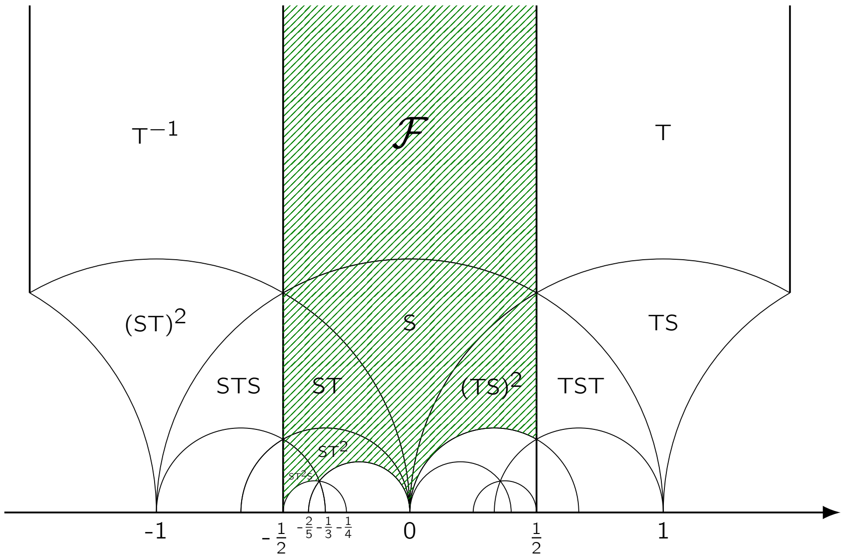


Figure 2.4 – The fundamental region for the congruence subgroup $\Gamma_0(4)$.

Fundamental domain for $\Gamma_0(8)$

We study now the case of $\Gamma_0(8)$. In principle, we could do a similar computation to the one we did for the $N = 2$ and $N = 4$. We recall that

$$\begin{array}{ccccccc}
\mathbb{P}^1(\mathbb{Z}/4\mathbb{Z}) = \{(0:1), & (1:1), & (-2:1), & (-1:1), & (1:0), & (1:2)\} & \\
\downarrow \qquad\quad \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
\mathbb{P}^1(\mathbb{Z}/8\mathbb{Z}) \supset \{(0:1), & (1:1), & (6:1), & (7:1), & (1:0), & (1:2)\} &
\end{array}$$

which means we should study the action of $(2:1), (3:1), (4:1), (5:1), (3:2), (1:4)$.

We have seen that this strategy presents a stage in which we proceed by attempts, namely the research for an optimal representation of the matrix. We will describe then a slightly different approach. Note that $\Gamma_0(4)$ has $\Gamma_0(8)$ as one of its subgroups of index 2 meaning that for each

$$[\Gamma_0(8)\backslash \mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(4)\backslash \mathrm{SL}_2(\mathbb{Z})] = 2$$

Thus, if M is a generator of $\Gamma_0(4)/\Gamma_0(8)$, then $\mathbb{P}^1(\mathbb{Z}/8\mathbb{Z}) = \mathbb{P}^1(\mathbb{Z}/4\mathbb{Z}) + M \cdot \mathbb{P}^1(\mathbb{Z}/4\mathbb{Z})$.
Finding a matrix $M \in \Gamma_0(4) \setminus \Gamma_0(8)$ is not hard. The only foresight we will have is to look for a matrix M which sends $\tau \in \mathcal{F}_{\Gamma_0(4)}$ in the region $|\mathrm{Re}(\tau)| \leq \frac{1}{2}$.

Lemma 2.9. *The matrix* $M = \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}$ *has the desired property.*

Proof. We consider the action of a generic matrix of $\mathrm{SL}_2(\mathbb{Z})$ on $\tau = s + ti$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d} = \frac{(ac|\tau|^2 + (ad+bc)s + bd) + ((ad-bc)t)i}{|c\tau + d|^2}$$

whose real part is

$$\mathrm{Re}(M \cdot \tau) = \frac{(ac|\tau|^2 + (ad+bc)s + bd)}{|c\tau + d|^2} = \frac{(-4|\tau|^2 + s)}{|-4\tau + 1|^2} = \frac{(-4|\tau|^2 + s)}{16|\tau|^2 - 8s + 1}$$

where we are substituting the values for M. Now

$$\mathrm{Re}(M \cdot \tau) = \frac{-(4|\tau|^2 - s)}{16|\tau|^2 - 8s + 1} = \frac{-(4|\tau|^2 - 2s + s)}{4(4|\tau|^2 - 2s) + 1} = -\frac{1}{4}\frac{4(4|\tau|^2 - 2s) + 1 + 4s - 1}{4(4|\tau|^2 - 2s) + 1} = -\frac{1}{4}\frac{h + 4s - 1}{h}$$

or

$$\mathrm{Re}(M \cdot \tau) = -\frac{1}{4} - \frac{1}{4}\frac{4s - 1}{h}$$

for $h = 4(4|\tau|^2 - 2s) + 1 = 16|\tau|^2 - 8s + 1 = (4s-1)^2 + 16t^2$.

We note that all the points of $\mathcal{F}_{\Gamma_0(4)}$ are outside the circle of center 1/3 and radius 1/3 thus

$$\left(s - \frac{1}{3}\right)^2 + t^2 \geq \frac{1}{9} \Rightarrow t^2 \geq \frac{2s}{3} - s^2 \implies h \geq \frac{8s}{3} + 1$$

At the same time

$$\left(s + \frac{1}{5}\right)^2 + t^2 \geq \frac{1}{25} \Rightarrow t^2 \geq -\frac{2s}{5} - s^2 \implies h \geq -\frac{72s}{5} + 1$$

To minimize h we choose to consider the region inside the unit circle; we will return on the study of $M\mathcal{F}$ at a later stage.

$$s^2 + t^2 \leq 1 \Rightarrow t^2 \leq 1 - s^2 \implies h \leq 17 - 8s$$

Thus

$$\begin{cases} \dfrac{4s-1}{\frac{8s}{3}+1} \geq \dfrac{4s-1}{h} \geq \dfrac{4s-1}{17-8s} \\ \dfrac{4s-1}{17-8s} \geq \dfrac{4s-1}{h} \geq \dfrac{4s-1}{\frac{8s}{3}+1} \\ \dfrac{4s-1}{17-8s} \geq \dfrac{4s-1}{h} \geq \dfrac{4s-1}{-\frac{72s}{5}+1} \end{cases} \Rightarrow \begin{cases} -\dfrac{1}{4} - \dfrac{4s-1}{\frac{32s}{3}+4} \leq \mathrm{Re}(M\cdot\tau) \leq -\dfrac{1}{4} - \dfrac{4s-1}{68-32s} & \text{if } s \geq \frac{1}{4} \\ -\dfrac{1}{4} - \dfrac{4s-1}{68-32s} \leq \mathrm{Re}(M\cdot\tau) \leq -\dfrac{1}{4} - \dfrac{4s-1}{\frac{32s}{3}+4} & \text{if } 0 \leq s < \frac{1}{4} \\ -\dfrac{1}{4} - \dfrac{4s-1}{68-32s} \leq \mathrm{Re}(M\cdot\tau) \leq -\dfrac{1}{4} - \dfrac{4s-1}{-\frac{288s}{5}+4} & \text{if } s < 0 \end{cases}$$

And the following picture implies that $-\frac{1}{2} \leq \mathrm{Re}(M\cdot\tau) \leq \frac{1}{2}$.

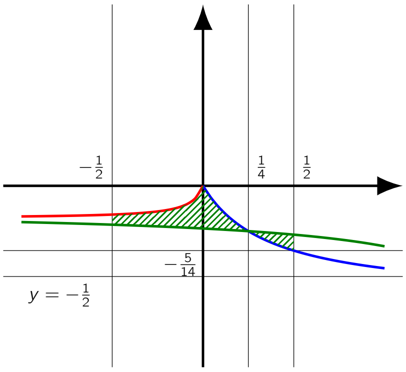


Figure 2.5 – Bound for the real part of $M\cdot\tau$.

A straightforward computation shows that $M\cdot\mathcal{F}$ is also in the desired region. □

Lemma 2.10. *The matrix M decomposes as $M = ST^4S$*

Hence, we have to study the action of ST^4S on the 6 regions of composing $\mathcal{F}_{\Gamma_0(4)}$:

$$ST^4S\cdot\mathcal{F}, \quad ST^4S\cdot S\cdot\mathcal{F} = ST^4\cdot\mathcal{F}, \quad ST^4S\cdot ST\cdot\mathcal{F} = ST^5\cdot\mathcal{F}, \quad ST^4S\cdot ST^2\cdot\mathcal{F} = ST^6\cdot\mathcal{F}$$

$$ST^4S\cdot ST^2S\cdot\mathcal{F} = ST^6S\cdot\mathcal{F}, \quad ST^4S\cdot(TS)^2\cdot\mathcal{F} = ST^4S\cdot(TS)^{-1}\cdot\mathcal{F} = ST^3\cdot\mathcal{F}$$

$$\begin{array}{llll}
ST^4S\cdot\mathcal{F}: & \infty \to -\dfrac{1}{4} & \dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{9}{42}+\dfrac{\sqrt{3}}{42}i & -\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{7}{26}+\dfrac{\sqrt{3}}{26}i \\
ST^4\cdot\mathcal{F}: & \infty \to 0 & \dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{9}{42}+\dfrac{\sqrt{3}}{42}i & -\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{7}{26}+\dfrac{\sqrt{3}}{26}i \\
ST^5\cdot\mathcal{F}: & \infty \to 0 & \dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{11}{62}+\dfrac{\sqrt{3}}{62}i & -\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{9}{42}+\dfrac{\sqrt{3}}{42}i \\
ST^6\cdot\mathcal{F}: & \infty \to -0 & \dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{13}{86}+\dfrac{\sqrt{3}}{86}i & -\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{11}{62}+\dfrac{\sqrt{3}}{62}i \\
ST^6S\cdot\mathcal{F}: & \infty \to -\dfrac{1}{6} & \dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{11}{62}+\dfrac{\sqrt{3}}{62}i & -\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{13}{86}+\dfrac{\sqrt{3}}{86}i \\
ST^3\cdot\mathcal{F}: & \infty \to 0 & \dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{7}{26}+\dfrac{\sqrt{3}}{26}i & -\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i \to -\dfrac{5}{14}+\dfrac{\sqrt{3}}{14}i
\end{array}$$

Remark. We have the following

$-\dfrac{5}{14}+\dfrac{\sqrt{3}}{14}i$ is the intersection between $\left(X+\dfrac{1}{5}\right)^2+Y^2=\dfrac{1}{25}$ and $\left(X+\dfrac{3}{8}\right)^2+Y^2=\dfrac{1}{64}$.

$-\dfrac{7}{26}+\dfrac{\sqrt{3}}{26}i$ is the intersection between $\left(X+\dfrac{1}{7}\right)^2+Y^2=\dfrac{1}{49}$ and $\left(X+\dfrac{3}{8}\right)^2+Y^2=\dfrac{1}{64}$.

$-\dfrac{9}{42}+\dfrac{\sqrt{3}}{42}i$ is the intersection between $\left(X+\dfrac{1}{9}\right)^2+Y^2=\dfrac{1}{81}$ and $\left(X+\dfrac{5}{24}\right)^2+Y^2=\dfrac{1}{576}$.

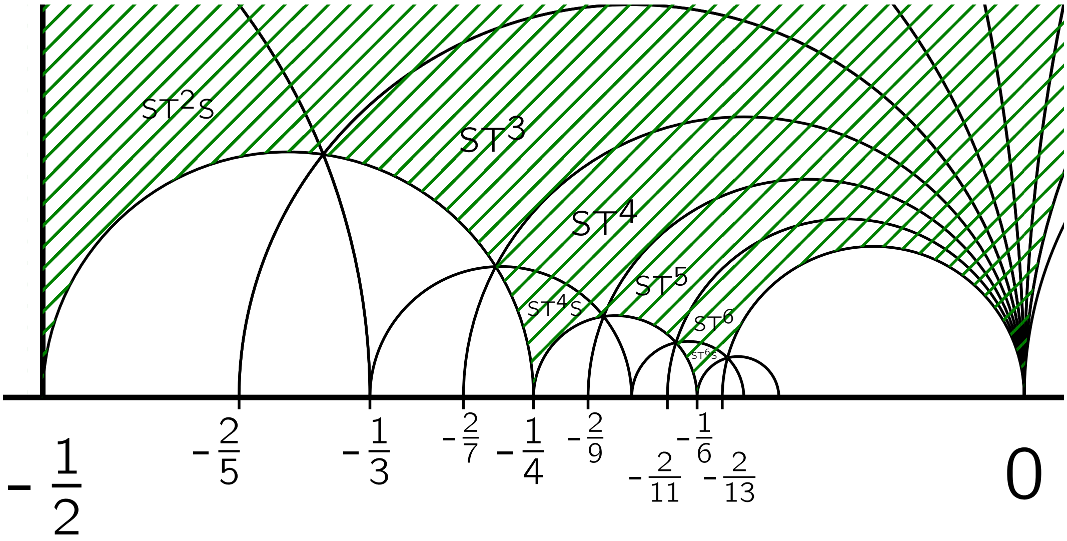


Figure 2.6 – A zoom in the fundamental region for $\Gamma_0(8)$.

$-\dfrac{11}{62}+\dfrac{\sqrt{3}}{62}i$ is the intersection between $\left(X+\dfrac{1}{11}\right)^2+Y^2=\dfrac{1}{121}$ and $\left(X+\dfrac{5}{24}\right)^2+Y^2=\dfrac{1}{576}$.
$-\dfrac{13}{86}+\dfrac{\sqrt{3}}{86}i$ is the intersection between $\left(X+\dfrac{1}{13}\right)^2+Y^2=\dfrac{1}{169}$ and $\left(X+\dfrac{7}{48}\right)^2+Y^2=\dfrac{1}{2304}$.

Below is a picture of the complete region.

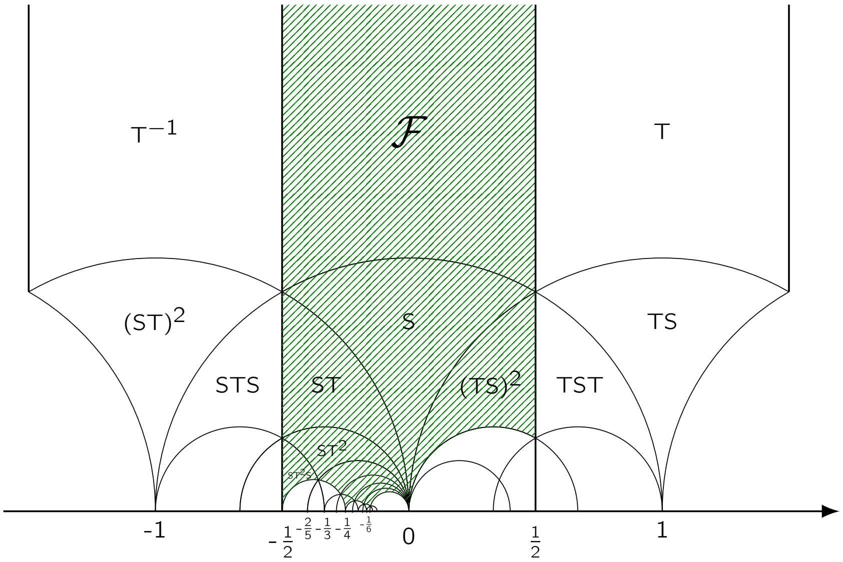
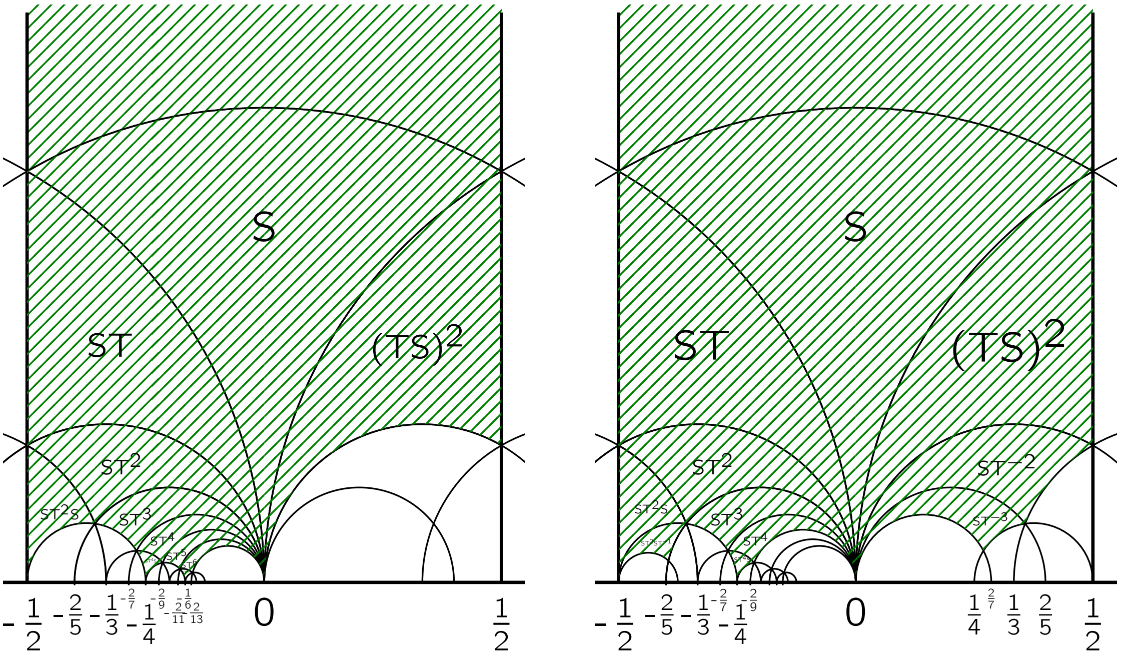


Figure 2.7 – The fundamental region for the congruence subgroup $\Gamma_0(8)$.

Let us try to make this last picture a little bit more legible. We observe that $ST^5=\begin{pmatrix}0 & -1\\ 1 & 5\end{pmatrix}$ represents $(5:1)$, $ST^6=\begin{pmatrix}0 & -1\\ 1 & 6\end{pmatrix}$ represents $(3:2)$ and $ST^6S=\begin{pmatrix}-1 & 0\\ 6 & -1\end{pmatrix}$ represents $(2:1)$.
We can easily see that the matrix $ST^{-8}S\in\Gamma_0(8)$ sends ST^5 to ST^{-3} and ST^6 to ST^{-2}. The two regions $ST^{-2}\cdot\mathcal{F}$ and $ST^{-3}\cdot\mathcal{F}$ are still in the strip $|\mathrm{Re}(\tau)|\leq 1/2$ and they are the symmetric to ST^2 and ST^2 with respect to the y-axis. Further, ST^{-2} (respectively ST^{-3}) still represents $(3:2)$ (respectively $(5:1)$).
Concerning the last matrix ST^6S, we observe that it could be sent to $ST^{-2}S$ but the image of ∞ after it is $1/2$. This is exactly $1+(-1/2)$ and we would like not to have points which are connected by T in our

fundamental region. Note that the new region should be conjugated to ST^2S by a matrix in $\Gamma_0(4)$ Thus, we could check for the matrices $ST^2S \cdot T$, $ST^2S \cdot T^{-1}$ which corresponds to two adjacent regions. It turns out that ST^2ST^{-1} is a good choice.



(a) The previous region

(b) A second region

Figure 2.8 – Two fundamental regions for $\Gamma_0(8)$.

We would like to point out that there exists an algorithm implemented in C that, given $0 < N \leq 50$, constructs the fundamental region for $\Gamma_0(N)$ ([Ver2]). For $\Gamma_0(8)$, in particular, it gives the one in Figure 2.8(b); this is because the authors of [Ver2] intended to have all the regions as big as possible.

Some remarks on the fundamental domain for $\Gamma_0(2^n)$

Given a region $A \cdot \mathcal{F}$, the algorithm in [Ver1] (whose code is described in [Ver2]) consists in checking all the adjacent regions $AS \cdot \mathcal{F}$, $AT \cdot \mathcal{F}$, $AT^{-1} \cdot \mathcal{F}$, $AST \cdot \mathcal{F}$, $AST^{-1} \cdot \mathcal{F}$. The diagram in Figure 2.9 is a graphic representation of the general situation.

Now suppose we have already constructed the fundamental region for $\Gamma_0(2^n)$ and we want to find the one for $\Gamma_0(2^n)$. We also suppose that we started from $\mathcal{F}_{\Gamma_0(8)}$ of Figure 2.7.

Lemma 2.11. *The matrix* $M = ST^{2^n}S = \begin{pmatrix} 1 & 0 \\ -2^n & 1 \end{pmatrix}$ *sends all the points of* $\mathcal{F}_{\Gamma_0(2^n)}$ *in the strip* $|\mathrm{Re}(\tau)| \leq \frac{1}{2}$.

Proof. This is the same kind of computation that we already did in Lemma 2.9 but we could also use a different approach. All the regions of $\mathcal{F}_{\Gamma_0(2^{n-1})}$ will be associated to matrices of the form I, S, $(TS)^2$, ST^k, ST^kS with $0 < k \leq 2^{n+1} - 2$ (easy to see by induction). Now

$$\begin{aligned}
ST^{2^n}S \cdot I &= ST^{2^n}S \\
ST^{2^n}S \cdot S &= ST^{2^n} \\
ST^{2^n}S \cdot (TS)^2 &= ST^{2^n-1} \\
ST^{2^n}S \cdot ST^k &= ST^{2^n+k} \\
ST^{2^n}S \cdot ST^kS &= ST^{2^n+k}S
\end{aligned}$$

And with the use of Figures 2.6 and 2.9 we deduce that all the corresponding region are in the correct strip. □

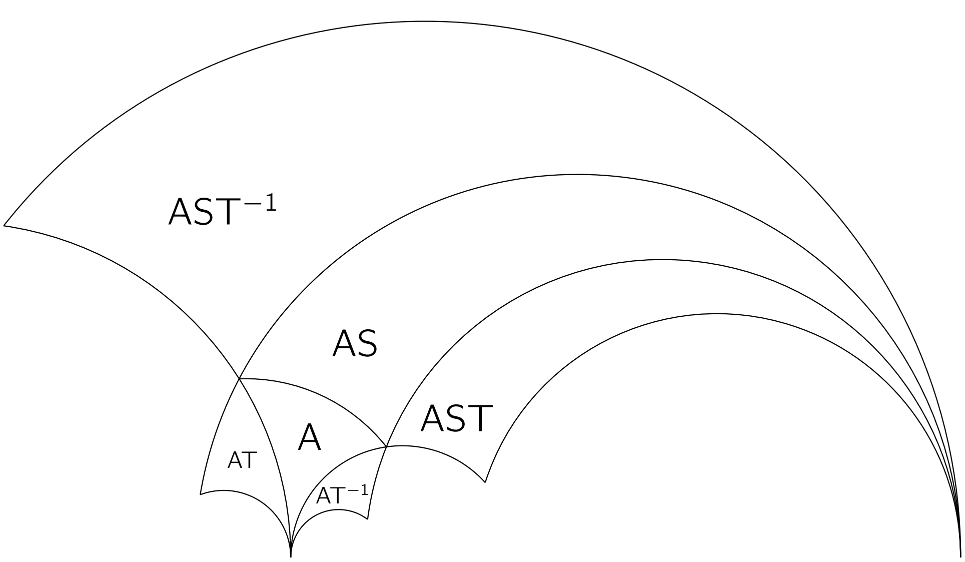


Figure 2.9 – Adjacent regions to $A \cdot \mathcal{F}$.

Remark. If the boundary of a region is of the form $\left(X + \frac{1}{r}\right)^2 + Y^2 = \frac{1}{r^2}$, then all the points on it are of the form $a + \sqrt{-a^2 - \frac{2a}{r}}i$ and our matrix sent them to the upper semicircle of center $\frac{1}{r} - \frac{2^{n+1}}{r^2}$ and radius $\frac{2^{n+1}}{r^2} - \frac{1}{r}$.

Lemma 2.12. *The lower circular bound edge of the region ST^k has equation $\left(X + \frac{1}{2k+1}\right)^2 + Y^2 = \frac{1}{(2k+1)^2}$.*

Proof. We carry on a simple computation

$$\begin{pmatrix} 0 & -1 \\ 1 & k \end{pmatrix} \begin{pmatrix} \frac{1}{2} + ai \\ 2 \end{pmatrix} = \frac{-1}{\frac{1}{2} + ai + k} = \frac{-2}{2k + 1 + 2ai} = \frac{2(-2k - 1 + 2ai)}{4k^2 + 4k + 1 + 4a^2} = \frac{-4k - 2 + 4ai}{4k^2 + 4k + 1 + 4a^2}$$

and now

$$\frac{16k^2 + 16k + 4}{(4k^2 + 4k + 1 + 4a^2)^2} - \frac{8k + 4}{r\,(4k^2 + 4k + 1 + 4a^2)} + \frac{16a^2}{(4k^2 + 4k + 1 + 4a^2)^2} = 0$$

implies

$$\frac{16k^2 + 16k + 4 + 16a^2}{(4k^2 + 4k + 1 + 4a^2)^2} = \frac{8k + 4}{r\,(4k^2 + 4k + 1 + 4a^2)}$$

$$\frac{4(4k^2 + 4k + 1 + 4a^2)}{(4k^2 + 4k + 1 + 4a^2)^2} = \frac{8k + 4}{r\,(4k^2 + 4k + 1 + 4a^2)} \Longrightarrow 4 = \frac{8k + 4}{r}$$

from which

$$r = \frac{8k + 4}{4} = 2k + 1$$

□

Lemma 2.13. *The bounds of the region $ST^{2k}S$ are the circles*

- $\left(X + \dfrac{2k - 1}{(2k - 2)2k}\right)^2 + Y^2 = \dfrac{1}{((2k - 2)2k)^2}$ *on the left.*
- $\left(X + \dfrac{2k + 1}{2k(2k + 2)}\right)^2 + Y^2 = \dfrac{1}{(2k(2k + 2))^2}$ *on the right.*
- $\left(X + \dfrac{2k}{(2k - 1)(2k + 1)}\right)^2 + Y^2 = \dfrac{1}{((2k - 1)(2k + 1))^2}$ *above.*

Proof. The left side bound is the circles connecting the images of ∞ under the actions of the matrices $ST^{2k-2}S$ and $ST^{2k}S$, i.e., the circle connecting $-\frac{1}{2k-2}$ and $-\frac{1}{2k}$. This has radius $\frac{1}{2}\left(\frac{1}{2k-2} + \frac{1}{2k}\right) = \frac{1}{(2k-1)2k}$ and center $-\frac{1}{2k} - \frac{1}{(2k-1)2k} = \frac{2k-1}{(2k-2)2k}$. The same computation can be repeated for the right edge.

Finally, the top bound is given by the circle connecting $ST^{2k-1}S\cdot\infty$ and $ST^{2k+1}S\cdot\infty$. □

We can now give a general description of the fundamental region of $\Gamma_0(2^n)$. This is composed by the regions corresponding to I, S, $(TS)^2$,

$$ST^k \text{ for } 1\leq k\leq 2^n-2$$

(all the triangles with vertex in 0 and lower radius $\frac{1}{2(2^n-2)+1}$) and

$$ST^{2k}S \text{ for } 1\leq k\leq 2^{n-1}-1$$

(all the triangles with vertex in $-\frac{1}{2k}$).

2.1.5 The cusps of $X_0(N)$

We recall that $X(\Gamma)$ is the compactification of $Y(\Gamma)$ and it is defined by considering the extended quotient

$$X(\Gamma)=\Gamma\backslash\mathbb{H}^*=Y(\Gamma)\cup\Gamma\backslash(\mathbb{Q}\cup\{\infty\})$$

The points in $\Gamma(\mathbb{Q}\cup\{\infty\})$ are called the cusps of $X(\Gamma)$.

Some good references are [CS2, §6.3], [DI, §9.1], [DS, §2.4, 3.8], [Man2, §2], [Miy, §1.7,1.8,4.2] and [Shi, §1.3-1.6].

Proposition 2.14. *For any congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$, the modular curve $X(\Gamma)$ has finitely many cusps.*

Proof. One easily checks that the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{Q})$ is transitive and that the stabilizer of ∞ in $\mathrm{SL}_2(\mathbb{Z})$ is

$$\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)=\left\{\pm\begin{pmatrix}1&b\\0&1\end{pmatrix}\;\middle|\; b\in\mathbb{Z}\right\}$$

This shows that we have a bijection

$$\mathrm{SL}_2(\mathbb{Z})/\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)\xrightarrow{\sim}\mathbb{P}^1(\mathbb{Q})$$

Now we know that

$$\mathrm{Cusps}(\Gamma)=\Gamma\backslash\mathbb{P}^1(\mathbb{Q})=\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)$$

which says that there is a surjective map

$$\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})\twoheadrightarrow\mathrm{Cusps}(\Gamma)$$

and the set on the left is finite since

$$\Gamma(n)\subseteq\Gamma\subseteq\mathrm{SL}_2(\mathbb{Z})\quad\text{some } n\in\mathbb{Z}$$

and $[\mathrm{SL}_2(\mathbb{Z}):\Gamma(n)]$ is finite □

This proof has a very important consequence:

Corollary 2.15. *The cusps of Γ correspond to the orbit in the action of $\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)$ on $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$.*

In the case of $\Gamma=\Gamma_0(N)$ we have already described these orbits in section 2.1.2 and we know that representatives for them are of the form

$$\left\{(a:b)\;\middle|\; b|N,\ a\text{ is the lift of an element of }\left(\frac{\mathbb{Z}}{(b,N/b)\,\mathbb{Z}}\right)^{\times}\text{ in }\frac{\mathbb{Z}}{N\mathbb{Z}}\text{ such that }(a,b,N/b)=1\right\}$$

There is also another way of obtaining Corollary 2.15. The way we computed the genus of the modular curve $X_0(N)$ was by considering the map $X_0(N)\to X(1)$ which is a covering of surfaces. The cusps of $X_0(N)$ are therefore the preimages of the cusps of $X(1)$ and

Lemma 2.16. *The modular curve $X(1)=\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^*$ has one cusp, namely ∞.*

Proof. This is clear from the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{Q})$ being transitive. □

Finally, by Proposition 1.33 we know that the preimage of ∞ is given by $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/\langle T\rangle$ and therefore we obtain again the previous corollary. Note that there is another way of describing the cusps.

Proposition 2.17. *Let $s = (a : b)$ and $s' = (a' : b')$ be elements of $\mathbb{P}^1(\mathbb{Q})$ with $(a : b) = (a' : b') = 1$. Then*

$$\Gamma_0(N)s' = \Gamma_0(N)s \iff (ya' : b') = (a + jc : yc) \bmod N$$

for some $y, j \in \mathbb{Z}$ with $(y, N) = 1$.

Proof. This is Proposition 3.8.3 in [DS]. □

From the discussion following the proposition in [DS, §3] one can obtain the same description of $\mathrm{Cusps}(\Gamma_0(N))$ and their number:

$$\sum_{d|N} \varphi\left(\left(d, \frac{N}{d}\right)\right)$$

Example. Let us take $N = 2^n$.

$$\mathrm{Cusps}(\Gamma_0(N)) = \left\{(a : 2^k) \,\middle|\, 1 \leq k \leq n-1,\ 1 \leq a \leq j_k - 1, \left(a, 2^k\right) = 1\right\} \cup \{[0]\,, \infty\}$$

where

$$j_k = \begin{cases} 2^k & \text{if } 1 \leq k \leq \lfloor n/2 \rfloor \\ 2^{n-k} & \text{if } \lfloor n/2 \rfloor < k < n \end{cases}$$

Example (Cusps for $N = 2^4 = 16$)**.** We have the following

b \ a	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	·	∞	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
2	·	⊗	·	⊗	·	⊗	·	⊗	·	·	·	·	·	·	·	·
3	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4	·	⊖	·	⊘	·	·	·	·	·	·	·	·	·	·	·	·
5	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
6	·	⊕	·	·	·	·	·	·	·	·	·	·	·	·	·	·
7	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
8	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
9	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
10	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
11	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
14	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
15	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

where $\infty, \times, \otimes, \ominus, \oslash, \oplus$ represent the 6 cusps and $\cdot$ represents an invalid choice of a and b.

Let now $\mathfrak{c}$ be a cusp for a congruence subgroup Γ. We denote by s the representative in $\mathbb{P}^1(\mathbb{Q})$ of the corresponding Γ-orbit. We denote by γ_s the matrix of $\mathrm{SL}_2(\mathbb{Z})$ such that $\gamma_s \cdot \infty = s$. Now for any $\gamma \in \Gamma$

$$\begin{aligned} \gamma \in \mathrm{Stab}_\Gamma(s) &\iff \gamma \cdot s = s \\ &\iff \gamma\gamma_s \cdot \infty = \gamma_s \cdot \infty \\ &\iff \gamma_s^{-1}\gamma\gamma_s \cdot \infty = \infty \\ &\iff \gamma_s^{-1}\gamma\gamma_s \in \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \end{aligned}$$

Which says

$$\mathrm{Stab}_\Gamma(s) = \Gamma \cap \gamma_s \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)\gamma_s^{-1}$$

Therefore, we obtain an injective map

$$\mathrm{Stab}_\Gamma(s)\backslash\gamma_s\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)\gamma_s^{-1} \hookrightarrow \Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$$

which in turns gives us that $\mathrm{Stab}_\Gamma(s)$ is of finite index in $\gamma_s\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)\gamma_s^{-1}$. We define

$$H_\mathfrak{c} = \gamma_s^{-1}\Gamma\gamma_s \cap \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)$$

and we know that $H_\mathfrak{c}$ is of finite index in $\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)$.

Lemma 2.18. *$H_\mathfrak{c}$ does not depend on the choice of s and γ_s.*

Definition. Let $\mathfrak{c}$ be a cusp for Γ and let s be a representative for the corresponding Γ-orbit in $\mathbb{P}^1(\mathbb{Q})$. The width of $\mathfrak{c}$ is

$$h_\mathfrak{c} = \left[\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) : \{\pm 1\}\, H_\mathfrak{c}\right]$$

We say that $\mathfrak{c}$ is irregular if $\gamma_s^{-1}\mathrm{Stab}_\Gamma(s)\gamma_s$ is generated by $\begin{pmatrix} -1 & h \\ 0 & -1 \end{pmatrix}$ for some $h \geq 1$ and regular otherwise.

Proposition 2.19. *Let $\mathfrak{c} \in \mathrm{Cusps}\left(\Gamma_0(N)\right)$ be represented by $s = (a : b) \in \mathbb{P}^1(\mathbb{Q})$ with $(a, b) = 1$, then the width of $\mathfrak{c}$ for $\Gamma_0 = (N)$ is equal to*

$$h_\mathfrak{c} = \frac{N}{(b^2, N)} = \frac{N/b}{(b, N/b)}$$

Proof. Let us consider a cusp $\mathfrak{c} \in \mathrm{Cusps}\left(\Gamma_0(N)\right)$ and a representative of its $s = (a : b) \in \mathbb{P}^1(\mathbb{Q})$. The matrix $\gamma_s \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ sends s to ∞.

We start by describing the stabilizer of s recalling that $\gamma_s \in \mathrm{SL}_2(\mathbb{Z})$ and therefore $ad - bc = 1$:

$$\begin{aligned}
\mathrm{Stab}_{\Gamma_0(N)}(s) &= \Gamma_0(N) \cap \gamma_s\mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty)\gamma_s^{-1} \\
&= \Gamma_0(N) \cap \left\{ \pm \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \;\middle|\; m \in \mathbb{Z} \right\} \\
&= \Gamma_0(N) \cap \left\{ \pm \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} d - mb & ma - c \\ -b & a \end{pmatrix} \;\middle|\; m \in \mathbb{Z} \right\} \\
&= \Gamma_0(N) \cap \left\{ \pm \begin{pmatrix} ad - mab - bc & ma^2 - ac + ac \\ bd - mb^2 - bd & mab - bc + ad \end{pmatrix} \;\middle|\; m \in \mathbb{Z} \right\} \\
&= \Gamma_0(N) \cap \left\{ \pm \begin{pmatrix} 1 - mab & ma^2 \\ -mb^2 & 1 + mab \end{pmatrix} \;\middle|\; m \in \mathbb{Z} \right\} \\
&= \left\{ \pm \begin{pmatrix} 1 - mab & ma^2 \\ -mb^2 & 1 + mab \end{pmatrix} \;\middle|\; m \equiv 0 \mod \frac{N}{(b^2, N)} \right\}
\end{aligned}$$

where the last equality comes from the fact that b is a divisor of N and $\Gamma_0(N)$ consists of matrices having $a_{2,1} \equiv 0$ modulo N.

Conjugating again by γ_s gives

$$\begin{aligned}
H_\mathfrak{c} &= \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \cap \gamma_s^{-1}\mathrm{Stab}_{\Gamma_0(N)}(s)\gamma_s \\
&= \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \cap \left\{ \pm \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} 1 - mab & ma^2 \\ -mb^2 & 1 + mab \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \;\middle|\; m \equiv 0 \bmod \frac{N}{(b^2, N)} \right\} \\
&= \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \cap \left\{ \pm \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} a - ma^2b + ma^2b & c - mabc + ma^2d \\ b - mab^2 + mab^2 & d - mb^2c + mabd \end{pmatrix} \;\middle|\; m \equiv 0 \bmod \frac{N}{(b^2, N)} \right\} \\
&= \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \cap \left\{ \pm \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} a & c - ma(bc - ad) \\ b & d - mb(bc - ad) \end{pmatrix} \;\middle|\; m \equiv 0 \bmod \frac{N}{(b^2, N)} \right\} \\
&= \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \cap \left\{ \pm \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} a & c + ma \\ b & d + mb \end{pmatrix} \;\middle|\; m \equiv 0 \bmod \frac{N}{(b^2, N)} \right\}
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \cap \left\{ \pm \begin{pmatrix} ad-bc & cd+mad-cd-mbc \\ ab-ab & -bc-mab+ad+mab \end{pmatrix} \;\middle|\; m \equiv 0 \bmod \frac{N}{(b^2,N)} \right\} \\
&= \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \cap \left\{ \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \;\middle|\; m \equiv 0 \bmod \frac{N}{(b^2,N)} \right\} \\
&= \left\{ \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \;\middle|\; m \equiv 0 \bmod \frac{N}{(b^2,N)} \right\}
\end{aligned}
$$

and this concludes the proof. □

Remark. Observe that from the description of $H_{\mathfrak{c}}$ we can also infer that all cusps for $\Gamma_0(N)$ are regular.

Lemma 2.20. *Let Γ be a congruence subgroup, and let $\overline{\Gamma}$ be the image of Γ in $\mathrm{PSL}_2(\mathbb{Z})$. Then*

$$\sum_{\mathfrak{c} \in \mathrm{Cusps}(\Gamma)} h_{\mathfrak{c}} = \left[\mathrm{PSL}_2(\mathbb{Z}) : \overline{\Gamma}\right] = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$$

Remark. If Γ is a normal congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, then $\gamma^{-1}\Gamma\gamma = \Gamma$ for all $\gamma \in \mathrm{SL}(\mathbb{Z})$. This implies that all cusps of Γ have the same width and either all are regular or all are irregular.

Example. We have seen that the cusps for $\Gamma_0(2^n)$ are

$$\mathrm{Cusps}(\Gamma_0(N)) = \left\{(a:2^k) \,\middle|\, 1 \le k \le n-1,\ 1 \le a \le j_k - 1,\ \left(a, 2^k\right) = 1\right\} \cup \{[0], \infty\}$$

where

$$j_k = \begin{cases} 2^k & \text{if } 1 \le k \le \lfloor n/2 \rfloor \\ 2^{n-k} & \text{if } \lfloor n/2 \rfloor < k < n \end{cases}$$

Their width is

$$h_{n,k} = \frac{2^n}{(2^{2k}, 2^n)} = \begin{cases} 1 & \text{if } 2k > n \\ 2^{n-2k} & \text{if } 2k \le n \end{cases}$$

Observe that

$$
\begin{aligned}
\sum_{\mathfrak{c} \in \mathrm{Cusps}(\Gamma_0(2^n))} h_{\mathfrak{c}} &= \sum_{k=1}^{n-1} h_{n,k} \cdot \varphi(j_k) + h_{[0]} + h_\infty \\
&= \sum_{k=1}^{\lfloor n/2 \rfloor} \varphi\left(2^k\right) \cdot 2^{n-2k} + \sum_{k=\lfloor n/2 \rfloor + 1}^{n-1} \varphi\left(2^{n-k}\right) + 2^n + 1 \\
&= \sum_{k=1}^{\lfloor n/2 \rfloor} 2^{k-1} \cdot 2^{n-2k} + \sum_{k=\lfloor n/2 \rfloor + 1}^{n-1} 2^{n-k-1} + 2^n + 1 \\
&= \sum_{k=1}^{\lfloor n/2 \rfloor} 2^{n-k-1} + \sum_{k=\lfloor n/2 \rfloor + 1}^{n-1} 2^{n-k-1} + 2^n + 1 \\
&= \sum_{k=2}^{n} 2^{n-k} + 2^n + 1 = \sum_{k=0}^{n-2} 2^k + 2^n + 1 \\
&= \left(\frac{1 - 2^{n-1}}{1-2}\right) + 2^n + 1 = 2^{n-1} - 1 + 2^n + 1 = 2^n + 2^{n-1} \\
&= 3 \cdot 2^{n-1} = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma]
\end{aligned}
$$

Example. Let us study the case $N = 16$.

```
sage: [Gamma0(16).cusp_width(c) for c in Gamma0(16).cusps()]
[16, 1, 1, 4, 1, 1]
```

Remark. The width of a cusp described the number of elements in the orbits corresponding to the cusp.

2.2 Defining equations of $X_0(N)$

The modular curves $X_0(N)$ have been studied for a long time since they provide a link between modular functions and elliptic curves. They are well known to parametrize pairs (E, ψ) of elliptic curves together with a cyclic isogeny of degree N and therefore their models encode much information and are of particular interest. There is a whole family of cryptographic protocols based on isogenies between elliptic curves; knowing how to compute isogenies linking a given elliptic curve to another one makes it easier to implement these systems.

In general, plane equations for modular curves are obtained by looking at two functions on $X_0(N)$ with nice properties and finding polynomial relations between them. These nice functions are called Modular functions. The problem of determining equations of modular curves has been addressed by many mathematicians: Galbraith [Gal1] uses different techniques based on the use of canonical projective embeddings. A similar approach was also used by Murabayashi [Mur] and Shimura [Shm]. Kohel [Koh2] used quaternion algebras to produce weight 2 cusp forms. The use of the Dedekind η-function appears in [Lig1] and has been eventually used by Enge and Schertz [ES], Yang [Yan2] and Kodrnja [Kod2]. Kodrnja [Kod1] also used embeddings in projective spaces and modular forms to find plane models for $X_0(N)$.

In this section we will try to give a complete background on modular functions and to show one way of finding a concrete model for $X_0(N)$.

2.2.1 Modular functions

We have a well defined $\mathbb{Z}$-periodic holomorphic map $q : \mathbb{H} \to \mathbb{D}$, from the upper half plane to the open unitary disc, defined by

$$q(\tau) = e^{2\pi i\tau} = e^{2\pi i(\mathrm{Re}(\tau)+i\mathrm{Im}(\tau))} = e^{-2\pi\mathrm{Im}(\tau)}e^{2\pi i\mathrm{Re}(\tau)} = e^{-2\pi\mathrm{Im}(\tau)}\left(\cos(2\pi\mathrm{Re}(\tau)) + i\sin(2\pi\mathrm{Re}(\tau))\right)$$

It bijectively maps each vertical strip [Sut3]

$$\mathbb{H}_n = \{\tau \in \mathbb{H} \mid n \le \mathrm{Re}(\tau) < n+1\} \quad \text{to} \quad \mathbb{D}_0 = \mathbb{D} \setminus \{0\}$$

and

$$\mathbb{H}_n^B = \{\tau \in \mathbb{H} \mid n \le \mathrm{Re}(\tau) < n+1,\ \mathrm{Im}(\tau) > B\} \quad \text{to} \quad e^{-2\pi B}\mathbb{D}_0$$

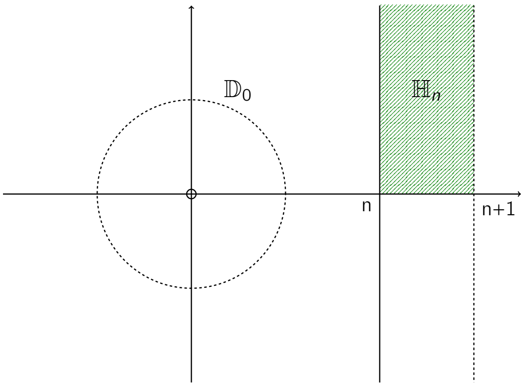


Figure 2.10 – The map q defines a bijection from $\mathbb{H}_n$ to $\mathbb{D}_0$.

Indeed

$$\lim_{\mathrm{Im}(\tau)\to+\infty} q(\tau) = 0 \quad \text{and} \quad \lim_{\mathrm{Im}(\tau)\to 0^+} q(\tau) = \cos(2\pi\mathrm{Re}(\tau)) + i\sin(2\pi\mathrm{Re}(\tau)) = \partial\mathbb{D}$$

Thus, if $f : \mathbb{H} \to \mathbb{C}$ is a meromorphic function of period 1, i.e., such that $f(\tau+1) = f(\tau)$ for any $\tau \in \mathbb{H}$ (we also say it is invariant under T), then it induces a map on the punctured disk such that

$$f(\tau) = f^*(q(\tau))$$

Definition. The q-expansion (or q-series) of $f(\tau)$ is obtained by composing the Laurent series expansion of

f^* at 0 with $q(\tau)$:

$$f(\tau) = f^*(q(\tau)) = \sum_{n=-\infty}^{+\infty} a_n q(\tau)^n = \sum_{n=-\infty}^{+\infty} a_n q^n$$

For f^* to be meromorphic at 0 we ask for the existence of a positive integer n_0 such that $z^{-n_0} f^*(z)$ is bounded near 0. In this case we write

$$f(\tau) = \sum_{n=n_0}^{+\infty} a_n q^n \qquad a_{n_0} \neq 0$$

and we call n_0 the order of f at 0. Further, we say that f is meromorphic at ∞ if f^* is meromorphic at 0.

Definition. Let f be a meromorphic function $f : \mathbb{H} \to \mathbb{C}$ which is Γ-invariant for some congruence subgroup Γ. The function $f(\tau)$ is said to be meromorphic at the cusps if $f(\gamma \cdot \tau)$ is meromorphic at ∞ for every $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

To check whether f is meromorphic at the cusps we need to consider a set of coset representatives $\gamma_1, \ldots, \gamma_n$ of $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$ (which will be finite since the index of Γ in $\mathrm{SL}_2(\mathbb{Z})$ is finite) and check the meromorphicity of f at $\gamma_i\infty$, $i = 1 \ldots, n$.

Remark. A Γ-invariant meromorphic function f satisfies

$$\lim_{\mathrm{Im}(\tau)\to\infty} f(\gamma\tau) = \lim_{\mathrm{Im}(\tau)\to\infty} f(\tau) \qquad \forall\, \gamma \in \Gamma$$

Hence, if f is meromorphic at the cusps, it determines a meromorphic function $g : X_\Gamma \to \mathbb{C}$ on the modular curve $X_\Gamma = \Gamma/\mathbb{H}^*$.

Definition. A modular function f for a congruence subgroup Γ is a complex valued function defined on the upper half plane such that

a. $f(\tau)$ is meromorphic on $\mathbb{H}$;

b. $f(\tau)$ is invariant under Γ, i.e., $f(\gamma \cdot \tau) = f(\tau)$ for all $\tau \in \mathbb{H}$ and every $\gamma \in \Gamma$;

c. $f(\tau)$ is meromorphic at the cusps.

Remark. Constant functions are modular functions for every congruence subgroup. Sums and products of modular functions for Γ are also modular functions for the same congruence subgroup Γ. Thus, the set of modular functions for Γ is a field, denoted $\mathbb{C}(\Gamma)$. This is a transcendental extension of $\mathbb{C}$.

Theorem 2.21. *$\mathbb{C}(\Gamma) \simeq \mathbb{C}(X_\Gamma)$, the function field of X_Γ as an algebraic curve over $\mathbb{C}$.*

Remark. If f is a modular function for a congruence subgroup Γ, it is a modular function for every congruence subgroup $\Gamma' \subseteq \Gamma$. This yields

$$\mathbb{C}(\Gamma) \subseteq \mathbb{C}(\Gamma')$$

and the corresponding inclusion of function fields $\mathbb{C}(X_\Gamma) \subseteq \mathbb{C}(X_{\Gamma'})$ induces a morphism of algebraic curves

$$X_{\Gamma'} \longrightarrow X_\Gamma$$

Modular functions for $\Gamma(1)$

Proposition 2.22. *The j-function*

$$j(\tau) = \frac{1728 g_2(\tau)^3}{\Delta(\tau)}$$

is a modular function for $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$.

Proof. See [Cox, III.11.A]. □

Remark. All the cusps are $\mathrm{SL}_2(\mathbb{Z})$-equivalent. Thus, showing that $j(\tau)$ is meromorphic at the cusps reduces to proving that it is meromorphic at ∞.

Lemma 2.23 ([Cox, Th. III.11.8])**.** *The j-function is a modular function for $\Gamma(1)$ with a simple pole at infinity. Its q-expansion is*

$$j(q) = \frac{1}{q} + 744 + 196884q + \ldots = \frac{1}{q} + \sum_{n=0}^{+\infty} a_n q^n \qquad a_n \in \mathbb{Z}$$

The j function defines a holomorphic bijection $j : Y(1) = \Gamma(1)/\mathbb{H} \to \mathbb{C}$ [Lan2, Th. 3.3.4] and extends to an isomorphism $X(1) \xrightarrow{\sim} \mathbb{P}^1(\mathbb{C}) \simeq S^1_{\mathbb{C}}$ which is holomorphic everywhere but at ∞ where it has a simple pole. This can be done easily by setting $j(\infty) = \infty$; indeed it can be proved that $\lim\limits_{\mathrm{Im}(\tau)\to\infty} j(\tau) = \infty$.

Lemma 2.24. *Every holomorphic modular function for $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$ is a polynomial in $j(\tau)$.*

Theorem 2.25. *Every modular function for $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$ is a rational function in $j(\tau)$, i.e., $\mathbb{C}(\Gamma(1)) = \mathbb{C}(j)$.*

Proof. See [Apo, §2.6]. □

We conclude (following Theorem 2.21) that the function field of the modular curve $X(1)$ is $\mathbb{C}(j)$. A discussion about this can be found in [DS, §3.2].

Modular functions for $\Gamma_0(N)$

We describe the function field of the modular curve $X_0(N)$. First we note that

Theorem 2.26. *The function $j_N(\tau) = j(N\tau)$ is a modular function for $\Gamma_0(N)$.*

More importantly, we obtain the following theorem [DS, Proposition 7.5.1]

Theorem 2.27. *The field of modular functions for the congruence subgroup $\Gamma_0(N)$ is an extension of $\mathbb{C}(j)$ of degree $n = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)]$ generated by $j_N(\tau)$.*

Theorem 2.28 ([DS, §7.7])**.** *The modular curve $X_0(N)$ is defined over $\mathbb{Q}$ and its function field over $\mathbb{Q}$ is $\mathbb{Q}(j, j_N)$.*

In proving Theorem 2.27, we come across the minimal polynomial of j_N over $\mathbb{C}(j)$; this is

$$\Phi_N(j(\tau), y) = \prod_{i=i}^{m} (x - j(N\gamma_i\tau))$$

for $\{\gamma_1, \ldots, \gamma_m\}$ a set of representatives of right cosets of $\Gamma_0(N)$ in $\Gamma(1)$. Since polynomials in $J(N\gamma_i\tau)$ are holomorphic rational functions in $j(\tau)$, they are polynomials in j (Lemma 2.24) meaning that $\Phi_N(j(\tau), y) \in \mathbb{C}[j, y]$. Substituting $j(\tau)$ with the variable x we obtain $\Phi_N(x, y) \in \mathbb{C}[x, y]$.

2.2.2 The modular polynomial

Definition. $\Phi_N(x, y)$ is called the classical modular polynomial of degree N. The equation $\Phi_N(x, y) = 0$ is called the classical modular equation.

Theorem 2.29 (Properties of the classical modular polynomial)**.** *Let N be a positive integer.*

(1) $\Phi_N(x, y) \in \mathbb{Z}[x, y]$*;*

(2) $\Phi_N(x, y)$ *is symmetric in x and y when $N > 1$, i.e., $\Phi_N(x, y) = \Phi_N(y, x)$;*

(3) $\Phi_N(x, y)$ *is irreducible as a polynomial in x (or y);*

(4) *if N is not a perfect square, then $\Phi_N(x, y)$ is a polynomial of degree greater than 1 whose leading coefficient is 1;*

(5) *(Kronecker congruence relation) if p is a prime, then $\Phi_N(x, y) \equiv (x^p - y)(x - y^p) \mod p$.*

Proof. See [Cox, §11.C] and [Lan2, §5.2]. □

Theorem 2.30. *The classical modular equation $\Phi_N(x, y) = 0$ is a model for the curve $X_0(N)$.*

Proof. See [Miln, §V.2]. □

Modular Polynomials play an essential role when dealing with elliptic curve isogenies. The reason why they are so important is enclosed in the following theorem:

Theorem 2.31. *Let E_0 and E_1 be two elliptic curves over a field k. There exists an isogeny $E_0 \to E_1$, possibly defined over some extension of k, with cyclic kernel of degree N if and only if the j-invariant j_0 of E_0 is a root of*

$$\Phi_N(x, j_1) = 0$$

The modular polynomial defines a correspondence in $X(1) \times X(1)$. The curve in $X(1) \times X(1)$ cut out by $\Phi_N(x, y)$ is a singular image of the modular curve $X_0(N)$ parametrizing pairs (E, φ) of an elliptic curve and a cyclic isogeny of degree N with domain E. Controlling modular polynomials gives an easy way to find isogenies between elliptic curves. Therefore they have been intensively used in cryptography and Elliptic curve based cryptosystems. The drawback is that they are not very handy. As the degree N grows, the size of the coefficients of the polynomial Φ_N increases dramatically. This motivates us to look for a different approach to the problem of finding models for $X_0(N)$.

Examples

To give an idea of how fast the coefficients of modular polynomials explode we write here few examples for very small degree.

$$\begin{aligned}\Phi_2(x, y) = x^3 - x^2 \cdot y^2 + 1488 \cdot x^2 \cdot y - 162000 \cdot x^2 + 1488 \cdot x \cdot y^2 + 40773375 \cdot x \cdot y + \\ + 8748000000 \cdot x + y^3 - 162000 \cdot y^2 + 8748000000 \cdot y - 157464000000000\end{aligned}$$

$$\begin{aligned}\Phi_3(x, y) = x^4 - x^3 \cdot y^3 + 2232 \cdot x^3 \cdot y^2 - 1069956 \cdot x^3 \cdot y + 36864000 \cdot x^3 + 2232 \cdot x^2 \cdot y^3 + \\ + 2587918086 \cdot x^2 \cdot y^2 + 8900222976000 \cdot x^2 \cdot y + 452984832000000 \cdot x^2 - 1069956 \cdot x \cdot y^3 + \\ + 8900222976000 \cdot x \cdot y^2 - 770845966336000000 \cdot x \cdot y + 1855425871872000000000 \cdot x + \\ + y^4 + 36864000 \cdot y^3 + 452984832000000 \cdot y^2 + 1855425871872000000000 \cdot y\end{aligned}$$

and $\Phi_5(x, y)$ has 38 monomials

$$\Phi_5(x, y) = x^6 + \ldots + 141359947154721358697753474691071362751004672000$$

A database of modular polynomials up to degree 300 can be found at [Sut2]. For a precise analysis of the growth of these coefficients one could refer to [BS2].

2.2.3 Dedekind eta function

We define the Dedekind eta-function as the infinite product

$$\eta(\tau) = q^{\frac{1}{24}} \prod_{n=1}^{+\infty} (1 - q^n) \qquad q = e^{2\pi i \tau}$$

The product converges normally for q in the unit disc and from section 2.2.1 we know that this is equivalent to $\tau \in \mathbb{H}$. This implies that η is a holomorphic function on $\mathbb{H}$.

The eta-function is one of the best known and most studied modular functions. Thanks to Euler's identity we know how to find its series expansion:

$$\eta(\tau) = \sum_{n=1}^{+\infty} \left(\frac{12}{n}\right) q^{\frac{n^2}{24}}$$

where $\left(\frac{12}{n}\right)$ denotes the Legendre-Kronecker symbol of quadratic reciprocity.

One of the reasons the Dedekind Eta-function is so attractive comes from the fact that raising it to the 24^{th}-power gives the Weierstrass modular discriminant [Kil, §5.2]:

$$\Delta(\tau) = \eta(\tau)^{24}$$

which is a modular form of weight 12. The latter is a remarkable function, having the properties that $\Delta(\tau)$ is precisely the discriminant of the elliptic curve E_τ in the Weierstrass form, it is non-vanishing on $\mathbb{H}$, it has a simple zero at the cusps and it is related to the j-function:

$$j(\tau) = \frac{E_4^3(\tau)}{\Delta(\tau)} \qquad \text{for} \qquad E_4(q) = 1 + 240\sum_{n=1}^{+\infty} \sigma_3(n)q^n \qquad \sigma_3(n) = \sum_{d|n} d^3$$

Further discussions on the relations with important quantities in number theory can be found in [Kno, §3.1-2].

We are interested in the transformation properties of the η-function. This will enable us to study its behavior at the cusps of $X_0(N)$. We find

Theorem 2.32. *The Dedekind η-function satisfies the following transformation properties:*

$$\eta(\tau+1) = e^{\frac{\pi i}{12}}\eta(\tau) \qquad \textit{for and} \qquad \eta\left(-\frac{1}{\tau}\right) = \sqrt{i\tau}\,\eta(\tau)$$

where we consider the principal branch of the square root.

Proof. See [Köh, §1.3] and [CS2, Theorem 5.8.1]. □

Note that the two transformations showed are the special ones. We already know that the two matrices

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

generates the modular group. From $\eta(T\cdot\tau) = e^{\pi i/12}\eta(\tau)$ and $\eta(S\cdot\tau) = e^{\pi i/4}(z)^{1/2}\eta(\tau)$, it follows that, in general, the η-function must satisfy

$$\eta(\gamma\cdot\tau) = \nu_\eta(\gamma)\,(c\tau+d)^{1/2}\,\eta(\tau)$$

for any matrix $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.
The map associating $\gamma \to \nu_\eta(\gamma)$ is called a multiplier system for η. A remarkable observation is that both $\nu_\eta(T)$ and $\nu_\eta(S)$ are 24^{th} roots of unity meaning that $\nu_\eta(\gamma)$ is a 24^{th} root of unity for any γ.

Theorem 2.33. *Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, the multiplier system of the η-function is given by*

$$\nu_\eta(\gamma) = \begin{cases} \left(\dfrac{d}{|c|}\right)\exp\left(\dfrac{2\pi i}{24}\left((a+d-3)c - bd(c^2-1)\right)\right) & \textit{if } c \textit{ is odd} \\ \epsilon(c,d)\left(\dfrac{d}{|c|}\right)\exp\left(\dfrac{2\pi i}{24}\left((a-2d)c - bd(c^2-1)+3d-3\right)\right) & \textit{if } c \textit{ is even} \end{cases}$$

where $\epsilon(c,d) = -1$ if $c \leq 0$ and $d < 0$ and $\epsilon(c,d) = 1$ otherwise.

Proof. It was first proved by Rademacher in 1931 (see chapter 9 of his book [Rad]) and then by Petersson in 1954. Other proofs can be found in [CS2, Theorem 5.8.1], [Kno, §4.1] and [Apo, §3.4]. □

2.2.4 Eta products and eta quotients

An η-product is an expression of the form

$$f(\tau) = \prod_m \eta(m\tau)^{a_m} = \prod_m \eta(q^m)^{a_m}$$

where m runs over a finite set of positive integers and the exponents $a_m \in \mathbb{Z}$.

Sometimes we might see the notation

$$f(\tau) = \prod_{m|N} \eta(m\tau)^{a_m}$$

where the index runs over the set of positive divisors of N, which is called the level of the η-product. Passing from one to the other is easy since the set of integers in the index is finite and we can therefore consider N as the lowest common multiple (of course we might need some exponents to be 0 as we might introduce new divisors).

The transformation properties of an η-product are of course strictly related to the ones for the η-function.

$$f(\gamma \cdot \tau) = \nu_f(\gamma)(c\tau + d)^a f(\tau)$$

for $a = \frac{1}{2}\sum_m a_m$.
For the sake of completeness we refer to [Köh, §2.1] and [CS2, Th. 5.9.2] for the computation of $\nu_f(\gamma)$:

$$\nu_f(\gamma) = \prod_{m|N} \left(\nu_\eta \begin{pmatrix} a & mb \\ c/m & d \end{pmatrix} \right)^{a_m}$$

In general, we will not be interested in the exact value of $\nu_f(\gamma)$ but in one special case: $\gamma = T$. We have already seen that $\eta(\tau + 1) = e^{\frac{2\pi i}{24}}\eta(\tau)$ meaning

$$f(\tau + 1) = \prod_{m|N} \eta(m(\tau + 1))^{a_m} = \prod_{m|N} e^{\frac{2\pi i}{24}\cdot m a_m}\eta(m\tau) = e^{\sum_{m|N}\frac{2\pi i\ m\ a_m}{24}} \prod_{m|N} \eta(m\tau)$$

and now we infer

$$\nu_f(T) = e^{2\pi i \frac{s}{t}} \qquad \text{where} \qquad \frac{s}{t} = \frac{1}{24}\sum_{m|N} m a_m \qquad (s,t) = 1$$

It turns out that s/t is the order of f at the cusp ∞. We will prove the general statement (that can be found in [Köh, §2.3] or [CS2, §5.9]).

Theorem 2.34. *Let $\eta_m(\tau) = \eta(m\tau)$ with $m \in \mathbb{Z}_{\geq 0}$, and let $\mathfrak{c} = \frac{a}{c} \in \mathbb{Q}$ be a reduced fraction with $c \neq 0$. We choose b and d such that the matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is in $\mathrm{SL}_2(\mathbb{Z})$. Then*

(a) *the expansion of η_m at the cusp $\mathfrak{c}$ is*

$$\eta_m(A\tau) = \nu_\eta(L)\left[\frac{g}{m}(c\tau + d)\right]^{\frac{1}{2}} \sum_{n=1}^{+\infty} \left(\frac{12}{n}\right) e^{\frac{2\pi i}{24m}(g^2 z + \alpha g)}$$

where $g = \gcd(c, m)$, $\alpha \in \mathbb{Z}$ and $L = \begin{pmatrix} x & y \\ u & v \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ for $x = \frac{ma}{\gcd(c,m)}$ and $u = \frac{c}{\gcd(c,m)}$.

(b) *The order of η_m at the cusp $\mathfrak{c}$ is*

$$\mathrm{ord}_{\mathfrak{c}}(\eta_m) = \frac{1}{24m}\gcd(c, m)^2$$

Proof. We know $\eta_m(\tau) = \eta(m\tau)$ from which

$$\eta_m(\gamma\tau) = \eta(m\gamma\tau) = \eta(\gamma_m\tau) \qquad \text{where} \qquad \gamma_m = \begin{pmatrix} ma & mb \\ c & d \end{pmatrix} \qquad \text{with} \qquad \det(\gamma_m) = m$$

Observe that $\gamma \cdot \infty = \frac{a}{c} = \mathfrak{c}$ which implies that the expansion of $\eta_m(\tau)$ at the cusp $\mathfrak{c}$ is given by the expansion of $\eta_m(\gamma\tau)$ at ∞. Hence, we look for a matrix $L \in \mathrm{SL}_2(\mathbb{Z})$ such that $L^{-1}\gamma_m$ has the lower left entry equal to 0.

$$L = \begin{pmatrix} x & y \\ u & v \end{pmatrix} \implies L^{-1} = \begin{pmatrix} v & -y \\ -u & x \end{pmatrix}$$

$$L^{-1}\gamma_m = \begin{pmatrix} v & -y \\ -u & x \end{pmatrix}\begin{pmatrix} ma & mb \\ c & d \end{pmatrix} = \begin{pmatrix} vma - yc & vmb - yd \\ xc - mau & xd - mub \end{pmatrix}$$

we impose $xc - mau = 0$ and we choose $x = \frac{ma}{\gcd(c,m)}$ and $u = \frac{c}{\gcd(c,m)}$; further, we set $g = \gcd(m, a)$. Since we want $L \in \mathrm{SL}_2(\mathbb{Z})$, we compute its discriminant:

$$1 = xv - uy = \frac{mav}{\gcd(c,m)} - \frac{cy}{\gcd(c,m)} \implies mav - cy = g$$

We conclude

$$L^{-1}\gamma_m = \begin{pmatrix} vma - yc & vmb - yd \\ xc - mau & xd - mub \end{pmatrix} = \begin{pmatrix} g & \alpha \\ 0 & m/g \end{pmatrix}$$

for $\alpha = mbv - yd \in \mathbb{Z}$. Now

$$\begin{aligned}
\eta_m(\gamma\tau) &= \eta(\gamma_m\tau) = \eta(LL^{-1}\gamma_m\tau) = \nu_\eta(L)\left[u(L^{-1}\gamma_m\tau) + v\right]^{\frac{1}{2}}\eta(L^{-1}\gamma_m\tau) = \\
&= \nu_\eta(L)\left[u\frac{g\tau+\alpha}{m/g} + v\right]^{\frac{1}{2}}\eta\left(\frac{g\tau+\alpha}{m/g}\right) = \\
&= \nu_\eta(L)\left[\frac{c\tau + \frac{c\alpha}{g}}{m/g} + v\right]^{\frac{1}{2}}\eta\left(\frac{g^2\tau+g\alpha}{m}\right) = \\
&= \nu_\eta(L)\left[\frac{g}{m}\left(c\tau + \frac{c\alpha}{g} + \frac{mv}{g}\right)\right]^{\frac{1}{2}}\eta\left(\frac{g^2\tau+g\alpha}{m}\right) = \\
&= \nu_\eta(L)\left[\frac{g}{m}(c\tau+d)\right]^{\frac{1}{2}}\eta\left(\frac{g^2\tau+g\alpha}{m}\right) = \\
&= \nu_\eta(L)\left[\frac{g}{m}(c\tau+d)\right]^{\frac{1}{2}}\sum_{n=1}^{+\infty}\left(\frac{12}{n}\right)e^{\frac{2\pi i}{24m}(g^2\tau+\alpha g)}
\end{aligned}$$

$$\boxed{\begin{array}{c} c\alpha+mv=mbvc-ycd+mv= \\ =mv(1+bc)-yca=mvad-ycd= \\ =d(mva-yc)=dg \end{array}}$$

This proves the first assertion. Now the first non-vanishing term in $(c\tau+d)^{-1/2}\cdot\eta_m(\gamma\tau)$ is a multiple of

$$e^{\frac{2\pi i}{24m}g^2\tau} = q^{\frac{g^2}{24m}}$$

proving point **(b)**. □

Corollary 2.35. *Let $f(\tau)$ be an eta product, and let $\mathfrak{c} = \frac{a}{c} \in \mathbb{Q}$, $\gcd(a,c) = 1$. Then the order of f at the cusp $\mathfrak{c}$ is*

$$\mathrm{ord}_{\mathfrak{c}}(f) = \frac{1}{24}\sum_{m|N}\frac{\gcd(c,m)^2}{m}a_m$$

Corollary 2.36. *An η-product $f(\tau)$ is holomorphic at $\mathfrak{c} = \frac{a}{c}$ if and only if the above sum is non-negative, and it vanishes at $\mathfrak{c}$ if and only if the sum is positive.*

Definition. An eta product which is holomorphic at all cusps is called a holomorphic eta product.

We are now interested in knowing when an η-product is a modular function. To handle this problem we have an essential tool. This criterion has been introduced by Newman in [New2] and then generalized by Ligozat in [Lig1].
First of all we note that we have a leading term $q^{\frac{1}{24}}$ in the expansion of the η-function and we need to kill it in order to have a function on $X_0(N)$. If we multiply all these terms together we get $q^{\frac{ma_m}{24}}$ which will motivate point **(ii)** of the criterion. We note that this is also the request of the meromorphicity at ∞. Property **(i)** comes from the fact that we need $f(\tau)$ to be invariant under $\Gamma_0(N)$ and therefore we need its weight to be 0 (the exponent of $(c\tau+d)$ in the transformation formula of f). Finally, the study of $\nu_f(\gamma)$ completes the criterion.

Theorem 2.37 (Ligozat-Newman Criterion)**.** *An η-product is a modular function for $\Gamma_0(N)$ if and only if*

(i) $\displaystyle\sum_{m|N} a_m = 0$;

(ii) $\displaystyle\sum_{m|N} m a_m \equiv 0 \mod 24$;

(iii) $\displaystyle\sum_{m|N} \frac{N}{m} a_m \equiv 0 \mod 24$;

(iv) $\displaystyle\prod_{m|N} \left(\frac{N}{m}\right)^{a_m} \in \mathbb{Q}^2$.

Corollary 2.38.A ([Ligozat-Newman Criterion for $N = 2^n$). *An η-product $f(\tau) = \prod_{k=0}^{n} \eta(2^k\tau)^{a_k}$ is a weight 0 modular form (a modular function) of level 2^n if and only if*

(i) $\displaystyle\sum_{0\leq k\leq n} a_k = 0$;

(ii) $\displaystyle\sum_{0\leq k\leq n} 2^k a_k \equiv 0 \mod 24$;

(iii) $\displaystyle\sum_{0\leq k\leq n} 2^{n-k} a_m \equiv 0 \mod 24$;

(iv) $\displaystyle\prod_{0\leq k\leq n} 2^{(n-k)a_k} \in \mathbb{Q}^2$.

Note that

$$\mathbb{Q}^2 \ni \prod_{m|N} 2^{(n-k)a_k} = 2^{\sum_{k=0}^{n} a_k(n-k)} \iff \sum_{k=0}^{n} a_k(n-k) \equiv 0 \mod 2 \iff \sum_{k=0}^{n} k a_k \equiv 0 \mod 2$$

the last equivalence keeping into account property **(i)**.

Corollary 2.38.B ([Ligozat-Newman Criterion for $N = 2^n$). *An η-product $f(\tau) = \prod_{k=0}^{n} \eta(2^k\tau)^{a_k}$ is a weight 0 modular form (a modular function) of level 2^n if and only if*

(i) $\displaystyle\sum_{0\leq k\leq n} a_k = 0$;

(ii) $\displaystyle\sum_{0\leq k\leq n} 2^k a_k \equiv 0 \mod 24$;

(iii) $\displaystyle\sum_{0\leq k\leq n} 2^{n-k} a_k \equiv 0 \mod 24$;

(iv) $\displaystyle\sum_{0\leq k\leq n} k a_k \equiv 0 \mod 2$.

Observe that if $f(\tau)$ is a modular function for $\Gamma_0(N)$ and since the width of the cusp $\mathfrak{c} = \frac{a}{c}$ on $\Gamma_0(N)$ is $h_{\mathfrak{c}} = \frac{N/c}{(c,N/c)}$ (Proposition 2.19), then the order of vanishing of f at $\mathfrak{c}$ is given by

$$\frac{N}{24c(N/c,c)} \sum_{m|N} \frac{(c,m)^2}{m} a_m$$

There are many interesting questions that can be asked about eta-functions and eta-products. We refer to [CS2, §5.9] for some theorems providing a sort of classification for interesting η-products.

Another useful fact about η-functions is the following:

Theorem 2.39 ([Ono, Th. 1.67]). *Every modular form on* $\mathrm{SL}_2(\mathbb{Z})$ *may be expressed as a rational function in $\eta(\tau)$, $\eta_2(\tau)$ and $\eta_4(\tau)$.*

In the remaining part of the section we will present some trivial lemmas on greatest common divisors that will enable us to better understand the behavior of the eta-products.

We define a matrix $A(N) = (\alpha_{c,m})_{c,m}$ encoding the data of an η-product $f(\tau) = \prod_{m|N} \eta_m(\tau)^{a_m}$ by

$$\alpha_{cm} = \frac{N}{c(N/c,c)} \cdot \frac{(c,m)^2}{m}$$

where the divisors c and m of N are taken in the natural order 1 to N: for the cusps we go from 0 (representing cusps $\bullet/1$) to ∞ (representing cusps $\bullet/N$) while for the η-function we go naturally from 1 to N. In the same spirit we define the column vector X by $X = (a_m)$. Thus, the study of eta-products reduced to the study of the system

$$A(N) \cdot X$$

together with the conditions **(i)**-**(iv)** of Theorem 2.37.

Remark. The columns of the matrix correspond to a fixed η-function η_m while the rows correspond to the cusps. The entry $\alpha_{c,m}$ corresponds to the order of the function η_m at the cusps $\frac{\bullet}{c}$ (multiplied by 24 for simplicity).

Example. For $N = 8$ we have the following orders of η-functions at each cusp:

	$\eta(\tau)$	$\eta_2(\tau)$	$\eta_4(\tau)$	$\eta_8(\tau)$
0	1/3	1/48	1/12	1/24
•/2	1/12	1/6	1/12	1/24
•/4	1/24	1/12	1/6	1/12
∞	1/24	1/12	1/6	1/3

and therefore the matrix $A(8)$ is

$$A(8) = \begin{pmatrix} 8 & 4 & 2 & 1 \\ 2 & 4 & 2 & 1 \\ 1 & 2 & 4 & 2 \\ 1 & 2 & 4 & 8 \end{pmatrix}$$

Lemma 2.40. *If* $\frac{\gcd(N,c)^2}{N} \in \mathbb{Z}$, *then*

$$\frac{N}{c \gcd(c, N/c)} = 1$$

This implies that, if the order of vanishing of f at ∞ (corresponding to the last row) as a function on the upper half plane is an integer, then f has the same order of vanishing at ∞ as a function on $X_0(N)$. In other words (as we already know) the width of ∞ is 1.

Proof. $\gcd(N, c)^2/N \in \mathbb{Z}$ means that $c^2/N \in \mathbb{Z}$. Hence $N \mid c^2$ and therefore $N/c \mid c$ meaning $\gcd(c, N/c) = N/c$ from which

$$\frac{N}{c \gcd(c, N/c)} = \frac{N}{cN/c} = 1$$

□

Lemma 2.41. *We have*

$$\frac{\gcd(m, c)^2}{m} \cdot \frac{N}{c \gcd(c, N/c)} \in \mathbb{Z}$$

This implies that the order of vanishing of our eta-products are integers, i.e., $A(N)$ has integral entries.

Proof. Since N, c and m are all integers we will suppose that their prime factorization is

$$N = p_1^{e_1} p_2^{e_2} \cdot \ldots \cdot p_k^{e_k} \qquad c = p_1^{f_1} p_2^{f_2} \cdot \ldots \cdot p_k^{f_k} \qquad m = p_1^{g_1} p_2^{g_2} \cdot \ldots \cdot p_k^{g_k}$$

Note that both c and m are divisors of N meaning that the primes involved are the same and for any $i = 1, \ldots, k$,

$$0 \leq g_i, f_i \leq e_i \quad \text{and} \quad e_i > 0$$

Now

$$\frac{\gcd(m, c)^2}{m} \cdot \frac{N}{c \gcd(c, N/c)} = \prod_{i=1}^{k} p_i^{e_i + 2\min(g_i, f_i) - g_i - f_i - \min(g_i, e_i - g_i)}$$

Looking at the gcd's, we divide the problem in 4 cases. Depending on the situation the exponent of p_i is shown in the following table

	$g_i \leq e_i - g_i$	$g_i \geq e_i - g_i$
$f_i \geq g_i$	$e_i + 2g_i - g_i - f_i - g_i =$ $= e_i - f_i \geq 0$	$e_i + 2g_i - g_i - f_i - e_i + g_i =$ $= 2g_i - f_i \underset{g_i \geq e_i - g_i}{\geq} e_i - f_i \geq 0$
$f_i \leq g_i$	$e_i + 2f_i - g_i - f_i - g_i =$ $= e_i + f_i - 2g_i \underset{e_i \geq 2g_i}{\geq} 0$	$e_i + 2f_i - g_i - f_i - e_i + g_i =$ $= f_i \geq 0$

In any case the exponent is bigger than or equal to 0. □

An interesting property of the matrix $A(N)$ is that if we multiply each row by $\gcd(c, N/c)$ (which is equivalent to multiplying the order of vanishing of the η_m's by $\frac{N}{c}$ instead of the width of the cusps) the resulting matrix is symmetric and we will denote it $A_{sym}(N)$.
In some applications, as studying holomorphic modular functions, this matrix might be of better use; if we need to solve the system $A(N) \cdot X \geq 0$, then multiplying the rows by some integer does not affect the result.

2.2.5 Maps on $X_0(N)$

In this section we describe some important maps defined on modular curves. The main references will be [Mcm1], [Gal1] and [Elk1] (for the three paragraphs respectively).

Maps between modular curves

The first thing we can study is how to relate one modular curve to another. In Section 1.3.3, we have used the fact that there is a canonical map $X_0(N) \to X(1)$ for any N. In the following we describe this morphism in more details and we generalize it.
It is clear that, if M divides N, then $\Gamma_0(N)$ is a subgroup of $\Gamma_0(M)$ and, therefore, we have a straightforward map on the quotients

$$\Gamma_0(N)\backslash\mathbb{H} \longrightarrow \Gamma_0(M)\backslash\mathbb{H}$$

given by $\Gamma_0(N)z \to \Gamma_0(M)z$ sending the $\Gamma_0(N)$-class of z to its $\Gamma_0(M)$-class. This map is usually indicated as π_1 and it is called forgetful map.

Definition. In this situation, we say that $X_0(N)$ lies above $X_0(M)$.

The map π_1 is a rational map between two modular curves. As we have seen in Section 1.1.2, every point on $X_0(M)$ but a finite number, has a fixed number n of points lying over it on $X_0(N)$. The number n is the degree of the map and, thanks to [DS, §3.1], we know that

$$\deg \pi_1 = [\Gamma_0(M) : \Gamma_0(N)]$$

We also know that at each of the exceptional points over which there are less than n points, the sum of the multiplicities of the points lying above it must be n in any case. This extra multiplicity known as ramification can occur only at the cusps 1.1.2. More in general:

Definition. Let M, N and d be positive integers such that $Md \mid N$. Then we define the moduli-theoretic map $\pi_d : X_0(N) \to X_0(M)$ by

$$\pi_d\,(E, C) = (E/C[d], C[Md]/C[d])$$

This is a moduli-theoretic map since points on the modular curve $X_0(N)$ are represented as pairs of an elliptic curve together with some extra (N-level) structure which comes in the form of a cyclic subgroup of order N.

Notation. $C[d]$ represents the d-torsion part of C.

π_1 is the same map we have defined at the beginning of the section: the forgetful map. The name comes from the fact that it leaves E unchanged and it simply discards (or forgets) some level structure.

Example. Suppose $N = pq$ for p and q two different primes. We get the following situation:

$$\begin{tikzcd}
 & & X_0(pq) \arrow[dddl, bend right, "\pi_1"'] \arrow[dddl, "\pi_q"] \arrow[dddr, bend left, "\pi_1"] \arrow[dddr, "\pi_p"'] & & \\
q+1 & & & & p+1 \\
 & & & & \\
 & X_0(p) \arrow[dddr, "\pi_p"] \arrow[dddr, bend right, "\pi_1"'] & & X_0(q) \arrow[dddl, "\pi_q"'] \arrow[dddl, bend left, "\pi_1"] & \\
p+1 & & & & q+1 \\
 & & & & \\
 & & X(1) & &
\end{tikzcd}$$

Atkin-Lehner involutions

In 1970 Atkin and Lehner introduced some important operators on $\Gamma_0(N)$ [AL]. Another reference will be [Lan3, Ch. VIII].

Definition. Suppose ℓ is a prime dividing N and let e be an integer such that $\ell^e \mid\mid N$, i.e., $\ell^e \mid N$ and $(\ell^e, N/\ell^e) = 1$. We can choose $a, b, c, d \in \mathbb{Z}$ such that $\ell^e ad - (N/\ell^e)bc = 1$ and we define

$$\mathcal{W}_\ell = \frac{1}{\sqrt{\ell^e}} \begin{pmatrix} \ell^e a & b \\ Nc & \ell^e d \end{pmatrix}$$

$\mathcal{W}_\ell$ is called Atkin-Lehner involution for $\Gamma_0(N)$. It is clearly a matrix in $\mathrm{SL}_2(\mathbb{R})$.

We have some straightforward properties of $\mathcal{W}_\ell$:

Lemma 2.42. *The following are true:*

(a) $\mathcal{W}_\ell^2 \equiv I \mod \Gamma_0(N)$ *(this is why we talk of an involution).*

(b) $\mathcal{W}_\ell$ *normalizes* $\Gamma_0(N)$*, i.e.,* $\mathcal{W}_\ell \Gamma_0(N) \mathcal{W}_\ell^{-1} = \Gamma_0(N)$*.*

(c) $\mathcal{W}_\ell \mathcal{W}_f \equiv \mathcal{W}_f \mathcal{W}_\ell \mod \Gamma_0(N)$*.*

Note that there might be many different choices for $\mathcal{W}_\ell$. For instance, we can always take $d = 1$ or $(c > 0)$:

$$\mathcal{W}_\ell = \frac{1}{\sqrt{\ell^e}} \begin{pmatrix} \ell^e a' & b' \\ Nc' & \ell^e \end{pmatrix}$$

In general, we do not fix a canonical choice since any two of them are equivalent up to $\Gamma_0(N)$-multiplication .

Lemma 2.43 ([AL, Lem. 8])**.** *For any two choices* $\mathcal{W}_\ell$ *and* $\mathcal{W}'_\ell$ *we have*

$$\mathcal{W}_\ell \Gamma_0(N) \mathcal{W}'_\ell = \Gamma_0(N)$$

For any composite number $n \mid N$ we define

$$\mathcal{W}_n = \prod_{\substack{\ell \mid n \\ \ell \text{ prime}}} \mathcal{W}_\ell$$

Remark. If n is a Hall-divisor of N (meaning that $n \mid\mid N$), then $\mathcal{W}_n$ has the same form of $\mathcal{W}_N$.

Remark. The definition of $\mathcal{W}_n$ only depends on the primes dividing n. This implies that there are many different divisors n of N giving the same involution $\mathcal{W}_n$; for example, if $N = 2^3 3^2 5$ then $\mathcal{W}_6 = \mathcal{W}_{12} = \mathcal{W}_{18} = \mathcal{W}_{24} = \mathcal{W}_{36} = \mathcal{W}_{72}$.

The only case in which we fix a canonical representation for the Atkin-Lehner involution is for $n = N$. In this case we set

$$\mathcal{W}_N = \frac{1}{\sqrt{N}} \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$$

By Lemma 2.42.**(b)**, we know that $\mathcal{W}_n$ normalizes $\Gamma_0(N)$. This implies that $\mathcal{W}_n$ defines an involution on the modular curve $X_0(N)$: if $\tau_1, \tau_2 \in \mathbb{H}^*$ are $\Gamma_0(N)$-equivalent, then there exists $\gamma \in \Gamma_0(N)$ such that $\tau_1 = \gamma\tau_2$. Then $\mathcal{W}_n \gamma \mathcal{W}_n^{-1} = \gamma' \in \Gamma_0(N)$ and, thus, $\mathcal{W}_n \tau_1 = \mathcal{W}_n \gamma \tau_2 = \gamma' \mathcal{W}_n \tau_2$ which says that also $\mathcal{W}_n \tau_1$ and $\mathcal{W}_n \tau_2$ are $\Gamma_0(N)$-equivalent.

Definition. We obtain an involution on $X_0(N)$, which is still called Atkin-Lehner involution. We will indicate it by ω_n. Clearly ω_n^2 acts as the identity on the modular curve.

We will conclude this paragraph with few consideration about the quotient $X_0(N)/\omega_n$. We observe that this quotient corresponds to the action of the group $G = \Gamma_0(N) \cup \mathcal{W}_\ell \Gamma_0(N)$ on the upper half plane:

$$X_0^{(n)}(N) = X_0(N)/\omega_n = G \backslash \mathbb{H}^*$$

from which we deduce that $X_0^{(n)}(N)$ is a Riemann surface. Since $\mathcal{W}_n$ is an involution, we have $[G : \Gamma_0(N)] = 2$ and then we get a map

$$\phi_{N,n} : X_0(N) \longrightarrow X_0(N)/\omega_n$$

which is a degree 2 meromorphic map ramified at those $\Gamma_0(N)$-orbits which are fixed by ω_n and mapping cusps to cusps.

Notation. If $n = N$ we write $X_0^+(N) = X_0(N)/\omega_N$.

Towers of modular curves

Let n be a positive integer, ℓ a prime number and K a field of characteristic different from ℓ. The modular curve $X_0(\ell^n)/K$ parametrizes sequences of ℓ-isogenies of length N, i.e., a point on it represents an elliptic curve together with a cyclic ℓ^n-isogeny (or, equivalently, a sequence of ℓ-isogenies) starting from it :

$$E_0 \to E_1 \to E_2 \to \ldots \to \ldots E_{n-1} \to E_n$$

with the property that $E_{i-1} \to E_{i+1}$ is a cyclic ℓ^2-isogeny for any $1 \leq i \leq n-1$ (i.e., it is not a cycle). Now, for each $k = 0, \ldots, n$ we obtain $n + 1 - k$ maps

$$\pi_i : X_0(\ell^n) \longrightarrow X_0(\ell^k) \qquad i \in \{1, \ell, \ldots, \ell^{n-k}\}$$

that extract subsequences of length k from the isogeny chain $E_1 \to \ldots \to E_{n+1}$, i.e., they return the cyclic ℓ^k-isogeny $E_s \to \ldots \to E_{s+k}$ where $s = \log_\ell(i)$. Each of these maps are of the form π_d we have described at the beginning of this section where the index corresponds to the divisor of ℓ^n. This means that any of these maps has degree ℓ^{n-k} unless $k = 0$ in which case it has degree $\ell^{n-1}(\ell + 1)$. Thus, we obtain a tower:

$$\begin{tikzcd}
X_0(\ell^n) \arrow[dr, "\pi_1", "\ell"'] & & & & & \\
& X_0(\ell^{n-1}) \arrow[dr, "\pi_1", "\ell"'] & & & & \\
& & \cdots \arrow[dr, "\pi_1", "\ell"'] & & & \\
& & & X_0(\ell^2) \arrow[dr, "\pi_1", "\ell"'] & & \\
& & & & X_0(\ell^1) \arrow[dr, "\pi_1", "\ell+1"'] & \\
& & & & & X(1)
\end{tikzcd}$$

Each curve $X_0(\ell^k)$ has an Atkin-Lehner involution ω_ℓ which takes a cyclic ℓ^k-isogeny to its dual, i.e., it reverses the sequence.

Lemma 2.44. *For any $n > m$ we have*

$$\omega_\ell^{(m)} \circ \pi_{\ell^i} = \pi_{\ell^{n-m-i}} \circ \omega_\ell^{(n)}$$

meaning that the following diagram commutes:

$$\begin{tikzcd}
X_0(\ell^n) \arrow[r, "\omega_\ell^{(n)}"] \arrow[d, "\pi_{\ell^i}"'] \arrow[dr, phantom, "\circlearrowright"] & X_0(\ell^n) \arrow[d, "\pi_{\ell^{n-m-i}}"] \\
X_0(\ell^m) \arrow[r, "\omega_\ell^{(m)}"'] & X_0(\ell^m)
\end{tikzcd}$$

In [Elk1], Elkies observes that knowing the equations for $X_0(\ell)$ and $X_0(\ell^2)$ together with the involutions $\omega_\ell^{(1)}$ and $\omega_\ell^{(2)}$ and the map $\pi_1 : X_0(\ell^2) \to X_0(\ell)$ suffices to make the whole tower explicit. In general, the explicit form of the Atkin Lehner involution can be deduced using compatibility relations similar to the one of Lemma 2.44. Note that in the easiest case $X_0(p) \to X(1)$ for a prime p, they reduce to $\pi_1 \circ \omega_p = \pi_p$.

$$\begin{tikzcd} X_0(p) \arrow[loop above, "\omega_p"] \arrow[d, bend right, "\pi_1"'] \arrow[d, bend left, "\pi_p"] \\ X(1) \end{tikzcd}
\qquad
\begin{tikzcd} & E \to F \arrow[rrr, bend left, "\omega_p"] \arrow[dl, bend right, "\pi_1"'] \arrow[dr, bend left, "\pi_p"] & & & F \to E \arrow[dl, bend right, "\pi_1"'] \arrow[dr, bend left, "\pi_p"] & \\ E & & F & F & & E \end{tikzcd}$$

2.2.6 Signature of $X_0^+(N)$

We have already introduced the notation

$$X_0^+(N) = X_0(N)/\omega_N$$

describing a modular curve with the following interpretation

Proposition 2.45 ([Gal3, Prop. 1])**.** *A non-cusp rational point on $X_0^+(N)$ corresponds to a pair $\{\psi : E \to F, \hat{\psi} : F \to E\}$*

More in general, we define the group

$$\Gamma_0^*(N) = \langle \Gamma_0(N) \cup \{\mathcal{W}_{N'}\}_{N'||N} \rangle$$

Lemma 2.43 tells us that $\Gamma_0^*(N)$ normalizes $\Gamma_0(N)$.

Lemma 2.46 ([AL, Lemma 9])**.** *The group $W(N) = \Gamma_0^*(N)/\Gamma_0(N)$ is abelian of type $(2, 2, \ldots, 2)$ with order $2^{\omega(n)}$ where $\omega(n)$ represents the number of distinct prime divisors of N.*

Remark. In particular, this lemma tells us that the degree of the quotient map $X_0(N) \to X_0^*(N)$ is $2^{\omega(N)}$.

We note $X_0^*(N) = X_0(N)/W(N) = \Gamma_0^*(N)\backslash\mathbb{H}^*$ and we observe that, for $N = p^r$, $X_0^*(p^r) = X_0^+(p^r)$.

In general, the action of the Atkin-Lehner involution ω_M on $X_0(N)$ (or of $\mathcal{W}_M$ on $\Gamma_0(N)\backslash\mathbb{H}^*$) is given by the modular description

$$(E, C) \longrightarrow (E/C[M], (C + E[M])/C[M])$$

Let us now focus on the modular tower $\{X_0(p^r)\}$; note that we can always define the degeneracy morphism induced by $z \to pz$ on the upper half plane:

$$\begin{aligned} \pi_p = \pi_{p^k, p^{k-2}} : X_0(p^k) &\longrightarrow X_0(p^{k-2}) \\ (E, C) &\longrightarrow (E/p^{k-1}C, pC \mod p) = (E/C[p], C[p^{k-1}]/C[p]) \end{aligned}$$

This induces a degeneracy morphism

$$X_0^+(p^k) \longrightarrow X_0^+(p^{k-2})$$

since $\pi_p \circ \omega_{p^k} = \omega_{p^{k-2}} \circ \pi_p$.

The degeneracy morphism shows that in many situations one could study only the cases p^2 and p^3 to recover information on the whole tower $\{X_0^+(p^r)\}$.

$$\begin{tikzcd}
 & & & & & & \cdots \arrow[dl] \\
 & & & & & X_0(p^5) \arrow[dl, "\pi"', "p"] \arrow[dr, "2"] \arrow[ddll, bend right, dotted, "\pi_p"'] & \\
 & & & & X_0(p^4) \arrow[dl, "\pi"', "p"] \arrow[dr, "2"] \arrow[ddll, bend right, dashed, "\pi_p"'] & & X_0^+(p^5) \arrow[ddll, bend left, dotted] \\
 & & & X_0(p^3) \arrow[dl, "\pi"', "p"] \arrow[dr, "2"] & & X_0^+(p^4) \arrow[ddll, bend left, dashed] & \\
 & & X_0(p^2) \arrow[dl, "\pi"', "p"] \arrow[dr, "2"] & & X_0^+(p^3) & & \\
 & X_0(p) \arrow[dl, "\pi"', "p+1"] & & X_0^+(p^2) & & & \\
X(1) & & & & & &
\end{tikzcd}$$

The goal for the rest of this section is to give a brief description of the cusps of $X_0^+(N)$ and to give a formula for its genus.

Lemma 2.47. *If N is square-free, then $X_0^*(N) \setminus Y_0^*(N)$ consists of a single cusp.*

The number of cusps of $X_0(N)$ is given by $\sum_{d|N} \varphi((d, N/d))$ (see §2.1.5). If N is square-free, we can write $N = p_1 p_2 \dots p_t$ with $p_i \neq p_j$ for $i \neq j$. This means that all the divisors of N are given by all the k-combinations of the p_i's for $k \in \{0, \dots, t\}$ (the 0-combination corresponds to 1). Thus, N has

$$\sum_{i=0}^{t} \binom{t}{i} = 2^t$$

divisors. Further, for each divisor d of N, $(d, N/d) = 1$ since any prime divisor appears only one time in the factorization. Thus, $\varphi((d, N/d)) = 1$ for any divisor d of N. In conclusion, $X_0(N)$ has number of cusps

$$\sum_{d|N} \varphi((d, N/d)) = \sum_{d|N} 1 = \#\{\text{divisors of } N\} = \sum_{i=0}^{t} \binom{t}{i} = 2^t$$

On the other hand, the map $X_0(N) \to X_0^+(N)$ has degree $2^{\omega(N)} = 2^t$.

The Lemma follows if we consider that none of the cusps of $X_0(N)$ are fixed by any of the Atkin-Lehner involutions:

Lemma 2.48 (Ogg)**.** *Let $N = N'N''$ be an integer such that $(N', N'') = 1$. $\mathcal{W}_{N'}$ has no fixed point at cusps (given $N' > 1$), except for the case $N' = 4$.*

In particular, for N square-free, there is no fixed cusp. Since points on $X_0(N)$ are ramified only if they are fixed by some Atkin-Lehner involution, we conclude that there must be only one cusp on $X_0^+(N)$.
Figure 2.11 represents the ramification diagram of the map $\pi^+ : X_0^+(p^4) \to X_0^+(p^2)$ for $p \neq 2$. As we can see there is no fixed cusp in the two quotient maps $\phi_{2,2} : X_0(p^2) \to X_0^+(p^2)$ and $\phi_{4,4} : X_0(p^4) \to X_0^+(p^4)$. The curve $X_0(p^2)$ has $\sum_{k=0}^{2} \varphi((p^k, p^{2-k})) = 1 + (p-1) + 1 = p + 1$ cusps: 0; ∞ and i/p for $i \in \{1, \dots, p-1\}$ (circled in blue). Going up in the tower $\{X_0(p^k)\}$ via the map $\pi_p : X_0(p^3) \to X_0(p^2)$, we see that 0 splits and has p cusps above: 0 and i/p for $i \in \{1, \dots, p-1\}$ while ∞ ramifies. Each $i/p \in X_0(p^2)$ also ramifies; above i/p we have i/p^2.
The next step is given by the map $\pi_1 : X_0(p^4) \to X_0(p^3)$. The situation switches: 0 and all the cusps i/p ramifies while ∞ splits; above ∞ we find ∞ and i/p^3 for $i \in \{1, \dots, p-1\}$. Finally, above each cusp i/p^2 of $X_0(p^3)$ we find p cusps j/p^2 for $j \in \left(\mathbb{Z}/p^2\mathbb{Z}\right)^*$ with $j \equiv i \mod p$. The map $\phi_{2,2}$ is a degree 2 map which identifies 0 and ∞ and all the pairs x/p and $-x^{-1}/p$. In particular, this means that $X_0^+(p^2)$ has $(p+1)/2$ cusps.
On $X_0^+(p^4)$, instead, the Atkin-Lehner involution identifies 0 and ∞, x/p and $-x^{-1}/p^3$ and x/p^2 with $-x^{-1}/p^2$. This tells us that there are $1 + (p-1) + p(p-1)/2 = p(p+1)/2$ cusps on $X_0^+(p^4)$.

Remark. We observe that we could also obtain $\pi^+ = \pi_p \circ \pi_1$ instead of $\pi_1 \circ \pi_p$.

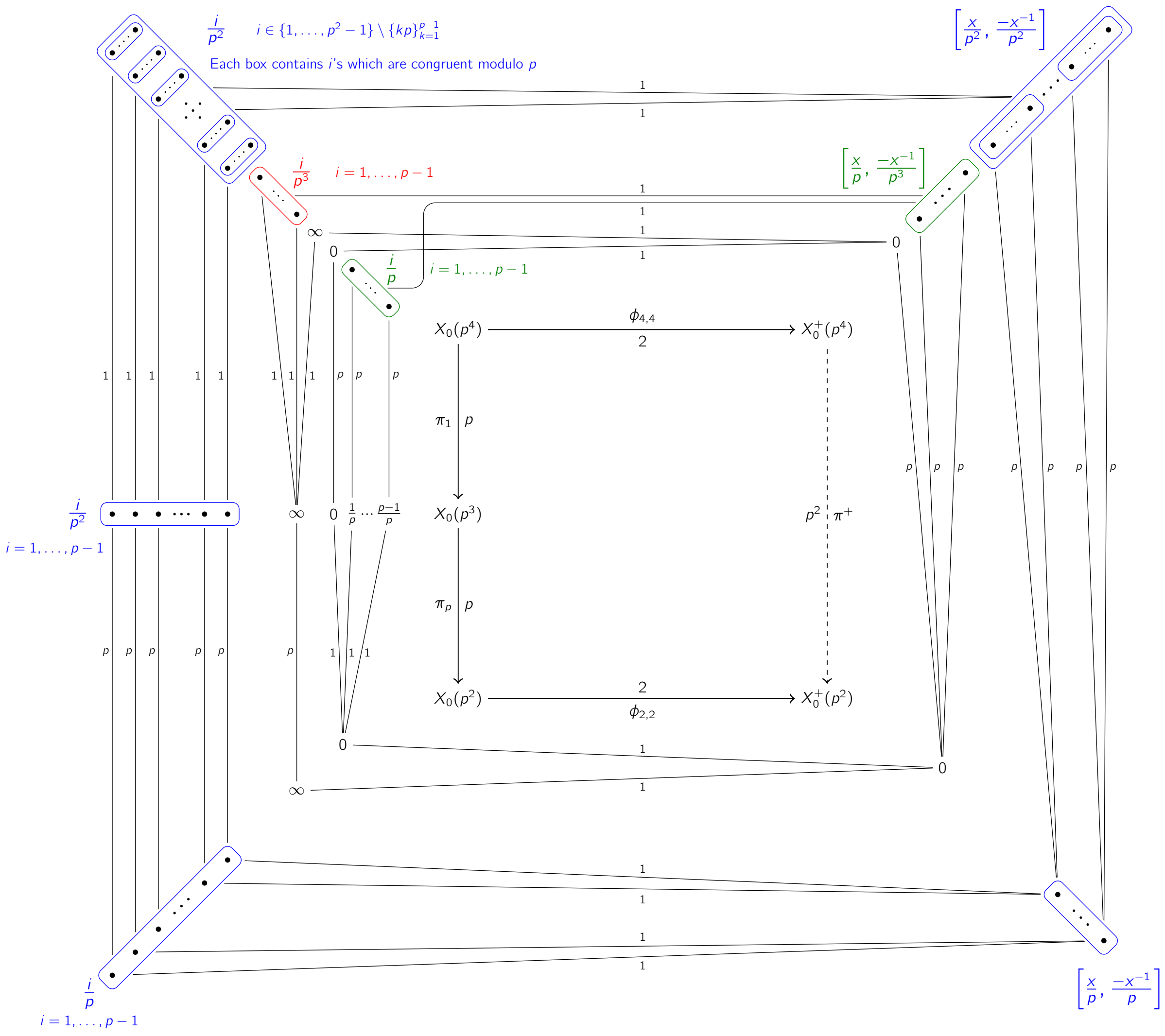

$\frac{i}{p^2}$ $i \in \{1,\dots,p^2-1\} \setminus \{kp\}_{k=1}^{p-1}$
Each box contains i's which are congruent modulo p
$\left[\frac{x}{p^2}, \frac{-x^{-1}}{p^2}\right]$
$\frac{i}{p^3}$ $i = 1,\dots,p-1$
$\left[\frac{x}{p}, \frac{-x^{-1}}{p^3}\right]$
∞
0
$\frac{i}{p}$ $i = 1,\dots,p-1$
$X_0(p^4)$
$\phi_{4,4}$
2
$X_0^+(p^4)$
π_1 p
$X_0(p^3)$
π_p p
$X_0(p^2)$
$\phi_{2,2}$
$X_0^+(p^2)$
p^2 π^+
$\frac{i}{p^2}$ $i = 1,\dots,p-1$
0 $\frac{1}{p} \cdots \frac{p-1}{p}$
1
p
$\frac{i}{p}$ $i = 1,\dots,p-1$
$\left[\frac{x}{p}, \frac{-x^{-1}}{p}\right]$

Figure 2.11 – An example of the ramification diagrams for Atkin-Lehner quotients.

Remark. The same diagram can be constructed for odd powers of p.

Remark. The case $p = 2$ is a special one since we have seen that $X_0(4) \to X_0^+(4)$ has a ramified cusp.

We conclude this section with a formula for computing the genus of Atkin-Lehner quotients. By Riemann-Hurwitz Formula 1.31 we know that

$$g(X_0(N)) = \frac{1}{2}\left(2 - 2d + 2dg(X_0^+(N)) + \sum_{P \in R}(e_p - 1)\right)$$

where d is the degree of the surjective map $\pi : X_0(N) \to X_0^+(N)$ and R is the set of its critical points. In our case $d = 2$ and $R = \{\text{fixed points of } \omega_N\}$. Further, since the degree of the map is 2, the ramification points have all ramification degree 2.

$$g(X_0^+(N)) = \frac{1}{4}\left(2g(X_0(N)) + 2 - \#\{\text{fixed points of } \omega_N\}\right)$$

which yields

$$g(X_0^+(N)) = \frac{g(X_0(N)) + 1}{2} - \frac{\#\{\text{fixed points of } \omega_N\}}{4}$$

Remark. We note that the same formula remains valid for any involution ω_M for $M \mid\mid N$.

Remark. For $X_0^*(N) = X_0(N)/W(N)$ the formula becomes

$$g(X_0^*(N)) = \frac{g(X_0(N)) + 1}{2^{\omega(n)-1}} - \frac{1}{2^{\omega(n)}} \sum_{1<d||N} \#\{\text{fixed points of } \omega_d\}$$

For simplicity we denote $\nu(N, d) = \#\{\text{points on } X_0(N) \text{ by } \omega_d\}$.

It remains to compute the number of fixed points of Atkin-Lehner involutions. We are mainly interested in the case $X_0^+(N)$, i.e., $d = N$.

Theorem 2.49 (Fricke). *For $N > 5$,*

$$\nu(N, N) = \begin{cases} h(-4N) & \text{if } N \not\equiv -1 \mod 4 \\ h(-4N) + h(-N) & \text{if } N \equiv -1 \mod 4 \end{cases}$$

Remark. We refer to [Klu] and [Ken] for the explicit formula in the general case. One could also look at the tables in [BT] for explicit numerical examples. Finally, in [Mor, Prop. 2.3] we find an equivalent description in the case $N = \ell$.

Corollary 2.50. *We have $g^+(p) = 0$ for $p \leq 31$ or $p \in \{41, 47, 59, 71\}$.*

Gonzalez and Lario propose an exhaustive list of integers N (not necessarily prime) for which $X_0^*(N)$ is of genus 0 and 1 [GL, Prop. 3.1/2]. For genus 2 Atkin-Lehner quotients we refer to [Has2, Rem. 1] and [Has1].

2.2.7 Equations for $X_0(N)$

A model for a modular curve is a scheme together with an isomorphism from its generic fiber (a $\mathbb{C}$-valued points) to a classical modular curve over $\mathbb{C}$. For this section we mainly follow [Liu] and [Har] for generalities on models of algebraic curves and [Mcm1] and [Yan2] with regards to models for modular curves specifically.

For arithmetic purposes and in view of explicit computations we need something more explicit. For this reason we embed this scheme into some projective space. This will enable us to exploit explicit calculations in terms of coordinate functions. We have seen in 1.1.3 that this can be done by choosing global sections of some invertible sheaf of the curve. This gives explicit immersions of our curve in the projective space.

In [Gal1] Galbraith used the canonical morphism which, as we have seen, comes from the invertible sheaf of holomorphic differentials. This is an injective morphism $X \to \mathbb{P}^{g-1}$ defined by $P \to [\omega_0(P) : \ldots : \omega_0 g - 1(P)]$ for a basis $\{\omega_i\}_{i=0}^{g-1}$ of the canonical line bundle Ω_X^1. If X is a modular curve for Γ, then this map turns out to be equivalent to $\tau \to [f_0(\tau) : \ldots : f_{g-1}(\tau)]$ for a basis $\{f_i\}_{i=0}^{g-1}$ of the space $S_2(\Gamma)$ of cusp forms of weight 2 on Γ. In other words the sheaf of holomorphic differentials is equivalent to the sheaf of cusp forms of weight 2 [Shi, Cor. 2.17]. To get a system of defining equations for the modular curve $X_0(N)$, we will search for linear relations between monomials in $f_0, \ldots, f_{g-1}$ [Shm, §2.2] and [Gal1, §3.1]. Another sheaf that can be used is the sheaf of holomorphic differentials with at most simple poles at the cusps; this is equivalent to the sheaf of holomorphic weight 2 forms on the modular curve.

Note that this method has some drawbacks: first of all it does not work for curves of genus 0 since the canonical embedding is not defined; but it turns out that it does not work for curves of genus 1 and 2 either because the canonical map does not provide enough information. Secondly, this method does not work for hyperelliptic modular curves since the canonical map is not injective. However there are other methods to deal with these curves.

Proposition 2.51 (Ogg). *There are 19 values of N for which the modular curve $X_0(N)$ is hyperelliptic.*

$g = 2$	$N = 22, 23, 26, 28, 29, 31, 37, 50$
$g = 3$	$N = 30, 33, 35, 39, 40, 41, 48$
$g = 4$	$N = 47$
$g = 5$	$N = 46, 59$
$g = 6$	$N = 71$

Systems of equations for hyperelliptic modular curves can be found using methods of Shimura [Shm], Galbraith [Gal1], Murabayashi [Mur] and Gonzalez [Gon]. More information can be found on the web page of [Gal1].

In general, the invertible sheaf we decide to work with is determined by the kind of functions or differentials that we have available. In the following we will work with η-products which are weight 0-forms with well known q-expansion and whose divisors are easy to compute (see 2.2.4); thus we will work directly with $\mathcal{L}(D)$. The reason why divisors are so important is that they provide a hint on the degree of the relations between the rational functions: the form of their divisors determines the form of the equation of the modular curve.

In the following we describe a method to solve the problem of finding equation for modular curves using algebraic relations between η-products. The nice feature of these equations is that they will inherit the desirable property of modular polynomials of encoding information about isogenies between elliptic curves. Further, we will see how to make explicit the maps π_d and ω_ℓ.

Definition. The η-products that we will use are called parameters on our modular curve.

There are some main advantages of using this method:

- First of all the η-products (or quotients) have divisor supported on the cusps and the ones we will use have poles only at infinity making all the computations relatively easier.

- All the equations will be plane curves, which is preferable when using them in practice.

- Even more importantly, the same method can be used to find equations for other modular curves, such as $X(N)$, $X_0(N)$, $X_0^+(N)$ or $X_1(N)$, regardless of the genus and the nature (as it was for the canonical embedding).

- We do not need the knowledge of the basis of cusps form as it was in the method of Galbraith [Gal1]. For computational purposes, much information can be found in Stein's database [Ste2].

- Finally, this method will provide a way of finding a basis for $S_2(\Gamma)$ [Yan2].

The drawback is that, in practice, as the genus grows, the computation can be really slow in terms of computer time.

Curves of genus 0

If $X_0(N)$ is of genus zero, then its function field will be generated by a single rational function. This is called an Hauptmodul for $\Gamma_0(N)$ and it is indicated by t_N. In this case, our modular curve $X_0(N) = \Gamma_0(N)\backslash\mathbb{H}^*$ can be identified with the Riemann sphere $S^1_{\mathbb{C}}$ and the Hauptmodul is the modular function allowing this identification.

Lemma 2.52. *The following results are known*

- *$X_0(N)$ has genus 0 if and only if $N \in \{1, 2 \ldots, 10, 12, 13, 16, 18, 25\}$;*

- *$X_1(N)$ has genus 0 if and only if $N \in \{1, 2 \ldots, 10, 12\}$;*

- *$X(N)$ has genus 0 if and only if $N \in \{1, 2, 3, 4, 5\}$;*

Hauptmoduls for the curves $X_0(N)$ in terms of η-quotients can be found in [Mai].

Curves of genus > 0

Let X is a modular curve of genus greater than 0 and $f \in K(X)$ a rational function. We have seen at the end of Section 1.1.1 that $\mathrm{div}(f)_\infty$ encodes the information about the poles of f; we define $\deg_\infty(f) = \deg(\mathrm{div}(f)_\infty)$ the total number of poles of f counted with their multiplicities.

Theorem 2.53 ([Yan2])**.** *Let x and y be two rational functions on X such that $(\deg_\infty(x), \deg_\infty(y)) = 1$. Then*

$$K(X) = \mathbb{C}(x, y)$$

and, therefore, a defining equation of x is of the form $F(x, y) = 0$ for $F \in \mathbb{C}[x, y]$ a polynomial of degree $\deg_\infty(y)$ in x and of degree $\deg_\infty(x)$ in y.

Proof. Let us fix $\deg_\infty(x) = n$ and $\deg_\infty(x) = m$

We know by Proposition 8.4 of [Ful] that $[K(X):\mathbb{C}(x)] = n$ and $[K(X):\mathbb{C}(y)] = m$.
This means that $[K(X):\mathbb{C}(x,y)]$ divides both m and n. Since by hypothesis $(n,m) = 1$, then $[K(X):\mathbb{C}(x,y)] = 1$ meaning that $K(X) = \mathbb{C}(x,y)$.

$$\begin{tikzcd}
 & K(X) \arrow[d, dash] \arrow[ddl, dash] \arrow[ddr, dash] & \\
 & \mathbb{C}(x,y) \arrow[dl, dash] \arrow[dr, dash] & \\
\mathbb{C}(x) \arrow[dr, dash] & & \mathbb{C}(y) \arrow[dl, dash] \\
 & \mathbb{C} &
\end{tikzcd}$$

□

Remark. Fulton [Ful] proves that $[K(X):\mathbb{C}(x)] = \deg_0(x)$. Using Corollary 1.5, we know that $\deg_0(x) = \deg_\infty(x)$. The reason behind the use of $\deg_0\infty(x)$ is that, as said at the beginning of the section, our functions will have poles only at infinity. Thus, we will only need to pick two functions whose order at ∞ are coprime.

A nice feature of the polynomial F is that it can be taken to be monic in x and y.

Theorem 2.54 ([Yan2])**.** *Suppose that x and y are rational functions on a modular curve X with a unique pole at infinity of order n and m respectively. Suppose that $(n,m) = 1$ and that the leading coefficients of the Fourier expansion of x and y are both 1. Then the polynomial F is of the form*

$$x^m - y^n + \sum_{\substack{a,b\geq 0\\ an+bm<nm}} \alpha_{a,b}x^a y^b$$

Explicit models

Having introduced all the necessary theory, we can finally describe the explicit process which enables us to find models for $X_0(N)$. The first step consists in the choice of the parameters. A parameter is a function which generates the function field and we have already seen that this will be an η-quotient having a pole only at infinity.

An η-quotient for N is an expression of the form 2.2.4:

$$f(\tau) = \prod_{m|N} \eta_m(\tau)^{a_m} = \prod_{i=1}^{k} \eta_i(\tau)^{a_i}$$

In the discussion following Theorem 2.39, we have seen that the matrix

$$A(N) = \left(\frac{N}{c(N/c,c)} \frac{(c,m)^2}{m} \right)_{c,m} = \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(k)} \end{pmatrix}$$

encodes the order of f at the cusps. Here A^i is the i-th row carrying the information about the orders of all the $\eta_m(\tau)$ cusp at the i-th cusp.
Therefore, a parameter on $X_0(N)$ is obtained solving the following integer programming problem

$$A(N)\cdot X \begin{pmatrix} \geq 0 \\ \vdots \\ \geq 0 \\ = -24 * \delta \end{pmatrix} \quad \text{together with conditions } \textbf{(i)-(iv)} \text{ of Theorem 2.37}$$

where the solution X gives the exponents a_m's.

$$\begin{cases} A^{(i)}\cdot X \geq 0 & \text{for all } i < k \\ A^{(k)}\cdot X = -24 * \delta \\ \text{Conditions } \textbf{(i)-(iv)} \text{ of Theorem 2.37} \end{cases}$$

δ is simply the order of the pole at infinity.

Remark. To approach this system of equations we will make intensive use of the software `lp_solve` [Lps].

Once that the parameters have been chosen we will focus on the maps π_d (forgetful maps) and ω_ℓ (Atkin Lehner involutions). The idea is to describe them in terms of the parameters. For the forgetful maps, this is done by comparing the q-expansions of the parameters on the two curves. The Atkin-Lehner involutions, instead, are usually deduced from their compatibility relations (Lemma2.44). For curves of genus greater than or equal to 1, their equations come from the algebraic relations between their two parameters (again these relations will be found comparing q-expansions).

Remark. The q-expansion of an η-quotient is relatively easy to compute.

Remark. In order to find algebraic relations between different series we will make use of the `magma`-function `AlgebraicRelations` contained in the Echidna Package [Ech].

2.3 Squares of isogenies

In this section we will describe a commutative square of isogenies of the form

$$\begin{tikzcd}
E_0 \arrow[r, "\ell"] \arrow[d, "q"'] & E_1 \arrow[d, "q"] \\
F_0 \arrow[r, "\ell"'] & F_1
\end{tikzcd}$$

We will focus on the case of isogenies of prime degrees $\ell \neq q$. In particular, we are interested in the following problem: given the triple (E_0, E_1, F_0) together with the degrees ℓ and q, find the bottom right vertex of the square, namely F_1.

We will then try to generalize the process to the case

$$\begin{tikzcd}
E_0 \arrow[r, "\ell"] \arrow[d, "q"'] & E_1 \arrow[r, "\ell"] \arrow[d, dashed] & E_2 \arrow[r, "\ell"] \arrow[d, dashed] & \cdots \arrow[r, "\ell"] & E_n \arrow[d, "q"] \\
F_0 \arrow[r, "\ell"'] & F_1 \arrow[r, "\ell"'] & F_2 \arrow[r, "\ell"'] & \cdots \arrow[r, "\ell"'] & F_n
\end{tikzcd}$$

where we are given an ℓ-isogeny chain at the top ($E_0 \to \ldots \to E_n$) together with a q-isogeny $E_0 \to F_0$ and we want to find F_n (or the ℓ-isogeny chain at the bottom).

In this section we will take a modular approach based on the moduli interpretation of curves $X_0(N)$. In section 1.2.5, we observed that isomorphism classes of elliptic curves are parametrized by points on $Y(1)$. Further, the construction of modular polynomials, and the fact that they provide models for $X_0(N)$ (see Section 2.2.2), permits one to infer that points on $Y_0(N)$ represent elliptic curves together with an N-isogeny. We formalize the previous observations in the following Theorem.

Theorem 2.55. *The points of $Y_0(N)$ are in bijection with the set of isomorphism classes of couples (E, C_N) of an elliptic curve E together with a cyclic subgroup C_N of order N. Since subgroups of Elliptic curves are in bijection with isogenies, this is equivalent to saying that points $Y_0(N)$ represent isomorphism classes of couples (E, ϕ_N) of an elliptic curve E together with an isogeny of degree N.*

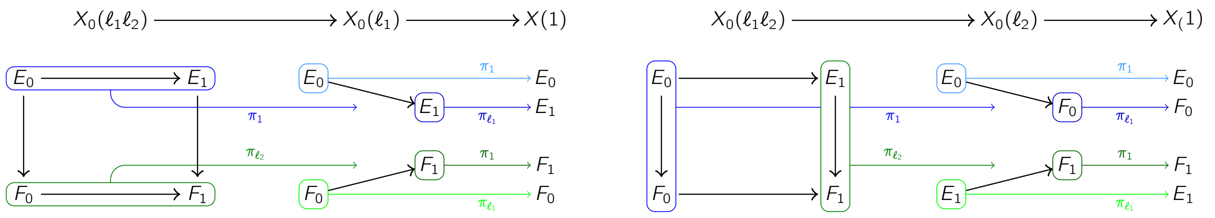
Proof. See [DS, Th. 1.5.1]. □

2.3.1 The modular curve $X_0(\ell_1\ell_2)$

Suppose we are given two primes $\ell_1 \neq \ell_2$. We observe that a square of isogenies is nothing but a data on the modular curve $X_0(\ell_1\ell_2)$. Therefore our object is a point on a modular curve and therefore it will be given by its coordinates (observe that by Lemma 2.52 there are only two cases for which $X_0(\ell_1\ell_2)$ has genus 0, i.e., $2 \cdot 3$ and $2 \cdot 5$). We obtain the following diagram (already described in Section 2.2.5)

$$\begin{tikzcd}
 & & X_0(\ell_1\ell_2) \arrow[dl, bend right, "\pi_1"'] \arrow[dl, "\pi_{\ell_2}"] \arrow[dr, bend left, "\pi_1"] \arrow[dr, "\pi_{\ell_1}"'] & & \\
\ell_2+1 & X_0(\ell_1) \arrow[dr, "\pi_{\ell_1}"] \arrow[dr, bend right, "\pi_1"'] & & X_0(\ell_2) \arrow[dl, "\pi_{\ell_2}"'] \arrow[dl, bend left, "\pi_1"] & \ell_1+1 \\
\ell_1+1 & & X(1) & & \ell_2+1
\end{tikzcd}$$

This has the following moduli interpretation



Remark. Note that once the maps $X_0(\ell_1) \to X(1)$ and $X_0(\ell_2) \to X(1)$ are chosen, the other maps $X_0(\ell_1\ell_2) \to X(\ell_1)$ and $X_0(\ell_1\ell_2) \to X(\ell_2)$ are uniquely determined; in fact, it is just a fiber product.

Remark. The Atkin Lehner involution acts on the square (E_0, E_1, F_0, F_1) by flipping the direction of the arrows

Example. Let us look at the case $X_0(14)$.

Both $X_0(2)$ and $X_0(7)$ have genus 0. In the Appendix A we see that $X_0(2)$ has parameter $t_2 = (\eta(\tau)/\eta_2(\tau))^{24}$ and the maps have description

$$\pi_1^*(j) = \frac{(t_2+256)^3}{t_2^2} \qquad \pi_2^*(j) = \frac{(t_2+16)^3}{t_2} \qquad \omega_2^*(t_2) = \frac{4096}{t_2}$$

Concerning $X_0(7)$ we find that the best choice for a parameter is $t_7 = (\eta(\tau)/\eta_7(\tau))^7$ and

$$\pi_1^*(j) = \frac{(t_7^2+245t_7+2401)^3(t_7^2+13t_7+49)}{t_7^7}$$

$$\pi_7^*(j) = \frac{(t_7^2+5t_7+1)^3(t_7^2+13t_7+49)}{t_7}$$

$$\omega_7^*(t_7) = \frac{49}{t_7}$$

Now, $X_0(14)$ has genus 1 and we find two η quotients

$$x_{14} = \frac{\eta_2(\tau)\eta_7(\tau)^7}{\eta_1(\tau)\eta_{14}(\tau)^7} \qquad y_{14} = \frac{\eta_2(\tau)^8\eta_7(\tau)^4}{\eta_1(\tau)^4\eta_{14}(\tau)^8}$$

giving the model for $X_0(14)$

$$y^2 - 5xy - 2y = x^3 - 3x^2 + 3x - 1$$

By looking at the algebraic relations between all these parameters, we find

$$\pi_1^*(t_2) = \frac{(y-7x+7)^2(x^2+40x+8-8y)}{y^3} \qquad \pi_7^*(t_2) = \frac{x^3(x^2-2x+1-y)}{y}$$

$$\pi_1^*(t_7) = \frac{y-7x+7}{x} \qquad \pi_2^*(t_7) = \frac{x^2-2x+1-y}{x}$$

Suppose we are given the following square of isogenies over the finite field $\mathbb{F}_{71}$.

$$\begin{tikzcd}
48 \arrow[r] \arrow[d] & 66 \arrow[d] \\
40 \arrow[r] & ??
\end{tikzcd}$$

There are several ways of completing it by finding the suitable bottom right corner:

(i) The easiest way consists in computing the gcd of the two modular polynomials $\Phi_2(40, x)$ and $\Phi_7(66, x)$ corresponding respectively to the bottom and the right isogenies. We find

$$\begin{cases} \Phi_2(40, x) = 0 \\ \Phi_7(66, x) = 0 \end{cases} \qquad \begin{cases} x^3 + 6x^2 + 35x = 0 \\ x^8 + 61x^7 + 45x^6 + 23x^5 + 11x^4 + 20x^3 + 43x^2 + 44x + 15 \end{cases}$$

whose solution is 17. We can easily check that this is the only possible j-invariant completing the square.

If we have to carry out this sort of computation only few times with similar parameters there is no doubt that this method is the most direct one. The drawback, as we have already mentioned at the end of section 2.2.2, is represented by the size of the modular polynomials. One way to get around this (as we will better describe in section 5.4.3) is to compute modular polynomials on modular curves other than $X(1)$ by comparing the parameter $t(q)$ and $t(q^p)$.

(ii) First of all we find that $t_2(q)$ and $t_2(q^7)$ satisfy a degree 8 symmetric polynomial $\Psi_7(x, y)$. This is still big in size but we already see some improvements: Φ_7 has 63 monomials while Ψ_7 has 51. The size of the coefficients also reduces by a factor of 3.

On the other hand, $t_7(q)$ and $t_7(q^2)$ satisfy the relation

$$\Psi_2(x, y) = x^3 - x^2y^2 - 8x^2y - 8xy^2 - 49xy + y^3$$

which is already much smaller than the one found in section 2.2.2. Note that Ψ_7 is a model for the image of $X_0(14)$ in the product $X_0(2) \times X_0(2)$ given by $(x, y) \mapsto (\pi_1(x, y), \pi_7(x, y))$. Besides, Ψ_2 represents the image of the map $X_0(14) \to X_0(7) \times X_0(7)$.

The upper side of the square is represented by the 2 isogeny corresponding to the point on $X_0(2)$ with coordinate t_2 obtained by

$$\begin{cases} \pi_1^*(j) = 48 \\ \pi_2^*(j) = 66 \end{cases} \qquad t_2 = 11$$

In the same way the left side of the isogeny square is a 7-isogeny corresponding to the point on $X_0(7)$ where the parameter t_7 has value

$$\begin{cases} \pi_1^*(j) = 41 \\ \pi_2^*(j) = 48 \end{cases}$$

since there are two 7-isogenies between 48 and 40 we get two possible values for t_7. At this stage we are not interested in distinguishing the two and we pick a random one $t_7 = 61 + 9\zeta_3$ where $\mathbb{F}_{71^2} = \mathbb{F}_{71}[\zeta_3]$.

We can now apply a 7-isogeny to t_2 and quotient out all the resulting isogenies not starting from 40:

$$\begin{cases} \Psi_7(t_2, x) = 0 \\ \pi_1^*(j) = 40 \end{cases} \qquad t_2' = 37$$

and the same for t_7 to which a 2-isogeny must coincide with the left side of the square

$$\begin{cases} \Psi_2(t_7, x) = 0 \\ \pi_1^*(j) = 66 \end{cases} \qquad t_7' = 58 + 55\zeta_3$$

We conclude by checking that both t_2' and t_7' represent isogenies to the same j-invariant by mapping them down to $X(1)$. We find $j = 17$ as before.

This strategy involves polynomials of smaller size but requires more computations. Nevertheless, we presented it because it will be of use in the next chapters where we will adapt it to modular curves with higher level. Although these two strategies work fine, they have a common disadvantage, namely the fact that all these computations have been done using specific numbers. We would like to construct precomputed rational functions that take as an input the two known isogenies and output the 4-th j-invariant.

(iii) As we said at the beginning of the section, the given square represents a point P on $X_0(14)$. The bottom 2-isogeny corresponds to $\pi_7(P)$ and the right 7-isogeny to $\pi_2(P)$. We can therefore look for relations between the two known $\pi_1(P) \in X_0(2), \pi_1(P) \in X_0(7)$ and one of the two $\pi_7(P) \in X_0(2), \pi_7(P) \in X_0(2)$. This amounts to studying the image of $X_0(14)$ in $X_0(2) \times X_0(2) \times X_0(7)$ or $X_0(2) \times X_0(7) \times X_0(7)$. In particular, we look for relations between $t_2(q), t_2(q^7)$ and $t_7(q)$ of degree 1 in $t_2(q^7)$. In our case we find

$$t_2' = \frac{39\left(t_2^3t_7^4 + 18t_2^3t_7^3 + 57t_2^3t_7^2 + 35t_2^3t_7 + t_2^3 + 39t_2^2t_7^5 + t_2^2t_7^4 + 27t_2^2t_7^3 + t_2^2t_7^2 + 42t_2^2t_7 + 28t_2^2 + 4t_2t_7^4 + 23t_2t_7^3 + 20t_2t_7^2 + 8t_2t_7 + 57t_2 + 10t_7^4\right)}{t_2^3t_7^3 + 49t_2^3t_7^2 + 66t_2^3t_7 + 52t_2^2t_7^4 + 21t_2^2t_7^3 + 28t_2^2t_7^2 + 39t_2^2t_7 + 45t_2^2 + 18t_2t_7^5 + 32t_2t_7^4 + 31t_2t_7^3 + 24t_2t_7^2 + 19t_2t_7 + 66t_2 + 55t_7^5 + 17t_7^4 + 11t_7^3}$$

and

$$t_7' = \frac{45t_2^2t_7^4 + 4t_2^2t_7^3 + 6t_2^2t_7^2 + 50t_2t_7^4 + 60t_2t_7^3 + 19t_2t_7^2 + 43t_2t_7 + 44t_2 + 22t_7^4}{t_2^2t_7^3 + 8t_2^2t_7^2 + 31t_2^2t_7 + 10t_2^2 + 70t_2t_7^4 + 57t_2t_7^3 + 47t_2t_7^2 + 36t_2t_7 + 6t_2 + 22t_7^3}$$

which once again gives $t_2' = 37$ and $t_7' = 58 + 55\zeta_3$. One could also compose these with π_2 and π_7 and get the j-invariant directly.

We conclude by noting that the above square corresponds to the point $(33+45\zeta_3, 57+41\zeta_3) \in X_0(14)$ and

$$\pi_7(x, y) = \pi_1(\omega_7(x, y)) = t_2'$$

2.3.2 The modular curve $X_0(\ell_1^n\ell_2)$

The same construction can be exploited for a rectangle of isogenies (a commutative diagram of isogenies of coprime degree). We picture below the modular tower. The two maps from $X_0(\ell_1^n\ell_2)$ to $X_0(\ell_1^n)$ send the rectangle to the two horizontal ℓ_1 isogeny chains at the top and at the bottom respectively.

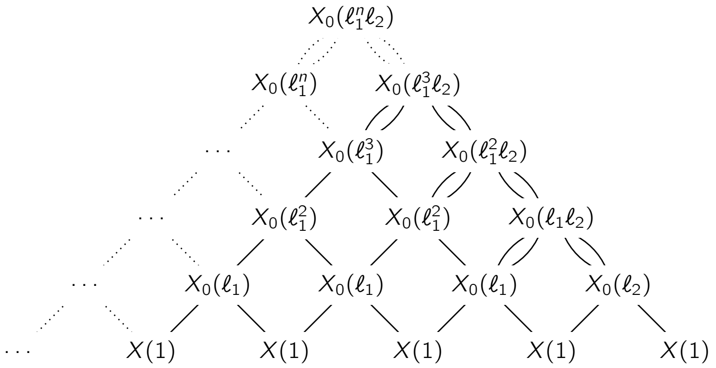


Figure 2.12 – Modular tower for $X_0(\ell_1^n\ell_2)$

The two maps $X_0(\ell_1^k\ell_2) \to X_0(\ell_1^{k-1}\ell_2)$ split the rectangle in the two unique sub-rectangles of length $k-1$.

$$\begin{tikzcd}
E_0 \arrow[r] \arrow[d] & E_1 \arrow[r] \arrow[d] & E_2 \arrow[r] \arrow[d] & E_3 \arrow[r] \arrow[d] & {} & {} \arrow[r] & E_{k-1} \arrow[r] \arrow[d] & E_k \arrow[d] \\
F_0 \arrow[r] & F_1 \arrow[r] & F_2 \arrow[r] & F_3 \arrow[r] & {} & {} \arrow[r] & F_{k-1} \arrow[r] & F_k
\end{tikzcd}$$

Example. Back to the previous example we suppose now to have the extended isogeny ladder (see section 5.3.2)

$$\begin{tikzcd}
48 \arrow[r] \arrow[d] & 66 \arrow[r] \arrow[d] & 41 \arrow[d] \\
40 \arrow[r] & 17 \arrow[r] & ??
\end{tikzcd}$$

We start by lifting the top 2-isogeny chain of length 2 to $X_0(4)$. Since there are two different two isogenies $66 \to 41$ we get two different points represented by the two parameters

$$t_4 = 4 + 68\zeta_3 \qquad s_4 = 7 + 3\zeta_3$$

The rectangle represents a point on $X_0(28)$, the two horizontal isogeny chains points on $X_0(4)$ and the vertical arrows points on $X_0(7)$.

The relation between $\pi_1(P) \in X_0(4)$, $\pi_7(P) \in X_0(4)$ and $\pi_1(P) \in X_0(7)$ gives the image of $X_0(28) \to X_0(4) \times X_0(4) \times X_0(7)$ and permits to recover a rational function which outputs the bottom 4-isogeny. We find $t_4' = 44$ and $s_4' = 64$ corresponding to the two chains $40 \to 17 \to 41$ and $40 \to 17 \to 24$ respectively.

Example. We can extend it even further

$$\begin{tikzcd}
48 \arrow[r] \arrow[d] & 66 \arrow[r] \arrow[d] & 41 \arrow[r] \arrow[d] & 17 \arrow[d] \\
40 \arrow[r] & 17 \arrow[r] & 24 \arrow[r] & ??
\end{tikzcd}$$

In order to lift the top isogeny chain, we can lift the 2 sub-chains of length 4 to $X_0(4)$. This way we ensure consistency with respect to the choice of the middle 2-isogeny. Then we lift the two to $X_0(8)$ finding

$$t_8 = 26 + 66\zeta_3$$

Note that, having fixed a choice for t_4 we only get one possibility for t_8. Once again, the image of $X_0(56)$ in $X_0(8) \times X_0(8) \times X_0(7)$ allows one to recover a rational function which outputs the bottom 8-isogeny.We find $t_8' = 44$ corresponding to the chain $40 \to 17 \to 24 \to 17$.

As showed in the example, one could construct rational functions completing the square by studying the image of $X_0(\ell_1\ell_2)$ in the product $X_0(\ell_1) \times X_0(\ell_1) \times X_0(\ell_2)$ given by the map

$$\begin{aligned} X_0(\ell_1\ell_2) &\longrightarrow X_0(\ell_1) \times X_0(\ell_1) \times X_0(\ell_2) \\ P &\longmapsto (\pi_1(P), \pi_{\ell_2}(P), \pi_1(P)) \end{aligned}$$

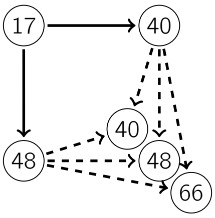
This image has a model coming from the algebraic relations between the parameter(s) on $X_0(\ell_1)$, their evaluation at q^{ℓ_2} and the parameter(s) on $X_0(\ell_2)$.

2.3.3 Obstructions to the completion of the square

Although the strategies employed in the completion of a rectangle of isogenies work perfectly in most of the cases, there are some subtleties we must pay attention to. For instance, if we look for the GCD between modular polynomials we might obtain a polynomial of degree greater than 1 resulting in roots with non-trivial multiplicities

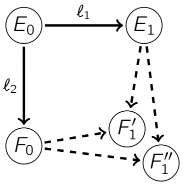
$$\begin{tikzcd}
17 \arrow[r] \arrow[d] & 40 \arrow[d, dashed, bend left] \arrow[d, dashed, bend right] \\
41 \arrow[r, dashed, bend left] \arrow[r, dashed, bend right] & 66
\end{tikzcd}$$

or multiple roots

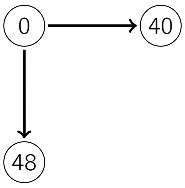


The former is due to the existence of multiple isogenies between two elliptic curves while the latter is more specific to the choice of parameters, namely the size of the prime. In fact, we have the following

Lemma 2.56. *The situation below, where there exist two or more different ways of completing a square, can only happen if there exists an endomorphism of E_1 (and F_0) of degree $\ell_1^2\ell_2^2$. Such endomorphisms exist if and only if there exist embeddings $\mathcal{O} \hookrightarrow \mathrm{End}(E_1)$ (and $\mathcal{O} \hookrightarrow \mathrm{End}(F_0)$) for a CM order of discriminant $|\Delta| \leq 4\ell_1^2\ell_2^2$.*



In any case, we are left with the problem of choosing among the possible solutions without being able to distinguish them. The same choice might be asked in the process of initializing the square, namely constructing the two known isogenies. In our example there exists multiple 2-isogenies between $j = 0$ and $j = 40$, between $j = 24$ and $j = 17$ or, once again, between $j = 40$ and $j = 66$ and almost all pairs of j-invariants have multiple 7-isogenies between them. Again, this can be overcome by increasing the size of the prime p, which reduces the probability of hitting problematic points, but we will always be left with some particular cases to deal with. In particular, the existence of extra automorphisms for some elliptic curves ($j = 0$ and 1728) , loops in the isogeny graphs and double edges require an ad-hoc study and prevent us from applying directly the methods above.
These pathological cases might also clash with the use of precomputed rational functions outputting the two missing sides of the square; if we try to complete the following square



we will note that both numerator and denominator of such a function are 0.

We will therefore add some extra piece of information to our construction so to avoid these situations. This extra data will come in the form of extra level-structure, namely we will work on modular curves covering $X_0(N)$. In doing so we will also gain another advantage: As it happens from $X(1)$ to $X_0(N)$, raising the level structure reduces the size of modular polynomials.

Chapter 3

Rigidification and higher level structures

In the previous chapters we have explored the arithmetic and geometric properties of the modular curves $X_0(N)$. These curves are associated with the congruence subgroups $\Gamma_0(N)$ and parametrize pairs (E, ϕ) of isomorphism classes of elliptic curves together with an isogeny of degree N. In Section 2.3 we have seen that, although this modular curve has some very nice description and seems to perfectly address the problem we were given to tackle, the existence of automorphisms of elliptic curves preserving the $\Gamma_0(N)$ structure poses a problem when trying to define a method to complete squares of isogenies. In particular, our choice to use "pathological" elliptic curves with j-invariant 0 or 1728, means that we have to deal with twists of these curves. This is a well known problem also referred to as "representability issue". It can be solved by adding some level structure and, by consequence, rigidifying the automorphisms of the elliptic curves.

In this chapter we will give a small introduction to the problem of representability. This will make us realize that the modular curves $X_1(N)$ and $X(N)$ are worth studying. We will therefore look at their signature by means of describing their cusps and elliptic points and finding some nice models to deal with them.

3.1 Modular curves with higher level structure

3.1.1 On Yoneda's lemma

Before attacking the representation problem we recall some facts form category theory. Let $\mathcal{C}$ be any category and let $X \in \mathcal{O}bj(\mathcal{C})$ be an object in it. We can construct the category $\mathbb{S}et^{\mathcal{C}}$ whose objects are functors $\mathcal{C} \to \mathbb{S}et$ and the morphisms are natural transformations. In the same exact way we can construct the category $\mathbb{S}et^{\mathcal{C}^{op}}$. We have two natural functors $Y : \mathcal{C} \to \mathbb{S}et^{\mathcal{C}}$ and $\mathcal{Y} : \mathcal{C} \to \mathbb{S}et^{\mathcal{C}^{op}}$ called Yoneda's covariant and contravariant functors

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "Y"] & \mathbb{S}et^{\mathcal{C}} \\
X \arrow[r] \arrow[d, "f"'] & h^X = \mathrm{Hom}_{\mathcal{C}}(X, *) \arrow[d, "h^f"] \\
Y \arrow[r] & h^Y = \mathrm{Hom}_{\mathcal{C}}(Y, *)
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathcal{C} \arrow[r, "\mathcal{Y}"] & \mathbb{S}et^{\mathcal{C}^{op}} \\
X \arrow[r] \arrow[d, "f"'] & h_X = \mathrm{Hom}_{\mathcal{C}}(*, X) \arrow[d, "h_f"] \\
Y \arrow[r] & h_Y = \mathrm{Hom}_{\mathcal{C}}(*, Y)
\end{tikzcd}$$

We will focus on the second one (the contravariant version). Observe that the morphism $h_f : h_X \to h_Y$ is defined by

$$\begin{aligned} h_f(W) : h_X(W) &\longrightarrow h_Y(W) \\ (\phi : W \to X) &\longrightarrow (f \circ \phi : W \to X \to Y) \end{aligned}$$

and it is a natural transformation since the following diagram commutes

$$\begin{tikzcd}
h_X(W) \arrow[d, "h_f(W)"'] & h_X(Z) \arrow[l, "h_X(\phi)"'] \arrow[d, "h_f(Z)"] \\
h_Y(W) & h_Y(Z) \arrow[l, "h_Y(\phi)"]
\end{tikzcd}$$

Definition. Let $h : \mathcal{C}^{op} \to \mathbb{S}et$ be a functor, we say that F is representable if there exists an object $X \in \mathcal{O}bj(\mathcal{C})$ such that $h \simeq h_X$.

There is an equivalent definition of representability which uses the notion of universal object. If $h : \mathcal{C}^{op} \to \mathbb{S}et$ is a functor, a universal object for h is a pair (X, ξ) where $X \in \mathcal{O}bj(\mathcal{C})$ and $\xi \in h(X)$ verifying the following universal property: for each other pair $(Y, \eta) \in \mathcal{O}bj(\mathcal{C}) \times F(Y)$, there is a unique morphism $f : Y \to X$ such that $F(f)(\xi) = \eta$.

Definition. Let $h : \mathcal{C}^{op} \to \mathbb{S}et$ be a functor, we say that F is representable if it admits a universal object.

Let X be an object of $\mathcal{C}$, $h : \mathcal{C}^{op} \to \mathbb{S}et$ be a functor and $f : Y \to X$ be a morphism. We can construct maps

$$\begin{array}{ccc} \mathrm{Hom}(h_X, h) & \xrightarrow{\alpha} & h(X) \\ (\tau : h_X \to F) & \longrightarrow & \tau(X)(\mathrm{Id}_X) \end{array} \qquad \begin{array}{ccc} h(X) & \xrightarrow{\beta} & \mathrm{Hom}(h_X, h) \\ \xi & \longrightarrow & \tau_\xi : h_X \to F \ \text{ s.t. } \tau_\xi(Y)(f) = h(f)(\xi) \end{array}$$

Lemma 3.1 (Yoneda's Lemma)**.** *α and β are bijections of sets and they are one the inverse of the other.*

Lemma 3.2 (Yoneda's Embedding)**.** *Let X and Y be two objects of $\mathcal{C}$. We have a bijection*

$$\mathrm{Hom}_{\mathcal{C}}(X, Y) \longrightarrow \mathrm{Hom}(h_X, h_Y)$$

which takes $f : X \to Y$ to $h_f : h_x \to h_y$ where, for W in $\mathcal{C}$, $h_f(W) : h_x(W) \to h_Y(W)$ is defined by $(g : W \to X) \longrightarrow (f \circ g : W \to X \to Y)$.

Corollary 3.3. *$X \simeq Y$ if and only if $h_x \simeq h_Y$.*

Example. We consider the functor

$$\begin{aligned} \Gamma : \mathrm{Sch} &\longrightarrow \mathbb{S}et \\ X &\longrightarrow \Gamma(X, \mathcal{O}_X) = \{\text{Global sections of } X\} \end{aligned}$$

Γ is representable.

Remark. For every ring R there is a unique ring homomorphism $\mathbb{Z} \to R$. Therefore, for every scheme X there exists a unique morphism of schemes $X \to \mathrm{Spec}(\mathbb{Z})$ ($\mathrm{Spec}(\mathbb{Z})$ is a final object in the category of schemes).

By [GW, Proposition 3.4] we have the following isomorphism

$$\mathrm{Hom}\left(X, \mathrm{Spec}(\mathbb{Z}[t])\right) \simeq \mathrm{Hom}\left(\mathbb{Z}[t], \Gamma(X, \mathcal{O}_X)\right)$$

which comes from the fact that every morphism of schemes $X \to \mathrm{Spec}(\mathbb{Z}[t])$ is determined by the induced map on the ring of global sections. Since $\mathrm{Hom}(\mathbb{Z}[t], A) \simeq A$ for any ring, we get

$$\mathrm{Hom}\left(X, \mathrm{Spec}(\mathbb{Z}[t])\right) \simeq \mathrm{Hom}\left(\mathbb{Z}[t], \Gamma(X, \mathcal{O}_X)\right) \simeq \Gamma(X, \mathcal{O}_X) = \mathcal{O}_X(X)$$

which implies $\mathrm{Hom}\left(X, \mathrm{Spec}(\mathbb{Z}[t])\right) \simeq \Gamma(X, \mathcal{O}_X)$ and, therefore, the representability of Γ.

3.1.2 The main issue with representability of $Y(1)$

We discuss now the moduli problem for $Y(1)$. For this section we will mainly follow the lecture by Snowden online at [Sno]. For a more profound approach one could refer to the classic book by Katz and Mazur [KM].

We have already defined the open modular curve $Y(1)$ as a Riemann surface and we have noted that its points are in bijection with the set of elliptic curves over $\mathbb{C}$.

$$\{\text{Elliptic Curves }/\mathbb{C}\}/\simeq \longleftrightarrow \mathbb{H}/\Gamma(1) \simeq \mathbb{C}$$

where the first map is a bijection and the isomorphism on the right is an isomorphism of Riemann surfaces. We would like to do this from an algebraic point of view; the algebraic moduli problem consists in describing the functor of points of $Y(1)$. If S is some scheme, then a map $S \to Y(1)$ should correspond to a family of elliptic curves over S.

Definition. An elliptic curve over S is a proper and smooth map $E \to S$ equipped with a section $0 \in E(S)$ such that each generic fiber is a geometrically connected genus 1 curve

Therefore, we define the functor

$$\begin{aligned} \Gamma(1) : \mathrm{Sch} &\longrightarrow \mathbb{S}et \\ S &\longrightarrow \left\{ \begin{matrix} \text{Isomorphism classes of} \\ \text{Elliptic Curves over } S \end{matrix} \right\} \end{aligned}$$

Definition. A fine moduli space for a functor $h : \mathcal{C}^{op} \to \mathbb{S}et$ is a universal object X representing h.

The idea is that, if we could prove that $\Gamma(1)$ is representable by an object, we would define $Y(1)$ to be this scheme. Unfortunately, it turns out that $\Gamma(1)$ is not representable.
The issue with this moduli problem is that elliptic curves have automorphisms and, therefore, the j-invariant, which provides the isomorphism between $Y(1)$ and $\mathbb{C}$, is not anymore a good detector of isomorphism classes when we work over other fields than $\mathbb{C}$ (in particular, it does not work over fields that are not algebraically closed). This causes $\Gamma(1)$ not to be a sheaf and therefore it cannot be representable.
There are quite few ways of overcoming this issue. One of this is to reduce our expectation and, instead of looking for a complex manifold (or a scheme) such as $Y(1)$, try to look for an orbifold (or its algebraic geometric counterpart, a Deligne Mumford stack [DM]) which enlarge the category so that the functor becomes representable. Another possibility is to study *coarse* moduli spaces instead of *fine* ones; this is a scheme which best approximates the representability of our functor but, instead of giving an isomorphism of functors between $\Gamma(1)$ and $\mathrm{Hom}(*, X)$ for some scheme X, only realizes a *universal* natural transformation between them.

The study of these two subjects goes beyond the goals of this thesis and has some drawbacks for what concerns our aim; thus, we will follow a different approach, namely rigidifying the isomorphisms classes of elliptic curves, by means of equipping them with some extra structure, in order to eliminate the automorphisms.

Definition. Let $N \geq 2$ be an integer and $E \to S$ an elliptic curve, with N invertible on S. A $\Gamma(N)$-structure (" full level N structure") on E is a pair of sections $(P, Q) \in E(S)[N]$ such that the map $(P, Q) : (\mathbb{Z}/N\mathbb{Z})^2_S \to E[N]$ is an isomorphism of group schemes over S. This is the same as asking P and Q to form a basis for the N torsion of E.

Proposition 3.4. *If $N \geq 3$ there is no non-trivial automorphism of an elliptic curve preserving a level N structure, i.e., for a pair $(E, \mathcal{B})$ with $\mathcal{B}$ a basis of the N-torsion of E there is not an automorphism of E (besides the identity) taking $\mathcal{B}$ to itself. Note that any automorphism of E will send $\mathcal{B}$ to another basis of $E[N]$ but only the identity will fix it.*

In the language of Katz and Mazur this proposition is equivalent to saying that the moduli problem associated to the functor $\Gamma(N)$ is *rigid*.

$$\begin{aligned} \Gamma(N) : \mathrm{Sch} &\longrightarrow \mathbb{S}et \\ S &\longrightarrow \left\{ \begin{matrix} \text{Isomorphism classes of pairs } (E, \mathcal{B}) \\ \text{with } E \text{ an elliptic curve over } S \text{ and} \\ \mathcal{B} \text{ a level } N \text{ structure} \end{matrix} \right\} \end{aligned}$$

Proof. We refer to [KM, Corollary 2.7.2]. □

This proposition gives us some chances of getting a representable functor. An immediate consequence is that the lack of automorphisms turns $\Gamma(N)$ into a sheaf for $N \geq 3$.

Theorem 3.5 ([KM, Corollary 4.7.2])**.** *Suppose $N \geq 3$. Then the fine moduli problem $\Gamma(N)i$ s representable by a smooth affine scheme $Y(N)$ over $\mathbb{Z}[1/N]$.*

The whole story could be adapted to the $\Gamma_1(N)$ case:

$$\Gamma_1(N) : \mathrm{Sch} \longrightarrow \mathbb{S}et$$
$$S \longrightarrow \left\{ \begin{array}{c} \text{Isomorphism classes of pairs } (E, \mathcal{B}) \\ \text{with } E \text{ an elliptic curve over } S \text{ and} \\ \mathcal{B} \text{ a } \Gamma_1(N) \text{ structure} \end{array} \right\}$$

where a $\Gamma_1(N)$-structure on E is a section $P \in E[N](S)$ of order N, i.e., a point of exact order N in $E(S)$, or a homomorphism $\phi : \mathbb{Z}/N\mathbb{Z} \to E[N](S)$.

Proposition 3.6 ([KM, Corollary 2.7.3])**.** *If $N \geq 4$ there is no non-trivial automorphism of an elliptic curve preserving a $\Gamma_1(N)$ structure.*

Theorem 3.7. *For $N \geq 4$, the functor $\Gamma_1(N)$ is representable by a smooth affine scheme $\mathcal{Y}_1(N)$ over $\mathbb{Z}[1/N]$.*

Unfortunately, $Y_0(N)$ with the usual interpretation is never representable since multiplication by -1 preserves the $\Gamma_0(N)$ structure.

3.1.3 Signature of $X(N)$

In this section we will describe the modular curve $X(N)$ by means of studying the degree of its covering of $X(1)$, the number of its elliptic points of period 2 and 3 and its set of cusps. We will mainly follow [CS2, Ch. 6], [New1], [Shi, Ch. 1] and [DS, Ch. 3]. We recall that

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N \right\}$$

Degree

Lemma 3.8. *Let $N \geq 3$. The degree of the map $X(N) \to X(1)$ is given by*

$$d = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)]/2 = \frac{N^3}{2} \prod_{p|N} \left(1 - \frac{1}{p^2}\right)$$

and $d = 6$ for $N = 2$.

Note that the division by 2 comes from the fact that for $N > 2$, $-I \notin \Gamma(N)$ (cf. beginning of §1.3.5).

Proof. The result follows from the short exact sequence [CS2, Prop. 6.2.4]

$$1 \longrightarrow \Gamma(N) \longrightarrow \mathrm{SL}_2(\mathbb{Z}) \longrightarrow \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) \longrightarrow 1$$

which comes from the study of the reduction map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$.
We can compute the cardinality of $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$ using the short exact sequence

$$1 \longrightarrow \mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}/p^n\mathbb{Z}) \longrightarrow (\mathbb{Z}/p^n\mathbb{Z})^\times \longrightarrow 1$$

knowing that $\#\mathrm{GL}_2(\mathbb{Z}/p^n\mathbb{Z}) = p^{4(n-1)}(p^2-1)(p^2-p)$. The Chinese remainder theorem permits one to conclude the proof. □

Cusps

Proposition 3.9 ([DS, Prop. 3.8.3])**.** *Let $s = (a : c)$ and $s' = (b : d)$ be elements of $\mathbb{P}^1(\mathbb{Q})$ with* $\gcd(a, c) = \gcd(b, d) = 1$. *They define the same cusp for $\Gamma(N)$, namely $\Gamma(N)s = \Gamma(N)s'$, if and only if $(a : c) \equiv \pm(b : d) \mod N$.*

Now we define a bijection

$$\varphi : (\mathbb{Z}^2)^\times / \sim_{\Gamma(N)} \longleftrightarrow \Gamma(N)/\mathbb{Q}$$

and describe an algorithm to enumerate all the cusps of $\Gamma(N)$ [New1, §2.3.1]. For $\overline{a}, \overline{c} \in \mathbb{Z}/N\mathbb{Z}$ we define $\gcd(\overline{a}, \overline{c})$ to be $\gcd(x, y)$ for (x, y) the smallest positive representatives of $\overline{a}$ and $\overline{c}$ respectively. We set

$$\left((\mathbb{Z}/N\mathbb{Z})^2\right)^\times = \left\{ \begin{bmatrix} \overline{a} \\ \overline{c} \end{bmatrix} \;\middle|\; \overline{a}, \overline{c} \in \mathbb{Z}/N\mathbb{Z},\ \gcd(\overline{a}, \overline{c}) = 1 \right\}$$

We get a map

$$\begin{aligned} \psi : \left((\mathbb{Z}/N\mathbb{Z})^2\right)^\times &\longrightarrow (\mathbb{Z}^2)^\times / \sim_{\Gamma(N)} \\ \begin{bmatrix} \overline{a} \\ \overline{c} \end{bmatrix} &\longmapsto \begin{bmatrix} a \\ c \end{bmatrix} \quad \text{where } \begin{bmatrix} a \\ c \end{bmatrix} \text{ is a lift of } \begin{bmatrix} \overline{a} \\ \overline{c} \end{bmatrix} \text{ to } (\mathbb{Z}^2)^\times \end{aligned}$$

Theorem 3.10 ([New1, Th. 2.10])**.** *The map ψ is independent of the choice of the lift and surjective.*

Lemma 3.11. $\begin{bmatrix} \overline{a} \\ \overline{c} \end{bmatrix} \neq \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix}$ *in $\left((\mathbb{Z}/N\mathbb{Z})^2\right)^\times$ have the same image via ψ if and only if* $\begin{bmatrix} \overline{a} \\ \overline{c} \end{bmatrix} = -\begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix}$.

Algorithm 3. Cusps of $\Gamma(N)$

Input: An integer N.
Output: A complete list of representatives for the cusps of $X(N)$

1. For integers $a, c \in \{0, \dots, N-1\}$ list all elements $\begin{bmatrix} a \\ c \end{bmatrix}$ such that $\gcd(a, c, N) = 1$.

2. Compute pairs $\left\{ \begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right\}$ such that $\begin{bmatrix} a \\ c \end{bmatrix} \equiv -\begin{bmatrix} x \\ y \end{bmatrix}$ mod N from the set of elements found in step **1**.

3. Choose one representative from each set computed in step **2**.

Example. We compute the cusps of $X(4)$.

1. $\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

2. $\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$

3. $\operatorname{Cusps}(X(4)) = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$

Theorem 3.12 ([CS2, Cor. 6.3.15])**.** *A system of representatives of the cusps for $\Gamma(N)$ is given by the set of $(a : b) \in \mathbb{P}^1(\mathbb{Q})$ constructed as follows: for each b such that $1 \leq b \leq N/2$ or $b = N$ and for each a_0 such that $0 \leq a_0 < N$ (or $0 \leq a_0 < N/2$ if $b = N/2$ or $b = N$) and $\gcd(a_0, b, N) = 1$, we choose an $a \equiv a_0$ (mod N) such that $\gcd(a, b) = 1$.*

Example. We compute once again the cusps of $X(4)$.

	$a_0 = 0$	$a_0 = 1$	$a_0 = 2$	$a_0 = 3$
$b = 1$	0	1	2	3
$b = 2$	✖	1		▨
$b = 4$	✖	1	▨	▨

$\operatorname{Cusps}(X(4)) = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$. Note that $\begin{bmatrix} 3 \\ 1 \end{bmatrix} \sim \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

We will now present two different approaches to count the number of cusps.

Lemma 3.13. *The number of cusps of* $\Gamma(N)$ *is*

$$\epsilon_\infty = \begin{cases} \dfrac{N^2}{2}\displaystyle\prod_{p|N}\left(1-\frac{1}{p^2}\right) & \text{If } N \geq 2 \\ 1 & \text{If } N = 1 \\ 3 & \text{If } N = 2 \end{cases}$$

Proof. Since $\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$, then all the cusps have same width. Then we can compute the width (the index of the stabilizer) of ∞

$$\Gamma(N) \cap \Gamma_\infty = \left\{ \begin{pmatrix} 1 & Nn \\ 0 & 1 \end{pmatrix} \,\middle|\, n \in \mathbb{Z} \right\}$$

Thus, $h_{\mathfrak{c}} = h_\infty = N$ for any cusp $\mathfrak{c} \in \mathrm{Cusps}(\Gamma(N))$. By Lemma 2.20 we get

$$\epsilon_\infty \cdot N = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] \implies \epsilon_\infty = \frac{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)]}{N}$$

which proves the claim. □

We now use a direct counting strategy to find a different (but equivalent) expression for ϵ_∞.

Lemma 3.14. *Let* N *be a positive integer and* d *a divisor of* N. *Let* $c \in \{1, \ldots, N-1\}$. *There are* $\phi(N/d)$ *values of* c *such that* $\gcd(N, c) = d$.

Proof. We set

$$P(M) = \{x \,|\, 1 \leq x \leq M-1, \ \gcd(x, M) = 1\} \simeq (\mathbb{Z}/M\mathbb{Z})^\times$$

$P(M)$ has cardinality $\phi(M)$. We construct a map

$$\begin{aligned} P(N/d) &\longrightarrow \{c \,|\, 1 \leq c \leq N-1, \ \gcd(c, N) = d\} \\ x &\longrightarrow xd \end{aligned}$$

If $\gcd(x, N/d) = 1$, then $\gcd(xd, N) = d$ which implies that the map is well defined. This map has a natural inverse

$$\begin{aligned} \{c \,|\, 1 \leq c \leq N-1, \ \gcd(c, N) = d\} &\longrightarrow P(N/d) \\ y &\longrightarrow y/d \end{aligned}$$

which again is well defined. This implies that

$$\#\{c \,|\, 1 \leq c \leq N-1, \ \gcd(c, N) = d\} = \#P(N/d) = \phi(N/d)$$

□

Theorem 3.15. $X(N)$ *has* 3 *cusps for* $N = 2$. *If* $N \geq 3$, *then*

$$\epsilon_\infty = \frac{1}{2}\sum_{d|N}\frac{N}{d}\phi(d)\phi(N/d)$$

Proof. We refer to Algorithm 3. In the first step we fix $a \in \{0, \ldots, N-1\}$ and count the number of c's in $\{0, \ldots, N-1\}$ such that $\gcd(a, c, N) = 1$. Suppose $\gcd(a, N) = d$, we need the c's such that $\gcd(c, d) = 1$ and they are $(N/d)\phi(d)$. The choices for a are $\phi(N/d)$ by the lemma before and, therefore, we get

$$\sum_{d|N}\frac{N}{d}\phi(d)\phi(N/d)$$

candidates. The grouping and successive choice in steps **2** and **3** yields the final formula

$$\epsilon_\infty = \frac{1}{2}\sum_{d|N} \frac{N}{d}\phi(d)\phi(N/d)$$

Indeed, the only elements such that $\begin{bmatrix} a \\ c \end{bmatrix} \equiv -\begin{bmatrix} a \\ c \end{bmatrix}$ (mod N) are $\begin{bmatrix} N/2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ N/2 \end{bmatrix}$ and $\begin{bmatrix} N/2 \\ N/2 \end{bmatrix}$. Clearly, if N is odd there are no such elements. If $N > 2$ is even, then none of these satisfy the gcd condition and, therefore they do not affect the computation. Finally, if $N = 2$ they represent all the cusps. $\epsilon_\infty(2) = 3$. □

Elliptic points

We recall that an elliptic point for Γ is a point with non-trivial stabilizer.

Theorem 3.16 ([Shi, Prop 1.39]). *For $N > 1$, the modular curve $X(N)$ has no elliptic point.*

Proof. We have seen (Prop 1.34) that every elliptic element of $\Gamma(1)$ is conjugate to one of the following

$$S^{\pm 1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{\pm 1} \qquad R^{\pm 1} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}^{\pm 1} \qquad W^{\pm 1} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}^{\pm 1}$$

and none of these is conjugate to $\pm I$ mod N for $N > 1$. Since $\Gamma(N)$ is normal in $\mathrm{SL}_2(\mathbb{Z})$, this concludes the proof. Indeed, suppose $\xi \in \epsilon_2$ or ϵ_3 and $A \in \mathrm{Stab}_\Gamma(\xi)$. If A is conjugate to $\Lambda \in \{S^{\pm 1}, R^{\pm 1}, W^{\pm 1}\}$, this would mean that $B^{-1}AB = \Lambda$ some $B \in \mathrm{SL}_2(\mathbb{Z})$ but the normality of $\Gamma(N)$ yields $\Lambda \in \Gamma(N)$ which is not true. □

Genus

It only remains to assemble all the information together

$$g(X(N)) = 1 + \frac{N^3}{24}\prod_{p|N}\left(1 - \frac{1}{p^2}\right) - \frac{N^2}{4}\prod_{p|N}\left(1 - \frac{1}{p^2}\right) = 1 + \frac{N^2}{4}\cdot\frac{N-6}{6}\cdot\prod_{p|N}\left(1 - \frac{1}{p^2}\right)$$

and $g(X(2)) = 0$.

3.1.4 Signature of $X_1(N)$

Before starting we remind that

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \mod N \right\}$$

is a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

Degree

Lemma 3.17. *The degree of the reduction map $X_1(N) \to X(1)$ is given by*

$$d = [SL_2(\mathbb{Z}) : \Gamma_1(N)]/2 = \frac{N^2}{2}\prod_{p|N}\left(1 - \frac{1}{p^2}\right)$$

for $N \geq 3$. Further $d(X_1(2) \to X(1)) = [SL_2(\mathbb{Z}) : \Gamma_1(2)] = 3$.

Proof. Again we will use a short exact sequence

$$1 \longrightarrow \Gamma(N) \longrightarrow \Gamma_1(N) \longrightarrow \mathbb{Z}/N\mathbb{Z} \longrightarrow 1$$
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto b \mod N$$

Thus

$$[SL_2(\mathbb{Z}):\Gamma_1(N)] = \frac{[SL_2(\mathbb{Z}):\Gamma(N)]}{[\Gamma(N):\Gamma_1(N)]} = \frac{N^3\prod_{p|N}\left(1-\frac{1}{p^2}\right)}{\#(\mathbb{Z}/N\mathbb{Z})} = \frac{N^3\prod_{p|N}\left(1-\frac{1}{p^2}\right)}{N} = N^2\prod_{p|N}\left(1-\frac{1}{p^2}\right)$$

In case $N \geq 3$ we divide by 2 since $-I \notin \Gamma_1(N)$. □

Cusps

As before, we will try to describe an algorithm enumerating the cusps of $\Gamma_1(N)$ following [New1]. We define an equivalence relation

$$\begin{bmatrix} a \\ c \end{bmatrix} \sim \begin{bmatrix} a' \\ c' \end{bmatrix} \iff \begin{bmatrix} a \\ c \end{bmatrix} \equiv \pm \begin{bmatrix} a'+jc \\ c' \end{bmatrix} \mod N \quad \text{for some } j \in \mathbb{Z}$$

By Proposition 3.8.3 in [DS] we know that this is equivalent to

$$\Gamma_1(N)\begin{bmatrix} a \\ c \end{bmatrix} = \Gamma_1(N)\begin{bmatrix} a' \\ c' \end{bmatrix}$$

which means that we can construct a bijection

$$\varphi : (\mathbb{Z}^2)^\times/\sim \longleftrightarrow \Gamma_1(N)\ \mathbb{Q}^\times$$

As before, we define a map

$$\begin{array}{rcl} \psi : \left((\mathbb{Z}/N\mathbb{Z})^2\right)^\times & \longrightarrow & (\mathbb{Z}^2)^\times/\sim \\ \begin{bmatrix} \overline{a} \\ \overline{c} \end{bmatrix} & \longmapsto & \begin{bmatrix} a' \\ c' \end{bmatrix} \text{ some lift} \end{array}$$

Theorem 3.18 ([New1, The. 2.17])**.** *ψ is independent of the choice of the lift, it is well defined and surjective.*

Lemma 3.19. *Fix $\overline{c} \in (\mathbb{Z}/N\mathbb{Z})$. Let $\begin{bmatrix} \overline{a} \\ \overline{c} \end{bmatrix}, \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix} \in \left((\mathbb{Z}/N\mathbb{Z})^2\right)^\times$ with lifts $\begin{bmatrix} a' \\ c' \end{bmatrix}$ and $\begin{bmatrix} x' \\ y' \end{bmatrix}$ in $(\mathbb{Z}^2)^\times$. If $d = \gcd(c;N)$, then $a \equiv x \mod d$ if and only if $\begin{bmatrix} a' \\ c' \end{bmatrix} \sim \begin{bmatrix} x' \\ y' \end{bmatrix}$*

From this discussion we deduce an algorithm:

Algorithm 4. Cusps of $\Gamma_1(N)$

Input: An integer N.
Output: A complete list of representatives for the cusps of $X_1(N)$

1. For all integers $c \in \{0,\dots,N-1\}$ define $d = \gcd(c;N)$. For all $a \in P(d)$ add $\begin{bmatrix} a \\ c \end{bmatrix}$ to the list.

2. Choose $\begin{bmatrix} a \\ c \end{bmatrix}$ from the list computed in step **1**. Let $d = \gcd(c,N)$; if possible, find a different element $\begin{bmatrix} x \\ y \end{bmatrix}$ from the list such that $y \equiv -c \mod N$ and $x \equiv -a \mod d$. Create the set $\left\{\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\right\}$ if such an element exists, otherwise create the set $\left\{\begin{bmatrix} a \\ c \end{bmatrix}\right\}$.

3. Choose representative from each set computed in step **2**.

Example. We list all the cusps of $X_1(10)$.

	$d = \gcd(c, N)$	$P(d)$	Candidates
$c = 0$	10	$\{1, 3, 7, 9\}$	$\begin{bmatrix}1\\0\end{bmatrix}^{\clubsuit}, \begin{bmatrix}3\\0\end{bmatrix}^{\spadesuit}, \begin{bmatrix}7\\0\end{bmatrix}^{\spadesuit}, \begin{bmatrix}9\\0\end{bmatrix}^{\clubsuit}$
$c = 1$	1	$\{0\}$	$\begin{bmatrix}0\\1\end{bmatrix}^{\blacklozenge}$
$c = 2$	2	$\{1\}$	$\begin{bmatrix}1\\2\end{bmatrix}^{\blacksquare}$
$c = 3$	1	$\{0\}$	$\begin{bmatrix}0\\3\end{bmatrix}^{\blacktriangle}$
$c = 4$	2	$\{1\}$	$\begin{bmatrix}1\\4\end{bmatrix}^{\bigstar}$
$c = 5$	5	$\{1, 2, 3, 4\}$	$\begin{bmatrix}1\\5\end{bmatrix}^{\bullet}, \begin{bmatrix}2\\5\end{bmatrix}^{\blacktriangledown}, \begin{bmatrix}3\\5\end{bmatrix}^{\blacktriangledown}, \begin{bmatrix}4\\5\end{bmatrix}^{\bullet}$
$c = 6$	2	$\{1\}$	$\begin{bmatrix}1\\6\end{bmatrix}^{\bigstar}$
$c = 7$	1	$\{0\}$	$\begin{bmatrix}0\\7\end{bmatrix}^{\blacktriangle}$
$c = 8$	2	$\{1\}$	$\begin{bmatrix}1\\8\end{bmatrix}^{\blacksquare}$
$c = 9$	1	$\{0\}$	$\begin{bmatrix}0\\9\end{bmatrix}^{\blacklozenge}$

where the symbols denote the equivalences found in step **2**. We conclude that

$$\mathrm{Cusps}(X_1(10)) = \left\{ \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}0\\3\end{bmatrix}, \begin{bmatrix}1\\4\end{bmatrix}, \begin{bmatrix}1\\5\end{bmatrix}, \begin{bmatrix}2\\5\end{bmatrix} \right\}$$

An equivalent description of the cusps of $\Gamma_1(N)$ is given by the following theorem

Theorem 3.20 ([CS2, Cor. 6.3.19])**.** *A system of representatives of the cusps for $X_1(N)$ is given by the set of $(a : b) \in \mathbb{P}^1(\mathbb{Q})$ constructed as follows: for each b such that $1 \leq b \leq N/2$ or $b = N$ and for each a_0 such that $0 \leq a_0 < \gcd(b, N)$ (or $0 \leq a_0 < \gcd(b, N)/2$ if $b = N/2$ or $b = N$) and $\gcd(a_0, b, N) = 1$, we choose an $a \equiv a_0 \pmod N$ such that $\gcd(a, b) = 1$.*

Example. We study what the representatives of the cusps of $X_1(10)$ look like using the theorem above.

b \ a_0	0	1	2	3	4	5
1	0	▨	▨	▨	▨	▨
2	✖	1	▨	▨	▨	▨
3	0	▨	▨	▨	▨	▨
4	✖	1	▨	▨	▨	▨
5	✖	1	2	▨	▨	▨
10	✖	1	✖	3	✖	▨

Once again we obtain the set

$$\mathrm{Cusps}(X_1(10)) = \left\{ \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}0\\3\end{bmatrix}, \begin{bmatrix}1\\4\end{bmatrix}, \begin{bmatrix}1\\5\end{bmatrix}, \begin{bmatrix}2\\5\end{bmatrix} \right\}$$

As we did for the modular curve $X(N)$ we would like to count the number of cusps.

Corollary 3.21. *The number of cusps of $\Gamma_1(N)$ is given by*

$$\#\mathrm{Cusps}(\Gamma_1(N)) = \begin{cases} \frac{1}{2}\sum_{d|N} \phi(d)\phi(N/d) & \text{If } N = 3 \text{or } N \geq 5 \\ n+1 & \text{If } N = 2^n,\ n \in \{0,1,2\} \end{cases}$$

Proposition 3.22. *Let $(a : c) \in \mathbb{P}^1(\mathbb{Q})$ with $\gcd(a,c) = 1$, then the width of a/b for $\Gamma_1(N)$ is equal to $N/\gcd(c,N)$ with the unique exception of $1/2$ of $\Gamma_1(4)$ which has width 1.*

Proof. This is proposition 6.3.20 in [CS2]. □

Elliptic points

Theorem 3.23. *There are no elliptic points on $X_1(N)$ for $N > 3$.*

Proof. The non-trivial elements of the stabilizers of an elliptic point of order 2 have trace 0 while those for elliptic points of order 3 have trace ± 1. Hence, elliptic elements for $\Gamma_1(N)$ can have trace in $\{-1,0,1\}$. This implies that, for $N > 3$, the curve $\Gamma_1(N)$ cannot have elliptic points, since its elements have trace $1 + Nk + 1 + Nk' = 2 + N(k+k')$. □

Remark. The modular curve $X_1(2)$ has a unique elliiptic point of order 2. Indeed, we know that $X_0(2) \to X(2)$ and $X_1(2) \to X(1)$ have the same degree and therefore $X_0(2) = X_1(2)$. This, together with the fact that the number of elliptic points of order 2 (respectively 3) for X_Γ is related with the number of orbits of S (respectively ST) on $\mathrm{SL}_2(\mathbb{Z})/\Gamma$ (see Sections 1.3.3 and 1.3.5), enables us to use the results of Section 2.1.2 to conclude $\epsilon_2(\Gamma_1(2)) = \epsilon_2(\Gamma_0(2)) = 1$, $\epsilon_3(\Gamma_1(2)) = 0$.

Remark. In the same way, one can note that $X_0(3) = X_1(3)$ and use the results of the previous chapter to conclude that $\epsilon_2(\Gamma_1(3)) = 0$ and $\epsilon_3(\Gamma_1(3)) = 1$..

Genus

We find

$$g(X_1(N)) = g_1(N) = \begin{cases} 1 + \frac{N^2}{24}\prod_{p|N}\left(1 - \frac{1}{p^2}\right) - \frac{\phi(N)}{4}\prod_{p^\alpha || N}\left(\alpha + 1 - \frac{\alpha - 1}{p}\right) & \text{If } N \geq 4 \\ 0 & \text{If } N = 2,3 \end{cases}$$

3.1.5 Cuspidal trees and fundamental domains

We recall that the degree of a map $\psi : X(\Gamma_1) \to X(\Gamma_2)$ equals the index of Γ_1 in Γ_2 (divided by 2 if $-I \in \Gamma_2 \setminus \Gamma_1$). In the previous sections we have shown that

$$\varphi : X(N) \longrightarrow X_1(N) \text{ has degree } N$$

$$\psi : X_1(N) \longrightarrow X_0(N) \text{ has degree } \frac{N}{2}\prod_{p|N}\left(1 - \frac{1}{p}\right)$$

$$\phi : X_0(N) \longrightarrow X(1) \text{ has degree } N\prod_{p|N}\left(1 + \frac{1}{p}\right)$$

A cuspidal tree is a stratified weighted diagram where

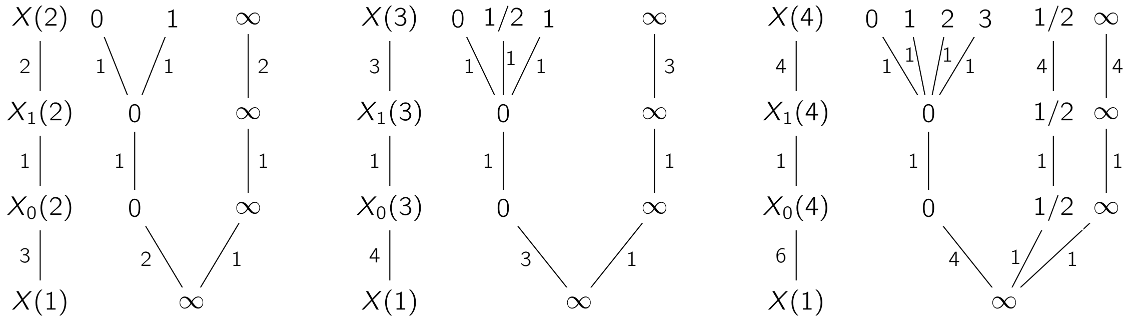
1. The vertices are sets of cusps of all modular curves involved.

2. There is an edge between two cusps if and only if there exists a non-trivial map taking one to the other.

3. The weights of the edges are the ramification degrees.

4. The stratification is the set of modular curves partially ordered by $X(\Gamma_1) > X(\Gamma_2)$ if $\Gamma_1 \subseteq \Gamma_2$.

In Appendix A, we describe some sort of cuspidal tree for models for $X_0(N)$. In that case we restrict our attention to the lower level of the stratification.



We would like now to describe the fundamental domain of the newly introduced modular curves. To do so we could "climb" the cuspidal tree and construct the fundamental domain of $X_1(N)$ starting from the knowledge of the one of $X_0(N)$ and then we will recover the one for the full level-N modular curve. Otherwise we could proceed by looking at coset representatives for $\Gamma_1(N)$ or $\Gamma(N)$ inside the full modular group $\mathrm{SL}_2(\mathbb{Z})$. We will mainly refer to Section 2.1.4 and to.[Kul].

Let us start by studying $X_1(5)$. First of all we look at the fundamental domain for $X_0(5)$.

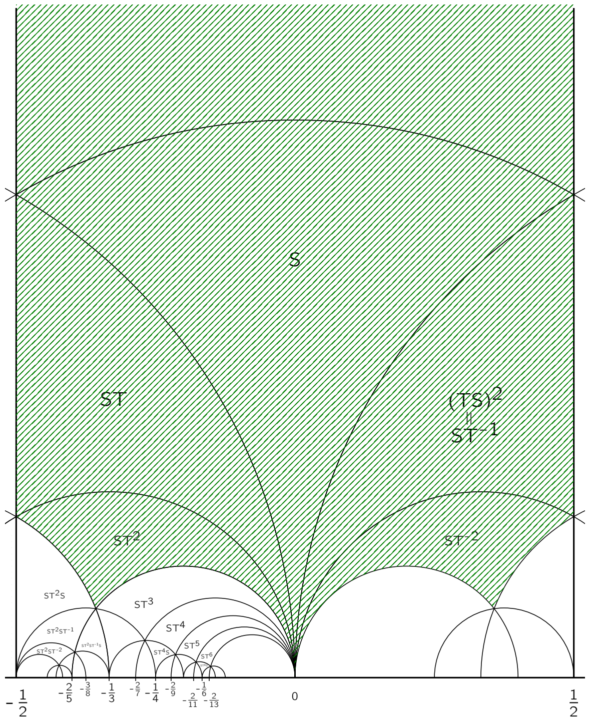


Figure 3.1 – The fundamental region for $\Gamma_0(5)$.

Now we look at the quotient $\Gamma_0(5)/\Gamma_1(5)$. This is well described by the short exact sequence

$$1 \longrightarrow \Gamma_1(N) \longrightarrow \Gamma_0(N) \longrightarrow (\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow 1$$

where the third map sends $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \to a \mod N$. We have to be careful because we have to take care of the quotient by $\{\pm I\}$. In case $N = 5$ we get 2 coset representatives

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

Hence, we have to study the behavior of $A = ST^{-2}ST^{-3}S$ on the fundamental domain of $X_0(5)$. A better decomposition is the one coming from another lift of $3 \in (\mathbb{Z}/5\mathbb{Z})^\times$, namely $B = \begin{pmatrix} -2 & -1 \\ 5 & 2 \end{pmatrix} = ST^2ST^{-2}S$; here "better" refers to the possibility of obtaining a connected fundamental domain.

With reference to Figure 2.9 we know that rotation counterclockwise in the triangles with vertex in the same cusp correspond to multiplication by T on the right while multiplying by S (still on the right) would result in moving to the only adjacent triangle with vertex in a different cusp.

Already looking in Figure 3.1 we know that the last two triangles ($ST^2ST^{-\delta}$ for $\delta = 3, 4$) will be really small and difficult to spot. For this reason and this reason only we will try to see if we could construct a more appealing picture.

We will therefore look for a matrix $\Lambda \in \Gamma_1(5)$ such that $ST^2ST^{-1}S = \Lambda ST^2ST^{-\delta}$ (the choice of the conjugate is not casual but comes from observing the output of Verrill's algorithm [Ver1] and [Ver2]).

$$ST^2ST^{-1}S = \Lambda ST^2ST^{-\delta} \Longrightarrow \Lambda = ST^2ST^{-1}ST^{\delta}ST^{-2}S$$

Hence, for $\delta = 3$, we find

$$\Lambda = \begin{pmatrix} -9 & -4 \\ 25 & 11 \end{pmatrix} \in \Gamma_1(5)$$

In the same way we find $\Lambda' \in \Gamma_1(5)$ such that $ST^{-2}S = \Lambda' ST^2ST^{-4}$.

$$-\Lambda' = -ST^{-2}ST^4ST^{-2}S = \begin{pmatrix} -9 & -4 \\ -20 & -9 \end{pmatrix}$$

Finally we have a picture for the fundamental domain of $X_1(5)$.

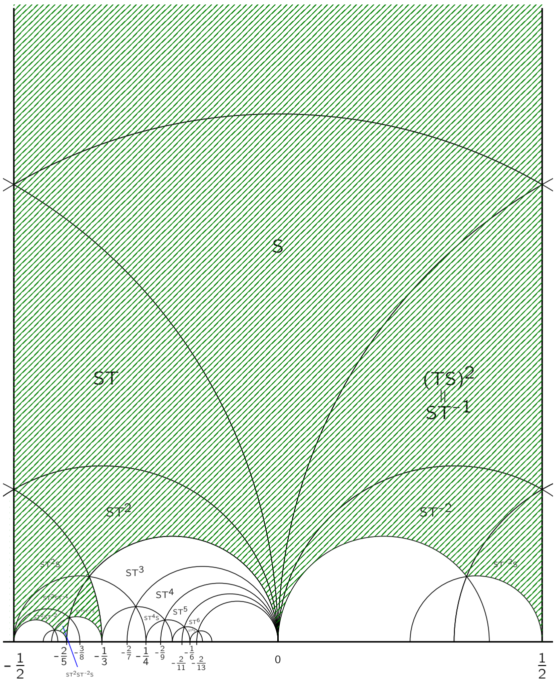


Figure 3.2 – The fundamental region for $\Gamma_1(5)$.

We conclude this section with the fundamental domain of $X(2)$. Note that $X_1(2) = X_0(2)$ and in the same way $X_1(3) = X_0(3)$ and $X_1(4) = X_0(4)$; this motivates the choice of looking at $X_1(5)$ before.

$\Gamma(2)$ is the kernel of the reduction map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$ and, as $\mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z}) = \Gamma_2(\mathbb{Z}/2\mathbb{Z})$ is isomorphic to S_3, it follows that $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(2)] = 6$, i.e., $\Gamma(2)$ can be written as disjoint union of 6 cosets.

Representatives of these cosets are

$$\Lambda_1 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \Lambda_2 = T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \Lambda_3 = S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$\Lambda_4 = TS = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \qquad \Lambda_5 = ST = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \qquad \Lambda_6 = TST^{-1} = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$$

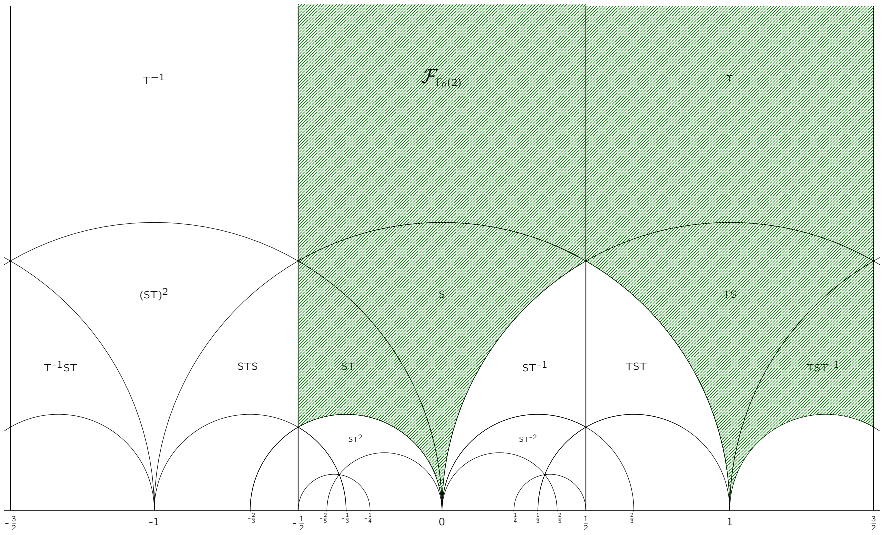


Figure 3.3 – The fundamental region for $\Gamma(2)$.

3.1.6 Generalized Dedekind eta function

Let Γ be a congruence subgroup; we denote by $\mathbb{Q}(\Gamma)$ the field of modular functions invariant under the action of Γ. If the modular curve $X(\Gamma) = \mathbb{H}/\Gamma$ has genus zero, then the function field $\mathbb{Q}(\Gamma)$ can be generated by a single function. Further, if Γ contains $\Gamma_0(N)$ for some N we could find a generator in the form of product of Dedekind η-functions. However, if Γ does not contain $\Gamma_0(N)$ Eta products are not sufficient.

In section 2.2.3 we have described the properties of Dedekind η-functions; following [Yan2], we will now define the generalized Dedekind η-function and study its transformation properties. Let

$$E_{g,h}(\tau) = q^{B(g/N)/2} \prod_{m=1}^{+\infty} \left(1 - e^{2\pi i h/N} q^{m-1+g/N}\right) \left(1 - e^{-2\pi i h/N} q^{m-g/N}\right)$$

for g and h not simultaneously congruent to 0 modulo N and

$$E_g(\tau) = q^{NB(g/N)/2} \prod_{m=1}^{+\infty} \left(1 - q^{(m-1)N+g}\right) \left(1 - q^{mN-g}\right)$$

for $g \not\equiv 0 \bmod N$. Here $B(x) = x^2 - x + 1/6$.

Remark. $E_{g,h}$ is called *generalized* Dedekind η-function because it reduces to η^2 when $g, h \equiv 0 \mod N$.

Proposition 3.24 ([Yan2, Prop. 1])**.** *The functions $E_{g,h}$ satisfy the following transformation properties:*

$$E_{g+N,h} = E_{-g,-h} = -\zeta_N E_{g,h}$$

$$E_{g,h+N} = E_{g,h}$$

If $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$*, then*

$$E_{g,h}(\gamma\tau) = \begin{cases} E_{g,h}(\tau + b) = e^{\pi i b B(g/N)} E_{g,bg+h}(\tau) & \textit{If } c = 0 \\ \epsilon(a,b,c,d) e^{\pi i \delta} E_{g',h'}(\tau) & \textit{If } c \neq 0 \end{cases}$$

where

$$\epsilon(a,b,c,d) = \begin{cases} e^{\pi i (bd(1-c^2)+c(a+d-3))/6} & \textit{If } c \textit{ is odd} \\ -ie^{\pi i (ac(1-d^2)+d(b-c+3))/6} & \textit{If } d \textit{ is odd} \end{cases}$$

$$\delta = \frac{g^2ab + 2ghbc + h^2cd}{N^2} - \frac{gb + h(d-1)}{N}$$

$$(g', h') = (g, h) \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Proposition 3.25. *The functions* E_g *satisfy the following transformation properties:*

$$E_{g+N} = E_{-g} = -E_g$$

And, if $\gamma \begin{pmatrix} a & b \\ cN & d \end{pmatrix} \in \Gamma_0(N)$*, then*

$$E_g(\gamma\tau) = \begin{cases} E_{g,h}(\tau + b) = e^{\pi i b N B(g/N)} E_g(\tau) & \textit{If } c = 0 \\ \epsilon(a, bN, c, d) e^{\pi i g^2 sb/N - gb} E_{ag}(\tau) & \textit{If } c \neq 0 \end{cases}$$

Proposition 3.26. *Let* $f(\tau) = \prod_g E_g(\tau)^{e_g}$ *for* $g, e_g \in \mathbb{Z}$ *with* $g \not\equiv 0 \mod N$*. If*

$$\sum_g e_g \equiv 0 \mod 12 \qquad \sum_g g e_g \equiv 0 \mod 2$$

then f *is invariant under the action of* $\Gamma(N)$*. If, in addition,*

$$\sum_g g^2 e_g \equiv 0 \mod 2N$$

then f *is a modular function on* $\Gamma_1(N)$*.*

Proposition 3.27 ([Yan1, Lemma 2])**.** *Let* $N \in \mathbb{Z}_{\geq 0}$ *and* $\gamma \in \mathrm{SL}_2(\mathbb{Z})$*.*

$$E_g(\gamma\tau) = \epsilon q^{\delta} + \textit{higher powers}$$

where $|\epsilon| = 1$*,*

$$\delta = \frac{(c,N)^2}{2N} P_2\left(\frac{ag}{(c,N)}\right)$$

where $P_2 = \{x\}^2 - \{x\} + 1/6$ *denotes the second Bernoulli polynomial for* $\{x\} = x - \lfloor x \rfloor$ *the fractional part of* x*.*

Proof. We know

$$1 - x^N = \prod_{h=0}^{N-1} (1 - \zeta_N^h x)$$

Thus,

$$E_g(\tau) = \prod_{h=0}^{N-1} E_{g,h}(\tau) \implies E_g(\gamma\tau) = \prod_{h=0}^{N-1} E_{g,h}(\gamma\tau) = \epsilon' \prod_{h=0}^{N-1} E_{ag+ch,bg+dh}(\tau) = \epsilon' \prod_{h=0}^{N-1} E_{g',h'}(\tau)$$

with $|\epsilon'| = 1$. And the leading term will be

$$\epsilon(a,b,c,d)^N e^{i\pi \sum_{h=0}^{N-1} \delta_{g,h}} q^{\sum_{h=0}^{N-1} B(g'/N)/2}$$

This means that the smallest term is the one of order

$$\lambda = \sum_{h=0}^{N-1} \frac{B(g'/N)}{2} = \frac{1}{2}\sum_{h=0}^{N-1} P_2\left(\frac{ag+ch}{N}\right)$$

by the transformation formulas in Proposition 3.24. Now P_2 is periodic (because $\{x\}$ is periodic) and so we can find its Fourier expansion.

Theorem (Fundamental Theorem on Convergence of Fourier Series). *Let $f(x)$ be a periodic function of period $\rho = 2L$ which is piece-wise differential on $[-L, L]$ (this means that it is continuous on $[-L, L]$ and there exists a finite number of points $x_0 = -L < x_1 < \ldots < x_n < L = x_{n+1}$ such that $f|_{(x_i, x_{i+1})}$ is everywhere differentiable). Then $f(x)$ has Fourier expansion*

$$f(x) = a_0 + \sum_{n=1}^{+\infty}\left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right)$$

where

$$a_0 = \frac{1}{2L}\int_{-L}^{L} f(x)dx \qquad a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}dx \qquad b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}dx$$

If x is a point of continuity of f, then the series converges in x to the value of $f(x)$.

Now $\{x\}$ has period 1 and so has P_2; further P_2 is piece-wise differential.

$$a_0 = \int_{-1/2}^{1/2} \{x\}^2 - \{x\} + \frac{1}{6}dx = 2\int_0^{1/2} x^2 - x + \frac{1}{6}dx = 2\left[\frac{x^3}{3} - \frac{x^2}{2} + \frac{x}{6}\right]_0^{1/2} = 0$$

$$\begin{aligned}
a_n =&4\int_0^{1/2}\left(x^2 - x + \frac{1}{6}\right)\cos(2n\pi x)dx = \\
=&4\left[\left(x^2 - x + \frac{1}{6}\right)\frac{\sin(2\pi n x)}{2\pi n}\right]_0^{1/2} - \frac{4}{2\pi n x}\int_0^{1/2}(2x-1)\sin(2\pi n x)dx = \\
=&\left[\frac{4}{2\pi n}\left(x^2 - x + \frac{1}{6}\right)\sin(2\pi n x) + \frac{4}{4\pi^2 n^2}(2x-1)\cos(2\pi n x) - \frac{4}{8\pi^3 n^3}\sin(2\pi n x)\right]_0^{1/2} = \frac{1}{n^2\pi^2}
\end{aligned}$$

Finally, $b_n = 0$ since P_2 is an odd function. Thus,

$$P_2(x) = \sum_{n=1}^{+\infty}\frac{\cos(2\pi n x)}{n^2\pi^2}$$

from which

$$\begin{aligned}
\lambda =&\frac{1}{2}\sum_{h=0}^{N-1} P_2\left(\frac{ag+ch}{N}\right) = \frac{1}{2}\sum_{h=0}^{N-1}\sum_{n=1}^{+\infty}\frac{\cos\left(\frac{2\pi n}{N}(ag+ch)\right)}{n^2\pi^2} = \\
=&\frac{1}{2\pi^2}\sum_{n=1}^{+\infty}\frac{1}{n^2}\sum_{h=0}^{N-1}\cos\left(\frac{2\pi n}{N}(ag+ch)\right)
\end{aligned}$$

Now

$$\begin{aligned}
&\sum_{h=0}^{N-1}\cos\left(\frac{2\pi n}{N}(ag+ch)\right) = \sum_{h=0}^{N-1}\cos\left(\frac{2\pi n}{N}ag + \frac{2\pi n}{N}ch\right) = \\
&= \cos\left(\frac{2\pi n}{N}ag\right)\sum_{h=0}^{N-1}\cos\left(\frac{2\pi n}{N}ch\right) - \sin\left(\frac{2\pi n}{N}ag\right)\sum_{h=0}^{N-1}\sin\left(\frac{2\pi n}{N}ch\right)
\end{aligned}$$

and

$$\begin{aligned}\sum_{h=0}^{N-1} \cos\left(\frac{2\pi n}{N} ch\right) &= \frac{1}{2}\left[\sum_{h=0}^{N-1} e^{2\pi i nch/N} + \sum_{h=0}^{N-1} e^{-2\pi i nch/N}\right] \\ &= \frac{1}{2}\left[\sum_{h=0}^{N-1}\left(e^{2\pi i nc/N}\right)^h + \sum_{h=0}^{N-1}\left(e^{-2\pi i nc/N}\right)^h\right] = \frac{1}{2}\left[\sum_{h=0}^{N-1}(\zeta_N^{cn})^h + \sum_{h=0}^{N-1}(\zeta_N^{-cn})^h\right] \\ &= \begin{cases}\frac{1}{2}[N+N] \\ \frac{1}{2}\left[\frac{\zeta_N^{cnN}-1}{\zeta_N^{cn}-1} + \frac{\zeta_N^{-cnN}-1}{\zeta_N^{-cn}-1}\right]\end{cases} = \begin{cases} N & \text{If } N \mid cn \\ 0 & \text{If } N \nmid cn\end{cases}\end{aligned}$$

Exactly in the same way

$$\sum_{h=0}^{N-1} \sin\left(\frac{2\pi n}{N} ch\right) = \frac{1}{2}\left[\sum_{h=0}^{N-1} e^{2\pi i nch/N} - \sum_{h=0}^{N-1} e^{-2\pi i nch/N}\right] = 0$$

since the minus sign reduces to zero the case $N \mid cn$ too. Hence,

$$\lambda = \frac{1}{2\pi^2}\sum_{n=1}^{+\infty}\frac{1}{n^2}\sum_{h=0}^{N-1}\cos\left(\frac{2\pi n}{N}(ag+ch)\right) = \frac{1}{2\pi^2}\sum_{\substack{n=1\\ N\mid nc}}^{+\infty}\frac{N}{n^2}\cos\left(\frac{2\pi n}{N}ag\right)$$

We observe that $N \mid nc \Leftrightarrow N/(N,c) \mid n$. Further, the first n for which $N \mid nc$ is $n = N/(N,c)$; indeed, if $c = p_1^{e_1} \cdot \ldots \cdot p_k^{e_k}$, $n = p_1^{g_1} \cdot \ldots \cdot p_k^{g_k}$ and $N = p_1^{f_1} \cdot \ldots \cdot p_k^{f_k}$ are the prime factorization of c, n and N, then $N \mid nc$ implies $e_i + g_i - f_i \geq 0$ for all i and, therefore, we take $g_i = 0$ if $e_i \geq f_i$ and $g_i = f_i - e_i$ if $f_i > e_i$. This says that $n = N/(N,c)$.

Thus, we write $mN/(N,c)$ and

$$\lambda = \frac{1}{2\pi^2}\sum_{\substack{n=1\\ N\mid nc}}^{+\infty}\frac{N}{n^2}\cos\left(\frac{2\pi n}{N}ag\right) = \frac{N}{2\pi^2}\sum_{m=1}^{+\infty}\frac{(N,c)^2}{m^2N^2}\cos\left(\frac{2\pi m}{(N,c)}ag\right) = \frac{(N,c)^2}{2\pi^2 N}\sum_{m=1}^{+\infty}\frac{1}{m^2}\cos\left(\frac{2\pi m}{(N,c)}ag\right)$$

and, remembering the Fourier expansion, we get

$$\lambda = \frac{(N,c)^2}{2\pi^2 N}\sum_{m=1}^{+\infty}\frac{1}{m^2}\cos\left(\frac{2\pi m}{(N,c)}ag\right) = \frac{(N,c)^2}{2N}P_2\left(\frac{ag}{(N,c)}\right)$$

□

3.1.7 Models for $X_1(N)$

We will now use Proposition 3.27 to show that we can construct modular functions with poles only at ∞ using generalized Dedekind η-products.

Let $N > 4$ (for $N = 2, 3, 4$ we have already seen that $X_1(N) = X_0(N)$). A cusp on $X_1(N)$ is said to be of first type if it lies above the cusp 0 on $X_0(p)$ for all $p \mid N$. We let

$$\mathcal{F}_1^0(N) = \left\{\begin{array}{c}\text{Group of functions on } X_1(N) \text{ whose} \\ \text{divisors have support within the} \\ \text{cusps of finite type}\end{array}\right\}$$

and $\mathcal{F}_1'(N)$ be the group generated by Generalized Dedekind η-products $\prod_{h=1}^{N-1} E_{0,h}^{e_h}$ satisfying the conditions

$$\sum_{h=1}^{N-1} h^2 e_h \equiv 0 \begin{cases}\text{mod } N \text{ if } N \text{ is odd} \\ \text{mod } 2N \text{ if } N \text{ is even}\end{cases} \qquad \text{and} \qquad \sum_{\substack{h\equiv\pm a \\ \text{mod } N/p}} e_h = 0 \quad \begin{array}{l}\text{for all } p \mid N \text{ and} \\ \text{for all residue classes } a \text{ in } \mathbb{Z}/p\mathbb{Z}\end{array}$$

Proposition 3.28 ([Yu, Th.s 2 and 4])**.** *$\mathcal{F}_1^0(N) = \mathcal{F}_1'(N)$ and they are of rank $\phi(N)/2 - 1$.*

In the same way (acting with the Atkin-Lehner involution) we define

$$\mathcal{F}_1^\infty(N) = \left\{ \begin{matrix} \text{Group of modular functions on } X_1(N) \\ \text{whose divisors have support within the} \\ \text{cusps lying above } \infty \in X_0(N) \end{matrix} \right\}$$

and $\mathcal{F}_1''(N)$ as the group generated by Generalized Dedekind η-products $\prod_{g=1}^{N-1} E_g(\tau)^{e_g}$ verifying

$$\sum_{g=1}^{N-1} g^2 e_g \equiv 0 \begin{cases} \text{mod } N \text{ if } N \text{ is odd} \\ \text{mod } 2N \text{ if } N \text{ is even} \end{cases} \qquad \text{and} \qquad \sum_{\substack{g \equiv \pm a \\ \text{mod } N/p}} e_g = 0 \quad \begin{matrix} \text{for all } p \mid N \text{ and} \\ \text{for all residue classes } a \text{ in } \mathbb{Z}/p\mathbb{Z} \end{matrix}$$

Proposition 3.29. *$\mathcal{F}_1^\infty(N) = \mathcal{F}_1''(N)$ and they are of rank $\phi(N)/2 - 1$.*

Proposition 3.30. *The group $\mathcal{F}_1^\infty(N)$ contains at least two functions that have poles only at infinity and such that their orders of poles are coprime.*

This will enable us to use Theorem 2.53 to find models for the curves $X_1(N)$.

Remark. Since $\Gamma_1(N)$ is normal in $\Gamma_0(N)$, from every modular function $f \in \mathbb{C}(X_1(N))$ we can recover a modular function on $X_0(N)$ by acting on f with coset representatives in $\Gamma_0(N)/\Gamma_1(N)$:

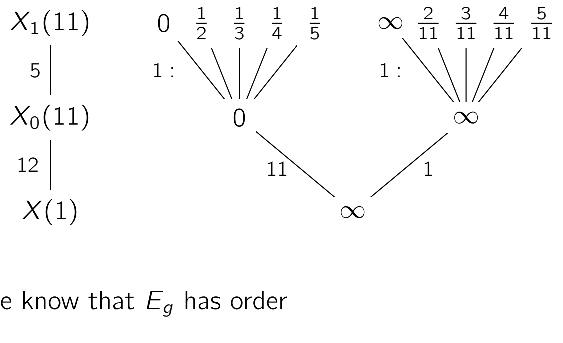
$$\sum_{\gamma \in \Gamma_0(N)/\Gamma_1(N)} f(\gamma\tau)$$

is a modular function on $X_0(N)$.

We define $W_k = E_{4k}/E_{2k}$. It is obvious that each W_k satisfies $\sum_g e_g = 1 - 1 \equiv 0 \mod 12$ and $\sum_g g e_g = 4k - 2k \equiv 0 \mod 2$. Thus, any product of W_k's will also satisfy these conditions. Hence, because of Proposition 3.26, $\prod_k W_k$ is a modular function on $X_1(N)$ if and only if

$$\sum_k k^2 e_k \equiv 0 \mod N \tag{3.1}$$

Example ($N = 11$)**.** We begin by studying the cuspidal tree for $N = 11$



Now, by Proposition 3.27, we know that E_g has order

$$\frac{(c,N)^2}{2N} P_2\left(\frac{ag}{(c,N)}\right)$$

at a cusp $\mathfrak{c} = a/c$. This implies that all the E_g's have same order at the cusps above $0 \in X_0(11)$. Hence, W_k can have zero or poles only at the cusps above $\infty \in X_0(11)$: $\mathfrak{c}_j = j/11$ for $j = 1, \ldots, 5$. In the following table we describe the orders $\nu_k(\mathfrak{c}_j)$ of W_k for $1 \leq k \leq 5$ (because of the rank of $\mathcal{F}_1^\infty$).

	$\mathfrak{c}_1$	$\mathfrak{c}_2$	$\mathfrak{c}_3$	$\mathfrak{c}_4$	$\mathfrak{c}_5$
$11\nu_1$	-5	2	10	-3	-4
$11\nu_2$	2	-3	-4	10	-5
$11\nu_3$	10	-4	2	-5	-3
$11\nu_4$	-3	10	-5	-4	2
$11\nu_5$	-4	-5	-3	2	10

This implies that, in order to get a function with only a pole at infinity $\mathfrak{c}_1$ of order δ we have to solve the following integer system

$$\begin{cases} -5x_1+2x_2+10x_3-3x_4-4x_5=-11\delta \\ 2x_1-3x_2-4x_3+10x_4-5x_5\geq 0 \\ 10x_1-4x_2+2x_3-5x_4-3x_5\geq 0 \\ -3x_1+10x_2-5x_3-4x_4+2x_5\geq 0 \\ -4x_1-5x_2-3x_3+2x_4+10x_5\geq 0 \\ x_1+4x_2+9x_3+16x_4+25x_5\equiv 0 \mod 11 \end{cases}$$

where the last line corresponds to the condition 3.1 which guarantees the modularity of the W_k-product. Using `lp_solve` we find two solutions for $\delta=2$ and 3 (which is expected since $X_1(N)$ is an elliptic curve). In particular, for $\delta=2$ we find the solution $(1,0,-2,-1,0)$ which gives

$$X=-\frac{W_1}{W_3^2W_4}=q^{-2}+2q^{-1}+4+5q+6q^2+5q^3+3q^4-q^5-6q^6+\dots$$

and for $\delta=3$ we get $(1,0,-2,0,2)$ yielding

$$Y=\frac{W_1W_5^2}{W_3^2}=q^{-3}+4q^{-2}+9q^{-1}+16+24q+30q^2+30q^3+21q^4+2q^5-26q^6+\dots$$

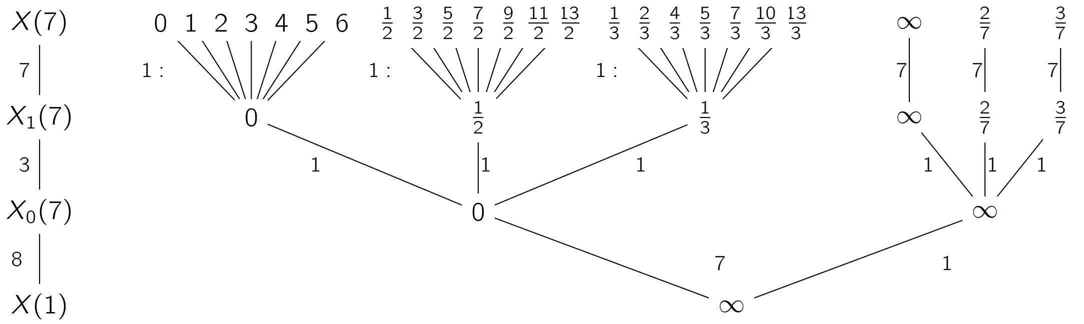
We find linear relations giving an equation for $X_1(11)$

$$Y^2-2XY+Y=X^3-2X+X$$

3.1.8 Models for $X(N)$

The procedure to find models for $X(N)$ follows exactly the one for $X_1(N)$. As noticed before the example for $X_1(11)$, if we define $W_k=E_{4k}/E_{2k}$, we observe that W_k satisfies $\sum_g e_g=1-1\equiv 0 \mod 12$ and $\sum_g ge_g=4k-2k\equiv 0 \mod 2$. Thus, thanks to Proposition 3.26, $\prod_k W_k$ is a modular function on $X(N)$ (Note that we do not need anymore Condition 3.1).

Example ($N=7$)**.** $X(7)$ is a genus 3 curve. Its cuspidal tree looks as follows:


$X(7)$
7
$X_1(7)$
3
$X_0(7)$
8
$X(1)$
0 1 2 3 4 5 6
$\frac{1}{2}$ $\frac{3}{2}$ $\frac{5}{2}$ $\frac{7}{2}$ $\frac{9}{2}$ $\frac{11}{2}$ $\frac{13}{2}$
$\frac{1}{3}$ $\frac{2}{3}$ $\frac{4}{3}$ $\frac{5}{3}$ $\frac{7}{3}$ $\frac{10}{3}$ $\frac{13}{3}$
∞ $\frac{2}{7}$ $\frac{3}{7}$
1 :
1 :
1 :
7 7 7
0
$\frac{1}{2}$
$\frac{1}{3}$
∞ $\frac{2}{7}$ $\frac{3}{7}$
1 1 1
1 1 1
0
∞
7
1
∞

We observe that E_g has order

$$\frac{(c,N)^2}{2N}P_2\left(\frac{ag}{(c,N)}\right)$$

at a cusp $\mathfrak{c}=a/c$. Now,

- for all $\mathfrak{c} \in \{0, \dots, 6\} \subseteq \mathrm{Cusps}(X(7))$, $\mathrm{ord}_{\mathfrak{c}} E_g = 0$;
- for all $\mathfrak{c} = i/2 \in \mathrm{Cusps}(X(7))$ with $i \in \{1, 3, 5, 7, 9, 11, 13\}$, $\mathrm{ord}_{\mathfrak{c}} E_g = 0$;
- for all $\mathfrak{c} = j/3 \in \mathrm{Cusps}(X(7))$ with $j \in \{1, 2, 4, 5, 7, 10, 13\}$, $\mathrm{ord}_{\mathfrak{c}} E_g = 0$;

i.e., all the E_g's have same order at the cusps above $0 \in X_0(7)$.

Remark. Note that this happens because the denominators of all the cusps above 0 are coprime with 7, i.e., the argument of P_2 is integral and $P_2(\mathbb{Z}) = 1/6$.

The W_k's can have zero or poles only at the cusps above $\infty \in X_0(11)$: $\infty, 2/7$ and $3/7$. We display the orders $\nu_k(\mathfrak{c}_j)$ of W_k for $\mathfrak{c}_j = j/7$, $1 \leq k \leq 3$ in the following table

	$\mathfrak{c}_1$	$\mathfrak{c}_2$	$\mathfrak{c}_3$
$7\nu_1$	-1	3	-2
$7\nu_2$	3	-2	-1
$7\nu_3$	-2	-1	3

Therefore, we have to solve the integer linear problem

$$\begin{cases} -x_1 + 3x_2 - 2x_3 = -7\delta \\ 3x_1 - 2x_2 - x_3 \geq 0 \\ -2x_1 - x_2 + 3x_3 \geq 0 \end{cases}$$

We find a function with a pole of order 3 at ∞

$$X = -\frac{1}{W_2^7} = q^{-3} + 7q^{-2} + 28q^{-1} + 77 + 154q + 217q^2 + 168q^3 - 97q^4 - 546q^5 + \dots$$

and another one with a pole of order 5:

$$Y = -\frac{1}{W_2^{13} W_3^2} = q^{-5} + 11q^{-4} + 68q^{-3} + 295q^{-2} + 980q^{-1} + 2583 + 5435q + 8868q^2 + 9964q^3 + \dots$$

The algebraic relation between these two functions yields an equation for the modular curve of full level 7 structure: $Y^3 = X^5 - 2X^3Y + XY^2$.

Otherwise, following [CKK], we can change the local parameter at ∞, taking $q_7 = e^{2\pi i/7}$ instead. This would transform the linear system in the following

$$\begin{cases} -x_1 + 3x_2 - 2x_3 = -\delta \\ 3x_1 - 2x_2 - x_3 \geq 0 \\ -2x_1 - x_2 + 3x_3 \geq 0 \end{cases}$$

which, for has solution

$$X = -W_1(q_7)W_3(q_7) = q_7^{-3} + q_7^4 + q_7^{11} - q_7^{25} - q_7^{32} + q_7^{46} + 2q_7^{53} + q_7^{60} + \dots$$

for $\delta = 3$ and

$$Y = -W_1(q_7)W_3(q_7) = q_7^{-5} + 2q_7^2 + 2q_7^9 + q_7^{16} - q_7^{23} - 3q_7^{30} - 2q_7^{37} + q_7^{44} + 5q_7^{51} + 6q_7^{58} + \dots$$

for $\delta = 5$. This yields a model for $X(7)$: $Y^3 = X^5 + XY$.

Note that, in general, from Proposition 3.24, we have

Lemma 3.31 ([Yan1, Cor. 1])**.** *Let (g, h) be pairs of integers, and suppose that $e_{g,h}$ are integers such that*

$$\sum_{(g,h)} e_{g,h} \equiv 0 \mod 12$$

and

$$\sum_{(g,h)} g^2 e_{g,h} \equiv \sum_{(g,h)} ghe_{g,h} \equiv \sum_{(g,h)} h^2 e_{g,h} \equiv 0 \mod 2N$$

Then the product $f(\tau) = \prod_{(g,h)} E_{g,h}(\tau)^{e_{g,h}}$ *is a modular function on* $\Gamma(N)$.

This says that $E_{g,h}^{12N}$ is a modular function on $\Gamma(N)$ with divisor supported on the cusps, see [KL, Th. II.1.2].

3.2 Non-split Cartan modular curves

In this section we study a different type of congruence subgroups called Cartan modular groups (and their normalizers).

We present here a toy example. Suppose p is an odd prime and $\lambda \in \mathbb{Z}/p\mathbb{Z}$ is not a square. Let Γ be the set of matrices in $SL_2(\mathbb{Z})$ defined by

$$\Gamma = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \, a \equiv d \text{ and } b \equiv \lambda c \mod p \right\}$$

One can easily verify that this is a congruence subgroup since it contains $\Gamma(p)$. Since λ is not a square in $\mathbb{F}_p$, then the polynomial $f(X) = X^2 - \lambda$ has no solution in $\mathbb{F}_p$. We construct the splitting field of $f(X) \in \mathbb{F}_p[X]$ by adjoining a root of $f(X)$ to $\mathbb{F}_p$ and we obtain $\mathbb{F}_p[\sqrt{\lambda}] = \mathbb{F}_{p^2}$.
Now consider the action of $\mathrm{SL}_2(\mathbb{Z})$ defined as usual

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(\tau) = \frac{a\tau + b}{c\tau + d}$$

This induces an action of $\mathrm{SL}_2(\mathbb{Z})$ on $\tau \in \mathbb{F}_{p^2} - \mathbb{F}_p$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(s\sqrt{\lambda} + t) = \frac{ac(t^2 - s^2\lambda) + (ad + bd)t + (ad - bc)\sqrt{\lambda}}{2dtc + c^2(t^2 - s^2\lambda)} \in \mathbb{F}_{p^2} - \mathbb{F}_p \qquad \text{since } \gamma \in SL_2(\mathbb{Z})$$

This action is transitive and such that $\mathrm{Stab}_{SL_2(\mathbb{Z})}(\sqrt{\lambda}) = \Gamma$, meaning that

$$\Gamma \backslash SL_2(\mathbb{Z}) \simeq \mathbb{F}_{p^2} - \mathbb{F}_p \qquad \Rightarrow \qquad \#\Gamma \backslash \mathrm{SL}_2(\mathbb{Z}) = p^2 - p = p(p-1)$$

. If we look at the action of S and on $\mathbb{F}_{p^2} - \mathbb{F}_p$, we get

$$\#S\text{-orbits of size } 1 = \begin{cases} 0 & \text{if } p \equiv 1 \mod 4 \\ 2 & \text{if } p \equiv 3 \mod 4 \end{cases}$$

The action of ST on $\mathbb{F}_{p^2} - \mathbb{F}_p$ yields

$$\#ST\text{-orbits of size } 1 = \begin{cases} 0 & \text{if } p \equiv 1 \mod 3 \\ 2 & \text{if } p \equiv 2 \mod 3 \end{cases}$$

Finally, the action of T on $\mathbb{F}_{p^2} - \mathbb{F}_p$ is given by $T(\tau) = \tau + 1$ and therefore T acts on $\mathbb{F}_{p^2} - \mathbb{F}_p$ by translation.

$$\#T\text{-orbits} = p - 1$$

Combining all of these produces the following genus formula:

$$g(X_\Gamma) = \frac{(p-6)(p-1) + a}{12} \qquad \text{where} \qquad a = \begin{cases} 12 & \text{if } p \equiv 1 \mod 12 \\ 4 & \text{if } p \equiv 5 \mod 12 \\ 6 & \text{if } p \equiv 7 \mod 12 \\ -2 & \text{if } p \equiv 11 \mod 12 \end{cases}$$

$X(\Gamma)$ is an example of a Cartan modular curve. In this section we generalize this construction to a generic N and we study the normalizers of these groups.

3.2.1 Quotients of the modular curve $X(N)$

Let N be a positive integer N. In the previous chapters we have studied the modular curve $X(N)$ with full level N-structure classifying elliptic curves with a basis of their N-torsion.

When studying modular curves one of the most appreciated approaches is to focus on their function fields. We know that $\mathbb{Q}(X(1)) = \mathbb{Q}(j)$. The field of functions of $X(N)$ is the field of modular functions of level N rational over its field of definition $\mathbb{Q}(\zeta_N)$ [Shi, Ch. 6] or [DS, Ch. 6]. Each function field $\mathbb{Q}(X(N)) = \mathbb{Q}(\zeta_N)(X(N)) = \mathbb{Q}(\zeta_N, E[N])$ is a Galois extension of $\mathbb{Q}(j) = \mathbb{Q}(X(1))$ with Galois group

$$\mathcal{Gal}(\mathbb{Q}(X(N))/\mathbb{Q}(j)) \simeq \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$$

which is defined up to an inner automorphism. Once this is fixed, we have

$$\mathcal{Gal}(\mathbb{Q}(X(N))/\mathbb{Q}(\zeta_N, j)) \simeq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$$

with $\mathcal{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \simeq (\mathbb{Z}/N\mathbb{Z})^\times$.

$$\begin{tikzcd}
\mathbb{Q}(X(N)) \arrow[d, dash, "\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}"'] \arrow[dd, dash, bend left=60, "\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}"] \\
\mathbb{Q}(\zeta_N)(X(1)) \arrow[d, dash, "(\mathbb{Z}/N\mathbb{Z})^\times"'] \\
\mathbb{Q}(X(1))
\end{tikzcd}$$

The group $G = \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ acts on $X(N)$ and the action is defined on non cuspidal points by

$$\begin{aligned} \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) \times X(N) &\longrightarrow X(N) \\ (\Lambda, (E, \phi)) &\longrightarrow (E, \Lambda \circ \phi) \end{aligned}$$

(where ϕ is an isomorphism $E[N] \to (\mathbb{Z}/N\mathbb{Z})^2$) and is then extended uniquely to the whole compactified curve $X(N)$. The quotient of this action is the j-line $X(1) = \mathbb{P}^1$ [Shi, Prop 6.6].

What happens if we quotient by a subgroup instead of the entire general linear group? If H is a subgroup of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ such that $\det(H) = (\mathbb{Z}/N\mathbb{Z})^\times$, then the corresponding modular curve $X_H = H \backslash X(N)$ is geometrically connected over $\mathbb{Q}$, see [DR] and [Roh, §1.2]. The quotient curve is described as the set of equivalence classes of elliptic curve enhanced with full N-level structure under the relation

$$(E, \phi) \sim (F, \psi) \Longleftrightarrow E \simeq E' \text{ and there exists } \sigma \in H \text{ s.t. } \phi = \psi^\sigma$$

In other words, $X_H = X(N)/H = X(N)/\sim$.

Note that the curve X_H can also be constructed via Galois theory; the subgroup H (provided it contains $-I$) corresponds to a subgroup of $\mathcal{Gal}(\mathbb{Q}(X(N))/\mathbb{Q}(X(1)))$ and by Galois correspondence we get an intermediate extension of $\mathbb{Q}(X(N))/\mathbb{Q}(X(1))$ to which we associate a curve which will be the desired X_H. If we denote $\det(H)$ is the image of H under the determinant map, then the curve X_H is defined over $\mathbb{Q}(\zeta_N)^{\det(H)}$ and so, if $\det(H) = (\mathbb{Z}/N\mathbb{Z})^\times$, then X_H is defined over $\mathbb{Q}$.

Otherwise, one can construct X_H in a more geometric way: let $H' = H \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ and let $\Gamma = \Gamma_H$ be a lift (a pullback via the reduction map) of H' to $\mathrm{SL}_2(\mathbb{Z})$; then the curve $H_\Gamma(\mathbb{C})$ is isomorphic to X_H.

Finally, we can construct X_H directly as a quotient of $X(N)$. We know that $X(N)$ classifies isomorphism classes of elliptic curves together with some full level N-structure.

3.2.2 Galois representations and elliptic curves

Let E be an elliptic curve over $\mathbb{Q}$; the absolute Galois group $\mathcal{G}_{\mathbb{Q}} = \mathcal{G}a\ell(\bar{\mathbb{Q}}/\mathbb{Q})$ acts on the torsion points of E component-wise. Identifying $E[N]$ with $(\mathbb{Z}/N\mathbb{Z})^2$ we obtain a representation

$$\rho_N : \mathcal{G}a\ell(\bar{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$$

When we restrict our attention to the prime case $N = p$, Serre conjectured that, for every elliptic curve over $\mathbb{Q}$ without complex multiplication, this representation is surjective for all but finitely many primes, called *exceptional* primes (Serre's Uniformity Conjecture) [Ser2]. This would prove that

$$\varprojlim_{m\in\mathbb{Z}} \rho_{E,m}(\mathcal{G}a\ell(\bar{\mathbb{Q}}/\mathbb{Q})) \subset \varprojlim_{m\in\mathbb{Z}} \mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z}) \simeq \mathrm{GL}_2(\hat{\mathbb{Z}})$$

has finite index (open image theorem). To study the upper bound for the exceptional primes on can study the image of the representation. If $\rho_p\left(\mathcal{G}a\ell(\bar{\mathbb{Q}}/\mathbb{Q})\right)$ is not $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$, then it has to be contained in one of its maximal subgroups. Maximal subgroups of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ are well known and they are

- Borel subgroups;
- Exceptional subgroups isomorphic to A_4, S_4 and A_5 (irregular case);
- Normalizers of split Cartan subgroups;
- Normalizers of non-split Cartan subgroups.

Using the moduli space description of modular curves it can be proved that non CM elliptic curves over $\mathbb{Q}$ such that $\rho_{E,N}(\mathcal{G}_{\mathbb{Q}})$ is conjugate to a subgroup of H for some maximal subgroup $H \leq \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ gives rise to a (non-cuspidal) $\mathbb{Q}$-rational point on X_H:

Theorem 3.32. *Let $H \leq \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ and suppose that $\det(H) = (\mathbb{Z}/N\mathbb{Z})^\times$. Then*

(i) *Y_H and its compactification X_H are defined over $\mathbb{Q}$.*

(i) *Every $Q \in Y_H(\mathbb{Q})$ is supported on some elliptic curve $E/\mathbb{Q}$ (i.e., that there is some E/K and an isomorphism $\alpha : E[N] \to (\mathbb{Z}/N\mathbb{Z})^2$ such that $Q = [(E,\alpha)]_H$).*

(iii) *If $Q \in Y_H(\mathbb{Q})$ and $j(Q) \neq 0, 1728$, then $Q = [(E,\alpha)]_H$ such that E is defined over $\mathbb{Q}$ and $\rho_{E,N}(\mathcal{G}_{\mathbb{Q}}) \subseteq H$ (up to conjugation). Conversely, if there is E is defined over $\mathbb{Q}$ and $\rho_{E,N}(\mathcal{G}_{\mathbb{Q}}) \subset H$ (up to conjugation) then $[(E,\alpha)] \in Y_H(\mathbb{Q})$ for a suitable α.*

Further, if $-I \notin H$, then the elliptic curve E describing Q above is unique, otherwise it can be replaced by any quadratic twist.

This implies that Serre's uniformity conjecture is equivalent to

Conjecture. *Let H be a proper subgroup of $\mathrm{GL}_2(\mathbb{F}_p)$ such that $\det(H) = \mathbb{F}_p^\times$. Then there exists a positive constant C such that, for all primes $p > C_H$, the only rational points on the modular curve $X_H(p)$ are the "expected" ones, namely the cusps and the CM points.*

If H is a Borel subgroup, then we denote the associated modular curve of level p by $X_0(p)$ while if H is isomorphic to a normalizer of a split (non-split) Cartan subgroup then we write $X_s^+(p)$ $(X_{ns}^+(p))$. This conjecture has been proved for 3 of the four cases needed; Serre himself dealt with the irregular case [Ser2].

Theorem 3.33 (Serre - Irregular case)**.** *If $p \geq 11$ then $X(\mathbb{Q}_p) = \emptyset$ for $X = X_{A_4}(p)$, $X_{S_4}(p)$, $X_{A_5}(p)$.*

Few years later Mazur settled the Borel case [Maz4].

Theorem 3.34 (Mazur - Borel case)**.** *If $p > 37$ then $X_0(p)(\mathbb{Q}) \subset \{cusps,\ CM\text{-}points\}$.*

In more recent times Bilu, Parent and Rebolledo proved Serre's uniformity conjecture in the case of H a subgroup of the normalizer of a split Cartan Subgroup.

Theorem 3.35 (Bilu, Parent and Rebolledo - Split Cartan case)**.** *If $p > 13$ then $X_s^+(p)(\mathbb{Q}) \subset \{cusps,\ CM\text{-}points\}$.*

Finally, Balakrishnan, Dogra, Müller, Tuitman and Vonk proved that in the split Cartan case one has $p \geq 13$ [Bal+].

The question about the non-split Cartan case remains opens but some cases are known: for $p < 11$ the non-split Cartan curves have genus zero and they have infinitely many rational points. Ligozat [Lig2] extended the result to level $p = 11$ and in 2020 Bilu, Bajolet and Matschke have proved that for $11 < p < 101$ $X_{ns}^{+}(p)$ has no integral points but the CM ones.

Before introducing the main characters of this chapter, Cartan subgroups, we will briefly recall how to construct expected CM-points on modular curves following [Maz3]. Let $E/\mathbb{C}$ be an elliptic curve with complex multiplication by an imaginary quadratic order $\mathrm{End}_{\mathbb{C}}(E) = \mathcal{O}_E$. We denote $\Delta_E = \mathrm{disc}(\mathcal{O}_E)$, $K = \mathrm{Frac}(\mathcal{O}_E)$ and $c = [\mathcal{O}_K : \mathcal{O}_E]$ the conductor of $\mathcal{O}_E$.

Remark. If E is isomorphic to the complex torus $\mathbb{C}/\Lambda$ for the lattice $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$ we have $\mathcal{O}_E = \mathbb{Z}[\tau]$.

For a prime p we know that $\mathcal{O}_E/p\mathcal{O}_E$ is a $\mathbb{F}_p$-algebra of dimension 2 which yields the three possibilities below:

- p ramifies which means that $p\mathcal{O}_E = \mathfrak{p}^2$ for a prime ideal $\mathfrak{p}$. In this case, if ϵ is a generator of $\mathfrak{p}$, then $\ker(\epsilon)$ is in $E[p]$ and it is a subgroup of order p; thus, the pair $(E, \ker(\epsilon))$ determines a point $Q \in X_0(p)$.

- p splits, i.e. $p\mathcal{O}_E = \mathfrak{p}\mathfrak{q}$ for a prime ideal $\mathfrak{p}$ and its conjugate $\mathfrak{q}$. In this case $\ker(\epsilon_1)$ and $\ker(\epsilon_2)$ are independent cyclic subgroups of order p in $E[p]$; thus, $(E, \ker(\epsilon_1), \ker(\epsilon_1))$ yields a point $Q \in X_s^{+}(p)$.

- p remains prime, i.e. $p\mathcal{O}_E = \mathfrak{p}$. In this case $\mathcal{O}_E/p\mathcal{O}_E$ is a field and we obtain a point on $Q \in X_{ns}^{+}(p)$.

Consequently, for each elliptic curve with complex multiplication E, we obtain a non-cuspidal point Q on one of the curves $X_0(N), X_s^{+}(p)$ or $X_{ns}^{+}(p)$ which is defined over a subfield of index two in the ray class field of $\mathcal{O}_E \otimes \mathbb{Q}$, with conductor equal to the conductor of $\mathcal{O}_E$. Thus, to get rational points on a modular curve we need that the associated elliptic curve E is defined over $\mathbb{Q}$. Now, since by the theory of complex multiplication we know that $[\mathbb{Q}(j_E) : \mathbb{Q}] = h_{\mathcal{O}_E}$ the class number of $\mathcal{O}_E$, we only get rational points from elliptic curves with CM by a class number one order. By [Cox, Theo. 7.30] we know that there are only 13 imaginary quadratic orders with class number 1, namely those of discriminant $\Delta = -3, -4, -7, -8, -11, -12, -16, -19, -27, -28, -43, -67, -163$. Hence, for each prime p we get 13 expected rational points on $X_0(N) \coprod X_s^{+}(p) \coprod X_{ns}^{+}(p)$ and, knowing that a prime ramifies in $\mathcal{O}$ if and only if it divides the discriminant, for $p > 163$ all the expected points land outside $X_0(p)$.

3.2.3 Non-split Cartan subgroups and their normalizers

In this section we finally introduce Cartan subgroups and their normalizers. The main references will be [Ser3, A.5] and [Bar2].

Let A be a finite free commutative $\mathbb{Z}/N\mathbb{Z}$ algebra of rank 2 with unit discriminant; if $p \mid N$, A/pA is either equal to $\mathbb{F}_p \times \mathbb{F}_p$ (split case) or $\mathbb{F}_{p^2}$ (non-split case). Let $A^{\times}$ be the multiplicative group of invertible elements of A; then we have a natural action of $A^{\times}$ on A, defined by multiplication, which defines an embedding $A^{\times} \hookrightarrow \mathrm{Aut}_{\mathbb{Z}/N\mathbb{Z}}(A) \simeq \mathrm{Aut}_{\mathbb{Z}/N\mathbb{Z}}\left((\mathbb{Z}/N\mathbb{Z})^2\right) = \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$.

Definition. A Cartan subgroup of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ is a subgroup arising as the image of such a group $A^{\times}$. If A is split (non-split) at p the subgroup will be called split (non-split) Cartan subgroup.

Remark. Since the non-split Cartan subgroups only depend on the basis chosen for A, all the non-split Cartan subgroups of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ are conjugated and so are their normalizers.

Let us take a quadratic order $\mathcal{O}$ with discriminant prime to N and suppose that every prime in the factorization of N is inert in R (i.e., $(\Delta_{\mathcal{O}} \mid p) = -1$); note that there are infinitely many such orders. The algebra $A = \mathcal{O}/N\mathcal{O}$ is a finite commutative $\mathbb{Z}/N\mathbb{Z}$-algebra of rank 2 with unit discriminant by construction.

The order $\mathcal{O}$ admits a $\mathbb{Z}$-basis of the form $\{1, \alpha\}$ where α is the zero of a monic irreducible polynomial $x^2 - ux + v$. The algebra $A = \mathcal{O}/N\mathcal{O}$ admits therefore a $\mathbb{Z}/N\mathbb{Z}$-basis of the form $\{1, \bar{\alpha}\}$ and the group $A^{\times}$ embeds in $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ via its regular representation. Its image will be denoted $C_{ns}(N)$ and it is a non-split Cartan subgroup.

Lemma 3.36. *The orders of the non-split Cartan subgroups of* $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ *are all equal to the order of* $A^\times$*:*

$$\#C_{ns}(N) = N^2 \prod_{p|N} \left(1 - \frac{1}{p^2}\right)$$

Proof. If $N = p^n$ is a power of primes, then

$$A = \frac{(\mathbb{Z}/p^n\mathbb{Z})\,[x]}{(f(x))}$$

with $\deg f = 2$. Then A is a local ring with maximal ideal $\mathfrak{m} = (p)$ and $A^\times \simeq \mathbb{F}_{p^2}^\times \times (1 + \mathfrak{m})$ and it is easy to check that $|\mathfrak{m}| = |1 + \mathfrak{m}| = p^{2(n-1)}$. Thus

$$\#A^\times = p^{2(n-1)} \cdot (p^2 - 1) = p^{2n} \left(1 - \frac{1}{p^2}\right)$$

We can easily extend the result to the general case by following [DD2]. □

Example. Let us study the case $N = p$. We choose a quadratic imaginary order with prime discriminant $q \neq p$. This means that we take $q \equiv 3 \mod 4$ and $\mathcal{O} = \mathbb{Z}[\sqrt{-q}] = \mathbb{Z}[x]/(x^2+q)$. We suppose $(q \mid p) = -1$.

$$\mathcal{O}/p\mathcal{O} \simeq \frac{\mathbb{F}_p[x]}{(x^2 + \bar{q})} \simeq \mathbb{F}_{p^2}$$

Now $\mathbb{F}_{p^2}^\times$ acts on $\mathbb{F}_{p^2}$ via multiplication. Let $\mathbb{F}_{p^2} = \mathbb{F}_p[\xi]$, then

$$\begin{aligned} \iota : \mathbb{F}_{p^2}^\times &\hookrightarrow \mathrm{GL}_2(\mathbb{F}_p) \\ \beta &\longrightarrow M(\beta) \end{aligned}$$

If $\beta = x + y\xi \in \mathbb{F}_{p^2}^\times$ then $\beta \cdot 1 = x + y\xi$ and $\beta \cdot \xi = x\xi - y\bar{q}$. Thus

$$M(\beta) = \begin{pmatrix} x & -\bar{q}y \\ y & x \end{pmatrix}$$

Thus the non-split Cartan subgroups of $\mathrm{GL}_2(\mathbb{F}_p)$ are all conjugate to

$$C_{ns}(p) = \left\{ \begin{pmatrix} a & \epsilon b \\ b & a \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F}_p) \,\middle|\, a^2 - \epsilon b^2 \neq 0 \right\}$$

Note that $C_{ns}(p)$ is the image of a cyclic group of order $p^2 - 1$ and, therefore, it is itself a cyclic group of order $p^2 - 1$.

Remark. If $p \equiv 3 \mod 4$ we can take $\epsilon = -1$ (Gaussian integers).

There is a geometric interpretation of non-split Cartan subgroups: the group $\mathrm{GL}_2(\mathbb{F}_p)$ acts on the right on $\mathbb{P}^1(\mathbb{F}_p)$ by $(x : y) \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (ax + cy : bx + dy)$. Any non-split Cartan subgroup of $\mathrm{GL}_2(\mathbb{F}_p)$ can be defined as the stabilizer of $(1 : \alpha) \in \mathbb{P}^1(\mathbb{F}_{p^2}) \setminus \mathbb{P}^1(\mathbb{F}_p)$ for a choice of α in $\mathbb{F}_{p^2} \setminus \mathbb{F}_p$.

Example. What happens for $N = p^n$ a power of a prime? Once again, we consider a quadratic order $\mathcal{O} = \mathbb{Z}[\sqrt{-q}]$ with q as above.

$$A = \mathcal{O}/p^2\mathcal{O} \simeq (\mathbb{Z}/p^2\mathbb{Z})\,[\alpha]$$

As noticed before, since $(x^2 + q)$ is irreducible modulo p, A is a local ring whose multiplicative group of units is isomorphic to $1 + pA$. Thus,

$$C_{ns}(p^n) = \left\{ \begin{pmatrix} a & \epsilon b \\ b & a \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}/p^n\mathbb{Z}) \,\middle|\, a^2 - \epsilon b^2 \not\equiv 0 \mod p \right\}$$

We may now describe the normalizer of $C_{ns}(N)$ inside $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$. For every prime p that divides N

there exists a unique ring automorphism σ_p of $\mathcal{O}/N\mathcal{O}$ such that

$$\sigma_p(\alpha) \equiv u - \alpha \mod p^{\nu_p(N)} \qquad \text{and} \qquad \sigma_p \equiv \mathrm{Id} \mod N/p^{\nu_p(N)}$$

We identify σ_p with an element S_p of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) \simeq \mathrm{Aut}(A)$. Since the automorphisms σ_p have order 2 and they all commute, the matrices S_p have the same properties.

Proposition 3.37. *The normalizer $C^+_{ns}(N)$ of the non-split Cartan subgroup C_{ns}, is the group*

$$\langle C_{ns}(N), S_p \mid \textit{for all } p \mid N\rangle$$

Proof. This is Proposition 2.3 in [Bar2]. □

We have a short exact sequence

$$1 \longrightarrow C_{ns}(N) \longrightarrow C^+_{ns}(N) \longrightarrow \langle S_p \,|\, \text{for all } p \mid N\rangle \longrightarrow 1$$

This gives

$$\#C^+_{ns}(N) = \#C_{ns}(N) \cdot \#\langle S_p\ ,\quad \forall p \mid N\rangle = N^2 2^{\omega(N)} \prod_{p|N}\left(1 - \frac{1}{p^2}\right)$$

where $\omega(N)$ is the prime function counting the number of prime divisors of N.

Example. In the prime case we have

$$C^+_{ns}(p) = \left\{G \in \mathrm{GL}_2(\mathbb{F}_p) \,\middle|\, GC_{ns}(p)G^{-1} = C_{ns}(p)\right\}$$

To find a description of the full normalizer of $C_{ns}(p)$ we note that $A = \mathcal{O}/p\mathcal{O} \simeq \mathbb{F}_{p^2} = \mathbb{F}_p[\xi]$ and we have two field automorphisms, namely the identity and σ which sends $\xi \to -\xi$. This is represented by the matrix $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ which happens to be in $C^+_{ns}(p)$ since

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}a & \epsilon b\\ b & a\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \begin{pmatrix}a & -\epsilon b\\ -b & a\end{pmatrix} \in C_{ns}(p)$$

Thus

$$C^+_{ns}(p) = \langle C_{ns}(p), \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\rangle = \left\{\begin{pmatrix}a & \epsilon b\\ b & a\end{pmatrix}, \begin{pmatrix}a & -\epsilon b\\ b & -a\end{pmatrix} \in \mathrm{GL}_2(\mathbb{F}_p) \,\middle|\, (a,b) \in \mathbb{F}_{p^2} \setminus \{(0,0)\}\right\}$$

We recover $\#C^+_{ns}(p) = 2\#C_{ns}(p) = 2(p^2-1)$.

3.2.4 The modular curves associated with to non-split Cartan subgroups

Let $X(N)$ be the modular curve of full level N. Its non-cuspidal points parametrize pairs (E,ϕ) where E is an elliptic curve over $\mathbb{C}$ and $\phi : E[N] \to \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ is an isomorphism of groups. Two such pairs (E,ϕ) and (F,φ) are isomorphic if there exists an isomorphism $\psi : E \to F$ such that the following diagram commutes:

$$\begin{tikzcd}
E[N] \arrow[r, "\psi"] \arrow[d, "\phi"'] & F[N] \arrow[d, "\varphi"] \\
\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} \arrow[r, equal] & \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}
\end{tikzcd}$$

As pointed out in Section 3.2.1, there is an action of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ on the curve $X(N)$. Now $C_{ns}(N)$ and $C^+_{ns}(N)$ are subgroups of $GL_2(\mathbb{Z}/N\mathbb{Z})$ whose image under the determinant map is the whole group $(\mathbb{Z}/N\mathbb{Z})^\times$ [Boo, §8]. Therefore, they define the modular curves

$$X_{ns}(N) = X(N)/C_{ns}(N) \qquad X^+_{ns}(N) = X(N)/C^+_{ns}(N)$$

The associated open non-cuspidal curves are indicated $Y_{ns}(N)$ and $Y^+_{ns}(N)$ as usual and their points correspond to isomorphism classes of pairs $(E, [\phi]_H)$ of an elliptic curve and an H-equivalence class of isomorphisms $\phi : E[N] \to \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ for $H = C_{ns}(N)$ or $C^+_{ns}(N)$.

There are natural covering morphisms

$$\begin{tikzcd} X_{ns}(N) \arrow[r, "\phi_1"] & X^+_{ns}(N) \arrow[r, "\phi_2"] & X(1) \end{tikzcd}$$

having degree $\deg \phi_1 = 2^{\omega(N)}$ and $\deg \phi_2 = \frac{N\phi(N)}{2^{\omega(N)}}$.

Remark. If H and H' are conjugate subgroups of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, then X_H is naturally isomorphic to $X_{H'}$ over $\mathbb{Q}$. Thus the modular curves associated to any non-split Cartan subgroup (normalizer of non-split Cartan subgroup) of level N are isomorphic to $X_{ns}(N)$ (respectively $X^+_{ns}(N)$) over $\mathbb{Q}$.

3.2.5 Coset representatives in $SL_2(\mathbb{Z})$

In order to better understand the modular curves $X_{ns}(N)$ and $X^+_{ns}(N)$ we describe the coset representatives of their associated congruence subgroups. We denote

$$C'_{ns}(N) = C_{ns}(N) \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) \qquad C'^+_{ns}(N) = C^+_{ns}(N) \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$$

and we write $\Gamma_{ns}(N)$ and $\Gamma'_{ns}(N)$ for their lifts to $\mathrm{SL}_2(\mathbb{Z})$.

As before, $\mathcal{O} = \mathbb{Z}[\xi]$ will be a quadratic order with basis $\{1, \xi\}$ such that $x^2 - ux + v$ is the minimal polynomial of ξ. We define $\mathcal{L}_c = \{M(y) \mid y \in (\mathbb{Z}/N\mathbb{Z})\,[\xi]^\times \text{ and } Nr(y) = c\}$ where $M(y)$ indicates the matrix associated with the multiplication by y. There is a natural identification

$$C'_{ns}(N) \longleftrightarrow \mathcal{L}_1 = \{M(y) \,|\, y \in (\mathbb{Z}/N\mathbb{Z})\,[\xi]^\times \text{ and } Nr(y) = 1\} \longleftrightarrow \mathcal{Y}_1 = \{y \in (\mathbb{Z}/N\mathbb{Z})\,[\xi]^\times \mid Nr(y) = 1\}$$

We define

$$\mathcal{Y}(N) = \bigcup_{a \in (\mathbb{Z}/N\mathbb{Z})^\times} \mathcal{Y}_a = \{y \in (\mathbb{Z}/N\mathbb{Z})\,[\xi]^\times \;\mid\; Nr(y) = a,\; a \in (\mathbb{Z}/N\mathbb{Z})^\times\}$$

Proposition 3.38 ([Bar2, Prop. 6.2])**.** *The matrices representing the linear maps that transform the basis* $\{1, \xi\}$ *as*

$$1 \to y^{-1} \qquad \xi \to \bar{y}(\xi + x)$$

where $x \in \mathbb{Z}/N\mathbb{Z}$ *and* $y \in \mathcal{Y}(N)$*, are coset representatives of* $C'_{ns}(N)$ *in* $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$*.*

Example ($N = p$)**.** We have

$$C_{ns}(p) = \left\{ \begin{pmatrix} a & \epsilon b \\ b & a \end{pmatrix} \,\middle|\, (a, b) \in \mathbb{F}_p^2 \setminus \{(0, 0)\} \right\}$$

$$C'_{ns}(p) = \left\{ \begin{pmatrix} a & \epsilon b \\ b & a \end{pmatrix} \,\middle|\, (a, b) \in \mathbb{F}_p^2 \setminus \{(0, 0)\},\; a - \epsilon b = 1 \right\} \longleftrightarrow \left\{ M(y) \,\middle|\, y \in (\mathbb{Z}/N\mathbb{Z})\,[\xi]^\times \text{ and } Nr(y) = 1 \right\}$$

Further, $\mathcal{Y}(p) = \{y_\alpha \mid a \in \mathbb{F}_p^\times\}$ where y_α is an element of $\mathbb{F}_p[\xi]^\times$ of norm α. As we did before, we consider $\xi = \sqrt{-q}$ for $q \neq p$ inert in $\mathbb{F}_p$.

$$Nr(y_a) = y_\alpha \bar{y_\alpha} = (a_\alpha + \xi b_\alpha)(a_\alpha - \xi b_\alpha) = a_\alpha^2 + q b_\alpha$$

We need to find matrices M such that

$$M \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = y_\alpha^{-1} = \frac{\bar{y_\alpha}}{y_\alpha \bar{y_\alpha}} = \begin{pmatrix} \frac{a_\alpha}{a_\alpha^2 + q b_\alpha^2} \\ -\frac{b_\alpha}{a_\alpha^2 + q b_\alpha^2} \end{pmatrix}$$

and

$$M \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \bar{y_\alpha}(\xi + x) = a_\alpha x + b_\alpha q + (a_\alpha - b_\alpha x)\xi = \begin{pmatrix} a_\alpha x - b_\alpha q \\ -a_\alpha + b_\alpha x \end{pmatrix}$$

Thus a system of coset representatives consists of the matrices

$$\begin{pmatrix} \frac{a_\alpha}{a_\alpha^2+qb_\alpha^2} & a_\alpha x + b_\alpha q \\ -\frac{b_\alpha}{a_\alpha^2+qb_\alpha^2} & a_\alpha - b_\alpha x \end{pmatrix} = \begin{pmatrix} \alpha^{-1}a_\alpha & a_\alpha x + b_\alpha q \\ -\alpha^{-1}b_\alpha & a_\alpha - b_\alpha x \end{pmatrix}$$

being $y_\alpha = a_\alpha + \xi b_\alpha$ an element of norm $a_\alpha^2 + qb_\alpha^2 = \alpha$. In other words, for every $\alpha \in \mathbb{F}_p^\times$ we choose an element $y_\alpha = (a_\alpha, b_\alpha) \in \mathcal{Y}_\alpha$ and the set

$$\mathcal{S} = \left\{ \begin{pmatrix} \alpha^{-1}a_\alpha & a_\alpha x + b_\alpha q \\ -\alpha^{-1}b_\alpha & a_\alpha - b_\alpha x \end{pmatrix} \,\middle|\, \alpha \in \mathbb{F}_p^\times,\ x \in \mathbb{F}_p \right\}$$

is a system of representatives for $C'_{ns}(p)$ in $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$.

We can repeat the same procedure for the normalizer of a non-split Cartan subgroup. We identify $C'^{+}_{ns}(p^r)$ with the group

$$C'_{ns}(p^r) \longleftrightarrow \mathcal{L}_1 \cup \mathcal{L}^\sigma = \{M(y) \,|\, y \in (\mathbb{Z}/p^r\mathbb{Z})[\xi]^\times \text{ and } Nr(y) = 1\} \cup \{M(y) \circ S_p \,|\, y \in (\mathbb{Z}/p^r\mathbb{Z})[\xi]^\times \text{ and } Nr(y) = -1\}$$

where S_p is the matrix associated to σ_p above. Once again, for every $\alpha \in (\mathbb{Z}/p^r\mathbb{Z})^\times / \pm 1$ we choose $y_\alpha \in (\mathbb{Z}/p^r\mathbb{Z})[\xi]^\times$ with norm $Nr(y_\alpha) = \alpha$.. We define the set

$$\mathcal{Y}_\pm(p^r) = \{y_\alpha \,|\, \alpha \in (\mathbb{Z}/p^r\mathbb{Z})^\times / \pm 1\}$$

Proposition 3.39 ([Bar2, Prop 6.3])**.** *The matrices representing the linear maps that transform the basis* $\{1, \xi\}$ *as*

$$1 \to y^{-1} \qquad \xi \to \bar{y}(\xi + x)$$

where $x \in \mathbb{Z}/p^r\mathbb{Z}$ *and* $y \in \mathcal{Y}_\pm(p^r)$*, are coset representatives of* $C'^{+}_{ns}(p^r)$ *in* $\mathrm{SL}_2(\mathbb{Z}/p^r\mathbb{Z})$*.*

Example. We will work out the structures just described in the case $N = 5$. $C_{ns}(5)$ has cardinality $5 * (5-1) = 24$ by Lemma 3.36 and it has elements of the form

$$\begin{pmatrix} a & -3b \\ b & a \end{pmatrix} \qquad (a, b) \in \mathbb{F}_p^2 \setminus \{(0,0)\}$$

Note that in this case we chose $\mathcal{O}$ to be the ring $\mathbb{Z}[\sqrt{-3}]$ inside Eisenstein integers. In the same way $C^{+}_{ns}(N)$ has order 48 by the short exact sequence after Proposition 3.37. To describe it we have to add to $C_{ns}(5)$ all the matrices $C_{ns}(5) \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ where we multiply the second column by -1.

We now intersect these two groups with $\mathrm{SL}_2(\mathbb{F}_5)$. Note that the matrices of $C_{ns}(N)$ ($C^{+}_{ns}(N)$) are equally divided in sets of matrices with same discriminant, i.e., for each $d \in (\mathbb{Z}/N\mathbb{Z})^\times$ the sets of matrices of $C_{ns}(N)$ ($C^{+}_{ns}(N)$) with discriminant d have the same cardinality [Baj, Prop. 2.11]. This means that $\#C'_{ns}(5) = \#C_{ns}(5)/\#\mathbb{F}_5^\times = 24/4 = 6$ and $\#C'^{+}_{ns}(5) = \#C^{+}_{ns}(5)/\#\mathbb{F}_5^\times = 48/4 = 12$.

$$C'_{ns}(5) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 2 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 3 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \right\}$$

$$C'^{+}_{ns}(5) = \left\{ \begin{matrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 2 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 3 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \\ \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -3 \\ 4 & -1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}, \begin{pmatrix} 4 & -2 \\ 1 & -4 \end{pmatrix}, \begin{pmatrix} 4 & -3 \\ 4 & -4 \end{pmatrix} \end{matrix} \right\}$$

In particular, this means that $C'_{ns}(5)$ has index $\#\mathrm{SL}_2(\mathbb{F}_5)/\#C'_{ns}(5) = 120/6 = 20$ and $C'^{+}_{ns}(5)$ has index $120/12 = 10$.

We construct now a system of representatives for the cosets of $C'_{ns}(5)$ in $\mathrm{SL}_2(\mathbb{F}_5)$. For each $\alpha \in \mathbb{F}_5^\times$ we choose $(x, y) \in \mathbb{F}_5^2 \setminus \{(0,0)\}$ such that $x^2 + 3y^2 = \alpha$ and we construct all the matrices of the form

$$\begin{pmatrix} \alpha^{-1}x & 3y + cx \\ -\alpha^{-1}y & x - cy \end{pmatrix} \qquad c \in \mathbb{F}_p$$

Here the complete list of coset representatives of $C'_{ns}(5)$ in $\mathrm{SL}_2(\mathbb{F}_5)$

	(x,y)	α^{-1} $-\alpha^{-1}$	$c=0$	$c=1$	$c=2$	$c=3$	$c=4$
$\alpha=1$	$(1,0)$	1 4	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}1&1\\0&1\end{pmatrix}$	$\begin{pmatrix}1&2\\0&1\end{pmatrix}$	$\begin{pmatrix}1&3\\0&1\end{pmatrix}$	$\begin{pmatrix}1&4\\0&1\end{pmatrix}$
$\alpha=2$	$(2,1)$	3 2	$\begin{pmatrix}1&3\\2&2\end{pmatrix}$	$\begin{pmatrix}1&0\\2&1\end{pmatrix}$	$\begin{pmatrix}1&2\\2&0\end{pmatrix}$	$\begin{pmatrix}1&4\\2&4\end{pmatrix}$	$\begin{pmatrix}1&1\\2&3\end{pmatrix}$
$\alpha=3$	$(0,1)$	2 3	$\begin{pmatrix}0&3\\3&0\end{pmatrix}$	$\begin{pmatrix}0&3\\3&4\end{pmatrix}$	$\begin{pmatrix}0&3\\3&3\end{pmatrix}$	$\begin{pmatrix}0&3\\3&2\end{pmatrix}$	$\begin{pmatrix}0&3\\3&1\end{pmatrix}$
$\alpha=4$	$(2,0)$	4 1	$\begin{pmatrix}3&0\\0&2\end{pmatrix}$	$\begin{pmatrix}3&2\\0&2\end{pmatrix}$	$\begin{pmatrix}3&4\\0&2\end{pmatrix}$	$\begin{pmatrix}3&1\\0&2\end{pmatrix}$	$\begin{pmatrix}3&3\\0&2\end{pmatrix}$

In order to find coset representatives for ${C'}^{+}_{ns}(5)$ we have to consider α modulo ± 1 which means that we have to take into account only the first two rows of the table above. For example, we see that $\begin{pmatrix}3&0\\0&2\end{pmatrix}$ is already in ${C'}^{+}_{ns}(5)$ while $\begin{pmatrix}3&2\\0&2\end{pmatrix}$ is in the same coset of $\begin{pmatrix}1&4\\0&1\end{pmatrix}$.

The last step would be the lifting of $C'_{ns}(N)$ and ${C'}^{+}_{ns}(p^r)$ to $\mathrm{SL}_2(\mathbb{Z})$ (pullback via the reduction map). Luckily, we have bijections

$$\left\{\begin{matrix}\text{Coset representatives}\\ \text{of } C'_{ns}(N) \text{ in } \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})\end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix}\text{Coset representatives}\\ \text{of } \Gamma_{ns}(N) \text{ in } \mathrm{SL}_2(\mathbb{Z})\end{matrix}\right\}$$

$$\left\{\begin{matrix}\text{Coset representatives}\\ \text{of } {C'}^{+}_{ns}(p^r) \text{ in } \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})\end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix}\text{Coset representatives}\\ \text{of } \Gamma^{+}_{ns}(N) \text{ in } \mathrm{SL}_2(\mathbb{Z})\end{matrix}\right\}$$

This implies that we only have to lift one by one the representatives of the cosets for $C'_{ns}(N)$ and ${C'}^{+}_{ns}(p^r)$. To do this we can use an algorithm [Baj, Alg. 1]:

Algorithm 5. Lift of a matrix in $\mathrm{SL}_2(\mathbb{F}_p)$ to $\mathrm{SL}_2(\mathbb{Z})$

Input: A matrix $\bar{M} \in \mathrm{SL}_2(\mathbb{F}_p)$.
Output: A matrix $N \in \mathrm{SL}_2(\mathbb{Z})$ such that $N \equiv N \bmod p$

1. We take M any lift (coordinate-wise) of $\bar{M}$.

2. By linear operations on lines and columns we find $U, V \in \mathrm{SL}_2(\mathbb{Z})$ such that $UMV = \begin{pmatrix}a&0\\0&b\end{pmatrix} \in \mathrm{SL}_2(\mathbb{F}_p)$.

Observe that this is always possible.

3. Let

$$W = \begin{pmatrix}b&1\\b-1&1\end{pmatrix} \qquad X = \begin{pmatrix}1&-b\\0&1\end{pmatrix}$$

4. Return

$$N = U^{-1}W^{-1}\begin{pmatrix}1&0\\1-a&1\end{pmatrix}X^{-1}V^{-1}$$

Example. In the example before we observe that the whole first row lifts naturally to $\mathrm{SL}_2(\mathbb{Z})$. The first interesting element is $M = \begin{pmatrix}1&3\\2&2\end{pmatrix}$. We find

$$\begin{pmatrix}1&0\\-2&1\end{pmatrix}\begin{pmatrix}1&3\\2&2\end{pmatrix}\begin{pmatrix}1&-3\\0&1\end{pmatrix} = \begin{pmatrix}1&0\\0&-4\end{pmatrix}$$

We set $W = \begin{pmatrix} -4 & 1 \\ -5 & 1 \end{pmatrix}$ and $X = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$. Finally

$$N = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 5 & -4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 7 & -13 \end{pmatrix}$$

and $N \in \mathrm{SL}_2(\mathbb{Z})$ reduces to M modulo p. We find a set of coset representatives for $\Gamma_{ns}(5)$:

$$\mathrm{Cosets}\left(\Gamma_{ns}(5)\right) = \left\{ \begin{array}{c} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & -2 \\ 7 & -13 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -3 \\ 7 & -20 \end{pmatrix}, \begin{pmatrix} 6 & 29 \\ 7 & 34 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \\ \begin{pmatrix} 10 & 33 \\ 3 & 10 \end{pmatrix}, \begin{pmatrix} -90 & 13 \\ -97 & 14 \end{pmatrix}, \begin{pmatrix} 10 & 33 \\ 13 & 43 \end{pmatrix}, \begin{pmatrix} 90 & -7 \\ 103 & -8 \end{pmatrix}, \begin{pmatrix} 90 & 23 \\ 43 & 11 \end{pmatrix} \\ \begin{pmatrix} 8 & -5 \\ 5 & -3 \end{pmatrix}, \begin{pmatrix} 13 & 47 \\ -5 & -18 \end{pmatrix}, \begin{pmatrix} -7 & -11 \\ -5 & -8 \end{pmatrix}, \begin{pmatrix} -7 & -4 \\ -5 & -3 \end{pmatrix}, \begin{pmatrix} 3 & -8 \\ -5 & -13 \end{pmatrix} \end{array} \right\}$$

and a set of coset representatives for $\Gamma'_{ns}(5)$:

$$\mathrm{Cosets}\left(\Gamma^+_{ns}(5)\right) = \left\{ \begin{array}{c} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & -2 \\ 7 & -13 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -3 \\ 7 & -20 \end{pmatrix}, \begin{pmatrix} 6 & 29 \\ 7 & 34 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \end{array} \right\}$$

3.2.6 Signature of $X_{ns}(N)$ and $X^+_{ns}(N)$

We start with an example and we will eventually describe a more general way of computing the genus of the modular curves associated to a (normalizer of a) non split Cartan subgroup.

Example. Let us try to work out the number of cusps and elliptic points of $X^+_{ns}(5)$. We recall that, noted $\epsilon_h = \#\{\text{elliptic points of order } h\}$ and $\nu_h = \#\phi^{-1}(P_h)$ for $h = 2, 3$, $P_2 = i \in X(1)$, $P_3 = \rho \in X(1)$ and $\phi : X^+_{ns}(5) \to X(1)$, then Section 1.3.5 gives

$$d - \epsilon_h = h(\nu_h - \epsilon_h)$$

and $\nu_\infty = \epsilon_\infty$. Let $\Gamma = \Gamma^+_{ns}(5)$. We also know that $\nu_2 = \Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/ < S >$, $\nu_3 = \Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/ < ST >$ and $\nu_\infty = \Gamma\backslash\mathrm{SL}_2(\mathbb{Z})/ < T >$ are all related to the number of orbits of the action of some matrices on the set of cosets of Γ, see proposition 1.33. From the example above we have a description of these cosets:

$$\mathrm{Cosets}\left(\Gamma^+_{ns}(5)\right) = \left\{ \begin{array}{c} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & -2 \\ 7 & -13 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -3 \\ 7 & -20 \end{pmatrix}, \begin{pmatrix} 6 & 29 \\ 7 & 34 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \end{array} \right\}$$

It is straightforward to observe that the first line is in the same orbit under the action of $\langle T \rangle$ and an easy computation shows that the second line form a second distinct orbit. Thus $\epsilon_\infty = \nu_\infty = 2$ and $X^+_{ns}(5)$ has 2 cusps.

A case by case computation shows that $\mathrm{Cosets}\left(\Gamma^+_{ns}(5)\right)$ consists of 6 S-orbits, 4 of which of size 2

$$\mathrm{Cosets}\left(\Gamma^+_{ns}(5)\right)/\langle S \rangle = \left\{ \begin{array}{c} \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ 7 & -13 \end{pmatrix}\right], \left[\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 6 & 29 \\ 7 & 34 \end{pmatrix}\right], \left[\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\right] \\ \left[\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}\right], \left[\begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}\right], \left[\begin{pmatrix} 1 & -3 \\ 7 & -20 \end{pmatrix}\right] \end{array} \right\}$$

This implies that $\nu_2 = 6$ and, by consequence, $X^+_{ns}(5)$ has $\epsilon_2 = 2$ elliptic points of order 2.

Finally, we study the orbits under the action of ST. We observe that the only stabilized element is

$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$, i.e., there exists a unique element $B \in \text{Cosets}\left(\Gamma_{ns}^{+}(5)\right)$ such that

$$BSTB^{-1} \in \Gamma_{ns}^{+}(5)$$

This means that all the other 9 elements are in orbits of size 3. This implies that $\nu_3 = 4$ and, therefore, $X_{ns}^{+}(5)$ has $\epsilon_3 = 1$ elliptic point of order 3.

Remark. Note that, since $\nu_h = (\epsilon_h(h-1)+d)/h$, then ϵ_h is the number of orbits of size 1 under the action of M_h where $M_2 = S$ and $M_3 = ST$.

In conclusion, $X_{ns}^{+}(5)$ has genus (corollary 1.32)

$$g_{ns}^{+}(5) = g(X_{ns}^{+}(5)) = \frac{1}{2}\left(2 + (10-2) + (10-6) + (10-4) - 20\right) = 0$$

We will now generalize all the ideas used in the example above. The main reference will be [Bar2]. We have already observed that $\deg(X(\Gamma_{ns}(N))) = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_{ns}(N)] = N\phi(N)$ and $\deg(X(\Gamma_{ns}^{+}(N))) = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_{ns}^{+}(N)] = N\phi(N)/2^{\omega(N)}$ since $\pm I \in \Gamma_{ns}(N), \Gamma_{ns}^{+}(N)$.

Proposition 3.40. *Let $N = p^r$ be a prime power. We have*

$$\epsilon_\infty\left(X_{ns}(p^r)\right) = \phi(p^r) \qquad \epsilon_\infty\left(X_{ns}^{+}(p^r)\right) = \frac{\phi(p^r)}{2}$$

Proof. As we have seen in the example, we have to study the number of orbits in $\mathrm{SL}_2(\mathbb{Z})/\Gamma_{ns}(p^r)$ under the action of $\langle T \rangle$. We do this with the help of the following.

Lemma 3.41. *Every matrix $T^a = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ with $a > 0$ such that $T^a \in G^{-1}\Gamma_{ns}(p^r)G$, for some coset representative G of $\Gamma_{ns}(p^r)$ in $\mathrm{SL}_2(\mathbb{Z})$, satisfies $a \equiv 0 \mod p^r$.*

Proof. We will use the definition of G as a linear map transforming the basis $\{1, \xi\}$ as

$$1 \to -y^{-1} \;\; y \in \mathcal{Y}(p^r) \qquad \xi \to \bar{y}(x+\xi) \;\; x \in (\mathbb{Z}/p^r\mathbb{Z})$$

Further, we can see matrices of $\Gamma_{ns}(p^r)$ modulo p^r as multiplication by k maps on $(\mathbb{Z}/p^r\mathbb{Z})[\xi]$ for $k \in (\mathbb{Z}/p^r\mathbb{Z})[\xi]^\times$ with $Nr(k) = 1$. We get a system

$$\begin{cases} -y^{-1} + a\bar{y}(x+\xi) \equiv -y^{-1}k \mod p^r \\ \bar{y}(x+\xi) \equiv \bar{y}(x+\xi)k \mod p^r \end{cases}$$

from which $k \equiv 1$ and $a \equiv 0 \bmod p^r$. □

This lemma implies that every cusp of $X_{ns}(p^r)$ has ramification degree p^r which means that $\epsilon_\infty = p^r\phi(p^r)/p^r = \phi(p^r)$. In the same way, the elliptic curve $X_{ns}^{+}(p^r)$ has $\phi(p^r)/2$ cusps. □

Proposition 3.42. *Let $N = p^r$ be a prime power. We have*

$$\epsilon_3\left(X_{ns}(p^r)\right) = \begin{cases} 2 & \textit{If } p \equiv 2 \mod 3 \\ 0 & \textit{Otherwise} \end{cases} \qquad \epsilon_3\left(X_{ns}^{+}(p^r)\right) = \begin{cases} 1 & \textit{If } p \equiv 2 \mod 3 \\ 0 & \textit{Otherwise} \end{cases}$$

Proof. As we have seen in the example, the number of elliptic points of order 3 corresponds to the number of orbits of size 1 in $\mathrm{SL}_2(\mathbb{Z})/\Gamma_{ns}(p^r)$ under the action of ST. Therefore, we have to count the number of coset representatives σ of $\Gamma_{ns}(p^r)$ in $\mathrm{SL}_2(\mathbb{Z})$ such that

$$\sigma \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} = \gamma\sigma \qquad \text{some } \gamma \in \Gamma_{ns}(p^r)$$

Once again, γ modulo p^r is the multiplication by k map on $(\mathbb{Z}/p^r\mathbb{Z})[\xi]$ for some $k \in (\mathbb{Z}/p^r\mathbb{Z})[\xi]^\times$ with norm 1. This means that we have to count the number of solutions $(x, y) \in \mathbb{Z}/p^r\mathbb{Z} \times \mathcal{Y}(p^r)$ of the system

$$\begin{cases} \bar{y}(x+\xi) \equiv -y^{-1}k \mod p^r \\ y^{-1} - \bar{y}(x+\xi) \equiv \bar{y}(x+\xi)k \mod p^r \end{cases}$$

Proposition 7.7 of [Bar2] tells us that this system has exactly two solutions if $p \equiv 2 \mod 3$ and 0 otherwise. Since the map $X_{ns}(p^r) \to X_{ns}^+(p^r)$ has degree 2 and it is unramified over elliptic points of order 3 we get $\epsilon_3(X_{ns}^+(p^r)) = \epsilon_3(X_{ns}(p^r))/2$. □

Exactly in the same way, by looking at the action of S, we obtain the number of elliptic points of order 2.

Proposition 3.43 ([Bar2, Prop. 7.10])**.** *Let $N = p^r$ be a prime power. We have*

$$\epsilon_2(X_{ns}(p^r)) = \begin{cases} 2 & \textit{If } p \equiv 3 \mod 4 \\ 0 & \textit{Otherwise} \end{cases} \qquad \epsilon_2(X_{ns}^+(p^r)) = \begin{cases} \frac{1}{2}p^r\left(1-\frac{1}{p}\right) & \textit{If } p \equiv 1 \mod 4 \\ 1+\frac{1}{2}p^r\left(1+\frac{1}{p}\right) & \textit{If } p \equiv 3 \mod 4 \\ 2^{r-1} & \textit{If } p = 2 \end{cases}$$

We define

$$\beta_2(N) = \begin{cases} 1 & \text{If } p \equiv 3 \text{ (mod 4) for all primes } p \mid N \\ 0 & \text{Otherwise} \end{cases}$$

$$\beta_3(N) = \begin{cases} 1 & \text{If } p \equiv 2 \text{ (mod 3) for all primes } p \mid N \\ 0 & \text{Otherwise} \end{cases}$$

Proposition 3.44. *The genus of the modular curves associated with a non-split Cartan subgroup of level N and its normalizer is given by*

$$g_{ns}(N) = g(X_{ns}(N)) = 1 + \frac{(N-6)\phi(N)}{12} - 2^{\omega(N)}\frac{\beta_2(N)}{4} - 2^{\omega(N)}\frac{\beta_3(N)}{3}$$

$$g_{ns}^+(N) = g(X_{ns}^+(N)) = 1 + \frac{(N-6)\phi(N)}{12 \cdot 2^{\omega(N)}} - \frac{\beta_3(N)}{3} - \frac{N}{4 \cdot 2^{\omega(N)}} \prod_{p \mid N} \begin{cases} \left(1-\frac{1}{p}\right) & \textit{If } p \equiv 1 \textit{ (mod 4)} \\ 1 & \textit{If } p = 2 \\ \left(1+\frac{1}{p}+\frac{2}{p^r}\right) & \textit{If } p \equiv 3 \textit{ (mod 4)} \end{cases}$$

By choosing an *optimal* set of coset representatives for $\Gamma_{ns}(N)$ or $\Gamma_{ns}^+(N)$ we can draw a connected fundamental domain for $X_{ns}(N)$ and $X_{ns}^+(N)$.

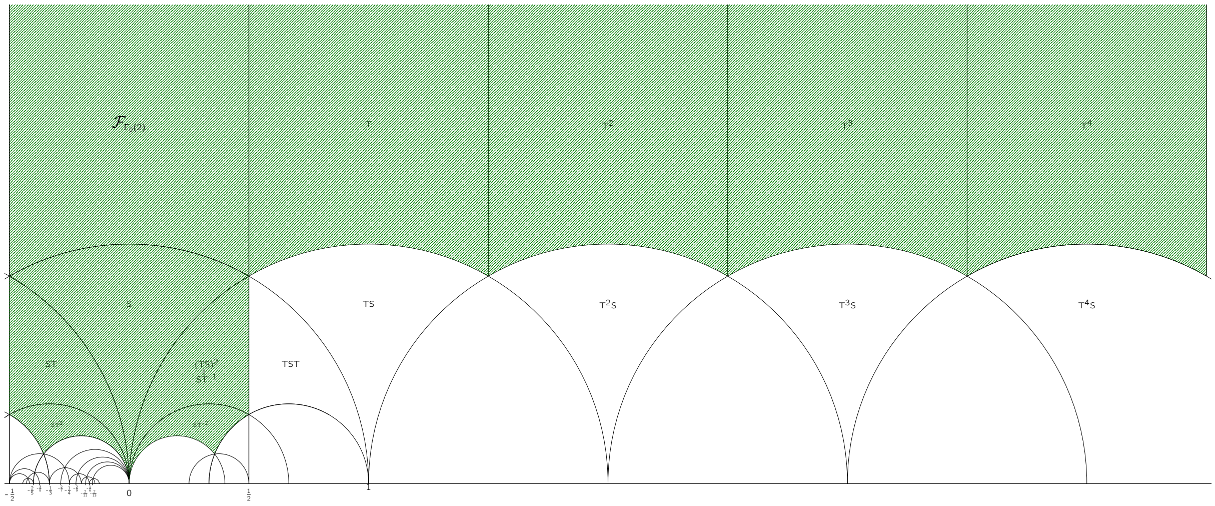


Figure 3.4 – The fundamental region for $\Gamma_{ns}^+(5)$.

3.2.7 Models for $X_{ns}(N)$ and $X_{ns}^{+}(N)$

The goal of this section is to describe the state of the art methods for constructing explicit models for the modular curves associated to non-split Cartan subgroups and their normalizers.

In the previous chapters we have found that there exist quite efficient methods to construct modular functions of the groups $\Gamma(N), \Gamma_1(N)$ and $\Gamma_0(N)$. For other congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ this is not the case. While for split Cartan curves we still have algorithms that generates modular functions, for the non-split case still little is known. This is probably due to the fact that, despite the curves $X_{ns}(N)$ and $X_{ns}^{+}(N)$ have the nice feature of being defined over $\mathbb{Q}$, their genera grows very rapidly and therefore computations are harder to exploit.

We know that $g_{ns}(p) = 0$ for $p = 2, 3, 5$ while $g_{ns}(7) = 1$. An equation for $X_{ns}(7)$ is given in [MS]. In [Dos+], Dose, Fernández and González give an equation for the genus 4 modular curve $X_{ns}(11)$ while the case $p = 13$ is studied in [DMS]. Concerning the curves $X_{ns}^{+}(p)$, we know that $g_{ns}^{+}(p) = 0$ for $p \leq 7$. Explicit models for the modular curve $X_{ns}^{+}(11)$ are given by Ligozat [Lig2]. More recently, Baran [Bar3] has computed an equation for the genus 3 curve associated to $C_{ns}^{+}(13)$. Finally, Mercuri and Schoof [MS] have presented numerical computations for $X_{ns}^{+}(17), X_{ns}^{+}(19)$ and $X_{ns}^{+}(23)$. Although the problem is of great interest, it is not the goal of this thesis. We will only need parameters for small non-split Cartan curves (for $N \leq 5$). Hence, we will devote the rest of the section to present some general results about modular functions.

We recall that lemma 3.31 tells us that $E_{g,h}^{12p}$ is a modular function for $\Gamma(p)$. If H is a subgroup of $\mathbb{F}_p^{\times}$ of index d containing -1, then

$$G_H = \{M \in C_{ns}^{+}(p) \mid \det(M) \in H\}$$

acts on $M_p = (\mathbb{Z}/p\mathbb{Z})^2 \setminus \{(0,0)\}$ and its action has exactly d orbits. We have a description of these orbits in terms of coset representatives for H in $\mathbb{F}_p^{\times}$: for any $a \in \mathbb{F}_p^{\times}/H$ we get an orbit

$$\mathcal{O}_a = \{(x,y) \in M_p \mid x^2 - \xi^2 y^2 \in aH\}$$

Therefore, if $\mathcal{O}$ is any of these orbits we define

$$E_{\mathcal{O}} = \prod_{(g,h)\in\mathcal{O}} E_{g,h}^{12p}$$

By construction, for every $\sigma \in \mathcal{Gal}\left(\mathbb{Q}(\zeta_p)(X(p))/\mathbb{Q}(\zeta_p)(X(p))^{G_H}\right)$ we have

$$E_{\mathcal{O}}^{\sigma} = \prod_{(g,h)\in\mathcal{O}} \left(E_{g,h}^{12p}\right)^{\sigma} = \prod_{(g,h)\in\mathcal{O}^{\sigma}} E_{g,h}^{12p} = E_{\mathcal{O}}$$

so that $E_{\mathcal{O}}$ is an element of

$$\mathbb{Q}(\zeta_p)(X(p))^{G_H} = \mathbb{Q}(\zeta_p)^{H}(X(p))$$

We will note $K = \mathbb{Q}(\zeta_p)^{H}$. This is the approach in [Baj]. Once again, we are interested in very specific curves. Therefore, following [Boo] and [Che] we will give another interpretation of this problem.

Let X be a projective non-singular algebraic curve defined over $\mathbb{Q}$. Suppose that X has genus 0 and at least one rational point (this is the case for $X_{ns}^{+}(N)$ for $N = 3, 4, 5$). Then, there exists an isomorphism defined over $\mathbb{Q}$

$$t : X \longrightarrow \mathbb{P}^1(\mathbb{Q})$$

which is unique up to automorphism. This will be our parameter (and it is sometimes called a uniformizer for X). Suppose that we have a morphism $\pi : X \to X(1)$. This induces an embedding of the function fields $\pi^* : \mathbb{Q}(X(1)) = \mathbb{Q}(j) \hookrightarrow \mathbb{Q}(t) = \mathbb{Q}(X)$. Thus, $\pi^*(j)$ can be expressed as

$$\pi^*(j) = \lambda \frac{P(t)}{Q(t)}$$

where $\lambda \in \mathbb{Q}$ and P, Q are rational polynomials. This relation is called a covering relation and characterize the uniformizer t relatively to j. Now, an explicit parametrization of X is a choice of uniformizers t and j (for X and $X(1)$ respectively) and an explicit covering relation for t.

Lemma 3.45 ([Che]). *Since $j(\rho) = 0$ and $j(\infty) = \infty$, we have*

$$P(T) = \prod_{z \in \pi^{-1}(\rho)} (t - t(z))^{e(z)} \qquad Q(T) = \prod_{z \in \pi^{-1}(\infty)} (t - t(z))^{e(z)}$$

where $e(z)$ is the ramification index of z over ρ or ∞. Further, if $z_0 = t^{-1}(\infty)$,

$$\lambda = j(z_0)\frac{Q(z_0)}{P(z_0)}$$

Using this lemma we could derive an explicit parametrization for $X_{ns}^{+}(3)$ and $X_{ns}^{+}(4)$.

Lemma 3.46. *There exists a uniformizer $t : X_{ns}^{+}(3) \to X(1)$ defined over $\mathbb{Q}$ verifying*

$$j = r^3$$

and such that $r(0) = 0$ and $r(\infty) = \infty$.

Proof. We recall that $X_{ns}^{+}(3)$ is of degree 3 over the j-line and it is ramified over ρ and ∞ with degree 3. This means (propositions 3.40 and 3.42) that we have only one cusp and one elliptic point of order 3 and they turn out to be rational; this is because $j(\rho)$ and $j(\infty)$ lie in $\mathbb{P}^1(\mathbb{Q})$ and the uniformizer is a rational function. Thus, the actino of $\mathcal{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ on $X_{ns}^{+}(3)$ sends the point above ∞ to a point over ∞ and a point above ρ still to a point above ρ. Therefore, by the uniqueness of these points, they have to be fixed by all the elements of $\mathcal{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ and so they have to be rational. Now, by the previous lemma,

$$j = \lambda\frac{(r - r(\rho))^3}{(r - r(\infty))^3}$$

By an automorphism of $\mathbb{P}^1$ we can assume $r(\rho) = 0$ and $r(\infty) = \infty$ and this imply that $j = \lambda r^3$. Finally, we note that λ has to be a cube since $j(\mathbb{Z}[i]) = 1728 = 12^3$ and 3 is unramified in $\mathbb{Z}[i]$ meaning that $\mathbb{Z}[i]$ gives rise to a point on $X_{ns}^{+}(3)$; thus, $r(\mathbb{Z}[i]) \in \mathbb{Q}$. Hence, by re-scaling, we can assume $\lambda = 1$ and $j = r^3$. □

Lemma 3.47 ([Che, Cor. 4.2]). *An elliptic curve E/K with $K \subseteq \bar{\mathbb{Q}}$ gives rise to a $\mathbb{Q}$-rational point on $X_{ns}^{+}(3)$ if and only if $j(E)$ is a cube in $\mathbb{Q}$.*

In the same way (it only requires some more work) we obtain

Lemma 3.48 ([Boo, Th. 51]). *There exists a choice of a uniformizer $t : X_{ns}^{+}(4) \to X(1)$ such that*

$$j = 2^{-14}t(t-1)^3 \qquad t(\sigma_1) = \frac{5}{4} + \frac{\sqrt{-2}}{4} \text{ and } t(\sigma_2) = \frac{5}{4} - \frac{\sqrt{-2}}{4}$$

where σ_1 and σ_2 are the two elliptic points of order 2 in $X_{ns}^{+}(4)$.

Lemma 3.49 ([Che, Cor. 5.3]). *There exists a choice of uniformizer $s : X_{ns}^{+}(5) \to X(1)$ verifying*

$$j = 5^3\frac{s(2s \pm 1)^3(2s^2 \pm 7s + 8)^3}{(s^2 + s - 1)^5}$$

such that $s(1)$ and $s(\infty)$ are roots of the quadratic polynomial $X^2 + X - 1$.

3.3 Weber modular curves

The goal of this section is to introduce some families of modular curves that will be of use in the next chapters. We will discuss the problem of automorphisms or double edges in isogeny graphs in the following chapters, but we have already seen in section 2.3 how these affect the construction of isogeny chains or isogeny rectangles. In the following, we are going to play with the following constraints:

- On the one hand, we want our curve to have low genus, preferably 0 so that its points are parametrized by a single value making computations more efficient.

- On the other hand, we would also like to raise the level structure in order to eliminate problems at vertices with extra automorphisms or double edges.
- Finally, increasing the level of the modular structure results in size reduction for modular polynomials.

We look therefore for the best compromise combining all these properties.

3.3.1 Rigidification of points

We have already seen that adding full level N structure, for $N \geq 3$, resolves all extra automorphisms (see proposition 3.4).

$$\begin{tikzcd}
 & X(2) \arrow[dl, dash, "2"'] \\
X_1(2) \arrow[d, equal] & \\
X_0(2) \arrow[dr, dash, "3"'] & \\
 & X(1)
\end{tikzcd}
\qquad
\begin{tikzcd}
 & X(p) \arrow[dl, dash, "p"'] \\
X_1(p) \arrow[d, dash, "(p-1)/2"', "(\mathbb{Z}/p\mathbb{Z})^{\times}\,/\,\{\pm 1\}"] & \\
X_0(p) \arrow[dr, dash, "p+1"'] & \\
 & X(1)
\end{tikzcd}$$

The curves $X(2)$, $X(3)$, $X(4)$ and $X(5)$ are genus zero examples while $X(6)$ has already genus 1. Generally, $g_0(p) = g(X_0(p)) \approx (p-1)/12 > 0$. We are mostly interested in $X(\Gamma)$ for a congruence subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ for $N = 2^a3^b$ to help with $\ell = 2$ or $\ell = 3$ isogeny chains.

		$a =$							
	$g_0(2^a3^b)$	0	1	2	3	4	5	6	
$b =$	0	0	0	0	0	0	1	3	
	1	0	0	0	1	3	9	21	
	2	0	0	1	5	13	33	73	
	3	1	4	10	25	55	121	253	
	4	4	16	37	85	$*$	$*$	$*$	↓
	5	19	$*$	$*$	$*$	$*$	$*$	$*$	keeps growing
	6	64	$*$	$*$	$*$	$*$	$*$	$*$	↓

So that the genus is zero for $N = 2^a3^b = 1, 2, 3, 4, 6, 8, 9, 12, 16, 18$.

One might also ask what happens if we relax a little bit our expectations: for instance, the modular curve associated with the normalizer of the non-split Cartan subgroup of level 3 does not have elliptic points of order 3 but only 3 of order 2. In order to eliminate double edges it suffices to move up to $X_0(N)$ where indeed we differentiate all possible isogeny kernels. Thus, we might also want to consider curves of the form $X(\Gamma(2) \cap \Gamma_0(3^b))$, $X(\Gamma(4) \cap \Gamma_0(3^b))$ or $X(\Gamma(3) \cap \Gamma_0(2^a))$. In order to do that, we start by considering the different curves involved and we eventually build the towers of modular curves on which they rely.

3.3.2 Adding level structure

When $N = 2, 3, 4$ or 6, the automorphism group $\Gamma_0(N^2)$ is conjugate in PSL to the level N principal congruence subgroup $\Gamma(N) \subseteq \mathrm{SL}_2(\mathbb{Z})$; hence, the corresponding modular curves $X_0(N^2)$ and $X(N)$ are isomorphic. Further, for any p, the quotient curve $X_0^+(p)$ is isomorphic to the curve $X_{sp}(p)$:

Lemma 3.50 ([BKX, Lem. 1])**.** *For $N \geq 3$, there exist morphisms $\pi_1 : X_1(N^2) \to X(N)$ and $\pi_0 : X(N) \to X_0(N^2)$ of degree $\phi(N)/2$ defined over $\mathbb{Q}$ such that the composition is the natural forgetful map and the following diagram commutes.*

$$\begin{tikzcd}
 & & X_1(N^2) \arrow[dddll, bend right] \arrow[d, "\pi_1"] & & \\
 & & X(N) \arrow[ddll, bend right, "f_1"'] \arrow[dl, "\pi_0"'] \arrow[dddrr, bend left] & & \\
 & X_0(N^2) \arrow[rr, leftrightarrow, "\omega_{N^2}"] \arrow[dd] & & X_0(N^2) \arrow[dd] & \\
X_1(N) \arrow[d, "\rho"'] & & & & \\
X_0(N) \arrow[r, leftrightarrow, "\omega_N"] \arrow[drr] & X_0(N) \arrow[dr] & & X_0(N) \arrow[r, leftrightarrow, "\omega_N"] \arrow[dl] & X_0(N) \arrow[dll] \\
 & & X(1) & &
\end{tikzcd}$$

where ω_N and ω_{N^2} are Atkin-Lehner involutions, and the maps with no name are the usual forgetful maps.

Remark. The map f_1 is the natural forgetful map $f_1 : X(N) \to X_1(N)$ given, in the moduli interpretation, by sending the pair (E, φ) to the pair $(E, \varphi^{-1}(1,1))$, since $\varphi^{-1}(1,1)$ is a point of exact order N.

This lemma implies that for $N = 3, 4, 6$, the modular curves $X(N)$ and $X_0(N^2)$ are identical over $\mathbb{Q}(\xi_N)$. This is analogous to the case of the curves $X_1(N)$ and $X_0(N)$ for $N = 3, 4, 6$.

The map $X(N) \to X_0(N^2)$ can be described on $Y(N)$ by sending the point $(E, \phi : E[N] \to \mathbb{Z}/N\mathbb{Z}\times\mathbb{Z}/N\mathbb{Z})$ on $Y(N)$ to the N^2 cyclic isogeny obtained by composing the dual N-isogeny $E \to E/F_1$ with the N-isogeny $E \to E/F_2$ where $F_1 = \phi^{-1}(\mathbb{Z}/N\mathbb{Z} \times \{1\})$ and $F_2 = \phi^{-1}(\{1\} \times \mathbb{Z}/N\mathbb{Z})$.

Corollary 3.51 ([BKX, Cor. 3])**.** *The curve $X(N)$ is isomorphic over $\mathbb{Q}$ to the fiber product of $X_1(N)$ and $X_0(N^2)$ over $X_0(N)$, with respect to the natural maps $X_1(N) \to X_0(N)$ and $X_0(N^2) \to X_0(N) \xrightarrow{\omega_N} X0(N)$ (composition of the natural map with the Atkin-Lehner involution).*

3.3.3 Modular curves $X(N)$ for $N = 2, 3, 4, 5$

The Modular curve $X(2)$

We know that any elliptic curve $E/\mathbb{C}$ has a Legendre model $y^2 = x(x-1)(x-\lambda)$, where $\lambda \in \mathbb{C} \setminus \{0, 1\}$ is the λ-invariant [Sil1, §III.1]. The λ-invariant can be chosen to be a single-valued function of $\tau \in \mathbb{H}$:

$$\lambda(q) = 2^4 \left(q_2 - 8q_2^2 + 44q_2^3 - 192q_2^4 + \dots\right) \qquad q_2 = q^{1/2} = e^{\pi\tau}$$

The λ-invariant is a Hauptmodul for $X(2)$ and the j-invariant can be expressed in terms of it by

$$j(\tau) = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda-1)^2}(\tau) = 2^4 \frac{(\lambda^2 + 14\lambda + 1)^3}{\lambda(\lambda-1)^4}(2\tau)$$

Since the group $\Gamma(2)$ and $\Gamma_0(4)$ are conjugate to each other in PSL by a 2-isogeny, the curves $X(2)$ and $X_0(4)$ are isomorphic and their canonical Hauptmoduls λ and t_4 are related by

$$\lambda(\tau) = \frac{16}{t_4 + 16}\left(\frac{\tau}{2}\right)$$

where $t_4 = (\eta(\tau)/\eta(4\tau))^8$ is the classical η-product for $X_0(4)$, see Appendix A.

The modular curve $X(3)$

Following [Cox, §12A], we can define the functions

$$\gamma_2(\tau) = j^{1/3}(\tau) \qquad \text{and} \qquad \gamma_3(\tau) = (j - 12)^{1/2}(\tau)$$

We Recall that in lemma 3.46, we proved that γ_2 is a Hauptmodul for $X_{ns}^+(3)$. By proposition 12.3 of [Cox, §12A], we obtain that $\gamma_2(3\tau)$ is a modular function for the group $\Gamma_0(9)$ which is conjugate to $\Gamma(3)$ in PSL.

By direct computation we get

$$\gamma_2(3\tau) = \frac{(t_9+3)(t_9+9)(t_9^2+27)}{t_9(t_9^2+9t_9+27)}$$

where $t_9 = (3\eta_9(\tau)/\eta_1(\tau))^3$ is the normalized Hauptmodul for $X_0(9)$, see [Mai].

Remark. We could use $t_9 = (\eta_1/\eta_9)^3$ as well; we would get the same algebraic relation with γ_2.

Any elliptic curve over $\mathbb{C}$ has cubic Hesse model

$$x^3+y^3+1-(\gamma+3)xy = 0 \tag{$*$}$$

where $\gamma \in \mathbb{C} \setminus \{0, 3(\xi_3-1), 3(\xi_3^2-1)\} = \mathbb{C} \setminus \left\{0, \frac{-9\pm 3\sqrt{3}}{2}\right\}$ [Hes] (see also [Cas1, Ch. 8]). We will study this family of curves in more details in Section 4.1.6.

The γ invariant, as it happens for the λ-invariant and the j-invariant may be taken to be a function on the upper half plane with the following algebraic relation with the j-invariant

$$j(\tau) = \frac{(\gamma+3)^3(\gamma+9)^3(\gamma^3+27)^3}{\gamma^3(\gamma^2+9\gamma+27)^3}(\tau) = \frac{(\gamma+3)^3(\gamma^3+9\gamma^2+27\gamma+3)^3}{\gamma(\gamma^2+9\gamma+27)}(3\tau)$$

Remark. To see this, one simply compute the j-invariant of the model $(*)$. First, we observe that $x^3+y^3+1 = 3Dxy$ is birationally equivalent to $y^2 = x^3-27D(D^3+8)x+54(D^6-20D^3-8)$ which is an elliptic curve of j-invariant

$$j = 1728\frac{4a^3}{4a^3+27b^2} = \left(\frac{3D(D^3+8)}{D^3-1}\right)^3$$

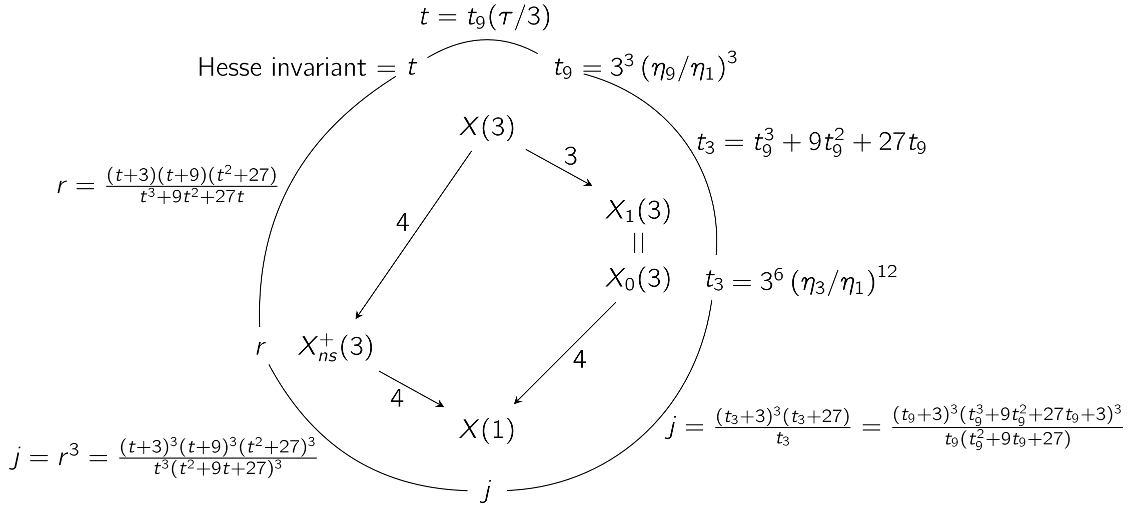
In our case $D = (\gamma+3)/3$ and therefore

$$j = \frac{(\gamma+3)^3(\gamma+9)^3(\gamma^3+27)^3}{\gamma^3(\gamma^2+9\gamma+27)^3}$$

Remark. We note the equivalence between the two relations found so far

$$j(\tau) = \frac{(\gamma+3)^3(\gamma+9)^3(\gamma^3+27)^3}{\gamma^3(\gamma^2+9\gamma+27)^3} \qquad \gamma_2(3\tau) = \frac{(t_9+3)(t_9+9)(t_9^2+27)}{t_9(t_9^2+9t_9+27)}$$

meaning that we could take $\gamma(\tau) = t_9(\tau/3)$. Since $t_9(\tau/3)$ is a Hauptmodul for $\Gamma(3)$, then the Hesse invariant γ is also a Hauptmodul for $\Gamma(3)$ and the Hesse model is associated to $\Gamma(3)$ exactly as the Legendre model is associated to $\Gamma(2)$ [Mai].

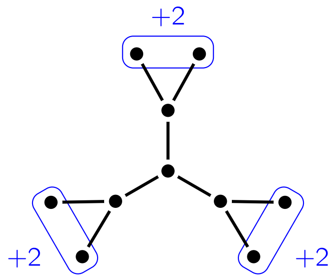

$t = t_9(\tau/3)$
Hesse invariant $= t$
$t_9 = 3^3\,(\eta_9/\eta_1)^3$
$X(3)$
3
$t_3 = t_9^3+9t_9^2+27t_9$
$r = \frac{(t+3)(t+9)(t^2+27)}{t^3+9t^2+27t}$
4
$X_1(3)$
$||$
$X_0(3)$
$t_3 = 3^6\,(\eta_3/\eta_1)^{12}$
r
$X_{ns}^+(3)$
4
4
$X(1)$
$j = \frac{(t_3+3)^3(t_3+27)}{t_3} = \frac{(t_9+3)^3(t_9^3+9t_9^2+27t_9+3)^3}{t_9(t_9^2+9t_9+27)}$
$j = r^3 = \frac{(t+3)^3(t+9)^3(t^2+27)^3}{t^3(t^2+9t+27)^3}$
j

The modular curve $X(4)$

The Legendre form $y^2 = x(x-1)(x-\lambda)$ provides a family of elliptic curves E_λ with an isomorphism $\mathbb{Z}/2\mathbb{Z} \times \mu_2 \to E_\lambda$ and therefore, as we have seen, the parameter λ provides a parametrization of $X(2)$, i.e., $k(\Gamma(2)) = k(\lambda)$. We can compute ψ_4/ψ_2, the quotient of the two division polynomials of E_λ with respect to 4 and 2, and obtain the primitive 4-division polynomial (encoding the points of exact order 4), see Chapter 4.

$$\deg \frac{\psi_4}{\psi_2} = \frac{16-4}{2} = \frac{E[4] \setminus E[2]}{\{\pm 1\}} = 6$$

This represents the number of curves at distance 2 from any chosen elliptic curve in the 2-isogeny graph:



The function field $k(X(4))$ is obtained by adjoining to $k(X(2)) = k(\lambda)$ roots of any two of the factors

$$\psi_4(x)/\psi_2(x) = (x^2-\lambda)(x^2-2x+\lambda)(x^2-2\lambda x+\lambda)$$

$$E[4]/E[2] \simeq (\mathbb{Z}/2\mathbb{Z})^2 \left\{ \begin{tikzcd} X(4) \arrow[dd, dash, "4"'] \arrow[dr, dash, "2"] & \\ & X\left(\Gamma(2) \cap \Gamma_0(4)\right) \\ X(2) \arrow[ur, dash, "2"'] & \end{tikzcd} \right.$$

Let $\mu^2 = \lambda$, then $\psi_4(x)/\psi_2(x) = (x-\mu)(x+\mu)(x^2-2x+\mu^2)(x^2-2\mu^2 x+\mu^2)$. The discriminants of the two quadratic polynomials are $-4(\mu^2-1)$ and $-4\mu^2(\mu^2-1)$ respectively; hence, if k contains $i = \sqrt{-1}$, then the splitting fields of f_1 and f_2 are the same.

$$k(X(4)) = k(\mu)\underbrace{[\sqrt{\mu^2-1}]}_{\substack{\text{conic} \\ \nu^2=\mu^2-1}} \supseteq k(\lambda) = k(\mu^2)$$

A parametrization of the conic gives a rational function (of degree 1) generating the function field of $X(4)$.

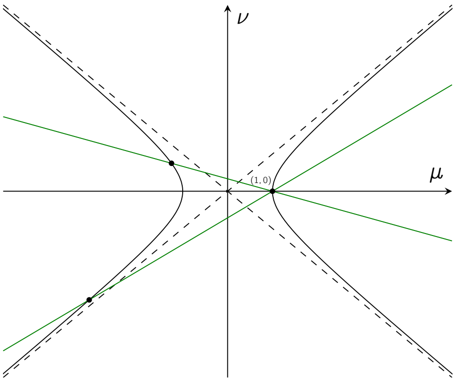


Figure 3.5 – Parametrization by $\mathbb{P}^1$ in (u,v): we fix the point $(1,0)$ and look at the system of lines passing through it. Each line intersects the conic in exactly one point and the slope of the line defines the parametrization.

A straightforward computation gives

$$\text{slope} = t = \frac{\nu}{\mu - 1} \implies t^2\mu - t^2 = \mu + 1 \implies \lambda = \mu^2 = \frac{(1+t^2)^2}{(1-t^2)^2}$$

We obtain a family of elliptic curves E_t with a prescribed isomorphism $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \to E_t[4]$; the new parameter t generates the function field of $X(4)$ and the relation above describes the map $X(4) \to X(2)$.

The modular curve $X(5)$

$\Gamma(5)$ is the congruence subgroup of full level five. It has index 60 in $\mathrm{SL}_2(\mathbb{Z})$ and signature $(g, \epsilon_\infty, \epsilon_2, \epsilon_3) = (0, 12, 0, 0)$. We have the following tower

$$\begin{tikzcd}
X(5) \arrow[d, dash, "5"'] & k(f) \arrow[d, dash, "5"'] \\
X_1(5) \arrow[d, dash, "2"'] & k(v) \arrow[d, dash, "2"'] \\
X_0(5) \arrow[d, dash, "6"'] & k(t_5) \arrow[d, dash, "6"'] \\
X(1) & k(j)
\end{tikzcd}$$

The Dedekind η-product $t_5 = (\eta_1(\tau)/\eta_5(\tau))^6$ is a parameter on $X_0(5)$ with maps down to the j-line

$$\pi_1^*(j) = \frac{(t_5^2 + 250t_5 + 3125)^3}{t_5^5} \qquad \text{and} \qquad \pi_1^*(j) = \frac{(t_5^2 + 10t_5 + 5)^3}{t_5}$$

If we now look at the elliptic curve $E_v : y^2 + (v+1)xy + vy = x^3 + vx^2$, we note that $P = (0, 0)$ is a 5-torsion point on E_v. Thus, the parameter v is a generator for the function field of $X_1(5)$. Note that $\Gamma_1(5)$ is a torsion free genus zero subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of index 12 and the lack of torsion implies that an Hauptmodul on $X_1(5)$, as v, provides an isomorphism with the projective line. The j-invariant of E

$$j(E_v) = -\frac{(v^4 + 12v^3 + 14v^2 - 12v + 1)^3}{v^5(v^2 + 11v - 1)}$$

provides a map to $X(1)$ and, by comparison, we get the relation

$$t_5 = -\frac{125v}{v^2 + 11v - 1}$$

Finally, $\Gamma(5)$ corresponds to a cyclic cover of degree 5 of $X_1(5)$ and, therefore, a parameter on it is a quintic root of v, $v = f^5$.

3.3.4 A new modular curve

For simplicity we will write $X_1 \otimes X_2$ for the normalization (non-singular model) of the fiber product $X_1 \times_{X(1)} X_2$ given two modular curves $X_1 \to X(1)$ and $X_2 \to X(1)$. We will be interested in the following situation

$$\begin{tikzcd}
 & & X = Y \otimes X(3) \arrow[dl, dash] \arrow[dr, dash] & \\
 & X_0(2) \otimes X_{ns}^+(3) = Y \arrow[dl, dash] \arrow[dr, dash] & & X(3) \arrow[dl, dash] \\
X_0(2) \arrow[dr, dash] & & X_{ns}^+(3) \arrow[dl, dash] & \\
 & X(1) & &
\end{tikzcd}$$

Let us start by looking at the full level three modular curve $X(3)$.

$$\begin{tikzcd}
& \overset{\overset{\text{Hesse invariant}}{t}}{X(3)} \arrow[dl, dash, "3"'] \arrow[ddr, dash, "4"] & \\
t_3\ \ X_0(3) \arrow[ddr, dash, "4"'] & & \\
& & X^+_{ns}(3)\ \ r \arrow[dl, dash, "3"] \\
& \underset{\underset{j\text{-invariant}}{j}}{X(1)} &
\end{tikzcd}$$

In Section 3.3.3 we have already described it and yet we will find another parameter on it in order to find more compact relations with the j-invariant. We take t_3, the Hauptmodul for $X_0(3)$ to be

$$t_3 = 27 + \left(\frac{\eta_1(\tau)}{\eta_3(\tau)}\right)^{12} \qquad \text{so that} \qquad j = \frac{t_3(t_3+216)^3}{(t_3-27)^3}$$

As usual, we consider r, the parameter on $X^+_{ns}(3)$ to be a cube root of the j-invariant. Thus, to pass from $X_0(3)$ to $X(3)$ we need to add a cube root of j. We take $t^3 = t_3$. Hence,

$$j = \frac{t^3(t^3+216)^3}{(t^3-27)^3}$$

and then

$$j = r^3 = \frac{t^3(t^3+216)^3}{(t^3-27)^3} \implies r = \frac{t(t^3+216)}{t^3-27}$$

We eventually look at the curve Y

$$\begin{tikzcd}
& \overset{u}{Y} \arrow[dl, dash, "3"'] \arrow[ddr, dash, "3"] & \\
s\ \ X_0(2) \arrow[ddr, dash, "3"] & & \\
& & X^+_{ns}(3)\ \ r \arrow[dl, dash, "3"', "j=r^3"] \\
& \underset{\underset{j\text{-invariant}}{j}}{X(1)} &
\end{tikzcd}$$

On $X_0(2)$ we define the classical Hauptmodul, see Appendix A

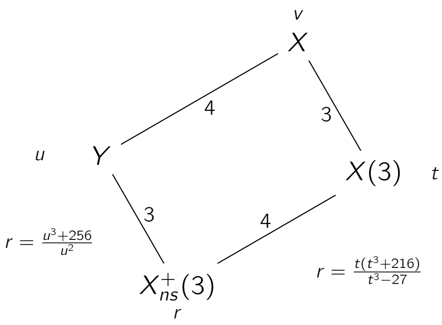
$$s = t_2 = \left(\frac{\eta_1(\tau)}{\eta_2(\tau)}\right)^{24} \qquad \text{such that} \qquad j = \frac{(s+256)^3}{s^2}$$

In order to define a parameter on Y we need to add a cube root of j. Thus we define $u^3 = s$ so that

$$j = \frac{(u^3+256)^3}{u^6} \implies r = \frac{u^3+256}{u^2}$$

Remark. Note that we can use the Atkin-Lehner copy of $X_0(2)$ in order to have $j = (s+16)^3/s$ and then take $u^3 = -s$ to ease the construction of the Weber modular curve, as we will see shortly.

Finally we find a generator v for the function field of X.



We have a complete description of the two maps

$$X(3) \longrightarrow X_{ns}^+(3) \qquad\qquad Y \longrightarrow X_{ns}^+(3)$$

$$t \longrightarrow r = \frac{t(t^3+216)}{t^3-27} \qquad\qquad u \longrightarrow r = \frac{u^3+256}{u^2}$$

Setting these equal, we find

$$\frac{t(t^3+216)}{t^3-27} = \frac{u^3+256}{u^2} \quad \Longrightarrow \quad (t^3+216)tu^2 - (u^3+256)(t^3-27) = 0$$

This is a (singular) model for X. Using `magma` we find out that X has genus 0. As we did in the previous section for $X(4)$, we can use the anti-canonical divisor (of degree 2) to obtain a plane conic model and, eventually, find a rational parametrization (given any point) by

$$\mathbb{P}^1 \longrightarrow X \qquad v \longmapsto (t, u) = (t(u), v(u))$$

Using `magma`, we find

$$t = \frac{v^3-2}{v} \qquad u = \frac{16(v^3+1)}{(v^3-8)v}$$

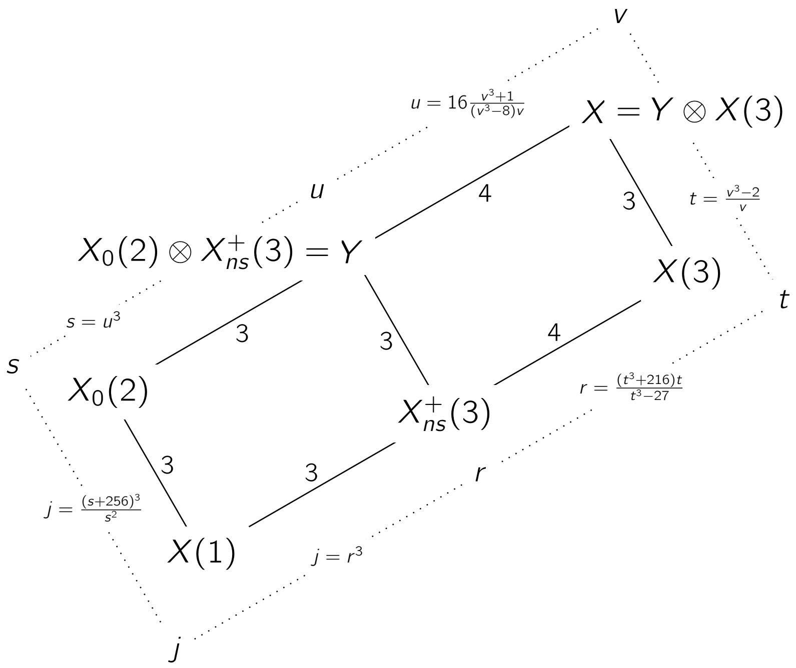

v
$u = 16\frac{v^3+1}{(v^3-8)v}$
$X = Y \otimes X(3)$
u
4
3
$t = \frac{v^3-2}{v}$
$X_0(2) \otimes X_{ns}^+(3) = Y$
X(3)
t
$s = u^3$
3
3
4
s
$X_0(2)$
$r = \frac{(t^3+216)t}{t^3-27}$
$X_{ns}^+(3)$
3
3
r
$j = \frac{(s+256)^3}{s^2}$
X(1)
$j = r^3$
j

Figure 3.6 – Constructing the modular curve X.

By successive approximations we can compute their q-expansions finding:

$$u(q) = q^{-\frac{1}{3}} - 8q^{\frac{2}{3}} + 28q^{\frac{5}{3}} - 64q^{\frac{8}{3}} + 134q^{\frac{11}{3}} - 288q^{\frac{14}{3}} + 568q^{\frac{17}{3}} + O(q^6)$$

$$t(q) = q^{-\frac{1}{3}} + 5q^{\frac{2}{3}} - 7q^{\frac{5}{3}} + 3q^{\frac{8}{3}} + 15q^{\frac{11}{3}} - 32q^{\frac{14}{3}} + 9q^{\frac{17}{3}} + O(q^6)$$
$$v(q) = q^{-\frac{1}{6}} + q^{\frac{1}{3}} + q^{\frac{5}{6}} - q^{\frac{4}{3}} - q^{\frac{11}{6}} + q^{\frac{17}{6}} + 2q^{\frac{10}{3}} - 2q^{\frac{13}{3}} - 3q^{\frac{29}{6}} + O(q^5)$$

Remark. We observe that the modular curve

$$X = X(3) \otimes X_0(2) = X(\Gamma_0(2) \cap \Gamma(3))$$

is associated with the group $\Gamma = \Gamma_0(2) \cap \Gamma(3)$ which has index two in $\Gamma(6)$.

3.3.5 Weber modular curves

Weber modular functions

We will need one more class of modular functions. These have been first introduced by Weber in [Web] and they are defined in terms of the Dedekind η-function (see §2.2.3). This new family of functions has the advantage of providing effective ways of calculating γ_2. We recall that

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{+\infty} (1 - q^n)$$

and that, since the modular discriminant does not vanish on the upper half plane, we may define a cube root of the j-invariant

$$\gamma_2(\tau) = \frac{E_4(q)}{\gamma_2(q)^{1/3}} = j(q)^{1/3}$$

Definition. The Weber modular functions $\mathfrak{f}$, $\mathfrak{f}_1$ and $\mathfrak{f}_2$ are defined to be

$$\mathfrak{f}(\tau) = \zeta_{48}^{-1} \frac{\eta(\frac{\tau+1}{2})}{\eta(\tau)} = q^{-\frac{1}{48}} \prod_{n=1}^{+\infty} \left(1 + q^{n-\frac{1}{2}}\right)$$

$$\mathfrak{f}_1(\tau) = \frac{\eta(\frac{\tau}{2})}{\eta(\tau)} = q^{-\frac{1}{48}} \prod_{n=1}^{+\infty} \left(1 - q^{n-\frac{1}{2}}\right)$$

$$\mathfrak{f}_2(\tau) = \sqrt{2}\frac{\eta(2\tau)}{\eta(\tau)} = \sqrt{2} q^{\frac{1}{24}} \prod_{n=1}^{+\infty} (1 + q^n)$$

These functions satisfy various identities, among which,

$$\mathfrak{f}_1(2\tau)\mathfrak{f}_2(\tau) = \mathfrak{f}(\tau)\mathfrak{f}_1(\tau)\mathfrak{f}_2(\tau) = \sqrt{2}$$

the latter showing that, unlike the case of the modular j-function, the values of these functions at singular moduli are “almost” algebraic units [GS].

Remark. It is easy to see that each of the functions $\mathfrak{f}$, $\mathfrak{f}_1$ and $\mathfrak{f}_2$ have rational Fourier expansions.

The definition of Weber functions may seem quite arbitrary at first but they arise in a natural way when studying the 2-torsion of the universal elliptic curve over $\mathbb{C}$.

Theorem 3.52. *Let $e_1 = \wp(\tau/2)$, $e_2 = \wp(1/2)$ and $e_1 = \wp((\tau+1)/2)$, then*

$$e_2 - e_1 = \pi^2 \eta(\tau)^4 \mathfrak{f}(\tau)^8$$
$$e_2 - e_3 = \pi^2 \eta(\tau)^4 \mathfrak{f}_1(\tau)^8$$
$$e_3 - e_1 = \pi^2 \eta(\tau)^4 \mathfrak{f}_2(\tau)^8$$

Corollary 3.53. *The following identities hold*

$$\gamma_2(\tau) = \frac{\mathfrak{f}(\tau)^{24} - 16}{\mathfrak{f}(\tau)^8} = \frac{\mathfrak{f}_1(\tau)^{24} + 16}{\mathfrak{f}_1(\tau)^8} = \frac{\mathfrak{f}_2(\tau)^{24} + 16}{\mathfrak{f}_2(\tau)^8}$$

$$\gamma_3(\tau) = \frac{(\mathfrak{f}(\tau)^{24}+8)(\mathfrak{f}_1(\tau)^8 - \mathfrak{f}_2(\tau)^8)}{\mathfrak{f}(\tau)^8}$$

From this we could also deduce an expression for j as a rational function in $\mathfrak{f}$, $\mathfrak{f}_1$ or $\mathfrak{f}_2$. Finally, as a consequence of the transformation laws of the Dedekind η-function, we have the following

Proposition 3.54. *The Weber functions satisfy the following transformation properties*

$$\begin{aligned} \mathfrak{f}(\tau+1) &= \zeta_{48}^{-1}\mathfrak{f}_1(\tau) & \qquad \mathfrak{f}\left(-\frac{1}{\tau}\right) &= \mathfrak{f}(\tau) \\ \mathfrak{f}_1(\tau+1) &= \zeta_{48}^{-1}\mathfrak{f}(\tau) & \qquad \mathfrak{f}_1\left(-\frac{1}{\tau}\right) &= \mathfrak{f}_2(\tau) \\ \mathfrak{f}_2(\tau+1) &= \zeta_{24}\mathfrak{f}_2(\tau) & \qquad \mathfrak{f}_2\left(-\frac{1}{\tau}\right) &= \mathfrak{f}_1(\tau) \end{aligned}$$

This proposition gives

Proposition 3.55. $\mathfrak{f}(\tau)^6$ *is a modular function for*

$$\Gamma_s(8) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \; b \equiv c \equiv 0 \mod 8 \right\}$$

Proposition 3.56. $\mathfrak{f}_1(8\tau)^6$ *is a modular function for* $\Gamma_0(32)$.

Proposition 3.57. $\mathfrak{f}_2(\tau)$ *is a modular function for* $\Gamma(24)$.

Note that f_2 was the first to be studied by Weber, who eventually introduced the related functions $\mathfrak{f}$ and $\mathfrak{f}_1$ satisfying

$$(X+\mathfrak{f}^8)(X-\mathfrak{f}_1^8)(X-\mathfrak{f}_2^8) = X^3 - \gamma_2 X + 16$$

Each of these functions generates an extension of degree 72 of $\mathbb{Q}(j)$. We construct therefore a new modular curve W with parameter $\mathfrak{f}_2$.

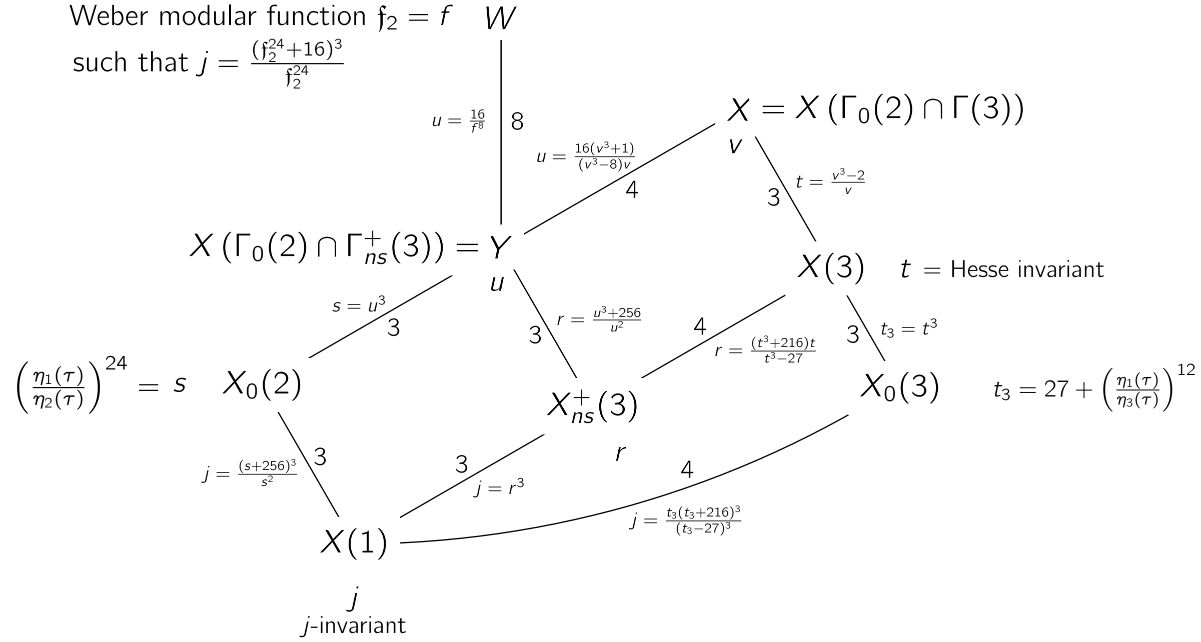

Weber modular function $\mathfrak{f}_2 = f$
such that $j = \frac{(\mathfrak{f}_2^{24}+16)^3}{\mathfrak{f}_2^{24}}$
W
$u = \frac{16}{f^8}$
8
$X = X(\Gamma_0(2) \cap \Gamma(3))$
v
$u = \frac{16(v^3+1)}{(v^3-8)v}$
4
3
$t = \frac{v^3-2}{v}$
$X(\Gamma_0(2) \cap \Gamma_{ns}^+(3)) = Y$
u
$X(3)$
t = Hesse invariant
$s = u^3$
3
3
$r = \frac{u^3+256}{u^2}$
4
$r = \frac{(t^3+216)t}{t^3-27}$
3
$t_3 = t^3$
$\left(\frac{\eta_1(\tau)}{\eta_2(\tau)}\right)^{24} = s$
$X_0(2)$
$X_{ns}^+(3)$
r
$X_0(3)$
$t_3 = 27 + \left(\frac{\eta_1(\tau)}{\eta_3(\tau)}\right)^{12}$
$j = \frac{(s+256)^3}{s^2}$
3
3
$j = r^3$
4
$j = \frac{t_3(t_3+216)^3}{(t_3-27)^3}$
$X(1)$
j
j-invariant

Figure 3.7 – Modular tower for the Weber modular curve W

We note that $\mathfrak{f}$ has a non-rational coefficient. If we consider the twist of $X_0(2)$ by the Atkin-Lehner involution ω_2, we can avoid this issue by working with $\mathfrak{f}(\tau)$ instead of $\mathfrak{f}_2$. Note that the same result can be obtained simply by replacing the Hauptmodul $s = (\eta_1/\eta_2)^{24}$ of $X_0(2)$ by its normalized version

$s = 2^{12}(\eta_2/\eta_1)^{24}$. In this case we also simplify some of the relations between the modular curves involved. Note that the choice $s = -u^3$ has been made in order to work with $\mathfrak{f}$ (instead of $\mathfrak{f}_1$).

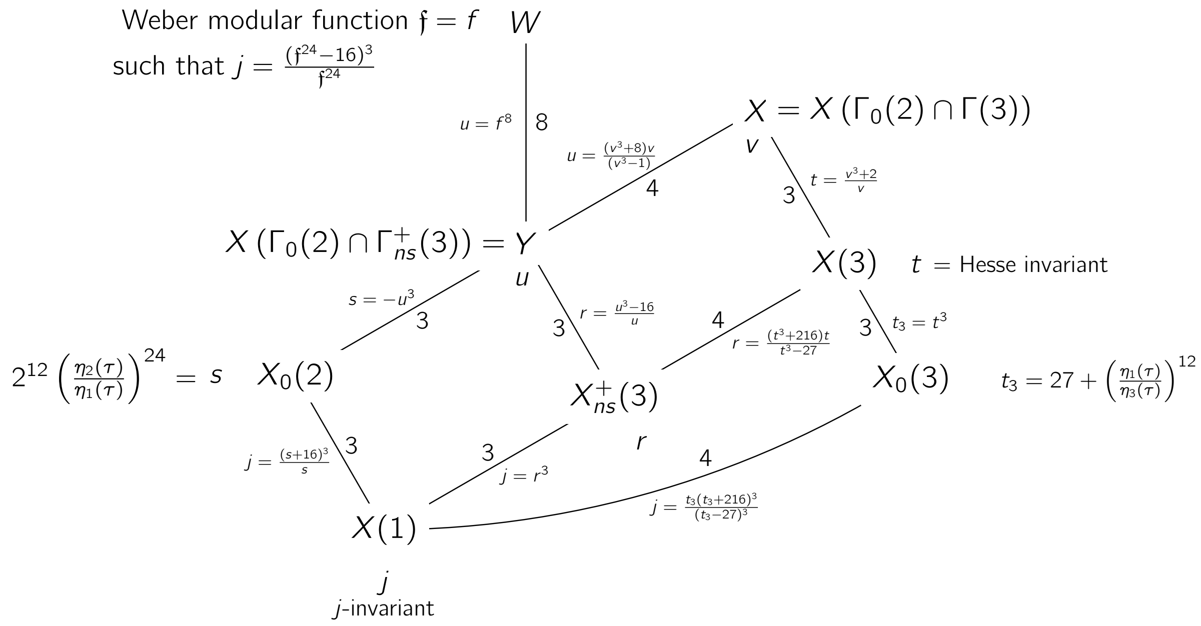

Weber modular function $\mathfrak{f} = f$
such that $j = \frac{(\mathfrak{f}^{24}-16)^3}{\mathfrak{f}^{24}}$
W
$u = f^8$
8
$X = X(\Gamma_0(2) \cap \Gamma(3))$
v
$u = \frac{(v^3+8)v}{(v^3-1)}$
4
$t = \frac{v^3+2}{v}$
3
$X(\Gamma_0(2) \cap \Gamma^+_{ns}(3)) = Y$
u
$X(3)$
t = Hesse invariant
$s = -u^3$
3
$r = \frac{u^3-16}{u}$
3
4
$r = \frac{(t^3+216)t}{t^3-27}$
3
$t_3 = t^3$
$2^{12}\left(\frac{\eta_2(\tau)}{\eta_1(\tau)}\right)^{24} = s$
$X_0(2)$
$X^+_{ns}(3)$
$X_0(3)$
$t_3 = 27 + \left(\frac{\eta_1(\tau)}{\eta_3(\tau)}\right)^{12}$
$j = \frac{(s+16)^3}{s}$
3
3
$j = r^3$
r
4
$j = \frac{t_3(t_3+216)^3}{(t_3-27)^3}$
$X(1)$
j
j-invariant

Figure 3.8 – Normalized modular tower for W

Once established that W and X cover Y we can construct $W \otimes X$ and find that $W \otimes X \to W$ is of degree

$$\deg(W \otimes X \to W) = \deg(X \to Y) = \deg\left(X(3) \to X^+_{ns}(3)\right) = 4$$

so that W does not cover any other intermediate curve. In other words, the two covers $W \to Y$ and $X \to Y$ are independent.

Lemma 3.58. *$W \otimes X \to W$ has genus* 21.

Weber curves

Setting $(\mathfrak{u}_0, \mathfrak{u}_1, \mathfrak{u}_2) = (\mathfrak{f}, \zeta_{16}\mathfrak{f}_1(\tau), \zeta_{16}^{-1}\mathfrak{f}_2(\tau))$ the triple $(\mathfrak{u}_0, \mathfrak{u}_1, \mathfrak{u}_2)$ satisfies the common relations

$$j = \frac{(\mathfrak{u}_0^{24} - 16)^3}{\mathfrak{u}^{24}} = \frac{(\mathfrak{u}_1^{24} - 16)^3}{\mathfrak{u}^{24}} = \frac{(\mathfrak{u}_2^{24} - 16)^3}{\mathfrak{u}^{24}},$$

and one verifies that the three orbits $\{\zeta_{24}^j \mathfrak{u}_i : j \in \mathbb{Z}/24\mathbb{Z}\}$ run over the 72 roots of the modular polynomial

$$(X^{24} - 16)^3 - j(q)X^{24} \in \mathbb{Q}(\zeta_{24})[\![q^{1/24}]\!],$$

and moreover, the functions $\mathfrak{u}_i$ satisfy transformations

$$(\mathfrak{u}_0, \mathfrak{u}_1, \mathfrak{u}_2) \circ S = (\mathfrak{u}_0, \zeta_8^{-1}\mathfrak{u}_2, \zeta_8\mathfrak{u}_1) \text{ and } (\mathfrak{u}_0, \mathfrak{u}_1, \mathfrak{u}_2) \circ T = (\zeta_{24}\mathfrak{u}_1, \zeta_{12}^{-1}\mathfrak{u}_0, \zeta_{24}\mathfrak{u}_2).$$

The map determined by the normalized Weber functions $(\mathfrak{u}_0^m : \mathfrak{u}_1^m : \mathfrak{u}_2^m : 1)$ determines a *Weber modular curve $\mathcal{W}_{3n}$* in $\mathbb{P}^3$

$$\mathcal{W}_{3n} : \begin{cases} X_0^n + X_1^n + X_2^n = 0, \\ X_0X_1X_2 = \sqrt{2}^m X_3^3 \end{cases}$$

for m and n such that $mn = 8$, with quotient Weber curve $\mathcal{W}_n$ defined as the image of $(\mathfrak{u}_0^{3m} : \mathfrak{u}_1^{3m} : \mathfrak{u}_2^{3m} : 1)$ in $\mathbb{P}^3$:

$$\mathcal{W}_n : \begin{cases} X_0^n + X_1^n + X_2^n = 48X_3^n, \\ X_0X_1X_2 = \sqrt{2}^{3m} X_3^3. \end{cases}$$

These defining relations follow directly from the relations $\mathfrak{f}^8 = \mathfrak{f}_1^8 + \mathfrak{f}_2^8$ and $\mathfrak{f}\,\mathfrak{f}_1\mathfrak{f}_2 = \sqrt{2}$, and the curves are equipped with maps $\mathcal{W}_{mn} \to \mathcal{W}_n$ for each product mn dividing 24. We will now show that Weber modular curves fit in the description of previous sections.

Modular group

To each factorization $mn = 24$, the Weber curve $\mathcal{W}_n$ in $\mathbb{P}^3$, defined by the triple of Weber functions $(\mathfrak{u}_0^m, \mathfrak{u}_1^m, \mathfrak{u}_2^m)$, comes equipped with an action of $\mathrm{PSL}_2(\mathbb{Z})$. We denote the kernel of the action by Γ_n, identifying the Weber curves with the modular curve $X(\Gamma_n)$. The action of $\mathrm{PSL}_2(\mathbb{Z})$ on Weber functions induces a representation in $\mathrm{GL}_3(\mathbb{Q}(\zeta_n))$ determined by the images of the generators S and T.

$$\begin{tikzcd} \mathrm{PSL}_2(\mathbb{Z}) \arrow[r] & \mathrm{GL}_3(\mathbb{Q}(\zeta_n)) \\ S, T \arrow[r, mapsto] & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \overline{\zeta}_8^m \\ 0 & \zeta_8^m & 0 \end{pmatrix}, \begin{pmatrix} 0 & \zeta_{24}^m & 0 \\ \overline{\zeta}_{24}^{2m} & 0 & 0 \\ 0 & 0 & \zeta_{24}^m \end{pmatrix} \end{tikzcd}$$

The image group is finite, whose kernel Γ_n is a normal congruence subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. We reduce the computation of Γ_n by first proving that $\Gamma_1 \cong \Gamma(2)$, then noting that $\Gamma_1 \subset \Gamma_3 \cap \Gamma_8 = \Gamma_{24}$, which reduces to determining Γ_3 and the chain $\Gamma_1 \subset \Gamma_2 \subset \Gamma_4 \subset \Gamma_8$.

Let $\boldsymbol{\mu}_m = \langle \zeta_m \rangle$ be the group of m-th roots of unity, and define the anti-diagonal group,

$$\nabla(\boldsymbol{\mu}_m^2) = \left\{ \left(\zeta_m^i, \zeta_m^j, \zeta_m^{-i-j}\right) \mid 0 \le i, j < m \right\},$$

and the diagonal group $\Delta(\boldsymbol{\mu}_m) = \left\{ \left(\zeta_m^i, \zeta_m^i, \zeta_m^i\right) \mid 0 \le i < m \right\}$, each of which acts by coordinate scaling on $\mathbb{A}^3$, and if $m \equiv 0 \bmod 3$, then $\nabla(\boldsymbol{\mu}_m^2) \cap \Delta(\boldsymbol{\mu}_m) = \Delta(\boldsymbol{\mu}_3)$, otherwise the intersection is trivial. For each divisor mn of 24, the action of the anti-diagonal group $\nabla(\boldsymbol{\mu}_m^2)$, extended to $\mathbb{P}^3$, stabilizes $\mathcal{W}_{mn}$.

Proposition 3.59. *For each product mn dividing 24, the group $\nabla(\boldsymbol{\mu}_m^2)$ acts on $\mathcal{W}_{mn}$ with quotient $\mathcal{W}_n$.*

- *If $m \not\equiv 0 \bmod 3$, then the group $\nabla(\boldsymbol{\mu}_m^2)$ acts faithfully and $\Gamma_{mn}/\Gamma_n \cong \nabla(\boldsymbol{\mu}_m^2)$.*
- *If $m \equiv 0 \bmod 3$, then group $\nabla(\boldsymbol{\mu}_m^2)$ acts with kernel $\Delta(\boldsymbol{\mu}_3)$, and $\Gamma_{mn}/\Gamma_n \cong \nabla(\boldsymbol{\mu}_m^2)/\Delta(\boldsymbol{\mu}_3)$.*

In particular, if m divides 8, the degree of $\mathcal{W}_{mn} \to \mathcal{W}_n$ is m^2 and the degree of $\mathcal{W}_{3n} \to \mathcal{W}_n$ is 3.

Proposition 3.60. *The Weber kernel group Γ_1 equals $\Gamma(2)$ and $\mathcal{W}_1 \cong X(2)$.*

Proposition 3.61. *The Weber kernel group Γ_3 equals $\Gamma(2) \cap \Gamma_{ns}^+(3)$, and for each n dividing 8*

$$\Gamma_{3n} = \Gamma_n \cap \Gamma_{ns}^+(3).$$

It thus suffices to characterize the groups Γ_n for n dividing 8.

Proposition 3.62. *The Weber kernel group Γ_2 equals $\Gamma(4)$ and $\mathcal{W}_2 = X(4)$.*

Proposition 3.63. *The Weber kernel group Γ_4 equals $\Gamma_s(8)$ and $\mathcal{W}_4 = X_s(8)$.*

It remains to characterize the group Γ_8 under which the triple of functions $(\mathfrak{u}_0^3, \mathfrak{u}_1^3, \mathfrak{u}_2^3)$ is invariant. This group is not the split Cartan subgroup $\Gamma_s(16)$, but we can show that

$$\Gamma(16) \subset \Gamma_8 \subset \Gamma_s(8) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : b \equiv c \equiv 0 \bmod 8 \right\},$$

and that the group $\Gamma_8/\Gamma(16)$ is cyclic of order 4 generated by

$$T^2 U^2 T^{-2} U^{-2} \equiv \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix} \bmod 16,$$

where $U = STS^{-1}$. The equality is easily verified in $\mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})$ and the word expression on the left maps to the identity under the above homomorphism to $\mathrm{GL}_3(\mathbb{Q}(\zeta_8))$, showing that the element is in the kernel of the action on $\mathcal{W}_8$. Moreover, the matrix on the right lifts to $\mathrm{SL}_2(\mathbb{Z})$. Given that the degree of $\mathcal{W}_8 \to \mathcal{W}_4 = X_s(8)$ is 4, and $X(16) \to X(8)$ is of degree 16, we obtain the following description of the kernel group Γ_8.

Proposition 3.64. *The Weber kernel group* Γ_8 *is the group generated by* $\Gamma(16)$ *and* $\begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}$.

In particular we note that Γ_8 contains the diagonal matrix in the center of $\mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})$:

$$\begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}^2 \equiv \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})$$

hence contains the subgroup

$$\Gamma(16, 8, 16) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ : \ a \equiv d \equiv 1 \bmod 8, b \equiv c \equiv 0 \bmod 16 \right\} \cdot$$

Given that $\Gamma_s(8)/\Gamma(16)$ is an abelian group:

$$\Gamma_s(8)/\Gamma(16) = \left\langle \begin{pmatrix} 13 & 0 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 8 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 8 & 1 \end{pmatrix} \right\rangle \cong C_4 \times V_4,$$

so that $\Gamma_s(8)/\Gamma(16, 8, 16) \cong C_2 \times V_4 = C_2^3$; we have the following diagram of modular curves between $X(16)$ and $X_s(8)$, where the curves represented by dots are intermediate quotients by the subgroups of $V_4 = \langle T^8, U^8 \rangle$.

$$\begin{tikzcd}
 & & X(16) \arrow[d, "\langle(\begin{smallmatrix} 9 & 0 \\ 0 & 9 \end{smallmatrix})\rangle"] & & \\
 & & X(16,8,16) \arrow[dll, bend right, shift right] \arrow[dll, bend right] \arrow[dll, bend right, shift left] \arrow[dl] \arrow[d, "\langle(\begin{smallmatrix} 13 & 0 \\ 0 & 5 \end{smallmatrix})\rangle"] \arrow[dr] \arrow[drr, bend left, "\langle(\begin{smallmatrix} 13 & 8 \\ 8 & 5 \end{smallmatrix})\rangle"] & & \\
\bullet\ \bullet\ \bullet \arrow[d, shift left] \arrow[d] \arrow[d, shift right] & X\big(\Gamma(16), (\begin{smallmatrix} 13 & 8 \\ 0 & 5 \end{smallmatrix})\big) \arrow[d, shift left] \arrow[d] \arrow[d, shift right] & X_s(16) \arrow[d, shift left] \arrow[d] \arrow[d, shift right] & X\big(\Gamma(16), (\begin{smallmatrix} 13 & 0 \\ 8 & 5 \end{smallmatrix})\big) \arrow[d, shift left] \arrow[d] \arrow[d, shift right] & \mathcal{W}_8 \arrow[d, shift left] \arrow[d] \arrow[d, shift right] \\
X(8) \arrow[drr, bend right, "\langle(\begin{smallmatrix} 5 & 0 \\ 0 & 5 \end{smallmatrix})\rangle"'] & \bullet\ \bullet\ \bullet \arrow[dr, shift left] \arrow[dr] \arrow[dr, shift right] & \bullet\ \bullet\ \bullet \arrow[d, shift left] \arrow[d] \arrow[d, shift right] & \bullet\ \bullet\ \bullet \arrow[dl, shift left] \arrow[dl] \arrow[dl, shift right] & \bullet\ \bullet\ \bullet \arrow[dll, bend left, shift left] \arrow[dll, bend left] \arrow[dll, bend left, shift right] \\
 & & X_s(8) & &
\end{tikzcd}$$

3.3.6 Fermat curves

Let $\mathcal{F}_n : X^n + Y^n + Z^n = 0$ be the n-th Fermat curve in $\mathbb{P}^2$. The hyperplane relation $X_0^n + X_1^n + X_2^n = 0$ in $\mathcal{W}_{3n}$ motivates the study of the relation with Fermat curves $\mathcal{F}_n$, in particular the projection $\pi : \mathcal{W}_{3n} \to \mathcal{F}_n$.

In order to obtain a cover of $\mathcal{W}_{3n}$, we also introduce the twists $\mathcal{F}_n^c : X^n + Y^n + cZ^n = 0$ of $\mathcal{F}_n$ by c. In particular, for $c = 16$ and $nm = 8$, the degree 9 quotient, $\mathcal{F}_{3n}^{16} \to \mathcal{F}_n$ given by $(X : Y : Z) \mapsto (X^3, Y^3, \sqrt{2}^m Z^3)$, maps through the projection $\mathcal{W}_{3n} \to \mathcal{F}_n$:

$$\begin{tikzcd}
\mathcal{F}_{3n}^{16} \arrow[r] & \mathcal{W}_{3n} \arrow[r] & \mathcal{F}_n \\
(X : Y : Z) \arrow[r, mapsto] & (X^3 : Y^3 : \sqrt{2}^m Z^3 : XYZ) \arrow[r, mapsto] & (X^3 : Y^3 : \sqrt{2}^m Z^3)
\end{tikzcd}$$

each of degree 3. In fact we can establish that the quotient $\mathcal{W}_{3n} \to \mathcal{W}_n$ induces an isomorphism of $\mathcal{F}_n$ to $\mathcal{W}_n$ such that the quotient of Weber curves is the composition of $\pi : \mathcal{W}_{3n} \to \mathcal{F}_n$ with the isomorphism $\mathcal{F}_n \to \mathcal{W}_n$.

$$\begin{tikzcd}
\mathcal{W}_{3n} \arrow[r] & \mathcal{F}_n \arrow[r] & \mathcal{W}_n \\
(X_0 : X_1 : X_2 : X_3) \arrow[r, mapsto] & (X_0 : X_1 : X_2) \arrow[r, mapsto] & (X_0^3 : X_1^3 : X_2^3 : \dfrac{X_0 X_1 X_2}{\sqrt{2}^m}) = (X_0^3 : X_1^3 : X_2^3 : X_3^3)
\end{tikzcd}$$

Remark. This shows that the curves $\mathcal{F}_n$, for n dividing 8, are modular, in the sense of being a moduli space for elliptic curves with a $2n$-level structure.

The curves $\mathcal{W}_n$ are generated by the triple of functions $(\mathfrak{f}, \mathfrak{f}_1, \mathfrak{f}_2)$, or rather their normalizations $(\mathfrak{u}_0^m, \mathfrak{u}_1^m, \mathfrak{u}_2^m)$, where $nm = 24$. The quotient to $\mathfrak{f}_i^m$ is a genus 0 quotient. For instance, for $m = 24$ (and $n = 1$), we have $\mathcal{W}_1 = X(2)$, but the quotient to $\mathfrak{f}^{24}$ is isomorphic to $X_0(2)$.

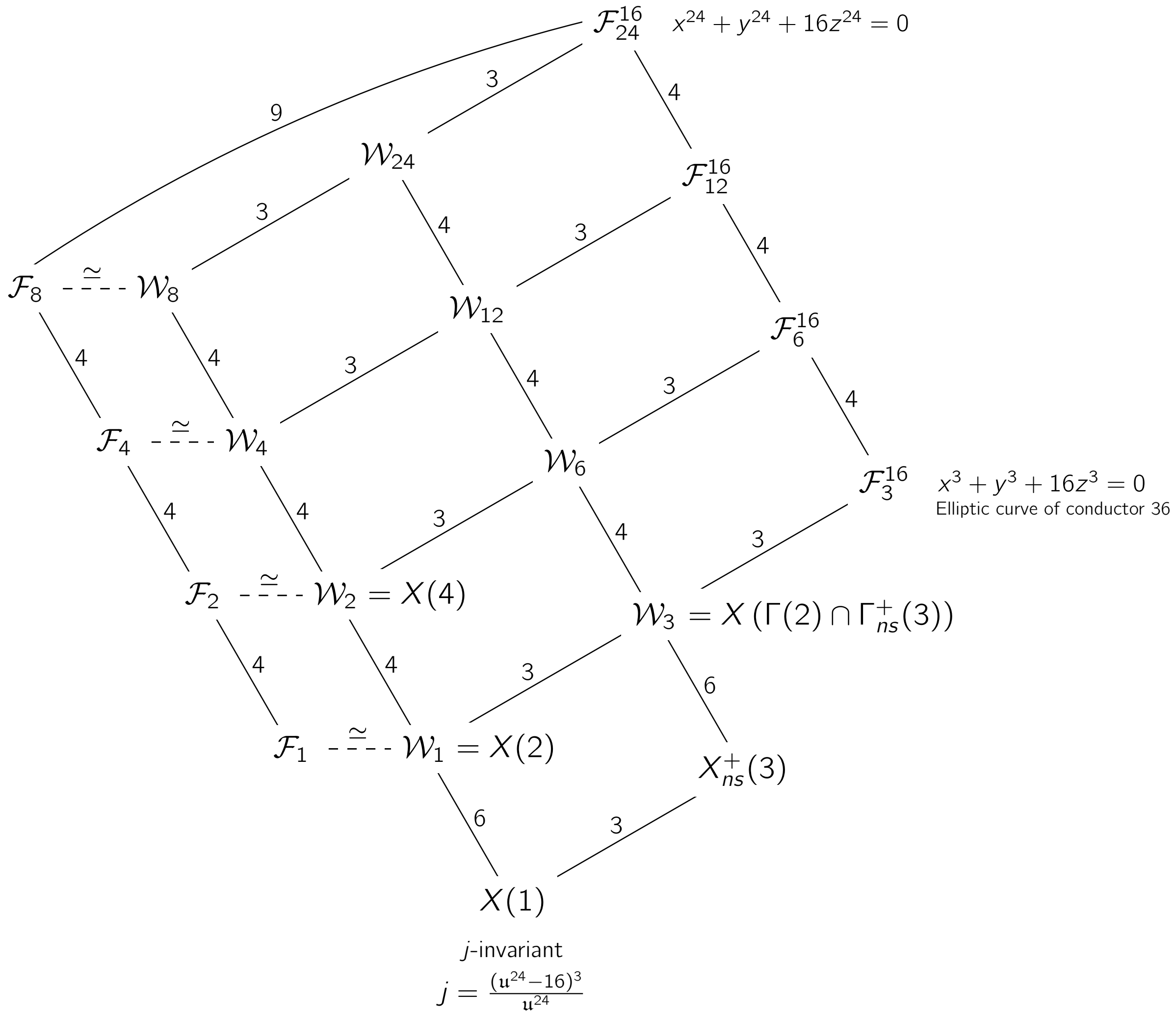

$\mathcal{F}_{24}^{16}$ $x^{24} + y^{24} + 16z^{24} = 0$
9
3
4
$\mathcal{W}_{24}$
$\mathcal{F}_{12}^{16}$
3
4
3
4
$\mathcal{F}_8$ $\simeq$ $\mathcal{W}_8$
$\mathcal{W}_{12}$
$\mathcal{F}_6^{16}$
4
4
3
4
3
4
$\mathcal{F}_4$ $\simeq$ $\mathcal{W}_4$
$\mathcal{W}_6$
$\mathcal{F}_3^{16}$ $x^3 + y^3 + 16z^3 = 0$
Elliptic curve of conductor 36
4
4
3
4
3
$\mathcal{F}_2$ $\simeq$ $\mathcal{W}_2 = X(4)$
$\mathcal{W}_3 = X\left(\Gamma(2) \cap \Gamma_{ns}^{+}(3)\right)$
4
4
3
6
$\mathcal{F}_1$ $\simeq$ $\mathcal{W}_1 = X(2)$
$X_{ns}^{+}(3)$
6
3
$X(1)$
j-invariant
$j = \frac{(\mathfrak{u}^{24}-16)^3}{\mathfrak{u}^{24}}$

Figure 3.9 – Normalized modular tower for W

Chapter 4

Division Polynomials and Isogenies of CM Curves

In previous chapters we have treated isogenies as moduli points. We now focus on a different approach which consists in describing isogenies by the polynomials cutting out their kernels. We start by giving a classification of different models of elliptic curves highlighting the natural level structure they carry. Eventually, we introduce the notion of division polynomials, i.e., those polynomials characterizing the n-torsion subscheme of an elliptic curve, and study their recursion formulæ. Finally, in view of the applications of next chapters, we describe how to adapt this construction to algebraic integers (following [Sat]) and then specialize to the Gaussian end Eisenstein case.

4.1 Models of elliptic curves

An elliptic curve is a projective non-singular genus 1 curve with a fixed distinguished base point O.

In view of a classification of different models for elliptic curves, we present here the structures up to which such a classification is made. To fix the notation, we will work with projective models to avoid divisions, except in the case of Weierstrass curves where working with an affine equations is not too much of a constraint since there is only one point missing, namely the point at infinity. A projective model for an elliptic curve is an embedding $\iota : E \to \mathbb{P}^r$. We will say that the model is projective normal if it is given by a complete linear system, i.e., if its coordinate ring is integrally closed [Har, Ex. II.3.18]. We let $X_0, \ldots, X_r$ denote the coordinate functions on $\mathbb{P}^r$. We obtain a surjective morphism of rings

$$\iota^* : k\left[\mathbb{P}^r\right] = k\left[X_0, \ldots, X_r\right] \longrightarrow k\left[E\right] = \frac{k\left[X_0, \ldots, X_r\right]}{I_E}$$

for I_E the defining ideal for the embedding.

Kohel observed that, using the completeness of the linear system defining the projective model, we can infer the existence of a point $S \in E(k)$ such that every hyperplane intersects E in $d = r+1$ points summing to S (counting their multiplicities) [Koh4].

On $\mathbb{P}^r$ we have a natural invertible sheaf $\mathcal{O}_{\mathbb{P}^r}(1)$ spanned by the global sections which are given by the homogeneous coordinates $\{X_0, \ldots, X_r\}$. Hence, for a projective model $\iota : E \to \mathbb{P}^r$ we obtain a sheaf on E given by the pullback $\mathcal{L} = \iota^*\mathcal{O}_{\mathbb{P}^r}(1)$ which is also invertible and is generated by the global sections $s_i = \iota^*(X_i)$. In the same way one can construct $\mathcal{L}^n = \iota^*\mathcal{O}_{\mathbb{P}^r}(n)$ where $\mathcal{O}_{\mathbb{P}^r}(n)$ is now generated by the monomials of degree n in the X_i's. Its space of global sections $\Gamma(E, \mathcal{L}^n)$ is a finite dimensional k vector space [Har, Ex. II.5.14] and

$$k[E] = \bigoplus_{n=0}^{+\infty} \Gamma(E, \mathcal{L}^n)$$

Now, if D is the divisor on E cut out by $X_0 = 0$, we can identify $\Gamma(E, \mathcal{L}^n)$ with the Riemann-Roch space associated to nD [Koh4]. This means that a projective model is determined by a line bundle $\mathcal{L}(D)$ on the curve which is given by a divisor D. Its space of global sections is the Riemann-Roch space $L(D)$.

Definition. The divisor D encoding the information on the projective model is called the embedding divisor. More precisely, its class is the embedding divisor class.

It turns out that if $\mathcal{L}(D)$ is an invertible sheaf, then D has to be linearly equivalent to $r(O)+(S)$ where S is the sum of all points (with their multiplicity) lying in the intersection of any hyperplane with E, as above. Thus, we will consider $D = r(O) + (S)$ as the embedding divisor [Koh3, Lemma 2].

For the Weierstrass model we can take $D = 3(O)$ which corresponds to $2(O) + (O)$ in the description above. Hence, we are taking the point S to be the identity element itself but, as we will see, this is not always the case. More in general, for quartic models, the only interesting divisors are $3(O) + S$ and $4(O)$. When working with Edwards curves it is natural to take $D = \sum_{i=0}^{3} iT \sim 3(O) + 2(S)$ with $S = 2T$.

Definition. A model is symmetric if $[-1]$ acts projectively linearly.

Remark. This is equivalent to the divisor embedding D being equivalent to $r(O) + (S)$ where $S \in E[2]$ is some 2 torsion point. In fact, if we have an affine model and its divisor is the divisor at infinity, the fact that -1 acts linearly means that S is either the point at infinity or a two torsion point.

Remark. By means of the choice of coordinate functions, $[-1]$ can be diagonalized into eigenspaces, which means that by -1 can be chosen to act by a swap of variables or a change of sign.

Since we are mostly interested in curves with efficient arithmetic, we will restrict to symmetric models. The divisor embedding provides the first invariant in our classification since there exists a linear map between two different models if and only if they have the same embedding divisor class [Koh4, Th 3.2]. A second criterion in the classification would be the level structure encoded by the model: we will be interested in elliptic curves with a given torsion subgroup $G \subset E(k)$; moreover, we want this subgroup to stabilize the embedding divisor D. This means that the pullback by the translation by P-map τ_P acts as the identity on D for all $P \in G$: $\tau_P^* D = D$. As a consequence G will act linearly on E [Koh3, Lemma 29].

Remark. This means that 3-torsion points act linearly on degree 3 models, for degree 4 models 4-torsion points act linearly (including 2 torsion points) and so on.

Finally, we will also look at the choice for the base point and at possible twists of the models.

Level 2 structure

We start by looking at families of elliptic curves parametrizing a 2-level structure. This means that they naturally come equipped with some 2-torsion structure. In particular, we will be interested in quartic models as two torsion acts linearly on these families.

Lemma 4.1 ([Koh3, Lemma 5])**.** *Let E be a projective normal degree d model of an elliptic curve. The translation by T map acts linearly if and only if T is in $E[d]$.*

This give a motivation to focus on quartic models in the framework of symmetric curves for which $[-1]$ acts linearly.

We will start by looking at a model encoding a $\Gamma_0(2)$ structure to eventually climb the modular tower adding a full level 2 structure, a $\Gamma_1(4)$-structure and, finally, a $\Gamma(4)$ structure.

4.1.1 Intersection of two quadrics

It is well known that any elliptic curve over k can be embedded in $\mathbb{P}^3$ as the intersection of two quadrics (in general defining a curve of genus 1). Cassels [CF, Ch. 7] shows that, provided the characteristic of k is not 2 or 3, a general elliptic curve $y^2 = x^3 + ax + b$ corresponds to the curve $\mathcal{C}$ in $\mathbb{P}^3$ defined by

$$\begin{cases} X_2^2 = bX_0^2 + aX_0X_1 + X_3X_1 \\ X_0X_3 = X_1^2 \end{cases}$$

via the embedding $(x, y) \to (1 : x : y : x^2)$. The identity element corresponds to a flex point where the tangent line has a multiplicity 2 intersection point. The group law is described geometrically in [Hus] and arithmetically in [CF].

Remark. Here we are taking the divisor D to be $4(O)$ and a basis for the Riemann-Roch space given by $\{1, x, y, x^2\}$.

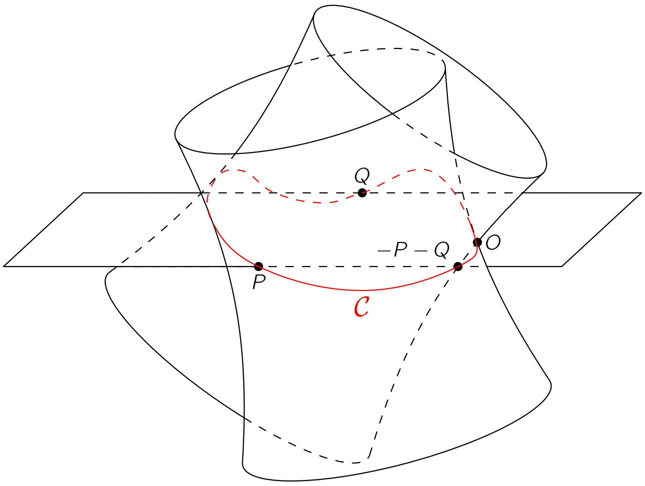


Figure 4.1 – Group law in the intersection of two quadrics

We let O be the identity element and, similarly to the chord-tangent law defined on the affine model of an elliptic curve, we let three points sum to O if they are coplanar to S, which equals O in this particular case. The negation of a point P will be the third intersection of the plane passing through O and P and containing the tangent line to the curve in O (Figure 4.1).

One of the advantages of this model is that the formulæ for adding points also work for doubling and has been therefore used in cryptography to prevent SPA-like attacks as in [BJ1] and [LS].

In practice, we will often be interested in elliptic curves with a 2-torsion point. In this case we can transform a Weierstrass model of E into the form

$$E : y^2 = x(x^2 + \alpha x + \beta)$$

with 2-torsion point $(0, 0)$. An embedding in $\mathbb{P}^3$ as a quartic curve $\mathcal{C}$ is obtained by the transformations

$$(x, y) \longmapsto (1 : x : y : x^2 + \alpha x + \beta)$$

or in projective coordinates by

$$(X : Y : Z) \longmapsto (Z^2 : XZ : YZ : X^2 + \alpha XZ + \beta Z^2)$$

giving defining polynomials for $\mathcal{C}$:

$$\begin{cases} X_0X_3 = \beta X_0^2 + \alpha X_0X_1 + X_1^2 \\ X_1X_3 = X_2^2 \end{cases}$$

Remark. The choice between x^2 and $x^2 + \alpha x + \beta$ as the fourth coordinate is justified by the projection to the last two coordinates:

$$(y : x^2 + \alpha x + \beta) = \big(y^2 : (x^2 + \alpha x + \beta)y\big) = (x : y)$$

versus the lack of cancellation between y and x^2:

$$(y : x^2) = (y^2 : x^2y) = (x^2 + \alpha x + \beta : xy)$$

As a result, the projection to the former two coordinates gives a map of degree 2 whereas the latter is of 3.

As we said above, the interest for looking at a level 2 structure on a quartic model is that the translation by 2-torsion points is linear:

$$(X_0 : X_1 : X_2 : X_3) \longmapsto (X_0 + \alpha/\beta X_1 - 1/\beta X_3 : -X_1 : X_2 : -X_3)$$

Remark. This curve carries a $\Gamma_0(2)$ structure which is given by its 2-torsion point. We can explicitly find the relations by α, β and the parameter s on $X_0(2)$ satisfying $j = (s+16)^3/s$. A simple `magma` computation gives $s = 256\beta/(\alpha^2-4\beta)$. Note that $t = \alpha^2/\beta$ is an invariant and therefore a twist consists in $(\alpha, \beta) \mapsto (\lambda\alpha, \lambda^2\beta)$.

4.1.2 Jacobi normal form

Assuming the characteristic of the base field is not 2, we let $J_{(a,b)}$ be the elliptic curve given by the intersection of the quadrics in $\mathbb{P}^3$

$$\begin{cases} X_2^2 = X_1^2 + aX_0^2 \\ X_3^2 = X_2^2 + bX_0^2 \\ X_1^2 = X_3^2 + cX_0^2 \end{cases}$$

where $a + b + c = 0$, with identity element $(0 : 1 : 1 : 1)$. The full 2-torsion part of $J_{(a,b)}$ is very explicit:

$$J_{(a,b)}[2] = \{(0:1:1:1)\ ,\ (0:-1:1:1)\ ,\ (0:1:-1:1)\ ,\ (0:1:1:-1)\}$$

The embedding divisor is equivalent to the formal sum of all the two torsion points, which is $4(O)$.

Theorem 4.2 ([Koh3, Th. 35]). *Let E/k be an elliptic curve with projective normal embedding in $\mathbb{P}^3$ such that $\mathcal{O}_E(1) \simeq \mathcal{L}(4(O))$ and $E(k)[2]$ is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$, then there exist $a, b \in k$ such that E is linearly isomorphic to $J_{(a,b)}$.*

Remark. The Jacobi curve $J_{(a,b)}$ comes equipped with the full 2-torsion subgroup. Thus, it carries a $\Gamma(2)$ structure. The relation with the λ invariant, which is a parameter for the modular curve $X(2)$, is given by a six-to-one cover [Sil1, §III.1]:

$$\lambda \in \left\{ -\frac{a}{b}, -\frac{b}{a}, \frac{a}{a+b}, \frac{b}{a+b}, -\frac{c}{a}, -\frac{c}{b} \right\}$$

This is perhaps not the most natural model on which to study polynomial maps, since the Kummer curve is a conic in $\mathbb{P}^2$. Indeed, the inversion map is given by

$$[-1](X_0 : X_1 : X_2 : X_3) = (-X_0 : X_1 : X_2 : X_3)$$

so that the Kummer curve is the projection to $(X_1 : X_2 : X_3)$ in $\mathbb{P}^2$ defined by

$$aX^2 + bY^2 + cZ^2 = 0$$

The Jacobi model is isomorphic to a curve in Legendre normal form

$$y^2 = x(x-1)(x-\lambda)$$

for which $\lambda = -1$ is in the family of Gaussian twists. Note that the λ is exactly the λ invariant above. An isomorphisms with the curve $(a, b, c) = (-1, -\lambda + 1, \lambda)$ is given by

$$(x, y) \longmapsto (2y : x^2 - \lambda : x^2 - 2x + \lambda : x^2 - 2\lambda x + \lambda)$$

The Kummer curve of J is parametrized by the Kummer line $\mathbb{P}^1$ of E by the map

$$(X : Z) \longmapsto (X^2 - 2\lambda XZ + \lambda Z^2 : X^2 - \lambda Z^2 : X^2 - 2XZ + \lambda Z^2)$$

where the zeros of the coordinate images are at the x-coordinate of the 4-torsion points.

Remark. The existence of a 4-torsion point shows that this model parametrizes something more than a $\Gamma(2)$-structure: it is indeed a point on $X(\Gamma_1(4) \cap X_0(2))$. Since there is no 4-torsion point over $k(\lambda)$, one need $\mu = \sqrt{\lambda}$ and then the cover $X(\Gamma_1(4) \cap X(2)) \to X(2)$ is given by $\mu \mapsto \lambda = \mu^2$.

Chudnovsky and Chudnovsky [CC] define a Jacobi form representing an elliptic curve as an intersection of two quadrics:

$$E_{J,\lambda} \begin{cases} X_0^2 + X_1^2 = X_3^2 \\ \lambda^2 X_0^2 + X_2^2 = X_3^2 \end{cases}$$

with base point $(0 : 1 : 1 : 1)$, which is a model for a curve in Jacobi normal form with $(a, b, c) = (1, -\lambda^2, \lambda^2 - 1)$ with map $(X_0 : X_1 : X_2 : X_3) \mapsto (X_0 : X_1 : X_3 : X_2)$ [Koh3].

Liardet and Smart [LS] obtained explicit formulæ for point addition on $E_{J,\lambda}$. Let $P_1 = (a_0 : a_1 : a_2 : a_3)$ and $P_2 = (b_0 : b_1 : b_2 : b_3)$ be two points on $E_{J,\lambda}$, then $P_1 + P_2 = P_3 = (c_0 : c_1 : c_2 : c_3)$ with

$$(c_0 : c_1 : c_2 : c_3) = (a_0b_1b_2a_3 + b_0a_1a_2b_3 : a_1b_1a_3b_3 - a_0b_0a_2b_2 : a_2b_2a_3b_3 - \lambda^2 a_0b_0a_1b_1 : (b_1a_3)^2 + (b_0a_2)^2)$$

4.1.3 Jacobi quartic form

Jacobi quartic curves over a field of characteristic different from 2 are defined by the model

$$J_k : y^2 = (1 - x^2)(1 - \kappa^2 x^2) \qquad \text{with } \kappa \neq 0, \pm 1$$

with base point $(0, 1)$ and 2-torsion point $(0, -1)$. These curves where introduced by Jacobi as they can be parameterized with Jacobi's elliptic functions. It is worth noting that its standard projective closure $Y^2Z^2 = \kappa^2X^4 - (\kappa^2 + 1)X^2Z^2 + Z^4$ in $\mathbb{P}^2$ is singular; the coordinates $(X : Y : Z)$ stands for equivalence classes of triplets where $(X_1 : Y_1 : Z_1) \sim (X_2 : Y_2 : Z_2)$ if there exists $t \in k^*$ such that $X_1 = tX_2$, $Y_1 = t^2Y_2$ and $Z_1 = tZ_2$.

As they have 2-torsion points, it is natural to ask if they provide a good parametrization of a $\Gamma_0(2)$-structure. Billet and Joye [BJ1] observed that if $E := y^2 = x^3 + ax + b$ has a rational 2-torsion point $(\theta, 0)$ (indeed any curve with even number of k-rational points), then it can be put in the more general affine form (usually referred to as extended Jacobi quartic form)

$$J_{\delta,\epsilon} : y^2 = \epsilon x^4 + 2\delta x^2 + 1 \tag{4.1}$$

with $\epsilon = -(3\theta^2 + 4a)/16$ and $\delta = -3\theta/4$.

Remark. In this context, "more general" means that it covers more isomorphism classes than the classic Jacobi model. By contrast, we descend the modular tower as J_k is a family with level $\Gamma_1(4) \cap \Gamma(2)$-structure. In fact, if μ is the square root of the λ-invariant introduced above, then $\kappa = (\mu - 1)/(\mu + 1)$.

This curves have identity element $(0, 1)$ and $(0, -1)$ is a point of order 2. Their discriminant is $\Delta = 256(\delta^2 - \epsilon)^2 \neq 0$ and the j invariant is given by $64(\delta^2 + 3\epsilon)^3/(\epsilon(\delta^2 - \epsilon)^2)$.

As for the classical Jacobi model, the projective closure $Y^2Z^2 = \epsilon X^4 + 2\delta X^2Z^2 + Z^4$ has a unique singular point $\Omega = (0 : 1 : 0)$. The blow-up resolving this singularity produces two points at infinity, noted Ω_1 and Ω_2. The divisor at infinity is $D = 2(\Omega_1) + 2(\Omega_2)$, and the Riemann–Roch space associated to it is generated by $\{1, x, y, x^2\}$. Thus the transformation $(x, y) \mapsto (1 : x : y : x^2)$ gives the projective normal closure $\mathcal{E} \subseteq \mathbb{P}^3$ with defining equations [Koh3]

$$\begin{cases} X_1^2 = X_0X_3 \\ X_2^2 = \epsilon X_3^2 + 2\delta X_0X_3 + X_0^2 \end{cases}$$

Liardet and Smart [LS] and Kohel [Koh3] considered the special case $\epsilon = 1$ which gives curves

$$\begin{cases} X_1^2 = X_0X_3 \\ X_2^2 = X_3^2 + 2\delta X_0X_3 + X_0^2 \end{cases}$$

The map $(X_0 : X_1 : X_2 : X_3) \mapsto (X_1 : X_2 : X_0 - X_3 : X_0 + X_3)$ defines an isomorphism to the Jacobi normal model defined in the previous section with $(a, b, c) = (-2(\delta + 1), 4, 2(\delta - 1))$.

The geometric interpretation of the addition law goes as shown in Figure 4.2 and it is the analogue of the chord-and-tangent rule for Weierstrass elliptic curves. The conic $\mathcal{C}$ passing through P_1, P_2 and O and tangent to $J_{\delta,\epsilon}$ at Ω has 8 intersections with our Jacobi quartic; this is a direct application of Bezout's Theorem as the number of intersection of a conic with a degree 4 model is $2 \cdot 4$. Since, it has multiplicity 4 at Ω and 3 intersections P_1, P_2 and O, it follows that $\mathcal{C}$ passes through a 4 point of the elliptic curve. We let P_1, P_2 and this third point to sum to O meaning that $\mathcal{C}$ crosses $J_{\delta,\epsilon}$ at $-P_3 = -(P_1 + P_2)$.

Remark. The negation of a point is $-(x, y) = (-x, y)$.

Proposition 4.3. *Given two points $P_1 = (a_0 : a_1 : a_2)$ and $P_2 = (b_0 : b_1 : b_2)$ on the Jacobi quartic $J_{\delta,\epsilon} : Y^2Z^2 = \epsilon X^4 + 2\delta X^2Z^2 + Z^4$, the conic $\mathcal{C}$ passing through P_1, P_2, O and tangent to $J_{\delta,\epsilon}$ at Ω has*

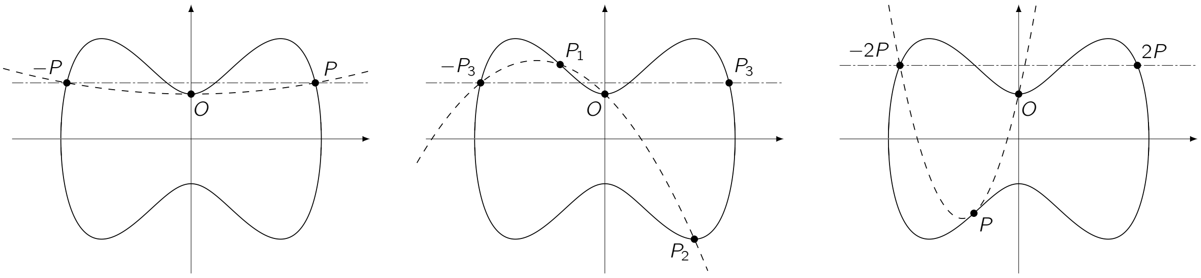


Figure 4.2 – Negation of a point, addition and doubling for Jacobi quartics

equation

$$\mathcal{C} : c_{X^2}X^2 + c_{YZ}(Z^2 - YZ) + c_{XZ}XZ = 0$$

where the coefficients $(c_{X^2} : c_{YZ} : c_{XZ}) \in \mathbb{P}^2$ *are given by*

I. *If* $P_1 \neq P_2$ *and* $P_1, P_2 \neq O$, *then*

$$\begin{cases} c_{X^2} = (a_2^2 - a_1a_2)(b_0b_2) - (b_2^2 - b_1b_2)(a_0a_2) \\ c_{YZ} = b_0^2a_0a_2 - a_0^2b_0b_2 \\ c_{XZ} = a_0^2(b_2^2 - b_1b_2) - b_0^2(a_2^2 - a_1a_2) \end{cases}$$

II. *if* $P_1 = P_2$, *then*

$$\begin{cases} c_{X^2} = 4a_2(\delta a_0^2 + a_2^2) - 2a_1a_2(a_1 + a_2) \\ c_{YZ} = -2a_0^2a_1 \\ c_{XZ} = -4\delta a_0^3 - 4a_0a_2^2 + 4a_0a_1a_2 \end{cases}$$

Remark. There exist different interpretations. One may refer, for instance to [Wan+, Th. 1] for a description in terms of cubics.

Corollary 4.4 ([His, §5.2.1])**.** *Assuming that* $(a_0 : a_1 : a_2) + (b_0 : b_1 : b_2) = (c_0 : c_1 : c_2)$ *we have*

$$\begin{cases} c_0 = (a_0b_1 - a_1b_0)(a_0^2b_2^2 - a_2^2b_0^2) \\ c_1 = (a_1b_1 - 2\delta a_0b_0)(a_0^2b_2^2 + a_2^2b_0^2) - 2a_0b_0(a_2^2b_2^2 + \delta a_0^2b_0^2) \\ c_2 = a_2b_2(a_0b_1 - a_1b_0)^2 \end{cases}$$

Further, if $2(a_0 : a_1 : a_2) = (c_0 : c_1 : c_2)$, *then*

$$\begin{cases} c_0 = 2a_0a_1(2a_2^2 - a_1^2 + 2\delta a_0^2) \\ c_1 = 2a_1^2(a_1^2 - 2\delta a_0^2) - (2a_2^2 - a_1^2 + 2\delta a_0^2)^2 \\ c_2 = (2a_2^2 - a_1^2 + 2\delta a_0^2)^2 \end{cases}$$

One could find the affine unified version of the addition law in [His+].

Remark. After embedding in $\mathbb{P}^3$ the cubic becomes an hyperplane and we recover the general description of Fig. 4.1, see [His, §5.2.2].

4.1.4 Edwards curves

In 2007 Edwards [Edw] introduced another quartic model for elliptic curves over a field of characteristic different from 2 and observed that formulæ for addition have interesting properties:

$$E_c : x^2 + y^2 = c^2\left(1 + x^2y^2\right)$$

However, elliptic curve over k can all be transformed into Edwards form only if k is algebraically closed. Over finite fields, only $\sim 1/4$ of isomorphism classes admits such a transformation to E_c for some c. Subsequently, Bernstein and Lange [BL] considered a twisted family of Edwards curves

$$E_d : x^2 + y^2 = 1 + dx^2y^2$$

Remark. Edwards curves define a family over $k(c) = k(X(4))$. The choice of Bernstein and Lange permits one to descend from $X(4)$ to $X_1(4)$ with $d = c^4$. This motivates the *twist* in the name as the image of $X(4)(k) \to X_1(4)(k)$ lies in a subset and therefore the choice of such a re-scaling allows one to cover a larger family of curves.

We can observe that E_d has a distinguished point $O = (0, 1)$ which plays the role of the identity element, and a point of order 2, $S = (0, -1)$. The points $(\pm 1, 0)$ have order 4 and they can be used to construct maps from Edwards curves to Weierstrass curves and to show that precisely those curves with a rational 4-torsion point are birationally equivalent to an Edwards curve.

Contrary to the prior models the divisor embedding is $3(O) + (S)$ and the Riemann-Roch space is then spanned by $\{1, x, y, xy\}$. The embedding $(x, y) \mapsto (1 : x : y : xy)$ defines the projective normal closure of E_d in $\mathbb{P}^3$ with identity $O = (1 : 0 : 1 : 0)$:

$$\begin{cases} X_0^2 + dX_3^2 = X_1^2 + X_2^2 \\ X_0X_3 = X_1X_2 \end{cases}$$

The addition law on Edwards curves has a geometric interpretation once again described in terms of divisors [Are+] and chord-and-tangent rule: given two points P_1 and P_2 there is a conic $\mathcal{C}$ passing through the 5 points $P_1, P_2, S, \Omega_1 = (1 : 0 : 0)$ and $\Omega_2 = (0 : 1 : 0)$ (the 2 points at infinity with multiplicity 2); the remaining intersection $\mathcal{C} \cap E_d$ is the point $-P_3$ and the associated divisors yield the equation $P_3 = P_1 + P_2$.

Proposition 4.5 ([Are+, Th. 1])**.** *Let E_d be an Edwards curve over k and let $P_1 = (a_0 : a_1 : a_2)$ and $P_2 = (b_0 : b_1 : b_2)$ be 2 affine points on E_d not necessarily distinct. Let $\mathcal{C}$ be the conic*

$$\mathcal{C} : c_{Z^2}(Z^2 + YZ) + c_{XY}XY + c_{XZ}XZ = 0$$

passing through P_1, P_2, S, Ω_1 and Ω_2. The coefficients of $\mathcal{C}$ are given by

I. *If $P_1 \neq P_2$, $P_1, P_2 \neq S$, then*

$$\begin{cases} c_{Z^2} = a_0b_0(a_1b_2 - b_1a_2) \\ c_{XY} = a_2b_2(a_0b_2 - b_0a_2 + a_0b_1 - b_0a_1) \\ c_{XZ} = b_0b_1a_2^2 - a_0a_1b_2^2 + a_1b_1(b_0a_2 - a_0b_2). \end{cases}$$

II. *If $P_1 \neq P_2 = S$, then $C_{Z^2} = -a_0$, $c_{XY} = a_2$, $c_{XZ} = a_2$.*

III. *If $P_1 = P_2$, then*

$$\begin{cases} c_{Z^2} = a_0a_2(a_2 - a_1) \\ c_{XY} = da_0^2a_1 - a_2^3 \\ c_{XZ} = a_2(a_2a_1 - a_0^2) \end{cases}$$

Proposition 4.6 ([His, §5.1.1])**.** *Let E_d be an Edwards curve over k. If $(a_0 : a_1 : a_2) + (b_0 : b_1 : b_2) = (c_0 : c_1 : c_2)$, then*

$$\begin{cases} c_0 = (a_0b_1 - a_1b_0)(a_0a_1b_2^2 + b_0b_1a_2^2) \\ c_1 = (a_1b_1 + a_0b_0)(a_0a_1b_2^2 - b_0b_1a_2^2) \\ c_2 = a_2b_2(a_0b_1 - a_1b_0)(a_1b_1 + a_0b_0) \end{cases}$$

Further, if $2(a_0 : a_1 : a_2) = (c_0 : c_1 : c_2)$, then

$$\begin{cases} c_0 = 2a_0a_1(2a_2^2 - a_1^2 - a_0^2) \\ c_1 = (a_1^2 - a_0^2)(a_1^2 + a_0^2) \\ c_2 = (a_1^2 + a_0^2)(2a_2^2 - a_1^2 - a_0^2) \end{cases}$$

The affine version of these addition formulæ can be found in [BL].

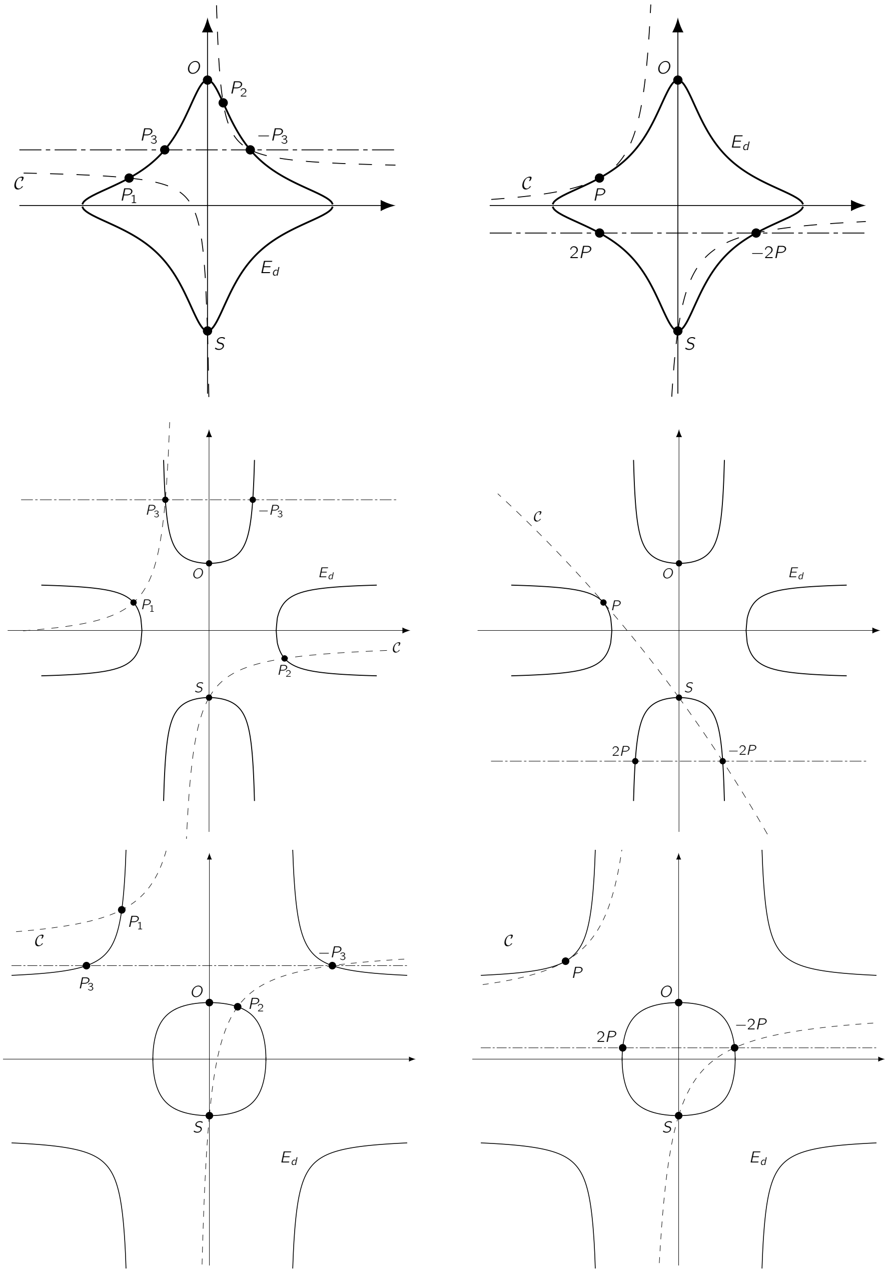


Figure 4.3 – Group law and doubling for Edwards curves

Remark. We can exploit addition laws as sum of points on a hyperplane in $\mathbb{P}^3$ following Hisil [HWC]: using the map $(x, y) \mapsto (1 : x : y : xy)$ described above, one could pass to the projective normal closure with identity element $(1 : 0 : 1 : 0)$. If $(a_0 : a_1 : a_2 : a_3) + (b_0 : b_1 : b_2 : b_3) = (c_0 : c_1 : c_2 : c_3)$, then

$$\begin{cases} c_0 = (a_0b_0 + da_3b_3)(a_0b_0 - da_3b_3) \\ c_1 = (a_0b_0 - da_3b_3)(a_1b_2 + a_2b_1) \\ c_2 = (a_0b_0 + da_3b_3)(a_2b_2 - a_1b_1) \\ c_3 = (a_1b_2 + a_2b_1)(a_2b_2 - a_1b_1) \end{cases}$$

Remark. Bernstein and Lange [BL] also defined twisted Edward curves as the 2 dimensional family

$$E_{a,d} : ax^2 + y^2 = 1 + dx^2y^2$$

These enlarge the family to include quadratic twists of Edwards curves.

4.1.5 Huff model

In 1948 Huff [Huf] introduced the curves

$$H_{a,b} : ax(1-y^2) = by(1-x^2) \qquad a, b \in k,\ a^2 - b^2 \neq 0$$

while studying a Diophantine problem. These curves have identity element $O = (0, 0)$, 3 points at infinity, namely $(1 : 0 : 0)$, $(0 : 1 : 0)$ and $(a : b : 0)$, which are exactly the 3 primitive 2-torsion points of E and 4 torsion points $(\pm 1 : \pm 1 : 1)$.

Remark. Huff curves are a family on $X(\Gamma(2) \cap \Gamma_1(4))$ parametrized by b/a. A descent to $X(2)$ is given by considering $ax(y^2 - d) = by(x^2 - d)$. On the other hand, the base extension $a^2 = b^2 - c^2$ gives a full level 4 structure $X(4) \simeq \mathbb{P}^1$, with isomorhism given by the parametrization of the conic

$$\begin{aligned} \mathbb{P}^1 &\longrightarrow X(4) : a^2 - b^2 = c^2 \\ (u : v) &\longmapsto (u^2 + v^2 : u^2 - v^2 : 2uv) \end{aligned}$$

These are not the ideal models to work with since, as of degree 3, we lose the interesting property of 4-torsion acting linearly. However, we mention here the equivalence with other known families.

Huff showed that $H_{a,b}/\mathbb{P}^2 : aX_1(X_0^2 - X_2^2) = bX_2(X_0^2 - X_1^2)$ is isomorphic to the Weierstrass curve

$$Y^2Z = X(X + a^2Z)(X + b^2Z)$$

via the transformations

$$\begin{aligned} \Upsilon &:(X_0 : X_1 : X_2) \mapsto (X : Y : Z) = \big(ab(aX_2 - bX_1) : ab(a^2 - b^2)X_0 : aX_1 - bX_2\big) \\ \Upsilon^{-1} &:(U : V : W) \mapsto (X_0 : X_1 : X_2) = \big(Y : b(X + a^2Z) : a(X + b^2Z)\big) \end{aligned}$$

There is an isomorphism between $H_{a,b}$ and the Jacobi normal curve

$$J_{(a^2, b^2 - a^2)} : \begin{cases} Z^2 = Y^2 + a^2X^2 \\ T^2 = Y^2 + (b^2 - a^2)X^2 \\ Y^2 = T^2 - b^2X^2 \end{cases}$$

via the map

$$(X : Y : Z : T) \longmapsto (X_0 : X_1 : X_2) = \big(a^2b^2(Z - T) : (a^2 - b^2)a^2b^2X : -(a^2 - b^2)Y - b^2Z + a^2T\big)$$

Further, we can describe an isomorphism with the twisted Edwards curve

$$E_{(a-b)/(a+b)} : X^2 - Z^2 = Y^2 - \left(\frac{a-b}{a+b}\right)^2 T^2 \qquad c = \frac{a-b}{a+b}$$

via the transformation

$$(X : Y : Z : T) \longmapsto (X_0 : X_1 : X_2) = \big(X + Z : Y - c^{-1}T : Y + c^{-1}T\big)$$

Level 3 structure

We focus now on degree 3 models on which the 3-torsion acts linearly.

4.1.6 Hessian curves

Contrary to previous models, Hessian curves parametrize elliptic curves with 3 torsion structure. We suppose the characteristic of the field is different form 2 and 3, and we used the distinguished 3-torsion point to write our curve in the form

$$H_d : X^3 + Y^3 + Z^3 = dXYZ$$

with $d \in k$ such that $d^3 \neq 1$.

Proposition 4.7 ([JQ, Prop. 1]). *H_d is isomorphic to the Weierstrass elliptic curve*

$$y^2 = x^3 - \frac{d(d^3+216)}{48}x + \frac{d^6 - 540d^3 - 5832}{864}$$

with discriminant and j-invariant given by

$$\Delta(H_d) = (d^3-27)^3 \neq 0 \qquad j(H_d) = \frac{d^3(d^3+216)^3}{(d^3-27)^3}$$

Remark. We obtain $d = t + 3$ for t the Hesse invariant on $X(3)$, see Section 3.3.3. Thus, the Hessian family parametrizes elliptic curves with full level 3-structure.

The identity element can be taken to be the point at infinity $O = (1 : -1 : 0)$ and the points $(0 : 1 : -1)$ and $(1 : 0 : -1)$ are two points of order 3.

We can interpret the group law on these curves geometrically as we did for the Weierstrass model, namely letting three colinear points sum to O. Note that this means that, as divisors, they are equivalent to $3(O)$: we consider two points P_1 and P_2 and the line r passing through them. The divisor of r is $\text{div}(r) = (P_1) + (P_2) + (-P_3) - 3(O)$. If r' is the line passing through the points $-P_3$ and P_3 then $\text{div}(r') = (P_3) + (-P_3) - 2(O)$ and this gives $P_3 = P_1 + P_2$.

Remark. The inverse of the point $(X : Y : Z)$ is the symmetric with respect to the first diagonal $-(X : Y : Z) = (Y : X : Z)$.

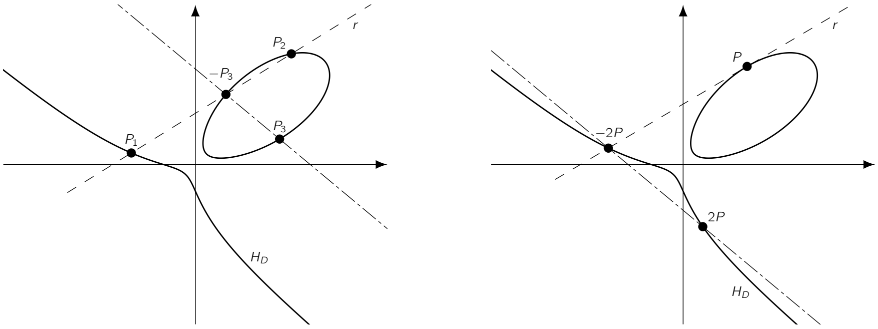


Figure 4.4 – Group law and doubling for Hessian curve H_D.

Proposition 4.8 ([GGX]). *Let $P_1 = (X_1 : Y_1 : Z_1)$ and $P_2 = (X_2 : Y_2 : Z_2)$ be two affine points on H_d. Define $P_3 = P_1 + P_2$. The line r passing through P_1, P_2 and $-P_3$ has equation*

$$c_X X + c_Y Y + c_Z Z = 0$$

where the coefficients are given by the following

I. *If $P_1 \neq P_2$, then*

$$\begin{cases} c_X = Y_1Z_2 - Z_1Y_2 \\ c_Y = Z_1X_2 - X_1Z_2 \\ c_Z = X_1Y_2 - Y_1X_2 \end{cases}$$

I. *If $P_1 = P_2$, then the tangent line has coefficients*

$$\begin{cases} c_X = 3X_1^2 - dY_1Z_1 \\ c_Y = 3Y_1^2 - dX_1Z_1 \\ c_Z = 3Z_1^2 - dX_1Y_1 \end{cases}$$

The explicit formulæ for addition and point doubling can be found using Cauchy-Desboves's formulæ [JQ].

Corollary 4.9. *Let $P_1 = (a_0 : a_1 : a_2)$ and $P_2 = (b_0 : b_1 : b_2)$ be two points on H_d, then*

$$P_1 + P_2 = P_3 = (c_0 : c_1 : c_2) \qquad \textit{where} \qquad \begin{cases} c_0 = a_1^2 b_0 b_2 - b_1^2 a_0 a_2 \\ c_1 = a_0^2 b_1 b_2 - b_0^2 a_1 a_2 \\ c_2 = a_2^2 b_1 b_0 - b_2^2 a_1 a_0 \end{cases}$$

$$2P_1 = P_3 = (c_0 : c_1 : c_2) \qquad \textit{where} \qquad \begin{cases} c_0 = a_1(a_2^3 - a_0^3) \\ c_1 = a_0(a_1^3 - a_2^3) \\ c_2 = a_2(a_0^3 - a_1^3) \end{cases}$$

Remark. In the literature, the definition of Hessian curve could present small differences in particular in the choice of the coefficient of the term XYZ which Hesse defined to be $6d$, while Joye and Quisquater [JQ] set as $3d$.

Remark. In [Ber+] the authors define a twisted Hessian curves $aX^3 + Y^3 + Z^3 = dXYZ$. It is worth noting that they chose a different base point element, namely $(0 : -1 : 1)$.

Remark. As we said, having two points of order 3, Hessian curves have a natural $\Gamma(3)$ structure. Twisted Hessian curves are more general in the sense that they represent a family over $k(X_0(3))$ and therefore cover a larger family of curves while parametrizing a lower level structure.

4.2 Division polynomials

The group law on an elliptic curve defines a morphism of varieties $E \times E \longrightarrow E$ for which we gave explicit rational functions and polynomials describing it.

Using these polynomials we can, in a natural way, construct the following maps extending to morphisms:

$$\begin{aligned} [n] : E &\longrightarrow E \\ P &\longrightarrow P + \ldots + P \end{aligned}$$

associating to a point P the sum of P with itself n times. The map $[n]$ is an endomorphism in $\mathrm{End}_k(E)$ and is therefore given by rational functions $\left(\frac{P_x(x,y)}{Q_x(x,y)}, \frac{P_y(x,y)}{Q_y(x,y)}\right)$. Recursively applying the addition law formulæ and the associativity of the group law, we can show that P_x, P_y, Q_x, Q_y verify stronger constraints [Sil1, Ex. 3.7]. In particular we obtain relatively prime polynomials ϕ_n, ψ_n and ω_n in $k[x, y]$ such that

$$\begin{aligned} [n] : E &\longrightarrow E \\ (x, y) &\longrightarrow (x_n, y_n) = \left(\frac{\phi_n(x, y)}{\psi_n(x, y)^2}, \frac{\omega_n(x, y)}{\psi_n(x, y)^3}\right) \end{aligned}$$

An alternative proof of this, using complex analysis and the Weierstrass $\wp$-function, could be found in [Was].

Definition. The polynomials ϕ_n, ψ_n and ω_n are the n-th division polynomials on E.

For its distinguished role, ψ_n is usually referred to as the n-th division polynomial since it cuts out the closed subscheme $E[n] \setminus \{O\}$ of degree $(n^2 - 1)$ on E, i.e., its roots correspond to the affine coordinates of points of order dividing n [Koh1]. These polynomials satisfy many recursion formulæ and can be related one another; in particular, we can express ϕ_n and ω_n in terms of ψ_n.

We present now explicit relations for these polynomials starting from an elliptic curve E in Weierstrass form. One of the reasons behind using this model is to lower the degree of ψ_n but the drawback is the appearance of the variable y for even n. Formulæ for elliptic curves in short Weierstrass form can be found in [Lan1] and [Sil1]; Morain [Mor] gives relations for ϕ_n and ω_n while Kohel [Koh1] and Lercier [Ler] deal with the general case.

$$\begin{aligned}
&\psi_0 = 0\\
&\psi_1 = 1\\
&\psi_2 = 2y + a_1x + a_3\\
&\psi_3 = 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8\\
&\psi_4 = \psi_2\left(2x^6 + b_2x^5 + 5b_4x^4 + 10b_6x^3 + 10b_8x^2 + (b_2b_8 - b_4b_6)x + b_4b_8 - b_6^2\right)\\
&\psi_{2m+1} = \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3 \qquad \text{for } m \geq 2\\
&\psi_{2m} = \psi_m\left(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2\right)/\psi_2 \qquad \text{for } m > 2
\end{aligned}$$

Remark. It can be shown that the numerator $\psi_m\left(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2\right)$ in the definition of ψ_{2m} is divisible by ψ_2^2.

Division polynomials are computed modulo the equation of the curve and we can therefore assume that the degree of the variable y never exceeds 1. Now, using the group law relation, we find [Cas1, Formulary]

$$\phi_r\psi_m^2 - \phi_m\psi_r^2 = \psi_{m-r}\psi_{m+r}$$

and this yields

$$\phi_0 = 1 \qquad \phi_1 = x \qquad \phi_n = x\psi_n - \psi_{n-1}\psi_{n+1}$$

Finally, in characteristic different from 2 we get

$$\omega_n = \frac{\psi_n + 2\psi_{n-1}^2 - \psi_{n-2}\psi_{n+1}^2}{2\psi_2} - \frac{(a_1\phi_n + a_3\psi_n^2)\psi_n}{2}$$

or

$$2\psi_n\omega_n = \psi_{2n} - (a_1\phi_n + a_3\psi_n^2)\psi_n^2$$

which comes from the action of the endomorphism $[n]$ on the invariant differential of E.

Lemma 4.10. *ψ_n and ϕ_n are polynomials in $\mathbb{Z}[x, \psi_2^2, \{a_i\}]$ if n is odd and in $\psi_2\mathbb{Z}[x, \psi_2^2, \{a_i\}]$ if n is even. Using the equation of the curve E, i.e., observing that these are polynomials in the set of global sections on the sheaf $\mathcal{O}_E$, we get $\psi_2^2 \in \mathbb{Z}[\{a_i\}, x]$ and therefore*

$$\phi_n, \psi_2^{-1}\psi_n \in \mathbb{Z}[\{a_i\}][x] \quad \textit{for all } n \in \mathbb{N}$$

We note in particular that on the Kummer line $\mathbb{P}^1$, determined by the projection $\pi : E \to \mathbb{P}^1$, the morphism $\pi \circ [n]$ is given by $(x : y : 1) \mapsto (\phi_n(x) : \psi_n(x, y)^2)$, inducing $\mathbb{P}^1 \to \mathbb{P}^1$, noting that $\psi_n(x, y)^2 = \psi_n^2(x)$ is a polynomial in x.

4.2.1 Division polynomials on the Kummer curve

Let $P \in E[n]$ be a n-torsion point and $x(P) \in \mathcal{K}_E$ the corresponding image on the Kummer line (see section 1.2.2). The torsion (or kernel) polynomial of $G = \langle P \rangle$ is given by

$$\psi_G(x) = \prod_{m \in \frac{(\mathbb{Z}/n\mathbb{Z})\setminus\{0\}}{\{\pm 1\}}} (x - [m](x(P)))$$

Note that $[m](x(P)) = x([m]P)$ is well defined as the multiplication by m map on the elliptic curve induces a map on the Kummer line (that we call $[m]$ as well).

$$\begin{tikzcd}
E \arrow[r, "{[m]}"] \arrow[d, "x"'] & E \arrow[d, "x"] \\
\mathcal{K}_E \arrow[r, "{[m]}"] & \mathcal{K}_E
\end{tikzcd}$$

Division polynomials are usually computed by recursion formulæ, as pointed out before, but this comes down to Kummer scalar multiplication where we take $x(P), x([m]P), x([m-1]P)$ and recover $x([m+1]P)$; this is a specialization of the differential addition law, see Algorithm 1.

Remark. Montgomery ladders ([Mon]) and Joye ladders ([Joy], [BJ2]) are algorithms performing scalar multiplication of group elements which are efficient on elliptic curves and Kummer lines [CS3].

Kummer scalar multiplication has a nice interpretation in terms of level structure: points on $X_1(n)$ parametrize elliptic curves with an n-torsion point P. Since (E,P) and $(E,-P)$ are in the same equivalence class, we can identify each point on $X_1(n)$ by $(E,x(P))$ and the map described above is just the descending map down to $X_0(n)$. The Galois group acting on the points of $X_1(n)$ is $(\mathbb{Z}/n\mathbb{Z})^\times/\{\pm1\}$ and correspond to the Kummer scalar multiplication.

$$\begin{tikzcd}
X_1(n) \arrow[d, dash, "(Z/nZ)^\times/\{\pm1\}"'] & {[E,P]=[E,x(P)]} \\
X_0(n) \arrow[d, dash] & {[E,G]=[E,\psi_G]} \\
X(1) & {[E]}
\end{tikzcd}$$

The Galois group is given by the
Kummer scalar multiplication
$[n]:(E,x(P))\mapsto(E,x([n]P))$

4.2.2 Quartic division polynomials

Let $E: y^2 = x(x^2+ax+b)$ be an elliptic curve with a two torsion point $(0,0)$.

Applying the embedding $\iota: E \hookrightarrow \mathbb{P}^3$ given on projective coordinates by $(x:y:1)\mapsto(1:x:y:x^2+ax+b)$, and denoting $\mathcal{C}$ its image, we obtain a commutative diagram:

$$\begin{tikzcd}
E \arrow[r, "{[n]}"] \arrow[d, "\iota"'] & E \arrow[d, "\iota"] \\
\mathcal{C} \arrow[r, "{[n]}"] & \mathcal{C}
\end{tikzcd}$$

which determines the division polynomials for $[n]:\mathcal{C}\to\mathcal{C}$:

$$(1:x:y:z)\longmapsto\left(1:\frac{\phi_n(x)}{\psi_n(x,y)^2}:\frac{\omega_n(x,y)}{\psi_n(x,y)^3}:\frac{\phi_n(x)^2+a\phi_n(x)\psi_n(x,y)^2+b\psi_n(x,y)^4}{\psi_n(x,y)^4}\right)$$

This gives the division polynomials for the level-2 quartic model as:

$$\left(\psi_n(x,y)^4:\phi_n(x)\psi_n(x,y)^2:\omega_n(x,y)\psi_n(x,y):\xi_n(x)\right)$$

where $\xi(x)=\phi_n(x)^2+a\phi_n(x)\psi_n(x,y)^2+b\psi_n(x,y)^4=\omega_n(x,y)^2/\phi_n(x)$. This shows that the division polynomials associated to the quartic model $\mathcal{C}$ are parametrized by the usual division polynomials (ψ_n,ϕ_n,ω_n) for the Weierstrass model.

Projection to the last two coordinates gives $(y:z)=(x:y)$. Since by construction x and y have a common zero, which implies that $y/x=z/y$ is a function of degree 2, we can ask if there is a natural simplification or recursion for the division polynomial projections

$$(\omega_n\psi_n:\xi_n)=(\phi_n\psi_n:\omega_n)$$

More in general, anytime we define a map from our curve to its Kummer line we can try to compute division polynomials there and then lift them back on our curve (see section 4.2.1). As we said in Section 1.2.2, the group law on an elliptic curve does not induce a group law on the Kummer line $\mathcal{K}_E$. Still, given two points P and Q, we can compute $P+Q$ once we know $P-Q$. For the generic Jacobi quartic $J_{a,b}$ we can use the map $(x,y)\mapsto y+x^2$ as done in [DKW1] or $(y+1)/x^2$ as proposed in [CV].

Moody [Moo2] used another technique to get recursion formulæ for Jacobi quartics. This comes from a similar approach introduced by Gauss while studying addition formulæon the lemniscate. Suppose we have $P=(x_P,y_P)$ and $Q=(x_Q,y_Q)$ two points on $J_{\delta,\epsilon}$. We let $(x_+,y_+)=(x_P,y_P)+(x_Q,y_Q)=P+Q$ and

$(x_-, y_-) = (x_P, y_P) - (x_Q, y_Q) = P - Q$. Using the addition formulæ 4.4 we obtain

$$x_+ + x_- = \frac{2x_P y_Q}{1-\epsilon(x_P x_Q)^2} \quad \text{and} \quad y_+ + y_- = \frac{2y_P y_Q(1+\epsilon(x_P x_Q)^2)}{(1-\epsilon(x_P x_Q)^2)^2}$$

We can now assume that $P = (x_n, y_n) = [n]Q = [n](x, y)$ and that we have already computed the coordinates of all integer multiples of Q up to n. Thus,

$$x_{n+1} + x_{n-1} = \frac{2x_n y}{1-\epsilon(x x_n)^2} \quad \text{and} \quad y_{n+1} + y_{n-1} = \frac{2y_n y(1+\epsilon(x_n x)^2)}{(1-\epsilon(x x_n)^2)^2}$$

and this permits to recover new recursion formulæ for Jacobi quartics.

Theorem 4.11 ([Moo2, Th. 3 and Lemma 1]). *Let $J_{\delta,\epsilon} : y^2 = h(x) = \epsilon x^4 - 2\delta x^2 + 1$ be a Jacobi quartic curve and $P = (x, y)$ a point on it. If (x_n, y_n) indicates the coordinates of the point $[n]P$ we have the following recursion formulæ:*

$$(x_n, y_n) = \left(xy\frac{f_n(x)}{g_n(x)}, \frac{p_n(x)}{g_n(x)^2} \right)$$

where $f_1 = 1, f_2 = -2, g_1 = 1, g_2 = \epsilon x^4 - 1$ and, for all $n > 1$,

$$f_{2n} = \frac{f_{2n-1}^2 - g_{2n-1}^2}{hf_{2n-2}^2} \quad \text{and} \quad f_{2n+1} = \frac{hf_{2n}^2 - g_{2n}^2}{f_{2n-1}^2}$$

$$g_{2n} = \frac{g_{2n-1}^2 - \epsilon x^4 f_{2n-1}^2}{f_{2n-2}^2} \quad \text{and} \quad g_{2n+1} = \frac{g_{2n}^2 - \epsilon h x^4 f_{2n}^2}{g_{2n-1}^2}$$

$$p_{2n} = \frac{2hp_{2n-1}(g_{2n-1}^2 + \epsilon x^4 f_{2n-1}^2) - p_{2n-2}g_{2n}^2}{g_{2n-2}^2} \quad \text{and} \quad p_{2n+1} = \frac{2p_{2n}(g_{2n}^2 + \epsilon x^4 hf_{2n}^2) - p_{2n-1}g_{2n+1}^2}{g_{2n-1}^2}$$

4.2.3 Jacobi division polynomials

The isomorphism $E \to J \subset \mathbb{P}^3$ is given by affine parametrization (see section 4.1.2)

$$(x : y : 1) \longmapsto (2y : x^2 - \lambda : x^2 - 2x + \lambda : x^2 - 2\lambda x + \lambda)$$

with inverse

$$(X_0 : X_1 : X_2 : X_3) \longmapsto (\lambda(X_2 - X_3) : \lambda(\lambda-1)X_0 : -(\lambda-1)X_1 + \lambda X_2 - X_3)$$

Recalling that the identity point on J is $(0 : 1 : 1 : 1)$, the vanishing of the first coordinate defines the kernel. The morphism $[n]$ is parametrized by the division polynomials for E:

$$(2\omega_n\psi_n : \phi_n^2 - \lambda\psi_n^4 : \phi_n^2 - 2\phi_n\psi_n^2 + \lambda\psi_n^4 : \phi_n^2 - 2\lambda\phi_n\psi_n^2 + \lambda\psi_n^4)$$

The same recursions apply to define the parametrizing triples $(\psi_n, \phi_n, \omega_n)$, however we need to supply their initialization in terms of $(1 : x_1 : x_2 : x_3)$ on the affine Jacobian normal form:

$$\begin{cases} x_1^2 - x_2^2 = 1 \\ x_1^2 - x_3^2 = \lambda \\ x_2^2 - x_3^2 = \lambda - 1 \end{cases}$$

Remark. This model is the Jacobi normal form of the "generic" Legendre curve. The case $j = 12^3$ corresponds to $\lambda = -1$, the Legendre curve $y^2 = x(x^2-1)$. In order to cover the quadratic twists $y^2 = x(x^2-u)$, the generic family can be twisted by u as well (setting $\lambda u = v$):

$$\begin{cases} x_1^2 - x_2^2 = u \\ x_1^2 - x_3^2 = v \\ x_2^2 - x_3^2 = v - u \end{cases}$$

isomorphic to the 2-parameter family curve $y^2 = x(x-u)(x-v)$, in which $v = -u$ is the twisted curves with $j = 1728$.

4.2.4 Edwards division polynomials

The Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ is isomorphic to the elliptic curve in short Weierstrass form

$$W : v^2 = u^3 - \frac{1+14d+d^2}{48}u - \frac{1-33d-33d^2+d^3}{864}$$

via the transformations

$$(x,y) \longmapsto \left(\frac{(5-d)+(1-5d)y}{12(1-y)}, \frac{(1-d)(1+y)}{4x(1-y)}\right) \qquad (u,v) \longmapsto \left(\frac{6u-(1+d)}{6v}, \frac{12u+d-5}{12u+1-5d}\right)$$

The composition with the projective normal closure $(x,y) \mapsto (1:x:y:xy)$ induces the $[n]$ morphism parametrized by the usual division polynomials. In [MHM], the authors deduce recursion formulæ directly from this isomorphism.

As we did in previous sections, we can work out different recursion formulæ: given two points $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$, we let $(x_+, y_+) = (x_P, y_P) + (x_Q, y_Q) = P + Q$ and $(x_-, y_-) = (x_P, y_P) - (x_Q, y_Q) = P - Q$. Using the addition formulæ 4.6, we obtain

$$x_+ + x_- = \frac{2x_Py_Q(1-dx_Q^2)}{(1-dx_P^2x_Q^2)} \qquad \text{and} \qquad y_+ + y_- = \frac{2(1-d)y_Py_Q}{1-d(y_P^2+y_Q^2)+dy_P^2y_Q^2}$$

The y-map associating to a point its y-coordinate is a quotient to the Kummer line and induces the following recursion formulæ

Proposition 4.12. *Let (x_n, y_n) be the coordinates of $[n](x,y)$. Then*

$$y_n = \begin{cases} \dfrac{P_n(y^2)}{Q_n(y^2)} & \text{if } n \text{ is even} \\ y\dfrac{P_n(y^2)}{Q_n(y^2)} & \text{if } n \text{ is odd} \end{cases}$$

where P_n, Q_n are integral polynomials defined by recursion:

$$P_1(t) = 1 \qquad Q_1(t) = 1 \qquad P_2(t) = -1 + 2t - dt^2 \qquad Q_2(t) = 1 - 2dt + dt^2$$

$$P_{n+1}(t) = \begin{cases} 2(1-d)Q_{n-1}P_nQ_n - P_{n-1}(Q_n^2 - dP_n^2 + dt(P_n^2 - Q_n^2)) & \text{if } n \text{ is even} \\ 2(1-d)tQ_{n-1}P_nQ_n - P_{n-1}(Q_n^2 - dt(P_n^2 + Q_n^2) + dt^2P_n^2) & \text{if } n \text{ is odd} \end{cases}$$

and

$$Q_{n+1}(t) = \begin{cases} Q_{n-1}(Q_n^2 - dP_n^2 + dt(P_n^2 - Q_n^2)) & \text{if } n \text{ is even} \\ Q_{n-1}(Q_n^2 - dt(P_n^2 + Q_n^2) + dt^2P_n^2) & \text{if } n \text{ is odd} \end{cases}$$

Remark. In [MMG, Th. 9.4], Moloney and McGuire find similar formulæ for x. However, this does not allow one to perform Kummer recursions as the polynomials involved have are not univariate.

Another interesting property of Edwards curves is that they are birationally equivalent to Montgomery curves which permits one to speed up some computation [CS3]. the Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ is indeed equivalent to

$$M : \frac{4}{1-d}v^2 = u^3 + 2\frac{1+d}{1-d}u^2 + u$$

via the transformations

$$(x,y) \longmapsto \left(\frac{1+y}{1-y}, \frac{1+y}{x(1-y)}\right) \qquad (u,v) \longmapsto \left(\frac{u}{v}, \frac{u-1}{u+1}\right)$$

this equivalence can be used to construct division polynomials for Montgomery curves using Montgomery ladders.

4.2.5 Hessian division polynomials

The arithmetic of the Hessian curve $H_d : x^3 + y^3 + 1 = 3xy$ is described in section 4.1.6. The distinguished point is the point at infinity $(1 : -1 : 0)$ and the negation is $-(x, y) = (y, x)$. This gives different possibilities for a degree 2 map down to $\mathbb{P}^1$. In [DKW1] the authors use $(x, y) \mapsto x + y$ while Farashahi and Joye [FJ] choose $(x, y) \mapsto x^3 + y^3$.

Another possibility would be to take the product of the coordinates $\sigma : (x, y) \mapsto xy$. This produces

$$\sigma(P+Q) = -\frac{(1-d\sigma(P))\,(1-d\sigma(Q))}{(\sigma(Q)-\sigma(P))^2} - \sigma(P-Q)$$

In [FFT] the authors take a slightly different approach. By using the complete addition formulæ of [FJ]

$$(x_P, y_P) + (x_Q, y_Q) = \left(\frac{y_Q - x_P y_P x_Q^2}{x_P x_Q^2 - y_Q y_P^2}, \frac{x_Q y_Q y_P^2 - x_P}{x_P x_Q^2 - y_Q y_P^2}\right)$$

they observed that $x_{P+Q}y_{P-Q}$ and $x_{P-Q}y_{P+Q}$ have quite a simple form and permit therefore to find explicit recursion formulæ.

4.3 Complex multiplication

4.3.1 Generalized division polynomials

Let E/k with complex multiplication by the ring of integers $\mathcal{O}_F$ of a quadratic imaginary field F. For any ideal $\mathfrak{a} \subset \mathcal{O}_F$, the group of $\mathfrak{a}$-torsion points on E is defined by

$$E[\mathfrak{a}] = \{P \in E \mid \alpha P = O \quad \text{for all } \alpha \in \mathfrak{a}\}$$

where this definition depends on the embedding $\mathcal{O}_F \hookrightarrow \mathrm{End}(E)$.

Definition. The generalized division polynomial attached to the ideal $\mathfrak{a}$ is defined by

$$\psi_{\mathfrak{a}}(x) = \prod_{P \in E[\mathfrak{a}]\setminus\{O\}} (x - x(P))$$

This generalizes the definition of standard division polynomials ψ_n for $n \in \mathbb{Z}$. From this definition we obtain a condition on the divisor:

$$\mathrm{div}(\psi_n) = \sum_{P\in E[n]\setminus\{O\}} ((P) - (O)) = \sum_{P\in E[n]} (P) - n^2(O) = \sum_{P\in E[n]\setminus\{O\}} (P) - (n^2-1)(O)$$

which shows that the degree of ψ_n is $\deg(\psi_n) = n^2 - 1$. If we proceed the other way around, this condition on the divisor defines division polynomials up to a constant multiple. The choice of this normalization condition is usually done by imposing $\psi_n^*\omega = n\omega$ for any invariant differential ω of E.

Satoh [Sat] tried to mimic this construction for any ideal $\mathfrak{a} \subset \mathcal{O}_F$. It turns out that there exists a rational function $\psi_{\mathfrak{a}}$ on E satisfying

$$\mathrm{div}(\psi_{\mathfrak{a}}) = \sum_{P\in E[\mathfrak{a}]} (P) - \mathrm{N}_{F/\mathbb{Q}}(\mathfrak{a})(O)$$

if and only if $\sum_{P\in E[\mathfrak{a}]} P = O$.

Definition. An ideal $\mathfrak{a} \subset \mathcal{O}_F$ is said to be unbiased if $\mathfrak{a} + 2\mathcal{O}_F = \mathcal{O}_F$ or $2\mathcal{O}_F$. An element $\alpha \in \mathcal{O}_F$ is unbiased if the principal ideal (α) is unbiased. A subgroup $G \subseteq E[n]$ is unbiased if $\sum_{P\in G} P = O$.

Remark. Note that an ideal $\mathfrak{a} \subset \mathcal{O}_F$ is unbiased if and only if $E[\mathfrak{a}]$ is unbiased, [Sat, Th. 2.5].

Lemma 4.13 ([Sat, Cor. 2.8])**.** *An algebraic integer $\alpha \in \mathcal{O}_F$ is unbiased if and only if either $2 \mid \alpha$ or $\mathrm{N}_{F/\mathbb{Q}}(\alpha)$ is odd. More explicitly, let d be a square free positive integer such that $F = \mathbb{Q}(\sqrt{-d})$ and let*

$$\omega = \begin{cases} \sqrt{-d} & \text{if } d \equiv 1, 2 \mod 4 \\ (1+\sqrt{-d})/2 & \text{if } d \equiv 3 \mod 4 \end{cases}$$

be a generator of the ring of integers $\mathbb{Z}[\omega]$ of F.

- *If $d \equiv 1 \mod 4$, $\alpha = a + b\omega$ is unbiased except for $a \equiv b \equiv 1 \mod 2$.*
- *If $d \equiv 2 \mod 4$, $\alpha = a + b\omega$ is unbiased except for $a \equiv 0$ and $b \equiv 1$ modulo 2.*
- *If $d \equiv 3 \mod 8$, every element $\alpha = a + b\omega$ of $\mathcal{O}_F$ is unbiased.*
- *If $d \equiv 7 \mod 8$, then $\alpha = a + b\omega$ is unbiased if and only if $b \equiv 0 \mod 2$.*

Proof. We characterize unbiased elements based on the splitting behavior of 2 in $\mathcal{O}_F$.

If 2 ramifies, then let ω be an unbiased generator of $\mathcal{O}_F$. The quotient to the residue field $\mathcal{O}_F \to \mathcal{O}_F/2\mathcal{O}_F \simeq \mathbb{F}_2$ sends the lattice of unbiased elements to 0, meaning that $\alpha = a + b\omega$ is unbiased except for $a \equiv b \equiv 1 \mod 2$. The distinction between the first two points is that $\sqrt{-d}$ is biased if $d \equiv 2 \mod 4$ and it is unbiased if $d \equiv 1 \mod 4$.

If 2 remains prime, then ω is unbiased since it has odd norm. In this case $\mathcal{O}_F/2\mathcal{O}_F \simeq \mathbb{F}_4$ and all the elements are unbiased.

If 2 splits both ω and $1+\omega$ are biased. By the quotient $\mathcal{O}_F \to \mathcal{O}_F/2\mathcal{O}_F \simeq \mathbb{F}_2 \times \mathbb{F}_2$ the biased elements are $(1,0)$ and $(0,1)$ which means that $\alpha = a + b\omega \in \mathcal{O}_F$ is unbiased if and only if $b \equiv 0 \mod 2$. □

For unbiased elements, we can define the α-th division polynomials as those functions such that

$$\operatorname{div}(\psi_\alpha) = \sum_{P \in \ker(\mathfrak{a})} (P) - \mathrm{N}_{F/\mathbb{Q}}(\mathfrak{a})(O)$$

together with the normalization condition

$$\psi_0(x) = 0 \qquad \psi_\alpha(x) = (-1)^{\mathrm{N}(\alpha)-1}\alpha x^{Nr(\alpha)-1} + \dots$$

In the following, we will write x_α to indicate the x-coordinate of αP. For two non-zero elements α, β such that $\alpha \pm \beta$ are unbiased, Satoh recovered the usual condition for division polynomials [CF, Appendix]

$$x_\alpha - x_\beta = \frac{\psi_{\alpha+\beta}\psi_{\alpha-\beta}}{\psi_\alpha^2\psi_\beta^2}$$

which permits one to show that the generalized division polynomials satisfy

$$\psi_\beta^2\psi_{\alpha+\gamma}\psi_{\alpha-\gamma} - \psi_\alpha^2\psi_{\beta+\gamma}\psi_{\beta-\gamma} = \psi_{\alpha+\beta}\psi_{\alpha-\beta}\psi_\gamma^2$$

This difference equation is again a generalization of the integral formula

$$\psi_n^2\psi_{m+1}\psi_{m-1} - \psi_m^2\psi_{n+1}\psi_{n-1} = \psi_{m+n}\psi_{m-n}$$

Such a relation was first studied by Ward [War3]; any sequence $(\psi_n)_{n\in\mathbb{N}}$ satisfying it together with the initial conditions $\psi_0 = 0, \psi_1 = 1, \psi_2\psi_3 \neq 0$ and $\psi_2 \mid \psi_4$ is called an *elliptic divisibility sequence*.

Remark. The *divisibility* in the name references to the divisibility property $\psi_m \mid \psi_n$ whenever $m \mid n$.

It was again Ward who studied this kind of recursions over the Gaussian integers while Durst [Dur] focused on the case of Eisenstein integers. Chudnovsky and Chudnovsky [CC] studied elliptic divisibility sequences indexed by endomorphism rings of elliptic curves [Str3].

In this framework, Satoh was able to prove the following recursion formulæ

Proposition 4.14 ([Sat, Prop. 3.8])**.** *Let $\alpha \in \mathcal{O}_F$ be unbiased. Then, there exist an elliptic divisibility sequence*

$$\psi_\alpha = \left(\psi_{\beta_1}^2\psi_{\beta_2}\psi_{\beta_3} - \psi_{\beta_4}^2\psi_{\beta_5}\psi_{\beta_6}\right)/\psi_\delta^t$$

where $\beta_1, \dots, \beta_6 \in \mathcal{O}_F$ and $\delta \in \{1, 2, \omega, 1+\omega, 1-\omega, 1+2\omega\}$ are unbiased elements and $t \in \{0, 1, 3\}$. Further, it satisfies the following condition: $\|\beta_i\| \leq \frac{1}{2}\|\alpha\| + 2$ for all i. Here $\|a + b\omega\| = \max(|a|, |b|)$

Now let $E/k : y^2 = x^3 + ax + b$ be an elliptic curve in short Weierstrass form, generalized division polynomials are defined over $\mathcal{O}_K$ where $K = F(a, b)$:

Corollary 4.15 ([Sat, Cor. 4.3])**.** *Let $\alpha \in \mathcal{O}_F$ be unbiased. If $\mathrm{N}_{F/\mathbb{Q}}(\alpha)$ is odd, then $\psi_\alpha \in \mathcal{O}_K[x]$ while $\psi_\alpha \in y\mathcal{O}_K[x]$ in case $\mathrm{N}_{F/\mathbb{Q}}(\alpha)$ is even.*

4.3.2 Division polynomials for the Gaussian integers

In this section we specialize the work of Satoh for the order $\mathbb{Z}[i]$ of Gaussian integers. In particular, since the case of unbiased elements has already been described by Satoh [Sat], we will mainly focus on biased elements.

We consider the elliptic curve $E : y^2 = x(x^2 + u)$ with j-invariant 1728; it has complex multiplication by the ring of Gaussian integers $\mathbb{Z}[i]$. By Corollary 2.8 of [Sat] we know that biased elements are of the form $a + bi$ with $a \equiv b \equiv 1 \bmod 2$.

Lemma 4.16. *We recall that an element $\alpha \in \mathbb{Z}[i]$ is biased if and only if one of the following equivalent conditions is satisfied:*

(1) $\alpha \equiv 1 + i$ *mod* $2\mathbb{Z}[i]$*;*

(2) $\mathrm{N}(\alpha) \equiv 2 \mod 4$*;*

(3) $\nu_{\mathfrak{p}}(\alpha) = 1$*;*

(4) *The $\mathfrak{p}$-adic expansion of α is* $(1 + i) + a_2(1 + i)^2 + \dots$

where $\mathfrak{p} = (1 + i)$ is the unique prime over 2 in $\mathbb{Z}[i]$

The key observation is that every element β of the form $(2k + 1) + (2k' + 1)i$ is divisible by $1 + i$:

$$(2k + 1) + (2k' + 1)i = (1 + i)(a + bi) \qquad \begin{cases} a - b = 2k + 1 \\ a + b = 2k' + 1 \end{cases} \implies \begin{cases} a = k + k' + 1 \\ b = k - k' \end{cases}$$

The new Gaussian integer $\alpha = (k+k'+1)+(k-k')i$ has the property of being unbiased since $k+k'+1 \not\equiv k-k'$ mod 2.

Remark. We know that an element ω of a ring of integers $\mathcal{O}_F$ is unbiased if and only if $2 \mid \omega$ or $Nr(\omega)$ is odd ([Sat, Cor. 2.6]). In our case, the second property holds since

$$Nr(\alpha) = k^2 + k'^2 + 1 + 2kk' + 2k + 2k' + k^2 + k'^2 - 2kk' = 2(k^2 + k'^2 + k + k') + 1$$

We therefore have $\langle\alpha\rangle + \langle 1 + i\rangle = \mathbb{Z}[i]$ and, by consequence, $E[\beta] = E[\alpha] \oplus E[1 + i]$.

Proposition 4.17. *With the notation as above, the division polynomials of β is given by $\psi_\beta^2 = (1+i)^2\psi_\alpha^2\phi_\alpha$.*

Proof using addition Formulæ. By Proposition 3.8 of [Sat] we know how to recursively construct the division polynomial ψ_α for the unbiased part. We therefore suppose to have the complete description

$$[\alpha] = \left(\frac{\phi_\alpha}{\psi_\alpha^2}, \frac{\omega_\alpha}{\psi_\alpha^3}\right)$$

Now

$$[\beta] = [(1 + i)\alpha] = [1 + i][\alpha] = [\alpha] + [i][\alpha] = \left(\frac{\phi_\alpha}{\psi_\alpha^2}, \frac{\omega_\alpha}{\psi_\alpha^3}\right) + \left(-\frac{\phi_\alpha}{\psi_\alpha^2}, i\frac{\omega_\alpha}{\psi_\alpha^3}\right)$$

By explicit addition formulæ[Sil1, Alg. III.2.3] we obtain

$$x_\beta = \left(\frac{\frac{\omega_\alpha}{\psi_\alpha^3} - i\frac{\omega_\alpha}{\psi_\alpha^3}}{\frac{\phi_\alpha}{\psi_\alpha^2} + \frac{\phi_\alpha}{\psi_\alpha^2}}\right)^2 = \left(\frac{(1 - i)\omega_\alpha}{2\psi_\alpha\phi_\alpha}\right)^2 = \frac{\omega_\alpha^2}{(1 + i)^2\psi_\alpha^2\phi_\alpha^2}$$

This implies that ψ_β divides $\psi_\alpha^2\phi_\alpha^2$.
On the other hand, we know that $[\alpha]P$ is still a point on the elliptic curve:

$$y_\alpha^2 = x_\alpha(x_\alpha^2 + u) \implies \frac{\omega_\alpha^2}{\psi_\alpha^3} = \frac{\phi_\alpha}{\psi_\alpha^2}\left(\frac{\phi^2}{\psi_\alpha^4} + u\right) \implies \omega_\alpha^2 = \phi_\alpha\left(\phi_\alpha^2 + u\psi_\alpha^4\right)$$

Thus,

$$x_\beta = \frac{\omega_\alpha^2}{(1 + i)^2\psi_\alpha^2\phi_\alpha^2} = \frac{\phi_\alpha\left(\phi_\alpha^2 + u\psi_\alpha^4\right)}{(1 + i)^2\psi_\alpha^2\phi_\alpha^2} = \frac{\phi_\alpha^2 + u\psi_\alpha^4}{(1 + i)^2\psi_\alpha^2\phi_\alpha}$$

and we cannot simplify further since ϕ_α and ψ_α are coprime. In the same fashion, we observe that

$$y_\beta = -\left(\frac{\frac{\omega_\alpha}{\psi_\alpha^3} - i\frac{\omega_\alpha}{\psi_\alpha^3}}{\frac{\phi_\alpha}{\psi_\alpha^2} + \frac{\phi_\alpha}{\psi_\alpha^2}}\right)^2 - \left(\frac{i\frac{\omega_\alpha}{\psi_\alpha^3}\frac{\phi_\alpha}{\psi_\alpha^2} + \frac{\omega_\alpha}{\psi_\alpha^3}\frac{\phi_\alpha}{\psi_\alpha^2}}{\frac{\phi_\alpha}{\psi_\alpha^2} + \frac{\phi_\alpha}{\psi_\alpha^2}}\right) = -\left(\frac{(1-i)\omega_\alpha}{2\phi_\alpha\psi_\alpha}\right)^3 - \left(\frac{(1+i)\omega_\alpha\phi_\alpha}{2\phi_\alpha\psi_\alpha^3}\right) =$$
$$= -\frac{\omega_\alpha^3}{(1+i)^3\phi_\alpha^3\psi_\alpha^3} - \frac{\omega_\alpha\phi_\alpha}{(1-i)\phi_\alpha\psi_\alpha^3} = -\frac{(1-i)\omega_\alpha^3 + (1+i)^3\omega_\alpha\phi_\alpha^3}{(1+i)^3(1-i)\phi_\alpha^3\psi_\alpha^3} = -(1-i)\omega_\alpha\frac{\omega_\alpha^2 - 2\phi_\alpha^3}{(1+i)^3(1-i)\phi_\alpha^3\psi_\alpha^3} =$$
$$= \omega_\alpha\frac{2\phi_\alpha^3 - \omega_\alpha^2}{(1+i)^3\phi_\alpha^3\psi_\alpha^3} = \omega_\alpha\phi_\alpha\frac{2\phi_\alpha^2 - \phi_\alpha^2 - u\psi_\alpha^4}{(1+i)^3\phi_\alpha^3\psi_\alpha^3} = \omega_\alpha\frac{\phi_\alpha^2 - u\psi_\alpha^4}{(1+i)^3\phi_\alpha^2\psi_\alpha^3}$$

□

Proof using composition. The advantage of using the composition $[\alpha]([1+i](x,y))$ and $[i+i]([\alpha](x,y))$ is that it is well defined on the x-coordinate (the Kummer line).

$$[1+i]\left(\frac{\phi_\alpha}{\psi_\alpha^2}, \frac{\omega_\alpha}{\psi_\alpha^3}\right) = \left(\frac{\frac{\phi_\alpha^2}{\psi_\alpha^4} + u}{(1+i)^2\frac{\phi_\alpha}{\psi_\alpha^2}}, \frac{\frac{\phi_\alpha^2}{\psi_\alpha^4} - u}{(1+i)^3\frac{\phi_\alpha^2}{\psi_\alpha^3}}\frac{\omega_\alpha}{\psi_\alpha^3}\right) = \left(\frac{\phi_\alpha^2 + u\psi_\alpha^4}{(1+i)^2\phi_\alpha\psi_\alpha^2}, \omega_\alpha\frac{\phi_\alpha^2 - u\psi_\alpha^4}{(1+i)^3\phi_\alpha^2\psi_\alpha^3}\right)$$

□

Definition. We define $\psi_\beta = (1+i)\psi_\alpha\phi_\alpha$.

Remark. The definition of $(\phi_{1+i}, \psi_{1+i}, \omega_{1+i})$ for $[1+i]$ requires a different multiplicity of the factor $(1+i)$ and x in the denominators:

$$[1+i](x,y) = \left(\frac{x^2+u}{(1+i)^2x}, \frac{(x^2-u)y}{(1+i)^3x^2}\right)$$

The same happens for the division polynomial of every unbiased element: $\psi_\beta = (1+i)\psi_\alpha\phi_\alpha$ but

$$[\beta](x,y) = \left(\frac{\phi_\alpha^2 + u\psi_\alpha^4}{(1+i)^2\psi_\alpha^2\phi_\alpha}, \omega_\alpha\frac{\phi_\alpha^2 - u\psi_\alpha^4}{(1+i)^3\phi_\alpha^2\psi_\alpha^3}\right)$$

where the unbiased part is represented by $(1+i)\psi_{1+i}$ and the biased part by ϕ_α.

We can infer something more observing the composition formulas for division polynomials [Str1, §2], [Sat, Lemm 3.5].

$$\phi_{(1+i)\alpha} = \left(\phi_\alpha \circ \frac{\phi_{i+1}}{\psi_{i+1}^2}\right)\psi_{i+1}^{2Nr(\alpha)} = \left(\phi_{i+1} \circ \frac{\phi_\alpha}{\psi_\alpha^2}\right)\psi_\alpha^4$$
$$\psi_{(1+i)\alpha}^2 = \left(\psi_\alpha^2 \circ \frac{\phi_{1+i}}{\psi_{1+i}^2}\right)\psi_{1+i}^{2Nr(\alpha)} = \left(\psi_{i+1}^2 \circ \frac{\phi_\alpha}{\psi_\alpha^2}\right)\psi_\alpha^4$$

Remark. Only even powers of ψ_{1+i} are involved and therefore we only have to deal with meromorphic functions.

Remark. Since ψ_α^2 has degree $Nr(\alpha) - 1$, then ψ_{1+i}^2 divides ψ_β^2. This, together with the fact that $\psi_\alpha \mid \psi_\beta$ confirms the result of Corollary 4.2 of [Str1].

Lemma 4.18. *The leading coefficient of ψ_α is α for every $\alpha \in \mathbb{Z}[i]$.*

Proof. For unbiased elements, this is the normalization condition used in [Sat, Def. 3.1] to define division polynomials. By the formula $\psi_\beta = (1+i)\psi_\alpha\phi_\alpha$ above we can extend the result to biased elements. □

The same result comes from the following observation. For any $\alpha = a+ib \in \mathbb{Z}[i]$ we have $[a+bi]\,P = \mathcal{O}$ if and only if $[a]\,P = -[ib]\,P$. In particular,

$$\left(\frac{\phi_a}{\psi_a^2}, \frac{\omega_a}{\psi_a^3}\right) = \left(-\frac{\phi_b}{\psi_b^2}, -i\frac{\omega_b}{\psi_b^3}\right)$$

Thus, $\psi_\alpha \mid \phi_a\psi_b^2 + \phi_b\psi_a^2$. Since the same holds for $\bar{\alpha}$ and no other Gaussian integer, we get

$$\psi_\alpha\psi_{\bar{\alpha}} = \phi_a\psi_b^2 + \phi_b\psi_a^2$$

Note that both a and ib are unbiased elements. Now,

$$\psi_\alpha \psi_{\bar{\alpha}} = \phi_a \psi_b^2 + \phi_b \psi_a^2 = \psi_b^2 \left(x\psi_a^2 - \psi_{a+1}\psi_{a-1} \right) + \psi_a^2 \left(x\psi_b^2 - \psi_{b+1}\psi_{b-1} \right)$$

We know [Sut1, Lemma 6.21]

$$\phi_n(x) = x^{n^2} + \dots \qquad \psi_n(x) = \begin{cases} nx^{\frac{n^2-1}{2}} + \dots & \text{if } n \text{ is odd} \\ y\left(nx^{\frac{n^2-2}{4}} + \dots\right) & \text{if } n \text{ is even} \end{cases}$$

Thus, denoted $\ell(P)$ the leading coefficient of $P(X) \in \mathbb{Z}[i][X]$, we have

$$\begin{aligned} \ell(\psi_\alpha)\ell(\psi_{\bar{\alpha}}) = \ell(\psi_b^2) \left[\ell(\psi_a^2) - \ell(\psi_{a+1})\ell(\psi_{a-1}) \right] + \ell(\psi_a^2) \left[\ell(\psi_b^2) - \ell(\psi_{b+1})\ell(\psi_{b-1}) \right] = \\ = b^2 \left[a^2 - (a+1)(a-1) \right] + a^2 \left[b^2 - (b+1)(b-1) \right] = b^2 + a^2 = Nr(\alpha) \end{aligned}$$

At the same time, composition formula says that, for unbiased elements α,

$$Nr(\alpha) = \ell(\psi_{Nr(\alpha)}) = \ell(\psi_\alpha)\ell(\psi_{\bar{\alpha}})\ell(\phi_\alpha)^{Nr(\alpha)}$$

implying $\ell(\phi_\alpha) = 1$; in particular, since all Gaussian primes are unbiased, ϕ_α is monic for any prime in $\mathbb{Z}[i]$.

Switching the roles of α and its conjugate, for Gaussian primes we get $\ell(\psi_\alpha) = \alpha$ or $\ell(\psi_\alpha) = \bar{\alpha}$. By composition, we extend the same result to all unbiased elements. Finally, given the formula for biased elements we reduce to the former and we extend the result to all Gaussian integers.

Lemma 4.19. *If α is an unbiased element of odd norm, then ψ_α is an even polynomial.*

Proof. We consider the composition $[i]\,[\alpha]$. Since both i and α are unbiased we can once more employ the composition formula.

$$\psi_{i\alpha}(x) = \psi_i \left(\frac{\phi_\alpha(x)}{\psi_\alpha(x)} \right) \psi_\alpha^{Nr(i)}(x) = i\ \psi_\alpha(x)$$

On the other hand, we can switch the roles and obtain

$$\psi_{i\alpha}(x) = \psi_\alpha \left(\frac{\phi_i(x)}{\psi_i(x)} \right) \psi_i^{Nr(\alpha)}(x) = \psi_\alpha(-x)\ i^{Nr(\alpha)}$$

Observe that $Nr(\alpha) = Nr(a+ib) = a^2 + b^2$. A famous theorem by Fermat (see [Hat, Ch. 6]) states that the values represented by the quadratic form $x^2 + y^2$ where x and y run over all integers are exactly the numbers of the form $m^2 p_1 \cdot \ldots \cdot p_k$ where m is an arbitrary integer and each p_i is either 2 or a prime congruent to 1 mod 4.

In particular this means that $Nr(\alpha)$ can be congruent to $0, 1$ or 2 mod 4 but not 3. By the assumptions made on α we conclude that $Nr(\alpha) \equiv 1$ mod 4 meaning

$$\psi_{i\alpha}(x) = \psi_\alpha(-x)\ i^{Nr(\alpha)} = i\psi_\alpha(-x)$$

Finally, comparing the two different expressions for $\psi_{i\alpha}(x)$, we obtain

$$i\psi_\alpha(-x) = i\psi_\alpha(x) \implies \psi_\alpha(-x) = \psi_\alpha(x)$$

□

Endomorphisms of E induce endomorphisms of the Kummer quotient $\mathbb{P}^1 = E/\{\pm 1\}$, but also the Kummer quotient $\mathbb{P}^1 = E/\{\pm 1, \pm i\}$ in x^2.

$$\begin{tikzcd} E \arrow[r] \arrow[d, "x"'] & E \arrow[d, "x"] \\ \mathbb{P}^1 \arrow[r] \arrow[d, "x^2"'] & \mathbb{P}^1 \arrow[d, "x^2"] \\ \mathbb{P}^1 \arrow[r] & \mathbb{P}^1 \end{tikzcd}$$

In particular, ψ_α is a polynomial in x^2 when α is unbiased.

Corollary 4.20. *If α is an unbiased element divisible by 2, then ψ_α^2 is an odd polynomial.*

Proof. We write $\alpha = 2^k\beta$ with $(\beta, 2) = 1$. We have to distinguish between 2 cases

(β unbiased) We will work on induction on k.

($k = 1$) By composition $\psi_{2\beta} = \psi_\beta \left(\phi_2/\psi_2^2\right) \psi_2^{2Nr(\beta)}$.
Suppose that we know the expansion of ψ_β^2

$$\psi_\beta^2 = \sum_{j=0}^{Nr(\beta)-1} a_j x^j$$

Thus,

$$\psi_{2\beta}^2 = \psi_2^{2Nr(\beta)} \sum_{j=0}^{Nr(\beta)-1} a_j \frac{\phi_2^j}{\psi_2^{2j}} = \sum_{j=0}^{Nr(\beta)-1} a_j \phi_2^j \psi_2^{2Nr(\beta)-2j} = \psi_2^2 \sum_{j=0}^{Nr(\beta)-1} a_j \phi_2^j \left(\psi_2^2\right)^{Nr(\beta)-j-1}$$

We know that ϕ_2 is an even polynomial ($\phi_2(x) = x^4 - 2ux^2 + u^2$) and $\psi_2^2(x) = 4x(x^2 + u)$ is odd.

By the lemma, we know that ψ_β^2 is even. Then, j runs over the even numbers meaning that $Nr(\beta) - j - 1$ is always even and, therefore, the sum is an even polynomial. Finally the multiplication by ψ_2^2 brings a factor of x which makes $\psi_\alpha = \psi_{2\beta}$ an odd polynomial.

(Induction step) We suppose that $\psi_{2^{k-1}\beta}^2$ is odd. With a little abuse of notation we look again at the composition formula above (with $2^{k-1}\beta$ in the role of β) then

$$\begin{aligned}\psi_{2^k\beta}^2 = \psi_2^{2Nr(2^{k-1}\beta)} \sum_{j=0}^{Nr(2^{k-1}\beta)-1} a_j \frac{\phi_2^j}{\psi_2^{2j}} = \sum_{j=0}^{Nr(2^{k-1}\beta)-1} a_j \phi_2^j \psi_2^{2Nr(2^{k-1}\beta)-2j} = \\ = \psi_2^2 \sum_{j=0}^{Nr(2^{k-1}\beta)-1} a_j \phi_2^j \left(\psi_2^2\right)^{Nr(2^{k-1}\beta)-j-1} = \psi_2^2 \sum_{j=0}^{Nr(2^{k-1}\beta)-1} a_j \phi_2^j \left(\psi_2^2\right)^{4^{k-1}Nr(\beta)-j-1}\end{aligned}$$

Now, by induction hypothesis, j ranges over all the odd integers in $\{0, \ldots, Nr(2^{k-1}\beta) - 1\}$ and therefore $4^{k-1}Nr(\beta) - j - 1$ is always even meaning that, again, the sum on the right is an even polynomial. As before, we conclude that ψ_α is odd.

(β biased) In this situation we know that $\beta = (1+i)\beta'$. Therefore, $\alpha = (1+i)2^k\beta'$ with β' unbiased. We have already found that $\psi_\alpha = \psi_{(1+i)2^k\beta'} = (1+i)\psi_{2^k\beta'}\phi_{2^k\beta'}$ and by the discussion above we know that $\psi_{2^k\beta'}$ is odd. It only remains to show that $\phi_{2^k\beta'}$ is even in which case the lemma would be proved.

We will discuss the subject in the next section and we will conclude the proof of the lemma there.

□

Remark. The extra factor of x in the case of an unbiased element α divisible by two comes indeed from the 2-torsion part of $E[\alpha]$; in other words, ψ_α/ψ_2 is even.

The numerator ϕ_α

The description of ϕ_α streams from the recursion formula

$$\phi_\delta \psi_\gamma^2 - \phi_\gamma \psi_\delta^2 = \psi_{\gamma+\delta}\psi_{\gamma-\delta}$$

outlined in [Cas1, Formulary] for integers and extended in [Sat, Prop. 3.6] for unbiased elements.

Remark. The proof of the above formula requires $\gamma + \delta$ and $\gamma - \delta$ to be unbiased.

We will distinguish two cases. In case k and k' have the same parity, then $\alpha \pm 1$ is unbiased. Hence,

$$\phi_\alpha = x\psi_\alpha^2 - \psi_{\alpha+1}\psi_{\alpha-1}$$

Otherwise ($k \not\equiv k'$ mod 2), $\alpha \pm 1$ is biased but $\alpha \pm i$ is unbiased. Thus,

$$\phi_\alpha = \psi_{\alpha+i}\psi_{\alpha-i} - x\psi_\alpha^2$$

Lemma 4.21. *If α is an unbiased element of odd norm, then ϕ_α is an odd polynomial.*

Proof. We will use the same strategy exploited in Lemma 4.19.

$$\phi_{i\alpha}(x) = \psi_\alpha^{2Nr(i)}(x)\phi_i\left(\frac{\phi_\alpha(x)}{\psi_\alpha(x)}\right) = \psi_\alpha(x)\frac{\phi_\alpha(x)}{\psi_\alpha(x)} = \phi_\alpha(x)$$

At the same time

$$\phi_{i\alpha}(x) = \psi_i^{2Nr(\alpha)}(x)\phi_\alpha\left(\frac{\phi_i(x)}{\psi_i(x)}\right) = i^{2Nr(\alpha)}\phi_\alpha(-x) = -\phi_\alpha(-x)$$

Thus $\phi_\alpha(-x) = -\phi(x)$. □

Corollary 4.22. *If α is an unbiased element divisible of the form $\alpha = 2^k\beta$ where β is unbiased with odd norm, then ϕ_α^2 is an even polynomial.*

Proof. As we did in Corollary 4.20, we can reduce to study the simplest case $k = 1$. If $\alpha = 2\beta$ then

$$\phi_{2\beta} = \psi_2^{2Nr(\beta)}\phi_\beta\left(\frac{\phi_2}{\psi_2^2}\right)$$

Since we already know that ϕ_β is odd, we can write $\phi_\beta = \sum_{j=0}^{(Nr(\beta)-1)/2} b_j x^{2j+1}$. Thus,

$$\phi_\alpha = \phi_{2\beta} = \psi_2^{2Nr(\beta)}\sum_{j=0}^{(Nr(\beta)-1)/2} b_j\left(\frac{\phi_2}{\psi_2^2}\right)^{2j+1} = \sum_{j=0}^{(Nr(\beta)-1)/2} b_j\phi_2^{2j+1}\psi_2^{2Nr(\beta)-2(2j+1)}$$

It is now easy to see that $2Nr(\beta) - 2(2j+1) \equiv 0$ mod 4 for all $j = 0, \dots, (Nr(\beta)-1)/2$.
Therefore, since all the factor of x carried by ψ_2^2 are squared, the entire polynomial ϕ_α is even. □

End of proof of Corollary 4.20. Since the product of an even and an odd polynomial is odd, this concludes the proof we left incomplete in the previous section □

Remark. We already knew that $[-1]$ stabilizes the torsion group $E[\alpha]$. In particular, we have $[-1]$ stabilizing the quotient $\phi_\alpha/\psi_\alpha^2$. Does this imply that it fixes both ψ_α and ϕ_α? A straightforward observation shows that $\phi_{-\alpha} = \phi_\alpha$ and $\psi_{-\alpha} = -\psi_\alpha$.

Remark. What happens with $[i]$? We have seen that $\phi_{i\alpha} = \phi_\alpha$ and $\psi_{i\alpha} = \psi_\alpha$. Indeed, $[i](x, y) = (-x, iy)$, so $\phi_{i\alpha}(x) = \phi_\alpha(x)$ if and only if ϕ_α is even and $\psi_{i\alpha}^2(x) = \psi_\alpha^2(-x) = \pm\psi_\alpha^2(x)$.

Gaussian multiplication

In conclusion, supposing that E is the elliptic curve $y^2 = x(x^2 + u)$, recursions for the division polynomial can be defined by

(i) the action on the invariant differential

$$[\alpha]^*\frac{dx_\alpha}{2y_\alpha} = \alpha\frac{dx}{2y}$$

(ii) composition law; in particular with $[1+i]$, noting that $[c+di] = [1+i][a+bi]$ where $(c, d) = (a-b, a+b)$ or equivalently $(a, b) = ((c+d)/2, (d-c)/2)$ when $a \equiv b$ mod 2.

(iii) addition laws in terms of $(x : y : z) = (x : y : x^2 + u)$.

An initial table of Gaussian division polynomials is below

$a+bi$	ψ_α	ϕ_α	ω_α
1	1	x	y
i	i	x	y
$1+i$	$(1+i)x$	x^2+u	$(x^2-u)y$
$1-i$	$(1-i)x$	x^2+u	$(x^2-u)y$
$1+2i$	$(1+2i)x^2+u$	$x^5+(2+4i)ux^3-(3-4i)u^2x$	
$1-2i$	$(1-2i)x^2+u$	$x^5+(2-4i)ux^3-(3+4i)u^2x$	

Table 4.1 – Initialization table for Gaussian division polynomials

4.3.3 Division polynomials for the Eisenstein integers

We now focus on the elliptic curve $y^2 = x^3 + u$ with j invariant 0 and complex multiplication by the ring of integers of $\mathbb{Q}(\sqrt{-3})$, namely the class number one order of Eisenstein integers $\mathbb{Z}[\omega]$ where $\omega^2+\omega+1 = 0$. By Lemma 4.13, we know that every element of $\mathbb{Z}[\omega]$ is unbiased. An initial table of Eisenstein division polynomials follows where the important relations are $1+\omega = -\omega^2 = -\overline{\omega}$ and $\omega(1-k\omega) = k+(k+1)\omega$

$a+bw$	ψ_α	ϕ_α	ω_α
1	1	x	y
ω	ω	x	y
$1+\omega$	$\omega+1$	x	y
$1-\omega$	$(1-\omega)x$	x^3+4u	$(x^3-8u)y$
$1+2\omega$	$(1+2\omega)x$	x^3+4u	$(x^3-8u)y$
$1-2\omega$	$(1-2\omega)x^3+4(\omega+1)u$	$x^7-4(1+12\omega)ux^4-16(2+3\omega)u^2x$	
$2+\omega$	$(2+\omega)x$	x^3+4u	$(x^3-8u)y$
$2-\omega$	$(2-w\omega)x^3-4(\omega+1)u$	$x^7+4(11+12\omega)ux^4+16(1+3\omega)u^2x$	

Table 4.2 – Initialization table for Eisenstein division polynomials

from which we infer the similarity between the division polynomials for $1-\omega$ and $1+2\omega$.

If we look at the map down to the Kummer line we observe that on E,

$$x_{P+Q} = (\frac{y_Q - y_P}{x_Q - x_P})^2 - x_P - x_Q \qquad x_{P-Q} = (\frac{-y_Q - y_P}{x_Q - x_P})^2 - x_P - x_Q$$

and therefore

$$x_{P+Q} - x_{P-Q} = (\frac{y_Q - y_P}{x_Q - x_P})^2 - (\frac{-y_Q - y_P}{x_Q - x_P})^2 = \frac{-4y_P y_Q}{(x_Q - x_P)^2}$$

Supposing $Q = [\omega](x,y)$ and $P = [\alpha](x,y)$ with $\alpha = a+b\omega$, then

$$x_{a+(b+1)\omega} = \frac{\phi_{\alpha-\omega}}{\psi^2_{\alpha-\omega}} - 4\frac{\psi_\alpha \omega_\alpha y}{(\phi_\alpha - \omega x \psi^2_\alpha)^2}$$

This yields

$$\psi_{\alpha+\omega} = \psi_{\alpha-\omega}(\phi_\alpha - \omega x \psi^2_\alpha)$$

In the same way

$$\psi_{\alpha+\omega} = \psi_{\alpha-1}(\phi_\alpha - x \psi^2_\alpha)$$

4.3.4 Edwards generalized division polynomials

Discriminant -3

The elliptic curve with j-invariant 0 has 6 automorphisms $\mathrm{Aut}(E_0) = \{[\pm 1], [\pm\zeta_3], [\pm\zeta_3^2]\}$ where ζ_3 is a primitive cube root of unity. In Weierstrass form, it has equation

$$E_0 : v^2 = u^3 + c$$

and the action of ζ_3 is given by $[\zeta_3](u,v) = (\zeta_3 u, v)$.

We want to describe this map on its Edwards model. First of all, we note that the general Edwards curve $E_d : x^2 + y^2 = 1 + dx^2y^2$ has j-invariant

$$j = \frac{16(1+14d+d^2)^3}{d(1-d)^4}$$

The Edwards model of E_0 has $d = -7 \pm 4\sqrt{3}$. There is an isomorphism between these two curves [Liu+]:

$$\begin{aligned}\varphi : E_{-7-4\sqrt{3}} &\longrightarrow W_0 \\ (x,y) &\longmapsto \left(c_1\frac{1+y}{1-y} + c_2, c_1\frac{1+y}{x(1-y)}\right)\end{aligned} \qquad \begin{aligned}\tilde{\varphi} : W_0 &\longrightarrow E_{-7-4\sqrt{3}} \\ (u,v) &\longmapsto \left(\frac{u-c_2}{v}, \frac{u-c_3}{u+c_4}\right)\end{aligned}$$

where

$$c_1 = \frac{1-d}{4} = 2+\sqrt{3} \qquad c_2 = \frac{1+d}{6} = -1-\frac{2}{3}\sqrt{3} \qquad c_3 = \frac{5-d}{12} = 1+\frac{1}{3}\sqrt{3} \qquad c_4 = \frac{1-5d}{12} = 3+\frac{5}{3}\sqrt{3}$$

In order to construct the rational maps describing the automorphism $[\zeta_3]$ we can use the equivalence to get to W_0, act by ζ_3 and finally apply the equivalence once again to find ourselves back to E_d.

$$\begin{aligned}
[\zeta_3](x,y) = \tilde{\varphi} \circ [\zeta_3] \circ \varphi(x,y) = & \\
&= \tilde{\varphi} \circ [\zeta_3]\left(c_1\frac{1+y}{1-y} + c_2, c_1\frac{1+y}{x(1-y)}\right) = \\
&= \tilde{\varphi}\left(\zeta_3\left[c_1\frac{1+y}{1-y} + c_2\right], c_1\frac{1+y}{x(1-y)}\right) = \\
&= \left(\frac{\zeta_3 c_1\frac{1+y}{1-y} + \zeta_3 c_2 - c_2}{\frac{c_1}{x}\frac{1+y}{1-y}}, \frac{\zeta_3 c_1\frac{1+y}{1-y} + \zeta_3 c_2 - c_1 - c_2}{\zeta_3 c_1\frac{1+y}{1-y} + \zeta_3 c_2 + c_1 - c_2}\right) = \\
&= \left(\frac{\zeta_3 c_1(1+y) + (\zeta_3 - 1)(1-y)c_2}{c_1(1+y)}x, \frac{\zeta_3 c_1(1+y) + c_2(\zeta_3-1)(1-y) - c_1(1-y)}{\zeta_3 c_1(1+y) + c_2(\zeta_3-1)(1-y) + c_1(1-y)}\right) = \\
&= \left(\frac{y(\zeta_3 c_1 - \zeta_3 c_2 + c_2) + \zeta_3 c_1 + \zeta_3 c_2 - c_2}{c_1(1+y)}, \frac{y(\zeta_3 c_1 - \zeta_3 c_2 + c_1 + c_2) + \zeta_3 c_1 + \zeta_3 c_2 - c_1 - c_2}{y(\zeta_3 c_1 - \zeta_3 c_2 - c_1 + c_2) + \zeta_3 c_1 + \zeta_3 c_2 + c_1 - c_2}\right) = \\
&= \left(x\frac{c_5 y + c_6}{y+1}, \frac{c_7 y + c_8}{y + c_9}\right)
\end{aligned}$$

where

$$\begin{cases}
c_5 = \dfrac{\zeta_3(5d-1) - 2d - 2}{3(d-1)} \\
c_6 = \dfrac{\zeta_3(d-5) + 2d + 2}{3(d-1)} \\
c_7 = \dfrac{\zeta_3(5d-1) + d - 5}{(\zeta_3 - 1)(5d-1)} \\
c_8 = \dfrac{d-5}{5d-1} \\
c_9 = \dfrac{\zeta_3(d-5) + 5d - 1}{(\zeta_3 - 1)(5d-1)}
\end{cases}$$

Since the cube root of unity ζ_3 satisfies the polynomial x^2+x+1, the action of $1+\zeta_3$ is given by $-\zeta_3^2 = \overline{\zeta}_3$.

$$[1+\zeta_3](x,y) = \left(-x\frac{\overline{c}_5 y + \overline{c}_6}{y+1}, -\frac{\overline{c}_7 y + \overline{c}_8}{y + \overline{c}_9}\right)$$

where we indicate by $\overline{c}_i$ the coefficient c_i with ζ_3 replaced by its conjugate.

Discriminant -4

The elliptic curve with j-invariant 1728 has Weierstrass form $W_{1728} : v^2 = u^3 + cu$ and it admits an endomorphism $[i]$, for i the primitive 4-th root of unity, given by

$$[i] : (u,v) \longmapsto (-u, iv)$$

The Edwards curve with j-invariant 1728 is defined by $E_{-1} : x^2 + y^2 = 1 - x^2y^2$. We can use the birational equivalence

$$\begin{aligned} \varphi : E_{-1} &\longrightarrow W_{1728} : v^2 = u^3 + 1/4u \\ (x, y) &\longmapsto \left(\frac{1}{2}\frac{1+y}{1-y}, \frac{1}{2}\frac{1+y}{x(1-y)}\right) \end{aligned} \qquad \begin{aligned} \tilde{\varphi} : W_{1728} &\longrightarrow E_{-1} \\ (u, v) &\longmapsto \left(\frac{u}{v}, \frac{u - 1/2}{u + 1/2}\right) \end{aligned}$$

to recover the action of i on E_{-1}.

$$[i](x, y) = \tilde{\varphi} \circ [i] \circ \varphi(x, y) = \tilde{\varphi}\left(-\frac{1}{2}\frac{1+y}{1-y}, \frac{i}{2}\frac{1+y}{x(1-y)}\right) = \left(ix, \frac{1}{y}\right)$$

This map fixes the identity O but it is not defined when $y = 0$, i.e., for points of the form $(\pm 1, 0)$ which are two primitive 4 torsion points.

In the same way, one could compute the image of $[1 + i]$:

$$\begin{aligned}
[1+i](x, y) &= \tilde{\varphi} \circ [1+i] \circ \varphi(x, y) = \\
&= \tilde{\varphi} \circ [1+i] \left(\frac{1}{2}\frac{1+y}{1-y}, \frac{1}{2}\ \frac{1+y}{x(1-y)}\right) = \\
&= \tilde{\varphi}\left(\frac{\frac{1}{4}\frac{(1+y)^2}{(1-y)^2} + \frac{1}{4}}{(1+i)^2\frac{1}{2}\frac{1+y}{1-y}}, \frac{\frac{1}{2x}\frac{1+y}{1-y}}{\frac{1+i}{2}\frac{1+y}{1-y}}\ \frac{\frac{1}{4}\frac{(1+y)^2}{(1-y)^2} - \frac{1}{4}}{(1+i)^2\frac{1}{2}\frac{1+y}{1-y}}\right) = \\
&= \tilde{\varphi}\left(\frac{(1+y)^2 + (1-y)^2}{2(1+i)^2(1+y)(1-y)}, \frac{1}{(1+i)x}\ \frac{(1+y)^2 - (1-y)^2}{2(1+i)^2(1+y)(1-y)}\right) = \\
&= \tilde{\varphi}\left(\frac{1+y^2}{(1+i)^2(1-y^2)}, \frac{1}{(1+i)x}\ \frac{2y}{(1+i)^2(1-y^2)}\right) = \\
&= \tilde{\varphi}\left(\frac{1}{(1+i)^2x^2}, \frac{1}{(1+i)^3x^3}\ \frac{2y}{1+y^2}\right) = \\
&= \left(\frac{(1+i)^3x^3}{(1+i)^2x^2}\ \frac{1+y^2}{2y}, \frac{\frac{1}{(1+i)^2x^2} - \frac{1}{2}}{\frac{1}{(1+i)^2x^2} + \frac{1}{2}}\right) = \\
&= \left(\frac{1+y^2}{(1-i)y}x, \frac{1-ix^2}{1+ix^2}\right) = \left(x\frac{1+y^2}{(1-i)y}, \frac{1+iy^2}{i+y^2}\right)
\end{aligned}$$

and the action of its conjugate

$$[1-i](x, y) = \left(\frac{1+y^2}{(1+i)y}x, \frac{1+ix^2}{1-ix^2}\right) = \left(x\frac{1+y^2}{(1+i)y}, \frac{1-iy^2}{-i+y^2}\right)$$

We can now start to deduce some recursion formulæ. Using the same technique as in section 4.2.4 we obtain

$$\begin{aligned}
[2+i](x, y) = (x_{2+i}, y_{2+i}) &= \left(\frac{2x_{1+i}y(1-dx^2)}{(1+x_{1+i}^2x^2)} - ix, \frac{4y_{1+i}y}{1+y_{1+i}^2+y^2-y_{1+i}^2y^2} - \frac{1}{y}\right) = \\
&= \left(-ix\frac{y^4+2(1+i)y^2+1}{y^4+2(-1-i)y^2+1}, \frac{(-1+2i)y^4+1}{y^5+(-1+2i)y}\right)
\end{aligned}$$

$$[2-i](x, y) = \left(-ix\frac{y^4+2(1-i)y^2+1}{y^4+2(-1+i)y^2+1}, \frac{(-1-2i)y^4+1}{y^5+(-1-2i)y}\right)$$

$$[1+2i](x, y) = \left(-x\frac{y^4+2(1-i)y^2+1}{y^4+2(-1+i)y^2+1}, y\frac{y^4+(-1-2i)}{(-1-2i)y^4+1}\right)$$

$$[1-2i](x, y) = \left(-x\frac{y^4+2(-1-i)y^2+1}{y^4+2(-1-i)y^2+1}, y\frac{y^4+(-1+2i)}{(-1+2i)y^4+1}\right)$$

Discriminant -7

The elliptic curve $W_{-3375} : v^2 = u^3 - 35u + 98$ has j-invariant -15^3 and its endomorphism ring contains the ring of integers of $\mathbb{Q}(\sqrt{-7})$ which is the class number one order $\mathbb{Z}[\alpha]$ generated by $\alpha = \frac{1+\sqrt{-7}}{2}$ whose minimal polynomial is $x^2 - x + 2$ [Sil2, §II.2].

$$\left[\frac{1+\sqrt{-7}}{2}\right](u,v) = \left(\alpha^{-2}\left(u - \frac{7(1-\alpha)^4}{u+\alpha^2-2}\right), \alpha^{-3}v\left(1 - \frac{7(1-\alpha)^4}{(u+\alpha^2-2)^2}\right)\right)$$

We find that the Edwards curve $E : x^2 + y^2 = 1 + dx^2y^2$ with coefficient $d = -48\sqrt{7} - 127$ also has j-invariant -15^3 and it is birationally equivalent to the curve $\tilde{W} : v^2 = u^3 - (240\sqrt{7}+635)u + (4048\sqrt{7}+10710)$ via the maps

$$\begin{aligned} \varphi : E &\longrightarrow \tilde{W} & \tilde{\varphi} : \tilde{W} &\longrightarrow E \\ (x,y) &\longrightarrow \left(\frac{c_1(1+y)}{1-y} + c_2, \frac{c_1(1+y)}{x(1-y)}\right) & (u,v) &\longmapsto \left(\frac{u-c_2}{v}, \frac{u-c_3}{u+c_4}\right) \end{aligned}$$

where $c_1 = 12\sqrt{7} + 32$, $c_2 = -8\sqrt{7} - 21$, $c3 = 4\sqrt{7} + 11$ and $c4 = 20\sqrt{7} + 53$.

Composing with the isomorphisms between W and $\tilde{W}$ (which are defined over a quadratic extension of $\mathbb{Q}(\sqrt{7})$), we recover the action of α on E.

$$[\alpha]\,(x,y) = \left(d_1 x\frac{(y+d_2)(y-d_3)}{(y-d_2)(y+1)}, d_4\frac{y^2+d_5y+d_3}{y^2-d_5y+d_3}\right)$$

with $d_1 = \frac{1}{7}(3\alpha+2)\sqrt{7} + (\alpha+1)$, $d_2 = \frac{1}{7}(3\alpha+2)\sqrt{7} - (\alpha+1)$, $d_3 = \frac{1}{7}(-16\alpha+8)\sqrt{7} + 6\alpha - 3$, $d_4 = \frac{1}{7}(3\alpha-5)\sqrt{7} - \alpha + 2$ and $d_5 = \frac{1}{7}(10\alpha+2)\sqrt{7} - 4\alpha$.

Discriminant -8

There exists a third curve that possesses an endomorphism of degree 2 [Sil2, Prop. 2.3.1], namely $W_{8000} : v^2 = u^3 + 4x^2 + 2x$ which has j-invariant 8000 and admits an embedding of the class number one order $\mathbb{Z}[\sqrt{-2}]$ of discriminant -8 into its endomorphism ring:

$$[\sqrt{-2}](u,v) = \left(\beta^{-2}\left(u + 4 + \frac{2}{u}\right), \beta^{-3}v\left(1 - \frac{2}{u^2}\right)\right) \qquad \beta = \sqrt{-2}$$

The Edwards model for j-invariant 20^3 elliptic curves is $E_{3-2\sqrt{2}} : x^2 + y^2 = 1 + (3 - 2\sqrt{2})x^2y^2$ with maps to $\tilde{W}_{8000} : v^2 = u^3 + (\frac{5}{6}\sqrt{2} - \frac{5}{4})u + (-\frac{49}{108}\sqrt{2} + \frac{35}{54})$

$$\begin{aligned} \varphi : E_{3-2\sqrt{2}} &\longrightarrow \tilde{W}_{8000} & \tilde{\varphi} : \tilde{W}_{8000} &\longrightarrow E_{3-2\sqrt{2}} \\ (x,y) &\longrightarrow \left(\frac{c_1(1+y)}{1-y} + c_2, \frac{c_1(1+y)}{x(1-y)}\right) & (u,v) &\longmapsto \left(\frac{u-c_2}{v}, \frac{u-c_3}{u+c_4}\right) \end{aligned}$$

The endomorphism $[\sqrt{-2}]$ on E is therefore

$$[\sqrt{-2}]\left(x\frac{y^2-(2\sqrt{2}+3)}{(-4i-\sqrt{-2})y}, \frac{y^2-(\sqrt{2}+1)}{(1-\sqrt{2})y^2-1}\right)$$

4.3.5 Hessian generalized division polynomials

Discriminant -3

In section 4.1.6 we have seen that the Hessian curve $H_d : X^3 + Y^3 + Z^3 = 3dXYZ$ has j-invariant $\left(3\frac{d(d^3+8)}{d^3-1}\right)^3$. Therefore, it has j invariant 0 if $d = 0, -2$. Setting $d = 0$ we get $H_0 : X^3 + Y^3 + Z^3 = 0$ which is birationally equivalent to the Weierstrass curve $W_0 : V^2 = U^3 - 16/27$ via the maps

$$\varphi : H_0 \longrightarrow W_0 \qquad\qquad \tilde{\varphi} : W_0 \longrightarrow H_0$$

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \longmapsto \begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \\ -4 & 4 & 0 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad \begin{pmatrix} U \\ V \\ W \end{pmatrix} \longmapsto \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0 & 3 & 4 \\ 0 & -3 & 4 \\ -6 & 0 & 0 \end{pmatrix} \begin{pmatrix} U \\ V \\ W \end{pmatrix}$$

Then,

$$[\zeta_3](X : Y : Z) = (X : Y : \zeta_3 Z)$$

Since

$$\begin{pmatrix} 0 & 3 & 4 \\ 0 & -3 & 4 \\ -6 & 0 & 0 \end{pmatrix} \begin{pmatrix} \zeta_3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 4 \\ -4 & 4 & 0 \\ -1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} -24 & 0 & 0 \\ 0 & -24 & 0 \\ 0 & 0 & -24\zeta_3 \end{pmatrix}$$

In the same way

$$[1 + \zeta_3](X : Y : Z) = (Y : X : \overline{\zeta}_3 Z)$$

Using addition formulæ 4.9 we find

$$\begin{aligned}[1 - \zeta_3](X : Y : Z) = (X : Y : Z) + (Y : X : \zeta_3 Z) &= \big(Y^3(\zeta_3 + 1) + Z^3 : X^3(\zeta_3 + 1) + Z^3 : (1 - \zeta_3^2)XYZ\big) \\ &= \big(Y^3 - \zeta_3 Z^3 : X^3 - \zeta_3 Z^3 : (1 - \zeta_3)XYZ\big)\end{aligned}$$

Discriminant -4

The Hessian curve with j-invariant 1728 has coefficient $d = 1 + \sqrt{3}$ and it is isomorphic to the Weierstrass curve $W_{1728} : V^2 = U^3 - (12 + 8\sqrt{3})U$ via the maps

$$\varphi : H_0 \longrightarrow W_0 \qquad \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \longmapsto \begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 4 + 2\sqrt{3} & 4 + 2\sqrt{3} & -2 - 2\sqrt{3} \\ 12 + 8\sqrt{3} & -12 - 8\sqrt{3} & 0 \\ -1 & -1 & -1 - \sqrt{3} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

$$\tilde{\varphi} : W_0 \longrightarrow H_0 \qquad \begin{pmatrix} U \\ V \\ W \end{pmatrix} \longmapsto \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1 + \sqrt{3} & 1 & -2 - 2 * \sqrt{3} \\ 1 + \sqrt{3} & -1 & -2 - 2 * \sqrt{3} \\ -2 & 0 & -8 - 4 * \sqrt{3} \end{pmatrix} \begin{pmatrix} U \\ V \\ W \end{pmatrix}$$

By consequence

$$[i](X : Y : Z) = (\zeta_3 X + \zeta_3^2 Y + Z : \zeta_3^2 X + \zeta_3 Y + Z : X + Y + Z)$$

Chapter 5

Class Groups Actions

5.1 CM groups

In this section we introduce a computational framework for computing actions of ray class groups of imaginary quadratic fields. We establish equivalences of categories that enable one to choose the best environment in which to work in different situations. In particular, we construct equivalences between quadratic forms, ideals in imaginary quadratic orders and points on the j-line $X(1)$. We eventually endow these objects with level structure.

5.1.1 The category of binary quadratic forms

We start by formalizing the notion of quadratic forms. The main references will be [Cox] and [Cas2] but some alternative books on the topic are [Hat] and [Conw] which take a more geometric point of view.

Definition. A binary quadratic form is a homogeneous polynomial of degree 2 in $\mathbb{Z}[x, y]$

$$F(x, y) = ax^2 + bxy + cy^2$$

We say that F is primitive if $\gcd(a, b, c) = 1$.

The discriminant of F is $\Delta = b^2 - 4ac$ and provides an important invariant in the study of quadratic forms. Forms with $\Delta > 0$ are called indefinite, those with $\Delta < 0$ are called definite and those with $\Delta = 0$ are said to be parabolic.

An integer n is said to be represented by a quadratic form F if there exist $(x_0, y_0) \in \mathbb{Z}^2$ such that $F(x_0, y_0) = n$. If (x_0, y_0) is coprime to n, we say that n is properly represented. The set of represented integers provides another invariant of F.

In view of the construction of an equivalence of categories with CM-points on the upper half plane, we equip the set of quadratic forms with an action of $\mathrm{SL}_2(\mathbb{Z})$. For a matrix

$$\gamma = \begin{pmatrix} s & t \\ u & v \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

we define $\gamma \cdot F(x, y) = F(sx + ty, ux + vy)$. One can check that this action preserves primitive forms and leave the discriminant unvaried.

Notation. We will usually denote $F(x, y) = ax^2 + bxy + xy^2$ by its coefficients (a, b, c).

The action of $\mathrm{SL}_2(\mathbb{Z})$ is translated into

$$\gamma \cdot \langle a, b, c\rangle = \langle a', b', c'\rangle \qquad \text{where} \quad \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} s^2 & su & u^2 \\ 2st & sv + tu & 2uv \\ t^2 & tv & v^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Two quadratic forms are said to be equivalent of they lie in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit; we write $(a, b, c) \sim (a', b', c')$.

Remark. Note that this action can be extended to the slightly larger group $\mathrm{GL}_2(\mathbb{Z})$. This has the advantage that quadratic forms in the same $\mathrm{GL}_2(\mathbb{Z})$-orbit represent the same integers. However, $\mathrm{SL}_2(\mathbb{Z})$ equivalence encodes more significant information.

Lemma 5.1. $\mathrm{SL}_2(\mathbb{Z})$*-equivalent quadratic forms represent properly the same set of integers.*

Remark. Note that the converse is not true.

The study of reduced forms differs depending on the sign of the discriminant. We specialize now on definite quadratic forms since these are the ones related to imaginary quadratic fields. Further, the theory of reduced forms is much richer in the definite case.

Definite quadratic forms

Definite quadratic forms split into two main classes: positive definite forms representing only positive integers and negative definite forms which only represent negative integers. Every orbit consists of either positive or negative definite quadratic forms and the operation $(a, b, c) \mapsto (-a, b, -c)$ permits one to pass from one to the other. Thus, they can be treated in the same way meaning there is no substantial difference between the two theories. Therefore, without loss of generality, we consider positive definite quadratic forms.

Definition. A positive definite quadratic form $F = (a, b, c)$ is reduced if

$$|b| \leq a \leq c \quad \text{and} \quad b \geq 0 \text{ if either } |b| = a \text{ or } a = c$$

A negative definite form is reduced if the corresponding positive definite form is reduced.

This definition has a geometric interpretation: we define the roots of F as the roots of $F(x, 1) \in \mathbb{Z}[x]$, namely $\tau_\pm = (-b \pm \sqrt{\Delta})/(2a)$. Then a form is reduced if an only if one of its roots lies in the fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on upper half plane $\mathbb{H}$, see Figure 2.1.

Reduced forms are distinguished objects in $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes as they form a complete set of representatives (once a fundamental domain has been chosen).

Theorem 5.2. *Every primitive positive definite quadratic form is equivalent to a unique reduced form.*

Corollary 5.3. *The number of equivalence classes of primitive positive definite quadratic forms of discriminant* Δ *is finite.*

Proof. For a reduced form $b^2 \leq a^2$ and $a \leq c$. Thus, $\Delta = b^2 - 4ac \leq -3a^2$ and this shows that there are a finite number of choices for a and then for b. The relation $\Delta = b^2 - 4ac$ determines c consequently. □

Definition. For a discriminant Δ we let $h_f(\Delta)$ denote the number of equivalence classes of primitive positive definite quadratic forms and call it the *(form) class number* of Δ.

The proof above yields an algorithm for reducing quadratic forms [Coh, §5.4.2].

Algorithm 6. Reduction of primitive positive definite quadratic forms

Input: A primitive positive definite form $F = (a, b, c)$
Output: The unique reduced form in the $\mathrm{SL}_2(\mathbb{Z})$-orbit of F

```
While $F$ is not reduced do:
    • If $a > c$ (or $c = a$ and $b < 0$) replace $F$ by $(c, -b, a) \sim (a, b, c)$.
    • If $|b| > a$ or $b = -a$, replace $F$ by $(a, b', c') \sim (a, b, c)$ where $b' \equiv b$ modulo $2a$, $-a < b' \leq a$ and
      $c' = ((b')^2 - \Delta)/(4a)$.
```

The two transformations above corresponds exactly to the action of S and T^k for $k = (b' - b)/(2a)$.

Example. Let $F = (31, 134, 151)$. Applying the Algorithm 6, we get

$$(31, 134, 151) \longmapsto (31, 10, 7) \longmapsto (7, -10, 31) \longmapsto (7, 4, 28)$$

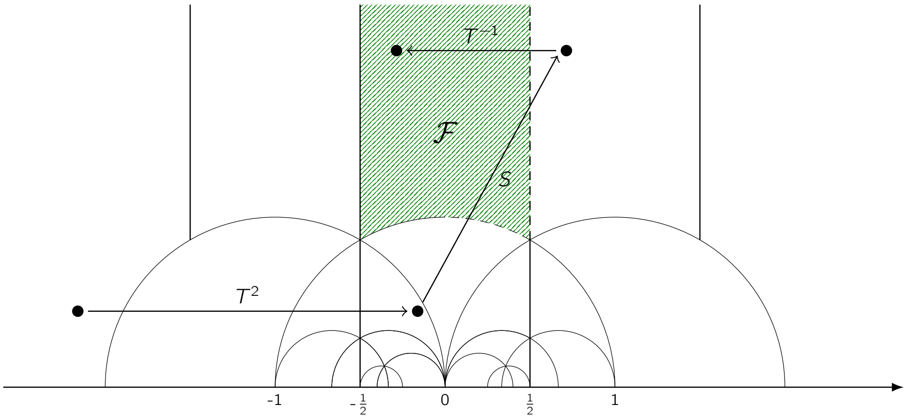


Figure 5.1 – Reduction algorithm for primitive positive definite quadratic forms

In practice the algorithm consists in acting on F with S and T until its root does not fall into the fundamental domain $\mathcal{F}$.

The proof of the corollary also allows one to construct algorithms that enumerates all reduced forms.

Algorithm 7. Enumeration of all primitive positive definite quadratic forms

Input: A discriminant $\Delta < 0$
Output: The set of reduced primitive quadratic forms of discriminant Δ

1. For a reduced form of discriminant Δ, we have $|b| \leq \sqrt{|\Delta|/3}$ and $b \equiv \delta \mod 2$. We list all possible values for b.

2. Next we note that a is a divisor of $(b^2 - \delta)/4$ and $|b| \leq a$. Since we need $a \leq c$ we need $a \leq \sqrt{(b^2 + \Delta)/4}$. We list all possible choices for a and compute the corresponding c's.

3. For each triple we check whether $(a, b, c) \sim (a, -b, c)$ and return one or both accordingly.

In the following table we enumerate and count reduced forms for small discriminants.

Δ	$h_f(\Delta)$	Reduced forms
-3	1	$(1, 1, 1)$
-4	1	$(1, 0, 1)$
-7	1	$(1, 1, 2)$
-8	1	$(1, 0, 2)$
-11	1	$(1, 1, 3)$
-12	1	$(1, 0, 1)$
-15	2	$(1, 1, 4), (2, 1, 2)$
-16	1	$(1, 0, 4)$
-19	1	$(1, 1, 5)$
-20	2	$(1, 0, 5), (2, 2, 3)$
-23	3	$(1, 1, 6), (2, \pm1, 3)$
-24	2	$(1, 0, 6), (2, 0, 3)$
-27	1	$(1, 1, 7)$
-28	1	$(1, 0, 7)$
-31	3	$(1, 1, 8), (2, \pm1, 4)$
-32	2	$(1, 0, 8), (3, 2, 3)$
-35	2	$(1, 1, 9), (3, 1, 3)$
-36	2	$(1, 0, 9), (2, 2, 5)$
-39	4	$(1, 1, 10), (2, \pm1, 5), (3, 3, 4)$
-40	2	$(1, 0, 10), (2, 0, 5)$
-43	1	$(1, 1, 11)$
-44	3	$(1, 0, 11), (3, \pm2, 4)$
-47	5	$(1, 1, 12), (2, \pm1, 6), (3, \pm1, 4)$
-48	2	$(1, 0, 12), (3, 0, 4)$
-51	2	$(1, 1, 13), (3, 3, 5)$
-52	2	$(1, 0, 13), (2, 2, 7)$
-55	4	$(1, 1, 14), (2, \pm1, 7), (4, 3, 4)$
-56	4	$(1, 0, 14), (2, 0, 7), (3, \pm2, 5)$

Table 5.1 – List of reduced forms and $h_f(\Delta)$ for small discriminants

Remark ([BM])**.** In a series of articles in the 30's, Deuring, Heilbronn and Siegel showed that $h_f(\Delta)$ grows as $\sqrt{\Delta}$ by proving that for every $\epsilon > 0$ there exists a positive constant C_ϵ such that $h_f(\Delta) \geq C_\epsilon|\Delta|^{\frac{1}{2}-\epsilon}$; unfortunately this constant is not effective. This implies that for any h there are finitely many discriminant with class number h, though it gives no explicit way of finding them.

A more effective result comes from the work of Goldfeld (1976), Gross and Zagier (1986) and Oesterlé; they related this problem to L-functions attached to elliptic curves, showing that the necessary elliptic curve exits, and found

$$h_f(\Delta) \geq C_E \log|\Delta|^{\frac{1}{2}-\epsilon} \qquad \text{for } C_E = 1/7000$$

Automorphisms

Definition. An automorphism of a form F is an element in the stabilizer of F under the action of $\mathrm{SL}_2(\mathbb{Z})$.

$$\mathrm{Aut}(F) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \gamma \cdot F = F\}$$

Proposition 5.4. *If F and G are two equivalent forms, i.e. $F = \gamma \cdot G$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, then* $\mathrm{Aut}(F) = \gamma\mathrm{Aut}(G)\gamma^{-1}$

If γ is an automorphism for a reduced quadratic form F, then it fixes its root τ in the fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$ which means that τ is an elliptic point. We know by 1.35 that there are only two such points and they are associated to the reduced forms $(1, 0, 1)$ and $(1, 1, 1)$.

It follows that the automorphism group of any definite quadratic form is isomorphic to $C_2 = \langle \pm I \rangle$, unless it is equivalent to one of the above two forms, in which case $\mathrm{Aut}(F) \simeq C_4$ or C_6 respectively.

Gauss composition and Bhargava cubes

We can now introduce an operation on the set of $\mathrm{SL}_2(\mathbb{Z})$-orbits, equivalently on the set of reduced forms. Note that the existence of such a group law is one of the reasons to prefer the $\mathrm{SL}_2(\mathbb{Z})$ action over the $\mathrm{GL}_2(\mathbb{Z})$ one.

Definition (Gauss, [Gau, Art. 235])**.** $H(x, y)$ is a direct composition of $F(x, y)$ and $G(x, y)$ if there exits two bilinear forms $B_i(x, y, z, w) = a_ixz + b_ixw + c_iyz + d_iyw$ for $i = 1, 2$ such that

$$F(x, y)G(z, w) = H(B_1(x, y, z, w), B_2(x, y, z, w)) \qquad \text{and} \qquad \begin{cases} a_1b_2 - a_2b_1 = F(1, 0) \\ a_1c_2 - a_2c_1 = G(1, 0) \end{cases}$$

This definition is not very satisfying since it does not make clear how to find the composition of any two forms, or whether any two such compositions are $\mathrm{SL}_2(\mathbb{Z})$-equivalent. Gauss [Gau, Art. 236-241] gave a positive answer to both questions and went on proving that this operation endows the set of equivalence classes of definite quadratic forms with a structure of abelian group, called the class group. However, we will not follow Gauss but we will take a different point of view; in 2004 Bhargava [Bha] reformulated this composition law in a more explicit and direct way.

Definition. A Bhargava cube is a $2 \times 2 \times 2$ cube with integers associated to its vertices.

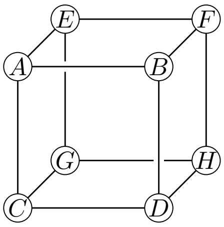


Figure 5.2 – A Bhargava cube

To any Bhargava cube we associate 3 quadratic forms

$$Q_1(x,y) = -\det\left(\begin{pmatrix} A & E \\ B & F \end{pmatrix} x + \begin{pmatrix} C & G \\ D & H \end{pmatrix} y\right)$$
$$Q_2(x,y) = -\det\left(\begin{pmatrix} A & C \\ E & G \end{pmatrix} x + \begin{pmatrix} B & D \\ F & H \end{pmatrix} y\right)$$
$$Q_3(x,y) = -\det\left(\begin{pmatrix} A & B \\ C & D \end{pmatrix} x + \begin{pmatrix} E & F \\ G & H \end{pmatrix} y\right)$$

These three forms all have the same discriminant and if any two of them are primitive so it the third one (in this case the cube is said to be projective). Further, it turns out that $Q_3(x,-y)$ is direct composition of $Q_1(x,y)$ and $Q_2(x,y)$ in the sense of Gauss.

Remark. Note that the cube carries an action of its symmetry group, isomorphic to $S_4 \times C_2$. Each rotation or reflection results in well defined transformations on the three forms associated to the cube: if $Q_i = (a_i, b_i, c_i)$ are the three quadratic forms giving rise to a Bhargava cube, we note $\hat{Q}_i = (c_i, b_i, a_i)$; then

- Rotation by $\pi/2$ around the vertical axis sends $(Q_1, Q_2, Q_3) \mapsto (-Q_1, \hat{Q}_3, -Q_2)$.
- Rotation by $2\pi/3$ around the diagonal AH sends $(Q_1, Q_2, Q_3) \mapsto (Q_3, Q_1, Q_2)$.
- Reflection switching from face and back face sends $(Q_1, Q_2, Q_3) \mapsto (-Q_1, -Q_2, \hat{Q}_3)$.

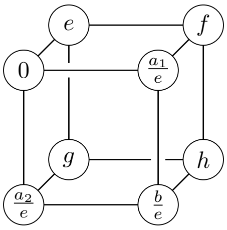
The first two generates the S_4 component while the last transformation corresponds to C_2.

Since Bhargava cubes represents triples of quadratic forms it is natural to ask how the action of $\mathrm{SL}_2(\mathbb{Z})$ affects it. Let $\mathcal{A}$ be a Bhargava cube in $(\mathbb{Z}^2)^{\otimes 3}$. A matrix

$$\gamma = \begin{pmatrix} s & t \\ u & v \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

acts on $\mathcal{A}$ by replacing the cube top face $\mathcal{F}_1$ and bottom face $\mathcal{F}_2$ by $s\mathcal{F}_1 + u\mathcal{F}_2$ and $t\mathcal{F}_1 + v\mathcal{F}_2$. This induces the action $\gamma \cdot (Q_1, Q_2, Q_3) = (\gamma \cdot Q_1, Q_2, Q_3)$. By defining similar actions on the left/right faces and front/back ones we get an action of $\mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{A}$.

Remark. The association $\mathcal{A} \mapsto (Q_1, Q_2, Q_3)$ is described above. Lemmermeyer [Lem, §3.4] gives an explicit formula for the inverse. Given two quadratic forms $Q_1 = (a_1, b_1, c_1)$ and $Q_2 = (a_2, b_2, c_2)$ of discriminant Δ, there is a cube $\mathcal{A}$ such that $\mathcal{A} \mapsto (Q_1, Q_2, Q_1Q_2)$:



where $b = (b_1 + b_2)/2$, $e = \gcd(a_1, a_2, b)$ and g, f, h are integral solutions to

$$\frac{a_1}{e} g - \frac{a_2}{e} f = \frac{b_1 - b_2}{2} \qquad h = \frac{fb - ec_2}{a_1}$$

If Q_3 is also known, one could find the suitable $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ which transforms Q_1Q_2 in Q_3 and use the transformation $\mathcal{A} \to \mathcal{A}^{\mathrm{Id} \times \mathrm{Id} \times \gamma}$.

5.1.2 The category of fractional ideals

We construct now the category of fractional ideals in quadratic fields and eventually specialize to the imaginary case. We will mainly follow [Cox] and [IK].

Let d be a square-free integer and $\mathbb{Q}(\sqrt{d})$ the corresponding quadratic number field of discriminant $\Delta_K = 4d$ except for the case $d \equiv 1 \mod 4$ in which case $\Delta_K = d$. We set $\omega = (t + \sqrt{\Delta_K})/2$ where $t \in \{0, 1\}$ is such that $\Delta_K \equiv t \mod 2$. It is well known that ω generates the ring of integers of K: $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\omega$.

The order of conductor f is the unique subring $\mathcal{O}$ of $\mathcal{O}_K$ of index f. It is of the form $\mathcal{O} = \mathbb{Z} + \mathbb{Z}f\omega$ and it has discriminant $f^2\Delta_K$.

Definition. A fractional ideal of an order $\mathcal{O}$ is a non-zero finitely generated $\mathcal{O}$-module, i.e., it is of the form $\alpha\mathfrak{a}$ for $\alpha \in K^\times$ and $\mathfrak{a}$ an ideal of $\mathcal{O}$.

Definition. A fractional ideal $\mathfrak{a}$ of $\mathcal{O}$ is proper if $\mathcal{O} = \{\alpha \in K \mid \alpha\mathfrak{a} \subseteq \mathfrak{a}\}$.

A fractional $\mathcal{O}$-ideal $\mathfrak{a}$ is invertible if there exists another fractional $\mathcal{O}$-ideal $\mathfrak{b}$ such that $\mathfrak{ab} = \mathcal{O}$. An important result show that proper fractional ideals are invertible and viceversa. It is therefore standard to refer to proper fractional ideals simply as ideals.

We define the Picard group of an order $\mathcal{O}$ as

$$\mathrm{Pic}(\mathcal{O}) = \frac{I(\mathcal{O})}{P(\mathcal{O})} = \frac{\{\text{Proper fractional } \mathcal{O}\text{-ideals}\}}{\{\text{Principal fractional } \mathcal{O}\text{-ideals}\}}$$

In the same way, when $\Delta_K > 0$, we define the narrow Picard group

$$\mathrm{Pic}^+(\mathcal{O}) = \frac{I(\mathcal{O})}{P^+(\mathcal{O})} = \frac{\{\text{Proper fractional } \mathcal{O}\text{-ideals}\}}{\{\text{Principal fractional } \mathcal{O}\text{-ideals generated by elements of positive norm}\}}$$

We call $\mathrm{Pic}(\mathcal{O})$ the (ideal) class group of Δ and we indicate it as $\mathcal{C}\ell(\Delta)$ or $\mathcal{C}\ell(\mathcal{O})$.

As we anticipated, the goal of the section is to construct equivalences of categories. We present here the first result in this direction, although we will make the construction explicit in Section 5.1.4.

Theorem 5.5 ([IK, Th. 6.7]). *Let K be an quadratic field of discriminant Δ_K and $\mathcal{O}$ the order of conductor f in $\mathcal{O}_K$.*

If K is a quadratic imaginary field, i.e., $\Delta_K < 0$, there exists a bijection between the set of equivalence classes of proper $\mathcal{O}$ ideals to the set of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of primitive positive definite quadratic forms with discriminant $f^2\Delta_K$.

If K is a real quadratic field, i.e., $\Delta_K > 0$, then there exist bijections between the set of narrow (respectively, wide) equivalence classes of proper $\mathcal{O}$-ideals and the set of $\mathrm{SL}_2(\mathbb{Z})$-equivalence (respectively $\mathrm{GL}_2(\mathbb{Z})$-equivalence) classes of primitive indefinite quadratic forms of discriminant $f^2\Delta_K$.

A principal ideal is generated by an element of positive norm if such element is positive for every real embedding of K. Then, if K is an imaginary quadratic field, there are no embeddings $K \hookrightarrow \mathbb{R}$, thus the condition above is trivially satisfied by all principal fractional ideals and then $\mathrm{Pic}^+(\mathcal{O}) = \mathrm{Pic}(\mathcal{O})$. For imaginary quadratic fields, we will extend the notation introduced before writing $\mathcal{C}\ell(\mathcal{O})$ to indicate the Picard group and we call it the class group of $\mathcal{O}$. This way we match the notation adopted for quadratic forms. The cardinality of the class group is called class number and it is indicated by $h(\mathcal{O})$.

From now on K will be an imaginary quadratic field. There are multiple ways of computing the class number of a given order. One of these comes from the study of a subclass of proper $\mathcal{O}$-ideals. We say that an ideal $\mathfrak{a} \subseteq \mathcal{O}$ is prime to the conductor of $\mathcal{O}$ if $\mathfrak{a} + f\mathcal{O} = \mathcal{O}$. These ideals are enough to describe the entire class group.

Theorem 5.6. *Given an order $\mathcal{O}$ of conductor f in $\mathcal{O}_K$, the maps*

$$\begin{gathered}\{\textit{fractional } \mathcal{O}\textit{-ideals coprime to } f\} \longleftrightarrow \{\textit{fractional } \mathcal{O}_K\textit{-ideals coprime to } f\}\\ \mathfrak{a} \longmapsto \mathfrak{a}_K\\ \mathfrak{b} \cap \mathcal{O} \longleftarrow\!\shortmid \mathfrak{b}\end{gathered}$$

are one the inverse of the other and induce an isomorphism between the class group of $\mathcal{O}$ and the set $I_{\mathcal{O}}(f)/P_{\mathcal{O},\mathbb{Z}}(f)$ where $I_{\mathcal{O}}(f)$ is the set of fractional $\mathcal{O}$-ideals coprime to f and $P_{\mathcal{O},\mathbb{Z}}(f)$ is the set of principal fractional $\mathcal{O}$-ideals coprime to f and congruent to an integer modulo f.

More precisely, we have a short exact sequence

$$1 \longrightarrow \frac{(\mathcal{O}_K/f\mathcal{O}_K)^\times}{\mathcal{O}_K^\times(\mathbb{Z}/f\mathbb{Z})^\times} \longrightarrow \mathcal{C}\ell(\mathcal{O}) \longrightarrow \mathcal{C}\ell(\mathcal{O}_K) \longrightarrow 1$$

This shows that the class group may be interpreted as an explicit quotient of the ray class group $\mathcal{C}\ell(f)$, and therefore is the Galois group of some finite abelian extension of K, the ring class field of $\mathcal{O}$.
The short exact sequence above yields the following formula

Theorem 5.7 ([Cox, Th. 7.24])**.** *Let $\mathcal{O}$ be the order of conductor f in $\mathcal{O}_K$.*

$$h(\mathcal{O}) = \frac{h(\mathcal{O}_K)f}{[\mathcal{O}_K^\times : \mathcal{O}^\times]} \prod_{p|f} \left(1 - \left(\frac{\Delta_K}{p}\right)\frac{1}{p}\right)$$

and $h(\mathcal{O})$ is an integer multiple of $h(\mathcal{O}_K)$.

5.1.3 The category of CM lattices or CM points

In Section1.2.5 we gave the general description of lattices. Here we focus on lattices associated to imaginary quadratic fields or, equivalently, to ordinary elliptic curves. We will mainly refer to [Sta] for the first part and to [Shi], [Lan2] and [Koh1] for the second one.

Assume that K is an imaginary quadratic field and suppose we fixed an embedding $K \hookrightarrow \mathbb{C}$. A complex lattice $\Lambda = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ is a K-lattice if $\omega_2/\omega_1 \in K$. Note that the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$ sends K to itself meaning that this definition does not depend on the choice of the basis.

Remark. The property of being a K-lattice is invariant under homothety.

Remark. K-lattices can be scaled to be sub-lattices of $\mathcal{O}_K$. This comes from the fact that (ω_1, ω_2) is a fractional ideal of $\mathcal{O}_K$ and then $d(\omega_1, \omega_2) \subseteq \mathcal{O}_K$ for some d in $\mathcal{O}_K$; thus, $d\Lambda \subseteq \mathcal{O}_K$. Hence, K-lattices are exactly those associated to $\tau \in K \hookrightarrow \mathbb{H}$. For this reason we will often identify the category of K-lattices with the one of CM-points consisting of points in the upper half plane coming from a quadratic imaginary field by defining $\tau = \omega_2/\omega_1$.

The most important invariant will be the order of Λ:

$$\mathrm{ord}(\Lambda) = \{x \in \mathbb{C} \mid x\Lambda \subseteq \Lambda\} = \{x \in K \mid x\Lambda \subseteq \Lambda\}$$

This object is invariant under homothety and $\mathrm{SL}_2(\mathbb{Z})$-action and does not depend on the choice of a basis for Λ. More importantly it provides a bridge to the theory of ideals. On can prove that $\mathrm{ord}(\Lambda)$ is an order in $\mathcal{O}_K$.

Theorem 5.8. *Fix an embedding of K in $\mathbb{C}$. There is a one to one-correspondence between K-lattices of order $\mathcal{O}$ up to homothety and fractional $\mathcal{O}$-ideals up to equivalence.*

The order of a K-lattice Λ_τ is obtained by embedding τ in $\mathbb{C}$, constructing its minimal polynomial $ax^2 + bx + c \in \mathbb{Z}[x]$ with $\gcd(a, b, c) = 1$ and taking $\mathbb{Z}[a\tau]$ [Lan2, Th. 8.1.5].

From this result we can deduce that quadratic forms and their roots are strongly related with K-lattices. This fits in the general idea of this chapter to relate all the objects described up to here. We will make these equivalences explicit in the next section. In the remaining, we will instead give a brief introduction to class field theory.

Class field theory

The idea behind class field theory is to construct a framework which permits one to describe (abelian) extensions of number fields. Let K be a number field, $\mathcal{O}_K$ the integral closure of $\mathbb{Z}$ in K and L/K a finite extension. For any prime $\mathfrak{p}$ of $\mathcal{O}_K$ and any prime $\mathfrak{P}$ of $\mathcal{O}_L$ above it, we get an extension of residue fields $\mathbb{F}_\mathfrak{P}/\mathbb{F}_\mathfrak{p}$ of degree $f_\mathfrak{P}$. We recall that

$$\sum_{\mathfrak{P}|\mathfrak{p}\mathcal{O}_L} e_\mathfrak{P} f_\mathfrak{P} = [L : K]$$

where $e_\mathfrak{P}$ is the exponent of $\mathfrak{P}$ in the prime factorization of $\mathfrak{p}\mathcal{O}_L$.

If L/K is a Galois extension, then the Galois group $\mathcal{G}al(L/K)$ acts transitively on the prime ideals above $\mathfrak{p}$; for any of them, say $\mathfrak{P}$, we can define the decomposition subgroup as

$$D(\mathfrak{P} \mid \mathfrak{p}) = \{\sigma \in \mathcal{G}al(L/K) \mid \mathfrak{P}^\sigma = \mathfrak{P}\}$$

Remark. Decomposition groups are the Galois groups of the associated local field extensions $L_{\mathfrak{P}}/K_{\mathfrak{p}}$.

There is a natural map

$$D(\mathfrak{P} \mid \mathfrak{p}) \longrightarrow \mathcal{G}al(\mathbb{F}_{\mathfrak{P}}/\mathbb{F}_{\mathfrak{p}})$$

which is a surjection of kernel $I(\mathfrak{P} \mid \mathfrak{p})$, the ramification group. If $\mathfrak{p}$ is unramified in L, then $D(\mathfrak{P} \mid \mathfrak{p}) \simeq \mathcal{G}al(\mathbb{F}_{\mathfrak{P}}/\mathbb{F}_{\mathfrak{p}})$ and then, for any $\mathfrak{P} \mid \mathfrak{p}\mathcal{O}_L$ there exists a unique $\sigma_{\mathfrak{P}} \in D(\mathfrak{P} \mid \mathfrak{p})$ mapping to $\mathrm{Frob}_{\mathfrak{p}} : x \mapsto x^{Nr(\mathfrak{p})}$, the Frobenius map on the residue field extension.

Suppose now that L/K is an abelian extension, then we denote $\sigma_{\mathfrak{p}} = \sigma_{\mathfrak{P}}$ for any prime $\mathfrak{P}$ above $\mathfrak{p}$. Since conjugation is trivial this is a well defined element and does not depend on the choice of $\mathfrak{P}$. We get a map

$$\begin{aligned} [\cdot, L/K] : I_K(\Delta_{L/K}) &\longrightarrow \mathcal{G}al(L/K) \\ \mathfrak{p} &\longmapsto \sigma_{\mathfrak{p}} \\ \mathfrak{a} = \prod \mathfrak{p}_i^{r_i} &\longmapsto \prod (\sigma_{\mathfrak{p}_i})^{r_i} \end{aligned}$$

extending multiplicatively the construction above to all ideals coprime to $\Delta_{L/K}$, the discriminant of L/K.

Definition. The homomorphism $[\cdot, L/K]$ is the Artin map.

The Artin map is surjective with kernel $N_{L/K}(I_L)P_K(\Delta_{L/K})$ where I_L is the group of non-zero fractional ideals of L [Sil2, Th. 3.2].

Ray class fields. For any ideal $\mathfrak{m}$ in $\mathcal{O}_K$ we note

$$\begin{aligned} I_K &= \{\text{Fractional ideals of } \mathcal{O}_K\} \\ I_K(\mathfrak{m}) &= \{\text{Fractional ideals of } \mathcal{O}_K \text{coprime to } \mathfrak{m}\} \\ P_K(\mathfrak{m}) &= \{\text{Principal ideals of } \mathcal{O}_K \text{coprime to } \mathfrak{m}\} \\ P_1(\mathfrak{m}) &= \{(\alpha) \in P_K(\mathfrak{m}) \text{ such that } \alpha \equiv 1 \bmod {}^*\mathfrak{m}\} \\ P_{\mathbb{Z}}(\mathfrak{m}) &= \{(\alpha) \in P_K(\mathfrak{m}) \text{ such that } \alpha/n \equiv 1 \bmod {}^*\mathfrak{m} \text{ for some } n \in \mathbb{Z}\} \end{aligned}$$

We define the ray class field of K (modulo $\mathfrak{m}$) as the finite abelian extension $K_{\mathfrak{m}}/K$ with the property that for any abelian extension L/K such that the conductor $\mathfrak{c}_{L/K}$ divides $\mathfrak{m}$, then $L \subseteq K_{\mathfrak{m}}$. Equivalently, $K_{\mathfrak{m}}$ is the largest unramified abelian extension of K such that $P_1(\mathfrak{m})$ is contained in the kernel of the the Artin map $[\cdot, K_{\mathfrak{m}}/K] : I(\mathfrak{m}) \to \mathcal{G}al(K_{\mathfrak{m}}/K)$.

Ray class fields provide an answer to the problem of describing abelian extension of K as any such extension lie in $K_{\mathfrak{m}}$ for some $\mathfrak{m}$. One could also prove that it suffices to consider ray class fields modulo an integer N, i.e., $\mathfrak{m} = N\mathcal{O}_K$.

Hilbert class field. We define the Hilbert class field of K as the maximal unramified extension of K.

From now on K will be an imaginary quadratic field. For a K-lattice $\Lambda = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ of order $\mathcal{O}$ in $\mathcal{O}_K$ we write $\tau = \omega_2/\omega_1 \in K \hookrightarrow \mathbb{H}$ and we denote

$$j(\Lambda) = j(\lambda\Lambda) = j(\mathcal{O}) = j(\tau)$$

where the last term is the value of the j-function at τ, see Lemma 2.23.

Theorem 5.9. *With the notation as above, $j(\mathcal{O}_K)$ is an algebraic integer and generates the Hilbert class field of K. If $\{\mathfrak{a}_i\}$ is a complete set of representatives of ideal classes of $\mathcal{O}_K$, then the values of $j(\mathfrak{a}_i)$ are the Galois conjugates of $j(\mathcal{O}_K)$ and the Artin map defines an isomorphism $\mathcal{C}\ell(\mathcal{O}_K) \to \mathcal{G}al(H/K)$.*

Proof. See [Lan2, Th. 10.1], [Sil2, Th. II.4.3] or [Cox, §11.D]. □

Definition. The Weber function associated to a lattice Λ is the complex function

$$h(z,\Lambda) = \begin{cases} \dfrac{g_2(\Lambda)^2}{\Delta(\Lambda)}\wp(z,\Lambda)^2 & \text{if } g_3(\Lambda)=0 \\ \dfrac{g_3(\Lambda)}{\Delta(\Lambda)}\wp(z,\Lambda)^3 & \text{if } g_2(\Lambda)=0 \\ \dfrac{g_2(\Lambda)g_3(\Lambda)}{\Delta(\Lambda)}\wp(z,\Lambda) & \text{otherwise} \end{cases}$$

where g_2, and g_3 are the normalized Eisenstein series, $\wp$ is the Weierstrass $\wp$-function (see Section 1.2.5) and Δ is the Weierstrass modular discriminant (see Section 2.2.3).

Theorem 5.10. *Let $K/\mathbb{Q}$ be a quadratic imaginary field and N a positive integer. Then,*

$$K(j(\mathcal{O}_K), h(1/N, \mathcal{O}_K))$$

is the ray class field for $N\mathcal{O}_K$.

Proof. See [Cox, Th. 11.39] or [Lan2, Th. 10.1.2] and [Sil2, Th. II.5.6] for the general case. □

Ring class fields. Generalizing the construction of the Hilbert class field we define the ring class field for and order $\mathcal{O}$ in $\mathcal{O}_K$ as the unique abelian extension of K with Galois group the class group $\mathcal{C}\ell(\mathcal{O})$.

Theorem 5.11. *Let K be a quadratic imaginary field and $\mathcal{O}$ an order of conductor f in K. Then $j(\mathcal{O})$ is an algebraic integer and it generates the ring class field $K_{\mathcal{O}}$ for the order $\mathcal{O}$. If $\{\mathfrak{a}_i\}$ is a complete set of representatives for the ideal classes of $\mathcal{O}$, then the values $j(\mathfrak{a}_i)$ are the Galois conjugates for $j(\mathcal{O})$. The Artin map defines an isomorphism $\mathcal{C}\ell(\mathcal{O}) \to \mathcal{G}a\ell(K_{\mathcal{O}}/K)$.*

Let $\mathcal{O}$ be an order in a quadratic imaginary field of discriminant Δ and class number $h(\mathcal{O}) = h(\Delta)$ and suppose that $\{\mathfrak{a}_i\}$ is a complete set of representatives for the ideal classes of $\mathcal{O}$; the polynomial

$$H_{\mathcal{O}}(X) = H_{\Delta}(X) = \prod_{i=1}^{h(\mathcal{O})} (X - j(\mathfrak{a}_i)) \in \mathbb{Z}[X]$$

is irreducible and generates $K_{\mathcal{O}}$ by the theorem above. We call H_{Δ}, the *Hilbert class polynomial.*

Idelic formulation of class field theory

We briefly recall here the language of adèles and idèles which gives a more concise formulation of many results in class field theory. For more details one can refer to [Sil2, §II.3].

Let K be a number field, for each place ν of K we denote K_ν the completion of K at ν and $\mathcal{O}_\nu$ its ring of integers.

Definition. The ring of adèles of K is defined as

$$\mathbf{A}_K = \prod_\nu{}' K_\nu$$

where the notation $\prod'$ means that $(a_\nu)_\nu \in \mathbf{A}_K$ must have $a_\nu \in \mathcal{O}_H\nu$ for all but finitely many ν. For any finite set of places S containing the infinite places, we define

$$\mathbf{A}_S = \prod_{\nu\in S} K_\nu \times \prod_{\nu\notin S} \mathcal{O}_H\nu$$

We can define a topology on $\mathbf{A}_K$ by setting $\prod_\nu \mathcal{O}_H\nu$ with its product topology to be open. We obtain a topology on $\mathbf{A}_S$ making it locally compact. Finally, $\mathbf{A}_K$ is the union of $\mathbf{A}_S$ for every finite set S of places of K containing the infinite places. This induces a topology on $\mathbf{A}_K$ as a topological ring.

The units in $\mathbf{A}_K$ form a multiplicative group, called the Idèle group and denoted $\mathbf{I}_K$. We endow it not with the subspace topology from $\mathbf{A}_K$ which is not sufficient to get a topological group, but imposing $\prod_\nu \mathcal{O}_H\nu^\times$

with its product topology, to be open. For an idèle $\mathfrak{s} \in \mathbf{I}_K$, we define the ideal of $\mathfrak{s}$ to be the fractional ideal $(\mathfrak{s})$ of K given by

$$(\mathfrak{s}) = \prod_{\mathfrak{p}} \mathfrak{p}^{\mathrm{ord}_{\mathfrak{p}} \mathfrak{s}_{\mathfrak{p}}}$$

The quotient $\mathbf{C}_K = \mathbf{I}_K / K^\times$ is called the idèle class group of K. The global Artin map is the surjective homomorphism

$$\begin{aligned} \mathbf{C}_K &\longrightarrow \mathcal{G}a\ell(K^{\mathrm{ab}}/K) \\ \mathfrak{s} &\longmapsto [\mathfrak{s}, K] \end{aligned}$$

with the property that for every finite abelian extension L/K such that no prime ramifying in L divides the ideal $(\mathfrak{s})$ of $\mathfrak{s}$, then

$$[\mathfrak{s}, K]|_L = [(\mathfrak{s}, L/K)]$$

where the map on the right is the classical Artin map, see [Sil2, Th. II.3.5].

The kernel of the global Artin map is the connected component of the identity. It becomes an isomorphism of topological groups when we take the inverse limit.

$$\hat{\mathbf{C}}_K \xrightarrow{\sim} \mathcal{G}a\ell(K^{\mathrm{ab}/K}) \qquad \text{for } \hat{\mathbf{C}}_K = \varprojlim_U \mathbf{C}_K/U$$

where the limit runs over every open subgroup of finite index. This map is compatible with the norm map

$$\mathrm{N}_{L/K} : \mathbf{A}_L \longrightarrow \mathbf{A}_K$$

since $\mathbf{A}_L \simeq \mathbf{A}_K \otimes L$. This gives a commutative diagram

$$\begin{tikzcd} \mathbf{C}_K \arrow[r, "{[\cdot, L]}"] \arrow[d, "\mathrm{N}_{L/K}"'] & \mathcal{G}a\ell(L^{\mathrm{ab}}/L) \arrow[d, "\mathrm{Res}"] \\ \mathbf{C}_K \arrow[r, "{[\cdot, K]}"'] & \mathcal{G}a\ell(K^{\mathrm{ab}}/K) \end{tikzcd}$$

the Artin map induces a correspondence between finite index open subgroups of $\mathbf{C}_K$ and finite index subgroups of $\mathcal{G}a\ell(K^{\mathrm{ab}}/K)$ which , in turns, is in bijection with finite extensions of K inside K^{ab}.

5.1.4 Equivalence of categories

As anticipated, in this section we are going to make explicit the connections between the objects described in the previous sections. We will start by providing maps connecting the sets of binary quadratic forms, ideals in imaginary quadratic fields and K-lattices or points on the upper half plane; eventually, we will define the corresponding categories providing a computational setup for the following.

Binary quadratic forms and ideal classes

In Theorem 5.5 we stated the existence of a bijection between the set of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of primitive positive definite quadratic forms of discriminant $\Delta = f^2\Delta_K$, $\mathcal{C}\ell(\Delta)$, and the ideal class group of the order $\mathcal{O}$ inside $\mathcal{O}_K$ of conductor f, $\mathcal{C}\ell(\mathcal{O})$.

To any primitive quadratic form $(a, b, c) = ax^2 + bxy + cy^2$ of discriminant Δ we attach the ideal in the order $\mathcal{O}$ of discriminant Δ given by

$$(a, \frac{-b+\sqrt{\Delta}}{2}) \subseteq \mathcal{O}$$

Conversely, we associate to any ideal $\mathfrak{a} = (\alpha, \beta)$ of $\mathcal{O}$ the quadratic form

$$\frac{Nr(\alpha x - \beta y)}{Nr(\mathfrak{a})}$$

Theorem 5.12. *The maps*

$$\begin{tikzcd}
\mathcal{Cl}(\mathcal{O}) \arrow[r] & \mathcal{Cl}(\Delta) & \mathcal{Cl}(\Delta) \arrow[r] & \mathcal{Cl}(\mathcal{O}) \\
\mathfrak{a} \arrow[r, mapsto] & \frac{Nr(\alpha x-\beta y)}{Nr(\mathfrak{a})} & (a,b,c) \arrow[r, mapsto] & (a, \frac{-b+\sqrt{\Delta}}{2})
\end{tikzcd}$$

are bijective and one the inverse of the other.

Binary quadratic forms and K-lattices

As we anticipated in Section 5.1.3, given a quadratic form $ax^2+bxy+cy^2$ one can associate its root $\tau=(-b+\sqrt{\Delta})/(2a)$ and this is a point in the upper half plane.

Each lattice Λ can be expressed in the form $\mathbb{Z}_{\geq 0}+\tau\mathbb{Z}$ with $\tau=\omega_2/\omega_1$. If the minimal polynomial of τ is ax^2+bx+c with $\gcd(a,b,c)=1$, we associate to it the quadratic form (a,b,c) which corresponds to $Nr(\omega_1 x-\omega_2 y)/Nr(\omega_1)$.

Theorem 5.13. *The maps*

$$\begin{tikzcd}
\mathcal{Cl}(\Delta) \arrow[r] & \{K\text{-lattices of order } \mathcal{O}\}/\sim & \{K\text{-lattices of order } \mathcal{O}\}/\sim \arrow[r] & \mathcal{Cl}(\Delta) \\
(a,b,c) \arrow[r, mapsto] & a\mathbb{Z}+\dfrac{-b\sqrt{\Delta}}{2}\mathbb{Z} & \omega_1\mathbb{Z}+\omega_2\mathbb{Z} \arrow[r, mapsto] & \dfrac{Nr(\omega_1 x-\omega_2 y)}{Nr(\omega_1)}
\end{tikzcd}$$

are bijective and one the inverse of the other.

Binary quadratic forms and K-lattices

The immediate consequence of Theorem 5.8 is that the identification between K-lattices of order $\mathcal{O}$ and ideal classes of $\mathcal{O}$ is given by a fixed embedding of K in $\mathbb{C}$ by considering ideals as lattices.

Theorem 5.14. *The maps*

$$\begin{tikzcd}
\mathcal{Cl}(\mathcal{O}) \arrow[r] & \{K\text{-lattices of order } \mathcal{O}\}/\sim & \{K\text{-lattices of order}\mathcal{O}\}/\sim \arrow[r] & \mathcal{Cl}(\mathcal{O}) \\
(\alpha,\beta) \arrow[r, mapsto] & \alpha\mathbb{Z}+\beta\mathbb{Z} & \omega_1\mathbb{Z}+\omega_2\mathbb{Z} \arrow[r, mapsto] & (\omega_1,\omega_2)
\end{tikzcd}$$

are bijective and one the inverse of the other.

Description of the categories

Let K be an imaginary quadratic field of discriminant Δ_K, and fix an embedding $K\hookrightarrow\mathbb{C}$, so that for each $\Delta=f^2\Delta_K$, we have a unique element $\sqrt{\Delta}\in\mathbb{H}_n\cap K$. In what follows we define three categories associated to complex multiplication structures and functors between them.

Matrix endomorphisms. Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ with $\Gamma(n)\subset\Gamma$. We define the matrix subring

$$\mathbb{M}(\Gamma)=\mathbb{Z}[\Gamma]\subseteq\mathbb{M}_2(\mathbb{Z})$$

such that $\mathbb{M}(\mathrm{SL}_2(\mathbb{Z}))=\mathbb{M}_2(\mathbb{Z})$. In general, we have $\mathbb{M}(\Gamma)\cap\mathrm{SL}_2(\mathbb{Z})=\langle -I_2,\Gamma\rangle$. The classical adjoint

$$()^*:\mathbb{M}_2(\mathbb{Z})\longrightarrow\mathbb{M}_2(\mathbb{Z})$$

determines an involution of the subrings $\mathbb{M}(\Gamma)$ taking $\gamma\in\Gamma$ to $\gamma^*=\gamma^{-1}$ and

$$\begin{pmatrix} s & t \\ u & v \end{pmatrix}\longrightarrow\begin{pmatrix} v & -t \\ -u & s \end{pmatrix}$$

Action on bases and forms. An element $\gamma\in\mathbb{M}(\Gamma)$ with $\det(\gamma)\neq 0$, of the form

$$\gamma=\begin{pmatrix} s & t \\ u & v \end{pmatrix}$$

acts on the left on pairs $\mathcal{B}=(\omega_1,\omega_2)$ via

$$\gamma\cdot\mathcal{B}=(s\omega_1+t\omega_2,\, u\omega_1+v\omega_2)$$

on binary quadratic forms f by (see Section 5.1.1)

$$\gamma \cdot f(x,y) = f(sx+ty, ux+vy)$$

and on K-points $\tau \in \mathbb{H} \cap K^\times$ by (see Sections 1.2.3 and 5.1.3)

$$\gamma \cdot \tau = \frac{s\tau + t}{u\tau + v}$$

This action extends to $\gamma = 0$ with $\gamma\mathcal{B} = (0,0)$, $\gamma \cdot f(x,y) = 0$ and $\gamma\tau = 0$,respectively.

Fractional ideal category. We first consider the category $\mathcal{I}$ whose objects are fractional ideals in K, and whose morphisms are

$$\mathrm{Hom}(\mathfrak{a}, \mathfrak{b}) = \{\lambda \in K \mid \lambda\mathfrak{a} \subseteq \mathfrak{b}\}$$

The degree of a morphism $\lambda \in \mathrm{Hom}(\mathfrak{a}, \mathfrak{b})$ is $|\mathfrak{b}/\lambda\mathfrak{a}|$, or 0 if $\lambda = 0$. The fact that $\mathfrak{c} = \mathrm{Hom}(\mathfrak{a}, \mathfrak{b})$ is itself a fractional ideal, lets us define a group action $\mathfrak{c} \cdot \mathfrak{a} = \mathfrak{b}$ on the subcategories $\mathcal{I}(\mathcal{O})$ of objects $\mathfrak{a}$ with $\mathcal{O} = \mathrm{End}(\mathfrak{a})$.

A fractional ideal admits a normalized quadratic map: $\mathbb{N} : \mathfrak{a} \to \mathbb{Z}$ given by $\alpha \mapsto \mathrm{N}(\alpha)/\mathrm{N}(\mathfrak{a})$, but the association of a binary quadratic form requires the additional structure of a basis. While this category permits us to introduce well-defined ideal class groups (including ring class groups), in order to define additional level structure it will be necessary to add additional structure on this category. With a view to the equivalence with binary quadratic forms, we introduce a category with basis, then show, in each of the equivalent categories, how to model class groups with a Γ level structure for any subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$.

Fractional ideals with basis. The categories which follow parametrize more data than just a fractional ideal class. This motivates us to introduce a category $\mathcal{I}_\infty$ of pairs $(\mathfrak{a}, \mathcal{B})$, whose objects are fractional ideals with oriented basis $\mathcal{B} = (\omega_1, \omega_2)$ such that $\omega_2/\omega_1 \in \mathbb{H}_n$. A morphism of

$$\mathrm{Hom}\left((\mathfrak{a}, \mathcal{B}), (\mathfrak{b}, \mathcal{B}')\right)$$

in $\mathcal{I}_\infty$ is the set of $\lambda \in K$ such that $\lambda\mathfrak{a} = \mathfrak{b}$ and $\lambda\mathcal{B} = \mathcal{B}'$.

Remark. This category can be extended to include morphisms $\mathbb{Z} \simeq \mathrm{End}\left((\mathfrak{a}, \mathcal{B})\right) \longrightarrow \mathbb{Z} \times \mathbb{M}(\{I_2\})$, whose morphisms (m, mI_2) give $m\mathcal{B} = mI_2\mathcal{B}'$.

Binary quadratic forms category. The second category $\mathcal{F}_\infty$ consists of integral binary quadratic forms $f = ax^2 + bxy + cy^2$, which are primitive ($\gcd(a,b,c) = 1$), of discriminant $\Delta = b^2 - 4ac = m^2\Delta_K$. There exists an identity morphism in $\mathrm{Hom}(f_1, f_2)$ if $f_1 = f_2$, otherwise $\mathrm{Hom}(f_1, f_2)$ is empty. In particular each form f of $\mathcal{F}$ represents a distinct isomorphism class.

Remark. This category can be extended to include morphisms $\mathrm{End}(f) = \mathbb{M}(\{I_2\}) = \mathbb{Z}I_2$.

Poincaré CM points The third category $\mathcal{Q}_\infty$ consists of objects in $\mathbb{H} \cap K^\times$, such that the set of morphisms $\mathrm{Hom}(\tau_1, \tau_2)$ contains the identity if $\tau_1 = \tau_2$ and is otherwise empty. In particular each object is in a distinct isomorphism class in bijection with $\mathbb{H} \cap K^\times$.

Remark. This category is equivalent to the previous two, but the endomorphisms of the objects are trivial, since the induced action of mI_2 is trivial.

Establishing the equivalences of the three categories $\mathcal{I}_\infty$, $\mathcal{F}_\infty$, and $\mathcal{Q}_\infty$

The categories $\mathcal{I}_\infty$, $\mathcal{F}_\infty$, and $\mathcal{Q}_\infty$ should be viewed as universal lifts of the CM points:

$$\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{Q} \xrightarrow{\sim} \{j(\tau) \,|\, \tau \in \mathcal{Q}\} \subseteq X(1)(K^{ab}).$$

In particular $\mathbb{Z} \cong \mathbb{M}(\{I_2\})$ is the injective limit of the inclusions

$$\mathbb{M}(\{I_2\}) = \varinjlim_{n} \mathbb{M}(\Gamma(n)) = \bigcap_{n \geq 1} \mathbb{M}(\Gamma(n)) \subset \mathbb{M}_2(\mathbb{Z}).$$

The categorial description of Theorems 5.12, 5.13 and 5.14 goes as follows:

Lemma 5.15. *There exists a functor $F : \mathcal{I}_\infty \to \mathcal{F}_\infty$ sending $(\mathfrak{a}, (\omega_1, \omega_2))$ to*

$$f(x, y) = \frac{\mathrm{N}(\omega_1 x - \omega_2 y)}{\mathrm{N}(\mathfrak{a})}$$

and conversely, a functor $I : \mathcal{F}_\infty \to \mathcal{I}_\infty$ sending $f(x, y) = ax^2 + bxy + cy^2$ of discriminant $D = b^2 - 4ac$ to

$$(\omega_1, \omega_2) = (a, (-b + \sqrt{D})/2).$$

The functors F and I are equivalences of the categories $\mathcal{I}_\infty$ and $\mathcal{F}_\infty$, inverse to one another.

Lemma 5.16. *There exists a functor $Q : \mathcal{F}_\infty \to \mathcal{Q}_\infty$ sending $f(x, y) = ax^2 + bxy + cy^2$ of discriminant $D = b^2 - 4ac$ to*

$$\tau = \frac{-b + \sqrt{D}}{2a} \in \mathbb{H} \cap K^*.$$

Conversely, there exists a functor $G : \mathcal{Q}_\infty \to \mathcal{F}_\infty$ sending an object τ to the form

$$f(x, y) = a(x - \tau y)(x - \overline{\tau} y),$$

where $a \in \mathbb{Z}$ is the unique positive integer such that $f(x, y)$ is primitive.

5.1.5 Level structures

In this section we are going to discuss how this categories can be provided of additional level structure.

The usual subgroups $\Gamma_0(N)$, $\Gamma_1(N)$ and $\Gamma(N)$ produce the following subrings

$$\mathbb{M}(\Gamma_0(N)) = \mathbb{Z}[\Gamma_0(N)] = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_2(\mathbb{Z}) \,\middle|\, c \equiv 0 \ \text{ modulo } N \right\}$$

$$\mathbb{M}(\Gamma_1(N)) = \mathbb{Z}[\Gamma_1(N)] = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_2(\mathbb{Z}) \,\middle|\, c \equiv 0 \text{ and } a \equiv d \ \text{ modulo } N \right\}$$

Finally,

$$\mathbb{M}(\Gamma(N)) = \mathbb{Z}[\Gamma(N)] = \mathbb{Z} + N\mathbb{M}_2(\mathbb{Z})$$

For a given embedding $\iota : \mathcal{O}_K \to \mathbb{M}_2(\mathbb{Z})$, and $\pi : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ for and N a power of a prime q,

$$\overline{\iota} : (\mathcal{O}_K/n\mathcal{O}_K)^* \longrightarrow \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$$

we obtain an associated split $\Gamma_s(N)$ or non-split $\Gamma_{ns}(N)$ Cartan subgroup

$$\pi^{-1}\big(\overline{\iota}\left((\mathcal{O}_K/n\mathcal{O}_K)^*\right)\big)$$

depending on whether q is split or inert in K.

The forgetful functor $\mathcal{I}_\infty \to \mathcal{I}$ sending $(\mathfrak{a}, \mathcal{B})$ to $\mathfrak{a}$ is the quotient by the action of $\mathrm{SL}_2(\mathbb{Z})$. Specifically, the set of objects mapping to $\mathfrak{a}$ are $(\mathfrak{a}, \gamma\mathcal{B})$ where $\mathcal{B}$ is any basis for $\mathfrak{a}$ and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. However, in $\mathcal{I}_\infty$, each such pair $(\mathfrak{a}, \gamma\mathcal{B})$ represents a distinct isomorphism class. This motivates the introduction of a category $\mathcal{I}(\mathrm{SL}_2(\mathbb{Z}))$, equipped with a richer structure of morphisms, which on objects is the quotient of $\mathcal{I}_\infty$ by $\mathrm{SL}_2(\mathbb{Z})$. Specifically, a morphism $(\mathfrak{a}, \mathcal{B})$ to $(\mathfrak{b}, \mathcal{B}')$ is a pair $(\lambda, \gamma) \in K \times \mathbb{M}_2(\mathbb{Z})$ such that

$$\lambda\mathfrak{a} \subseteq \mathfrak{b} \text{ and } \lambda\mathcal{B} = \gamma\mathcal{B}'$$

By construction $\mathcal{I}(\mathrm{SL}_2(\mathbb{Z}))$ is equivalent to $\mathcal{I}$, and the composition

$$\mathcal{O} = \mathrm{End}(\mathfrak{a}) \longrightarrow \mathrm{End}((\mathfrak{a}, \mathcal{B})) \subset \mathcal{O} \times \mathbb{M}_2(\mathbb{Z}) \longrightarrow \mathbb{M}_2(\mathbb{Z})$$

determines a homomorphism of rings $\mathcal{O} \to \mathbb{M}_2(\mathbb{Z})$. In particular, $-1 \in \mathcal{O}$ maps to $-I_2 \in \mathbb{M}_2(\mathbb{Z})$, and the automorphism groups

$$\langle(-1, -I_2)\rangle \subseteq \mathrm{Aut}((\mathfrak{a}, \mathcal{B})) = \mathrm{End}((\mathfrak{a}, \mathcal{B}))^*$$

with equality unless the endomorphism ring of $\mathfrak{a}$ admits extra automorphisms.

Example. We consider the automorphism groups in $\mathcal{I}(\mathrm{SL}_2(\mathbb{Z}))$. If $\mathcal{O} = \mathfrak{a} = \mathbb{Z}[i]$ with $i^2 = -1$, then the automorphism group has order 4:

$$\mathrm{Aut}((\mathbb{Z}[i], (1, i))) = \langle (i, S) \rangle \text{ where } S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and if $\mathcal{O} = \mathfrak{a} = \mathbb{Z}[\zeta_3]$, where $\zeta_3^2 + \zeta_3 + 1 = 0$, the automorphism group has order 6:

$$\mathrm{Aut}((\mathbb{Z}[\zeta_3], (1, \zeta_3))) = \langle \pm(\zeta_3, W) \rangle \text{ where } W = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$$

Every other ideal not homothetic to $\mathbb{Z}[i]$ or $\mathbb{Z}[\zeta_3]$ has trivial automorphism group $\{\pm(1, I_2)\}$. These examples exhibit the induced isomorphisms

$$\boldsymbol{\mu}_4 = \langle i \rangle \to \langle S \rangle \text{ and } \boldsymbol{\mu}_6 = \langle \pm\zeta_3 \rangle \to \langle \pm W \rangle,$$

with subgroups $\boldsymbol{\mu}_2 = \langle -1 \rangle \to \langle -I_2 = S^2 \rangle$.

In order to classify objects with level structure, we define equivalent categories $\mathcal{I}(\Gamma)$, $\mathcal{F}(\Gamma)$ and $\mathcal{Q}(\Gamma)$ for which the isomorphism classes are equivalence classes under the group $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$. We can interpret the categories $\mathcal{I}_\infty$, $\mathcal{F}_\infty$ and $\mathcal{Q}_\infty$ as projective limits of the categories $\mathcal{I}(\Gamma)$, $\mathcal{F}(\Gamma)$ and $\mathcal{Q}(\Gamma)$.

Ideals with Γ-structure. The set of morphisms $\mathrm{Hom}((\mathfrak{a}, \mathcal{B}), (\mathfrak{b}, \mathcal{B}'))$ in the category $\mathcal{I}(\Gamma)$ are pairs $(\lambda, \gamma) \in K \times \mathbb{M}(\Gamma)$ such that $\lambda\mathfrak{a} \subseteq \mathfrak{b}$ and $\lambda\mathcal{B} = \gamma\mathcal{B}'$. The degree of a morphism, $|\mathfrak{b}/\lambda\mathfrak{a}| = \det(\gamma)$, is well-defined, and the dual morphism is $(\lambda^{-1}\det(\gamma), \gamma^*)$.

Binary quadratic forms with Γ-structure. The category $\mathcal{F}(\Gamma)$ consists of primitive integral binary quadratic forms $f = ax^2 + bxy + cy^2$ of discriminant $D = b^2 - 4ac = m^2 D_K$. The set of morphisms $\mathrm{Hom}(f_1, f_2)$ consists of proper integral similarities $(n, \gamma) \in \mathbb{N} \times \mathbb{M}(\Gamma)$ such that

$$nf_1(x, y) = (f_2 \cdot \gamma)(x, y).$$

The integer $\det(\gamma) \geq 0$ is the degree of the morphism, and the dual morphism of (n, γ) is $(n^{-1}\det(\gamma)^2, \gamma^*)$.

N.B. The scalar n is uniquely determined by from γ and (f_1, f_2). In particular, if f_1 and f_2 are of conductors m_1 and m_2, i.e. their discriminants are respectively $m_1^2 D_K$ and $m_2^2 D_K$, then comparing the conductors of the transformed forms gives the relation $nm_1 = \det(\gamma)m_2$, and if $m_1 = m_2$, we have $n = \det(\gamma)$ and the dual morphism is (n, γ^*). This corresponds to a horizontal isogeny.

Poincaré CM points with Γ-structure. The category $\mathcal{Q}(\Gamma)$ consists of objects in $\mathbb{H} \cap K^\times$, whose set of morphisms $\mathrm{Hom}(\tau_1, \tau_2)$ consists of elements $\gamma \in \mathbb{M}(\Gamma)$, with $\gamma = 0$, or $\det(\gamma) > 0$ such that $\gamma(\tau_1) = \tau_2$. The degree of γ is defined to be $\det(\gamma)$ and the dual morphism is γ^*.

Example (Endomorphisms of τ)**.** . Let τ be an object of $\mathcal{Q}(\Gamma)$ satisfying the quadratic equation $a\tau^2 + b\tau + c = 0$. While τ represents the ideal class of any ideal $\mathfrak{a}$ with oriented basis of the form $(\omega_1, \omega_2) = (\omega_1, \tau\omega_1)$, we show that the endomorphism ring of $\mathrm{End}(\tau)$ is nevertheless nontrivial.

Coving and quotient functors. For each inclusion $\Gamma_2 \subseteq \Gamma_1$ in $\mathrm{SL}_2(\mathbb{Z})$ there exists a functor $\mathcal{I}(\Gamma_2) \to \mathcal{I}(\Gamma_1)$ compatible with the covering functor from $\mathcal{I}_\infty$ and forgetful functors to $\mathcal{I}$:

$$\begin{tikzcd}
\mathcal{I}_\infty \arrow[dr] \arrow[drr, bend left] & & & \\
& \mathcal{I}(\Gamma_2) \arrow[r] \arrow[drr, bend right] & \mathcal{I}(\Gamma_1) \arrow[dr] & \\
& & & \mathcal{I},
\end{tikzcd}$$

We identify $\mathcal{I}(\mathrm{SL}_2(\mathbb{Z})$ with $\mathcal{I}$ via the equivalence of categories induced by the forgetful functor to $\mathcal{I}$.

5.2 Isogeny graphs

Definition. Given an elliptic curve E defined over a field k and a finite set of primes S coprime to the characteristic of k, we associate the *isogeny graph* $G = G_S(E)$ such that

Vertices: $\overline{k}$-isomorphism classes of elliptic curves over $\overline{k}$ which are $\overline{k}$-isogenous to E.

Edges: Isogenies of degree $\ell \in S$, up to isomorphism.

If $S = \{\ell\}$, we call G the ℓ-isogeny graph of E and denote it $G_\ell(E)$.

Remark. Since we consider isomorphisms over an algebraic closure, nodes are parametrized by j-invariants.

Remark. In the literature one may find different definitions of an isogeny graphs. The differences are mainly due to the field over which curves, isomorphism and isogenies are considered. See for instance [Pan].

The isogeny graph is, a priori, a directed graph since isogenies have well defined domain and codomain. However, the existence of dual isogenies (see Section 1.2.3) shows that, in practice, one can think of these objects as undirected graphs by identifying each isogeny with its dual. The caution used in the last statement is due to the fact that G_S is only "nearly" an undirected graph. Vertices are defined up to $\overline{k}$-isomorphisms an edges from a given vertex are defined up to isomorphisms of the codomain, i.e., they are identified post composition with an automorphism of the second curve. This means that at curves with extra-automorphisms (out the usual ones ± 1), this identification may fail and therefore, at neighboring curves to the problematic ones, the number of outgoing and ingoing isogenies might differ.

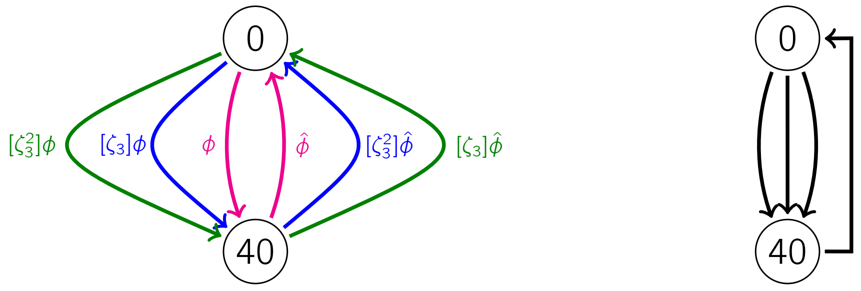
The next theorem shows that this occurs only in a handful of cases and, therefore, one can think of G_S as undirected without losing global information.

Theorem 5.17 ([Sil1, Th. III.10.1])**.** *Let E be an elliptic curve defined over k. The automorphism group of E is a finite group of cardinality dividing* 24*.*

$$\mathrm{Aut}(E) = \begin{cases} C_2 & \text{if } j(E) \neq 0, 1728 \\ C_4 & \text{if } j(E) = 1728 \text{ and } \mathrm{char}(k) \neq 2,3 \\ C_6 & \text{if } j(E) = 0 \text{ and } \mathrm{char}(k) \neq 2,3 \\ C_3 \rtimes C_4 & \text{if } j(E) = 0 = 1728 \text{ and } \mathrm{char}(k) = 3 \\ Q_4 \rtimes C_3 & \text{if } j(E) = 0 = 1728 \text{ and } \mathrm{char}(k) = 2 \end{cases} \qquad \#\mathrm{Aut}(E) = \begin{cases} 2 \\ 3 \\ 6 \\ 12 \\ 24 \end{cases}$$

where C_i is the cyclic group of order i and Q_4 represents the quaternion group $\{\pm 1, \pm i, \pm j, \pm k\}$.

Example. Let us consider $p = 71$. $E_0/\mathbb{F}_{71}$ has 6 automorphisms, $\mathrm{Aut}(E_0) = \{[\pm 1, \pm\zeta_3, \pm\zeta_3^2]\}$. There are three isogenies $E_0 \longrightarrow E_{40}$ and their duals differ by post composition with an automorphism of E_0 which means that they result in a single arrow in the isogeny graph (on the right).



An ℓ-isogeny graph Γ is equipped with an action of $\mathcal{G} = \mathcal{G}al(\overline{k}/k)$, with the vertex $[E]$ a fixed point, as follows. We have

$$E[\ell] = \{P \in E(\bar{k}) \mid \ell P = O\} \cong (\mathbb{Z}/\ell\mathbb{Z})^2.$$

The set cyclic subgroups is in bijection with $\mathbb{P}(E[\ell]) \cong \mathbb{P}^1(\mathbb{Z}/\ell\mathbb{Z})$, which in turn is in bijection with the set of ℓ-isogenies from E. The $\mathcal{G}$-action on $E[\ell]$ induces an action by $\mathcal{G}$ on the $\ell + 1$ cyclic subgroups. This action extends to paths without backtracking of length n, via the action on the cyclic subgroups G of order ℓ^n in

$$E[\ell^n] = \{P \in E(\bar{k}) \mid \ell^n P = O\} \cong (\mathbb{Z}/\ell^n\mathbb{Z})^2.$$

which are in bijection with $\mathbb{P}(E[\ell^n]) \cong \mathbb{P}^1(\mathbb{Z}/\ell^n\mathbb{Z})$. This determines a compatible Galois action on vertices $[E/G]$ and edges $\varphi : E/G_i \to E/G_{i+1}$ where $G_i \subset G_{i+1}$ is of index ℓ.

Thus, an ℓ-isogeny graph is $(\ell+1)$-regular for outgoing edges. The existence of curves of j-invariant 0 or 12^3 with additional automorphisms in the graph implies a reduced number of incoming edges at these vertices.

Remark. This regularity only appears on the algebraic closure. If we restrict to different fields of definition in constructing the isogeny graphs we will obtain more irregular subgraphs. A result of Atkin [Atk1; Atk2] shows that the number of distinct $\mathbb{F}_q$-rational ℓ-isogenies of E equals the number of linear factors of $\Phi_\ell(j(E), y)$, the ℓ-modular polynomial, over $\mathbb{F}_q$ and this is either 0,1,2 or $\ell+1$.

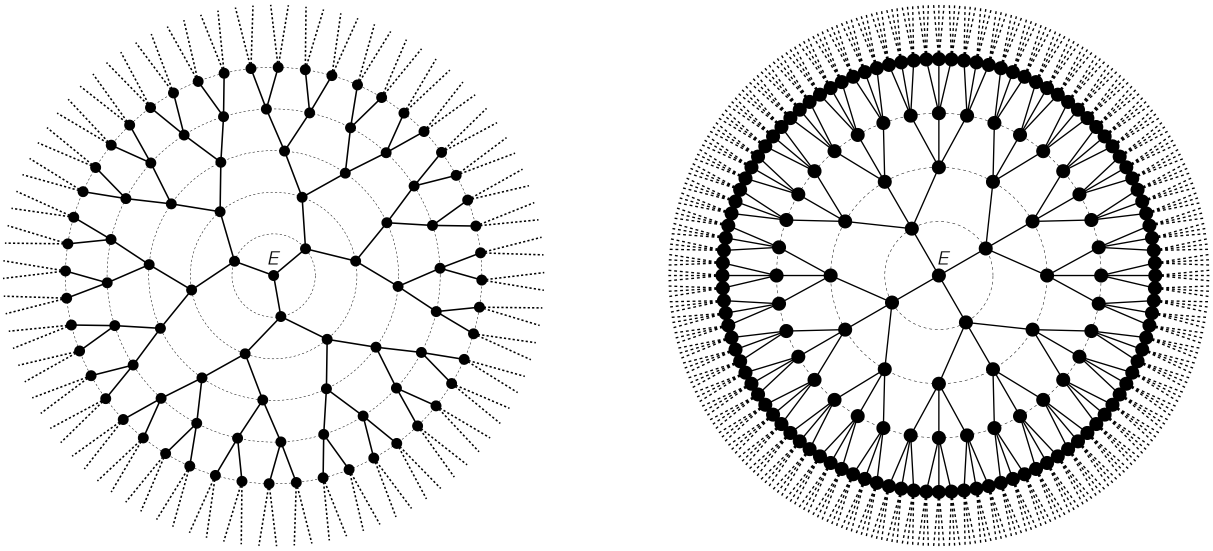


Figure 5.3 – Isogeny graphs $G_2(E)$ and $G_3(E)$ for an elliptic curve E without complex multiplication

Definition. We define an undirected quotient graph $\overline{G}_\ell(E)$ by identifying an isogeny $\varphi : E_0 \to E_1$ with its dual $\hat{\varphi} : E_0 \to E_1$, and if $\mathrm{Aut}(E_0) \neq \{\pm 1\}$ or $\mathrm{Aut}(E_1) \neq \{\pm 1\}$ the orbits

$$\mathrm{Aut}(E_1)\varphi\mathrm{Aut}(E_0) \text{ and } \mathrm{Aut}(E_0)\hat{\varphi}\mathrm{Aut}(E_1)$$

are identified with φ and $\hat{\varphi}$, respectively. Taking the quotient gives a bijective correspondence between edges and dual edges. On the contrary, without taking the quotient, there exist problematic vertices.

The action of the Galois group can be extended to the whole graph by the Galois action on the projective Tate module $\mathbb{P}(T_\ell(E)) \cong \mathbb{P}^1(\mathbb{Z}_\ell)$. In the same way we define the $\mathcal{Gal}$-action on $G_S(E)$ derived from the $\mathcal{Gal}$-set structure of $\mathbb{P}(T_S(E))$, where

$$T_S(E) = \prod_{\ell \in S} T_\ell(E).$$

The choice of base curve E distinguishes the Galois action on G from a conjugate Galois action derived from a twist of E.

If we fix a basis of $T_\ell(E)$ we get a representation

$$\mathcal{Gal}(\overline{k}/k) \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell)$$

permitting one to associate to each vertex of $G_\ell(E)$ a matrix in $\mathrm{GL}_2(\mathbb{Q}_\ell)$. One can prove that the matrix associated to $\phi : E_1 \to E_2$ is one of the form

$$\begin{pmatrix} \ell & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & a \\ 0 & \ell \end{pmatrix} \quad \text{where } 0 \leq a < \ell$$

and by composing them we obtain matrices for ℓ^n isogenies. The resulting structure is an infinite tree on $\mathrm{PSL}_2(\mathbb{Q}_\ell)/\mathrm{PSL}_2(\mathbb{Z}_\ell)$.

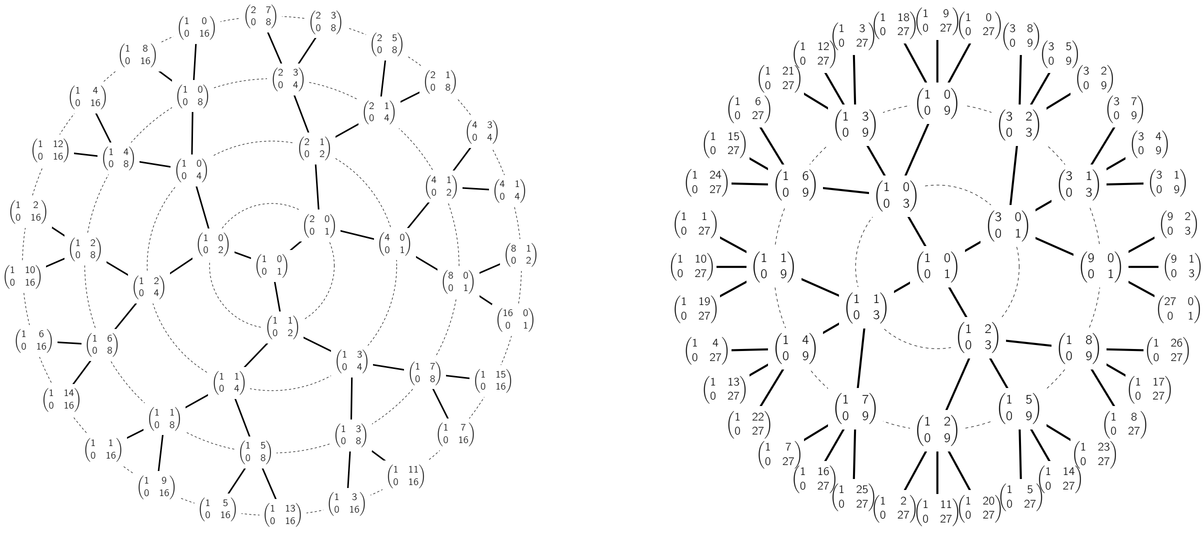

$\begin{pmatrix}1&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&0\\0&2\end{pmatrix}$
$\begin{pmatrix}2&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&1\\0&2\end{pmatrix}$
$\begin{pmatrix}1&0\\0&4\end{pmatrix}$
$\begin{pmatrix}2&1\\0&2\end{pmatrix}$
$\begin{pmatrix}4&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&3\\0&4\end{pmatrix}$
$\begin{pmatrix}1&1\\0&4\end{pmatrix}$
$\begin{pmatrix}1&2\\0&4\end{pmatrix}$
$\begin{pmatrix}1&0\\0&8\end{pmatrix}$
$\begin{pmatrix}2&3\\0&4\end{pmatrix}$
$\begin{pmatrix}2&1\\0&4\end{pmatrix}$
$\begin{pmatrix}4&1\\0&2\end{pmatrix}$
$\begin{pmatrix}8&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&7\\0&8\end{pmatrix}$
$\begin{pmatrix}1&3\\0&8\end{pmatrix}$
$\begin{pmatrix}1&5\\0&8\end{pmatrix}$
$\begin{pmatrix}1&1\\0&8\end{pmatrix}$
$\begin{pmatrix}1&6\\0&8\end{pmatrix}$
$\begin{pmatrix}1&2\\0&8\end{pmatrix}$
$\begin{pmatrix}1&4\\0&8\end{pmatrix}$
$\begin{pmatrix}1&0\\0&16\end{pmatrix}$
$\begin{pmatrix}2&7\\0&8\end{pmatrix}$
$\begin{pmatrix}2&3\\0&8\end{pmatrix}$
$\begin{pmatrix}2&5\\0&8\end{pmatrix}$
$\begin{pmatrix}2&1\\0&8\end{pmatrix}$
$\begin{pmatrix}4&3\\0&4\end{pmatrix}$
$\begin{pmatrix}4&1\\0&4\end{pmatrix}$
$\begin{pmatrix}8&1\\0&2\end{pmatrix}$
$\begin{pmatrix}16&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&15\\0&16\end{pmatrix}$
$\begin{pmatrix}1&7\\0&16\end{pmatrix}$
$\begin{pmatrix}1&11\\0&16\end{pmatrix}$
$\begin{pmatrix}1&3\\0&16\end{pmatrix}$
$\begin{pmatrix}1&13\\0&16\end{pmatrix}$
$\begin{pmatrix}1&5\\0&16\end{pmatrix}$
$\begin{pmatrix}1&9\\0&16\end{pmatrix}$
$\begin{pmatrix}1&1\\0&16\end{pmatrix}$
$\begin{pmatrix}1&14\\0&16\end{pmatrix}$
$\begin{pmatrix}1&6\\0&16\end{pmatrix}$
$\begin{pmatrix}1&10\\0&16\end{pmatrix}$
$\begin{pmatrix}1&2\\0&16\end{pmatrix}$
$\begin{pmatrix}1&12\\0&16\end{pmatrix}$
$\begin{pmatrix}1&4\\0&16\end{pmatrix}$
$\begin{pmatrix}1&8\\0&16\end{pmatrix}$
$\begin{pmatrix}1&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&0\\0&3\end{pmatrix}$
$\begin{pmatrix}3&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&1\\0&3\end{pmatrix}$
$\begin{pmatrix}1&2\\0&3\end{pmatrix}$
$\begin{pmatrix}1&0\\0&9\end{pmatrix}$
$\begin{pmatrix}1&3\\0&9\end{pmatrix}$
$\begin{pmatrix}1&6\\0&9\end{pmatrix}$
$\begin{pmatrix}3&2\\0&3\end{pmatrix}$
$\begin{pmatrix}3&1\\0&3\end{pmatrix}$
$\begin{pmatrix}9&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&8\\0&9\end{pmatrix}$
$\begin{pmatrix}1&5\\0&9\end{pmatrix}$
$\begin{pmatrix}1&2\\0&9\end{pmatrix}$
$\begin{pmatrix}1&7\\0&9\end{pmatrix}$
$\begin{pmatrix}1&4\\0&9\end{pmatrix}$
$\begin{pmatrix}1&1\\0&9\end{pmatrix}$
$\begin{pmatrix}1&3\\0&27\end{pmatrix}$
$\begin{pmatrix}1&18\\0&27\end{pmatrix}$
$\begin{pmatrix}1&9\\0&27\end{pmatrix}$
$\begin{pmatrix}1&0\\0&27\end{pmatrix}$
$\begin{pmatrix}3&8\\0&9\end{pmatrix}$
$\begin{pmatrix}3&5\\0&9\end{pmatrix}$
$\begin{pmatrix}3&2\\0&9\end{pmatrix}$
$\begin{pmatrix}3&7\\0&9\end{pmatrix}$
$\begin{pmatrix}3&4\\0&9\end{pmatrix}$
$\begin{pmatrix}3&1\\0&9\end{pmatrix}$
$\begin{pmatrix}9&2\\0&3\end{pmatrix}$
$\begin{pmatrix}9&1\\0&3\end{pmatrix}$
$\begin{pmatrix}27&0\\0&1\end{pmatrix}$
$\begin{pmatrix}1&26\\0&27\end{pmatrix}$
$\begin{pmatrix}1&17\\0&27\end{pmatrix}$
$\begin{pmatrix}1&8\\0&27\end{pmatrix}$
$\begin{pmatrix}1&23\\0&27\end{pmatrix}$
$\begin{pmatrix}1&14\\0&27\end{pmatrix}$
$\begin{pmatrix}1&5\\0&27\end{pmatrix}$
$\begin{pmatrix}1&20\\0&27\end{pmatrix}$
$\begin{pmatrix}1&11\\0&27\end{pmatrix}$
$\begin{pmatrix}1&2\\0&27\end{pmatrix}$
$\begin{pmatrix}1&25\\0&27\end{pmatrix}$
$\begin{pmatrix}1&16\\0&27\end{pmatrix}$
$\begin{pmatrix}1&7\\0&27\end{pmatrix}$
$\begin{pmatrix}1&22\\0&27\end{pmatrix}$
$\begin{pmatrix}1&13\\0&27\end{pmatrix}$
$\begin{pmatrix}1&4\\0&27\end{pmatrix}$
$\begin{pmatrix}1&19\\0&27\end{pmatrix}$
$\begin{pmatrix}1&10\\0&27\end{pmatrix}$
$\begin{pmatrix}1&1\\0&27\end{pmatrix}$
$\begin{pmatrix}1&24\\0&27\end{pmatrix}$
$\begin{pmatrix}1&15\\0&27\end{pmatrix}$
$\begin{pmatrix}1&6\\0&27\end{pmatrix}$
$\begin{pmatrix}1&21\\0&27\end{pmatrix}$
$\begin{pmatrix}1&12\\0&27\end{pmatrix}$

Figure 5.4 – Serre trees associated to $G_2(E)$ and $G_3(E)$ for an elliptic curve E without complex multiplication

In the next sections we will describe isogeny graphs for curves over finite fields. These has been studied by Mestre [Mes], Pizer [Piz3] and Kohel [Koh1]. By Theorem 1.20, we know that isogenous curves are always either both ordinary, or both supersingular and we will see that the structure of their isogeny graph is very different. We start by a preliminary classification.

Remark. Being ordinary or supersingular depends on the trace of the Frobenius endomorphism, see Theorem 1.21. The invariance of the trace of Frobenius in an isogeny class dates back to Hasse (1936) for curves, and Weil (1946) for abelian varieties in general.

Lemma 5.18. *Let E be an elliptic curve over k with endomorphism ring $\mathcal{O}$, and for a prime $\ell \neq \operatorname{char}(k)$ let $\overline{G}_\ell(E)$ be its undirected ℓ-isogeny graph.*

1. *If $\mathcal{O} = \mathbb{Z}$, then each component of $\overline{G}_\ell(E)$ is an infinite tree.*
2. *If $\mathcal{O}$ is an order in a CM field K, then each component $\overline{G}$ of $\overline{G}_\ell(E)$ is infinite and either*
   - *the prime ℓ is split in K and $\overline{G}$ has a unique cycle, or*
   - *the prime ℓ is ramified or inert in K and $\overline{G}$ is a tree.*
3. *If $\mathcal{O}$ is an order in a quaternion algebra, then $\overline{G}_\ell(E)$ is finite and connected.*

5.2.1 Ordinary isogeny graphs

We start by studying the ordinary case. Let E be an elliptic curve defined over a finite field k with complex multiplication by a quadratic imaginary field K, by which we mean that its endomorphism algebra is K. We denote Δ_K the discriminant of $\mathcal{O}_K$, the ring of integers of K. We know that any order $\mathcal{O}$ in K is completely determined by its conductor f, or equivalently by its discriminant $\Delta = f^2\Delta_K$. In general, if π is the Frobenius endomorphism of E we have the following chain of inclusions

$$\mathbb{Z}[\pi] \subseteq \operatorname{End}_k(E) \subseteq \mathcal{O}_K$$

The goal of the section is to describe the ℓ-isogeny graph $G_\ell(E)$. This consists of all elliptic curves over k with complex multiplication by K (up to $\overline{k}$ isomorphism) by Theorem 1.20. We therefore need to study this set and find a convenient way of describing isogenies.

Main theorem of class field theory I. Let $\mathcal{C}\ell(\mathcal{O}) \simeq \mathcal{C}\ell(\Delta)$ denote the class group of the order $\mathcal{O}$ with discriminant Δ and $h(\mathcal{O}) = h(\Delta)$ its cardinality, namely the class number of $\mathcal{O}$. Class field theory tells us that there exists a number field $H_{\mathcal{O}}$, called the Hilbert class field (or the ring class field) of K with Galois group isomorphic to the class group of $\mathcal{O}$, see Theorems 5.9 and 5.11.

Further, the class group of $\mathcal{O}$ is in bijection with the set of K-lattices of order $\mathcal{O}$ (see 5.8) which in turns is equivalent to the set of elliptic curves over $\mathbb{C}$ with endomorphism ring $\mathcal{O}$, see Section 1.2.5. Thus, for every fractional ideal $\mathfrak{a} \subseteq \mathcal{O}$ we associate the complex elliptic curve $\mathbb{C}/\mathfrak{a}$ corresponding to the lattice attached to $\mathfrak{a}$; this is defined over $H_{\mathcal{O}}$ and has endomorphism ring $\mathcal{O}$.

Theorem 5.19 (Deuring Reduction Theorem)**.** *Let E be an elliptic curve over a number field, with* $\mathrm{End}(E) \simeq \mathcal{O}$ *for and order $\mathcal{O}$ in a quadratic imaginary field K. Let $\mathfrak{P}$ be a place of $\overline{\mathbb{Q}}$ above a prime p in $\mathbb{Q}$, where E has non-degenerate reduction $\overline{E}$, i.e., $\overline{E}$ is again a smooth curve. Suppose that p splits completely in K, let f be the conductor of $\mathcal{O}$ and write $f = p^r f_0$ with $(p, f_0) = 1$. Then we have an isomorphism*

$$\mathrm{End}(E) \longrightarrow \mathrm{End}(\overline{E})$$
$$\lambda \longmapsto \overline{\lambda}$$

Proof. See [Lan2, Th. 13.4.12]. □

Over $\mathbb{F}_p$ this reduction theorem is even more meaningful. As above, let $\mathcal{O}$ be an order in an imaginary quadratic field K, $H_{\mathcal{O}}$ be the ring class field of $\mathcal{O}$ and p a prime which splits completely in $H_{\mathcal{O}}$. We will fix a prime $\mathfrak{P}$ of $H_{\mathcal{O}}$ lying above p, so that $\mathcal{O}_H/\mathfrak{P} \simeq \mathbb{F}_p$.

Theorem 5.20. *Let E be an elliptic curve over $H_{\mathcal{O}}$ which has good reduction at $\mathfrak{P}$ and endomorphism ring* $\mathrm{End}(E) \simeq \mathcal{O}$*. By the Deuring reduction theorem, $\overline{E}$ is an elliptic curve over $\mathbb{F}_p$. Then there exists $\lambda \in \mathcal{O}$ such that $p = \lambda\overline{\lambda}$ and $\#\overline{E}(\mathbb{F}_p) = p + 1 - (\lambda + \overline{\lambda})$. Further,* $\mathrm{End}(\overline{E}) = \mathcal{O}$ *and every elliptic curve over $\mathbb{F}_p$ with endomorphism ring isomorphic to $\mathcal{O}$ arises in this way.*

In other words, there exists an isomorphism (which depends on the choice of the prime $\mathfrak{p}$ above p in K) between fractional ideals classes in $\mathcal{C}\ell(\mathcal{O})$ and isomorphism classes of elliptic curves $E/\mathbb{F}_p$ with $\mathrm{End}(E) \simeq \mathcal{O}$.

The next theorem shows that the reduction process can be reversed also if we move away from $\mathbb{F}_p$.

Theorem 5.21 (Deuring Lifting Theorem)**.** *Let E be an elliptic curve in characteristic p, with a non-trivial endomorphism α. Then there exists an elliptic curve $\tilde{E}$ defined over a number field, an endomorphism $\tilde{\alpha}$ of $\tilde{E}$, and a non-degenerate reduction of $\tilde{E}$ at a place $\mathfrak{P}$ lying above p, such that we have an isomorphism $E \simeq \overline{(\tilde{E})}$ and α correspond to the reduction of $\tilde{\alpha}$ under this isomorphism.*

Proof. See [Lan2, Th. 13.5.14]. □

Main theorem of class field theory II. With the notation as above we can state a beautiful result telling us how the Galois group acts on j-invariants of CM elliptic curves. Let K/Q be an imaginary quadratic field and let $E/\mathbb{C}$ be an elliptic curve with $End^0(E) \simeq K$.

Theorem 5.22. *Let $\sigma \in$* $\mathrm{Aut}(\mathbb{C})$ *and $\mathfrak{s} \in \mathbf{I}_K$ be an idèle corresponding to the restriction of σ to K^{ab} via the Artin map, $[\mathfrak{s}, K] = \sigma|_{K^{\mathrm{ab}}}$. Fix an analytic isomorphism $\xi : \mathbb{C}/\mathfrak{a} \xrightarrow{\sim} E(\mathbb{C})$ for a fractional ideal $\mathfrak{a}$ of K. Then there exists a unique complex analytic isomorphism*

$$\xi_{\mathfrak{s}} : \mathbb{C}/\mathfrak{s}^{-1}\mathfrak{a} \longrightarrow E^{\sigma}$$

such that the following diagram commutes

$$\begin{tikzcd}
K/\mathfrak{a} \arrow[r, "\xi"] \arrow[d, "\mathfrak{s}^{-1}"'] & E \arrow[d, "\sigma"] \\
K/\mathfrak{s}^{-1}\mathfrak{a} \arrow[r, "\xi_{\mathfrak{s}}"'] & E^{\sigma}
\end{tikzcd}$$

Proof. See [Shi, Th. 5.4]. □

Principal homogeneous space. The results above bring us to the following theorem. Let k be a finite field, K a quadratic imaginary field with ring of integers $\mathcal{O}_K$ and $\mathcal{O}$ and order in $\mathcal{O}_K$. Let us denote $\mathrm{Ell}_k(\mathcal{O})$ the set of k-isomorphism classes of elliptic curves over k with endomorphism ring isomorphic to $\mathcal{O}$.In defining $\mathrm{Ell}_k(\mathcal{O})$ we suppose that every elliptic curve $E \in \mathrm{Ell}_k(\mathcal{O})$ comes equipped with a fixed isomorphism $\iota_E : \mathcal{O} \to \mathrm{End}(E)$.

Theorem 5.23 ([Wat1, Th. 4.5] and [SchR, Th. 4.5])**.** *If $p =$* $\mathrm{char}(k) > 0$ *is not inert in $\mathcal{O}$, then the set* $\mathrm{Ell}_k(\mathcal{O})$ *is a principal homogeneous space for the group $\mathcal{C}\ell(\mathcal{O})$.*

Being a principal homogeneous space (or a torsor) means that there is an action

$$\begin{aligned} * : \mathcal{C}\!\ell(\mathcal{O}) \times \mathrm{Ell}_k(\mathcal{O}) &\longrightarrow \mathrm{Ell}_k(\mathcal{O}) \\ (\mathfrak{a}, E) &\longmapsto \mathfrak{a} * E \end{aligned}$$

and this action is free and transitive; thus, for every fixed elliptic curve E, it induces a bijection

$$\begin{aligned} \mathcal{C}\!\ell(\mathcal{O}) &\longrightarrow \mathrm{Ell}_k(\mathcal{O}) \\ [\mathfrak{a}] &\longmapsto [\mathfrak{a} * E] \end{aligned}$$

The action $*$ is defined by the composition of theorems 5.19, 5.21 and 5.22. More explicitly, for a fractional ideal $\mathfrak{a}$, let us define

$$E[\mathfrak{a}] = \{P \in E(\overline{k}) \mid \alpha P = O \text{ for all } \alpha \in \mathfrak{a}\}$$

This is a finite subgroup of E of cardinality $\mathrm{N}(\mathfrak{a})$ and therefore induces the separable isogeny $\phi_{\mathfrak{a}} : E \to E/E[\mathfrak{a}]$. Then $\mathfrak{a} * E$ corresponds to the isomorphism class of $E/E[\mathfrak{a}]$.

All these results answer the question about the description of Elliptic curves over k with endomorphism ring an order in K. It remains to study relations between them.

The isogeny volcano

Let us fix an elliptic curve E over a finite field k with endomorphism ring an order $\mathcal{O}$ in a quadratic imaginary field K of discriminant Δ_K. We denote f the conductor of $\mathcal{O}$ and $\Delta = f^2\Delta_K$ its discriminant. Finally, let ℓ be a prime.

Following the terminology of Kohel [Koh1], if $\mathcal{O}$ is maximal at ℓ we say that E lies at the surface (relative to ℓ), if $\mathcal{O}$ has index divisible by ℓ^t and not by ℓ^{t+1}, then we say that E lies at depth or level t and if the index of $\mathbb{Z}[\pi]$ in $\mathcal{O}$ is not divisible by ℓ we say that E lies at the floor.

Definition. An ℓ-volcano is an undirected graph consisting of a cycle (which may be as well reduced to a single node) where each node is the root of a complete tree of same length. The cycle is called surface or crater, the level (or depth) of a node is the distance from the surface and the floor consists of nodes at maximal distance from the crater. Edges between nodes at the same depth are called horizontal, those from a certain level to a higher one are called ascending and, those from a level to a lower depth are said to be descending.

Kohel shows that CM isogeny graphs are, in fact, volcanoes.

Theorem 5.24 ([Koh1, Prop. 21])**.** *Let E_1 and E_2 be ordinary elliptic curves over k and $\phi : E_1 \to E_2$ an isogeny between them of prime degree $\ell \neq \operatorname{char} k$. Then $\mathcal{O}_1 = \operatorname{End}(E_1)$ contains $\mathcal{O}_2 = \operatorname{End}(E_2)$ or $\mathcal{O}_2$ contains $\mathcal{O}_1$ and the index of one in the other divides ℓ.*

With the terminology above, we say that

- ϕ is *horizontal* if $\mathcal{O}_1 = \mathcal{O}_2$;
- ϕ is *ascending* if $\mathcal{O}_1 \subset \mathcal{O}_2$ with index ℓ
- ϕ is *descending* if $\mathcal{O}_2 \subset \mathcal{O}_1$ with index ℓ

Theorem 5.25. *The ℓ-isogeny graph $G_\ell(E)$ of an ordinary elliptic curve E is an ℓ-volcano.*

It only remains to study the size of the crater (how many elliptic curves lie at the surface), the height of it and its degree.

Horizontal isogenies. Horizontal isogenies come from the action of a fractional ideal $\mathfrak{a}$ of $\mathcal{O}$ as described above [Koh1, Prop. 22]. If ℓ divides the index of $\mathcal{O}$ in $\mathcal{O}_K$ then no such isogeny exists. Otherwise, if we are on the crater, the number of isogenies to curves with the same endomorphism ring is

$$1 + \left(\frac{\Delta_{\mathcal{O}}}{\ell}\right) = \begin{cases} 0 & \text{if } \ell \text{ is inert in } \mathcal{O} \\ 1 & \text{if } \ell \text{ is ramified in } \mathcal{O} \\ 2 & \text{if } \ell \text{ splits in } \mathcal{O} \end{cases}$$

Therefore the crater consists of either

(a) a single node with no horizontal isogenies;

(b) two (isomorphism classes of) elliptic curves with one horizontal isogeny between them or a single node with a horizontal loop;

(c) a cycle of isomorphism classes of elliptic curves.

Note that if ℓ splits as $\ell\mathcal{O} = \mathfrak{l}\bar{\mathfrak{l}}$, the isogenies ϕ_ℓ and $\phi_{\bar{\mathfrak{l}}}$ are one the dual of the other.

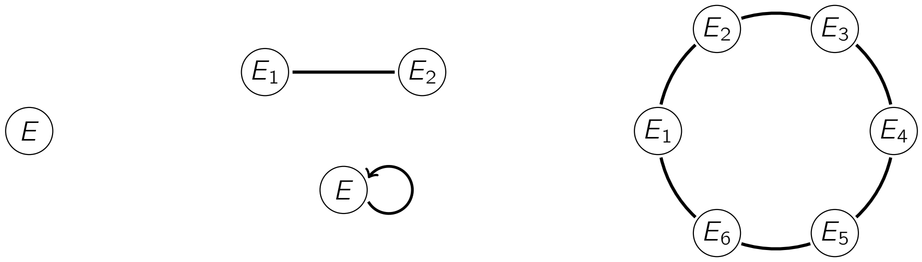


Figure 5.5 – Possible crater structures in an ℓ-volcano.

Number of connected components. Instead of studying the isogeny graph of a fixed elliptic curve one may look at the isogeny graph of $\mathrm{Ell}_k(\mathcal{O},\pi)$, the set of elliptic curves over k with endomorphism ring $\mathcal{O}$ and Frobenius endomorphism π. This consists in the union of the isogeny graph of all elliptic curve over k with complex multiplication by $\mathcal{O}$ and same trace of Frobenius,

$$G_\ell(\mathcal{O},\pi) = \bigcup_{E\in\mathrm{Ell}_k(\mathcal{O},\pi)} G_\ell(E)$$

It is clear that this is the union of a certain number of volcanoes. Once again, their number depends on the splitting behavior of ℓ in the ℓ-maximal order $\mathcal{O}$. By the classification of horizontal isogenies above,

(a) if ℓ is inert in $\mathcal{O}$, then there are $h(\mathcal{O})$ distinct ℓ-isogeny volcanoes of with surface in $\mathrm{Ell}_k(\mathcal{O})$;

(b) if ℓ ramifies in $\mathcal{O}$, i.e., $\ell\mathcal{O} = \mathfrak{l}^2$, then there are $h(\mathcal{O})$ or $h(\mathcal{O})/2$ distinct ℓ-isogeny volcanoes with surface in $\mathrm{Ell}_k(\mathcal{O})$ depending on whether $\mathfrak{l}$ is principal or not;

(c) if ℓ splits in $\mathcal{O}$ as the product of $\mathfrak{l}$ and $\bar{\mathfrak{l}}$, then there are $h(\mathcal{O})/n$ distinct ℓ-isogeny volcanoes with surface in $\mathrm{Ell}_k(\mathcal{O})$ of size n, where n is the order of $\mathfrak{l}$ in $\mathcal{C}\ell(\mathcal{O})$.

If there is more than one connected component, we refer to the union of distinct volcanoes with the term *cordillera*.

Vertical isogenies. Let $\mathcal{O}$ be an imaginary quadratic order and $\mathcal{O}' = \mathbb{Z} + \ell\mathcal{O}$ be the index ℓ suborder. Sutherland [Sut4] proved that the vertical isogenies from $\mathrm{Ell}_k(\mathcal{O}')$ to $\mathrm{Ell}_k(\mathcal{O})$ correspond to the surjective map of Section 5.1.2

$$\mathcal{C}\ell(\mathcal{O}') \to \mathcal{C}\ell(\mathcal{O})$$

In general, there are 0 or 1 ascending ℓ-isogenies from E depending on whether $\ell \mid [\mathcal{O}_K : \mathcal{O}]$.

Floor of rationality. As we have already stated, $\mathbb{Z}[\pi] \subseteq \mathrm{End}(E)$ which means that the ℓ-isogeny volcano has depth $\nu_\ell([\mathcal{O}_K : \mathbb{Z}[\pi]])$. A set of Lemmas in [Fou] and [FM] conclude the description of the ℓ-isogeny volcano: it is a volcano truncated at the level of $\mathbb{Z}[\pi]$, see [Koh1, §4.2].

Lemma 5.26. *Let E, E_1 and E_2 be elliptic curves over k with endomorphism rings $\mathcal{O}, \mathcal{O}_1$ and $\mathcal{O}_2$ in K, respectively.*

(a) *If $\mathbb{Z}[\pi]$ is maximal at ℓ, then all ℓ-isogenies are horizontal.*

(b) *If $\ell \mid [\mathcal{O}_K : \mathbb{Z}[\pi]]$ but $\ell \nmid [\mathcal{O} : \mathbb{Z}[\pi]]$, i.e., E lies at the floor of rationality, then the only isogeny from E is ascending.*

(c) *If there is a descending isogeny $\phi : E_1 \to E_2$ and $\ell \mid [\mathcal{O} : \mathbb{Z}[\pi]]$, any isogeny ψ from E_2 different from $\hat{\phi}$ is descending.*

$\nu_\ell([\mathcal{O}_K : \mathcal{O}])$	$\nu_\ell([\mathcal{O} : \mathbb{Z}[\pi]])$	Position of E	$\left(\frac{\Delta}{\ell}\right)$	$\left(\frac{\Delta_\pi}{\ell}\right)$	Isogenies	$\mathcal{N}_\ell(E)$
$\ell \nmid [\mathcal{O}_K : \mathcal{O}]$	$\ell \nmid [\mathcal{O} : \mathbb{Z}\pi]$	Crater=Floor	-1	-1	none	0
			+1	+1	2 horizontal	2
			0	0	1 horizontal	1
	$\ell \mid [\mathcal{O} : \mathbb{Z}[\pi]]$	Crater	-	0	$1 + \left(\frac{\Delta}{\ell}\right)$ horizontal $\ell - \left(\frac{\Delta}{\ell}\right)$ descending	$\ell + 1$
$\ell \mid [\mathcal{O}_K : \mathcal{O}]$	$\ell \mid [\mathcal{O} : \mathbb{Z}[\pi]]$	Side	0	0	1 ascending ℓ descending	$\ell + 1$
	$\ell \nmid [\mathcal{O} : \mathbb{Z}[\pi]]$	Floor	0	0	1 ascending	1

Table 5.2 – Number and type of isogenies from E depending on its endomorphism ring $\mathrm{End}(E) \simeq \mathcal{O}$

All these information can be resumed in the following table where $\mathcal{N}_\ell(E)$ denotes the number of linear factors of the modular polynomial $\Phi_\ell(X, j(E))$.

In other words, the crater is a regular graph of degree at most 2 and it is composed of all possible horizontal isogenies; its size is given by the behavior of ℓ in the ℓ-maximal order $\mathcal{O}$. Moving along the side of the volcano, i.e., descending, we encounter elliptic curves with smaller and smaller endomorphism ring; they all have one neighbor in the higher level and ℓ descending isogenies corresponding to the surjective morphism $\mathcal{Cl}(\mathbb{Z} + \ell\mathcal{O}) \to \mathcal{Cl}(\mathcal{O})$; thus they all have degree $\ell + 1$. We reach the floor when ℓ does not divide the index of $\mathbb{Z}[\pi]$ in the last order. Here we have nodes of degree 1 with only ascending isogenies.

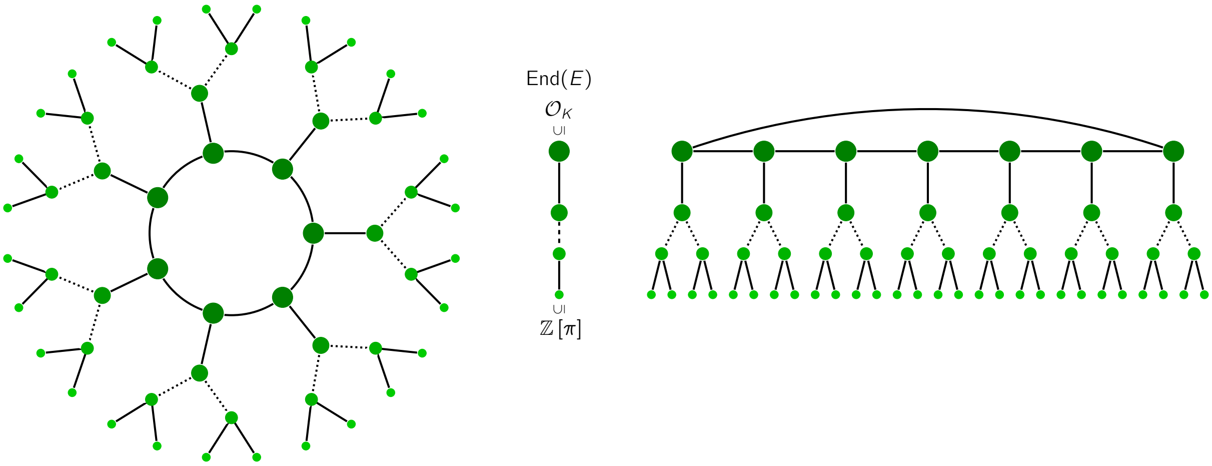


Figure 5.6 – Aerial and side view of a 2-isogeny volcano.

Special vertices. All we have said up to her still holds in case our volcano contains elliptic curves with extra automorphisms $j = 0, 1728$ except for some minor modification. First of all we note that these special curve can only appear on the crater. Further,

(Case $j = 0$) If level 1 is non empty, then it contains

$$\frac{1}{3}\left(\ell - \left(\frac{-3}{\ell}\right)\right)$$

curves each of which has 3 incoming isogenies from $j = 0$ but only one going back up;

(Case $j = 1728$) If level 1 is non empty, then it contains

$$\frac{1}{2}\left(\ell - \left(\frac{-1}{\ell}\right)\right)$$

curves each of which has 2 incoming isogenies from $j = 1728$ but only one going back up.

5.2.2 Supersingular isogeny graphs

We focus now on supersingular elliptic curves. Before describing their isogeny graphs we describe their endomorphism rings, which are isomorphic to orders in a quaternion algebra.

Quaternion algebras

A quaternion algebra $\mathfrak{A}$ over $\mathbb{Q}$ is a central simple $\mathbb{Q}$-algebra over $\mathbb{Q}$ of dimension 4. This means that $\mathfrak{A}$ is a $\mathbb{Q}$-algebra with no non-trivial 2-sided ideals and center $\mathbb{Q}$. The isomorphism $\mathbb{Q} \to Z(\mathfrak{A})$ makes $\mathfrak{A}$ into a $\mathbb{Q}$-vector space of dimension 4. Wedderburn structure theorem implies that a quaternion algebra $\mathfrak{A}$ over $\mathbb{Q}$ is either a division algebra over $\mathbb{Q}$, if all its elements have inverses, or it is isomorphic to $\mathbb{M}_2(\mathbb{Q})$, the matrix algebra over $\mathbb{Q}$. As a consequence, for every non-central element $\alpha \in \mathfrak{A}$, the ring $K = \mathbb{Q}(\alpha)$ is a quadratic extension of $\mathbb{Q}$. Any quaternion algebra over $\mathbb{Q}$ is therefore isomorphic to $K + K\beta$ where K is a quadratic extension of $\mathbb{Q}$ and β is a non-central element of $\mathfrak{A}$ such that $\beta^2 = b \in \mathbb{Q}^\times$ and $\beta\alpha = \sigma(\alpha)\beta$ for every $\alpha \in K$, where σ is the non-trivial automorphism of K. Equivalently, $\mathfrak{A} \simeq \mathbb{Q}\langle\alpha, \beta\rangle$ for α, β non-central elements of $\mathfrak{A}$ such that $\alpha^2, \beta^2 \in \mathbb{Q}^\times$ and $\alpha\beta = -\beta\alpha$.

To each element in $\mathfrak{A}$ we can associate its conjugate $\overline{\alpha} = \sigma(\alpha)$ in $\mathbb{Q}(\alpha)$; this defines an involution on $\mathfrak{A}$. We define the reduced norm and the reduced trace by

$$\begin{aligned} \mathrm{N} : \mathfrak{A} &\longrightarrow \mathbb{Q} \\ \alpha &\longmapsto \alpha\overline{\alpha} \end{aligned} \qquad\qquad \begin{aligned} \mathrm{Tr} : \mathfrak{A} &\longrightarrow \mathbb{Q} \\ \alpha &\longmapsto \alpha + \overline{\alpha} \end{aligned}$$

Further, for elements $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathfrak{A}$, we define their discriminant as

$$\Delta(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \det\left(\mathrm{Tr}(\alpha_i\alpha_j)\right)_{1\leq i,j\leq 4}$$

A quaternion algebra $\mathfrak{A}$ is said to ramify (or to be ramified) at a prime p if $\mathfrak{A}_p = \mathfrak{A} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is a division algebra and it is defined to split at p if $\mathfrak{A}_p \simeq \mathbb{M}_2(\mathbb{Q}_p)$. We say that $\mathfrak{A}$ is ramified (or splits) at ∞ if $\mathfrak{A}_p\infty = \mathfrak{A} \otimes_{\mathbb{Q}} \mathbb{R}$ is a division algebra (respectively, $\mathfrak{A}_\infty \simeq \mathbb{M}_2(\mathbb{R})$). We define $\mathfrak{A}_{p,\infty}$ as the quaternion algebra ramified only at p and infinity.

Lattices, orders and ideals. Let $\mathfrak{A}$ be a quaternion algebra over $\mathbb{Q}$. As we did in dimension 2 we define a lattice in $\mathfrak{A}$ to be a finitely generated $\mathbb{Z}$-module which contains a basis for $\mathfrak{A}$ over $\mathbb{Q}$. An order of a $\mathfrak{A}$ will be a lattice which is also a subring containing 1. We will denote $\mathfrak{O}$ an order in a quaternion algebra. An order consists of elements with integral norm and trace. However, unlikely the quadratic case, the set of all such elements might not be a ring. In fact, a quaternion algebra has more than one maximal order.

Given an order $\mathfrak{O}$ of $\mathfrak{A}$, we define a fractional left (right) ideal I to be a lattice in $\mathfrak{A}$ such that $\alpha I \subseteq I$ ($I\alpha \subseteq I$) for all $\alpha \in \mathfrak{O}$. I is integral if it is contained in $\mathfrak{O}$. The reduced norm $\mathrm{N}(I)$ of an ideal I is the fractional ideal of $\mathbb{Z}$ generated by the reduced norms of its elements. Finally, the inverse ideal is $I^{-1} = \{\alpha \in \mathfrak{A} \mid I\alpha I \subseteq I\}$. For an ideal I, we define its left and right orders as

$$\mathfrak{O}_L(I) = \mathrm{ord}_L(I) = \{\alpha \in \mathfrak{A} \mid \alpha I \subseteq I\} \qquad (\mathfrak{O}_R(I) = \mathrm{ord}_R(I) = \{\alpha \in \mathfrak{A} \mid I\alpha \subseteq I\})$$

An important invariant of quaternion orders is represented by their discriminant. It is defined as follows: let I be the ideal generated by all $\Delta(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ for $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathfrak{O}$; one can prove that it is the square of an ideal in $\mathbb{Z}$. We define $\Delta(\mathfrak{O})$ as this $\mathbb{Z}$-ideal (the square root of I).

Lemma 5.27 ([Vig]). *Let $\mathfrak{O}$ and $\mathfrak{O}'$ be two orders of $\mathfrak{A}$. If $\mathfrak{O}' \subseteq \mathfrak{O}$, then $\Delta(\mathfrak{O}')$ divides $\Delta(\mathfrak{O})$ and $\Delta(\mathfrak{O}') = \Delta(\mathfrak{O})$ if and only if $\mathfrak{O}' = \mathfrak{O}$. An order is maximal if and only if its discriminant is the product of the finite primes of $\mathbb{Q}$ ramifying in $\mathfrak{A}$.*

Ideal classes. Two ideals I, J are said to be left (right) equivalent if there exists $\alpha \in \mathfrak{A}^\times$ such that $I = \alpha J$ ($I = J\alpha$). Let $\mathfrak{O}$ be an order in $\mathfrak{A}$, the left ideal classes of $\mathfrak{O}$ are the right-equivalence classes of ideals with left order equal to $\mathfrak{O}$. Once again the right ideal classes of $\mathfrak{O}$ are left-equivalence classes of ideals with right order $\mathfrak{O}$.

Lemma 5.28. *The inverse map $I \mapsto I^{-1}$ establishes a bijection between right and left ideal classes of an order $\mathfrak{O}$.*

Therefore, there is a well defined notion of class number $h(\mathfrak{O})$ as the number of right or left equivalence classes of $\mathfrak{O}$.

We say that two orders $\mathfrak{O}$ and $\mathfrak{O}'$ are of the same type if there exists an element $\alpha \in \mathfrak{A}$ such that $\alpha\mathfrak{O}\alpha^{-1} = \mathfrak{O}'$, i.e., they are conjugate by α. Note that this is equivalent to the existence of a principal ideal I such that $\mathfrak{O}_l(I) = \mathfrak{O}$ and $\mathfrak{O}_R(I) = \mathfrak{O}'$. Being of the same type is an equivalence relation and we define the type number $t(\mathfrak{A})$ of $\mathfrak{A}$ as the number of conjugacy classes of maximal orders.

Lemma 5.29. *Orders of the same type have the same number of left (or right) ideal classes.*

Corollary 5.30. *All maximal orders have the same class number. We define it as the class number of $\mathfrak{A}$.*

Proof. See [Vig, Th. III.5.4] or [Rei, §26]. □

An equivalence of categories

We will now establish the equivalence of categories between supersingular elliptic curves over finite fields and quaternion orders; as in the ordinary case this will enables us to exploit properties of one to infer properties of the other though the non-commutativity of the quaternion world will prevent us from obtaining the same nice structure of ordinary graphs.

Deuring [Deu] proved the existence of a bijection between the set of supersingular j-invariants in characteristic p and the class number of a $\mathfrak{O}$ for any of the types of maximal orders $\mathfrak{O}$ in $\mathfrak{A}_{p,\infty}$, the unique quaternion algebra ramified only at p and infinity. The already cited Theorem of Waterhouse [Wat1, Th. 4.5] exploits this correspondence in terms of kernel ideals.

Theorem 5.31. *Given a type of maximal order, there exist one or two supersingular j-invariants such that the corresponding endomorphism ring is of the given type depending on whether the prime ideal $\mathfrak{P}$ over $\mathfrak{p}$ is principal or not.*

In his thesis, Kohel [Koh1] gives a functorial description of this correspondence. We define $\mathcal{S}_k$ as the category of supersingular elliptic curves over a field k of $q = p^r$ elements. The objects are pairs (E, π) where E is a supersingular elliptic curve over k and π is the Frobenius endomorphism. A morphism is a map $\psi : (E_1, \pi_1) \to (E_2, \pi_2)$ such that $\psi \circ \pi_1 = \pi_2 \circ \psi$.

Now let $\mathfrak{O}$ be a maximal order in $\mathfrak{A}_{p,\infty}$ containing an element of reduced norm q. The second category is $\mathcal{M}_{\mathfrak{O},q}$ consisting of projective right modules of rank 1 over $\mathfrak{O}$. The objects are pairs (I, ϕ) where I is a projective $\mathfrak{O}$-module of rank 1 and ϕ is an endomorphism of I of norm q. A morphism is $\psi : (I_1, \phi_1) \to (I_2, \phi_2)$ such that $\psi : I_1 \to I_2$ is a homomorphism such that $\psi \circ \phi_1 = \phi_2 \circ \psi$. We can think of $\mathfrak{O}$ as $\mathfrak{O}_R(I)$ for some ideal I; then projective orders of rank 1 are left ideals of $\mathfrak{O}$, see [Voi, §20.3].

Then there is an equivalence of categories established by the functor $\mathbf{I}$ from $\mathcal{S}_k$ to $\mathcal{M}_{\mathfrak{O},q}$ such that

$$(E, \pi) \rightsquigarrow (\mathbf{I}(E), \mathbf{I}(\pi)) = (\mathrm{End}(E), \mu_\pi)$$

where μ_π is the endomorphism of $\mathrm{End}(E)$ given by composition with π.

Supersingular j-invariants

In this section we are going to make explicit use of this correspondence to study isogeny graphs of supersingular elliptic curves. Unfortunately, the lack of a commutative action of the class group on the set of j-invariants prevents one to obtain the rigid structure of the volcano. Nevertheless, supersingular curves still have some interesting properties.

Field of definition. One of the characterization of supersingular elliptic curves is that they have no torsion p-points in characteristic p, see Theorem 1.21. Then the endomorphism $[p] = \hat{\pi}\pi$ has trivial kernel an therefore $\hat{\pi}$ has trivial kernel and it is therefore purely inseparable. Now we can decompose it as $\hat{\pi}_{\mathrm{sep}}\pi$ by Proposition 1.14, where we factor through the separable isogeny $\hat{\pi}$. We get $[p] = \hat{\pi}_{\mathrm{sep}}\pi^2$ and, since $\hat{\pi}_{\mathrm{sep}}$ is an isomorphism, then π^2 is separable of degree p^2. This means that π^2 fixes the j-invariant of E which is then defined over $\mathbb{F}_{p^2}$; therefore, every supersingular elliptic curve is isomorphic to one defined over $\mathbb{F}_{p^2}$.

By a similar argument, one can prove that the subset of supersingular elliptic curves defined over $\mathbb{F}_p$ consists of those whose endomorphism ring contains an element with minimal polynomial $x^2 + p$. If $j \neq 0, 1728$ this element can only be the Frobenius endomorphism (up to sign). Ibukiyama [Ibu] gave an explicit description (on the quaternion side) of maximal orders containing such an element.

Although j-invariants are always defined over a quadratic extension of $\mathbb{F}_p$, in general the isogenies will still be defined over $\overline{\mathbb{F}}_p$. Concerning the field of definition of endomorphisms of a supersingular elliptic curve, we can once again resort to [Wat1, Th. 4.5].

We conclude this brief overview by looking at special curves $j = 0, 1728$. These are supersingular in characteristic p if and only if p does not split in $\mathbb{Q}(\sqrt{-3})$, respectively $\mathbb{Q}(i)$, [Lan2, Th. 13.4.12] and [Sil1, Eg V.4.4-5]; equivalently, if and only if $p \not\equiv 1$ modulo 3, respectively 4.

Number of curves. The number of supersingular j-invariants over $\mathbb{F}_{p^2}$ is given by [Sil1, Th. V.4.1]

$$S_{p^2} = \left\lfloor \frac{p-1}{12} \right\rfloor + \begin{cases} 0 & \text{If } p \equiv 1 \mod 12 \\ 1 & \text{If } p \equiv 5, 7 \mod 12 \\ 2 & \text{If } p \equiv 11 \mod 12 \end{cases}$$

The subset of j-invariants defined over $\mathbb{F}_p$ has cardinality

$$S_p = \begin{cases} h(-4p)/2 & \text{If } p \equiv 1 \mod 4 \\ 2h(-p) & \text{If } p \equiv 3 \mod 8 \\ h(-p) & \text{If } p \equiv 7 \mod 8 \end{cases}$$

where $h(\Delta)$ is the class number of an imaginary quadratic order of discriminant Δ, see [Cox, Th. 14.18].

Deuring correspondence. We will now make explicit the connection between ideals and orders in a quaternion algebra and supersingular j-invariants.

Theorem 5.32 ([Deu])**.** *Let E be a supersingular elliptic curve over $\mathbb{F}_{p^2}$. Then* $\mathrm{End}^0(E) \simeq \mathfrak{A}_{p,\infty}$*. Further,*

(a) *The number of isomorphism classes of supersingular elliptic curves over $\mathbb{F}_{p^2}$ equals the class number $h(p)$ of $\mathfrak{A}_{p,\infty}$.*

(b) *There is a bijection*

$$\begin{tikzcd} \left\{\begin{array}{c}\textit{Isomorphism classes of}\\ \textit{Supersingular elliptic}\\ \textit{curves over } \mathbb{F}_{p^2}\end{array}\right\}/\mathcal{G}al(\overline{\mathbb{F}}_p/\mathbb{F}_p) \arrow[r, leftrightarrow] & \left\{\begin{array}{c}\textit{Maximal orders}\\ \textit{of } \mathfrak{A}_{p,\infty}\end{array}\right\}/\textit{Type} \\ E \arrow[r, mapsto] & \mathrm{End}(E) \end{tikzcd}$$

which means that the number of $\mathcal{G}al(\overline{\mathbb{F}}_p/\mathbb{F}_p)$-conjugacy classes of j-invariants of supersingular elliptic curves over $\mathbb{F}_{p^2}$ is the type number $t(p)$ of $\mathfrak{A}_{p,\infty}$.

(c) *Let $\mathfrak{O}$ be a maximal order of $\mathfrak{A}_{p,\infty}$ and $\{I_i\}$ be a set of left ideal class representatives for $\mathfrak{O}$. There is a bijection between the set of supersingular j invariants j_i over $\mathbb{F}_{p^2}$ and the set of maximal orders $\{\mathfrak{O}_R(I_i)\}$ such that* $\mathrm{End}(j_i) \simeq \mathfrak{O}_R(I_i)$*.*

Remark. If E is defined over $\mathbb{F}_{p^2} \setminus \mathbb{F}_p$ there exists two ideal classes $[I]$ and $[J]$ such that $\mathrm{End}(E) \simeq \mathfrak{O}_R(I) \simeq \mathfrak{O}_R(J)$ and this corresponds to the existence of the Galois conjugate E^σ such that $\mathrm{End}(E) \simeq \mathrm{End}(E^\sigma)$ (here σ is the non-trivial automorphism of $\mathbb{F}_{p^2}$); therefore, without fixing a set of representatives for left ideal classes, the correspondence of Theorem 5.32.c is only defined up to conjugation, see [Bel, Th. 2.2.4].

We will now look at the relation between ideals and isogenies. Let E be a supersingular elliptic curve over $\mathbb{F}_{p^2}$ with endomorphism $\mathrm{End}(E)$; we fix an isomorphism $\iota : \mathfrak{O} \to \mathrm{End}(E)$ for a maximal order $\mathfrak{O}$ in $\mathfrak{A}_{p,\infty}$. For a left $\mathfrak{O}$-ideal I one can define

$$E[I] = \left\{ P \in E(\overline{\mathbb{F}}_p) \,\middle|\, \alpha(P) = O \text{ for all } \alpha \in \iota(I) \right\} = \bigcap_{\alpha \in \iota(I)} E[\alpha]$$

This induces is a unique separable isogeny $\phi_I : E \to E/E[I]$ of degree $\mathrm{N}(I)$, up to isomorphism of the codomain.

We say that an isogeny $\phi : E_1 \to E_2$ is normalized if $\phi^*\omega_2 = \omega_1$ where ω_i is the invariant differential of E_i. A normalized isogeny $\phi_I : E \to E/E[I]$ induces an isomorphism

$$\begin{aligned} i_I : \mathfrak{A}_{p,\infty} &\longrightarrow \mathrm{End}^0(E) = \mathrm{End}(E) \otimes \mathbb{Q} \\ \alpha &\longmapsto \phi_I i(\alpha) \hat{\phi}_I \otimes \deg(\phi_I)^{-1} \end{aligned}$$

Lemma 5.33. *The endomorphism ring of $E/E[I]$ is isomorphic to the maximal order $\mathfrak{O}_R(I)$.*

Cervino [Cer2; Cer1] proposes an explicit algorithm that given, a prime number p, returns a list of all the supersingular j-invariants together with a $\mathbb{Z}$-basis for the maximal order $\mathfrak{O}$ of $\mathfrak{A}_{p,\infty}$ isomorphic to $\operatorname{End}(E)$. This is based on the correspondence between quaternionic orders and ternary quadratic forms, see [Voi, Ch. 22].

Supersingular isogeny graphs

Despite the rich theory of quaternion algebras, the fact that the quaternion ideal classes do not form a group prevents us to obtain the nice structure of the volcano. In fact, they usually do not have a regular shape and looks quite complicated Nevertheless, supersingular isogeny graphs still have good probabilistic properties that make them good candidates for applications. We provide here a short summary of graph theory. Most of the following only applies to finite graphs and we state it here to motivate the interest in studying supersingular isogeny graphs.

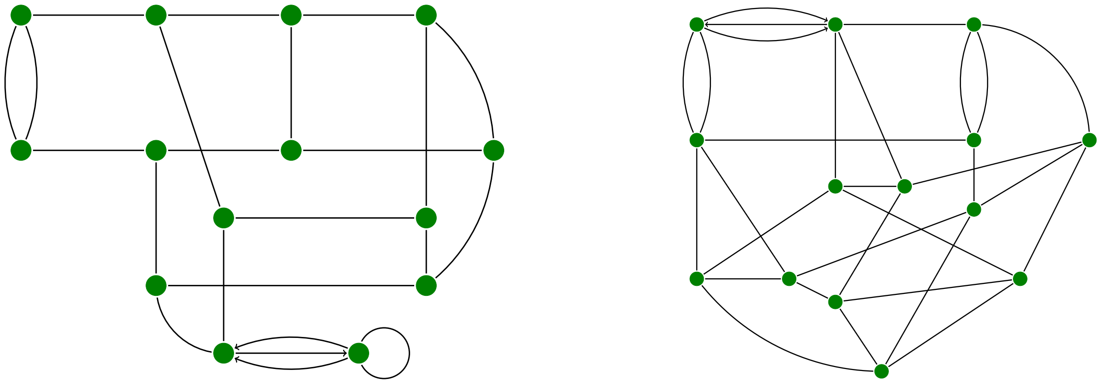


Figure 5.7 – The supersingular 2 and 3-isogeny graphs over $\overline{\mathbb{F}}_{163}$

Graph theory. Given a graph G we say that it is connected if there is a path connecting any two vertices. The distance between two vertices is the minimal length of a path between them, i.e., the length of the shortest path. If G is connected, its diameter, noted $\operatorname{diam}(G)$, is the longest distance between any two of its vertices. Finally, the adjacency matrix of G with vertex set $\mathcal{V} = \{v_1, \dots, v_n\}$ is the $n \times n$ matrix $A(G) = (a_{i,j})_{i,j}$ such that $a_{i,j} = 1$ if there is an edge between v_i and v_j, and 0 otherwise. For undirected graphs, the adjacency matrix is symmetric; thus it has n real eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$. A k-regular graph G has the property that $k \geq \lambda_1 \geq \dots \geq \lambda_n \geq -k$. We say that it is an expander graph if there exists $\epsilon > 0$ such that $(1-\epsilon)k \geq \lambda_2 \geq \dots \lambda_n \geq -(1-\epsilon)k$. If we can further reduce the bound to $2\sqrt{k-1}$, we say that the graphs has the Ramanujan property.

Expander graphs have small diameters (bounded by $O(\log(n))$) and random paths of generate any vertex with probability that becomes close to uniform as their length approaches the diameter.

Theorem 5.34. *Let p be a prime. If $\ell \neq p$ is a prime number, the ℓ-isogeny graph of supersingular curves over $\mathbb{F}_{p^2}$ is connected, $(\ell+1)$-regular, and has the Ramanujan property.*

Proof. See [Piz3, Th. 1]. □

Special j-invariants. Looking at Figure 5.7 we notice some curious behavior. We have already explained what causes a different number of incoming and outgoing isogenies; we may therefore ask ourselves what determines loops or multiple edges between two nodes. Cycles corresponds to endomorphisms of degree ℓ and they happen at j invariants that satisfy the modular polynomial $\Phi_\ell(X, X)$; hence, they are bounded in number by $\deg(\Phi_\ell(X, X))$ On the other hand a j-invariants j_1 admits two isogenies to some j_2 if it is a root of the resultant in Y of Φ_ℓ and $\partial_Y \Phi_\ell$ [Arp+, Lemma 2.4]:

$$\operatorname{Res}_Y\left(\Phi_\ell, \frac{\partial}{\partial Y}\Phi_\ell\right)$$

Arpin [Arp1] resumes conditions on the prime p which produce supersingular graphs with no loops or multiple edges for 2 and 3 isogeny graphs.

The isogeny graph over $\mathbb{F}_p$. Supersingular isogeny graphs do not have very regular shapes, however one can find more rigid structures if focusing on subgraphs; in turns, the study of these smaller graphs may provide information about the whole picture. Delfs and Galbraith [DG] studied the subgraph consisting of j invariants and isogenies all defined over $\mathbb{F}_p$; Adj, Ahmadi and Menezes [AAM], instead work with the whole set of j invariants but only consider isogenies defined over $\mathbb{F}_{p^2}$. In [Arp+], the authors consider the set of $\mathbb{F}_p$-rational j invariants but allow all isogenies between them and they call the resulting graph the *spine*.

Remark. We point out here that the restriction to a category of k-isogenies is equivalent to imposing an orientation by $\mathbb{Z}[\pi]$ in the sense of Section 5.3.1.

The main reason to restrict set of edges to those isogenies defined over $\mathbb{F}_p$ comes from the following result which is a direct consequence of Theorems 1.19 and 1.17.

Theorem 5.35 ([DG, Th. 2.1])**.** *Let $p > 5$ be a prime and $q = p^e$. There exist an isogeny class of supersingular elliptic curves defined over $\mathbb{F}_q$ with trace of Frobenius t if and only if one of the following holds:*

(a) *e is even and either*

- **(i)** $t = \pm 2\sqrt{q}$
- **(ii)** $p \not\equiv 1 \bmod 3$ *and* $t = \pm\sqrt{q}$;
- **(iii)** $p \not\equiv 1 \bmod 4$ *and* $t = 0$.

(b) *e is odd and $t = 0$.*

In the case a(i), $\mathrm{End}^0_{\mathbb{F}_q}(E)$ *is the quaternion algebra $\mathfrak{A}_{p,\infty}$ over $\mathbb{Q}$, π is a rational integer and* $\mathrm{End}_{\mathbb{F}_q}(E)$ *is a maximal order in $\mathfrak{A}_{p,\infty}$. In the other three cases* $\mathrm{End}^0_{\mathbb{F}_q}(E) \simeq K = \mathbb{Q}(\pi)$ *is an imaginary quadratic field and* $\mathrm{End}_{\mathbb{F}_q}(E)$ *is an order in it with conductor prime to p.*

In particular, the restriction of the endomorphism ring to $\mathbb{F}_p$ is an order in a quadratic imaginary field. Since $\pi^2 + p = 0$ we must have $K = \mathbb{Q}(\sqrt{-p})$. Since

$$\mathbb{Z}[\pi] = \mathbb{Z}[\sqrt{-p}] \subseteq \mathrm{End}_{\mathbb{F}_q}(E) \subseteq \mathcal{O}_K$$

and $\mathbb{Z}[\sqrt{-p}]$ has conductor either 1 (if $p \equiv 1 \bmod 4$)or 2 (if $p \equiv 3 \bmod 4$) in $\mathcal{O}_K$, then $\mathrm{End}_{\mathbb{F}_q}(E)$ can only be

- $\mathbb{Z}[\sqrt{-p}] = \mathcal{O}_K$ if $p \equiv 1 \bmod 4$.
- $\mathbb{Z}[\sqrt{-p}] = ZZ + 2\mathcal{O}_K$ or $\mathcal{O}_K = \mathbb{Z}[(1+\sqrt{-p})/2]$ if $p \equiv 3 \bmod 4$.

The resulting graph $G_\ell(\mathbb{F}_p)$ is a cordillera (of volcanoes) of height 1 or 2 depending on p modulo 4 [DG].

5.2.3 Supersingular isogeny graphs with level structure

We recall that an isogeny graph $G = G_S(E)$, of an elliptic curve E/k is a graph whose vertices are elliptic curves $\bar{k}$-isogenous to E, and whose directed edges are isogenies of prime degree $\ell \in S$.

If $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup of level N, then we define $G_S(E,\Gamma)$ as the graph whose vertices are pairs $(E, \Gamma(P,Q))$ where $\Gamma(P,Q)$ is the orbit of an ordered basis (P,Q) of $E[N]$ such that the Weil pairing $e_n(P,Q) = \zeta_N$ is a fixed root of unity in $\bar{k}$, and whose edges are isogenies of prime degree $\ell \in S$. The orbit is defined with respect to the left action of Γ, given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(P,Q) = (aP + bQ, cP + dQ).$$

We identify $G_S(E)$ with $G_S(E, \mathrm{SL}_2(\mathbb{Z}))$ and for each inclusion $\Gamma_1 \subset \Gamma_2$ we obtain a morphism of graphs

$$G_S(E,\Gamma_1) \to G_S(E,\Gamma_2).$$

We can identify the vertices of $G_S(E,\Gamma)$ with points on the modular curve $X(\Gamma)$ and edges with points on the modular curve $X(\Gamma\cap\Gamma_0(\ell))$ for ℓ in S different from the level of Γ, and otherwise $X(\Gamma\cap\Gamma_0(\ell^i))$, where i is the smallest exponent such that Γ is not contained in $\Gamma_0(\ell^i)$. When $S=\ell$, we will write simply $G_\ell(E)$ and $G_\ell(E,\Gamma)$. We stress the fact that adding level structure gives a covering graph, which maps surjectively down to the $\Gamma(1)$-structure one.

Supersingular fields of definition. We will now study the field of definition of supersingular invariants on modular curves with particular focus on Weber modular curves, see Section 3.3.5.

Theorem 5.36. *For any positive integer N, the supersingular invariants on the modular curve $X_0(N)$ are defined over $\mathbb{F}_{p^2}$, and if $p\equiv\pm1 \bmod N$, then the supersingular invariants also split over $\mathbb{F}_{p^2}$ on $X_1(N)$.*

Proof. For any elliptic curve E in the isogeny class of a curve over $\mathbb{F}_p$, the full endomorphism ring $\mathfrak{O}$ is defined over $\mathbb{F}_{p^2}$. Since the action of $\mathfrak{O}/N\mathfrak{O}\cong\mathbb{M}_2(\mathbb{Z}/N\mathbb{Z})$ on the $E[N]$ is defined over $\mathbb{F}_{p^2}$, it follows that the Galois action on $E[N]$, which commutes with $\mathfrak{O}/N\mathfrak{O}$, acts through the center $(\mathbb{Z}/N\mathbb{Z})^*$ of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, and more precisely, Frobenius acts as $-p$ on $E[N]$. Consequently, the lines are Galois stable and every cyclic N-isogeny is defined over $\mathbb{F}_{p^2}$. In view of the action of Frobenius, if $p\equiv\pm1 \bmod N$, the Galois action on the N-torsion of E or its twist is trivial, so the supersingular moduli are defined in $\mathbb{F}_{p^2}$. □

Remark. Equivalently, for $X_0(N)$ we can state that every supersingular j-invariant j_0 splits completely under the map $X_0(N)\to X(1)$, or that the polynomial $\Phi_N(x,j_0)$ splits completely, where $\Phi_N(x,y)$ is the classical modular polynomial. For $X_1(N)$, the splitting of the supersingular points is recognized by the factorization of the N-division polynomial ψ_N.

As a consequence, the split Cartan modular curve $X_s(N)$, defined by the congruence subgroup

$$\Gamma_s(N)=\left\{\begin{pmatrix}a&b\\c&d\end{pmatrix}\ \middle|\ b\equiv c\equiv 0 \bmod N\right\},$$

parametrizing elliptic curves with a disjoint pair of cyclic N-isogenies, also splits the supersingular moduli.

Corollary 5.37. *For any positive integer N, then the supersingular invariants on the split Cartan modular curve $X_s(N)$ are defined over $\mathbb{F}_{p^2}$. In particular, if $p\equiv\pm1 \bmod N$, then the supersingular invariants on the modular curve $X(N)$ are defined over $\mathbb{F}_{p^2}$.*

Proof. The first statement follows from the splitting of N-isogenies over $\mathbb{F}_{p^2}$. In addition if $p\equiv\pm1 \bmod N$, the points of each kernel are fixed, hence a basis is defined over $\mathbb{F}_{p^2}$ (up to twist). □

Remark. For the levels N in $\{1,2,3,4,6\}$, the unit group $(\mathbb{Z}/N\mathbb{Z})^*/\{\pm1\}$ is trivial so the supersingular points split for all p. This corresponds to the geometric equalities $X_1(N)=X_0(N)$ and $X(N)=X_s(N)$.

The Weber moduli are functions on $X(48)$ which map through $\mathcal{W}_{24}$. To show the splitting of supersingular points on $\mathcal{W}_{24}$ it suffices to prove it for $\mathcal{W}_3$ and $\mathcal{W}_8$. However, $X(6)$ covers $\mathcal{W}_3$, so the supersingular moduli on $\mathcal{W}_3$ split over $\mathbb{F}_{p^2}$ by the previous theorem. To prove that they split on $\mathcal{W}_8$ it is necessary to consider the factorization

$$\begin{tikzcd}
X(16,8,16) \arrow[r] \arrow[d, "\left\langle\left(\begin{smallmatrix}13&8\\8&5\end{smallmatrix}\right)\right\rangle"'] & X(8) \arrow[d, "\left\langle\left(\begin{smallmatrix}5&0\\0&5\end{smallmatrix}\right)\right\rangle"] \\
\mathcal{W}_8 \arrow[r] & X_s(8),
\end{tikzcd}$$

where $X(16,8,16)$ is the quotient of $X(16)$ by the diagonal matrix group $\langle\pm9I_2\rangle\subset\mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z})/\{\pm1\}$.

The supersingular points split in $X_s(8)$ by the previous theorem. On the other hand, for the classes $p \bmod 8$ in the coset $\{\pm5\}\subset(\mathbb{Z}/8\mathbb{Z})^*/\{\pm1\}$ form an obstruction to lifting supersingular points to $X(8)$ over $\mathbb{F}_{p^2}$. Clearly, since $\langle 9I_2\rangle\subset\Gamma(16,8,16)/\Gamma(16)$, for the primes p such that $p \bmod 16$ lie in the kernel

$$\langle -1,9\rangle=\{\pm1,\pm9\}\subset(\mathbb{Z}/16)^*/\{\pm1\}\longrightarrow(\mathbb{Z}/8)^*/\{\pm1\},$$

the supersingular invariants in $X(16,8,16)$ split over $\mathbb{F}_{p^2}$. It remains to show that the obstruction vanishes also on the coset $\{\pm3,\pm5\}$. However, this follows since the subgroup of $\Gamma_8/\Gamma(16)$ surjects on the diagonal subgroup of $\Gamma_s(8)/\Gamma(8)$:

$$\left\langle\begin{pmatrix}13&8\\8&5\end{pmatrix}\right\rangle\longrightarrow\left\langle\begin{pmatrix}5&0\\0&5\end{pmatrix}\right\rangle$$

under $\mathrm{SL}_2(\mathbb{Z}/16\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/8\mathbb{Z})$, corresponding to the fact that $\mathcal{W}_8$ does not factor through $X(8)$. This establishes the following theorem.

Theorem 5.38. *The supersingular Weber invariants on* $\mathcal{W}_{24}$ *are defined over* $\mathbb{F}_{p^2}$.

Remark. A point (u_0, u_1, u_2) on $\mathcal{W}_{24}$ over the j-invariant j_0 consists of a triple of common roots of the polynomial $(x^{24} - 16)^3 - j_0 x^{24}$, and the set roots is precisely $\{\, \zeta_{24}^i u_j \mid 0 \leq i < 24, 0 \leq j < 3 \,\}$. The property that j_0 splits completely under $\mathcal{W}_{24} \to X(1)$ over $\mathbb{F}_{p^2}$ is equivalent to this polynomial splitting completely over $\mathbb{F}_{p^2}$.

Endomorphism rings. Let Γ be a congruence subgroup. We define the endomorphism ring of the pair $(E, \Gamma(P, Q))$ as the subring of $\mathrm{End}(E)$

$$\mathrm{End}(E, \Gamma(P, Q)) = \{\alpha \in \mathrm{End}(E) \mid \alpha(\Gamma(P, Q)) \subseteq \Gamma(P, Q)\}$$

In case $\Gamma = \Gamma_0(N)$, it is well known that the $\mathrm{End}(E, \Gamma(P, Q))$ is an Eichler order of level N; in fact, on the quaternion side, an Eichler order is an equivalent data to two maximal orders with a connecting ideal of norm N, see [KLPT] and [Arp2].

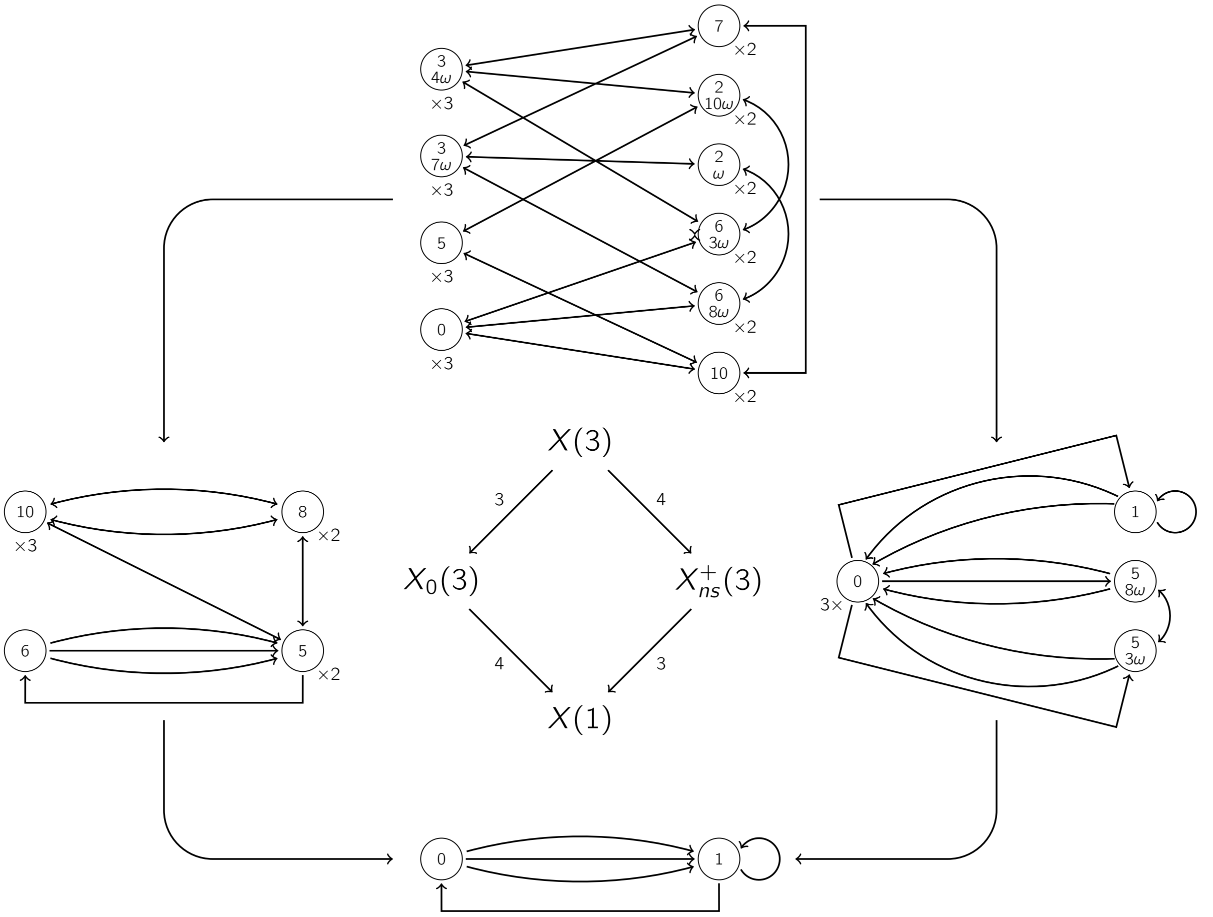


Figure 5.8 – Adding level 3 structure to the supersingular isogeny graph over $\mathbb{F}_{11^2} = \mathbb{F}_{11}[i]$ for i a root of $x^2 + 1$. We use the Hesse invariant on $X(3)$, the classical η quotient $(\eta_1/\eta_3)^{12}$ on $X_0(3)$ and the cube root of the j invariant on $X^+_{ns}(3)$.
The cover $X(3) \longrightarrow X_0(3)$ has the following description: $\{0\} \mapsto 6$, $\{5, 3+4\omega, 3+7\omega\} \mapsto 10$, $\{10, 6+3\omega, 6+8\omega\} \mapsto 5$ and $\{7, 2+\omega, 2+10\omega\} \mapsto 8$.
In $X(3) \to X^+_{ns}(3)$ we find $\{0, 5, 3+4\omega, 3+7\omega\} \mapsto 0$, $\{6+8\omega, 2+10\omega\} \mapsto 5+3\omega$, $\{6+3\omega, 2+\omega\} \mapsto 5+8\omega$ and $\{7, 10\omega\} \mapsto 1$.
Concerning $X_0(3), X^+_{ns}(3) \to X(1)$, the coverings have to be read column by column.

The change of the endomorphism ring results in a rigidification of the isogeny graph. In Figure 5.8 we observe how adding level structure progressively eliminates loops and multiple edges.

5.3 Orientations, isogeny chains, and ladders

We introduce now a category of supersingular elliptic curves oriented by an imaginary quadratic order $\mathcal{O}$, and derive properties of the associated oriented and non-oriented supersingular ℓ-isogeny graphs. In other words we enhance the category of supersingular curves with an extra piece of information that come in the form of an embedding $\mathcal{O} \hookrightarrow \mathrm{End}(E)$. This permits one to derive a faithful group action on a subset of oriented supersingular curves, equipped with a forgetful map to the set of non-oriented supersingular curves. Further, it allows to impose compatible actions of the class groups of the suborders of this quadratic order on the descending isogeny chains and therefore on the isogeny volcano of oriented curves.

5.3.1 Orientations

Suppose that E is a supersingular elliptic curve over a finite field k of characteristic p, and denote by $\mathrm{End}(E)$ the full endomorphism ring. We assume moreover that k contains $\mathbb{F}_{p^2}$ and E is in an isogeny class such that $\mathrm{End}_k(E) = \mathrm{End}(E)$. We denote by $\mathrm{End}^0(E)$ the $\mathbb{Q}$-algebra $\mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q} = \mathfrak{A}_{p,\infty}$. Let K be a quadratic imaginary field of discriminant Δ_K with maximal order $\mathcal{O}_K$. Then there exists an embedding $\iota : K \to \mathrm{End}^0(E)$ if and only if p is inert or ramified in $\mathcal{O}_K$, and there exists an order $\mathcal{O} \subseteq \mathcal{O}_K$ such that $\iota(\mathcal{O}) = \iota(K) \cap \mathrm{End}(E)$.

Definition. A *K-orientation* on a supersingular elliptic curve E/k is a homomorphism $\iota : K \hookrightarrow \mathrm{End}^0(E)$. An *$\mathcal{O}$-orientation* on E is a K-orientation such that the image of the restriction of ι to $\mathcal{O}$ is contained in $\mathrm{End}(E)$. We write $\mathrm{End}((E,\iota))$ for the order $\mathrm{End}(E) \cap \iota(K)$ in $\iota(K)$. An $\mathcal{O}$-orientation is *primitive* if ι induces an isomorphism of $\mathcal{O}$ with $\mathrm{End}((E,\iota))$.

Let $\phi : E \to F$ be an isogeny of degree ℓ. A K-orientation $\iota : K \hookrightarrow \mathrm{End}^0(E)$ determines a K-orientation $\phi_*(\iota) : K \hookrightarrow \mathrm{End}^0(F)$ on F, defined by

$$\phi_*(\iota)(\alpha) = \frac{1}{\ell}\,\phi \circ \iota(\alpha) \circ \hat{\phi}.$$

Conversely, given K-oriented elliptic curves (E, ι_E) and (F, ι_F) we say that an isogeny $\phi : E \to F$ is *K-oriented* if $\phi_*(\iota_E) = \iota_F$, i.e. if the orientation on F is induced by ϕ. The restriction to K-oriented isogenies determines a category of K-oriented elliptic curves, hence of K-oriented isomorphism classes, and a subcategory of $\mathcal{O}$-oriented elliptic curves.

If E admits a primitive $\mathcal{O}$-orientation by an order $\mathcal{O}$ in K, $\phi : E \to F$ is an isogeny then F admits an induced primitive $\mathcal{O}'$-orientation for an order $\mathcal{O}'$ satisfying

$$\mathbb{Z} + \ell\mathcal{O} \subseteq \mathcal{O}' \text{ and } \mathbb{Z} + \ell\mathcal{O}' \subseteq \mathcal{O}.$$

We say that an isogeny $\phi : E \to F$ is an $\mathcal{O}$-oriented isogeny if $\mathcal{O} = \mathcal{O}'$.

The introduction of an orientation permits one to recover the terminology and the approach used in the ordinary case, see 5.2.1. If ℓ is prime, as direct analogue of [Koh1, Prop. 4.2.23], one of the following holds:

- $\mathcal{O} = \mathcal{O}'$ and we say that ϕ is *horizontal*,
- $\mathcal{O} \subset \mathcal{O}'$ with index ℓ and we say that ϕ is *ascending*,
- $\mathcal{O}' \subset \mathcal{O}$ with index ℓ and we say that ϕ is *descending*.

Moreover if the discriminant of $\mathcal{O}$ is Δ, then there are exactly $\ell - \left(\frac{\Delta}{\ell}\right)$ descending isogenies. If $\mathcal{O}$ is maximal at ℓ, then there are $\left(\frac{\Delta}{\ell}\right) + 1$ horizontal isogenies, and if $\mathcal{O}$ is non-maximal at ℓ, then there is exactly one ascending ℓ-isogeny and no horizontal isogenies.

For an oriented class (E, ι) with endomorphism ring $\mathcal{O} = \mathrm{End}((E,\iota))$, we define (E,ι) to be at the *surface* (or depth 0) if $\mathcal{O}$ is ℓ-maximal, and to be at *depth n* if the valuation at ℓ of $[\mathcal{O}_K : \mathcal{O}]$ is n. In the next section we introduce ℓ-isogeny chains linking oriented curves at the surface to oriented curves at depth n.

The oriented graph $G_S(E,\iota)$ is the graph whose vertices are K-oriented isomorphism classes, with fixed base vertex (E,ι), and whose edges are K-oriented ℓ-isogenies for ℓ in S.

5.3.2 Isogeny chains and ladders

Let E_0/k be a fixed supersingular elliptic curve, equipped with an $\mathcal{O}$-orientation, and let $\ell \neq p$ be a prime.

Definition. We define an *ℓ-isogeny chain* of length n from E_0 to E to be a sequence of ℓ-isogenies:

$$E_0 \xrightarrow{\phi_0} E_1 \xrightarrow{\phi_1} E_2 \xrightarrow{\phi_2} \dots \xrightarrow{\phi_{n-1}} E_n = E.$$

We say that the ℓ-isogeny chain is *without backtracking* if $\ker(\phi_{i+1} \circ \phi_i) \neq E_i[\ell]$ for each $i = 0, \dots, n-1$, and say that the isogeny chain is *descending* (or *ascending*, or *horizontal*) if each ϕ_i is descending (or ascending, or horizontal, respectively).

Since the dual isogeny of ϕ_i, up to isomorphism, is the only isogeny ϕ_{i+1} satisfying $\ker(\phi_{i+1} \circ \phi_i) = E_i[\ell]$, an isogeny chain is without backtracking if and only if the composition of two consecutive isogenies is cyclic. Moreover, we can extend this characterization in terms of cyclicity to the entire ℓ-isogeny chain.

Lemma 5.39. *The composition of the isogenies in an ℓ-isogeny chain is cyclic if and only if the ℓ-isogeny chain is without backtracking.*

Remark. If an isogeny ϕ is descending, then the unique ascending isogeny from $\phi(E)$, up to isomorphism, is the dual isogeny $\hat{\phi}$, satisfying $\hat{\phi}\phi = [\ell]$. As an immediate consequence, a descending ℓ-isogeny chain is automatically without backtracking, and an ℓ-isogeny chain without backtracking is descending if and only if ϕ_0 is descending.

Remark. An ℓ-isogeny chain corresponds to a path in the underlying ℓ-isogeny graph $G_\ell(E)$. The concept of backtracking, however, is more subtle in the ℓ-isogeny graph with level structure Γ. In particular, an ℓ-isogeny chain with $\phi_{i+1} \circ \phi_i = E_i[\ell]$ to $E_{i+2} = E_i$ may induce a nontrivial automorphism of Γ-orbits in $E_i[N]$. One of the interests of introducing level structure on graphs is to avoid backtracking, loops and cycles of order 2 in an ℓ-isogeny graphs.

Suppose that (E_i, ϕ_i) is an ℓ-isogeny chain, with E_0 equipped with an $\mathcal{O}_K$-orientation $\iota_0 : \mathcal{O}_K \to \mathrm{End}(E_0)$. For each i, let $\iota_i : K \to \mathrm{End}^0(E_i)$ be the induced K-orientation on E_i; we note $\mathcal{O}_i = \mathrm{End}(E_i) \cap \iota_i(K)$ with $\mathcal{O}_0 = \mathcal{O}_K$ and $\Delta_i = \mathrm{discr}(\mathcal{O}_i)$ with $\Delta_0 = \Delta_K$. In particular, if (E_i, ϕ_i) is a descending ℓ-chain, then ι_i induces an isomorphism

$$\iota_i : \mathbb{Z} + \ell^i \mathcal{O}_K \longrightarrow \mathcal{O}_i.$$

Let q be a prime different from p and ℓ that splits in $\mathcal{O}_K$, let $\mathfrak{q}$ be a fixed prime over q. For each i we set $\mathfrak{q}_{(i)} = \iota_i(\mathfrak{q}) \cap \mathcal{O}_i$, and define

$$C_i = E_i[\mathfrak{q}_{(i)}] = \{P \in E_i[q] \mid \psi(P) = 0 \text{ for all } \psi \in \mathfrak{q}_{(i)}\}.$$

We define $F_i = E_i/C_i$, and let $\psi_i : E_i \to F_i$, an isogeny of degree q. By construction, it follows that $\phi_i(C_i) = C_{i+1}$ for all $i = 0, \dots, n-1$. In particular, if (E_i, ϕ_i) is a descending ℓ-ladder, then ι_i induces an isomorphism

$$\iota_i : \mathbb{Z} + \ell^i \mathcal{O}_K \longrightarrow \mathcal{O}_i.$$

The isogeny $\psi_0 : E_0 \to F_0 = E/C_0$ gives the following diagram of isogenies:

$$\begin{tikzcd}
E_0 \arrow[r, "\phi_0"] \arrow[d, "\psi_0"'] & E_1 \arrow[r, "\phi_1"] & E_2 \arrow[r, "\phi_2"] & \cdots \arrow[r, "\phi_{n-1}"] & E_n \\
F_0 & & & &
\end{tikzcd}$$

and for each $i = 0, \dots, n-1$ there exists a unique $\phi_i' : F_i \to F_{i+1}$ with kernel $\psi_i(\ker(\phi_i))$ such that the following diagram commutes:

$$\begin{tikzcd}
C_i \subseteq E_i \arrow[r, "\phi_i"] \arrow[d, "\psi_i"'] & E_{i+1} \supseteq C_{i+1} \arrow[d, "\psi_{i+1}"] \\
F_i \arrow[r, "\phi_i'"'] & F_{i+1}
\end{tikzcd}$$

The isogenies $\psi_i : E_i \to F_i$ induce orientations $\iota_i' : \mathcal{O}_i' \to \mathrm{End}(F_i)$. This construction motivates the following definition.

Definition. An *ℓ-ladder* of length n and degree q is a commutative diagram of ℓ-isogeny chains (E_i, ϕ_i) and (F_i, ϕ_i') of length n connected by q-isogenies $(\psi_i : E_i \to F_i)$:

$$\begin{tikzcd}
\overset{E_0}{\bullet} \arrow[r, "\phi_0"] \arrow[d, "\psi_0"'] & \overset{E_1}{\bullet} \arrow[r, "\phi_1"] \arrow[d, "\psi_1"'] & \overset{E_2}{\bullet} \arrow[r, "\phi_2"] \arrow[d, "\psi_2"'] & \cdots \arrow[r, "\phi_{n-1}"] & \overset{E_n}{\bullet} \arrow[d, "\psi_n"] \\
\underset{F_0}{\bullet} \arrow[r, "\phi_0'"'] & \underset{F_1}{\bullet} \arrow[r, "\phi_1'"'] & \underset{F_2}{\bullet} \arrow[r, "\phi_2'"'] & \cdots \arrow[r, "\phi_{n-1}'"'] & \underset{F_n}{\bullet}
\end{tikzcd}$$

We also refer to an ℓ-ladder of degree q as a *q-isogeny* of ℓ-isogeny chains, which we express as $\psi : (E_i, \phi_i) \to (F_i, \phi_i')$.

We say that an ℓ-ladder is ascending (or descending, or horizontal) if the ℓ-isogeny chain (E_i, ϕ_i) is ascending (or descending, or horizontal, respectively). We say that the ℓ-ladder is *level* if ψ_0 is a horizontal q-isogeny. If the ℓ-ladder is descending (or ascending), then we refer to the length of the ladder as its *depth* (or, respectively, as its *height*).

Lemma 5.40. *An ℓ-ladder $\psi : (E_i, \phi_i) \to (F_i, \phi_i')$ of oriented elliptic curves is level if and only if* $\mathrm{End}((E_i, \iota_i))$ *is isomorphic to* $\mathrm{End}((F_i, \iota_i'))$ *for all* $0 \le i \le n$. *In particular, if the ℓ-ladder is level, then (E_i, ϕ_i) is descending (or ascending, or horizontal) if and only if (F_i, ϕ_i') is descending (or ascending, or horizontal).*

Remark. In the sequel we will assume that E_0 is oriented by a maximal order $\mathcal{O}_K$. In Section 5.3.3 we investigate using the effective horizontal isogenies of E_0 to derive an effective class group action, and introduce a modular version of this action in Section 5.3.4. Walking down a descending isogeny chain, each elliptic curve will be oriented by an order of decreasing size and the final elliptic curve, which will be our final object of study, will have an orientation by an order of large index in $\mathcal{O}_K$ with action by a large class group.

Since the supersingular ℓ-isogeny graph is connected, every supersingular elliptic curve admits an ℓ-isogeny chain back to a curve oriented by any given maximal order $\mathcal{O}_K$, so such a construction exists for any supersingular elliptic curve.

5.3.3 Oriented curves and class group action

As before, we let K be an imaginary quadratic field, $\mathcal{O}_K$ its maximal order and $\mathcal{O} \subseteq \mathcal{O}_K$ an arbitrary order. Let $\mathfrak{A}$ denote a quaternion algebra in which K embeds and $\mathfrak{O}$ an arbitrary maximal order. By the hypothesis that K embeds in $\mathfrak{A}$, there exist unique primes $\mathfrak{p} \subset \mathcal{O}$ and $\mathfrak{P} \subset \mathfrak{O}$ over p. We recall that $\mathfrak{O}$ is locally at p a non-commutative discrete valuation ring with residue field $\mathfrak{O}/\mathfrak{P}$ isomorphic to $\mathbb{F}_{p^2}$. In what follows we assume that E is a supersingular elliptic curve over $\overline{\mathbb{F}}_p$, and suppose that $\mathfrak{O} = \mathrm{End}(E)$ and $\mathfrak{A} = \mathfrak{A}_{p,\infty} = \mathrm{End}^0(E) = \mathrm{End}(E) \otimes \mathbb{Q}$, and that E admits a primitive orientation by $\mathcal{O}$ (which is consequently p-maximal). We denote by σ the arithmetic Frobenius map on $\overline{\mathbb{F}}_p$, and its induced map on $\mathbb{F}_{p^2}$:

$$\begin{tikzcd}
\overline{\mathbb{F}}_p \arrow[r, "\sigma"] \arrow[d] & \overline{\mathbb{F}}_p \arrow[d] \\
\mathbb{F}_{p^2} \arrow[r, "\sigma"] & \mathbb{F}_{p^2}
\end{tikzcd}$$

and the Frobenius p-isogeny $\pi_p : E \longrightarrow E^\sigma$, where E^σ is sometimes denoted $E^{(p)}$.

Class group action

Let $\mathrm{SS}(p)$ denote the set of supersingular elliptic curves over $\overline{\mathbb{F}}_p$ up to isomorphism, and let $\mathrm{SS}_{\mathcal{O}}(p)$ be the set of $\mathcal{O}$-oriented supersingular elliptic curves up to K-isomorphism over $\overline{\mathbb{F}}_p$, and denote the subset of primitive $\mathcal{O}$-oriented curves by $\mathrm{SS}_{\mathcal{O}}^{pr}(p)$.

$$\mathrm{SS}_{\mathcal{O}}^{pr}(p) = \{\text{primitive oriented supersingular elliptic curve}/\overline{\mathbb{F}}_p\}/\simeq$$

An element of $\mathrm{SS}_{\mathcal{O}}^{pr}(p)$ consists of the data of

- a supersingular elliptic curve $E/\overline{\mathbb{F}}_p$,
- a primitive orientation $\iota : \mathcal{O} \hookrightarrow \mathrm{End}(E)$.

The additional structure of a p-orientation is a homomorphism $\rho : \mathcal{O} \longrightarrow \overline{\mathbb{F}}_p$. We note that $\mathrm{End}(E)$ is equipped with a p-orientation $\rho : \mathrm{End}(E) \hookrightarrow \overline{\mathbb{F}}_p$ given by its action on the 1-dimensional vector space of invariant differentials, with kernel $\mathfrak{P}$:

$$\alpha^* \omega_E = \rho(\alpha)\omega_E \text{ for all } \alpha \in \mathrm{End}(E).$$

For this fixed $\mathrm{End}(E)/\mathfrak{P} \hookrightarrow \mathbb{F}_{p^2} \subseteq \overline{\mathbb{F}}_p$, a p-orientation on $\mathcal{O}$ is determined by ι, as it is the unique choice of reduction such that $\omega_E \circ \iota(\alpha) = \rho(\alpha)\omega$ for all $\alpha \in \mathcal{O}$. :

$$\begin{tikzcd} \mathcal{O} \arrow[r] \arrow[rrr, bend right, "\rho"'] & \mathcal{O}/\mathfrak{p} \arrow[r, "\iota"] & \mathrm{End}(E)/\mathfrak{P} \arrow[r, phantom, "\subseteq"] & \overline{\mathbb{F}}_p. \end{tikzcd}$$

Remark. On the quaternion algebra side these are called normalized optimal embeddings, see [Bel]. We are distinguishing between the two optimal embedding which are one the complex conjugate of the other. Indeed, there are two conjugate choices for ρ and they correspond to primitive orientations on a curve E and its twist E^σ: ρ and $\overline{\rho}$ determine two points (E, ι) and $(E^{(\sigma)}, \iota^{(\sigma)})$ in $\mathrm{SS}_{\mathcal{O}}^{pr}$ where $\iota^{(\sigma)}$ is the Frobenius conjugate of ι.

If we look at it the other way around, once we fix a choice for ρ then we get a choice for ι: out of the two embeddings with the same image (conjugate to one another) we pick the one which is normalized by the choice of ρ.

Remark. What Onuki [Onu], and others following him, do is to start from the set $\mathcal{E\ell\ell}(\mathcal{O})$ of isomorphism classes of elliptic curves over $H_{\mathcal{O}}$ —the ring class field of $\mathcal{O}$— with CM by $\mathcal{O}$, and eventually take their reduction modulo p

$$\begin{aligned} \rho : \mathcal{E\ell\ell}(\mathcal{O}) &\longrightarrow \mathrm{SS}_{\mathcal{O}}^{pr}(p) \\ E &\longrightarrow (\tilde{E}, \iota_E) = (E \bmod p, [\cdot] \bmod p) \end{aligned}$$

where $[\cdot]$ is the isomorphism $\mathcal{O} \to \mathrm{End}(E)$. Here lies the choice of the reduction in the sense that this isomorphism can be chosen so to be normalized or not.

We denote by $\mathrm{SS}_{\mathcal{O}}(\rho)$ the set of oriented supersingular elliptic curves with ρ induced by ι and $\mathrm{End}(E)/\mathfrak{P} \hookrightarrow \overline{\mathbb{F}}_p$, and $\mathrm{SS}_{\mathcal{O}}(\overline{\rho})$ the opposite p-orientation class. In the notation of Belding [Bel, § 2.3.2], the set $\mathrm{SS}_{\mathcal{O}}^{pr}(\rho)$ consists of *normalized* optimal embeddings.

Remark. The p-orientation ρ restricts to the same subset $\mathrm{SS}_{\mathcal{O}}^{pr}(\rho)$ of $\mathrm{SS}_{\mathcal{O}}^{pr}(p)$ as the image of the canonical lift:

$$\rho : \mathcal{E\ell\ell}(\mathcal{O}) \longrightarrow \mathrm{SS}_{\mathcal{O}}^{pr}(p),$$

in the notation of Onuki [Onu].

Remark. The Frobenius automorphism σ induces an isomorphism $\mathrm{SS}_{\mathcal{O}}(\rho) \to \mathrm{SS}_{\mathcal{O}}(\overline{\rho})$ taking (E, ι) to (E^σ, ι^σ), noting that $\overline{\rho} = \sigma \circ \rho$.

Remark. Since K embeds in $\mathfrak{A}$, the prime p is either ramified or inert (if p splits in K, then $\mathrm{SS}_{\mathcal{O}}^{pr}(p)$ is empty, [Onu, Th 3.2] and [ACL+2, Prop. 2.19]). In the former case, $\rho = \bar{\rho}$, hence

$$\mathrm{SS}_{\mathcal{O}}(p) = \mathrm{SS}_{\mathcal{O}}(\rho) = \mathrm{SS}_{\mathcal{O}}(\bar{\rho}).$$

This follows since the image of ρ lies in $\mathbb{F}_p$:

$$\begin{tikzcd} \mathcal{O} \arrow[r] \arrow[rr, bend right, "\rho = \overline{\rho}"'] & \mathcal{O}/\mathfrak{p} \arrow[r] & \mathbb{F}_p \subseteq \overline{\mathbb{F}}_p. \end{tikzcd}$$

In the latter case, $\mathrm{SS}_{\mathcal{O}}(p)$ decomposes into the disjoint union of the set of normalized oriented curves and its conjugate:

$$\mathrm{SS}_{\mathcal{O}}(p) = \mathrm{SS}_{\mathcal{O}}(\rho) \cup \mathrm{SS}_{\mathcal{O}}(\bar{\rho}) = \mathrm{SS}_{\mathcal{O}}(\rho) \cup \mathrm{SS}_{\mathcal{O}}(\rho)^\sigma.$$

With this notation for $\mathrm{SS}_{\mathcal{O}}(\rho)$, distinguished from $\mathrm{SS}_{\mathcal{O}}(p)$, we restate the theorem from [CK1]

Theorem 5.41. *$\mathrm{SS}^{pr}_{\mathcal{O}}(\rho)$ is a torsor for $\mathcal{C}\ell(\mathcal{O})$.*

N.B.[1] This result is restated as [Onu, Th. 3.4] and follows from [Bel]; theorem 3.3 in [Onu] implies the equality

$$\mathrm{SS}^{pr}_{\mathcal{O}}(p) = \mathrm{SS}^{pr}_{\mathcal{O}}(\rho) \cup \mathrm{SS}^{pr}_{\mathcal{O}}(\bar{\rho}).$$

In [ACL+2, Prop. 4.2], the fact that these two orbits are disjoint when p is inert in $\mathcal{O}$ is proved.

In this context the class group action is given by

$$(\mathcal{C}\ell(\mathcal{O}) \rtimes \langle\sigma\rangle) \times \mathrm{SS}^{pr}_{\mathcal{O}}(p) \longrightarrow \mathrm{SS}^{pr}_{\mathcal{O}}(p)$$

By lifting σ to complex conjugation, the dihedral group $\mathcal{C}\ell(\mathcal{O}) \rtimes \langle\sigma\rangle$, can be identified with the Galois group $\mathcal{G}a\ell(H_{\mathcal{O}}/\mathbb{Q})$ acting on the set of canonical lifts $\mathcal{E}\ell\ell(\mathcal{O})$.

Remark. We note that $\mathcal{C}\ell(\mathcal{O})$ acts on $\mathrm{SS}_{\mathcal{O}}(\bar{\rho})$ via the conjugate action (with respect to that on $\mathrm{SS}_{\mathcal{O}}(\rho)$), giving a dihedral group action.

The action of the class group is given as in the ordinary case

$$\begin{tikzcd}
\mathcal{C}\ell(\mathcal{O}) \times \mathrm{SS}_{\mathcal{O}}(\rho) \arrow[r] & \mathrm{SS}_{\mathcal{O}}(p) \\
([\mathfrak{a}]\,, E) \arrow[r, mapsto] & {[\mathfrak{a}] \cdot E = E/E[\mathfrak{a}]}
\end{tikzcd}$$

where $\mathfrak{a}$ is any representative ideal coprime to the index $[\mathcal{O}_K : \mathcal{O}]$ so that the isogeny $E \to E/E[\mathfrak{a}]$ is horizontal. In particular, for fixed primitive $\mathcal{O}$-oriented E, we hence obtain a bijection of sets:

$$\begin{tikzcd}
\mathcal{C}\ell(\mathcal{O}) \arrow[r] & \mathrm{SS}^{pr}_{\mathcal{O}}(\rho) \\
{[\mathfrak{a}]} \arrow[r, mapsto] & {[\mathfrak{a}] \cdot E}
\end{tikzcd}$$

For any ideal class $[\mathfrak{a}]$ and generating set $\{\mathfrak{q}_1, \dots, \mathfrak{q}_r\}$ of small primes, coprime to $[\mathcal{O}_K : \mathcal{O}]$, we can find an identity $[\mathfrak{a}] = [\mathfrak{q}_1^{e_1} \cdot \ldots \cdot \mathfrak{q}_r^{e_r}]$, in order to compute the action via a sequence of low-degree isogenies.

Oriented supersingular isogeny graphs

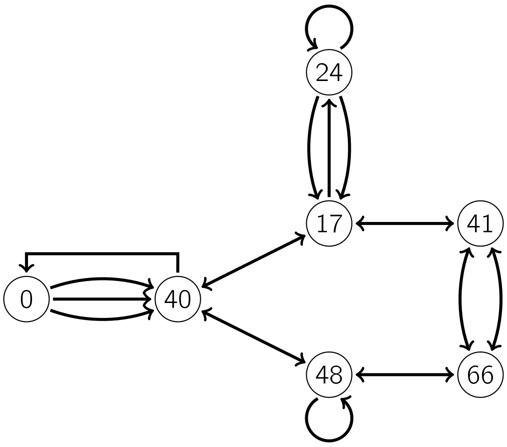
For an ordinary ℓ-isogeny isogeny graph $G_\ell(E)$, the points defined over $\mathbb{F}_{p^n}$ are determined by the condition $\mathbb{Z}[\pi^n] \subseteq \mathrm{End}(E)$. Since the class numbers of orders $\mathcal{O}$ in K are unbounded, the previous theorem implies that the oriented supersingular graphs are infinite. While all supersingular curves and isogenies can be defined over $\mathbb{F}_{p^2}$, we can use the inclusion of an order $\mathcal{O} \subset \mathrm{End}(E)$ to restrict to a finite subgraph.

Corollary 5.42. *Let (E, ι) be a K-oriented elliptic curve. The ℓ-isogeny graph $G_\ell(E, \iota)$ is an infinite graph which is the union of the finite subgraphs whose vertices are restricted to $\mathrm{SS}_{\mathcal{O}}(p)$ for an order $\mathcal{O}$ in K.*

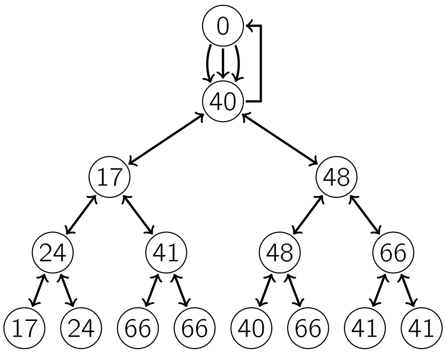
The subrings $\mathcal{O}_n = \mathbb{Z} + \ell^n\mathcal{O}$ are a linearly ordered family which serve to bound the depth of K-oriented curves relative to a curve at the surface with orientation by an ℓ-maximal order $\mathcal{O}$.

Example. Let $E_0/\overline{\mathbb{F}}_{71}$ be the supersingular elliptic curve with $j(E) = 0$, oriented by the order $\mathcal{O}_K = \mathbb{Z}[\omega]$, where $\omega^2 + \omega + 1 = 0$. The unoriented 2-isogeny graph is the finite graph:

[1] With respect to our original article [CK1] there is a change of notation from $\mathrm{SS}_{\mathcal{O}}(p)$ to $\mathrm{SS}_{\mathcal{O}}(\rho)$. This is due to an omission in our original work where we implicitly assumed that $\mathrm{SS}_{\mathcal{O}}(p)$ respected the p-orientation, but failed to state this in our definition. Since there are now multiple citations of our work, rather than change the definition, we introduce a new notation $SS_{\mathcal{O}}(\rho)$ for oriented curves respecting the p-orientation (referred to as a normalized orientation in Belding's thesis [Bel].



The orientation by $K = \mathbb{Q}[\omega]$ differentiates vertices in the descending paths from E_0, determining a lift to an infinite graph shown here to depth 4:



Consider the descending path along vertex j-invariants $(0, 40, 17, 41, 66)$, and let $\mathfrak{p}_7$ be a prime over the split prime 7. Since $\Delta_K = -3$ and $\Delta_1 = \mathrm{disc}(\mathcal{O}_1) = -12$ are of class number one, $\mathfrak{p}_7 \sim 1$, and the 7-isogenous chain is likewise of the form $(0, 40, \dots)$.

At depth 2, the class number of $\mathcal{O}_2$ of discriminant -48 is 2, and a Minkowski reduction of $\mathfrak{p}_7$ is equivalent to a prime $\mathfrak{p}_3$ over 3. In particular, this prime is non-principal of order 2, so the image chain extends $(0, 40, 48, \dots)$.

At depth 3, the class number of $\mathcal{O}_3$ is 4, and $\mathfrak{p}_7 \sim \bar{\mathfrak{p}}_7$ are primes of order 2 in the class group, hence the two 7-isogenies are to the same chain $(0, 40, 48, 48, \dots)$. Finally at depth 4 we differentiate the two primes $\mathfrak{p}_7$ and $\bar{\mathfrak{p}}_7$ in $\mathcal{O}_4$ each of order 4. The two extensions $(0, 40, 48, 48, 66)$ and $(0, 40, 48, 48, 40)$, each of which corresponds to one of the primes over 7. For a choice of prime $\mathfrak{p}_7$ we have thus determined the following ladder inducing the action of $\mathfrak{p}_7$ on the ℓ-isogeny chain.

$$\begin{tikzcd}
0 \arrow[r] \arrow[d] & 40 \arrow[r] \arrow[d] & 17 \arrow[r] \arrow[d] & 41 \arrow[r] \arrow[d] & 66 \arrow[d] \\
0 \arrow[r] & 40 \arrow[r] & 48 \arrow[r] & 48 \arrow[r] & 66
\end{tikzcd}$$

The forgetful map to unoriented isogeny graphs

In this section we address the extent of non-injectivity of the forgetful map from oriented curves in the infinite oriented supersingular ℓ-isogeny graphs to the finite supersingular graph.

By Theorem 5.41, we have a bijection (isomorphism of sets with $\mathcal{C\ell}(\mathcal{O})$-action):

$$\mathcal{C\ell}(\mathcal{O}) \cong \mathrm{SS}_{\mathcal{O}}^{pr}(p) \subseteq \mathrm{SS}_{\mathcal{O}}(p)$$

determined by any choice of base point. On the other hand, for a descending chain of imaginary quadratic orders of index ℓ,

$$\mathcal{O}_K = \mathcal{O}_0 \supset \mathcal{O}_1 \supset \cdots \supset \mathcal{O}_i \supset \cdots$$

determined by a descending ℓ-isogeny chain, the class numbers satisfy the geometric growth $h(\mathcal{O}_{i+1}) = \ell h(\mathcal{O}_i)$ for all $i \geq 1$. In particular, the inclusion $\mathcal{O}_{i+1} \subset \mathcal{O}_i$ determines an inclusion $\mathrm{SS}_{\mathcal{O}_i}(\rho) \subset \mathrm{SS}_{\mathcal{O}_{i+1}}(\rho) = \mathrm{SS}_{\mathcal{O}_i}(\rho) \cup \mathrm{SS}^{pr}_{\mathcal{O}_{i+1}}(\rho)$. Consequently we have an unbounded chain of sets

$$\mathrm{SS}_{\mathcal{O}_K}(\rho) \subset \mathrm{SS}_{\mathcal{O}_1}(\rho) \subset \cdots \subset \mathrm{SS}_{\mathcal{O}_i}(\rho) \subset \cdots$$

equipped with forgetful maps $\mathrm{SS}_{\mathcal{O}_i}(\rho) \to \mathrm{SS}(p)$ sending the $\mathcal{O}_i$-isomorphism class $[(E, \mathcal{O}_i)]$ to the isomorphism class $[E]$ determined by the j-invariant $j(E)$.

This motivates the questions of when the map $\mathrm{SS}_{\mathcal{O}_i}(\rho) \to \mathrm{SS}(p)$ and its restriction to $\mathrm{SS}^{pr}_{\mathcal{O}_i}(\rho)$ are injective, and when these maps are surjective. We adopt the notation $H(p)$ for the cardinality $|\mathrm{SS}(p)|$ of supersingular curves, denote by X_i the image of $\mathrm{SS}_{\mathcal{O}_i}(\rho)$ in $\mathrm{SS}(p)$ and write Y_i for the image of $\mathrm{SS}^{pr}_{\mathcal{O}_i}(\rho)$. Moreover we write $\lambda_i = \log_p(|\Delta_i|)$ where $\Delta_i = \ell^{2i}\Delta_K = \Delta(\mathcal{O}_i)$. With this notation, Tables 5.3-5.6 give tables of values for $|Y_i|$, $|X_i|$, and λ_i, for primes of $10, 11, 12$ and 13 bits respectively, depicting the boundary line for injectivity at $\lambda_i = 1$ and the critical line for surjectivity at $\lambda_i = 2$. We conclude this section with a general proposition, which follows from the following algebraic lemma, in order to justify the injectivity bound.

Lemma 5.43. *Let α_1 and α_2 be elements of a maximal quaternion order in a quaternion algebra over $\mathbb{Q}$ ramified at a prime p. Set $\Delta_i = \mathrm{disc}(\mathbb{Z}[\alpha_i])$ for $i \in \{1, 2\}$, and define ω to be the commutator $[\alpha_1, \alpha_2] = \alpha_1\alpha_2 - \alpha_2\alpha_1$. Then ω satisfies $\mathrm{Tr}(\omega) = 0$, $\mathrm{N}(\omega) = (\Delta_1\Delta_2 - T^2)/4$ where $T = 2\mathrm{Tr}(\alpha_1\alpha_2) - \mathrm{Tr}(\alpha_1)\mathrm{Tr}(\alpha_2)$, and $\mathrm{N}(\omega) \equiv 0 \bmod p$.*

Proof. The equality $\mathrm{Tr}(\omega) = 0$ follows from the relation $\mathrm{Tr}(\alpha_1\alpha_2) = \mathrm{Tr}(\alpha_2\alpha_1)$ and linearity of the reduced trace. The expression for the reduced norm $\mathrm{N}(\omega)$ is an elementary calculation. The congruence $\mathrm{N}(\omega) = 0 \bmod p$ holds since the unique maximal ideal $\mathfrak{P}$ over p in the quaternion order is the subset of elements α with $\mathrm{N}(\alpha) \equiv 0 \bmod p$, and the quotient by $\mathfrak{P}$ is isomorphic to the (commutative) finite field $\mathbb{F}_{p^2}$. Hence $\alpha_1\alpha_2 \equiv \alpha_2\alpha_1 \bmod \mathfrak{P}$ which implies $\omega \bmod \mathfrak{P} = 0$, from which $\mathrm{N}(\omega) \equiv 0 \bmod p$ holds. □

Proposition 5.44. *Let $\mathcal{O}$ be an imaginary quadratic order of discriminant Δ and p a prime which is inert in $\mathcal{O}$. If $|\Delta| < p$, then the map $\mathrm{SS}_{\mathcal{O}}(\rho) \to \mathrm{SS}(p)$ is injective.*

Proof. If the map is not injective, there exists a supersingular elliptic curve $E/\overline{\mathbb{F}}_p$, such that $\mathrm{End}(E)$ admits distinct embeddings $\iota_i : \mathcal{O} = \mathbb{Z}[\alpha] \to \mathrm{End}(E)$, for $i \in \{1, 2\}$. Let $\alpha_i = \iota_i(\alpha)$ and set $\omega = [\alpha_1, \alpha_2]$. By the previous lemma, we have

$$\mathrm{N}(\omega) = \frac{\Delta^2 - T^2}{4} \equiv 0 \bmod p.$$

Since p is prime, and $T \equiv \Delta \bmod 2$, we have either $|\Delta| - |T| \equiv 0 \bmod 2p$ or $|\Delta| + |T| \equiv 0 \bmod 2p$. Moreover, since $\mathrm{End}(E)$ is an order in a definite quaternion algebra, we have $\mathrm{N}(\omega) > 0$, hence $|T| < |\Delta|$. It follows that $2p \leq |\Delta| + |T| \leq 2|\Delta|$, and hence $p \leq |\Delta|$. As a consequence, we conclude that if the map is injective, then $|\Delta| < p$. □

Remark. See also [Kan, Th. 2′] for a similar result on the quaternion side.

$p = 1013$

i	$h(O_i)$	$\lvert Y_i\rvert$	$\lvert X_i\rvert$	$H(p)$	λ_i
1	1	1	1	85	0.3590
2	2	2	3	85	0.5593
3	4	4	7	85	0.7596
4	8	8	15	85	0.9599
5	16	16	29	85	1.1603
6	32	26	47	85	1.3606
7	64	43	66	85	1.5609
8	128	70	82	85	1.7612
9	256	79	85	85	1.9615
10	512	83	85	85	2.1618

$p = 1019$

i	$h(O_i)$	$\lvert Y_i\rvert$	$\lvert X_i\rvert$	$H(p)$	λ_i
1	1	1	1	86	0.3587
2	2	2	3	86	0.5588
3	4	4	7	86	0.7590
4	8	8	15	86	0.9591
5	16	15	30	86	1.1593
6	32	29	49	86	1.3594
7	64	46	69	86	1.5595
8	128	64	81	86	1.7597
9	256	83	84	86	1.9598
10	512	86	86	86	2.1600

Table 5.3 – Sizes of images of oriented classes mapping to supersingular curves for primes of 10 bits

$p = 2027$

i	$h(O_i)$	$\|Y_i\|$	$\|X_i\|$	$H(p)$	λ_i
1	1	1	1	170	0.3263
2	2	2	3	170	0.5084
3	4	4	7	170	0.6904
4	8	8	15	170	0.8725
5	16	16	31	170	1.0546
6	32	30	57	170	1.2366
7	64	55	98	170	1.4187
8	128	92	144	170	1.6007
9	256	136	166	170	1.7828
10	512	165	169	170	1.9649
11	1024	166	170	170	2.1469

$p = 2039$

i	$h(O_i)$	$\|Y_i\|$	$\|X_i\|$	$H(p)$	λ_i
1	1	1	1	171	0.3260
2	2	2	3	171	0.5080
3	4	4	7	171	0.6899
4	8	8	15	171	0.8718
5	16	15	30	171	1.0537
6	32	29	56	171	1.2357
7	64	54	94	171	1.4176
8	128	87	130	171	1.5995
9	256	132	157	171	1.7814
10	512	155	169	171	1.9634
11	1024	169	171	171	2.1453

Table 5.4 – Sizes of images of oriented classes mapping to supersingular curves for primes of 11 bits

$p = 4079$

i	$h(O_i)$	$\|Y_i\|$	$\|X_i\|$	$H(p)$	λ_i
1	1	1	1	341	0.2988
2	2	2	3	341	0.4656
3	4	4	7	341	0.6323
4	8	8	15	341	0.7991
5	16	16	31	341	0.9658
6	32	31	62	341	1.1326
7	64	61	113	341	1.2993
8	128	111	196	341	1.4661
9	256	180	276	341	1.6328
10	512	258	326	341	1.7996
11	1024	318	340	341	1.9663
12	2048	340	341	341	2.1331

$p = 4091$

i	$h(O_i)$	$\|Y_i\|$	$\|X_i\|$	$H(p)$	λ_i
1	1	1	1	342	0.2987
2	2	2	3	342	0.4654
3	4	4	7	342	0.6321
4	8	8	15	342	0.7988
5	16	16	31	342	0.9655
6	32	30	59	342	1.1322
7	64	59	110	342	1.2989
8	128	107	182	342	1.4656
9	256	186	263	342	1.6323
10	512	266	326	342	1.7990
11	1024	314	341	342	1.9657
12	2048	339	342	342	2.1323

Table 5.5 – Sizes of images of oriented classes mapping to supersingular curves for primes of 12 bits

$p = 8147$

i	$h(O_i)$	$\|Y_i\|$	$\|X_i\|$	$H(p)$	λ_i
1	1	1	1	680	0.2759
2	2	2	3	680	0.4298
3	4	4	7	680	0.5838
4	8	8	15	680	0.7377
5	16	16	31	680	0.8916
6	32	32	63	680	1.0456
7	64	64	123	680	1.1995
8	128	118	225	680	1.3535
9	256	218	369	680	1.5074
10	512	364	533	680	1.6613
11	1024	530	650	680	1.8153
12	2048	644	675	680	1.9692
13	4096	677	680	680	2.1232

$p = 8171$

i	$h(O_i)$	$\|Y_i\|$	$\|X_i\|$	$H(p)$	λ_i
1	1	1	1	682	0.2758
2	2	2	3	682	0.4297
3	4	4	7	682	0.5836
4	8	8	15	682	0.7375
5	16	16	31	682	0.8914
6	32	32	63	682	1.0452
7	64	64	123	682	1.1991
8	128	121	236	682	1.3530
9	256	223	399	682	1.5069
10	512	385	582	682	1.6608
11	1024	527	662	682	1.8147
12	2048	631	678	682	1.9686
13	4096	680	681	682	2.1225

Table 5.6 – Sizes of images of oriented classes mapping to supersingular curves for primes of 13 bits

5.3.4 Modular isogenies

In this section we consider the way in which we effectively represent and compute isogenies. With the view to oriented isogenies, we focus on horizontal isogenies with kernel $E[\mathfrak{q}]$, where E is a primitive $\mathcal{O}$-oriented elliptic curve and $\mathfrak{q}$ a prime ideal of $\iota(\mathcal{O})$. In what follows we suppress ι and identify $\mathcal{O}$ with $\iota(\mathcal{O})$.

Effective endomorphism rings and isogenies

We say a subring of $\mathrm{End}(E)$ is effective if we have explicit polynomial or rational functions which represent its generators. The subring $\mathbb{Z}$ in $\mathrm{End}(E)$ is thus effective. Examples of effective imaginary quadratic subrings $\mathcal{O} \subset \mathrm{End}(E)$, are the subring $\mathcal{O} = \mathbb{Z}[\pi]$ generated by Frobenius, for either an ordinary elliptic curve, or a supersingular elliptic curve defined over $\mathbb{F}_p$, or an elliptic curve obtained by CM construction for an order $\mathcal{O}$ of small discriminant (in absolute value).

In the Couveignes [Cou] or the Rostovtsev-Stolbunov [RS] constructions, or in the CSIDH protocol [Cas+], one works with the ring $\mathcal{O} = \mathbb{Z}[\pi]$. The disadvantage is that for large finite fields, the class group of $\mathcal{O}$ is large and the primes $\mathfrak{q}$ in $\mathcal{O}$ have no small degree elements. For large p and small q, the smallest degree element of a prime $\mathfrak{q}$ of norm q is the endomorphism $[q]$, of degree q^2. The division polynomial $\psi_q(x)$, which cuts out the torsion group $E[q]$, is of degree $(q^2-1)/2$. Consequently factoring $\psi_q(x)$ to find the kernel polynomial (see Kohel [Koh1, Chapter 2]) of degree $(q-1)/2$ for $E[\mathfrak{q}]$ is relatively expensive. As a result, in the SIDH protocol [DJP], the ordinary protocol of De Feo, Smith, and Kieffer [DKS], or the CSIDH protocol [Cas+], the curves are chosen such that the points of $E[\mathfrak{q}]$ are defined over a small degree extension κ/k, particularly $[\kappa/k] \in \{1,2\}$, and working with rational points in $E(\kappa)$.

In the OSIDH protocol outlined in Chapter 6, we propose the use of an effective CM order $\mathcal{O}_K$ of class number 1. In particular every prime $\mathfrak{q}$ of norm q is generated by an endomorphism of the minimal degree q. For example we may take $\mathcal{O}_K$ to be the Eisenstein or Gaussian integers of discriminant -3 or -4, generated by an automorphism. The kernel polynomial of degree $(q-1)/2$ can be computed directly without need for a splitting field for $E[\mathfrak{q}]$, and the computation of a generator isogeny is a one-time precomputation. Using an analog of the construction of division polynomials, the computation of the kernel polynomial requires $O(q)$ field operations.

Push forward isogenies

The extension of an isogeny (or, as we will see in the next section, of an endomorphism) of E_0 to an ℓ-isogeny chain (E_i, ϕ_i) reduces to the construction of a ladder. At each step we are given $\phi_i : E_i \to E_{i+1}$ and $\psi_i : E_i \to F_i$ of coprime degrees, and need to compute

$$\psi_{i+1} : E_{i+1} \to F_{i+1} \text{ and } \phi_i' : F_i \to F_{i+1}.$$

Rather than working with elliptic curves and isogenies, we construct the oriented graphs directly as points on a modular curve linked by modular correspondences defined by modular polynomials.

Modular curves and isogenies

The use of modular curves for efficient computation of isogenies has an established history (see Elkies [Elk2] and Chapter 2). For this purpose we represent isogeny chains and ladders as finite sequences of points on the modular curve $\mathcal{X} = X(1)$ preserving the relations given by a modular equation.

We recall that the modular curve $X(1) \cong \mathbb{P}^1$ classifies elliptic curves up to isomorphism, and the function j generates its function field. The family of elliptic curves

$$E : y^2 + xy = x^3 - \frac{36}{(j-1728)}x - \frac{1}{(j-1728)}$$

covers all isomorphism classes $j \neq 0, 12^3$ or ∞, such that the fiber over $j_0 \in k$ is an elliptic curve of j-invariant j_0. The curves $y^2 + y = x^3$ and $y^2 = x^3 + x$ deal with the cases $j = 0$ and $j = 1728$.

The modular polynomial $\Phi_m(X, Y)$ defines a correspondence in $X(1) \times X(1)$ such that $\Phi_m(j(E), j(E')) = 0$ if and only if there exists a cyclic m-isogeny ϕ from E to E', possibly over some extension field. The curve in $X(1) \times X(1)$ cut out by $\Phi_m(X, Y) = 0$ is a singular image of the modular curve $X_0(m)$ parametrizing such pairs (E, ϕ).

Remark. The modular curve $X(1)$ can be replaced by any genus 0 modular curve $\mathcal{X}$ parametrizing elliptic curves with level structure. Lifting the modular polynomials back to $\mathcal{X}$ of higher level (but still genus 0) has an advantage of reducing the coefficient size of the corresponding modular polynomials $\Phi_m(X, Y)$, see chapter 3.

In the case of CSIDH, the authors use $\mathcal{X} = X_0(4)$, with a modular function $a \in k(X_0(4))$ to parametrize the family of curves

$$E : y^2 = x(x^2 + ax + 1),$$

together with a cyclic subgroup $C \subset E$ of order 4, whose generators are cut out by $x = 1$. The map $\mathcal{X} \to X(1)$ is given by

$$j = \frac{2^8(a^2-3)^3}{(a-2)(a+2)}.$$

The approach via modular isogenies of this section can be adapted as well to the CSIDH protocol.

Definition. A *modular ℓ-isogeny chain* of length n over k is a finite sequence $(j_0, j_1, \dots, j_n)$ in k such that $\Phi_\ell(j_i, j_{i+1}) = 0$ for $0 \le i < n$. A *modular ℓ-ladder* of length n and degree q over k is a pair of modular ℓ-isogeny chains

$$(j_0, j_1, \dots, j_n) \text{ and } (j'_0, j'_1, \dots, j'_n),$$

such that $\Phi_q(j_i, j'_i) = 0$.

Clearly an ℓ-isogeny chain (E_i, ϕ_i) determines the modular ℓ-isogeny chain $(j_i = j(E_i))$, but the converse is equally true.

Proposition 5.45. *If $(j_0, \dots, j_n)$ is a modular ℓ-isogeny chain over k, and E_0/k is an elliptic curve with $j(E_0) = j_0$, then there exists an ℓ-isogeny chain (E_i, ϕ_i) such that $j_i = j(E_i)$ for all $0 \le i \le n$.*

Given any modular ℓ-isogeny chain (j_i), elliptic curve E_0 with $j(E_0) = j_0$, and isogeny $\psi_0 : E_0 \to F_0$, it follows that we can construct an ℓ-ladder $\psi : (E_i, \phi_i) \to (F_i, \phi'_i)$ and hence a modular ℓ-isogeny ladder. In fact the ℓ-ladder can be efficiently constructed recursively from the modular ℓ-isogeny chain $(j_0, \dots, j_n)$ and $(j'_0, \dots, j'_n)$, by solving the system of equations

$$\Phi_\ell(j'_i, Y) = \Phi_q(j_{i+1}, Y) = 0 \qquad \text{for } Y = j'_{i+1}$$

Remark. The modular polynomial $\Phi_q(X, Y)$ is degree $q + 1$ in X and Y. The evaluation at $X = j \in \mathbb{F}_{p^2}$ requires $O(q^2)$ field multiplications. The subsequent gcd requires $O(\ell q)$ operations, and these operations are repeated to depth n.

Now let $\Delta < 0$ be the discriminant of $\mathcal{O} \overset{m}{\subseteq} \mathcal{O}_K$. We define $\mathrm{CM}(\Delta) \subseteq X(1)_{\mathbb{Q}} = j$-line, the subscheme of CM-points of discriminant $\mathrm{discr}(\mathcal{O}) = \Delta$.

$$\begin{tikzcd}
\mathrm{Spec}(\mathbb{Z}[j]) & \subseteq X(1)_{\mathbb{Z}} & X(1)_{\mathbb{Q}} \arrow[l] \arrow[d] & \supseteq \mathrm{CM}(\Delta) \arrow[d, "p \nmid m,\quad \left(\frac{\Delta}{p}\right) = \left(\frac{\Delta_K}{p}\right) \neq 1"] \\
 & & X(1)_{\mathbb{F}_p} \arrow[ul, bend left] & \supseteq \mathrm{SS}(\mathbb{F}_p) \\
 & & \mathrm{Spec}(\mathbb{F}_p[j]) \arrow[u, phantom, sloped, "\subseteq"] \arrow[uull, bend left, "\substack{\text{Reduction}\\ \mathbb{Z}[j]\to\mathbb{F}_p[j]}"]
\end{tikzcd}$$

Remark. In our situation $\Delta = \ell^{2n}\Delta_K$ with $p \nmid m = \ell^n$.

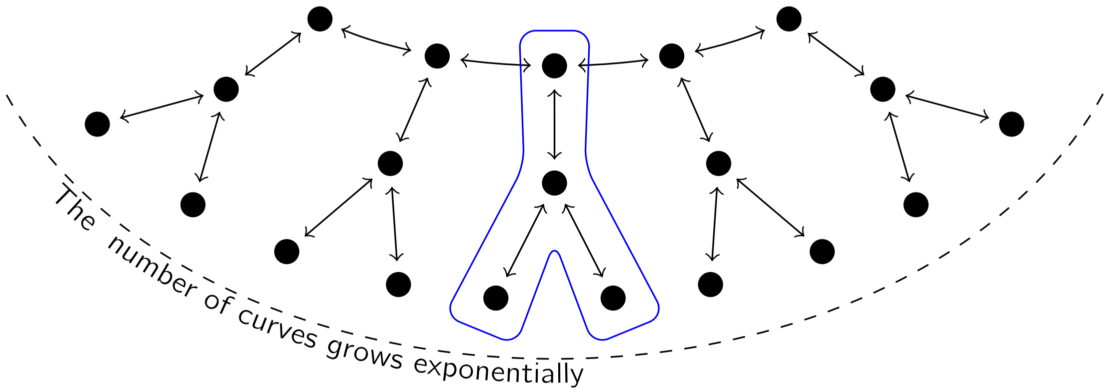
On the divisors, this gives the following picture

$$\begin{tikzcd}
\mathrm{Div}\left(\mathrm{CM}(\Delta)_{\mathbb{Q}}\right) \arrow[r] \arrow[d, phantom, sloped, "\subseteq"] & \mathrm{Div}\left(\mathrm{SS}(\mathbb{F}_p)_{\mathbb{F}_p}\right) \arrow[d, phantom, sloped, "\subseteq"] \\
\mathrm{Div}\left(X(1)_{\mathbb{Q}}\right) \arrow[r] & \mathrm{Div}\left(X(1)_{\mathbb{F}_p}\right)
\end{tikzcd}$$

For a prime q, we get a Hecke operator

$$T_q : P \longrightarrow \sum_{\underbrace{\Phi_q(P,Q) = 0}_{q+1 \text{ elements}}} Q$$

associating to an elliptic curve all its neighboring curves in the isogeny graph.



Level structure. More in general, a modular ℓ-isogeny chain is determined by a set of (supersingular) moduli points on a modular curve $\mathcal{X} = X(\Gamma)/\mathbb{F}_p$ for some $\Gamma \subset \Gamma(N)$, and edge relations given by points in the cover,

$$\mathcal{X}(\Gamma_0(\ell)) \longrightarrow \mathcal{X} \times \mathcal{X}$$

over a given pair of moduli points, or by $\mathcal{X}(\Gamma_0(\ell^{t+1})) \to \mathcal{X} \times \mathcal{X}$ when $\Gamma \subset \Gamma_0(\ell^t)$. Here $\mathcal{X}(\Gamma_0(\ell))$ is the modular curve

$$\mathcal{X}(\Gamma_0(\ell)) = X(\Gamma_0(\ell) \cap \Gamma)$$

Remark. When $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, the moduli points are the j-invariants of the supersingular points, and the image of $\mathcal{X}(\Gamma_0(\ell)$ is given by the classical modular polynomial $\Phi_\ell(x, y)$.

When working with a level structure Γ, the oriented points are associated to a ray class group $\mathcal{C}\ell(\mathcal{O}_K, \Gamma)$ preserving the Γ-structure and ℓ-isogeny chains with ray class groups $\mathcal{C}\ell(\mathcal{O}_n, \Gamma)$. When representing the ideal classes by binary quadratic forms (or lattices in $\mathbb{C}$), the equivalence class is determined by a form or lattice with basis up to equivalence by Γ rather than by the full group $\mathrm{PSL}_2(\mathbb{Z})$, see Section 5.1.

5.4 Initialization of ladders

In what follows we characterize the initialization phase of a ladder construction, i.e., the construction of q-isogenies of ℓ-chains, for level one ($\Gamma = \mathrm{PSL}_2(\mathbb{Z})$). One could note, however, that an ℓ-isogeny chain is essentially a level-$\Gamma_0(\ell^n)$ point (over a fixed on $X(1)$). The lift to a level-Γ structure preserves the local structure, but we will be able to separate points sooner (i.e. the 2-torsion of $\mathcal{C}\ell(\mathcal{O}_m)$ may lift to non 2-torsion point in $\mathcal{C}\ell(\mathcal{O}_m, \Gamma)$).

We consider an elliptic curve E_0/k ($k = \mathbb{F}_{p^2}$) with an $\mathcal{O}_K$-orientation by an effective ring $\mathcal{O}_K$; most of the times $\mathcal{O}_K$ will be of small class number or even of class number 1, e.g. $j = 0$ or $j = 12^3$ (for which $\mathcal{O}_K = \mathbb{Z}[\zeta_3]$ or $\mathbb{Z}[i]$), We also fix a small prime ℓ, and a descending ℓ-isogeny chain from E_0 to $E = E_n$. The $\mathcal{O}_K$-orientation on E_0 and ℓ-isogeny chain induces isomorphisms

$$\iota_i : \mathbb{Z} + \ell^i \mathcal{O}_K \to \mathcal{O}_i \subset \mathrm{End}(E_i),$$

and we set $\mathcal{O} = \mathcal{O}_n$.

For a small prime q which splits in $\mathcal{O}_K$, we push forward a horizontal q-isogeny $\phi_0 \in \mathrm{End}(E_0)$, to a q-isogeny $\psi : (E_i, \phi_i) \to (F_i, \phi_i')$.

If the class number of $\mathcal{O}_K$ is small, then there is the non-negligible probability (which in fact is inversely proportional to the size of the class group by the Dirichlet density theorem) that the primes above q are principal. Principal ideals lie in the identity class of $\mathcal{C}\ell(O_K)$ and correspond therefore to an endomorphism, giving a trivial action. This is even more evident if the class number of $\mathcal{O}_K$ is 1 in which case any horizontal q-isogeny $\psi_0 : E_0 \to F_0$ is, up to isomorphism $F_0 \cong E_0$, an endomorphism.

By sending $\mathfrak{q} \subset \mathcal{O}_K$ to $\psi_0 : E_0 \to F_0 = E_0/E_0[\mathfrak{q}] \cong E_0$, and pushing forward to $\psi_n : E_n \to F_n$, we obtain the effective action of $\mathcal{C}\ell(\mathcal{O})$ on ℓ-isogeny chains of length n from E_0.

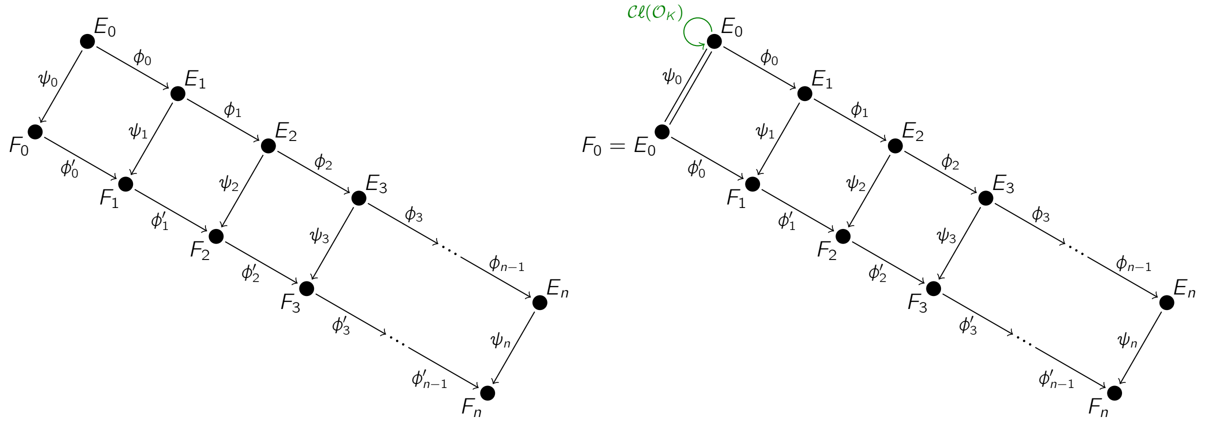


Figure 5.9 – The general case is on the left and the class number 1 case on the right.

In order to realize the class group action, it suffices to replace the ℓ-ladder of Figure 5.9 with its modular version below.

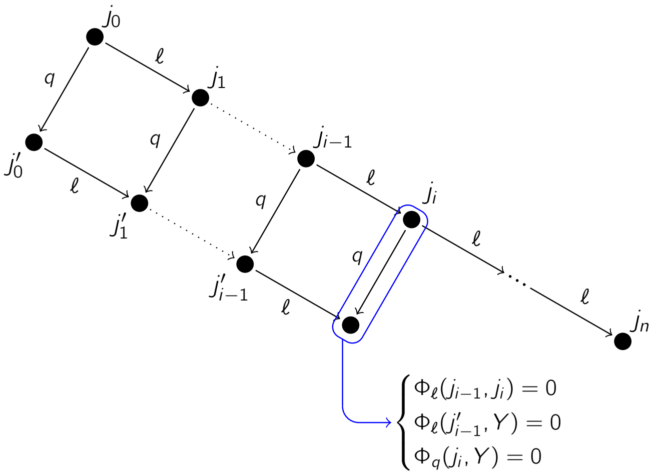


Figure 5.10 – Constructing a ladder.

At the first index for which $j'_i = j(E_i/E_i[\mathfrak{q}_i])$ is different from $j''_i = j(E_i/E_i[\bar{\mathfrak{q}}_i])$, that is, $[\mathfrak{q}_i] \neq [\bar{\mathfrak{q}}_i]$ in $\mathcal{C}\ell(\mathcal{O}_i)$, we can solve iteratively for j'_{i+1} from j'_i and j_{i+1} using the equations:

$$\Phi_\ell(j'_i, Y) = \Phi_q(j_{i+1}, Y) = 0.$$

The action of primes $\mathfrak{q}$ through $\mathcal{C}\ell(\mathcal{O})$ can be precomputed by its action on these initial segments which permits us to separate the action of $\mathfrak{q}$ and $\bar{\mathfrak{q}}$, hence assures a unique solution to the above system.

5.4.1 Action of conjugate ideal classes

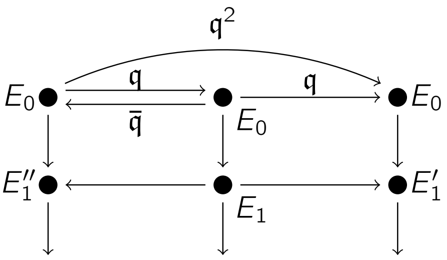
In this section, we give an answer to the question of which distance one should go before being able to separate the action of $\mathfrak{q}$ and $\bar{\mathfrak{q}}$.

Lemma 5.46. *If $\ell^{2i}|\Delta_K| > 4q$ then $\mathfrak{q}$ is not principal in $\mathcal{O}_i$, i.e., it acts non-trivially on E.*

Proof. We know that the primes above q are principal if and only if q is represented by the definite quadratic forms in the trivial class $\left[(1, 0, |\Delta_K|\ell^{2i}/4)\right]$ or $\left[(1, \ell^i, \ell^{2i}(1 + |\Delta_K|)/4)\right]$ if Δ_K is odd. □

Remark. Since $|\Delta_K|$ is minimal for 3, we can further take $\ell^{2i} > q$ in the previous lemma.

In the same way one can look at conjugate classes of ideals.



With reference to the previous picture, $E_i' \neq E_i''$ if and only if $\mathfrak{q}^2 \cap \mathcal{O}_i$ is not principal and the probability that a random ideal in $\mathcal{O}_i$ is principal is $1/h(\mathcal{O}_i)$. In fact, we can do better:

Lemma 5.47. *If $\ell^{2i} > q^2$ then $\mathfrak{q}^2$ is not principal and therefore $\mathfrak{q}$ and $\overline{\mathfrak{q}}$ act differently on E.*

Proof. We write $\mathcal{O}_K = \mathbb{Z}[\omega]$ and we observe that if $\mathfrak{q}^2$ was principal, then

$$q^2 = \mathrm{N}(\mathfrak{q}^2) = \mathrm{N}(a + b\ell^i\omega)$$

since it would be generated by an element of $\mathcal{O}_i = \mathbb{Z} + \ell^i\mathcal{O}_K$. Now

$$\mathrm{N}(a + b\ell^i) = a^2 \pm abt\ell^i + b^2s\ell^{2i} \qquad \text{where} \qquad \omega^2 + t\omega + s = 0$$

Thus, as soon as $\ell^{2i} > q^2$ we are guaranteed that $\mathfrak{q}^2$ is not principal. □

Remark. More generally, if $\mathfrak{a}$ is a non-primitive ideal, i.e., it is not contained in $m\mathcal{O}_i$ for any integer m, with $\ell^{2i} > Nr(\mathfrak{a})$, then $\mathfrak{a}$ is non principal.

The structure of oriented isogeny graphs (of level one) depends only on the class groups $\mathcal{C}\ell(\mathcal{O}_n)$ (at level n) and the quotient maps $\mathcal{C}\ell(\mathcal{O}_n) \to \mathcal{C}\ell(\mathcal{O}_{n-1})$. The quotient maps determine the edges of the ℓ-isogeny graph (between level n and $n-1$) and the class of the prime ideals over $q \neq \ell$ in $\mathcal{C}\ell(\mathcal{O}_n)$ determine edges between vertices at level n.

We assume we are given a descending modular ℓ-isogeny chain, beginning with an initial moduli point associated to a CM point with CM order $\mathcal{O}_K$. In order to initialize a q-ladder, at small distance m from the initial point, we can identify a reduced ideal class in $\mathcal{C}\ell(\mathcal{O}_m)$ which gives the same

We illustrate the procedure with an example.

Example. For discriminant $\Delta_K = -3$, and $\ell = 2$, we give in table form the first index m for which a given split prime q splits into classes outside of the 2-torsion subgroup of $\mathcal{C}\ell(\mathcal{O}_m)$. We note that for all $n \geq 3$, that

$$\mathcal{C}\ell(\mathcal{O}_n) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z}$$

and in particular, $\mathcal{C}\ell(\mathcal{O}_n)[2]$ consist of the classes of binary quadratic forms

$$\left\{\left(1, 0, |\Delta_K|\ell^{2(n-1)}\right), \left(|\Delta_K|, 0, \ell^{2(n-1)}\right), \left(\ell^2, \ell^2, n_1\right), \left(\ell^2|\Delta_K|, \ell^2|\Delta_K|, n_2\right)\right\}$$

where $\ell^4 - 4\ell^2 n_1 = \ell^4|\Delta_K|^2 - 4\ell^2|\Delta_K|n_2 = -\ell^{2n}|\Delta_K|$, whence

$$n_1 = 1 + \ell^{2(n-2)}|\Delta_K| \text{ and } n_2 = |\Delta_K| + \ell^{2(n-2)}.$$

For $n = 3$, the form $(12, 12, 7)$ reduces to $(7, 2, 7)$ and the reduced representatives are:

$$\{(1, 0, 48), (3, 0, 16), (4, 4, 13), (7, 2, 7)\}.$$

but for $n \geq 4$, since $12 < n_2$, the forms

$$\{(1, 0, 3 \cdot 4^{n-1}), (3, 0, 4^{n-1}), (4, 4, n_1), (12, 12, n_2)\}$$

are reduced.
The class group $\mathcal{C}\ell(\ell^{2n}\Delta_K)$ acts through its quotients to each level m and we can represent the oriented volcano by reduced forms, see Figure 5.11.

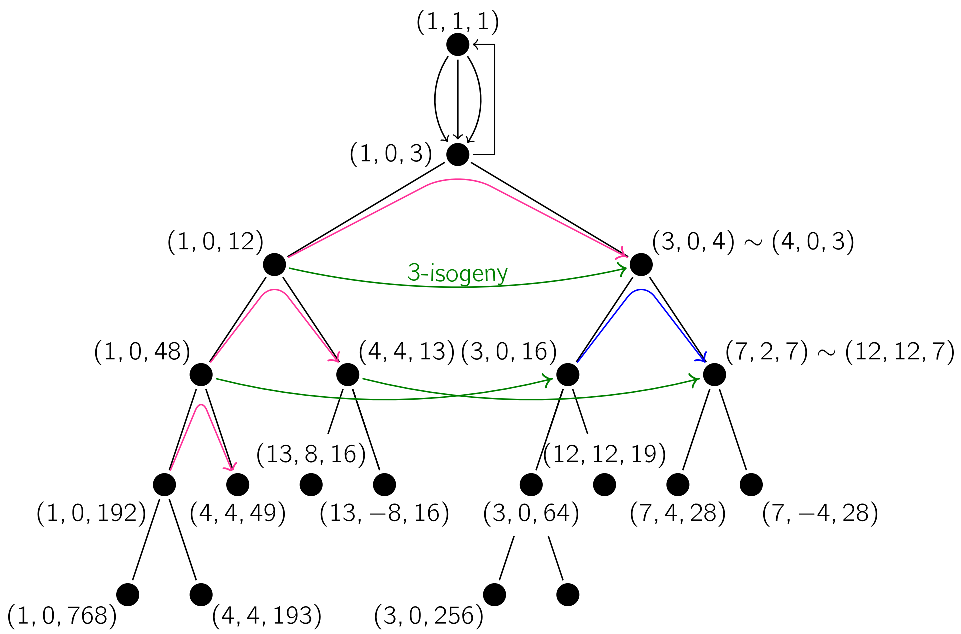


Figure 5.11 – An oriented volcano for reduced quadratic forms. It is constructed in the following way: once a descending isogeny chain is fixed we represent its elements with the identity in the class group of the corresponding level (the left side). The remaining vertices encode the quotients $\mathcal{C}\ell(\mathcal{O}_n) \to \ldots \to \mathcal{C}\ell(\mathcal{O}_2) \to \mathcal{C}\ell(\mathcal{O}_1) \to \mathcal{C}\ell(\mathcal{O}_K)$.

As we said, the initialization phase of a ladder is problematic at 2-torsion elements:

$$\begin{tikzcd}
\langle (3,0,4)\,,\underbrace{(4,4,193)}_{\text{kernel}}\rangle = \mathcal{C}\ell(\mathcal{O}_4)[2] \arrow[d] \\
\langle (3,0,64)\,,\underbrace{(4,4,49)}_{\text{kernel}}\rangle = \mathcal{C}\ell(\mathcal{O}_3)[2] \arrow[d] \\
\langle (3,0,16)\,,\underbrace{(4,4,13)}_{\text{kernel}}\rangle = \mathcal{C}\ell(\mathcal{O}_2)[2] \arrow[d] \\
\langle (3,0,4)\rangle = \mathcal{C}\ell(\mathcal{O}_1)[2] \arrow[d, "\text{prime above 3}"'] \\
\{1\} = \mathcal{C}\ell(\mathcal{O}_K)[2] = \mathcal{C}\ell(\mathcal{O}_K)
\end{tikzcd}$$

The class of $\left(3,0,2^{2(n-1)}\right)$ represents primes $q = 3x^2 + 4^{n-1}y^2$ that are either 3 or bigger than 4^{n-1}. These are not problematic since any such prime lifts to a 2-torsion element, in the class of the unique 3-isogeny (they are the green arrows in Figure 5.11). On the other hand, the class of $(4,4,c_n)$ where $c_n = 3{\cdot}4^{n-2}+1$ need to be resolved into the two 4 torsion classes in $\mathcal{O}_{n+1}$. Again, these primes are $q = 4x^2+4xy+c_ny^2 \geq 3{\cdot}4^{n-2}+1$ (magenta lines in Figure 5.11). Finally, the product class $\left(3,0,4^{n-1}\right)\cdot(4,4,c_n) = (12,12,d_n)$ represents primes bigger than $d_n = 4^{n-2}+3$. These are still 2-torsion elements at level $n-1$ and therefore need to be lifted to the next level. If we are concerned with primes $q \leq 1024$, then $n = 7$ should suffice to have non-representative in a 2-torsion class, see Table 5.7.

In the table that follows, for each split prime q in $\mathcal{O}_K = \mathbb{Z}[\omega]$ we give the level m_1 at which the primes above q are note principal, the first index m_2 such that the pair of primes over q in $\mathcal{O}_m$ lie outside the 2-torsion subgroup, the associated binary quadratic form f_m of this ideal, its reduced form $[f_m]$, and the 2-torsion class in $\mathcal{C}\ell(\mathcal{O}_{m-1})$ in which it lies.

$\Delta_K = -3$, $\ell = 2$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
7	2	4	$(7,4,28)$	$[(7,4,28)]$	$[(7,2,7)]$
13	3	4	$(13,8,16)$	$[(13,8,16)]$	$[(4,4,13)]$
19	2	5	$(19,14,43)$	$[(19,14,43)]$	$[(12,12,19)]$
31	2	4	$(31,10,7)$	$[(7,4,28)]$	$[(7,2,7)]$
37	3	4	$(37,34,13)$	$[(13,-8,16)]$	$[(4,4,13)]$
43	2	5	$(43,14,19)$	$[(19,-14,43)]$	$[(12,12,19)]$
61	3	4	$(61,56,16)$	$[(13,-8,16)]$	$[(4,4,13)]$
67	2	6	$(67,24,48)$	$[(48,-24,67)]$	$[(12,12,67)]$
73	4	5	$(73,40,16)$	$[(16,-8,49)]$	$[(4,4,49)]$
79	2	4	$(79,38,7)$	$[(7,4,28)]$	$[(7,2,7)]$
97	4	5	$(97,56,16)$	$[(16,8,49)]$	$[(4,4,49)]$
103	2	4	$(103,46,7)$	$[(7,-4,28)]$	$[(7,2,7)]$
109	3	4	$(109,70,13)$	$[(13,8,16)]$	$[(4,4,13)]$
127	2	4	$(127,116,28)$	$[(7,4,28)]$	$[(7,2,7)]$
139	2	6	$(139,120,48)$	$[(48,-24,67)]$	$[(12,12,67)]$
151	2	4	$(151,134,31)$	$[(7,-4,28)]$	$[(7,2,7)]$
157	3	4	$(157,86,13)$	$[(13,-8,16)]$	$[(4,4,13)]$
163	2	5	$(163,158,43)$	$[(19,-14,43)]$	$[(12,12,19)]$
181	3	4	$(181,104,16)$	$[(13,8,16)]$	$[(4,4,13)]$
193	5	6	$(193,8,16)$	$[(16,-8,193)]$	$[(4,4,193)]$
199	2	4	$(199,174,39)$	$[(7,4,28)]$	$[(7,2,7)]$

Table 5.7 – The behavior of primes in $\Delta_K = -3$.

Tables 5.8-5.9 present some additional data for class number 1 cases.

$\Delta_K = -3$, $\ell = 3$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
7	2	2	$(7,3,9)$	$[(7,3,9)]$	$[(1,1,7)]$
13	2	2	$(13,11,7)$	$[(7,3,9)]$	$[(1,1,7)]$
19	2	2	$(19,17,7)$	$[(7,-3,9)]$	$[(1,1,7)]$
31	2	2	$(31,25,7)$	$[(7,3,9)]$	$[(1,1,7)]$
37	2	2	$(37,33,9)$	$[(7,-3,9)]$	$[(1,1,7)]$
43	2	2	$(43,31,7)$	$[(7,-3,9)]$	$[(1,1,7)]$
61	3	3	$(61,3,9)$	$[(9,-3,61)]$	$[(1,1,61)]$
67	3	3	$(67,15,9)$	$[(9,3,61)]$	$[(1,1,61)]$
73	3	3	$(73,21,9)$	$[(9,-3,61)]$	$[(1,1,61)]$
79	2	2	$(79,51,9)$	$[(7,-3,9)]$	$[(1,1,7)]$
97	2	2	$(97,57,9)$	$[(7,3,9)]$	$[(1,1,7)]$
103	3	3	$(103,39,9)$	$[(9,-3,61)]$	$[(1,1,61)]$
109	2	2	$(109,53,7)$	$[(7,3,9)]$	$[(1,1,7)]$

Table 5.8 – The case $\Delta_K = -3$ and $\ell = 3$.

5.4.2 Extending the ladder

After the initialization phase it only remains to extend an initial segment of a ladder to the full ladder. Ultimately, this is the same problem of completing a square (or a rectangle) of isogenies, see Section 2.3, and therefore we can use any of the methods described there. However, we present here a simplification we might have when working with small ℓ. When working with small primes, the most efficient way of completing a ladder is by solving the system of modular equations

$$\Phi_\ell(j', Y) = \Phi_q(j_{i+1}, Y)$$

$\Delta_K = -4$, $\ell = 2$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
5	2	3	(5, 2, 13)	[(5, 2, 13)]	[(4, 4, 5)]
13	2	3	(13, 2, 5)	[(5, −2, 13)]	[(4, 4, 5)]
17	3	4	(17, 8, 16)	[(16, −8, 17)]	[(4, 4, 17)]
29	2	3	(29, 18, 5)	[(5, 2, 13)]	[(4, 4, 5)]
37	2	3	(37, 22, 5)	[(5, −2, 13)]	[(4, 4, 5)]
41	3	4	(41, 40, 16)	[(16, −8, 17)]	[(4, 4, 17)]
53	2	3	(53, 50, 13)	[(5, −2, 13)]	[(4, 4, 5)]
61	2	3	(61, 54, 13)	[(5, 2, 13)]	[(4, 4, 5)]
73	4	5	(73, 24, 16)	[(16, 8, 65)]	[(4, 4, 65)]
89	4	5	(89, 40, 16)	[(16, −8, 65)]	[(4, 4, 65)]
97	3	4	(97, 72, 16)	[(16, −8, 17)]	[(4, 4, 17)]
101	2	3	(101, 42, 5)	[(5, −2, 13)]	[(4, 4, 5)]
109	2	3	(109, 92, 20)	[(5, 2, 13)]	[(4, 4, 5)]
113	4	5	(113, 56, 16)	[(16, 8, 65)]	[(4, 4, 65)]

$\Delta_K = -4$, $\ell = 3$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
5	1	2	(5, 4, 17)	[(5, 4, 17)]	[(2, 2, 5)]
13	2	2	(13, 12, 9)	[(9, 6, 10)]	[(1, 0, 9)]
17	1	2	(17, 4, 5)	[(5, −4, 17)]	[(2, 2, 5)]
29	1	2	(29, 16, 5)	[(5, 4, 17)]	[(2, 2, 5)]
37	2	2	(37, 34, 10)	[(9, −6, 10)]	[(1, 0, 9)]
41	1	3	(41, 6, 18)	[(18, −6, 41)]	[(2, 2, 41)]
53	1	3	(53, 30, 18)	[(18, 6, 41)]	[(2, 2, 41)]
61	2	2	(61, 46, 10)	[(9, 6, 10)]	[(1, 0, 9)]
73	2	2	(73, 48, 9)	[(9, 6, 10)]	[(1, 0, 9)]
89	1	2	(89, 78, 18)	[(5, −4, 17)]	[(2, 2, 5)]
97	3	3	(97, 24, 9)	[(9, −6, 82)]	[(1, 0, 81)]
101	1	3	(101, 66, 18)	[(18, 6, 41)]	[(2, 2, 41)]
109	2	2	(109, 60, 9)	[(9, −6, 10)]	[(1, 0, 9)]
113	1	2	(113, 44, 5)	[(5, −4, 17)]	[(2, 2, 5)]

$\Delta_K = -7$, $\ell = 2$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
11	2	4	(11, 10, 43)	[(11, 10, 43)]	[(11, 6, 11)]
23	2	5	(23, 10, 79)	[(23, 10, 79)]	[(23, 18, 23)]
29	3	4	(29, 8, 16)	[(16, −8, 29)]	[(4, 4, 29)]
37	3	4	(37, 24, 16)	[(16, 8, 29)]	[(4, 4, 29)]
43	2	4	(43, 10, 11)	[(11, −10, 43)]	[(11, 6, 11)]
53	3	4	(53, 40, 16)	[(16, −8, 29)]	[(4, 4, 29)]
67	2	4	(67, 34, 11)	[(11, 10, 43)]	[(11, 6, 11)]
71	2	6	(71, 56, 112)	[(71, 56, 112)]	[(28, 28, 71)]
79	2	5	(79, 10, 23)	[(23, −10, 79)]	[(23, 18, 23)]
107	2	4	(107, 54, 11)	[(11, −10, 43)]	[(11, 6, 11)]
109	3	4	(109, 72, 16)	[(16, −8, 29)]	[(4, 4, 29)]
113	4	5	(113, 8, 16)	[(16, −8, 113)]	[(4, 4, 113)]
127	2	6	(127, 86, 71)	[(71, 56, 112)]	[(28, 28, 71)]

$\Delta_K = -7$, $\ell = 3$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
11	1	1	(11, 5, 2)	[(2, −1, 8)]	[(1, 1, 2)]
23	1	1	(23, 11, 2)	[(2, 1, 8)]	[(1, 1, 2)]
29	1	1	(29, 13, 2)	[(2, −1, 8)]	[(1, 1, 2)]
37	1	2	(37, 5, 4)	[(4, 3, 36)]	[(4, 1, 4)]
43	1	2	(43, 11, 4)	[(4, −3, 36)]	[(4, 1, 4)]
53	1	1	(53, 19, 2)	[(2, 1, 8)]	[(1, 1, 2)]
67	2	2	(67, 61, 16)	[(9, −3, 16)]	[(1, 1, 16)]
71	1	1	(71, 47, 8)	[(2, −1, 8)]	[(1, 1, 2)]
79	2	2	(79, 67, 16)	[(9, 3, 16)]	[(1, 1, 16)]
107	1	1	(107, 77, 14)	[(2, 1, 8)]	[(1, 1, 2)]
109	1	3	(109, 67, 22)	[(22, 21, 63)]	[(7, 7, 22)]
113	1	1	(113, 29, 2)	[(2, −1, 8)]	[(1, 1, 2)]
127	2	2	(127, 103, 22)	[(9, −3, 16)]	[(1, 1, 16)]

$\Delta_K = -8$, $\ell = 2$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
11	1	2	(11, 2, 3)	[(3, −2, 11)]	[(3, 2, 3)]
17	2	3	(17, 10, 9)	[(9, 8, 16)]	[(4, 4, 9)]
19	1	2	(19, 10, 3)	[(3, 2, 11)]	[(3, 2, 3)]
41	3	4	(41, 24, 16)	[(16, 8, 33)]	[(4, 4, 33)]
43	1	2	(43, 42, 11)	[(3, −2, 11)]	[(3, 2, 3)]
59	1	2	(59, 52, 12)	[(3, −2, 11)]	[(3, 2, 3)]
67	1	2	(67, 26, 3)	[(3, −2, 11)]	[(3, 2, 3)]
73	2	3	(73, 46, 9)	[(9, 8, 16)]	[(4, 4, 9)]
83	1	2	(83, 72, 16)	[(3, 2, 11)]	[(3, 2, 3)]
89	2	3	(89, 72, 16)	[(9, 8, 16)]	[(4, 4, 9)]
97	2	3	(97, 78, 17)	[(9, −8, 16)]	[(4, 4, 9)]
107	1	2	(107, 34, 3)	[(3, 2, 11)]	[(3, 2, 3)]
113	3	4	(113, 72, 16)	[(16, −8, 33)]	[(4, 4, 33)]
131	1	2	(131, 38, 3)	[(3, −2, 11)]	[(3, 2, 3)]

$\Delta_K = -8$, $\ell = 3$

q	m_1	m_2	f_m	$[f_m]$	$[f_{m-1}]$
11	1	2	(11, 10, 17)	[(11, 10, 17)]	[(2, 0, 9)]
17	1	2	(17, 10, 11)	[(11, −10, 17)]	[(2, 0, 9)]
19	2	2	(19, 6, 9)	[(9, −6, 19)]	[(1, 0, 18)]
41	1	2	(41, 34, 11)	[(11, 10, 17)]	[(2, 0, 9)]
43	2	2	(43, 30, 9)	[(9, 6, 19)]	[(1, 0, 18)]
59	1	2	(59, 58, 17)	[(11, −10, 17)]	[(2, 0, 9)]
67	2	2	(67, 42, 9)	[(9, −6, 19)]	[(1, 0, 18)]
73	2	2	(73, 70, 19)	[(9, −6, 19)]	[(1, 0, 18)]
83	1	3	(83, 12, 18)	[(18, −12, 83)]	[(2, 0, 81)]
89	1	3	(89, 24, 18)	[(18, 12, 83)]	[(2, 0, 81)]
97	2	2	(97, 82, 19)	[(9, 6, 19)]	[(1, 0, 18)]
107	1	2	(107, 84, 18)	[(11, −10, 17)]	[(2, 0, 9)]
113	1	3	(113, 48, 18)	[(18, −12, 83)]	[(2, 0, 81)]
131	1	3	(131, 60, 18)	[(18, 12, 83)]	[(2, 0, 81)]

Table 5.9 – Additional cases.

which consists in the GCD of two modular polynomials, see Figure 5.10. These can be reduced in order to simplify the computation; we illustrate the idea with $\ell = 2$.

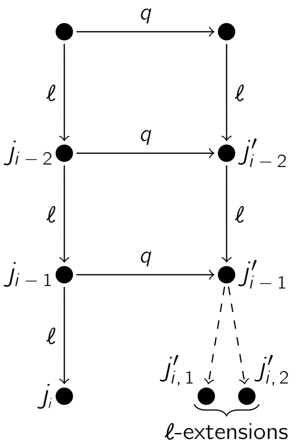


Figure 5.12 – Extending a ladder

Given an initial segment of a ladder, the i-th step consist in a choice between two possible extensions determined by a quadratic polynomial $\phi_\ell(Y)$with roots j'_{i-2} and j'_{i-1} (since we want to avoid backtracking), see Figure 5.12. The choice between one of them is determined by the horizontal isogenies and, therefore, by a degree $q+1$ equation $\phi_q(Y) = 0$ which is determined by j_i. The simplification comes from the fact that we do not need to compute the whole ϕ_q but it suffices to work with $\phi_q(Y) \bmod \phi_\ell(Y)$. In fact,

$$\Phi_q(X, Y) \equiv \phi_q(Y) \mod (X - j_i, \phi_\ell(Y))$$

By the Chinese Remainder Theorem this is the same as

$$\big(\Phi_q(j_i, j'_{i,1}), \Phi_q(j_i, j'_{i,2})\big) \mod (X - j_i, Y - j'_{i,1}) \times (X - j_i, Y - j'_{i,2})$$

since $\phi_\ell(Y) = (Y - j'_{i,1})(Y - j'_{i,2})$.

$$\begin{tikzcd}
\dfrac{k[X,Y]}{(X - j_i, \phi_\ell(Y))} \arrow[d] \arrow[dr] & & & \Phi_q(X,Y) \arrow[d, mapsto] \arrow[dr, mapsto] & \\
\dfrac{k[Y]}{(\phi_\ell(Y))} \arrow[r] & \dfrac{k[Y]}{(Y - j'_{i,1})} \times \dfrac{k[Y]}{(Y - j'_{i,2})} \simeq k^2 & & \Phi_q(j_i, Y) \arrow[r, mapsto] & \big(\Phi_q(j_i, j'_{i,1}), \Phi_q(j_i, j'_{i,2})\big)
\end{tikzcd}$$

Assuming that one of $\Phi_q(j_i, j'_{i,1})$ or $\Phi_q(j_i, j'_{i,2})$ is non-zero, we have identified in the other the correct root j'_i for which $\Phi_q(j_i, j'_i) = 0$. This is what we would have found by

$$\gcd\big(\Phi_\ell(j'_{i-1}, Y)/(Y - j_{i-2}), \Phi_q(j_i, Y)\big)$$

but we can avoid computing the polynomial $\Phi_q(j_i, Y) = \phi_q(Y)$.

5.4.3 Weber isogeny ladders

So far we have constructed ladders on $X(1)$, by means of representing the nodes by j-invariants. We now show how we can increase the level structure to obtain benefits in the construction of a ladder, both in the initialization and in the extension phase. First of all, with an ℓ-level structure, the extension of ℓ-isogenies by modular correspondences allows one to automatically remove the dual isogeny (backtracking): there are ℓ rather than $\ell + 1$ extensions and we do not need to quotient out the last root. Secondly, adding level structure helps rigidifying the automorphism group of elliptic curves and this eliminates problems such as problematic vertices with a different number of ingoing and outgoing edges. The rigidification of the automorphisms also shorten the distance to which we need to go in order to differentiate 2 points, namely two torsion elements of $\mathcal{Cl}(\mathcal{O})$ may lift to non 2-torsion point in $\mathcal{Cl}(\mathcal{O},\Gamma)$. Last but not least, q modular polynomials of higher level are smaller in terms of numbers of monomials and coefficient size.

The best known reduction in coefficient size as well as in sparsity of coefficients is obtained for the Weber function $\mathfrak{f}$ of level 48, see Section 3.3.5,

$$\mathfrak{f}(\tau) = \zeta_{48}^{-1}\frac{\eta\left(\frac{\tau+1}{2}\right)}{\eta(\tau)},$$

which generates a degree-72 cover of the j-line, given by

$$j = \frac{(\mathfrak{f}^{24}-16)^3}{\mathfrak{f}^{24}}.$$

The modular polynomials with respect to $\mathfrak{f}$ are the integral polynomials $\Phi_q(x,y)$ such that

$$\Phi_q(\mathfrak{f}(\tau),\mathfrak{f}(q\tau)) = 0.$$

Although the Weber function does not generate the full modular curve $X(48)$, which has genus 2689, it still satisfies a transformation giving the following symmetry of its induced modular polynomials.

Lemma 5.48. *The modular functions $\Phi_q(x,y)$ of prime level q with respect to the Weber function satisfies the transformation:*

$$\Phi_q(\zeta_{24}x,\zeta_{24}^{q}y) = \zeta_{24}^{q+1}\Phi_q(x,y),$$

with respect to a primitive 24-th root of unity ζ_{24}.

Asymptotically, modular polynomials have q^2 monomials, but in a practical range the sparseness is dictated by this transformation (on the order of $q^2/24$ monomials) makes these polynomials attractive for constructing isogeny invariants.

$$\begin{aligned}
&\Phi_5(x,y) = x^6 - x^5y^5 + 4xy + y^6\\
&\Phi_7(x,y) = x^8 - x^7y^7 + 7x^4y^4 - 8xy + y^8\\
&\Phi_{11}(x,y) = x^{12} - x^{11}y^{11} + 11x^9y^9 - 44x^7y^7 + 88x^5y^5 - 88x^3y^3 + 32xy + y^{12}
\end{aligned}$$

For example, the modular polynomial $\Phi_{71}(x,y)$ has exactly $3\cdot 71 = 213$ nonzero coefficients, ignoring the symmetry $\Phi_q(x,y) = \Phi_q(y,x)$, which implies an even smaller number of distinct coefficients.

In the interest of constructing ℓ-isogeny chains, especially for $\ell = 2$ or $\ell = 3$, we note that the 48-level structure gives the modular polynomials $\Phi_2(x,y)$ and $\Phi_3(x,y)$ a particular form. We descend the 2-level structure by setting $t = -\mathfrak{f}^8$, so that

$$j = \left(\frac{t^3+16}{t}\right)^3$$

With respect to this function, we obtain the modular polynomial:

$$\Psi_2(x,y) = (x^2-y)y + 16x$$

and the Weber modular polynomial $\Phi_2(x,y) = -\Psi_2(-x^8,-y^8)$ remains irreducible.

Remark. More correctly, the modular polynomial $\Psi_2(x,y)$ satisfies $\Psi_2(\mathfrak{f}_1^8(\tau),\mathfrak{f}_1^8(2\tau)) = 0$, where $\mathfrak{f}_1$ is the conjugate Weber function

$$\mathfrak{f}_1^8(\tau) = -\mathfrak{f}(\tau+3)^8 = \left(\frac{\eta(\frac{\tau}{2})}{\eta(\tau)}\right)^8$$

and hence $\Psi_2(\mathfrak{f}^8(\tau),\mathfrak{f}^8(2\tau-3)) = 0$. Nevertheless, this modular relation describes a 2-isogeny relation of the underlying curves, extending the parametrized 2-isogeny to a 4-isogeny, and will be used for defining our 2-isogeny chains.

A similar descent of the 3-level to the function $r = \mathfrak{f}^3$, gives the modular polynomial

$$\Psi_3(x,y) = x^4 - x^3y^3 + 8xy + y^4,$$

such that $\Psi_3(r(\tau), r(3\tau)) = 0$, for which $\Phi_3(x,y) = \Psi_3(x^3,y^3)$ is irreducible. For a given supersingular Weber invariant, these relations determine orbits under multiplication by ζ_8 or ζ_3, but in view of the global relation of Lemma 5.48, the lift to the orbit can be chosen to be compatible with isogeny relations of other prime degrees.

Weber initialization

In what follows we denote by u a supersingular value of the Weber function, $r = u^3$, $t = -u^8$ and $s = t^3$. This gives the following relations with the j-invariant:

$$j = \frac{(u^{24}-16)^3}{u^{24}} = \frac{(r^8-16)^3}{r^8} = \left(\frac{t^3+16}{t}\right)^3 = \frac{(s+16)^3}{s}.$$

While we will only evoke the elliptic curves associated to Weber invariants, we note that such a curve can be viewed as a fiber in the family:

$$y^2 + xy = x^3 - \frac{1}{u^{24}-64}x$$

over u on the Weber curve $\mathcal{X}$.

In the table below we give associated values for the CM j-invariants and s-invariants (where $s = -u^{24}$) for the discriminants of the first class number one CM orders, and their index 2 orders.

D	j_0	s_0	t_0
-3	0	-2^4	$-(\sqrt[3]{2})^4$
-4	12^3	2^3	2
-7	-15^3	-1	-1
-8	20^3	2^6	2^2
-12	$2^4 15^3$	-2^8	$-(\sqrt[3]{2})^8$
-16	66^3	2^9	2^3
-28	255^3	-2^{12}	-2^4
-32	j_1	t_1^3	$2^3(\sqrt{2}+1)$

In what follows we describe the initialization of a Weber ℓ-isogeny chain for $\ell = 2$. In view of the previous form of modular polynomials at 2 we use $t = -u^8$ and the modular polynomial

$$\Psi_2(x,y) = (x^2-y)y + 16x$$

to construct the 2-isogeny graph on t-values.

Remark. Associated to a t-value t_i is a u-value u_i satisfying $t_i = -u_i^8$, which was can obtain by extracting three square roots (making arbitrary choices of signs, the result is determined up to an 8-th root of unity). It remains to be determined whether the sequence (u_i) of Weber values is sufficient, or whether a sequence of points

$$U_i = (u_1,u_2,u_3)_i \in \mathcal{W}_{24}(\mathbb{F}_{p^2})$$

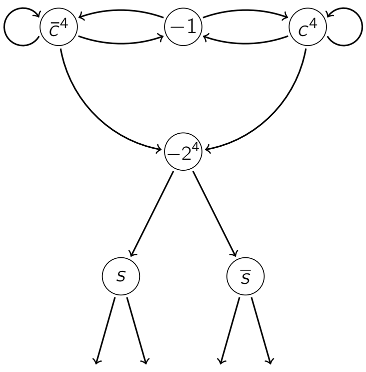
should be computed such that $u_1^8 + u_2^8 + u_3^8 = 0$ and $u_1u_2u_3 = \sqrt{2} \in \mathbb{F}_{p^2}$.

Discriminant -7. The fundamental discriminant -7 is an interesting starting point as the endomorphism ring is small enough to be effective — generated by an endomorphism of degree 2 — while avoiding any pathologies associated with the extra automorphisms for $D = -3$ and $D = -4$. In fact, choosing $p = 1 \bmod 12$ we can assure that no supersingular point has extra automorphisms.

Let $t_0 = -1$ and let c be a root of $x^2 - x + 2$ in $\mathbb{F}_{p^2} = \mathbb{F}_p[c]$, with conjugate $\bar{c} = 1 - c = 2/c$. We note that c^4 and $\bar{c}^4$ are also t-values over $j = -15^3$, and since $\Psi_2(-1,c^4) = \Psi_2(-1,\bar{c}^4) = 0$, the two extensions correspond to the horizontal 2-isogenies (endomorphisms of the underlying elliptic curve). Subsequently, we find

$$\Psi_2(c^4,c^4) = \Psi_2(c^4,-2^4) = 0.$$

The former enters a cycle of degree-2 endomorphisms, while the latter induces a descending isogeny, as depicted in the 2-isogeny graph below.



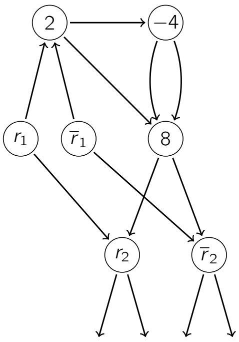
This suggests the following initialization of the 2-isogeny chain of t-values $(t_0, t_1, t_2, \dots)$ beginning with $(-1, c^4, -2^4, \dots)$. Successive solutions to $\Psi_2(t_i, t_{i+1}) = 0$ are necessarily descending with respect to the orientation by $\mathbb{Q}(\sqrt{-7})$. A random choice of root t_{i+1} of $\Psi_2(t_i, x)$ completes this initialization of the 2-isogeny chain to any desired depth n.

The above discussion, applicable in any characteristic, leaves open the question of the field of definition of the u-values, in particular whether $u_1 = \sqrt[8]{-c^4}$ is in $\mathbb{F}_{p^2}$. We give an affirmative response to this question in general in Theorem 5.38 which follows.

Discriminant -4**.** The t-invariants over $j = 12^3$ fall in two orbits of points, $\{2, 2\omega, 2\omega^2\}$ of multiplicity 2, and $\{-4, -4\omega, -4\omega^2\}$ of multiplicity 1. These points at the surface are linked by a 2-isogeny (the degree 2 endomorphism by $1 + i$), since $\Psi_2(2, -4) = 0$ and to 2-depth 1, to $t = 8$:

$$\Psi_2(2, 8) = \Psi_2(-4, 8) = 0.$$

Given that $\Psi_2(\omega x, \omega^2 y) = \omega \Psi_2(x, y)$, the choice of representative in the orbit gives rise to one of three distinct components of the 2-isogeny graph, with the component of 2 depicted below.



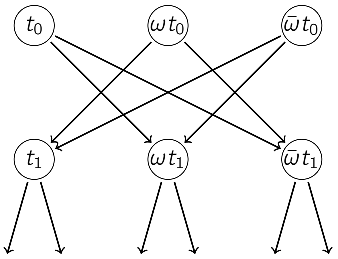
This suggests the initialization $(t_0, t_1, t_2, \dots) = (2, 8, 8c, \dots)$ of the 2-isogeny chain, where c is a root of $x^2 - 8x - 2$, extended by random selection of a root t_{i+1} of $\Psi_2(t_i, x)$.

Remark. The full 2-isogeny graph has ascending edges from the depth one points $t_1' = -4 + 3\sqrt{2}$ and its conjugate $t_1'' = -4 - 3\sqrt{2}$ to $t_0 = 2$, as well as a descending isogeny from each of t_1' and t_1'' to depth 2. In general, if an isogeny is descending its only extension to a 2-isogeny chain is descending (since this is a property of the underlying j-invariant chain).

Discriminant -3**.** In what follows we will give preference to the discriminant -3, which we treat next. Let $t_0 = -(\sqrt[3]{2})^4 = -2\sqrt[3]{2}$. Then $\{t_0, t_0\omega, t_0\omega^2\}$ is the set of t-values over $j = 0$, each of multiplicity 3, where as above $\omega^2 + \omega + 1 = 0$. Setting $t_1 = -t_0^2$, we verify that

$$\Psi_2(t_0, t_1\omega) = \Psi_2(t_0, t_1\omega^2) = 0,$$

Since 2 is inert, every path from t_0 is descending, so we may initialize the 2-isogeny chain with $(t_0, t_1\omega)$. The graph of descending isogenies from the surface vortex appears as follows.



As above there are additional t-invariants at each depth greater than 0 which admit ascending and descending isogenies. At depth 1, we have $\{t_1', t_1'\omega, t_1'\omega^2\}$ and $\{t_1'', t_1''\omega, t_1''\omega^2\}$, which ascending to the surface points $\{t_0, t_0\omega, t_0\omega^2\}$ and descending (in bijection) with the points at depth 2. Any descending isogenies must rejoin this graph of descending isogenies from the surface.

Lifting 2-isogeny chains from t-invariants to Weber points. The modular curve $\mathcal{W}_3$ has affine model

$$x_0 + x_1 + x_2 = 0,\ x_0x_1x_2 = 16,$$

from which we derive $x_0x_1^2 + x_0^2x_1 + 16 = 0$, and projects to the t-value $t = -x_0$. Let (y_0, y_1, y_2) be an affine point 2-isogenous to (x_0, x_1, x_2). The two possible extensions, determined by the modular polynomial $\Psi_2(x, y) = (x^2 - y)y + 16x$ on $(-x_0, -y_0)$ are given by the solutions

$$y_0 = x_0x_1 \text{ or } y_0 = x_0x_2.$$

We may assume that $y_0 = x_0x_1$, determined by the prior choice of point (x_0, x_1, x_2), as opposed to (x_0, x_2, x_1). The neighboring 2-isogenous point (y_0, y_1, y_2) is determined by a choice of solution to the quadratic equation:

$$y_1^2 + y_0y_1 + x_2.$$

noting that the other root is $y_2 = -y_0 - y_1$. The third coordinate is redundant, given the dependence on (y_0, y_1). The lifts to points of $\mathcal{W}_{24}$ are choices of points (u_0, u_1, u_2) and (v_0, v_1, v_2) such that

$$(u_0^8, u_1^8, u_2^8) = (x_0, x_1, x_2),\ u_0u_1u_2 = \sqrt{2},$$

and

$$(v_0^8, v_1^8, v_2^8) = (y_0, y_1, y_2),\ v_0v_1v_2 = \sqrt{2},$$

for a fixed element $\sqrt{2}$ of $\mathbb{F}_{p^2}$. The values of (u_0, u_1) and (v_0, v_1) can be arbitrarily chosen, up to a power of ζ_8, after which the product relation determines u_2 and v_2.

Chapter 6

OSIDH

By imposing the data of an orientation by an imaginary quadratic ring $\mathcal{O}$, we obtain an augmented category of supersingular curves on which the class group $\mathcal{C}\ell(\mathcal{O})$ acts faithfully and transitively. This idea is already implicit in the CSIDH protocol [Cas+], in which supersingular curves over $\mathbb{F}_p$ are oriented by the Frobenius subring $\mathbb{Z}[\pi] \cong \mathbb{Z}[\sqrt{-p}]$. In contrast we consider an elliptic curve E_0 oriented by a CM order $\mathcal{O}_K$ of class number one. To obtain a nontrivial group action, we consider descending ℓ-isogeny chains in the ℓ-volcano, on which the class group of an order $\mathcal{O}$ of large index ℓ^n in $\mathcal{O}_K$ acts. The map from an ℓ-isogeny chain to its terminal node forgets the structure of the orientation, giving rise to a generic curve in the supersingular isogeny graph. Within this general framework we define now a new oriented supersingular isogeny Diffie-Hellman (OSIDH) protocol, which has fewer restrictions on the proportion of supersingular curves covered and on the torsion group structure of the underlying curves. Moreover, the group action can be carried out effectively solely on the sequences of modular points (such as j-invariants) on a modular curve, thereby avoiding expensive isogeny computations, and is further amenable to speed-up by precomputations of endomorphisms on the base curve E_0.

The OSIDH protocol is an application of the ideas developed so far rather than an effective cryptographic protocol. In fact, while SIDH [DJP] or CSIDH [Cas+] have undergone specific refinement to practical cryptography, OSIDH is more of a framework in which we do not propose specific parameters.

6.1 Isogeny based cryptographic protocols

Secure public key cryptography is based on the existence of a trapdoor one-way public function $f : X \to Y$, such that the secret trapdoor permits efficient inversion of the function. The determination of the trapdoor, or more generally of a preimage for the function, is supposed to be computationally difficult, i.e., the computational cost to find x from $y = f(x)$ is assumed to be prohibitively expensive compared to the cost of evaluating $f(x)$. Typical cryptosystems are RSA [RSA] and ElGamal [ElG], based on arithmetic in the unit groups of the finite rings $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. The security of these cryptosystems relies on the difficulty of the problem of factoring integers and of the discrete logarithm problem. Due to the existence of special subexponential algorithms for these problems, the unit group of $\mathbb{F}_p$ in ElGamal can be replaced by the group of points $E(\mathbb{F}_p)$ of an elliptic curve over a finite field $\mathbb{F}_p$, for which no such subexponential algorithm is known. As such, one can exploit the differential between polynomial time group operations (for encryption), and exponential algorithms (for cryptanalytic attacks).

The advent of a quantum computer permits polynomial time quantum algorithms for factorization and discrete logarithms, including on elliptic curves, eliminating the (sub)exponential differential between encryption and cryptanalytic attacks, see Shor's algorithm [Sho]. Suitable replacements for the classical (pre-quantum) hard problems have been proposed, to address the new post-quantum axioms for computing. These include lattice-based problems (the closest vector problem) and problems from coding theory (the decoding problem - finding the nearest codeword). The latest member of these candidates for quantum hard problems is a supersingular isogeny path problem on elliptic curves. This problem had previously served as a basis for a new hash function, and was recently exploited as the basis of a key exchange algorithm, the supersingular elliptic curve isogeny Diffie-Hellman protocol [DJP].

The supersingular isogeny path problem takes place in the geometric category of supersingular elliptic curves. In the equivalent category of left ideals for a quaternion order, the analogous path problem has been shown to be polynomial time in practice [KLPT]. An analogous situation holds between the discrete logarithms problem (DLP) in finite fields (for which the best known algorithms are subexponential) and the DLP in an additive abelian group (which is trivial). Rather than rendering the supersingular isogeny path problem polynomial time, this result identifies the supersingular endomorphism ring problem (see Kohel [Koh1]) as a (potentially hard) computational problem. A solution can be used to pull back the quaternion ideal isogeny path problem to the supersingular isogeny path problem. Nevertheless, the best known solution to this problem remains exponential.

In this section we give a very brief overview of the most known cryptographic protocols based on supersingular isogeny graphs isogeny graphs. As we said above, these are based on the following hard problem

Problem 1 (Path finding problem)**.** Given two supersingular elliptic curves E_1 and E_2 in a supersingular isogeny graph $G_\ell(p)$, find a path from E_1 to E_2, namely a chain of ℓ-isogenies.

We stress the fact that this path exists because of Theorem 5.34. Further, this path is clearly not unique. If we can construct two such paths, the composition of one with the dual of the other gives an endomorphisms of E_1 and by repeating the process multiple times we end up with enough linearly independent endomorphisms to have a $\mathbb{Z}$-basis of $\mathrm{End}(E)$.

Problem 2 (Endomorphism ring computation)**.** Given a prime p and a supersingular j-invariant in $\mathbb{F}_{p^2}$, compute its endomorphism ring.

Mathematically, these two problems are equivalent under Deuring correspondence but, as already stated, its constructive version is still exponential in practice, [PL].

6.1.1 Couveignes-Rostovtsev-Stolbunov

We start by describing the (historically) first key exchange protocol based on isogeny graphs. This works in the ordinary realm but we include it here because, besides being the first one, it is very neat and, therefore, it permits one to better understand all the others.

This protocol first appeared in a preprint by Couveignes [Cou] in 1997 but was later rediscovered, turned into practice and publicized by Rostovtsev and Stolbunov [RS] in 2006. The protocol consists in constructing horizontal isogeny paths in an ordinary isogeny graph.

Definition. Given a group G acting freely on a set X and a subset $S \subseteq G$ closed under inversion and not containing the identity, the Schreier graph (S, X) is the graph whose set of vertices is X and edges are determined by the action of elements of S, i.e., there is an edge connecting x_1 and x_2 in X if and only if $s \cdot x_1 = x_2$ for some $s \in S$.

We have a well established free and transitive action of the class group of a quadratic imaginary order $\mathcal{O}$ on the set of ordinary elliptic curves with endomorphism ring isomorphic to $\mathcal{O}$, see Section 5.2.1. We can thus consider a subset S of $\mathcal{C}\ell(\mathcal{O})$ closed under inversion and construct the Schreier graph $(S, \mathrm{Ell}_k(\mathcal{O}))$. For example, if $\mathfrak{l}$ is a prime above a split prime $\ell \in \mathbb{Z}$, we can construct $(\{\mathfrak{l}, \bar{\mathfrak{l}}\}, \mathrm{Ell}_k(\mathcal{O}))$ consisting of the horizontal ℓ-isogenies between the curves in $\mathrm{Ell}_k(\mathcal{O})$. If S contains a set of generators of the class group, then the graph $(S, \mathrm{Ell}_k(\mathcal{O}))$ is expander and has the right properties to construct a cryptographic protocol.

The public parameters consist of a large prime power $q = p^r$ and an elliptic curve E defined over $\mathbb{F}_q$. The Discriminant Δ_π of the Frobenius endomorphism π of E is computed and a set of primes $S = \{\ell_1, \dots, \ell_n\}$ splitting in $\mathbb{Z}[\pi]$ is chosen. Each prime determines two directions $\mathfrak{l}$ and $\bar{\mathfrak{l}}$ associated to the roots of

$$x^2 - t_\pi x + q = (x - \lambda_i)(x - \mu_i) \bmod \ell_i$$

One of the two is chosen to be the positive direction relative to ℓ_i.

At this point both parties pick a random path in the form of an ideal I_A (respectively I_B) which correspond to an isogeny ϕ_A (respectively ϕ_B). They apply it to the shared curve E and exchange the ending vertices. Finally, both reapply their secret path to the elliptic curve received from the other party. Because of the commutativity of the class group, the following diagram commutes

$$\begin{tikzcd} E \arrow[r, "\phi_A"] \arrow[d, "\phi_B"'] & E_A \arrow[d, "\phi_B"] \\ E_B \arrow[r, "\phi_A"'] & E_{AB} \end{tikzcd}$$

Therefore, they each end up computing the same elliptic curve which plays the role of shared secret.

Remark. The action of an ideal is computed by solving the explicit isogeny problem, see [Bos+], [BCL] and [Gal4].

Couveignes-Rostovtsev-Stolbunov key exchange scheme		
PUBLIC DATA:	An elliptic curve E over a large finite field k. The discriminant Δ_π of the Frobenius Endomorphism of E. A finite set $S = \{\ell_i\}_i$ of primes splitting in $\mathbb{Z}[\pi]$. A Frobenius eigenvalue λ_i for each ℓ_i.	
	ALICE	**BOB**
Choose a random path	ϕ_A	ϕ_B
Compute the image	$E_A = \phi_A(E)$	$E_B = \phi_B(E)$
Exchange data		
	E_B	E_A
Compute shared secret	$E_{BA} = \phi_A(E_B)$	$E_{AB} = \phi_B(E_A)$
SHARED SECRET: Final elliptic curve $E_{AB} = E_{BA}$		

6.1.2 SIDH

The nice structure of the Couveignes-Rostovtsev-Stolbunov protocol is due to the existence of an abelian group acting on the set of elliptic curves. However, the same action makes it vulnerable to a subexponential attack on a quantum computer under the approach of Childs, Jao and Soukharev [CJS] who adapted quantum algorithms by Regev [Reg] and Kuperberg [Kup1]. Further, the Couveignes-Rostovtsev-Stolbunov scheme is very slow even considering the speed-ups proposed by De Feo, Kieffer, and Smith [DKS]. For these reasons, the supersingular isogeny graphs seems a promising direction of investigation since they lack a commutative torsor. Supersingular isogeny graphs also have another nice feature: one prime suffices to make G_ℓ an expander graph, see Theorem 5.34.

In 2011, in response to the NIST call for applications for post-quantum cryptographic candidates, De Feo, Jao and Plût [DJP] proposed a cryptographic key-exchange protocol based on supersingular isogeny graphs which is believed to be quantum resistant. Although it might look as an attempt to mimic Couveignes-Rostovtsev-Stolbunov protocol, the resulting scheme is closer to the Hash function proposed by Charles, Goren, and Lauter [CGL] in 2006. In fact, the lack of a commutative group action on the set $SS(p)$ prevents one form obtaining a commutative diagram as before, see Section 6.1.1. The solution they proposed was to share an extra piece of information along with the data about the ending vertex. This extra structure comes in the form of the action of the secret isogeny on some torsion point. We refer to the original article [DJP] for the details and we simply highlight here the main strategy.

The setting is the following: Alice and Bob choose one small prime each, say ℓ_A and ℓ_B, and they agree on a prime p of the form $p = \ell_A^{e_A}\ell_B^{e_B} f \pm 1$. Note that in general the two small primes can be chosen to be 2 and 3 and the correction factor f to be 1. The particular shape of p permits them to pick a supersingular curve E over $\mathbb{F}_{p^2}$ with cardinality $(p \pm 1)^2$ (see Theorem 1.17) and with two subgroups $E[\ell_A^{e_A}]$ and $E[\ell_B^{e_B}]$ of coprime order. Basis $\{P_A, Q_A\}$ and $\{P_B, Q_B\}$ for $E[\ell_A^{e_A}]$ and $E[\ell_B^{e_B}]$ are fixed.

The key idea is that Alice and Bob choose random path in the ℓ_A-isogeny graph (respectively ℓ_B-isogeny graph) by picking random secret subgroups

$$\langle A\rangle = \langle [m_A]P_A + [n_A]Q_A\rangle \qquad \text{and} \qquad \langle B\rangle = \langle [m_B]P_B + [n_B]Q_B\rangle$$

of respective orders $\ell_A^{e_A}$ and $\ell_B^{e_B}$, and compute the secret isogeny paths

$$\phi_A : E \longrightarrow E_A = E/\langle A\rangle \qquad \text{and} \qquad \phi_B : E \longrightarrow E_B = E/\langle B\rangle$$

At this point they would like to share E_A and E_B and to repeat the procedure as in Couveignes-Rostovtsev-Stolbunov Protocol. As already said, however, they need to pass along the action of their secret path on the basis of the torsion group used by the other, namely Alice has to compute $\{\phi_A(P_B), \phi_A(Q_B)\}$ and transmit the information to Bob who, in turn, needs to compute $\{\phi_B(P_A), \phi_B(Q_A)\}$ and share the result. Alice can now compute

$$\phi'_A : E_B \longrightarrow E_{BA} = E_B/\langle\phi_B(A)\rangle$$

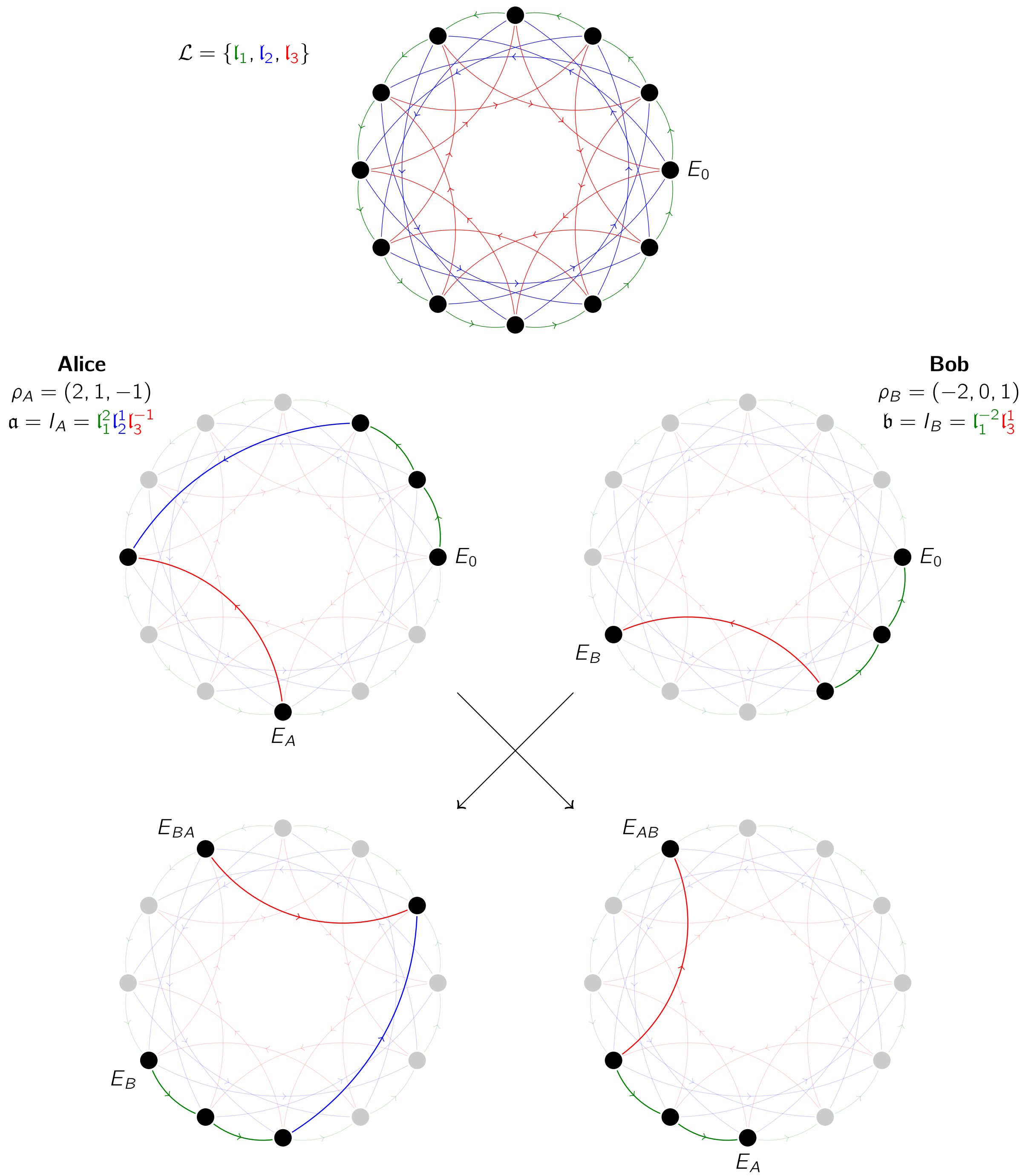

$\mathcal{L} = \{\mathfrak{l}_1, \mathfrak{l}_2, \mathfrak{l}_3\}$
E_0
Alice
$\rho_A = (2, 1, -1)$
$\mathfrak{a} = I_A = \mathfrak{l}_1^2\mathfrak{l}_2^1\mathfrak{l}_3^{-1}$
Bob
$\rho_B = (-2, 0, 1)$
$\mathfrak{b} = I_B = \mathfrak{l}_1^{-2}\mathfrak{l}_3^1$
E_0
E_A
E_0
E_B
E_{BA}
E_B
E_{AB}
E_A

Figure 6.1 – Graphic representation of the Couveignes-Rostovtsev-Stolbunov protocol.

and Bob computes

$$\phi'_B : E_A \longrightarrow E_{AB} = E_A/\langle \phi_A(B) \rangle$$

Because of $\langle A, B \rangle = \langle B, A \rangle$ we now have the following commutative diagram which guarantees that the two parties end up sharing the same elliptic curve:

$$\begin{tikzcd}
E \arrow[r, "\phi_A"] \arrow[d, "\phi_B"'] & E/\langle A \rangle \arrow[d, "\phi'_B"] \\
E/\langle B \rangle \arrow[r, "\phi'_A"'] & E/\langle A, B \rangle
\end{tikzcd}$$

Remark. The need for some extra public information is due to the fact that the lack of an abelian torsor on the set of supersingular elliptic curves translates to the loss of a canonical way of labeling the edges of the supersingular isogeny graph, i.e., finding a positive and a negative direction. However, as noted in the introduction, sharing torsion points opens the doors to the attacks of Petit [Pet], Castryck and Decru [CD], Maino and Martindale [MM] and Robert [Rob] which have wrecked the security of SIDH.

One can consult at [DeF2] for more details or [Cos] for a toy example.

SIDH protocol

PUBLIC DATA:	Two small primes ℓ_A and ℓ_B and a large prime $p = \ell_A^{e_A}\ell_B^{e_B} f \mp 1$. A supersingular elliptic curve E over $\mathbb{F}_{p^2}$ of cardinality $(p \pm 1)^2$. A basis $\{P_A, Q_A\}$ for $E[\ell_A^{e_A}]$ A basis $\{P_B, Q_B\}$ for $E[\ell_B^{e_B}]$	
	ALICE	**BOB**
Choose a random secret	$\langle A\rangle = \langle [m_A]P_A + [n_A]Q_A\rangle$	$\langle B\rangle = \langle [m_B]P_B + [n_B]Q_B\rangle$
Compute secret path	$\phi_A : E \to E_A = E/\langle A\rangle$	$\phi_B : E \to E_B = E/\langle B\rangle$
Compute the image	$\{\phi_A(P_B), \phi_A(Q_B)\}$	$\{\phi_B(P_A), \phi_B(Q_A)\}$
Exchange data		
	$E_B, \phi_B(P_A), \phi_B(Q_A)$	$E_A, \phi_A(P_B), \phi_A(Q_B)$
Compute shared secret	$\phi'_A : E_B \to E_B/\langle \phi_B(A)\rangle$	$\phi'_B : E_A \to E_A/\langle \phi_A(B)\rangle$
SHARED SECRET: Final elliptic curve $E/\langle A, B\rangle$		

6.1.3 CSIDH

In 2018, Castryck, Lange, Martindale, Panny and Renes [Cas+] proposed another approach to supersingular isogeny based cryptography. They idea is to directly adapt the Couveignes-Rostovtsev-Stolbunov protocol to supersingular curves. In order to overcome the lack of rigid structure in the supersingular realm they propose to restrict the attention to supersingular elliptic curves that are defined over a prime field $\mathbb{F}_p$ instead of the whole set $\mathrm{SS}(p)$. Further, instead of the full endomorphism ring, they only consider the subring of $\mathbb{F}_p$-rational endomorphisms, which is a quadratic imaginary order $\mathcal{O}$, see Section 5.2.2 and Theorem 5.35.

We will not discuss here the technical details and all the improvements proposed in [Cas+] but we will limit ourselves to describing the protocol. The global parameters are a large prime p of the form $p = 4 \cdot \ell_1 \cdot \ell_2 \cdot \ldots \cdot \ell_n - 1$ where the ℓ_i's are small distinct odd primes, and the supersingular elliptic curve $E_0 : y^2 = x^3 + x$ over $\mathbb{F}_p$. The subring of $\mathrm{End}(E)$ consisting of those endomorphisms which are defined over $\mathbb{F}_p$ is $\mathcal{O} = \mathbb{Z}[\pi]$. Because of the shape of p, the primes ℓ_i split in $\mathcal{O}$ as $\ell_i\mathcal{O} = \mathfrak{l}_i\bar{\mathfrak{l}}_i$ where $\mathfrak{l}_i = (\ell_i, \pi - 1)$ and $\bar{\mathfrak{l}}_i = (\ell_i, \pi + 1)$ since $\pi^2 - 1 \equiv 0 \bmod \ell_i$. The private key consist of an n-tuple $(e_1, \ldots, e_n)$ of integers in a range $[-m, \ldots, m]$ which corresponds to the ideal class $[\mathfrak{a}] = [\mathfrak{l}_1^{e_1} \cdot \ldots \cdot \mathfrak{l}_n^{e_n}]$ where $\mathfrak{l}_i = (\ell_i, \pi - 1)$. Both parties apply their secret key to the elliptic curve E_0 and obtain an elliptic curve in Montgomery form $[\mathfrak{a}]E_0 : y^2 = x^3 + Ax^2 + x$; they share the coefficient A of the final elliptic curve and re-apply their secret ideal action on the elliptic curve received from the other. The secret key is the coefficient of the Montgomery form of the shared curve obtained by acting on E_0 via the two secret ideals. This is the same for both parties since $\mathcal{C}\ell(\mathcal{O})$ is commutative.

CSIDH protocol

PUBLIC DATA:	A large prime $p = 4\ell_1 \cdot \ldots \cdot \ell_n - 1$ for small distinct odd primes ℓ_i. The supersingular elliptic curve $E_0 : y^2 = x^3 + x$ over $\mathbb{F}_p$. The endomorphism ring of E_0 over $\mathbb{F}_p$ is $\mathcal{O} = \mathbb{Z}[\pi]$. The prime ℓ_i splits in $\mathcal{O}$ as $\ell_i\mathcal{O} = \mathfrak{l}_i\bar{\mathfrak{l}}_i = (\ell_i, \pi - 1)(\ell_i, \pi + 1)$	
	ALICE	**BOB**
Key generation	$(e_1, \ldots, e_n) \in$ $[-m, \ldots, m]^n$	$(f_1, \ldots, f_n) \in [-m, \ldots, m]^n$
Compute secret path	$\mathfrak{a} = [\mathfrak{l}_1^{e_1}, \ldots, \mathfrak{l}_n^{e_n}]$	$\mathfrak{b} = [\mathfrak{l}_1^{f_1}, \ldots, \mathfrak{l}_n^{f_n}]$
Compute the image	$\phi_A : E \to [\mathfrak{a}] \cdot E = E/\mathfrak{a}$	$\phi_B : E \to [\mathfrak{b}] \cdot E = E/\mathfrak{b}$
Exchange data		
	E_B	E_A
Compute shared secret	$\phi_A : E_B \to [\mathfrak{a}] \cdot E_B = E_B/\mathfrak{a}$	$\phi_B : E_A \to [\mathfrak{b}] \cdot E_A = E_A/\mathfrak{b}$
SHARED SECRET: Final elliptic curve $[\mathfrak{a}][\mathfrak{b}] \cdot E$		

6.2 The OSIDH protocol

The choice of considering curves over $\mathbb{F}_p$ and not $\mathbb{F}_{p^2}$ in CSIDH was motivated by the need of recovering a class group action. By imposing the data of an orientation we can obtain the same goal without limiting the size of the vertex space. However, we point out once again that while SIDH and CSIDH have well established choices of parameters and have been adapted to other cryptographic purposes besides key-exchange schemes, OSIDH remains a general framework of mainly theoretical interest. The motivation behind it is multifold; on one hand it represents a direct application of the idea of orienting a supersingular elliptic curve; In fact, we recall that the main difficulty in using supersingular isogeny graphs for cryptographic schemes is the lack of a canonical way of distinguishing between its edges. Further, OSIDH provides a general way of approaching the study of supersingular isogeny graphs; as we will see it is a generalization of CSIDH which takes place at a different level of an isogeny volcano (CSIDH happen at the crater). Finally, it permits to overcome some of the restriction imposed by prior protocols.

The idea of SIDH is to fix a large prime number p of the form $p = \ell_A^{e_A}\ell_B^{e_B} f \pm 1$ for a small cofactor f and to let the two parties Alice and Bob take random walks (i.e., isogenies chains) of length e_A (or e_B) in the ℓ_A-isogeny graph (or the ℓ_B-isogeny graph, respectively) on the set of supersingular j-invariants defined over $\mathbb{F}_{p^2}$, see Section 6.1.2. In order to have the two key spaces of similar size $\ell_A^{e_A} \approx \ell_B^{e_B}$, we need to take $\ell_A^{e_A} \approx \ell_B^{e_B} \approx \sqrt{p}$. Since the total number of supersingular j-invariants is around $p/12$, this implies that, for each party, the space of choices for the secret key is limited to $1/\sqrt{p}$ of the whole set of supersingular j-invariants over $\mathbb{F}_{p^2}$. In other words, in choosing their secrets, Alice and Bob can go only "halfway" around the graph from the starting vertex j_0.

On the other hand, the peculiarity of CSIDH is that it works with curves defined over $\mathbb{F}_p$ and restricts the endomorphism rings of such curves to the commutative subring consisting of $\mathbb{F}_p$-rational endomorphisms. Starting from this setup, the scheme is an adaptation of the Couveignes and Rostovtsev-Stolbunov idea. Observe that the choice of looking at curves defined over $\mathbb{F}_p$, instead of $\mathbb{F}_{p^2}$, limits the key spaces for Alice and Bob to $\#\mathcal{Cl}(\mathbb{Z}[\sqrt{-p}])$ supersingular points. For a given p, this is the same order of magnitude, $O(\sqrt{p}\log(p))$, as for SIDH, but the class group is transitive on this subset.

The new cryptographic protocol we are about to introduce, OSIDH, is defined over an arbitrarily large subset of oriented supersingular elliptic curves over $\mathbb{F}_{p^2}$, which combines features of SIDH and CSIDH, and permits one to cover an arbitrary proportion of all isomorphism classes of supersingular elliptic curves.

Finally, a feature shared by SIDH and CSIDH is that the isogenies are constructed as quotients of rational torsion subgroups: the secret path of length e_A in the ℓ_A-isogeny graph corresponds to a secret cyclic subgroup $\langle A\rangle \subseteq E[\ell^{e_A}]$ where A is a rational $\ell_A^{e_A}$-torsion point on E. The need for rational points imposes limits on the choice of the prime p and, thus, of the finite field we work on. In contrast OSIDH relies on constructions that can be carried out only with the use of modular polynomials hence avoiding conditions on the rational torsion subgroup.

In summary, an orientation provides a class group action on lifts of an arbitrarily large subset of supersingular points. Exploiting an effective subring $\mathcal{O}$ of the full endomorphism ring we obtain an effective action by the class group of this subring on the isogeny volcano (*whirlpool*). This approach generalizes the class group action of CSIDH where supersingular elliptic curves are oriented by the commutative subring $\mathbb{Z}[\pi]$ generated by Frobenius $\pi = \sqrt{-p}$. To avoid subexponential (or polynomial) time reductions, in the OSIDH protocol, as detailed in Section 6.2.3, the orientation and associated class group action is hidden in the intermediate data exchanged by Alice and Bob. This gives a protocol for which the best known attacks at present are fully exponential.

6.2.1 Some geometric definitions

We start by introducing some geometric definitions that will help in understanding the general construction.

Instead of considering the union of different isogeny graphs as in Couveignes-Rostovtsev-Stolbunov, we focus on a fixed prime ℓ and we think of the other primes as acting on the ℓ-isogeny graph. The resulting object is the union of ℓ-isogeny volcanoes mixing under the action of $\mathcal{Cl}(\mathcal{O})$. This action stabilizes the subgraph at the surface (the craters) and preserves descending paths. This view is consistent with the construction of orientations by ℓ-isogeny chains (paths in the ℓ-isogeny graph) anchored at the surface, with action of the class group determined by ladders.

Definition. A *vortex* is defined to be an ℓ-isogeny subgraph whose vertices are isomorphism classes of $\mathcal{O}$-oriented elliptic curves with ℓ-maximal endomorphism ring, equipped with the action of $\mathcal{Cl}(\mathcal{O})$. A *whirlpool* is defined to be a complete ℓ-isogeny graph of K-oriented elliptic curves whose subgraphs of $\mathcal{O}$-oriented classes are acted on by $\mathcal{C}\ell(\mathcal{O})$. More precisely, if $\mathcal{O}$ is an ℓ-maximal order and E an $\mathcal{O}$-oriented elliptic curve, a *vortex* at ℓ is the ℓ-isogeny subgraph $G_\ell(E, \mathcal{O})$ of $G_\ell(E, K)$ equipped with the action by $\mathcal{C}\ell(\mathcal{O})$. A *whirlpool* the union of the subgraphs $G_\ell(E, \mathcal{O}_n)$ in $G_\ell(E, K)$ equipped with the compatible actions of $\mathcal{C}\ell(\mathcal{O}_n)$.

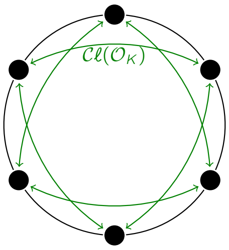


Figure 6.2 – A *vortex* consists of ℓ-isogeny cycles at the surface acted on by the class group $\mathcal{C}\ell(\mathcal{O})$ of an ℓ-maximal order $\mathcal{O}$.

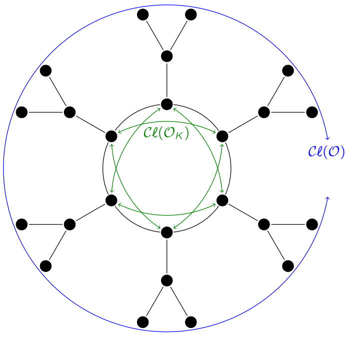

$\mathcal{C}\ell(\mathcal{O}_K)$
$\mathcal{C}\ell(\mathcal{O})$

Figure 6.3 – A *whirlpool* is an ℓ-isogeny graph equipped with compatible actions on its subgraphs by $\mathcal{C}\ell(\mathcal{O})$. The depicted 4-regular graph arises from $\ell = 3$, and the cycle length is the order of a prime over ℓ in the ℓ-maximal order.

The underlying graph of a whirlpool is composed of multiple connected components, with the class group acting transitively on components with the same ℓ-maximal order of its vortex. The existence of multiple components of ℓ-volcanoes is studied in [Mir+] and [FM], where the set of ℓ-volcanoes is called an ℓ-cordillera. A general whirlpool can be depicted as in Figure 6.4, as an ℓ-cordillera (black lines) acted on by the class group, as represented by colored arrows.

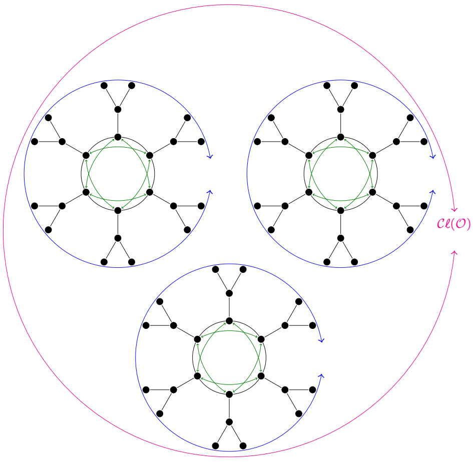


Figure 6.4 – An ℓ-isogeny graph of a whirlpool may have multiple components. The action depicts the subgraph acted on by a class group $\mathcal{C}\ell(\mathcal{O})$ of order 18, in which $\ell = 3$ has order six, such as for discriminants -1691, -2291, and -2747.

Whirlpool examples. We give examples of both ordinary and supersingular whirlpool structures of ℓ-isogeny graphs with induced class group actions.

Let $E/\mathbb{F}_{353}$ be an ordinary elliptic curve with 344 rational points, and consider the subgraph of $\Gamma_2(E)$ of curves defined over $\mathbb{F}_{353}$. The ring $\mathbb{Z}[\pi]$ generated by Frobenius π has index 2 in the maximal order $\mathcal{O}_K \simeq \mathbb{Z}[\sqrt{-82}]$ of class number 4. The set of j-invariants of such curves at the surface is $\{160, 230, 270, 298\}$, and the j-invariants of curves at depth 1 are $\{66, 182, 197, 236, 253, 264, 304, 330\}$.

This graph, depicted in Figure 6.5, consists of two 2-volcanoes, and hence the whirlpool consists of two components permuted by the transitive action of $\mathcal{C}\ell(\mathbb{Z}[\pi])$.

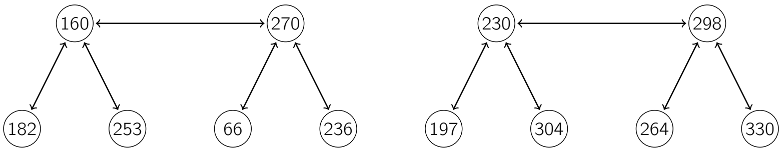


Figure 6.5 – A 2-cordillera.

Figure 6.6 represents the whirlpool, with blue lines indicating the 7-isogenies and red lines corresponding to the 13-isogenies.

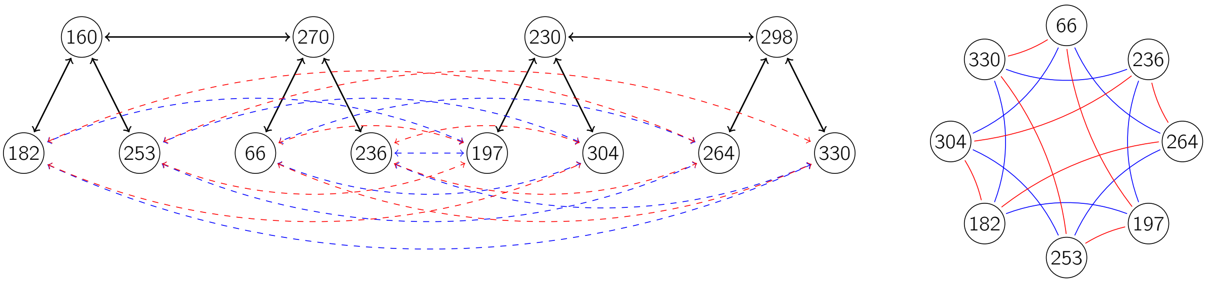


Figure 6.6 – A whirlpool with two components.

In the remaining part of this section we describe the local structure of oriented isogeny graphs and their associated class group actions. Given an order $\mathcal{O}$ in an imaginary quadratic field K, for simplicity of notation we write $\mathcal{O}(M)$ for the order $\mathbb{Z} + M\mathcal{O}$ of index M, and for fixed prime ℓ we write $\mathcal{O}_n$ for $\mathcal{O}(\ell^n)$. Moreover we denote the kernel

$$U(\mathcal{O}, M) = \ker(\mathcal{C}\ell(\mathcal{O}(M)) \to \mathcal{C}\ell(\mathcal{O}))$$

which will be identified with the stabilizer subgroup of an isomorphism class of a curve oriented by $\mathcal{O}$.

Remark. We recall that

$$1 \longrightarrow \frac{\mathcal{O}_K/M\mathcal{O}_K}{\mathcal{O}_K^\times(\mathbb{Z}/M\mathbb{Z})^\times} \longrightarrow \mathcal{C}\ell\left(\mathcal{O}(M)\right) \longrightarrow \mathcal{C}\ell(\mathcal{O}_K) \longrightarrow 1$$

In order to discuss the local neighborhood of a graph, and conduct a breadth-first searches, we introduce the notion of an ℓ-isogeny cloud around E.

Definition. An *ℓ-isogeny cloud of radius r* at E is a subgraph of $G_\ell(E)$, whose paths from E extend to length r.

Suppose that ℓ is a prime, $\mathcal{O}$ an ℓ-maximal order in K, and E an $\mathcal{O}$-oriented elliptic curve. Set $\mathcal{O}_r = \mathcal{O}(\ell^r)$ for all positive integers r. The subgraph of descending ℓ-isogenies in the ℓ-isogeny cloud of radius r at E in $G_\ell(E, K)$ admits an action of $U(\mathcal{O}, \ell^r)$. In view of the exact sequence of class groups,

$$1 \to U(\mathcal{O}, \ell^r) \to \mathcal{C}\ell(\mathcal{O}_r) \to \mathcal{C}\ell(\mathcal{O}) \to 1,$$

we consider the corresponding decomposition into subgraphs.

Definition. As before, let $\mathcal{O}$ be an ℓ-maximal order and E an $\mathcal{O}$-oriented elliptic curve. An *eddy* at E is the subgroup of ℓ-isogenies descending from E, equipped with the compatible actions of $U(\mathcal{O}, \ell^n)$. The restriction of the whirlpool (or eddy) to $G_\ell(E, \mathcal{O}_r)$ is called the whirlpool (or eddy) of depth r.

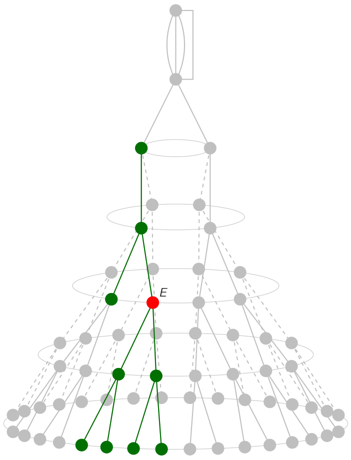


Figure 6.7 – A cloud in an oriented isogeny volcano.

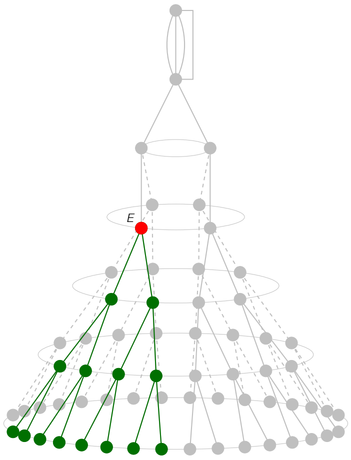


Figure 6.8 – An eddie in an oriented isogeny volcano.

Even at a non-split prime ℓ, in which every ℓ-isogeny is descending, the distinction between cloud and eddy, a set and G-set, respectively, is nontrivial. The structure of a cloud can be constructed from ℓ-isogeny relations, on which the G-set structure can be enumerated, but an effective G-set structure on the eddy is entirely determined by its transitive action on one descending path. An effective action of $U(\mathcal{O}, \ell^r)$ yields a compression from an ℓ-cloud of order ℓ^r elements to a ℓ-isogeny chain of length r.

6.2.2 A first naive protocol

We first describe a simplified version of OSIDH as intermediate step. The reason for doing that is twofold. On one hand it permits us to observe how the notions introduced so far lead to a cryptographic protocol, and on the other hand it highlights the critical security considerations and identifies the computationally hard problems on which the security is based.

As described at the beginning of the section, we fix a maximal order $\mathcal{O}_K$ in a quadratic imaginary field K of small discriminant Δ_K and a large prime p such that $\left(\frac{\Delta_K}{p}\right) \neq 1$. Further, the two parties agree on an elliptic curve E_0 with effective maximal order $\mathcal{O}_K$ embedded in the endomorphism ring and a descending ℓ-isogeny chain:

$$E_0 \longrightarrow E_1 \longrightarrow E_2 \longrightarrow \cdots \longrightarrow E_n.$$

Each constructs a power smooth horizontal endomorphism ψ of E_0 as the product of generators of small principal ideals in $\mathcal{O}_K$. A power smooth isogeny, for which the prime divisors and exponents of its degree are bounded, ensures that ψ can be efficiently extended to a ladder.

Remark. In practice, we will fix $\mathcal{O}_K$ to be either the Eisenstein integers $\mathbb{Z}[\zeta_3]$ or the Gaussian integers $\mathbb{Z}[\zeta_4](=\mathbb{Z}[i])$. Since the ladder is descending, we have that $\mathrm{End}((E_i, \iota_i)) \simeq \mathbb{Z} + \ell^i\mathcal{O}_K$ for all $i = 0, \dots, n$.

Alice privately chooses a horizontal power smooth endomorphism $\psi_A = \psi_0 : E_0 \to F_0 = E_0$, and pushes it forward to an ℓ-ladder of length n:

$$\begin{tikzcd}
E_0 \arrow[r, "\phi_0"] \arrow[d, "\psi_A"'] & E_1 \arrow[r, "\phi_1"] \arrow[d] & E_2 \arrow[r, "\phi_2"] \arrow[d] & \cdots \arrow[r, "\phi_{n-1}"] & E_n \arrow[d] \\
F_0 \arrow[r, "\phi'_0"'] & F_1 \arrow[r, "\phi'_1"'] & F_2 \arrow[r, "\phi'_2"'] & \cdots \arrow[r, "\phi'_{n-1}"'] & F_n
\end{tikzcd}$$

By Lemma 5.40, this ℓ-ladder is level, hence $\mathrm{End}((E_i, \iota_i)) = \mathrm{End}((F_i, \iota'_i))$.

The ℓ-isogeny chain (F_i) is sent to Bob, who chooses a horizontal smooth endomorphism ψ_B, and sends the resulting ℓ-isogeny chain (G_i) to Alice. Each applies (and, eventually, push forward) the private endomorphism to obtain $(H_i) = \psi_B \cdot (F_i) = \psi_A \cdot (G_i)$, and $H = H_n$ is the shared secret.

In the following picture the blue arrows correspond to the orientation chosen throughout by Alice while the red ones represent the choice made by Bob.

$$\begin{tikzcd}
 & H_0 \arrow[rr] & & H_1 \arrow[rr] & & H_2 \arrow[rr] & & \cdots \arrow[rr] & & H_n \\
F_0 \arrow[ur, red] \arrow[rr, blue] & & F_1 \arrow[ur, red] \arrow[rr, blue] & & F_2 \arrow[ur, red] \arrow[rr, blue] & & \cdots \arrow[rr, blue] & & F_n \arrow[ur, red] & \\
 & G_0 \arrow[uu, blue] \arrow[rr, red] & & G_1 \arrow[uu, blue] \arrow[rr, red] & & G_2 \arrow[uu, blue] \arrow[rr, red] & & \cdots \arrow[rr, red] & & G_n \arrow[uu, blue] \\
E_0 \arrow[uu, blue] \arrow[ur, red] \arrow[rr] & & E_1 \arrow[uu, blue] \arrow[ur, red] \arrow[rr] & & E_2 \arrow[uu, blue] \arrow[ur, red] \arrow[rr] & & \cdots \arrow[rr] & & E_n \arrow[uu, blue] \arrow[ur, red] &
\end{tikzcd}$$

Naive OSIDH protocol

PUBLIC DATA:	A prime p and a supersingular elliptic curve E_0 over $\mathbb{F}_{p^2}$. An order $\mathcal{O}$ of class number 1 orienting E_0 A descending ℓ-isogeny chain $E_0 \to E_1 \to \cdots \to E_n$	
	ALICE	**BOB**
Choose a smooth endomorphism of E_0 in $\mathcal{O}_K$	$\begin{tikzcd} E_0 \arrow[loop left, blue] \arrow[dr, equal] & \\ & F_0 \end{tikzcd}$	$\begin{tikzcd} E_0 \arrow[loop left, red] \arrow[dr, equal] & \\ & G_0 \end{tikzcd}$
Push it forward to depth n	$\underbrace{F_0 \to F_1 \to \cdots \to F_n}_{\psi_A}$	$\underbrace{G_0 \to G_1 \to \cdots \to G_n}_{\psi_B}$
Exchange data	(G_i)	(F_i)
Compute shared secret	Compute $\psi_A \cdot (G_i)$	Compute $\psi_B \cdot (F_i)$
SHARED SECRET: Final elliptic curve H_n of the shared chain $E_0 \to H_1 \to \cdots \to H_n$		

This naive protocol reveals too much information and is susceptible to attack by computing the endomorphism rings of the end curves $\mathrm{End}(E_n)$, $\mathrm{End}(F_n)$, and $\mathrm{End}(G_n)$. In general, the problem of computing an isogeny between two supersingular elliptic curves E and F knowing $\mathrm{End}(E)$ is broadly equivalent to the task of computing $\mathrm{End}(F)$ [Gal+; Eis+]. Kohel's algorithm [Koh1], and the refinement of Galbraith [Gal2], compute several paths in the isogeny graph to find isogenies $F \to F$. Thus, as noted in [Gal+], computing $\mathrm{End}(F)$ can be reduced to finding an endomorphism $\phi : F \to F$ that is not in $\mathbb{Z}[\pi]$.

Remark. Observe that in SIDH and CSIDH the endomorphism ring of the starting elliptic curve is known since the shared initial curve is chosen to have special form. In OSIDH the situation changes: we need to find an isogeny starting from E_n, and not the curve E_0 for which we have an explicit description of the endomorphism ring. However, knowing $\mathrm{End}(E_0)$, we can deduce at each step

$$\mathbb{Z} + \ell\mathrm{End}(E_i) \simeq \mathbb{Z} + \phi_i \mathrm{End}(E_i)\hat{\phi}_i \subset \mathrm{End}(E_{i+1})$$

and thus we obtain the inclusion $\mathbb{Z} + \ell^n\mathrm{End}(E_0) \hookrightarrow \mathrm{End}(E_n)$.

Note that, in general, knowing the existence of a copy of an imaginary quadratic order inside the maximal order of a quaternion algebra does not guarantee the knowledge of the embedding as there might be many [Eic2, p. II.5]. In this case, from the knowledge of a subring $\mathbb{Z}+\ell\mathrm{End}(E_i)$ of finite index ℓ^3 we can reconstruct $\mathrm{End}(E_{i+1})$ step-by-step from the ℓ-isogeny chain $E_0 \to E_1 \to \ldots \to E_n$, and hence compute $\mathrm{End}(E_n)$.

In the naive protocol we also share the full isogeny chain (F_i) (or their j-invariant sequence), which allows an adversary to deduce the oriented endomorphism ring

$$\mathbb{Z} + \ell^n \mathcal{O}_K \hookrightarrow \mathrm{End}(F_n)$$

of the terminal elliptic curve $F = F_n$. This gives enough information to deduce $\mathrm{Hom}(E, F)$ and construct a representative smooth ideal in $\mathcal{C}\ell(\mathcal{O})$ sending E to F.

We observe that there is another approach to this problem which uses only properties of the ideal class group. Suppose we have a K-descending ℓ-isogeny chain $E_0 \longrightarrow E_1 \longrightarrow \ldots \longrightarrow E_n$ with

$$\mathrm{End}(E_0) \supsetneq \mathcal{O}_K = \mathcal{O}_0 \supset \mathcal{O}_1 \supset \ldots \supset \mathcal{O}_n \simeq \mathbb{Z} + \ell^n \mathcal{O}_K$$

This induces a sequence at the level of class groups

$$\mathcal{C}\ell(\mathcal{O}_n) \longrightarrow \cdots \longrightarrow \mathcal{C}\ell(\mathcal{O}_i) \longrightarrow \cdots \longrightarrow \mathcal{C}\ell(\mathcal{O}_K) \simeq \{1\}$$

In particular, there exists a surjection

$$\mathcal{C}\ell(\mathcal{O}_{i+1}) \simeq \frac{\left(\mathcal{O}_K/\ell^{i+1}\mathcal{O}_K\right)^\times}{\overline{\mathcal{O}}_K^\times \left(\mathbb{Z}/\ell^{i+1}\mathbb{Z}\right)^\times} \longrightarrow \frac{\left(\mathcal{O}_K/\ell^{i}\mathcal{O}_K\right)^\times}{\overline{\mathcal{O}}_K^\times \left(\mathbb{Z}/\ell^{i}\mathbb{Z}\right)^\times} \simeq \mathcal{C}\ell(\mathcal{O}_i)$$

whose kernel is easily described. First, the map $\psi : \mathcal{C}\ell(\mathcal{O}_1) \to \mathcal{C}\ell(\mathcal{O}_K)$ has kernel

$$\begin{cases} \mathbb{F}_{\ell^2}^\times / \mathbb{F}_\ell^\times & \text{of order } \ell + 1 \quad \text{if } \ell \text{ is inert} \\ \left(\mathbb{F}_\ell^\times \times \mathbb{F}_\ell^\times\right) / \mathbb{F}_\ell^\times & \text{of order } \ell - 1 \quad \text{if } \ell \text{ splits} \\ \left(\mathbb{F}_\ell[\xi]\right)^\times / \mathbb{F}_\ell^\times & \text{of order } \ell \quad\quad\;\; \text{if } \ell \text{ is ramified} \end{cases}$$

where $\xi^2 = 0$ (see [Cox, §7.D] and [Neu, §12]). Thereafter, for each $i > 1$, the surjection $\mathcal{C}\ell(\mathcal{O}_{i+1}) \to \mathcal{C}\ell(\mathcal{O}_i)$ has cyclic kernel of order ℓ by virtue of the class number formula, and hence we have a short exact sequence

$$1 \to \mathbb{Z}/\ell\mathbb{Z} \to \mathcal{C}\ell(\mathcal{O}_{i+1}) \to \mathcal{C}\ell(\mathcal{O}_i) \to 1$$

If we have already constructed some representative for ψ_A modulo $\ell^i \mathcal{O}_K$, we can lift it to find ψ_A mod $\ell^{i+1}\mathcal{O}_K$ from ℓ possible preimages. For each candidate lift ψ_A mod $\ell^{i+1}\mathcal{O}_K$, we search for an smooth representative

$$\psi_A \equiv \psi_1^{e_1} \psi_2^{e_2} \cdot \ldots \cdot \psi_t^{e_t} \bmod \ell^{i+1}\mathcal{O}_K$$

with $\deg(\psi_j) = q_j$ small. The candidate smooth lift can be applied to E_{i+1} and the correct lift is that which sends E_{i+1} to F_{i+1} in the ℓ-isogeny chain (see Figure 6.9). This yields an algorithm involving multiple instances of the discrete logarithm problem in a group of order ℓ as in Pohlig-Hellman algorithm [PH] and in the generalization of Teske [Tes].

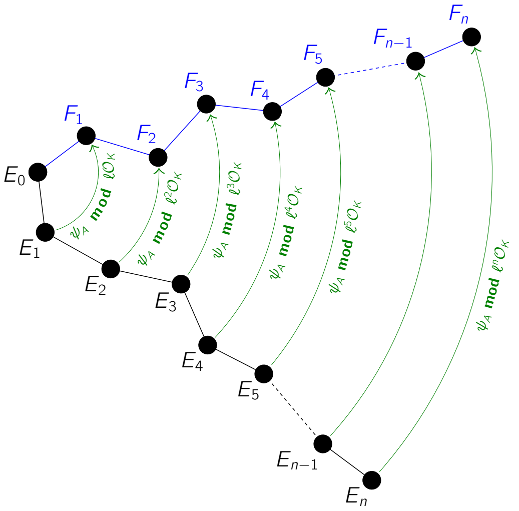

F_n
F_{n-1}
F_5
F_3
F_4
F_1
F_2
E_0
E_1
E_2
E_3
E_4
E_5
E_{n-1}
E_n
ψ_A mod $\ell\mathcal{O}_K$
ψ_A mod $\ell^2\mathcal{O}_K$
ψ_A mod $\ell^3\mathcal{O}_K$
ψ_A mod $\ell^4\mathcal{O}_K$
ψ_A mod $\ell^5\mathcal{O}_K$
ψ_A mod $\ell^n\mathcal{O}_K$

Figure 6.9 – Reconstruction of Alice's secret key

We sketch here a constructive attack to this naïve protocol, see also [DD1, §3].

Algorithm 8. Attack to the naïve OSIDH protocol

Input: Two oriented isogeny chains $(E_k, \iota_k)_{0\le k\le n}$ and $(F_k, \iota'_k)_{0\le k\le n}$
Output: An ideal class $[\mathfrak{a}] \in \mathcal{C}\ell(\mathcal{O}_n)$ such that $\mathfrak{a}\cdot(E_k, \iota_k)_{0\le k\le n} = (F_k, \iota'_k)_{0\le k\le n}$

- Compute a basis $\{\mathfrak{q}_1, \dots, \mathfrak{q}_{h(\mathcal{O}_n)}\}$ for $\mathcal{C}\ell(\mathcal{O}_n)$, see [DD1, Alg. 1].
- At the first index i_0 where the chains diverge find an ideal $\mathfrak{a}_{i_0}$ such that $\mathfrak{a}_{i_0}\cdot(E_k, \iota_k)_{0\le k\le i_0} = (F_k, \iota'_k)_{0\le k\le i_0}$, i.e., an ideal $\mathfrak{a}_{i_0}$ such that $\mathfrak{a}_{i_0}\cdot(E_{i_0}, \iota_{i_0}) = (F_{i_0}, \iota'_{i_0})$ such that $a_{i_0}\cap\mathcal{O}_{i_0-1} \in P(\mathcal{O}_{i_0-1})$ is principal.
- Express $\mathfrak{a}_{i_0}$ in terms of the basis of $\mathcal{C}\ell(\mathcal{O}_n)$: $a_{i_0} = \displaystyle\prod_{j=1}^{h(\mathcal{O}_n)} \mathfrak{q}_j^{e_{i_0,j}}$
- For $i \in [i_0, \dots, n-1]$ do
  - Compute the kernel of the reduction surjective map $\mathcal{C}\ell(\mathcal{O}_{i+1}) \to \mathcal{C}\ell(\mathcal{O}_i)$, see [DD1, Alg. 2]. The ideal $\mathfrak{a}_{i+1}$ such that $\mathfrak{a}_{i+1}\cdot(E_k, \iota_k)_{0\le k\le i+1} = (F_k, \iota'_k)_{0\le k\le i+1}$ will be written as

    $$a_{i+1} = \prod_{j=1}^{h(\mathcal{O}_n)} \mathfrak{q}_j^{e_{i+1,j}}$$

    since $[a_{i+1}\cap\mathcal{O}_i] = [a\cap\mathcal{O}_i] = [a_i\cap\mathcal{O}_i]$, then $\mathfrak{a}_{i+1} = \mathfrak{a}_i\cdot\mathfrak{b}_i$ for some $\mathfrak{b}_i = \prod\mathfrak{q}_j^{f_j}$ such that $\mathfrak{b}_i\cap\mathcal{O}_i$ is principal. Further $e_{i+1,j} = e_{i,j} + f_j$.
  - Look for $\mathfrak{b}_i$ in the finite kernel of $\mathcal{C}\ell(\mathcal{O}_{i+1}) \to \mathcal{C}\ell(\mathcal{O}_i)$ described above.
  - Reduce the exponents of $\mathfrak{a}_i\cdot\mathfrak{b}_i$ by reduction algorithms in the relation lattice, see [DD1, §3.3].

An example

We present here a toy example with only one prime involved. We fix $q = p^2 = 10007^2$. Observe that $p \equiv 2 \mod 3$ which tells us that the elliptic curve $E_0 : y^2 = x^3 + 1$ of j-invariant 0 is supersingular. We consider the embedding $\mathbb{Z}[\omega] \hookrightarrow \mathrm{End}(E_0)$ where $\omega^2 + \omega + 1 = 0$. We refer to the following picture:

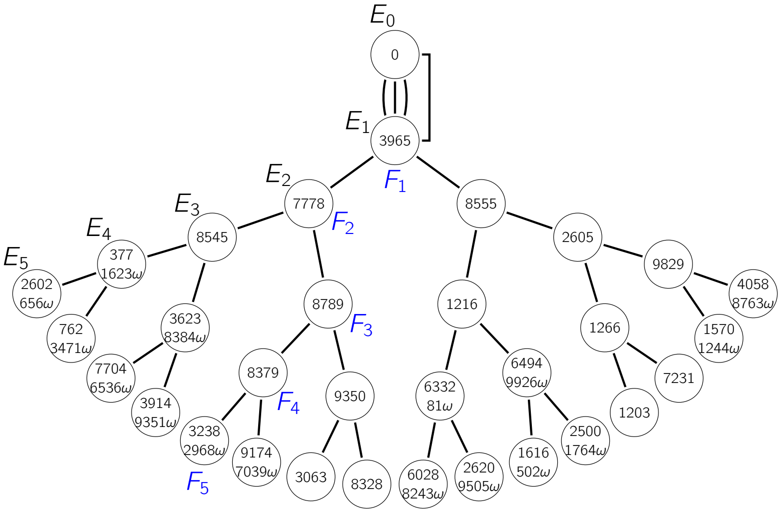


Figure 6.10 – Key reconstruction for the naïve OSIDH protocol.

We observe that the first two steps after E_0 in the sequences $\{E_i\}_i$ and $\{F_i\}_i$ are equal: $E_0 = F_0$,

$E_1 = F_1$ and $E_2 = F_2$. The first instance where the two isogeny chains differ occurs at the third step. Using modular polynomials, we immediately find that there exists a 13-isogeny between E_3 and F_3.

An easy computation shows that:

$$\begin{aligned}
13\mathbb{Z}[\omega] &= \mathfrak{q}_0\bar{\mathfrak{q}}_0 &&\text{where}\quad \mathfrak{q}_0 = (13, \omega+4) \text{ and } \bar{\mathfrak{q}}_0 = (13, \omega+10)\\
13\mathbb{Z}[2\omega] &= \mathfrak{q}_1\bar{\mathfrak{q}}_1 &&\text{where}\quad \mathfrak{q}_1 = (13, 2\omega+8) \text{ and } \bar{\mathfrak{q}}_1 = (13, 2\omega+7)\\
13\mathbb{Z}[4\omega] &= \mathfrak{q}_2\bar{\mathfrak{q}}_2 &&\text{where}\quad \mathfrak{q}_2 = (13, 4\omega+3) \text{ and } \bar{\mathfrak{q}}_2 = (13, 4\omega+1)\\
13\mathbb{Z}[8\omega] &= \mathfrak{q}_3\bar{\mathfrak{q}}_3 &&\text{where}\quad \mathfrak{q}_3 = (13, 8\omega+6) \text{ and } \bar{\mathfrak{q}}_3 = (13, 8\omega+2)
\end{aligned}$$

Observe that $\mathfrak{q}_3$ and $\bar{\mathfrak{q}}_3$ acts in the same way since $\mathfrak{q}_3^2 = (7+8\omega)$ is principal (see the Remark immediately before Subsection 5.1). Following the same approach described in the Remark, it is already possible to see that in the next step $\mathfrak{q}_4$ and $\bar{\mathfrak{q}}_4$ will not act in the same way; indeed, $\ell^{2i} = 16^2 > 13^2$. Equivalently, this can be see using modular polynomials since a quick computation shows that there are two horizontal 13-isogenies from E_4.

$$13\mathbb{Z}[16\omega] = \mathfrak{q}_4\bar{\mathfrak{q}}_4 \qquad \text{where} \qquad \mathfrak{q}_4 = (13, 16\omega+12) \text{ and } \bar{\mathfrak{q}}_4 = (13, 16\omega+4)$$

Using the action of the class group on the set of $\mathcal{O}_4$-oriented elliptic curves, we see that the isogeny from E_4 to F_4 corresponds to the ideal $\mathfrak{q}_4$. Methods to compute such an action can be found in[Bos+], [BCL], [DKS] and [Gal4].

In this particular situation, we could use the following strategy:

Algorithm 9. Action of an ideal $\left[(m, a+b\ell^i w)\right] \in \mathcal{C}\ell(\mathbb{Z}+\ell^i\mathcal{O}_K)$ lying over m on the set of primitive $\mathcal{O}$-oriented elliptic curves $\mathrm{SS}^{pr}_{\mathcal{O}}(p)$.

Input: The j-invariants of two elliptic curves E and E' over $\mathbb{F}_{p^2}$ known to be m-isogenous.
Output: The ideal $[\mathfrak{a}] \in \{[\mathfrak{m}], [\overline{\mathfrak{m}}]\}$ such that $[\mathfrak{a}] * j(E) = j(E')$.

```
1. Compute m-division polynomial $\psi_m(x)$.
2. Factor $\psi_m(x)$ and find the factor $f(x)$ corresponding to the desired isogeny $\phi : E \to E'$.
3. Pick a root of $f$, i.e., a $m$-torsion point $P$ lying in the kernel of $\phi$.
4. Set $m\mathcal{O} = \mathfrak{m}\overline{\mathfrak{m}} = (m, a+b\ell^i w)(m, a'+b'\ell^i w)$.
5. If $[a]\,P + [b]\cdot\left[\ell^i w\right]\,P = O_E$
       Return $\mathfrak{m}$.
   Else
       Return $\overline{\mathfrak{m}}$.
```

Remark. Points **1.** and **2.** can be replaced by Elkies algorithm [Elk2].

Computing the action of ω on oriented curves. The only non-trivial part is to compute the action of ω on points of E. In our situation, however, there is an elegant solution. Suppose that our elliptic curves lie at depth i in the 2-isogeny volcano; this implies that there is a sequence of i 2-isogenies to the surface (crater), i.e., $j = 0$. Such a sequence is the dual of the public path from E_0 to E: if the public key is

$$E_0 \xrightarrow{\phi_0} E_1 \xrightarrow{\phi_1} E_2 \xrightarrow{\phi_2} \cdots \xrightarrow{\phi_{n-1}} E_n$$

then the isogeny from $E = E_i$ to level 0 is given by $\hat{\phi}_0 \circ \hat{\phi}_1 \circ \hat{\phi}_2 \circ \cdots \circ \hat{\phi}_{i-1}$.

Now the computation becomes easier because we know exactly how ω acts on the elliptic curve $y^2 = x^3+1$ of j-invariant 0 ([Sil2, p. II.2.2.2]):

$$[\omega]\,(x, y) = (\tilde{\omega}x, y)$$

Finally, the action of $\ell^i\omega$ on E will be given by the composition

$$\underbrace{\phi_{i-1i} \circ \cdots \circ \phi_2 \circ \phi_1 \circ \phi_0}_{\substack{i \text{ 2isogenies}\\ \text{going back to } E}} \circ [\omega] \circ \underbrace{\hat{\phi}_0 \circ \hat{\phi}_1 \circ \hat{\phi}_2 \circ \cdots \circ \hat{\phi}_{i-1}}_{\substack{i \text{ 2-isogenies}\\ \text{going to the surface}}}$$

Remark. Observe that this is the exact definition of the $\mathcal{O}_i$-orientation induced on E_i along the isogeny chain $E_0 \to E_1 \to \ldots \to E_i$.

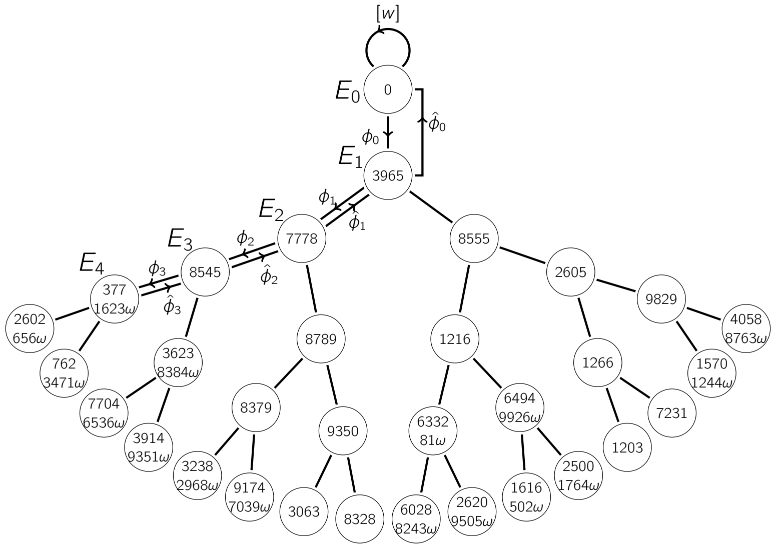


Figure 6.11 – The action of ω on oriented curves.

Remark. There is a second way of exploiting the action of $[\omega]$ on E_0. It consists in computing the generalized division polynomial of $[\omega]_{E_0}$ and then push it forward through the chain of 2-isogenies from E_0 to E. In our situation this is not too hard since the 2-division polynomial describing isogenies has limited size.

Finally, we study the situation at depth 5:

$$13\mathbb{Z}[32\omega] = \mathfrak{q}_5\bar{\mathfrak{q}}_5 \qquad \text{where} \qquad \mathfrak{q}_5 = (13, 32\omega + 11) \text{ and } \bar{\mathfrak{q}}_5 = (13, 32\omega + 8)$$

Modular polynomials show that there is no 13-isogeny between j_{E_5} and j_{F_5} but we can construct a 13-isogeny chain:

$$\begin{tikzcd} 2602+656\omega \arrow[r] & 9174+7039\omega \arrow[r] & 7704+6536\omega \arrow[r] & 8328 \arrow[r] & 762+3471\omega \arrow[r] & 3238+2968\omega \end{tikzcd}$$

Having found that $[\mathfrak{q}]^5 \cdot E_5 = F_5$ we study the order of $\mathfrak{q}_i$ in $\mathcal{C}\ell(\mathcal{O}_i)$; clearly $[\mathfrak{p}_0] = [1]$ and $[\mathfrak{q}_1] = [1]$, i.e., $\mathfrak{q}_0$ and $\mathfrak{q}_1$ are principal, since the class number of $\mathcal{O}_K$ and $\mathcal{O}_1$ is one: $h(\mathcal{O}_K) = h(\mathcal{O}_1) = 1$. In particular,

$$\mathfrak{q}_0 = (13, \omega + 4) = (\omega + 4) \text{ since } (\omega + 4)(3 - \omega) = 13$$

$$\mathfrak{q}_1 = (13, 2\omega + 8) = (-4\omega - 3) \quad \text{since} \quad \begin{cases} (-4\omega - 3)(4\omega + 1) = 13 \\ (-4\omega - 3)(2\omega) = (2\omega + 8) \end{cases}$$

We also observe that $[\mathfrak{q}_2] = [1] \in \mathcal{C}\ell(\mathcal{O}_2)$:

$$\mathfrak{q}_2 = (13, 4\omega + 3) = (4\omega + 3) \text{ since } (4\omega + 3)(-4\omega - 1) = 13$$

This is confirmed by the fact that (13) acts trivially on $\mathrm{SS}^{pr}_{\mathcal{O}_2}(10007)$ and, in fact, $E_2 = F_2$.

The situation changes when we approach the third step: the equation

$$\mathrm{N}(8\beta\omega + \alpha) = \alpha^2 - 8\alpha\beta + 64\beta^2 = 13$$

does not have integral solutions. This implies that $\mathfrak{q}_3$ and its conjugate cannot be principal. On the other hand, we have already seen that $[\mathfrak{q}_3]^2 = [1]$ in $\mathcal{C}\ell(\mathcal{O}_3)$ supported by the evidence that $\mathfrak{p}_3$ and $\bar{\mathfrak{q}}_3$ acts identically on $\mathrm{SS}^{pr}_{\mathcal{O}_3}(10007)$.

Finally, we get to the interesting part; a straightforward computation shows that $[\mathfrak{q}_4]$ has order 4 in $\mathcal{C}\ell(\mathcal{O}_4)$:

$$\mathfrak{q}_4^4 = (176\omega + 15)$$

In the same way, $[\mathfrak{q}_5]$ has order 8 in $\mathcal{C}\ell(\mathcal{O}_5)$, which is reflected in the fact that there exist a chain of 13-isogenies of length 8 from E_5 to itself (and no shorter chain).

Having seen that $[\mathfrak{q}_5]$ has order 8 in $\mathcal{C}\ell(\mathcal{O}_5)$, we conclude that the secret key chosen to obtain the sequence $\{F_i\}_{i=0}^5$ is the $\mathcal{O}_5$-ideal $\mathfrak{q}_5^5$.

$$\begin{tikzcd}
\mathcal{C}\ell(\mathcal{O}_5) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_4) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_3) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_2) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_1) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_K) \\
\cup & \cup & \cup & \cup & \cup & \cup \\
{[\mathfrak{q}_5]^5} \arrow[r] & {[\mathfrak{q}_4]^5} \arrow[r] \arrow[d, equal] & {[\mathfrak{q}_3]} \arrow[r] & {[\mathfrak{q}_2]} \arrow[r] \arrow[d, equal] & {[\mathfrak{q}_1]} \arrow[r] \arrow[d, equal] & {[\mathfrak{q}_0]} \arrow[d, equal] \\
 & {[\mathfrak{q}_4]} & & {[1]} & {[1]} & {[1]}
\end{tikzcd}$$

Note that we could also work with $\bar{\mathfrak{p}}$:

$$\begin{tikzcd}
\mathcal{C}\ell(\mathcal{O}_5) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_4) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_3) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_2) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_1) \arrow[r] & \mathcal{C}\ell(\mathcal{O}_K) \\
\cup & \cup & \cup & \cup & \cup & \cup \\
{[\bar{\mathfrak{q}}_5]^3} \arrow[r] & {[\bar{\mathfrak{q}}_4]^3} \arrow[r] & {[\bar{\mathfrak{q}}_3]^3} \arrow[r] \arrow[d, equal] & {[\bar{\mathfrak{q}}_2]} \arrow[r] \arrow[d, equal] & {[\bar{\mathfrak{q}}_1]} \arrow[r] \arrow[d, equal] & {[\bar{\mathfrak{q}}_0]} \arrow[d, equal] \\
 & & {[\bar{\mathfrak{q}}_3]} & {[1]} & {[1]} & {[1]}
\end{tikzcd}$$

In conclusion, this naïve protocol is insecure because two parties share the knowledge of the entire chains (F_i) and (G_i). The question becomes: how can we avoid sharing the ℓ-isogeny chains while still giving the other party enough information to carry out their isogeny walk?

6.2.3 A more secure version

We now detail how to send enough public data to compute the isogenies ψ_A and ψ_B on $G = G_n$ and $F = F_n$, respectively, without revealing the ℓ-isogeny chains (F_i) and (G_i). The setup remains the same with a public choice of $\mathcal{O}_K$-oriented elliptic curve E_0 and ℓ-isogeny chain

$$E_0 \to E_1 \to \cdots \to E_n.$$

Moreover, a set of primes $\mathfrak{q}_1, \ldots, \mathfrak{q}_t$ (above $q_1, \ldots, q_t$) splitting in $\mathcal{O}_K$ is fixed.

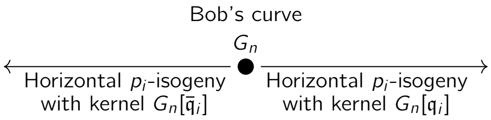
The first step consists of choosing the secret keys; these are represented by a sequence of integers $(e_1, \ldots, e_t)$ such that $|e_i| \leq r$. The bound r is taken so that the number $(2r+1)^t$ of curves that can be reached is sufficiently large. This choice of integers enables Alice to compute a new elliptic curve

$$F_n = \frac{E_n}{E_n\left[\mathfrak{q}_1^{e_1} \cdots \mathfrak{q}_t^{e_t}\right]}$$

by means of constructing the following commutative diagram

$$\begin{tikzcd}
 & E_0 \arrow[d, equal] & & E_0 \arrow[d, equal] & E_0 \arrow[d, equal] & & E_0 \arrow[d, equal] & & E_0 \arrow[d, equal] & & E_0 \arrow[d, equal] \\
E_0\bullet \arrow[r] \arrow[d] & \overset{\frac{E_0}{E_0[\mathfrak{q}_1]}}{\bullet} \arrow[r] \arrow[d] & \cdots \arrow[r] & \overset{\frac{E_0}{E_0\left[\mathfrak{q}_1^{e_1}\right]}}{\bullet} \arrow[r] \arrow[d] & \overset{\frac{E_0}{E_0\left[\mathfrak{q}_1^{e_1}\mathfrak{q}_2^{1}\right]}}{\bullet} \arrow[r] \arrow[d] & \cdots \arrow[r] & \overset{\frac{E_0}{E_0\left[\mathfrak{q}_1^{e_1}\mathfrak{q}_2^{e_2}\right]}}{\bullet} \arrow[r] \arrow[d] & \cdots \arrow[r] & \overset{\frac{E_0}{E_0\left[\mathfrak{q}_1^{e_1}\ldots\mathfrak{q}_{t-1}^{e_{t-1}}\right]}}{\bullet} \arrow[r] \arrow[d] & \cdots \arrow[r] & \overset{\frac{E_0}{E_0\left[\mathfrak{q}_1^{e_1}\ldots\mathfrak{q}_t^{e_t}\right]}}{\bullet} F_0 \arrow[d] \\
E_1\bullet \arrow[r] \arrow[d] & \bullet \arrow[r] \arrow[d] & \cdots \arrow[r] & \bullet \arrow[r] \arrow[d] & \bullet \arrow[r] \arrow[d] & \cdots \arrow[r] & \bullet \arrow[r] \arrow[d] & \cdots \arrow[r] & \bullet \arrow[r] \arrow[d] & \cdots \arrow[r] & \bullet F_1 \arrow[d] \\
\vdots \arrow[d] & \vdots \arrow[d] & & \vdots \arrow[d] & \vdots \arrow[d] & & \vdots \arrow[d] & & \vdots \arrow[d] & & \vdots \arrow[d] \\
E_n\bullet \arrow[r] & \underset{F_n^{(1)}}{\bullet} \arrow[r] & \cdots \arrow[r] & \underset{F_n^{(e_1)}}{\bullet} \arrow[r] & \underset{F_n^{(e_1,1)}}{\bullet} \arrow[r] & \cdots \arrow[r] & \underset{F_n^{(e_1,e_2)}}{\bullet} \arrow[r] & \cdots \arrow[r] & \underset{F_n^{(e_1,\ldots,e_{t-1})}}{\bullet} \arrow[r] & \cdots \arrow[r] & \underset{F_n^{(e_1,\ldots,e_t)}}{\bullet} F_n
\end{tikzcd}$$

At this point the idea is to exchange curves F_n and G_n and to apply the same process again starting from the elliptic curve received from the other party. Unfortunately, this is not enough to get to the same final elliptic curve. Once Alice receives the unoriented curve G_n computed by Bob she also needs additional information for each prime $\mathfrak{q}_i$:


Bob's curve
G_n
Horizontal p_i-isogeny with kernel $G_n[\bar{\mathfrak{q}}_i]$
Horizontal p_i-isogeny with kernel $G_n[\mathfrak{q}_i]$

but she has no information as to which directions — out of $q_i + 1$ total q_i-isogenies — to take as $\mathfrak{q}_i$ and $\bar{\mathfrak{q}}_i$. For this reason, once that they have constructed their elliptic curves F_n and G_n, they precompute, for each prime $\mathfrak{q}_i$, the q_i-isogeny chains coming from $\bar{\mathfrak{q}}_i^j$ (denoted by the class $\mathfrak{q}_i^{-j}$) and $\mathfrak{q}_i^j$:

$$F_{n,i}^{(-r)} \leftarrow \cdots \leftarrow F_{n,i}^{(-1)} \leftarrow F_n \rightarrow F_{n,i}^{(1)} \rightarrow \cdots \rightarrow F_{n,i}^{(r-1)} \rightarrow F_{n,i}^{(r)}$$

and

$$G_{n,i}^{(-r)} \leftarrow \cdots \leftarrow G_{n,i}^{(-1)} \leftarrow G_n \rightarrow G_{n,i}^{(1)} \rightarrow \cdots \rightarrow G_{n,i}^{(r-1)} \rightarrow G_{n,i}^{(r)}$$

Now Alice obtains from Bob the curve G_n and, for each i, the horizontal q_i-isogeny chains determined by the isogenies with kernels $G_n[\mathfrak{q}_i^j]$. With this information Alice can take e_1 steps in the $\mathfrak{q}_1$-isogeny chain and push forward all the $\mathfrak{q}_i$-isogeny chains for $i > 1$.

Remark. We recall that pushing forward means constructing a ladder which transmits all the information about the commutative action of $\mathfrak{q}_i^{e_i}$ in the class group.

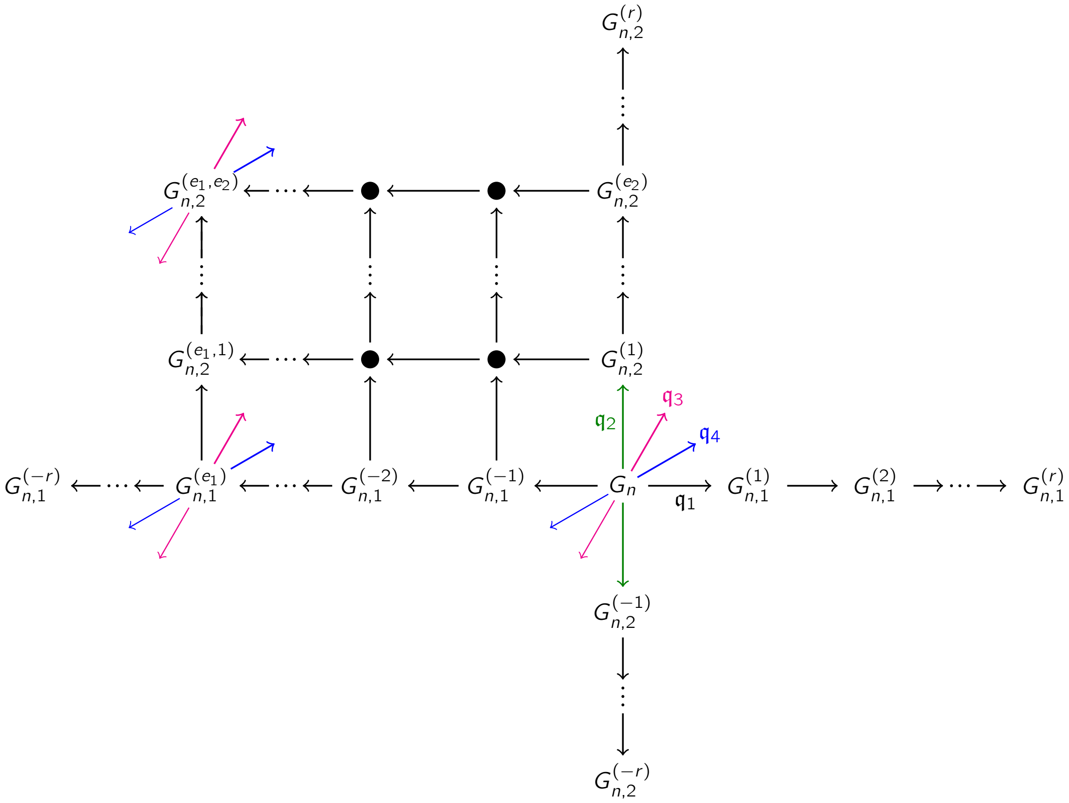


Figure 6.12 – Ladder construction in OSIDH protocol.

Alice repeats the process for all the $\mathfrak{q}_i$'s every time pushing forward the isogenies for the primes with index strictly bigger than i. Finally, she obtains a new elliptic curve

$$H_n = \frac{E_n}{E_n\left[\mathfrak{q}_1^{e_1+d_1} \cdots \mathfrak{q}_t^{e_t+d_t}\right]}$$

Bob follows the same process with the public data received from Alice, in order to compute the same curve H_n. Recall that, in the naive protocol, Alice and Bob compute the group action on the full ℓ-isogeny chains:

$$\begin{tikzcd}
E_0 \arrow[r] & E_1 \arrow[r] & E_2 \arrow[r] \arrow[d, "\text{Alice}"] & \cdots \arrow[r] & E_n \arrow[r, "\text{Bob}"] & E_0 \arrow[r] & G_1 \arrow[r] & G_2 \arrow[r] \arrow[d, "\text{Alice}"] & \cdots \arrow[r] & G_n \\
E_0 \arrow[r] & F_1 \arrow[r] & F_2 \arrow[r] & \cdots \arrow[r] & F_n \arrow[r, "\text{Bob}"] & E_0 \arrow[r] & H_1 \arrow[r] & H_2 \arrow[r] & \cdots \arrow[r] & H_n
\end{tikzcd}$$

In the refined OSIDH protocol, Alice and Bob share sufficient information to determine the curve H_n without knowledge of the other party's ℓ-isogeny chain (G_i) and (F_i), nor the full ℓ-isogeny chain (H_i) from the base curve E_0.

OSIDH protocol

PUBLIC DATA:	A prime p and a supersingular elliptic curve E_0 over $\mathbb{F}_{p^2}$. An order $\mathcal{O}$ of class number 1 orienting E_0 A descending ℓ-isogeny chain $E_0 \to E_1 \to \cdots \to E_n$ A set of splitting primes $\mathfrak{q}_1, \dots, \mathfrak{q}_t \subseteq \mathcal{O} = \operatorname{End}(E_n) \cap K \hookrightarrow \mathcal{O}_K$	
	ALICE	**BOB**
Choose integers in an interval $[-r, r]$	$(e_1, \dots, e_t)$	$(d_1, \dots, d_t)$
Construct an isogenous curve	$F_n = \dfrac{E_n}{E_n\left[\mathfrak{q}_1^{e_1} \cdots \mathfrak{q}_t^{e_t}\right]}$	$G_n = \dfrac{E_n}{E_n\left[\mathfrak{q}_1^{d_1} \cdots \mathfrak{q}_t^{d_t}\right]}$
Precompute all directions $\forall\, i$	$F_n \to F_{n,i}^{(1)} \to \cdots \to F_{n,i}^{(r)}$	$G_n \to G_{n,i}^{(1)} \to \cdots \to G_{n,i}^{(r)}$
... and their conjugates	$\underbrace{F_{n,i}^{(-r)} \leftarrow \cdots \leftarrow F_{n,i}^{(-1)} \leftarrow F_n}$	$\underbrace{G_{n,i}^{(-r)} \leftarrow \cdots \leftarrow G_{n,i}^{(-1)} \leftarrow G_n}$
Exchange data	G_n+directions	F_n+directions
Compute shared data	Takes e_i steps in $\mathfrak{q}_i$-isogeny chain & push forward information for all $j > i$.	Takes d_i steps in $\mathfrak{q}_i$-isogeny chain & push forward information for all $j > i$.

SHARED SECRET: Final elliptic curve H_n

$$H_n = \frac{F_n}{F_n\left[\mathfrak{q}_1^{d_1} \cdots \mathfrak{q}_t^{d_t}\right]} = \frac{G_n}{G_n\left[\mathfrak{q}_1^{e_1} \cdots \mathfrak{q}_t^{e_t}\right]} = \frac{E_n}{E_n\left[\mathfrak{q}_1^{e_1+d_1} \cdots \mathfrak{q}_t^{e_t+d_t}\right]}.$$

Remark. We can read this scheme using the terminology of Section 6.2.1. After the choice of the secret key, we observe a vortex: Alice (respectively Bob) acts on an isogeny crater (that in the case of $\mathcal{O}_K = \mathbb{Z}[\zeta_3]$ or $\mathbb{Z}[i]$ consists of a single point) with the primes $\mathfrak{q}_1^{e_1} \cdot \ldots \cdot \mathfrak{q}_t^{e_t}$ (respectively $\mathfrak{q}_1^{d_1} \cdot \ldots \cdot \mathfrak{q}_t^{d_t}$).

This action is eventually transmitted along the ℓ-isogeny chain and we get a whirlpool. We can think of the isogeny volcano as rotating under the action of the secret keys and the initial ℓ-isogeny path transforming into the two secret isogeny chains.

6.2.4 Security considerations

Let $\mathcal{O}$ be an order of the form $\mathcal{O} = \mathcal{O}_n(M) = \mathcal{O}_K(\ell^n M) = \mathbb{Z} + M\mathcal{O}_n$. In order to ensure security of the system, we have seen that the data giving the orientation must remain hidden. A second consideration is the proportion of curves attained by the action of the class group $\mathcal{Cl}(\mathcal{O})$, and by the private walks ψ_A and ψ_B of Alice and Bob in that class group. The size of the orbit of $\mathcal{Cl}(\mathcal{O})$ is controlled by the chain length n and conductor M, and the number of curves attained by the private walks is further limited by the prime power data, up to exponent bounds, which we allow ourselves to transmit.

We note that $\mathcal{Cl}(\mathcal{O}) \simeq \mathrm{SS}_{\mathcal{O}}^{pr}(\rho)$ and define I the exponents space $I_1 \times \ldots \times I_t \subseteq \mathbb{Z}^t$ where $I_j = [-r_j, r_j]$. The security of OSIDH depends on the injectivity and surjectivity of the following maps

$$I = \prod_{i=1}^{t} [-r_i, r_i] \longrightarrow \mathrm{SS}_{\mathcal{O}}^{pr}(\rho) \longrightarrow \mathrm{SS}(p)$$

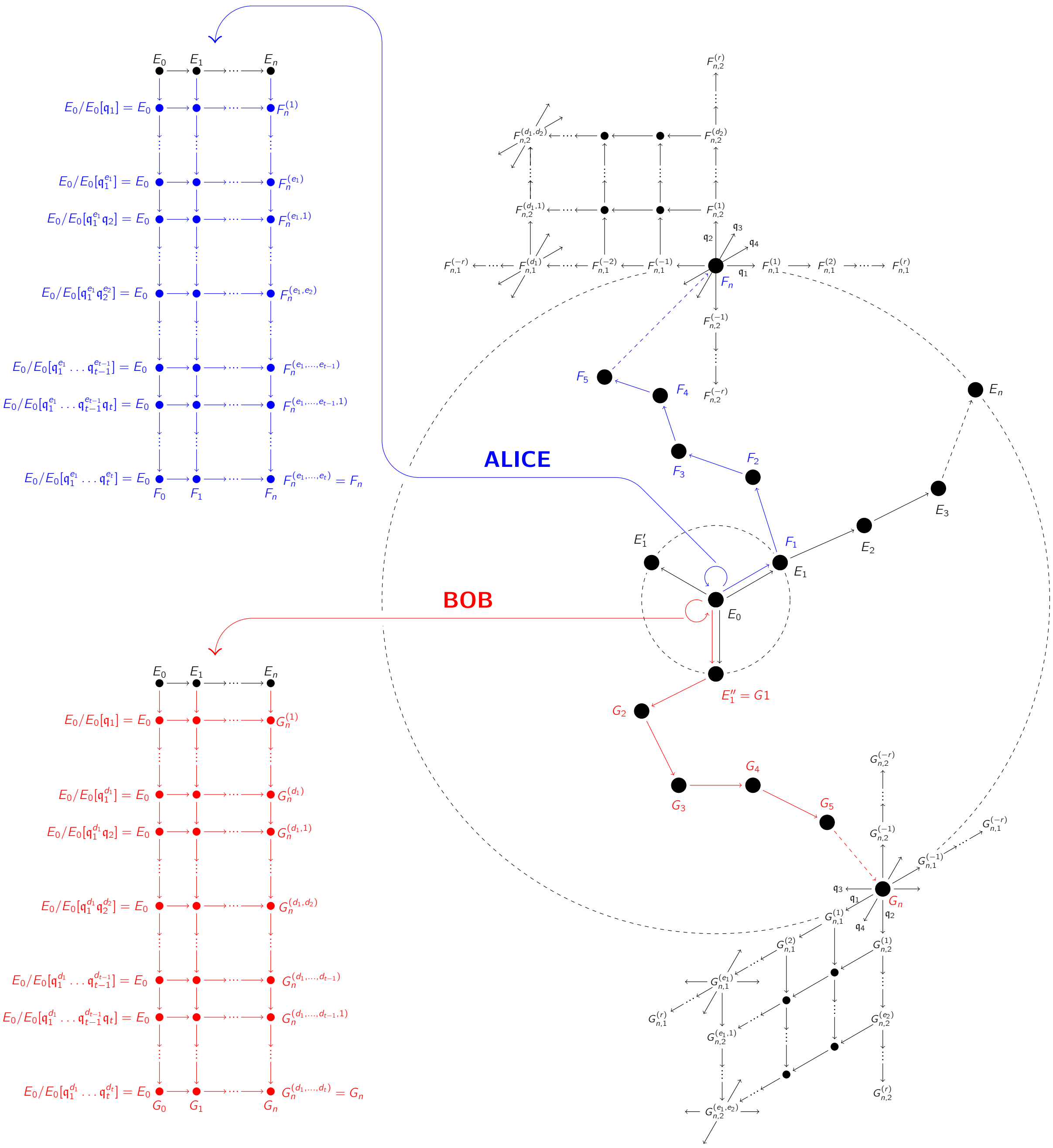

ALICE
BOB
E_0
E_1
E_2
E_3
E_n
E_1'
$E_1'' = G1$
F_1
F_2
F_3
F_4
F_5
F_n
G_2
G_3
G_4
G_5
G_n
$E_0/E_0[\mathfrak{q}_1] = E_0$
$E_0/E_0[\mathfrak{q}_1^{e_1}] = E_0$
$E_0/E_0[\mathfrak{q}_1^{e_1}\mathfrak{q}_2] = E_0$
$E_0/E_0[\mathfrak{q}_1^{e_1}\mathfrak{q}_2^{e_2}] = E_0$
$E_0/E_0[\mathfrak{q}_1^{e_1}\dots\mathfrak{q}_{t-1}^{e_{t-1}}] = E_0$
$E_0/E_0[\mathfrak{q}_1^{e_1}\dots\mathfrak{q}_{t-1}^{e_{t-1}}\mathfrak{q}_t] = E_0$
$E_0/E_0[\mathfrak{q}_1^{e_1}\dots\mathfrak{q}_t^{e_t}] = E_0$
$F_n^{(1)}$
$F_n^{(e_1)}$
$F_n^{(e_1,1)}$
$F_n^{(e_1,e_2)}$
$F_n^{(e_1,\dots,e_{t-1})}$
$F_n^{(e_1,\dots,e_{t-1},1)}$
$F_n^{(e_1,\dots,e_t)} = F_n$
F_0
F_1
F_n
$E_0/E_0[\mathfrak{q}_1^{d_1}] = E_0$
$E_0/E_0[\mathfrak{q}_1^{d_1}\mathfrak{q}_2] = E_0$
$E_0/E_0[\mathfrak{q}_1^{d_1}\mathfrak{q}_2^{d_2}] = E_0$
$E_0/E_0[\mathfrak{q}_1^{d_1}\dots\mathfrak{q}_{t-1}^{d_{t-1}}] = E_0$
$E_0/E_0[\mathfrak{q}_1^{d_1}\dots\mathfrak{q}_{t-1}^{d_{t-1}}\mathfrak{q}_t] = E_0$
$E_0/E_0[\mathfrak{q}_1^{d_1}\dots\mathfrak{q}_t^{d_t}] = E_0$
$G_n^{(1)}$
$G_n^{(d_1)}$
$G_n^{(d_1,1)}$
$G_n^{(d_1,d_2)}$
$G_n^{(d_1,\dots,d_{t-1})}$
$G_n^{(d_1,\dots,d_{t-1},1)}$
$G_n^{(d_1,\dots,d_t)} = G_n$
G_0
G_1
G_n
$F_{n,2}^{(r)}$
$F_{n,2}^{(d_2)}$
$F_{n,2}^{(1)}$
$F_{n,2}^{(d_1,d_2)}$
$F_{n,2}^{(d_1,1)}$
$F_{n,1}^{(-r)}$
$F_{n,1}^{(d_1)}$
$F_{n,1}^{(-2)}$
$F_{n,1}^{(-1)}$
$F_{n,1}^{(1)}$
$F_{n,1}^{(2)}$
$F_{n,1}^{(r)}$
$F_{n,2}^{(-1)}$
$F_{n,2}^{(-r)}$
$\mathfrak{q}_1$
$\mathfrak{q}_2$
$\mathfrak{q}_3$
$\mathfrak{q}_4$
$G_{n,2}^{(-r)}$
$G_{n,2}^{(-1)}$
$G_{n,1}^{(-r)}$
$G_{n,1}^{(-1)}$
$G_{n,1}^{(1)}$
$G_{n,1}^{(2)}$
$G_{n,2}^{(1)}$
$G_{n,1}^{(e_1)}$
$G_{n,1}^{(r)}$
$G_{n,2}^{(e_1,1)}$
$G_{n,2}^{(e_2)}$
$G_{n,2}^{(e_1,e_2)}$
$G_{n,2}^{(r)}$

Figure 6.13 – Graphic representation of OSIDH

In order to analyze the security of the protocol we state the following bounds and study their related impact. We first consider the properties of the map $\mathrm{SS}^{pr}_{\mathcal{O}}(p) \to \mathrm{SS}(p)$.

Supersingular covering bound. We deal with the problem of covering a reasonable number of elliptic curves in $\mathrm{SS}(p)$. We say that the map $\mathcal{C}\ell(\mathcal{O}) \simeq \mathrm{SS}^{pr}_{\mathcal{O}}(p) \longrightarrow \mathrm{SS}(p)$ is λ-*surjective* if $\#\mathcal{C}\ell(\mathcal{O}) \geq p^{\lambda}$ where λ is the *logarithmic covering radius*. We get

$$\lambda \log_{\ell}(p) \leq n + \log_{\ell}(M) + \log_{\ell}(h(\mathcal{O}_K)) \tag{SCB}$$

Remark. In SIDH the intermediate cloud covers $O(\sqrt{p})$ supersingular elliptic curves, so the logarithmic covering radius is $1/2$. By varying the conductor $M\ell^n$, we determine the proportion of curves covered.

Supersingular injectivity bound. How can one insure the injectivity of the map $\mathrm{SS}_{\mathcal{O}}^{pr}(p) \to \mathrm{SS}(p)$? We set

$$n + \log_\ell(M) + \frac{1}{2}\log_\ell(|\Delta_K|) \leq \frac{1}{2}\log_\ell(p) \tag{SIB}$$

Lemma 6.1. *If (SIB) holds, then the map $\mathrm{SS}_{\mathcal{O}}^{pr}(p) \to \mathrm{SS}(p)$ is injective.*

Proof. As in Proposition 5.44, the failure of the injectivity implies the existence of two embeddings of $\mathcal{O}$ in the endomorphism ring $\mathrm{End}(E)$. This implies the existence of T such that $(\Delta^2 - T^2)/4 = mp$, see Lemma 5.43. Hence $2|\Delta| \geq |\Delta| + T \geq p$ and, therefore, if injectivity fails we must have $p \leq |\Delta|$. □

Remark. Comparing bounds (SCB) and (SIB), one sees that the transition from injectivity to non-injectivity happens around the logarithmic covering radius $\lambda = \frac{1}{2}$ while surjectivity is only possible for $\lambda \geq 1$. Further, (SCB) does not guarantee surjectivity but only provides an upper bound on the number of classes covered. This incompatibility however, is not problematic as injectivity is not really necessary for security.

We will now focus on the first map $I \to \mathcal{C}\ell(\mathcal{O})$.

Minkowski norm bound. The set of elements obtained by random walks should contain no cycle; thus,

$$\sum_{i=1}^{t} r_i \log_\ell(q_i) \leq n + \log_\ell(M) + \frac{1}{2}\log_\ell(|\Delta_K|/4) \tag{MNB}$$

Lemma 6.2. *If (MNB) holds, then the map $I \to \mathrm{SS}_{\mathcal{O}}^{pr}(p)$ is injective.*

Proof. A collision in the map $I \to \mathcal{C}\ell(\mathcal{O})$ implies the relation

$$\prod_{j=1}^{t} \mathfrak{q}_j^{c_j} \sim \prod_{j=1}^{t} \mathfrak{q}_j^{d_j} \text{ hence } (\alpha) = \prod_{j=1}^{t} \mathfrak{q}_j^{s_j}$$

where $c_j, d_j \in [-r_j, r_j]$ with $s_j = c_j - d_j$, implying $|s_j| \leq 2r_j$, hence

$$\log_\ell(Nr(\alpha)) = \sum_{j=1}^{t} |s_j| \log_\ell(q_j) \geq \log_\ell(|\Delta|/4) = 2n + 2\log_\ell(M) + \log_\ell(|\Delta_K|/4) \qquad \square$$

Class group covering bound. We now consider λ-surjectivity of the map $I \to \mathrm{SS}_{\mathcal{O}}^{pr}(p)$. In order to have a uniform element of $\mathcal{C}\ell(\mathcal{O})$ it is desirable to be able to reach all elements of $\mathcal{C}\ell(\mathcal{O})$. A necessary condition for surjectivity is that the cardinality of walks ending points is at least the class number of $\mathcal{O}$. Adding a parameter λ, we say that $I \to \mathcal{C}\ell(\mathcal{O})$ is λ-surjective if $\#I \geq h(\mathcal{O})^\lambda$. Taking the logarithm, this gives

$$\sum_{i=1}^{t} \log_\ell(2r_i + 1) \geq \lambda\left(n + \log_\ell(M) + \log_\ell(h(\mathcal{O}_K))\right) \tag{CGCB}$$

Remark. In adaptations of the OSIDH protocol, see Section 6.4.1, the class group surjectivity for $\mathcal{C}\ell(\mathcal{O}_n)$ is achieved with the weaker bound

$$\sum_{i=1}^{t} \log_\ell(2r_i + 1) \geq n + \log_\ell(h(\mathcal{O}_K))$$

and surjectivity for $\mathcal{C}\ell(\mathcal{O})$ is achieved by a random walk at conductor M, giving a random element of the kernel $\mathcal{C}\ell(\mathcal{O}) \to \mathcal{C}\ell(\mathcal{O}_n)$.

6.2.5 Parameter selection

Taking into account the above bounds and their security implications, we describe optimal choices for the parameters of OSIDH: the chain length, the degree of the walks and the exponents of the private walks.

Chain length. Suppose that (E_i) is an isogeny chain of length n, from a supersingular elliptic curve E_0 oriented by $\mathcal{O}_K$ of class number one, and consider

$$\mathrm{Hom}(E_0, E_n) = \phi\mathcal{O}_K + \psi\mathcal{O}_K.$$

As a quadratic module with respect to the degree map, its determinant is p^2. If the chain is of sufficient length n such that E_n represents a general curve in $\mathrm{SS}(p)$, then a set of reduced basis elements ϕ and ψ satisfies

$$\deg(\phi) \approx \deg(\psi) \approx \sqrt{p}.$$

Now suppose that $\phi : E_0 \to E_n$ is the isogeny giving the ℓ-isogeny chain. If $\deg(\phi) = \ell^n$ is less than $\sqrt{p}$, then $\phi\mathcal{O}_K$ is a submodule generated by short isogenies, and E_n is special. We conclude that we must choose n to be at least $\log_\ell(p)/2$ in order to avoid an attack which seeks to determine $\phi\mathcal{O}_K$ as a distinguished submodule of low degree isogenies. Using the *supersingular covering bound*, we extend this argument to consider the logarithmic proportion λ of supersingular elliptic curves we can reach. In order to cover p^λ supersingular curves, out of $|\mathrm{SS}(p)| = p/12 + \varepsilon_p$, $\deg(\phi)$ must be such that

$$|\mathcal{C}\!\ell(\mathcal{O})| = \left|\frac{(\mathcal{O}_K/\ell^n\mathcal{O}_K)^*}{\mathcal{O}_K^*(\mathbb{Z}/\ell^n\mathbb{Z})^*}\right| \approx \ell^n = \deg(\phi) \approx p^\lambda.$$

In particular, choosing $\lambda = 1$, we find that $n = \log_\ell(p)$ is the critical length for reaching all supersingular curves, which is a rather mild lower bound on n.

Degree of private walks. Suppose now that $E = E_n$ is a generic supersingular curve and F another. Without an $\mathcal{O}_K$-module structure, we have a basis $\{\psi_1, \psi_2, \psi_3, \psi_4\}$ such that

$$\mathrm{Hom}(E, F) = \mathbb{Z}\psi_1 + \mathbb{Z}\psi_2 + \mathbb{Z}\psi_3 + \mathbb{Z}\psi_4.$$

Assuming that E and F are generic relative to one another, a reduced basis satisfies $\deg(\psi_i) \approx \sqrt{p}$, as above. Thus the private walk ψ_A should satisfy

$$\log_p(\deg(\psi_A)) \geq \frac{1}{2}$$

in order that $\mathbb{Z}\psi_A$ is not a distinguished submodule of $\mathrm{Hom}(E, F)$. This critical distance is the maximal that can be attained by the SIDH protocol.

As above, another measure of the generality of ψ_A is the number of curves that can be reached by different choices of the isogeny ψ_A. For a fixed degree m, the number of curves which can be attained is

$$|\mathbb{P}(E[m])| = |PP^1(\mathbb{Z}/m\mathbb{Z})| \approx m.$$

For the SIDH protocol, on has $\ell_A^{n_A} \approx \ell_B^{n_B} \approx \sqrt{p}$, and only $\sqrt{p}$ curves out of $p/12$ can be reached.

In the CSIDH or OSIDH protocols, the degree of the isogeny is not fixed. The total number of isogenies of any degree d up to m is

$$\sum_{d=1}^{m} |\mathbb{P}(E[d])| \approx m^2,$$

but the choice of ψ_A is restricted to a subset of $\mathcal{O}$-oriented isogenies in $\mathcal{C}\!\ell(\mathcal{O})$. Such isogenies are restricted to a class proportional to m. Specifically, in the OSIDH construction, if we let $S_m \subset \mathcal{O}_K$ be the set of endomorphisms of degree up to m, and consider the map

$$S_m \subset \mathcal{O}_K \longrightarrow \frac{(\mathcal{O}_K/\ell^n\mathcal{O}_K)^*}{\mathcal{O}_K^*(\mathbb{Z}/\ell^n\mathbb{Z})^*} \simeq \mathcal{C}\!\ell(\mathcal{O}).$$

Since $|S_m| \approx m$, to cover a subset of p^λ classes, we need $\log_p(\deg(\psi_A)) \geq \lambda$.

Private walk exponents. In practice, rather than bounding the degree, for efficient evaluation one fixes a subset of small split primes, and the space of exponent vectors is bounded. The instantiation CSIDH-512 (see [Cas+]) uses a prime of 512 bits such that for each of 74 primes one has a choice of 11 exponents in $[-5, 5]$. This gives 256 bits of freedom which is of the order of magnitude to cover $h(-p) \approx \sqrt{p}$ classes (up to logarithmic factors). In this instance the class number $h(-p)$ was computed in [BKV]

For the general OSIDH construction, we choose exponent vectors $(e_1, \dots, e_t)$ in the space $I_1 \times \cdots \times I_t \subset \mathbb{Z}^t$, where $I_j = [-r_j, r_j]$, defining ψ_A with kernel

$$\ker(\psi_A) = E\left[\mathfrak{q}_1^{e_1} \cdots \mathfrak{q}_t^{e_t}\right].$$

In order to avoid revealing any cycles we want to respect the *Minkowski norm bound* while, in order to cover as many classes as possible, we need the *Supersingular surjective hypothesis* to be verified. This second condition yields $p^\lambda < \prod_{j=1}^{t}(2r_j+1) < |\mathcal{C}\ell(\mathcal{O})| \approx \ell^n$. For fixed $r = r_j$, this gives

$$n > t\log_\ell(2r+1) > \lambda\log_\ell(p).$$

Setting $\lambda = 1$, $\ell = 2$ and $\log_\ell(p) = 256$, the parameters $t = 74$ and $r = 5$ give critical values as in CSIDH-512, with group action mapping to the full set of supersingular points $\mathrm{SS}(p)$.

6.3 Attacks on OSIDH

In proposing OSIDH [CK1] the authors had two main security concerns: avoiding sharing too much information and the length of the initial chain (and, by consequence, the number of attained elliptic curves). The second problem has been immediately recognized as the main key in the protocol. If, on one hand, one would like to be able to reach all possible oriented curve at a certain depth, the progressive loss of injectivity of the projection map $\mathrm{SS}^{pr}_{\mathcal{O}_n}(\rho) \to \mathrm{SS}(p)$ poses a serious issue. This subtle clash between the ideal situation and the practical solution was solved by relaxing expectations on the forgetful map above which, in the first instance of OSIDH, needed to be only effectively injective, meaning that one could expect to be computationally hard to find cycles. Indeed, it is well known that the possibility of constructing endomorphisms of an elliptic curve is equivalent to the construction of an isogeny path form said elliptic curve [Wes2].
In the specific case of OSIDH, the protocol requires Alice and Bob to share information about the $\mathfrak{q}_j$ horizontal chains on F_n and G_n. Onuki [Onu] claimed that the knowledge of such actions would give enough information to recover the secret chains $(F_i, \iota_i^{(A)})_{0\le i\le n}$ and $(G_i, \iota_i^{(B)})_{0\le i\le n}$ and therefore expose the secret keys, see Algorithm 8.

6.3.1 The attack of Onuki

The idea in [Onu, §6.3] is an adaptation of Petit's attack on SIDH [Pet] and consists in showing that the knowledge of a K-oriented endomorphism $\iota_n^{(A)}(\beta)$ (respectively $\iota_n^{(B)}(\beta)$) with $\beta \in \mathcal{O}_n \setminus \mathcal{O}_{n+1}$ permits one to recover the curve F_{n-1} (respectively G_{n-1}) and by recursion the whole chain $(F_i, \iota_i^{(A)})_{0\le i\le n}$ (respectively $(G_i, \iota_i^{(B)})_{0\le i\le n}$). We will sketch here the idea, taking the point of view of an attacker trying to recover Alice's secret key; in particular we will write ι_i instead of $\iota_i^{(A)}$.

Let us fix a generator ω for $\mathcal{O}_K = \mathbb{Z}[\omega]$. Then $\ell^n\omega$ generates $\mathcal{O}_n = \mathbb{Z}+\ell^n\mathcal{O}_K$. We then write $\beta = a + b\ell^n\omega$ for $a, b \in \mathbb{Z}$ and $(\ell, b) = 1$ since $\beta \notin \mathcal{O}_{n+1}$. Knowing $\iota_n(\beta)$ permits one to recover $\iota_n(b\ell^n\omega)$ since $\iota_n(a) = [a]$ and this is enough to construct the curve F_{n-1} thanks to [DD1, Lemma 2]. This lemma uses the diagram below, and therefore the equality $\iota_n(b\ell^n\omega) = \phi_{n-1} \circ \iota_{n-1}(b\ell^{n-1}\omega) \circ \hat{\phi}_{n-1}$ to prove that

$$\ker\left(\iota_n(b\ell^n\omega)\right) \cap F_n[\ell] = \ker\left(\hat{\phi}_{n-1}\right)$$

Therefore, evaluating $\iota_n(b\ell^n\omega)$ on $F_n[\ell]$ permits one to recover $\ker(\hat{\phi}_{n-1})$ and this is the input needed to construct $\hat{\phi}_{n-1}$ and F_{n-1} using Velu's formulæ [Vél].

$$\begin{tikzcd}
\vdots \arrow[r] & \vdots \arrow[d, "\phi_{n-3}"] \\
F_{n-2} \arrow[u, "\hat{\phi}_{n-3}"] \arrow[r, "\iota_{n-2}(b\ell^{n-2}\omega)"] & F_{n-2} \arrow[d, "\phi_{n-2}"] \\
F_{n-1} \arrow[u, "\hat{\phi}_{n-2}"] \arrow[r, "\iota_{n-1}(b\ell^{n-1}\omega)"] & F_{n-1} \arrow[d, "\phi_{n-1}"] \\
F_n \arrow[u, "\hat{\phi}_{n-1}"] \arrow[r, "\iota_n(b\ell^n\omega)"] & F_n
\end{tikzcd}$$

Note that at this point one knows F_{n-1} and the directions at F_n for all the primes $\mathfrak{q}_i$:

$$F_{n,i}^{(-r)} = [\mathfrak{q}_i]^{-r} \cdot F_n \longleftarrow \cdots \longleftarrow F_{n,i}^{(-1)} = [\mathfrak{q}_i]^{-1} \cdot F_n \longleftarrow F_n \longrightarrow F_{n,i}^{(1)} = [\mathfrak{q}_i] \cdot F_n \longrightarrow \cdots \longrightarrow F_{n,i}^{(r)} = [\mathfrak{q}_i]^{r} \cdot F_n$$

By pushing these horizontal isogenies along the dual isogeny $\hat{\phi}_{n-1}$ we obtain the directions at F_{n-1}

$$F_{n-1,i}^{(-r)} = [\mathfrak{q}_i]^{-r} \cdot F_{n-1} \longleftarrow \cdots \longleftarrow F_{n-1} \longrightarrow \cdots \longrightarrow F_{n-1,i}^{(r)} = [\mathfrak{q}_i]^{r} \cdot F_{n-1}$$

After this polynomial time reduction, the only remaining question is how to find such an endomorphism $\iota_n(\beta)$. Onuki [Onu] proposes to find β such that $\beta\mathcal{O}_n = I \cdot J$ where I and J are two ideals such that

$$I = \prod_{i=1}^{t} (\mathfrak{q}_j \cap \mathcal{O}_n)^{e_j} \qquad\qquad e_1, \ldots, e_t \in [-r, \ldots, r] \cap \mathbb{Z}$$

and J has smooth norm. Then there exists an horizontal isogeny $I \cdot (F_n, \iota_n) \to IJ \cdot (F_n, \iota_n)$ with kernel $I \cdot F_n[J]$ that can be found with a meet-in-the-middle exhaustive search strategy. The desired endomorphism $\iota_n(\beta)$ will be the compositum of $F_n \to I \cdot F_n$, that can be deduced from the knowledge of the directions at F_n, with $I \cdot (F_n, \iota_n) \to IJ \cdot (F_n, \iota_n)$.

The attack of Onuki requires a huge number of computations (2^{100} with parameters $n = 256, t = 74, \ell = 2$ and $r = 5$) and remains exponential. The high cost and complexity motivated Dartois and De Feo [DD1] to propose a different method to compute $\iota_n(\beta)$ based on a lattice reduction.

6.3.2 The attack of Dartois and De Feo

The idea of Dartois and De Feo reduces to finding a short enough non-zero vector of bounded norm in the relation lattice

$$L_n := \left\{ (e_1, \ldots, e_t) \in \mathbb{Z}^t \;\middle|\; \prod_{j=1}^{t} [\mathfrak{q}_j \cap \mathcal{O}_n]^{e_j} = [1] \text{ in } \mathcal{C}\ell(\mathcal{O}_n) \right\}$$

where $(e_1, \ldots, e_t)$ is the exponent vector of the private walk in the space $I_1 \times \ldots \times I_t \subseteq \mathbb{Z}^t$ such that $I_j = [-r_j, r_j]$ defining ψ_A as $\ker(\psi_A) = E\,[\mathfrak{q}_1^{e_1} \cdot \ldots \cdot \mathfrak{q}_t^{e_t}]$, see [CK1, §6].

Some (not so effective) countermeasures

Dartois and De Feo [DD1, §5] analyze some possible countermeasures to their attack and come to the conclusion that none of them would make OSIDH really efficient. We sketch here both their ideas.

Increasing the parameters. The first approach would be to increase the size of the parameters and therefore of the relation lattice above. The drawback is that this slows down drastically the protocol while reducing it to a lattice problem for which cryptographic schemes normally have much faster implementation.

Increasing the class group size. The idea of playing with the length of the public chain was already present in the first appearance of OSIDH but left aside to prioritize the surjectivity of the forgetful map $\mathrm{SS}_{\mathcal{O}_n}^{pr}(p) \to \mathrm{SS}(p)$, i.e., the possibility to reach as many isomorphism classes as possible. At the same time, once the public chain is fixed, Alice and Bob work in a well determined class group and a second walk comes into play, namely their private chain. The length of this path should verify the *supersingular covering bound* (SCB) and the *Minkowski norm bound* (MNB). As for the first constraint, since the size of the class group is around $h_n \approx \ell^n$, it suffices to take $\prod_{j=1}^{t}(2r_i + 1) < h_n$ as in [CK1, §6].

To theoretically prove that the product is not principal, would require $a = \prod_j q^{r_j}$ to be such that the binary quadratic form $ax^2 + bxy + cy^2$ with $\Delta = b^2 - 4ac = \Delta_K \ell^{2n}$ remains reduced ($|b| < a < c$), which would imply a bound

$$\prod_{j=1}^{t}(2r_i + 1) < \sqrt{|\Delta|/3} = \ell^n \sqrt{|\Delta_K|/3},$$

which makes (MNB) a stronger constraint. If we set $r_i = r/\log(q_i)$ then (MNB) implies roughly $t \log_\ell(2r) < n$ while SCB implies $tr < n$.

As pointed out in [DD1], one could relax the condition on the surjectivity and the corresponding inequality to resist Dartois-De Feo attack. Assuming that the relations lattice L is a random lattice, one get that the shortest infinity norm is:

$$\lambda_1^{+\infty}(L) \geq \left(1 - \frac{\log\log(t)}{t}\right) \frac{h(O_n)^{1/t}}{2}$$

Hence, we simply have to make sure that:

$$(1 - \log\log(t)/t)h(O_n)^{1/t}/2 > 2r$$

which yields $n > t\log_\ell(4r)$.

Nevertheless, this turns out to be not enough because even if $\lambda_1 > 2r$, we could still find an endomorphism by exhaustive search as explained in [DD1, §4.3].

In OSIDH [CK1], we proposed $n = 256$ which gave a coverage of $\mathcal{C}\ell(\mathcal{O}_n)$ by $\prod(2r+1)$ classes. Recognizing that we cannot attain all $\mathcal{C}\ell(\mathcal{O}_n)$ classes, but only $\prod(2r_i+1)$ of them, we impose the condition

$$\log_\ell\left(\prod_{i=1}^{t}(2r_i+1)\right) \geq \lambda\log_\ell(p) \quad \text{which gives} \quad 2\sum_{i=1}^{t} r_i > \sum_{i=1}^{t}\log_\ell(2r_i+1) \geq \lambda\log_\ell(p)$$

Thus, setting the bound $B = r_i\log_\ell(q_i)$, (SCB) is replaced by

$$2B\left(\sum_{i=1}^{t}\frac{1}{\log_\ell(q_i)}\right) \geq \lambda\log_\ell(p)$$

In order to protect against short vector attacks in the class group it is more critical to have

$$\prod_{i=1}^{t}(2r_i+1) < h(\mathcal{O}_n)$$

Therefore, we can replace (MNB) by $n \geq tB$.

Example. For $\lambda = 1$, $\log_\ell(p) = 256$, $n = 1024 = 2^{10}$, $\ell = 2$, $B = 16$ and $t = 64$ we get

$$128 = \frac{1}{2}\log_\ell(p) < 136 \approx B\sum_{i=1}^{t}(\log_\ell(q_i))^{-1} \leq tB \leq n = 1024$$

6.4 Expanding the OSIDH protocol

These prior attacks and analyses motivate the idea of enlarging the class group without touching the key space using *clouds*. In this section we propose two approaches to augments $\mathcal{C}\ell(\mathcal{O}_n(M))$ in a way that no effective data is transmitted for a third party to compute cycle relations. In both cases, it comes down to an extension of the initial chain by the two parties separately.

6.4.1 Modular OSIDH protocol

The OSIDH protocol (Section 6.2) made exclusive use of the class group action at split primes in $\mathcal{O}$. In this work we extend the protocol to include descent in the eddies at non-split primes (inert or ramified) or at large primes which are not cost-effective for use for longer isogeny walks.

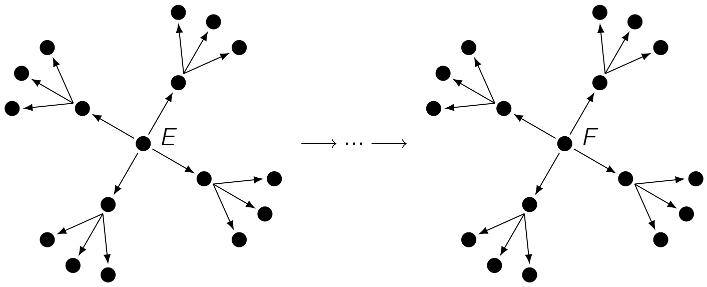
Isogeny computation. The standard practice in supersingular elliptic curve-base cryptosystems, such as SIDH [DJP], CSIDH [Cas+] and SQISign [De+], is to use torsion points of smooth order, whose existence is assured by the factorization of $p^2 - 1$. This practice is justified by the analysis of De Feo, Kieffer and Smith [DKS], who describe algorithms for isogeny computation with modular curves and torsion points. In particular, for the modular curve approach they describe explicit algorithms, "Elkies first step"' and "Elkies walk" and analyze their complexity. In conclusion, the alternative "Vélu walk" using torsion points is found more efficient.

By means of precomputed modular isogeny walks (chains and clouds), in this protocol we replace the Elkies steps with pushforward algorithms which require only gcd's of modular polynomials to "compose" isogenies, at the cost of memory. These algorithms are described below. We use Weber modular polynomials, whose sparseness and small coefficient size renders the construction of their specializations from bivariate to univariate polynomials more efficient.

The use of a torsion point of order q^r or a basis of the q^r torsion would provide an efficient method of compressing the precomputed q-isogeny chains and q-isogeny clouds, respectively. This would be of particular interest for the larger primes, and the use of the modular approach for small primes would vastly simplify the sieving step for finding suitable primes.

Pushforward of chains. The pushforward of a q-isogeny chain by an ℓ-isogeny, both from a curve E, was described in [CK1] in terms of construction of isogeny ladders. This process is iterated to push forward the q-isogeny chain along a ℓ-isogeny chain $E = F_0 \to \cdots \to F_n = F$ to create the image q-isogeny chain from F.

Pushforward of clouds. The pushforward of a q-isogeny cloud by an ℓ-isogeny chain involves pushing forward each of the $q + 1$ neighboring q-isogenies along the isogeny chain, iterating to depth r. This is represented graphically in the following image (for $q = 3$).



The ladder construction requires taking gcd's with polynomials of degree $\ell + 1$ and degree $q + 1$. As the matching of the $q + 1$ neighbors progresses, the degree of the latter polynomial decreases, but the total cost remains proportional to the size of the cloud.

The protocol

We describe the new modular OSIDH key exchange protocol using horizontal isogenies at split primes and descents at non-split primes. The protocol is broken down into three phases of parameter setup, key generation, and finally key exchange. Only the final stage requires computation in real time, as the first two steps concern the initialization of the protocol and systems of public-private keys.

The protocol requires setting up a set $\mathcal{P}_s$ of split primes in $\mathcal{O}_K$ for exchange of horizontal isogeny chains, shared by parties A and B, and two sets $\mathcal{P}_A$ and $\mathcal{P}_B$ of non-split prime powers for the exchange of isogeny clouds. The depth n at ℓ of the initial descent, and the conductors M_A and M_B, products of the prime powers $\mathcal{P}_A$ and $\mathcal{P}_B$, respectively, determine the conductors of the exchanged oriented orders $\mathcal{O}_n(M_A)$ and $\mathcal{O}_n(M_B)$. In the following section we give particular parameters choices for the sets $\mathcal{P}_A$ and $\mathcal{P}_B$ in order to assure that $M_A \approx M_B$.

Phase 0 - Parameter setup. The initialization phase consists of fixing a discriminant Δ_K and prime ℓ, and constructing a common public isogeny chain (of moduli) from a given supersingular curve $E_0/\mathbb{F}_{p^2}$ with CM by the order of discriminant Δ_K. The terminus E_n of this isogeny chain is denoted E.

$$E_0 \longrightarrow E_1 \longrightarrow \ldots \longrightarrow E_n = E$$

The protocol is based on a prior specification of disjoint prime sets $\mathcal{P}_s$ of split primes, and $\mathcal{P}_A$ and $\mathcal{P}_B$ which partition the non-split (inert or ramified) and larger split primes. In the case of split or ramified primes, only descending isogenies are used, which have the effect of increasing the class group size.

We will first look at the set $\mathcal{P}_s$: at this stage we need to precompute the initial directions in the class group (this is the analog of Elkies first step) at split primes - this step was also present in the original OSIDH protocol [CK1]. This means that, to each prime and exponent pair (q, r) in $\mathcal{P}_s$, q-isogeny chains of length $2r$ are constructed from E_0 for each prime $\mathfrak{q}$ over q, and pushed forward to E. One direction is declared positive and the other negative, so that the concatenated chains are in bijection with $[-2r, 2r]$.

Secondly, we let the two parties precompute labeled clouds at non-split primes including ℓ. For each prime and exponent pair (q, r) in $\mathcal{P}_A$ and $\mathcal{P}_B$ the q-isogeny eddy of depth r is constructed around E_0 and pushed forward to E.

This initialization data is made public, after which each party in an exchange can play the role of A or B, initializing one or more public and private key data sets as follows.

Phase I - Key generation. Following the setup procedure, A and B will compute their secret key.

On one side, A begins with $F = E$. As in the original instantiation of OSIDH we define $I_j = [-r_i, r_i]$; for each prime q_i in $\mathcal{P}_s$, A chooses a random $s_i \in I_i$. For $j = 1$ up to t, she constructs the q_j-isogeny walk of length s_j from the current F, relabeling the remaining curves in the interval $I_j + s_j = [-r_j + s_j, r_j + s_j] \subset [-2r_j, 2r_j]$ to I_j, and pushing forward the curves in the intervals I_i for each $i < j$ and the intervals $I_i + s_i$ for each $i > j$. She also pushes forward the q-clouds at each prime q in $\mathcal{P}_A$ and $\mathcal{P}_B$.

A should then compute the non-split key: for each prime and exponent pair (q_j, r_j) in $\mathcal{P}_A$, A chooses a random walk of length r_j in the cloud to a new curve F and pushes forward the remaining unused q-clouds for q in $\mathcal{P}_A$ as well as all q in $\mathcal{P}_B$ to F.

The data F and q-isogeny chains at primes q in $\mathcal{P}_s$ and q-clouds at primes q in $\mathcal{P}_B$ constitute A's public key. The indices s_j, for $1 \le j \le t$ and the isogeny walks in $\mathcal{P}_A$ form her private key.

In parallel, B constructs an equivalent public key on a curve G with data for the primes in $\mathcal{P}_s$ and in $\mathcal{P}_A$, saving his private key. The public key data for A and B can be certified by a certification authority.

Phase II - Key establishment. A obtains the public key data from B (or a certification authority) and reconstructs her isogeny walk from G using B's data for $\mathcal{P}_s$ and $\mathcal{P}_A$.

B obtains the public key data from A (or a certification authority) and reconstructs his isogeny walk from F using A's data for $\mathcal{P}_s$ and $\mathcal{P}_B$.

The curve H resulting from A's and B's random walks serves as secret key.

6.4.2 Parameter selection

In this section we make specific proposals for the sets $\mathcal{P}_s$, $\mathcal{P}_A$ and $\mathcal{P}_B$, with respect to the maximal order $\mathcal{O}_K$ of discriminant $\Delta_K = -3$ and prime $\ell = 2$, and analyze the security consequences. We defer the question of size of p, which does not impact the size of the class group or attacks in that class group, noting only that n should be of the same order of magnitude as $\log_\ell(p)$ to ensure that $|\mathcal{C}\ell(\mathcal{O}_n)| \approx p$, hence that the initial curve $E = E_n$ can be a generic curve in the supersingular graph. The use of a 256 bit prime was specified for the original OSIDH protocol [CK1].

In order to have a uniform contribution for each prime, we choose the maximum exponent r of a split prime q in $\mathcal{P}_s$ such that q^r is bounded by a bit bound B_s per prime power, i.e., we set $r = \lfloor B_s / \log_2(q) \rfloor$. For primes $\mathfrak{q}_j$ over q_j, we recall that a collision in the class group of $\mathcal{O}_n(M)$,

$$\prod_{j=1}^{t} \mathfrak{q}_j^{a_j} \sim \prod_{j=1}^{t} \mathfrak{q}_j^{b_j},$$

with $|a_j|, |b_j| \le r_j$, gives rise to an endomorphism α, with exponents $s_j = |a_j - b_j|$ bounded by $2r_j$, up to replacing $\mathfrak{q}_j$ with $\bar{\mathfrak{q}}_j$, and with norm satisfying

$$\mathrm{N}(\alpha) = \prod_{j=1}^{t} q_j^{s_j} > \frac{|\Delta|}{4} = M^2 \ell^{2n} \frac{1 + |\Delta_K|}{4}.$$

Hence, if we choose t, B_s, M and n such that the Minkowski norm bound (MNB)

$$\sum_{j=1}^{t} r_j \log_\ell(q_j) \le tB_s \le \log_\ell(M) + n$$

holds, no such endomorphism exists in $\mathcal{O}_n(M)$, and the lattice-based class group attack of Dartois-De Feo [DD1] does not apply.

Security considerations. As a consequence of the analysis of security against lattice-based class group attack, the number t of split primes, and the product tB_s in particular, should be strictly controlled, as well as the size of the conductors M (equal to M_A or M_B). In the example below we begin with $t = 10$, and a bit bound $B_s = 32$ (or 24), giving $tB_s = 320$ bits (or 240 bits). We also use eddies at larger split primes which contribute equally a term $\log_2(q)$ to both sides of the inequality, but allow for an increase in the number of vertices reachable in the supersingular isogeny graph.

Parameter sets. In what follows we set $\Delta_K = -3$ and $\ell = 2$, and propose parameter sets $\mathcal{P}_s$, $\mathcal{P}_A$ and $\mathcal{P}_B$ out of the primes up to 163 with an analysis of their security constraints. We subsequently consider the security implications of modifying B_s and moving split primes into the sets $\mathcal{P}_A$ or $\mathcal{P}_B$ to use for descents (in the eddy).

Ten split primes. For a parameter set, we consider $\mathcal{P}_s$ consisting of the first 10 split primes q in the maximal order $\mathcal{O}_K$ of discriminant -3, with a bound $r = \lfloor B_s / \log_2(q) \rfloor$, with $B_s = 32$, such that we take a walk whose length is in the interval $[-r, r]$ — a positive value is with respect to a prime $\mathfrak{q}$ over q and a negative value is with respect to $\bar{\mathfrak{q}}$.

	$q:$	7	13	19	31	37	43	61	67	73	79
$\mathcal{P}_s:$	$r:$	11	8	7	6	6	6	5	5	5	5
	$\#:$	23	17	15	13	13	13	11	11	11	11

The third line is the number $2r + 1$ of curves reached by a uniformly random length walk. This gives a logarithmic contribution of

$$\sum_{j=1}^{10} \log_2(2r_j + 1) = 37.4569...$$

to the entropy of the random walk. On the other hand, the logarithmic norm, which we must bound is:

$$\sum_{j=1}^{10} r_j \log_2(q_j) = 306.2115...(< 320 = 32 \cdot 10).$$

We partition the remaining primes up to 163 into sets $\mathcal{P}_A$ and $\mathcal{P}_B$, with a radius for the cloud (or eddy), as follows:

	$q:$	2	11	17	41	47	59	83	101	103	109	131	149	151	157
$\mathcal{P}_A:$	$r:$	7	2	1	1	1	1	1	1	1	1	1	1	1	1
	$\#:$	128	132	18	42	48	60	84	102	102	108	132	150	150	156

including the split primes 103, 109, 151 and 157, and

	$q:$	3	5	23	29	53	71	89	97	107	113	127	137	139	163
$\mathcal{P}_B:$	$r:$	4	3	1	1	1	1	1	1	1	1	1	1	1	1
	$\#:$	81	150	24	30	54	72	90	96	108	114	126	138	138	162

including the split primes 97, 127, 139 and 163.

The specification of the descending isogenies in $\mathcal{P}_A$ (by B) at the split primes leaks the horizontal directions for these primes, giving an additional contribution of

$$\log_2(103) + \log_2(109) + \log_2(151) + \log_2(157) = 27.9877..$$

bits to the logarithmic norm, and symmetrically, the descending isogenies given in $\mathcal{P}_B$ (by A) leak an additional

$$\log_2(97) + \log_2(127) + \log_2(139) + \log_2(163) = 28.0562...$$

bits to the logarithmic norm. In both cases the maximal logarithmic norm which can be attained is less than 335 bits. These prime sets each contribute a $\log_2(M)$ of 90 bits, such that n must be at least 244 to defeat the lattice-based class group attack.

The third line of the tables for $\mathcal{P}_A$ and $\mathcal{P}_B$ are the numbers $m(q, r)$ of curves at distance r:

$$m(q, r) = \left(q - \left(\frac{\Delta_K}{q}\right)\right) q^{r-1},$$

and for each set we have $\sum \log_2(m(q, r)) \approx \log_2(M) \approx 90$ bits contributed to the number of curves reached by the random isogeny walk. Together with the contribution of split primes, this gives approximately 128 bits (out of $\log_2(p)$ bits in $|\mathrm{SS}(p)|$) to the *supersingular covering bound* (SCB).

Remark. While the lattice-based class group attack is rendered ineffective for $n = 256$ by the logarithmic

Minkowski norm bound (MNB) (including split primes in $\mathcal{P}_A$ or $\mathcal{P}_B$ with $r = 1$),

$$\sum_{j=1}^{t} r_j \log_2(q_j) \approx 344 \leq \log_2(M) + n \approx 90 + 256 = 356,$$

with respect to the class group $\mathcal{C}\ell(\mathcal{O}_n(M))$, the margin of security of 12 bits is insufficient. It suffices to construct a putative cycle, with respect to the class group $\mathcal{C}\ell(\mathcal{O}_m(M))$ for $m + 12 < n$, and construct the corresponding isogeny in the class group $\mathcal{C}\ell(\mathcal{O}_n(M))$, and carry out an exhaustive search up to radius 12 in the ℓ-isogeny graph to find a collision. This suggests a more moderate bound B_s like 24, which gives the split prime table:

	$q:$	7	13	19	31	37	43	61	67	73	79
$\mathcal{P}_s:$	$r:$	8	6	5	4	4	5	4	3	3	3
	$\#:$	17	13	11	9	9	9	9	7	7	7

which results in a more secure 120-bit margin of security:

$$\sum_{j=1}^{t} r_j \log_2(q_j) \approx 236 \leq \log_2(M) + n \approx 356.$$

On the other hand the random walk at split primes contributes only 32 bits to the number of curves (modular invariants) attainable in the graph, for a total of 122 bits.

Two split primes. For comparison we consider the use of just two shared split primes with a bit bound $B_s = 64$ which allows for longer walks at the split primes.

	$q:$	7	13
$\mathcal{P}_s:$	$m:$	22	17
	$\#:$	45	35

The longer walks, with $B_s = 64$ contribute 125 bits to the maximal norm of a product, while giving $35 \cdot 45 = 1575$ possible isogenies.

The remaining 36 primes up to 163 can be partitioned into sets $\mathcal{P}_A$ and $\mathcal{P}_B$ as follows:

	$q:$	2	11	17	31	37	41	47	59	67	73	83	101	103	109	131	149	151	157
$\mathcal{P}_A:$	$r:$	7	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	$\#:$	128	132	18	30	36	42	48	60	66	72	84	102	102	108	132	150	150	156

including the split primes $31, 37, 67, 73, 103, 109, 151, 157$, and

	$q:$	3	5	19	23	29	43	53	61	71	79	89	97	107	113	127	137	139	163
$\mathcal{P}_B:$	$r:$	4	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	$\#:$	81	150	18	24	30	42	54	60	72	78	90	96	108	114	126	138	138	162

including the split primes $19, 43, 61, 79, 97, 127, 139, 163$. These split primes each contribute 50 bits to the maximal norm of a product ideal, and each set $\mathcal{P}_A$ and $\mathcal{P}_B$ contributes 112 bits to the conductor M ($= M_A$ or M_B). As a result we have a bound of $125 + 50 = 175$ bits on logarithmic norms, compared to $\log_2(M) = n = 112 + 256 = 368$ coming from the discriminant bounds.

Remark. We emphasize the disparity between 1575 curves reachable by 7- and 13-isogenies, a contribution of just 10 bits, compared to their maximal degree (norm of a product ideal) of 125 bits. These products will be complemented with eight other split primes in $\mathcal{P}_A$ or $\mathcal{P}_B$ (see below), but it is unlikely the set of $1575 \cdot 3^8$ possible products (a 23-bit number) will contain a principal ideal in a class group of 360 bits. This improbability of a collision with the principal class is ignored in favor of the stronger absolute bound on norms of endomorphisms, which gives a provable zone of exclusion.

Conclusion. By the same logic one could consider reducing to zero split primes. However, the number of curves which can be reached will be fewer, since the 45 or 35 isogeny neighbors which can be used by both A and B, which will be replaced by an eddy at 7 or 13 (in $\mathcal{P}_A$ or $\mathcal{P}_B$) of fewer than 1575 neighbors, since to depth 2 this contribute 42 and 156 neighbors respectively. Extending to depth 3 would explode the size of the cloud beyond reasonable limits (passing at most 162 neighbors per prime up to 163).

6.4.3 Security considerations revisited

The parameter choices described for OSIDH [CK1] where presented with two competing considerations, see Section 6.2.4. The K-orientation must remain hidden and a large logarithmic covering radius λ of total supersingular curves should be reached by the class group, *supersingular covering bound* (SCB):

$$\lambda \leq \log_p \left(|\mathcal{C}\ell(\mathcal{O}_n)|\right)$$

so that $\mathcal{C}\ell(\mathcal{O}_n) \to \mathrm{SS}(p)$ has large image. Secondly, the *Minkowski norm bound* (MNB) implies the non-existence of $I \cap \ker\left(\mathbb{Z}^t \to \mathcal{C}\ell(\mathcal{O}_n)\right)$, see Lemma 6.2.

In CSIDH-512 [Cas+], the proposed parameters $n = 256$, $t = 74$, and $r = 5$ were supposed to sample class group elements uniformly; this means that the authors were more interested in insuring surjectivity (CGCB) rather than injectivity (MNB) of $I \to \mathcal{C}\ell(\mathcal{O})$,

$$\log_\ell(|I|) = t\log_\ell(2r+1) \approx \log_\ell\left(|\mathcal{C}\ell(\mathcal{O}_n)|\right) \approx n.$$

Unfortunately, to provably preclude collisions in OSIDH, one needs a stronger bound on the norms of elements:

$$\sum_{j=1}^{t} r_j \log_\ell(q_j) < n + \log_\ell(M)$$

which clearly fails for $(t = 74, r_j = 5, n = 256)$. In fact these parameters permitted the full class group computation for the CSIDH-512, resulting in the CSI-FiSh protocol [BKV]. In the setting of CSIDH-512, the class group and its order was unknown and $O(t)$ cycles are needed to determine the class group. In the setting of OSIDH, the class group is known, and only one short vector is needed to determine the K-orientation.

Onuki [Onu, §6.3] recalls the design objectives from OSIDH [CK1, §5.1] that the knowledge of the K-orientation breaks the cryptosystem, giving an ascending walk up the ℓ-isogeny chain, and proposed an exponential meet-in-the-middle attack. Dartois and De Feo [DD1] carry out the class group attack to determine a cycle hence breaking this parameter set. Combining a class group analysis in $\mathcal{C}\ell(\mathcal{O}_{n-r})$ and meet-in-the-middle attack to depth r makes this approach even more significant.

The norm bound (MNB) suggests using a uniform bound B_s on $r_j\log_\ell(q_j)$ rather than the exponents r_j. This gives

$$\lambda\log_\ell(p) \leq \sum_{i=1}^{t}\log_\ell(2r_j+1) \leq \sum_{j=1}^{t} r_j\log_\ell(q_j) \leq tB_s < n + \log_\ell(M)$$

used in the parameters selection of Section 6.4.1 and for which $(t = 64, B_s = 16, n = 1024)$ represent a choice of parameters ensuring injectivity of $I \to \mathcal{C}\ell(\mathcal{O})$, see Section 6.3.2.

Bibliography

[AAM] G. Adj, O. Ahmadi, and A. Menezes. "On isogeny graphs of supersingular elliptic curves over finite fields". In: *Finite Fields and Their Applications* 55 (2019), pp. 268–283.

[ACL+1] S. Arpin, M. Chen, K.E. Lauter, R. Scheidler, K.E. Stange, and H.T.N. Tran. "Orienteering with one endomorphism". In: (2022). url: `https://arxiv.org/abs/2201.11079`.

[ACL+2] S. Arpin, M. Chen, K.E. Lauter, R. Scheidler, K.E. Stange, and H.T.N. Tran. "Orientations and cycles in supersingular isogeny graphs". In: (2022). url: `https://arxiv.org/abs/2205.03976`.

[AL] A.O.L Atkin and J. Lehner. "Hecke Operators on $\Gamma_0(m)$." In: *Mathematische Annalen* 185 (1970), pp. 134–160.

[Apo] T.M. Apostol. *Modular Functions and Dirichlet Series in Number Theory*. Graduate Texts in Mathematics. Springer New York, 1997.

[Are+] C. Arene, T. Lange, M. Naehrig, and C. Ritzenthaler. "Faster Computation of the Tate Pairing". In: *Journal of Number Theory* (Jan. 2010).

[Arp1] S. Arpin. *Good Primes Supersingular* 2, 3*-Isogeny Graphs*. Notes online. 2019. url: `http://math.colorado.edu/~saar7867/GoodPrimes.pdf`.

[Arp2] S. Arpin. *Adding Level Structure to Supersingular Elliptic Curve Isogeny Graphs*. 2022.

[Arp+] S. Arpin, C. Camacho-Navarro, K.E. Lauter, J. Lim, K. Nelson, T. Scholl, and J. Sotáková. *Adventures in Supersingularland*. 2019. url: `https://arxiv.org/abs/1909.07779`.

[Atk1] A.O.L Atkin. *The Number of Points on an Elliptic Curve Modulo a Prime (I)*. Email. 1991. url: `https://www.lix.polytechnique.fr/Labo/Francois.Morain/AtkinEmails/19910614.txt`.

[Atk2] A.O.L Atkin. *The Number of Points on an Elliptic Curve Modulo a Prime (II)*. Email. 1992. url: `https://www.lix.polytechnique.fr/Labo/Francois.Morain/AtkinEmails/19920319.txt`.

[Baj] A. Bajolet. "Aspects numèriques de l'analyse diophantienne". PhD Thesis. Universitè Bordeaux 1, 2012.

[Bal+] J.S. Balakrishnan, N. Dogra, J.S. Müller, J. Tuitman, and J. Vonk. "Explicit Chabauty-Kim for the split Cartan modular curve of level 13". In: *Annals of Mathematics* 189.3 (2019), pp. 885–944.

[Bar1] B. Baran. "A modular curve of level 9 and the class number one problem". In: *Journal of Number Theory* 129.3 (Mar. 2009), pp. 715–728.

[Bar2] B. Baran. "Normalizers of non-split Cartan subgroups, modular curves, and the class number one problem". In: *Journal of Number Theory* 130.12 (2010), pp. 2753–2772.

[Bar3] B. Baran. "An exceptional isomorphism between modular curves of level 13". In: *Journal of Number Theory* 145 (2014), pp. 273–300.

[Barb] A. Barbon. "Algebraic Brill-Noether Theory". Master Thesis. Radboud University Nijmegen, Vrije Universiteit Amsterdam and Universiteit Van Amsterdam, 2014.

[BBM] A. Bajolet, Y. Bilu, and M. Matschke. *Computing integral points on* $X_{ns}^{+}(p)$. 2020. arXiv: `1212.0665 [math.NT]`.

[BCL] R. Bröker, D. Charles, and K.E. Lauter. "Evaluating Large Degree Isogenies and Applications to Pairing Based Cryptography". In: *Pairing-Based Cryptography – Pairing 2008*. Ed. by S.D. Galbraith and K.G. Paterson. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 100–112.

[BD] P. Bruin and S. Dahmen. *Modular Forms*. Course notes online at http://www.few.vu.nl/~sdn249/modularforms16/Notes.pdf.

[Bel] J. Belding. "Number Theoretic Algorithms For Elliptic Curves". PhD Thesis. University of Maryland, College Park, 2008.

[Bel+] J. Belding, R. Bröker, A. Enge, and K. Lauter. "Computing Hilbert Class Polynomials". In: *ANTS-VIII - Eighth Algorithmic Number Theory Symposium*. Ed. by A. J. van der Poorten and A. Stein. Vol. 5011. Lecture Notes in Computer Science. Banff, Canada: Springer-Verlag, 2008, pp. 282–295. url: https://hal.inria.fr/inria-00246115.

[Ben] P. Benioff. "The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines". In: *Journal of Statistical Physics* 22 (1980), pp. 563–591.

[Ber+] D.J. Bernstein, C. Chuengsatiansup, D. Kohel, and T. Lange. "Twisted Hessian Curves". In: *Progress in Cryptology – LATINCRYPT 2015*. Ed. by K. Lauter and F. Rodríguez-Henríquez. Springer International Publishing, 2015, pp. 269–294.

[Bha] M. Bhargava. "Higher composition laws. I. A new view on Gauss composition, and quadratic generalizations". In: *Annals of Mathematics* 159 (1 2004), pp. 217–250.

[BJ1] O. Billet and M. Joye. "The Jacobi Model of an Elliptic Curve and Side-Channel Analysis". In: *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*. Ed. by M. Fossorier, T. Høholdt, and A. Poli. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 34–42.

[BJ2] E. Brier and M. Joye. "Weierstraß Elliptic Curves and Side-Channel Attacks". In: *Public Key Cryptography - PKC 2003, 6th International Workshop on Theory and Practice in Public Key Cryptography, Miami, FL, USA, January 6-8, 2003, Proceedings*. Vol. 2274. Lecture Notes in Computer Science. Springer, 2002, pp. 335–345.

[BKV] W. Beullens, T. Kleinjung, and F. Vercauteren. "CSI-FiSh: Efficient Isogeny Based Signatures Through Class Group Computations". In: *Advances in Cryptology – ASIACRYPT 2019*. Ed. by S.D. Galbraith and S. Moriai. Cham: Springer International Publishing, 2019, pp. 227–247.

[BKX] F. Bars, A. Kontogeorgis, and X. Xarles. "Bielliptic and Hyperelliptic modular curves $X(N)$ and the group $\mathrm{Aut}(X(N))$". In: *Acta Arithmetica* 161.3 (July 2013), pp. 283–299.

[BL] D.J. Bernstein and T. Lange. "Faster Addition and Doubling on Elliptic Curves". In: *Advances in Cryptology – ASIACRYPT 2007*. Ed. by Kaoru Kurosawa. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 29–50.

[BLS] R. Bröker, K. Lauter, and A. Sutherland. "Modular polynomials via isogeny volcanoes". In: *Mathematics of Computation* 81 (2012), pp. 1201–1231.

[BM] A. Bhand and M.R. Murty. "Class Numbers of Quadratic Fields". In: *Hardy-Ramanujan Journal* Volume 42 - Special Commemorative volume in honor of Alan Baker - 2019 (May 2020), pp. 17–25. url: https://hal.archives-ouvertes.fr/hal-02554226.

[Boo] J. Booher. *Modular Curves and the Class Number One Problem*. Expository Notes. Available online at https://www.math.canterbury.ac.nz/~j.booher/expos/class_number_one.pdf. 2014.

[Bos+] A. Bostan, F. Morain, B. Salvy, and É. Schost. "Fast algorithms for computing isogenies between elliptic curves". In: *Mathematics of Computation* 77.263 (Sept. 2008), pp. 1755–1778.

[BP] Y. Bilu and P. Parent. "Serre's uniformity problem in the split Cartan case". In: *Annals of Mathematics* 173 (Aug. 2011), pp. 569–584.

[BPR] Y. Bilu, P. Parent, and M. Rebolledo. *Rational points on $X_0^+(p^r)$*. 2011.

[BS1] G. Bisson and A.V. Sutherland. "Computing the endomorphism ring of an ordinary elliptic curve over a finite field". In: *Journal of Number Theory*. Elliptic Curve Cryptography 131.5 (2011), pp. 815–831. url: https://hal.inria.fr/inria-00383155.

[BS2] R. Bröker and A.V. Sutherland. "An explicit height bound for the classical modular polynomial". In: *The Ramanujan Journal* 22.3 (2010), pp. 293–313.

[BT] P. Bayer and A. Travesa. *Corbes modulars: Taules*. UB-UAB-UPC, Barcelona. Notes del Seminari de Teoria de Nombres de Barcelona. 1992.

[Cas1] J.W.S. Cassels. *Lectures on Elliptic Curves*. London Mathematical Society Student Texts. Cambridge University Press, 1991.

[Cas2] J.W.S. Cassels. *Rational Quadratic Forms*. Dover Books on Mathematics. Dover Publications, 2008.

[Cas+] W. Castryck, T. Lange, C. Martindale, L. Panny, and J. Renes. "CSIDH: An Efficient Post-Quantum Commutative Group Action". In: *Advances in Cryptology - ASIACRYPT 2018*. Ed. by T. Peyrin and S. Galbraith. Cham: Springer International Publishing, 2018, pp. 395–427.

[CC] D.V Chudnovsky and G.V Chudnovsky. "Sequences of Numbers Generated by Addition in Formal Groups and New Primality and Factorization Tests". In: *Adv. Appl. Math.* 7.4 (Dec. 1986), pp. 385–434.

[CD] W. Castryck and T. Decru. *An efficient key recovery attack on SIDH (preliminary version)*. Cryptology ePrint Archive, Paper 2022/975. 2022. url: `https://eprint.iacr.org/2022/975`.

[Cer1] J.M. Cerviño. *Supersingular elliptic curves and maximal quaternionic orders*. url: `https://www.uni-math.gwdg.de/tschinkel/SS04/cervino.pdf`.

[Cer2] J.M. Cerviño. *On the Correspondence between Supersingular Elliptic Curves and maximal quaternionic Orders*. 2004. url: `https://arxiv.org/abs/math/0404538`.

[CF] J.W.S. Cassels and E.V. Flynn. *Prolegomena to a Middlebrow Arithmetic of Curves of Genus 2*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1996.

[CGL] D. Charles, E. Goren, and K.E. Lauter. *Cryptographic hash functions from expander graphs*. Cryptology ePrint Archive, Paper 2006/021. 2006. url: `https://eprint.iacr.org/2006/021`.

[CH] T. Couveignes J.-M. and Henocq. "Action of Modular Correspondences around CM Points". In: *Algorithmic Number Theory*. Ed. by C. Fieker and D. Kohel. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 234–243.

[Che] I. Chen. "On Siegel's Modular Curve of Level 5 and the Class Number One Problem". In: *Journal of Number Theory* 74 (1999), pp. 278–297.

[CJS] A. Childs, D. Jao, and V. Soukharev. "Constructing elliptic curve isogenies in quantum subexponential time". In: *Journal of Mathematical Cryptology* 8.1 (2014), pp. 1–29.

[CK1] L. Colò and D. Kohel. "Orienting supersingular isogeny graphs". In: *Journal of Mathematical Cryptology* 14 (2020), pp. 414–437.

[CK2] L. Colò and D. Kohel. *On the modular OSIDH protocol*. To appear. 2022.

[CKK] B. Cho, N. Kim, and J. Koo. "Affine models of the modular curves $X(p)$ and its application". In: *The Ramanujan Journal* 24 (Feb. 2011), pp. 235–257.

[Coh] H. Cohen. *A Course in Computational Algebraic Number Theory*. Graduate Texts in Mathematics. Springer-Verlag, 1993.

[Conr] K. Conrad. *Hensel's Lemma*. Online at `https://kconrad.math.uconn.edu/blurbs/gradnumthy/hensel.pdf`.

[Conw] J.H. Conway. *The Sensual (Quadratic) Form*. Mathematical Association of America, 1997.

[Cos] C. Costello. *Supersingular isogeny key exchange for beginners*. Cryptology ePrint Archive, Paper 2019/1321. 2019. url: `https://eprint.iacr.org/2019/1321`.

[Cou] J.M. Couveignes. "Hard Homogeneous Spaces". In: *IACR Cryptology ePrint Archive* 2006 (2006), p. 291. url: `https://eprint.iacr.org/2006/291`.

[Cox] D.A. Cox. *Primes of the Form x^2+ny^2: Fermat, Class Field Theory, and Complex Multiplication*. Monographs and textbooks in pure and applied mathematics. Wiley, 1989.

[CS1] M. Chenu and B. Smith. "Higher-degree supersingular group actions". In: *Mathematical Cryptology* (2021). url: `https://hal.inria.fr/hal-03288075`.

[CS2] H. Cohen and F. Strömberg. *Modular Forms*. Graduate Studies in Mathematics. American Mathematical Society, 2017.

[CS3] C. Costello and B. Smith. "Montgomery curves and their arithmetic: The case of large characteristic fields". In: *Journal of Cryptographic Engineering* 8 (Mar. 2017), pp. 227–240.

[CV] W. Castryck and F. Vercauteren. “Toric forms of elliptic curves and their arithmetic”. In: *Journal of Symbolic Computation* 46.8 (2011), pp. 943–966.

[Dan] H-B. Daniels. “Siegel Functions, Modular Curves, and Serre’s Uniformity Problem”. PhD Thesis. University of Connecticut, 2013.

[DD1] P. Dartois and L. De Feo. “On the security of OSIDH”. In: (2021). url: `https://eprint.iacr.org/2021/1681`.

[DD2] I. Del Corso and R. Dvornicich. *Finite groups of units of finite characteristic rings*. 2017. arXiv: `1607.00965 [math.RA]`.

[De+] L. De Feo, D. Kohel, A. Leroux, C. Petit, and B. Wesolowski. “SQISign: Compact Post-Quantum Signatures from Quaternions and Isogenies”. In: *26th Annual International Conference on the Theory and Application of Cryptology and Information Security (Asiacrypt 2020)*. Springer, Aug. 2020.

[DeF1] L. De Feo. “Exploring Isogeny Graphs,” Habilitation á diriger des recherches. Université de Versailles Saint-Quentin-en-Yvelines, 2018. url: `https://defeo.lu/hdr/#manuscript`.

[DeF2] L. De Feo. *Mathematics of Isogeny Based Cryptography*. 2017. url: `https://arxiv.org/abs/1711.04062`.

[Deu] M. Deuring. “Die Typen der Multiplikatorenringe elliptischer Funktionenkörper”. In: *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 14 (1941), pp. 197–272.

[DG] C. Delfs and S. Galbraith. “Computing isogenies between supersingular elliptic curves over $\mathbb{F}_p$”. In: *Designs, Codes and Cryptography* 78 (Oct. 2013), pp. 425–440.

[DH] W. Diffie and M.E. Hellman. “New Directions in Cryptography”. In: *IEEE Transactions on Information Theory* 22.6 (Nov. 1976), pp. 644–654.

[DI] F. Diamond and J. Im. “Modular forms and modular curves”. In: *Seminar on Fermat’s Last Theorem* (1995), pp. 39–133.

[DJ] D. Jao and L. De Feo. “Towards Quantum-Resistant Cryptosystems from Supersingular Elliptic Curve Isogenies”. In: *Post-Quantum Cryptography*. Ed. by B.-Y. Yang. Springer Berlin Heidelberg, 2011, pp. 19–34.

[DJP] L. De Feo, D. Jao, and J. Plût. “Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies”. In: *Journal of Mathematical Cryptology* 8.3 (2014), pp. 209–247.

[DKS] L. De Feo, J. Kieffer, and B. Smith. “Towards Practical Key Exchange from Ordinary Isogeny Graphs”. In: *Advances in Cryptology – ASIACRYPT 2018*. Ed. by T. Peyrin and S. Galbraith. Cham: Springer International Publishing, 2018, pp. 365–394.

[DKW1] R. Drylo, T. Kijko, and M. Wronski. “Determining Formulas Related to Point Compression on Alternative Models of Elliptic Curves”. In: *Fundamenta Informaticæ* 169 (2019), pp. 285–294.

[DKW2] R. Drylo, T. Kijko, and M. Wronski. “Efficient Montgomery-like formulas for general Huff’s and Huff’s elliptic curves and their applications to the isogeny-based cryptography”. In: *IACR Cryptol. ePrint Arch.* (2020). `https://eprint.iacr.org/2020/526.pdf`.

[DM] P. Deligne and D. Mumford. “The irreducibility of the space of curves of given genus”. In: *Publications Mathématiques de l’IHÉS* 36 (1969), pp. 75–109.

[DMS] V. Dose, P. Mercuri, and C. Stirpe. “Cartan modular curves of level 13”. In: *arXiv: Number Theory* (2017).

[Dos+] V. Dose, J. Fernández, J. González, and R. Schoof. “The automorphism group of the non-split Cartan modular curve of level 11”. In: *Journal of Algebra* 417 (Nov. 2014), pp. 95–102.

[DR] P. Deligne and M. Rapoport. “Les Schémas de Modules de Courbes Elliptiques.” In: *Modular Functions of One Variable II*. Ed. by P. Deligne and W. Kuijk. Vol. 349. Lecture Notes in Mathematics. Springer, Berlin, Heidelberg, 1973.

[DS] F. Diamond and J. Shurman. *A First Course in Modular Forms*. Graduate Texts in Mathematics. Springer-Verlag New York, 2005.

[Dur] L.K. Durst. “The Apparition Problem for Equianharmonic Divisibility Sequences”. In: *Proceedings of the National Academy of Sciences of the United States of America*. Vol. 38. Apr. 1952, pp. 330–333.

[Ech] D. Kohel. *Echidna Algorithm*. Version 5.0. url: `https://www.i2m.univ-amu.fr/perso/david.kohel/alg/index.html`.

[Edw] H.M. Edwards. “A normal form for elliptic curves”. In: *Bulletin of the American Mathematical Society* 44 (2007), pp. 393–422.

[Eic1] M. Eichler. “Zur Zahlentheorie der Quaternionen-Algebren”. In: *Journal für die reine und angewandte Mathematik* 195 (1955), pp. 127–151.

[Eic2] M. Eichler. “The Basis Problem for Modular Forms and the Traces of the Hecke Operators”. In: *Modular Functions of One Variable I*. Ed. by W. Kuijk. Berlin, Heidelberg: Springer Berlin Heidelberg, 1973, pp. 75–152.

[Eis+] K. Eisenträger, S. Hallgren, K.E. Lauter, T. Morrison, and C. Petit. “Supersingular Isogeny Graphs and Endomorphism Rings: Reductions and Solutions”. In: *Advances in Cryptology – EUROCRYPT 2018*. Ed. by J.B. Nielsen and V. Rijmen. Cham: Springer International Publishing, 2018, pp. 329–368.

[ElG] T. ElGamal. “A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms”. In: *Advances in Cryptology*. Ed. by G.R. Blakley and D. Chaum. Berlin, Heidelberg: Springer Berlin Heidelberg, 1985, pp. 10–18.

[Elk1] N.D. Elkies. “Explicit modular towers”. In: *Proceedings of the Thirty-fifth annual Allerton conference on communication, control and computing*. Ed. by T. Basar and A. Vardy. Univ. of Illinois at Urbana-Champaign. 1997, pp. 23–32.

[Elk2] N.D. Elkies. “Elliptic and modular curves over finite fields and related computational issues”. In: *Computational Perspectives in Number Theory: Conference in Honor of A.O.L. Atkin*. Ed. by D.A. Buell and J.T. Teitelbaum. American Mathematical Society. 1998, pp. 21–76.

[ES] A. Enge and R. Schertz. “Modular curves of composite level”. In: *Acta Arithmetica* 118.2 (2005), pp. 129–141. url: `http://eudml.org/doc/278252`.

[Fey] R.P. Feynman. “Simulating physics with computers”. In: *International journal of theoretical physics* 21.6/7 (1982), pp. 467–488.

[FFT] P.B. Fouazou Lontouo, E. Fouotsa, and D. Tieudjo. “Division polynomials on the Hessian model of elliptic curves”. In: *Applicable Algebra in Engineering, Communication and Computing* (Nov. 2020), pp. 1–16.

[FJ] R.R. Farashahi and M. Joye. “Efficient Arithmetic on Hessian Curves”. In: *Public Key Cryptography – PKC 2010*. Ed. by P.Q. Nguyen and D. Pointcheval. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 243–260.

[FM] M. Fouquet and F. Morain. “Isogeny Volcanoes and the SEA Algorithm”. In: *Algorithmic Number Theory*. Ed. by C. Fieker and D. Kohel. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 276–291.

[Fou] M. Fouquet. “Anneau d’endomorphismes et cardinalité des courbes elliptiques : aspects algorithmiques”. PhD Thesis. École Polytechnique, 2001.

[Ful] W. Fulton. *Algebraic Curves: An Introduction to Algebraic Geometry*. Advanced book classics. Addison-Wesley Pub. Co., Advanced Book Program, 1989.

[Gal1] S.D. Galbraith. “Equations For Modular Curves”. PhD Thesis. University of Oxford, St. Cross College, Mathematical Institute, 1996. url: `https://www.math.auckland.ac.nz/~sgal018/thesis.html`.

[Gal2] S.D. Galbraith. “Constructing Isogenies between Elliptic Curves Over Finite Fields”. In: *LMS Journal of Computation and Mathematics* 2 (1999), pp. 118–138.

[Gal3] S.D. Galbraith. “Rational points on $X_0^+(N)$ and quadratic $\mathbb{Q}$-curves”. In: *J. de la Theorie des Nombres de Bordeaux* 14 (2002), pp. 205–219.

[Gal4] S.D. Galbraith. *Mathematics of Public Key Cryptography*. 1st. USA: Cambridge University Press, 2012.

[Gal+] S.D. Galbraith, C. Petit, B. Shani, and Y.B. Ti. "On the security of supersingular isogeny cryptosystems". In: *Advances in Cryptology – ASIACRYPT 2016*. Ed. by J.H. Cheon and T. Takagi. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 63–91.

[Gat] A. Gathmann. *Algebraic Geometry*. TU Kaiserslautern. Course notes online at `https://www.mathematik.uni-kl.de/~gathmann/class/alggeom-2014/alggeom-2014.pdf`. 2014.

[Gau] C.F. Gauss. *Disquisitiones Arithmeticae*. Trans. by A.A. Clarke. Lipsiae: In commissis apud Gerh. Fleischer, 1801.

[Gee] A.C.P. Gee. "Class fields by Shimura reciprocity". PhD Thesis. University of Amsterdam, 2001.

[GGX] H. Gu, D. Gu, and W. Xie. "Efficient Pairing Computation on Elliptic Curves in Hessian Form". In: *Proceedings of the 13th International Conference on Information Security and Cryptology*. ICISC'10. Seoul, Korea: Springer-Verlag, 2010, pp. 169–176.

[GL] R.J. González and J.C. Lario. "Rational and Elliptic Parametrizations of $\mathbb{Q}$-Curves". In: *Journal of Number Theory* 72.1 (1998), pp. 13–31.

[Gon] R.J. Gonzalez. "Equations of hyperelliptic modular curves". In: *Annales de l'Institut Fourier* 41.4 (1991), pp. 779–795.

[Gou] F. Gouvea. *p-adic Numbers: An Introduction*. Universitext. Springer Berlin Heidelberg, 2003.

[GS] A.C.P. Gee and P. Stevenhagen. "Generating Class Fields Using Shimura Reciprocity". In: *Algorithmic Number Theory: Third International Symposium, ANTS-III, Portland, Orgeon, USA, June 21-25, 1998, Proceedings*. Ed. by J.P. Buhler, R.L. Dobrusin, and P.A. Schweitzer. Springer Science & Business Media, 1998, pp. 441–453.

[GV] S.D. Galbraith and F. Vercauteren. "Computational problems in supersingular elliptic curve isogenies". In: *Quantum Information Processing* 17 (2018), pp. 1–22.

[GW] U. Görtz and T. Wedhorn. *Algebraic Geometry: Part I: Schemes. With Examples and Exercises*. Advanced Lectures in Mathematics. Vieweg+Teubner Verlag, 2010.

[Har] R. Hartshorne. *Algebraic Geometry*. Graduate Texts in Mathematics. Springer-Verlag, 1977.

[Has1] Y. Hasegawa. "Table of quotient curves of modular curves $X_0(N)$ with genus 2". In: *Proc. Japan Acad. Ser. A* 71 (1995), pp. 235–239.

[Has2] Y. Hasegawa. "Hyperelliptic modular curves $X_0^*(N)$". In: *Acta Arithmetica* 81 (1997), pp. 369–385.

[Hat] A. Hatcher. *Topology of Numbers*. Available online at `https://pi.math.cornell.edu/~hatcher/TN/TNpage.html`. 2020.

[Hes] O. Hesse. "Über die Elimination der Variabeln aus drei algebraischen Gleichungen vom zweiten Grade mit zwei Variabeln". In: *Journal für die reine und angewandte Mathematik* 10 (1844), pp. 68–96.

[His] H. Hisil. "Elliptic Curve, Group Law and Efficient Computation". PhD Thesis. Queensland University of Technology, 2010.

[His+] H. Hisil, K.K.H. Wong, G. Carter, and E. Dawson. "Jacobi Quartic Curves Revisited". In: *Proceedings of the 14th Australasian Conference on Information Security and Privacy*. ACISP '09. Brisbane, Australia: Springer-Verlag, 2009, pp. 452–468.

[Huf] G.B. Huff. "Diophantine problems in geometry and elliptic ternary forms". In: *Duke Mathematical Journal* 15.2 (1948), pp. 443–453.

[Hus] D. Husemöller. *Elliptic Curves*. Graduate Texts in Mathematics. Springer New York, 2006.

[HWC] H. Hisil, K.K.H. Wong, and E. Carter G. and Dawson. "Twisted Edwards Curves Revisited". In: *Advances in Cryptology - ASIACRYPT 2008*. Ed. by J. Pieprzyk. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 326–343.

[Ibu] T. Ibukiyama. "On maximal orders of division quaternion algebras over the rational number field with certain optimal embeddings". In: *Nagoya Mathematical Journal* 88 (1982), pp. 181–195.

[IK] T. Ibukiyama and M. Kaneko. "Quadratic Forms and Ideal Theory of Quadratic Fields". In: *Bernoulli Numbers and Zeta Functions*. Tokyo: Springer Japan, 2014, pp. 75–93.

[Joy] M. Joye. "Highly Regular Right-to-Left Algorithms for Scalar Multiplication". In: *Cryptographic Hardware and Embedded Systems - CHES 2007*. Ed. by P. Paillier and I. Verbauwhede. Springer Berlin Heidelberg, 2007, pp. 135–147.

[JQ] M. Joye and J-J Quisquater. "Hessian Elliptic Curves and Side-Channel Attacks". In: *Cryptographic Hardware and Embedded Systems — CHES 2001*. Ed. by Ç.K. Koç, D. Naccache, and C Paar. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 402–410.

[JTV] M. Joye, M. Tibouchi, and D. Vergnaud. "Huff's Model for Elliptic Curves". In: *Algorithmic Number Theory*. Ed. by G. Hanrot and E. Morain F. and Thomé. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 234–250.

[Kan] M. Kaneko. "Supersingular j-invariants as singular moduli mod p". In: *Osaka Journal of Mathematics* 26.4 (1989), pp. 849–855.

[Ken] A. Kenku. *Atkin-Lehner involutions and class number residuality*. 1977.

[Kil] L.J.P. Kilford. *Modular Forms: A Classical and Computational Introduction*. Imperial College Press, 2008.

[KL] D. Kubert and S. Lang. *Modular Units*. Grundlehren der mathematischen Wissenschaften. Springer, 1981.

[KLPT] D. Kohel, K.E. Lauter, C. Petit, and J.P. Tignol. "On the quaternion ℓ-isogeny path problem". In: *LMS Journal of Computation and Mathematics* 17.A (2014), pp. 418–432.

[Klu] P. G. Kluit. "On the normalizer of $\Gamma_0(N)$". In: *Modular Functions of one Variable V*. Ed. by D.B. Serre J-P. and Zagier. Springer Berlin Heidelberg, 1977, pp. 239–246.

[KM] N.M. Katz and B. Mazur. *Arithmetic Moduli of Elliptic Curves*. Annals of Mathematics Studies. Princeton University Press, 1985.

[Kno] M.I. Knopp. *Modular Functions in Analytic Number Theory*. AMS Chelsea Publishing Series. AMS Chelsea Pub., American Mathematical Society, 2008.

[Kob] N. Koblitz. "Elliptic Curve Cryptosystems". In: *Mathematics of Computation* 48.177 (Jan. 1987), pp. 203–209.

[Kod1] I. Kodrnja. "Eta-quotients and Embeddings of $X_0(N)$ in the Projective Plane". In: *The Ramanujan Journal* 46 (2018).

[Kod2] I. Kodrnja. "On a simple model of $X_0(N)$". In: *Monatshefte für Mathematik* 186 (Apr. 2018), pp. 653–661.

[Koh1] D. Kohel. "Endomorphism rings of elliptic curves over finite fields". PhD Thesis. University of California at Berkeley, 1996.

[Koh2] D. Kohel. *Computing modular curves via quaternions*. Unpublished note based on a talk given at the fourth Computational Algebraic Number Theory (CANT) Conference, Sidney. Dec. 1997. url: `http://iml.univ-mrs.fr/~kohel/pub/sydney.pdf`.

[Koh3] D. Kohel. "Addition law structure of elliptic curves". In: *Journal of Number Theory* 131.5 (2011), pp. 894–919. url: `https://www.sciencedirect.com/science/article/pii/S0022314X10002672`.

[Koh4] D. Kohel. "The geometry of efficient arithmetic on elliptic curves". In: *Algorithmic Arithmetic, Geometry, and Coding Theory, AMS Contemporary Mathematics*. Vol. 637. 2015, pp. 95–110.

[Köh] G. Köhler. *Eta Products and Theta Series Identities*. Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelberg, 2011.

[Kul] Ravi S. Kulkarni. "An Arithmetic-Geometric Method in the Study of the Subgroups of the Modular Group". In: *American Journal of Mathematics* 113.6 (1991), pp. 1053–1133.

[Kup1] G. Kuperberg. "A Subexponential-Time Quantum Algorithm for the Dihedral Hidden Subgroup Problem". In: *SIAM Journal on Computing* 35.1 (July 2005), pp. 170–188.

[Kup2] G. Kuperberg. "Another Subexponential-time Quantum Algorithm for the Dihedral Hidden Subgroup Problem". In: *8th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2013)*. Ed. by S. Severini and F. Brandao. Vol. 22. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013, pp. 20–34.

[Lan1] S. Lang. *Elliptic Curves: Diophantine Analysis*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer-Verlag, 1978.

[Lan2] S. Lang. *Elliptic Functions*. Graduate texts in mathematics. Springer, 1987.

[Lan3] S. Lang. *Introduction to Modular Forms*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2001.

[LB] J. Love and D. Boneh. "Supersingular curves with small non-integer endomorphisms". In: *ANTS XIV - Proceedings of the Fourteenth Algorithmic Number Theory Symposium*. Vol. 4. Open Book Series 1. Mathematical Sciences Publishers, 2020, pp. 7–22.

[Lem] F. Lemmermeyer. *Binary quadratic forms; an elementary approach to the arithmetic of elliptic and hyperelliptic curves*. Online at `http://www.rzuser.uni-heidelberg.de/~hb3/publ/bf.pdf`. 2010.

[Len] H.W. Lenstra, Jr. "Complex Multiplication Structure of Elliptic Curves". In: *Journal of Number Theory* 56.2 (1996), pp. 227–241.

[Ler] R. Lercier. "Algorithmique des courbes elliptiques dans les corps finis". PhD Thesis. École Polytechnique, 1997.

[Lig1] G. Ligozat. "Courbes modulaires de genre 1". In: *Mémoires de la Sociéé Mathématique de France* 43 (1975), pp. 5–80.

[Lig2] G. Ligozat. "Courbes modulaires de niveau 11". In: *Modular Functions of one Variable V*. Ed. by J.P. Serre and D.B. Zagier. Berlin, Heidelberg: Springer Berlin Heidelberg, 1977, pp. 149–237.

[Liu] Q. Liu. *Algebraic Geometry and Arithmetic Curves*. Oxford Graduate Texts in Mathematics. OUP Oxford, 2006.

[Liu+] Z. Liu, J. Großschädl, Z. Hu, K. Järvinen, H. Wang, and I. Verbauwhede. "Elliptic Curve Cryptography with Efficiently Computable Endomorphisms and Its Hardware Implementations for the Internet of Things". In: *IEEE Transactions on Computers* 66.5 (2017), pp. 773–785.

[Lps] M. Berkelaar, K. Eikland, and P. Notebaert. *lp_solve, Open source (Mixed-Integer) Linear Programming system*. 2004. url: `http://lpsolve.sourceforge.net/5.5/`.

[LS] P-Y Liardet and N.P. Smart. "Preventing SPA/DPA in ECC Systems Using the Jacobi Form". In: *Proceedings of the Third International Workshop on Cryptographic Hardware and Embedded Systems*. CHES '01. Berlin, Heidelberg: Springer-Verlag, 2001, pp. 391–401.

[Mag] W. Bosma, J. Cannon, and C. Playoust. "The Magma algebra system. I. The user language". In: *Journal of Symbolic Computation* 24 (1997). Computational algebra and number theory (London, 1993). url: `http://magma.maths.usyd.edu.au/magma/`.

[Mai] R.S. Maier. "On Rationally Parametrized Modular Equations". In: *Journal of the Ramanujan Mathematical Society* 24.1 (2009), pp. 1–73.

[Man1] Y.I. Manin. *Vychislimoe i nevychislimoe (Computable and Non-computable)*. Moscow: Sovetskoe radio, 1980.

[Man2] Y.I. Manin. "Parabolic Points and Zeta-Functions of Modular Curves". In: (English Version:) Selected Papers. World Scientific, Singapore, 1996, pp. 268–290.

[Maz1] B. Mazur. "Courbes elliptiques et symboles modulaires". In: *Séminaire Bourbaki* 1971/72, exposés 400-417 (1973), pp. 277–294.

[Maz2] B. Mazur. "Modular curves and the Eisenstein ideal". In: *Publications Mathématiques de l'IHÉS* 47 (1977), pp. 33–186.

[Maz3] B. Mazur. "Rational points on modular curves". In: *Modular Functions of one Variable V*. Ed. by J.P. Serre and D.B. Zagier. Springer Berlin Heidelberg, 1977, pp. 107–148.

[Maz4] B. Mazur. "Rational Isogenies of Prime Degree." In: *Inventiones mathematicae* 44 (1978), pp. 129–162.

[Mcm1] K.W. Mcmurdy. "A Splitting Criterion for Galois Representations Associated to Exceptional Modular Forms". PhD Thesis. University of California, Berkeley, 2001.

[Mcm2] K.W. Mcmurdy. “Explicit Parametrizations of Ordinary and Supersingular Regions of $X_0(p^n)$”. In: *Modular Curves and Abelian Varieties*. Ed. by Quer J. Cremona J.E. Lario J.C. and Ribet K.A. Vol. 224. Progress in Mathematics. Birkhäuser, 2011, pp. 165–179.

[Mer] P. Mercuri. “Rational Points on Modular Curves”. PhD Thesis. University Sapienza, Roma, 2014.

[Mes] J.-F. Mestre. “La méthode des graphes. Exemples et applications”. In: *Proceedings of the international conference on class numbers and fundamental units of algebraic number fields (Katata, 1986)*. Nagoya University, 1986.

[MHM] R. Moloney, L. Hitt, and G. McGuire. *Division polynomials for twisted Edwards curves*. 2008. url: `https://arxiv.org/abs/0809.2182`.

[Mill] V.S. Miller. “Use of Elliptic Curves in Cryptography”. In: *Advances in Cryptology — CRYPTO ’85 Proceedings*. Ed. by H.C. Williams. Springer Berlin Heidelberg, 1986, pp. 417–426.

[Miln] J.S. Milne. *Elliptic Curves*. Kea books. BookSurge Publishers, 2006.

[Mir+] J. Miret, D. Sadornil, J. Tena, R. Tomas, and M. Valls. “Isogeny cordillera algorithm to obtain cryptographically good elliptic curves”. In: *Fifth Australasian Information Security Workshop (Privacy Enhancing Technologies) (AISW 2007)*. Ed. by L. Brankovic and C. Steketee. Vol. 68. CRPIT. Ballarat, Australia: ACS, 2007, pp. 153–157.

[Miy] T. Miyake. *Modular Forms*. Springer-Verlag Berlin, 1989.

[MM] L. Maino and C. Martindale. *An attack on SIDH with arbitrary starting curve*. Cryptology ePrint Archive, Paper 2022/1026. 2022. url: `https://eprint.iacr.org/2022/1026`.

[MMG] R. Moloney and G. McGuire. “Two Kinds of Division Polynomials For Twisted Edwards Curves”. In: *Applicable Algebra in Engineering Communication and Computing* 22 (Dec. 2011), pp. 321–345.

[Mom] F. Momose. “Rational Points on the Modular Curve $X_0^+(p^r)$”. In: *Journal of the Faculty of Science, University of Tokyo, Section 1A Mathematics* 33.3 (1986), pp. 441–466.

[Mon] P.L. Montgomery. “Speeding the Pollard and Elliptic Curve Methods of Factorization”. In: *Mathematics of Computation* 48.177 (1987), pp. 243–264.

[Moo1] D. Moody. *Divison Polynomials for Alternate Models of Elliptic Curves*. Cryptology ePrint Archive, Report 2010/630. `https://ia.cr/2010/630`. 2010.

[Moo2] D. Moody. “Division Polynomials for Jacobi Quartic Curves”. In: ISSAC ’11. San Jose, California, USA: Association for Computing Machinery, 2011, pp. 265–272.

[Mor] F. Morain. “Calcul du nombre de points sur une courbe elliptique dans un corps fini : aspects algorithmiques”. In: *Journal de Théorie des Nombres de Bordeaux* 7.1 (1995), pp. 255–282.

[MS] P. Mercuri and R. Schoof. “Modular forms invariant under non-split Cartan subgroups”. In: *Mathematics of Computation* 89 (Nov. 2019).

[Mur] N. Murabayashi. “On normal forms of modular curves of genus 2”. In: *Osaka J. Math.* 29.2 (1992), pp. 405–418.

[Neu] J. Neukirch. *Algebraische Zahlentheorie*. Masterclass. Springer Berlin Heidelberg, 2006.

[New1] S. Newberg. “Data of Modular Curves”. Master Thesis. California State University San Marcos, 2017.

[New2] M. Newman. “Construction and Application of a Class of Modular Functions”. In: *Proceedings of the London Mathematical Society* s3-7.1 (1957), pp. 334–350.

[NIST] National Institute of Standards and Technology. *Post-quantum cryptography standardization*. December 2016. url: `https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Post-Quantum-Cryptography-Standardization`.

[NX] H. Niederreiter and C. Xing. *Algebraic Geometry in Coding Theory and Cryptography*. USA: Princeton University Press, 2009.

[Ogg] A.P. Ogg. “Hyperelliptic modular curves”. In: *Bulletin de la Société Mathématique de France* 102 (1974), pp. 449–462.

[Ono] K. Ono. *The Web of Modularity: Arithmetic of the Coefficients of Modular Forms and q-series*. Vol. 102. CBMS Regional Conference Series in Mathematics. American Mathematical Society, 2004.

[Onu] H. Onuki. “On oriented supersingular elliptic curves”. In: *Finite Fields and Their Applications* 69 (2021).

[Pan] L. Panny. “Cryptography on Isogeny Graphs”. PhD Thesis. Technische Universiteit Eindhoven, 2021.

[Pet] C. Petit. “Faster Algorithms for Isogeny Problems Using Torsion Point Images”. In: *Advances in Cryptology – ASIACRYPT 2017*. Ed. by T. Takagi and T. Peyrin. Cham: Springer International Publishing, 2017, pp. 330–353.

[PH] S. Pohlig and M. Hellman. “An Improved Algorithm for Computing Logarithms over GF(p) and Its Cryptographic Significance”. In: *IEEE Transactions on Information Theory* 24.1 (Sept. 1978), pp. 106–110.

[Piz1] A.K. Pizer. “The action of the canonical involution on modular forms of weight 2 on $\Gamma_0(M)$”. In: *Mathematische Annalen* 226 (1977), pp. 99–116.

[Piz2] A.K. Pizer. “An algorithm for computing modular forms on $\Gamma_0(N)$”. In: *Journal of Algebra* 64.2 (1980), pp. 340–390.

[Piz3] A.K. Pizer. “Ramanujan graphs and Hecke operators”. In: *Bulletin (New Series) of the American Mathematical Society* 23.1 (1990), pp. 127–137.

[PL] C. Petit and K.E. Lauter. *Hard and Easy Problems for Supersingular Isogeny Graphs*. Cryptology ePrint Archive, Paper 2017/962. 2017. url: `https://eprint.iacr.org/2017/962`.

[Rad] H. Rademacher. *Topics in Analytic Number Theory*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 1973.

[Reg] O. Regev. *A Subexponential Time Algorithm for the Dihedral Hidden Subgroup Problem with Polynomial Space*. 2004. url: `https://arxiv.org/abs/quant-ph/0406151`.

[Rei] I. Reiner. *Maximal Orders*. London Mathematical Society monographs series: London Mathematical Society. Clarendon Press, 2003.

[Rob] D. Robert. *Breaking SIDH in polynomial time*. Cryptology ePrint Archive, Paper 2022/1038. 2022. url: `https://eprint.iacr.org/2022/1038`.

[Roh] D.F. Rohrlich. “Modular Curves, Hecke Correspondences, and L-Functions”. In: *Modular Forms and Fermat’s Last Theorem*. Ed. by G. Cornell, J.H. Silverman, and G. Stevens. Springer, 1997. Chap. III.

[RS] A. Rostovtsev and A. Stolbunov. *Public-key cryptosystem based on isogenies*. 2006. url: `https://eprint.iacr.org/2006/145`.

[RSA] R.L. Rivest, A. Shamir, and L. Adleman. “A Method for Obtaining Digital Signatures and Public-Key Cryptosystems”. In: *Communications of the ACM* 21 (1978), pp. 120–126.

[Sage] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 8.2)*. 2018. url: `http://www.sagemath.org`.

[SAS] M. Shi, A. Alahmadi, and P. Solé. *Codes and Rings: Theory and Practice*. Pure and Applied Mathematics. Elsevier Science, 2017.

[Sat] T. Satoh. “Generalized division polynomials”. In: *MATHEMATICA SCANDINAVICA* 94.2 (June 2004), pp. 161–184.

[SchB] B. Schöeneberg. *Elliptic Modular Functions: An Introduction*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1974.

[SchR] R. Schoof. “Nonsingular plane cubic curves over finite fields”. In: *Journal of Combinatorial Theory, Series A* 46.2 (1987), pp. 183–211.

[Ser1] J.P. Serre. *Cours d’arithmétique*. Presses Universitaires de France, Paris, 1970.

[Ser2] J.P. Serre. “Propriètès Galoisiennes des Points d’Ordre Fini des Courbes Elliptiques”. In: *Invent. Math.* 15 (Dec. 1971), pp. 259–331.

[Ser3] J.P. Serre. *Lectures on the Mordell-Weil Theorem*. Aspects of Mathematics. Amer Mathematical Society, 1997.

[Sha] I.G. Shafarevich. *Basic Algebraic Geometry 1*. 3rd ed. Springer Berlin Heidelberg, 2013.

[Shi] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Functions*. Kanô memorial lectures. Princeton University Press, 1971.

[Shm] M. Shimura. "Defining Equations of Modular Curves $X_0(N)$". In: *Tokyo Journal of Mathematics* 18.2 (Dec. 1995), pp. 443–456.

[Sho] P.W. Shor. "Algorithms for quantum computation: discrete logarithms and factoring". In: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*. 1994, pp. 124–134.

[Sil1] J.H. Silverman. *The Arithmetic of Elliptic Curves*. Graduate Texts in Mathematics. Springer New York, 2009.

[Sil2] J.H. Silverman. *Advanced Topics in the Arithmetic of Elliptic Curves*. Graduate Texts in Mathematics. Springer New York, 2010.

[Sma] N. P. Smart. "The Hessian Form of an Elliptic Curve". In: *Cryptographic Hardware and Embedded Systems — CHES 2001*. Ed. by Ç.K. Koç, D. Naccache, and C Paar. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 118–125.

[Sno] A. Snowden. *Lecture 14: Modular curves over Q*. University of Michigan. 2013. url: `http://www-personal.umich.edu/~asnowden/teaching/2013/679/L14.html`.

[Soh] G. Sohn. "Division polynomials for Huff's model of an elliptic curve". In: *Proceedings of the Jangjeon Mathematical Society* 16.1 (2013), pp. 115–124.

[Sta] K.E. Stange. *Quadratic forms, lattices, and ideal classes*. Notes online. University of Colorado. 2021. url: `http://math.colorado.edu/~kstange/teaching-resources/numthy/quad-forms-class-gp.pdf`.

[Ste1] W.A. Stein. *Modular Forms, a Computational Approach*. Graduate studies in mathematics. American Mathematical Society, 2007.

[Ste2] W.A. Stein. "The Modular Forms Database". In: (2012). `http://wstein.org/Tables`.

[Str1] M. Streng. *Notes on CM Division Polynomials*. Notes online at `https://www.math.leidenuniv.nl/~streng/divpol.pdf`.

[Str2] M. Streng. "Elliptic divisibility sequences with complex multiplication". Master's thesis. Utrecht University, 2006.

[Str3] M. Streng. "Divisibility sequences for elliptic curves with complex multiplication". In: *Algebra & Number Theory* 2.2 (2008), pp. 183–208.

[Sut1] A. Sutherland. *Isogeny kernels and division polynomials*. Course notes online at `https://math.mit.edu/classes/18.783/2017/LectureNotes6.pdf`.

[Sut2] A. Sutherland. *Modular polynomials*. Online Database at `https://math.mit.edu/~drew/ClassicalModPolys.html`.

[Sut3] A. Sutherland. *The modular equation*. Course notes online at `http://www.few.vu.nl/~sdn249/modularforms16/Notes.pdf`.

[Sut4] A.V. Sutherland. "Isogeny volcanoes". In: (2012). url: `http://arxiv.org/abs/1208.5370`.

[Tat] J. Tate. "Endomorphisms of Abelian Varieties over Finite Fields." In: *Inventiones mathematicae* 2 (1966), pp. 134–144.

[Tes] E. Teske. "The Pohlig–Hellman Method Generalized for Group Structure Computation". In: *Journal of Symbolic Computation* 27.6 (1999), pp. 521–534.

[TY1] J. Top and N. Yui. "Explicit Equations of Some Elliptic Modular Surfaces". In: *Rocky Mountain Journal of Mathematics* 37.2 (2007), pp. 663–687.

[TY2] F.T. Tu and Y. Yang. "Defining equations of $X_0(2^{2^n})$". In: *Osaka Journal of Mathematics* 46.1 (2009), pp. 105–113.

[Vél] J. Vélu. "Isogénies entre courbes elliptiques". In: *Comptes rendus hebdomadaires des séances de l'Académie des sciences: Sciences chimiques, Série A*. 273. July 1971, pp. 238–241.

[Ver1] H. Verrill. "Algorithm for Drawing Fundamental Domains". In: (2001). Online at `https://math.mit.edu/classes/18.783/2017/lectures.html`.

[Ver2] H. Verrill. *Fundamental Domain drawer*. 2001. url: `https://wstein.org/Tables/fundomain/`.

[Vig] M.F. Vignéras. *Arithmétique des Algèbres de quaternions*. Lecture notes in mathematics. Springer-Verlag, 1980.

[Voi] J. Voight. *Quaternion Algebras*. Graduate Texts in Mathematics. Springer-Verlag, 2021. url: `https://math.dartmouth.edu/~jvoight/quat-book.pdf`.

[Von] J.B. Vonk. *MATH 596: Topics in Algebra and Number Theory*. Preliminary notes online. McGill University. 2017. url: `https://pub.math.leidenuniv.nl/~vonkjb/teaching/2017_math596/PartialNotes.pdf`.

[Wan+] H. Wang, K. Wang, L. Zhang, and B. Li. "Pairing Computation on Elliptic Curves of Jacobi Quartic Form". In: *IACR Cryptol. ePrint Arch.* 2010 (2010), p. 475.

[War1] M. Ward. "Memoir on elliptic divisibility sequences". In: *Amer. J. Math.* 70 (1948), pp. 31–74.

[War2] M. Ward. "The law of repetition of primes in an elliptic divisibility sequence". In: *Duke Mathematical Journal* 15.4 (Dec. 1948), pp. 941–946.

[War3] M. Ward. "Arithmetical properties of polynomials associated with the lemniscate elliptic functions". In: *Proceedings of the National Academy of Sciences of the United States of America*. Vol. 36. June 1950, pp. 359–362.

[Was] L.C. Washington. *Elliptic Curves: Number Theory and Cryptography*. 2nd ed. Chapman & Hall/CRC, 2008.

[Wat1] W.C. Waterhouse. "Abelian varieties over finite fields". In: *Annales scientifiques de l'École Normale Supérieure* Série 4, Tome 2.4 (1969), pp. 521–560.

[Wat2] M.E. Watkins. "Class numbers of imaginary quadratic fields". In: *Mathematics of Computation* 73 (2004), pp. 907–938.

[Web] H. Weber. *Lehrbuch der algebra*. Vol. 3. Chelsea, 1961.

[Wes1] B. Wesolowski. "Orientations and the Supersingular Endomorphism Ring Problem". In: Trondheim, Norway: Springer-Verlag, 2022, pp. 345–371.

[Wes2] B. Wesolowski. "The supersingular isogeny path and endomorphism ring problems are equivalent". In: *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (2022), pp. 1100–1111.

[WF] H. Wu and R. Feng. "Elliptic curves in Huff's model". In: *Wuhan University Journal of Natural Sciences* 17 (2012), pp. 473–480.

[Yan1] Y Yang. "Transformation Formulas for Generalized Dedekind Eta Functions". In: *Bulletin of the London Mathematical Society* 36.5 (2004), pp. 671–682.

[Yan2] Y. Yang. "Defining equations of modular curves". In: *Advances in Mathematics* 204.2 (2006), pp. 481–508.

[Yu] J. Yu. "A Cuspidal Class Number Formula for the Modular Curves $X_1(N)$." In: *Mathematische Annalen* 252 (1980), pp. 197–216.

Glossary of notation

$\mathbb{A}^n$	Affine space of dimension n
$\mathbf{A}_K$	Ring of adèles of K
$\mathfrak{A}_{p,\infty}$	The quaternion algebra ramified at p and ∞
$\mathrm{Aut}(E)$	Automorphism group of the elliptic curve E
$\mathcal{C}\ell(\Delta)$	Form class group of discriminant Δ
$\mathcal{C}\ell(\mathcal{O})$	Class group of the order $\mathcal{O}$
$\mathbf{C}_K$	Idèle class group of K
$\mathrm{Cusps}(\Gamma)$	Set of cusps of the congruence subgroup Γ
Δ	Discriminant
$\mathrm{Div}(X)$	Group of divisors of X
$\mathrm{Div}^0(X)$	Group of degree zero divisors of X
$E_k(\tau)$	normalized Eisenstein series of degree k
$\mathcal{E}\ell\ell_K$	Set of isomorphism classes of elliptic curves over K
$\mathcal{E}\ell\ell(\mathcal{O})$	Set of isomorphism classes of elliptic curves over the ring class field of $\mathcal{O}$
$\mathrm{End}(E)$	Endomorphism ring of the elliptic curve E
$\mathrm{End}^0(E)$	Endomorphism algebra of the elliptic curve E
$\mathrm{EDiv}(X)$	Monoid of effective divisors of X
$\eta(\tau)$	Dedekind eta-function
$\mathfrak{f}, \mathfrak{f}_1, \mathfrak{f}_2$	Weber modular functions
$\mathbb{F}_q$	The finite field with q-elements
$\Gamma(1)$	The modular group
$\mathcal{G}a\ell(L/K)$	Galois group of the field extension L/K
$G_k(\Lambda)$	The Eisenstein series of weight k associate to the lattice Λ
$G_\ell(E)$	The ℓ-isogeny graph of the elliptic curve E
$G_\ell(E, \Gamma)$	The ℓ-isogeny graph of the elliptic curve E with Γ-level structure
$G_\ell(E, \iota)$	The ℓ-isogeny graph of the elliptic curve E with $\mathcal{O}$-orientation ι
$\mathrm{GL}_2(R)$	General linear group of degree 2 over the ring R
$\mathbb{H}$	The Poincare upper-half plane
$\mathbb{H}^*$	The extended upper-half plane
$H_\Delta, H_\mathcal{O}$	Hilbert class polynomial
$h(\Delta)$	Form class number of discriminant Δ
$h(\mathcal{O})$	Class number of the order $\mathcal{O}$
$\mathrm{Hom}(E_1, E_2)$	Group of isogenies between the elliptic curves E_1 and E_2
$\mathbf{I}_K$	Idèle group of K
$j, j_E, j(E)$	j-Invariant of an elliptic curve E
$\overline{K}$	Algebraic closure of the field K
$K[x_1, \ldots, x_n]$	Polynomial ring over K
$K[X]$	Coordinate ring of the variety X
$K(X)$	Function field of the variety X
K_X	Canonical divisor of the variety X
$\mathcal{K}_E$	Kummer line of E
$L(D)$	Riemann-Roch space associated to D
$\mathcal{L}$	Set of lattices over $\mathbb{C}$
$\mathcal{L}(D)$	Complete linear system of D

$\ell(D)$	Dimension of $L(D)$
$\mathrm{M}_2(R)$	Group of 2×2 matrices over the ring R
$\mathcal{O}$	Order in a quadratic imaginary field
$\mathfrak{O}$	Order in a quaternion algebra
$\mathbb{P}^n$	Projective space of dimension n
$\mathrm{Pic}(X)$	The Picard group of X
$\mathrm{Pic}^0(X)$	Picard group of degree-zero divisor classes on X
$\mathrm{Prin}(X)$	Subgroup of principal divisors of X
Φ_N	Classical modular polynomial of degree N
ψ_n	The n-th division polynomial
π, π_E, π_q	Frobenius endomorphism of an elliptic curve E over $\mathbb{F}_q$
$\wp$	The Weierstrass $\wp$-function
$\mathrm{PSL}_2(R)$	Projective special linear group of degree 2 over the ring R
$\mathrm{SL}_2(R)$	Special linear group of degree 2 over the ring R
$\mathrm{SS}(p)$	Set of isomorphism classes of supersingular elliptic curves over $\overline{\mathbb{F}}_p$
$\mathrm{SS}_{\mathcal{O}}(p)$	Isomorphism classes of $\mathcal{O}$-oriented supersingular elliptic curves over $\overline{\mathbb{F}}_p$
$\mathrm{SS}_{\mathcal{O}}^{pr}(p)$	Primitive $\mathcal{O}$-oriented supersingular elliptic curves over $\overline{\mathbb{F}}_p$
$\mathrm{SS}_{\mathcal{O}}^{pr}(\rho)$	Primitive $\mathcal{O}$-oriented supersingular elliptic curves over $\overline{\mathbb{F}}_p$ with p-orientation ρ
$T_\ell(E)$	The ℓ-adic Tate module attached to the elliptic curve E
$X(K)$	Set of K-rational point of X
$X(\Gamma)$	Compact modular curve of level structure Γ
$X_0^+(N)$	Atkin-Lehner quotient of $X_0(N)$
$X_{ns}(N)$	Modular curve associated to a non-split Cartan subgroup of level N
$X_{ns}^+(N)$	Modular curve associated to the normalizer of a non-split Cartan group of level N
$Y(\Gamma)$	Affine modular curve of level structure Γ
$\mathcal{W}_n$	Weber modular curve of
ω_n	Atkin-Lehner involution of degree n
ω_E	Invariant differential on the elliptic curve E
Ω_E	Sheaf of differentials on the elliptic curve E
Ω_X^1	Line bundle of holomorphic differentials on X

Appendices

A Models for modular curves

Models for $X_0(2^N)$

Model for $X_0(2)$

$X_0(2)$ is curve of genus 0. We get

$$A(2) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

we look for the easiest parameter (the one with a single pole at ∞) by solving

$$\begin{cases} 2a_1 + a_2 \geq 0 \\ a_1 + 2a_2 = -24 \\ \textbf{(i) } a_1 + a_2 = 0 \\ \textbf{(ii) } a_1 + 2a_2 \equiv 0 \mod 24 \\ \textbf{(iii) } 2a_1 + a_2 \equiv 0 \mod 24 \\ \textbf{(iv) } a_1 + 2a_2 \equiv 0 \mod 2 \end{cases}$$

We obtain the parameter

$$t_2 = \left(\frac{\eta(\tau)}{\eta_2(\tau)}\right)^{24} = q^{-1} - 24 + 276q - 2048q^2 + 11202q^3 + \dots$$

which has divisor

$$(t_2) = (0) - (\infty)$$

From the divisor it is easy to observe that $t_2(0) = 0$ and $t_2(\infty) = \infty$ and these are called coordinates of t_2 at the cusps.

We have two maps $\pi_1, \pi_2 : X_0(2) \to X(1)$ both of degree 3. An element on $X_0(2)$ can be described either as a pair (E, C) of an elliptic curve together with a cyclic subgroup of order 2 or as a cyclic isogeny $E \to F$. In the first case, the map π_1 sends (E, C) to E (forgetful map) and π_2 sends (E, C) to E/C. On the other hand, π_1 sends $E \to F$ to E and π_2 sends the same isogeny to F. Clearly this two description are equivalent.

We are interested on the action of these two maps on the parameters of $X_0(2)$ and $X(1)$, namely t_2 and the j-function. Thus, we look for algebraic relations between t_2 and j. In particular, we would like to express $j = j(t_2)$ as a function in t_2. For this reason, we look for a relation of degree 1 in j; we obtain the polynomial

$$P(X, Y) = XY^2 - Y^3 - 768Y^2 - 196608Y - 16777216 = 0$$

which has solution (j, t_2).

Inverting the relations with respect to j, we conclude that

$$\pi_1^*(j) = \frac{t_2^3 + 768t_2^2 + 196608t_2 + 16777216}{t_2^2} = \frac{(t_2 + 256)^3}{t_2^2}$$

Concerning the second map π_2 we have to compare the q-expansion of t_2 with the q-expansion of $j(q^2)$

since π_2 has the effect $q \to q^2$ on the canonical q-expansion at ∞. We find

$$\pi_2^*(j) = \frac{(t_2 + 16)^3}{t_2}$$

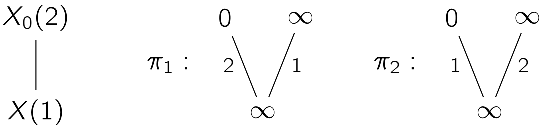
For completeness, we will also study the formula for the Atkin-Lehner involution ω_2. We have seen that ω_2 simply switches the cusps (0) and (∞). Thus $\omega_2(t_2) = \frac{c}{t_2}$ for some constant c. In order to recover it, we recall that

$$\pi_1 \circ \omega_2 = \pi_2 \Longrightarrow \omega_2^* \circ \pi_1^* = \pi_2^*$$

and substituting the expressions that we already know, we get

$$\omega_2(t_2) = \frac{4096}{t_2}$$

We conclude with a picture describing the ramification of the forgetful maps at the cusps



Model for $X_0(4)$

Let's take another step up in the tower $X_0(2^n)$. $X_0(4)$ is again a curve of genus 0 and

$$A(4) = \begin{pmatrix} 4 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

Solving the linear integer programming problem for $\delta = 1$ yields

$$t_4 = \left(\frac{\eta(\tau)}{\eta_4(\tau)}\right)^8 = q^{-1} - 8 + 20q - 62q^3 + 216q^5 - 641q^7 + \dots$$

with divisor

$$(t_4) = (0) - (\infty)$$

The coordinates at 0 and ∞ are easily deduced from the divisor: $t_4(0) = 0$ and $t_4(\infty) = \infty$. The coordinates at $\frac{1}{2}$ are slightly more difficult to find. We can compute them starting from a different parameter whose divisor is more helpful: for instance, we can look for a modular function on $X_0(4)$ with divisor $(\frac{1}{2})-(\infty)$. We find

$$t_4' = \frac{\eta_2(\tau)^{24}}{\eta(\tau)^8 \eta_4(\tau)^{16}} \qquad (t_4') = (\frac{1}{2}) - (\infty)$$

We can easily find, using `magma`, that $t_4 = t_4' - 16$ from which

$$t_4\left(\frac{1}{2}\right) = t_4'\left(\frac{1}{2}\right) - 16 = -16$$

We have two maps $\pi_1, \pi_2 : X_0(4) \to X_0(2)$. We can think of points on $X_0(4)$ both as elliptic curves with a cyclic subgroup of order 4 or as sequence of 2-isogenies of length 2. In the first setting, the map π_1 sends (E, C) to $(E, C[2])$ while π_2 sends (E, C) to $(E/C[2], C/C[2])$. If we think at the isogeny chain $E \to F \to G$, π_1 outputs the first isogeny $E \to F$ while π_2 gives the second one $F \to G$.

At this point, it appears clear that, in order to obtain an explicit description of π_1^* we have to compare the q-expansions of t_4 and t_2 while π_2^* comes from the comparison of the expansions of t_4 and $t_2(q^2)$.

We find

$$\pi_1^*(t_2) = \frac{t_4^2}{t_4 + 16} \qquad \pi_2^*(t_2) = t_4(t_4 + 16)$$

Note that there are also 3 maps $\pi_1, \pi_2, \pi_4 : X_0(4) \to X(1)$ which give the three elliptic curves involved in the isogeny chain. From the comparison between $t_4(q)$ and $j(q)$, $j(q^2)$, $j(q^4)$ respectively, we obtain

$$\pi_1^*(j) = \frac{t_4^6 + 768t_4^5 + 208896t_4^4 + 23068672t_4^3 + 855638016t_4^2 + 12884901888t_4 + 68719476736}{t_4^5 + 16t_4^4}$$

$$\pi_2^*(j) = \frac{t_4^6 + 48t_4^5 + 1536t_4^4 + 28672t_4^3 + 393216t_4^2 + 3145728t_4 + 16777216}{t_4^4 + 32t_4^3 + 256t_4^2}$$

$$\pi_4^*(j) = \frac{t_4^6 + 48t_4^5 + 816t_4^4 + 5632t_4^3 + 13056t_4^2 + 12288t_4 + 4096}{t_4^2 + 16t_4}$$

If we simplify the expressions we get

$$\pi_1^*(j) = \frac{(t_4^2 + 256t_4 + 4096)^3}{t_4^4(t_4 + 16)} \qquad \pi_2^*(j) = \frac{(t_4^2 + 16t_4 + 256)^3}{t_4^2(t_4^2 + 32t_4 + 256)} \qquad \pi_4^*(j) = \frac{(t_4^2 + 16t_4 + 16)^3}{t_4(t_4 + 16)}$$

and we can immediately see that they correspond to the compositions

$$\pi_1 : X_0(4) \xrightarrow{\pi_1} X_0(2) \xrightarrow{\pi_1} X(1)$$

$$\pi_2 : X_0(4) \xrightarrow{\pi_1} X_0(2) \xrightarrow{\pi_2} X(1) = X_0(4) \xrightarrow{\pi_2} X_0(2) \xrightarrow{\pi_1} X(1)$$

$$\pi_4 : X_0(4) \xrightarrow{\pi_2} X_0(2) \xrightarrow{\pi_2} X(1)$$

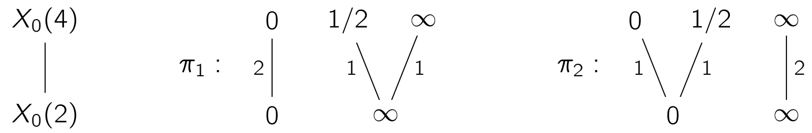
Remark. The equality in the second equation can be explained intuitively: π_2 associates to $E \to F \to G$, the second elliptic curve F. Since F is the central curve, it is involved in both the 2-isogenies. What the second line is saying is simply that the second elliptic curve of the first 2-isogeny is equal to the first elliptic curve of the second 2-isogeny.

Finally, we study Atkin-Lehner involutions. Since $4 = \ell^2$ there is only one such involution, namely ω_4. This switches the cusps 0 and ∞ meaning that it is of the form c/t_4 for some constant c. From lemma 2.44, we know that $\pi_1 \circ \omega_4 = \omega_2 \circ \pi_2$ where $\pi_1, \pi_2 : X_0(4) \to X_0(2)$, $\omega_2 : X_0(2) \to X_0(2)$ and $\omega_4 : X_0(4) \to X_0(4)$ and we get

$$\omega_4^*(t_4) = \frac{256}{t_4}$$

Remark. The Atkin-Lehner involution has the effect of switching the order of the isogeny (chain) as seen at the end of section 2.2.5

As before, we conclude the section with the ramification diagrams



Remark. The ramification diagram for $X_0(4) \to X(1)$ can be obtained by the multiplicativity of the ramification indices.

Model for $X_0(8)$

The curve $X_0(8)$ is also of genus 0. The associated matrix is

$$A(8) = \begin{pmatrix} 8 & 4 & 2 & 1 \\ 2 & 4 & 2 & 1 \\ 1 & 2 & 4 & 2 \\ 1 & 2 & 4 & 8 \end{pmatrix}$$

Once again, we look for a modular function having poles only at infinity and we find

$$t_8 = \frac{\eta(\tau)^4\eta_4(\tau)^2}{\eta_2(\tau)^2\eta_8(\tau)^4} = q^{-1} - 4 + 4q + 2q^3 - 8q^5 - q^7 + 20q^9 + \dots \qquad (t_8) = (0) - (\infty)$$

There are 4 cusps on $X_0(8)$. The coordinates of t_8 at 0 and ∞ are

$$t_8(0) = 0 \qquad \text{and} \qquad t_8(\infty) = \infty$$

Concerning the cusp $\mathfrak{c}_2 = 1/2$ we compare t_8 with

$$t_8' = \frac{\eta_2(\tau)^{10}}{\eta(\tau)^4\eta_4(\tau)^2\eta_8(\tau)^4}$$

which has divisor $(t_8') = (1/2) - (\infty)$. We find that $t_8 - t_8' + 8 = 0$ from which $t_8(1/2) = t_8'(1/2) - 8 = -8$.

Finally, for $\mathfrak{c}_4 = 1/4$ we use the η-quotient

$$t_8'' = \frac{\eta_4(\tau)^{12}}{\eta_2(\tau)^4\eta_8(\tau)^8} \qquad\qquad (t_8'') = (1/4) - (\infty)$$

which is related to t_8 by the relation $t_8 - t_8'' + 4 = 0$. Thus $t_8''(1/4) = -4$.

Contrary to what we did for $X_0(4)$ we will not study all the maps $X_0(8) \to X(1)$ or $X_0(8) \to X_0(2)$ since they start becoming too many. More importantly, they can be easily deduced by composition.

We will then study only $\pi_1, \pi_2 : X_0(8) \to X_0(4)$. The description of these maps is similar to the one we explored before. Starting from a sequence of 2-isogenies $E_0 \to E_1 \to E_2 \to E_3$ (a point on $X_0(8)$), these maps select the only two subsequences of length 2, respectively $E_0 \to E_1 \to E_2$ and $E_1 \to E_2 \to E_3$.

By comparing the q-expansions of t_8 and t_4 we find

$$\pi_1^*(t_4) = \frac{t_8^2}{t_8 + 4}$$

Now we observe that the two sequences above $E_0 \to E_1 \to E_2$ and $E_1 \to E_2 \to E_3$ are 2-isogenous meaning that the explicit form of π_2 can be found comparing $t_8(q)$ with $t_4(q^2)$ where the exponent 2 encodes indeed the 2-isogeny relation

$$\pi_2^*(t_4) = t_8(t_8 + 4)$$

Finally, we study the Atkin-Lehner involution. We can observe that the only involution on $X_0(8)$ is ω_8. On the cusps, this has the effect of switching 0 with ∞ and 1/2 with 1/4. From the divisor of t_8, we see once again that ω_8 must be of the form c/t_8. Thanks to the compatibility relations we obtain

$$\omega_8^*(t_8) = \frac{32}{t_8}$$

$$\begin{tikzcd}
X_0(8) \arrow[d, dash] \\
X(4)
\end{tikzcd}
\qquad
\pi_1 : \begin{tikzcd}
0 \arrow[d, dash, "2"'] & 1/2 \arrow[d, dash, "2"'] & 1/4 \arrow[dr, dash, "1"'] & & \infty \arrow[dl, dash, "1"] \\
0 & 1/2 & & \infty &
\end{tikzcd}
\qquad
\pi_2 : \begin{tikzcd}
0 \arrow[dr, dash, "1"'] & & 1/2 \arrow[dl, dash, "1"] & 1/4 \arrow[d, dash, "2"] & \infty \arrow[d, dash, "2"] \\
& 0 & & 1/2 & \infty
\end{tikzcd}$$

Remark. The first diagram can be obtained just by observing how the cusps are conjugated under $\Gamma_0(4)$. For the second one, we can use once again the compatibility relations.

Model for $X_0(16)$

The curve $X_0(16)$ has genus 0.

$$A(16) = \begin{pmatrix} 16 & 8 & 4 & 2 & 1 \\ 4 & 8 & 4 & 2 & 1 \\ 1 & 2 & 4 & 2 & 1 \\ 1 & 2 & 4 & 8 & 4 \\ 1 & 2 & 4 & 8 & 16 \end{pmatrix}$$

We select the parameter

$$t_{16} = \frac{\eta(\tau)^2\eta_8(\tau)}{\eta_2(\tau)\eta_{16}(\tau)^2} \qquad\qquad (t_{16}) = (0) - (\infty)$$

The coordinates are listed in the following table

Cusp $\mathfrak{c}$	Coordinates $t_{16}(\mathfrak{c})$	Eta-products with divisor $(\mathfrak{c})-(\infty)$
0	0	t_{16}
1/2	-4	$\dfrac{\eta_2(\tau)^5\eta_8(\tau)}{\eta(\tau)^2\eta_4(\tau)^2\eta_{16}^2}$
1/4 and 3/4	Galois conjugates; roots of $t_{16}^2+4t_{16}+8=0$	$\dfrac{\eta_4(\tau)^{10}}{\eta_2(\tau)^4\eta_8(\tau)^2\eta_{16}(\tau)^4}$ $(\mathfrak{c})-2(\infty)$
1/8	-2	$\dfrac{\eta_8(\tau)^6}{\eta_4(\tau)^2\eta_{16}^4}$
∞	∞	t_{16}

We encountered a problem when studying cusps 1/4 and 3/4, namely that there is no η-quotients with divisor $(\mathfrak{c})-(\infty)$. Trying to increase δ, we find $t'_{16}=\frac{\eta_4(\tau)^{10}}{\eta_2(\tau)^4\eta_8(\tau)^2\eta_{16}(\tau)^4}$. The drawback is that now the algebraic relation with t_{16} cannot be linear anymore. Since the pole at ∞ is of order two we get a polynomial of degree 2 in t_{16}. Hence, coordinates at these cusps are given as roots of a quadratic polynomial.

Finally, we give a description for the maps $\pi_1,\pi_2 : X_0(16)\to X_0(8)$.

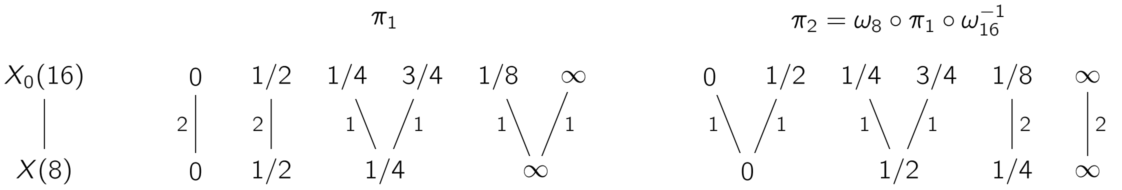
$$\pi_1^*(t_8)=\frac{t_{16}^2}{t_{16}+2}$$

$$\pi_2^*(t_8)=t_{16}(t_{16}+4)$$

Once again, since t_{16} has a simple divisor and ω_{16} switches 0 and ∞, the Atkin-Lehner involution has a simple form:

$$\omega_{16}=\frac{8}{t_{16}}$$

Here the ramification diagrams for $X_0(16)\to X_0(8)$:



Model for $X_0(32)$

The situation changes when we pass to the next level of the modular chain $X_0(2^n)$. $X_0(32)$ is a curve of genus 1, an elliptic curve. This means that, using Theorem 2.53, we will look for two parameters $x=x_{32}$ and $y=y_{32}$ with coprime pole divisor degrees.

Let's start by looking at the matrix

$$A(32)=\begin{pmatrix}32&16&8&4&2&1\\8&16&8&4&2&1\\2&4&8&4&2&1\\1&2&4&8&4&2\\1&2&4&8&16&8\\1&2&4&8&16&32\end{pmatrix}$$

Since we expect $X_0(32)$ to be an elliptic curve we want x to have order 2 at infinity and y to have order 3. We find

$$x=\frac{\eta_2(\tau)^2\eta_{16}(\tau)}{\eta_4(\tau)\eta_{32}(\tau)^2}\qquad (x)=(0)+(1/2)-2(\infty)$$

$$y=\frac{\eta(\tau)^2\eta_{16}(\tau)^4}{\eta_2(\tau)\eta_8(\tau)\eta_{32}(\tau)^4}\qquad (x)=2(0)+(1/16)-3(\infty)$$

Thanks to the particular choice of parameters we find a very special algebraic relation between x and y:

$$x^3 + 2x^2 - 4xy - y^2 - 8y = 0 \quad \Rightarrow \quad y^2 + 4xy + 8y = x^3 + 2x^2$$

in which we easily recognize an elliptic curve. The Weierstrass model for $X_0(32)$ is therefore

$$X_0(32) : y^2 = x^3 + 5184x$$

This elliptic curve has j-invariant 1728.

Cusp $\mathfrak{c}$	Coordinates $(x_{32}(\mathfrak{c}), y_{32}(\mathfrak{c}))$	Eta-products used for comparison
0	$(0, 0)$	(x_{32}, y_{32})
1/2	$(0, -8)$	(x_{32}, x_{32})
1/4 and 3/4	Galois conjugates; roots of $(x^3 + 64\ ,\ y^2 - 8y + 32)$	$\dfrac{\eta_4(\tau)^5\eta_{16}(\tau)}{\eta_2(\tau)^2\eta_8(\tau)^2\eta_{32}(\tau)^2} \quad (\mathfrak{c}_4) - 2(\infty)$
1/8 and 3/8	Galois conjugates; roots of $(x^2 + 4x + 8 = 0\ ,\ y^2 + 16)$	$\dfrac{\eta_8(\tau)^4\eta_{16}(\tau)^2}{\eta_4(\tau)^2\eta_{32}(\tau)^4} \quad (\mathfrak{c}_8) + (1/16) - 3(\infty)$
1/16	$(-2, 0)$	(y_{32}, y_{32})
∞	(∞, ∞)	(x_{32}, y_{32})

The next step is to study the maps $\pi_1, \pi_2 : X_0(32) \to X_0(16)$. Their pullback will be rational functions in x and y: $\pi^*(t_16) \in \mathbb{Q}(x_{32}, y_{32})$.

We compare the q-expansions of $x_{32}(q)$ and $y_{32}(q)$ with the ones of $t_{16}(q)$ and $t_{16}(q^2)$ and we get

$$\pi_1^*(t_{16}) = \frac{y}{x+2} \qquad \pi_2^*(t_{16}) = x$$

Remark. The degrees seems to agree with our expectations since x has a pole of order 2 at infinity and so does $t_{16}(q^2)$. At the same time $t_{16}(q)$ has a single pole at ∞ and the same happens for y/x.

We can easily verify that these maps have the correct degree:

```
> E:=EllipticCurve([4,2,8,0,0]);  E;
Elliptic Curve defined by y^2 + 4*x*y + 8*y = x^3 + 2*x^2 over Rational Field
> F<x,y>:=FunctionField(E);  F;
Function Field of Elliptic Curve defined by y^2 + 4*x*y + 8*y = x^3 + 2*x^2 over
Rational Field
> MinimalPolynomial(y);
$.1^2 + (4*x + 8)*$.1 - x^3 - 2*x^2
```

this recovers the equation of the elliptic curve.

```
> u:=y/(x+2);
> Degree(u);
2
> v:=x;
> Degree(v);
2
```

therefore showing that π_1 and π_2 have indeed degree 2.

Finally, we try to make explicit the Atkin-Lehner involution ω_{32}. Observe that its pull-back is of the form

$$\omega_{32}^*(x_{32}, y_{32}) = ((g_1(x_{32}, y_{32}).g_1(x_{32}, y_{32})))$$

The main remark is that ω_{32} switches the cusps 0 and ∞, 1/2 and 1/16, 1/4 and 3/8, 3/4 and 1/8.

Thus, in order to find $g_1(x, y)$ we look for a modular function on $X_0(32)$ whose divisor equals the divisor of x_{32} acted on by ω_{32}, i.e., we look for $x'(q)$ such that $(x') = -2(0) + (1/16) + (\infty)$. Concerning $g_2(x, y)$, we need a sort of inverse of y, namely a modular function y' with divisor $(y') = -3(0) + (1/2) + 2(\infty)$.

Solving the linear integer programming problem we find

$$x' = \frac{\eta_2(\tau)\eta_{16}(\tau)^2}{\eta(\tau)^2\eta_8(\tau)} \qquad (x') = -2(0) + (1/16) + (\infty)$$

$$y' = \frac{\eta_2(\tau)^4\eta_{32}(\tau)^2}{\eta(\tau)^4\eta_4(\tau)\eta_{16}(\tau)} \qquad (y') = -3(0) + (1/2) + 2(\infty)$$

We find

$$g_1(x, y) = c_1\frac{x+2}{y} \qquad \text{and} \qquad g_2(x, y) = c_2\frac{4x+y+8}{xy}$$

Remark. Note that we could somehow predict the shape of g_1 using the compatibility relations since $\pi_2(t_{16}) = x$.

Once again, the constants c_1 and c_2 are recovered using the same relations.

$$\begin{tikzcd}
X_0(32) \arrow[r, "\omega_{32}"] \arrow[d, "\pi_1"'] \arrow[dr, phantom, "\circlearrowleft"] & X_0(32) \arrow[d, "\pi_2"] \\
X_0(16) \arrow[r, "\omega_{16}"'] & X_0(16)
\end{tikzcd}$$

Then $\pi_2 \circ \omega_{32} = \omega_{16} \circ \pi_1$ from which $\omega_{32}^* \circ \pi_2^* = \pi_1^* \circ \omega_{16}^*$. Now

$$\omega_{32}^*\left(\pi_2^*\left(t_{16}\right)\right) = \pi_1^*\left(\omega_{16}^*\left(t_{16}\right)\right)$$

$$\omega_{32}^*\left(x\right) = \pi_1^*\left(\frac{8}{t_{16}}\right)$$

$$c_1\frac{x+2}{y} = \frac{8}{\frac{y}{x+2}}$$

We conclude that $c_1 = 8$.

In order to find the other constant, we use again compatibility relations but with a different order:

$$\pi_1 \circ \omega_{32} = \omega_{16} \circ \pi_2 \implies \omega_{32}^* \circ \pi_1^* = \pi_2^* \circ \omega_{16}^*$$

this means

$$\omega_{32}^*\left(\pi_1^*(t_{16})\right) = \pi_2^*\left(\omega_{16}^*(t_{16})\right)$$

$$\omega_{32}^*\left(\frac{y}{x+2}\right) = \pi_2^*\left(\frac{8}{t_{16}}\right)$$

$$\frac{g_2(x, y)}{g_1(x, y) + 2} = \frac{8}{x}$$

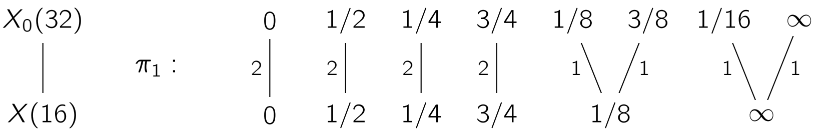
$$c_2\frac{4x+y+8}{xy}\,\frac{y}{8x+16+2y} = \frac{8}{x}$$

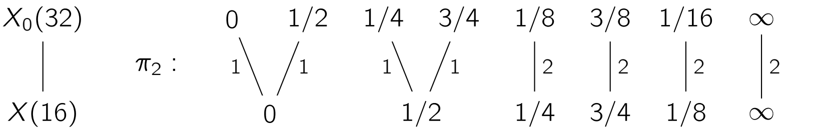
$$c_2\frac{4x+y+8}{8x+2y+16} = 8 \implies \frac{c_2}{2} = 8 \implies c_2 = 16$$

We conclude that

$$\omega_{32}^*(x, y) = \left(\frac{8(x+2)}{y}, \frac{16(4x+y+8)}{xy}\right)$$

Remark. It is sometimes useful to study the action of the Atkin-Lehner involution on the equation defining the curve. This might help to find the constants. For an example we refer to [Mcm2, End of §2]





Remark. For π_1 it suffices to study the inequivalent cusps of $X_0(32)$ that become equivalent in $X_0(16)$, i.e., they are in the same $\Gamma_0(16)$-orbit.

Model for $X_0(64)$

The last example will be $N = 2^6$. From the table at the end of Section 2.1.3, $X_0(64)$ is a genus 3 curve.

Theorem A.3 ([Maz1])**.** *Let $X_0(N)$ be the classical modular curve.*

- *It has genus 1 if and only if $N = 11, 14, 15, 17, 19, 20, 21, 24, 27, 32, 36, 49$.*
- *It has genus 2 if and only if $N = 22, 23, 26, 28, 29, 31, 37, 50$.*

Solving the linear problem associated to the matrix

$$A = \begin{pmatrix} 64 & 32 & 16 & 8 & 4 & 2 & 1 \\ 16 & 32 & 16 & 8 & 4 & 2 & 1 \\ 4 & 8 & 16 & 8 & 4 & 2 & 1 \\ 1 & 2 & 4 & 8 & 4 & 2 & 1 \\ 1 & 2 & 4 & 8 & 16 & 8 & 4 \\ 1 & 2 & 4 & 8 & 16 & 32 & 16 \\ 1 & 2 & 4 & 8 & 16 & 32 & 64 \end{pmatrix}$$

we find the two generators of the function field of $X_0(64)$:

$$x_{64} = \frac{\eta_{16}(\tau)^2\eta_{32}(\tau)}{\eta_8(\tau)\eta_{64}(\tau)^2} \qquad (x_{64}) = \sum_{i=1,3}(i/16) + (1/32) - 3(\infty)$$

$$y_{64} = \frac{\eta_{32}(\tau)^6}{\eta_{16}(\tau)^2\eta_{64}(\tau)^4} \qquad (y_{64}) = 4(1/32) - 4(\infty)$$

By looking at the behavior of x_{64} and y_{64}, we find a model for $X_0(64)$

$$x^4 - y^3 - 4y = 0$$

Cusp $\mathfrak{c}$	Coordinates $(x_{32}(\mathfrak{c}), y_{32}(\mathfrak{c}))$	Eta-products used for comparison
0	$(2, 2)$	$\dfrac{\eta_4(\tau)^2\eta_{32}(\tau)}{\eta_8(\tau)\eta_{64}(\tau)^2}$ $(0) + (1/2) + (\mathfrak{c}_4) - 4(\infty)$
$1/2$	$(2, 2)$	$\dfrac{\eta_4(\tau)^2\eta_{32}(\tau)}{\eta_8(\tau)\eta_{64}(\tau)^2}$ $(0) + (1/2) + (\mathfrak{c}_4) - 4(\infty)$
$1/4$ and $3/4$	$(2, 2)$	$\dfrac{\eta_4(\tau)^2\eta_{32}(\tau)}{\eta_8(\tau)\eta_{64}(\tau)^2}$ $(0) + (1/2) + (\mathfrak{c}_4) - 4(\infty)$
$1/8$, $3/8$, $5/8$, $7/8$	Galois conjugates with $y = -2$ and $x^4 - 4x + 8 = 0$	$\dfrac{\eta_8(\tau)^5\eta_{32}(\tau)}{\eta_4(\tau)^2\eta_{16}(\tau)^2\eta_{64}(\tau)^2}$ $(\mathfrak{c}_8) - 4(\infty)$
$1/16$ and $3/16$	$(0, 0)$	(x_{32}, x_{32})
$1/32$	$(0, 0)$	(x_{32}, y_{32})
∞	(∞, ∞)	(x_{32}, y_{32})

As usual, we have two maps $\pi_1, \pi_2 : X_0(64) \to X_0(32)$ defined by

$$\pi_1^*(x_{32}, y_{32}) = \left(\frac{x_{64}^2 - 2y_{64}}{y_{64}}, \frac{-2x_{64}^2 + x_{64}y_{64} + 2x_{64}}{y_{64}}\right)$$

$$\pi_2^*(x_{32}, y_{32}) = \left(y_{64} - 2, x_{64}^2 - 2y_{64}\right)$$

We also deduce that

$$\omega_{64}^*(x_{64}, y_{64}) = \left(c_1\frac{2 - y_{64}}{y_{64} - 2x_{64} + 2}, c_2\frac{2x_{64} + y_{64} + 2}{y_{64} - 2x_{64} + 2}\right)$$

From the compatibility relation $\omega_{64}^* \circ \pi_2^* = \pi_1^* \circ \omega_{32}^*$ we find

$$\omega_{64}^*\left(\pi_2^*(x_{32})\right) = \pi_1^*\left(\omega_{32}^*(x_{32})\right)$$

$$\omega_{64}^*\left(y - 2\right) = \pi_1^*\left(\frac{x_{32} + 16}{y_{32}}\right)$$

$$c_2\frac{2x + y + 2}{y - 2x + 2} - 2 = \frac{8x^2}{-2x^2 + xy + 2x} \implies c_2(2x + y + 2) + 4x - 2y - 4 = 8x \implies c_2 = 2$$

In the same way we find

$$\begin{array}{ccc}
\omega_{64}^*\left(\pi_2^*(x_{32})\right) & = & \pi_1^*\left(\omega_{32}^*(x_{32})\right) \\
\| & & \| \\
\omega_{64}^*\left(x^2 - 2y\right) & & \pi_1^*\left(\dfrac{64x_{32} + 16y_{32} + 128}{x_{32}y_{32}}\right) \\
\| & & \| \\
c_1^2\dfrac{(2 - y)^2}{(y - 2x + 2)^2} - 4\dfrac{y + x + 2}{y - 2x + 2} & & \dfrac{\frac{64x^2 - 128y}{y} + \frac{-32x^2 + 16xy + 32x}{y} + 128}{\frac{x^2 - 2y}{y} \cdot \frac{-2x^2 + xy + 2x}{y}} \\
\| & & \| \\
c_1^2\dfrac{(2 - y)^2}{(y - 2x + 2)^2} - 4\dfrac{y + x + 2}{y - 2x + 2} & = & \dfrac{32x^2y + 16xy^2 + 32xy}{(x^2 - 2y)(-2x^2 + xy + 2x)}
\end{array}$$

Thus,

$$c_1^2\frac{(2 - y)^2}{y - 2x + 2} = 4(y + x + 2) + y\frac{32x + 16y + 32}{x^2 - 2y}$$

$$c_1^2\frac{(2 - y)^2}{y - 2x + 2} = \left(4 + \frac{16y}{x^2 - 2y}\right)(y + 2x + 2)$$

$$c_1^2\frac{(2 - y)^2}{y - 2x + 2} = \frac{4(x^2 + 2y)}{x^2 - 2y}(y + 2x + 2)$$

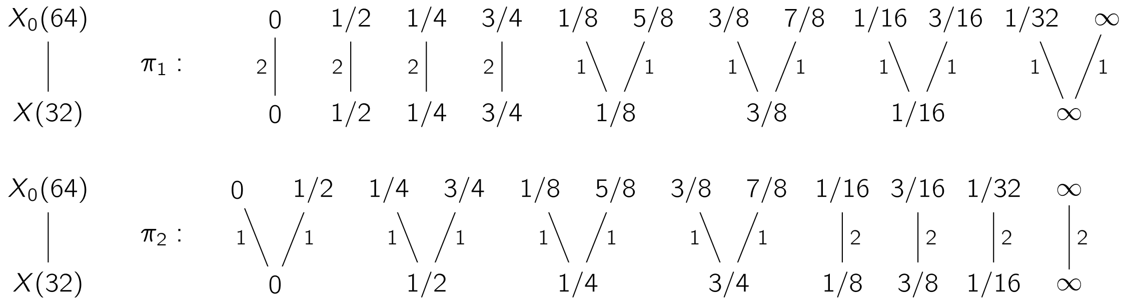
and, finally,

$$
\begin{aligned}
c_1^2 = 4\,\frac{(y+2x+2)(y-2x+2)(x^2+2y)}{(y-2)^2(x^2-2y)} &= \\
&= 4\,\frac{(y^2+4y+4-4x^2)(x^2+2y)}{(y^2-4y+4)(x^2-2y)} = \\
&= 4\,\frac{y^2x^2+4yx^2+4x^2+2y^3+8y^2+8y-4x^4-8x^2y}{y^2x^2-4yx^2+4x^2-2y^3+8y^2-8y} = \\
&= 4\,\frac{y^2x^2-4yx^2+4x^2-2y^3+4y^3+8y^2-8y+16y-4x^4}{y^2x^2-4yx^2+4x^2-2y^3+8y^2-8y} = \\
&= 4\left(1+\frac{4y^3+16y-4x^4}{\dots}\right) = 4
\end{aligned}
$$

since on the last numerator we recognize the equation of $X_0(64)$. We conclude

$$
\omega_{64}^*(x_{64},y_{64}) = \left(2\frac{2-y_{64}}{y_{64}-2x_{64}+2}, 2\frac{2x_{64}+y_{64}+2}{y_{64}-2x_{64}+2}\right)
$$

The ramification diagrams look as follows



For some applications, as computing a stable model for our modular curve, it might be useful to study the quotient curve $X_0^+(N)$ [Mcm2]; we refer to section 2.2.6 for details. The computation of a model for the Atkin-Lehner quotient of a modular curve provides also another equation for $X_0(N)$. We will see the procedure with $X_0(64)$.

In order to construct a model for $X_0^+(64) = X_0(64)/\omega_{64}$ we need to choose two functions on $X_0(64)$ which are invariant under the Atkin-Lehner involution. We select

$$
f = \frac{x^2-2y}{y-2x+2} = \frac{\eta_2(\tau)^3\eta_{32}(\tau)^3}{\eta(\tau)^2\eta_4(\tau)\eta_{16}(\tau)\eta_{64}(\tau)^2}
$$

with divisor $(f) = -2(0)+2(1/2)+2(1/32)-2(\infty)$, and

$$
g = \frac{x(y-2)}{y-2x+2} = \frac{\eta_2(\tau)\eta_4(\tau)^2\eta_{16}(\tau)^2\eta_{32}(\tau)}{\eta(\tau)^2\eta_8(\tau)^2\eta_{64}(\tau)^2}
$$

whose divisor is $(g) = -3(0)+(1/2)+\sum_{i=1,3}(i/4)+\sum_{i=1,3}(i/16)+(1/32)-3(\infty)$.

From the genus formula (end of section 2.2.6), we know that $X_0^+(64)$ is a genus one curve, i.e., an elliptic curve. By linear algebra we argue that f and g satisfy the following algebraic relation:

$$
g^2+2fg = f^3+2f^2+2f
$$

We have described an elliptic curve of j-invariant 1728 which corresponds to a quotient of $X_0^+(64)$. We observe that there is an algebraic relation

$$
x^2+x(2f-g+2)+2g = 0 \qquad (*)
$$

implying that $X_0(64)$ is of degree two on the elliptic curve. This means that the latter is the whole $X_0^+(64)$ and that f and g generate the full function field of $X_0^+(64)$. Further, the two equations above give another model for $X_0(64)$.

Remark. We can also check that equation $(*)$ describes a genus 3 curve (and therefore it must be $X_0(64)$) by counting the ramification points. A simple way of doing this is by describing their f and g coordinates. From $(*)$ we get the discriminant

$$(2f-g+2)^2-4(2g)=0 \quad \Rightarrow \quad 4f^2+g^2+4-4fg+8f-4g-8g=0 \quad \Rightarrow$$

$$\Rightarrow \quad 4f^2+f^3+2f^2-2fg+2f+4-4fg+8f-12g=0 \quad \Rightarrow \quad f^3+6f^2+10f+4=6g(f+2)$$

This gives the g-coordinate:

$$g=\frac{f^3+6f^2+10f+4}{6f+12}$$

Now, the f-coordinate is obtained using the algebraic relation between f and g:

$$f^6-12f^5-64f^4-40f^3+76f^2-112f-272=0$$

which gives

$$f^4-16f^3-4f^2+40f-68=0$$

This implies that there are 4 ramification points and, thanks to the Riemann-Hurwitz formula 1.31 we argue that the genus g of $(*)$ is

$$2g=2+(2g(X_0^+(64))-2)\cdot d+\sum(e_P-1)=2+\sum(e_P-1)=2+4 \quad \Rightarrow \quad g=3$$

where d is the degree of the map.

Remark. As genus 1 curves, $X_0^+(64)$ and $X_0(32)$ are isomorphic (they are two elliptic curves with the same j-invariant) but this isomorphism is not modular. The two curves have not the same moduli interpretation and they are not isomorphic as modular curves.

Defining equations for $X_0(2^{2n})$

In [TY2], Fang-Ting Tu and Yifan Yang propose a recursive definition for models of $X_0(2^{2n})$. Let

$$\theta_2(\tau)=\sum_{n\in\mathbb{Z}}q^{(2n+1)^2/8}=2\frac{\eta(2\tau)^2}{\eta(\tau)}$$

$$\theta_3(\tau)=\sum_{n\in\mathbb{Z}}q^{n^2/2}=\frac{\eta(\tau)^5}{\eta(\tau/2)^2\eta(2\tau)^2}$$

$$\theta_4(\tau)=\sum_{n\in\mathbb{Z}}(-1)^nq^{n^2/2}=\frac{\eta(\tau/2)^2}{\eta(\tau)}$$

be the Jacobi theta functions.

Theorem A.4 ([TY2, Th. 1])**.** *Let $P_6(x,y)=x^4-y^3-4y$ and for $n\geq 7$ define polynomials $P_n(x,y)$ recursively by*

$$P_n(x,y)=P_{n-1}\left(\frac{x}{\sqrt{y}},\frac{\sqrt{y^2+4}}{\sqrt{x}}\right)P_{n-1}\left(\frac{x}{\sqrt{y}},-\frac{\sqrt{y^2+4}}{\sqrt{x}}\right)y^{2^{n-5}}$$

Then $P_{2n}(x,y)=0$ is a defining equation of the modular curve $X_0(2^{2n})$ for $n\geq 3$.

To be more precise, for $n\geq 1$, let

$$x_n=\frac{\theta_2(8\tau)}{\theta_2(2^{n-1}\tau)} \qquad y_n=\frac{2\theta_3(2^{n-1}\tau)}{\theta_2(2^{n-1}\tau)}$$

Then,

1. *For $n\geq 2$, we have $x_{n-1}=x_n/\sqrt{y_n}$ and $y_{n-1}=\sqrt{(y_n^2+4)/y_n}$.*

2. *For $n\geq 6$, $P_n(x,y)$ is irreducible over $\mathbb{C}$.*

3. *When n is an even integer greater than 4, x_n and y_n are modular functions on $\Gamma_0(2^n)$ holomorphic everywhere except for a pole of order $2^{n-4}-1$ and 2^{n-4}, respectively, at ∞. Thus, they generate the field of modular functions on $X_0(2^n)$ and they satisfy the algebraic relation $P_n(x_n, y_n) = 0$.*

Modular tower for $X_0(48)$

We want to construct a model for $X_0(48)$ and the maps down the tower

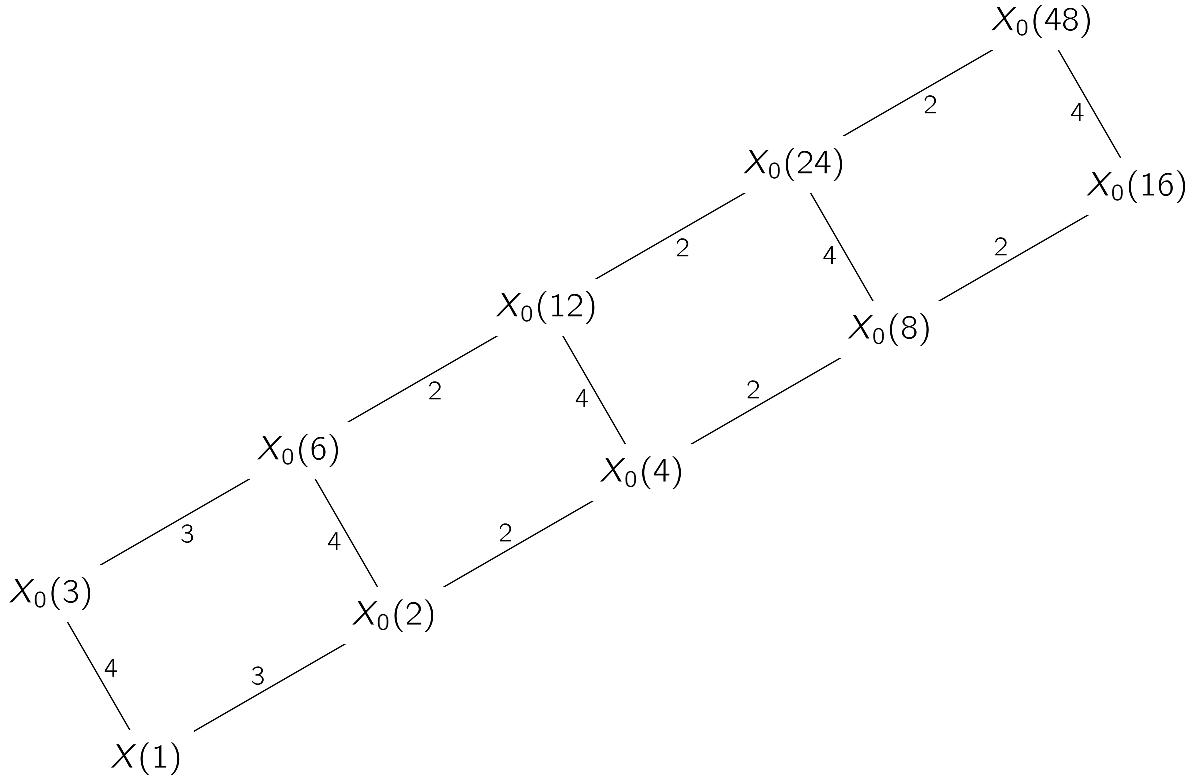


Figure A.2 – Modular tower for $X_0(48)$

We have already studied the descent $X_0(16) \to X(1)$. We focus therefore on the top chain.

Model for $X_0(3)$

$X_0(3)$ is a genus 0 modular curve with parameter

$$t_3 = \left(\frac{\eta_1(\tau)}{\eta_3(\tau)}\right)^{12} \qquad (t_3) = (0) - (\infty)$$

It has reduction maps down to $X(1)$ given by

$$\pi_1^*(j) = \frac{(t_3+27)(t_3+243)^3}{t_3^3} \qquad \pi_3^*(j) = \frac{(t_3+27)(t_3+3)^3}{t_3}$$

Finally, the Atkin-Lehner involution acts on t_3 as

$$\omega_3^*(t_3) = \frac{729}{t_3}$$

Model for $X_0(6)$

The modular curve $X_0(6)$ is again of genus 0 and it has Hauptmodul

$$t_6 = \frac{\eta_1(\tau)^5\eta_3(\tau)}{\eta_2(\tau)\eta_6(\tau)^5} \qquad (t_6) = (0) - (\infty)$$

This curve has composite level structure resulting in 4 maps.
We have $\pi_1, \pi_3 : X_0(6) \to X_0(2)$ with pullback

$$\pi_1^*(t_2) = \frac{t_6^3(t_6+8)}{(t_6+9)^3} \qquad \pi_3^*(t_2) = \frac{t_6(t_6+8)^3}{t_6+9}$$

and $\pi_1, \pi_2 : X_0(6) \to X_0(3)$ with pullback

$$\pi_1^*(t_3) = \frac{t_6^2(t_6+9)}{(t_6+8)^2} \qquad \pi_2^*(t_3) = \frac{t_6(t_6+9)^2}{t_6+8}$$

The Atkin-Lehner involutions act as

$$\omega_6^*(t_6) = \frac{72}{t_6} \qquad \omega_2^*(t_6) = \frac{-8(t_6+9)}{t_6+8} \qquad \omega_3^*(t_6) = \frac{-9(t_6+8)}{t_6+9}$$

Model for $X_0(12)$

$X_0(12)$ is a modular curve of genus 0 and parameter

$$t_{12} = \frac{\eta_1(\tau)^3\eta_4(\tau)\eta_6(\tau)^2}{\eta_2(\tau)^2\eta_3(\tau)\eta_{12}(\tau)^3} \qquad (t_{12}) = (0) - (\infty)$$

We have two maps $\pi_1, \pi_3 : X_0(12) \to X_0(4)$

$$\pi_1^*(t_4) = \frac{t_{12}^3(t_{12}+4)}{(t_{12}+3)^3} \qquad \pi_3^*(t_4) = \frac{t_{12}(t_{12}+4)^3}{t_{12}+3}$$

and two maps $\pi_1, \pi_2 : X_0(12) \to X_0(6)$

$$\pi_1^*(t_6) = \frac{t_{12}^2}{t_{12}+2} \qquad \pi_2^*(t_6) = t_{12}(t_{12}+6)$$

The 3 Atkin-Lehner involutions are described by

$$\omega_{12}^*(t_{12}) = \frac{12}{t_{12}} \qquad \omega_4^*(t_{12}) = \frac{-4(t_{12}+3)}{t_{12}+4} \qquad \omega_3^*(t_{12}) = \frac{-2(t_{12}+4)}{t_{12}+3}$$

Model for $X_0(24)$

$X_0(24)$ is a genus 1 curve with modular functions generating its function field

$$x_{24} = \frac{\eta_6(\tau)^3\eta_8(\tau)}{\eta_2(\tau)\eta_{24}(\tau)^3} \quad \text{and} \quad y_{24} = \frac{\eta_4(\tau)\eta_8(\tau)^2\eta_{12}(\tau)^5}{\eta_2(\tau)\eta_6(\tau)\eta_{24}(\tau)^6}$$

with divisors

$$(x_{24}) = (1/3) + (1/6) - 2(\infty) \quad \text{and} \quad (y_{24}) = (1/4) + (1/8) + (1/12) - 3(\infty)$$

The relation between the two functions gives a model for $X_0(24)$:

$$y^2 = x^3 - x^2 - 4x + 4$$

Regarding the maps down the modular tower we find $\pi_1, \pi_3 : X_0(24) \to X_0(8)$

$$\pi_1^*(t_8) = \frac{y_{24}(x_{24}+2)}{(x_{24}-1)^2} - 4 \qquad \pi_3^*(t_8) = \frac{y_{24}(x_{24}-2)}{x_{24}-1} - 4$$

and $\pi_1, \pi_2 : X_0(24) \to X_0(12)$

$$\pi_1^*(t_{12}) = \frac{y_{24}}{x_{24}-2} - 3 \qquad \pi_2^*(t_{12}) = x_{24} - 4$$

We now want to construct the Atkin-Lehner involutions. Looking at the action on the cusps, we look for two functions x'_{24} and y'_{24} with divisors

$$(x'_{24}) = -2(0) + (1/4) + (1/8) \qquad \text{and} \qquad (y'_{24}) = -3(0) + (1/2) + (1/3) + (1/6)$$

we find

$$x'_{24} = \frac{\eta_3(\tau)\eta_4(\tau)^3}{\eta_1(\tau)^3\eta_{12}(\tau)} \qquad \text{and} \qquad y'_{24} = \frac{\eta_2(\tau)^5\eta_3(\tau)^2\eta_6(\tau)}{\eta_1(\tau)^6\eta_4(\tau)\eta_{12}(\tau)}$$

These functions satisfy the following relations

$$x'_{24} = \frac{y_{24}}{y_{24} - 3x_{24} + 6} \qquad \text{and} \qquad y'_{24} = \frac{3y_{24}^2 + 7x_{24}y_{24} + 18x_{24} - 12}{3y_{24}^2 - 11x_{24}y_{24} + 44y_{24} - 42x_{24} + 60}$$

so that

$$\omega_{24}^*(x_{24}, y_{24}) = \left(c_1 \frac{y_{24}}{y_{24} - 3x_{24} + 6}, c_2 \frac{3y_{24}^2 + 7x_{24}y_{24} + 18x_{24} - 12}{3y_{24}^2 - 11x_{24}y_{24} + 44y_{24} - 42x_{24} + 60} \right)$$

The two constants are deduced by the usual commutative diagrams

$$\omega_{24}^*(\pi_2^*(t_{12})) = \pi_1^*(\omega_{12}^*(t_{12})) \qquad \text{and} \qquad \omega_{24}^*(\pi_1^*(t_{12})) = \pi_2^*(\omega_{12}^*(t_{12}))$$

The first one yields $c_1 = 4$ while the second one gives $c_2 = 6$.

It remains to describe the Atkin-Lehner involutions ω_3 and ω_8. We recall that they represent the flipping of the rectangle of isogenies with respect to the two axis (vertical and horizontal).

Once again, we start by looking at their action on the cusps; using [Ogg, Prop. 2], we obtain the following picture

$$\begin{tikzcd}
(1/8) & (1/4) & (\infty) & (1/2) & (1/12) & (1/1) & (1/6) & (1/3) \\
(1/1) \arrow[u, "\omega_8"] \arrow[d, "\omega_3"'] & (1/2) \arrow[u] \arrow[d] & (1/3) \arrow[u] \arrow[d] & (1/4) \arrow[u] \arrow[d] & (1/6) \arrow[u] \arrow[d] & (1/8) \arrow[u] \arrow[d] & (1/12) \arrow[u] \arrow[d] & (\infty) \arrow[u] \arrow[d] \\
(1/3) & (1/6) & (1/1) & (1/12) & (1/2) & (\infty) & (1/4) & (1/8)
\end{tikzcd}$$

For ω_8 we then need two functions x''_{24} and y''_{24} with divisors

$$(x''_{24}) = -2(1/3) + (1/12) + (\infty) \qquad \text{and} \qquad (y''_{24}) = (1/1) + (1/2) - 3(1/3) + (1/6)$$

We find

$$x''_{24} = \frac{\eta_1(\tau)\eta_{12}(\tau)^3}{\eta_3(\tau)^3\eta_4(\tau)} \qquad \text{and} \qquad y''_{24} = \frac{\eta_1(\tau)^2\eta_2(\tau)\eta_6(\tau)^5}{\eta_3(\tau)^6\eta_4(\tau)\eta_{12}(\tau)}$$

so that

$$x''_{24} = \frac{x_{24} - 2}{x_{24} + y_{24} - 2} \qquad \text{and} \qquad y''_{24} = \frac{x_{24}^2 - 4x_{24}}{x_{24}^2 + 2x_{24} + 2y_{24} - 4}$$

For ω_3 we look for two functions x'''_{24} and y'''_{24} with divisors

$$(x'''_{24}) = (1/1) + (1/2) - 2(1/8) \qquad \text{and} \qquad (y'''_{24}) = (1/4) - 3(1/8) + (1/12) + (\infty)$$

$$x'''_{24} = \frac{\eta_1(\tau)\eta_{12}(\tau)^3}{\eta_3(\tau)^3\eta_4(\tau)} \qquad \text{and} \qquad y'''_{24} = \frac{\eta_1(\tau)^2\eta_2(\tau)\eta_6(\tau)^5}{\eta_3(\tau)^6\eta_4(\tau)\eta_{12}(\tau)}$$

They can be expressed as

$$x_{24}''' = \frac{x_{24}-4}{x_{24}-1} \qquad \text{and} \qquad y_{24}''' = \frac{y_{24}}{x_{24}^2-2x_{24}+1}$$

By composing $\omega_3 \circ \omega_8 = \omega_{24}$ and $\omega_8 \circ \omega_3 = \omega_{24}$, we find the constants defining the involutions.
In conclusion, we get

$$\omega_{24}^*(x_{24}, y_{24}) = \left(4\frac{y_{24}}{y_{24}-3x_{24}+6}, 6\frac{3y_{24}^2+7x_{24}y_{24}+18x_{24}-12}{3y_{24}^2-11x_{24}y_{24}+44y_{24}-42x_{24}+60}\right)$$

$$\omega_8^*(x_{24}, y_{24}) = \left(4\frac{x_{24}-2}{x_{24}+y_{24}-2}, 2\frac{x_{24}^2-4x_{24}}{x_{24}^2+2x_{24}+2y_{24}-4}\right)$$

$$\omega_3^*(x_{24}, y_{24}) = \left(\frac{x_{24}-4}{x_{24}-1}, 3\frac{y_{24}}{(x_{24}-1)^2}\right)$$

Model for $X_0(48)$

Finally, we get to study the top of our tower, $X_0(48)$: this is a hyperelliptic curve of genus 3.

We find two η products describing its function field:

$$x_{48} = \frac{\eta_4(\tau)\eta_{16}(\tau)^2\eta_{24}(\tau)^3}{\eta_8(\tau)^3\eta_{12}(\tau)\eta_{48}(\tau)^2} \qquad \text{and} \qquad x_{48} = \frac{\eta_4(\tau)\eta_{16}(\tau)^2\eta_{24}(\tau)^3}{\eta_8(\tau)^3\eta_{12}(\tau)\eta_{48}(\tau)^2}$$

with divisors

$$(x_{48}) = -(1/8)+(1/16)+(1/24)-(\infty) \qquad \text{and} \qquad (y_{48}) = 4(1/16)-4(\infty)$$

The two satisfy the following hyperelliptic equation of the form $y^2 + H(x)y = F(x)$

$$y^2 - y(x^4+1) = 3x^4$$

with the change of variables $y' = 2y - H(x)$ we get a model for $X_0(48)$ of the form $y^2 = P(X)$.

$$y^2 = x^8 + 14x^4 + 1$$

We are interested in the maps down the tower; we get $\pi_1, \pi_3 : X_0(48) \to X_0(16)$

$$\pi_1^*(t_{16}) = \frac{x_{48}y_{48}+3x_{48}-2y_{48}}{y_{48}} \qquad \pi_3^*(t_{16}) = \frac{y_{48}-2x_{48}-1}{x_{48}}$$

and $\pi_1, \pi_2 : X_0(48) \to X_0(24)$

$$\pi_1^*(x_{24}, y_{24}) = \left(\frac{x_{48}^2+y_{48}}{x_{48}^2}, \frac{x_{48}^2y_{48}+3x_{48}^2+y_{48}^2-y_{48}}{x_{48}(y_{48}-1)}\right) \qquad \pi_2^*(x_{24}, y_{24}) = \big(y_{48}+1, x_{48}^2(y_{48}+3)\big)$$

Finally, we want to describe Atkin-Lehner involutions.

Remark. It is worth observing that $X_0(48)$ is one of the only two hyperelliptic modular curves whose hyperelliptic involution is not of Atkin-Lehner type. We recall that the hyperelliptic involution on $y^2+H(x)y = F(x)$ is the map $(x, y) \mapsto (x, -y - H(x))$.

Let us look at the action of ω_{16} and ω_3 on the cusps. We find

$$\begin{array}{ccccccccccccc}
 & (1/16) & (1/8) & (\infty) & (3/4) & (1/4) & (1/24) & (1/2) & (7/12) & (1/12) & (1/1) & (1/16) & (1/3) \\
\omega_{16} & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\
 & (1/1) & (1/2) & (1/3) & (1/4) & (3/4) & (1/6) & (1/8) & (1/12) & (7/12) & (1/16) & (1/24) & (\infty) \\
\omega_3 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 & (1/3) & (1/6) & (1/1) & (1/12) & (7/12) & (1/2) & (1/24) & (1/4) & (3/4) & (\infty) & (1/8) & (1/16)
\end{array}$$

We find

$$\omega_{48}^*(x_{48}, y_{48}) = \left(\frac{x_{48}+1}{x_{48}-1}, -3\frac{2x_{48}^2 - x_{48}y_{48} + 3x_{48} - y_{48} + 1}{6x_{48}^2 + x_{48}y_{48} - 3x_{48} - 3y_{48} + 3}\right)$$

$$\omega_{16}^*(x_{48}, y_{48}) = \left(\frac{1-x_{48}}{1+x_{48}}, \frac{6x_{48}^2 + x_{48}y_{48} - 3x_{48} - 3y_{48} + 3}{2x_{48}^2 - x_{48}y_{48} + 3x_{48} - y_{48} + 1}\right)$$

$$\omega_3^*(x_{48}, y_{48}) = \left(-\frac{1}{x_{48}}, -3\frac{1}{y_{48}}\right)$$

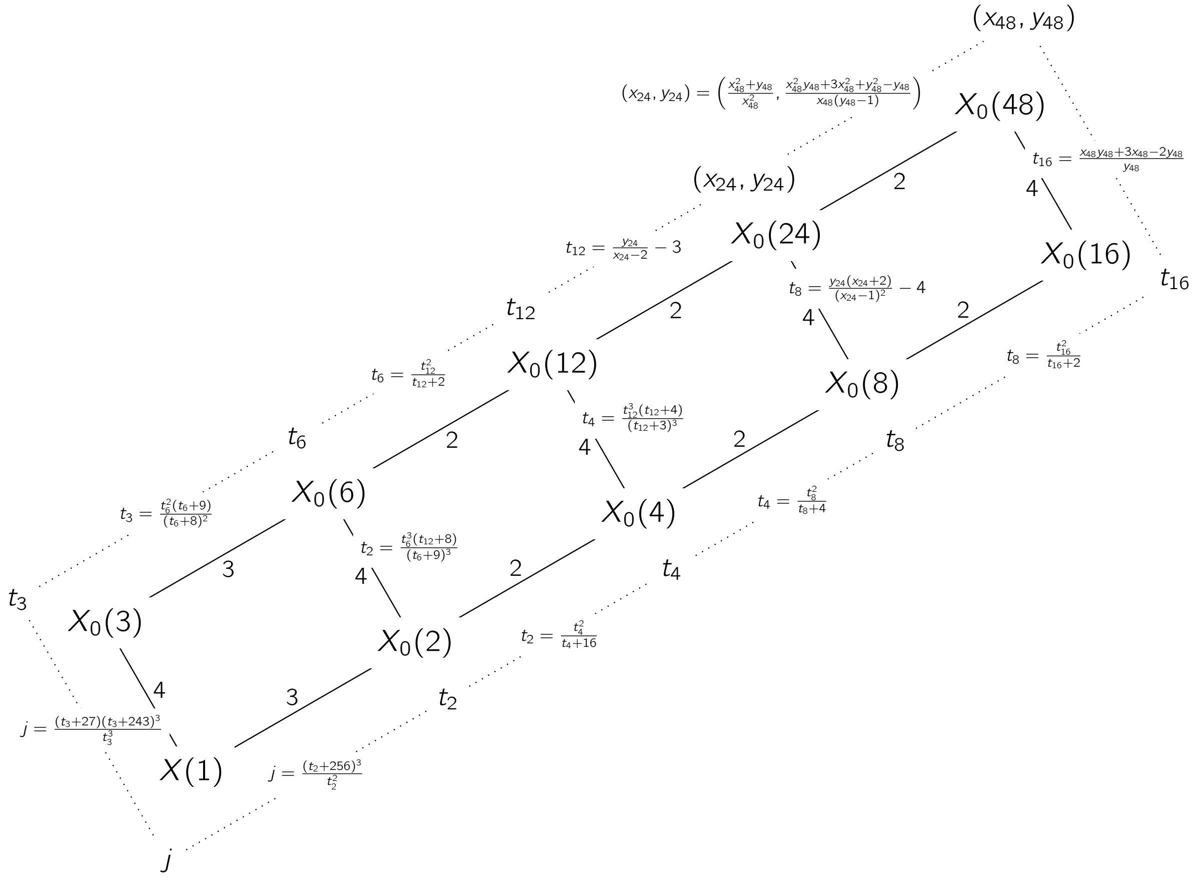

(x_{48}, y_{48})
$(x_{24}, y_{24}) = \left(\frac{x_{48}^2+y_{48}}{x_{48}^2}, \frac{x_{48}^2y_{48}+3x_{48}^2+y_{48}^2-y_{48}}{x_{48}(y_{48}-1)}\right)$
$X_0(48)$
$t_{16} = \frac{x_{48}y_{48}+3x_{48}-2y_{48}}{y_{48}}$
(x_{24}, y_{24})
2
4
$t_{12} = \frac{y_{24}}{x_{24}-2} - 3$
$X_0(24)$
$X_0(16)$
t_{16}
$t_8 = \frac{y_{24}(x_{24}+2)}{(x_{24}-1)^2} - 4$
t_{12}
2
4
2
$t_6 = \frac{t_{12}^2}{t_{12}+2}$
$X_0(12)$
$X_0(8)$
$t_8 = \frac{t_{16}^2}{t_{16}+2}$
$t_4 = \frac{t_{12}^3(t_{12}+4)}{(t_{12}+3)^3}$
t_6
2
4
2
t_8
$X_0(6)$
$t_4 = \frac{t_8^2}{t_8+4}$
$t_3 = \frac{t_6^2(t_6+9)}{(t_6+8)^2}$
$X_0(4)$
$t_2 = \frac{t_6^3(t_{12}+8)}{(t_6+9)^3}$
3
4
2
t_4
t_3
$X_0(3)$
$X_0(2)$
$t_2 = \frac{t_4^2}{t_4+16}$
4
3
t_2
$j = \frac{(t_3+27)(t_3+243)^3}{t_3^3}$
$X(1)$
$j = \frac{(t_2+256)^3}{t_2^2}$
j

Figure A.3 – Modular tower for $X_0(48)$ with maps

Modular tower for $X_0(36)$

Let's have a look at what happens when we add an extra 3-level structure

Model for $X_0(9)$

This is a genus 0 curve with parameter

$$t_9 = \left(\frac{\eta_1}{\eta_9}\right)^3 \qquad (t_9) = (0) - (\infty)$$

and maps to $X_0(9) \to X_0(3)$

$$\pi_1^*(t_3) = \frac{t_9^3}{t_9^2 + 9t_9 + 27} \quad \text{and} \quad \pi_3^*(t_3) = t_9(t_9^2 + 9t_9 + 27)$$

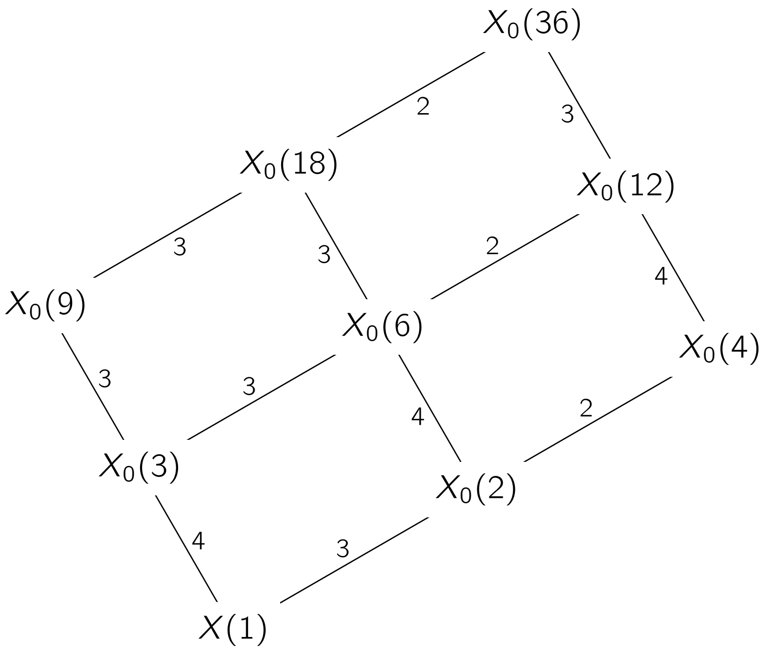


Figure A.4 – Modular tower for $X_0(36)$

The Atkin-Lehner involution is

$$\omega_9^*(t_9) = \frac{27}{t_9}$$

Model for $X_0(18)$

$X_0(18)$ is a genus 0 curve with parameter

$$t_{18} = \frac{\eta_1^2\eta_6(\tau)\eta_9(\tau)}{\eta_2\eta_3(\tau)\eta_{18}(\tau)^2} \qquad (t_{18}) = (0) - (\infty)$$

and moduli-theoretic maps $\pi_1, \pi_3 : X_0(18) \to X_0(6)$

$$\pi_1^*(t_6) = \frac{t_{18}^3}{t_{18}^2 + 3t_{18} + 3} \quad \text{and} \quad \pi_3^*(t_6) = t_{18}(t_{18}^2 + 6t_{18} + 12)$$

and $\pi_1, \pi_2 : X_0(18) \to X_0(9)$

$$\pi_1^*(t_9) = \frac{t_{18}^2(t_{18}+3)}{(t_{18}+2)^2} \quad \text{and} \quad \pi_2^*(t_9) = \frac{t_{18}(t_{18}+3)^2}{t_{18}+2}$$

The Atkin-Lehner involutions act as

$$\omega_{18}^*(t_{18}) = \frac{6}{t_{18}}$$
$$\omega_2^*(t_{18}) = -2\frac{t_{18}+3}{t_{18}+2}$$
$$\omega_9^*(t_{18}) = -3\frac{t_{18}+2}{t_{18}+3}$$

Model for $X_0(36)$

Finally, $X_0(36)$ has genus 1. It has two modular functions

$$x_{36} = \frac{\eta_2(\tau)^2\eta_{12}(\tau)\eta_{18}(\tau)}{\eta_4(\tau)\eta_6(\tau)\eta_{36}(\tau)^2} \quad \text{and} \quad y_{36} = \frac{\eta_1(\tau)^2\eta_9(\tau)\eta_{12}(\tau)\eta_{18}(\tau)}{\eta_2(\tau)\eta_3(\tau)\eta_{36}(\tau)^3}$$

with divisors

$$(x_{36}) = (0) + (1/2) - 2(\infty) \quad \text{and} \quad (y_{36}) = 2(0) + (1/9) - 3(\infty)$$

yielding the model

$$y^2 + 4xy + 6y = x^3 + 2x^2$$

The maps down the tower are given by $\pi_1, \pi_3 : X_0(36) \to X_0(12)$

$$\pi_1^*(t_{12}) = \frac{x_{36}y_{36} - x_{36}^2 + 3y_{36}}{x_{36}^2 + 3x_{36} + 3} \quad \text{and} \quad \pi_3^*(t_{12}) = 2x_{36} + y_{36}$$

and $\pi_1, \pi_2 : X_0(36) \to X_0(18)$

$$\pi_1^*(t_{18}) = \frac{y_{36}}{x_{36} + 2} \quad \text{and} \quad \pi_2^*(t_{18}) = x_{36}$$

The Atkin-Lehner involutions are

$$\omega_{36}^*(x_{36}, y_{36}) = \left(6\frac{y_{36}+2}{x_{36}}, 12\frac{x_{36}^2 + 2x_{36} - y_{36}}{y_{36}^2}\right)$$

$$\omega_4^*(x_{36}, y_{36}) = \left(2\frac{x_{36}(3y_{36} - x_{36}^2)}{y_{36}^2}, 4\frac{x_{36}+3}{2x_{36} + y_{36} + 4}\right)$$

$$\omega_9^*(x_{36}, y_{36}) = \left(-3\frac{(x_{36}+2)}{x_{36}+3}, 3\frac{x_{36}^2 + 2x_{36} - y_{36}}{(x_{36}+3)^2}\right)$$

Models for $X_0(2p)$

We present here a table with a choice for the parameters on $X_0(2p)$.
Notation. The η-product $\eta_1(\tau)^{a_1}\eta_2(\tau)^{a_2}\eta_p(\tau)^{a_p}\eta_{2p}(\tau)^{a_{2p}}$ is represented by $[a_1, a_2, a_p, a_{2p}]$.

$p \equiv 1$ mod 24

Lemma A.5. *If $p = 24k + 1 \equiv 1$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*
$a = [-1, 1, 1, -1]$ *with* $\deg_\infty(a) = (p-1)/12 = 2k$.
$b = [6, -2, 18, -22]$ *with* $\deg_\infty(b) = (13p-1)/12 = 24k + 1$.
$c = [7, -3, 17, -21]$ *with* $\deg_\infty(c) = (25p-1)/24 = 25k + 1$.

If k is even, then (a, c) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.
If k is odd, then (a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.

$p \equiv 5$ mod 24

Lemma A.6. *If $p = 24k + 5 \equiv 5$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*
$a = [-1, 1, 5, -5]$ *with* $\deg_\infty(a) = (p-1)/4 = 6k + 1$.
$b = [4, -4, 4, -4]$ *with* $\deg_\infty(b) = (p+1)/3 = 8k + 2$.

(a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.

$p \equiv 7$ mod 24

Lemma A.7. *If $p = 24k + 7 \equiv 7$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*
$a = [3, -3, 3, -3]$ *with* $\deg_\infty(a) = (p+1)/4 = 6k + 2$.
$b = [-4, 8, 4, -8]$ *with* $\deg_\infty(b) = (p-1)/2 = 12k + 3$.

(a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.

$p \equiv 11$ mod 24

Lemma A.8. *If $p = 24k + 11 \equiv 11$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*
$a = [2, -2, 2, -2]$ *with* $\deg_\infty(a) = (p+1)/6 = 4k + 2$.
$b = [-4, 8, 4, -8]$ *with* $\deg_\infty(b) = (p-1)/2 = 12k + 5$.

p	$g\left(X_0(2p)\right)$	x_{2p}	y_{2p}	$[\deg_\infty(x_{2p}), \deg_\infty(y_{2p})]$
3	0	$[5, -1, 1, -5]$	$--$	$[1]$
5	0	$[3, -1, 1, -3]$	$--$	$[1]$
7	1	$[3, -3, 3, -3]$	$[-4, 8, 4, -8]$	$[2, 3]$
11	2	$[2, -2, 2, -2]$	$[-4, 8, 4, -8]$	$[2, 5]$
13	2	$[-2, 2, 2, -2]$	$[-2, 4, 2, -4]$	$[2, 3]$
17	3	$[1, 1, -1, -1]$	$[0, -2, 8, -6]$	$[4, 7]$
19	4	$[1, -1, 5, -5]$	$[-4, 4, 4, -4]$	$[5, 6]$
23	5	$[1, -1, 1, -1]$	$[-4, 8, 4, -8]$	$[2, 11]$
29	6	$[-1, 1, 5, -5]$	$[4, -4, 4, -4]$	$[7, 10]$
31	7	$[3, -3, 3, -3]$	$[-4, 8, 4, -8]$	$[8, 15]$
37	8	$[-2, 2, 2, -2]$	$[-5, -7, 7, -5]$	$[6, 19]$
41	9	$[1, 1, -1, -1]$	$[0, -2, 8, -6]$	$[10, 17]$
43	10	$[1, -1, 5, -5]$	$[-4, 4, 4, -4]$	$[11, 14]$
47	11	$[1, -1, 1, -1]$	$[-4, 8, 4, -8]$	$[4, 23]$
53	12	$[-1, 1, 5, -5]$	$[4, -4, 4, -4]$	$[13, 18]$
59	14	$[2, -2, 2, -2]$	$[-4, 8, 4, -8]$	$[10, 29]$
61	14	$[-2, 2, 2, -2]$	$[-5, -7, 7, -5]$	$[10, 31]$
67	16	$[1, -1, 5, -5]$	$[-4, 4, 4, -4]$	$[17, 22]$
71	17	$[1, -1, 1, -1]$	$[-4, 8, 4, -8]$	$[6, 35]$
73	17	$[-1, 1, 1, -1]$	$[6, -2, 18, -22]$	$[6, 79]$
79	19	$[3, -3, 3, -3]$	$[-4, 8, 4, -8]$	$[20, 39]$
83	20	$[2, -2, 2, -2]$	$[-4, 8, 4, -8]$	$[14, 41]$
89	21	$[1, 1, -1, -1]$	$[0, -2, 8, -6]$	$[22, 37]$
97	23	$[-1, 1, 1, -1]$	$[7, -3, 17, -21]$	$[8, 101]$
101	24	$[-1, 1, 5, -5]$	$[4, -4, 4, -4]$	$[25, 34]$

Table A.2 – Eta quotients on $X_0(2p)$

(a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.

$p \equiv 13$ mod 24

Lemma A.9. *If $p = 24k + 13 \equiv 13$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*

$a = [-2, 2, 2, -2]$ *with* $\deg_\infty(a) = (p-1)/6 = 4k + 2$.
$b = [5, -7, 7, -5]$ *with* $\deg_\infty(b) = (p+1)/2 = 12k + 7$.
$b = [-2, 4, 2, -4]$ *with* $\deg_\infty(b) = (p-1)/4 = 6k + 3$.

(a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime. In case $p = 13$ we can lower the degree of the second parameter by choosing (a, c).

$p \equiv 17$ mod 24

Lemma A.10. *If $p = 24k + 17 \equiv 17$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*

$a = [1, 1, -1, -1]$ *with* $\deg_\infty(a) = (p-1)/4 = 6k + 4$.
$b = [0, -2, 8, -6]$ *with* $\deg_\infty(b) = (5p-1)/12 = 10k + 7$.

(a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.

$p \equiv 19$ mod 24

Lemma A.11. *If $p = 24k + 19 \equiv 19$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*

$a = [1, -1, 5, -5]$ *with* $\deg_\infty(a) = (p+1)/4 = 6k + 5$.
$b = [-4, 4, 4, -4]$ *with* $\deg_\infty(b) = (p-1)/3 = 8k + 6$.

(a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.

$p \equiv 19$ mod 24

Lemma A.12. *If $p = 24k + 23 \equiv 23$ mod 24, the following are η-quotients for $\Gamma_0(2p)$*
$a = [1, -1, 1, -1]$ *with* $\deg_\infty(a) = (p + 1)/12 = 2k + 2$.
$b = [-4, 8, 4, -8]$ *with* $\deg_\infty(b) = (p - 1)/2 = 12k + 11$.

(a, b) is a possible choice of parameters for $X_0(2p)$ since their degrees $\deg_\infty$ are coprime.

Models for $X_0(3p)$

We present here a table with a possible choice for the parameters on $X_0(3p)$.

p	$g\,(X_0(3p))$	x_{3p}	y_{3p}	$[\deg_\infty(x_{3p}), \deg_\infty(y_{3p})]$
5	0	$[2, -2, 2, -2]$	$[4, 2, 4, -10]$	$[2, 5]$
7	1	$[1, 1, -1, -1]$	$[0, 2, 6, -8]$	$[2, 5]$
11	3	$[1, -1, 1, -1]$	$[-2, 8, 4, -10]$	$[2, 11]$
13	3	$[-1, 1, 1, -1]$	$[7, -1, 5, -11]$	$[2, 15]$
17	5	$[2, -2, 2, -2]$	$[-1, 7, 5, -11]$	$[6, 19]$
19	5	$[1, 1, -1, -1]$	$[0, -2, 6, -4]$	$[6, 11]$
23	7	$[1, -1, 1, -1]$	$[6, 6, 6, -18]$	$[4, 45]$
29	9	$[2, -2, 2, -2]$	$[4, 2, 4, -10]$	$[10, 31]$
31	9	$[1, 1, -1, -1]$	$[0, 2, 6, -8]$	$[10, 23]$
37	11	$[-1, 1, 1, -1]$	$[7, -1, 5, -11]$	$[6, 43]$
41	13	$[2, -2, 2, -2]$	$[-1, 7, 5, -11]$	$[14, 47]$
43	13	$[1, 1, -1, -1]$	$[0, -2, 6, -4]$	$[14, 25]$
47	15	$[1, -1, 1, -1]$	$[6, 6, 6, -18]$	$[8, 93]$
53	17	$[2, -2, 2, -2]$	$[4, 2, 4, -10]$	$[8, 57]$
59	19	$[1, -1, 1, -1]$	$[-2, 8, 4, -10]$	$[10, 63]$
61	19	$[-1, 1, 1, -1]$	$[7, -1, 5, -11]$	$[10, 71]$
67	21	$[1, 1, -1, -1]$	$[0, -2, 6, -4]$	$[22, 39]$
71	23	$[1, -1, 1, -1]$	$[-6, 18, 18, -30]$	$[12, 211]$
73	23	$[-1, 1, 1, -1]$	$[5, -1, 7, -11]$	$[12, 79]$
79	25	$[1, 1, -1, -1]$	$[0, 2, 6, -8]$	$[26, 59]$
83	27	$[1, -1, 1, -1]$	$[-2, 8, 4, -10]$	$[14, 89]$
89	29	$[2, -2, 2, -2]$	$[-1, 7, 5, -11]$	$[30, 103]$
97	31	$[-1, 1, 1, -1]$	$[5, -1, 7, -11]$	$[16, 105]$
101	33	$[2, -2, 2, -2]$	$[4, 2, 4, -10]$	$[34, 109]$

Table A.3 – Eta quotients on $X_0(3p)$

$p \equiv 1$ mod 12

Lemma A.13. *If $p = 12k + 1 \equiv 1$ mod 12, the following are η-quotients for $\Gamma_0(3p)$*
$a = [-1, 1, 1, -1]$ *with* $\deg_\infty(a) = (p - 1)/6 = 2k$.
$b = [7, -1, 5, -11]$ *with* $\deg_\infty(b) = (7p - 1)/6 = 14k + 1$.
$c = [5, -1, 7, -11]$ *with* $\deg_\infty(c) = (13p - 1)/12 = 13k + 1$.

Therefore, if $p \equiv 1 \mod 12$, then a choice for parameters for $X_0(3p)$ could be (a, b) since their degrees $\deg_\infty$ are coprime.
In case $p \equiv 1 \mod 24$ we can lower the degree of the second parameter by choosing (a, c) instead. If $p \equiv 13 \mod 24$ $\deg_\infty(a)$ and $\deg_\infty(c)$ are not coprime anymore as they are both even.

$p \equiv 5$ mod 12

Lemma A.14. *If $p = 12k + 5 \equiv 5$ mod 12, the following are η-quotients for $\Gamma_0(3p)$*
$a = [2, -2, 2, -2]$ *with* $\deg_\infty(a) = (p + 1)/3 = 4k + 2$.

$b = [-1, 7, 5, -11]$ *with* $\deg_\infty(b) = (7p-5)/6 = 14k+5$.
$c = [4, 2, 4, -10]$ *with* $\deg_\infty(c) = (13p-5)/12 = 13k+5$.

Therefore, if $p \equiv 5 \mod 12$, a choice of parameters for $X_0(3p)$ could be (a, b) since their degrees $\deg_\infty$ are coprime.
In case $p \equiv 5 \mod 24$ we can lower the degree of the second parameter by choosing (a, c) instead. If $p \equiv 17 \mod 24$, or k is odd $\deg_\infty(a)$ and $\deg_\infty(c)$ are not coprime anymore as they are both even.

$p \equiv 7$ mod 12

Lemma A.15. *If* $p = 12k+7 \equiv 7$ *mod* 12*, the following are* η*-quotients for* $\Gamma_0(3p)$
$a = [1, 1, -1, -1]$ *with* $\deg_\infty(a) = (p-1)/3 = 4k+2$.
$b = [0, 2, 6, -8]$ *with* $\deg_\infty(b) = (3p-1)/4 = 9k+5$.
$c = [0, -2, 6, -4]$ *with* $\deg_\infty(c) = (13p-5)/12 = 7k+4$.

Therefore, if $p \equiv 7 \mod 24$, a choice of parameters for $X_0(3p)$ would be (a, b) since their degrees $\deg_\infty$ are coprime.
Instead, if $p \equiv 19 \mod 24$, a choice of parameters for $X_0(3p)$ would be (a, c).

$p \equiv 11$ mod 12

Lemma A.16. *If* $p = 12k+11 \equiv 11$ *mod* 12*, the following are* η*-quotients for* $\Gamma_0(3p)$
$a = [1, -1, 1, -1]$ *with* $\deg_\infty(a) = (p+1)/6 = 2k+2$.
$b = [6, 6, 6, -18]$ *with* $\deg_\infty(b) = 2p-1 = 24k+21$.
$c = [-2, 8, 4, -10]$ *with* $\deg_\infty(c) = (13p-11)/12 = 13k+11$.

If $p \equiv 11, 23 \mod 36$ a choice of parameters for $X_0(3p)$ could be (a, b) since their degrees $\deg_\infty$ are coprime.
In case $p \equiv 11 \mod 24$ we can lower the degree of the second parameter by choosing (a, c) instead.

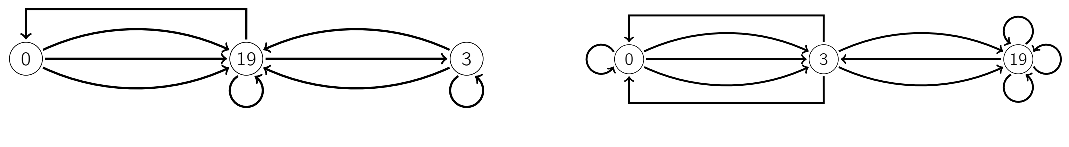
B Isogeny graphs

We present here some pictures of supersingular isogeny graphs.

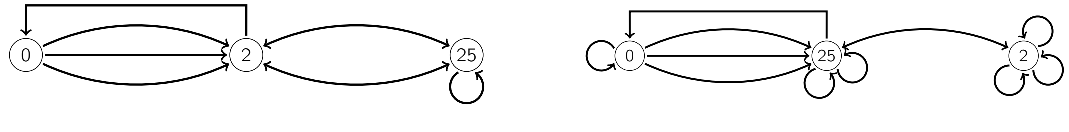
Supersingular isogeny graphs

The following figures represent the 2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_p$ for p up to 71.

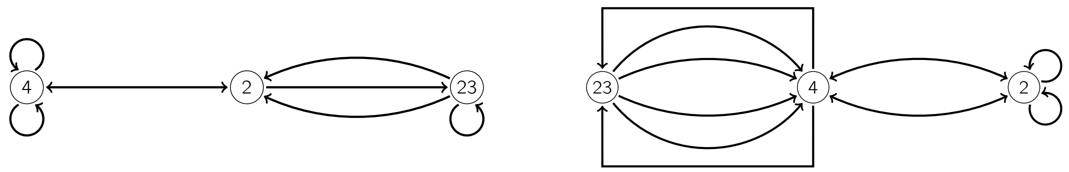
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{23}$



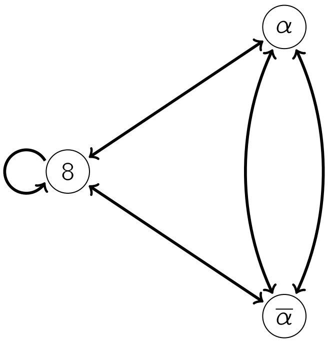
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{29}$

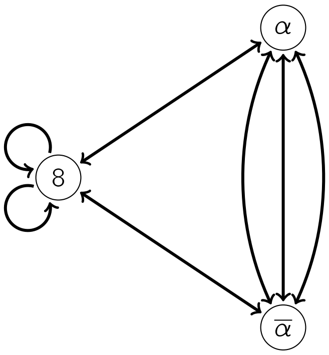


2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{31}$

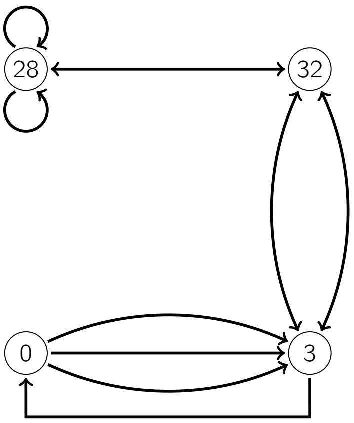


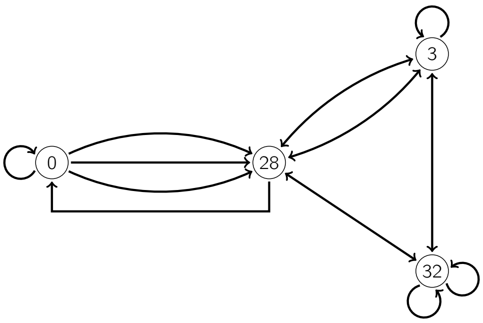
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{37}$



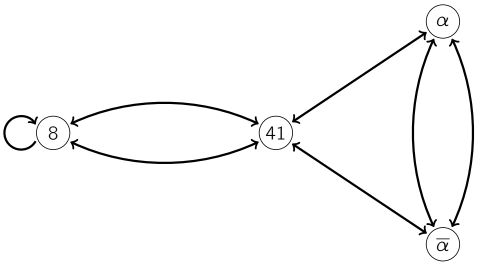


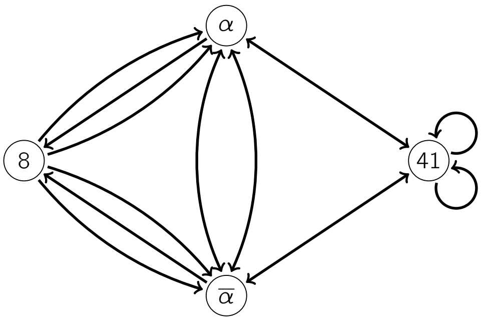
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{41}$



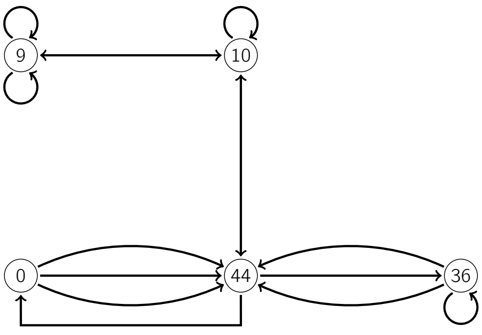


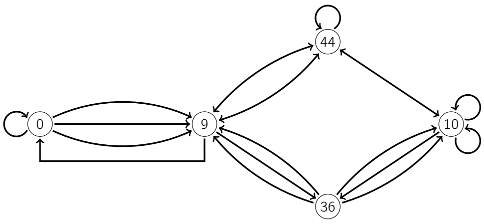
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{43}$



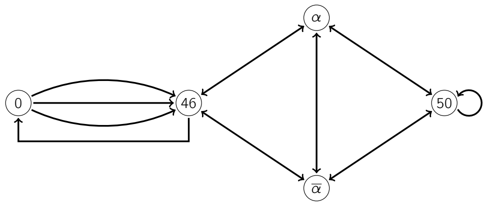


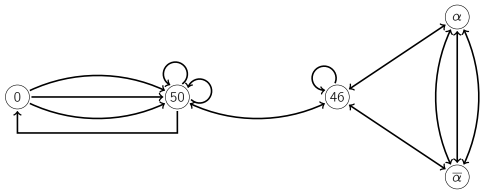
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{47}$



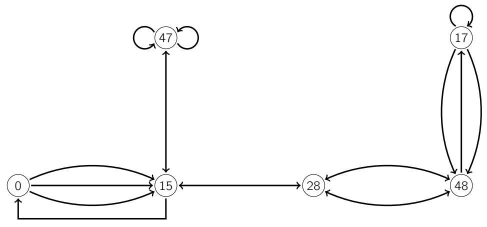


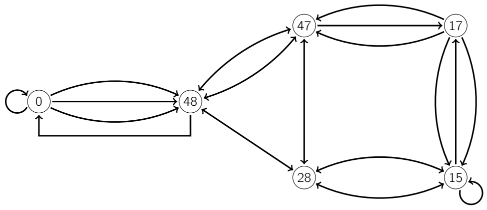
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{53}$



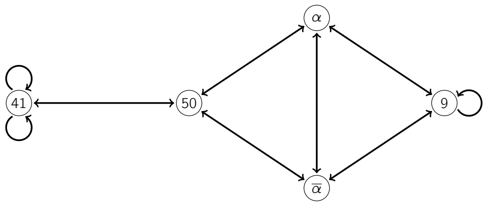


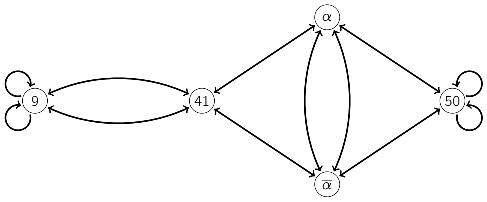
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{59}$



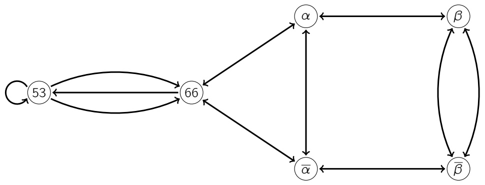


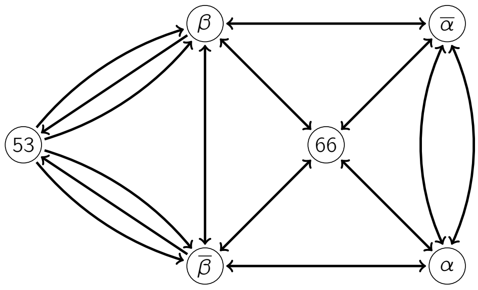
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{61}$



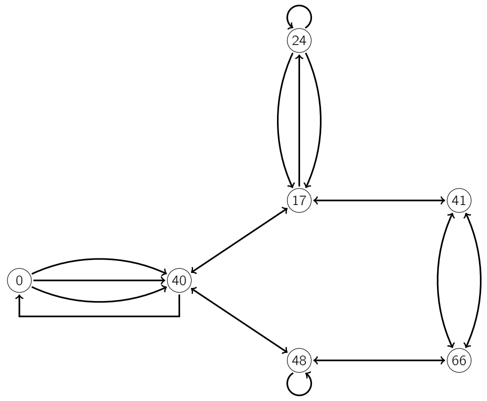


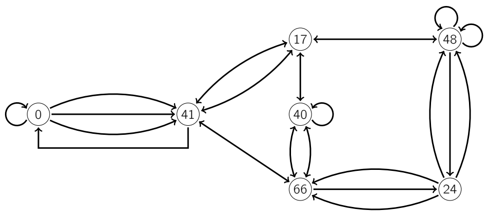
2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{67}$





2 and 3 supersingular isogeny graphs over $\overline{\mathbb{F}}_{71}$

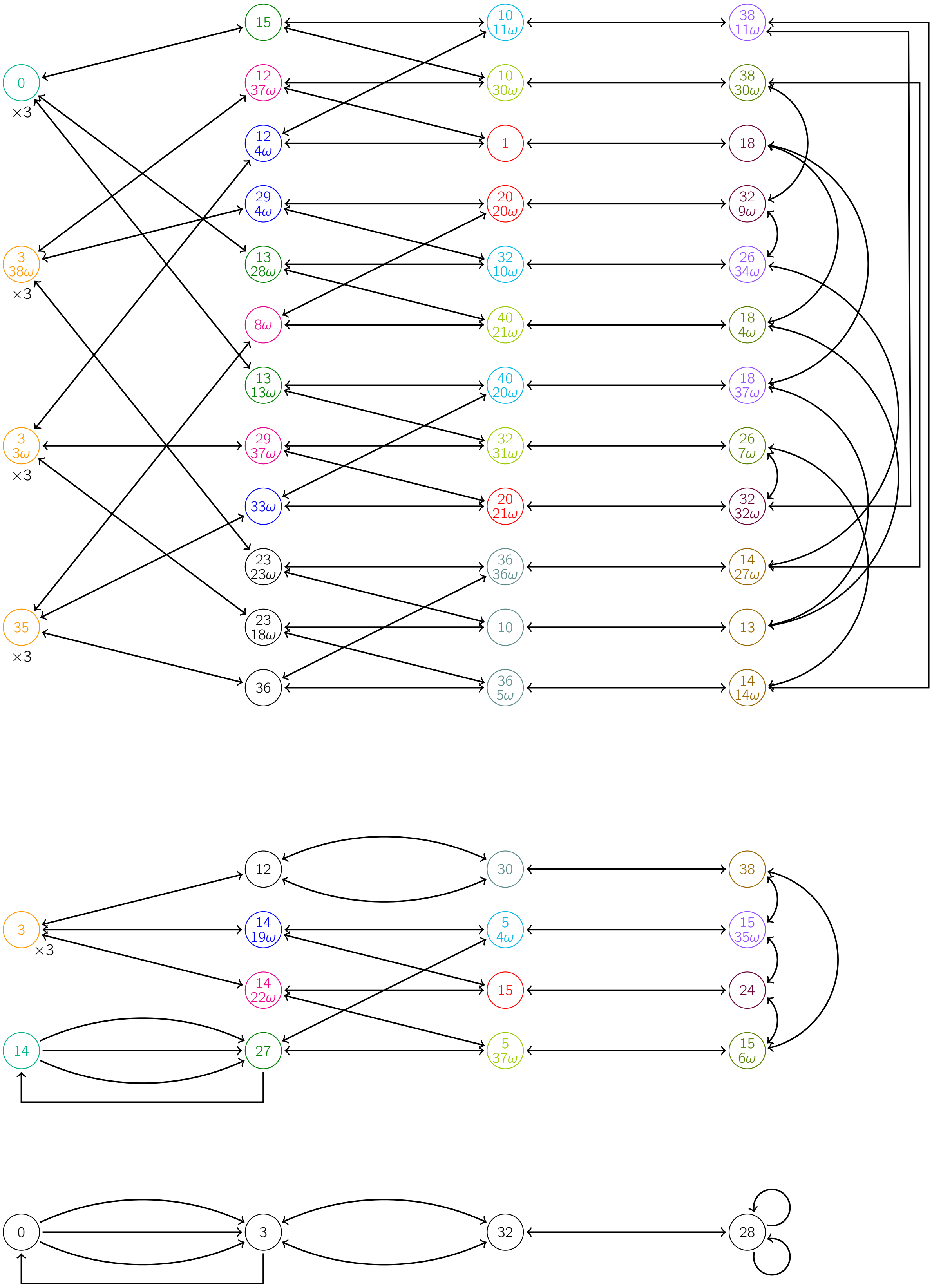




Supersingular isogeny graphs with level structure

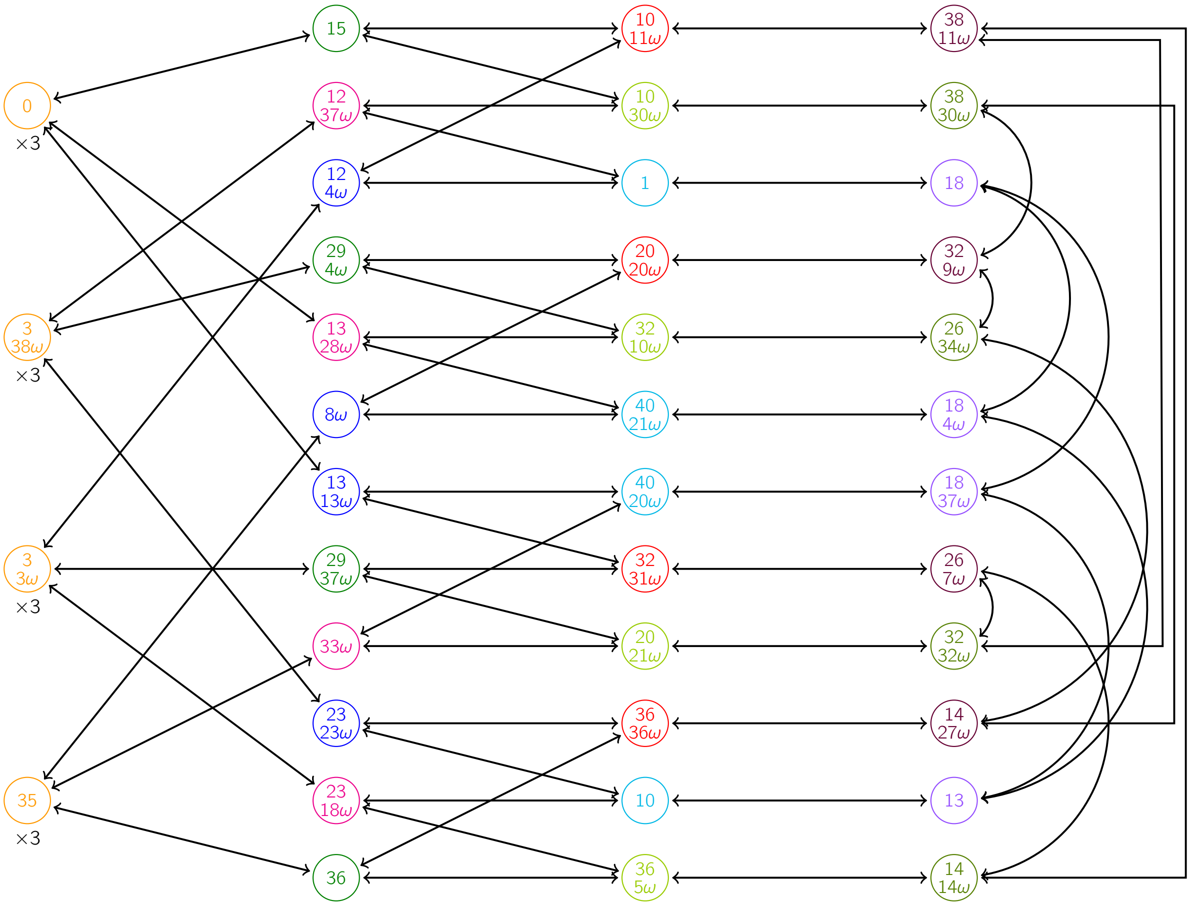
2-isogeny graphs over $\overline{\mathbb{F}}_{41}$ with $\Gamma_0(3)$ and $\Gamma(3)$ structure

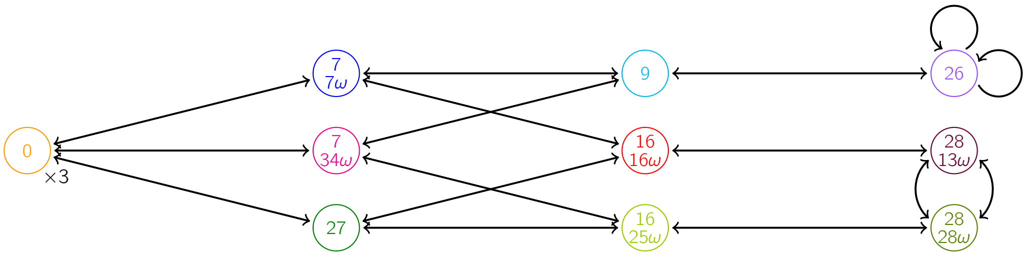
In the picture below we present the cover $G_2(41, \Gamma(3)) \to G_2(41, \Gamma_0(3)) \to G_2(41, \Gamma(1))$. The colors represent the images of the map $X(3) \to X_0(3)$.

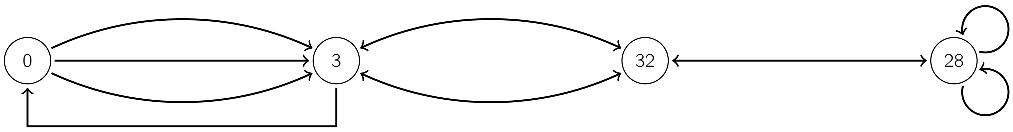


2-isogeny graphs over $\overline{\mathbb{F}}_{41}$ with $\Gamma_{ns}^{+}(3)$ and $\Gamma(3)$ structure

We picture now the cover $G_2\left(41, \Gamma(3)\right) \to G_2\left(41, \Gamma_{ns}^{+}(3)\right) \to G_2\left(41, \Gamma(1)\right)$. Once again, the colors represent the images of the map $X(3) \to X_{ns}^{+}(3)$.

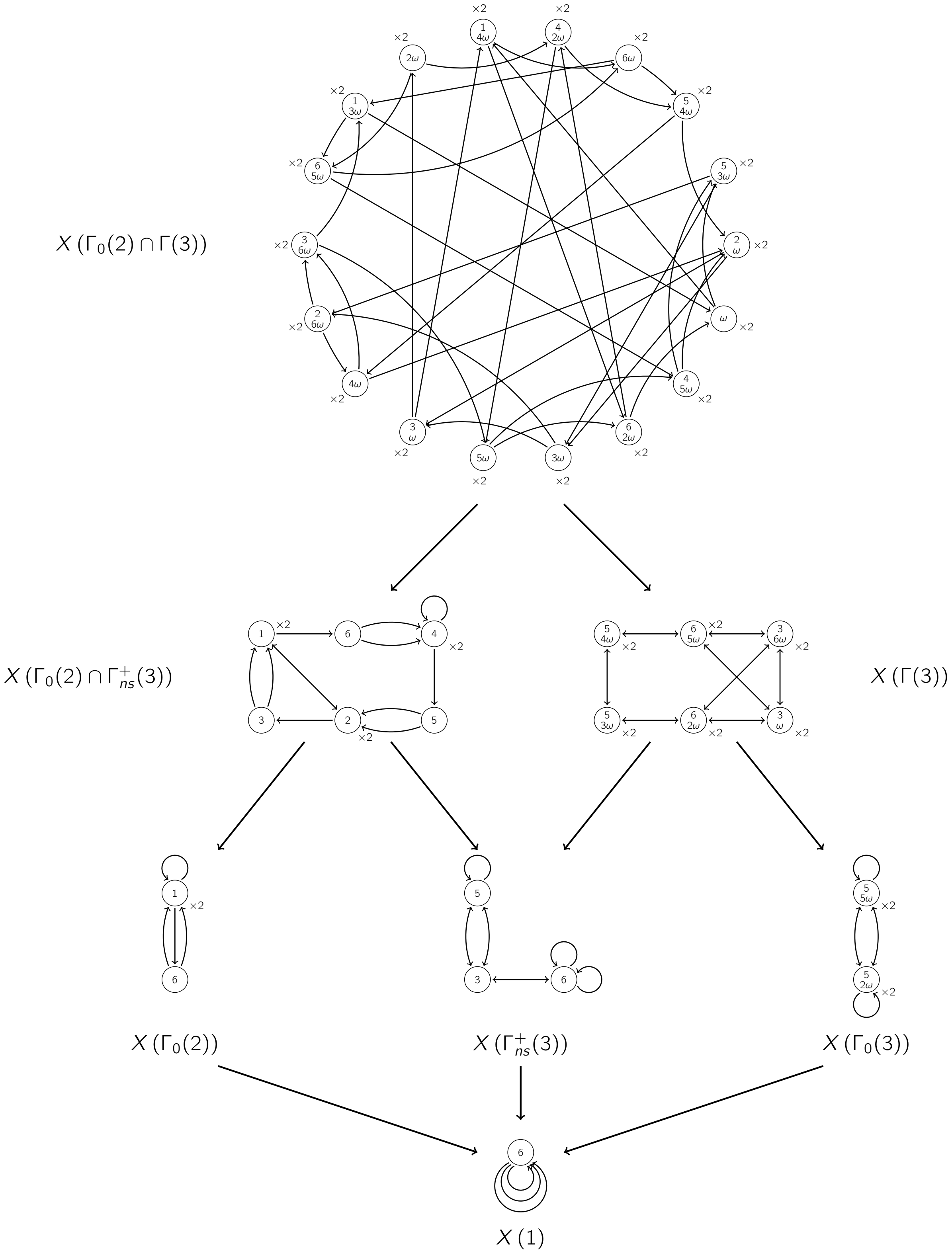






2-isogeny graphs over $\overline{\mathbb{F}}_7$ with $\Gamma_0(2) \cap \Gamma(3)$ structure

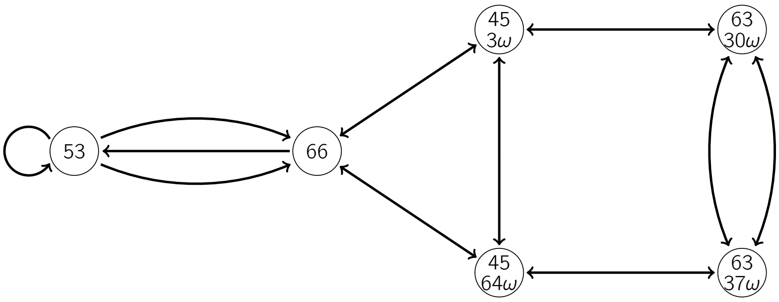
Following Figure 3.6, we draw below the corresponding covering graphs over $\overline{\mathbb{F}}_7$.



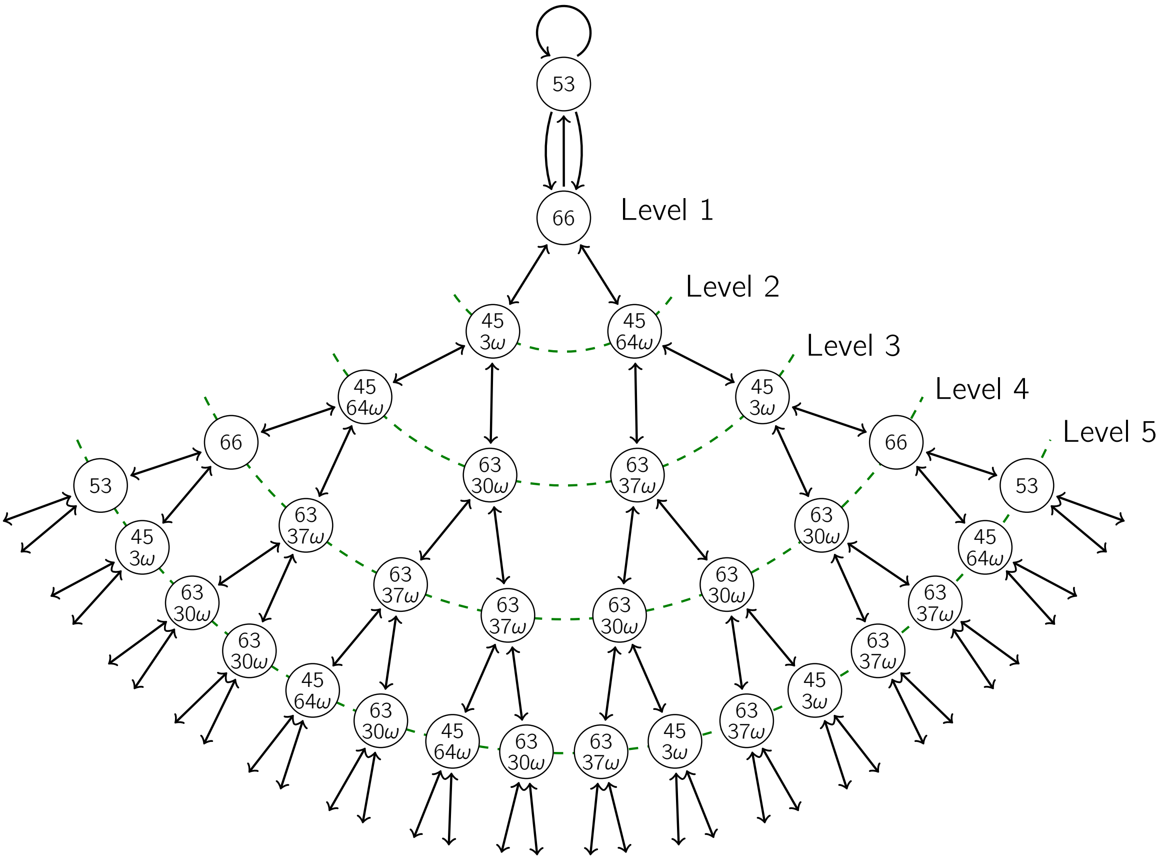
Oriented supersingular isogeny graphs

Orienting the supersingular isogeny graph over $\overline{\mathbb{F}}_{67}$ by $\mathbb{Q}(\sqrt{-1})$

The unoriented supersingular 2-isogeny graph over $\overline{\mathbb{F}}_{67}$; we note $\mathbb{F}_{67^2} = \mathbb{F}_{67}[\omega]$ where $\omega^2 + 1 = 0$.

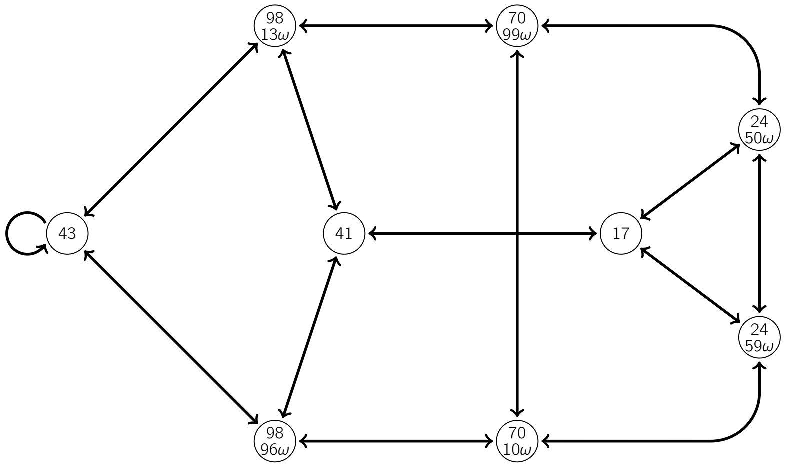


We endow the above graph with the orientation by $\mathcal{O}_K$ where $K = \mathbb{Q}(\sqrt{-1})$ and we get the following infinite volcano.

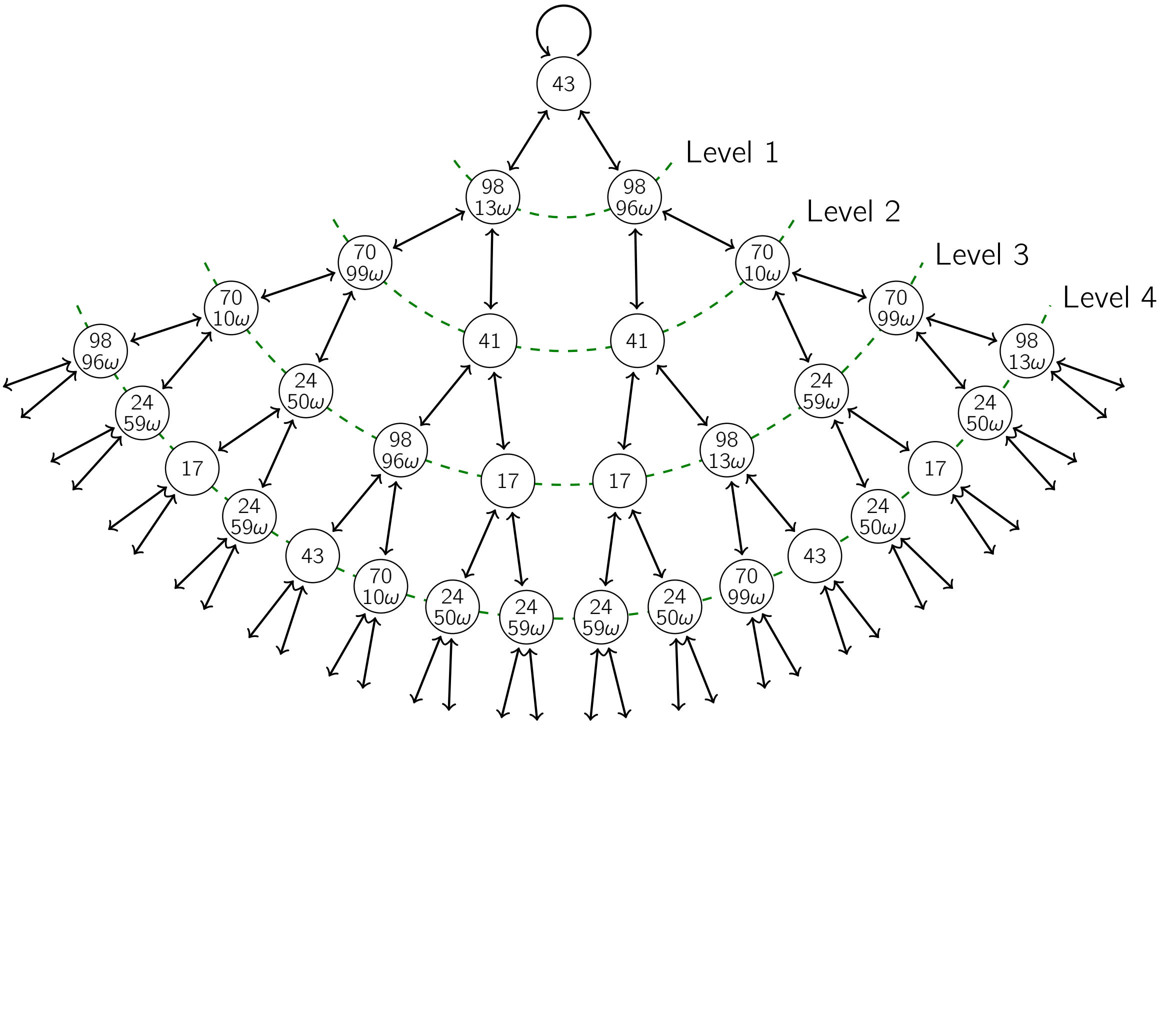


Orienting the supersingular isogeny graph over $\overline{\mathbb{F}}_{109}$ by $\mathbb{Q}(\sqrt{-2})$

The unoriented supersingular 2-isogeny graph over $\overline{\mathbb{F}}_{109}$; we note $\mathbb{F}_{109^2} = \mathbb{F}_{109}[\omega]$ where $\omega^2 + 2 = 0$.

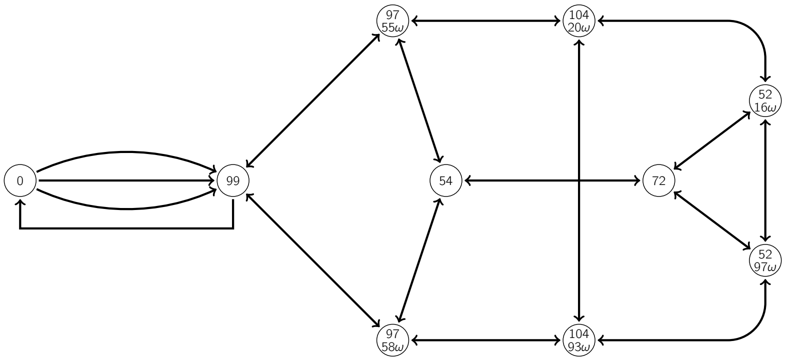


We orient the above graph by $\mathcal{O}_K$ where $K = \mathbb{Q}(\sqrt{-2})$ and we get the following infinite volcano.

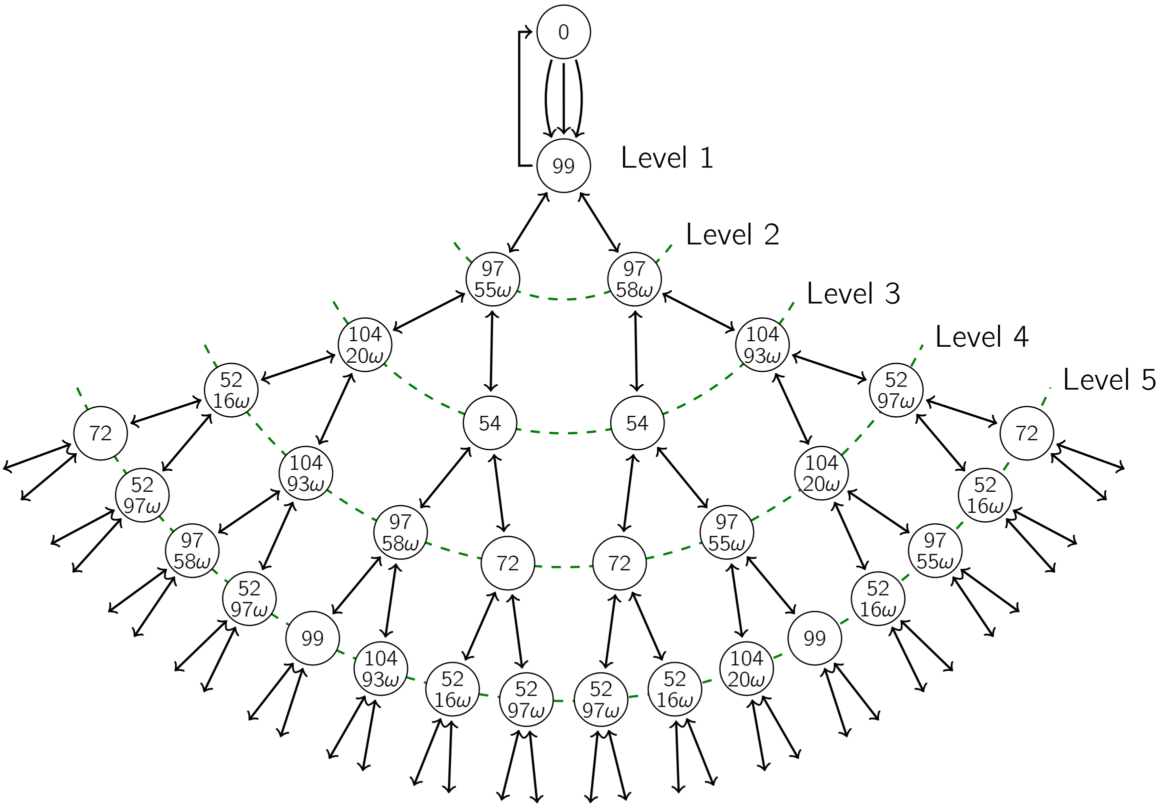


Orienting the supersingular isogeny graph over $\overline{\mathbb{F}}_{113}$ by $\mathbb{Q}(\sqrt{-3})$

The unoriented supersingular 2-isogeny graph over $\overline{\mathbb{F}}_{113}$; we note $\mathbb{F}_{113^2} = \mathbb{F}_{113}[\omega]$ where $\omega^2 + 3 = 0$.

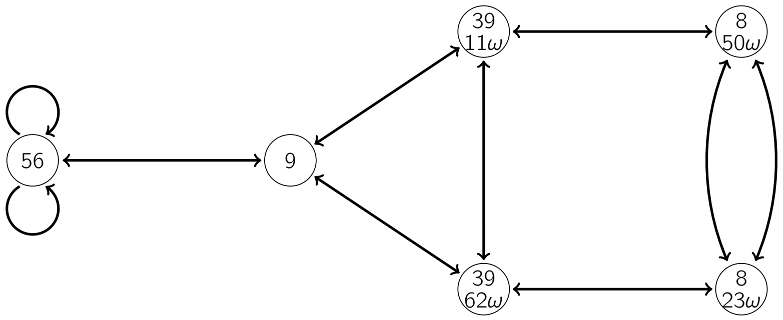


The orientation of the above graph by $\mathcal{O}_K$ (where $K = \mathbb{Q}(\sqrt{-3})$) results in the following infinite volcano.

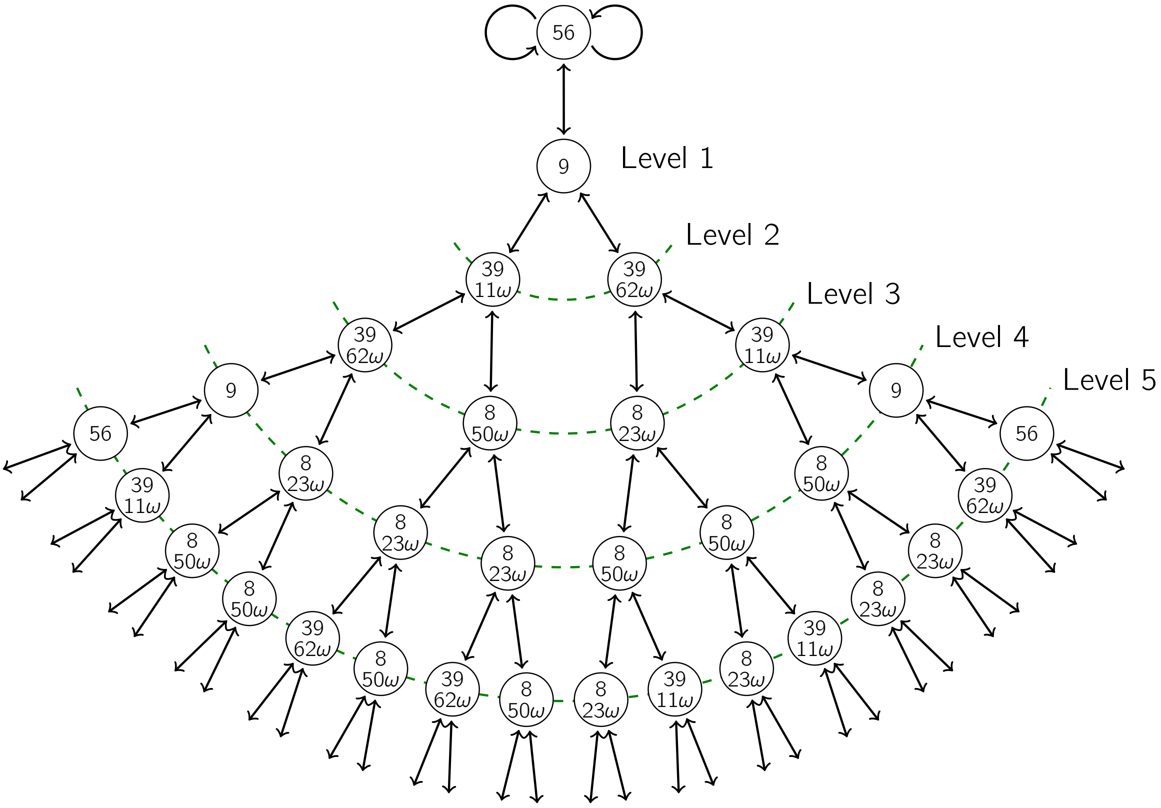


Orienting the supersingular isogeny graph over $\overline{\mathbb{F}}_{73}$ by $\mathbb{Q}(\sqrt{-7})$

The unoriented supersingular 2-isogeny graph over $\overline{\mathbb{F}}_{73}$; we note $\mathbb{F}_{73^2} = \mathbb{F}_{73}[\omega]$ where $\omega^2 + 5 = 0$.

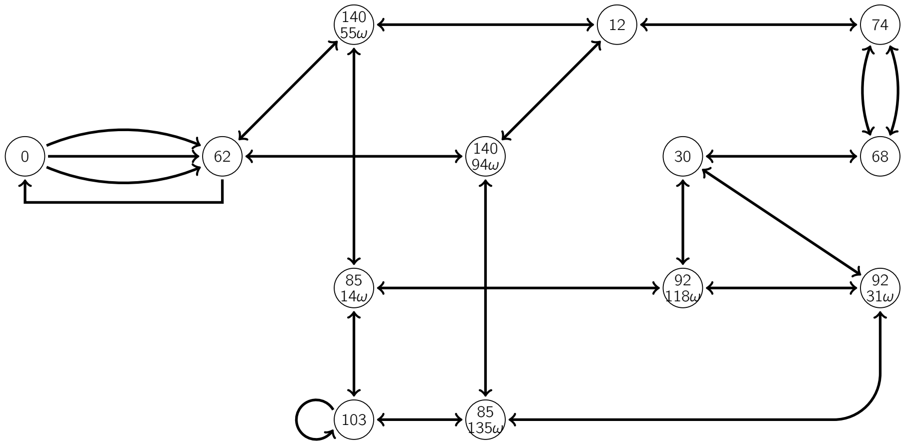


We endow the above graph with the orientation by $\mathcal{O}_K$ where $K = \mathbb{Q}(\sqrt{-7})$ and we get the following infinite volcano.

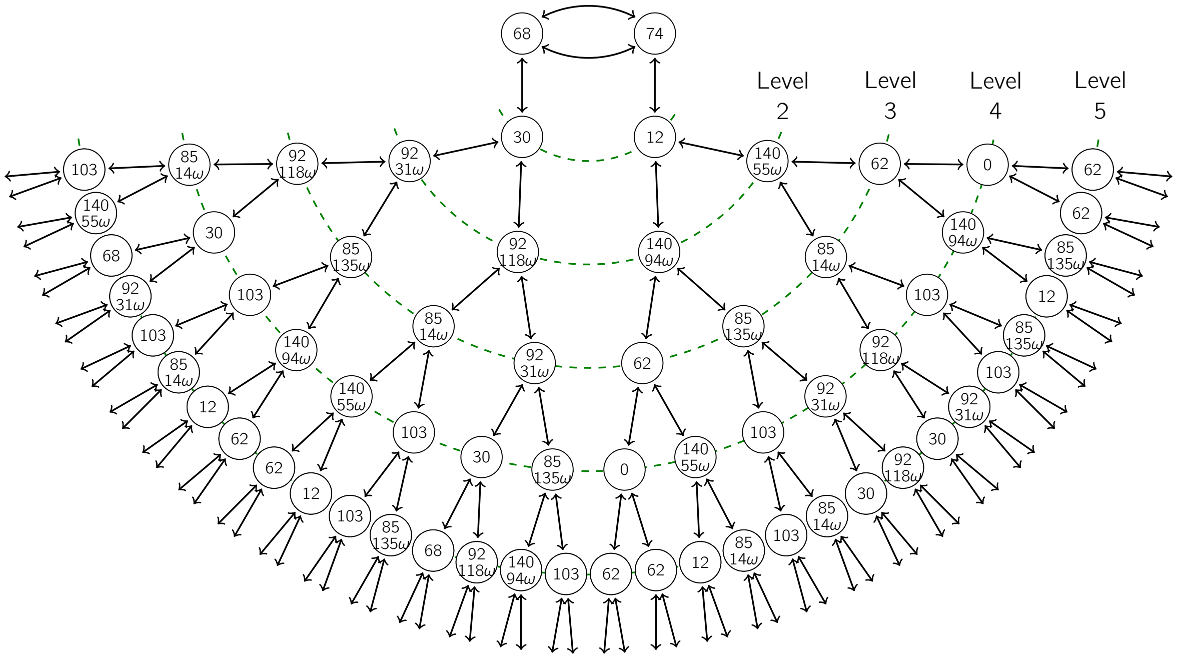


Orienting the supersingular isogeny graph over $\overline{\mathbb{F}}_{149}$ by $\mathbb{Q}(\sqrt{-15})$

The unoriented supersingular 2-isogeny graph over $\overline{\mathbb{F}}_{149}$; we note $\mathbb{F}_{149^2} = \mathbb{F}_{149}[\omega]$ where $\omega^2 + 2 = 0$.



The orientation of the above graph by $\mathcal{O}_K$ (where $K = \mathbb{Q}(\sqrt{-15})$) produces the following infinite volcano.



68
74
Level 2
Level 3
Level 4
Level 5
30
12
103
85 14ω
92 118ω
92 31ω
140 55ω
62
0