Sato–Tate Groups of Generic Superelliptic Curves

Nir Elber

2024–2025

Contents

I encourage my fellow graduates to think about what they knew when they started out here, and how much has been layered on top of that since then. I think you'll find it's more than you had imagined.

—Miles Kretschmer [Kre23]

Chapter 0
Introduction

What we didn't do is make the construction at all usable in practice!
This time we will remedy this.

—Kiran S. Kedlaya, [Ked21]

0.1 Overview

Over the past few decades, there has been a growing interest in understanding how the geometry of a space (such as a smooth projective variety) affects its arithmetic.

One example of these effects arises in the form of the Sato–Tate conjecture, which takes an abelian variety A over $\mathbb{Q}$ and predicts the distribution of the point-counts $\#A(\mathbb{F}_p)$ (suitably interpreted) as the primes p varies. Here, one finds that the "geometric" invariant $\mathrm{End}_{\mathbb{C}}(A)$ essentially determines the desired distribution. We refer to section 3.1 for a more precise discussion, but approximately speaking, the point is that one expects a "motivic monodromy group" to control this distribution, and the motivic monodromy group can be computed either in a geometric situation over $\mathbb{C}$ or understood via such point-counts in an arithmetic situation.

To be slightly more explicit, there are various monodromy groups at play: in the complex analytic situation, there is the Mumford–Tate group $\mathrm{MT}(A)$, and in the ℓ-adic situation, there is the ℓ-adic monodromy group $G_\ell(A)$. There are conjectural relations between these, and these conjectures codify the interplay between geometry and arithmetic; for example, the Mumford–Tate conjecture predicts that $\mathrm{MT}(A)_{\mathbb{Q}_\ell} = G_\ell(A)^\circ$. Ultimately, to understand point-counts, one becomes interested in the groups $G_\ell(A)$, but this group is difficult to compute directly, so it is frequently profittable to compute $\mathrm{MT}(A)$ instead and then use one of the aforementioned conjectures.

In this article, we are interested in the effect of so-called "exceptional" geometry on arithmetic, continuing the work of [GGL24]. The exceptional geometry we are interested in concerns exceptional Hodge classes, which are Hodge classes on A (or a power of A) which are not generated by an endomorphism of A or the polarization of A. The absence of such classes gives control of the geometry of A and hence makes $\mathrm{MT}(A)$ and $G_\ell(A)$ easy to compute. As another application, in the absence of exceptional classes, one knows the Hodge conjecture for all powers of A, so exceptional geometry is in some sense "the enemy" of proving the Hodge conjecture.

0.1.1 Fermat Curves

Roughly speaking, most abelian varieties do not support exceptional classes, so it requires some effort to find abelian varieties with exceptional classes in nature (and then prove and study their existence!). In [GGL24], Gallese, Goodson, and Lombardo are able to control exceptional classes in the Jacobians of the hyperelliptic "Fermat" curves

$$y^2 = x^N + 1$$

as $N \geq 1$ varies over positive integers. Namely, they are able to write down an algorithm which computes the groups MT and G_ℓ for moderately sized N (say, $N \leq 100$), and they are able to prove general results in certain cases (such as N prime). It is still true that some N fail to support exceptional classes, such as when N is a prime, but composite N frequently support exceptional geometry, which must be understood to execute the computation.

The present article can be considered a continuation of the work of [GGL24]. For example, the authrors there remark that their methods should be able to be used to compute MT and G_ℓ for the Jacobians of quotients of the smooth projective Fermat curve

$$X_N \colon X^N + Y^N + Z^N = 0,$$

which includes the hyperelliptic curves $y^2 = x^N + 1$ above. This is carried out in section 4.2; we note that the main theorem is Theorem 4.33, where we provide an explicit description of the Galois action on (absolute) Hodge classes in terms of Galois action on certain explicitly computed periods, but we will not give the statement in the introduction because it is somewhat technical.

> **Remark 0.1.** As an aside, we note that the authors of [GGL24] recourse to more general Fermat hypersurfaces
>
> $$X_0^N + X_1^N + \cdots + X_m^N = 0.$$
>
> in order to understand powers of the Fermat curve X_N. This theory rests on somewhat technical algebraic geometry due to Deligne [Del18, Section 7]. In this article, we rebuild the thoery of [GGL24] while only handling powers of X_N directly, allowing us to avoid Deligne's algebraic geometry. The key point is that a careful analysis of the Künneth isomorphism allows one to gain the same level of control on the Hodge classes of a power of X_N as one would get with embedding in a Fermat hypersurface. This is carried out in section 4.2.1.

Having access to more general quotients allows us to see more geometry. To explain one example, we recall the definition of $G_\ell(A)$. Given an abelian variety A defined over a number field K, one can use the Galois action on the Tate module $V_\ell A$ of A to define a Galois representation

$$\rho_\ell \colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(V_\ell A).$$

Here, $V_\ell A$ turns out to be a vector space ver $\mathbb{Q}_\ell$ of dimension $2 \dim A$. We then define $G_\ell(A)$ to be the smallest algebraic $\mathbb{Q}_\ell$-subgroup containing the image of ρ_ℓ. The Mumford–Tate conjecture explains that one expects to recover $G_\ell(A)^\circ$ from the complex geometry of A, so it becomes interesting to understand the quotient $G_\ell(A)/G_\ell(A)^\circ$, which we note is finite because $G_\ell(A)$ is an algebraic group. In light of the definition of $G_\ell(A)$, we see that we are interested in the pre-image $\rho_\ell^{-1}(G_\ell(A)^\circ)$; this needs to be a finite-index open subgroup of $\mathrm{Gal}(\overline{K}/K)$, so there is a finite extension K_A^{conn} of K such that $\rho_\ell(\sigma) \in G_\ell(A)^\circ$ if and only if σ fixes K_A^{conn}.

In [GGL24, Theorem 7.1.1], the authors find that their hyperelliptic curves $y^2 = x^N + 1$ all have K_A^{conn} to be a multiquadratic extension of $\mathbb{Q}(\zeta_N)$, and they provide an algorithm to compute it. Further, they find that the prime-power case will always have $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_N)$. One can now ask if one can hope for such control for general quotients of the Fermat curve. Well, [Del18, Theorem 7.15] explains that the extension $K_A^{\mathrm{conn}}/\mathbb{Q}(\zeta_N)$ should always be abelian. However, it turns out that one cannot hope for much more than this.

Example 0.2. In Proposition 4.79, we show that the Jacobian of the superelliptic curve

$$y^9 = x\left(x^2+1\right),$$

which is a quotient of the Fermat curve $X^{18}+Y^{18}+Z^{18}=0$, has $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_{18}, \sqrt[18]{432})$, which is a degree-$18$ cyclic extension of $\mathbb{Q}(\zeta_{18})$.

Example 0.3. The Jacobian of the previous example is not simple. At the cost of having slightly higher dimension, one can show something similar for the Jacobian of $y^{11} = x^2\left(x^2+1\right)$, but now this Jacobian is simple.

In section 4.2, we work out the example curve $y^9 = x^3+1$ in detail. Here, one does find exceptional classes, but we still have $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$.

To work with more examples, we need a finer understanding of the periods of Fermat curves [GGL24, Sections 6.3–6.4]. This is accomplished in section 4.3. In short, it turns out that we need to understand the algebraicity properties of certain products of Γs, and these products can be understood in terms of the (combinatorial) theory of distributions. In [GGL24], the authors only work with the periods which can come from the hyperelliptic Fermat curve, but we work with periods of the full Fermat curve. Here is an example of what we can prove.

Theorem 4.77. Let K_A^{conn} be the connected monodromy field of the Jacobian A of the Fermat curve X_N, and define the field

$$K_N = \mathbb{Q}(i,\zeta_{2N})\left(\{p^{p/N} : \text{prime } p \mid N\}\right).$$

(a) We have $K_N \subseteq K_A^{\mathrm{conn}}(i,\zeta_{2N})$.

(b) The extension $K_A^{\mathrm{conn}}(i,\zeta_{2N})/K_N$ is multiquadratic.

(c) If N is odd or divisible by 4, then

$$\log_2[K_A^{\mathrm{conn}}(i,\zeta_{2N}) : K_N] \leq 2^{\omega(N)-1}-1,$$

where $\omega(N)$ is the number of distinct prime factors of N.

Remark 0.4. It would be interesting to know if the upper bound in (c) is sharp. This seems to be unknown unless $\omega(N)=1$.

0.1.2 Beyond CM

One aspect of these Fermat curves is that they have so many automorphisms (given by multiplying X or Y by an Nth root of unity) that their Jacobians have complex multiplication. Complex multiplication aides the computation in a few key ways: in this case, $\mathrm{MT}(A)$ and $G_\ell(A)$ are both tori, thus making them much easier to control. For example, the Mumford–Tate conjecture is known in this case, and there exist algorithms to compute $\mathrm{MT}(A)$ from certain combinatorial data attached to A.

As such, to the author's knowledge, the literature does not have an example computation of $G_\ell(A)$ when A does not have complex multiplication and is not fully of Lefschetz type.[1] In this article, we work out such an example. Admittedly, we do not go far from complex multiplication: where complex multiplication would require $\mathrm{End}(A)\otimes_{\mathbb{Z}}\mathbb{Q}$ to contain a CM field of dimension $2\dim A$, we work with certain abelian varieties A

[1] Roughly speaking, "fully of Lefschetz type" means that all Hodge classes on A can be explained by endomorphisms and the polarization. In type III, it turns out that these classes do imply the existence of an exceptional class, which is the difference between not supporting exceptional cycles and being "fully of Lefschetz type."

such that $\mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$ contains a CM field of dimension $\dim A$. Our limitations are rather technical, and we expect that one can do much better.

As an example difficulty, let's focus on computing $\mathrm{MT}(A)$. Recall that $\mathrm{MT}(A)$ is a connected reductive algebraic group defined over $\mathbb{Q}$, so we can split up its computation into computing the derived subgroup $\mathrm{MT}(A)^{\mathrm{der}}$ and the neutral component $Z(\mathrm{MT}(A))^\circ$ of the torus. In section 2.2, we explain how the current arguments used to understand $\mathrm{MT}(A)$ for A with complex multiplication can be used to compute $Z(\mathrm{MT}(A))^\circ$. To explain this result, we pick up some notation: set $E \coloneqq Z(\mathrm{End}(A))$, and then one can diagonalize the action of E on $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A(\mathbb{C}), \mathbb{C})$ to produce a piece of combinatorial data called the "signature" $\Phi\colon \mathrm{Hom}(E, \mathbb{C}) \to \mathbb{Z}_{\geq 0}$; for brevity, we will set $\Sigma_E \coloneqq \mathrm{Hom}(E, \mathbb{C})$. It turns out that one can embed $Z(\mathrm{MT}(A))^\circ$ into the torus $\mathrm{T}_E \coloneqq \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E}$, and our first main result explains how to recover this subtorus.

Corollary 2.77. Fix an abelian variety A over $\mathbb{C}$ such that $Z(\mathrm{End}(A))$ equals a CM algebra E, and define $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. Let $\Phi\colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ be the signature defined in Lemma 2.72. Then $Z(\mathrm{MT}(V))^\circ \subseteq \mathrm{T}_E$ has cocharacter group equal to the smallest saturated Galois submodule of $\mathrm{X}_*(\mathrm{T}_E) = \mathbb{Z}[\Sigma_E^\vee]$ containing

$$\sum_{\sigma \in \Sigma_E} \Phi(\sigma)\sigma^\vee.$$

Remark 0.5. In fact, a careful reading of the arguments in section 2.2 reveal that we are actually able to compute an explicit power of $Z(\mathrm{MT}(A))$, which technically contains more information. For example, one could provide a sufficient condition for $Z(\mathrm{MT}(A))$ being disconnected.

It remains to compute $\mathrm{MT}(A)^{\mathrm{der}}$. Under certain simplifying hypotheses given above, we work this out in Proposition 2.152, which we restate below for convenience. Here $\mathrm{L}(A)$ is the Lefschetz group, which is intuitively what $\mathrm{MT}(A)$ would be in the absence of exceptional classes.

Proposition 2.152. Fix a geometrically simple abelian variety A over a number field K. Suppose that $E = Z(\mathrm{End}_{\overline{K}}(A))$ equals a CM field such that $\dim A = \dim E$. Letting Φ be the corresponding signature, we further suppose that $\Phi(\sigma) = 1$ for exactly two $\sigma \in \Sigma_E$. Then we show the Mumford–Tate conjecture holds for A, and

$$\mathrm{MT}(A)^{\mathrm{der}} = \mathrm{L}(A)^{\mathrm{der}}.$$

The argument proving Proposition 2.152 achieves something slightly stronger, but it is technical to state and not required for our application. In short, the idea of the proof is to upgrade the fact that the real Lie groups $\mathrm{SU}(2,0)$ and $\mathrm{SU}(1,1)$ are not isomorphic using the Galois action.

Now that we understand $\mathrm{MT}(A)$, we would like to upgrade this to an understanding of $G_\ell(A)$. After the Mumford–Tate conjecture, we (roughly speaking) need to understand the quotient $G_\ell(A)/G_\ell(A)^\circ$, whch section 2.4.3 explains that this amounts to computing the Galois action on certain "Tate classes." Thus, the trick is to not look at a particular Galois representation ρ_ℓ but instead a family of them. We can engineer everything so that generic members of the family satisfy the properties needed for the rest of the present subsection to go through. Then our last trick is ensure that some special members of the family are quotients of a Fermat curve, where we know the Galois action! In this way, we can "transport" the understanding of the Galois action afforded by the Fermat curves to a generic curve. Here is the toy result we are able to prove.

Theorem 4.39. For given $\lambda \in \mathbb{Q}(\zeta_9) \setminus \{0, 1\}$, define A to be the Jacobian of the proper curve $\widetilde{C}$ with affine chart $y^9 = x(x-1)(x-\lambda)$. Suppose that A does not have complex multiplication. Then we show $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$, and we compute $\mathrm{ST}(A)$.

0.2 Odds and Ends

In this section, we explain some existential properties of this article.

0.2.1 What Is in This Article

Let's take a moment to explain the layout. For the most part, the exposition is arranged topically (by chapter, then section), and we have done our best to remove situations which would require forward references.

As a very brief overview, chapter 1 explains all the Hodge theory we will require and then ends by introducing the category of motives using Deligne's theory of absolute Hodge classes. Chapter 2 explains everything we will want to know about abelian varieties, detailing in particular properties and computational aspects of the ℓ-adic representation. Chapter 3 motivates and states the Sato–Tate conjecture for abelian varieties and then indicates some tools used in the proof of some of the known cases. Lastly, Chapter 4 explains how to compute the ℓ-adic monodromy group and the periods of the Fermat curves.

Because there are certain subsections whose purpose may not immediately be clear in a linear read of the article, we also take a moment to explain some of the stories present in the exposition.

- Computation: a major goal of the thesis is to compute ℓ-adic monodromy groups. Proposition 2.86 explains how to compute the center of the Mumford–Tate group, which the Mumford–Tate conjecture relates to ℓ-adic monodromy. Our needed case of the Mumford–Tate conjecture is given in Proposition 2.152; we note that Lemma 1.62 is a key input. From here, Proposition 2.159 explains how to treat the disconnected parts of the group, and this is the discussion used in the example discussions of sections 3.1.3 and 4.2.4. Our most complicated example is given in Proposition 4.79, where we require the algorithmic discussion of Fermat periods discussed in (the rather painful) section 4.3.

- Motives: after explaining what is required about Hodge structures in section 1.1, we may put in quite a bit of effort in section 1.3 to define a category of motives. These notions are then used to define the motivic Galois group in section 2.4.4, which helps contextualize our monodromy groups (see Example 1.143 and Remark 2.120) and the Mumford–Tate conjecture (see Conjecture 2.168). Motives are then used in the proof of that the Mumford–Tate conjecture implies the Algebraic Sato–Tate conjecture in Theorem 3.23.

- Complex multiplication: complex multiplication is defined for abelian varieties in section 2.1.5, where it serves as a basic case for many of conjectures and computations; for example, the Fermat Jacobians have complex multiplication. We point out that the Mumford–Tate conjecture is proven for abelian varieties in Example 2.145 by combining Propositions 2.86 and 2.143; these propositions also explain how to the ℓ-adic monodromy group in practice. A notable input is the Fundamental theorem of complex multiplication, a version of which is stated in Theorem 2.138. Another one of its applications is to prove the Sato–Tate conjecture in this case, which is done in Theorem 3.117.

0.2.2 What Is Not in This Article

What follows are some topics which potentially fit in with the theme of the current article, but the author did not find adequate time to think through them in detail and write down their details. Any reader is encouraged to email the author if they have ideas or want further explanation.

1. A discussion of rigid cohomology and Kedlaya's algorithm to compute Frobenius matrices.
   (a) This would allow us to computationally verify Theorem 4.33.
   (b) This would allow us to form p-adic analogues of many parts of our computation, such as Proposition 2.152.
   (c) In some cases, one could supplement a p-adic approximation (e.g., via the Cartier–Manin matrix) with the Fundamental theorem of complex multiplication to be able to compute Frobenius matrices.

2. Vertical Sato–Tate considerations.
   (a) It may be possible to prove a vertical Sato–Tate result for arbitrary Shimura curves, such as the one considered in this article, imitating [Del80, Theorem 3.5.3] or [Kat88, Theorem 3.6].

(b) Computation of the center part of the relevant monodromy group requires the computation of the Frobenius at at least one point, such as a special point. As such, one could apply computations from 1(c) above.

3. More on Fermat periods.

(a) It would be interesting to lower-bound the degree of the connected monodromy field of the Fermat curve (as an extension of its endomorphism field). Note that Proposition 4.73 provides an upper bound.

(b) In [GGL24, Theorem 9.3.13], the authors prove a weak version of a Gross–Koblitz formula over $\mathbb{Q}$. It should be possible to work with arbitrary characters α of constant weight using Theorem 4.33 (and the idea of Lemma 4.76).

0.2.3 Acknowledgements

The author is indebted to many people for the existence of this article. Most importantly, the author is extremely grateful to his advisor Yunqing Tang for many, many patient and enlightening conversations, for example by suggesting the key ideas that went into the main results. The author would also like to thank Sug Woo Shin and Yiannis Sakellaridis for the opportunity to speak about the Sato–Tate conjecture and helpful conversations to this end. Additionally, the author thanks Hannah Larson for help understanding Hurtwitz spaces. None of this mathematics would have been possible without their constant encouragement.

Most of the research in this article was conducted at the University of California, Berkeley, and the remainder was conducted at Johns Hopkins University. Simply put, this article would not be possible without the extremely welcoming mathematics communities at both places. In particular, the author would like to thank Jad Damaj, Sophie McCormick, Julie Shields, and Toan Pham for diverting conversations; and the author would like to thank Sam Goldberg, Mitch Majure, and Justin Wu for productive conversations. The author also thanks Miles Kretschmer for permission to use the quote in the table of contents.

There are also ennumerable people who have helped the author get to where he is today who were not directly involved in the creation of this article. To name just a few, the author thanks his parents Ron Elber and Virginia Yip and his sister Nurit Elber for believing in his blooming academic carreer. And most importantly, the author is forever in debt to Hui Sun for constant support and companionship. Without her, the author would be without soul.

Chapter 1
A Little Hodge Theory

Once we explicitely know a Mumford-Tate group, we can let it work for us.

—Moonen [Moo, (5.5)]

In this chapter, we define the notion of a Hodge structure as well as some related groups (the Mumford–Tate group and the Hodge group). Our exposition follows Moonen's unpublished notes [Moo; Moo99] and Lombardo's master's thesis [Lom13, Chapter 3]. Throughout, we find motivation from geometry (and in particular the cohomology of complex varieties), but we will review cohomology only later.

1.1 Hodge Structures

Cohomology of a variety frequently comes with some extra structure. On the étale site, we will later get significant utility of the fact that étale cohomology is a Galois representaion. On the analytic site, the corresponding structure is called a "Hodge structure."

1.1.1 Definition and Basic Properties

Here is our defintion.

Definition 1.1 (Hodge structure)**.** A $\mathbb{Q}$-*Hodge structure* is a finite-dimensional vector space $V \in \mathrm{Vec}_{\mathbb{Q}}$ such that $V_{\mathbb{C}}$ admits a decomposition

$$V_{\mathbb{C}} = \bigoplus_{p,q\in\mathbb{Z}} V_{\mathbb{C}}^{p,q}$$

where $V_{\mathbb{C}}^{p,q} = \overline{V_{\mathbb{C}}^{q,p}}$. For fixed $m \in \mathbb{Z}$, if $V_{\mathbb{C}}^{p,q} \neq 0$ unless $p+q=m$, we say that V is *pure of weight* m. We let $\mathrm{HS}_{\mathbb{Q}}$ denote the category of $\mathbb{Q}$-Hodge structures, where a morphism of Hodge structures is a linear map preserving the decomposition over $\mathbb{C}$. In the sequel, it may be helpful to note that one can bring this definition down to $\mathbb{Z}$ as well.

Example 1.2. We give the "Tate twist" $\mathbb{Q}(1) \coloneqq 2\pi i\mathbb{Q}$ a Hodge structure of weight -2 where the only nonzero entry is $\mathbb{Q}(1)^{-1,-1} = \mathbb{Q}(1)$.

Example 1.3. Given a complex projective smooth variety X, the Betti cohomology $\mathrm{H}^n_{\mathrm{B}}(X,\mathbb{Q})$ admits a Hodge structure via the comparison isomorphisms: we find that

$$\mathrm{H}^n_{\mathrm{B}}(X,\mathbb{C}) \simeq \bigoplus_{p+q=n} \mathrm{H}^{p,q}(X),$$

where $\mathrm{H}^{p,q}(X) \coloneqq \mathrm{H}^q(X,\Omega^p_{X/\mathbb{C}})$. This construction is even functorial: a morphism of complex projective smooth varieties $\varphi\colon X \to Y$ induces a morphism of Hodge structures $\varphi^*\colon \mathrm{H}^n_{\mathrm{B}}(Y,\mathbb{Q}) \to \mathrm{H}^n_{\mathrm{B}}(X,\mathbb{Q})$.

Perhaps one would like to check that the category $\mathrm{HS}_{\mathbb{Q}}$ is abelian. The quickest way to do this is to realize $\mathrm{HS}_{\mathbb{Q}}$ as a category of representations of some group. The relevant group is the Deligne torus.

Notation 1.4 (Deligne torus)**.** Let $\mathbb{S} \coloneqq \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\,\mathbb{G}_{m,\mathbb{C}}$ denote the Deligne torus. We also let $w\colon \mathbb{G}_{m,\mathbb{R}} \to \mathbb{S}$ denote the *weight cocharacter* given by $w(r) \coloneqq r \in \mathbb{C}$ on $\mathbb{R}$-points.

Remark 1.5. One can realize $\mathbb{S}$ more concretely as

$$\mathbb{S}(R) = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in \mathrm{GL}_2(R) : a^2+b^2 \in R^\times \right\},$$

where R is an $\mathbb{R}$-algebra. Indeed, there is a ring isomorphism from $R \otimes_{\mathbb{R}} \mathbb{C}$ to $\left\{ \left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right] : a,b \in R \right\}$ by sending $1\otimes 1 \mapsto \left[\begin{smallmatrix} 1 & \\ & 1 \end{smallmatrix}\right]$ and $1 \otimes i \mapsto \left[\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right]$. For example, one can define two characters $z,\overline{z}\colon \mathbb{S}_{\mathbb{C}} \to \mathbb{G}_{m,\mathbb{C}}$ given by $z\colon \left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right] \mapsto a+bi$ and $\overline{z}\colon \left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right] \mapsto a-bi$ so that $(z,\overline{z})$ is an isomorphism $\mathbb{S}_{\mathbb{C}} \to \mathbb{G}^2_{m,\mathbb{C}}$. Thus, the character group $X^*(\mathbb{S})$ is a free $\mathbb{Z}$-module of rank 2 with basis $\{z,\overline{z}\}$, and the action of complex conjugation $\iota \in \mathrm{Gal}(\mathbb{C}/\mathbb{R})$ simply swaps z and $\overline{z}$.

Example 1.6. The following cocharacters of $\mathbb{S}$ will be helpful.

- We define the *weight cocharacter* $w\colon \mathbb{G}_{m,\mathbb{R}} \to \mathbb{S}$ given by $w(r) \coloneqq r \in \mathbb{C}$ on $\mathbb{R}$-points.
- We define the *miniscule cocharacter* $\mu\colon \mathbb{G}_{m,\mathbb{C}} \to \mathbb{S}_{\mathbb{C}}$ given by $\mu(z) \coloneqq (z,1)$ on $\mathbb{C}$-points.

Here is the relevance of $\mathbb{S}$ to Hodge structures.

Lemma 1.7. Fix some $V \in \mathrm{Vec}_{\mathbb{Q}}$. Then a Hodge structure on V has equivalent data to a representation $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$.

Proof. Remark 1.5 informs us that the character group $X^*(\mathbb{S})$ of group homomorphisms $\mathbb{S} \to \mathbb{G}_m$ is a rank-2 free $\mathbb{Z}$-module generated by $z\colon \left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right] \mapsto a+bi$ and $\overline{z}\colon \left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right] \mapsto a-bi$ on $\mathbb{C}$-points.[1] Without too many details, upon passing to the Hopf algebra, one is essentially looking for units in $\mathbb{R}\left[a,b,\left(a^2+b^2\right)^{-1}\right]$, of which there are not many. Note that there is a Galois action by $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ on these two characters $\{z,\overline{z}\}$, given by swapping them. Let $\iota \in \mathrm{Gal}(\mathbb{C}/\mathbb{R})$ denote complex conjugation, for brevity.

Now, a representation $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$ must have $V_{\mathbb{C}}$ decompose into eigenspaces according to the characters $X^*(\mathbb{S})$, so one admits a decomposition

$$V_{\mathbb{C}} = \bigoplus_{\chi \in X^*(\mathbb{S})} V^{\chi}_{\mathbb{C}}.$$

However, one also needs $V^{\iota\chi}_{\mathbb{C}} = \overline{V^{\chi}_{\mathbb{C}}}$ because ι swaps $\{\chi,\iota\chi\}$. By Galois descent, this is enough data to (conversely) define a representation $h\colon \mathbb{S} \to \mathrm{Gal}(V)_{\mathbb{R}}$.

[1] Alternatively, note one has an isomorphism $(\mathbb{C}\otimes_{\mathbb{R}}\mathbb{C})^\times \cong \mathbb{C}^\times \times \mathbb{C}^\times$ by sending $(z,w) \mapsto z\otimes w$. Then these two characters are $(z,w)\mapsto z$ and $(z,w)\mapsto w$.

To relate the previous paragraph to Hodge structures, we recall that $X^*(\mathbb{S})$ is a rank-2 free $\mathbb{Z}$-module, so write $\chi_{p,q} \coloneqq z^{-p}\overline{z}^{-q}$ so that $\iota\chi_{p,q} = \chi_{q,p}$. Setting $V_{\mathbb{C}}^{p,q} \coloneqq V_{\mathbb{C}}^{\chi_{p,q}}$ now explains how to relate the previous paragraph to a Hodge structure, as desired. ■

Remark 1.8. The weight of a Hodge structure on some $V \in \mathrm{HS}_{\mathbb{Q}}$ can be read off of h as follows: note the weight cocharacter $h \circ w$ equals the $(-m)$th power map if and only if the weight is m.

Thus, we see immediately the category $\mathrm{HS}_{\mathbb{Q}}$ is abelian. Additionally, representation theory explains how to take tensor products and duals.

Example 1.9. We see that $V \in \mathrm{HS}_{\mathbb{Q}}$ has $V^\vee$ inherit a Hodge structure by setting $(V^\vee)^{p,q} \coloneqq (V^{-p,-q})^\vee$.

Example 1.10. We are now able to define the Tate twists $\mathbb{Q}(n) \coloneqq \mathbb{Q}(1)^{\otimes n}$, where negative powers indicates taking a dual. In particular, one can check that $\mathbb{Q}(n) \otimes \mathbb{Q}(m) = \mathbb{Q}(n+m)$ for any $n, m \in \mathbb{Z}$.

Notation 1.11. For any Hodge structure $V \in \mathrm{HS}_{\mathbb{Q}}$ and integer $m \in \mathbb{Z}$, we may write

$$V(m) \coloneqq V \otimes \mathbb{Q}(m).$$

We conclude this section by explaining one important application of Hodge structures.

Definition 1.12 (Hodge class)**.** Fix a $\mathbb{Q}$-Hodge structure V. A *Hodge class* of V is an element of $V \cap V^{0,0}$.

Remark 1.13. Looking at the construction in the proof of Lemma 1.7, we see that $v \in V$ is a Hodge class if and only if it is fixed by the corresponding representation $h \colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$.

Example 1.14. Fix a complex projective smooth variety X of dimension n and some even nonnegative integer $2p \geq 0$. Then one has Hodge classes given by elements of

$$\mathrm{H}_{\mathrm{B}}^{2p}(X, \mathbb{Q}) \cap \mathrm{H}^{p,p}(X)(p).$$

Now, any algebraic subvariety $Z \subseteq X$ of codimension k defines a linear functional on $\mathrm{H}_{\mathrm{dR}}^{2n-2k}(X, \mathbb{C})$ defined by

$$\omega \mapsto \int_Z \omega,$$

which one can check is supported on $\mathrm{H}^{k,k}$. Thus, by Poincaré duality, one finds that Z produces a Hodge cycle in $\mathrm{H}_{\mathrm{B}}^{2k}(X, \mathbb{Q})$.

In light of the above example, one has the following conjecture.

Conjecture 1.15 (Hodge)**.** Fix a complex projective smooth variety X. Then any Hodge class can be written as a linear combination of classes arising from algebraic subvarieties.

Remark 1.16. Here are some remarks on what is known about the Hodge conjecture, though it is admittedly little in this level of generality.

- The Hodge classes in $\mathrm{H}^2_{\mathrm{B}}(X)(1)$ come from algebraic subvarieties.
- The cup product of any two classes arising from algebraic subvarieties continues to be Hodge and arises from algebraic subvarieties.

For example, if one can show that all Hodge classes are cup products of Hodge classes of codimension 1 on a variety X, then one knows the Hodge conjecture for X.

We are not interested in proving (cases of) the Hodge conjecture in this thesis, so we will not say much more.

1.1.2 Polarizations

Here is an important example of a morphism of Hodge structures.

Definition 1.17 (polarization)**.** Fix a Hodge structure $V \in \mathrm{HS}_{\mathbb{Q}}$ pure of weight m given by the representation $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$. A *polarization* on V is a morphism $\varphi\colon V \otimes V \to \mathbb{Q}(-m)$ of Hodge structures such that the induced bilinear form on $V_{\mathbb{R}}$ given by

$$\langle v, w\rangle \coloneqq (2\pi i)^m \varphi(h(i)v \otimes w)$$

is symmetric and positive-definite. If V admits a polarization, we may say that V is *polarizable*, and we let $\mathrm{HS}^{\mathrm{pol}}_{\mathbb{Q}} \subseteq \mathrm{HS}_{\mathbb{Q}}$ be the full subcategory of polarizable $\mathbb{Q}$-Hodge structures.

Remark 1.18. The positive-definiteness condition on $\langle \cdot, \cdot \rangle$ implies that φ is non-degenerate. Indeed, one may check non-degeneracy upon base-changing to $\mathbb{R}$ (because this is equivalent to inducing an isomorphism of vector spaces $V \to V^\vee$, which can be checked by fixing some $\mathbb{Q}$-bases and computing a determinant). Then we see that $\langle \cdot, \cdot \rangle$ being non-degenerate implies that

$$\varphi(v \otimes w) = (2\pi i)^{-m} \langle h(-i)v, w\rangle$$

is non-degenerate because $h(-i)\colon V \to V$ is an isomorphism of vector spaces (because $h(-i)^4 = \mathrm{id}_V$).

Remark 1.19. The symmetry condition on $\langle \cdot, \cdot \rangle$ implies a symmetry or alternating condition on φ. Indeed, we compute

$$\begin{aligned}
\varphi(v \otimes w) &= (2\pi i)^{-m} \langle h(-i)v, w\rangle \\
&= (2\pi i)^{-m} \langle w, h(-i)v\rangle \\
&= \varphi(h(i)w \otimes h(-i)v) \\
&= h_{\mathbb{Q}(-m)}(i)\varphi(w \otimes h(-1)v) \\
&= 1\varphi\left(w \otimes (-1)^m w\right) \\
&= (-1)^m \varphi(w \otimes v).
\end{aligned}$$

Thus, φ is symmetric when m is even, and φ is alternating when m is odd.

Let's give some constructions of polarizable Hodge structures.

Example 1.20. It will turn out that $\mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$ of any abelian variety A (over $\mathbb{C}$) is polarizable, explaining the importance of this notion for our application. Because we are reviewing abelian varieties in chapter 2, we will not say more here.

Example 1.21. If V is polarizable and pure of weight m, then any Hodge substructure $W \subseteq V$ is still polarizable (and pure of weight m). Indeed, one can simply restrict the polarization to W, and all the checks go through. For example, positive-definiteness of $\langle \cdot, \cdot \rangle$ means $\langle v, v \rangle > 0$ for all nonzero $v \in V$, so the same will be true upon restricting to W.

Example 1.22. If V and W are polarizable and pure of weight m, then $V \oplus W$ is also polarizable. Indeed, letting φ and ψ be polarizations on V and W respectively, we see that $(\varphi \oplus \psi)$ defined by

$$(\varphi \oplus \psi)((v, w), (v', w')) \coloneqq \varphi(v, v') + \psi(w, w')$$

succeeds at being a polarization: certainly it is a morphism of Hodge structures to $\mathbb{Q}(-m-n)$, and one can check that the corresponding bilinear form on $V \oplus W$ simply splits into a sum of the forms on V and W and is therefore symmetric and positive-definite.

Example 1.23. If V and W are polarizable and pure of weights m and n respectively, then $V \otimes W$ is also polarizable. Indeed, as in Example 1.22, let φ and ψ be polarizationson V and W respectively, and then we find that $(\varphi \otimes \psi)$ can be defined on pure tensors by

$$(\varphi \otimes \psi)(v \otimes w, v' \otimes w') \coloneqq \varphi(v, v')\psi(w, w').$$

One checks as before that this gives a polarization on $V \otimes W$: we certainly have a morphism of Hodge structures, and the corresponding bilinear form is the product of the bilinear forms on V and W and is therefore symmetric and positive-definite.

Example 1.24. If V is polarizable and pure of weight m with polarization φ, and $W \subseteq V$ is a Hodge substructure (which is polarizable by Example 1.21), then we claim $W^\perp$ (taken with respect to $\langle \cdot, \cdot \rangle$) is also a Hodge substructure and hence polarizable by Example 1.21. Well, for any $w' \in W_{\mathbb{R}}^\perp$ and $z \in \mathbb{S}(\mathbb{R})$, we must check that $h(z)w' \in W_{\mathbb{R}}^\perp$. For this, we note that any $w \in W$ has

$$\begin{aligned} \langle w, h(z)w' \rangle &= (2\pi i)^{-m} \varphi(h(i)w \otimes h(z)w') \\ &= h_{\mathbb{Q}(-m)}(1/z)(2\pi i)^{-m} \varphi(h(i/z)w \otimes w') \\ &= h_{\mathbb{Q}(-m)}(1/z)\langle h(i/z)w, w' \rangle \\ &= 0, \end{aligned}$$

where the last equality holds because $W \subseteq V$ is a Hodge substructure.

Note that one does not expect any Hodge substructure to have a complement, so Example 1.24 is a very important property of polarizations.

1.1.3 The Albert Classification

The presence of a polarization places strong restrictions on the endomorphisms of a Hodge structure. To explain how this works, we begin by reducing to the irreducible case: given a polarizable Hodge structure $V \in \mathrm{HS}_{\mathbb{Q}}$, we begin by noting that V can be decomposed into irreducible Hodge substructures

$$V = \bigoplus_{i=1}^{N} V_i^{\oplus m_i},$$

where V_i is an irreducible Hodge structure (i.e., an irreducible representation of $\mathbb{S}$) and $m_i \geq 0$ is some nonnegative integer. Then standard results on endomorphisms of representations tell us that

$$\mathrm{End}_{\mathrm{HS}}(V) = \bigoplus_{i=1}^{N} M_{m_i}(\mathrm{End}_{\mathrm{HS}}(V_i)),$$

and Schur's lemma implies that $\mathrm{End}_{\mathrm{HS}}(V_i)$ is a division algebra. The point of the above discussion is that we may reduce our discussion of endomorphisms to irreducible Hodge structures. We remark that polarizability of V implies that irreducible Hodge substructures continue to be polarizable by Example 1.21.

We are thus interested in classifying what algebras may appear as $\mathrm{End}_{\mathrm{HS}}(V)$ for irreducible Hodge structures $V \in \mathrm{HS}_{\mathbb{Q}}$. To this end, we note that $\mathrm{End}_{\mathrm{HS}}(V)$ comes with some extra structure.

Definition 1.25 (Rosati involution)**.** Let φ be a polarization on a Hodge structure $V \in \mathrm{HS}_{\mathbb{Q}}$. The *Rosati involution* is the function $(\cdot)^\dagger \colon \mathrm{End}_{\mathbb{Q}}(V) \to \mathrm{End}_{\mathbb{Q}}(V)$ defined by

$$\varphi(dv \otimes w) = \varphi(v \otimes d^\dagger w)$$

for all $d \in \mathrm{End}_{\mathrm{HS}}(V)$ and $v, w \in V$.

Remark 1.26. In light of Remark 1.18, we see that $d^\dagger$ is simply the adjoint of $d \colon V \to V$ associated to φ viewed as a non-degenerate bilinear pairing. For example, we immediately see that $(\cdot)^\dagger$ induces a well-defined linear operator $\mathrm{End}_{\mathbb{Q}}(V) \to \mathrm{End}_{\mathbb{Q}}(V)$.

Here are the important properties of the Rosati involution.

Lemma 1.27. Fix a Hodge structure $V \in \mathrm{HS}_{\mathbb{Q}}$ pure of weight m with polarization φ and associated Rosati involution $(\cdot)^\dagger$.

(a) If $d \in \mathrm{End}_{\mathrm{HS}}(V)$, then $d^\dagger \in \mathrm{End}_{\mathrm{HS}}(V)$.

(b) Anti-involution: for any $d, e \in \mathrm{End}_{\mathbb{Q}}(V)$, we have $d^{\dagger\dagger} = d$ and $(de)^\dagger = e^\dagger d^\dagger$.

(c) Positive: for any nonzero $d \in \mathrm{End}_{\mathbb{Q}}(V)$, we have $\operatorname{tr} dd^\dagger > 0$.

Proof. We show the claims in sequence.

(a) This follows because φ is a morphism of Hodge structures. Formally, we would like to check that $d^\dagger$ commutes with the action of $\mathbb{S}$. Let $h \colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$ be the representation corresponding to the Hodge structure. Well, for any $g \in \mathbb{S}(\mathbb{C})$ and $v, w \in V$, we compute

$$\begin{aligned}
\varphi(v \otimes d^\dagger h(g)w) &= \varphi(dv \otimes h(g)w) \\
&= h_{\mathbb{Q}(-m)}(g)\varphi\left(h(g^{-1})dv \otimes w\right) \\
&\overset{*}{=} h_{\mathbb{Q}(-m)}(g)\varphi\left(dh(g^{-1})v \otimes w\right) \\
&= h_{\mathbb{Q}(-m)}(g)\varphi\left(h(g^{-1})v \otimes d^\dagger w\right) \\
&= \varphi(v \otimes h(g)d^\dagger w)
\end{aligned}$$

where $\overset{*}{=}$ holds because d is a morphism of Hodge structures. The non-degeneracy of φ given in Remark 1.18 now implies that $d^\dagger h(g) = h(g)d^\dagger$, so we are done.

(b) This is a purely formal property of adjoints.

(c) The point is to reduce this to the case where V is a matrix algebra over $\mathbb{R}$ and $(\cdot)^\dagger$ is the transpose. Indeed, this positivity can be checked after a base-change to $\mathbb{R}$. As such, we let $\langle\cdot,\cdot\rangle$ be the symmetric positive-definite bilinear form assocated to φ defined by

$$\langle v, w\rangle := (2\pi i)^{-m}\varphi(h(i)v \otimes w)$$

for any $v, w \in V_{\mathbb{R}}$. We thus see that $(\cdot)^\dagger$ is also the adjoint operator with respect to $\langle\cdot,\cdot\rangle$: we know

$$(2\pi i)^{-m}\langle h(i)dv, w\rangle = (2\pi i)^{-m}\langle h(i)v, d^\dagger w\rangle$$

for any $v, w \in V_{\mathbb{R}}$, which is equivalent to always having $\langle dv, w\rangle = \langle v, d^\dagger w\rangle$. Now, we may fix an orthonomal basis of $V_{\mathbb{R}}$ with respect to $\langle\cdot,\cdot\rangle$ so that $\mathrm{End}_{\mathbb{R}}(V_{\mathbb{R}})$ is identified with $M_n(\mathbb{R}^{\dim V})$ and $(\cdot)^\dagger$ is identified with the transpose. Then $\operatorname{tr} dd^{\intercal}$ is the sum of the squares of the matrix entries of d and is therefore positive when d is nonzero. ■

We are now ready to state the Albert classification, which classifies division algebras over $\mathbb{Q}$ equipped with a positive anti-involution.

Theorem 1.28 (Albert classification)**.** Let D be a division algebra over $\mathbb{Q}$ equipped with a Rosati involution $(\cdot)^\dagger\colon D \to D$. Further, let F be the center of D, and let $F^\dagger$ be the subfield fixed by $(\cdot)^\dagger$. Then D admits exactly one of the following types.

- Type I: D is a totally real number field so that $D = F = F^\dagger$, and $(\cdot)^\dagger$ is the identity.
- Type II: D is a totally indefinite quaternion division algebra over F where $F = F^\dagger$, and $(\cdot)^\dagger$ corresponds to the transpose on $D \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$.
- Type III: D is a totally definite quaternion division algebra over F where $F = F^\dagger$, and $(\cdot)^\dagger$ corresponds to the canonical involution on $D \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{H}$ (where $\mathbb{H}$ is the quaternions).
- Type IV: D is a division algebra over F, where F is a totally imaginary quadratic extension of $F^\dagger$, and $(\cdot)^\dagger$ is the complex conjugation automorphism of F. In other words, F is a CM field, and $F^\dagger$ is the maximal totally real subfield.

Proof. This is a rather lengthy computaion. We refer to [Mum74, Section 21, Application I]. ■

1.2 Monodromy Groups

In this section, we define the Mumford–Tate group and the Hodge group.

1.2.1 The Mumford–Tate Group

We are now ready to define the Mumford–Tate group. Intuitively, it is the monodromy group of the associated representation of a Hodge structure.

Definition 1.29 (Mumford–Tate group)**.** For some $V \in \mathrm{HS}_{\mathbb{Q}}$, the *Mumford–Tate group* $\mathrm{MT}(V)$ is the smallest algebraic $\mathbb{Q}$-group containing the image of the corresponding representation $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$.

Remark 1.30. Because $\mathbb{S}$ is connected, we see that h is also connected. Namely, $\mathrm{MT}(V)^\circ \subseteq \mathrm{MT}(V)$ will be an algebraic $\mathbb{Q}$-group containing the image of h if $\mathrm{MT}(V)$ does too, so equality is forced.

Example 1.31. Suppose that $V \in \mathrm{HS}_{\mathbb{Q}}$ is pure of weight m.

- If $m = 0$, then we claim that $\mathrm{MT}(V) \subseteq \mathrm{SL}(V)$. It is enough to check that h outputs into $\mathrm{SL}(V)$.
- If $m \neq 0$, then we claim that $\mathrm{MT}(V)$ contains $\mathbb{G}_{m,\mathbb{Q}}$. It is enough to check that $\mathrm{MT}(V)_{\mathbb{C}}$ contains $\mathbb{G}_{m,\mathbb{C}}$. Well, for any $z \in \mathbb{C}$ $h(z, \overline{z})$ acts on the component $V^{p,q} \subseteq V_{\mathbb{C}}$ by $z^{-p}z^{-q} = z^{-m}$, so $\mathrm{MT}(V)_{\mathbb{C}}$ must contain the scalar z^{-m} for all $z \in \mathbb{C}$. The conclusion follows.

Because Hodge structures are defined after passing to $\mathbb{C}$, it will be helpful to have a definition of $\mathrm{MT}(V)$ as a monodromy group corresponding to a morphism over $\mathbb{C}$.

Lemma 1.32. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$, and let $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$ be the corresponding representation. Then $\mathrm{MT}(V)$ is the smallest algebraic $\mathbb{Q}$-subgroup of $\mathrm{GL}(V)$ such that $\mathrm{MT}(V)_{\mathbb{C}}$ contains the image of $h_{\mathbb{C}} \circ \mu$.

Proof. Let M' be the smallest algebraic $\mathbb{Q}$-subgroup of $\mathrm{GL}(V)$ containing $h_{\mathbb{C}} \circ \mu$. We want to show that $M' = M$.

- To show $M' \subseteq \mathrm{MT}(V)$, we must show that $\mathrm{MT}(V)_{\mathbb{C}}$ contains the image of $h_{\mathbb{C}} \circ \mu$. Well, $\mathrm{MT}(V)_{\mathbb{R}}$ contains the image of h, so $\mathrm{MT}(V)_{\mathbb{C}}$ contains the image of $h_{\mathbb{C}}$, which contains the image of $h_{\mathbb{C}} \circ \mu$.
- Showing $\mathrm{MT}(V) \subseteq M'$ is a little harder. We must show that M' contains the image of $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$. It is enough to check that M' contains the image of $h_{\mathbb{C}}$ because then we can descend everything to $\mathbb{R}$, and because $\mathbb{C}$ is algebraically closed, we see that $\mathbb{C}$-points are certainly dense enough so that it is enough to chek that $M'(\mathbb{C})$ contains the image $h(\mathbb{S}(\mathbb{C}))$.

  The point is that M' is defined over $\mathbb{Q}$, so $M'_{\mathbb{C}}$ is stable under the action of complex conjugation, which we denote by ι. Similarly, h being defined over $\mathbb{R}$ guarantees that it commutes with complex conjugation. In particular, we already know that M' contains the points of the form $h(z, 1)$ for $(z, 1) \in \mathbb{S}(\mathbb{C})$. Thus, we see that M' also contains the points

  $$\iota(h(z, 1)) = h(\iota(z, 1)) = h(1, z)$$

  because everything is defined over $\mathbb{R}$. (This last equality follows by tracking through the action of ι on $\mathbb{S}(\mathbb{C})$.) We conclude that M' contains $h(z, w)$ for any $(z, w) \in \mathbb{S}(\mathbb{C})$, so we are done. ∎

Roughly speaking, the point of the group $\mathrm{MT}(V)$ is that $\mathrm{MT}(V)$ is an algebraic $\mathbb{Q}$-group remembering everything one wants to know about the Hodge structure. One way to rigorize this is as follows.

Proposition 1.33. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$. Suppose $T \in \mathrm{HS}_{\mathbb{Q}}$ can be written as

$$T = \bigoplus_{i=1}^{N} \left(V^{\otimes m_i} \otimes (V^{\vee})^{\otimes n_i}\right),$$

where $m_i, n_i \geq 0$ are nonnegative integers. Then $W \subseteq T$ is a Hodge substructure if and only if the action of $\mathrm{MT}(V)$ on T stabilizes W.

Proof. For each $W \in \mathrm{HS}_{\mathbb{Q}}$, we let h_W denote the corresponding representation. In the backwards direction, we note that $\mathrm{MT}(V)$ stabilizing W implies that $h(s)$ stabilizes $W_{\mathbb{R}}$ for any s. We can thus view $W_{\mathbb{R}} \subseteq T_{\mathbb{R}}$ as a subrepresentation of $\mathbb{S}$, so taking eigenspaces reveals that W can be given the structure of a Hodge substructure of T.

The converse will have to use the construction of T. Indeed, suppose that $W \subseteq T$ is a Hodge substructure, and let $M \subseteq \mathrm{GL}(V)$ be the smallest algebraic $\mathbb{Q}$-group stabilizing $W \subseteq T$. We would like to show that $\mathrm{MT}(V) \subseteq M$. By definition of $\mathrm{MT}(V)$, it is enough to show that h factors through $M_{\mathbb{R}}$, meaning we must show that $h(s)$ stabilizes W for each $s \in \mathbb{S}$. Well, $h(s)$ will act by characters on the eigenspaces $W^{p,q}_{\mathbb{C}} \subseteq W_{\mathbb{C}}$, so $h(s)$ does indeed stabilize W. ∎

Corollary 1.34. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$. Suppose $T \in \mathrm{HS}_{\mathbb{Q}}$ can be written as

$$T = \bigoplus_{i=1}^{N} \left(V^{\otimes m_i} \otimes (V^\vee)^{\otimes n_i} \right),$$

where $m_i, n_i \geq 0$ are nonnegative integers. Then $t \in T$ is a Hodge class if and only if it is fixed by $\mathrm{MT}(V)$.

Proof. We apply Proposition 1.33 to $\mathbb{Q}(0) \oplus T$. Then we note that $\mathrm{span}_{\mathbb{Q}}\{(1,t)\} \subseteq \mathbb{Q}(0) \oplus T$ is a Hodge substructure if and only if it is preserved by $\mathrm{MT}(V)$. We now tie each of these to the statement.

- On one hand, we see that being a one-dimensional Hodge substructure implies that $(1,t)$ must have bidegree (p,p) for some $p \in \mathbb{Z}$, but we have to live in $(0,0)$ because our 1 lives in $\mathbb{Q}(0)$. Thus, this is equivalent to being a Hodge class.
- On the other hand, being preserved by $\mathrm{MT}(V)$ implies that $\mathrm{MT}(V)$ acts by scalars on $(1,t)$, but $\mathrm{MT}(V)$ acts trivially on $\mathbb{Q}(0)$, so all the relevant scalars must be 1. Thus, this is equivalent to being fixed by $\mathrm{MT}(V)$. ∎

We thus see that understanding the Mumford–Tate group is important from the perspective of the Hodge conjecture (Conjecture 1.15). It will be helpful to note that this characterizes $\mathrm{MT}(V)$ in some cases.

Proposition 1.35. Fix a field K of characteristic 0. Let $H \subseteq \mathrm{GL}_{n,K}$ be a reductive subgroup. Suppose H' is the algebraic $\mathbb{Q}$-subgroup of $\mathrm{GL}_{n,K}$ defined by fixing all H-invariants occuring in any tensor representation

$$T = \bigoplus_{i=1}^{N} \left(V^{\otimes m_i} \otimes (V^\vee)^{\otimes n_i} \right),$$

where $m_i, n_i \geq 0$ are nonnegative integers. Then $H = H'$.

Proof. Note $H \subseteq H'$ is automatic, so the main content comes from proving the other inclusion. Proving this would step into the (rather deep) theory of algebraic groups, which we will avoid. Instead, we will mention that the key input is Chevalley's theorem, which asserts that any subgroup H of G is the stabilizer of some line in some representation of G. We refer to [Del18, Proposition 3.1]; see also [Mil17, Theorem 4.27]. ∎

Corollary 1.36. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ such that $\mathrm{MT}(V)$ is reductive. Then $\mathrm{MT}(V)$ is exactly the algebraic $\mathbb{Q}$-subgroup of $\mathrm{GL}(V)$ fixing all Hodge classes.

Proof. Corollary 1.34 explains that the Hodge classes are exactly the vectors fixed by $\mathrm{MT}(V)$, so this follows from Proposition 1.35. ∎

Remark 1.37. Corollary 1.36 is true without a reductivity assumption (see [Del18, Proposition 3.4]), but we will not need this in our applications. (On the other hand, one does not expect Proposition 1.35 to be true without any assumptions on H.) Namely, we will be interested in abelian varieties, whose Hodge structures are polarizable by Example 1.20, and we will shortly see that this implies that $\mathrm{MT}(V)$ is reductive in Lemma 1.44.

1.2.2 The Hodge Group

In computational applications, it will be frequently be easier to compute a smaller monodromy group related to $\mathrm{MT}(V)$.

Definition 1.38 (Hodge group)**.** Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight. Then the *Hodge group* $\mathrm{Hg}(V)$ is the smallest algebraic $\mathbb{Q}$-subgroup $\mathrm{GL}(V)$ containing the image of $h|_{\mathbb{U}}$, where $\mathbb{U} \subseteq \mathbb{S}$ is defined as the kernel of the norm character $z\overline{z}\colon \mathbb{S} \to \mathbb{G}_{m,\mathbb{R}}$.

Remark 1.39. Even though z and $\overline{z}$ are only defined as characters $\mathbb{S}_{\mathbb{C}} \to \mathbb{G}_{m,\mathbb{C}}$, the norm character $z\overline{z}$ is defined as a character $\mathbb{S} \to \mathbb{G}_{m,\mathbb{R}}$ because it is fixed by complex conjugation. For example, we see that

$$\mathbb{U}(\mathbb{R}) = \{z \in \mathbb{C} : |z| = 1\}.$$

Thus, we see that $\mathbb{U}$ stands for "unit circle." While we're here, we remark that $\mathbb{U}(\mathbb{C}) \subseteq \mathbb{S}(\mathbb{C})$ is identified with the subset $\{(z, 1/z) : z \in \mathbb{C}^\times\}$.

Remark 1.40. The same argument as in Remark 1.30 shows that the connectivity of $\mathbb{U}$ implies the connectivity of $\mathrm{Hg}(V)$.

Intuitively, $\mathrm{Hg}(V)$ removes the scalars that might live in $\mathrm{MT}(V)$ by Example 1.31. These scalars are an obstruction to $\mathrm{MT}(V)$ being a semisimple group, and we will see in Proposition 2.67 that $\mathrm{Hg}(V)$ will thus frequently succeed at being semisimple. Let's rigorize this discusison.

Lemma 1.41. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ pure of weight m, and let $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$ be the corresponding representation.

(a) We have $\mathrm{Hg}(V) \subseteq \mathrm{SL}(V)$.

(b) Thus,

$$\mathrm{MT}(V) = \begin{cases} \mathrm{Hg}(V) & \text{if } m = 0, \\ \mathbb{G}_{m,\mathbb{Q}} \mathrm{Hg}(V) & \text{if } m \neq 0, \end{cases}$$

where the almost direct product in the second case is given by embedding $\mathbb{G}_{m,\mathbb{Q}} \to \mathrm{GL}(V)$ via scalars.

Proof. We show the claims in sequence.

(a) It is enough to check that $\mathrm{SL}(V)$ contains the image of $h|_{\mathbb{U}}$. In other words, we want to check that $\det h(z) = 1$ for all $z \in \mathbb{U}(\mathbb{R})$. By extending scalars, it is enough to compute the determinant as an operator on $V_{\mathbb{C}}$. For this, we note that $h(z)$ acts on the component $V^{p,q} \subseteq V_{\mathbb{C}}$ by the scalar $z^{-p}\overline{z}^{-q}$, so the determinant of $h(z)$ acting on $V^{p,q} \oplus V^{q,p}$ is

$$\left(z^{-p}\overline{z}^{-q}\right)^{\dim V^{p,q}} \cdot \left(z^{-q}\overline{z}^{-p}\right)^{\dim V^{q,p}} = (z\overline{z})^{-(p+q)\dim V^{p,q}}$$

because $\dim V^{p,q} = \dim V^{q,p}$. This simplifies to $(z\overline{z})^{-\frac{1}{2}m \dim(V^{p,q} \oplus V^{q,p})}$ because V is pure of weight m, so the result follows by summing over all pairs (p, q).[2]

(b) Before doing anything serious, we remark that $\mathbb{G}_{m,\mathbb{Q}} \mathrm{Hg}(V)$ is in fact an almost direct product. Namely, we should check that the intersection $\mathbb{G}_{m,\mathbb{Q}} \cap \mathrm{Hg}(V)$ is finite (even over $\mathbb{C}$). Well, by (a), $\mathrm{Hg}(V) \subseteq \mathrm{SL}(V)$. Thus, it is enough to notice that $\mathbb{G}_{m,\mathbb{Q}} \cap \mathrm{SL}(V)$ is finite because V is finite-dimensional over $\mathbb{C}$: over $\mathbb{C}$,

[2] If m is even, this argument does not work verbatim for the component $(m/2, m/2)$. Instead, one can compute the determinant of $h(z)$ acting on $V^{m/2,m/2}$ directly as $(z\overline{z})^{-\frac{1}{2}m \dim V^{m/2,m/2}}$.

the intersection consisits of scalar matrices $\lambda \operatorname{id}_V$ such that $\lambda^{\dim V} = 1$, so the intersefction is the finite algebraic group $\mu_{\dim V}$.

We now proceed with the argument. Because $\mathbb{U} \subseteq \mathbb{S}$, we of course have $\operatorname{Hg}(V) \subseteq \operatorname{MT}(V)$, and if $m \neq 0$, then Example 1.31 implies that $\mathbb{G}_{m,\mathbb{Q}} \subseteq \operatorname{MT}(V)$ so that $\mathbb{G}_{m,\mathbb{Q}} \operatorname{Hg}(V) \subseteq \operatorname{MT}(V)$. It is therefore enough to check the given equalities after base-changing to $\mathbb{R}$. Namely, using Lemma 1.32, we should check that $\operatorname{Hg}(V)(\mathbb{C})$ contains the image of $h_{\mathbb{C}} \circ \mu$ when $m = 0$, and $\mathbb{C}^\times \operatorname{Hg}(V)(\mathbb{C})$ contains the image of $h_{\mathbb{C}} \circ \mu$ when $m \neq 0$. Well, for any $z \in \mathbb{C}^\times$, we may write $z = re^{i\theta}$ where $r \in \mathbb{R}^+$ and $\theta \in \mathbb{R}$. Then we compute

$$\begin{aligned} h(\mu(z)) &= h(z, 1) \\ &= h\left(re^{i\theta}, 1\right) \\ &= h\left(\sqrt{r}e^{i\theta/2}, \sqrt{r}e^{-i\theta/2}\right) h\left(\sqrt{r}e^{i\theta/2}, \frac{1}{\sqrt{r}e^{i\theta/2}}\right). \end{aligned}$$

Now, $h\left(\sqrt{r}e^{i\theta/2}, \sqrt{r}e^{-i\theta/2}\right)$ is a scalar as computed in Example 1.31, and $\left(\sqrt{r}e^{i\theta/2}, \frac{1}{\sqrt{r}e^{i\theta/2}}\right)$ lives in $\mathbb{U}(\mathbb{C}) = \{(z, w) : zw = 1\}$. Thus, we see that $h(\mu(z))$ is certainly contained in $\mathbb{C}^\times \operatorname{Hg}(V)(\mathbb{C})$, completing the proof in the case $m \neq 0$. In the case where $m = 0$, the scalar $h\left(\sqrt{r}e^{i\theta/2}, \sqrt{r}e^{-i\theta/2}\right)$ is actually the identity, so we see that $h(\mu(z)) \in \operatorname{Hg}(V)(\mathbb{C})$. ■

It is worthwhile to note that there is also a tensor characterization of $\operatorname{Hg}(V)$.

Proposition 1.42. Fix $V \in \operatorname{HS}_{\mathbb{Q}}$ of pure weight. Suppose $T \in \operatorname{HS}_{\mathbb{Q}}$ is of pure weight n and can be written as

$$T = \bigoplus_{i=1}^{N} \left(V^{\otimes m_i} \otimes (V^\vee)^{\otimes n_i}\right),$$

where $m_i, n_i \geq 0$ are nonnegative integers. Then $W \subseteq T$ is a Hodge substructure if and only if the action of $\operatorname{Hg}(V)$ on T stabilizes W.

Proof. Of course, if $W \subseteq T$ is a Hodge substructure, then W is preserved by the action of $\operatorname{MT}(V)$, so W will be preserved by the action of $\operatorname{Hg}(V) \subseteq \operatorname{MT}(V)$.

Conversely, if $\operatorname{Hg}(V)$ stabilizes W, then we would like to show that $W \subseteq T$ is a Hodge substructure, which by Proposition 1.33 is the same as showing that $\operatorname{MT}(V)$ stabilizes W. For this, we use Lemma 1.41, which tells us that $\operatorname{MT}(V) \subseteq \mathbb{G}_{m,\mathbb{Q}} \operatorname{Hg}(V)$. Namely, because $\operatorname{Hg}(V)$ already stabilizes W, it is enough to note that of course the scalars $\mathbb{G}_{m,\mathbb{Q}}$ stabilize the subspace $W \subseteq T$. ■

Corollary 1.43. Fix an irreducible Hodge structure $V \in \operatorname{HS}_{\mathbb{Q}}$ of pure weight. Observe that the inclusion $\operatorname{Hg}(V) \subseteq \operatorname{GL}(V)$ makes V into a representation of $\operatorname{Hg}(V)$. Then V is irreducible as a representation of $\operatorname{Hg}(V)$.

Proof. By Proposition 1.42, a $\operatorname{Hg}(V)$-submodule is a Hodge substructure, but there are no nonzero proper Hodge substructures because V is an irreducible Hodge structure. ■

1.2.3 Bounding with Known Classes

Here, we use endomorphisms and the polarization to bound the size of $\operatorname{MT}(V)$ and $\operatorname{Hg}(V)$.

Lemma 1.44. Fix a polarizable Hodge structure $V \in \operatorname{HS}_{\mathbb{Q}}$ of pure weight. Then $\operatorname{MT}(V)$ and $\operatorname{Hg}(V)$ are reductive.

Proof. By [Mil17, Corollary 19.18], it is enough to find faithful semisimple representations of $\mathrm{MT}(V)$ and $\mathrm{Hg}(V)$. We claim that the inclusions $\mathrm{MT}(V) \subseteq \mathrm{GL}(V)$ and $\mathrm{Hg}(V) \subseteq \mathrm{GL}(V)$ provide this representation: certainly this representation is faithful, and it is faithful because any subrepresentation is a Hodge substructure by Propositions 1.33 and 1.42. ■

Lemma 1.45. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$. Let $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ be the endomorphism algebra of V. Then $\mathrm{MT}(V)$ is an algebraic $\mathbb{Q}$-subgroup of

$$\mathrm{GL}_D(V) \coloneqq \{g \in \mathrm{GL}(V) : g \circ d = d \circ g \text{ for all } d \in D\}.$$

Proof 1. Noting that $\mathrm{GL}_D(V)$ is an algebraic $\mathbb{Q}$-group (it is a subgroup of $\mathrm{GL}(V)$ cut out by the equations given by commuting with a basis of D), it is enough to show that $\mathrm{GL}_D(V)$ contains the image of the representation $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$. Well, by definition D consists of morphisms commuting with the action of $\mathbb{S}$, so the image of h must commute with D. ■

Proof 2. Motivated by Corollary 1.36, one expects to find Hodge classes corresponding to the condition of commuting with D. Well, there is a canonical isomorphism $V \otimes V^\vee \to \mathrm{End}_{\mathbb{Q}}(V)$ of $\mathbb{S}$-representations, so by tracking through how representations of $\mathbb{S}$ correspond to Hodge structures, we see that $f\colon V \to V$ preserves the Hodge structure if and only if it is fixed by $\mathbb{S}$, which is equivalent to the corresponding element $f \in V \otimes V^\vee$ being fixed by $\mathbb{S}$, which is equivalent to f being a Hodge class by Remark 1.13. This completes the proof of the lemma upon comparing with Corollary 1.34. ■

Remark 1.46. Of course, we also have $\mathrm{Hg}(V) \subseteq \mathrm{GL}_D(V)$ because $\mathrm{Hg}(V) \subseteq \mathrm{MT}(V)$.

Lemma 1.47. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ pure of weight m with polarization φ. Then $\mathrm{MT}(V)$ is an algebraic $\mathbb{Q}$-subgroup of

$$\mathrm{GSp}(\varphi) \coloneqq \{g \in \mathrm{GL}(V) : \varphi(gv \otimes gw) = \lambda(g)\varphi(v \otimes w) \text{ for fixed } \lambda(g) \in \mathbb{Q}\}.$$

Proof 1. Once again, we note that $\mathrm{GSp}(\varphi)$ is an algebraic $\mathbb{Q}$-group cut out by equations of the form

$$\varphi(gv \otimes gw)\varphi(v' \otimes w') = \varphi(v \otimes w)\varphi(gv' \otimes gw')$$

as $v, w, v', w' \in V$ varies over a basis. Thus, it is enough to check that $\mathrm{GSp}(\varphi)$ contains the image of $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$. Well, for any $z \in \mathbb{S}(\mathbb{R})$, we note that

$$\varphi(h(z) \otimes h(z)w) = h_{\mathbb{Q}(-m)}(z)\varphi(v \otimes w)$$

for any $v, w \in V_{\mathbb{R}}$ because φ is a morphism of Hodge structures. ■

Proof 2. Once again, Corollary 1.36 tells us to expect the polarization to produce a Hodge class corresponding to the above equations cutting out $\mathrm{MT}(V)$.

This construction is slightly more involved. We begin by constructing two Hodge classes.

- Note $\varphi\colon V \otimes V \to \mathbb{Q}(-m)$ is a morphism of Hodge structrures, so it is an $\mathbb{S}$-invariant map and hence given by an $\mathbb{S}$-invariant element of $V^\vee \otimes V^\vee(-m)$. Thus, $\varphi \in V^\vee \otimes V^\vee(-m)$ is a Hodge class by Remark 1.13.
- Because φ is non-degenerate, it induces an isomorphism $V(m) \to V^\vee$. Now, $\mathrm{End}_{\mathbb{Q}}(V)$ is canonically isomorphic to $V \otimes V^\vee$, which we now see is isomorphic (via φ) to $V \otimes V(m)$. We let $C \in V \otimes V(m)$ be the image of $\mathrm{id}_V \in \mathrm{End}_{\mathbb{Q}}(V)^{\mathbb{S}}$ in $V \otimes V(m)$, which we note is a Hodge class again by Remark 1.13. (Here, C stands for "Casimir.")

In total, we see that we have produced a Hodge class $C \otimes \varphi$. It remains to show that $g \in \mathrm{GL}(V)$ fixing $C \otimes \varphi$ implies that $g \in \mathrm{GSp}(\varphi)$, which will complete the proof by Corollary 1.34.

Well, suppose $g(C \otimes \varphi) = C \otimes \varphi$. Note $g(C \otimes \varphi) = gC \otimes g\varphi$, which can only equal $C \otimes \varphi \in (V \otimes V) \otimes_{\mathbb{Q}} (V^\vee \otimes V^\vee)$ if there is a scalar $\lambda \in \mathbb{Q}^\times$ such that $gC = \lambda C$ and $g\varphi = \lambda^{-1}\varphi$. This second condition amounts to requiring

$$\varphi\left(g^{-1}v \otimes g^{-1}w\right) = \lambda^{-1}\varphi(v \otimes w)$$

for any $v, w \in V$, which rearranges into $g \in \mathrm{GSp}(\varphi)$. ■

Remark 1.48. The construction given in the above proof is described in [GGL24, Remark 8.3.4]. They also show the converse claim that any $g \in \mathrm{GSp}(\varphi)$ fixes $C \otimes \varphi$.

To see this, one has to do an explicit computation with C. For this, let $\{v_1, \dots, v_n\}$ be a basis of V, and $\{v_1^*, \dots, v_n^*\}$ be the dual basis of $V(m)$ taken with respect to φ. Then $C = \sum_{i=1}^n v_i \otimes v_i^*$. Similarly, we see that $\{gv_1, \dots, gv_n\}$ is a basis of V with a dual basis $\{(gv_1)^*, \dots, (gv_n)^*\}$ so that $C = \sum_{i=1}^n (gv_i) \otimes (gv_i)^*$. Now, on one hand, if g has multiplier λ, then $g\varphi = \lambda^{-1}\varphi$. On the other hand, $\varphi(gv_i, gv_j^*) = \lambda 1_{i=j}$, so $(gv_i)^* = \lambda^{-1} gv_i^*$, which allows us to compute $gC = \lambda C$. In total, $g(C \otimes \varphi) = C \otimes \varphi$.

Remark 1.49. One can check that $\mathrm{GSp}(\varphi)$ does not depend on the choice of polarization. Roughly speaking, the point is that the choice of a different polarization amounts to some choice of an element in $D^\times$ which we can track through.

In light of the above two lemmas, we pick up the following notation.

Notation 1.50. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ pure of weight m with $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ and polarization φ. Then we define

$$\mathrm{GSp}_D(\varphi) \coloneqq \mathrm{GL}_D(V) \cap \mathrm{GSp}(\varphi).$$

By Lemmas 1.45 and 1.47, we see that $\mathrm{MT}(V) \subseteq \mathrm{GSp}_D(\varphi)$.

Remark 1.51. In "most cases," we expect that generic Hodge structures V should have the equality $\mathrm{MT}(V) = \mathrm{GL}_D(V)$, and if V admits a polarization φ, then we expect the equality $\mathrm{MT}(V) = \mathrm{GSp}_D(\varphi)$. To rigorize this intuition, one must discuss Shimura varieties, which we will avoid doing for now.

We can also apply Lemmas 1.45 and 1.47 to bound $\mathrm{Hg}(V)$.

Notation 1.52. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ pure of weight m with $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ and polarization φ. Then we define

$$\mathrm{Sp}(\varphi) \coloneqq \{g \in \mathrm{GL}(V) : \varphi(gv \otimes gw) = \varphi(v \otimes w)\},$$

and

$$\mathrm{Sp}_D(\varphi) \coloneqq \mathrm{GL}_D(V) \cap \mathrm{Sp}(\varphi).$$

Remark 1.53. Let's explain why $\mathrm{Hg}(V) \subseteq \mathrm{Sp}_D(\varphi)$. By Lemma 1.45, we see that $\mathrm{Hg}(V) \subseteq \mathrm{MT}(V) \subseteq \mathrm{GL}_D(V)$, so it remains to check that $\mathrm{Hg}(V) \subseteq \mathrm{Sp}(\varphi)$. Proceeding as in Lemma 1.47, it is enough to check that the image of $h|_{\mathbb{U}}$ lives in $\mathrm{Sp}(\varphi)_{\mathbb{R}}$, for which we note that any $z \in \mathbb{U}(\mathbb{R})$ has

$$\varphi(h(z)v \otimes h(z)w) = h_{\mathbb{Q}(-m)}(z)\varphi(v \otimes w),$$

but $h_{\mathbb{Q}(-m)}(z) = |z|^{-2m} \mathrm{id}_{\mathbb{Q}(-m)}$ is the identity because $z \in \mathbb{U}(\mathbb{R})$.

Thus far, our tools have been upper-bounding $\mathrm{MT}(V)$ and $\mathrm{Hg}(V)$. Here is a tool which sometimes provides a lower bound.

Lemma 1.54. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight, and let $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ be the endomorphism algebra of V. Then

$$D = \mathrm{End}_{\mathbb{Q}}(V)^{\mathrm{MT}(V)} = \mathrm{End}_{\mathbb{Q}}(V)^{\mathrm{Hg}(V)}.$$

Proof. As discussed in the second proof of Lemma 1.45, the Hodge calsses of $\mathrm{End}_{\mathbb{Q}}(V) \cong V \otimes V^{\vee}$ are exactly the endomorphisms of the Hodge structure, so the first equality follows from Corollary 1.34.

The second equality is purely formal: note that the scalar subgroup $\mathbb{G}_{m,\mathbb{Q}} \subseteq \mathrm{GL}(V)$ acts trivially on $V \otimes V^{\vee} \cong \mathrm{End}_{\mathbb{Q}}(V)$. Thus, we use Lemma 1.41 to compute

$$\begin{aligned}\mathrm{End}_{\mathbb{Q}}(V)^{\mathrm{Hg}(V)} &= \mathrm{End}_{\mathbb{Q}}(V)^{\mathbb{G}_{m,\mathbb{Q}}\,\mathrm{Hg}(V)} \\ &= \mathrm{End}_{\mathbb{Q}}(V)^{\mathbb{G}_{m,\mathbb{Q}}\,\mathrm{MT}(V)} \\ &= \mathrm{End}_{\mathbb{Q}}(V)^{\mathrm{MT}(V)},\end{aligned}$$

as required. ∎

Remark 1.55. To understand Lemma 1.54 as providing a lower bound, note that if $\mathrm{MT}(V)$ is "too small," then there will be many invariant elements in $\mathrm{End}_{\mathbb{Q}}(V)^{\mathrm{MT}(V)}$, perhaps exceeding D. On the other hand, the upper bound $\mathrm{MT}(V) \subseteq \mathrm{GL}_D(V)$ corresponds to the inequality $D \subseteq \mathrm{End}_{\mathbb{Q}}(V)^{\mathrm{MT}(V)}$.

1.2.4 Sums

For later use in computations, it will be helpful to have a few remarks on computing the Mumford–Tate and Hodge groups of a sum. Here the Hodge group really shines: given two Hodge structures $V_1, V_2 \in \mathrm{MT}(V)$ pure of nonzero weight, Lemma 1.41 tells us that $\mathrm{MT}(V_1)$ and $\mathrm{MT}(V_2)$ and $\mathrm{MT}(V_1 \oplus V_2)$ are all equal to some smaller group times scalars. It will turn out to be reasonable to hope that

$$\mathrm{Hg}(V_1 \oplus V_2) \stackrel{?}{=} \mathrm{Hg}(V_1) \times \mathrm{Hg}(V_2),$$

but then the introduction of scalars makes the hope $\mathrm{MT}(V_1 \oplus V_2) \stackrel{?}{=} \mathrm{MT}(V_1) \times \mathrm{MT}(V_2)$ unreasonable!

With this in mind, let's begin to study Hodge groups of sums of Hodge structures.

Lemma 1.56. Fix Hodge structures $V_1, \dots, V_k \in \mathrm{Hg}_{\mathbb{Q}}$ pure of the same weight.

(a) The subgroup $\mathrm{Hg}(V_1 \oplus \cdots \oplus V_k) \subseteq \mathrm{GL}(V_1 \oplus \cdots \oplus V_k)$ is contained in $\mathrm{Hg}(V_1) \times \cdots \times \mathrm{Hg}(V_k) \subseteq \mathrm{GL}(V_1 \oplus \cdots \oplus V_k)$.

(b) For each i, the projection map $\mathrm{pr}_i \colon \mathrm{Hg}(V_1 \oplus \cdots \oplus V_k) \to \mathrm{Hg}(V_i)$ is surjective.

Proof. For each i, let h_i denote the representations of $\mathbb{S}$ corresponding to the Hodge structures V_i, and let $h \coloneqq (h_1, \dots, h_k)$ be the representation $\mathbb{S} \to \mathrm{GL}(V)$ where $V \coloneqq V_1 \oplus \cdots \oplus V_k$. We show the claims in sequence.

(a) We must show that $\mathrm{Hg}(V_1) \times \cdots \times \mathrm{Hg}(V_k)$ contains the image of $h|_{\mathbb{U}}$. Well, for any $z \in \mathbb{U}(\mathbb{R})$ and index i, we see that $h_i(z) \in \mathrm{Hg}(V_i)$, so

$$h(z) = \mathrm{diag}(h_1(z), \dots, h_k(z))$$

lives in $\mathrm{Hg}(V_1) \times \cdots \times \mathrm{Hg}(V_k)$, as required.

(b) Fix an index i. It is enough to show that smallest algebraic $\mathbb{Q}$-group containing the image of pr_i also contains the image of $h_i|_{\mathbb{U}}$. Well, by definition of h, we see that h_i is equal to the composite

$$\mathbb{S} \stackrel{h}{\to} \mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_k) \stackrel{\mathrm{pr}_i}{\to} \mathrm{GL}(V_i),$$

from which the claim follows. ∎

Remark 1.57. All the claims in Lemma 1.56 are true if Hg is replaced by MT everywhere. One simply has to replace $\mathbb{U}$ with $\mathbb{S}$ in the proof.

Lemma 1.56 makes $\mathrm{Hg}(V_1 \oplus V_2) \stackrel{?}{=} \mathrm{Hg}(V_1) \times \mathrm{Hg}(V_2)$ appear to be a reasonable expectation. However, we note that we cannot in general expect this to be true: roughly speaking, there may be Hodge cycles on $V_1 \oplus V_2$ which are not seen on just V_1 or V_2. Here is a degenerate example.

Example 1.58. Fix a Hodge structure $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight, and let $n \geq 1$ be a positive integer. Letting $h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}$ be the corresponding representation, we get another Hodge structure $h^n\colon \mathbb{S} \to \mathrm{GL}\left(V^{\oplus n}\right)$. We claim that the diagonal embedding of $\mathrm{Hg}(V)$ into $\mathrm{GL}(V)^n \subseteq \mathrm{GL}\left(V^{\oplus n}\right)$ induces an isomorphism

$$\mathrm{Hg}(V) \to \mathrm{Hg}\left(V^{\oplus n}\right).$$

On one hand, we note that $\mathrm{Hg}\left(V^{\oplus n}\right)$ lives inside the diagonal embedding of $\mathrm{Hg}(V)$: note $\mathrm{Hg}\left(V^{\oplus n}\right) \subseteq \mathrm{Hg}(V)^n$ by Lemma 1.56, and $\mathrm{Hg}\left(V^{\oplus n}\right)$ must live inside the diagonal embedding of $\mathrm{GL}(V) \subseteq \mathrm{GL}\left(V^{\oplus n}\right)$ becuase all components of $h^n\colon \mathbb{S} \to \mathrm{GL}\left(V^{\oplus n}\right)_{\mathbb{R}}$ are equal. On the other hand, the surjectivity of the projections $\mathrm{Hg}\left(V^{\oplus n}\right) \to \mathrm{Hg}(V)$ from Lemma 1.56 implies that $\mathrm{Hg}\left(V^{\oplus n}\right)$ must equal the diagonal embedding of $\mathrm{Hg}(V)$ (instead of merely being contained in it).

One can upgrade this example as follows.

Lemma 1.59. Fix Hodge structures $V_1, \ldots, V_k \in \mathrm{Hg}_{\mathbb{Q}}$ pure of the same weight, and let $m_1, \ldots, m_k \geq 1$ be positive integers. Then the diagonal embeddings $\Delta_i\colon \mathrm{GL}(V_i) \to \mathrm{GL}\left(V_i^{\oplus m_i}\right)$ induce an isomorphism

$$\mathrm{Hg}(V_1 \oplus \cdots \oplus V_k) \to \mathrm{Hg}\left(V_1^{\oplus m_1} \oplus \cdots \oplus V_k^{\oplus m_k}\right).$$

Proof. We proceed in steps. The proof is a direct generalization of the one given in Example 1.58. For each i, let $h_i\colon \mathbb{S} \to \mathrm{GL}(V_i)_{\mathbb{R}}$ be the representation corresponding to the Hodge structure, and set $h := (h_1^{m_1}, \ldots, h_k^{m_k})$.

1. We claim that $\mathrm{Hg}\left(V_1^{\oplus m_1} \oplus \cdots \oplus V_k^{\oplus m_k}\right)$ lives in the image of $(\Delta_1, \ldots, \Delta_k)$. Indeed, the image is some algebraic $\mathbb{Q}$-subgroup of $\mathrm{GL}\left(V_1^{\oplus m_1} \oplus \cdots \oplus V_k^{\oplus m_k}\right)$, so we would like to check that this algebraic $\mathbb{Q}$-subgroup contains the image of $h|_{\mathbb{U}}$. Well, for any $z \in \mathbb{U}(\mathbb{R})$, we see that

   $$h(z) = (\Delta_1(h_1(z)), \ldots, \Delta_k(h_k(z)))$$

   lives in the image of $(\Delta_1, \ldots, \Delta_k)$.

2. For each i, let H_i be the projection of $\mathrm{Hg}\left(V_1^{\oplus m_1} \oplus \cdots \oplus V_k^{\oplus m_k}\right)$ onto one of the V_i components as in Lemma 1.56; the choice of V_i component does not matter by the previous step. By Lemma 1.56, we see that $H_i = \mathrm{Hg}(V_i)$. However, the previous step now requires

   $$\mathrm{Hg}\left(V_1^{\oplus m_1} \oplus \cdots \oplus V_k^{\oplus m_k}\right) = \Delta_1(H_1) \times \cdots \times \Delta_k(H_k),$$

   so we are done. ∎

Remark 1.60. As usual, this statement continues to be true for MT replacing Hg. One can either see this by applying Lemma 1.41 or by redoing the proof with $\mathbb{S}$ replacing $\mathbb{U}$.

The point of the lemma is that we can reduce our computation of Hodge groups to Hodge structures which are the sum of pairwise non-isomoprhic irreducible Hodge strucutures. Let's make a few remarks about this situation for completeness. Let $V_1, \ldots, V_k$ be pairwise non-isomorphic irreducuble Hodge structures which are pure of the same weight, and set $V := V_1 \oplus \cdots \oplus V_k$. Here are some remarks on $\mathrm{Hg}(V_1 \times \cdots \times V_k)$, summarizing everything we have done so far.

- We know that $\mathrm{Hg}(V) \subseteq \mathrm{Hg}(V_1) \times \cdots \times \mathrm{Hg}(V_k)$.
- We know that the projections of $\mathrm{Hg}(V)$ onto each factor $\mathrm{Hg}(V_i)$ are surjective.
- For each i, we may view V_i as a representation of $\mathrm{Hg}(V_i)$ via the inclusion $\mathrm{Hg}(V_i) \subseteq \mathrm{GL}(V_i)$. Then Corollary 1.43 tells us that V_i is an irreducible representation of $\mathrm{Hg}(V_i)$.
- One can also apply Lemma 1.54 to the full space V to see that

$$\begin{aligned}\mathrm{End}_{\mathrm{Hg}(V)}(V) &= \mathrm{End}_{\mathrm{HS}}(V) \\ &= \prod_{i=1}^{k} \mathrm{End}_{\mathrm{HS}}(V_i) \\ &= \prod_{i=1}^{k} \mathrm{End}_{\mathrm{Hg}(V_i)}(V_i).\end{aligned}$$

The following results take the above situation and provides some criteria to have

$$\mathrm{Hg}(V) \stackrel{?}{=} \mathrm{Hg}(V_1) \times \cdots \times \mathrm{Hg}(V_k).$$

Before stating the lemma, we remark that all groups in sight are connected by Remark 1.40, and we already have one inclusion by Lemma 1.56, so it suffices to pass to an algebraic closure and work with Lie algebras instead of the Lie groups. The following lemma is essentially due to Ribet [Rib76, pp. 790–791].

Lemma 1.61 (Ribet)**.** Work over an algebraically closed field of characteristic 0. Let $V_1, \ldots, V_k$ be finite-dimensional vector spaces, and let $\mathfrak{g}$ be a Lie subalgebra of $\mathfrak{gl}(V_1) \times \cdots \times \mathfrak{gl}(V_k)$. For each index i, let $\mathrm{pr}_i \colon (\mathfrak{gl}(V_1) \times \cdots \times \mathfrak{gl}(V_k)) \to \mathfrak{gl}(V_i)$ be the ith projection, and set $\mathfrak{g}_i \coloneqq \mathrm{pr}_i(\mathfrak{g})$. Suppose the following.

(i) Each $\mathfrak{g}_i$ is nonzero and simple.

(ii) For each pair (i, j) of distinct indices, the projection map $(\mathrm{pr}_i, \mathrm{pr}_j) \colon \mathfrak{g} \to \mathfrak{g}_i \times \mathfrak{g}_j$ is surjective.

Then $\mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_k$.

Proof. We proceed by induction on k. If $k \in \{0, 1\}$, then there is nothing to say. For the induction, we now assume that $k \geq 2$ and proceed in steps.

1. For our set-up, we let J be the kernel of $\mathrm{pr}_k \colon \mathfrak{g} \to \mathfrak{g}_n$. By definition, $J \subseteq \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_k$ takes the form $I \oplus 0$ for some subspace $I \subseteq \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$. Formally, one may let I be the set of vectors v such that $(v, 0) \in J$ and argue for the equality $J = I \oplus 0$ because all vectors in J take the form $(v, 0)$.

   The main content of the proof goes into showing that I is actually an ideal. To set ourselves up to prove this claim, let $\mathfrak{n} \subseteq \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$ denote its normalizer. We would like to show that $\mathfrak{n} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$, for which we use the inductive hypothesis.

2. For each pair of distinct indices $i, j < k$, we claim that the projection $(\mathrm{pr}_i, \mathrm{pr}_j) \colon \mathfrak{n} \to \mathfrak{g}_i \times \mathfrak{g}_j$ is surjective. Well, choose $X_i \in \mathfrak{g}_i$ and $X_j \in \mathfrak{g}_j$, and we need to find an element in $\mathfrak{n}$ with X_i and X_j at the correct coordinates.

   To begin, we note that (ii) yields some $(X_1, \ldots, X_k) \in \mathfrak{g}$ such that with the correct $X_i \in \mathfrak{g}_i$ and $X_j \in \mathfrak{g}_j$ coordinates. We would like to show that $X \coloneqq (X_1, \ldots, X_{k-1})$ lives in $\mathfrak{n}$, which will complete this step. Well, select any $Y \coloneqq (Y_1, \ldots, Y_{k-1})$ in I, and we see $(Y, 0) \in J$, so

$$[(X, X_k), (Y, 0)] = ([X, Y], 0)$$

   lives in J too (recall J is an ideal), so we conclude $[X, Y] \in I$. We conclude that X normalizes I, so $X \in \mathfrak{n}$.

3. We take a moment to complete the proof that $I \subseteq \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$ is an ideal. It is enough to check that the normalizer $\mathfrak{n}$ of I in $\mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$ equals all of $\mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$. For this, we use the inductive hypothesis. The previous step shows that $\mathfrak{g}_i = \mathrm{pr}_i(\mathfrak{n})$ for each i, and we know by (i) that each $\mathfrak{g}_i$ is already nonzero and simple. Lastly, the previous step actually checks condition (ii) for the inductive hypothesis, completing the proof that $\mathfrak{n} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$.

4. We claim $I = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$. Because $I \subseteq \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$ is an ideal of a sum of simple algebras, we know that
$$I = \bigoplus_{i \in S} \mathfrak{g}_i$$
for some subset $S \subseteq \{1, \ldots, k-1\}$ of indices. Thus, to achieve the equality $I \stackrel{?}{=} \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_{k-1}$, it is enough to check that each projection $\mathrm{pr}_i \colon I \to \mathfrak{g}_{k-1}$ is surjective. Unravelling the definition of I, it is enough to check that each $X_i \in \mathfrak{g}_i$ has some $(X_1, \ldots, X_k) \in \mathfrak{g}$ with the correct X_i coordinate and $X_k = 0$. This last claim follows from hypothesis (ii) of $\mathfrak{g}$!

5. We now finish the proof of the lemma. Certainly $\mathfrak{g} \subseteq \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_k$, so it is enough to compute dimensions to prove the equality. By the short exact sequence
$$0 \to J \to \mathfrak{g} \to \mathfrak{g}_n \to 0,$$
it is enough to show that $\dim J = \dim \mathfrak{g}_1 + \cdots + \dim \mathfrak{g}_{k-1}$. However, this follows from the previous step because $\dim J = \dim I$. ∎

In practice, it is somewhat difficult to check (ii) of Lemma 1.61. Here is an automation.

Lemma 1.62 (Moonen–Zarhin)**.** Work over an algebraically closed field of characteristic 0. Let $V_1, \ldots, V_k$ be finite-dimensional vector spaces, and let $\mathfrak{g}$ be a Lie subalgebra of $\mathfrak{gl}(V_1) \times \cdots \times \mathfrak{gl}(V_k)$. For each index i, let $\mathrm{pr}_i \colon (\mathfrak{gl}(V_1) \times \cdots \times \mathfrak{gl}(V_k)) \to \mathfrak{gl}(V_i)$ be the ith projection, and set $\mathfrak{g}_i := \mathrm{pr}_i(\mathfrak{g})$. Suppose the following.

(i) Each $\mathfrak{g}_i$ is nonzero and simple.

(ii) Fix a simple Lie algebra $\mathfrak{l}$, and define $I(\mathfrak{l}) := \{i : \mathfrak{g}_i \cong \mathfrak{l}\}$. If $\#I(\mathfrak{l}) > 1$, we require the following to hold.

- All automorphisms of $\mathfrak{l}$ are inner.
- One can choose isomorphisms $\mathfrak{l} \to \mathfrak{g}_i$ for each $i \in I(\mathfrak{l})$ such that the representations $\mathfrak{l} \to \mathfrak{g}_i \to \mathfrak{gl}(V_i)$ are all isomorphic.
- The diagonal inclusion
$$\prod_{i \in I(\mathfrak{l})} \mathrm{End}_{\mathfrak{g}_i}(V_i) \to \mathrm{End}_{\mathfrak{g}} \left(\bigoplus_{i \in I(\mathfrak{l})} V_i \right)$$
is surjective.

Then $\mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_k$.

Proof. We will show that (ii) in the above lemma implies (ii) of Lemma 1.61, which will complete the proof. We will proceed by contraposition in the following way. Fix a pair (i, j) of distinct indices, and we are interested in the map $(\mathrm{pr}_i, \mathrm{pr}_j) \colon \mathfrak{g} \to \mathfrak{g}_i \times \mathfrak{g}_j$. Supposing that $(\mathrm{pr}_i, \mathrm{pr}_j)$ fails to be surjective (which is a violation of (ii) of Lemma 1.61), we will show that (ii) cannot be true. In particular, we will assume the first two points of (ii) and show then that the third point of (ii) is false.

Roughly speaking, we are going to use the first two points of (ii) to find an $\mathfrak{h}$ and then produce an endomorphism of $\bigoplus_{i \in I(\mathfrak{h})} V_i$ which does not come from gluing together endomorphisms of the V_is. Having stated the outline, we proceed with the proof in steps.

1. We claim that the image $\mathfrak{h}$ of the map $(\mathrm{pr}_i, \mathrm{pr}_j)\colon \mathfrak{g} \to \mathfrak{g}_i \times \mathfrak{g}_j$ is the graph of an isomorphism $\mathfrak{g}_i \to \mathfrak{g}_j$. For this, we use the hypothesis that $(\mathrm{pr}_i, \mathrm{pr}_j)$ fails to be surjective. Well, we claim that the projections $\mathfrak{h} \to \mathfrak{g}_i$ and $\mathfrak{h} \to \mathfrak{g}_j$ are isomorphisms, which implies that $\mathfrak{h}$ is the graph of the composite isomorphism

$$\mathfrak{g}_i \leftarrow \mathfrak{h} \to \mathfrak{g}_j.$$

   By symmetry, it is enough to merely check that $\mathfrak{h} \to \mathfrak{g}_i$ is an isomorphism. On one hand, $\mathfrak{h} \to \mathfrak{g}_i$ is surjective because $\mathrm{pr}_i\colon \mathfrak{g} \to \mathfrak{g}_i$ is surjective by construction of $\mathfrak{g}_i$. On the other hand, the kernel of the projection $\mathfrak{h} \to \mathfrak{g}_i$ will be an ideal of $\mathfrak{h}$ of the form $0 \oplus I$ where $I \subseteq \mathfrak{g}_j$ is some subspace. In fact, becasue the projection $\mathfrak{h} \to \mathfrak{g}_j$ is also surjective, we see that $I \subseteq \mathfrak{g}_j$ must be an ideal, so the simplicity of $\mathfrak{g}_j$ grants two cases.

   - If $I = 0$, then $\mathrm{pr}_i\colon \mathfrak{h} \to \mathfrak{g}_i$ becomes injective and is thus an isomorphism, completing this step.
   - If $I = \mathfrak{g}_j$, then $\mathfrak{h}$ fits into a short exact sequence

     $$0 \to (0 \oplus \mathfrak{g}_j) \to \mathfrak{h} \to \mathfrak{g}_i \to 0,$$

     so $\dim \mathfrak{h} = \dim(\mathfrak{g}_i \oplus \mathfrak{g}_j)$, implying the inclusion $\mathfrak{h} \subseteq \mathfrak{g}_i \oplus \mathfrak{g}_j$ is an equality. However, this cannot be the case because we assumed that $(\mathrm{pr}_i, \mathrm{pr}_j)\colon \mathfrak{g} \to \mathfrak{g}_i \to \mathfrak{g}_j$ fails to be surjective!

2. We construct an isomorphism of $\mathfrak{g}$-representations $V_i \to V_j$. For this, we use the first two points of (ii). Let's begin by collecting some data.

   - The previous step informs us that $\mathfrak{g}_i \cong \mathfrak{g}_j$. In fact, because this isomorphism is witnessed by the projections $\mathrm{pr}_i\colon \mathfrak{g} \to \mathfrak{g}_i$ and $\mathrm{pr}_j\colon \mathfrak{g} \to \mathfrak{g}_j$, we see that we are granted an isomorphism $f\colon \mathfrak{g}_i \to \mathfrak{g}_j$ such that $\mathrm{pr}_j = f \circ \mathrm{pr}_i$.
   - We now let $\mathfrak{l}$ be a simple Lie algebra isomorphic to both(!) $\mathfrak{g}_i$ and $\mathfrak{g}_j$. The second point of (ii) grants isomorphisms $f_i\colon \mathfrak{l} \to \mathfrak{g}_i$ and $f_j\colon \mathfrak{l} \to \mathfrak{g}_j$ of Lie algebras and an isomorphism $d\colon V_i \to V_j$ of $\mathfrak{l}$-representations.

   We now construct our isomorphism from d. Because d is only an isomorphism of $\mathfrak{l}$-representations, we are only granted that $(X_1, \dots, X_k) \in \mathfrak{g}$ satisfies $f(X_i) = X_j$ and hence

$$\begin{aligned} d\left((f_i f_j^{-1} f)(X_i) v_i\right) &= d\left(f_i\left(f_j^{-1} f(X_i)\right) v_i\right) \\ &= f_j\left(f_j^{-1} f(X_i)\right) d(v_i) \\ &= X_j d(v_i) \end{aligned}$$

   for all $v_i \in V_i$. We would be done if we could remove the pesky automorphism $f_i f_j^{-1} f\colon \mathfrak{g}_i \to \mathfrak{g}_i$. This is possible because all automorphisms of $\mathfrak{g}_i \cong \mathfrak{l}$ are inner (!), so one may simply "change bases" to remove the inner automorphism. Explicitly, find $a \in \mathrm{GL}(V_i)$ such that $f_i f_j^{-1} f(X) = aXa^{-1}$ for all $X \in \mathfrak{g}_i$, and then we define $e \coloneqq d \circ a$. Then we find that any $v_i \in V_i$ has

$$\begin{aligned} e(X_i v_i) &= d\left(aX_i a^{-1} \cdot av\right) \\ &= d\left((f_i f_j^{-1} f)(X_i) \cdot av\right) \\ &= X_j d(av) \\ &= X_j e(v). \end{aligned}$$

3. We complete the proof. The previous step provides a morphism $e\colon V_i \to V_j$ of $\mathfrak{g}$-representations. We thus note that the composite

$$\bigoplus_{i' \in I(\mathfrak{l})} V_{i'} \twoheadrightarrow V_i \xrightarrow{e} V_j \hookrightarrow \bigoplus_{i' \in I(\mathfrak{l})} V_{i'}$$

   is an endomorphism which does not come from the diagonal inclusion of $\prod_{i \in I(\mathfrak{l})} \mathrm{End}_{\mathfrak{g}_i}(V_i)$. This completes the proof by showing that the third point of (ii) fails to hold. ■

Remark 1.63. We should remark on some history. Lemma 1.61 is due to Ribet [Rib76, pp. 790–791], but the given formulation is due to Moonen and Zarhin [MZ95, Lemma 2.14]. In the same lemma, Moonen and Zarhin prove Lemma 1.62, and they seem to be the first to recognize the utility of this lemma for computing Hodge groups. For example, Lombardo includes this result in his master's thesis [Lom13, Lemma 3.3.1] and includes a generalized version in another paper as [Lom16, Lemma 3.7], where it is used to compute Hodge groups of certain products of abelian varieties.

Remark 1.64. Let's explain how Lemma 1.62 is typically applied, which is admittedly somewhat different from the application used in this thesis. In the generic case, one expects (i), for example if $\mathrm{Hg}(V) = \mathrm{Sp}_D(\varphi)^\circ$ for D of Types I–III as in Remark 1.51. In this case, one can also check the first condition of (ii) by a direct computation, the second condition of (ii) has no content, and the third condition of (ii) comes from Lemma 1.54. For more details, we refer to (for example) the applications given in [Lom13; Lom16].

1.2.5 The Lefschetz Group

For motivational reasons, we mention the Lefschetz group $\mathrm{L}(V)$, which contains $\mathrm{Hg}(V)$ but is more controlled. Here is our definition.

Definition 1.65 (Lefschetz group)**.** Fix a polarizable Hodge structure $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight. Then we define

$$\mathrm{L}(V) \coloneqq \mathrm{Sp}_D(\varphi),$$

where $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$, and φ is a polarization.

Thus, Remark 1.53 that $\mathrm{Hg}(V) \subseteq \mathrm{L}(V)$.

Remark 1.66. Let's interpret $\mathrm{L}(V)$ geometrically. Roughly speaking, $\mathrm{L}(V)$ is a form of $\mathrm{Hg}(V)$ which only keeps track of endomorphisms and the polarization instead of keeping track of all Hodge classes. As such, we generically expect $\mathrm{Hg}(V) = \mathrm{L}(V)$ to hold, but we do not expect it to hold always. (Technically, there are generic cases when we do not expect this equality; for example, if V is irreducible of Type III in ths sense of the Albert classificaiotn Theorem 1.28, then $\mathrm{L}(V)$ is not connected, so we cannot have equality.) Furthermore, when $\mathrm{Hg}(V) = \mathrm{L}(V)$, we expect to have strong control on the Hodge classes of V; for example, the Hodge conjecture is known in many such cases [Mur84, Theorem 3.1].

Computationally, one reason why $\mathrm{L}(V)$ is more controlled is that it is much easier to compute. For example, L behaves well in sums.

Lemma 1.67. Fix pairwise non-isomorphic irreducible polarizable Hodge structures $V_1, \dots, V_k$ of the same pure weight, and let $m_1, \dots, m_k \geq 1$ be integers. Then the diagonal embeddings $\Delta_i \colon \mathrm{GL}(V_i) \to \mathrm{GL}\left(V_i^{\oplus m_i}\right)$ induce an isomorphism

$$\mathrm{L}(V_1) \times \cdots \times \mathrm{L}(V_k) \to \mathrm{L}\left(V_1^{\oplus m_1} \oplus \cdots \oplus V_k^{\oplus m_k}\right).$$

Proof. The main idea is to compute some endomorphism algebras and polarizations. We proceed in steps. Set $V \coloneqq V_1^{\oplus m_1} \oplus \cdots \oplus V_k^{\oplus m_k}$ for brevity.

1. We work with endomorphisms. We may view Hodge structures as $\mathbb{S}$-representations, whereupon we find that
$$\mathrm{End}_{\mathrm{HS}}(V) = \mathrm{End}_{\mathrm{HS}}(V_1)^{m_1 \times m_1} \times \cdots \times \mathrm{End}_{\mathrm{HS}}(V_k)^{m_k \times m_k}.$$
In particular, we see that any f commuting with $\mathrm{End}_{\mathrm{HS}}(V)$ implies that f must preserve each $V_i^{\oplus m_i}$ (because there is a separate algebra $\mathrm{End}_{\mathrm{HS}}\left(V_i^{\oplus m_i}\right)$ for each i). Further, $f|_{V_i^{\oplus m_i}}$ must come from the

diagonal embedding $\mathrm{End}_{\mathrm{HS}}(V_i) \to \mathrm{End}_{\mathrm{HS}}\left(V_i^{\oplus m_i}\right)$ because $\mathrm{End}_{\mathrm{HS}}(V_i)^{m_i \times m_i}$ may swap any of the m_i copies of V_i.

We conclude that f commutes with endomorphisms implies that

$$f = (\Delta_1 f_1, \dots, \Delta_k f_k),$$

where $\Delta_i \colon \mathrm{End}(V_i) \to \mathrm{End}\left(V_i^{\oplus m_i}\right)$ is the diagonal embedding, and each f_i commutes with $\mathrm{End}_{\mathrm{HS}}(V_i)$. Conversely, the computation of $\mathrm{End}_{\mathrm{HS}}(V)$ above allows us to conclude that any f in the above form commutes with $\mathrm{End}_{\mathrm{HS}}(V)$.

2. We work with the polarization. Choose polarizations $\varphi_1, \dots, \varphi_k$ on $V_1, \dots, V_k$ (respectively), and we note that these polarizations glue into a polarization φ on V. With this choice of polarization, we see that $f = (\Delta_1 f_1, \dots, \Delta_k f_k)$ as in the previous step preserves φ if and only if each factor $\Delta_i f_i$ preserves the polarization $\varphi|_{V_i^{\oplus m_i}}$, which is equivalent to f_i preserving the polarization φ_i. In total, we thus see that $f \in \mathrm{L}(V)$ if and only $f_i \in \mathrm{L}(V_i)$ for each i, so we are done. ∎

Lemma 1.67 tells us that we can always reduce the computation of the Lefschetz group to irreducible components. In this way, it now suffices to compute $\mathrm{L}(V)$ by working with V according to the Albert classification (Theorem 1.28). All these computations are recorded in [Mil99, Section 2]. Because we will only be interested in Type IV in the sequel, we will only record the part of this computation we need for completeness.

Lemma 1.68. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight with $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ and polarization φ. Suppose $D = F$ is a CM field. Then

$$\mathrm{L}(V)_{\mathbb{C}} \cong \mathrm{GL}_{[V:F]}(\mathbb{C})^{\frac{1}{2}[F:\mathbb{Q}]}.$$

Proof. We proceed in steps. Let $F^\dagger \subseteq F$ be the maximal totally real subfield, and choose embeddings $\rho_1, \dots, \rho_{e_0} \colon F^\dagger \hookrightarrow \mathbb{R}$, where $e_0 \coloneqq \frac{1}{2}[F:\mathbb{Q}]$. For each i, we will let σ_i and τ_i be complex conjugate embeddings $F \hookrightarrow \mathbb{C}$ restricting to ρ_i.

1. We begin by explaining the exponent $e_0 = \frac{1}{2}[F:\mathbb{Q}]$. Note V is a free $F^\dagger$-module of rank $[V:F]$, so $V_{\mathbb{R}}$ is a free module over

$$F^\dagger \otimes \mathbb{R} = \prod_{i=1}^{e_0} F^\dagger_{\rho_i},$$

where $F^\dagger_\rho = \mathbb{R}$ refers to the $F^\dagger \otimes \mathbb{R}$ module where F acts by ρ. The above decomposition of $F \otimes \mathbb{R}$ implies a decomposition

$$V_{\mathbb{R}} = V_1 \oplus \cdots \oplus V_{e_0},$$

where each V_i of a vector space over $F^\dagger_{\rho_i}$, all the same dimension.

We now understand the effect of endomorphisms and the polariaztion on our decomposition. Thus, we see that $f \colon V_{\mathbb{R}} \to V_{\mathbb{R}}$ commutes with $F^\dagger \otimes \mathbb{R}$ if and only if f preserves each factor V_i (due to the decomposition of $F^\dagger \otimes \mathbb{R}$) and commute with the action of $F^\dagger_{\rho_i}$ on each V_i. Similarly, we see that the polarization φ makes the V_is orthogonal: for each $d \in F^\dagger$, we see that any $v_i \in V_i$ and $v_j \in V_j$ has

$$\begin{aligned}
\rho_i(d)\varphi(v_i, v_j) &= \varphi(dv_i, v_j) \\
&= \varphi(v_i, \overline{d} v_j) \\
&= \varphi(v_i, d v_j) \\
&= \rho_j(d)\varphi(v_i, v_j),
\end{aligned}$$

so $i \neq j$ implies that $\varphi(v_i, v_j) = 0$. Thus, we see that φ must restrict to non-degenerate skew-symmetric bilinear forms on each V_i individually. In total, $f \colon V_{\mathbb{R}} \to V_{\mathbb{R}}$ preserves φ if and only if $f|_{V_i}$ preserves $\varphi|_{V_i}$ for each i. In total, we see that

$$\mathrm{L}(V)_{\mathbb{R}} = \mathrm{Sp}_{F \otimes_{\rho_1} \mathbb{R}}(\varphi|_{V_1}) \times \cdots \times \mathrm{Sp}_{F \otimes_{\rho_k} \mathbb{R}}(\varphi|_{V_{e_0}}).$$

2. It remains to show that $\mathrm{Sp}_{F\otimes_{\rho_i}\mathbb{R}}(\varphi|_{V_i})_{\mathbb{C}}$ is isomorphic to $\mathrm{GL}_{[V:F]}(\mathbb{C})$; here, note $[V:F]=[V_i:F_{\rho_i}^\dagger]$. For this, we abstract the situation somewhat: suppose that a vector space V over $\mathbb{R}$ has been equipped with an action by $\mathbb{C}\subseteq\mathrm{End}_{\mathbb{R}}(V)$, and furthermore, φ is a skew-Hermitian form on V. Then we want to show $\mathrm{Sp}_{\mathbb{C}}(\varphi)_{\mathbb{C}}\cong\mathrm{GL}_{[V:\mathbb{R}]}(\mathbb{C})$.

   The trick is that we can keep track of commuting with the action of $\mathbb{C}$ on V by merely commuting with the action of $i\in\mathbb{C}$. Thus, let $J\colon V\to V$ be this map, which satisfies $J^2=-1$. Now, the action of $J_{\mathbb{C}}$ on $V_{\mathbb{C}}$ must diagonalize into eigenspaces $V_i\oplus V_{-i}$ with eigenvalues i and $-i$ respectively; note that we must have $\dim V_i=\dim V_{-i}$ in order for the characateristic polynomial of J to have real coefficients. The point is that $f\in\mathrm{End}(V_{\mathbb{C}})$ commutes with the action of $\mathbb{C}$ if and only if it commutes with the action of J, which we can see is equivalent to f preserving the decomposition $V_i\oplus V_{-i}$.

   We now study the polarization φ. Note that φ vanishes on $V_{\pm i}\oplus V_{\pm i}$: for any $v,v'\in V_{\pm i}$, we see that

$$\begin{aligned}\pm i\varphi(v,v')&=\varphi(Jv,v')\\&=\varphi(v,-Jv')\\&=\mp i\varphi(v,v'),\end{aligned}$$

   from which $\varphi(v,v')=0$ follows. For example, this implies that any $f\in\mathrm{End}(V_{\mathbb{C}})$ commuting with the J-action will automatically preserve φ on $V_{\pm i}\times V_{\pm i}$. Additionally, we see that φ must restrict to a non-degenerate bilinear form on $V_i\times V_{-i}$.

   We are now ready to claim that restriction defines an isomorphism $\mathrm{Sp}_{\mathbb{C}}(\varphi)_{\mathbb{C}}\to\mathrm{GL}_{\mathbb{C}}(V_i)$. This restriction does actually output to $\mathrm{GL}_{\mathbb{C}}(V_i)$ because $g\in\mathrm{Sp}_{\mathbb{C}}(\varphi)_{\mathbb{C}}$ must preserve the decomposition $V_i\oplus V_{-i}$. To see the injectivity, we note that preserving φ requires

$$\varphi(v,gw)=\varphi\left(g^{-1}v,w\right)$$

   for all $v\in V_i$ and $w\in V_{-i}$; thus, the non-degeneracy of φ implies that $g\in\mathrm{Sp}_{\mathbb{C}}(\varphi)_{\mathbb{C}}$ is uniquely determined by its action on V_i. Conversely, for the surjectivity, we see that we can take any element in $\mathrm{GL}(V_i)$ and use the previous sentence to extend it uniquely to an element of $\mathrm{Sp}_{\mathbb{C}}(\varphi)_{\mathbb{C}}$. ■

1.3 Absolute Hodge Classes

We now discuss the main application of Hodge structures: cohomology. This will allow us to discuss absolute Hodge classes. Our exposition an abbreviated form [Del18].

1.3.1 Some Cohomology Theories

In this subsection, we will give a lighting introduction to the cohomology theories that we will use. We begin with sheaf cohomology.

Definition 1.69 (sheaf cohomology)**.** Fix a topological space X. Then the category $\mathrm{Ab}(X)$ of abelian sheaves on X has enough injectives. Given a sheaf $\mathcal{F}$ on X, we then may define the *sheaf cohomology* as the abelian groups

$$\mathrm{H}^i(X,\mathcal{F})\coloneqq\mathrm{R}^i\Gamma(X,\mathcal{F}),$$

where $\Gamma\colon\mathrm{Ab}(X)\to\mathrm{Ab}$ is the global-sections functor. Explicitly, one can compute these cohomology groups by taking the cohomology of an acyclic resolution of $\mathcal{F}$.

This allows us a quick definition of Betti cohomology.

Definition 1.70 (Betti cohomology)**.** Fix a topological space X and a ring R. Then we define the *Betti cohomology* of X with coefficients in R as $\mathrm{H}^i(X,\underline{R})$, where $\underline{R}$ denotes the constant sheaf R.

It will be helpful to a more geometric description of $\mathrm{H}^\bullet_{\mathrm{B}}$.

Definition 1.71 (singular homology, singular cohomology)**.** Fix a topological space X and a ring R. For each $n \geq 0$, we define the n-simplex $\Delta^n \subseteq \mathbb{R}^{n+1}$ as the set of points $(t_0, \dots, t_n) \subseteq [0,1]^{n+1}$ summing to 1. Then we define the complex $S_\bullet(X, R)$ as having entries which are the free R-module with basis given by the maps $\Delta_\bullet \to X$ and boundary morphism given by $\partial \colon S_n(X,R) \to S_{n-1}(X,R)$ given by

$$\partial(\sigma) \coloneqq \sum_{i=0}^{n} (-1)^i \sigma([0, \dots, \widehat{i}, \dots, n])$$

for $\sigma \colon \Delta_n \to X$, where $[0, \dots, \widehat{i}, \dots, n]$ denotes the $(n-1)$-simplex with vertices $\{0, \dots, \widehat{i}, \dots, n\}$. Then we define the *singular homology* $\mathrm{H}^{\mathrm{B}}_i(X,R)$ as the homology of this complex. We now define *singular cohomology* as the cohomology of the dual cocomplex $S^\bullet(X,R)$.

Remark 1.72. The universal coefficient theorem shows that singular homology and cohomology are dual if R is a principal ideal domain, such as $\mathbb{Z}$ or a field.

Our notation suggests that singular cohomology should be Betti cohomology, so we check this.

Theorem 1.73. Fix a topological manifold X. For any field K, there is a canonical isomorphism

$$\mathrm{H}^i(S^\bullet(X,K)) \to \mathrm{H}^i(X, \underline{K}).$$

Proof. The idea is to replace $S^\bullet(X,K)$ with a complex of sheaves $\mathcal{S}^\bullet(X,K)$, and then one finds that this complex is an acyclic resolution of $\underline{K}$. The requirement that X be a topological manifold helps because it allows us to reduce local checks on X to the case of a unit ball. ∎

We now add smoothness to our manifolds, which allows us to define de Rham cohomology.

Definition 1.74 (de Rham cohomology)**.** Fix a smooth manifold X of dimension n. For each $i \geq 0$, we let $\Omega^i_{X_\infty}$ be the sheaf of smooth differential n-forms on X. Then we define *de Rham cohomology* $\mathrm{H}^i_{\mathrm{dR}}(X, \mathbb{R})$ to be the cohomology of the complex

$$0 \to \Omega^0_{X_\infty} \xrightarrow{d} \Omega^2_{X_\infty} \xrightarrow{d} \cdots \xrightarrow{d} \Omega^n_{X_\infty} \to 0,$$

where d denotes the de Rham differential.

We once again have a comparison isomorphism.

Theorem 1.75. Fix a smooth manifold X. For each i, there is a functorial perfect paring $\mathrm{H}^{\mathrm{B}}_i(X,\mathbb{R}) \times \mathrm{H}^i_{\mathrm{dR}}(X,\mathbb{R}) \to \mathbb{R}$ given by

$$\langle \sigma, \omega \rangle \coloneqq \int_\sigma \omega$$

for each smooth map $\sigma \colon \Delta^i \to X$.

We next upgrade to complex Kähler manifolds. For example, one can upgrade our de Rham cohomology to use holomorphic differential forms instead of smooth differential forms, and the cohomology does not change. The key benefit of the complex manifold situation is that the de Rham cohomology gains a Hodge structure.

Theorem 1.76. Fix a compact complex Kähler manifold X. For each $n \geq 0$, there is a decomposition

$$\mathrm{H}^i_{\mathrm{dR}}(X, \mathbb{C}) = \bigoplus_{p+q=n} \mathrm{H}^{pq}(X),$$

where $\mathrm{H}^{pq}(X) \coloneqq \mathrm{H}^p(X, \Omega^q_X)$.

For our last setting, let X be a smooth projective variety over a field K. Here, there are multiple ways to form Betti cohomology.

Notation 1.77. Fix a smooth projective variety over a field K. For any embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, we define Betti cohomology relative to σ as

$$\mathrm{H}^i_\sigma(X, R) \coloneqq \mathrm{H}^i_{\mathrm{B}}(X_\sigma(\mathbb{C}), R)$$

for any ring R. Frequently, we will have fixed once and for all an embedding of K into $\mathbb{C}$, so we may abbreviate $\mathrm{H}^i_\sigma(X, R)$ to just $\mathrm{H}^i_{\mathrm{B}}(X, R)$.

Similarly, one is now able to define de Rham cohomology for X, though we do make a moment to remark that there is a theory of algebraic de Rham cohomology that is able to work in greater generality.

Working with varieties gives access to another cohomology theory we will need.

Definition 1.78. Fix a smooth projective variety X over a field K. For some étale sheaf $\mathcal{F}$, we are able to define the étale cohomology $\mathrm{H}^i(X, \mathcal{F})$ in the same way as sheaf cohoomology. In particular, for any prime ℓ which is nonzero in K, we define the *ℓ-adic cohomology* by

$$\mathrm{H}^i_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell) \coloneqq \left(\varprojlim \mathrm{H}^i_{\text{ét}}(X_{\overline{K}}, \underline{\mathbb{Z}/\ell^\bullet\mathbb{Z}})\right) \otimes_{\mathbb{Z}} \mathbb{Q}$$

Importantly, we note that étale cohomology has the natural action by $\mathrm{Gal}(\overline{K}/K)$. As usual, there is a comparison isomorphism.

Theorem 1.79. Fix a smooth projective variety X over $\mathbb{C}$. Then there is a natural isomorphism

$$\mathrm{H}^i_{\mathrm{B}}(X, \mathbb{Q}_\ell) \to \mathrm{H}^i_{\text{ét}}(X, \mathbb{Q}_\ell).$$

We may find it convenient to glue our cohomology theories together.

Notation 1.80. Fix a smooth projective variety X over a field K with an embedding $\sigma\colon K \hookrightarrow \mathbb{C}$. Then we define

$$\mathrm{H}^i_{\mathbb{A}}(X) \coloneqq \mathrm{H}^i_{\mathrm{dR}}(X, \mathbb{R}) \times \left(\varprojlim_n \mathrm{H}^i_{\text{ét}}(X_{\overline{K}}, \underline{\mathbb{Z}/n\mathbb{Z}}) \otimes_{\mathbb{Z}} \mathbb{Q}\right).$$

We note that there are natural projections π_∞ onto $\mathrm{H}^i_{\mathrm{dR}}(X, \mathbb{R})$ and π_ℓ onto $\mathrm{H}^i_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$.

Remark 1.81. One can realize this as a restricted direct product

$$\mathrm{H}^i_{\mathrm{dR}}(X, \mathbb{R}) \times \prod_\ell \left(\mathrm{H}^i_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell), \mathrm{H}^i_{\text{ét}}(X_{\overline{K}}, \mathbb{Z}_\ell)\right),$$

which provides some motivation for the $\mathbb{A}$ in the notation.

The above cohomology theories have mostly worked in characteristic 0, but it is notable that étale cohomology has managed to work with $\mathrm{char}\, K \nmid \ell$. This does leave a gap in positive characteristic, where we may

want to define a cohomology "at" $\operatorname{char} K = \ell$. This is achieved by crystalline cohomology; we will not use crystalline cohomology very much, but we go ahead and give some of its basic properties. We refer to [Ill94] for a more thorough review.

Definition 1.82. Fix a smooth projective variety X over a perfect field K of positive characteristic p. Let W be the ring of Witt vectors over K, and define $W_n \coloneqq W/p^n W$ for each $n \geq 0$. Then we define the *crystalline cohomology* $\mathrm{H}^\bullet_{\mathrm{crys}}(X, W)$ as

$$\mathrm{H}^i_{\mathrm{crys}}(X, W) \coloneqq \varprojlim \mathrm{H}^i_{\mathrm{crys}}(X, W_n),$$

where the right-hand cohomology is the cohomology on the crystalline site of X over W_n.

Remark 1.83. Because X has an absolute Frobenius, the crystalline cohomology groups come with the action by an absolute Frobenius.

Theorem 1.84. Fix a smooth projective variety X over a perfect field K of positive characteristic p, and set $W \coloneqq W(K)$. If there is a smooth projective variety Z over W such that $Z_K = X$, then there is a natural isomorphism

$$\mathrm{H}^\bullet_{\mathrm{crys}}(X, W) \to \mathrm{H}^\bullet_{\mathrm{dR}}(Z, W).$$

Thus far, we have defined many cohomology theories, so it is worthwhile to explain why one may expect them to somehow be related to one another. We have already mentioned a few comparison isomorphisms, but it also turns out that they all have other properties which tie them together. For example, they all have a cup product, which turn the collection of cohomology groups $\mathrm{H}^i(X)$ into a graded commutative ring $\mathrm{H}^\bullet(X)$. There is also some functoriality: for a map $f\colon X \to Y$ of spaces, there is always an induced pullback map

$$f^*\colon \mathrm{H}^\bullet(Y) \to \mathrm{H}^\bullet(X),$$

which turns out to be a homomorphism of graded algebras.

As a more complicated example, there is a Künneth formula: for any of the above cohomology theories H defined on a space X and Y, there is an isomorphism

$$\mathrm{H}^n(X \times Y) = \bigoplus_{i+j=n} \mathrm{H}^i(X) \otimes \mathrm{H}^j(Y).$$

Of course, it is a major theorem among each of our cohomology theories that the Künneth formula is satisfied, which we will not prove.

There is also a notion of Poincaré duality. To explain Poincaré duality, we need some twists.

Definition 1.85 (Tate twist)**.** We define our Tate twists as follows.

- If X is a topological manifold, then the Tate twist $\mathbb{Q}_{\mathrm{B}}(1)$ is the $\mathbb{Q}$-vector space $2\pi i\mathbb{Q}$.
- If X is a smooth manifold, then the Tate twist $\mathbb{R}_{\mathrm{dR}}(1)$ is simply $\mathbb{R}$. It has a Hodge structure of pure of weight -2 concentrated in bidegree $(-1, -1)$.
- If X is a smooth projective variety over a field K, then the Tate twist $\mathbb{Q}_\ell(1)$ for any prime ℓ (nonzero in K) is the Galois representation $(\varprojlim \mu_{\ell^\bullet}) \otimes_{\mathbb{Z}} \mathbb{Q}$.

Notation 1.86. For any cohomology theory H defined on a space X, we may write

$$\mathrm{H}^i(X)(n) \coloneqq \mathrm{H}^i(X) \otimes \mathrm{T}^{\otimes n},$$

where T denotes the Tate twist, and $i \geq 0$ and $n \in \mathbb{Z}$. If $n \leq 0$, then we take the dual.

Now, for any of these cohomology theories H over a field F defined on a space X of equidimension d, Poincaré duality provides a perfect pairing

$$\mathrm{H}^i(X) \otimes \mathrm{H}^{2d-i}(X)(d) \to F$$

for each index i. Once again, it is a major theorem among each of our cohomology theories above that Poincaré duality is satisfied.

1.3.2 Weil Cohomology Theories

It will be worth our time to encode everything we need that the above cohomology theories have in common. In essence, we are asking for a formalism of a cohomology theory, which is known as a Weil cohomology theory. Approximately speaking, a Weil cohomology theory is a cohomology theory with the minimum amount of data to prove the Lefschetz trace formula without too much pain. Our exposition here follows [SP, Tag 0FFG]. Throughout, we freely use facts about intersection theory and Chow groups because the author is too ignorant to provide a suitable review of these notions; everything we need can be found in [Ful98].

Throughout, we fix a base field K and a coefficient field F. We require $\operatorname{char} F = 0$, but we do not require K to be algebraically closed. These hypotheses will not be repeated!

Notation 1.87. Let $\mathcal{P}(K)$ denote the category of smooth projective varieties over K, with morphisms given by regular maps.

Here is the data we will be working with.

Definition 1.88 (Weil cohomology datum)**.** A *Weil cohomology datum* consists of the following data.

- A one-dimensional F-vector space $F(1)$.
- A contravariant functor $\mathrm{H}^\bullet$ from $\mathcal{P}(K)$ to the category of $\mathbb{Z}$-graded commutative F-algebras. We will write the product as a cup $\cup$.
- For $X \in \mathcal{P}(K)$ of equidimension d, there is a trace map $\int_X : \mathrm{H}^{2d}(X)(d) \to F$.
- For $X \in \mathcal{P}(K)$, there is a cycle class map $\mathrm{cl}_X : \mathrm{CH}^i(X) \to \mathrm{H}^{2i}(X)(i)$, which is required to be a group homomorphism.

Frequently, we will call $\mathrm{H}^\bullet$ alone the Weil cohomology datum, leaving the other inputs implied.

In short, $F(1)$ is the Tate twist, $\mathrm{H}^\bullet$ are the vector spaces one usually remembers with Weil cohomology theories, $\int_X$ keeps track of Poincaré duality, and cl_X relates cohomology to geometry.

In order to keep us thinking "cohomologically," we use some special notation.

Notation 1.89. Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F.

- For any F-vector space V, we write $V(n) \coloneqq V \otimes F(1)^{\otimes n}$. Here, negative exponents denote duals.
- If $f : X \to Y$ is a regular map, we let $f^* : \mathrm{H}^\bullet(Y) \to \mathrm{H}^\bullet(X)$ denote the induced ring homomorphism.

Remark 1.90. In the sequel, we may note that $f * (\alpha \cup \beta) = f^*\alpha \cup f^*\beta$ without comment: indeed, this follows because f^* is a ring homomorphism! Similarly, we may use the fact that $(g \circ f)^* = f^* \circ g^*$, which follows because the functor $\mathrm{H}^\bullet$ is contravariant.

Now, a Weil cohomology datum is going to be required to satisfy many axioms. Before going further, let's summarize them.

- We need a Künneth formula to ensure that products of varieties go to products in graded algebras.
- We need Poincaré duality, for example to define pushfowards. This adds some coherence to the cycle class maps.
- To add some geometric input to the picture, we need some coherence of our cycle class maps.
- Lastly, we will need another axiom to ensure that, for example, H is only supported in nonnegative indices.

Let's begin with the Künneth formula.

Definition 1.91 (Künneth formula)**.** Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F. Then $\mathrm{H}^\bullet$ satisfies the *Künneth formula* if and only if it satisfies the following for all $X, Y \in \mathcal{P}(K)$.

(a) Künneth formula: the map

$$\begin{aligned} \mathrm{H}^\bullet(X) \otimes \mathrm{H}^\bullet(Y) &\to \mathrm{H}^\bullet(X \times Y) \\ \alpha \otimes \beta &\mapsto \mathrm{pr}_1^* \alpha \cup \mathrm{pr}_2^* \beta \end{aligned}$$

is an isomorphism of graded F-algebras. We may write $\alpha \boxtimes \beta \coloneqq \mathrm{pr}_1^* \alpha \cup \mathrm{pr}_2^* \beta$.

(b) Fubini's theorem: if X and Y have equidimension d and e, respectively, then

$$\int_{X \times Y} (\alpha \boxtimes \beta) = \int_X \alpha \cdot \int_Y \beta$$

for any $\alpha \in \mathrm{H}^{2d}(X)(d)$ and $\beta \in \mathrm{H}^{2e}(Y)(e)$.

Remark 1.92. It is worth recalling the grading on the tensor product of two graded vector spaces: if V and W are $\mathbb{Z}$-graded vector spaces, then $(V \otimes W)$ has a grading given by

$$(V \otimes W)_n = \bigoplus_{i+j=n} V_i \otimes W_j.$$

In particular, we see that satisfying the Künneth formula implies that there is a canonical isomorphism

$$\bigoplus_{i+j=n} \mathrm{H}^i(X) \otimes \mathrm{H}^j(Y) \to \mathrm{H}^n(X \times Y).$$

It is worth noting that the Künneth formula has good functoriality properties.

Lemma 1.93. Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F satisfying the Künneth formula. Given morphisms $f\colon X \to X'$ and $g\colon Y \to Y'$ in $\mathcal{P}(K)$, we have

$$(f \times g)^* = f^* \otimes g^*.$$

Proof. Note that these are both automatically ring maps $\mathrm{H}^\bullet(X' \times Y') \to \mathrm{H}^\bullet(X \times Y)$. By the Künneth formula, it is enough to check this on elements of the form $\alpha \boxtimes \beta = \mathrm{pr}_1^* \alpha \cup \mathrm{pr}_2^* \beta$, where $\alpha \in \mathrm{H}^\bullet(X)$ and $\beta \in \mathrm{H}^\bullet(Y)$. Well, we note

$$(f \times g)^* \mathrm{pr}_1^* \alpha = f^* \alpha,$$

and similarly $(f \times g)^* \mathrm{pr}_2^* \beta = g^* \beta$. Combining completes the proof. ∎

We now move on to Poincaré duality.

Definition 1.94 (Poincaré duality)**.** Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F. Then $\mathrm{H}^\bullet$ satisfies *Poincaré duality* if and only if it satisfies the following for all $X \in \mathcal{P}(K)$ of equidimension d.

(a) Finite type: we have $\dim_F \mathrm{H}^i(X) < \infty$ for all $i \in \mathbb{Z}$.

(b) Poincaré duality: for each index i, the composite

$$\mathrm{H}^i(X) \times \mathrm{H}^{2d-i}(X)(d) \xrightarrow{\cup} \mathrm{H}^{2d}(X)(d) \xrightarrow{\int_X} F$$

is a perfect pairing of vector spaces over F.

Remark 1.95. Notably, our definition allows cohomology to be supported in negative degrees! We will remedy this later in Lemma 1.120 when we have a full definition of a Weil cohomology theory.

An important feature of Poincaré duality is that it lets us define the pushforward.

Notation 1.96. Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F satisfying Poincaré duality. If $f\colon X \to Y$ is a regular map of smooth projective varieties of equidimensions d and e respectively, we define the index-i pushforward

$$f_*\colon \mathrm{H}^{2d-i}(X)(d) \to \mathrm{H}^{2e-i}(Y)(e)$$

as the transpose of the pullback f^* under Poincaré duality.

Remark 1.97. Explicitly, given $\alpha \in \mathrm{H}^{2d-i}(X)(d)$, then $f_*\alpha \in \mathrm{H}^{2e-i}(X)(e)$ is defined as the unique element such that

$$\int_X (f^*\beta \cup \alpha) = \int_Y (\beta \cup f_*\alpha)$$

for all $\beta \in \mathrm{H}^i(Y)$. For example, if $\alpha \in \mathrm{H}^{2d}(X)(d)$, we may choose $\beta = 1$ to see that $\int_X \alpha = \int_Y f_*\alpha$.

Remark 1.98. The pushforward construction is functorial: given maps $f\colon X \to Y$ and $g\colon Y \to Z$, we check that $(g \circ f)_* = g_* \circ f_*$. Well, we already know that $(g \circ f)^* = f^* \circ g^*$ by functoriality of $\mathrm{H}^\bullet$, so this follows by taking the transpose along Poincaré duality.

Remark 1.99. If $\dim X = \dim Y$, then f_* preserves the grading. Further, we can undo the twisting to see that f_* becomes a graded linear map $f_*\colon \mathrm{H}^\bullet(X) \to \mathrm{H}^\bullet(Y)$.

We know that $f^*(\alpha \cup \beta) = f^*\alpha \cup f^*\beta$. We would like a similar way to compute f_* on products. This is not quite possible, but one can do something.

Lemma 1.100 (Projection formula)**.** Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F satisfying Poincaré duality. If $f\colon X \to Y$ is a regular map of smooth projective varieties of equidimensions d and e respectively, then

$$f_*(f^*\beta \cup \alpha) = \beta \cup f_*\alpha$$

for each $\alpha \in \mathrm{H}^{2d-i}(X)(d)$ and $\beta \in \mathrm{H}^j(Y)$.

Proof. We unravel the definition, following Remark 1.97. Indeed, for any $\beta' \in \mathrm{H}^{i-j}(X)$ has

$$\int_X f^*\beta' \cup (f^*\beta \cup \alpha) = \int_Y \beta' \cup (\beta \cup f_*\alpha)$$

by definition of $f_*\alpha$. ∎

Remark 1.101. This projection formula is expected on the level of cycles: for $\alpha \in \mathrm{CH}(X)$ and $\beta \in \mathrm{CH}(Y)$, one has $f_*(f^*\beta \cdot \alpha) = \beta \cdot f_*\alpha$ for any proper map $f\colon X \to Y$.

Lemma 1.102. Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F satisfying the Künneth formula and Poincaré duality. Given $X, Y \in \mathcal{P}(K)$ which are equidimensional of dimensions d and e respectively, then

$$\mathrm{pr}_{2*}(\alpha \boxtimes \beta) = \left(\int_X \alpha\right)\beta$$

for any $\alpha \in \mathrm{H}^{2d}(X)(d)$ and $\beta \in \mathrm{H}^\bullet(Y)(e)$.

Proof. It is enough to consider the case where β is homogeneous, so say $\beta \in \mathrm{H}^{2d-j}(Y)(e)$. Then we must check that

$$\int_{X\times Y} \mathrm{pr}_2^*\beta' \cup (\alpha \boxtimes \beta) \stackrel{?}{=} \int_Y \beta' \cup \left(\int_X \alpha\right)\beta$$

for any $\beta' \in \mathrm{H}^j(Y)$. Well, $\beta' \cup (\alpha \boxtimes \beta) = \alpha \boxtimes (\beta'\beta)$, so this follows from the Künneth formula. ∎

Our last collection of coherence assumptions on $\mathrm{H}^\bullet$ is for the cycle class maps.

Definition 1.103 (cycle coherence)**.** Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F satisfying Poincaré duality. Then $\mathrm{H}^\bullet$ satisfies *cycle coherence* if and only if it satisfies the following.

(a) Pullbacks: if $f\colon X \to Y$ is a regular map of smooth projective varieties, then $\mathrm{cl}_X(f^!\beta) = f^*\,\mathrm{cl}_X(\beta)$ for any $\beta \in \mathrm{CH}^\bullet(Y)$.

(b) Pushforwards: if $f\colon X \to Y$ is a regular map of smooth equidimensional projective varieties, then $\mathrm{cl}_Y(f_*\alpha) = f_*\,\mathrm{cl}_X(\alpha)$ for any $\alpha \in \mathrm{CH}^\bullet(X)$.

(c) Cup products: given $\alpha, \alpha' \in \mathrm{CH}^\bullet(X)$, we have $\mathrm{cl}_X(\alpha \cdot \alpha') = \mathrm{cl}_X(\alpha) \cup \mathrm{cl}_X(\alpha')$.

(d) Non-degeneracy: we have $\int_{\mathrm{Spec}\,K} \mathrm{cl}_{\mathrm{Spec}\,K}([\mathrm{Spec}\,K]) = 1$.

We now have enough axioms to start proving some results, so let's give a name for our current stopping point.

Definition 1.104 (pre-Weil cohomology theory)**.** Fix a Weil cohomology datum $\mathrm{H}^\bullet$ over K with coefficients in F satisfying Poincaré duality. Then $\mathrm{H}^\bullet$ is a *pre-Weil cohomology theory* if and only if $\mathrm{H}^\bullet$ satisfies the Künneth formula, Poincaré duality, and cycle coherence.

As we start to move into proving things, it is worth keeping track of the following idea.

Idea 1.105. To prove something about all Weil cohomology theories, one proves something "motivic" (i.e., "geometric") and then does linear algebra.

We will point out the various places we use motivic input; typically, one can see it as where we apply anything about cycle class maps. As an example, let's compute the cohomology of the point.

Example 1.106. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Then the cohomology ring $\mathrm{H}^\bullet(\operatorname{Spec} K)$ is supported in degree 0, and

$$\int_{\operatorname{Spec} K} : \mathrm{H}^0(\operatorname{Spec} K) \to F$$

is an isomorphism of algebras over F.

Proof. Our pieces of motivic input will be that $\operatorname{Spec} K \times \operatorname{Spec} K = \operatorname{Spec} K$ and that $[\operatorname{Spec} K] \cdot [\operatorname{Spec} K] = [\operatorname{Spec} K]$ in $\mathrm{CH}^0(\operatorname{Spec} K)$.

Note $\operatorname{Spec} K \times \operatorname{Spec} K \cong \operatorname{Spec} K$, so $\dim_F \mathrm{H}^\bullet(\operatorname{Spec} K \times \operatorname{Spec} K) = \dim_F \mathrm{H}^\bullet(\operatorname{Spec} K)$. Thus, the Künneth formula requires $\dim_F \mathrm{H}^\bullet(\operatorname{Spec} K) \in \{0, 1\}$. However, the non-degeneracy part of cycle coherence forces $\mathrm{H}^0(\operatorname{Spec} K) \neq 0$, so we conclude $\dim_F \mathrm{H}^\bullet(\operatorname{Spec} K) = 1$. Now, Poincaré duality tells us that $\dim_F \mathrm{H}^i(X) = \dim_F \mathrm{H}^{-i}(X)$ for all $i \in \mathbb{Z}$, so $\mathrm{H}^\bullet$ must be supported in degree 0.

It remains to show that $\int_{\operatorname{Spec} K} : \mathrm{H}^0(\operatorname{Spec} K) \to F$ is an isomorphism of algebras. This map is certainly an F-linear map of one-dimensional F-vector spaces, so it takes the form $a \mapsto a \int_{\operatorname{Spec} K} 1$ where $1 \in \mathrm{H}^0(\operatorname{Spec} K)$ is the unit. It thus suffices to check that $\int_{\operatorname{Spec} K} 1 = 1$. Well, cycle coherence requires $\int_{\operatorname{Spec} K} \mathrm{cl}_{\operatorname{Spec} K}([\operatorname{Spec} K]) = 1$, so we would like to show $\mathrm{cl}_{\operatorname{Spec} K}([\operatorname{Spec} K]) = 1$. For this, we note that

$$[\operatorname{Spec} K] \cdot [\operatorname{Spec} K] = [\operatorname{Spec} K],$$

so cycle coherence forces $\mathrm{cl}_{\operatorname{Spec} K}([\operatorname{Spec} K]) \in \{0, 1\}$, and zero it is not permitted by non-degeneracy. ∎

Corollary 1.107. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. If $X \in \mathcal{P}(K)$, then $\mathrm{cl}_X([X]) = 1$.

Proof. Let $p_X : X \to \operatorname{Spec} K$ be the structure map. Then we have some motivic input $[Y] = p_Y^*([\operatorname{Spec} K])$, so cycle coherence tells us that

$$\mathrm{cl}_Y([Y]) = p_Y^*(\mathrm{cl}_{\operatorname{Spec} K}([\operatorname{Spec} K])),$$

from which $\mathrm{cl}_Y([Y]) = 1$ follows by Example 1.106. ∎

We can also check that our cohomology is sufficiently nontrivial.

Proposition 1.108. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. If $X \in \mathcal{P}(K)$ is nonempty, then $\mathrm{H}^0(X) \neq 0$.

Proof. Throughout, for $Y \in \mathcal{P}(K)$, the structure morphism is denoted by $p_Y : Y \to \operatorname{Spec} K$. The proof has two steps.

1. We show that $\mathrm{H}^\bullet(X) \neq 0$ if X is nonempty and irreducible. It suffices to show that $\mathrm{H}^\bullet$ has some nonzero functional, for which we use points. Because X is smooth, it has a closed point $x \in X$ with residue field $\kappa(x)$ finite and separable over K; let $i : \{x\} \to X$ denote the inclusion. Then $(p_X \circ i) : \{x\} \to \operatorname{Spec} K$ is given by the inclusion $K \hookrightarrow \kappa(x)$, from which we can compute

   $$(p_X)_* i_* [x] = [\kappa(x) : K] \cdot [\operatorname{Spec} K].$$

   (At the level of intersection theory, one can see this by passing to the algebraic closure, whereupon x splits into $[\kappa(x) : K]$ distinct geometric points.) This provides our geometric input. Then cycle class coherence and Corollary 1.107 show that

   $$(p_X)_*(\mathrm{cl}_X(i_*[x])) = [\kappa(x) : K].$$

   Because F has characteristic 0, we see that the right-hand is nonzero, so $\mathrm{cl}_X(i_*[x]) \neq 0$, so $\mathrm{H}^\bullet(X) \neq 0$.

2. We reduce to the irreducible case. Suppose X is nonempty, and let $X' \subseteq X$ be an irreducible component. We would like to show that $1 \neq 0$ in $\mathrm{H}^\bullet(X)$. Well, there is a ring map $\mathrm{H}^\bullet(X) \to \mathrm{H}^\bullet(X')$ given by the inclusion, so it is actually enough to check that $1 \neq 0$ in $\mathrm{H}^\bullet(X')$. This has been done in the previous step. ■

Example 1.109. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Then $\mathrm{H}^\bullet(\varnothing) = 0$.

Proof. For any $X \in \mathcal{P}(K)$, our geometric input is that $\varnothing \times X = \varnothing$, from which the Künneth formula requires

$$\dim_F \mathrm{H}^\bullet(\varnothing) \cdot \dim_F \mathrm{H}^\bullet(X) = \dim_F \mathrm{H}^\bullet(\varnothing).$$

Now, we choose X to be nonempty of dimension at least 1 (for example, $X = \mathbb{P}^1_K$), then Proposition 1.108 shows $\mathrm{H}^0(X) \neq 0$, from which Poincaré duality yields $\dim_F \mathrm{H}^\bullet(X) \geq 2$. Plugging this in to the above equality gives $\mathrm{H}^0(X) \dim_F \mathrm{H}^\bullet(\varnothing) = 0$, from which the result follows. ■

In the sequel, we will also want more general control over unions.

Proposition 1.110. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Given $X, Y \in \mathcal{P}(K)$, let $i_1 \colon X \to X \sqcup Y$ and $i_2 \colon Y \to X \sqcup Y$ denote the canonical inclusions. Then the map

$$\begin{array}{ccc} \mathrm{H}^\bullet(X \sqcup Y) & \to & \mathrm{H}^\bullet(X) \times \mathrm{H}^\bullet(Y) \\ \gamma & \mapsto & (i_1^*\gamma \ , \ i_2^*\gamma) \end{array}$$

is an isomorphism.

Proof. If $X = \varnothing$ or $Y = \varnothing$, then the other inclusion is an isomorphism, and there is nothing to do. Let the given map be denoted i. Ultimately, the difficulty in this proof arises from the fact that there is no canonical inverse map, so we will have to apply various tricks to put ourselves in situations where we have approximations.

Quickly, we note that i is a product of algebra maps and hence an algebra map, so the main content comes from checking that this is a bijection. We will check injectivity and surjectivity, both in two steps. Let's start with injectivity.

1. We show that i is injective if X and Y are equidimensional with $\dim X = \dim Y$. This hypothesis will be used to allow us to think of pushforwards along i_1 and i_2 at the level of the full graded vector spaces, as in Remark 1.99. In particular, we will show that

   $$\gamma \stackrel{?}{=} i_{1*}i_1^*\gamma + i_{2*}i_2^*\gamma$$

   for any $\gamma \in \mathrm{H}^\bullet(X \sqcup Y)$; injectivity follows because this shows that $(\alpha, \beta) \sqcup i_{1*}\alpha + i_{2*}\beta$ is a one-sided inverse for i.

   By the projection formula (Lemma 1.100), it is enough to check that

   $$1 \stackrel{?}{=} i_{1*}1 + i_{2*}1,$$

   from which one can apply $\gamma \cup -$. Well, by Corollary 1.107, this is equivalent to asking for

   $$\mathrm{cl}_{X \sqcup Y}([X \sqcup Y]) = i_{1*}\,\mathrm{cl}_X([X]) + i_{2*}\,\mathrm{cl}_Y([Y]),$$

   We now see that this has motivic input given by the equation $[X \sqcup Y] = [X] + [Y]$, from which the result follows after using cycle coherence.

2. We show that i is injective in the general case. This will require a geometric trick. Given X and a positive integer $d > \dim X$, we will construct X' of dimension d for which there is an embedding $j_X \colon X \to X'$ and a projection $q_X \colon X' \to X$ such that $q_X \circ j_X = \mathrm{id}_X$. If we choose d to exceed $\max\{\dim X, \dim Y\}$ and apply the same construction to Y, then we can conclude as follows. The diagrams
$$\begin{array}{ccc} X \sqcup Y & \longleftarrow & X, Y \\ {\scriptstyle q_X \sqcup q_Y}\uparrow & & {\scriptstyle q_X}\uparrow \; \uparrow{\scriptstyle q_Y} \\ X' \sqcup Y' & \longleftarrow & X', Y' \end{array} \qquad \begin{array}{ccc} \mathrm{H}^\bullet(X \sqcup Y) & \longrightarrow & \mathrm{H}^\bullet(X) \times \mathrm{H}^\bullet(Y) \\ {\scriptstyle (q_X \sqcup q_Y)^*}\downarrow & & {\scriptstyle q_X^*}\downarrow \; \downarrow{\scriptstyle q_Y^*} \\ \mathrm{H}^\bullet(X' \sqcup Y') & \longrightarrow & \mathrm{H}^\bullet(X') \times \mathrm{H}^\bullet(Y') \end{array}$$
commute (the right diagram is induced from the left by functoriality), and the bottom row of the right diagram is injective by the previous step. Now, $q_\bullet \circ i_\bullet = \mathrm{id}_\bullet$, so $i_\bullet^* \circ q_\bullet^* = \mathrm{id}_\bullet^*$, meaning that the vertical $q_\bullet^*$s in the right diagram are all injective. Thus, the diagonal morphism of the right diagram is injective, so its top morphism is injective as well.

   It remains to construct X'. Decompose X into irreducible components $\{X_1, \dots, X_n\}$, and we note that the smoothness of X implies that its irreducible components are connected components as well. Thus, $X = X_1 \sqcup \cdots \sqcup X_n$, allowing us to define
$$X' \coloneqq \left(X_1 \times \mathbb{P}_K^{d-\dim X_1}\right) \sqcup \cdots \sqcup \left(X_n \times \mathbb{P}_K^{d-\dim X_n}\right).$$
   Choosing a point of the projective spaces gives an inclusion $X \hookrightarrow X'$, and there is an obvious projection $X' \twoheadrightarrow X$ by getting rid of the projective spaces.

We now turn to the surjectivity. It would be wonderful if the one-sided inverse in the first step also showed surjectivity (even in the case $\dim X = \dim Y$), but this only works once we know that the maps $\mathrm{H}^\bullet(X \sqcup Y) \to \mathrm{H}^\bullet(X)$ and $\mathrm{H}^\bullet(X \sqcup Y) \to \mathrm{H}^\bullet(Y)$ are surjective. We will have to expend some effort for this.

3. Suppose that there is a morphism $f \colon Y \to X$. Then we show that the map $i_1^* \colon \mathrm{H}^\bullet(X \sqcup Y) \to \mathrm{H}^\bullet(X)$ is surjective. Indeed, the inclusion $i_1 \colon X \subseteq X \sqcup Y$ admits a section $s \colon X \sqcup Y \to X$ by sending all of Y along f. Thus, $s \circ i_1 = \mathrm{id}_X$, meaning $i_1^* \circ s^* = \mathrm{id}_X^*$, so i_1^* is surjective.

4. We show that the map $i_1^* \colon \mathrm{H}^\bullet(X \sqcup Y) \to \mathrm{H}^\bullet(X)$ is always surjective. This requires a trick: all objects among F-vector spaces are faithfully falt, so we may check surjectivity after applying $- \otimes \mathrm{H}^\bullet(Z)$ for any Z. By the Künneth formula, we see that we are reduced to checking if
$$i_1^* \colon \mathrm{H}^\bullet((X \times Z) \sqcup (Y \times Z)) \to \mathrm{H}^\bullet(X \times Z)$$
   is surjective. In light of the previous step, we are tasked with finding Z such that there is a map $(Y \times Z) \to (X \times Z)$. Well, X is nonempty and smooth, so it has some closed point $x \in X$ with separable residue field $\kappa(x)$; then there is a map $Y_{\kappa(x)} \to X_{\kappa(x)}$ given by mapping all of Y to x.

5. We show that the map i is surjective. We are not going to use an assumption like $\dim X = \dim Y$; instead, we interface directly with $e_X \coloneqq \mathrm{cl}_{[X \sqcup Y]}([X])$ and $e_Y \coloneqq \mathrm{cl}_{[X \sqcup Y]}([Y])$.

   By the previous step, the map $i_1^* \mathrm{H}^\bullet(X \sqcup Y) \to \mathrm{H}^\bullet(X)$ is surjective, as is i_2^* by symmetry. Thus, it suffices to show that i surjects onto elements of the form $(i_1^*\gamma, i_2^*\delta)$. Well, we claim that
$$\begin{cases} i_1^*(e_X \cup \gamma + e_Y \cup \delta) \overset{?}{=} i_1^*\gamma, \\ i_2^*(e_X \cup \gamma + e_Y \cup \delta) \overset{?}{=} i_2^*\delta. \end{cases}$$
   Indeed, because i_1^* and i_2^* are ring homomorphisms, it is enough to note that $i_1^* e_X = e_X$ and $i_1^* e_Y = 0$ by cycle coherence for the first equality, and $i_2^* e_X = 0$ and $i_2^* e_Y = e_Y$ by cycle coherence for the second equality. ∎

Remark 1.111. If X and Y are equidimensional with $\dim X = \dim Y$, then the first step shows that there is a canonical inverse given by
$$(\alpha, \beta) \mapsto i_{1*}\alpha + i_{2*}\beta.$$
Importantly, these pushforwards really only make sense in the equidimensional case!

Corollary 1.112. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Suppose $X, Y \in \mathcal{P}(K)$ are equidimensional of dimension d. For any $\alpha \in \mathrm{H}^{2d}(X \sqcup Y)(d)$, we have

$$\int_{X \sqcup Y} \alpha = \int_X i_1^* \alpha + \int_Y i_2^* \alpha.$$

Proof. By Remark 1.111, we see that $\alpha = i_{1*} i_1^* \alpha + i_{2*} i_2^* \alpha$. Thus, for example, we compute $\int_{X \sqcup Y} i_{1*} i_1^* \alpha$ is

$$\int_{X \sqcup Y} (1 \cup i_{1*} i_1^* \alpha) = \int_X (1 \cup i_1^* \alpha),$$

which is $\int_X i_1^* \alpha$. Adding together a similar computation for $i_2^* \alpha$ completes the argument. ∎

As an application, we can now fairly easily compute the cohomology of multiple points.

Example 1.113. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Suppose $X \in \mathcal{P}(K)$ is zero-dimensional. Then $\mathrm{H}^\bullet(X)$ is supported in degree 0, and $\mathrm{H}^0(X)$ is a separable algebra over F of dimension equal to the degree of $X \to \operatorname{Spec} K$. Further, $\int_X \colon \mathrm{H}^0(X) \to F$ is the trace.

Proof. For psychological reasons, we quickly reduce to the case where X is a closed point. By decomposing X into irreducible components (which are connected components by smoothness) and using Proposition 1.110, it suffices to show the various claims in the case that X is irreducible (indeed, the conclusion is closed under taking disjoint unions). Thus, we may assume that X is irreducible.

Because X is zero-dimensional, the structure morphism $X \to \operatorname{Spec} K$ is finite, so X is affine; we write $X = \operatorname{Spec} L$. Because X is smooth and hence étale, we see that L must be a finite-dimensional separable algebra over K. In fact, L must be a field extension of K because X is irreducible. Let M be a Galois closure of the separable extension L/K. Roughly speaking, the idea of the proof is to run all of our checks after extending up to M. We proceed in steps.

1. We explain how to base-change to M. Well, there is an isomorphism

$$\begin{array}{ccc} L \otimes M & \to & \displaystyle\prod_{\sigma \in \mathrm{Hom}_K(L,M)} M \\ a \otimes b & \mapsto & (\sigma(a) b)_\sigma \end{array}$$

because L/K is separable. This translates into the motivic input $X \times \operatorname{Spec} M = \bigsqcup_{\sigma \in \mathrm{Hom}_K(L,M)} \operatorname{Spec} M$, which induces an isomorphism

$$\begin{array}{ccc} \mathrm{H}^\bullet(X) \otimes \mathrm{H}^\bullet(\operatorname{Spec} M) & \to & \mathrm{H}^\bullet(\operatorname{Spec} M)^{\mathrm{Hom}_K(L,M)} \\ \alpha \otimes \beta & \mapsto & (\sigma^* \alpha \cup \beta)_\sigma \end{array}$$

by the Künneth formula and Proposition 1.110.

2. We check that $\mathrm{H}^\bullet(X)$ is concentrated in degree 0, and $\mathrm{H}^0(X)$ is an algebra over F of dimension equal to the degree of the structure morphism $X \to \operatorname{Spec} K$. (Note that this degree is $[L : K]$.) Well, taking dimensions on both sides of the last map in step 1 (and noting $\dim_F H^\bullet(\operatorname{Spec} M) \geq \dim_F \mathrm{H}^0(\operatorname{Spec} F) > 0$ by Proposition 1.108), we find that

$$\dim_F \mathrm{H}^\bullet(X) = \dim_F \mathrm{H}^0(X) = [L : K].$$

The needed claims follow.

3. We check that $\mathrm{H}^0(X)$ is separable over F. Well, $\mathrm{H}^0(Y)$ is faithfully flat over F because it is a finite-dimensional separable algebra over F by what we already know. Further, separability can be checked after a faithfully flat extension, so checking the separability of $\mathrm{H}^0(X)$ over F can be seen by checking the separabiility of

$$\mathrm{H}^0(X) \otimes \mathrm{H}^0(Y) = \mathrm{H}^0(Y)^{\mathrm{Hom}_K(L,M)}$$

over $\mathrm{H}^0(Y)$, which is now clear.

4. We show that $\int_X : \mathrm{H}^0(X) \to F$ is the trace. The main point is to compare the traces on $X \times \operatorname{Spec} M$ and $\bigsqcup_{\sigma \operatorname{Hom}_K(L,M)} \operatorname{Spec} M$. Fix some $\alpha \in \mathrm{H}^0(X)$, and we would like to compute $\int_X \alpha$. On one hand, Lemma 1.102 gives $\int_X \alpha = \mathrm{pr}_{2*}(\alpha \boxtimes 1)$, but alternatively one can see via our explicit isomorphism that

$$\mathrm{pr}_{2*}(\alpha \boxtimes 1) = \sum_{\sigma \in \operatorname{Hom}_K(L,M)} \sigma^* \alpha.$$

Indeed, for any $\beta \in \mathrm{H}^\bullet(\operatorname{Spec} M)$, we see $\sum_\sigma \int_{\operatorname{Spec} M} (\beta \cup \sigma^* \alpha) = \int_{X \times \operatorname{Spec} M} \mathrm{pr}_2^* \beta \cup (\alpha \boxtimes 1)$, where we have used Corollary 1.112. It remains to check that $\alpha \mapsto \sigma^* \alpha$ amounts to the full set of homomorphisms $\mathrm{H}^0(X) \to \overline{F}$. Well, upon choosing some map $\iota : \mathrm{H}^0(\operatorname{Spec} M) \to \overline{F}$, we see that there is an isomorphism

$$\begin{aligned} \mathrm{H}^0(X) \otimes \overline{F} &\to F^{\operatorname{Hom}_K(L,M)} \\ \alpha \otimes \beta &\mapsto (\tau(\sigma^* \alpha) \cup \beta)_\sigma \end{aligned}$$

which completes the proof because $\mathrm{H}^0(X) \otimes \overline{F}$ is supposed to be isomorphic to $\overline{F}^{\operatorname{Hom}(\mathrm{H}^0(X), \overline{F})}$ via this sort of map. ∎

Corollary 1.114. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Given $X \in \mathcal{P}(K)$ and some zero-dimensional cycle $Z \subseteq X$, we have

$$\deg [Z] = \int_X \mathrm{cl}_X([Z]).$$

Proof. We may adjust Z so that it is smooth divisor. Letting $i : Z \to X$ denote the inclusion, we get the motivic input that $[Z] = i_*[Z]$, so $\mathrm{cl}_X([Z]) = i_* 1$ by Corollary 1.107 and cycle coherence. It follows that

$$\int_X \mathrm{cl}_X([Z]) = \int_Z 1$$

by Remark 1.97. We now use Example 1.113 to compute the right-hand side: because $\int_X : \mathrm{H}^0(Z) \to F$ is the trace, its evaluation on 1 is the dimension $\dim_F \mathrm{H}^0(Z)$, which we know to be the degree of $Z \to \operatorname{Spec} K$. This completes the proof. ∎

Now that we've done work with our pre-Weil cohomology theories, let's introduce our last axiom.

Definition 1.115 (Weil cohomology theory)**.** Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Then $\mathrm{H}^\bullet$ is a *Weil cohomology theory* if and only if the induced map

$$\mathrm{H}^0(\operatorname{Spec} \Gamma(X, \mathcal{O}_X)) \to \mathrm{H}^0(X)$$

is an isomorphism for all $X \in \mathcal{P}(K)$.

Remark 1.116. Let's explain where this map comes from. There is a natural map $X \to \operatorname{Spec} \Gamma(X, \mathcal{O}_X)$; for example, this exists already on the level of locally ringed spaces, though one could alternatively define it by gluing together maps on affine open subschemes. However, we must check $\operatorname{Spec} \Gamma(X, \mathcal{O}_X) \in \mathcal{P}(K)$: certainly $\Gamma(X, \mathcal{O}_X)$ is some finite-dimensional K-algebra, so the issue is separability. For this, we base-change to $\overline{K}$, noting

$$\Gamma(X, \mathcal{O}_X)_{\overline{K}} = \Gamma(X_{\overline{K}}, \mathcal{O}_{X_{\overline{K}}})$$

because cohomology is stable under base change. The right-hand side is a product of fields because $X_{\overline{K}}$ is still a proper variety, so it follows that $\Gamma(X, \mathcal{O}_X)$ is separable and hence smooth over K.

It is certainly desirable to have $\mathrm{H}^0(\operatorname{Spec} \Gamma(X, \mathcal{O}_X)) \to \mathrm{H}^0(X)$ be an isomorphism. Let's explain some of its applications.

Lemma 1.117. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. For any $X \in \mathcal{P}(K)$ of equidimension d, the space $\mathrm{H}^{2d}(X)(d)$ is generated by classes of points as an $\mathrm{H}^0(X)$-module.

Proof. If $X = \varnothing$, there is nothing to do, so we assume that X is nonempty. By Proposition 1.110, we may assume that X is irreducible. Define $L \coloneqq \Gamma(X, \mathcal{O}_X)$ for brevity; because X is irreducible, L is a field, and we know that it is finite separable over K.

Now, for each closed point $x \in X$ (which we assume to have residue field $\kappa(x)$ to be separable over L), let $i\colon \{x\} \to X$, and we would like to check that the class $\mathrm{cl}_X([x]) \in \mathrm{H}^{2d}(X)(d)$ generates as a module over $\mathrm{H}^0(X) = \mathrm{H}^0(\operatorname{Spec} L)$. Quickly, note that $\mathrm{cl}_X([x]) = i_*1$ by Corollary 1.107 and cycle coherence. As such, we want to show that the map $\mathrm{H}^0(X) \to \mathrm{H}^{2d}(X)(d)$ given by $\alpha \mapsto (\alpha \cup i_*1)$ is surjective. Now, Lemma 1.100 explains $\alpha \cup i_*1 = i_*i^*\alpha$, so we might as well show that the map $i_*\colon \mathrm{H}^0(\{x\}) \to \mathrm{H}^{2d}(X)(d)$ is surjective.

Continuing, it is enough to check that the transpose $i^*\colon \mathrm{H}^0(X) \to \mathrm{H}^0(\{x\})$ is injective. Now, let $p\colon X \to \operatorname{Spec} L$ be the canonical projection, and then $p^*\colon \mathrm{H}^0(\operatorname{Spec} L) \to \mathrm{H}^0(X)$ is an isomorphism! Thus, it is enough to show that $i^*p^*\colon \mathrm{H}^0(\operatorname{Spec} L) \to \mathrm{H}^0(\{x\})$ is injective. There are a few ways to conclude, but here is one using Example 1.113: it is enough to check injectivity after faithfully flat base change, so we may check injectivity after tensoring with the separable K-algebra $\mathrm{H}^0(\operatorname{Spec} M)$, where M is some Galois closure of $L\kappa(x)/K$. Then both $\mathrm{H}^0(\operatorname{Spec} L)$ and $\mathrm{H}^0(\{x\})$ split up into products of $\mathrm{H}^0(\operatorname{Spec} M)$, from which the injectivity follows. ∎

Remark 1.118. It turns out that the conclusion of the lemma also implies that $\mathrm{H}^0(\operatorname{Spec}\Gamma(X, \mathcal{O}_X)) \to \mathrm{H}^0(X)$ is an isomorphism, but we will not need this. We refer to [SP, Tag `0FI0`].

Lemma 1.119. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. If $f\colon X \to Y$ is a finite map of equidimensional varieties of dimension d with Y geometrically irreducible, then $f_*f^* = (\deg f)$.

Proof. We begin with a couple reductions.

- It is enough to check that $f_*f^* = (\deg f)$ on homogeneous elements of $\mathrm{H}^\bullet(Y)$, and in fact, it is enough to merely check equality of traces on elements in $\mathrm{H}^{2d-i}(Y)(d)$. Indeed, to check that $f_*f^*\beta = (\deg f)\beta$ for any $\beta \in \mathrm{H}^{2d-i}(Y)(d)$, Remark 1.97 explains that it is enough to check
$$\int_X f^*\beta' \cup f^*\beta \overset{?}{=} \int_Y \beta' \cup (\deg f)\beta$$
for all $\beta' \in \mathrm{H}^i(Y)$. This now follows by applying $\int_Y \circ (f_*f^*) = (\deg f)\int_Y$ to $\beta' \cup \beta \in \mathrm{H}^{2d}(Y)(d)$; in particular, recall $\int_Y \circ f_* = \int_X$ by Remark 1.97.
- We show that it is enough to check the equality $\int_X \circ f^* = (\deg f)\int_Y$ on the image of $\mathrm{cl}_X\colon \mathrm{CH}^d(Y) \to \mathrm{H}^{2d}(Y)(d)$. Because Y is geometrically irreducible, we see that $\Gamma(Y, \mathcal{O}_Y) = K$ (this can be checked after passing to the algebraic closure), so $\mathrm{H}^{2d}(Y)(d)$ is isomorphic to $\mathrm{H}^0(Y)$ (by Poincaré duality), which is isomorphic to $\mathrm{H}^0(\operatorname{Spec} K)$ (because this is a Weil cohomology theory), which is simply F (by Example 1.106). It is thus enough to check the result at a single vector in $\mathrm{H}^{2d}(Y)(d)$, such as the class of a point (which is nonzero by Lemma 1.117).

As such, our "motivic" input will come from checking $\int_X \circ f^* = (\deg f)\int_Y$ on classes of points: because f is finite, any $q \in Y$ has
$$f^*[q] = \sum_{p \in f^{-1}(\{q\})} m_p \cdot [p],$$
where m_p is a multiplicity satisfying $\sum_p m_p[\kappa(p) : K] = \deg f$. Then passing this through cl_X (and using cycle coherence), followed by applying $\int_X$ (and Corollary 1.114) completes this check. ∎

Lemma 1.120. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. For any $X \in \mathcal{P}(K)$ of dimension d, the graded algebra $\mathrm{H}^\bullet(X)$ is supported in degrees $[0, 2d]$.

Proof. By Proposition 1.110, it is enough to check this in the case that X is irreducible. Then X has equidimension d, so Poincaré duality implies that it is enough to show that $\mathrm{H}^\bullet(X)$ is supported in nonnegative degrees.

We will show that $\mathrm{H}^\bullet(X)$ is supported in nonnegative degrees by an awkward contraposition: we will show that any pre-Weil cohomology theory $\mathrm{H}^\bullet$ admitting some $Y \in \mathcal{P}(Y)$ with $\mathrm{H}^\bullet(Y)$ supported at a negative index must fail to be a Weil cohomology theory. By replacing Y with $Y \times Y$ and using the Künneth formula, we may assume that $\mathrm{H}^{-2n}(Y) \neq 0$ for some $n > 0$. We now set $X \coloneqq Y \times \mathbb{P}^n_K$, so the Künneth formula gives

$$\mathrm{H}^0(X) = \bigoplus_{i \in \mathbb{Z}} \mathrm{H}^i(Y) \otimes \mathrm{H}^{-i}(\mathbb{P}^n_K)$$

For example, $\mathrm{H}^0(X)$ contains the summands $\mathrm{H}^0(Y) \subseteq \mathrm{H}^0(X)$ and $\mathrm{H}^{-2n}(Y) \otimes \mathrm{H}^{2n}(\mathbb{P}^n_K)$, so

$$\dim_F \mathrm{H}^0(X) > \dim_F \mathrm{H}^0(Y).$$

(Note $\mathrm{H}^{2n}(\mathbb{P}^n_K)$ is nonzero by Proposition 1.108 and Poincaré duality.) However, $\Gamma(X, \mathcal{O}_X) = \Gamma(Y, \mathcal{O}_Y)$: a global section is a map to $\mathbb{A}^1$, and the only maps $\mathbb{P}^n_K \to \mathbb{A}^1$ are constants anyway. Thus, it is impossible to have both $\mathrm{H}^0(X) \cong \mathrm{H}^0(\Gamma(X, \mathcal{O}_X))$ and $\mathrm{H}^0(Y) \cong \mathrm{H}^0(\Gamma(Y, \mathcal{O}_Y))$! ∎

We have now cobbled together enough of a theory of Weil cohomology. Let's work towards an application: the Lefschetz trace formula. After everything we've done, this proof is purely formal. Our exposition follows [Mil13, Section 25].

Given a regular map $f \colon X \to X$, the Lefschetz trace formula computes the intersection number $\Gamma_f \cdot \Delta$ in terms of cohomology. Thus, our proof will begin by understanding the graph Γ_f.

Lemma 1.121. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. For any regular map $f \colon X \to Y$ of equidimensional projective varieties and $\beta \in \mathrm{H}^\bullet(Y)$, we have

$$\mathrm{pr}_{1*}\left(\mathrm{cl}_{X \times Y}([\Gamma_f]) \cup \mathrm{pr}_2^* \beta\right) = f^* \beta.$$

Proof. Our motivic input is that $[\Gamma_f] = (\mathrm{id}_X, f)_*([X])$, by definition. Then cycle coherence and Corollary 1.107 shows $\mathrm{cl}_{X \times Y}([\Gamma_f]) = (\mathrm{id}_X, f)_* 1$. Thus, the projection formula (Lemma 1.100) implies

$$\mathrm{pr}_{1*}(\mathrm{cl}_{X \times Y}([\Gamma_f]) \cup \mathrm{pr}_2^* \beta) = \mathrm{pr}_{1*}(\mathrm{id}_X, f)_*(\mathrm{id}_X, f)^* \mathrm{pr}_2^* \beta.$$

Functoriality reveals this is $f^* \beta$. ∎

Lemma 1.122. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. For equidimensional $X \in \mathcal{P}(K)$ with $d \coloneqq \dim X$, let $\{e_{ij}\}_{1 \le j \le \beta_i}$ be a basis of $\mathrm{H}^i(X)$ for each i; further, choose a dual basis $\{e^\vee_{2d-i,j}\}_{1 \le j \le \beta_i}$ of $\mathrm{H}^{2d-i}(X)(d)$ so that $\int_X (e^\vee_{2d-i,j} \cup e_{ij'}) = 1_{j=j'}$ for each j and j'. Then any regular map $f \colon X \to X$ admits a decomposition

$$\mathrm{cl}_{X \times X}([\Gamma_f]) = \sum_{\substack{i \in \mathbb{Z} \\ 1 \le j \le \beta_i}} f^* e_{ij} \boxtimes e^\vee_{2d-i,j}.$$

Proof. Note that the $e^\vee_{2d-i,j}$s exist by Poincaré duality. Now, the Künneth formula tells us that $\mathrm{H}^d(X \times X)(d) = \bigoplus_{i \in \mathbb{Z}} \mathrm{H}^i(X) \otimes \mathrm{H}^{2d-i}(X)(d)$, so $\mathrm{cl}_{X \times X}([\Gamma_f])$ admits some decomposition

$$\mathrm{cl}_{X \times X}([\Gamma_f]) = \sum_{\substack{i \in \mathbb{Z} \\ 1 \le j \le \beta_i}} \alpha_{ij} \boxtimes e^\vee_{2d-i,j},$$

where $\alpha_{ij} \in \mathrm{H}^i(X)$ is some class. We would like to show $\alpha_{ij} = f^* e_{ij}$. To extract out the needed coefficients, we need to cup with a basis vector and apply the pairing. As such, we compute

$$\mathrm{pr}_{1*}\left(\mathrm{cl}_{X\times X}([\Gamma_f]) \cup \mathrm{pr}_2^* e_{ij}\right) = \sum_{\substack{i\in\mathbb{Z}\\ 1\le j\le \beta_i}} \mathrm{pr}_{1*}\left(\alpha_{ij} \boxtimes (e^\vee_{2d-i,j} \cup e_{ij})\right),$$

which collapses down to α_{ij} by Lemma 1.102 and construction of the $e^\vee_{2d-i,j}$s. We now complete the proof by recognizing the left-hand side as $f^* e_{ij}$ by Lemma 1.121. ■

Example 1.123. Taking $f = \mathrm{id}_X$ shows that the diagonal $\Delta \subseteq X \times X$ has a decomposition

$$\mathrm{cl}_{X\times X}([\Delta]) = \sum_{\substack{i\in\mathbb{Z}\\ 1\le j\le \beta_i}} e_{ij} \boxtimes e^\vee_{2d-i,j}.$$

Remark 1.124. It may appear that Lemma 1.122 needs some finiteness condition like Lemma 1.120, but our proof actually shows that all but finitely many of the $f^* e_{ij}$ are allowed to vanish.

We are now ready for the proof.

Theorem 1.125 (Lefschetz trace formula)**.** Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. For equidimensional $X \in \mathcal{P}(K)$ and endomorphism $f\colon X \to X$, we have

$$\deg([\Gamma_f] \cdot [\Delta]) = \sum_{i=0}^{2d} (-1)^i \operatorname{tr}\left(f^*; \mathrm{H}^i(X)\right).$$

Proof. This proof is essentially a direct computation. By Corollary 1.114, we see that

$$\deg([\Gamma_f] \cdot [\Delta]) = \int_{X\times X} \mathrm{cl}_{X\times X}([\Gamma_f]) \cup \mathrm{cl}_{X\times X}([\Delta]),$$

where we have quietly also used cycle coherence. We now fix a basis $\{e_{ij}\}_{ij}$ of $\mathrm{H}^\bullet(X)$ and a dual basis $\{e_{2d-i,j}\}_{ij}$ of $\mathrm{H}^{2d-\bullet}(X)$ as in Lemma 1.122. Then Lemma 1.122 (and a reversed Example 1.123) allows us to compute this as

$$\deg([\Gamma_f] \cdot [\Delta]) = \sum_{\substack{i,i'\in\mathbb{Z}\\ 1\le j,j'\le \beta_i}} \int_{X\times X} \left(f^* e_{ij} \boxtimes e^\vee_{2d-i,j}\right) \cup \left((-1)^{i'} e^\vee_{2d-i',j'} \boxtimes e_{i'j'}\right).$$

By expanding out $\alpha \boxtimes \beta = \mathrm{pr}_1^* \alpha \cup \mathrm{pr}_2^* \beta$ and rearranging, we may rewrite the right-hand side as

$$\deg([\Gamma_f] \cdot [\Delta]) = \sum_{\substack{i,i'\in\mathbb{Z}\\ 1\le j,j'\le \beta_i}} (-1)^{i+ii'} \int_{X\times X} (f^* e_{ij} \cup e^\vee_{2d-i',j'}) \boxtimes (e^\vee_{2d-i,j} \boxtimes e_{i',j'}),$$

which by the Künneth formula is

$$\deg([\Gamma_f] \cdot [\Delta]) = \sum_{\substack{i,i'\in\mathbb{Z}\\ 1\le j,j'\le \beta_i}} (-1)^{i+ii'} \int_X (f^* e_{ij} \cup e^\vee_{2d-i',j'}) \int_X (e^\vee_{2d-i,j} \cup e_{i',j'}).$$

Now, the right-hand integral is $1_{i=i'} 1_{j=j'}$ by construction of our dual basis, so we are left with

$$\deg([\Gamma_f] \cdot [\Delta]) = \sum_{\substack{i\in\mathbb{Z}\\ 1\le j\le \beta_i}} \int_X (f^* e_{ij} \cup e^\vee_{2d-i,j}).$$

Because technically $\{e_{ij}\}_j$ and $\{(-1)^i e^\vee_{2d-i,j}\}_j$ are the dual bases with $\int_X (e_{ij} \cup (-1)^i e^\vee_{2d-i,j'}) = 1_{j=j'}$, we see that the right-hand integral collapses down to $(-1)^i \operatorname{tr}(f^*; \mathrm{H}^i(X))$. This completes the proof upon using Lemma 1.120 to restrict the sum to $i \in [0, 2d]$. ■

Remark 1.126. Technically, this argument works for pre-Weil cohomology theories, provided we sum over all $i \in \mathbb{Z}$ instead of $i \in [0, 2d]$.

Let's apply some of the theory we built to do one last calculation.

Example 1.127. Fix a pre-Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Then

$$\mathrm{H}^i(\mathbb{P}^1_K) = \begin{cases} F & \text{if } i = 0, \\ F(-1) & \text{if } i = 2, \\ 0 & \text{else.} \end{cases}$$

Proof. The main claim is that $\dim_F \mathrm{H}^\bullet(\mathbb{P}^1_K) = 2$. Quickly, let's explain why the main claim completes the proof. Certainly $\mathrm{H}^0(\mathbb{P}^1_K) \neq 0$ by Proposition 1.108, so $\mathrm{H}^2(\mathbb{P}^1_K)(1) \neq 0$ by Poincaré duality as well, which provides the lower bound $\dim_F \mathrm{H}^\bullet(\mathbb{P}^1_K) \geq 2$. If we were to have equality, then we must have $\mathrm{H}^\bullet(\mathbb{P}^1_K) = \mathrm{H}^0(\mathbb{P}^1_K) \oplus \mathrm{H}^2(\mathbb{P}^1_K)$, and $\mathrm{H}^0(\mathbb{P}^1_K) = F$ and $\mathrm{H}^2(\mathbb{P}^1_K)(1) = F$ become forced.

We now prove the main claim. It remains to show $\dim_F \mathrm{H}^\bullet(\mathbb{P}^1_K) \leq 2$. Technically, Theorem 1.125 will not be enough for our purposes because the Euler characteristic includes a $-\dim_F \mathrm{H}^1(X)$ term. Our motivic input is that the cycle class $[\Delta]$ in $\mathbb{P}^1_K \times \mathbb{P}^1_K$ is equal to $\mathrm{pr}_1^*[\infty] + \mathrm{pr}_2^*[\infty]$, where $\infty \in \mathbb{P}^1_K$ is a point at infinity. Indeed, consider the function $f \colon \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^1$ given by $f(x, y) \coloneqq x - y$. Then f has zero-set given by Δ and poles given by $\{\infty\} \times \mathbb{P}^1_K$ and $\mathbb{P}^1 \times \{\infty\}$, so

$$\operatorname{div} f = \mathrm{pr}_1^*[\infty] + \mathrm{pr}_2^*[\infty] - \Delta$$

must be a trivial divisor class. We conclude that

$$\mathrm{cl}_{\mathbb{P}^1_K \times \mathbb{P}^1_K}([\Delta]) = \mathrm{cl}_{\mathbb{P}^1_K}([\infty]) \boxtimes 1 + 1 \boxtimes \mathrm{cl}_{\mathbb{P}^1_K}([\infty]).$$

Now, Example 1.123 shows that the left-hand side has no expression in terms of fewer than $\dim_F \mathrm{H}^\bullet(X)$ total pure tensors, so we concldue that $\dim_F \mathrm{H}^\bullet(X) \leq 2$! ■

1.3.3 Tannakian Formalism

It was frequently apparent from our discussion of Weil cohomology theories that proofs frequently have some geometric component, from which some algebraic calculations derived an interesting result. As such, we are motivated to look for a conjectural category where we can run such geometric calculations. Of course, it would be lovely to work directly with $\mathcal{P}(K)$ (or $\mathcal{P}(K)^{\mathrm{op}}$) directly, but this is a pretty bad category; for example, it is very far from abelian.

Instead, we will attempt to "close up" the category $\mathcal{P}(K)$ in various ways to produce a well-behaved category. In this subsection, we will make rigorous what we mean by "well-behaved": we are hoping for (neutral) Tannakian categories. Our exposition follows [DM12] and [And04, Chapters 2 and 6].

Warning 1.128. We will not need any proofs from the theory of Tannakian formalism, so we will not provide them.

Intuitively, a Tannakian category is one that looks like the category $\mathrm{Rep}_F(G)$ of finite-dimensional representations of an affine F-group G. An important property of $\mathrm{Rep}_F(G)$ is the ability to take tensor products, so we codify how useful tensor products are.

Definition 1.129 (monoidal)**.** A *monoidal category* or $\otimes$*-category* is a category $\mathcal{C}$ equipped with a bifunctor $\otimes\colon \mathcal{C}\times\mathcal{C}\to\mathcal{C}$ and identity object $1\in\mathcal{C}$ with the following identities.

- Associativity: there is a natural isomorphism $\alpha\colon ((-\otimes-)\otimes-)\Rightarrow(-\otimes(-\otimes-))$.
- Identity: there are natural isomorphisms $(1\otimes-)\Rightarrow-$ and $(-\otimes 1)\Rightarrow 1$.

These isomorphisms satisfy certain coherence properties ensuring that one can associate and apply identity naturally in any suitable situation.

In fact, $\mathrm{Rep}_F(G)$ has a symmetry property.

Definition 1.130 (symmetric monoidal)**.** A *symmetric monoidal category* is a monoidal category $\mathcal{C}$ further equipped with a symmetry isomorphism $(-\otimes-)\Rightarrow(-\otimes-)$ such that the composite

$$(A\otimes B)\to(B\otimes A)\to(A\otimes B)$$

is the identity.

The reason we restricted $\mathrm{Rep}_F(G)$ to finite-dimensional representations is so that we can take duals.

Definition 1.131 (rigid)**.** A *rigid* symmetric monoidal category is a symmetric monoidal category $\mathcal{C}$ further equipped with a natural isomorphism $(-)^\vee\colon\mathcal{C}\to\mathcal{C}^{\mathrm{op}}$ such that each $A\in\mathcal{C}$ makes $(-\otimes A^\vee)$ is left adjoint to $(-\otimes A)$, and $(A^\vee\otimes-)$ is right adjoint to $(A\otimes-)$.

Remark 1.132. Rigidity allows one to define an internal hom by $\mathrm{Hom}(X,Y)\coloneqq X^\vee\otimes Y$. For example, one may define the trace tr_X as the composite

$$\mathrm{End}(X)=X^\vee\otimes X\to 1,$$

where the second map is canonically given by the adjunction. With a trace, one can also define a rank by $\mathrm{rank}\,X\coloneqq\mathrm{tr}_X(\mathrm{id}_X)$.

Lastly, $\mathrm{Rep}_F(G)$ has a forgetful functor to Vec_F, akin to the forgetful functor $\mathrm{Set}(G)\to\mathrm{Set}$ which appears in Grothendieck's Galois theory (used to define the étale fundamental group).

Definition 1.133 (fiber functor)**.** Fix an abelian rigid symmetric monoidal category $\mathcal{C}$ such that $F'\coloneqq\mathrm{End}(1)$ is a field. A *fiber functor* is a faithful exact $\otimes$-functor $\omega\colon\mathcal{C}\to\mathrm{Vec}_F$ for some finite field extension F of F. If $F=F'$, then we say that $\mathcal{C}$ is *neutral Tannakian over* F.

What is remarkable is that it turns out that one can recover the affine F-group G from the (forgetful) fiber functor $\omega\colon\mathrm{Rep}_F(G)\to\mathrm{Vec}_F$ as "$\underline{\mathrm{Aut}}^\otimes(\omega)$." Explicitly, for an F-algebra R, an element of $\underline{\mathrm{Aut}}^\otimes(\omega)(R)$ is a collection of automorphisms $(g_X)_{X\in\mathrm{Rep}_F(G)}$ where g_X is an R-linear automorphism of $\omega(X)\otimes_F R$, and these automorphisms are natural in G-linear maps $X\to Y$.

This process can in general recover a group G from a neutral Tannakian category.

Theorem 1.134. Fix a neutral Tannakian category $\mathcal{C}$ over a field F equipped with fiber functor $\omega\colon\mathcal{C}\to\mathrm{Vec}_F$.

(a) The functor $\underline{\mathrm{Aut}}^\otimes(\omega)$ (defined analogously as above) is represented by an affine F-group G.

(b) The fiber functor ω then upgrades to a $\otimes$-equivalence $\mathcal{C}\to\mathrm{Rep}_F(G)$.

Proof. See [DM12, Theorem 2.11]. ∎

In fact, a careful review of the proof reveals that one can do away with many hypotheses on $\mathcal{C}$.

Theorem 1.135. Suppose that $\mathcal{C}$ is an essentially small F-linear category equipped with an F-linear symmetric monoidal functor $\otimes\colon \mathcal{C}\times\mathcal{C}\to\mathcal{C}$. Further, suppose that there is an exact faithful functor $\omega\colon \mathcal{C}\to \mathrm{Vec}_F$ satisfying the following.

(i) $\omega(X\otimes Y)=\omega(X)\otimes\omega(Y)$ for all $X,Y\in\mathcal{C}$.

(ii) The functor ω preserves the commutativity and associativity coherences.

(iii) The functor ω sends the unit 1 to $F\in\mathrm{Vec}_F$, and ω preserves the unit coherences.

(iv) Each $X\in\mathcal{C}$ such that $\dim_F\omega(X)=1$ has some object $Y\in\mathcal{C}$ such that $X\otimes Y\cong 1$.

Then $\mathcal{C}$ is neutral Tannakian, and ω is a fiber functor.

Proof. See [Mil17, Theorem 9.24]. ∎

Let's see some examples.

Example 1.136. Of course, $\mathrm{Rep}_F(G)$ is a neutral Tannakian category for any affine group G over F, where the fiber functor is given by the forgetful functor $\omega\colon \mathrm{Rep}_F(G)\to\mathrm{Vec}_F$.

Example 1.137. For any profinite group G and field F, the category $\mathrm{Rep}_F G$ of continuous representations of G succeeds at being neutral Tannakian. The fiber functor is still the forgetful functor.

Example 1.138. The category GrVec_F of $\mathbb{Z}$-graded vector spaces is a neutral Tannakian category, where the fiber functor is the forgetful functor. In fact, by diagonalizing, we can see that a graded vector space has exactly the same data as a representation of $\mathbb{G}_{m,F}$, where the graded piece in degree $d\in\mathbb{Z}$ corresponds to the eigenvector with eigenvalue $T\mapsto T^d$.

Example 1.139. The category $\mathrm{HS}_{\mathbb{R}}$ of real Hodge structures is Tannakian. Indeed, Lemma 1.7 explains that a real Hodge structure corresponds to a representation of the Deligne torus $\mathbb{S}=\mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_{m,\mathbb{C}}$. In fact, one can check (e.g., with Theorem 1.135) that the category $\mathrm{HS}_{\mathbb{Q}}$ of rational Hodge structures continues to be a Tannakian category.

Example 1.140. If $\mathcal{C}$ is a neutral Tannakian category over a field F with fiber functor ω, and $\mathcal{D}$ is an abelian rigid symmetric monoidal category equippd with a faithful exact $\otimes$-functor $\mathcal{D}\to\mathcal{C}$, then the composite

$$\mathcal{D}\to\mathcal{C}\xrightarrow{\omega}\mathrm{Vec}_F$$

becomes a fiber functor for $\mathcal{D}$, thereby making $\mathcal{D}$ neutral Tannakian.

For more examples, we pass to subcategories.

Definition 1.141 ($\otimes$-subcategory)**.** Fix an abelian rigid symmetric monoidal category $\mathcal{C}$. Then the *full $\otimes$-subcategory* generated by a subset $S\subseteq\mathcal{C}$ of objects, denoted $\langle S\rangle^{\otimes}$ is the smallest full abelian rigid monoidal subcategory.

Remark 1.142. One can see (e.g., via Example 1.140) that a fiber functor for $\mathcal{C}$ will induce a fiber functor for a full abelian rigid monoidal subcategory.

Example 1.143. Given a rational Hodge structure V, we claim that the Mumford–Tate group $\mathrm{MT}(V)$ is exactly the group corresponding to the subcategory $\langle V\rangle^{\otimes} \subseteq \mathrm{HS}_{\mathbb{Q}}$. Indeed, we can see that $\langle V\rangle^{\otimes}$ consists of the Hodge substructures W of large tensors T which look like

$$T := \bigoplus_{i=1}^{N} \left(V^{\otimes m_i} \otimes (V^{\vee})^{\otimes n_i}\right),$$

but Proposition 1.33 explains that $W \subseteq T$ is a rational Hodge substructure if and only if W is a subrepresentation of $\mathrm{MT}(V)$. This implies $\langle V\rangle^{\otimes} \subseteq \mathrm{Rep}_{\mathbb{Q}}(\mathrm{MT}(V))$, and this embedding is essentially surjective because all representations of $\mathrm{MT}(V)$ can be generated by the (faithful) standard representation V [Mil17, Theorem 4.14].

The above example is in fact extremely important: it is the guiding principle behind what a monodromy group is. In particular, this idea of monodromy group is akin to the definition of a fundamental group as the automorphism group of the category of covering spaces, and it is akin to defining the étale fundamental group as the automorphism group of the category of finite étale covering spaces. Let's codify this intuition into some notation.

Notation 1.144. Fix a neutral Tannakian category $\mathcal{C}$ over a field F. Given a fiber functor $\omega\colon \mathcal{C} \to \mathrm{Vec}_F$, we set $G_\omega := \underline{\mathrm{Aut}}^{\otimes}\omega$ to be the corresponding group. For any subset $S \subseteq \mathcal{C}$, we define $G_\omega(S)$ to be the group corresponding to the tensor subcategory $\langle S\rangle^{\otimes}$.

Remark 1.145. If $S \subseteq T$, then $\langle T\rangle^{\otimes} \subseteq \langle S\rangle^{\otimes}$, so we induce a surjection $G_\omega(T) \twoheadrightarrow G_\omega(S)$.

While we're discussing (neutral) Tannakian categories, we take a moment to define some useful language. Because we will be interested in constructing a useful neutral Tannakian category from the category $\mathcal{P}(K)^{\mathrm{op}}$, it will be helpful to have a notion of some gradings and "Tate twist" in our category.

Definition 1.146 (grading)**.** Fix a field F. A *$\mathbb{Z}$-grading* on an F-linear abelian symmetric monoidal category $\mathcal{C}$ is a homomorphism $\mathbb{G}_m \to \underline{\mathrm{Aut}}^{\otimes}(\mathrm{id}_{\mathcal{C}})$.

Remark 1.147. The data of the homomorphism $w\colon \mathbb{G}_m \to \underline{\mathrm{Aut}}^{\otimes}(\mathrm{id}_{\mathcal{C}})$ is equivalent to the data of a homomorphism $\mathbb{G}_m \to \mathrm{Aut}_{\mathcal{C}}\, X$ for each object $X \in \mathcal{C}$ which is functorial in X and respects tensor products, where the latter means that $w(t)(X \otimes Y) = w(t)(X) \otimes w(t)(Y)$ for any $t \in \mathbb{G}_m$ and $X, Y \in \mathcal{C}$. By diagonalizing the $\mathbb{G}_m$-action in the usual way, we see that this is equivalent to producing a functorial $\mathbb{Z}$-grading on each object $X \in \mathcal{C}$ (say, $X = \bigoplus_{n\in\mathbb{Z}} X_n$) which also preserves tensor products, in that

$$(X \otimes Y)_n = \bigoplus_{i+j=n} X_i \otimes Y_j.$$

This particular grading of the tensor product arises from the (diagonalization) identification $\mathrm{Rep}_F\, \mathbb{G}_m = \mathrm{GrVec}_F$.

Definition 1.148 (Tate triple)**.** Fix a field F. A *Tate triple* is a triple $(\mathcal{C}, w, \mathrm{T})$ of a neutral Tannakian category $\mathcal{C}$ over F, a *weight $\mathbb{Z}$-grading* $w\colon \mathbb{G}_m \to \underline{\mathrm{Aut}}^{\otimes}(\mathrm{id}_{\mathcal{C}})$, and an invertible object $\mathrm{T} \in \mathcal{C}$ (called the Tate twist) whose induced $\mathbb{Z}$-grading is supported in degree -2. A morphism of Tate triples is a tensor functor preserving the grading and Tate twist.

Notation 1.149. Fix a Tate triple $(\mathcal{C}, w, \mathrm{T})$ over a field F. For any object $X \in \mathcal{C}$ and integer $n \in \mathbb{Z}$, we may write $X(n) := X \otimes \mathrm{T}^{\otimes n}$.

Example 1.150. The category $\mathrm{HS}_{\mathbb{Q}}$ of rational Hodge structures is aready neutral Tannakian. Continuing, we note that all Hodge structures already come with a functorial weight grading which preserves tensor products. (Explicitly, for a Hodge structure V, the decomposition $V_{\mathbb{C}} = \bigoplus_{i,j} V^{ij}$ may define the grading by $V_n \coloneqq \bigoplus_{i+j=n} V$.) This becomes a Tate triple after defining the Tate twist $\mathrm{T} \coloneqq \mathbb{Q}(1)$.

Remark 1.151. Because T is invertible, we see that $\langle \mathrm{T} \rangle^{\otimes}$ simply has quotients of objects of the form $\bigoplus_i \mathrm{T}^{\otimes n_i}$. Because T has pure nonzero weight, we see that $\langle \mathrm{T} \rangle^{\otimes}$ admits a fully faithful functor to GrVec_F with essential image in fact equivalent to GrVec_F. We conclude that $G_\omega(T) = \mathbb{G}_m$.

It is helpful to have some more concrete ways to understand G from its Tannakian category. Here are a few incarnations of this by "functoriality."

Proposition 1.152. Fix a morphism $f\colon G \to G'$ of affine F-groups G, and let $\omega\colon \mathrm{Rep}_F(G') \to \mathrm{Rep}_F(G)$ be the corresponding functor.

(a) Suppose $\mathrm{Rep}_F(G)$ is semisimple and that F has characteristic 0. Then f is faithfully flat if and only if the following holds: for given $X' \in \mathrm{Rep}_F(G')$, every subobject of $\omega(X')$ is isomorphic to $\omega(Y')$ for some subobject Y' of X'.

(b) Then f is a closed embedding if and only if every object $X \in \mathrm{Rep}_F(G)$ is isomorphic to a subquotient of $\omega(X')$ for some $X' \in \mathrm{Rep}_F(G')$.

Proof. Combine [DM12, Remark 2.29] with [DM12, Proposition 2.21]. ∎

Proposition 1.153. Fix an affine F-group G.

(a) Then G is finite if and only if there is an object X such that every object of $\mathrm{Rep}_F(G)$ is a subquotient of $X^{\oplus n}$ for some nonnegative n.

(b) Then G is algebraic (namely, finite type over F) if and only if $\mathrm{Rep}_F(G)$ equals $\langle X \rangle^{\otimes}$ for some object X.

Proof. See [DM12, Proposition 2.20]. ∎

Proposition 1.154. Fix a field F of characteristic 0 and an affine F-group G. Then $G^\circ \subseteq G$ is a projective limit of reductive F-groups if and only if $\mathrm{Rep}_F(G)$ is semisimple.

Proof. See [DM12, Remark 2.28]. ∎

Example 1.155. The category of polarizable Hodge structures is semisimple, so its corresponding affine group is pro-reductive by Proposition 1.154.

1.3.4 Chow Motives

In this subsection, we explain how to (conjecturally!) turn the category $\mathcal{P}(K)^{\mathrm{op}}$ into a neutral Tannakian category. Roughly speaking, we are looking for a graded, neutral Tannakian category $\mathcal{M}(K)$ such that each object $X \in \mathcal{P}(K)$ gives rise to an object $h(X) \in \mathcal{M}(K)$. In fact, each $h(X)$ should also spawn objects

$$h^0(X), h^1(X), \dots, h^{2d}(X) \in \mathcal{M}(K),$$

where $d \coloneqq \dim X$. Additionally, regular maps $f\colon X \to Y$ should produce pullback maps $f^*\colon h(Y) \to h(X)$ which respect the grading.

However, it turns out to be desirable to have access to more maps than just these pullbacks. A basic deficiency is that arbitrary regular maps cannot be added together. Here is one incarnation of this: for any $i \in \mathbb{Z}$, it is natural to expect the composite

$$h(X) \twoheadrightarrow h^i(X) \hookrightarrow h(X)$$

to be an endomorphism of $h(X)$,[3] but this map cannot come from an endomorphism $f\colon X \to X$ in general.

Example 1.156. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. Then there is no endomorphism $f\colon \mathbb{P}^1_K \to \mathbb{P}^1_K$ such that $f^*\colon \mathrm{H}^\bullet(\mathbb{P}^1_K) \to \mathrm{H}^\bullet(\mathbb{P}^1_K)$ equals the composite

$$\mathrm{H}^\bullet(\mathbb{P}^1_K) \twoheadrightarrow \mathrm{H}^2(\mathbb{P}^1_K) \hookrightarrow \mathrm{H}^\bullet(\mathbb{P}^1_K).$$

Proof. There are two cases for an endomorphism $f\colon \mathbb{P}^1_K \to \mathbb{P}^1_K$.

- If f is a constant map to a point $x \in \mathbb{P}^1_K$, then f factors into $i \circ p$, where $p\colon \mathbb{P}^1_K \to \{x\}$ is some projection and $i\colon \{x\} \hookrightarrow \mathbb{P}^1_K$ is some inclusion. It follows that $f^* = p^* \circ i^*$ must factor through $\mathrm{H}^\bullet(\{x\})$. However, $\mathrm{H}^\bullet(\{x\})$ is supported in degree 0 by Example 1.113, so the image of f^* must also be supported in degree 0, so we are done.
- If f is non-constant, then it is a finite map of some degree $\deg f$. Then Lemma 1.119 explains that $f_* f^*$ is multiplication by $\deg f$, so it is not possible for f^* to be zero in degree 0 and the identity in degree 2. ∎

To "linearize" our regular maps, we use correspondences.

Definition 1.157 (correspondence)**.** Given X and Y in $\mathcal{P}(K)$, we define *correspondences* as the cycles in $\mathrm{Corr}(X, Y) \coloneqq \mathrm{CH}(X \times Y)$. For $\gamma \in \mathrm{Corr}(X, Y)$, we define $\gamma^*\colon \mathrm{CH}(Y) \to \mathrm{CH}(X)$ by

$$\gamma^*(\beta) \coloneqq \mathrm{pr}_{1*}\left(\gamma \cdot \mathrm{pr}_2^*\,\beta\right)$$

Example 1.158. Let's explain why this is a reasonable definition of γ^*: if $f\colon X \to Y$ is a regular map, then $[\Gamma_f] \in \mathrm{Corr}(X, Y)$, and Lemma 1.121 shows that our pullback (on cohomology) satisfies

$$f^*\beta = \mathrm{pr}_{1*}\left([\Gamma_f] \cdot \mathrm{pr}_2^*\,\beta\right).$$

Thus, we have expanded our regular maps to include sums and differences, but our new expansion needs a notion of composition.

Definition 1.159. Given $X, Y, Z \in \mathcal{P}(K)$ and $\gamma \in \mathrm{Corr}(X, Y)$ and $\delta \in \mathrm{Corr}(Y, Z)$, we define the composite $(\delta \circ \gamma) \in \mathrm{Corr}(X, Z)$ by

$$(\delta \circ \gamma) \coloneqq \mathrm{pr}_{13*}\left(\mathrm{pr}_{12}^*\,\gamma \cdot \mathrm{pr}_{23}^*\,\delta\right).$$

Here are some basic checks.

Notation 1.160. Given $\gamma \in \mathrm{Corr}(X, Y)$, we define $\gamma^\intercal \in \mathrm{Corr}(Y, X)$ to be $\mathrm{sw}^*\,\gamma = \mathrm{sw}_*\,\gamma$, where $\mathrm{sw}\colon X \times Y \to Y \times X$ is the isomorphism swapping the two coordinates.

[3] Such an endomorphism is called a "Künneth projector."

Lemma 1.161. Fix a ground field K, and choose $W, X, Y, Z \in \mathcal{P}(K)$.

(a) The operation $\circ$ is $\mathbb{Z}$-bilinear.

(b) Associativity: given $\gamma \in \mathrm{Corr}(W, X)$ and $\delta \in \mathrm{Corr}(X, Y)$ and $\varepsilon \in \mathrm{Corr}(Y, Z)$, we have $\varepsilon \circ (\delta \circ \gamma) = (\varepsilon \circ \delta) \circ \gamma$.

(c) Function composition: given $\gamma \in \mathrm{Corr}(X, Y)$ and $f \colon W \to X$ and $h \colon Z \to Y$, we have

$$[\Gamma_h^\intercal] \circ \gamma \circ [\Gamma_f] = (f, h)^* \gamma.$$

(d) Functoriality: given $\gamma \in \mathrm{Corr}(X, Y)$ and $\delta \in \mathrm{Corr}(Y, Z)$, we have $(\delta \circ \gamma)^* = \gamma^* \circ \delta^*$.

Proof. All these proofs are basically direct computation with the projection formula and base-change of cycles. Throughout this proof, we may write things like $\mathrm{pr}_{ABC,AC}$ or pr_{AC} for the projection $A \times B \times C \to A \times C$.

(a) Pullbacks are ring homomorphisms, and multiplication is $\mathbb{Z}$-bilinear, so this follows from the definition of $\circ$.

(b) By a direct expansion, we see that $\varepsilon \circ (\delta \circ \gamma)$ is

$$\mathrm{pr}_{WYZ,WZ*} \left(\mathrm{pr}^*_{WYZ,WY} \mathrm{pr}_{WXY,WY*} (\mathrm{pr}^*_{WXY,WX} \gamma \cdot \mathrm{pr}^*_{WXY,XY} \delta) \cdot \mathrm{pr}^*_{WYZ,YZ} \varepsilon \right).$$

By base-change, we see that $\mathrm{pr}^*_{WYZ,WY} \mathrm{pr}_{WXY,WY*} = \mathrm{pr}_{WXYZ,WYZ*} \mathrm{pr}^*_{WXYZ,WXY}$, so the projection formula allows us to collapse the above into

$$\mathrm{pr}_{WXYZ,WZ*} \left(\mathrm{pr}^*_{WXYZ,WX} \gamma \cdot \mathrm{pr}^*_{WXYZ,XY} \delta \cdot \mathrm{pr}^*_{WXYZ,YZ} \varepsilon \right).$$

A symmetric argument shows that this is also equal to $(\varepsilon \circ \delta) \circ \gamma$.

(c) Note that the expression $[\Gamma_h^\intercal] \circ \gamma \circ [\Gamma_f]$ makes sense because we already checked associativity. For clarity, we will show this in two parts.

- We show that $\gamma \circ [\Gamma_f] = (f, \mathrm{id}_Y)^* \gamma$. Well, $[\Gamma_f] = (\mathrm{id}_W, f)_* [W]$, so $\gamma \circ [\Gamma_f]$ is

$$\mathrm{pr}_{WY*} \left(\mathrm{pr}^*_{WX} (\mathrm{id}_W, f)_* [W] \cdot \mathrm{pr}^*_{XY} \gamma \right).$$

Now, base-change implies that $\mathrm{pr}^*_{WXY,WX} (\mathrm{id}_W, f)_* = (\mathrm{id}_W, f, \mathrm{id}_Y)_* \mathrm{pr}^*_{WX,W}$, so the projection formula shows that this equals

$$\mathrm{pr}_{WY*} (\mathrm{id}_W, f, \mathrm{id}_Y)_* \left(\mathrm{pr}^*_W [W] \cdot (\mathrm{id}_W, f, \mathrm{id}_Y)^* \mathrm{pr}^*_{XY} \gamma \right)$$

Functoriality and the fact that $[W]$ is the unit for the intersection product finishes.

- We show that $[\Gamma_h^\intercal] \circ \gamma = (\mathrm{id}_X, h)^* \gamma$. This proof is the same. Note that $[\Gamma_h^\intercal] = (h, \mathrm{id}_Z)_* [Z]$, so $[\Gamma_h^\intercal] \circ \gamma$ is

$$\mathrm{pr}_{XZ*} \left(\mathrm{pr}^*_{XY} \gamma \cdot \mathrm{pr}^*_{YZ} (h, \mathrm{id}_Z)_* [Z] \right).$$

Now, base-change implies that $\mathrm{pr}^*_{XYZ,YZ} (h, \mathrm{id}_Z)_* = (\mathrm{id}_X, h, \mathrm{id}_Z)_* \mathrm{pr}^*_{XZ,Z}$, so the projection formula shows that this equals

$$\mathrm{pr}_{XZ*} (\mathrm{id}_X, h, \mathrm{id}_Z)_* \left((\mathrm{id}_X, h, \mathrm{id}_Z)^* \mathrm{pr}^*_{XY} \gamma \cdot \mathrm{pr}^*_Z [Z] \right).$$

The same sort of functoriality and fact that $[Z]$ is the multiplicative unit finishes.

Combining the above two points completes the proof.

(d) Choose $\alpha \in \mathrm{CH}(Z)$, and we must show that $(\delta \circ \gamma)^* \alpha = \gamma^* \delta^* \alpha$.

On one hand, $\gamma^* \delta^* \alpha$ is

$$\mathrm{pr}_{XY,X*} \left(\gamma \cdot \mathrm{pr}_{XY,Y}^* \, \mathrm{pr}_{YZ,Y*} (\delta \cdot \mathrm{pr}_{YZ,Z}^* \, \alpha) \right).$$

Now, base-change gives $\mathrm{pr}_{XY,Y}^* \, \mathrm{pr}_{YZ,Y*} = \mathrm{pr}_{XYZ,Y*} \, \mathrm{pr}_{XYZ,YZ}^*$, so we may use the projection formula to collapse the above expression into

$$\mathrm{pr}_{XYZ,X*} \left(\mathrm{pr}_{XYZ,XY}^* \, \gamma \cdot \mathrm{pr}_{XYZ,YZ}^* \, \delta \cdot \mathrm{pr}_{XYZ,Z}^* \, \alpha \right)$$

after a little functoriality.

On the other hand, $(\delta \circ \gamma)^* \alpha$ is

$$\mathrm{pr}_{XZ,X*} \left(\mathrm{pr}_{XYZ,XZ*} (\mathrm{pr}_{XYZ,XY}^* \, \gamma \cdot \mathrm{pr}_{XYZ,YZ}^* \, \delta) \cdot \mathrm{pr}_{XZ,Z}^* \, \alpha \right).$$

An application of the projection formula reveals this to be

$$\mathrm{pr}_{XYZ,X*} \left(\mathrm{pr}_{XYZ,XY}^* \, \gamma \cdot \mathrm{pr}_{XYZ,YZ}^* \, \delta \cdot \mathrm{pr}_{XYZ,Z}^* \, \alpha \right),$$

so we are done. ■

Example 1.162. Letting γ be the diagonal class in (c), we see that $f \colon X \to Y$ and $g \colon Y \to Z$ will have

$$[\Gamma_f^\intercal] \circ [\Gamma_g^\intercal] = [\Gamma_{g \circ f}^\intercal].$$

Thus, $\mathcal{P}(K)$ with correspondences for its morphisms produces a $\mathbb{Z}$-linear category. We will not show it now (because we do not need it), but this category admits sums given by $\sqcup$, and it is a symmetric monoidal category where the tensor product is given by $\times$.

Quickly, it is worthwhile to note that we ought to not work with all correspondences for our morphisms because many "shift degree" in a way that the graph of a regular map would not.

Notation 1.163. Given X and Y in $\mathcal{P}(K)$, subdivide X into $\bigcup_{d \geq 0} X_d$,w here X_d is the union of d-dimensional irreducible components. For each $i \in \mathbb{Z}$, we define

$$\mathrm{Corr}^i(X, Y) \coloneqq \bigoplus_{d \geq 0} \mathrm{CH}^{d+i}(X_d \times Y).$$

Example 1.164. If $f \colon Y \to X$ is a regular map and $\dim X = d$, then $[\Gamma_f^\intercal]$ is a class of codimension d in $X \times Y$. If X is no longer equidimensional, then we still have $[\Gamma_f^\intercal] \in \mathrm{Corr}^0(X, Y)$ by construction.

Remark 1.165. Because pullback preserves codimension, and pushforward preserves dimension, we see that $\circ$ defines an operation

$$\circ \colon \mathrm{Corr}^j(Y, Z) \times \mathrm{Corr}^i(X, Y) \to \mathrm{Corr}^{i+j}(X, Z)$$

for any $i, j \in \mathbb{Z}$. Indeed, by dividing everything into connected components, we may assume that everything in sight is connected. Then $\gamma \in \mathrm{Corr}^i(X, Y)$ and $\delta \in \mathrm{Corr}^j(Y, Z)$ makes $\mathrm{pr}_{12}^* \, \gamma \cdot \mathrm{pr}_{23}^* \, \delta$ have codimension $i + j + \dim X + \dim Y$ and hence dimension $\dim Z - (i + j)$, so the pushforward has codimension $(i + j) + \dim X$.

Thus, we may want to consider a category $\mathcal{C}_{\mathbb{Q}}(K)$ where the objects are given by $h(X)$ for any $X \in \mathcal{P}(K)$, and the morphisms are given by

$$\mathrm{Mor}_{\mathcal{C}_{\mathbb{Q}}(K)}(h(X), h(Y)) \coloneqq \mathrm{Corr}^0(X, Y)_{\mathbb{Q}}.$$

(The composition is well-defined by Remark 1.165.) The category $\mathcal{C}_{\mathbb{Q}}(K)$ already has some desirable properties. We know that it is $\mathbb{Q}$-linear (by Lemma 1.161), and there is already a canonical faithful contravariant functor $h\colon P(K)^{\mathrm{op}} \to \mathcal{C}_{\mathbb{Q}}(K)$ given by sending $X \mapsto h(X)$ and a morphism $f\colon Y \to X$ to $[\Gamma_f^{\intercal}] \in \mathrm{Corr}^0(X,Y)$ (by Examples 1.162 and 1.164). Here are two more easy checks.

Lemma 1.166. The category $\mathcal{C}_{\mathbb{Q}}(K)$ is additive. In fact, $h(X) \times h(Y) \cong h(X \sqcup Y)$.

Proof. The empty product is $h(\varnothing)$. As for products of two objects, after undoing the transposition, we need to show that the inclusions induce a natural isomorphism

$$\mathrm{Corr}^0(X \sqcup Y, -) \overset{?}{\simeq} \mathrm{Corr}^0(X,-) \times \mathrm{Corr}^0(Y,-),$$

which amounts to checking

$$\bigoplus_{d,e\geq 0} \mathrm{CH}^{d+e}((X_d \sqcup Y_e) \times -) \overset{?}{\simeq} \bigoplus_{d\geq 0} \mathrm{CH}^d(X_d \times -) \oplus \bigoplus_{e\geq 0} \mathrm{CH}(Y_e \times -),$$

where X_d is the union of the d-dimensional irreducible components of X (with Y_e defined analogously). Well, $(X \sqcup Y) \times -$ is isomorphic to $(X \times -) \sqcup (Y \times -)$ already as schemes, so this follows because any cycle on a disjoint union can be uniquely decomposed into a cycle on either part. In other words, we see that the inclusions induce a natural isomorphism

$$\mathrm{CH}(X \times -) \times \mathrm{CH}(Y \times -) \to \mathrm{CH}((X \times -) \sqcup (Y \times -)),$$

so we are done after tracking that the codimensions pass through correctly on each irreducible component ∎

Lemma 1.167. The category $\mathcal{C}_{\mathbb{Q}}(K)$ admits the structure of a symmetric monoidal category with unit $h(\mathrm{pt})$ and product $h(X) \otimes h(Y) \coloneqq h(X \times Y)$.

Proof. In fact, $\mathcal{P}(K)$ is already a symmetric monoidal category with unit pt and product $\times$. We already have commutativity and associativity constraints induced by the universal property of the fiber product, and there is a canonical isomorphism $X \times \mathrm{pt} \to X$. The various coherences required for $\times$ here are automatically satisfied by the universal property of the fiber product. ∎

Thus, we can see that $\mathcal{C}_{\mathbb{Q}}(K)$ is pretty close to our category of motives, but it has two key failures at being neutral Tannakian.

- The category $\mathcal{C}_{\mathbb{Q}}(K)$ fails to be abelian. Glaringly, there are many correspondences which fail to have kernels.
- The category $\mathcal{C}_{\mathbb{Q}}(K)$ fails to be rigid. Namely, we want to have duals, which by an expected Poincaré duality axiom, more or less amounts to adding a Tate twist.

We are going to handle each of these concerns individually. To begin, we will not add all kernels and cokernels or even all kernels; it turns out that it will be enough to merely add kernels of idempotents. This is a rather explicit construction in pure category theory.

Definition 1.168 (Karoubian)**.** A $\mathbb{Q}$-linear category $\mathcal{C}$ is *Karoubian* or *pre-abelian* if and only if any $X \in \mathcal{C}$ and idempotent $p\colon X \to X$ admits a kernel.

Remark 1.169. Because $p\colon X \to X$ is idempotent, we see that $(1-p)\colon X \to X$ is also an idempotent. As such, we claim that

$$X \stackrel{?}{=} \ker(p) \oplus \ker(1-p).$$

Indeed, this follows by writing out what it means to be a direct sum in an additive category and noting that the relevant equations are satisfied because $1 = p + (1-p)$ and $p(1-p) = (1-p)p = 0$. In particular, we see that $p\colon X \to X$ factors through $\ker(1-p)$, and $(1-p)\colon X \to X$ factors through $\ker(p)$.

Lemma 1.170. Fix a $\mathbb{Q}$-linear category $\mathcal{C}$, and define the category $\mathrm{Split}(\mathcal{C})$ to be the category whose objects are pairs (X,p) where $X \in \mathcal{C}$ and $p\colon X \to X$ is idempotent, and morphisms are given by

$$\mathrm{Mor}_{\mathrm{Split}(\mathcal{C})}((X,p),(Y,q)) \coloneqq q \circ \mathrm{Mor}_{\mathcal{C}}(X,Y) \circ p.$$

Then $\mathrm{Split}(\mathcal{C})$ is $\mathbb{Q}$-linear and Karoubian. Further, any $\mathbb{Q}$-linear functor $F\colon \mathcal{C} \to \mathcal{D}$ to a Karoubian category factors uniquely through $\mathrm{Splic}(\mathcal{C})$.

Proof. We have many checks. Intuitively, the point is that (X,p) should be the image of the idempotent of $p\colon X \to X$; in particular, because $1 = p + (1-p)$, the object (X,p) should be the kernel of the idempotent $(1-p)$.

1. We check that $\mathrm{Split}(\mathcal{C})$ makes sense as an additive $\mathbb{Q}$-linear category. Note $\mathrm{Mor}((X,p),(Y,q))$ is a $\mathbb{Q}$-subspace of $\mathrm{Mor}(X,Y)$, and with composition defined as usual, we still have identity morphisms (where $p \in \mathrm{Mor}((X,p),(X,p))$ behaves as an identity), and composition is well-defined and $\mathbb{Q}$-bilinear by construction.

   While we're here, we note that there is a $\mathbb{Q}$-linear faithful functor $h\colon \mathcal{C} \to \mathrm{Split}(\mathcal{C})$ sending objects X to $(X,1)$ and morphisms to themselves.

2. We check that $\mathrm{Split}(\mathcal{C})$ is Karoubian. Well, let $pfp\colon (X,p) \to (X,p)$ be some idempotent, and we need to show that this map has a kernel. For brevity, we $q \coloneqq pfp$, and we note that $pq = q = qp$ because p is itself idempotent. Now, $q\colon X \to X$ is already some endomorphism, and q and hence $(1-q)$ are idempotent by hypothesis, so $(X,1-q)$ is an object in $\mathrm{Split}(\mathcal{C})$ which we expect to be the kernel. Note that there is a canonical map $(X,p-q) \to (X,p)$ given by $p(p-q) = p-q$.

   It remains to check that we have actually constructed the kernel. Suppose we have some morphism $pgr\colon (Z,r) \to (X,p)$ such that $pqgr = 0$. We would like this pgr to factor uniquely through $(X,1-q)$. Namely, we are looking for some unique $(p-q)g'r\colon (Z,r) \to (X,p-q)$ such that

   $$p(p-q)g'r = pgr.$$

   Certainly $g = g'$ works because $pqgr = 0$ by hypothesis; on the other hand, if some other g' has $p(1-q)g'r = p(1-q)gr$, then we note that $(1-p)(p-q)g'r = (1-p)(p-q)gr$ as well because $(1-p)(p-q) = 0$, so summing gives $(p-q)g'r = (p-q)gr$.

3. Suppose that $F\colon \mathcal{C} \to \mathcal{D}$ is a $\mathbb{Q}$-linear functor to a Karoubian category, which we would like to uniquely factor through h. Well, we will simply describe how to extend the functor F on $\mathcal{C}$ to a functor G on $\mathrm{Split}(\mathcal{C})$. For each $(X,p) \in \mathrm{Split}(\mathcal{C})$, we must determine $G((X,p)) \in \mathcal{D}$; well, G needs to be an additive functor, so Remark 1.169's decomposition

   $$(X,1) = (X,p) \oplus (X,1-p)$$

   shows that $G((X,p))$ must be the kernel of $F(1-p)\colon FX \to FX$ (which is equivalently the image of Fp). (This provides uniqueness up to some natural isomorphism.) Continuing, any morphism $qfp\colon (X,p) \to (Y,q)$ must factor through the aforementioned decompositions,[4] and therefore must

[4] Explicitly, the morphism $(X,p) \to (Y,q)$ can be expressed as a composite $(X,p) \hookrightarrow X \xrightarrow{f} Y \to (Y,q)$, whose behavior upon being passed through G is now forced by F.

be sent to the induced map on $G(qfp)$. Lastly, we ought to check that this functor is well-defined: well, G sends identities to identities by construction, and the relevant uniqueness in place provides functoriality. ■

Definition 1.171 (Karoubian envelope)**.** Given a $\mathbb{Q}$-linear category $\mathcal{C}$, we define the $\mathbb{Q}$-linear additive category $\mathrm{Split}(\mathcal{C})$ of Lemma 1.170 to be the *Karoubian envelope*.

Remark 1.172. If $\mathcal{C}$ is additive, then $\mathrm{Split}(\mathcal{C})$ is also: the direct sum of (X,p) and (Y,q) can simply be given by $(X\oplus Y,(p,q))$. Indeed, note a pair of morphisms $rfp\colon (X,p)\to(Z,r)$ and $rgq\colon (Y,q)\to(Z,r)$ amount to the same data as a single morphism $(rfp,rgq)\colon (X\oplus Y,(p,q))\to(Z,r)$.

Remark 1.173. If $\mathcal{C}$ admits a symmetric monoidal structure given by $\otimes$, then $\mathrm{Split}(\mathcal{C})$ does as well, where we define

$$(X,p)\otimes(Y,q)\coloneqq(X\otimes Y,p\otimes q).$$

The relevant coherences for $\otimes$ all lift from $\mathcal{C}$ to $\mathrm{Split}(\mathcal{C})$.

Example 1.174. Let's exhibit the sort of decompositions we can exhibit in $\mathrm{Split}(\mathcal{C})$. Suppose that we have a "projection" $p\colon X\to Y$ in $\mathcal{C}$ with a section $s\colon Y\to X$, meaning that $ps=\mathrm{id}_Y$. Then we note that $sp\colon X\to X$ is an idempotent, and we claim that $(X,sp)\cong(Y,\mathrm{id}_Y)$, meaning that Y is not a sub-object of X in $\mathrm{Split}(\mathcal{C})$! To show this, we note that $p=psp$ is a map $(X,sp)\to(Y,\mathrm{id}_Y)$, and $s=sps$ is a map $(Y,\mathrm{id}_Y)\to(X,sp)$, and we know $ps=\mathrm{id}_Y$ and $sp=\mathrm{id}_{(X,sp)}$.

Thus, to make $\mathcal{C}_{\mathbb{Q}}(K)$ more abelian, we can take its Karoubian envelope. This produces the category of effective Chow motives.

Definition 1.175 (effective Chow motives)**.** Fix a ground field K. The category $\mathrm{ChMot}^{+}_{\mathbb{Q}}(K)$ of *effective Chow motives* is the Karoubian envelope of $\mathcal{C}_{\mathbb{Q}}(K)$.

Remark 1.176. Because $\mathcal{C}_{\mathbb{Q}}(K)$ is $\mathbb{Q}$-linear, additive, and symmetric monoidal, the same holds for its Karoubian envelope $\mathrm{ChMot}^{+}_{\mathbb{Q}}(K)$ (see Remarks 1.172 and 1.173), but effective Chow motives now succeed at being Karoubian. Notably, the canonical functor $h\colon \mathcal{P}(K)^{\mathrm{op}}\to\mathcal{C}_{\mathbb{Q}}(K)$ extends to $\mathrm{ChMot}^{+}_{\mathbb{Q}}(K)$, and we may write the effective Chow motive $(h(X),p)$ as simply $ph(X)$.

Example 1.177 (Künneth projector)**.** It is a standard conjecture that there is a correspondence $h(X)\to h(X)$ giving rise to the Künneth projections, so $h^i(X)$ can be defined as the image.

As our standard example, let's begin computing the motive of $\mathbb{P}^1$: by Example 1.127, we are expecting $h(\mathrm{pt})$ and some other piece given by a Tate twist.

Lemma 1.178. Fix a ground field K. Suppose some irreducible $X\in\mathcal{P}(K)$ has a K-rational point $\infty\in X(K)$. Then $h(\mathrm{pt})$ is a sub-object of $h(X)$.

Proof. We use Example 1.174. Consider the structure morphism $p\colon X\to\mathrm{pt}$ and the embedding $i\colon \mathrm{pt}\to X$. Then $pi=\mathrm{id}_{\mathrm{pt}}$, so $h(i)\circ h(p)=\mathrm{id}_{h(\mathrm{pt})}$ by functoriality, so the result follows. ■

Next up, we would like to add in a Tate twist to recover our rigidity. Namely, we would like to have duals. For example, Lemma 1.178 tells us that $h(\mathbb{P}^1)$ decomposes as

$$h(\mathbb{P}^1)=h(\mathrm{pt})\oplus\mathrm{L}$$

for some effective Chow motive L, which is expected to be the dual of the Tate twist by Example 1.127. Thus, to ensure that L has a dual, we must add in its inverse! Note that once we have Tate twists, Poincaré duality tells us that we expect all of our Chow motives to have duals. We are now ready to define the category of Chow motives.

Definition 1.179 (Chow motives)**.** Fix a ground field K. The category $\mathrm{ChMot}_{\mathbb{Q}}(K)$ of *Chow motives* is defined as the category of triples (X, p, i) where $X \in \mathcal{C}_{\mathbb{Q}}(K)$ and $p \in \mathrm{Corr}^0(X, X)$ is an idempotent and $i \in \mathbb{Z}$, where morphisms are given by

$$\mathrm{Hom}_{\mathrm{ChMot}_{\mathbb{Q}}(K)}((X, p, i), (Y, q, j)) \coloneqq q \circ \mathrm{Corr}^{j-i}(X, Y)_{\mathbb{Q}} \circ p.$$

For brevity, we define the *Tate motive* $\mathrm{T} \coloneqq (\mathrm{pt}, \mathrm{id}, 1)$ and the *Lefschetz motive* $\mathrm{L} \coloneqq (\mathrm{pt}, \mathrm{id}, -1)$.

Remark 1.180. As usual, we remark that composition makes sense by Remark 1.165 and is $\mathbb{Q}$-linear by Lemma 1.161.

Remark 1.181. The canonical faithful, essentially surjective, $\mathbb{Q}$-linear functor $h\colon \mathcal{P}(K)^{\mathrm{op}} \to \mathcal{C}_{\mathbb{Q}}(K)$ extends to $\mathrm{ChMot}_{\mathbb{Q}}(K)$ by $X \mapsto (X, \Delta_X, 0)$, where $\Delta_X \subseteq X \times X$ is the diagonal. (The idea is that the "degree-0" part of our Chow motives simply recovers the effective Chow motives.) As such, we may write the Chow motive (X, p, i) as $ph(X)(i)$.

Remark 1.182. One may alternatively define Chow motives by taking $\mathcal{C}_{\mathbb{Q}}(K)$, first adding in Tate twists by considering pairs (X, i) where $i \in \mathbb{Z}$, and then taking the Karoubian envelope. We have not done this because the intermediate category of pairs (X, i) is not obviously additive: for example, how should one add $(\mathrm{pt}, 0)$ and $(\mathrm{pt}, 1)$?

Remark 1.183. Note that $\mathrm{ChMot}_{\mathbb{Q}}(K)$ continues to be Karoubian. The point is that an idempotent q of some triple (X, p, i) will have $q \in \mathrm{Hom}((X, p), (X, p))$ anyway, so letting $(X, p) = \ker(q) \oplus \mathrm{im}(q)$ be the sum of Remark 1.169 in the category of effective Chow motives, we see

$$(X, p, i) = (\ker(q), i) \oplus (\mathrm{im}(q), i)$$

by shifting the Tate twist by i everywhere, so we conclude that $(\ker(q), i)$ is the kernel of $q\colon (X, p, i) \to (X, p, i)$.

Here are our basic checks on this category.

Lemma 1.184. The category $\mathrm{ChMot}_{\mathbb{Q}}(K)$ admits the structure of a symmetric monoidal category with unit $h(\mathrm{pt})(0)$.

Proof. Unsurprisingly, we define

$$(X, p, i) \otimes (Y, q, j) \coloneqq (X \times Y, p \times q, i + j).$$

Then one can simply repeat the proof of Lemma 1.167, carrying around commutativity and associativity of addition in $\mathbb{Z}$ to upgrade the commutativity and associativity constraints. ∎

Remark 1.185. This is not actually the correct symmetric monoidal structure! In short, the problem is the commutativity constraint does not take into account the fact that $h(X)$ should behavae as a graded commutative algebra. Explicitly, given any Weil cohomology theory $\mathrm{H}^\bullet$, we would like the commutativity constraint $h(X) \otimes h(Y) \to h(Y) \otimes h(X)$ to be given by

$$\mathrm{H}^\bullet(X) \otimes \mathrm{H}^\bullet(Y) = \mathrm{H}^\bullet(X \times Y) \xrightarrow{\mathrm{sw}} \mathrm{H}^\bullet(Y \times X) = \mathrm{H}^\bullet(Y) \otimes \mathrm{H}^\bullet(X).$$

But sw needs to be an isomorphism of graded commutative rings, so the map $\mathrm{H}^i(X) \otimes \mathrm{H}^j(Y) \to \mathrm{H}^j(Y) \otimes \mathrm{H}^i(X)$ needs to have the sign $(-1)^{ij}$.

Example 1.186. We now see that $ph(X)(i) = ph(X) \otimes \mathrm{T}^i$, thus explaining why we might view the category of Chow motives as simply the category of effective Chow motives extended by the Tate twist $\mathrm{T} = h(\mathrm{pt})(1)$.

Example 1.187. Fix a ground field K. Then

$$h\left(\mathbb{P}^1_K\right) = h(\mathrm{pt}) \oplus \mathrm{L}.$$

In particular, L is an effective Chow motive.

Proof. We imitate Example 1.127. For brevity, set $X \coloneqq \mathbb{P}^1_K$. Upon choosing a point $\infty \in X$, we recall from Example 1.127 that we had a "motivic" input

$$[\Delta_X] = [\infty \times X] + [X \times \infty],$$

where $\Delta_X \subseteq X \times X$ is the diagonal. Notably, $[\Delta_X] = \mathrm{id}_{h(X)}$, so the above is a decomposition of the identity. In fact, it is a decomposition into idempotents: for example, $[\infty \times X]$ is $[\Gamma_i^\intercal]$, where $i \colon X \to X$ is the constant map sending all points to ∞, so the equality $i \circ i = i$ implies that $[\Gamma_i^\intercal]$ is an idempotent by Example 1.162. It follows that $[X \times \infty]$ is the orthogonal idempotent.

Now, Lemma 1.178 tells us that $h(\mathrm{pt})$ is already a sub-object of $h(X)$, and in fact the proof shows that $h(\mathrm{pt})$ is in fact the image of $h(i) \colon h(X) \to h(X)$; in other words, $h(\mathrm{pt})$ is isomorphic to $(X, [\infty \times X])$. Thus, it remains to check that

$$\mathrm{L} \overset{?}{\cong} (X, [X \times \infty]).$$

In fact, we suspect that L should be the image of $[X \times \mathrm{pt}] \in \mathrm{Corr}^0(X, X)$. Indeed, $[X \times \mathrm{pt}]$ is an element of

$$\mathrm{Hom}_{\mathrm{ChMot}_{\mathbb{Q}}(K)}((X, [X \times \infty], 0), (\mathrm{pt}, \mathrm{id}, -1)) = \mathrm{CH}^0(X \times \mathrm{pt}) \circ [X \times \infty]$$

because $[X \times \mathrm{pt}] \circ [X \times \infty] = [X \times \mathrm{pt}]$ by a direct calculation of the composition. On the other hand, $[\mathrm{pt} \times \infty]$ is an element of

$$\mathrm{Hom}_{\mathrm{ChMot}_{\mathbb{Q}}(K)}((\mathrm{pt}, \mathrm{id}, -1), (X, [X \times \infty], 0)) = [X \times \infty] \circ \mathrm{CH}^1(\mathrm{pt} \times X)$$

because $[X \times \infty] \circ [\mathrm{pt} \times x] = [\mathrm{pt} \times \infty]$ for any $x \in X(K)$. It remains to calculate $[X \times \mathrm{pt}] \circ [\mathrm{pt} \times \infty] = [\mathrm{pt} \times \mathrm{pt}]$ and $[\mathrm{pt} \times \infty] \circ [X \times \mathrm{pt}] = [X \times \infty]$ are both their respective identities, so we are done. ∎

Lemma 1.188. The category $\mathrm{ChMot}_{\mathbb{Q}}(K)$ is additive.

Proof. The empty product is $h(\varnothing)$. We exhibit our sums in two steps.

1. Copying the proof of Lemma 1.166 with an appropriate degree change shows $ph(X)(i) \times qh(Y)(i) = (p \sqcup q)h(X \sqcup Y)(i)$, so the main problem is dealing with degree shifts. (In degree $i = 0$, we already knew this from Remark 1.176.) To be slightly more explicit, after decomposing X and Y as $X = \bigsqcup_{d \geq 0} X_d$ and $Y = \bigsqcup_{e \geq 0} Y_e$ into equidimensional pieces, we find

$$\mathrm{Hom}_{\mathrm{ChMot}_{\mathbb{Q}}(K)}((X \sqcup Y, p \sqcup q, i), (Z, r, j)) = \bigoplus_{d \geq 0} r \circ \mathrm{CH}^{d+j-i}((X_d \sqcup Y_d) \times Z) \circ (p \sqcup q),$$

   which then decomposes into cycles on X and Y individually as in the proof of Lemma 1.166.

2. We now reduce to the previous case. For any Chow motives (X, p, i) and (Y, q, j), we note that there is an integer n large enough so that $(X, p, i) \otimes \mathrm{L}^{\otimes n}$ and $(Y, q, j) \otimes \mathrm{L}^{\otimes n}$ are both effective: for example, $(X, p, i-n)$ becomes effective as soon as $i - n$ is nonpositive, for then we get $(X, p, 0) \otimes \mathrm{L}^{-(i-n)}$, which is effective by Example 1.187. Thus, we may define the sum of (X, p, i) and (Y, q, j) as

$$\big((X, p, i) \otimes \mathrm{L}^{\otimes n} \oplus (Y, q, j) \otimes \mathrm{L}^{n}\big) \otimes \mathrm{T}^{\otimes n}.$$

   The fact that L and T are inverses shows that this is in fact a valid sum.[5] ■

Thus, we have built a $\mathbb{Q}$-linear, additive, and Karoubian category $\mathrm{ChMot}_{\mathbb{Q}}(K)$ of Chow motives. The remaining properties are only conjectural.

Conjecture 1.189 (Grothendieck)**.** The category $\mathrm{ChMot}_{\mathbb{Q}}(K)$ is a semisimple neutral Tannakian category.

Remark 1.190. It turns out that any pre-Weil cohomology theory $\mathrm{H}^\bullet \colon \mathcal{P}(K)^{\mathrm{op}} \to \mathrm{GrVec}_F$ extends to a unique $\mathbb{Q}$-linear symmetric monoidal functor $\mathrm{ChMot}_{\mathbb{Q}}(K) \to \mathrm{GrVec}_F$, fulfilling a prophecy from the start of this subsection. In fact, one expects this functor to be a fiber functor for our neutral Tannakian category! We will not need this fact, and the proof is rather involved, so we will not prove it. Instead, we refer to [SP, Proposition `0FHM`], and we note that Theorem 1.210 proves a version of this in the next section.

1.3.5 Motives from Absolute Hodge Cycles

The goal of the present subsection is to build a concrete category $\mathrm{Mot}_{\mathbb{Q}}(K)$ of motives which we can prove satisfies the required properties (namely, it is semisimple neutral Tannakian and lives in a Tate triple) and is conjecturally equivalent to $\mathrm{ChMot}_{\mathbb{Q}}(K)$. The idea is to add in more correspondences to $\mathrm{Corr}(X, Y)$. For example, the previous subsection repeatedly asked for an idempotent $h(X) \to h(X)$ whose image is $h^i(X)$, but the existence of such correspondences in $\mathrm{Corr}^0(X, X)$ is still conjectural. Thus, we will want the category $\mathrm{Mot}_{\mathbb{Q}}(K)$ to admit such correspondences.

In particular, isntead of having $\mathrm{Corr}(X, Y)$ be made up of algebraic cycle classes, we will use absolute Hodge classes, following [Del18]. For motivation, we want the Hodge classes on a complex Kähler manifold X to be elements of the cohomology group $\mathrm{H}^{2i}_{\mathrm{dR}}(X, \mathbb{C})(i)$ of bidegree $(0, 0)$ and satisfying some rationality condition. The definition of an absolute Hodge class comes from trying to be agnostic about the embedding of the base field of X.

[5] One can see that $\mathrm{Hom}(- \otimes \mathrm{T}, -) \simeq \mathrm{Hom}(-, - \otimes \mathrm{L})$ already on the level of correspondences.

Definition 1.191 (absolute Hodge class)**.** Fix a smooth projective variety X over a field K algebraic over $\mathbb{Q}$. An *absolute Hodge class* is an element t of some $\mathrm{H}^{2i}_{\mathbb{A}}(X_{\overline{K}})(i)$ if and only if it satisfies the following properties.

- $\pi_\infty(t)$ lives in the component $(0,0)$ of $\mathrm{H}^{2i}_{\mathrm{dR}}(X,\mathbb{C})$.
- For each embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, the element t is in the image of the embedding $\mathrm{H}^{2i}_{\mathrm{B}}(X,\mathbb{Q})(i)$ into $\mathrm{H}^{2i}_{\mathbb{A}}(X)(i)$.

We denote the collection of these absolute Hodge classes by $C^i_{\mathrm{AH}}(X_{\overline{K}})$ or $C^i_{\mathrm{AH}}(X)$.

Remark 1.192. Deligne [Del18, Section 2] gives a definition for smooth projective varieties defined over a general field of characteristic 0. The above definition makes sense essentially verbatim for any field K of characteristic 0 and finite transcendence degree because then one has access to embeddings into $\mathbb{C}$. For the general case, one must argue that any class with sufficient rationality properties will descend to a field of finite transcendence degree and that the choice of this descent does not matter.

Example 1.193. Any algebraic class $\gamma \in \mathrm{CH}^i(X)$ produces cycle classes in the various cohomology theories. Because γ ought to arise rationally (over K) because it already produces a cycle class in Betti cohomology, we see that taking the corresponding cycle class in $\mathrm{H}^{2i}_{\mathbb{A}}(X)(i)$ successfully produces an absolute Hodge class. Note that the Hodge conjecture would imply that all absolute Hodge classes arise in this way.

Remark 1.194. Here is a notable advantage of working with absolute Hodge classes over typical Hodge classes: there is an action of $\mathrm{Gal}(\overline{K}/K)$ on $\mathrm{H}^{2i}_{\mathbb{A}}(X_{\overline{K}})(i)$ given by the pullback of the action on $X_{\overline{K}}$, but this Galois action may very well permute the image of $\mathrm{H}^{2i}_{\mathrm{B}}(X,\mathbb{Q})(i)$ for a given $\sigma\colon K \hookrightarrow \mathbb{C}$. Indeed, $\tau \in \mathrm{Gal}(\overline{K}/K)$ has $\tau^*\mathrm{H}_\sigma = \mathrm{H}_{\tau\sigma}$. As such, the space of Hodge classes is not obviously a Galois representation, but the space of absolute Hodge classes is!

We are ready to (re)define our correspondences in terms of absolute Hodge classes.

Notation 1.195. Fix a field K algebraic over $\mathbb{Q}$. For any $X, Y \in \mathcal{P}(K)$, we define

$$\mathrm{Corr}_{\mathrm{AH}}(X,Y) \coloneqq C_{\mathrm{AH}}(X \times Y).$$

Upon decomposing X into equidimensional components as $\bigsqcup_{d\geq 0} X_d$, we may set the degree-i component as

$$\mathrm{Corr}^i_{\mathrm{AH}}(X,Y) \coloneqq \bigoplus_{d\geq 0} C^{i+d}_{\mathrm{AH}}(X_d \times Y).$$

It is worthwhile to describe these correspondences cohomologically.

Definition 1.196 (absolute Hodge correspondence)**.** Fix a field K algebraic over $\mathbb{Q}$ and $X, Y \in \mathcal{P}(K)$. Then an *absolute Hodge correspondence of degree* i is a triple $((f_\ell)_\ell, f_{\mathrm{dR}}, (f_\sigma)_\sigma)$ as follows.

- For each prime ℓ, the element f_ℓ is a Galois-invariant graded homomorphism $\mathrm{H}^\bullet_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell) \to \mathrm{H}^\bullet_{\text{ét}}(Y_{\overline{K}}, \mathbb{Q}_\ell)(i)$.
- The element f_{dR} is a graded homomorphism $\mathrm{H}^\bullet_{\mathrm{dR}}(X, \mathbb{C}) \to \mathrm{H}^\bullet_{\mathrm{dR}}(Y, \mathbb{C})(i)$ preserving the Hodge structure.
- For each embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, the element f_σ is a graded homomorphism $\mathrm{H}^\bullet_\sigma(X) \to \mathrm{H}^\bullet_\sigma(Y)(i)$. Further, we require f_ℓ and f_{dR} to agree with f_σ after applying the suitable comparison isomorphism (Theorems 1.75 and 1.79).

Lemma 1.197. Fix a field K algebraic over $\mathbb{Q}$ and $X, Y \in \mathcal{P}(K)$. The group $\mathrm{Corr}^i_{\mathrm{AH}}(X, Y)$ is isomorphic to the vector space of absolute Hodge correspondences of degree i.

Proof. This is [DM12, Proposition 6.1]. We go ahead and decompose $X = \bigsqcup_{d \geq 0} X_d$, where X_d is equidimensional of dimension d. The point is to describe how a correspondence should give rise to a morphism in cohomology. To be explicit, our correspondences are just some classes in $\bigoplus_d \mathrm{H}^{2i+2d}_{\mathbb{A}}(X_d \times Y)(i+d)$, which the Künneth formula and Poincaré duality tell us give rise to elements in

$$\begin{aligned}\mathrm{H}^{2i+2d}_{\mathbb{A}}(X_d \times Y)(i+d) &= \bigoplus_{p+q=2i+2d} \mathrm{H}^p_{\mathbb{A}}(X_d)(d) \otimes \mathrm{H}^q_{\mathbb{A}}(Y)(i) \\ &= \bigoplus_p \mathrm{H}^p_{\mathbb{A}}(X_d)^\vee \otimes \mathrm{H}^{p+2i}_{\mathbb{A}}(Y)(i) \\ &= \mathrm{Hom}\left(\mathrm{H}^\bullet_{\mathbb{A}}(X_d), \mathrm{H}^\bullet_{\mathbb{A}}(Y)(i)\right).\end{aligned}$$

This explains how $\mathrm{Corr}^p_{\mathrm{AH}}(X, Y)$ embeds into the group of tuples $((f_\ell), f_{\mathrm{dR}})$. (Note that the f_σ are uniquely determined if they exist by the nature of the comparison isomorphisms.) It remains to characterize the image, so pick up some $f \in \mathrm{Corr}^i_{\mathrm{AH}}(X, Y)$, and we must describe what the image tuple must look like. Here are our checks.

- Note f is a Hodge cycle by definition, so it must be in the $(0,0)$ component in all the above equalities, eventually causing the induced map f_{dR} on de Rham cohomology to preserve the Hodge structures.
- Because our $f \in \mathrm{Corr}^i_{\mathrm{AH}}(X, Y)$ is required to be absolutely Hodge, it will come from a rational element $f_\sigma \in \mathrm{H}^{2i+2d}_\sigma(X \times Y)(i)$ for each embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, from which the above equalities explain how to produce morphisms $f_\sigma\colon \mathrm{H}_\sigma(X) \to \mathrm{H}_\sigma(X)(i)$. This explains why the f_σ exist.
- Lastly, because f arises rationally, it must be a Galois-invariant class, so because the equalities above are Galois-invariant at each ℓ, we conclude that the f_ℓs are Galois-invariant at the end.

Conversely, given an absolute Hodge correspondence $((f_\ell), f_{\mathrm{dR}}, (f_\sigma)_\sigma)$, we may go backwards to produce $f \in \bigoplus_d \mathrm{H}^{2i+2d}_{\mathbb{A}}(X_d \times Y)(i+d)$, and the above checks are all reversible and thus tell us that the provided f is an absolute Hodge class. ∎

Intuitively, if one can canonically produce a class for all of our known cohomology theories, we receive an absolute Hodge class. Here are a few examples.

Example 1.198 (Künneth projectors)**.** For any pre-Weil cohomology $\mathrm{H}^\bullet$ and $X \in \mathcal{P}(K)$ of dimension d and index $i \in [0, 2d]$, the various projections

$$\mathrm{H}^\bullet(X) \twoheadrightarrow \mathrm{H}^i(X) \hookrightarrow \mathrm{H}^\bullet(X)$$

assemble into an absolute Hodge correspondence. Indeed, this follows from properties of each cohomology theory and their comparison isomorphisms. We call this absolute Hodge correspondence π_i, and we may identify it with an element in $\mathrm{Corr}^0(X, X)$ by Lemma 1.197.

Example 1.199 (Poincaré duality)**.** Fix a field K algebraic over $\mathbb{Q}$ and some $X \in \mathcal{P}(K)$ which is equidimensional of dimension d. Poincaré duality provides a perfect pairing

$$\mathrm{H}^{2d}_{\mathbb{A}}(X \times X) = \bigoplus_i \mathrm{H}^i_{\mathbb{A}}(X) \otimes \mathrm{H}^{2d-i}_{\mathbb{A}}(X) \to \mathrm{H}^{2d}_{\mathbb{A}}(X) \xrightarrow{\int_X} \mathrm{H}^0_{\mathbb{A}}(\mathrm{pt})(-d),$$

which lives in Betti cohomology and is compatible for all of our cohomology theories. Thus, this perfect pairing arises from some absolute Hodge class $\psi \in \mathrm{Corr}^{-d}(X \times X, \mathrm{pt})$.

Example 1.200 (Hodge involution)**.** Fix a field K algebraic over $\mathbb{Q}$ and some $X \in \mathcal{P}(K)$ which is equidimensional of dimension d. For each index i, there is $^* \in \mathrm{Corr}_{\mathrm{AH}}(X, X)$ such that the degree-$(-i)$ component induces an isomorphism

$$\mathrm{H}^i_{\mathbb{A}}(X) \to \mathrm{H}^{2d-i}_{\mathbb{A}}(X)(d-i).$$

Proof. This is the main content of [DM12, Proposition 6.2]. We use the Hard Lefschetz theorem [GH94, p. 122], whose statement we now recall. Upon choosing a projective embedding for X, we may find a generic hyperplane whose intersection L with X is smooth of codimension 1. As such, L produces a cycle class $\ell \in \mathrm{H}^2_{\mathbb{A}}(X)(1)$. Then the Hard Lefschetz theorem asserts that the cup-product map

$$\ell^i \colon \mathrm{H}^{d-i}_{\mathbb{A}}(X) \to \mathrm{H}^{d+i}_{\mathbb{A}}(X)(i)$$

is an isomorphism for all $i \leq d$. As an application, we are able to deduce the Lefschetz decomposition: note that ℓ^i being an isomorphism implies that $\ell^{i+1} \colon \mathrm{H}^{d-i}_{\mathbb{A}}(X) \to \mathrm{H}^{d+i+2}_{\mathbb{A}}(X)(i+1)$ is the first time one can see a kernel, so we define the primitive cohomology

$$\mathrm{H}^{d-i}_{\mathbb{A}}(X)_{\mathrm{prim}} \coloneqq \ker\left(\ell^{i+1} \colon \mathrm{H}^{d-1}_{\mathbb{A}}(X) \to \mathrm{H}^{d+i+2}_{\mathbb{A}}(X)(i+1)\right)$$

as precisely this kernel. We now claim that

$$\mathrm{H}^{d-i}_{\mathbb{A}}(X) \overset{?}{=} \mathrm{H}^{d-i}_{\mathbb{A}}(X)_{\mathrm{prim}} \oplus \ell \mathrm{H}^{d-i-2}_{\mathbb{A}}(X)(-1)$$

for each $i \leq d$. Indeed, note the left-exact sequence

$$0 \to \mathrm{H}^{d-i}_{\mathbb{A}}(X)_{\mathrm{prim}} \to \mathrm{H}^{d-i}_{\mathbb{A}}(X) \to \mathrm{H}^{d+i+1}_{\mathbb{A}}(X)(i+1)$$

in fact is surjective on the right due to the Hard Lefschetz theorem providing a splitting map

$$\mathrm{H}^{d+i+1}_{\mathbb{A}}(X)(i+1) \xleftarrow{\sim} \mathrm{H}^{d-i-1}_{\mathbb{A}}(X)(-1) \overset{\ell}{\subseteq} \mathrm{H}^{d-i}_{\mathbb{A}}(X).$$

Applying our claim inductively reveals that

$$\mathrm{H}^{d-i}_{\mathbb{A}}(X) = \bigoplus_{j \geq 0} \ell^j \mathrm{H}^{d-i-2j}_{\mathbb{A}}(X)_{\mathrm{prim}}(-j)$$

for each $i \leq d$. Applying the Hard Lefschetz theorem once more grants the equality

$$\mathrm{H}^{d+i}_{\mathbb{A}}(X) = \bigoplus_{j \geq 0} \ell^{i+j} \mathrm{H}^{d-i-2j}_{\mathbb{A}}(X)_{\mathrm{prim}}(-j),$$

but we can synthesize the prior two assertions into the single Lefschetz decomposition

$$\mathrm{H}^{i}_{\mathbb{A}}(X) = \bigoplus_{\substack{j \geq 0 \\ i-2j \leq d}} \ell^{j} \mathrm{H}^{i-2j}_{\mathbb{A}}(X)_{\mathrm{prim}}(-j).$$

We are now ready to define our operator *: for $x \in \mathrm{H}^i_{\mathbb{A}}(X)$, this Lefschetz decomposition lets us expand $x = \sum_j \ell^j x_j$ for $x_j \in \mathrm{H}^{i-2j}_{\mathbb{A}}(X)_{\mathrm{prim}}(-j)$, and then we define

$$^*x \coloneqq \sum_{\substack{j \geq 0 \\ i-2j \leq d}} (-1)^{(i-2j)(i-2j+1)/2} \ell^{d-i+j} x_j$$

so that $^*x \in \mathrm{H}^{2d-i}_{\mathbb{A}}(X)(d-i)$. This operator * is defined compatibly for all of our cohomology theories, so it produces an absolute Hodge correspondence and so comes from an absolute Hodge class by Lemma 1.197. Additionally, we see that * merely rearranges the Lefschetz decomposition up to a sign, so it is an isomorphism. ■

Remark 1.201. The Hodge–Riemann relations [GH94, p. 123] show that the induced composite

$$\mathrm{H}^i_\sigma(X) \otimes \mathrm{H}^i_\sigma(X) \to \mathrm{H}^i_\sigma(X) \otimes \mathrm{H}^{2d-i}_\sigma(X)(d-i) \to \mathrm{H}^0_\sigma(\mathrm{pt})(-i)$$

is a polarization of Hodge structures. We remark that one can sum this polarization over different Xs, so its existence (coming from an absolute Hodge class) no longer requires that X is equidimensional.

We now repeat the story of the previous section to construct a category of motives from absolute Hodge classes. Let's take a moment to quickly review the constructions.

- Pullbacks: any $\gamma \in \mathrm{Corr}_{\mathrm{AH}}(X, Y)$ gives rise to a morphism $\gamma^* \colon C_{\mathrm{AH}}(Y) \to C_{\mathrm{AH}}(X)$ given by

  $$\gamma^*(\beta) \coloneqq \mathrm{pr}_{1*}(\gamma \cup \mathrm{pr}_2^* \beta),$$

  where we are using the $\cup$ product structure which exists on $\mathrm{H}^\bullet_{\mathbb{A}}$.

- Composition: any $\gamma \in \mathrm{Corr}_{\mathrm{AH}}(X, Y)$ and $\delta \in \mathrm{Corr}_{\mathrm{AH}}(Y, Z)$ can be composed via

  $$\delta \circ \gamma \coloneqq \mathrm{pr}_{13*}\left(\mathrm{pr}_{12}^* \gamma \cup \mathrm{pr}_{23}^* \delta\right).$$

  The exact same proof as in Lemma 1.161 (replacing the use of the projection formula with Lemma 1.100 and base-change with base-change formulae in our cohomology theories) establishes $\mathbb{Q}$-linearity and associativity of $\circ$ and that $[\Gamma_f^\intercal] \circ [\Gamma_g^\intercal] = [\Gamma_{g \circ f}^\intercal]$. The same calculation aas in Remark 1.165 shows that $\circ$ is in fact a morphism of $\mathbb{Z}$-graded groups.

  While we're here, we note that $(\delta \circ \gamma)^* = \gamma^* \circ \delta^*$ allows one to see that we may as well just compose the corresponding absolute Hodge correspondences.

- We may now define a category $\mathcal{C}_{\mathrm{AH}}(K)$ whose objects are given by $h(X)$ for $X \in \mathcal{P}(K)$ and morphisms given by correspondences in degree 0. Then we still have a faithful, essentially surjective, additive functor $h \colon \mathcal{P}(K)^{\mathrm{op}} \to \mathcal{C}_{\mathrm{AH}}(K)$. The same arguments as in Lemmas 1.166 and 1.167 show that $\mathcal{C}_{\mathrm{AH}}(K)$ is additive (with $h(X) \times h(Y) = h(X \sqcup Y)$) and symmetric monoidal (with $h(X) \otimes h(Y) = h(X \otimes Y)$).

- We are now ready to define the category of effective motives as $\mathrm{Mot}^+_{\mathbb{Q}}(K) \coloneqq \mathrm{Split}(\mathcal{C}_{\mathrm{AH}}(K))$, which is now also Karoubian. For example, one can use the idempotents π_i from Example 1.198 to define

  $$h^i(X) \coloneqq (h(X), \pi_i).$$

- Lastly, by adding in Tate twists, we may define the category of motives $\mathrm{Mot}_{\mathbb{Q}}(K)$ as the category of triples (X, p, i) where $X \in \mathcal{P}(K)$ and $p \in \mathrm{Corr}^0_{\mathrm{AH}}(X, X)$ is an idempotent and $i \in \mathbb{Z}$. Here, morphisms are given by
$$\mathrm{Hom}_{\mathrm{Mot}_{\mathbb{Q}}(K)}((X, p, i), (Y, q, j)) \coloneqq q \circ \mathrm{Corr}^{j-i}_{\mathrm{AH}}(X, Y) \circ p.$$
This category is still $\mathbb{Q}$-linear, and the argument of Remark 1.183 shows that it is still Karoubian. We continue to set $\mathrm{T} \coloneqq (\mathrm{pt}, \mathrm{id}, 1)$ and $\mathrm{L} \coloneqq (\mathrm{pt}, \mathrm{id}, -1)$ to be the Tate and Lefschetz motives respectively, and we remark that the exact same argument as in Example 1.187 shows that L is an effective motive. As such, the argument of Lemma 1.188 verifies that $\mathrm{Mot}_{\mathbb{Q}}(K)$ is additive.

Remark 1.202. Later on, it will be useful to note that any embedding $K \subseteq K'$ of fields gives rise to a fully faithful base-change functor $\mathrm{Mot}_{\mathbb{Q}}(K) \to \mathrm{Mot}_{\mathbb{Q}}(K')$. To check that this functor is fully faithful, we are implicitly using Remark 1.192: we need to know that extending K does not actually affect the rational subspace of absolute Hodge classes. By construction, we can also see that this functor is linear, and it will preserve the symmetric monoidal structure of Proposition 1.209 once we get there.

Our present goal is to show that $\mathrm{Mot}_{\mathbb{Q}}(K)$ is a neutral Tannakian category, for which we will use Theorem 1.135; later, we will also want to place $\mathrm{Mot}_{\mathbb{Q}}(K)$ in a Tate triple. Let's begin by showing that $\mathrm{Mot}_{\mathbb{Q}}(K)$ is semisimple abelian. Here is a general test which explains how to do this upgrading.

Lemma 1.203. Let $\mathcal{C}$ be a $\mathbb{Q}$-linear, additive, Karoubian category. Suppose that $\mathrm{End}_{\mathcal{C}}(X)$ is a finite-dimensional semisimple algebra for all $X \in \mathcal{C}$. Then $\mathcal{C}$ is a semisimple abelian category.

Proof. This is [Jan92, Lemma 2]. We proceed in steps.

1. We note that any object $X \in \mathcal{C}$ is a sum of finitely many indecomposable objects. Indeed, $\mathrm{End}_{\mathcal{C}}(X)$ is a semisimple algebra, so Wedderburn's theorem allows us to write it as a product
$$\mathrm{End}_{\mathcal{C}}(X) \cong M_{n_1}(A_1) \times \cdots \times M_{n_k}(A_k)$$
of matrix algebras over division algebras. Expanding $\mathrm{End}_{\mathcal{C}}(X)$ out as a product like this produces an idempotent decomposition of id_X, so Remark 1.169 (recall $\mathcal{C}$ is Karoubian!) shows
$$X \cong X_1 \oplus \cdots \oplus X_k,$$
where $X_\bullet$ is the image of the idempotent in $\mathrm{End}_{\mathcal{C}}(X)$ which corresponds to the identity in $M_{n_\bullet}(A_\bullet)$; in particular, $\mathrm{End}_{\mathcal{C}}(X_\bullet) = M_{n_\bullet}(A_\bullet)$. (We can see this on the level of the construction of $\mathrm{Split}(\mathcal{C})$, which must be canonically equivalent to $\mathcal{C}$.) Next, we let $Y_\bullet$ be the projection of $X_\bullet$ along the idempotent in $M_{n_\bullet}(A_\bullet)$ which is the elementary matrix E_{11}. The idempotent decomposition $1_{n_\bullet} = E_{11} + \cdots + E_{n_\bullet n_\bullet}$ can be plugged into Remark 1.169 to show
$$X_\bullet \cong Y_\bullet^{n_\bullet}.$$
We now have $\mathrm{End}_{\mathcal{C}}(Y_\bullet) = A_\bullet$. Because $A_\bullet$ is a division algebra, it has no idempotents other than 0 and 1, so $Y_\bullet$ must be indecomposable.

2. The main claim is that $X \cong Y$ if and only if $\mathrm{Hom}_{\mathcal{C}}(X, Y) \neq 0$ for any indecomposable $X, Y \in \mathcal{C}$. Let's quickly explain why the main claim implies the result.

   - We check that every morphism has a kernel and cokernel. Using the previous step, we may suppose that our morphism f is between the objects $\bigoplus_{i=1}^n X_i^{\oplus k_i}$ and $\bigoplus_{i=1}^n X_i^{\oplus \ell_i}$ for some indecomposables $X_\bullet$ and sequences $k_\bullet$ and $\ell_\bullet$ of nonnegative integers. But the hypothesis implies that the different indecomposables have no interaction with each other, so
$$\mathrm{Hom}_{\mathcal{C}}\left(\bigoplus_{i=1}^n X_i^{\oplus k_i}, \bigoplus_{i=1}^n X_i^{\oplus \ell_i}\right) = \bigoplus_{i=1}^n M_{\ell_i \times k_i}(\mathrm{End}_{\mathcal{C}} X_i),$$

so we can realize f as an n-tuple of matrices over division algebras. Doing some row-reduction (which amounts to changing bases of the $X_i^{\oplus k_i}$s and $X_i^{\oplus \ell_i}$s) lets us put the matrix form of f into a row-reduced Echelon form, from which one can read off a kernel and cokernel for f as one does for vector spaces.

- We check that every monomorphism is a kernel; the check that every epimorphism is a cokernel is essentially the same. As in the previous point, we may write our morphism f as some map $f\colon \bigoplus_{i=1}^n X_i^{\oplus k_i} \to \bigoplus_{i=1}^n X_i^{\oplus \ell_i}$ in a matrix form, which we may put into row-reduced Echelon form. Note then that all diagonal entries of all matrices must be nonzero, for otherwise f has a nontrivial kernel, so f will fail to be a monomorphism. It follows from the row-reduced Echelon form that $k_i \leq \ell_i$ for each i, and f is simply embedding $X_i^{\oplus k_i}$ into the first k_i coordinates of $X_i^{\oplus \ell_i}$. In particular, f will then be the kernel of projection

$$\bigoplus_{i=1}^n X_i^{\oplus \ell_i} \twoheadrightarrow \bigoplus_{i=1}^n X_i^{\oplus (\ell_i - k_i)}$$

  away from these coordinates.
- We check that $\mathcal{C}$ is semisimple. By the previous step, it is enough to check that every indecomposable object $X \in \mathcal{C}$ is actually simple. Well, any nontrivial map $X' \to X$ must have quotient 0. Indeed, after decomposing X' into indecomposables, we may assume that X' is indecomposable. But now the main claim implies $X' \cong X$, so because $\mathrm{End}_{\mathcal{C}}(X)$ is a division algebra (by the previous step) and so the map $X' \to X$ is an isomorphism.

3. It remains to prove the main claim. Certainly $X \cong Y$ implies $\mathrm{Hom}_{\mathcal{C}}(X,Y) \neq 0$, so we merely must show the converse. As such, suppose that $\mathrm{Hom}_{\mathcal{C}}(X,Y) \neq 0$. Observe that we will be done as soon as we know that there are $f\colon X \to Y$ and $g\colon Y \to X$ such that $gf \neq 0$ or $fg \neq 0$; take $gf \neq 0$ because the other case is similar. Well, because X is indecomposable, $\mathrm{End}_{\mathcal{C}}(X)$ is a division algebra (see the first step), so $gf \in \mathrm{End}\,\mathcal{C}(X)$ has an inverse, so $f\colon X \to Y$ has a left inverse given by $g' \coloneqq (gf)^{-1}g$. Thus, Example 1.174 tells us that Y decomposes into $X = \mathrm{im}\, fg'$ plus another object $\mathrm{im}(1 - fg')$, but then $X \cong Y$ is forced because Y is indecomposable.

   It remains to show that such $f\colon X \to Y$ and $g\colon Y \to X$ exist. This will require a trick. As in the first step, we may view $\mathrm{End}_{\mathcal{C}}(X \oplus Y)$ as some algebra 2×2 matrices

$$\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a \in \mathrm{End}_{\mathcal{C}}(X), b \in \mathrm{Hom}_{\mathcal{C}}(Y,X), c \in \mathrm{Hom}_{\mathcal{C}}(X,Y), \mathrm{Hom}_{\mathcal{C}}(Y,Y) \right\}.$$

   Now, consider the subgroup

$$N \coloneqq \left\{ \begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix} : c \in \mathrm{Hom}_{\mathcal{C}}(X,Y) \right\}.$$

   This subgroup N is nonzero and nilpotent, so because $\mathrm{End}_{\mathcal{C}}(X \oplus Y)$, it cannot be an ideal! Thus, we must be able to find morphisms such that

$$\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ c_2 & 0 \end{bmatrix} \begin{bmatrix} a_3 & b_3 \\ c_3 & d_3 \end{bmatrix} \notin N$$

   A quick calculation shows that this matrix is $\left[\begin{smallmatrix} b_1c_2a_3 & b_1c_2b_3 \\ d_1c_1a_3 & d_1c_2b_3 \end{smallmatrix}\right]$, so $b_1c_2 \neq 0$ or $c_2b_3 \neq 0$, as needed. ∎

Remark 1.204. We needed to assume that $\mathcal{C}$ was additive in order to be able to write down the sum $X \oplus Y$. This seems to be the only place where we need to use the existence of arbitrary finite sums.

Thus, we would like to check that $\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M)$ is a finite-dimensional semisimple algebra for each $M \in \mathrm{Mot}_{\mathbb{Q}}(K)$. Finite-dimensionality is easy.

Lemma 1.205. Fix a field K algebraic over $\mathbb{Q}$. For any $M \in \mathrm{Mot}_{\mathbb{Q}}(K)$, we have

$$\dim_{\mathbb{Q}} \mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M) < \infty.$$

Proof. Write $M = (X, p, i)$, and then

$$\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M) \subseteq \mathrm{Corr}^0_{\mathrm{AH}}(X, X)$$

by construction, so we are reduced to checking that $\dim_{\mathbb{Q}} C_{\mathrm{AH}}(X) < \infty$ for any $X \in \mathcal{P}(K)$. Well, for any fixed index i and embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, the space $C_{\mathrm{AH}}(X)$ is contained in the image of $\mathrm{H}^{2i}_{\mathrm{B}}(X)(i)$ in $\mathrm{H}^{2i}_{\mathbb{A}}(X)(i)$, and $\dim_{\mathbb{Q}} \mathrm{H}^{2i}_{\mathrm{B}}(X)(i) < \infty$ by properties of $\mathrm{H}^\bullet_{\mathrm{B}}$. ∎

To check that $\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M)$ is semisimple will require a trick: we will use polarizations.

Lemma 1.206. Fix a $\mathbb{Q}$-algebra A. Suppose that there is an involution $(\cdot)^\dagger\colon A^{\mathrm{op}} \to A$ such that $aa^\dagger \neq 0$ for all nonzero $a \in A$. Then A is semisimple.

Proof. We will show that any nonzero two-sided ideal $I \subseteq A$ fails to be nilpotent. Define the function $N\colon (I \setminus \{0\}) \to (I \setminus \{0\})$ by

$$N(a) \coloneqq aa^\dagger$$

We are given that N is well-defined. Note that $N(a)^\dagger = N(a)$ for each a, so N becomes squaring on its image. We conclude that all iterated squares of any $b \in \mathrm{im}\, N$ continue to be nonzero, so $\mathrm{im}\, N \subseteq I \setminus \{0\}$ fails to be nilpotent. ∎

Lemma 1.207. Fix a field K algebraic over $\mathbb{Q}$. For any $M \in \mathrm{Mot}_{\mathbb{Q}}(K)$, the algebra $\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M)$ is semisimple.

Proof. We proceed in steps.

1. We reduce to the case of M of the form $h(X)$. Indeed, we may write $M = (X, p, i)$, from which we find that
$$\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M) = p \circ \mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(h(X)) \circ p.$$
Now, if we know that $\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(h(X))$ is semisimple, we may use Wedderburn's theorem (finite-dimensionality follows from Lemma 1.205) to write it as a product
$$\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(h(X)) = M_{n_1}(A_1) \times \cdots \times M_{n_k}(A_k)$$
of matrix algebras of division algebras. Our idempotent p can now be viewed as some tuple of idempotent matrices in the $M_{n_\bullet}(A_\bullet)$s. After base-changing from $\mathbb{Q}$ to $\mathbb{C}$, we see that each of these matrices can be upper-triangularized and is thus diagonalizable with eigenvalues in $\{0, 1\}$ because p is an idempotent; by searching for this eigenbasis over $\mathbb{Q}$, we see that p is still diagonalizable over $\mathbb{Q}$. It follows that $\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M)$ is isomorphic to a product of submatrix algebras from the given product, so it continues to be semisimple.

2. We show that $\mathrm{End}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(h(X)) = \mathrm{Corr}^0(X, X)$ is semisimple. We will use Lemma 1.206. For each i, let ψ_i be the polarization of $\mathrm{H}^i_\sigma(X)$ defined in Remark 1.201 by using the Hodge involution and Poincaré duality. Polarizations are perfect pairings, so any $\gamma \in \mathrm{Corr}^0_{\mathrm{AH}}(X, X)$ induces a pullback map $\gamma^* \in \mathrm{End}(\mathrm{H}^\bullet_{\mathbb{A}}(X))$, which then must have a unique transpose map $(\gamma^\dagger)^* \in \mathrm{End}(\mathrm{H}^\bullet_{\mathbb{A}}(X))$ satisfying
$$\psi_i(\gamma^* \alpha_i, \beta_i) = \psi_i\left(\alpha_i, (\gamma^\dagger)^* \beta_i\right)$$

for any $i \in \mathbb{Z}$ and $\alpha_i, \beta_i \in \mathrm{H}^i_{\mathbb{A}}(X)$. The uniqueness (plugged into Lemma 1.197) shows that $(\gamma^\dagger)^*$ arises rationally and is compatible with all of our cohomology theories, so it comes from an element in $\mathrm{Corr}^0_{\mathrm{AH}}(X, X)$.

The ambient uniqueness shows that $\gamma \mapsto \gamma^\dagger$ is $\mathbb{Q}$-linear, involutive, and we can see that $(\gamma\delta)^\dagger = \delta^\dagger\gamma^\dagger$ by a computation with the uniqueness. To apply Lemma 1.206, it remains to check that $\gamma\gamma^\dagger \neq 0$ for each nonzero γ. It is enough to find $\alpha, \beta \in \mathrm{H}^\bullet_{\mathbb{A}}(X)$ such that

$$\psi\left(\alpha, (\gamma\gamma^\dagger)^*\beta\right) = \psi\left(\gamma^*\alpha, \gamma^*\beta\right)$$

is nonzero. It is enough to check this on the de Rham component where ψ becomes a polarization, and then we may as well base-change everything from $\mathbb{Q}$ to $\mathbb{R}$. In particular, we may take $\alpha \neq 0$ and $\beta \coloneqq \sqrt{-1}\alpha$ (where $\sqrt{-1}$ acts on $\mathrm{H}_\sigma(X)_{\mathbb{R}}$ via the Hodge structure), so the fact that γ^*_{dR} is a morphism of Hodge structures shows that the above value will be positive by the positive-definiteness of ψ. ■

Proposition 1.208. The category $\mathrm{Mot}_{\mathbb{Q}}(K)$ is a $\mathbb{Q}$-linear, semisimple, abelian category.

Proof. The category $\mathrm{Mot}_{\mathbb{Q}}(K)$ is already $\mathbb{Q}$-linear, additive, and Karoubian essentially by its construction, so we may plug Lemmas 1.205 and 1.207 into Lemma 1.203. ■

We have completed our first major check leading up to the application of Theorem 1.135 showing that $\mathrm{Mot}_{\mathbb{Q}}(K)$ is a neutral Tannakian category. Next up, we will show that $\mathrm{Mot}_{\mathbb{Q}}(K)$ has a symmetric monoidal structure.

Proposition 1.209. Fix a field K algebraic over $\mathbb{Q}$. The category $\mathrm{Mot}_{\mathbb{Q}}(K)$ has a symmetric monoidal structure.

Proof. Repeating the proof of Lemma 1.184, we may simply define

$$(X, p, i) \otimes (Y, q, j) \coloneqq (X \times Y, p \times q, i + j).$$

For example, we can see that the unit should be given by $(\mathrm{pt}, \mathrm{id}, 0)$. The associativity coherence will be induced by the associativity of the fiber product (and addition in $\mathbb{Z}$), but Remark 1.185 explains that we should be slightly careful with the commutativity coherence. Because we have Künneth projectors (Example 1.198), we may expand

$$ph(X)(i) = \bigoplus_n ph^n(X)(i) \qquad \text{and} \qquad qh(Y)(j) = \bigoplus_m qh^m(Y)(j),$$

so we define the commutativity constraint $(X, p, i) \otimes (Y, q, j) \to (Y, q, j) \otimes (X, p, i)$ to be the obvious signs multiplied by the sign $(-1)^{mn}$ on each of the above graded pieces. ■

And let's complete the proof.

Theorem 1.210. Fix a field K algebraic over $\mathbb{Q}$. The category $\mathrm{Mot}_{\mathbb{Q}}(K)$ is neutral Tannakian. In fact, for each embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, the Betti cohomology functor $\mathrm{H}^\bullet_\sigma$ induces a fiber functor ω_σ.

Proof. We use Theorem 1.135 with $\omega = \mathrm{H}^\bullet_\sigma$. Explicitly, $\mathrm{H}^\bullet_\sigma$ is extended to $\mathrm{Mot}_{\mathbb{Q}}(K)$ by

$$\omega^\bullet_\sigma((X, p, i)) \coloneqq p_\sigma \mathrm{H}^\bullet_\sigma(X)(i),$$

where the notation p_σ comes from viewing p as an absolute Hodge correspondence via Lemma 1.197. Functoriality for absolute Hodge correspondences grants functoriality for $\mathrm{H}^\bullet_\sigma$.[6]

[6] Formally, one ought to appeal to Lemma 1.170 and then explain functoriality with the Tate twist by hand. We will not bother.

Proposition 1.208 has shown that $\mathrm{Mot}_{\mathbb{Q}}(K)$ is already a $\mathbb{Q}$-linear abelian and semisimple category, and Proposition 1.209 gives it the structure of a symmetric monoidal category. Continuing, we note that the functor $\mathrm{H}^\bullet_\sigma$ is certainly $\mathbb{Q}$-linear and faithful (see Lemma 1.197), and $\mathrm{H}^\bullet_\sigma$ is exact because $\mathrm{Mot}_{\mathbb{Q}}(K)$ is already semisimple and $\mathrm{H}^\bullet_\sigma$ preserves sums because it is additive.

We now must check (i)–(iv) of Theorem 1.135. For (i), the Künneth formula explains why $\mathrm{H}^\bullet_\sigma$ preserves products. For (ii), the construction of the symmetric monoidal structure explains that $\mathrm{H}^\bullet_\sigma$ successfully preserves the commutativity and associativity constraints; we refer to Remark 1.185 to explain why $\mathrm{GrAlg}_{\mathbb{Q}}$ requires the sign in the commutativity constraint. Additionally, for (iii), we note $\mathrm{H}^\bullet_\sigma(\mathrm{pt}) = \mathbb{Q}$, and one can check that the unit constraints are all preserved by $\mathrm{H}^\bullet_\sigma$ because they are all given by the canonical isomorphism $\mathrm{pr}_1 \colon X \times \mathrm{pt} \to X$.

Lastly, for (iv), it remains to understand the objects $(X,p,i) \in \mathrm{Mot}_{\mathbb{Q}}(K)$ such that $\dim_{\mathbb{Q}} \mathrm{H}^\bullet_\sigma((X,p,i)) = 1$. We may as well assume that $i = 0$ because it will not affect the dimension, and $(X,p,0)$ admits an inverse if and only if $(X,p,i) = (X,p,0) \otimes \mathrm{T}^{\otimes i}$ admits an inverse. Upon decomposing X into equidimensional pieces as $X = \bigsqcup_d X_d$ where X_d is equidimensional of dimension d, we see that Poincaré duality (via Example 1.199) gives a morphism

$$h(X) \otimes \underbrace{\bigoplus_{d \geq 0} h(X_d)(d)}_{M' :=} \to \mathrm{pt}$$

which produces the Poincaré duality pairing upon applying $\mathrm{H}^\bullet_\sigma$ (or $\mathrm{H}^\bullet_{\mathbb{A}}$). Now, setting $q := 1 - p$ allows a decomposition $h(X) = ph(X) \oplus qh(X)$. Letting p' and q' be the dual maps (on $\mathrm{H}^\bullet_{\mathbb{A}}$ or $\mathrm{H}^\bullet_\sigma$s) via Poincaré duality, we see that they produce absolute Hodge correspondences by the coherences, so we receive a dual decomposition $M' = p'M' \oplus q'M'$. Namely, the induced map $ph(X) \otimes p'M' \to \mathrm{pt}$ will induce a perfect pairing

$$p_\sigma \mathrm{H}^\bullet_\sigma(X) \otimes p'_\sigma \mathrm{H}^\bullet_\sigma(M') \to \mathrm{H}^0_\sigma(\mathrm{pt}).$$

For example, this implies that $\dim_{\mathbb{Q}} p'_\sigma \mathrm{H}^\bullet_\sigma(M') = 1$. Lastly, because $\mathrm{H}^\bullet_\sigma$ is faithful, we conclude that the induced map $ph(X) \otimes p'M' \to \mathrm{pt}$ is an isomorphism. This completes the check (iv) of Theorem 1.135 and thus the proof. ∎

Remark 1.211. The fiber functor $\omega_\sigma \colon \mathrm{Mot}_{\mathbb{Q}}(K)$ in fact factors through $\mathrm{HS}_{\mathbb{Q}}$. To begin, note $p_\sigma \mathrm{H}^\bullet_\sigma(X)(i)$ is a rational Hodge structure because $\mathrm{H}^\bullet_\sigma(X)$ and T are, and p_σ is an endomorphism of Hodge structures (because p_{dR} is by Lemma 1.197). Furthermore, any morphism $f \colon (X,p,i) \to (Y,q,j)$ of motives arises from an absolute Hodge correspondence, which does induce a morphism of rational Hodge structures upon passing through ω_σ because f_{dR} preserves Hodge structuers (by Lemma 1.197).

Remark 1.212. One can repeat this proof for ℓ-adic or de Rham cohomology, provided that we base-change $\mathrm{Mot}_{\mathbb{Q}}(K)$ to the corresponding F-linear category $\mathrm{Mot}_F(K)$, where F is the coefficient field. In particular, each prime ℓ has $\mathrm{H}^\bullet_{\text{ét}}$ induce a fiber functor $\omega_\ell \colon \mathrm{Mot}_{\mathbb{Q}_\ell}(K) \to \mathrm{Vec}_{\mathbb{Q}_\ell}$. But now, ω_ℓ actually factors through $\mathrm{Rep}_{\mathbb{Q}_\ell} \mathrm{Gal}(\overline{K}/K)$: the proof is the same as in Remark 1.211, where the main point is that ℓ-adic cohomology produces Galois representations, and our absolute Hodge correspondences specialize to Galois-invariant maps by their definition.

Remark 1.213. We remark that ω_ℓ is naturally isomorphic to $(\cdot)_{\mathbb{Q}_\ell} \circ \omega_\sigma$. Indeed, this follows from the fact that the comparison isomorphism Theorem 1.79 is an isomorphism of Weil cohomology theories, so we can see (by hand, via the constructions suggested in Theorem 1.210) that the comparison isomorphism induces a natural isomorphism $(\cdot)_{\mathbb{Q}_\ell} \circ \omega_\sigma \Rightarrow \omega_\ell$.

While we're here, we remark that we can upgrade these things to Tate triples.

Corollary 1.214. Fix a field K algebraic over $\mathbb{Q}$. The Künneth decompositions induce a $\mathbb{Z}$-grading w on $\mathrm{Mot}_{\mathbb{Q}}(K)$, thus making $(\mathrm{Mot}_{\mathbb{Q}}(K), w, \mathrm{T})$ into a Tate triple.

Proof. We already know $\mathrm{Mot}_{\mathbb{Q}}(K)$ is neutral Tannakian by Theorem 1.210, and we are going put T in weight -2, so the main content of the argument arises from defining the weight grading. For any effective motive $ph(X) \in \mathrm{Mot}_{\mathbb{Q}}(K)$, we claim that

$$ph(X) \stackrel{?}{=} \bigoplus_{i \in \mathbb{Z}} ph^i(X).$$

Indeed, p is induced by an absolute Hodge correspondence $h(X) \to h(X)$, so p has degree 0, meaning that all the induced maps on cohomology preserve the degree. Thus, the map $\bigoplus_{i \in \mathbb{Z}} ph^i(X) \to ph(X)$ is an isomorphism on each of our cohomology theories, so its inverse also succeeds at being an absolute Hodge correspondence because the uniqueness of the inverse provides the needed compatibility. The equality follows.

Our weight grading is now given by the decomposition

$$ph(X)(n) = \bigoplus_{i \in \mathbb{Z}} ph^{i+2n}(X)(n).$$

(In particular, T sits in weight -2.) Here are the needed checks on this grading.

- Functorial: a morphism $ph(X)(n) \to qh(Y)(m)$ of motives arises from an absolute Hodge correspondence γ of degree $m - n$. Such an absolute Hodge correspondence arises from graded maps $p\mathrm{H}^\bullet(X)(n) \to q\mathrm{H}^\bullet(Y)(m)$ on our cohomology. We conclude that our absolute Hodge correspondences preserve the Künneth projectors (we are implicitly using some functoriality) and thus the gradings.
- Tensor: given two motives $ph(X)(n)$ and $qh(Y)(m)$, their tensor product has been given by

  $$ph(X)(n) \otimes qh(Y)(m) = (p \times q)h(X \times Y)(n+m).$$

  The Künneth isomorphism for our cohomology theories upgrades to an absolute Hodge correspondence by its compatibility, thereby ensuring

  $$ph(X)(n) \otimes qh(Y)(m) = \bigoplus_{i,j} ph^i(X)(n) \otimes qh^j(Y)(m).$$

  Thus, for any k, the degree-k piece on the right-hand side is given by

  $$(ph(X)(n) \otimes qh(Y)(m))_k = \bigoplus_{i+j=k} ph^{i+2n}(X)(n) \otimes qh^{j+2m}(Y)(m),$$

  as required. ∎

Remark 1.215. In fact, for any embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, the functor ω_σ is a morphism of Tate triples $(\mathrm{Mot}_{\mathbb{Q}}(K), w, \mathrm{T}) \to (\mathrm{HS}_{\mathbb{Q}}, w, \mathbb{Q}(1))$. Of course, T goes to $\mathbb{Q}(1)$, so it remains to check that ω_σ preserves the weight gradings. But this is basically by construction: for any motive $ph(X)(n)$, we have

$$p_\sigma \mathrm{H}^\bullet_\sigma(X)(n) = \bigoplus_{i \in \mathbb{Z}} p_\sigma \mathrm{H}^{i-2n}_\sigma(X)(n)$$

because p_σ is a morphism of rational Hodge structures.

Remark 1.216. The category $\operatorname{im} \omega_\ell \subseteq \mathrm{Rep}_{\mathbb{Q}_\ell} \mathrm{Gal}(\overline{K}/K)$ now has an induced weight grading by simply porting over the weight grading from $\mathrm{Mot}_{\mathbb{Q}}(K)$. Noting that $\omega_\ell(\mathrm{T}) = \mathbb{Q}_\ell(1)$ by construction of ω_ℓ, we find that $\omega_\ell\colon \mathrm{Mot}_{\mathbb{Q}}(K) \to \operatorname{im} \omega_\ell$ upgrades to a morphism of Tate triples.

We have thus completed the main content of the present subsection. Of course, even though we have found that $\mathrm{Mot}_{\mathbb{Q}}(K)$ is neutral Tannakian, this does not make it easy to understand; for example, it is highly nonobvious what the corresponding affine group should be. We close this section with the easiest nontrivial subset of this question.

Definition 1.217 (Artin motive)**.** Fix a field K algebraic over $\mathbb{Q}$. The category $\mathrm{Mot}^0_{\mathbb{Q}}(K)$ of *Artin motives* is the full $\otimes$-subcategory

$$\langle h(X) : \dim X = 0 \rangle^{\otimes}.$$

Example 1.218. Fix a field K algebraic over $\mathbb{Q}$. The functor $\mathrm{Mot}^0_{\mathbb{Q}}(K) \to \mathrm{Rep}_{\mathbb{Q}} \mathrm{Gal}(\overline{K}/K)$ defined by extending $h(X) \mapsto \mathrm{Mor}(X(\overline{K}), \mathbb{Q})$ is an equivalence.

Proof. This is [DM12, Proposition 6.17]. For brevity, we will set $G \coloneqq \mathrm{Gal}(\overline{K}/K)$. We proceed in steps.

1. Define the category $\mathcal{C}^0_{\mathrm{AH}}(K)$ as the full subcategory of $\mathcal{C}_{\mathrm{AH}}(K)$ given by 0-dimensional varieties. Let's begin by defining a fully faithful functor $\omega \colon \mathcal{C}^0_{\mathrm{AH}}(K) \to \mathrm{Rep}_{\mathbb{Q}} G$ on objects. Well, for any choice of embedding $\sigma \colon K \hookrightarrow \mathbb{C}$, we note that

   $$\mathrm{H}^\bullet_\sigma(X) = \mathrm{Mor}(X(\overline{K}), \mathbb{Q}),$$

   and this embedding is independent of the choice of σ: we are simply getting a copy of $\mathbb{Q}$ in degree 0 for each geometric point. Note that the right-hand side is a permutation representation of a quotient of G (note $\#X(\overline{K}) < \infty$ because X is proper and zero-dimensional), so this does in fact produce an object in $\mathrm{Rep}_{\mathbb{Q}} G$.

2. We explain why the functor $\omega \colon \mathcal{C}^0_{\mathrm{AH}}(K) \to \mathrm{Rep}_{\mathbb{Q}} G$ is well-defined and fully faithful. Well, for $X, Y \in \mathcal{C}^0_{\mathrm{AH}}(K)$, an absolute Hodge correspondence f in $\mathrm{Corr}^0_{\mathrm{AH}}(X, Y)$ amounts to a special map $\mathrm{H}^\bullet_{\mathbb{A}}(X) \to \mathrm{H}^\bullet_{\mathbb{A}}(Y)$ satisfying some properties and arising from Betti cohomology. By the previous paragraph, arising from Betti cohomology is equivalent to saying that f arises from a linear map

   $$r(f) \colon \mathrm{Mor}(X(\overline{K}), \mathbb{Q}) \to \mathrm{Mor}(Y(\overline{K}), \mathbb{Q}).$$

   As for the extra properties, we note that the de Rham part f_{dR} automatically preserves the relevant Hodge structure because everything is already supported in degree $(0,0)$, and we note that f_ℓ being Galois-invariant is equivalent to $r(f)$ being Galois-invariant. We conclude that r induces an isomorphism

   $$\mathrm{Corr}^0_{\mathrm{AH}}(X, Y) \to \mathrm{Mor}_G \left(\mathrm{Mor}(X(\overline{K}), \mathbb{Q}), \mathrm{Mor}(Y(\overline{K}), \mathbb{Q}) \right).$$

3. Now, $\mathrm{Rep}_{\mathbb{Q}} G$ is Karoubian (indeed, it is abelian), so ω uniquely extends to the Karoubian envelope $\mathrm{Split}\left(\mathcal{C}^0_{\mathrm{AH}}(K)\right)$ of $\mathcal{C}^0_{\mathrm{AH}}(K)$. We claim that the essential image $\mathrm{im}\, h \subseteq \mathrm{Mot}_{\mathbb{Q}}(K)$ of $\mathrm{Split}\left(\mathcal{C}^0_{\mathrm{AH}}(K)\right)$ is exactly $\mathrm{Mot}^0_{\mathbb{Q}}(K)$. For this, we should show that $\mathrm{im}\, h$ is already a right abelian symmetric monoidal subcategory.

   Well, the same argument as in Proposition 1.208 explains that $\mathrm{Split}\left(\mathcal{C}^0_{\mathrm{AH}}(K)\right)$ and hence $\mathrm{im}\, h$ is semisimple abelian. Further, the construction of the symmetric monoidal structure in Proposition 1.209 explains that $\mathrm{im}\, h$ is also closed udner $\otimes$. Lastly, the proof of Theorem 1.210 shows that the dual of $h(X)$ is $h(X)(\dim X) = h(X)$ (with the perfect pairing given by Poincaré duality), so $\mathrm{im}\, h$ is rigid.

4. The previous steps have shown that the fiber functor ω_σ of $\mathrm{Mot}^0_{\mathbb{Q}}(K)$ upgrades to a fully faithful functor $\omega_\sigma \colon \mathrm{Mot}^0_{\mathbb{Q}}(K) \to \mathrm{Rep}_{\mathbb{Q}} G$. It remains to show that this last functor is essentially surjective.

   To begin, we claim that the representation $\mathrm{Mor}(S, \mathbb{Q})$ is in the essential image, for any $S \in \mathrm{FinSet}(G)$. Indeed, Grothendieck's theory of the étale fundamental group establishes that $\pi_1^{\text{ét}}(\mathrm{Spec}\, K) = G$ (essentially reformulating Galois theory), meaning that taking geometric points produces an equivalence of categories from the category of finite étale covers of $\mathrm{Spec}\, K$ to the category $\mathrm{FinSet}(G)$. Namely, there is some smooth projective zero-dimensional scheme X over $\mathrm{Spec}\, K$ such that $X(\overline{K}) \cong S$ as G-sets, implying that

   $$\omega_\sigma(h(X)) \cong \mathrm{Mor}(S, \mathbb{Q}).$$

   Thus, $\mathrm{Mor}(S, \mathbb{Q})$ is in our essential image.

It remains to show that the representations $\mathrm{Mor}(S, \mathbb{Q}) \in \mathrm{Rep}_{\mathbb{Q}} G$ generate the category. Indeed, any representation V of G has an open stabilizer $H \subseteq G$, so V descends to a representaiton of G/H. But G/H is a finite group, so $\mathrm{Rep}_{\mathbb{Q}} G/H$ is generated by the regular representation, which is a permutation representation, thereby completing the proof; explicitly, we have $V \in \langle \mathbb{Q}[G/H] \rangle^{\otimes}$. ■

Chapter 2
Abelian Varieties

Hold tight to your geometric motivation as you learn the formal structures which have proved to be so effective in studying fundamental questions

—Ravi Vakil [Vak23]

In this chapter, we gather together all the results about abelian varieties we need. Many of the results in the earlier sections discussed here can be found in any reasonable text on abelian varieties such as [Mum74; Mil08; EGM]. Results in the later sections are more specialized, and we will provide references when appropriate. Ultimately, our goal is to define ℓ-adic monodromy groups, explain why one might care about them, and indicate how one might compute them.

2.1 Definitions and Constructions

In this section, we set up the theory of abelian varieties rather quickly. We will usually only indicate proofs that work in the complex analytic situation because the general theory usually requires intricate algebraic geometry.

2.1.1 Starting Notions

Let's begin with a definition.

Definition 2.1 (abelian variety)**.** Fix a ground scheme S. An *abelian scheme* A over S is a smooth projective geometrically integral group scheme over S. An *abelian variety* A is an abelian scheme over a field.

Remark 2.2. Throughout, we will work with abelian varieties instead of abelian schemes as much as possible. However, one should be aware that many of the results generalize.

Here, a group variety refers to a group object in the category of varieties over K.

Remark 2.3. With quite a bit of work, one can weaken the hypotheses of being an abelian variety quite significantly. For example, arguments involving group varieties are able to show that being connected and geometrically reduced implies geometrically integral, and it is a theorem that one can replace projectivity with properness. See [SP, Remark 0H2U] for details.

Here are the starting examples.

Example 2.4 (elliptic curves)**.** Any (smooth) cubic equation cuts out a genus-1 curve in $\mathbb{P}^2$. If the curve has points defined over K, this defines an elliptic curve, which can be shown to be an abelian variety. The interesting part comes from defining the group structure. One way to do this is to show that the map $E \to \mathrm{Pic}^0_{E/K}$ given by $x \mapsto [x] - [\infty]$ is an isomorphism of schemes and then give E the group structure induced by $\mathrm{Pic}^0_{E/K}$. (Here, $\mathrm{Pic}^0_{E/K}$ is the moduli space of line bundles over E of degree 0. Smoothness of the curve makes this in bijection with divisors of degree 0.)

Example 2.5. Fix a positive integer $g \geq 0$. If $\Lambda \subseteq \mathbb{C}^g$ is a polarizable sublattice, then $\mathbb{C}^g/\Lambda$ defines an abelian variety over $\mathbb{C}$. Here, polarizable means that there is an alternating map $\varphi\colon \Lambda \times \Lambda \to \mathbb{Z}$ such that the pairing

$$\langle x, y\rangle \coloneqq \psi_{\mathbb{R}}(x, iy)$$

on $\Lambda_{\mathbb{R}}$ is symmetric and positive-definite. (As worked out in [Mil20b, Section I.2], this is equivalent data to a polarization on the Hodge structure $\Lambda = \mathrm{H}^{\mathrm{B}}_1(A, \mathbb{Z})$.) The requirement of polarizability is used to show that the quotient $\mathbb{C}^g/\Lambda$ is actually projective; see [Mum74, Section 3, Theorem].

It is notable that we have not required our abelian varieties A to actually be abelian even though (notably) both examples above are abelian. Indeed, abelian varieties are always abelian groups, which follows from an argument using the Rigidity theorem. We will not give this argument in full because we will not use it, but we state a useful corollary.

Proposition 2.6. Let $\varphi\colon A \to B$ be a smooth map of abelian varieties over a field K. Then φ is the composition of a homomorphism and a translation.

Proof. By composing with a translation, we may assume that $\varphi(0) = 0$. Then one applies the Rigidity theorem to the map $\widetilde{\varphi}\colon A \times A \to B$ defined by

$$\widetilde{\varphi}(a, a') \coloneqq \varphi(a + a') - \varphi(a) - \varphi(a')$$

to find that $\widetilde{\varphi}$ is constantly 0, completing the proof. See [Mil08, Corollary I.1.2] for details. ∎

Corollary 2.7. The group law on an abelian variety A is commutative.

Proof. The inversion map $i\colon A \to A$ on an abelian variety sends the identity to itself, so Proposition 2.6 tells us that i must be a homomorphism. It follows that the group law is commutative. ∎

In particular, we find that morphisms between abelian varieties are rather strutured: we are allowed to basically only ever consider homomorphisms!

It will turn out that considering abelian varieties up to isomorphism is too strong for most purposes, so we introduce the following definition.

Definition 2.8 (isogeny)**.** A morphism $\varphi\colon A \to B$ of abelian varieties over a field K is an *isogeny* if and only if it is a homomorphism satisfying any one of the following equivalent conditions.

(a) φ is surjective with finite kernel.

(b) $\dim A = \dim B$, and φ is surjective.

(c) $\dim A = \dim B$, and φ has finite kernel.

(d) φ is finite, flat, and surjective.

The *degree* of the isogeny is $\#\ker\varphi$ (thought of as a group scheme).

Remark 2.9. Let's briefly indicate why (a)–(d) above are equivalent; see [Mil08, Proposition 7.1] for details. A spreading out argument combined with the homogeneity of abelian varieties implies that

$$\dim B = \dim A + \dim \varphi^{-1}(\{b\})$$

for any b in the image of φ; this gives the equivalence of (a)–(c). Of course (d) implies (a) (one only needs the finiteness and surjectivity); to show (a) implies (d), we note flatness follows by "miracle flatness" because all fibers have equal dimension, and finiteness follows because finite kernel upgrades to quasi-finiteness.

Intuitively, an isogeny is a "squishy isomorphism."

Example 2.10. Any dominant morphism of elliptic curves sending the identity to the identity is an isogeny.

Example 2.11. In the complex analytic setting, an isogeny of two abelian varieties $A = \mathbb{C}^g/\Lambda$ and $B = \mathbb{C}^g/\Lambda'$ amounts (up to change of basis) an inclusion of lattices $\Lambda' \subseteq \Lambda$.

Example 2.12. Fix any abelian variety A. For any nonzero integer n, the multiplication-by-n endomorphism $[n]_A\colon A \to A$ is an isogeny. To see this, note that it is enough to check that $A[n] \coloneqq \ker[n]_A$ is finite. In the complex analytic situation where $A = \mathbb{C}^g/\Lambda$, this follows because $\frac{1}{n}\Lambda/\Lambda$ is finite; in general, one must show that $A[n] \coloneqq \ker[n]_A$ is zero-dimensional, which is somewhat tricky. See [SP, Lemma 0BFG] for details. We remark that one can compute $\deg[n]_A = d^{2\dim A}$, which is again not so hard to see in the complex analytic situation.

Motivated by the complex analytic setting (and the "squishy isomorphism" intuition), one might hope that one can recover weak-ish inverses for isogenies. This turns into an important property of abelian varieties.

Lemma 2.13. Fix an isogeny $\varphi\colon A \to B$ of abelian varieties of degree d. Then there exists an "inverse isogeny" $\beta\colon B \to A$ such that

$$\begin{cases} \alpha \circ \beta = [d]_B, \\ \beta \circ \alpha = [d]_A. \end{cases}$$

Proof. By some theory regrading group scheme quotients, it is enough to check that φ factors through $[d]_A$, which holds because $\ker\varphi$ has order d as a group scheme and thus vanishes under $[d]_A$. ∎

Remark 2.14. As usual, we remark that the above lemma is easier to see in the complex analytic situation, but the key point of trying to factor through $[d]_A$ remains the same.

Lemma 2.13 motivates the following definition, and it codifies our intuition viewing isogenies as squishy isomorphisms.

Definition 2.15 (isogeny category)**.** Fix a field K. We define the *isogeny category* of abelian varieties over K as having objects which are abelian varieties over K, and a morphism $A \to B$ in the isogeny category is an element of $\mathrm{Hom}_K(A, B)_{\mathbb{Q}}$.

We close our discussion of isogenies with one last remark on the size of kernels.

Remark 2.16. If $\varphi\colon X \to Y$ is a finite separable morphism of varieties, then a spreading out argument shows that the number of geometric points in a general fiber of φ equals the degree of φ. Applied to isogenies, the homogeneity of abelian varieties is able to show that the number of geometric points in the fiber of any separable isogeny equals the degree.

Example 2.17. Here is an application of Remark 2.16: if $\mathrm{char}\, K \nmid n$, then one can show that $A[n]$ has $n^{2\dim A}$ geometric points. Again, this is not so hard to see in the complex analytic setting. The hypothesis $\mathrm{char}\, K \nmid n$ is needed to show that $[n]_A$ is separable; in general, the argument is trickier and can (for example) use some intersection theory [Mil08, Theorem I.7.2].

Now that we have a reasonable category, one can ask for decompositions. Here is the relevant result and definition.

Theorem 2.18 (Poincaré reducibility)**.** Fix an abelian subvariety B of an abelian variety A defined over a field K. Then there is another abelian subvariety $B' \subseteq A$ such that the multiplication map induces an isogeny $B \times B' \to A$.

Proof. As usual, we argue only in the complex analytic case. Here write $A = V/\Lambda$ for complex affine space V, and we find that $B = W/(\Lambda \cap W)$ for some subspace $W \subseteq V$. Now, the polarization induces a Hermitian form on V, so we can define $W' \coloneqq W^{\perp}$ so that $B' \coloneqq W'/(\Lambda \cap W')$ will do the trick. For more details, see [Mil20b, Theorem 2.12] for more details. ∎

Definition 2.19 (simple)**.** Fix a field K. An abelian variety A over K is *simple* if and only if it is irreducible in the isogeny category.

Remark 2.20. Theorem 2.18 implies that any abelian variety can be decomposed uniquely into a product of simple abelian varieties, of course up to isogeny and permutation of factors.

2.1.2 The Jacobian

In this thesis, the abelian varieties of interest to us will be Jacobians. There are a few approaches to their definition, which we will not show are equivalent, but we refer to [Mil08, Chapter III] for details. The most direct definition is as a moduli space.

Definition 2.21 (Jacobian)**.** Fix a smooth proper curve C over a field K such that $C(K)$ is nonempty. Then the *Jacobian* $\mathrm{Jac}\, C$ is the group variety $\mathrm{Pic}^0_{C/K}$, where $\mathrm{Pic}^0_{C/K}$ is the moduli space of line bundles on C with degree 0.

Remark 2.22. We will not check that we have defined an abelian variety, nor that we have even defined a scheme. There are interesting questions regarding the representability of moduli spaces, which we are omitting a discussion of. Milne provides a reasonably direct construction in [Mil08, Section III.1], but we should remark that one expects representability to be true in a broader context. In particular, there are formal ways to check (say) properness of $\mathrm{Pic}^0_{C/K}$, from which it does follow that we have defined an abelian variety.

Remark 2.23. One can actually weaken the smoothness assumption on C to merely being "compact type." This is occasionally helpful when dealing with moduli spaces because it allows us to work a little within the boundary of the moduli space of curves.

Remark 2.24. Notably, Example 2.4 tells us that the Jacobian of a curve is E itself.

Note that the assumption $C(K) \neq \varnothing$ allows us to choose some point $\infty \in C(K)$ and then define a map $C(K) \to \operatorname{Jac} C$ by $p \mapsto [p] - [\infty]$. This map turns out to be a regular closed embedding [Mil08, Proposition 2.3]. It is psychologically grounding to see that this map is universal in some sense.

Proposition 2.25. Fix a smooth proper curve C over a field K such that $C(K) \neq \varnothing$. Choose $\infty \in C(K)$, and consider the map $\iota\colon C \to \operatorname{Jac} C$ given by $\iota(p) \coloneqq [p] - [\infty]$. For any abelian variety A over K and smooth map $\varphi\colon C \to A$ such that $\varphi(\infty) = 0$, there exists a unique map $\widetilde{\varphi}\colon \operatorname{Jac} C \to A$ making the following diagram commute.

$$\begin{array}{ccc} C & \xrightarrow{\iota} & \operatorname{Jac} C \\ & \searrow^{\varphi} & \downarrow{\scriptstyle \widetilde{\varphi}} \\ & & A \end{array}$$

Proof. We will not need this, so we won't even point in a direction of a proof. We refer to [Mil08, Proposition III.6.1]. ∎

It is worthwhile to provide a complex analytic construction of the Jacobian. Given a curve C, line bundles are in bijection with divisor classes, and divisor classes of degree 0 can all be written in the form $\sum_{i=1}^{k}([P_i]-[Q_i])$ for some points $P_1, Q_1, \dots, P_k, Q_k \in C(\mathbb{C})$. One can take such a divisor and define a linear functional on $\mathrm{H}^1(C, \Omega^1_C)$ by

$$\omega \mapsto \sum_{i=1}^{k} \int_{Q_i}^{P_i} \omega.$$

The construction of this linear functional is not technically well-defined up to divisor class; instead, one can check that changing the divisor class adjusts the linear functional exactly by the choice of a cycle in $\mathrm{H}^{\mathrm{B}}_1(C, \mathbb{Z})$ embedded into $\mathrm{H}^1(C, \Omega^1_C)^\vee$ via the integration pairing. In this one way, one finds that

$$\operatorname{Jac} C(\mathbb{C}) = \frac{\mathrm{H}^1(C, \Omega^1_C)^\vee}{\mathrm{H}^{\mathrm{B}}_1(C, \mathbb{Z})}.$$

In particular, we have realized $\operatorname{Jac} C$ explicitly as a complex affine space modulo some lattice, exactly as in Example 2.5. (One sees that $\operatorname{rank}_{\mathbb{Z}} \mathrm{H}^{\mathrm{B}}_1(C, \mathbb{Z}) = \dim_{\mathbb{R}} \mathrm{H}^1(C, \Omega^1_C)^\vee$ by the Betti-to-de Rham comparison isomorphism.) This construction makes it apparent that

$$\mathrm{H}^{\mathrm{B}}_1(\operatorname{Jac} C(\mathbb{C}), \mathbb{Z}) \cong \mathrm{H}^{\mathrm{B}}_1(C, \mathbb{Z}).$$

This is in fact a general property.

Proposition 2.26. Fix a smooth proper curve C over a field K such that $C(K) \neq \varnothing$. Choose $\infty \in C(K)$, and consider the map $\iota\colon C \to \operatorname{Jac} C$ given by $\iota(p) \coloneqq [p] - [\infty]$. Then the induced map

$$\iota^*\colon \mathrm{H}^1(\operatorname{Jac} C) \to \mathrm{H}^1(C)$$

is an isomorphism, where H is any of the Weil cohomology theories of section 1.3.1.

Proof. The proof requires analyzing each cohomology theory individually. Above we outlined the proof when H is Betti cohomology, and we note that the result follows for de Rham cohomology by the comparison isomorphism. ∎

Corollary 2.27. Fix a smooth proper curve C over a field K such that $C(K) \neq \varnothing$. Then $\dim \operatorname{Jac} C$ equals the genus of the curve C.

Proof. Again, this is easy to see in the complex analytic case from the explicit construction. In general, one can read off the dimension of an abelian variety A from $\dim \mathrm{H}^1(A)$ and then apply Proposition 2.26. ∎

2.1.3 The Dual

Even though we will technically not need it, we take a moment to discuss duality and polarizations of abelian varieties; we do want to understand these notions so that we can make sense of the Weil pairing. Motivated by the utility of the Picard group in defining the Jacobian, we make the following definition.

Definition 2.28 (dual abelian variety)**.** Fix an abelian variety A over a field K. Then we define the *dual abelian variety* $A^\vee$ as the group scheme $\operatorname{Pic}^\circ_{A/K}$ over K.

Remark 2.29. As usual, we will not check that $A^\vee$ is an abelian variety or even a scheme, but it is. (The ingredients that go into these arguments will not be relevant for us.) We refer to [EGM, Chapter 6] for these arguments, in addition to the useful fact that $\dim A = \dim A^\vee$.

Remark 2.30. It is worthwhile to note that, in the complex analytic situation, there already is a notion of a dual abelian variety. If $A = V/\Lambda$ is an abelian variety, then $A^\vee = V^*/\Lambda^*$, where V^* is the vector space of conjugation-semilinear functionals $V^* \to \mathbb{C}$, and Λ^* consists of the functionals which are integral on Λ. It is rather tricky to explain how this definition relates to the one above, so we will not do so and instead refer to [Ros86, Section 4].

It is worth our time to explain why this is called duality. To begin, there is a duality for morphisms.

Lemma 2.31. Fix a homomorphism $f\colon A \to B$ of abelian varieties over a field K. Then there is a dual homomorphism $f^\vee\colon B^\vee \to A^\vee$.

Proof. We define the homomorphism on geometric points. Then a point of $B^\vee(\overline{K})$ is a line bundle $\mathcal{L}$ on $B_{\overline{K}}$, which we can pull back to a line bundle $f^*\mathcal{L}$ on $A_{\overline{K}}$, which is a point of $A^\vee(\overline{K})$. ∎

Lemma 2.32. Fix an abelian variety A over a field K. Then there is a canonical isomorphism $A \to A^{\vee\vee}$.

Proof. We sketch the construction of the map and refer to [EGM, Theorem 7.9] for details. Because $A^\vee$ is a moduli space of line bundles, there is a universal Poincaré line bundle $\mathcal{P}_A$ on $A \times A^\vee$. Unravelling the definition of $A^\vee$, we see that morphisms $S \to A^\vee$ correspond to line bundles on $A \times S$. Turning this around, we thus see that we can view $\mathcal{P}_A$ as a family of line bundles on $A^\vee$ parameterized by A and thus providing a map $A \to A^{\vee\vee}$. This map is the required isomorphism. ■

Most of the utility one achieves from the dual is that it allows us to the complex-analytic notion of a polarization into algebraic geometry. As in Remark 2.30, we view $A = V/\Lambda$ as a complex torus, and the dual abelian variety $A^\vee$ can be realized concretely as some V^*/Λ^*. Now, a polarization of A refers to a polarization of $\Lambda = \mathrm{H}_1^{\mathrm{B}}(A, \mathbb{Z})$, which as mentioned in Example 2.5 has equivalent data to an alternating form $\psi\colon \Lambda \otimes \Lambda \to \mathbb{Z}$ such that the bilinear form

$$\langle x, y\rangle \coloneqq \psi_{\mathbb{R}}(x, iy)$$

on $\Lambda_{\mathbb{R}}$ is symmetric and positive-definite. But now we see that this choice of ψ determines a map $A \to A^\vee$ given by $v \mapsto \psi(v, \cdot)$.

Thus, we would like our polarizations some kind of map $A \to A^\vee$. However, we need to keep track of all the adjectives that ψ had in order to make this an honest definition. For example, perhaps we want to keep track of the constraint that ψ is alternating. To do so, we use cohomology. We will shortly explain in Theorem 2.98 that the cup product provides an isomorphism $\wedge^2 \mathrm{H}^1(A, \mathbb{Z}) \to \mathrm{H}^2(A, \mathbb{Z})$, which induces an isomorphism

$$\mathrm{Hom}_{\mathbb{Z}}\left(\wedge^2 \Lambda, \mathbb{Z}\right) \cong \mathrm{H}^2(A, \mathbb{Z})$$

upon taking duals. Thus, ψ being an alternating form can be traced backed to it coming from a class in $\mathrm{H}^2(A, \mathbb{Z})$.

Continuing, perhaps we want to keep track of the constaint that $\langle \cdot, \cdot \rangle$ is symmetric. This is equivalent to having $\psi_{\mathbb{R}}(ix, iy) = \psi(x, y)$, which turns out to be equivalent to $\psi_{\mathbb{C}} \in \mathrm{H}^2(A, \mathbb{C})$ living in the $(1,1)$ component. Well, it turns out that the exponential short exact sequence

$$0 \to \mathbb{Z} \overset{2\pi i}{\to} \mathcal{O}_A \overset{\exp}{\to} \mathcal{O}_A^\times \to 0$$

induces a (first Chern class) map $c_1 : \mathrm{H}^1(A, \mathcal{O}_A^\times) \to \mathrm{H}^2(A, \mathbb{Z})$, which is an isomorphism onto the $(1,1)$ component. Thus, the condition that $\langle \cdot, \cdot \rangle$ is symmetric can be traced back to $\psi_{\mathbb{C}}$ coming from a class in $\mathrm{H}^1(A, \mathcal{O}_A^\times)$, which has equivalent data to a line bundle $\mathcal{L}$.

Lastly, it turns out that positive-definiteness of $\langle \cdot, \cdot \rangle$ corresponds to the $\mathcal{L}$ being ample. On the other hand, given a line bundle $\mathcal{L}$ on A, we remark that there already is a natural way to construct a map $A \to A^\vee$ from a line bundle. This gives our definition.

Definition 2.33 (polariaztion)**.** Fix an abelian variety A over a field K. A *polarization* is a morphism $\varphi\colon A \to A^\vee$ such that there is an ample line bundle $\mathcal{L}$ on $A_{\overline{K}}$ giving the equality

$$\varphi(x) = t_x^*\mathcal{L} \otimes \mathcal{L}^{-1}$$

for any $x \in A_{\overline{K}}$. We say that φ is *principal* if and only if it is an isomorphism, and we say that A is a *pricipally polarized*.

Remark 2.34. It turns out that the construction of the above map does correspond to the map $A \to A^\vee$ defined complex-analytically.

Remark 2.35. It turns out that polarizations are isogenies.

Remark 2.36. Here is the sort of thing that one can do with this definition. One may also want to define a Rosati involution on $\mathrm{End}(A)_{\mathbb{Q}}$, analogous to the Rosati involution on polarized Hodge structures. Well, given a (principal) polarization $\varphi\colon A \to A^\vee$, we can define a Rosati involution $(\cdot)^\dagger$ on $\mathrm{End}(A)_{\mathbb{Q}}$ by sending any $f \in \mathrm{End}(A)_{\mathbb{Q}}$ to

$$f^\dagger \coloneqq \varphi^{-1} \circ f^\vee \circ \varphi.$$

If λ is a principal polarization, then this Rosati involution descends to $\mathrm{End}(A)$. One can check that $(\cdot)^\dagger$ continues to be a positive anti-involution, but it is not easy; see for example [EGM, Theorem 12.26]. This allows us to apply the Albert classification Theorem 1.28 to our situation.

Example 2.37. For any smooth proper curve C such that $C(K) \neq \varnothing$, it turns out that the Jacobian $\mathrm{Jac}\, C$ is principally polarized. It is not too hard to describe the line bundle which gives the polarization: let $\iota\colon C \to \mathrm{Jac}(C)$ be an embedding given be one of the points in $C(K)$, and then the line bundle is given by the divisor

$$\underbrace{f(C) + \cdots + f(C)}_{g-1},$$

where g is the genus of C. See [EGM, Theorem 14.23] or [Mil08, Theorem 6.6] for more details.

Analogous to the complex-analytic setting $A = V/\Lambda$, we may still want to be able to define an alternating form on $\Lambda = \mathrm{H}_1^{\mathrm{B}}(A, \mathbb{Z})$. We will achieve a satisfying version of this in Lemma 2.113, but for now, let us point that this is not immediately obvious how to do this because we currently have no analogue for Λ in the general setting. However, we note that the alternating form Λ is able to induce an alternating form on V, and we can access a dense subset of V by taking torsion. Thus, for now, we will aim to provide a pairing

$$A[n](K^{\mathrm{sep}}) \times A[n](K^{\mathrm{sep}}) \to \mathbb{Z}/n\mathbb{Z}$$

for each integer n such that $\mathrm{char}\, K \nmid n$. Unwinding how we took a polarization to a map $A \to A^\vee$, we note that we may as well define the above map using a polarization $\varphi\colon A \to A^\vee$ by instead defining a pairing

$$A[n](K^{\mathrm{sep}}) \times A^\vee[n](K^{\mathrm{sep}}) \to \mathbb{Z}/n\mathbb{Z}$$

and then pre-composing with $A \to A^\vee$. More generally, given an isogeny $f\colon A \to B$, we will be able to show that there is a perfect pairing

$$(\ker f) \times (\ker f^\vee) \to \mathbb{G}_m,$$

upon which we find the desired pairing by taking $f = [n]_A$ and taking K^{sep}-points.

Proposition 2.38 (Weil pairing)**.** Fix an isogeny $f\colon A \to B$ of abelian varieties over K. Then there is a perfect pairing

$$(\ker f) \times (\ker f^\vee) \to \mathbb{G}_m.$$

Proof. We provide an explicit construction of the pairing on K^{sep}-points, but we will not check that it is perfect, for which we refer to [Ton15, Theorem 8.1.3]. Select $x \in (\ker f)(K^{\mathrm{sep}})$ and $y^\vee \in (\ker f^\vee)(K^{\mathrm{sep}})$. The point $y^\vee$ corresponds to a line bundle $\mathcal{L}$ on $B^\vee_{K^{\mathrm{sep}}}$. Being in the kernel of f grants a trivialization $\sigma\colon f^*\mathcal{L} \to \mathcal{O}_{A_{K^{\mathrm{sep}}}}$, which is unique up to a scalar. Now, note that $t_a^* f^*\mathcal{L} = f^* t_{f(a)}^* \mathcal{L} = f^*\mathcal{L}$ because $a \in \ker f$, so there is another trivialization of $f^*\mathcal{L}$ given by $t_a^*\beta\colon \mathcal{L} \to \mathcal{O}_{A_{K^{\mathrm{sep}}}}$. We now define our Weil pairing as $t_a^*\beta \circ \beta^{-1}$, which we realize as an element of $\mathbb{G}_m(K^{\mathrm{sep}})$ by noting that $t_a^*\beta \circ \beta^{-1}$ is an automorphism of $\mathcal{O}_{A_{K^{\mathrm{sep}}}}$ and is therefore a scalar. ∎

Corollary 2.39. Fix an abelian variety A over a field K, and let $\varphi\colon A \to A^\vee$. For each positive integer n, there is a Galois-invariant perfect symplectic pairing

$$e_\varphi\colon A[n](K^{\mathrm{sep}}) \times A[n](K^{\mathrm{sep}}) \to \mu_n.$$

Furthermore, for any positive integer m, the following diagram commutes.

$$\begin{array}{ccc} A[nm](K^{\mathrm{sep}}) & \times & A[nm](K^{\mathrm{sep}}) \xrightarrow{e_\varphi} \mu_{mn} \\ \downarrow{\scriptstyle m} & & \downarrow{\scriptstyle m} \quad \downarrow{\scriptstyle m} \\ A[n](K^{\mathrm{sep}}) & \times & A[n](K^{\mathrm{sep}}) \xrightarrow{e_\varphi} \mu_n \end{array}$$

Proof. We described above how to construct the pairing from the one given in Proposition 2.38 by setting $f = [n]_A$ and then using the polarization φ. The remaining properties of e_φ (such as Galois-invariance) can be checked using the explicit construction given in Proposition 2.38. ∎

2.1.4 Applying Hodge Theory

We now explain the utility of chapter 1 to our application. Here is the main result.

Theorem 2.40 (Riemann)**.** The functor $A \mapsto \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$ provides an equivalence of categories between the isogeny category of abelian varieties defined over $\mathbb{C}$ and the category of polarizable $\mathbb{Q}$-Hodge structures V such that $V_{\mathbb{C}} = V^{0,1} \oplus V^{1,0}$.

Proof. Writing $A = \mathbb{C}^g/\Lambda$ for a polarizable lattice Λ, we see that the given functor takes A to $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$. It is thus not hard to see that this functor is fully faithful. To see that it is essentially surjective, we begin with any polarizable $\mathbb{Q}$-Hodge structure V and find a polarizable sublattice Λ in order to produce the desired abelian variety A/Λ. Admittedly, most of the work for this theorem was already done in Example 1.20 when we showed that the previous sentence actually gives an abelian variety! ∎

The moral of the story is that we can keep track of abelian varieties A over $\mathbb{C}$ by only keeping track of their Hodge structures $\mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. With this in mind, we allow ourselves the following notation.

Notation 2.41. Fix an abelian variety A over $\mathbb{C}$. Then we define the *Mumford–Tate group* of A to be

$$\mathrm{MT}(A) \coloneqq \mathrm{MT}\left(\mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})\right).$$

We define $\mathrm{Hg}(A)$ and $\mathrm{L}(A)$ similarly.

Here is the main corollary of Theorem 2.40 that we will want.

Corollary 2.42. Fix an abelian variety A over $\mathbb{C}$. Then the natural map

$$\mathrm{End}_{\mathbb{C}}(A) \otimes_{\mathbb{Z}} \mathbb{Q} \to \mathrm{End}_{\mathbb{Q}}\left(\mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})\right)^{\mathrm{MT}(A)}$$

is an isomorphism.

Proof. By Lemma 1.54, we see that the right-hand side is simply $\mathrm{End}_{\mathrm{HS}}\left(\mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})\right)$. The result now follows from Theorem 2.40. ∎

As another aside, we go ahead and restate the Albert classification (Theorem 1.28) for our abelian varieties.

Proposition 2.43. Fix a simple abelian variety A of dimension g, defined over a field K of characteristic 0, and set $D \coloneqq \mathrm{End}_K(A)_{\mathbb{Q}}$ and $E \coloneqq Z(D)$. Letting $(\cdot)^\dagger$ be the Rosati involution on D, we also let $E^\dagger$ be the $(\cdot)^\dagger$-invariants of E. Further, set $d \coloneqq \sqrt{[D:E]}$ and $e \coloneqq [E:\mathbb{Q}]$ and $e_0 \coloneqq [E^\dagger:\mathbb{Q}]$. Then we have the following table of restrictions on (g, d, e, e_0).

Type	e	d	Restriction
I	e_0	1	$e \mid g$
II	e_0	2	$2e \mid g$
III	e_0	2	$2e \mid g$
IV	$2e_0$	d	$e_0 d^2 \mid g$

Proof. Recall that D is amenable to the Albert classification as discussed in Remark 2.36. The middle two columns follow from the discussion of the various types; for example, in Type I, we see $d = 1$ because $D = E$, and $e = e_0$ because E is totally real. To receive the dimension restrictions, we note that some descent argument allows us to reduce to the case where $K = \mathbb{C}$, where we receive an inclusion $D \subseteq \mathrm{End}(\mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q}))$ by Theorem 2.40.[1] This is an inclusion of division $\mathbb{Q}$-algebras, so we see that $\dim_{\mathbb{Q}} D \mid 2g$; this implies

$$d^2 e \mid 2g,$$

which rearranges into the required restrictions. ∎

Remark 2.44. The requirement that $\mathrm{char}\, E = 0$ is necessary in the table; the restrictions are somewhat different (and weaker!) in positive characteristic.

While we're here, we state the main theorem of [Del18] on absolutely Hodge cycles.

Theorem 2.45 (Deligne)**.** Fix an abelian vareity A defined over a number field K. Then all Hodge classes on A are absolutely Hodge.

2.1.5 Complex Multiplication

Even though it is not strictly necessary for our exposition, we take a moment to discuss some theory surrounding complex multiplication. We refer to [Mil20b] throughout for more details. The relevance of this discussion to us mostly arises because we have defined analogous notions in sections 2.2.2 and 2.2.3.

Intuitively, complex multiplication simply means that an abelian variety has many endomorphisms. To explain this properly, we note that the endomorphism algebra of a simple abelian variety A is a division $\mathbb{Q}$-algebra described in Proposition 2.43; if we drop the assumption that A is simple, then it could be a product of matrix algebras of such division $\mathbb{Q}$-algebras. This motivates the following definition to properly account for such matrix algebras.

Definition 2.46 (reduced degree)**.** Write a semisimple algebra D over a field K as a product $D_1 \times \cdots \times D_k$ of simple algebras. Then we define the *reduced degree* as

$$[D:K]_{\mathrm{red}} \coloneqq \sum_{i=1}^{k} \sqrt{[D_i : E_i]} \cdot [D_i : K],$$

where $E_i \coloneqq Z(D_i)$ for each i

[1] It is still possible to get an inclusion like this in general. It requires a discussion of the ℓ-adic representations, which we engage in later.

Remark 2.47. It is not technically obvious that $[D_i : F_i]$ is a square, but this follows from the theory of central simple algebras. Roughly speaking, one can show that $D_i \otimes \overline{D_i} \cong M_n(\overline{D_i})$ for some $n \geq 0$, from which the result follows; see [Mil20a, Corollary IV.2.16].

Remark 2.48. Given an inclusion $B \subseteq \mathrm{End}_K(V)$, one receives a bound

$$[B : K]_{\mathrm{red}} \leq [V : K].$$

Roughly speaking, this follows by breaking up B into simple pieces (which are matrix algebras of division algebras) and then looking for these pieces in $\mathrm{End}_K(V)$. See [Mil20b, Proposition I.1.2]

In light of the previous remark, we are now able to make a definition.

Definition 2.49 (complex multiplication)**.** Fix an abelian variety A over a field K. Then A has *complex multiplication over* K if and only if

$$[\mathrm{End}_K(A)_{\mathbb{Q}} : \mathbb{Q}]_{\mathrm{red}} = 2 \dim A.$$

Namely, we see that $2 \dim A$ is as large as possible by Remark 2.48, by taking V to be H^1 for some Weil cohomology H.[2]

Remark 2.50. The key benefit of the reduced degree is that it is additive: given abelian varieties A and A', we claim

$$[\mathrm{End}(A \oplus A')_{\mathbb{Q}} : \mathbb{Q}]_{\mathrm{red}} \stackrel{?}{=} [\mathrm{End}(A)_{\mathbb{Q}} : \mathbb{Q}]_{\mathrm{red}} + [\mathrm{End}(A')_{\mathbb{Q}} : \mathbb{Q}]_{\mathrm{red}}.$$

Indeed, by breaking everything into simple pieces, we may assume that A and A' are both powers of a simple abelian variety. If they are powers of different simple abelian varieties, then this is a direct computation. Otherwise, they are powers of the same simple abelian variety, in which case all central simple algebras in sight are matrix algebras over the same division algebra, and the result follows by another computation.

Remark 2.51. A computation with Proposition 2.43 shows that a simple abelian variety A has complex multiplication only in Type IV when $d = 1$; i.e., we require $\mathrm{End}_K(A)$ to be a CM field. Combining this with Remark 2.50, we find that an abelian variety A has complex multiplication if and only if each of its factors does.

Remark 2.52. If an abelian variety A with complex multiplication is a sum of non-isomorphic simple abelian varieties, then its endomorphism algebra is simply a product of CM fields. In general, one can show that it is still the case that any abelian variety A with complex multiplication has a CM algebra of degree $2 \dim A$ contained in its endomorphism algebra. However, this requires a little structure theory of semisimple algebras; see [Mil20b, Proposition 3.6].

Complex multiplication places strong constraints on the Mumford–Tate group.

Proposition 2.53. Fix an abelian variety A over $\mathbb{C}$. Then A has complex multiplication if and only if $\mathrm{MT}(A)$ is a torus.

Proof. We show the two implications separately.

[2] Outside the complex-analytic case, it may look like one wants to use the ℓ-adic result Theorem 2.126 over a general field. However, it turns out to be enough to merely achieve the injectivity of the map Theorem 2.126, which is easier.

- In one direction, if A has complex multiplication, then Remark 2.52 grants a CM algebra $E \subseteq \mathrm{End}_{\mathbb{C}}(A)_{\mathbb{Q}}$ with $[E:\mathbb{Q}] = 2\dim A$. Then $\mathrm{H}^1_{\mathrm{B}}(A,\mathbb{Q})$ is a free module over E of rank 1, so we see that $\mathrm{GL}_F\left(\mathrm{H}^1_{\mathrm{B}}(A,\mathbb{Q})\right)$ is isomorphic to T_F. We conclude by Lemma 1.45.
- In the other direction, suppose $\mathrm{MT}(A)$ is a torus. Find a maximal torus T containing $\mathrm{MT}(A)$. Then Corollary 2.42 tells us that
$$\mathrm{End}_{\mathbb{C}}(A)_{\mathbb{Q}} = \mathrm{End}_{\mathbb{Q}}\left(\mathrm{H}^1_{\mathrm{B}}(A,\mathbb{Q})\right)^{\mathrm{MT}(A)},$$
which then contains $\mathrm{End}_{\mathbb{Q}}\left(\mathrm{H}^1_{\mathrm{B}}(A,\mathbb{Q})\right)^{T}$. However, the latter is a commutative semisimple $\mathbb{Q}$-algebra of dimension $2g$: it suffices to check this after base-changing to $\mathbb{C}$, whereupon we may identify T with the diagonal torus, from which the claim follows. This completes the proof. ■

One benefit of complex multiplication is that it lets move difficult geometric questions into combinatorial ones. To see this, we need to define the following combinatorial gadget.

Definition 2.54. Fix an abelian variety A with complex multiplication defined over $\mathbb{C}$, and set $V := \mathrm{H}^1_{\mathrm{B}}(A,\mathbb{Q})$. Choose a CM algebra $E \subseteq \mathrm{End}_{\mathbb{C}}(A)_{\mathbb{Q}}$ with $\dim E = 2\dim A$. Then we define the *CM type* $\Phi\colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ of A to be the CM signature (E,Φ) given by
$$V^{1,0} \cong \bigoplus_{\sigma\in\Sigma_F} \mathbb{C}_\sigma^{\Phi(\sigma)}.$$
Note that $\mathrm{H}^1_{\mathrm{B}}(A,\mathbb{Q})$ is then a one-dimensional E-vector space, so $\mathrm{im}\,\Phi \subseteq \{0,1\}$, so we can realize Φ as a subset of $\mathrm{Hom}(E,\mathbb{C})$.

Remark 2.55. Note that we are not requiring $E = Z(\mathrm{End}_{\mathbb{C}}(A)_{\mathbb{Q}})$, though this is automatically the case when the simple components of A all have multiplicity 1. Of course, there still is a CM signature coming from the case $E = Z(\mathrm{End}_{\mathbb{C}}(A)_{\mathbb{Q}})$.

Remark 2.56. There is a still a way to recover the CM type even when A is not defined over $\mathbb{C}$. For example, one can note that H^{10} is supposed to be the Lie algebra $\mathrm{Lie}\,A$, so one can instead recover Φ from the E-action on $\mathrm{Lie}\,A$.

Remark 2.57. One can read the simplicity of A off of the CM type (E,Φ). To begin, one needs E to be a field for A to be simple. Now that E is a field, we know that $A \sim B^r$ where B is an abelian variety with complx multiplication; say that it has CM type (E',Φ'). Then the Hodge structure on A is determined by the Hodge structure on B. Tracking this through as in [Lan11, Theorem 3.6] shows that A is simple if and only if any Galois extension $L/\mathbb{Q}$ of E has that
$$\{\sigma \in \mathrm{Gal}(L/\mathbb{Q}) : \Phi\sigma = \Phi\} = \mathrm{Gal}(L/E),$$
where Φ is being suitably thought of as an element of $\mathbb{Z}[\mathrm{Hom}(E,L)]$.

Remark 2.58. It turns out that there is (essentially) exactly one abelian variety with CM type (E,Φ), up to isogeny over the algebraic closure. See [Mil20b, Proposition 3.12].

Remark 2.58 tells us that we are basically allowed to only pay attention to the CM type in the theory of complex multiplication.

2.2 The Center of MT

In this section, we begin with a computational tool to compute $\mathrm{MT}(A)$ for an abelian variety A. This discussion is somewhat involved, so we will spend a full section here.

Let's begin with some motivation. Fix an abelian variety A. In the application of this thesis, we will use Lemma 1.62 to compute $\mathrm{Hg}(A)^{\mathrm{der}}$: note $\mathrm{Hg}(A)^{\mathrm{der}}$ is semisimple and hence its Lie algebra can be written as the sum of simple Lie algebras which may be amenable to the lemma. Because $\mathrm{Hg}(A)$ is reductive by Lemma 1.44, it remains to compute the center $Z(\mathrm{Hg}(A))$; recall $\mathrm{Hg}(A)$ is connected by Remark 1.40, so we may as well compute the connected component $Z(\mathrm{Hg}(A))^\circ$. As usual, the same discussion holds for $\mathrm{MT}(A)$, but we note that $Z(\mathrm{MT}(A))^\circ$ tends to be nontrivial because usually $\mathbb{G}_{m,\mathbb{Q}} \subseteq \mathrm{MT}(A)$ by Example 1.31.

In Proposition 2.67, we find that $Z(\mathrm{Hg}(A))^\circ$ is trivial unless A has irreducible factors of Type IV in the sense of the Albert classification (Theorem 1.28). As such, we spend the rest of the section focusing on computations in Type IV. Computations are well-understood when V comes from an abelian variety with complex multiplication, so the main contribution here is that these arguments generalize with only slight modifications.

2.2.1 General Comments

In this subsection, we will mostly work with general polarizable Hodge structures V.

Lemma 2.59. Fix $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight, and set $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ with $E \coloneqq Z(D)$. Viewing D as a $\mathbb{Q}$-group, we have

$$Z(\mathrm{Hg}(V)) \subseteq \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E},$$

where $\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E}$ embeds into $\mathrm{GL}(V)$ via the D-action on V.

Proof. Here, E is a product of number fields because it is a commutative semisimple $\mathbb{Q}$-algebra. Recall from Lemma 1.54 that

$$D = \mathrm{End}_{\mathbb{Q}}(V)^{\mathrm{Hg}(V)},$$

which upgrades to an equality of algebraic subgroups of $\mathrm{End}_{\mathbb{Q}}(V)$ because $\mathbb{Q}$-points are dense in these algebraic groups by combining [Mil17, Corollary 17.92] and [Mil17, Definition 12.59]. In particular, we see that $\mathrm{Hg}(V)$ commutes with $D^\times$, so the double centralizer theorem enforces $Z(\mathrm{Hg}(V)) \subseteq D^\times$ even as algebraic groups. However, $Z(\mathrm{Hg}(V))$ now commutes fully with $D^\times$, so in fact $Z(\mathrm{Hg}(V)) \subseteq Z(D)^\times$, which is what we wanted. ∎

Remark 2.60. One also has $Z(\mathrm{MT}(V)) \subseteq \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E}$ because $\mathrm{MT}(V) \subseteq \mathbb{G}_{m,\mathbb{Q}} \mathrm{Hg}(V)$ by Lemma 1.41, and the scalars $\mathbb{G}_{m,\mathbb{Q}}$ already live in $\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E}$.

Lemma 2.59 is that it places the center $Z(\mathrm{Hg}(V))$ in an explicit torus $\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E}$. Subgroups of tori are well-understood by (co)character groups, so this puts us in good shape. This torus will be important enough to have its own notation.

Notation 2.61. Fix a commutative semisimple $\mathbb{Q}$-algebra E (i.e., a product of number fields). Then we define the torus

$$\mathrm{T}_E \coloneqq \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{m,E}.$$

Remark 2.62. Writing E as a product of number fields $E_1 \times \cdots \times E_k$, we find

$$\mathrm{T}_E = \mathrm{T}_{E_1} \times \cdots \times \mathrm{T}_{E_k}$$

because $E = E_1 \times \cdots \times E_k$ is an equality of $\mathbb{Q}$-algebras.

Remark 2.63. Let's compute the character group $\mathrm{X}^*(\mathrm{T}_E)$. By Remark 2.62, it's enough to do this computation when E is a field. The choice of a primitive element $\alpha \in E$ with minimal monic polynomial $f(x)$ yields an isomorphism $E \cong \mathbb{Q}[x]/(f(x))$. Upon base-changing to $\overline{\mathbb{Q}}$, we get a ring isomorphism

$$E \otimes_{\mathbb{Q}} \overline{\mathbb{Q}} \cong \prod_{i=1}^{n} \frac{\overline{\mathbb{Q}}[x]}{(x-\alpha_i)},$$

where $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ are the roots of $f(x)$ in $\overline{\mathbb{Q}}$. Each root α_i provides a unique embedding $E \hookrightarrow \overline{\mathbb{Q}}$, so we see that $(\mathrm{T}_E)_{\overline{\mathbb{Q}}} \cong \mathbb{G}_{m,\overline{\mathbb{Q}}}^n$, where the n maps $(\mathrm{T}_E)_{\overline{\mathbb{Q}}} \to \mathbb{G}_{m,\overline{\mathbb{Q}}}$ are given by the embedding $\sigma_i \colon E \hookrightarrow \overline{\mathbb{Q}}$ defined by $\sigma_i(\alpha) \coloneqq \alpha_i$. In total, we find that $\mathrm{X}^*(\mathrm{T}_E)$ is a free $\mathbb{Z}$-module spanned by the embeddings $\Sigma_E \coloneqq \{\sigma_1, \dots, \sigma_n\}$, and it has the natural Galois action. Dually, $\mathrm{X}_*(\mathrm{T}_E)$ has the dual basis $\Sigma_E^\vee = \{\sigma_1^\vee, \dots, \sigma_n^\vee\}$.

In the light of the above remark, we will want the following notation.

Notation 2.64. Given a number field E, we let Σ_E denote the collection of embeddings $E \hookrightarrow \overline{\mathbb{Q}}$. Given a product of number fields $E \coloneqq E_1 \times \cdots \times E_k$, we define $\Sigma_E \coloneqq \Sigma_{E_1} \sqcup \cdots \sqcup \Sigma_{E_k}$.

The point of the above notation is that $\mathrm{X}^*(\mathrm{T}_E) = \mathbb{Z}[\Sigma_E]$ as Galois modules.

It is possible to upgrade Lemma 2.59 in the presence of a polarization.

Lemma 2.65. Fix a polarizable $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight, and set $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ with $E \coloneqq Z(D)$. Then

$$Z(\mathrm{Hg}(V)) \subseteq \left\{ g \in \mathrm{T}_E : gg^\dagger = 1 \right\},$$

where $(\cdot)^\dagger$ is the Rosati involution.

Proof. As usual, everything in sight upgrades to algebraic groups. Let φ be a polarization. Fix some $g \in \mathrm{Hg}(V)$; note that Lemma 2.59 implies $g \in \mathrm{T}_E$, so it makes sense to write down $g^\dagger$.

Now, by the non-degeneracy of φ, it is enough to show that

$$\varphi\left(gg^\dagger v \otimes w\right) \stackrel{?}{=} \varphi(v \otimes w)$$

for any $v, w \in V$. Well, the definition of $(\cdot)^\dagger$ tells us that the left-hand side equals $\varphi\left(g^\dagger v \otimes g^\dagger w\right)$, which equals $\varphi(v \otimes w)$ because $\mathrm{Hg}(V) \subseteq \mathrm{Sp}(\varphi)$ by Remark 1.53. ∎

Once again, this torus is important enough to earn its own notation.

Notation 2.66. Fix a commutative semisimple $\mathbb{Q}$-algebra E with involution $(\cdot)^\dagger$. Then we define the torus

$$\mathrm{U}_E \coloneqq \left\{ g \in \mathrm{T}_E : xx^\dagger = 1 \right\}.$$

Here is an application of Lemma 2.65.

Proposition 2.67. Fix polarizable $V \in \mathrm{HS}_{\mathbb{Q}}$ of pure weight. Suppose that V has no irreducible Hodge substructures with endomorphism algebra of Type IV in the sense of the Albert classification (Theorem 1.28). Then $Z(\mathrm{Hg}(V))$ is finite, and $\mathrm{Hg}(V)$ is semisimple.

Proof. Quickly, recall from Lemma 1.44 that $\mathrm{Hg}(V)$ is reductive, so the finitness of $Z(\mathrm{Hg}(V))$ implies that $Z(\mathrm{Hg}(V))^\circ = 1$ and thus $\mathrm{Hg}(V) = \mathrm{Hg}(V)^{\mathrm{der}}$, making $\mathrm{Hg}(V)$ is semisimple. (See also [Mil17, Proposition 19.10].) As such, we will focus on the first claim.

Set $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ with $E \coloneqq Z(D)$ so that $\mathrm{Hg}(V) \subseteq \mathrm{U}_E$ by Lemma 2.65. It is therefore enough to check that U_E is finite. Well, E is a product of number fields, and upon comparing with Theorem 1.28, we see that avoiding Type IV implies that E is a product of totally real fields. Totally real fields have only two units, so finiteness of U_E follows. ■

Thus, we see that we have pretty good control outside of Type IV factors, so we will spend the rest of this section on Type IV. For some applications outside Type IV, see (for example) [Lom16].

2.2.2 Type IV: The Signature

The arguments in the next two subsections are motivated by the computation of [Yu15, Lemma 4.2] and [Yan94, Proposition 1.1]. For this subsection, A is an abelian variety over $\mathbb{C}$ whose irreducible factors are of Type IV in the sense of the Albert classification (Theorem 1.28). Note that $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$ is a Hodge structure concentrated in $V^{0,1}$ and $V^{1,0}$, so we do so.

By assumption, we know that $D \coloneqq \mathrm{End}_{\mathrm{HS}}(V)$ is a division algebra over its center $E \coloneqq Z(D)$, where E is a CM algebra (i.e., a product of CM fields), and the Rosati involution $(\cdot)^\dagger$ restricts to complex conjugation on E. In particular, $E^\dagger$ is the product of the maximal totally real subfields of E.

The basic approach of this subsection is that Lemma 2.59 requires $Z(\mathrm{Hg}(A))^\circ \subseteq \mathrm{T}_E$, and one can compute subtori using the machinery of (co)character groups. In particular, we recall that $\mathrm{X}^*(\Sigma_E) = \mathbb{Z}[\Sigma_E]$ and $\mathrm{X}_*(\Sigma_E) = \mathbb{Z}[\Sigma_E^\vee]$ as Galois modules. We will need a way to work directly with the Hodge structure on V. It will be described by the following piece of combinatorial data. Recall that a CM algebra is a product of CM fields.

Definition 2.68 (signature)**.** Fix a CM algebra E, and recall that Σ_E is the set of homomorphisms $E \hookrightarrow \overline{\mathbb{Q}}$. Then a *signature* is a function $\Phi\colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ such that the sum

$$\Phi(\sigma) + \Phi(\overline{\sigma})$$

is constant with respect to $\sigma \in \Sigma_E$; here, $\overline{\sigma}$ denotes the complex conjugate embedding to σ. We may call the pair (E, Φ) a *CM signature*.

Remark 2.69. One can also view Φ as an element of $\mathbb{Z}[\Sigma_E]$ as

$$\Phi \coloneqq \sum_{\sigma \in \Phi} \Phi(\sigma)\sigma.$$

Remark 2.70. The case that $\Phi(\sigma) + \Phi(\overline{\sigma})$ always equals 1 corresponds to Φ being a CM type.

Remark 2.71. If we expand E as a product of CM fields $E = E_1 \times \cdots \times E_k$, then $\Sigma_E = \Sigma_{E_1} \sqcup \cdots \sqcup \Sigma_{E_k}$. Thus, we see that a signature of E has only a little more data than a signature on each of the $\Sigma_{E_\bullet}$s individually; in particular, one should make sure that $\Phi(\sigma) + \Phi(\overline{\sigma})$ remains equal across the different fields.

The idea is that we can keep track of a signature with a Hodge structure.

Lemma 2.72. Fix an abelian variety A over $\mathbb{C}$ such that $\mathrm{End}(A)$ contains a CM algebra E, and define $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. Then the function $\Phi\colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ defined by

$$V^{1,0} \cong \bigoplus_{\sigma \in \Sigma_E} \mathbb{C}_\sigma^{\Phi(\sigma)}$$

is a signature, which we will call the induced signature. This is an isomorphism of E-representations, where $\mathbb{C}_\sigma$ is a complex E-representation via the embedding σ.

Proof. In short, the condition that $\Phi(\sigma) + \Phi(\overline{\sigma})$ is constant comes from the condition $V^{0,1} = \overline{V^{1,0}}$. To see this, note that V is a free module over E, so $V_{\mathbb{C}}$ is a free module over $E \otimes \mathbb{C}$ of finite rank. As such, we may set $d \coloneqq [V : E]$ so that $V \cong E^d$ as E-representations, and then we find

$$V_{\mathbb{C}} \cong \bigoplus_{\sigma \in \Sigma_E} \mathbb{C}_\sigma^d.$$

Now, $V_{\mathbb{C}} = V^{1,0} \oplus V^{0,1}$, and because E acts by endomorphisms of Hodge structures, we get a well-defined action of E on $V^{1,0}$ and $V^{0,1}$ individually. In particular, the definition of Φ also grants

$$V^{0,1} \cong \bigoplus_{\sigma \in \Sigma_E} \mathbb{C}_\sigma^{d-\Phi(\sigma)}$$

as E-representations, so

$$\overline{V^{0,1}} \cong \bigoplus_{\sigma \in \Sigma_E} \mathbb{C}_{\overline{\sigma}}^{d-\Phi(\sigma)}$$

To complete the proof, we note that $V^{0,1} = \overline{V^{1,0}}$ continues to be true as E-representations, so we must have $\Phi(\sigma) = d - \Phi(\overline{\sigma})$ for all σ. The result follows. ∎

Of course, we cannot expect this signature Φ to remember everything about the Hodge structure. For example, if $\mathrm{End}(A)$ contains a larger CM algebra E' than E, then the signature induced by E' knows more about the Hodge structure than the one induced by E. However, in "generic cases," this signature is expected to suffice. For our purposes, we will take generic to mean that there are no more endomorphisms than the ones promised by E; i.e., this explains why we will assume $Z(\mathrm{End}(A)) = E$ in the sequel.

We now relate our signature to cocharacters of $Z(\mathrm{Hg}(A))^\circ$. For this, it will be helpful to realize $Z(\mathrm{Hg}(A))$ as some kind of monodromy group. The trick is to consider the determinant.

Lemma 2.73. Fix an abelian variety A over $\mathbb{C}$ such that $Z(\mathrm{End}(A))$ equals an algebra E, and define $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. Then $Z(\mathrm{Hg}(A))^\circ$ equals the largest algebraic $\mathbb{Q}$-subgroup of T_E containing the image of $(\det_E \circ h) \colon \mathbb{U} \to (\mathrm{T}_E)_{\mathbb{R}}$.

Proof. The point is that taking the determinant will kill $\mathrm{Hg}(A)^{\mathrm{der}}$ because $\mathrm{Hg}(A) \subseteq \mathrm{GL}_E(V)$. There are two inclusions we must show.

- We show that $Z(\mathrm{Hg}(A))^\circ$ contains the image of $(\det E \circ h|_{\mathbb{U}})$. Well, $\mathrm{Hg}(A)$ contains the image of $h|_{\mathbb{U}}$, so it is enough to show that $Z(\mathrm{Hg}(A))^\circ$ contains the image of $\det_E \colon \mathrm{Hg}(A) \to \mathrm{T}_E$. For this, we recall that $\mathrm{Hg}(A)$ is connected (by Remark 1.40), so

  $$\mathrm{Hg}(A) = Z(\mathrm{Hg}(A))^\circ \, \mathrm{Hg}(A)^{\mathrm{der}}.$$

  Note that $\det_E$ is simply $(\cdot)^{\dim_E V}$ on the torus $Z(\mathrm{Hg}(V))^\circ$, so that piece surjects onto $Z(\mathrm{Hg}(A))^\circ$! Thus, it is enough to check that $\det_E \colon \mathrm{Hg}(A)^{\mathrm{der}} \to \mathrm{T}_E$ is trivial, which is true by the definition of the derived subgroup upon noting that $\det_E$ is a homomorphism with abelian target.
- Suppose that $T \subseteq \mathrm{T}_E$ contains the image of $(\det_E \circ h|_{\mathbb{U}})$. Then we claim that T contains $Z(\mathrm{Hg}(A))^\circ$. Let $H \subseteq \mathrm{GL}_E(V)$ be the pre-image of T under $\det_E \colon \mathrm{GL}_E(A) \to \mathrm{T}_E$. Then H must contain the image of $h|_{\mathbb{U}}$, so it contains $\mathrm{Hg}(A)$ by defintion. In particular, H contains $Z(\mathrm{Hg}(A))^\circ$! Now, T contains $\det_E(H)$, so T contains $\det_E(Z(\mathrm{Hg}(A))^\circ)$, but the previous point check remarked that this simply equals $Z(\mathrm{Hg}(A))^\circ$, so we are done. ∎

Proposition 2.74. Fix an abelian variety A over $\mathbb{C}$ such that $Z(\mathrm{End}(A))$ equals a CM algebra E, and define $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. Let $\Phi \colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ be the induced signature. Then the induced representation $(\det_E \circ h) \colon \mathbb{U} \to (\mathrm{T}_E)_{\mathbb{R}}$ sends the generator of $\mathrm{X}_*(\mathbb{U})$ to

$$-\sum_{\sigma \in \Sigma_E} (\Phi(\sigma) - \Phi(\overline{\sigma}))\sigma^\vee.$$

Proof. This boils down to computing the map $\det_E \circ h|_{\mathbb{U}}$. We proceed in steps.

1. To set ourselves up, recall that
$$\mathbb{U}_{\mathbb{C}} = \{(z, 1/z) : z \in \mathbb{G}_{m,\mathbb{C}}\},$$
so one has an isomorphism cocharacter $z^\vee \colon \mathbb{G}_{m,\mathbb{C}} \to \mathbb{U}_{\mathbb{C}}$ given by $z^\vee \mapsto z \mapsto (z, 1/z)$. Thus, we have left to show that
$$\det_E \circ h_{\mathbb{C}} \circ z^\vee \stackrel{?}{=} -\sum_{\sigma \in \Sigma_E} (\Phi(\sigma) - \Phi(\overline{\sigma}))\sigma^\vee.$$
We may check this equality on geometric points.

2. We describe the map $h_{\mathbb{C}} \colon \mathbb{S}_{\mathbb{C}} \to \mathrm{GL}(V)_{\mathbb{C}}$. By definition, $h(z, w) \in \mathrm{GL}(V)$ acts by z^{-1} on $V^{1,0}$ and by w^{-1} on $V^{0,1}$. Thus, the definition of Φ grants that $h(z, w)$ diagonalizes. To be more explicit, for each $\sigma \in \Sigma_E$, we define $V_\sigma^{p,q}$ to be the σ-eigenspace for the E-action on $V^{p,q} \subseteq V_{\mathbb{C}}$. Then we see that $h(z, w)$ acts on $V_\sigma^{1,0}$ by the scalar z^{-1} and on $V^{0,1}$ by the scalar w^{-1}.

3. We describe the map $(\det_E \circ h_{\mathbb{C}}) \colon \mathbb{S}_{\mathbb{C}} \to (\mathrm{T}_E)_{\mathbb{C}}$. Realizing geometric points in $(\mathrm{T}_E)_{\mathbb{C}}$ as tuples in $\mathbb{C}^{\Sigma_E}$, we see that $\det_E$ simply takes the determinant of the matrix $h_{\mathbb{C}}(z, w)|_{V_\sigma}$ to the σ-component in $(\mathrm{T}_E)_{\mathbb{C}}$. (One finds this by tracking through the definition of $\det_E$ as a morphism of algebraic groups.) As such, we see that
$$\det h_{\mathbb{C}}(z, w)|_{V_\sigma} = z^{-\Phi(\sigma)} w^{-\Phi(\overline{\sigma})}$$
because Φ is a signature.

4. We complete the proof. The previous step shows that $(\det_E \circ h_{\mathbb{C}} \circ z^\vee)(z)$ goes to the element
$$\left(z^{-\Phi(\sigma)+\Phi(\overline{\sigma})}\right)_{\sigma \in \Sigma(E)} \in \mathbb{C}^{\Sigma_E}.$$
This completes the proof upon noting that the cocharacter $\sigma^\vee \colon \mathbb{G}_{m,\mathbb{C}} \to \mathrm{T}_E$ simply maps into the σ-component of $\mathbb{C}^{\Sigma_E}$ on geometric points. ∎

Remark 2.75. Notably, the given element sums to 0, which corresponds to the fact that $\mathrm{Hg}(A) \subseteq \mathrm{SL}(V)$ as seen in Lemma 1.41. Indeed, by diagonalizing the E-action on V, we see that $(\mathrm{T}_E \cap \mathrm{SL}(V))^\circ$ consists of the $g \in \mathrm{T}_E$ such that the product of the elements in g equals 1.

Proposition 2.74 easily translates into a computation of the cocharacter group $\mathrm{X}_*(\mathrm{Hg}(A))^\circ$. In the next few results, saturated simply means that the induced quotient is torsion-free.

Corollary 2.76. Fix an abelian variety A over $\mathbb{C}$ such that $Z(\mathrm{End}(A))$ equals a CM algebra E, and define $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. Let $\Phi \colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ be the induced signature. Then $Z(\mathrm{Hg}(A))^\circ \subseteq \mathrm{T}_E$ has cocharacter group equal to the smallest saturated Galois submodule of $\mathrm{X}_*(\mathrm{T}_E) = \mathbb{Z}[\Sigma_E^\vee]$ containing
$$\sum_{\sigma \in \Sigma_E} (\Phi(\sigma) - \Phi(\overline{\sigma}))\sigma^\vee.$$

Proof. This is immediate from combining Lemma 2.73 and Proposition 2.74 with the equivalence of categories X_* between algebraic tori and Galois modules. See [Mil17, Theorem 12.23] for the proof that X^* is an equivalence, which is similar. ∎

Corollary 2.77. Fix an abelian variety A over $\mathbb{C}$ such that $Z(\mathrm{End}(A))$ equals a CM algebra E, and define $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. Let $\Phi \colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ be the signature defined in Lemma 2.72. Then $Z(\mathrm{MT}(V))^\circ \subseteq \mathrm{T}_E$ has cocharacter group equal to the smallest saturated Galois submodule of $\mathrm{X}_*(\mathrm{T}_E) = \mathbb{Z}[\Sigma_E^\vee]$ containing
$$\sum_{\sigma \in \Sigma_E} \Phi(\sigma)\sigma^\vee.$$

Proof. This follows from Corollary 2.76. By Lemma 1.41, it is enough to add in the cocharacter given by the scalars $\mathbb{G}_{m,\mathbb{Q}} \to \mathrm{T}_E$, which is $\sum_{\sigma\in\Sigma_E} \sigma^\vee$. Thus, the fact that Φ is a signature implies that

$$\sum_{\sigma\in\Sigma_E} \Phi(\sigma)\sigma^\vee$$

certainly lives in $\mathrm{X}_*(\mathrm{MT}(A)) \subseteq \mathrm{X}_*(\mathrm{T}_E)$.

Conversely, if X is some saturated Galois submodule containing $\sum_{\sigma\in\Sigma_E} \Phi(\sigma)\sigma^\vee$, then we would like to show that $\mathrm{X}_*(\mathrm{MT}(A)) \subseteq X$. Well, X is a Galois submodule, so it must contain the complex conjugate element $\sum_{\sigma\in\Sigma_E} \Phi(\overline{\sigma})\sigma^\vee$. On one hand, this then sums with the given element to produce

$$\sum_{\sigma\in\Sigma_E} \sigma^\vee \in X$$

because X is saturated. On the other hand, we can take a difference to see that

$$\sum_{\sigma\in\Sigma_E} (\Phi(\sigma) - \Phi(\overline{\sigma}))\sigma^\vee \in X.$$

We conclude that X contains the cocharacter of the scalars $\mathbb{G}_{m,\mathbb{Q}} \subseteq \mathrm{T}_E$ and the cocharacter lattice of $Z(\mathrm{Hg}(A))^\circ \subseteq \mathrm{T}_E$, so we conclude that X must also contain the cocharacter lattice of $Z(\mathrm{MT}(A))^\circ$. ∎

Remark 2.78. One can also prove the above corollary by following the proof of Corollary 2.76. For example, this approach provides a monodromy interpretation of $Z(\mathrm{MT}(A))^\circ$ analogous to Lemma 2.73. Here, one replaces the generator of $\mathrm{X}_*(\mathbb{U})$ with the cocharacter $\mu \in \mathrm{X}_*(\mathbb{S})$, and one finds that $\det_E \circ h_{\mathbb{C}}$ sends μ to $\sum_{\sigma\in\Sigma_E} \Phi(\sigma)\sigma^\vee$. One is then able to prove statements analogous to Proposition 2.74 and Corollary 2.76.

Let's pause for a moment with an explanation of how one can use Corollary 2.77 to compute $Z(\mathrm{MT}(A))^\circ \subseteq \mathrm{T}_E$. The approach for $Z(\mathrm{Hg}(A))^\circ$ is similar but only a little more complicated.

We will only compute over a Galois extension $L/\mathbb{Q}$ containing all factors of E. In this case, the E-action on V_L diagonalizes, so one can identify $(\mathrm{T}_E)_L \subseteq \mathrm{GL}(V)_L$ as the diagonal torus for some basis of V_L. In particular, for each $\sigma \in \Sigma_E$, the cocharacter $\sigma^\vee$ corresponds to one of the standard cocharacters for the diagonal torus of $\mathrm{GL}(V)_L$. Now, Corollary 2.77 tells us that $\mathrm{X}_*(Z(\mathrm{MT}(A))^\circ) \subseteq \mathrm{X}_*(\mathrm{T}_E)$ equals the saturation of the sublattice spanned by the vectors

$$g\left(\sum_{\sigma\in\Sigma_E} \Phi(\sigma)\sigma^\vee\right) = \sum_{\sigma\in\Sigma_E} \Phi(\sigma)(g\sigma)^\vee,$$

where g varies over $\mathrm{Gal}(L/E)$. By computing a basis of the saturation of this sublattice, we get a family of 1-parameter subgroups of the diagonal torus of $\mathrm{GL}(V)_L$ which together generate $Z(\mathrm{MT}(A))^\circ$. This more or less computes $Z(\mathrm{MT}(A))^\circ$.

2.2.3 Type IV: The Reflex

In the sequel, we will be most interested in equations cutting out $Z(\mathrm{MT}(A))^\circ \subseteq \mathrm{T}_E$. One could imagine proceeding as above to compute $Z(\mathrm{MT}(A))^\circ \subseteq \mathrm{T}_E$ via 1-parameter subgroups and then afterwards finding the desired equations. This is somewhat computationally intensive, so instead we will turn our attention to computing character groups. As in [Yu15, Lemma 4.2], this will require a discussion of the reflex.

Definition 2.79 (reflex signature)**.** Fix CM fields E and E^* and CM signatures (E,Φ) and (E^*,Φ^*). We say that these CM signatures are *reflex* if and only if there is a Galois extension $L/\mathbb{Q}$ containing E and E^* such that each $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ has

$$\Phi(\sigma|_E) = \Phi^*\left(\sigma^{-1}|_{E^*}\right).$$

In this situation, we may call (E^*,Φ^*) a *reflex signature* for (E,Φ).

Remark 2.80. We check that (E, Φ) and (E^*, Φ^*) does not depend on the choice of Galois extension L. Indeed, suppose that we have another Galois extension $L'/\mathbb{Q}$ containing E and E^*; let L'' be a Galois extension containing both L and L'. By symmetry, it is enough to check that (E, Φ) are reflex with respect to L if and only if they are reflex with respect to L''. Well, for any $\sigma \in \mathrm{Gal}(L''/\mathbb{Q})$, we see that $\Phi(\sigma|_E) = \Phi^*\left(\sigma^{-1}|_{E^*}\right)$ is equivalent to $\sigma|_L \in \mathrm{Gal}(L/\mathbb{Q})$ satisfying $\Phi(\sigma|_L|_E) = \Phi^*\left(\sigma|_L^{-1}|_{E^*}\right)$, so we are done after remarking that restriction $\mathrm{Gal}(L''/\mathbb{Q}) \to \mathrm{Gal}(L/\mathbb{Q})$ is surjective.

Remark 2.81. We check that reflex signatures certainly exist: one can choose any Galois closure L of E and then define $\Phi^*\colon \mathrm{Gal}(L/\mathbb{Q}) \to \mathbb{Z}_{\geq 0}$ by $\Phi^*(\sigma) \coloneqq \Phi\left(\sigma^{-1}|_L\right)$.

Remark 2.82. In the theory of abelian varieties with complex multiplication, it is customary to make E^* as small as possible, which makes it unique. This is useful for moduli problems. However, this is not our current interest, and we are not requiring that the reflex signature be unique because it will be convenient later to take large extensions.

The point of introducing the reflex is that it provides another monodromy interpretation of $Z(\mathrm{MT}(A))^\circ$. To achieve this, we need the reflex norm.

Definition 2.83 (reflex norm)**.** Fix CM fields E and E^* and reflex CM signatures (E, Φ) and (E^*, Φ^*). Then we define the *reflex norm* as the map $\mathrm{N}_{\Phi^*}\colon E^* \to \overline{\mathbb{Q}}$ by

$$\mathrm{N}_{\Phi^*}(x) \coloneqq \prod_{\sigma \in \Sigma_{E^*}} \sigma(x)^{\Phi^*(\sigma)}.$$

Note that this is a character in $\mathrm{X}^*(\mathrm{T}_{E^*})$.

Technically, this definition does not require us to remember that (E^*, Φ^*) is reflex to (E, Φ), but we will want to know this in the following checks.

Lemma 2.84. Fix CM fields E and E^* and reflex CM signatures (E, Φ) and (E^*, Φ^*).

(a) If (E_1^*, Φ_1^*) is a CM signature restricting to (E^*, Φ^*), then (E, Φ) and (E_1^*, Φ_1^*) are still reflex, and

$$\mathrm{N}_{\Phi_1^*} = \mathrm{N}_{\Phi^*} \circ \mathrm{N}_{E_1^*/E^*}.$$

(b) The image of N_{Φ^*} lands in E.

Proof. Here, "restricting" simply means that E_1^* contains E^* and $\Phi_1^*(\sigma) = \Phi^*(\sigma|_{E^*})$ for all $\sigma \in \Sigma_{E_1^*}$.

(a) That (E, Φ) and (E_1^*, Φ_1^*) are still reflex follows from the definition: choose a Galois extension L containing E and E_1^*, and then each $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ has

$$\begin{aligned}\Phi(\sigma|_E) &= \Phi^*\left(\sigma^{-1}|_{E^*}\right) \\ &= \Phi_1^*\left(\sigma^{-1}|_{E_1^*}\right).\end{aligned}$$

To check the equality of reflex norms, we extend each $\sigma \in \Sigma_{E^*}$ to some $\widetilde{\sigma} \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and then we

directly compute

$$\begin{aligned}\mathrm{N}_{\Phi^*}\left(\mathrm{N}_{E_1^*/E^*}(x)\right) &= \prod_{\sigma\in\Sigma_{E^*}} \sigma\left(\mathrm{N}_{E_1^*/E^*}(x)\right)^{\Phi^*(\sigma)} \\ &= \prod_{\substack{\sigma\in\Sigma_E^* \\ \tau\in\mathrm{Hom}_{E^*}(E_1^*,\overline{\mathbb{Q}})}} \widetilde{\sigma}\tau(x)^{\Phi^*(\sigma)} \\ &= \prod_{\substack{\sigma\in\Sigma_E^* \\ \tau\in\mathrm{Hom}_{E^*}(E_1^*,\overline{\mathbb{Q}})}} \widetilde{\sigma}\tau(x)^{\Phi_1^*(\widetilde{\sigma}\tau)} \\ &= \mathrm{N}_{\Phi_1^*}(x),\end{aligned}$$

where the last step holds by noting that $\widetilde{\sigma}\circ\tau$ parameterizes Σ_{E^*}.

(b) We begin by reducing to the case where $E^*/\mathbb{Q}$ is Galois. Indeed, the previous step tells us that extending E^* merely passes to a norm subgroup of E^*, but norm subgroups are Zariski dense in T_{E^*}, so it suffices to check the result on such norm subgroups. Thus, we may assume that $E^*/\mathbb{Q}$ is Galois, contains E, and thus $\Phi^*(\sigma) = \Phi\left(\sigma^{-1}|_E\right)$. Now, for any $g\in\mathrm{Gal}(E^*/E)$, we see $\Phi^*(\sigma) = \Phi^*\left(g^{-1}\sigma\right)$, so

$$\begin{aligned}g\left(\mathrm{N}_{\Phi^*}(x)\right) &= \prod_{\sigma\in\mathrm{Gal}(E^*/\mathbb{Q})} g\sigma(x)^{\Phi^*(\sigma)} \\ &= \prod_{\sigma\in\mathrm{Gal}(E^*/\mathbb{Q})} \sigma(x)^{\Phi^*\left(g^{-1}\sigma\right)} \\ &= \mathrm{N}_{\Phi^*}(x),\end{aligned}$$

as required. ∎

At long last, we move towards our monodromy interepretation using the reflex. The following argument generalizes [Yu15, Lemma 4.2].

Lemma 2.85. Fix reflex CM signatures (E,Φ) and (E^*,Φ^*). Suppose that E^* contains E and is Galois over $\mathbb{Q}$. For each $g\in\mathrm{Gal}(E^*/\mathbb{Q})$, the reflex norm $\mathrm{N}_{\Phi^*}\colon \mathrm{T}_{E^*}\to\mathrm{T}_E$ sends the cocharacter $g^\vee\in\mathrm{X}_*(\mathrm{T}_{E^*})$ to

$$\mathrm{X}_*\left(\mathrm{N}_{\Phi^*}\right)\left(g^\vee\right) = \sum_{\sigma\in\Sigma_E}\Phi(\sigma)(g\sigma)^\vee.$$

Proof. Notably, N_{Φ^*} outputs to T_E by Lemma 2.84. To begin, we expand

$$\mathrm{X}_*\left(\mathrm{N}_{\Phi^*}\right)\left(g^\vee\right) = \sum_{\sigma\in\Sigma_{E^*}}\Phi^*(\sigma)\mathrm{X}_*(\sigma)(g^\vee).$$

We now check $\mathrm{X}_*(\sigma)(g^\vee) = \left(g\sigma^{-1}\right)^\vee$: for any $\tau\in\mathrm{X}^*(\mathrm{T}_{E^*})$, we compute the perfect pairing

$$\langle\tau,\mathrm{X}_*(\sigma)(g^\vee)\rangle = \langle\tau\sigma,g^\vee\rangle,$$

which is the indicator function for $\tau\sigma = g$ and hence equals $\left\langle\cdot,(g\sigma^{-1})^\vee\right\rangle$. We are now able to write

$$\mathrm{X}_*\left(\mathrm{N}_{\Phi^*}\right)\left(g^\vee\right) = \sum_{\sigma\in\Sigma_{E^*}}\Phi^*(\sigma)\left(g\sigma^{-1}\right)^\vee.$$

Replacing σ with σ^{-1}, we are done upon recalling $\Phi^*\left(\sigma^{-1}\right) = \Phi(\sigma|_E)$ and collecting terms which together restrict to the same embedding of E. ∎

Proposition 2.86. Fix an abelian variety A over $\mathbb{C}$ such that $Z(\mathrm{End}(A))$ equals a CM algebra $E = E_1 \times \cdots \times E_k$, and define $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$. Let $\Phi\colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ be the induced signature, which we decompose as $\Phi = \Phi_1 \sqcup \cdots \sqcup \Phi_k$ where $(E_\bullet, \Phi_\bullet)$ is a CM signature for all $E_\bullet$. Suppose E^* is a CM field equipped with CM signatures $\Phi_1^*, \ldots, \Phi_k^*$ such that (E_i, Φ_i) and (E^*, Φ_i^*) are reflex for all i. Then $Z(\mathrm{MT}(A))^\circ \subseteq \mathrm{T}_E$ is the image of

$$(\mathrm{N}_{\Phi_1^*}, \ldots, \mathrm{N}_{\Phi_k^*})\colon \mathrm{T}_{E^*} \to \mathrm{T}_E.$$

Proof. Note that norms are surjective on these algebraic tori, so Lemma 2.84 tells us that the image of N_{Φ^*} will not change if we pass to an extension of E^*. As such, we will go ahead and assume that E^* contains E and is Galois over $\mathbb{Q}$.

In light of Corollary 2.77, it is enough to show that the image of $\mathrm{X}_*(\mathrm{N}_{\Phi^*})$ (which we note is already a Galois submodule) has saturation equal to the smallest saturated Galois submodule of $\mathrm{X}_*(\mathrm{T}_E)$ containing $\sum_{\sigma \in \Sigma_E} \Phi(\sigma)\sigma^\vee$. This follows from the computation of Lemma 2.85 upon letting g vary over $\mathrm{Gal}(E^*/\mathbb{Q})$. ∎

Let's explain how Proposition 2.86 is applied to compute equations cutting out $Z(\mathrm{MT}(A))^\circ \subseteq \mathrm{T}_E$, where $E = E_1 \times \cdots \times E_k$ is a CM algebra. As before, we will only compute over an extension $L = E^*$ of E which is Galois over $\mathbb{Q}$; let $\Phi_1^*, \ldots, \Phi_k^*$ be the signatures on L making (L, Φ_i^*) and (E_i, Φ_i) reflex for each i. Note, we know that $(\mathrm{T}_E)_L \subseteq \mathrm{GL}(V)_L$ may embed as a diagonal torus.

An equation cutting out $Z(\mathrm{MT}(A))^\circ_L$ in the (subtorus of the) diagonal torus $(\mathrm{T}_E)_L \subseteq \mathrm{GL}(V)_L$ then becomes a character of $(\mathrm{T}_E)_L$ which is trivial on $Z(\mathrm{MT}(A))^\circ$. In other words, these equations are given by the kernel of

$$\mathrm{X}^*(\mathrm{T}_E) \to \mathrm{X}^*(Z(\mathrm{MT}(A))^\circ).$$

We now use Proposition 2.86. We know that $Z(\mathrm{MT}(A))^\circ \subseteq \mathrm{T}_E$ is the image of $(\mathrm{N}_{\Phi_1^*}, \ldots, \mathrm{N}_{\Phi_k^*})\colon \mathrm{T}_L \to \mathrm{T}_E$, so the kernel of the above map is the same as the kernel of

$$\mathrm{X}^*\left((\mathrm{N}_{\Phi_1^*}, \ldots, \mathrm{N}_{\Phi_k^*})\right) : \mathrm{X}^*(\mathrm{T}_E) \to \mathrm{X}^*(\mathrm{T}_L).$$

To compute this kernel cleanly, note Lemma 2.85 computes $\mathrm{X}_*\left(\mathrm{N}_{\Phi_i^*}\right)$ for each i, so we see $\mathrm{X}^*\left(\mathrm{N}_{\Phi_i^*}\right)$ can be computed as the transpose of the matrix of $\mathrm{X}_*\left(\mathrm{N}_{\Phi^*}\right)$. Attaching these matrices together gives a matrix representation for the above map, and we get our equations by computing the kernel of this matrix.

Remark 2.87. In practice, one can expand $V = V_1 \oplus \cdots \oplus V_k$ into irreducible Hodge substructures and then work with $E \coloneqq E_1 \times \cdots \times E_k$ where $E_i \coloneqq Z(\mathrm{End}_{\mathrm{HS}}(V_i))$ for each i. Technically speaking, E may only embed into E "diagonally" because some $V_\bullet$s may be isomorphic to each other. However, this does not really affect anything we do because we may as well work with the image of $Z(\mathrm{MT}(A))^\circ$ under the inclusion $\mathrm{T}_E \subseteq \mathrm{T}_E$. Working with T_E is more convenient because it can actually be identified with the diagonal torus of $\mathrm{GL}(V)_E$ instead of merely a diagonally embedded subtorus.

2.3 The ℓ-Adic Representation

In this subsection, we now define the ℓ-adic representation and give some of its basic properties.

2.3.1 The Cohomology of Abelian Varieties

Fix an abelian variety A over a field K. In this section, we will compute the Weil cohomology ring $\mathrm{H}^\bullet(A)$, for many Weil cohomology theories $\mathrm{H}^\bullet$ defined over K with coefficients in F. As usual, $\mathrm{char}\, F = 0$. More precisely, we will show that $\dim_F \mathrm{H}^1(A) = 2\dim A$ implies that the cup product defines an isomorphism

$$\wedge^\bullet \mathrm{H}^1(A) \to \mathrm{H}^\bullet(A).$$

As is usual with our discussions of Weil cohomology, our argument will have a linear algebraic component and a motivic component; the "motivic" component will be this equality $\dim_F \mathrm{H}^1(A) = 2\dim A$. Our exposition follows [EGM, Corollary 13.32].

Let's begin with the linear algebraic component. Our exposition follows [Hat01, Section 3.C]. The key point is that the group structure on A will endow $\mathrm{H}^\bullet(A)$ with extra structure: $\mathrm{H}^\bullet(A)$ becomes a Hopf algebra. The following definition is a bit non-standard, but it will suffice for our purposes.

Definition 2.88 (Hopf algebra)**.** Fix a field F. A *graded Hopf algebra* $H^\bullet$ over F is a $\mathbb{Z}_{\geq 0}$-graded commutative algebra over F equipped with graded F-algebra homomorphisms $e\colon H^\bullet \to F$ (called the co-unit) and $m\colon H^\bullet \to H^\bullet \otimes H^\bullet$ (called the co-multiplication). Further, e and m are required to satisfy the following.

(a) Co-identity: $(e \otimes \mathrm{id}) \circ m$ and $m \circ (e \otimes \mathrm{id})$ both equal to $\mathrm{id}\colon H^\bullet \to H^\bullet$.

(b) Co-associativity: we have $(m \otimes \mathrm{id}) \circ m = m \circ (m \otimes \mathrm{id})$ as maps $H^\bullet \to H^\bullet \otimes H^\bullet \otimes H^\bullet$.

If the structure map $F \to H^0$ is an isomorphism, then $H^\bullet$ is *connected*.

It turns out that we can get by with less information. For our purposes, we will really only need the following fact about the co-multiplication.

Lemma 2.89. Let $H^\bullet$ be a connected, graded Hopf algebra over a field F.

(a) The co-unit $e\colon H^\bullet \to F$ is the inverse of $F \to H^0$ in degree 0 and vanishes in higher degrees.

(b) For each $\alpha \in H^n$ with $n > 0$, we have

$$m(\alpha) - (\alpha \otimes 1 + 1 \otimes \alpha) \in \bigoplus_{i,j>0} H^i \otimes H^j.$$

Proof. We show each part in separately.

(a) Because e is a homomorphism of graded F-algebras, e automatically vanishes in positive degrees. As for degree 0, we already know that the structure map $F \to H^0$ is an isomorphism, so $e\colon H^0 \to F$ must be its inverse because it maps the "basis vector" $1 \in H^0$ back to $1 \in F$.

(b) This follows from the co-identity axiom. To begin, the grading structure on $H^\bullet \otimes H^\bullet$ implies that we may write

$$m(\alpha) = \sum_{i,j=0}^{\infty} \alpha_i \otimes \alpha'_j,$$

where $\alpha_i, \alpha'_i \in H^i$ for each i. Thus, applying $(e \otimes \mathrm{id})$ to this expression reveals

$$\alpha = \sum_{j=0}^{\infty} \alpha_0 \otimes \alpha'_j.$$

We conclude that α'_j automatically vanishes except at degree n, where $\alpha = \alpha_0 \alpha'_n$. A symmetric argument (using $(\mathrm{id} \otimes e) \circ m = \mathrm{id}$) shows that α_i vanishes except at degree n, where $\alpha = \alpha'_0 \alpha_n$. We conclude that

$$m(\alpha) - (\alpha \otimes 1 + 1 \otimes \alpha) = \sum_{i,j=1}^{\infty} \alpha_i \otimes \alpha'_j,$$

as required. ∎

Here are some basic examples.

Example 2.90. If $A^\bullet$ and $B^\bullet$ are graded commutative Hopf algebras over F, then $A^\bullet \otimes B^\bullet$ is as well.

Proof. Let e_A and m_A be the co-unit and co-multiplication for $A^\bullet$, respectively; define e_B and m_B analogously for $B^\bullet$. Now, we can define $e\colon A^\bullet \otimes B^\bullet \to F$ by $e_A \otimes e_B$, and we define $m\colon (A^\bullet \otimes B^\bullet) \to (A^\bullet \otimes B^\bullet)^{\otimes 2}$ on pure homogeneous tensors by $m(a \otimes b) \coloneqq (-1)^{\deg(a)\deg(b)} m(a) \otimes m(b)$, where we have identified

$$(A^\bullet \otimes B^\bullet)^{\otimes 2} = (A^\bullet)^{\otimes 2} \otimes (B^\bullet)^{\otimes 2}$$

by swapping the middle two entries. Here are our checks.

- Homomorphisms: note e is a homomorphism because it is the tensor product of two homomorphisms (e_A and e_B). Similarly, m is also a tensor product of two homomorphisms (m_A and m_B) but now followed up with a swap
$$B^\bullet \otimes A^\bullet \to A^\bullet \otimes B^\bullet,$$
which we can see is also a homomorphism of graded algebras. (The sign is present to account for graded commutativity!)
- Co-identity: this follows by taking the tensor product of the co-identity axioms for A and B and then swapping to correct the order of the factors. We won't write out these manipulations.
- Co-associativity: the same discussion as for co-identity applies. ∎

Example 2.91. Let V be a graded vector space over F with $\operatorname{char} F \neq 2$, supported in positive degree.

(a) If V is supported in even degree, then the symmetric algebra $S^\bullet V$ is a graded commutative Hopf algebra.

(b) If V is supported in odd degree, then the exterior algebra $\wedge^\bullet V$ is a graded commutative Hopf algebra.

Proof. In both cases, let the given algebra be $A^\bullet$, and then the co-unit $e\colon A^\bullet \to F$ is defined in degree 0 by $\mathrm{id}\colon A^0 \to F$ and vanishing in higher degrees. Additionally, the co-multiplication is defined by $m\colon A^\bullet \to A^\bullet \otimes A^\bullet$ is defined by extending $m(v) \coloneqq (1 \otimes v) + (v \otimes 1)$ for any $v \in V$. Notably, given vectors $v_1, \dots, v_n \in V$, we see that we have to define

$$m(v_1 \otimes \cdots \otimes v_n) \coloneqq \prod_{i=1}^{n} (v_i \otimes 1 + 1 \otimes v_i).$$

It remains to run our checks.

- Connected: because $V^0 = 0$, we have $A^\bullet V = F$ in both cases.
- Homomorphisms: in both cases, e can be described as the quotient of the functional on the tensor algebra $T^\bullet V$ which just sends all the generators to 0. This functional $T^\bullet V \to F$ is a homomorphism of graded F-algebras, so e is as well.

  It remains to check that m is a well-defined homomorphism. Once again, m begins its life as a graded linear map $T^\bullet V \to T^\bullet V \otimes T^\bullet V$ on the tensor algebra, given by the above formula on pure tensors. We now go down to $A^\bullet$ in cases.

  - If V is supported in even degrees, then we consider the quotient map $T^\bullet V \to S^\bullet V \otimes S^\bullet V$. For any $v, w \in V$, we can compute that $m(v \otimes w)$ and $m(w \otimes v)$ are both equal to $2(vw \otimes 1 + 1 \otimes vw)$ by commutativity, so we find that this descends to a well-defined graded linear map $S^\bullet V \to S^\bullet V \otimes S^\bullet V$. As for multiplicativity, it is now enough to check multiplicativity on generators, where it follows by definition.
  - If V is supported in odd degrees, then we consider the quotient map $T^\bullet V \to \wedge^\bullet V \to \wedge^\bullet V$. Once again, for any $v \in V$, we can compute that $m(v \otimes v) = 2(v \otimes v)$, which vanishes because $(v \otimes 1)(1 \otimes v) = -(1 \otimes v)(v \otimes 1)$ forces $v \otimes v = 0$. Thus, we descend to a graded linear map $\wedge^\bullet V \to \wedge^\bullet V \otimes \wedge^\bullet V$, and multiplicativity follows because it is true on generators by definition.

- Co-identity: it is enough to check the equality of two maps $A^\bullet \to A^\bullet$ on generators. By symmetry, it will be enough to check that $(e \otimes \mathrm{id}) \circ m = \mathrm{id}$. Well, for any $v \in V$, we see that $(e \otimes \mathrm{id})(m(v))$ is

$$(e \otimes \mathrm{id})(v \otimes 1 + 1 \otimes v) = v.$$

- Co-associativity: again, it is enough to check the equality of two maps $A^\bullet \to (A^\bullet)^{\otimes 3}$ on generators. As such, for any $v \in V$, we compute that $(m \otimes \mathrm{id})(m(v))$ is

$$(m \otimes \mathrm{id})(1 \otimes v + v \otimes 1) = 1 \otimes 1 \otimes v + 1 \otimes v \otimes 1 + v \otimes 1 \otimes 1,$$

  which by a similar argument is the same as $(\mathrm{id} \otimes m)(m(v))$. ■

Remark 2.92. Fix graded vector spaces V and W supported in odd positive degree. (There is an analogous remark for positive even degree.) Then the inclusions provide a canonical graded linear map

$$\wedge^\bullet V \otimes \wedge^\bullet W \to \wedge^\bullet (V \oplus W).$$

(Explicitly, this map sends $v \otimes 1 \mapsto v$ and $1 \otimes w \mapsto w$.) If V and W are finite-dimensional, then one can see on graded components that this graded linear map restricts to a bijection of bases, so this is an isomorphism. In fact, this map can be quickly checked to be multiplicative, and in fact it is an isomorphism of graded commutative Hopf algebras.

Unsurprisingly, here is our main example.

Example 2.93. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over K with coefficients in F. For any abelian variety A, the graded F-algebra $\mathrm{H}^\bullet(A)$ has the structure of a connected, graded commutative Hopf algebra over F.

Proof. It suffices to define the co-unit and the co-multiplication, and then we need to check the required properties. Here is the data.

- Co-unit: the identity map $e\colon \operatorname{Spec} K \to A$ of our abelian variety defines a pullback $e^*\colon \mathrm{H}^\bullet(A) \to \mathrm{H}^\bullet(\operatorname{Spec} K)$. Because $\mathrm{H}^\bullet(\operatorname{Spec} K) = F$ by Example 1.106, we may let e^* be our co-unit.
- Co-multiplication: the multiplication map $m\colon A \times A \to A$ defines a pullback $m^*\colon \mathrm{H}^\bullet(A) \to \mathrm{H}^\bullet(A \times A)$. This becomes our co-multiplication as soon as we identify $\mathrm{H}^\bullet(A \times A) = \mathrm{H}^\bullet(A) \otimes \mathrm{H}^\bullet(A)$ via the Künneth formula.

And here are our checks.

- Co-identity: by symmetry, it will be enough to check $(e^* \otimes \mathrm{id}_{\mathrm{H}^\bullet(A)}) \circ m^* = \mathrm{id}_{\mathrm{H}^\bullet(A)}$. This comes from the identity law on the abelian variety, which tells us $m \circ (e \times \mathrm{id}_A) = \mathrm{id}_A$. Indeed, this implies that

$$(e \times \mathrm{id}_A)^* \circ m^* = \mathrm{id}_A^*\,.$$

  We can see that $\mathrm{id}_A^* = \mathrm{id}_{\mathrm{H}^\bullet(A)}$, so we are done as soon as we note that $(e \times \mathrm{id}_A)^* = e^* \otimes \mathrm{id}_A^*$ by Lemma 1.93.

- Co-associativity: this follows from the associativity law on abelian varieties. Indeed, we know that

$$m \circ (m \times \mathrm{id}_A) = m \circ (\mathrm{id}_A \times m),$$

  so taking pullbacks gives

$$(m \times \mathrm{id}_A)^* \circ m^* = (\mathrm{id}_A \times m)^* \circ m^*.$$

  We are done after plugging in Lemma 1.93.

- Connected: this follows because A is proper and geometrically irreducible. Indeed, this implies that $\Gamma(A, \mathcal{O}_A) = K$, so the fact that $\mathrm{H}^\bullet$ is a Weil cohomology theory enforces a structure isomorphism $\mathrm{H}^0(\operatorname{Spec} K) \to \mathrm{H}^0(A)$. But $\mathrm{H}^0(K) = F$ by Example 1.106, so we are done. ■

The benefit to having given ourselves extra structure is that it severely cuts down on the possibilities for the ring structure of $\mathrm{H}^\bullet(A)$.

Theorem 2.94 (Hopf)**.** Let $H^\bullet$ be a connected, graded commutative Hopf algebra over F, where $\operatorname{char} F = 0$. Suppose that $\dim_F H^i < \infty$ for all i. Then the F-algebra $H^\bullet$ is isomorphic to a tensor product of exterior and symmetric power algebras.

Proof. We follow [Hat01, Theorem 3.C.4]. We proceed in steps.

1. Let's set up some generators. Because $\dim_F H^i < \infty$ for all i, we may find a countable list $\{x_1, x_2, \ldots\}$ of generators of $H^\bullet$ as an F-algebra. By decomposing these generators into homogeneous componenents, we may assume that our generators are homogeneous of positive degree. Additionally, the finite-dimensional constraint implies that we may rearrange our generators so that $\deg x_1 \le \deg x_2 \le \cdots$.

2. Our proof is going to proceed by induction, so let's set this up. For each $n \ge 0$, set $H_n^\bullet$ to be the connected, graded commutative F-algebra generated by the elements $\{x_1, \ldots, x_n\}$. For example, for each n, $H_n^\bullet$ has all the needed generators in degree less than $\deg x_n$, so

$$H_n^i \subseteq H_n^\bullet$$

for each $i < \deg x_n$. Quckly, we claim that $H_n^\bullet \subseteq H^\bullet$ is a Hopf subalgebra. For example, one can simply restrict the co-identity e to $H_n^\bullet$, and one can use Lemma 2.89 to see that m also restricts to $H_n^\bullet$: by induction, it is enough to check $m(x_n) \in H_n^\bullet$, which is true because $m(x_n)$ only uses the terms x_n and ones of strictly smaller degree! The co-identity and co-associativity axioms now hold by restriction.

We will use induction to show that $H_n^\bullet$ is a tensor product of exterior and symmetric power algebras for each $n \ge 0$. Because $H^\bullet = \bigcup_n H_n^\bullet$, the conclusion will then follow for $H^\bullet$ because tensor products commute with colimits. As for our induction, we quickly note that $H_0^\bullet = F$, so there is noting to show for our base case $n = 0$.

3. We explain the main claim in the inductive step. Suppose $H_n^\bullet$ is a product of exterior and symmetric power algebras, and we want to show the same for $H_{n+1}^\bullet$. If $x_{n+1} \in H_n^\bullet$, then $H_{n+1}^\bullet = H_n^\bullet$, and there is nothing to do. Otherwise, we let V be the one-dimensional vector space spanned by x_{n+1}. Then there is a canonical map $A^\bullet V \to H_{n+1}^\bullet$ of F-algebras, where

$$A^\bullet V \coloneqq \begin{cases} \wedge^\bullet V & \text{if } \deg x_{n+1} \text{ is even,} \\ S^\bullet V & \text{if } \deg x_{n+1} \text{ is odd.} \end{cases}$$

Indeed, there is certainly a canonical map $T^\bullet V \to H_{n+1}^\bullet$ sending $x_{n+1} \mapsto x_{n+1}$, but $T^\bullet V = S^\bullet V$, so we are done in the even-degree case. In the odd-degree case, it remains to note that $x_{n+1}^2 = 0$ if $\deg x_{n+1}$ is odd, so our map descends to $\wedge^\bullet V$. Now, because $H_{n+1}^\bullet$ is generated by $H_n^\bullet$ and x_{n+1}, there is a canonical surjection

$$p \colon H_n^\bullet \otimes A^\bullet V \to H_{n+1}^\bullet.$$

The main claim of this proof is that this map is injective, hence an isomorphism, which completes the inductive step and hence the proof.

4. We define a linear map which will help us "take derivatives." Let $I \subseteq H_{n+1}^\bullet$ be the ideal generated by $H_n^\bullet$ and x_{n+1}^2, and we will be interested in the map $f \colon H_{n+1}^\bullet \to H_{n+1}^\bullet \otimes H_{n+1}^\bullet / I$ defined by the composite

$$H_{n+1}^\bullet \xrightarrow{m} H_{n+1}^\bullet \otimes H_{n+1}^\bullet \twoheadrightarrow H_{n+1}^\bullet \otimes H_{n+1}^\bullet / I.$$

For example, any $\alpha \in H_n^\bullet$ vanishes in the quotient, so it goes to $\alpha \otimes 1$ (where we have quietly used Lemma 2.89). Also, x_{n+1} goes to $x_{n+1} \otimes 1 + 1 \otimes x_{n+1}$ because the remaining terms given in Lemma 2.89 all live in I. We conclude that a generic element $\sum_{i=0}^\infty \alpha_i x_{n+1}^i$ of $H_{n+1}^\bullet$ is mapped to

$$\sum_{i=0}^\infty (\alpha_i \otimes 1)(x_{n+1} \otimes 1 + 1 \otimes x_{n+1})^i = \sum_{i=0}^\infty \alpha_i x_{n+1}^i \otimes 1 + \sum_{i=1}^\infty i\alpha_i x_{n+1}^{i-1} \otimes x_i.$$

5. We show that p is injective $\deg x_{n+1}$ is odd. In this case, a generic element of $H_n^\bullet \otimes \wedge^\bullet V$ looks like $\alpha_0 \otimes 1 + \alpha_1 \otimes x_{n+1}$ for some $\alpha_0, \alpha_1 \in H_n^\bullet$. If this element lived in $\ker p$, then $\alpha_0 + \alpha_1 x_{n+1} = 0$, and we will show that $\alpha_0 = \alpha_1 = 0$. Indeed, passing the relation $\alpha_0 + \alpha_1 x_{n+1} = 0$ through f, we find that $\alpha_1 = 0$, so $\alpha_0 = 0$ follows.

6. We show that p is surjective when $\deg x_{n+1}$ is even. In this case, a generic element of $H_n^\bullet \otimes S^\bullet V$ looks like $\beta \coloneqq \sum_{i=0}^d \alpha_i \otimes x_{n+1}^i$. We will show that $\beta \in \ker p$ implies $\beta = 0$ by induction on d. Indeed, given $\beta \in \ker p$, we can pass the equality $\sum_{i=0}^d \alpha_i x_{n+1}^i = 0$ through f to see

$$\sum_{i=1}^d i\alpha_i x_{n+1}^{i-1} = 0.$$

This is some element with strictly smaller x_{n+1}-degree, so we see that $i\alpha_i = 0$ for $i \in \{1, \dots, d\}$. Thus, $\beta = \alpha_0$, but now $\alpha_0 = p(\beta) = 0$ as well. ■

Remark 2.95. This proof does not use the co-associativity axiom anywhere.

Corollary 2.96. Let $H^\bullet$ be a connected, graded commutative Hopf algebra over F, where $\operatorname{char} F = 0$. If $\dim_F H^\bullet < \infty$, then $H^\bullet$ is isomorphic (as an F-algebra) to $\wedge^\bullet V$ for some graded vector space V supported in odd degrees.

Proof. Because $\dim_F H^\bullet < \infty$, we see that Theorem 2.94 forces $H^\bullet$ to be a finite tensor product of exterior and symmetric power algebras, but symmetric power algebras are infinite-dimensional and hence also disallowed. The result now follows from Remark 2.92. ■

And here is our application of this linear algebraic input.

Proposition 2.97. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over a field K with coefficients in F. For any abelian variety A over K, if $\dim_F \mathrm{H}^1(A) = 2 \dim A$, then the cup product defines an isomorphism

$$\wedge^\bullet \mathrm{H}^1(A) \to \mathrm{H}^\bullet(A)$$

of graded commutative F-algebras.

Proof. We proceed in steps. Set $g \coloneqq \dim A$ for brevity.

1. By Example 2.93, we find that $\mathrm{H}^\bullet(A)$ is a connected, graded commutative Hopf algebra over F, and Poincaré duality (and Lemma 1.120) tell us that $\dim_F \mathrm{H}^\bullet(A) < \infty$, so Corollary 2.96 kicks in to provide with an isomorphism

$$\mathrm{H}^\bullet(A) \cong \wedge^\bullet V_1 \otimes \wedge^\bullet V_3 \otimes \cdots$$

of F-algebras, where V_i is some finite-dimensional graded vector space over F supported in degree i. Because $\mathrm{H}^\bullet(A)$ is only supported in degrees $i \in [0, 2g]$, we see immediately that $V_i = 0$ for $i > 2g$.

2. For each i, set $d_i \coloneqq \dim_F V_i$. We provide a relation among the d_is. Indeed, $\dim_F \mathrm{H}^0(A) = 1$ because A is geometrically irreducible (so that $\Gamma(A, \mathcal{O}_A) = K$, where we are quietly using Example 1.106), so $\dim_F \mathrm{H}^{2g}(A) = 1$ as well by Poincaré duality. The only way to get a one-dimesional vector space out of our tensor product of exterior powers is to have
$$\mathrm{H}^{2g}(A) \cong \wedge^{d_1} V_1 \wedge \cdots \wedge^{d_{2g-1}} V_{2g-1},$$
so $d_1 + \cdots + d_{2g-1} = 2g$ follows.

3. We complete the proof. Because we have assumed $d_1 = 2g$, we see that $d_i = 0$ for all other i. Now, checking degree 1 reveals our isomorphism must send $\mathrm{H}^1(A) \cong V^1$, so inverting this produces our required isomorphism $\wedge^\bullet \mathrm{H}^1(A) \to \mathrm{H}^\bullet(A)$ of F-algebras. ■

Theorem 2.98. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over a field K with coefficients in F, among those defined in section 1.3.1. For any abelian variety A over K, the cup product defines an isomorphism
$$\wedge^\bullet \mathrm{H}^1(A) \to \mathrm{H}^\bullet(A)$$
of graded commutative F-algebras.

Proof. By Proposition 2.97, we must show that $\dim_F \mathrm{H}^1(A) = 2g$, where $g \coloneqq \dim A$. We proceed in cases.

- Suppose that A is defined over $\mathbb{C}$, and we show $\dim_{\mathbb{Q}} \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q}) = 2g$; note this gives $\dim_{\mathbb{R}} \mathrm{H}^1_{\mathrm{dR}}(A, \mathbb{R}) = 2g$ as well by the comparison isomorphism in Theorem 1.75. We proceed as in [Mil20b, Proposition 2.6]. Write $A = \mathbb{C}^g/\Lambda$ for a lattice Λ. Fixing some index p, we will show that the cup product defines an isomorphism
$$\dim_{\mathbb{Q}} \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Z}) = 2g.$$
Well, we note that A is homeomorphic to $\left(S^1\right)^{2g}$, so the Künneth formula allows us to reduce the question to S^1, where the result is true by a direct computation.
- It remains to handle ℓ-adic cohomology. In this case, we must show that $\dim_{\mathbb{Q}_\ell} \mathrm{H}^1_{\text{ét}}(A, \mathbb{Q}_\ell) = 2g$. In the following section, we will show that $\mathrm{H}^1_{\text{ét}}(A, \mathbb{Q}_\ell)$ is dual to the ℓ-adic Tate module $T_\ell A$, which can be directly computed to be $2g$-dimensional. ■

Remark 2.99. One does have $\dim_F \mathrm{H}^1(A) = 2 \dim A$ for any Weil cohomology theory $\mathrm{H}^\bullet$, but this requires more significant motivic input (and possibly more linear algebraic input) than we would like to introduce here. We refer to [EGM, Corollary 13.32] for a proof.

Because we are able to prove a theorem for many cohomology theories, it should not be surprising that we can show a motivic variant as well.

Definition 2.100. Let $\mathcal{C}$ be an abelian symmetric monoidal category. For any object $X \in \mathcal{C}$ and some $n \geq 0$, we consider the natural action of S_n on $X^{\otimes n}$. Then we define $\mathrm{Sym}^n X$ as the eigenspace of the trivial character $S_n \to \{\pm 1\}$. Further, we define $\mathrm{Sym}^\bullet X \coloneqq \bigoplus_{n \geq 0} \wedge^n X$ provided that this sum exists.

Remark 2.101. Equivalently, if $\mathcal{C}$ is $\mathbb{Q}$-linear, we may define $\wedge^n X$ as the kernel of the idempotent
$$\frac{1}{n!} \sum_{\sigma \in S_n} \sigma \in \mathbb{Q}[S_n]$$
acting on X. This is a definition which also works for symmetric monoidal, Karoubian categories.

Example 2.102. If $\mathcal{C} = \mathrm{Rep}_F(G)$ for some affine group G over F, then $\mathrm{Sym}^n V$ and $\mathrm{Sym}^\bullet V$ are exactly the expected objects for any $V \in \mathrm{Rep}_F(G)$.

Example 2.103. One must be careful with $\mathrm{Mot}_{\mathbb{Q}}(K)$ because the symmetric monoidal structure is set up to be graded commutative. For any $X \in \mathcal{P}(K)$, we in fact claim that

$$\mathrm{H}^\bullet_\sigma\left(\mathrm{Sym}^\bullet h^1(X)\right) = \wedge^\bullet \mathrm{H}^1_\sigma(X),$$

where the right-hand exterior product is the usual exterior power of vector spaces. The point is that the exterior power of vector spaces will actually be the symmetric power for H^1_σ in the category of graded vector spaces because the grading adds a sign for every transposition. The result follows by properties of the fiber functor $\mathrm{H}^\bullet_\sigma$ (see Theorem 1.210).

Corollary 2.104. Fix an abelian variety A over a field K algebraic over $\mathbb{Q}$. Then the cup product defines an isomorphism

$$\mathrm{Sym}^\bullet h^1(A) \to h(A)$$

of motives in $\mathrm{Mot}_{\mathbb{Q}}(K)$.

Proof. Quickly, we refer to Example 2.103 to explain why we are taking the symmetric power instead of the exterior power in the statement. The map $\wedge^\bullet \mathrm{H}^1(A) \to \mathrm{H}^\bullet(A)$ is defined for any Weil cohomology theory $\mathrm{H}^\bullet$, so upon noting the compatibility of the cup product, Lemma 1.197 explains that there is an absolute Hodge correspondence giving rise to the map $\mathrm{Sym}^\bullet h^1(A) \to h(A)$ which specializes to the cup-product map for any of our cohomology theories.

To show that this map is an isomorphism on motives, it is enough to explain how to construct the inverse absolute Hodge correspondence. Well, Theorem 2.98 does promise that the cup-product map does have a (unique) inverse on each cohomology theory, which will be compatible among our cohomology theories by the ambient uniqueness, so Lemma 1.197 promises that we have an inverse on the level of absolute Hodge classes. ∎

Remark 2.105. As in Remark 2.99, we note that one can actually exhibit this isomorphism on the level of the Chow motives $\mathrm{ChMot}_{\mathbb{Q}}(K)$. Once again, this requires more motivic input than we would like to introduce, so we merely refer to [EGM, Theorem 13.47]. To give a taste for why one might expect this to be difficult, we note that the statement requires one to define $h^1(A)$, so one has to explain why Künneth projectors exist for abelian varieties.

Remark 2.106. Corollary 2.104 explains that the tensor subcategory $\langle h(A)\rangle^\otimes \subseteq \mathrm{Mot}_{\mathbb{Q}}(K)$ may in fact merely be generated by $h^1(A)$ and $\mathbb{Q}(1)$.

2.3.2 The Construction

A priori, an abelian variety A gives rise to many ℓ-adic Galois representations via each of its cohomology groups $\mathrm{H}^\bullet_{\text{ét}}(A_{K^{\mathrm{sep}}}, \mathbb{Q}_\ell)$. However, by Theorem 2.98, we see that one can understand all cohomology groups of A by merely understanding $\mathrm{H}^1_{\text{ét}}(A_{K^{\mathrm{sep}}}, \mathbb{Z}_\ell)$. Analogous to the complex analytic case, we will be able to work with the dual "homology group" more concretely.

Let's spend some time giving a more elementary description of $\mathrm{H}^1_{\text{ét}}(A_{K^{\mathrm{sep}}}, \mathbb{Z}_\ell)$. We refer to [EGM, Corollary 10.38] and the surrounding discussion for an alternate approach. After passing to finite level and twisting, we reduce to the following lemma.

Lemma 2.107. Fix an abelian variety A over a field K, and let n be an integer coprime to $\operatorname{char} k$. Then there is an isomorphism

$$\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mu_n) \to A^\vee(K^{\text{sep}})[n]$$

of Galois representations which is functorial in n.

Proof. We will need to use some facts about étale cohomology, which we will not prove and instead refer to [Del77]. We proceed in steps, for clarity.

1. The main point is to use the "Kummer" exact sequence

   $$0 \to \mu_n \to \mathbb{G}_m \xrightarrow{n} \mathbb{G}_m \to 1$$

   of étale sheaves on $A_{K^{\text{sep}}}$ [Del77, Kummer Theory II.2.5]. In brief, the left-exactness here can be checked on sections (essentially by definition of μ_n), and the surjectivity on the right is checked on stalks: for any local section $s \in \mathrm{H}^0_{\text{ét}}(U; \mathbb{G}_m)$ where $U \subseteq A_{k^{\text{sep}}}$ is some affine étale open neighborhood, one can write $U = \operatorname{Spec} A$ and pass to the étale cover $\operatorname{Spec} A[x]/(x^n - s) \twoheadrightarrow U$ to find a pre-image for s.

   While we're here, we note that the Kummer exact sequence admits some functoriality: for any two integers n and d both coprime to $\operatorname{char} k$, there is a morphism

   $$\begin{array}{ccccccccc} 1 & \longrightarrow & \mu_n & \longrightarrow & \mathbb{G}_m & \xrightarrow{n} & \mathbb{G}_m & \longrightarrow & 1 \\ & & \downarrow & & \| & & \downarrow{\scriptstyle d} & & \\ 1 & \longrightarrow & \mu_{nd} & \longrightarrow & \mathbb{G}_m & \xrightarrow{nd} & \mathbb{G}_m & \longrightarrow & 1 \end{array}$$

   of short exact sequences of étale sheaves.

2. Taking étale cohomology of the Kummer sequence gives the exact sequence

   $$K^{\text{sep}\times} \xrightarrow{n} K^{\text{sep}\times} \to \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mu_n) \to \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mathbb{G}_m) \xrightarrow{n} \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mathbb{G}_m)$$

   (The left two cohomology groups are computed using the properness of A.) The left map is surjective, so we see that there is an isomorphism

   $$\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mu_n) \to \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mathbb{G}_m)[n].$$

   Note that this isomorphism is functorial in n due to the functoriality of the Kummer exact sequences given above.

3. It remains to show that $\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mathbb{G}_m)[n] \cong A^\vee[n]$. Well, we actually note that

   $$\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mathbb{G}_m) \cong \operatorname{Pic} A_{K^{\text{sep}}}$$

   by [Del77, Proposition II.2.3], so the result follows upon recalling $A^\vee = \operatorname{Pic}^0 A$. Let's say a little about the construction of the above isomorphism: as in Zariski cohomology, one finds that $\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mathbb{G}_m)$ classifies $\mathbb{G}_m$-torsors on $A_{K^{\text{sep}}}$, which can be checked to be the same thing as a line bundle in the étale topology. Lastly, a descent argument shows that a line bundle in the étale topology is the same as a line bundle in the Zariski topology. ∎

Proposition 2.108. Fix an abelian variety A over a field K, and let ℓ be prime such that $\operatorname{char} K \nmid \ell$. Then there is a Galois-equivariant isomorphism

$$\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}; \mathbb{Z}_\ell) \cong \left(\varprojlim A^\vee(K^{\text{sep}})[\ell^\bullet]\right)(1).$$

Proof. This follows by taking a limit of the isomorphisms of Lemma 2.107 constructed at finite level. ∎

We are now allowed to define the Tate module.

Definition 2.109 (Tate module)**.** Fix an abelian variety A over a field K, and suppose ℓ is a prime such that $\operatorname{char} K \nmid \ell$. Then we define the ℓ-adic Tate module as

$$T_\ell A \coloneqq \varprojlim A\left[\ell^\bullet\right](K^{\mathrm{sep}}),$$

and we define the rational ℓ-adic Tate module as $V_\ell A \coloneqq T_\ell A \otimes_{\mathbb{Z}} \mathbb{Q}$.

Remark 2.110. Intuitively, $T_\ell A$ should be thought of as an ℓ-adic stand-in for $\mathrm{H}_1(A)$. This will be explained by Proposition 2.112 below.

The discussion above suggests that $T_\ell A$ should be a free $\mathbb{Z}_\ell$-module of rank 2. Let's check this directly. By taking limits, it is enough to show the following.

Lemma 2.111. Fix an abelian variety A over a field K, and suppose ℓ is a prime such that $\operatorname{char} K \nmid \ell$. For each $\nu \geq 0$, there is a group isomorphism

$$A\left[\ell^\nu\right](K^{\mathrm{sep}}) \cong \mathbb{Z}/\ell^{2\nu \dim A}\mathbb{Z}.$$

Proof. The two groups have the same size by Example 2.17, so the result follows for $\nu \in \{0,1\}$ automatically. For $\nu \geq 2$, we induct using the short exact sequence

$$0 \to A[\ell](K^{\mathrm{sep}}) \to A\left[\ell^{\nu+1}\right](K^{\mathrm{sep}}) \xrightarrow{\ell} A\left[\ell^\nu\right](K^{\mathrm{sep}}) \to 0$$

and some cardinality arguments. For example, one can finish by applying the classification of finite abelian groups. ∎

Proposition 2.112. Fix an abelian variety A over a field K, and let ℓ be prime such that $\operatorname{char} K \nmid \ell$. Then there is a Galois-equivariant isomorphism

$$\mathrm{H}^1_{\text{ét}}(A_{K^{\mathrm{sep}}}; \mathbb{Z}_\ell) \cong (T_\ell A)^\vee.$$

Proof. Note that the functor $(-)^\vee$ passes through the inverse limit $T_\ell A$ because the internal modules of $T_\ell A$ are free by Lemma 2.111. Thus, by taking limits, it is enough to exhibit a natural isomorphism

$$\mathrm{H}^1_{\text{ét}}(A_{K^{\mathrm{sep}}}; \mathbb{Z}/n\mathbb{Z}) \overset{?}{\cong} A[n](K^{\mathrm{sep}})^\vee,$$

where n is some integer coprime to $\operatorname{char} K$. This isomorphism will follow from Lemma 2.107 after some massaging. By moving the Tate twist around (and noting that all modules in sight are free over $\mathbb{Z}/n\mathbb{Z}$), it is enough to exhibit an isomorphsm

$$A^\vee[n](K^{\mathrm{sep}}) \to A[n](K^{\mathrm{sep}})^\vee \otimes \mu_n$$

of Galois modules, or equivalently,

$$A^\vee[n](K^{\mathrm{sep}}) \to \operatorname{Hom}(A[n](K^{\mathrm{sep}}, \mu_n).$$

We now see that this isomorphism is induced by the Weil pairing (Corollary 2.39), which we recall is Galois-invariant and perfect. ∎

One benefit of a more concrete object is that it is easier to work with directly. For example, we can now find a perfect pairing on $\mathrm{H}^1_{\text{ét}}(A_{K^{\mathrm{sep}}}, \mathbb{Z}_\ell)$.

Lemma 2.113. Fix an abelian variety A over a field K, and suppose ℓ is a prime such that $\operatorname{char} K \nmid \ell$. Choose a polarization $\varphi\colon A \to A^\vee$. Then the Weil pairing induces a Galois-invariant perfect symplectic pairing

$$e_\varphi\colon \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Z}_\ell) \otimes_{\mathbb{Q}_\ell} \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Z}_\ell) \to \mathbb{Z}_\ell(-1).$$

Proof. By taking duals, it is enough to induce a Galois-invariant perfect symplectic pairing

$$e_\varphi\colon T_\ell A \otimes_{\mathbb{Q}_\ell} T_\ell A \to \mathbb{Z}_\ell(1).$$

This follows by taking a limit of the Weil pairing given in Corollary 2.39. Recall that $\mathbb{Z}_\ell(1)$ is the Galois representation $\varprojlim \mu_{\ell^\bullet}$. ∎

One can also see the Galois action more explicitly: being careful about the Galois action on cohomology and the Tate module, we see that the induced Galois representation

$$\rho_\ell\colon \operatorname{Gal}(K^{\text{sep}}/K) \to \operatorname{GL}(T_\ell A)$$

is simply given by the Galois action on the points in the limit $A\left[\ell^\bullet\right](K^{\text{sep}})$.

2.3.3 The ℓ-Adic Monodromy Group

Now that we have a representation, we may as well define a monodromy group.

Definition 2.114 (ℓ-adic monodromy group)**.** Fix an abelian A over a field K, and suppose ℓ is a prime such that $\operatorname{char} K \nmid \ell$. Then the *$\ell$-adic monodromy group* $G_\ell(A)$ is the smallest algebraic $\mathbb{Q}_\ell$-group containing the image of the Galois representation

$$\operatorname{Gal}(K^{\text{sep}}/K) \to \operatorname{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Q}_\ell)\right).$$

Remark 2.115. By taking duals, we see that one produces an isomorphic Galois representation by working with $T_\ell A$ instead. Note this dual is not very expensive: by using the Weil pairing of Lemma 2.113, we can remove the dual in exchange for a twist, writing

$$\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Z}_\ell) \cong T_\ell A(-1).$$

Remark 2.116. Unlike $\mathrm{MT}(A)$ and $\mathrm{Hg}(A)$, we do not expect $G_\ell(A)$ to be connected in general. However, being an algebraic $\mathbb{Q}_\ell$-group, it will only have finitely many connected components. Thus, we see that the pre-image of $G_\ell(A)^\circ$ in $\operatorname{Gal}(K^{\text{sep}}/K)$ is an open subgroup of finite index, so there is a unique minimal field extension K_A^{conn}/K such that $G_\ell(A_{K_A^{\text{conn}}}) = G_\ell(A)^\circ$. Thus, our group becomes connected, only at the cost of a field extension.

Remark 2.117. For a finite extension K' of K, Remark 2.116 explains that we may easily have $G_\ell(A) \neq G_\ell(A_{K'})$, but we now remark that $G_\ell(A)^\circ = G_\ell(A_{K'})^\circ$. Well,

$$\rho_\ell(\operatorname{Gal}(K^{\text{sep}}/K')) \subseteq \rho_\ell(\operatorname{Gal}(K^{\text{sep}}/K))$$

is some finite-index subgroup, so $G_\ell(A_{K'}) \subseteq G_\ell(A)$ is a finite-index subgroup (upon taking the closure). It follows these groups must have the same connected component; for example, one can pass to $\mathbb{C}$ and then see that a closed subgroup of a Lie group with smaller dimension necessarily has infinite index due to being able to continuously translate the smaller subgroup.

The interesting geometric objects arising from Hodge theory were the Hodge classes, which Remark 1.13 explains were exactly the vectors fixed by the group action. Analagously, we pick up the following definition.

Definition 2.118 (Tate class)**.** Fix an abelian A over a field K, and suppose ℓ is a prime such that $\operatorname{char} K \nmid \ell$. Then a *Tate class* is a vector of some tensor construction

$$\bigoplus_{i=1}^{k} \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Q}_\ell)^{\otimes n_i} \otimes \mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Q}_\ell)^{\vee \otimes m_i}(p_i),$$

where the $n_\bullet$s, $m_\bullet$s, and $p_\bullet$s are some nonnegative integers, fixed by the action of $\mathrm{Gal}(K^{\text{sep}}/K)$

Remark 2.119. We remark as in Proposition 1.33 that a subspace V as above is fixed by the Galois action if and only if it is fixed by the indcued action by $G_\ell(A)$. Indeed, the subset of $\mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Q}_\ell)\right)$ fixing V is some algebraic $\mathbb{Q}_\ell$-subgroup, so if it contains the image of $\mathrm{Gal}\left(K^{\text{sep}}/K\right)$, then it contains $G_\ell(A)$. We also take a moment to note that Proposition 1.35 explains that one can now cut out $G_\ell(A)$ by requiring it to hold all the Tate classes invariant, as discussed in Corollary 1.36.

Remark 2.120. The same argument as in Example 1.143 explains that $G_\ell(A)$ is the algebraic group corresponding to the tensor subcategory

$$\left\langle \mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell) \right\rangle^{\otimes} \subseteq \mathrm{Rep}_{\mathbb{Q}_\ell}(\mathrm{Gal}(\overline{K}/K)).$$

Notably, the application of Proposition 1.33 is replaced by the discussion in Remark 2.119.

Analogous to Conjecture 1.15, one has a Tate class, which we will only state for abelian varieties.

Conjecture 2.121 (Tate)**.** Fix an abelian variety A over a number field K, and fix a prime number ℓ. Then any Tate class can be written as a $\mathbb{Q}_\ell$-linear combination of classes arising from algebraic subvarieties of powers of A.

Remark 2.122. Of course, there are Tate classes and there is a Tate conjecture for more general varieties.

We conclude this section with a few bounds on the ℓ-adic monodromy group, analogous to the discussion for Mumford–Tate groups in section 1.2.3. Let's begin with endomorphisms.

Lemma 2.123. Fix an abelian variety A over a field K, and suppose ℓ is a prime such that $\operatorname{char} K \nmid \ell$. Set $D \coloneqq \mathrm{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q}$. Then

$$G_\ell(A) \subseteq \left\{ g \in \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Q}_\ell)\right) : g \circ d = d \circ g \text{ for all } d \in D \right\}.$$

Proof. We proceed as in Lemma 1.45. The right-hand group is an algebraic $\mathbb{Q}_\ell$-group, so it suffices to check that it contains the image of $\mathrm{Gal}(K^{\text{sep}}/K)$. Well, for any $g \in \mathrm{Gal}(K^{\text{sep}}/K)$, we see that

$$g \circ d = d \circ g$$

is an equality which holds on the level of endomorphisms of A because d is defined over K (which g fixes). ∎

Lemma 2.124. Fix an abelian variety A over a field K, and suppose ℓ is a prime such that $\operatorname{char} K \nmid \ell$. Choose a polarization $\varphi \colon A \to A^\vee$. Then there is a perfect symplectic pairing e_φ such that

$$G_\ell(A) \subseteq \left\{ g \in \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{K^{\text{sep}}}, \mathbb{Q}_\ell)\right) : e_\varphi(gv \otimes gw) = \lambda(g) e_\varphi(v \otimes w) \text{ for fixed } \lambda(g) \in \mathbb{Q}_\ell \right\}.$$

Proof. We proceed as in Lemma 1.47. The right-hand group is an algebraic $\mathbb{Q}_\ell$-group, so it suffices to check that it contains the image of $\mathrm{Gal}(K^{\mathrm{sep}}/K)$. Well, for any $g \in \mathrm{Gal}(K^{\mathrm{sep}}/K)$, we see that

$$e_\varphi(gv \otimes gw) = ge_\varphi(v \otimes w)$$

by the Galois-invariance of Lemma 2.113. Now, we note that $\mathrm{Gal}(K^{\mathrm{sep}}/K)$ acts on $\mathbb{Q}_\ell(-1)$ through the cyclotomic character, so the right-hand side equals a scalar $\lambda(g)$ times $e_\varphi(v \otimes w)$, so we are done. ∎

Remark 2.125. There are of course alternate proofs of Lemmas 2.123 and 2.124 by finding Tate classes and then appealing to Remark 2.119. One uses the same classes constructed in the alternate proofs of Lemmas 1.45 and 1.47.

Lastly, we would like to recover the bound of Corollary 2.42 on endomorphisms, sharpening Lemma 2.123. However, the proof is not so easy: the proof of Corollary 2.42 had to translate endomorphisms of the Hodge structure back to endomorphisms of the abelian variety via Theorem 2.40. Recovering the equivalence of Theorem 2.40 is rather difficult: this result is due to Faltings [Fal86, Theorem 3], in his proof of Mordell's conjecture.

Theorem 2.126 (Faltings)**.** Fix an abelian variety A over a number field K, and suppose ℓ is a prime. Then the induced map

$$\mathrm{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q}_\ell \to \mathrm{End}_{\mathrm{Gal}(\overline{K}/K)}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$$

is an isomorphism.

We will definitely not attempt to summarize a proof here, but we will remark that it is not even totally obvious that this map is injective! Speaking from experience, this makes for a reasonable topic for a final term paper in a first course in algebraic geometry.

Remark 2.127. Via the isomorphism

$$\mathrm{End}_{\mathbb{Q}_\ell}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right) \cong \mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell) \otimes \mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)^\vee,$$

we see that Theorem 2.126 can be viewed as asserting that all the Tate classes in the above space arise from endomorphisms of A. This verifies Conjecture 2.121.

Remark 2.128. We have snuck in the hypothesis that K is a number field into the statement of Theorem 2.126. It is also true for finite fields, where it is due to Tate [Tat66]. However, it is not expected to be true in general!

We are now able to provide a satisfying analogue to Lemma 1.54.

Corollary 2.129. Fix an abelian variety A over a number field K, and suppose ℓ is a prime. Then the natural map

$$\mathrm{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q}_\ell \to \mathrm{End}_{G_\ell(A)}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$$

is an isomorphism.

Proof. Remark 2.127 explains that the endomorphisms of A are exactly the Tate classes, so the result follows from the discussion in Remark 2.119. ∎

Remark 2.130. The above corollary allows us to prove the following analogue of Proposition 2.53 (by the same proof!): A has CM defined over a number field K if and only if $G_\ell(A)$ is a torus.

While we're here, we remark on another property of $G_\ell(A)$ due to Faltings.

Theorem 2.131 (Faltings). Fix an abelian variety A over a number field K, and suppose ℓ is a prime. Then $G_\ell(A)$ is reductive.

Proof. By [Mil17, Corollary 19.18], it is enough to find a faithful semisimple representation of $G_\ell(A)$. As in Lemma 1.44, we see that the inclusion

$$G_\ell(A) \subseteq \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$$

is semisimple by [Fal86, Theorem 3], so we are done. ∎

Remark 2.132. Over finite fields, Tate [Tat66] has proven that the Galois representation $\mathrm{H}^1_{\text{ét}}(A_{\overline{\mathbb{F}_q}}, \mathbb{Q}_\ell)$ is semisimple. Because the Galois group is (topologically) generated by the Frobenius, this amounts to checking that the endomorphism Frob_q has semisimple action.

To finish up our discussion of computational tools for $G_\ell(A)$, we repeat the results Lemmas 1.56 and 1.59 for our new context. Their proofs are exactly the same, replacing $\mathbb{U}$ (or $\mathbb{S}$) with $\mathrm{Gal}(\overline{K}/K)$ and then making the same minimality arguments for our monodromy groups.

Lemma 2.133. Fix abelian varieties $A_1, \dots, A_k$ over a field K.

(a) The subgroup
$$G_\ell(A_1 \times \cdots \times A_k) \subseteq \mathrm{GL}(\mathrm{H}^1_{\text{ét}}((A_1 \times \cdots \times A_k)_{K^{\text{sep}}}, \mathbb{Q}_\ell))$$
is contained in $G_\ell(A_1) \times \cdots \times G_\ell(A_k)$.

(b) For each i, the projection map $\mathrm{pr}_i \colon G_\ell(A_1 \times \cdots \times A_k) \to G_\ell(A_i)$ is surjective.

Lemma 2.134. Fix abelian varieties $A_1, \dots, A_k$ over a field K, and let $m_1, \dots, m_k \geq 1$ be positive integers. Then the diagonal embeddings $\Delta_i \colon \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{i,K^{\text{sep}}}, \mathbb{Q}_\ell)\right) \to \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A^{m_i}_{i,K^{\text{sep}}}, \mathbb{Q}_\ell)\right)$ induce an isomorphism
$$G_\ell(A_1 \times \cdots \times A_k) \to G_\ell\left(A_1^{m_1} \times \cdots \times A_k^{m_k}\right).$$

2.4 Computational Tools

In this section, we give some tools to compute the ℓ-adic representation and the ℓ-adic monodromy group in particular.

2.4.1 The Fundamental Theorem of Complex Multiplication

Before continuing, we give essentially the only class of examples in which one is able to imagine computing the ℓ-adic representation. For this subsection, we will let A be an abelian variety of dimension g defined over a number field K with complex multiplication by an order $\mathcal{O}$ of a CM number field E. We let Φ denote the CM type, which we now think of as a subset of Σ_E, and we let (E^*, Φ^*) be a reflex CM type; we may as well descend (E^*, Φ^*) to be as small as possible. Our exposition closely follows [Con04, Section 3]. It is slightly beyond the scope of our current discussion to give a precise statement of the Fundamental theorem of complex multiplication; instead, we will work with the following consequence.

Ultimately, we are interested in computing the Galois action of $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ on the Tate module of A. In order to avoid fixing a prime ℓ, we pick up the following notation.

Notation 2.135. Fix an abelian variety A defined over a field L of characteristic 0. Then we define the adelic Tate modules $\widehat{T}_f(A) \coloneqq \prod_\ell T_\ell(A)$ and $\widehat{V}_f(A) \coloneqq T_f(A) \otimes_{\mathbb{Z}} \mathbb{Q}$.

Remark 2.136. Note that $\widehat{T}_f A$ is a free $\widehat{\mathbb{Z}}$-module of rank $2g$, and $\widehat{V}_f A$ is a free $\mathbb{A}_{\mathbb{Q},f}$-module of rank $2g$.

Thus, we are interested in the Galois representation

$$\rho_A \colon \mathrm{Gal}(\overline{\mathbb{Q}}/K) \to \mathrm{GL}(\widehat{V}_f A).$$

Suitably interpreted, this Galois representation will turn out to be the reflex norm, up to a root of unity.

Because A has complex multiplication by E defined over K, the image of ρ_A commutes with the action of K on $\widehat{V}_f A$, so ρ_A actually factors through $\mathrm{GL}_{E \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}}(V_f A)$. By looking factor-by-factor (on each ℓ), we see that this target is contained in $E \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$ because K is its own commutator. Thus, ρ_A factors as a Galois representation

$$\rho_A \colon \mathrm{Gal}(\overline{\mathbb{Q}}/K) \to (E \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}})^\times,$$

where the embedding $E \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}} \hookrightarrow \mathrm{GL}(T_f A)$ is given by the action of E on A. We take a moment to note that this target is $(E \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}})^\times = (E \otimes_{\mathbb{Q}} \mathbb{A}_{\mathbb{Q},f})^\times = \mathbb{A}_{E,f}^\times$. Anyway, because the target is now abelian, we see that ρ_A factors through

$$\rho_A \colon \mathrm{Gal}(K^{\mathrm{ab}}/K) \to \mathbb{A}_{E,f}^\times.$$

Artin reciprocity provides a canonical map $\mathrm{Art}_K \colon \mathbb{A}_K^\times \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$, so we may as well work with the composite $\overline{\rho}_A$

$$\mathbb{A}_K^\times \to \mathrm{Gal}(K^{\mathrm{ab}}/K) \xrightarrow{\rho_A} \mathbb{A}_{E,f}^\times.$$

We take a moment to remark that we may as well work with a quotint of $\mathbb{A}_K^\times$.

Remark 2.137. By [NSW08, Corollary 8.2.2], we note that Art_K is surjective with kernel containing $K^\times$ and $\mathbb{A}_{K,\infty}^\times \subseteq \mathbb{A}_K^\times$. To see $\mathbb{A}_{K,\infty}^\times$ is in the kernel, we need to know that K has no real places, which holds because K contains the CM field E^* because E^* is the field of definition for the endomorphisms of A.

For example, this implies that $\overline{\rho}_A$ factors through $\mathbb{A}_K^\times \twoheadrightarrow \mathbb{A}_{K,f}^\times$.

It is this induced map $\overline{\rho}_A$ which will essentially turn out to be the reflex norm. Here is our statement of the Fundamental theorem of complex multiplication, which we will not prove.

Theorem 2.138 (Fundamental)**.** Fix an abelian variety A with complex multiplication by (E, Φ) defined over a number field K. Then there is a continuous homomorphism $\lambda \colon \mathbb{A}_{K,f}^\times \to E^\times$ such that any $s_f \in \mathbb{A}_{K,f}^\times$ has

$$\rho_A(\mathrm{Art}_K\, s_f) = \lambda(s_f)\, \mathrm{N}_\Phi(\mathrm{N}_{K/E^*}(s_f))^{-1}.$$

Here, λ is continuous where $E^\times$ has been given the discrete topology.

Remark 2.139. Technically, the definition of N_{K/E^*} depends on a choice of reflex E^* inside K, which depends on a choice of embedding $K \hookrightarrow \overline{\mathbb{Q}}$. However, it turns out that the composite $\mathrm{N}_\Phi \circ \mathrm{N}_{K/E^*}$ does not depend on the choice of embedding $L \hookrightarrow \overline{\mathbb{Q}}$. We will not need this, so we will not show it; we remark that this is essentially shown in Lemma 2.84.

Remark 2.140. Theorem 2.138 is frequently cited as merely a corollary of the Fundamental theorem, and the Fundamental theorem is indeed a more precise statement about the Galois action on A. However, the precise statement of the usual Fundamental theorem is rather technical, and we will not need it, so we will be happy merely using Theorem 2.138 in this article.

Let's give a few properties of this mysterious character λ for future use.

Proposition 2.141. Fix an abelian variety A with complex multiplication by (E, Φ) defined over a number field K. Define the continuous character $\lambda\colon \mathbb{A}_{K,f}^\times \to E^\times$ as in Theorem 2.138.

(a) For $s_f \in \mathbb{A}_{K,f}^\times$, the fractional ideal generated by $\mathrm{N}_\Phi(\mathrm{N}_{K/E^*}(s_f))$ is $\lambda(s_f)\mathcal{O}_E$.

(b) Choose a prime $\mathfrak{P}$ of K. Then A has good reduction at $\mathfrak{P}$ if and only if λ is trivial on $\mathcal{O}_\mathfrak{P}^\times \subseteq \mathbb{A}_{K,f}^\times$.

(c) Choose a prime $\mathfrak{P}$ of K with uniformizer $\varpi_\mathfrak{P} \in \mathcal{O}_\mathfrak{P}^\times$. Suppose A has good reduction at $\mathfrak{P}$. Then $\lambda(\varpi_\mathfrak{P}) \in \mathcal{O}_K$, and it agrees with the $q_\mathfrak{P}$-Frobenius endomorphism on the reduction $A_{\kappa(\mathfrak{P})}$ (where $q_\mathfrak{P} \coloneqq \#\kappa(\mathfrak{P})$).

Proof. We show these parts one at a time. For (b) and (c), it will help to fix a rational prime ℓ not lying under $\mathfrak{P}$.

(a) For each finite prime $\mathfrak{P}$ of K, we must show $\mathrm{N}_\Phi(\mathrm{N}_{K/E^*}(s_f))$ and $\lambda(s_f)$ have the same $\mathfrak{P}$-valuations. Equivalently, we would like to check that

$$u(s_f) \coloneqq \lambda(s_f)\,\mathrm{N}_\Phi(\mathrm{N}_{K/E^*}(s_f))^{-1} \overset{?}{\in} \prod_\mathfrak{P} \mathcal{O}_{E,\mathfrak{P}}^\times.$$

Well, by Theorem 2.138, we see that $u(s_f)$ acts on the Tate module $\widehat{V}_f A$ as $\rho_A(\mathrm{Art}_L\, s_f)$, which is notably an automorphism fixing the integral sublattice $\widehat{T}_f A \subseteq \widehat{V}_f A$. We conclude that $u(s_f)$ is a unit at all finite places.

(b) We use the Néron–Ogg–Shafaverich criterion [ST68], which tells us that A has good reduction at $\mathfrak{P}$ if and only if $\rho_A\colon \mathrm{Gal}(\overline{\mathbb{Q}}/K) \to V_\ell A$ is trivial on the inertia subgroup $I_\mathfrak{P}$. The Artin map $\mathrm{Art}_K\colon \mathbb{A}_{K,f}^\times \to \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ is surjective, and the image of $\mathcal{O}_\mathfrak{P}^\times \subseteq \mathbb{A}_{K,f}^\times$ is precisely $I_\mathfrak{P}$, so A has good reduction at $\mathfrak{P}$ if and only if the composite

$$\mathcal{O}_\mathfrak{P}^\times \subseteq \mathbb{A}_{K,f}^\times \twoheadrightarrow \mathrm{Gal}(K^{\mathrm{ab}}/K) \to \mathbb{A}_{E,f}^\times \twoheadrightarrow \mathbb{A}_{E,\ell}^\times$$

is trivial. Well, by Theorem 2.138, we see that this composite is λ multiplied by the reflex norm. The image of the reflex norm on $\mathcal{O}_\mathfrak{P}^\times \subseteq \mathbb{A}_{K,f}^\times$ will land away from $\mathbb{A}_{K,\ell}^\times \subseteq \mathbb{A}_{E,f}^\times$, so it does not affect whether this composite is trivial. Thus, we conclude that the composite is trivial if and only if $\lambda|_{\mathcal{O}_\mathfrak{P}^\times}$ is trivial.

(c) Quickly, observe that $\lambda(\varpi_\mathfrak{P}) \in \mathcal{O}_E^\times$ follows from agreeing with the Frobenius on the reduction. Indeed, agreeing with the Frobenius on the reduction implies that $\lambda(\varpi_\mathfrak{P})$ is the root of the characteristic polynomial of the Frobenius, which is monic with coefficients in $\mathbb{Z}$.

It remains to check agreement with the Frobenius on the reduction. The computation of the composite used in the proof of (b) explains that the Galois action of $\mathrm{Frob}_\mathfrak{P} = \mathrm{Art}_K(\varpi_\mathfrak{P})$ on $V_\ell A$ is given by $\lambda(\varpi_\mathfrak{P})_\ell \in \mathbb{A}_{E,\ell}^\times$. Thus, the action of $\lambda(\varpi_\mathfrak{P})$ on the Tate module $T_\ell A_{\kappa(\mathfrak{P})}$ of the reduction also agrees with the action of the Frobenius, which lifts to an equality of actions on the actual reduction $A_{\kappa(\mathfrak{P})}$ because passing to the Tate module is faithful (see Remark 2.128). ∎

Remark 2.142. For (c), one may want to say that $\lambda(\varpi_\mathfrak{P})$ is a characteristic-0 lifting of the Frobenius endomorphism on the reduction. However, if we do not have $\mathcal{O}_E \subseteq \mathrm{End}_K(A)$, then we cannot actually guarantee this lifting.

Here is an example application of Theorem 2.138.

Proposition 2.143. Fix an abelian variety A over a number field K with complex multiplication by a CM algebra $E = E_1 \times \cdots \times E_k$. Let $\Phi\colon \Sigma_E \to \mathbb{Z}_{\geq 0}$ be the induced signature, which we decompose as $\Phi = \Phi_1 \sqcup \cdots \sqcup \Phi_k$ where $(E_\bullet, \Phi_\bullet)$ is a CM signature for all $E_\bullet$. Extend K to be a CM field equipped with CM signatures $\Phi_1^*, \dots, \Phi_k^*$ such that (E_i, Φ_i) and (K, Φ_i^*) are reflex for all i. Then $G_\ell(A)^\circ \subseteq \mathrm{T}_E$ is the Zariski closure of the image of

$$(\mathrm{N}_{\Phi_1^*}, \dots, \mathrm{N}_{\Phi_k^*})\colon \mathrm{T}_K \to \mathrm{T}_E.$$

Proof. We follow [Yu15, Lemma 4.1]. We quickly explain why extending K does not actually affect the conclusion. On one hand, $G_\ell(A)^\circ$ is independent of extending K by Remark 2.117; on the other hand, passing to an extension cannot change the closure of the image of the reflex norms by Lemma 2.84 and the fact that norms of field extensions have Zariski dense image.

Technically, the rest of this subsection has dealt with simple abelian varieties, so we must do some work to handle the given CM algebra E. We may decompose $A = A_1 \times \cdots \times A_k$, where A_i is simple and has complex multiplication by (E_i, Φ_i). Define λ_i for A_i as in Theorem 2.138. Then we see ρ_A outputs to $\mathrm{T}_E = \mathrm{T}_{E_1} \times \cdots \times \mathrm{T}_{E_i}$, where the ith component is just given by ρ_{A_i}.

Recall from Remark 2.137 that the Artin map $\mathrm{Art}_K\colon \mathbb{A}_{K,f}^\times / K^\times \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$ is surjectve, so it is enough to compute the image of $\rho_{A,\ell} \circ \mathrm{Art}_K$, where the ℓ signifies that we are working with the ℓ-component. In particular, for $s_f \in \mathbb{A}_{K,f}^\times$, Theorem 2.138 implies that

$$(\rho_{A_i,\ell} \circ \mathrm{Art}_K)(s_f) = \lambda_i(s_f)(\mathrm{N}_{\Phi_i^*})^{-1}(s_\ell),$$

so

$$(\rho_{A,\ell} \circ \mathrm{Art}_K)(s_f) = \left(\lambda_1(s_f)(\mathrm{N}_{\Phi_1^*})^{-1}(s_\ell), \dots, \lambda_k(s_f)(\mathrm{N}_{\Phi_k^*})^{-1}(s_\ell)\right).$$

We may as well compress the right-hand side into $\lambda(s_f)\,\mathrm{N}_{\Phi^*}(s_f)_\ell^{-1}$, where λ and N_{Φ^*} output to k-tuples in T_E. The above equality explains where the image of the reflex norm is going to come from. We now have two inclusions; let $\mathrm{T} \subseteq \mathrm{T}_E$ denote the Zariski closure of N_{Φ^*}.

- We claim that $G_\ell(A)^\circ \subseteq \mathrm{T}$. Note $\ker\lambda \subseteq \mathbb{A}_{K,f}^\times$ is open by continuity of the λ_is, so strong approximation implies that $K^\times \backslash \mathbb{A}_{K,f}^\times / \ker\lambda$ is finite.[3] Thus, $\mathrm{im}\,\lambda/\mathrm{T}$ is finite, and we conclude that $G_\ell(A)$ is contained in T multiplied by some finite group $\mathrm{im}\,\lambda/\mathrm{T}$. Finite groups are disconnected, so $G_\ell(A)^\circ \subseteq \mathrm{T}$ follows.

- We claim that $\mathrm{T} \subseteq G_\ell(A)$. Again, $\ker\lambda \subseteq \mathbb{A}_{K,f}^\times$ is open, so

$$\rho_{A,\ell}(\mathrm{Art}_K(\ker\lambda)) = \mathrm{N}_{\Phi^*}(\ker\lambda_\ell).$$

  Now, $\ker\lambda_\ell$ is Zariski dense in $\mathbb{Q}_\ell^\times$, so the right-hand side is Zariski dense in T. The inclusion follows. ∎

2.4.2 The Mumford–Tate Conjecture

Over the next few subsections, we will explain some tools used to compute $G_\ell(A)$. In this subsection, we will discuss $G_\ell(A)^\circ$. Suppose that A is defined a number field K.

A motivic perspective would have us hope that all the monodromy groups attached to A are essentially the same. However, as explained in Remark 2.116, we only expect $G_\ell(A)$ to be connected after an extension K. Thus, for example, one can only hope that $\mathrm{MT}(A)$ knows about $G_\ell(A)^\circ$; this now makes sense because $G_\ell(A)^\circ$ is independent of the base field K by Remark 2.117. We may now state the following conjecture.

[3] In this case, this reduces to finiteness of the class number h_K: being open means $\ker\lambda$ is commensurable with $\widehat{\mathcal{O}}_K^\times$, and $K^\times \backslash \mathbb{A}_{K,f}^\times / \widehat{\mathcal{O}}_K^\times$ is isomorphic to the class group as a pointed set.

Conjecture 2.144 (Mumford–Tate)**.** Fix an abelian variety A over a number field K. For all primes ℓ, we have

$$\mathrm{MT}(A)_{\mathbb{Q}_\ell} = G_\ell(A)^\circ$$

as subgroups of $\mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A, \mathbb{Q}_\ell)\right)$. Here, $\mathrm{MT}(A)$ is embedded into this group by the Betti-to-étale comparison isomorphism.

Our work in chapter 1 provides many tools for computing $\mathrm{MT}(A)$, so Conjecture 2.144 would allow us to translate this knowledge into a computation of $G_\ell(A)^\circ$.

Even though Conjecture 2.144 is not fully proven, there is a lot known. Let's review a little.

Example 2.145. If A is an absolutely simple abelian variety with complex multiplication by a CM algebra E, then both $G_\ell(A)^\circ$ and $\mathrm{MT}(A)$ equal the Zariski closure of the image of a suitably defined reflex norm to T_F. For the Mumford–Tate group, this is Proposition 2.86; for the ℓ-adic monodromy group, this is Proposition 2.143.

Remark 2.146. The Mumford–Tate conjecture for abelian varieties with complex multiplication is quite old: it is originally due to Pohlmann [Poh68, Theorem 4], but Ribet in [Rib04] has pointed out that the result is a corollary of results due to Shimura and Tanimaya [ST61], and [Yu15] has recently explicated this argument.

Let's move on to some more general results. For example, both groups are reductive by Lemma 1.44 and Theorem 2.131. Additionally, Theorem 2.126 provides a suitable analogue of Theorem 2.40, telling us that both groups $\mathrm{MT}(A)$ and $G_\ell(A)$ cut out endomorphisms in $\mathrm{End}(A)$.

As a philosophical check, one can show that $G_\ell(A)^\circ$ "contains" the Hodge structure morphism; the following result is due to Sen [Sen73, Theorem 1].

Theorem 2.147 (Sen)**.** Fix an abelian variety A over a number field K. Define the operator Φ as acting by mutltiplication-by-i on each eigenspace

$$\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)_{\mathbb{C}_\ell}(i),$$

where the (i)th eigenspace acts by ith power of the cyclotomic character. Then $\mathrm{Lie}\, G_\ell(A)^\circ$ is the smallest Lie algebra containing Φ.

Continuing, one inclusion of Conjecture 2.144 is known, due to Deligne [Del18, Corollary 6.2].

Theorem 2.148 (Deligne)**.** Fix an abelian variety A over a number field K. For all primes ℓ, we have

$$G_\ell(A)^\circ \subseteq \mathrm{MT}(A)_{\mathbb{Q}_\ell}.$$

In particular, it becomes enough to compare numerical invariants of the two groups (such as rank) to argue for an equality. For example, the following independence result is due to Larsen and Pink [LP95, Theorem 4.3].

Theorem 2.149 (Larsen–Pink)**.** Fix an abelian variety A over a number field K. If $\mathrm{MT}(A)_{\mathbb{Q}_\ell} = G_\ell(A)^\circ$ holds for any prime ℓ, then it holds for all primes ℓ.

One even knows that the centers of the groups coincide, due to Vaisu [Vas07, Theorem 1.3.1].

Theorem 2.150 (Vaisu)**.** Fix an abelian variety A over a number field K. For each prime ℓ, we have

$$Z(\mathrm{MT}(A))^\circ_{\mathbb{Q}_\ell} = Z(G_\ell(A))^\circ.$$

Vaisu [Vas07] has in fact shown quite a bit about the Mumford–Tate conjecture; see in particular [Vas07, Theorem 1.3.4].

Much is known about products, especially products with restricted endomorphism types. By combining [Ich91; Lom16], one is able to compute both $\mathrm{MT}(A)$ and $G_\ell(A)^\circ$ for many abelian varieties of Types I–III and control contributions coming from Type IV; this permits a proof of the Mumford–Tate conjecture for products of abelian varieties of dimension at most 3. More generally, the following result is due to Commelin [Com18, Theorem 1.2].

Theorem 2.151 (Commelin)**.** Fix abelian varieties A and B over a number field K. If the Mumford–Tate conjecture holds for both A and B, then it holds for $A \times B$.

To give a taste for how some of these results are proven, we show the following, which follows from [Vas07, Theorem 1.3.4].

Proposition 2.152. Fix a geometrically simple abelian variety A over a number field K. Suppose that $E = Z(\mathrm{End}_{\overline{K}}(A))$ equals a CM field such that $\dim A = \dim E$. Letting Φ be the corresponding signature, we further suppose that $\Phi(\sigma) = 1$ for exactly two $\sigma \in \Sigma_E$. Then we show the Mumford–Tate conjecture holds for A, and

$$\mathrm{MT}(A)^{\mathrm{der}} = \mathrm{L}(A)^{\mathrm{der}}.$$

Proof. For special ℓ, we will actually compute $\mathrm{MT}(A)^{\mathrm{der}}$ and $G_\ell(A)^{\circ,\mathrm{der}}$ "simultaneously" to show that they are equal to the suitable version of $\mathrm{GSp}_E(\varphi)^{\mathrm{der}}$ or $\mathrm{GSp}_E(e_\varphi)^{\mathrm{der}}$. By adding in what we know about the centers from Theorem 2.150 (and the independence of ℓ given in Theorem 2.149), the Mumford–Tate conjecture follows for A. The outline is to base-change to $\mathbb{C}$, where the Lie algebra of $\mathrm{L}(A)^{\mathrm{der}}$ becomes a product of $\mathfrak{sl}_2(\mathbb{C})$s, from which we can appeal to Lemma 1.62.

Before beginning the computation, we set up some notation. In practice, it will be convenient to only write down the computation for $\mathrm{MT}(A)^{\mathrm{der}}$, but we will indicate along the way the changes that need to be made for $G_\ell(A)^{\circ,\mathrm{der}}$. Now, for brevity, set $V \coloneqq \mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q})$ so that $\mathrm{Hg}(A) = \mathrm{Hg}(V)$ and $\mathrm{L}(A) = \mathrm{L}(V)$; we remark that V is a free module over E of rank 2.

Continuing with the set-up, we recall some part of the computation from Lemma 1.68. Fix a polarization φ on V. Then let $\rho_1, \dots, \rho_{e_0}$ be the embeddings of $E_i^\dagger$ into a Galois closure $M^\dagger$, which is the totally real subfield of the Galois closure M of E. Then we admit a decomposition

$$V_{M^\dagger} = V_1 \oplus \cdots \oplus V_{e_0}$$

so that

$$\mathrm{L}(V)_{M^\dagger} = \mathrm{Sp}_{E \otimes_{\rho_1} M^\dagger}(\varphi|_{V_1}) \times \cdots \times \mathrm{Sp}_{E \otimes_{\rho_{e_0}} M^\dagger}(\varphi|_{V_{e_0}}).$$

We now also recall from Lemma 1.68 that each $\mathrm{Sp}_{E \otimes_{\rho_i} M^\dagger}(\varphi|_{V_i})_M$ is isomorphic to $\mathrm{GL}_2(M)$; in particular, this group is connected. In particular, to achieve this decomposition, we diagonalize the induced action of M on V_i and then projects onto one of the eigenspaces.

Now, we would like to show that the inclusion

$$\mathrm{Hg}(V)^{\mathrm{der}}_M \subseteq \mathrm{Sp}_{E \otimes_{\rho_1} M^\dagger}(\varphi|_{V_1})_M \times \cdots \times \mathrm{Sp}_{E \otimes_{\rho_{e_0}} M^\dagger}(\varphi|_{V_{e_0}})_M$$

is an isomorphism, where the last group is embedded in $\mathrm{GL}(V)_M$. All groups involved are connected, so we may check this inclusion on the level of the Lie algebra, so we would like for the inclusion

$$\mathrm{Lie}\,\mathrm{Hg}(V)^{\mathrm{der}}_M \subseteq \mathrm{Sp}_{E \otimes_{\rho_1} M^\dagger}(\varphi|_{V_1})_M \times \cdots \times \mathrm{Sp}_{E \otimes_{\rho_{e_0}} M^\dagger}(\varphi|_{V_{e_0}})_M$$

is surjective. For this, we use Lemma 1.62. Here are our checks; for brevity, set $\mathfrak{hg}(V) \coloneqq \mathrm{Lie}\,\mathrm{Hg}(V)_M$, and let $\mathfrak{sl}_2(M)_i$ be the factor $\mathrm{Lie}\,\mathrm{Sp}_{E \otimes_{\rho_i} M^\dagger}(\varphi|_{V_i})^{\mathrm{der}}_M$, which we note is isomorphic to $\mathfrak{sl}_2(M)$.

(i) We claim that $\mathfrak{hg}(V)^{\mathrm{der}}$ surjects onto $\mathfrak{sl}_2(M)_i$, which we note is nonzero and simple. Because the $\mathfrak{hg}(V)$ is semisimple, its image in $\mathfrak{sl}_2(M)_i$ continues to be reductive.

Now, reductive subgroups of $\mathfrak{sl}_2(M)$ are either tori of all of $\mathfrak{sl}_2(M)$, so we merely need to check that the image cannot be a torus. If the image in some $\mathfrak{sl}_2(M)_i$ is a torus, then because the Galois action $\mathrm{Gal}(M/\mathbb{Q})$ permutes the decomposition of V into $\{V_i\}_i$ (but will fix $\mathrm{Hg}(V)$), so we see that the image in $\mathfrak{sl}_2(M)_i$ will continue to be a torus for all i. Explicitly, we note that the image of $\mathrm{Hg}(V)$ in $\mathrm{Sp}_{E\otimes_{\rho_i} M^\dagger}(\varphi|_{V_i})_M$ needs to be preserved under $\mathrm{Gal}(M/\mathbb{Q})$, so if the projection is commutative in one factor, then it is commutative in all factors because the $\mathbb{Q}$-points are dense. In particular, $\mathrm{Hg}(V)$ must be a torus, so A has complex multiplication by Proposition 2.53, which is a contradiction to its definition.

(ii) The first point of (ii) is automatic from the construction. The second point follows because all the $\mathfrak{sl}_2(M)_i$s include as the standard representation into $\mathfrak{gl}(V_i)$.

For the last point, we use the Galois action together with the hypothesis on the signature. Arguing as in the proof of Lemma 1.62, it is really enough to check the $(V_i)_M$s are non-isomorphic as $\mathfrak{hg}(V)_M^{\mathrm{der}}$-modules. To make sense of the signature, we choose an embedding $\varepsilon\colon M \to \mathbb{C}$, and then Lemma 2.72 grants a signature Φ_ε from the decomposition of V_ε into $E\otimes_\varepsilon \mathbb{C}$-eigenspaces: explicitly, for each embedding $\sigma \in \mathrm{Hom}(E, M)$, we find

$$\Phi_\varepsilon(\sigma) = \dim(V_\sigma)_\varepsilon^{1,0},$$

where $(\cdot)^{1,0}$ signifies that we are taking the eigenspace where $i \in \mathbb{C}$ acts by i^{-1}. However, the choice of a different embedding ε will permute the V_σs in sight.

To explain how the signature is now used, we note that if $\{\Phi_\varepsilon(\sigma), \Phi_\varepsilon(\overline{\sigma})\} \neq \{\Phi_\varepsilon(\tau), \Phi_\varepsilon(\overline{\tau})\}$ for two embeddings $\sigma, \tau \in \mathrm{Hom}(E, M)$ where $\rho_i = \sigma|_{E^\dagger}$ and $\rho_j = \tau|_{E^\dagger}$, then we must have $V_i \not\cong V_j$ as $\mathfrak{hg}(V)_\varepsilon^{\mathrm{der}}$-modules. Indeed, unwrapping the definition of the signature, we know that the projection of $\mathfrak{hg}(V)_{\mathbb{R}}$ (where the embedding $M^\dagger \hookrightarrow \mathbb{R}$ is given by the restriction of ε) into $\mathfrak{gl}_4(\mathbb{R})$ is

$$\mathfrak{so}(\Phi_\varepsilon(\sigma), \Phi_\varepsilon(\overline{\sigma})).$$

To see this, note that this a semisimple algebra of the correct rank, so it is enough remark that the image of $\mathfrak{hg}(V)^{\mathrm{der}}$ must land in the above Lie subalgebra by tracking the action of $h(i)$. (One should use Theorem 2.147 in the ℓ-adic computation.) Thus, we are now able to remark that $\mathfrak{so}(\Phi_\varepsilon(\sigma), \Phi_\varepsilon(\overline{\sigma})) \not\cong \mathfrak{so}(\Phi_\varepsilon(\tau), \Phi_\varepsilon(\overline{\tau}))$.

To complete the proof, the hypothesis implies there exists exactly one pair $\{\sigma_0, \overline{\sigma}_0\}$ of embeddings $E \hookrightarrow \mathbb{C}$ such that $\Phi_\varepsilon(\sigma_0) = \Phi_\varepsilon(\overline{\sigma}_0) = 1$. Thus, for any two distinct embeddings $\sigma, \tau \in \mathrm{Hom}(E, M)$, we can choose ε so that $\varepsilon\sigma = \sigma_0$ but $\varepsilon\tau \neq \sigma_0$ and apply the previous paragraph. ∎

Remark 2.153. This argument is inspired by [Zar83, Remark 1.9.4], where "changing the embedding" is used similarly to conclude that the Hodge group is large.

2.4.3 Computing ℓ-Adic Monodromy

The previous subsection explains that one expects to be able to compute $G_\ell(A)^\circ = \mathrm{MT}(A)$. We now explain how to use a computation of $G_\ell(A)^\circ$ to compute $G_\ell(A)$ in full. The idea is to use the Galois action on Tate classes. Our exposition follows [GGL24, Sections 8.1–8.2]. We begin with some notation.

Notation 2.154. Fix an abelian variety A defined over a field K, and let ℓ be a prime such that $\mathrm{char}\, K \nmid \ell$. We will write $V := \mathrm{H}^1_{\acute{e}t}(A_{\overline{K}}, \mathbb{Q}_\ell)$. For each $n \geq 0$, we define W_n to be the spce of Tate classes in the nth tensor power, writing

$$W_n := \left(V^{\otimes n} \otimes V^{\vee\otimes n}\right)^{G_\ell(A)^\circ}.$$

We also write $W := \bigoplus_{n\geq 0} W_n$ for brevity.

Remark 2.155. Because A is an abelian variety, one has a polarization $V \otimes V \to \mathbb{Q}_\ell(1)$, so we see that one can replace W_n with

$$\left(V^{\otimes 2n}(n)\right)^{G_\ell(A)^\circ}.$$

Roughly speaking, the point is that the spaces $W_\bullet$ of Tate classes are able to keep track of $G_\ell(A)^\circ$.

Lemma 2.156. Fix an abelian variety A defined over a field K, and let ℓ be a prime such that $\operatorname{char} K \nmid \ell$, and define V and $W_\bullet$ as in Notation 2.154.

(a) If $G \subseteq \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$ fixes W, then $G \subseteq G_\ell(A)^\circ$.

(b) There is a finite-dimensional subspace $W' \subseteq W$ such that $G \subseteq \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$ fixes W' if and only if $G \subseteq G_\ell(A)^\circ$.

Proof. This essentially follows from Proposition 1.35.

(a) Recall $G_\ell(A)^\circ$ is reductive by Theorem 2.131. Thus, by Proposition 1.35, we know that if $G \subseteq \mathrm{GL}(V)$ fixes every $G_\ell(A)^\circ$-invariant in any

$$\bigoplus_{i=1}^{k} \left(V^{\otimes m_i} \otimes V^{\vee \otimes n_i}\right),$$

then $G \subseteq G_\ell(A)^\circ$. However, we claim that all $G_\ell(A)^\circ$-invariants in the above space can be found in W, which will complete the proof. Indeed, by Theorem 2.150, we see that the scalars $\mathbb{G}_{m,\mathbb{Q}_\ell}$ can be found in $G_\ell(A)^\circ$; however, these scalars act by the character $z \mapsto z^{m_i - n_i}$ on $V^{\otimes m_i} \otimes V^{\vee \otimes n_i}$, so any $G_\ell(A)^\circ$-invariant subspace must then have $m_i = n_i$.

(b) The above argument provides countably many equations (in the form of invariant tensors) which cut out $G_\ell(A)^\circ$. However, any algebraic subgroup of $\mathrm{GL}(V)$ will be cut out by finitely many equations, so we can choose W' to be the span of any such subset of finitely many defining equations. ∎

Remark 2.157. The proof of (b) in fact gives an effective way to compute the subspace W': simply write down enough tensor elements to cut out $G_\ell(A)^\circ \subseteq \mathrm{GL}(V)$.

We would now like to upgrade from $G_\ell(A)^\circ$ to $G_\ell(A)$.

Lemma 2.158. Fix an abelian variety A defined over a field K, and let ℓ be a prime such that $\operatorname{char} K \nmid \ell$, and define V and $W_\bullet$ as in Notation 2.154. For each $n \geq 0$, the subspace W_n is stabilized by $G_\ell(A)$.

Proof. We already know that $G_\ell(A)^\circ$ acts trivially on W_n, so this will follow purely formally from the fact that $G_\ell(A)^\circ$ is a normal subgroup of $G_\ell(A)$.

We would like to show that each $g \in G_\ell(A)$ stabilizes W_n. Well, W_n exactly consists of the $G_\ell(A)^\circ$-invariants inside $V^{\otimes n} \otimes V^{\vee \otimes n}$, so it suffices to show that gW_n is stabilized by $G_\ell(A)^\circ$. Well, for any $g_0 \in G_\ell(A)^\circ$, we see that

$$g_0 g W_n = g \cdot g^{-1} g_0 g W_n,$$

so we conclude by noting that $g^{-1} g_0 g \in G_\ell(A)^\circ$ because $G_\ell(A)^\circ \subseteq G_\ell(A)$ is a normal subgroup. ∎

Combining the above two lemmas, we see that we get a faithful representation

$$G_\ell(A)/G_\ell(A)^\circ \to \mathrm{GL}(W).$$

This faithful representation allows us to compute $G_\ell(A)$: we are looking for elements of $\mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$ which produce the automorphisms of W seen in the image of the above faithful representation. Tracking through this sort of reasoning produces our main result.

Proposition 2.159. Fix an abelian variety A defined over a field K, and let ℓ be a prime such that $\operatorname{char} K \nmid \ell$, and define V and $W_\bullet$ as in Notation 2.154. Then $G_\ell(A)$ equals the group

$$\bigcup_{\sigma \in \mathrm{Gal}(\overline{K}/K)} \{g \in \mathrm{GL}(V) : g|_W = \sigma|_W\}.$$

In fact, each set in the union is a connected component of $G_\ell(A)$.

Proof. We begin by noting that $\mathrm{Gal}(\overline{K}/K)$ does in fact preserve W: indeed, one has a composite

$$\mathrm{Gal}(\overline{K}/K) \to G_\ell(A) \to \mathrm{GL}(W),$$

where the first map is well-defined by the definition of $G_\ell(A)$, and the second map is well-defined by summing Lemma 2.158.

Now, we have two inclusions to show.

- Suppose $g \in G_\ell(A)$. Then we must find $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ such that $g|_W = \sigma|_W$. Well, $G_\ell(A)$ is by definition the Zariski closure of the image of $\mathrm{Gal}(\overline{K}/K)$ in $\mathrm{GL}(V)$, so the open subset $gG_\ell(A)^\circ$ of $G_\ell(A)$ must contain $\sigma|_V$ for some $\sigma \in \mathrm{Gal}(\overline{K}/K)$. Now, $G_\ell(A)^\circ$ acts trivially on W, so we see that $g|_W = \sigma|_W$.
- Suppose $g \in \mathrm{GL}(V)$ satisfies $g|_W = \sigma|_W$. Then we would like to show that $g \in G_\ell(A)$. The argument in the previous point grants $g_0 \in G_\ell(A)$ such that $g_0|_V = \sigma|_V$, so in particular, $g|_W = g_0|_W$. Thus, gg_0^{-1} acts trivially on W, so $gg_0^{-1} \in G_\ell(A)^\circ$, so it follows that $g \in G_\ell(A)$.

Lastly, it remains to discuss connected components. Well, note that $g, g' \in G_\ell(A)$ live in the same connected component if and only if $g'g^{-1} \in G_\ell(A)$, which is equivalent to $g'g^{-1}$ acting trivially on W, which is equivalent to $gG_\ell(A)^\circ = g'G_\ell(A)^\circ$. ∎

Remark 2.160. A careful reading of the above proof shows that we only needed the following facts about W: it is stable under $G_\ell(A)$, and $g \in \mathrm{GL}(V)$ lives in $G_\ell(A)^\circ$ if and only if it fixes W. Thus, we see that we can replace W with any $G_\ell(A)$-subrepresentation $W' \subseteq W$ which cuts out $G_\ell(A)^\circ$ in the sense of Lemma 2.158. This allows us to make W' quite small (e.g., finite-dimensional).

Remark 2.161. It is worth comparing Proposition 2.159 with the twisted Lefschetz group, defined in [BK15, Definition 5.2]. Roughly speaking, the twisted Lefschetz group is simply the construction of Proposition 2.159 with W replaced by the subspace of W generated by endomorphisms and the polarization; see [GGL24, Remark 8.3.5] for precise discussion of the relation. In this way, one expects the twisted Lefschetz group to equal $G_\ell(A)$ in generic cases, but Remark 2.160 explains that one may need to remember more Hodge classes in exceptional cases.

Proposition 2.159 suggests that one can find representatives of each connected component in $G_\ell(A)$ by looping over all $\sigma \in \mathrm{Gal}(\overline{K}/K)$ and finding some $g \in \mathrm{GL}(V)$ such that $g|_W = \sigma|_W$. This is currently not so computable because $\mathrm{Gal}(\overline{K}/K)$ is an infinite group, and W is an infinite-dimensional vector space. Remark 2.160 explains how to replace W with a finite-dimensional subrepresentation, so it remains to explain how to reduce $\mathrm{Gal}(\overline{K}/K)$ to a finite quotient.

Definition 2.162 (connected monodromy field)**.** Fix an abelian variety A defined over a field K, and let ℓ be a prime such that $\operatorname{char} K \nmid \ell$. Then we define the *connected monodromy field* K_A^{conn} so that the open subgroup $\mathrm{Gal}(\overline{K}/K_A^{\mathrm{conn}})$ is the pre-image of the connected component $G_\ell(A)^\circ$ in the Galois representation

$$\mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right).$$

Remark 2.163. Note that such a field K_A^{conn} exists and is finite over K by Galois theory: note $G_\ell(A)^\circ \subseteq G_\ell(A)$ is a finite-index subgroup (because the quotient is a discrete algebraic group), so the pre-image $U \subseteq \mathrm{Gal}(\overline{K}/K)$ of $G_\ell(A)^\circ$ similarly must be open and finite index and hence closed and finite index.

Thus, we see that the Galois reprentation to $\mathrm{GL}(W)$ factors through the finite group $\mathrm{Gal}(K_A^{\mathrm{conn}}/K)$. In this way, we are able to reduce the computation suggested by Proposition 2.159 from the infinite group $\mathrm{Gal}(\overline{K}/K)$ to the finite quotient $\mathrm{Gal}(K_A^{\mathrm{conn}}/K)$.

Remark 2.164. Let's describe how one might compute K_A^{conn} in practice. By combining the definition of K_A^{conn} with Lemma 2.156, we see that $\mathrm{Gal}(\overline{K}/K_A^{\mathrm{conn}})$ is the kernel of the representation

$$\mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(W),$$

so one could imagine computing the open subgroup $\mathrm{Gal}(\overline{K}/K_A^{\mathrm{conn}})$ by computing the above representation. As usual, we remark that Lemma 2.156 allows us to replace W with a finite-dimensional subrepresentation W' "cutting out" $G_\ell(A)^\circ$.

2.4.4 The Motivic Galois Group

In this last subsection, we recast some of our monodromy discussions motivically. The Mumford–Tate conjecture is more or less an assertion that there should really only be one monodromy group for an abelian variety. This indicates that there should be a motivic version of this conjecture. Here is one formulation, using our category of motives.

Definition 2.165. Fix a motive M over an algebraic extension K of $\mathbb{Q}$.

- For a fixed embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, we define the *Mumford–Tate group* $\mathrm{MT}(M)$ as the Mumford–Tate group of the rational Hodge stucture $\mathrm{H}_\sigma(M)$. (See Remark 1.211.)
- For each prime ℓ, we define the *ℓ-adic monodromy group* as the smallest algebraic subgroup containing the image of
$$\mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}\left(\omega_\ell(M)\right),$$
where $\omega_\ell\colon \mathrm{Mot}_{\mathbb{Q}}(K) \to \mathrm{Rep}_{\mathbb{Q}_\ell}\mathrm{Gal}(\overline{K}/K)$ is given by ℓ-adic cohomology. (See Remark 1.212.)

Remark 2.166. The same arguments as in Example 1.143 and Remark 2.120 show that $\mathrm{MT}(M)$ is the algebraic group attached to the subcategory $\langle \mathrm{H}_\sigma^\bullet(M)\rangle^\otimes \subseteq \mathrm{HS}_{\mathbb{Q}}$, and $G_\ell(M)$ is the algebraic group attached to the subcategory $\langle \omega_\ell(M)\rangle^\otimes \subseteq \mathrm{Rep}_{\mathbb{Q}_\ell}\mathrm{Gal}(\overline{K}/K)$.

Example 2.167. Fix an abelian variety A. Because $\langle \mathrm{H}_\sigma^\bullet(A)\rangle^\otimes = \langle \mathrm{H}_\sigma^1(A)\rangle^\otimes$ (by Theorem 2.98) we see $\mathrm{MT}(h(A)) = \mathrm{MT}(A)$. The same argument shows $G_\ell(h(A)) = G_\ell(A)$.

Conjecture 2.168 (Motivic Mumford–Tate)**.** Fix a motive M over a number field K. For each prime ℓ, we have

$$\mathrm{MT}(M)_{\mathbb{Q}_\ell} = G_\ell(M)^\circ.$$

(More precisely, these are isomorphic via the embeddings of Remarks 2.179 and 2.180.)

Example 2.169. Let's prove the conjecture when M is an Artin motive. On one hand, $H^\bullet_\sigma(M_{\overline{K}})$ has Hodge structure concentrated in bidegree $(0,0)$, so $\mathrm{MT}(M)$ is trivial. On the other hand, $G_\ell(M)$ is algebraic (by its construction) and a quotient of the profinite group $\mathrm{Gal}(\overline{K}/K)$ (by Example 1.218) and thus finite. We conclude $G_\ell(M)^\circ$ is trivial and thus agrees with $\mathrm{MT}(M)_{\mathbb{Q}_\ell}$.

However, a motivic formulation not only tells us what to expect more generally, but it will tell us what the more general monodromy group attached to a motive should be. Following Remark 2.166, we are motivated to define a motivic monodromy group as follows.

Definition 2.170 (motivic Galois group)**.** Fix an algebraic extension K of $\mathbb{Q}$.

- For a set of motives $S \subseteq \mathrm{Mot}_{\mathbb{Q}}(K)$, we define the *motivic Galois group* $G_{\mathrm{mot},K}(S)$ to be the algebraic group associated with the tensor subcategory $\langle S\rangle \subseteq \mathrm{Mot}_{\mathbb{Q}}(K)$.
- If $S = \{M\}$ is a singleton, we may write $G_{\mathrm{mot},K}(M)$.
- Further, if $M = h(X)$, we may write $G_{\mathrm{mot},K}(X)$.

We will omit the subscripted field K from the notation as much as possible. If we want to specify the fiber functor ω_σ (for an embedding $\sigma\colon K \hookrightarrow \mathbb{C}$) in this notation, we may write G_σ instead of G_{mot}.

Example 2.171. Let M be an Artin motive. In this case, Example 1.218 explains that we may identify M with the Galois representation $\omega_\sigma(M)$. Then the same argument as in Remark 2.120 shows that $G_{\mathrm{mot}}(M)$ is exactly the image of the structure map

$$\mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(\omega_\sigma(M)).$$

Remark 2.172. For an abelian variety A, Remark 2.106 explains why $G_{\mathrm{mot}}(A) = G_{\mathrm{mot}}\left(h^1(A)\right)$: the tensor categories of these motives are the same!

Remark 2.173. For an abelian variety A of dimension g, we claim that $G_{\mathrm{mot}}(A) \subseteq \mathrm{GL}_{2g,\mathbb{Q}}$. Indeed, choosing an embedding $\omega\colon K \hookrightarrow \mathbb{C}$ will induce a group homomorphism

$$\underline{\mathrm{Aut}}^{\otimes}\omega_\sigma|_{\langle h(A)\rangle^\otimes} \to \mathrm{GL}\left(\mathrm{H}^1_\sigma(A)\right)$$

given by explaining how a given $\otimes$-automorphism of ω_σ acts on $\omega_\sigma\left(h^1(A)\right)$. The corresponding functor $\mathrm{Rep}_{\mathbb{Q}}\mathrm{GL}\left(\mathrm{H}^1_\sigma(A)\right) \to \langle h(A)\rangle^\otimes$ simply takes the tensor generator $\mathrm{H}^1_\sigma(A)$ back to the tensor generator $h^1(A)$, so the above group homomorphism is an embedding by Proposition 1.154.

Remark 2.174. In fact, $G_{\mathrm{mot}}(A) \subseteq \mathrm{GSp}_{2g,\mathbb{Q}}$. For a given polarization $A \to A^\vee$, the induced Weil pairing and polarization on Hodge structures assemble into an absolute Hodge correspondence $h^1(A) \otimes h^1(A) \to \mathrm{L}$. Now, each $g \in G_{\mathrm{mot}}(A)$ must commute with this absolute Hodge correspondence, which means (on the Betti realization, say) that g preserves the induced perfect pairing on $\mathrm{H}^1_\sigma(A)$ up to a scalar given by the action of g on L.

For example, one expects that $G_{\mathrm{mot}}(M)^\circ = \mathrm{MT}(M)$ and $G_{\mathrm{mot}}(M)_{\mathbb{Q}_\ell} = G_\ell(M)$, but we cannot expect to be able to prove these equalities easily because they together imply the (Motivic) Mumford–Tate conjecture. Fortunately, we will be able to prove the former equality when M is an abelian variety, and we will then be able to show that the latter equality is equivalent to the Mumford–Tate conjecture. This is the goal of the present section.

Let's begin with the equality $G_{\mathrm{mot}}(A)^\circ = \mathrm{MT}(A)$.

Lemma 2.175. Let A be an abelian variety defined over an algebraic extension K of $\mathbb{Q}$. Then $G_{\mathrm{mot}}(A_{\overline{K}}) = \mathrm{MT}(A)$.

Proof. This is [DM12, Proposition 6.22(a)]. Fix an embedding $\sigma\colon K \hookrightarrow \mathbb{C}$. Then Remark 1.211 explains that the fiber functor $\omega_\sigma\colon \mathrm{Mot}_{\mathbb{Q}}(K) \to \mathrm{HS}_{\mathbb{Q}}$ is faithful, but because all Hodge classes on abelian varieties are in fact absolute Hodge classes, we will be able to show that the restricted functor

$$\omega_\sigma\colon \langle h(A)\rangle^{\otimes} \to \mathrm{HS}_{\mathbb{Q}}$$

is fully faithful. Indeed, $\langle h(A)\rangle^{\otimes}$ is made of quotients of objects which look like $\bigoplus_i h(A)^{n_i} \otimes (h(A)^\vee)^{m_i}$, but Poincaré duality (in Theorem 1.210) explains $h(A)^\vee = h(A)(\dim A)$, so we may work with quotients of objects which look like

$$\bigoplus_i h(A)^{n_i}(m_i \dim A).$$

But then correspondences between such quotients may as well be lifted up to absolute Hodge classes on disjoint unions of powers of A, which are the same as Hodge classes by Theorem 2.45, so we may unwind our correspondences to merely be given by Hodge classes! This shows that ω_σ is fully faithful on the subcategory $\langle h(A)\rangle^{\otimes}$.

To finish the proof, we see that the induced functor

$$\omega_\sigma\colon \langle h(A)\rangle^{\otimes} \to \langle \mathrm{H}^\bullet_\sigma(A)\rangle^{\otimes}$$

is an equivalence (it is essentially surjective by construction), so the groups given by Tannakian reconstruction must be isomorphic. ∎

Remark 2.176. In fact, the proof shows that we expect to have $G_{\mathrm{mot}}(M_{\overline{K}}) = \mathrm{MT}(M)$, but we only know achieve this once we know that all Hodge classes in $\langle M\rangle^{\otimes}$ are absolute Hodge. Nonetheless, Proposition 1.152 explains that the proof may take ω_σ and produce an embedding $\mathrm{MT}(M) \to G_{\mathrm{mot}}(M_{\overline{K}})$ for any motive M.

Lemma 2.177. Fix any set S of motives over an algebraic extension K of $\mathbb{Q}$, and let Γ be the Tannakian group of the category $\langle S\rangle^{\otimes} \cap \mathrm{Mot}^0_{\mathbb{Q}}(K)$. Then there is an exact sequence

$$1 \to G_{\mathrm{mot}}(S_{\overline{K}}) \to G_{\mathrm{mot}}(S) \to \Gamma \to 1.$$

Proof. This is [DM12, Proposition 6.23]. Throughout this argument, we are fixing an algebraic closure $\overline{K}$ and K along with a frequently implicit embedding $\iota\colon K \subseteq \overline{K}$. We will also need to choose an embedding $\sigma\colon \overline{K} \hookrightarrow \mathbb{C}$. Anyway, we proceed in steps.

1. We describe the left map. There is a natural functor $\langle S\rangle^{\otimes} \to \langle S_{\overline{K}}\rangle^{\otimes}$ given by base-changing our motives (along ι). By construction, Proposition 1.152 explains that the relevant group homomorphism $p\colon G_{\mathrm{mot}}(S_{\overline{K}}) \to G_{\mathrm{mot}}(S)$ is an embedding.

   It will be worthwhile to explicate this map somewhat: given some $g \in G_{\mathrm{mot}}(S_{\overline{K}})$, we note that g is really an automorphism of the $\otimes$-functor ω_σ on $\langle S_{\overline{K}}\rangle^{\otimes}$. But then g induces an automorphism on $\langle S\rangle^{\otimes}$ (and hence an element $i(g) \in G_{\mathrm{mot}}(S)$) as

   $$\omega_{\sigma\iota}(M) = \omega_\sigma(M_{\overline{K}}) \xrightarrow{g} \omega_\sigma(M_{\overline{K}}) = \omega_{\sigma\iota}(M).$$

   Namely, because g is already an automorphism of $\otimes$-functors, we see that $i(g)$ is as well.

2. We describe the right map. There is a fully faithful embedding $\langle S\rangle^{\otimes} \cap \mathrm{Mot}^0_{\mathbb{Q}}(K) \subseteq \langle S\rangle^{\otimes}$, so our Tannakian formalism (Proposition 1.152) induces an embedding $p\colon G_{\mathrm{mot}}(S) \to \Gamma$. As in the previous point, we may view p as restricting an automorphism of the $\otimes$-functor $\omega_{\sigma\iota}$ from $\langle M\rangle^{\otimes}$ to the subcategory $\langle S\rangle^{\otimes} \cap \mathrm{Mot}^0_{\mathbb{Q}}(K)$.

3. The above remarks have already provided exactness of our sequence on the left and right. It remains to show exactness at $G_{\mathrm{mot}}(S)$. One of these checks is easier: we start by showing that $p \circ i$ is trivial. Namely, for any $g \in G_{\mathrm{mot}}(S_{\overline{K}})$, we must show that $i(g)$ fixes $\omega_{\sigma\iota}(M)$ for any $M \in \langle S\rangle^{\otimes} \cap \mathrm{Mot}_{\mathbb{Q}}(K)$. Upon unwinding the definition of $i(g)$, we see that we would like to check that g fixes $\omega_{\sigma}(M_{\overline{K}})$. It will be enough to check that any $\otimes$-automorphism of ω_{σ} acting on $\langle S_{\overline{K}}\rangle^{\otimes} \cap \mathrm{Mot}^0_{\mathbb{Q}}(\overline{K})$ is trivial, but this is not hard: this category is just $\langle h(\operatorname{Spec}\overline{K})\rangle^{\otimes}$, and any $\otimes$-automorphism will fix the unit.

4. We finish showing exactness in the middle. Suppose $g \in G_{\mathrm{mot}}(S)$ goes to the identity in Γ, and we want to show that $g \in \operatorname{im} i$. The main point is to show that $g_M \in \operatorname{Aut}\omega_{\sigma\iota}(M)$ only depends on $M_{\overline{K}}$.

   For a moment, choose two motives $M, N \in \mathrm{Mot}_{\mathbb{Q}}(K)$, which we will assume to be isomorphic after base-change to $\overline{K}$ in a moment. Observe that $\mathrm{Hom}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M_{\overline{K}}, N_{\overline{K}})$ is some subspace of absolute Hodge classes, so it is a Galois representation by Remark 1.194.[4] It follows that we may view $\mathrm{Hom}_{\mathrm{Mot}_{\mathbb{Q}}(K)}(M_{\overline{K}}, N_{\overline{K}})$ as an Artin motive in $\mathrm{Mot}^0_{\mathbb{Q}}(K)$ via Example 1.218, so g acts trivially on this motive. This means that the action of g fixes the relevant absolute Hodge correspondences, which causes the diagram

$$\begin{array}{ccc} \omega_{\sigma\iota}(M_{\overline{K}}) & \xrightarrow{g_M} & \omega_{\sigma\iota}(M_{\overline{K}}) \\ {\scriptstyle \omega_{\sigma\iota}(f)}\downarrow & & \downarrow{\scriptstyle \omega_{\sigma\iota}(f)} \\ \omega_{\sigma\iota}(N_{\overline{K}}) & \xrightarrow{g_N} & \omega_{\sigma\iota}(N_{\overline{K}}) \end{array}$$

   to commute for any $f\colon M_{\overline{K}} \to N_{\overline{K}}$. For example, upon taking f to be an isomorphism, we are left with the statement that g_M and g_N are the same automorphism.

   As such, we may define $\overline{g} \in \underline{\mathrm{Aut}}^{\otimes}\omega_{\sigma}$ by $\overline{g}_{M_{\overline{K}}} \coloneqq g_M$, which the previous paragraph promises is well-defined. (These motives generate our category, so $\overline{g}$ can be uniquely extended to kernels and tensor products because it is already a linear $\otimes$-automorphism where it is defined.) Then $i(\overline{g}) = g$ by construction. ∎

Proposition 2.178. Fix an abelian variety A over an algebraic extension K of $\mathbb{Q}$. Then

$$G_{\mathrm{mot}}(A)^{\circ} = \mathrm{MT}(A).$$

Proof. Plugging the equality $\mathrm{MT}(A) = G_{\mathrm{mot}}(A_{\overline{K}})$ of Lemma 2.175 into Lemma 2.177 yields the exact sequence

$$1 \to \mathrm{MT}(A) \to G_{\mathrm{mot}}(A) \to \Gamma \to 1,$$

where Γ is some quotient of $\mathrm{Gal}(\overline{K}/K)$ by Example 1.218. In particular, Γ is thus a quotient of a profinite group and an algebraic group $G_{\mathrm{mot}}(A)$ by Proposition 1.153, so Γ must be finite.

Now, on one hand, $\mathrm{MT}(A)$ is connected by Remark 1.30, so $\mathrm{MT}(A) \subseteq G_{\mathrm{mot}}(A)^{\circ}$ follows. On the other hand, Γ is discrete, so $G_{\mathrm{mot}}(A)^{\circ}$ must be contained in the kernel of the right-hand projection, which is exactly $\mathrm{MT}(A)$ by exactness. The result follows. ∎

Remark 2.179. Continuing from Remark 2.176, we see that this proof shows $G_{\mathrm{mot}}(M)^{\circ} = \mathrm{MT}(M)$ for an arbitrary motive as soon as we know that all Hodge classes are absolutely Hodge, and one can always construct an embedding $\mathrm{MT}(M) \to G_{\mathrm{mot}}(M)^{\circ}$.

We now turn to the second equality $G_{\mathrm{mot}}(M)_{\mathbb{Q}_\ell} = G_\ell(M)$, which is called a "motivic analogue of the Tate conjecture" in [CC22].

[4] Fixing a degree via Tate twists and taking idempotent subspaces are both Galois-invariant operations, so the subspace of absolute Hodge classes continues to be Galois-invariant.

Remark 2.180. One can construct a candidate isomorphism for Conjecture 2.181. The comparison isomorphism (in the form of Remark 1.213) shows that the fiber functors $\mathrm{Mot}_{\mathbb{Q}}(K) \to \mathrm{Vec}_{\mathbb{Q}_\ell}$ defined by $M \mapsto \omega_\sigma(M)_{\mathbb{Q}_\ell}$ and $M \mapsto \omega_\ell(M)$ are naturally isomorphic. This induces a morphism $\langle M \rangle^\otimes \to \langle \omega_\ell(M) \rangle^\otimes$ of neutral Tannakian categories, which then induces the desired map $G_\ell(M) \to G_{\mathrm{mot}}(M)$. Example 1.140 explains that this map is an embedding.

Conjecture 2.181. Fix a motive M over a number field K. For each prime ℓ, the canonical map

$$G_\ell(M) \to G_{\mathrm{mot}}(M)_{\mathbb{Q}_\ell}$$

of Remark 2.180 is an isomorphism.

Example 2.182. Let's prove the conjecture when M is an Artin motive. Well, the comparison isomorphism Remark 1.213 explains that there is an isomorphism $\omega_\sigma(M)_{\mathbb{Q}_\ell} \to \omega_\ell(M)$ of Galois representations, so we are done as soon as we compare Example 2.171 with the definition of $G_\ell(M)$.

Intuitively, one should expect Conjecture 2.181 to follow by independently comparing identity components and component groups. Proposition 2.178 indicates that comparing the identity components will require some input from the Mumford–Tate conjecture, but luckily, we can compare the component groups less conjecturally.

Lemma 2.183. Fix a motive M over a number field K. Then the canonical map $G_\ell(M) \to G_{\mathrm{mot}}(M)$ of Remark 2.180 induces a surjection

$$\pi_0 G_\ell(M) \to \pi_0 G_{\mathrm{mot}}(M)_{\mathbb{Q}_\ell}.$$

Proof. The idea is that finite groups should correspond to Artin motives, where the conjecture is already known by Example 2.182. Let's begin by finding the relevant Artin motive: the quotient map $G_{\mathrm{mot}}(M) \twoheadrightarrow \pi_0 G_{\mathrm{mot}}(M)$ induces an embedding

$$\mathrm{Rep}_{\mathbb{Q}} \pi_0 G_{\mathrm{mot}}(M) \hookrightarrow \langle M \rangle^\otimes.$$

The left-hand category has a tensor generator (e.g., take the regular representation of the finite group $\pi_0 G_{\mathrm{mot}}(M)$), so the essential image has a tensor generator $N \in \langle M \rangle^\otimes$. To see that this is an Artin motive, we note that $\pi_0 G_{\mathrm{mot}}(M)$ is a quotient of the motivic Galois group of $\langle M \rangle^\otimes \cap \mathrm{Mot}^0_{\mathbb{Q}}(K)$ by Lemma 2.177, so we must have $N \in \mathrm{Mot}^0_{\mathbb{Q}}(K)$.

Let's explain why N is the Artin motive we are looking for: by the construction of N, the category $\langle N \rangle^\otimes$ is equivalent to $\mathrm{Rep}_{\mathbb{Q}} \pi_0 G_{\mathrm{mot}}(M)$, so $G_{\mathrm{mot}}(N) = \pi_0 G_{\mathrm{mot}}(M)$. We are now ready to complete the proof: the commutative diagram

$$\begin{array}{ccc} \langle N \rangle^\otimes & \hookrightarrow & \langle M \rangle^\otimes \\ {\scriptstyle \omega_\ell}\downarrow & & \downarrow{\scriptstyle \omega_\ell} \\ \langle \omega_\ell(N) \rangle^\otimes & \hookrightarrow & \langle \omega_\ell(M) \rangle^\otimes \end{array}$$

induces a commutative diagram

$$\begin{array}{ccc} G_\ell(M) & \twoheadrightarrow & G_\ell(N) \\ \downarrow & & \downarrow \\ G_{\mathrm{mot}}(M)_{\mathbb{Q}_\ell} & \twoheadrightarrow & G_{\mathrm{mot}}(N)_{\mathbb{Q}_\ell} \end{array}$$

where the right-hand arrow is in fact an isomorphism by Example 2.182. We conclude that the induced map $G_\ell(M) \to \pi_0 G_{\mathrm{mot}}(M)$ is surjective, so the claim follows. ∎

Proposition 2.184. Fi an abelian variety A over a number field K. Then Conjecture 2.181 for A is equivalent to the Mumford–Tate conjecture for A.

Proof. This is part of [CC22, Theorem]. Remark 2.180 induces a morphism

$$\begin{array}{ccccccccc} 1 & \longrightarrow & G_\ell(A)^\circ & \longrightarrow & G_\ell(A) & \longrightarrow & \pi_0 G_\ell(A) & \longrightarrow & 1 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 1 & \longrightarrow & G_{\mathrm{mot}}(A)^\circ_{\mathbb{Q}_\ell} & \longrightarrow & G_{\mathrm{mot}}(A)_{\mathbb{Q}_\ell} & \longrightarrow & \pi_0 G_{\mathrm{mot}}(A)_{\mathbb{Q}_\ell} & \longrightarrow & 1 \end{array}$$

of short exact sequences. Quickly, we note that the left map is injective because the middle map is injective, and the right map is surjective by Lemma 2.183. Additionally, we note that the canonical map $\mathrm{MT}(A) \to G_{\mathrm{mot}}(A)^\circ$ is an isomorphism by Proposition 2.178.

Before continuing, we note that the Mumford–Tate conjecture (in the form Conjecture 2.144) is equivalent to the induced map $G_\ell(A)^\circ \to \mathrm{MT}(A)_{\mathbb{Q}_\ell}$ being an isomorphism. Indeed, perhaps one can be worried that the map constructed in Conjecture 2.144 is not this map, but it is: the embedding $\mathrm{MT}(A) \to \mathrm{GL}\left(\mathrm{H}^1_{\mathrm{B}}(A)\right)$ simply asks how $\mathrm{MT}(A)$ should act on the vector space $\mathrm{H}^1_{\mathrm{B}}(A)$ and thus factors through $G_{\mathrm{mot}}(A)$ by Remarks 2.173 and 2.176. Similarly, the embedding $G_\ell(A) \to \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$ again factors through $G_{\mathrm{mot}}(A)$ via the discussion of Remark 2.180.

We now show both directions of the proposition independently.

- Given the Mumford–Tate conjecture, the snake lemma now implies that the left and right maps being surjective implies that the middle map is surjective, thereby proving Conjecture 2.181.
- Given Conjecture 2.181, we see that the left map is an isomorphism because taking identity components is functorial. ∎

Chapter 3
The Sato–Tate Conjecture

Well, there's only one way I've ever seen something substantial proved about an arithmetic L-function—and that's to relate it to automorphic forms.

—Richard Taylor, [Tay07]

The classical application of monoromy groups is to the Sato–Tate conjecture, which we now discuss.

3.1 The Statement

In this section, we state the Sato–Tate conjecture, and then we explain how it can be numerically verified.

3.1.1 The Weil Conjectures

The Sato–Tate conjecture is about counting points on an abelian variety A over finite fields $\mathbb{F}_q$ as q varies. In this subsection, we will describe the Weil conjectures because they explain why these point-counts ought to be related to cohomology.

The main character of our story is a zeta function.

Definition 3.1. Fix a variety X over a finite field $\mathbb{F}_q$. Then its *zeta function* is the formal power series

$$Z_X(T) \coloneqq \exp\left(\sum_{m=1}^{\infty} \#X\left(\mathbb{F}_{q^m}\right) \frac{T^m}{m}\right).$$

Remark 3.2. It is not important why this is the precise definition of the ζ-function, but we remark that there is a general definition of ζ_X when X is a scheme of finite type over $\mathbb{Z}$, given by

$$\zeta_X(s) = \prod_{\text{closed } x \in X} \frac{1}{1 - \mathrm{N}(x)^{-s}},$$

where $\mathrm{N}(x) \coloneqq \#\kappa(x)$; for example, for a number field K, $\zeta_{\operatorname{Spec} \mathcal{O}_K}(s) = \zeta_K(s)$. A purely formal argument can verify that $\zeta_X(s) = Z_X(q^{-s})$. Roughly speaking, one merely has to write $\log \zeta_X(s)$ out as a sum over closed points x and then group the terms by $\mathrm{N}(x)$.

The Weil conjectures [Wei49, p. 507], now theorems due to Deligne [Del74; Del80], assert the following properties for $Z_X(T)$ when X is smooth, geometrically irreducible, projective variety over $\mathbb{F}_q$ of dimension d.

1. Rationality: $Z_X(T) \in \mathbb{Q}(T)$.
2. Functional equation: there is a sign $\pm$ such that
$$Z_X\left(\frac{1}{q^d T}\right) = \pm q^{d\chi/2} T^\chi Z_X(T),$$
where χ is a suitably defined Euler characteristic $\deg(\Delta \cdot \Delta)$.
3. Riemann hypothesis: there is a factorization
$$Z_X(T) = \frac{P_1(T) \cdots P_{2n-1}(T)}{P_2(T) \cdots P_{2d}(T)}$$
so that $P_1(T) = 1 - T$, $P_{2d}(T) = 1 - q^d T$, and each $P_i(T)$ has $P(0) = 1$ and has roots which all take the form $1/\alpha_i$ where α_i is an algebraic integer of magnitude $q^{i/2}$.
4. Betti numbers: if X admits a smooth projective integral model $\mathcal{X}$, then
$$\deg P_i = \dim_{\mathbb{Q}} \mathrm{H}^i_{\mathrm{B}}(\mathcal{X}(\mathbb{C}), \mathbb{Q})$$
for each i.

All of these conjectures are proven using cohomological methods. We will not explain the proofs of all of them here, but we will show everything except the Riemann hypothesis. Our exposition follows [Har77, Appendix C.4].

In short, the proofs we provide will take some cohomological input and then do a little linear algebra to prove a theorem. For example, to explain how cohomology enters our picture via the Lefschetz trace formula (Theorem 1.125), we need the following linear algebraic lemma.

Lemma 3.3. Let $f \colon V \to V$ be an endomorphism of a vector space. Then there is an equality of formal power series
$$-\log\det(1 - fT; V) = \sum_{m=1}^{\infty} \operatorname{tr}(f^m; V) \frac{T^m}{m}.$$

Proof. Taking traces and determinants is immune to base-change by a field, so we may assume that V is a vector space over an algebraically closed field. Then we may choose a basis of V so that the matrix representing f is upper-triangular. Now, we set $n \coloneqq \dim V$, and we let $\{\lambda_1, \ldots, \lambda_n\}$ be the diagonal entries of f. Then
$$-\log\det(1 - fT; V) = \sum_{i=1}^{n} -\log(1 - \lambda_i T),$$
and
$$\operatorname{tr}(f^m; V) = \sum_{i=1}^{n} \lambda^i m,$$
so we conclude the proof using the formal power series expansions $-\log(1 - \lambda_i T) = \sum_{m \geq 1} \lambda_i^m T^m / m$. ∎

Proposition 3.4. Fix a Weil cohomology theory $\mathrm{H}^\bullet$ over $\mathbb{F}_q$ with coefficients in F. Choose a smooth projective variety X over a finite field $\mathbb{F}_q$ of equidimension d. Letting $\mathrm{Frob}_q \colon X \to X$ be the q-power (absolute) Frobenius endomorphism, we have

$$Z_X(T) = \prod_{i=0}^{2d} \det\left(1 - \mathrm{Frob}_q^* T; \mathrm{H}^i(X)\right)^{(-1)^{i+1}}.$$

Proof. We proceed in steps.

1. For expositional reasons, we isolate the "motivic" input in this proof: we claim the diagonal $\Delta \colon X \to X \times X$ and the graph Γ of Frob_q intersect transversely. Because Γ and Δ have the correct dimensions to intersect transversely, it is enough to check their Zariski tangent spaces are disjoint on the intersection. On one hand, for $x \in X(\mathbb{F}_q)$, we see that
$$T_{(x,x)}\Delta = \{(v, v) : v \in T_x X\}$$
because Δ is the image of $(\mathrm{id}_X, \mathrm{id}_X) \colon X \to X \times X$. On the other hand, the action of $\mathrm{Frob}_q \colon X \to X$ on the Zariski cotangent space is an endomorphism $d(\mathrm{Frob}_q)_x$ of $\mathfrak{m}_x/\mathfrak{m}_x^2$ which can be computed to vanish: the derivative of the q-power map in positive characteristic will vanish! Thus,
$$T_{(x,x)}\Gamma = \{(v, 0) : v \in T_x X\}$$
because Γ is the image of $(\mathrm{id}_X, \mathrm{Frob}_q) \colon X \to X \times Y$. This tangent spaces are in fact disjoint.

2. We now claim that $\#X(\mathbb{F}_q) = \deg(\Gamma \cdot \Delta)$, where $\Delta \colon X \to X \times X$ is the diagonal. To begin, by embedding X into some projective space, where the action of the Frobenius can be seen as taking q-powers on points, we see that $X(\mathbb{F}_q)$ is exactly the set of (geometric) points fixed by f. However, the set of $\mathbb{F}_q$-points fixed by the Frobenius is count is simply $\Gamma(\mathbb{F}_q) \cap \Delta(\mathbb{F}_q)$, whose cardinality will equal $\deg(\Gamma \cdot \Delta)$ because Γ and Δ intersect transversely.

3. We are now in a position to use the Lefschetz trace formula, so the remainder of the proof is a calculation. Indeed, by Theorem 1.125, we may take powers of Frob_q in the previous two steps to see that $\#X(\mathbb{F}_{q^m})$ equals
$$\deg(\Delta \cdot \Gamma_{\mathrm{Frob}_q^m}) = \sum_{i=0}^{2d} (-1)^i \operatorname{tr}\left((\mathrm{Frob}_q^m)^*; \mathrm{H}^i(X)\right).$$
Thus, summing over all m, we see that
$$\sum_{m=1}^{\infty} \#X(\mathbb{F}_{q^m}) \frac{T^m}{m} = \sum_{i=0}^{2d} (-1)^{i+1} \log\det(1 - \mathrm{Frob}_q^*; \mathrm{H}^i(X))$$
by Lemma 3.3. Taking exponentials completes the proof. ∎

Remark 3.5 (Betti numbers)**.** In the Riemann hypothesis conjecture, one takes
$$P_i(T) \coloneqq \det\left(1 - \mathrm{Frob}_q^* T; \mathrm{H}^i(X)\right).$$
Thus, taking $\mathrm{H}^\bullet$ to be ℓ-adic cohomology, the Betti–étale comparison isomorphism Theorem 1.79 establishes the Betti numbers conjecture.

Example 3.6 (Riemann hypothesis, $i = 0$)**.** If X is a smooth, geometrically irreducible, projective variety over $\mathbb{F}_q$, then $\Gamma(X, \mathcal{O}_X) = \mathbb{F}_q$. (One can check this by base-changing to the algebraic closure.) But now the q-power Frobenius acts trivially on $\mathbb{F}_q$, so we conclude $P_0(T)$ should equal
$$\det\left(1 - \mathrm{Frob}_q^* T; \mathrm{H}^0(X)\right) = 1 - T.$$

Thus, we see that computing $Z_X(T)$ is "as easy as" computing the characteristic polynomial of a Frobenius. At the very least, this perspective will have a lot of theoretical power. Let's start with the rationality conjecture.

Lemma 3.7. Let $F \subseteq F'$ be a field extension. Then $F[[T]] \cap F'(T) = F(T)$.

Proof. Of course, $F(T) \subseteq F[[T]] \cap F'(T)$, so the difficult inclusion is the reverse.

Suppose $f(T) = \sum_{i=0}^{\infty} a_i T^i$ is in $F[[T]]$. Then $f(T)$ lives in $F(T)$ if and only if one can find a polynomial $g(T) \in F[T]$ such that $f(T)g(T) \in F[T]$. In other words, we need to find finitely many coefficients $\{b_0, \dots, b_n\}$ such that the coefficient

$$\sum_{i=0}^{n} a_{N-i} b_i$$

of T^N in $f(T)\sum_{i=0}^{n} b_i T^i$ vanishes for N large enough. However, this is equivalent to the purely linear-algebraic condition that there is are $n, N > 0$ for which the subspace

$$V_{n,N} \coloneqq \operatorname{span}_F\{(a_i, \dots, a_{i+n}) : i > N\}$$

of F^{n+1} vanishes on a nontrivial functional. In other words, we are asking for some $n, N > 0$ for which the inclusion $V_{n,N} \subseteq F^{n+1}$ is proper.

However, linear algebraic conditions can be checked after field extensions, so we are basically done. Indeed, $f(T) \notin F(T)$ is equivalent to having $V_{n,N} = F^{n+1}$ for all $n, N > 0$. However, such an equality of vector spaces can be checked after base-changing to F', so this is equivalent to $V_{n,N} \otimes_F F' = (F')^{n+1}$ for all $n, N > 0$, which is equivalent to $f(T) \notin F'[[T]]$ by the previous paragraph! ∎

Theorem 3.8 (Rationality)**.** Suppose there is a Weil cohomology theory $\mathrm{H}^\bullet$ over $\mathbb{F}_q$ with coefficients in F. Then $Z_X(T) \in \mathbb{Q}(T)$ for any smooth projective variety X over $\mathbb{F}_q$.

Proof. By construction, $Z_X(T) \in \mathbb{Q}[[T]]$, and Proposition 3.4 shows that $Z_X(T) \in F(T)$, so we are done by Lemma 3.7. ∎

Remark 3.9. This proof says nothing about the rationality of the polynomials $\det(1 - f^*T; \mathrm{H}^i(X))$, which by the Riemann hypothesis (and Remark 3.5) are expected to be rational and with very controlled roots.

We now turn to the functional equation. This will come from "dualizing" $Z_X(T)$ via Poincaré duality. As such, we will want to understand the dual of $f^* \colon \mathrm{H}^\bullet(X) \to \mathrm{H}^\bullet(X)$ and how this affects characteristic polynomials.

Lemma 3.10. Let X be a scheme over $\mathbb{F}_q$ of finite type and equidimension d. Then the q-power Frobenius $\mathrm{Frob}_q \colon X \to X$ is finite of degree q^d.

Proof. Note that Frob_q is finite because it is affine, finite type, universally closed (hence proper), and quasifinite by definition. As for its degree, it is enough to compute the degree affine-locally, so we may assume that $X = \operatorname{Spec} A$. Now, Noether normalization provides some finite map $A \to \mathbb{A}^d_{\mathbb{F}_q}$. Because the Frobenius Frob_q will commute with any morphism over $\mathbb{F}_q$, we see that it is then enough to compute the degree of the Frobenius on $\mathbb{A}^d_{\mathbb{F}_q}$. But this can be done directly: writing $\mathbb{A}^d_{\mathbb{F}_q} = \operatorname{Spec} \mathbb{F}_q[X_1, \dots, X_d]$, the degree of the Frobenius equals the degree of the extension

$$[\mathbb{F}_q(X_1, \dots, X_d) : \mathbb{F}_q(X_1^q, \dots, X_d^q)],$$

which is simply q^d. ∎

Lemma 3.11. Let V and W be an n-dimensional vector spaces over F equipped with a perfect pairing $\langle\cdot,\cdot\rangle\colon V\times W\to F$. Given endomorphisms f and g of V and W, respectively, satisfying $\langle f(v),g(w)\rangle=\lambda\langle v,w\rangle$ for some nonzero λ, we have $\det(f;V)\det(g;W)=\lambda^n$ and

$$\det\left(1-\frac{f}{\lambda T};V\right)=\frac{(-1)^n\det(f;V)}{\lambda^nT^n}\det(1-gT;W).$$

Proof. Calculation of the characteristic polynomial is invariant under extending F, so we may assume that F is algebraically closed. Setting $n\coloneqq\dim V$, we may now choose a basis $\{v_1,\dots,v_n\}$ for V making the matrix representing f upper-triangular; then the dual basis $\{w_1,\dots,w_n\}$ of W will make g lower-triangular. Indeed, we have that $\langle f(v_i),w_j\rangle=0$ for $i<j$, which implies that $\langle v_i,g(w_j)\rangle=0$ for $i<j$ as well.

In fact, we can relate the diagonal entries of f and g: because f is upper-triangular while g is lower-triangular, we see that $\langle f(v_i),g(w_i)\rangle$ equals

$$\langle v_i,g(w_i)\rangle\langle f(v_i),w_i\rangle.$$

Thus, if the diagonal entries of f are $\{\lambda_1,\dots,\lambda_n\}$, then the diagonal entries of g are $\{\lambda/\lambda_1,\dots,\lambda/\lambda_n\}$. For example, multiplying together all these entries reveals $\det(f;V)\det(g;W)=\lambda^n$. Additionally, we see that $\det(1-gT;W)$ is the product

$$\prod_{i=1}^n\left(1-\frac{\lambda T}{\lambda_i}\right)=\frac{(-1)^n\lambda^nT^n}{\det(f;V)}\prod_{i=1}^n\left(1-\frac{\lambda_i}{\lambda T}\right),$$

and now the product on the right-hand side is $\det(1-f/\lambda T;V)$. The result follows after some rearrangement. ∎

Theorem 3.12 (Functional equation)**.** Suppose there is a Weil cohomology theory $\mathrm{H}^\bullet$ over $\mathbb{F}_q$ with coefficients in F. Choose a smooth, geometrically irreducible, projective variety X over a finite field $\mathbb{F}_q$ of d. Then there is a sign $\pm$ such that

$$Z_X\left(\frac{1}{q^dT}\right)=\pm q^{d\chi/2}T^\chi Z_X(T),$$

where $\chi\coloneqq\deg(\Delta\cdot\Delta)$.

Proof. Let $\mathrm{Frob}_q\colon X\to X$ denote the q-power Frobenius. The idea is to use Poincaré duality to relate $\mathrm{H}^i(X)$ with $\mathrm{H}^{2d-i}(X)$. By Lemma 3.10, the degree of Frob_q is q^d, so Lemma 1.119 shows that $(\mathrm{Frob}_q)_*\mathrm{Frob}_q^*=q^d$. Unwrapping the definitions, this is saying that

$$\int_X(\mathrm{Frob}_q^*\alpha\cup\mathrm{Frob}_q^*\alpha')=q^d\int_X(\alpha\cup\alpha')$$

for any $\alpha\in\mathrm{H}^i(X)$ and $\alpha'\in\mathrm{H}^{2d-i}(X)(d)$. By Poincaré duality, this trace pairing is perfect, so Lemma 3.11 implies

$$\det\left(1-\frac{\mathrm{Frob}_q^*}{q^dT};\mathrm{H}^i(X)\right)=\frac{(-1)^{\beta_i}\det(\mathrm{Frob}_q^*;\mathrm{H}^i(X))}{q^{d\beta_i}T^{\beta_i}}\det\left(1-\mathrm{Frob}_q^*T;\mathrm{H}^{2d-i}(X)(d)\right),$$

where $\beta_i\coloneqq\dim_F\mathrm{H}^i(X)$. Note that the twist (d) will not change the characteristic polynomial, so we may ignore it. Now, by taking the (signed) product over all i, we see that

$$Z_X\left(\frac{1}{q^dT}\right)=\left((-1)^\chi q^{d\chi}T^\chi\prod_{i=0}^{2d}\det\left(\mathrm{Frob}_q^*;\mathrm{H}^i(X)\right)^{(-1)^{i+1}}\right)Z_X(T),$$

where $\chi = \sum_{i=0}^{2d} (-1)^i \beta_i$, which is $\chi = \deg(\Delta \cdot \Delta)$ by Theorem 1.125. It remains to compute the product on the right-hand side. Another application of Lemma 3.11 shows

$$\det\left(\mathrm{Frob}_q^*; \mathrm{H}^i(X)\right) \det\left(\mathrm{Frob}_q^*; \mathrm{H}^{2d-i}(X)\right) = q^{d\beta_i},$$

so the square of the product is $q^{-d\chi}$. Thus, our product is $\pm q^{d\chi/2}$, and the proof is complete upon plugging this in. ∎

Remark 3.13. The proof explains that the sign $\pm$ is $(-1)^\chi \operatorname{sgn} \det\left(\mathrm{Frob}_q^*; \mathrm{H}^d(X)\right)$.

Example 3.14 (Riemann hypothesis, $i = 2d$)**.** The proof shows that $P_{2d}(T) = \det\left(1 - \mathrm{Frob}_q^* T; \mathrm{H}^{2d}(X)\right)$ is

$$-\frac{(-1)^{\beta_1} q^{d\beta_1} T^{\beta_1}}{\det\left(\mathrm{Frob}_q^*; \mathrm{H}^0(X)\right)} P_0\left(\frac{1}{q^d T}\right).$$

Because X is geometrically irreducible, $\Gamma(X, \mathcal{O}_X) = \mathbb{F}_q$, so $\mathrm{H}^0(X) = F$ by Example 1.106, so $\beta_1 = 1$. Now, Example 3.6 explains that $P_0(T) = 1 - T$, so $P_{2d}(T) = 1 - q^d T$ follows.

It remains to prove the Riemann hypothesis conjecture. This is much too difficult to be done here in any amount of detail, but we will mention how one might do this for abelian varieties A over $\mathbb{F}_q$. By Theorem 2.98, it really amounts to checking the required properties of the roots of

$$P_1(T) = \det\left(1 - \mathrm{Frob}_q^* T; \mathrm{H}^1(A)\right).$$

Approximately speaking, one wants to find a duality among the roots of $P_1(T)$: for example, if α is a root, then there should be another root $\overline{\alpha}$ with $\alpha\overline{\alpha} = q$. After some work with linear algebra, this eventually boils down to the following fact.

Proposition 3.15. Fix an abelian variety A over a finite field $\mathbb{F}_q$, and consider the induced Frobenius endomorphism Frob_q. Then

$$\mathrm{Frob}_q \circ \mathrm{Frob}_q^\dagger = [q]_A.$$

Proof. We refer to [Mil08, Lemma III.1.2]. ∎

3.1.2 The Sato–Tate Group

In this section, we will define the Sato–Tate group and state the Sato–Tate conjecture. Our exposition loosely follows [Sut19]. Fix an abelian variety A defined over a number field K, and choose a prime ℓ. We also let $\rho_\ell \colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$ denote the associated Galois representation.

Intuitively, the Sato–Tate conjecture asserts that the Frobenius elements $\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ equidistribute in $G_\ell(A)$ as $\mathfrak{p}$ varies over the maximal ideals of $\mathcal{O}_K$. This conjecture does not make sense verbatim, so we will have to work a bit to write down something formal. Consider the following points.

- To begin, we note that $\mathrm{Frob}_\mathfrak{p}$ only makes sense as a conjugacy class, and it only makes sense as a conjugacy class when ρ_ℓ vanishes on the relevant inertia subgroup of $\mathrm{Gal}(\overline{K}/K)$.

  Two remarks are thus in order. First, to vanish on the inertia subgroup, we must exclude a finite set of primes $\mathfrak{p}$ where A has bad reduction. (We are using the Néron–Ogg–Shafarevich criterion [BLR90, Theorem 5].) Second, we will simply regard $\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ as a conjugacy class as well. Thus, we really want to say that conjugacy classes equidistribute in a suitable space of conjugacy classes.

- It turns out that $\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ is not a totally random element of $G_\ell(A)$. Indeed, by Proposition 3.15, we see that the multiplier of $\mathrm{Frob}_\mathfrak{p}$ acting on $\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)$ equals $\mathrm{N}(\mathfrak{p})$. Thus, we would like to rescale $\mathrm{Frob}_\mathfrak{p}$ back down by $1/\sqrt{\mathrm{N}(\mathfrak{p})}$.

Once again, this requires two remarks. First, after rescaling, we will be working in the smaller subgroup

$$G_\ell^1(A) \coloneqq G_\ell(A) \cap \mathrm{Sp}(e_\varphi),$$

where φ is a choice of polarization on A. Second, the rescaling cannot happen in $\mathbb{Q}_\ell$ because $\mathbb{Q}_\ell$ does not have enough square roots. As such, we must choose an embedding $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$, allowing us to consider the elements $\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ in the complex Lie group $G_\ell^1(A)_\iota(\mathbb{C})$.[1]

- Another piece of structure to keep track of is that $\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ is semisimple (see Remark 2.132). This means that the subgroup topological generated by $\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ (which we now see has all eigenvalues equal to 1 after the normalization in the previous step) will be compact! A standard result in the structure theory of complex Lie groups is that they have maximal compact subgroups unique up to conjugacy, so one can find an element in our conjugacy class $\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ in any given maximal compact subgroup of $G_\ell^1(A)_\iota(\mathbb{C})$.

With the above preparations, we are now ready to state the Sato–Tate conjecture.

Definition 3.16 (Sato–Tate group)**.** Fix an abelian variety A defined over a number field K, and choose a prime ℓ and an embedding $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$. Then we define the *Sato–Tate group* $\mathrm{ST}(A)$ to be a maximal compact subgroup of the complex Lie group $G_\ell^1(A)_\iota$, where $G_\ell^1(A)$ is the subset of $G_\ell(A)$ with multiplier equal to 1.

Conjecture 3.17 (Sato–Tate)**.** Fix an abelian variety A defined over a number field K, and choose a prime ℓ and an embedding $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$. For each nonzero prime ideal $\mathfrak{p}$ of K such that A has good reduction at $\mathfrak{p}$, choose the conjugacy class $x_\mathfrak{p} \in \mathrm{Conj}(\mathrm{ST}(A))$ containing the conjugacy class $\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})$. Then the conjugacy classes $\{x_\mathfrak{p}\}$ equidistribute with respect to the pushforward of the Haar measure along $\mathrm{ST}(A) \to \mathrm{Conj}(\mathrm{ST}(A))$.

The relevance of the Sato–Tate conjecture for us is that it will let us numerically check that we have the correct ℓ-adic monodromy group; precisely how this is done will be explained in the subsequent subsections.

We will spend the rest of the present subsection making some remarks about Conjecture 3.17.

Remark 3.18. Not much is known about Conjecture 3.17. Roughly speaking, all known proofs prove something akin to modularity for not just the Galois representation attached to A but also its symmetric powers (and maybe more!).

- If A has complex multiplication, then this essentially follows from the Fundamental theorem of complex multiplication.
- For elliptic curves, the state of the art is [Bar+14; Bar+11], where the Sato–Tate conjecture is proven for elliptic curves over totally real and CM fields.
- These potential automorphy techniques were extended to some classes of abelian varieties by Johansson in [Joh17, Theorem 1].

One obnoxious defect of Conjecture 3.17 is that we must make choices regarding ℓ and ι. The choice ι is not so egregious because everything ought to descend to something algebraic, but it is quite unclear that $\mathrm{ST}(A)$ and even $G_\ell^1(A)$ does not depend crucially on ℓ. One expects $G_\ell(A)^\circ$ to not depend on ℓ by the Mumford–Tate conjecture (Conjecture 2.144). The relevant conjecture for the full group $G_\ell(A)$ is the Algebraic Sato–Tate conjecture [BK15, Conjecture 2.1].

[1] Another reason for passing to $\mathbb{C}$ is that groups in $\mathbb{C}$ have access to a good measure theory.

Conjecture 3.19 (Algebraic Sato–Tate)**.** Fix an abelian variety A defined over a number field K. Then there exists an algerbaic subgroup $\mathrm{AST}(A) \subseteq \mathrm{GL}_{2g}(\mathbb{Q})$ such that

$$\mathrm{AST}(A)_{\mathbb{Q}_\ell} = G^1_\ell(A)$$

for all primes ℓ.

This conjecture, being similar in spirit to the Mumford–Tate conjecture, has quite a bit known. For example, Banaszak and Kedlaya have shown this conjecture for products of abelian varieties of dimensions at most 3 [BK15, Theorem 6.11]. Roughly speaking, their proof boils down to the fact that one has $\mathrm{Hg}(A) = \mathrm{L}(A)^\circ$ in these small dimensions, which permits a direct computation of $\mathrm{AST}(A)$ along the lines of Proposition 2.159 (see Remark 2.160).

Remarkably, Farfán and Commelin have shown that the Algebraic Sato–Tate conjecture is implied by the Mumford–Tate conjecture in [CC22]. We will spend the rest of this subsection explaining their proof. Because we are interested in exhibiting a monodromy group related to all ℓ-adic groups, we are motivated to relate our conjectural AST to the motivic Galois group. Thus, we want to use a construction from our Tannakian formalism.

Notation 3.20. Fix a Tate triple $(\mathcal{C}, w, \mathrm{T})$ over a field F, and choose a fiber functor $\omega\colon \mathcal{C} \to \mathrm{Vec}_F$. If G_ω is the corresponding affine group, we let G^1_ω denote the kernel of the canonical map $G_\omega \twoheadrightarrow G_\omega(\mathrm{T})$. If $S \subseteq \mathcal{C}$ is a subset, we will write $G^1_\omega(S)$ for the image of G^1_ω in G_ω.

The following lemma aides our computation of G^1_ωs.

Lemma 3.21. Fix a Tate triple $(\mathcal{C}, w, \mathrm{T})$ over a field F, and choose a fiber functor $\omega\colon \mathcal{C} \to \mathrm{Vec}_F$. Given a subset $S \subseteq \mathcal{C}$, let m be the smallest positive integer such that $\mathrm{T}^{\otimes m} \in \langle S\rangle^\otimes$ and 0 if there is no such integer. Then $G^1_\omega(S)$ is the kernel of the canonical map $G_\omega(S) \twoheadrightarrow G^1_\omega\left(\mathrm{T}^{\otimes m}\right)$.

Proof. This is in [CC22, Section 2]. Let K denote the kernel of the canonical map $G^1_\omega(S) \twoheadrightarrow G^1_\omega\left(\mathrm{T}^{\otimes m}\right)$, which is faithfully flat by Proposition 1.152. Then K will fit into the following morphism

$$\begin{array}{ccccccccc} 1 & \longrightarrow & G^1_\omega & \longrightarrow & G_\omega & \twoheadrightarrow & G_\omega(\mathrm{T}) & \longrightarrow & 1 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 1 & \longrightarrow & K & \longrightarrow & G_\omega(S) & \twoheadrightarrow & G_\omega\left(\mathrm{T}^{\otimes m}\right) & \longrightarrow & 1 \end{array}$$

of short exact sequences. Here, the commutativity (and faithfully flatness) of the top-right square can be seen categorically, from which the left arrow is induced. Similarly, the construction of $G^1_\omega(S)$ permits the morphism

$$\begin{array}{ccccccccc} 1 & \longrightarrow & G^1_\omega & \longrightarrow & G_\omega & \longrightarrow & G_\omega(\mathrm{T}) & \longrightarrow & 1 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 1 & \longrightarrow & G^1_\omega(S) & \longrightarrow & G_\omega(S) & \longrightarrow & G_\omega(S)/G^1_\omega(S) & \longrightarrow & 1 \end{array}$$

of short exact sequences, where the right arrow is induced.

The main claim is that the surjection

$$G_\omega \twoheadrightarrow G_\omega(\mathrm{T}) \twoheadrightarrow G_\omega(S)/G^1_\omega(S)$$

in fact factors through $G_\omega(\mathrm{T}^{\otimes m})$. In fact, any $g \in G_\omega$ is the identity on $\langle \mathrm{T}^{\otimes m}\rangle^\otimes = \langle S\rangle^\otimes \cap \langle \mathrm{T}\rangle^\otimes$ if and only if $g|_{\langle S\rangle^\otimes}$ admits an extension to G^1_ω, which is equivalent to $g|_{\langle S\rangle^\otimes} \in G^1_\omega(S)$. Thus, we see that the induced surjection $G_\omega(\mathrm{T}^{\otimes m}) \twoheadrightarrow G_\omega(S)/G^1_\omega(S)$ is also an embedding and therefore an isomorphism. We conclude that the kernel K of the surjection $G_\omega(S) \twoheadrightarrow G_\omega(\mathrm{T}^{\otimes m})$ must in fact by $G^1_\omega(S)$. ∎

Example 3.22. We show that $G_\ell^1(A)$ is the kernel of the canonical map $G_\ell(A) \twoheadrightarrow G_\ell(\mathrm{T})$. (As in Remark 2.174, the polarization realizes T^{-1} as a quotient of $h^1(A) \otimes h^1(A)$.) Well, for any $g \in G_\ell(A)$, the argument of Remark 2.174 shows that g preserves a choice of Weil pairing

$$\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell) \otimes \mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell) \to \mathbb{Q}_\ell(-1),$$

so g has multiplier equal to 1 if and only if g fixes $\omega_\ell\left(\mathrm{T}^{-1}\right) = \mathbb{Q}_\ell(-1)$.

We are now ready for our main result.

Theorem 3.23 (Farfán–Commelin)**.** Fix an abelian variety A defined over a number field K. If A satisfies the Mumford–Tate conjecture (Conjecture 2.144) that $G_\ell(A)^\circ = \mathrm{MT}(A)$ for all primes ℓ, then A satisfies the Algebraic Sato–Tate conjecture (Conjecture 3.19). In fact, for all primes ℓ,

$$G^1_{\mathrm{mot}}(A)_{\mathbb{Q}_\ell} = G^1_\ell(A).$$

Proof. This is part of [CC22, Theorem]; its argument is analogous to Proposition 2.184. The main point is that Remark 2.180 induces a morphism

$$\begin{array}{ccccccccc} 1 & \longrightarrow & G^1_\ell(A) & \longrightarrow & G_\ell(A) & \longrightarrow & G_\ell(\mathrm{T}) & \longrightarrow & 1 \\ & & \vdots & & \downarrow & & \downarrow & & \\ 1 & \longrightarrow & G^1_{\mathrm{mot}}(A)_{\mathbb{Q}_\ell} & \longrightarrow & G_{\mathrm{mot}}(A)_{\mathbb{Q}_\ell} & \longrightarrow & G_{\mathrm{mot}}(\mathrm{T})_{\mathbb{Q}_\ell} & \longrightarrow & 1 \end{array}$$

of short exact sequences, where $G^1_{\mathrm{mot}}(A)$ is defined in the obvious way, and the horizontal maps are well-defined because $\mathrm{T} \in \langle h(A) \rangle^\otimes$ (see Remark 2.174). Indeed, the commutativity of the right square can be seen categorically on the level of fiber functors, and then the left map is induced.

We now complete the proof. Remark 1.151 explains that $G_{\mathrm{mot}}(\mathrm{T}) = \mathbb{G}_{m,\mathbb{Q}}$ and $G_\ell(\mathrm{T}) = \mathbb{G}_{m,\mathbb{Q}_\ell}$ simply by identifying the relevant catregories with GrVec. Thus, the right arrow is an isomorphism, so the middle arrow is an isomorphism if and only if the left one is. We are now done because the middle arrow being an isomorphism is equivalent to the Mumford–Tate conjecture for A by Proposition 2.184. ∎

Remark 3.24. In fact, the proof above shows that the canonical map $G^1_\ell(A) \to G^1_{\mathrm{mot}}(A)_{\mathbb{Q}_\ell}$ being an isomrphism is equivalent to the Mumford–Tate conjecture for A. In other words, a sufficiently precise version of the Algebraic Sato–Tate conjecture is equivalent to the Mumford–Tate conjecture.

3.1.3 Some Examples

In this subsection, we compute some basic Sato–Tate groups. The general outline is to compute the Hodge or Mumford–Tate groups first, check the Mumford–Tate conjecture to get G_ℓ°, and then compute some Galois action to get G_ℓ. We begin with some elliptic curves.

Example 3.25 (no complex multiplication)**.** Consider the elliptic curve $E\colon y^2 = x^3 + x + 1$ over $\mathbb{Q}$. One can compute that $\mathrm{End}_{\mathbb{C}}(E) = \mathbb{Z}$, so E does not have complex multiplication. Thus, $\mathrm{Hg}(E) \subseteq \mathrm{SL}_{2,\mathbb{Q}}$ needs to be a connected reductive subgroup which is not a torus (see Proposition 2.53); however, the only Lie subalgebras of $\mathfrak{sl}_2(\mathbb{C})$ are either commutative or all of $\mathfrak{sl}_2(\mathbb{C})$, so we conclude that $\mathrm{Hg}(E) = \mathrm{SL}_{2,\mathbb{Q}}$. Thus, $\mathrm{MT}(E) = \mathrm{GL}_{2,\mathbb{Q}}$.

The same computation (with Remark 2.130) allows us to conclude that $G_\ell(E) = \mathrm{GL}_{2,\mathbb{Q}_\ell}$ for all primes ℓ, thus proving the Mumford–Tate conjecture (Conjecture 2.144) in this case. We thus find $G^1_\ell(E) = \mathrm{SL}_{2,\mathbb{Q}_\ell}$, so upon choosing $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$, we see that $G^1_\ell(E)_\iota = \mathrm{SL}_{2,\mathbb{C}}$, so choosing a maximal compact subgroup finds $\mathrm{ST}(E) = \mathrm{SU}_2$.

Example 3.26 (complex multiplication)**.** Consider the elliptic curve $E\colon y^2 = x^3 + 1$ over $\mathbb{Q}(\zeta_3)$. Then we see that $\mathrm{End}_{\mathbb{C}}(E) = \mathbb{Z}[\zeta_3]$, where ζ_3 acts by $(x, y) \mapsto (\zeta_3 x, y)$, so E has complex multiplication. Thus, $\mathrm{Hg}(E) \subseteq \mathrm{SL}_{2,\mathbb{Q}(\zeta_3)}$ is a torus (by Proposition 2.53), but it cannot be trivial (by Corollary 2.42), so we conclude that $\mathrm{Hg}(E)$ is the diagonal torus of $\mathrm{SL}_{2,\mathbb{Q}(\zeta_3)}$.

For primes ℓ which split completely in $\mathbb{Q}(\zeta_3)$, the same computation (with Remark 2.130 and Corollary 2.129) where ℓ splits completely in $\mathbb{Q}_\ell$ reveals $G_\ell(E) = \mathbb{G}^2_{m,\mathbb{Q}_\ell}$ equals the diagonal torus in $\mathrm{GL}_{2,\mathbb{Q}(\zeta_3)}$, proving the Mumford–Tate conjecture (Conjecture 2.144) in this case. We thus find $G^1_\ell(E) \cong \mathbb{G}_{m,\mathbb{Q}_\ell}$, so upon choosing $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$, we see that $G^1_\ell(E) \cong \mathbb{G}_{m,\mathbb{Q}_\ell}$, so choosing a maximal compact subgroup finds $\mathrm{ST}(E) \cong \mathrm{U}_1$.

Example 3.27 (potential complex multiplication)**.** Consider the elliptic curve $E\colon y^2 = x^3+1$ but now over $\mathbb{Q}$. Example 3.26 computed that $\mathrm{MT}(E) \cong \mathbb{G}_{m,\mathbb{Q}}$ and $G_\ell(E)^\circ = \mathbb{G}_{m,\mathbb{Q}_\ell}$ (for primes $\ell \equiv 1 \pmod 3$). In this case, we see that there are endomorphisms not defined over $\mathbb{Q}$ and hence not fixed by $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, so $K_E^{\mathrm{conn}} \neq \mathbb{Q}$; instead, these endomorphisms are defined over $K_E^{\mathrm{conn}} = \mathbb{Q}(\zeta_3)$. We thus see that $G_\ell(E) \subseteq \mathrm{GL}_{2,\mathbb{Q}_\ell}$ normalizes its index-2 subgroup $G_\ell(E)^\circ$ (which is the diagonal torus), so $G_\ell(E)$ must be the diagonal torus together with the nontrivial Weyl element in $\mathrm{GL}_{2,\mathbb{Q}_\ell}$, which we write as $\mathbb{G}^2_{m,\mathbb{Q}_\ell} \rtimes S_2$. We thus find $G^1_\ell(E) \cong \mathbb{G}_{m,\mathbb{Q}_\ell} \rtimes S_2$, so $\mathrm{ST}(E) \cong \mathrm{U}_1 \rtimes S_2$.

Remark 3.28. In the above example, we appealed to the fact that the only elements normalizing the diagonal torus are the Weyl elements, which is a bit ad-hoc and will not work in higher dimensions. Roughly speaking, Proposition 2.159 provides the machine which works in higher dimensions, where we know that the Galois representation now factors through $\mathrm{Gal}(\mathbb{Q}(\zeta_3)/\mathbb{Q})$, and we are allowed to replace W with merely $W_1 \oplus W_2$, which can be computed to be generated by the endomorphisms and polarization.

We take a moment to remark that the above examples generalize to work with all elliptic curves, doing casework on having no complex multiplication, complex multiplication, and potential complex multiplication.

We now introduce the main example of the present thesis.

Proposition 3.29. Fix $\lambda \in \mathbb{C} \setminus \{0, 1\}$, and define A to be the Jacobian of the normalization of the proper curve C with affine chart $y^9 = x(x-1)(x-\lambda)$. If A does not have complex multiplication, then

$$\begin{cases} \mathrm{MT}(A)^{\mathrm{der}}_{\mathbb{C}} \cong \mathrm{SL}_2(\mathbb{C})^3 \\ Z(\mathrm{MT}(A))^\circ_{\mathbb{C}} \cong \mathbb{G}_m^4. \end{cases}$$

We use this to compute $\mathrm{ST}(A_K)$ if $\lambda \in K$ and K contains K_A^{conn}.

Proof. We proceed in steps.

1. To begin, we do some prelimarinary algebraic geometry, along the lines of [Moo10, Section 1]. The curve C comes equipped with a natural map $x\colon C \to \mathbb{P}^1$, with Galois with cyclic Galois group μ_9, where μ_9 acts on C by multiplication of the y-coordinate. As such, a computation with the Riemann–Hurwitz formula reveals that the genus is $g = 7$, so $\dim A = 7$. From here, we can find the differentials

$$\left\{ \frac{dx}{y^4}, \frac{dx}{y^5}, \frac{dx}{y^6}, \frac{dx}{y^7}, \frac{dx}{y^8}, \frac{x\,dx}{y^7}, \frac{x\,dx}{y^8} \right\}$$

are all holomorphic on C, and they are linearly independent, so we see that this is a basis of the space of differentials in $\mathrm{H}^0(C, \Omega^1_{C/\mathbb{C}}) = \mathrm{H}^0(A, \Omega^1_{A/\mathbb{C}})$. We remark that the above is also an eigenbasis for the induced μ_9-action on $\mathrm{H}^0(A, \Omega^1_{A/\mathbb{C}})$.

2. We decompose A into pieces. Note that C projects onto the elliptic curve $C_0\colon y^3 = x(x-1)(x-\lambda)$ via the map $(x,y) \mapsto \left(x, y^3\right)$, so C_0 is a factor of A. One can see that the basis of differentials of C_0 is given by dx/y^2, which pulls back to the differential dx/y^6 on A. In this way, we see that the quotient $A_1 \coloneqq A/C_0$ will have $\mathrm{H}^0(A_1, \Omega^1_{A_1/\mathbb{C}})$ have a basis given by

$$\left\{\frac{dx}{y^4}, \frac{dx}{y^5}, \frac{dx}{y^7}, \frac{dx}{y^8}, \frac{x\,dx}{y^7}, \frac{x\,dx}{y^8}\right\}.$$

Note that we do not yet know if A_1 is simple!

3. We compute some endomorphism algebras. Note C_0 has $\mu_3 \subseteq \mathrm{Aut}(C_0)$ where ζ_3 acts by multiplication on the y-coordinate, so C_0 has complex multiplication by $F_0 \coloneqq \mathbb{Q}(\zeta_3)$.

We conclude this step by showing that A_1 is simple. This will follow from the fact that A does not have complex multiplication. Note the μ_9-action on A fixes C_0 (we can be seen on the level of the Hodge structure), so it must also fix A_1, so we see $\mathbb{Q}(\zeta_9) \subseteq \mathrm{End}_{\mathbb{C}}(A_1)_{\mathbb{Q}}$. Thus, A_1 contains an isotypic component B^r (where B is simple) such that

$$\mathbb{Q}(\zeta_9) \subseteq \mathrm{End}_{\mathbb{C}}(B^r) = M_r\left(\mathrm{End}_{\mathbb{C}}(\mathrm{H}^1_{\mathrm{B}}(B,\mathbb{C}))\right).$$

As such, we set $D \coloneqq \mathrm{End}_{\mathbb{C}}(B)$ and $F \coloneqq Z(D)$ so that $d \coloneqq \sqrt{[D:F]}$ and $e \coloneqq [F:\mathbb{Q}]$ satisfy $6 \mid rde$ (because $\mathbb{Q}(\zeta_9)$ is contained in a maximal subfield of $M_r(D)$) and $r^2d^2e \leq 2\dim A_1 = 12$. If we had $r^2d^2e = 12$, then A_1 would have complex multiplication, which contradicts the fact that A does not have complex multiplication. Thus, we must instead have $rde = r^2d^2e = 6$, which implies that $r = d = 1$ and so $A_1 = B$ with $\mathrm{End}_{\mathbb{C}}(A_1)$ given exactly by $F_1 \coloneqq \mathbb{Q}(\zeta_9)$.

4. We compute some signatures. We begin with C_0. Letting $\tau_i \in \mathrm{Gal}(\mathbb{Q}(\zeta_3)/\mathbb{Q})$ be given by $\tau_i(\zeta_3) \coloneqq \zeta_3^i$ for $i \in \{1,2\}$, we see that the signature $\Phi_0\colon \mathrm{Gal}(\mathbb{Q}(\zeta_3)/\mathbb{Q}) \to \mathbb{Z}_{\geq 0}$ of E_0 is thus given by $\Phi_0(\tau_1) = 1$ and $\Phi_0(\tau_2) = 0$ because the second step provided an (eigen)basis of $\mathrm{H}^{10}(C_0) = \mathrm{H}^0(C_0, \Omega^1_{C_0/\mathbb{C}})$.

We next consider A_1. The second step provided a basis of $\mathrm{H}^{10}(A_1) = \mathrm{H}^0(A_1, \Omega^1_{A_1/\mathbb{C}})$. As such, we define $\sigma_i \in \mathrm{Gal}(\mathbb{Q}(\zeta_9)/\mathbb{Q})$ to be the automorphism given by $\sigma_i(\zeta_9) \coloneqq \zeta_9^i$ for $i \in \{1,2,4,5,7,8\}$, and we are able to compute that our signature $\Phi_1\colon \mathrm{Gal}(\mathbb{Q}(\zeta_9)/\mathbb{Q}) \to \mathbb{Z}_{\geq 0}$ is given by

$$\Phi(\sigma_i) = \begin{cases} 0 & \text{if } i \in \{7,8\}, \\ 1 & \text{if } i \in \{4,5\}, \\ 2 & \text{if } i \in \{1,2\}. \end{cases}$$

5. We compute $\mathrm{MT}(A)^{\mathrm{der}}$; note that this equals $\mathrm{Hg}(A)^{\mathrm{der}}$ by Lemma 1.41. By Lemma 1.56, we have an inclusion

$$\mathrm{Hg}(A) \to \mathrm{Hg}(C_0) \oplus \mathrm{Hg}(A_1)$$

which surjects onto each factor. Now, C_0 has complex multiplication, so $\mathrm{Hg}(C_0)$ is a torus by Proposition 2.53, so $\mathrm{Hg}(A)^{\mathrm{der}}$ has trivial projection onto $\mathrm{Hg}(C_0)$. We conclude that the above inclusion upgrades into an isomorphism $\mathrm{Hg}(A)^{\mathrm{der}} \to \mathrm{Hg}(A_1)^{\mathrm{der}}$.

To compute $\mathrm{Hg}(A_1)^{\mathrm{der}}$, we use Proposition 2.152 to see that this equals $\mathrm{L}(A_1)^{\mathrm{der}}$, so we complete this step by noting that $\mathrm{L}(A_1)^{\mathrm{der}}_{\mathbb{C}} \cong \mathrm{SL}_2(\mathbb{C})^3$ by the computation in Lemma 1.68.

6. We compute $Z(\mathrm{MT}(A))^\circ_{\mathbb{C}}$. We use Proposition 2.86 and in particular the discussion following the proof. Indeed, set $L \coloneqq \mathbb{Q}(\zeta_9)$, which we note is a Galois extension of $\mathbb{Q}$ containing F_0F_1. Then we note that $Z(\mathrm{MT}(A))^\circ \subseteq \mathrm{T}_F$, where $F \coloneqq F_0 \times F_1$ has $(\mathrm{T}_F)_L$ embedded into $\mathrm{GL}\left(\mathrm{H}^1_{\mathrm{B}}(A,L)\right)$ as a subtorus of the diagonal torus. Explicitly, we can choose an F-eigenbasis of $\mathrm{H}^1_{\mathrm{B}}(A,L) = \mathrm{H}^1_{\mathrm{B}}(C_0,L) \oplus \mathrm{H}^1_{\mathrm{B}}(A_1,L)$ as

$$\{u_1, u_2, v_1, v_1', v_2, v_2', v_4, v_4', v_5, v_5', v_7, v_7', v_8, v_8'\},$$

where the subscript partially indicates the F-eigenvalue. (For technical reasons, we will want to know that $\{v_i, v_i'\}$ is a dual basis for $\{v_{9-i}, v_{9-i}'\}$ according to the polarization.) Then we see that $(\mathrm{T}_F)_L \subseteq \mathrm{GL}\left(\mathrm{H}^1_{\mathrm{B}}(A, L)\right)$ embeds as

$$\{\mathrm{diag}(\mu_1, \mu_2, \lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_4, \lambda_4, \lambda_5, \lambda_5, \lambda_7, \lambda_7, \lambda_8, \lambda_8) : \mu_\bullet, \lambda_\bullet \in \mathbb{G}_{m,L}\}.$$

The discussion following Proposition 2.86 explains that equations cutting out $Z(\mathrm{MT}(A))^\circ_L \subseteq (\mathrm{T}_F)_L$ can be viewed as elements of the kernel of the map

$$\mathrm{X}^*\left((\mathrm{N}_{\Phi_0^*}, \mathrm{N}_{\Phi_1^*})\right) : \mathrm{X}^*(\mathrm{T}_F) \to \mathrm{X}^*(\mathrm{T}_L).$$

Using the established bases for these lattices, we see that our map can be written as the matrix

$$\begin{array}{c} \\ \sigma_1 \\ \sigma_2 \\ \sigma_4 \\ \sigma_5 \\ \sigma_7 \\ \sigma_8 \end{array} \begin{array}{c} \begin{array}{cc|cccccc} \mu_1 & \mu_2 & \lambda_1 & \lambda_2 & \lambda_4 & \lambda_5 & \lambda_7 & \lambda_8 \end{array} \\ \left[\begin{array}{cc|cccccc} 1 & 0 & 2 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 & 0 & 2 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 2 & 2 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 2 & 2 \end{array}\right] \end{array}.$$

Then one can compute a basis of the kernel of the matrix, which tells us that $Z(\mathrm{MT}(A))^\circ_L \subseteq (\mathrm{T}_F)_L$ is cut out by the equations

$$\begin{aligned} \lambda_1\lambda_8 &= \lambda_2\lambda_7, \\ \lambda_1\lambda_8 &= \lambda_4\lambda_5, \\ \mu_1\mu_2\lambda_7 &= \lambda_5\lambda_8, \\ \lambda_1\lambda_4\lambda_7 &= \lambda_2\lambda_5\lambda_8. \end{aligned}$$

Thus, we see that $Z(\mathrm{MT}(A))^\circ_{\mathbb{C}} \cong \mathbb{G}^4_{m,\mathbb{C}}$ with isomorphism given by the cocharacters $(\mu_1, \lambda_1, \lambda_4, \lambda_8)$.

7. We use the previous steps to compute $G^1_\ell(A)$ when ℓ splits completely in K_A^{conn}. Recall we notably know the Mumford–Tate conjecture that $G_\ell(A)^\circ = \mathrm{MT}(A)_{\mathbb{Q}_\ell}$ by Proposition 2.152. Thus, we choose ℓ to split completely in K_A^{conn} so that $\mathbb{Q}(\zeta_9) \subseteq \mathbb{Q}_\ell$, allowing us to engage in the diagonalization of the previous step. For example, the computation in Lemma 1.68 reveals that the isomorphism between $\mathrm{L}(A)^{\mathrm{der}}$ and SL_2^3 is defined over L (indeed, one merely needs to be able to take L-eigenspaces), so we find that

$$G_\ell(A)^{\mathrm{der}} = \left\{\mathrm{diag}\left(1_2, g_1, g_2, g_4, g_4^{-\intercal}, g_2^{-\intercal}, g_1^{-\intercal}\right) : g_1, g_2, g_3 \in \mathrm{SL}_{2,\mathbb{Q}_\ell}\right\}.$$

Continuing, we add in the equation $\det g = 1$ to the equations cutting out $Z(G_\ell(A_L))^\circ \subseteq (\mathrm{T}_F)_{\mathbb{Q}_\ell}$ given in the previous step. This reveals that $Z\left(G^1_\ell(A_L)\right)^\circ \subseteq (\mathrm{T}_F)_{\mathbb{Q}_\ell}$ is cut out by the equations

$$\begin{aligned} \mu_1\mu_2 &= 1, \\ \lambda_1\lambda_8 &= 1, \\ \lambda_2\lambda_7 &= 1, \\ \lambda_4\lambda_5 &= 1, \\ \lambda_2 &= \lambda_1\lambda_4. \end{aligned}$$

In particular, we see that $Z\left(G^1_\ell(A)\right)^\circ \cong \mathbb{G}^3_{m,\mathbb{Q}_\ell}$ given by the cocharacters $(\mu_1, \lambda_1, \lambda_4)$. In total, we find $G^1_\ell(A) \subseteq \mathrm{GL}_{14,\mathbb{Q}_\ell}$ equals

$$\left\{\mathrm{diag}\left(\mu_1, \mu_1^{-1}, \lambda_1 g_1, \lambda_1\lambda_4 g_2, \lambda_4 g_4, \lambda_4^{-1} g_4^{-\intercal}, \lambda_1^{-1}\lambda_4^{-1} g_2^{-\intercal}, \lambda_1^{-1} g_1^{-\intercal}\right) : \mu_\bullet, \lambda_\bullet \in \mathbb{G}_{m,\mathbb{Q}_\ell}, g_\bullet \in \mathrm{SL}_{2,\mathbb{Q}_\ell}\right\}.$$

8. At last, we compute $\mathrm{ST}(A_K)$ where K contains K_A^{conn}. By Theorem 3.23, we see that $\mathrm{ST}(A)$ does not depend on the choice ℓ, so we may as well choose ℓ to split completely in K_A^{conn}. Then we simply base-change the result of the previous step to $\mathbb{C}$, and then we may take maximal compact subgroups to see ST is

$$\left\{\mathrm{diag}\left(\mu_1, \mu_1^{-1}, \lambda_1 g_1, \lambda_1\lambda_4 g_2, \lambda_4 g_4, \lambda_4^{-1} g_4^{-\intercal}, \lambda_1^{-1}\lambda_4^{-1} g_2^{-\intercal}, \lambda_1^{-1} g_1^{-\intercal}\right) : \mu_\bullet, \lambda_\bullet \in \mathrm{U}_1, g_\bullet \in \mathrm{SU}_2\right\}.$$

(It is not too hard to see that the product of maximal compact subgroups continues to be a maximal compact subgroup.) This completes the computation. ∎

Remark 3.30. Note that $\mathrm{MT}(A) \neq \mathrm{L}(A)$ because the centers are different! This continues to be visible in the Sato–Tate group: the first four equations $\mu_1\mu_2 = \lambda_1\lambda_8 = \lambda_2\lambda_7 = \lambda_4\lambda_5 = 1$ can be explained by the polarization (see Lemma 2.65), but the last equation $\lambda_2 = \lambda_1\lambda_4$ corresponds to an exceptional Hodge class not generated by endomorphisms or the polarization.

Remark 3.31. Up to squaring, one can replace the equation $\mu_1\mu_2\lambda_7 = \lambda_5\lambda_8$ with the equation $\lambda_1\lambda_8 = \mu_1^2\mu_2^2$, thus making it clear that it arises from the polarization. Note this squaring is not too much of an issue because we had to take a determinant in Remark 2.78 anyway; in particular, by looking at the end result of the computation, we do see that $\mathrm{MT}(A)$ contains the diagonalizable group cut out by our equations where we have done the replacement with $\lambda_1\lambda_8 = \mu_1^2\mu_2^2$.

The hypothesis that A fails to have CM is necessary, as we will see in the following two examples.

Proposition 3.32. Define A to be the Jacobian of the proper curve C with affine chart $y^9 = x^3 - 1$. Then $\mathrm{MT}(A)_{\mathbb{C}}$ is a torus isomorphic to $\mathbb{G}_{m,\mathbb{C}}^4$. We use this to compute $\mathrm{ST}(A_K)$ where K contains K_A^{conn}.

Proof. We proceed in steps, following Proposition 3.29.

1. To begin, we once again note that C has genus 7, so A has dimension 7, and we have a basis of holomorphic differentials given by

$$\left\{\frac{dx}{y^4}, \frac{dx}{y^5}, \frac{dx}{y^6}, \frac{dx}{y^7}, \frac{dx}{y^8}, \frac{x\,dx}{y^7}, \frac{x\,dx}{y^8}\right\}.$$

   This time around, we see that $\mu_3 \times \mu_9$ acts on C by coordinate-wise multiplication on $(x, y) \in C$.

2. We decompose A into pieces.

   - Note C projects onto $C_0 \colon y^3 = x^3 - 1$ by $(x, y) \mapsto (x, y^3)$. (This is the quotient of C by $\mu_3 \times 1$.) We see that C_0 is an elliptic curve, and it has complex multiplicatino by μ_3; for example, μ_3 can act by multiplication on y. One can compute that C_0 has a basis of holomorphic differentials given by dx/y^2, which pulls back to the differential dx/y^6 on C.
   - Note C projects onto the proper curve C_1 with affine chart $y^9 = x^3(x-1)$ by $(x, y) \mapsto (x^3, xy)$, so A has $A_1 := \mathrm{Jac}\, C_1$ as a factor.[2] (This is the quotient of C by $\mu_3 \subseteq \mu_3 \times \mu_9$ embedded by $\zeta \mapsto (\zeta, \overline{\zeta})$.) One can compute that C_1 is genus 3 using the Riemann–Hurwitz formula, and then we can compute that it has a basis of holomorphic differentials given by $\{x^2\,dx/y^8, x^2\,dx/y^7, x\,dx/y^5\}$, which pull back to the differentials $\{dx/y^8, x\,dx/y^7, dx/y^5\}$ on C (up to a scalar).

     Note that C_1 has an action by μ_9 by multiplying on the y-coordinate, so $\mathbb{Q}(\zeta_9) \subseteq \mathrm{End}_{\mathbb{C}}(A_1)_{\mathbb{Q}}$. However, $\dim A_1 = 3$, so we see that A_1 has complex multiplication. We will check that A_1 is simple shortly.

[2] Technically, we should take normalizations everywhere. We will omit these normalizations.

- Note C projects onto the proper curve C_2 with affine chart $y^9 = x^6(x-1)$ by $(x,y) \mapsto \left(x^3, x^2y\right)$, so A has $A_2 := \operatorname{Jac} C_2$ as a factor. (This is the quotient of C by $\mu_3 \subseteq \mu_3 \times \mu_9$ embedded by $\zeta \mapsto (\zeta, \zeta)$.) One can compute that C_2 has genus 3 using the Riemann–Hurwitz formula, and then we can compute that it has a basis of holomorphic differentials given by $\left\{x^5\,dx/y^8, x^4\,dx/y^7, x^2\,dx/y^4\right\}$, which pull back to the differentials $\left\{x\,dx/y^8, dx/y^7, dx/y^4\right\}$ on C (up to a scalar).
  Note that C_2 has an action by μ_9 by multiplying on the y-coordinate, so $\mathbb{Q}(\zeta_9) \subseteq \operatorname{End}_{\mathbb{C}}(A_2)_{\mathbb{Q}}$. However, $\dim A_2 = 3$, so we see that A_2 has complex multiplication. We will check that A_2 is simple shortly.

We spend a moment checking that A is isogenous to $C_0 \times A_1 \times A_2$. The above computations have provided a map $C_0 \times A_1 \times A_2 \to A$, so it is enough to check that this is an isomorphism after base-changing to $\mathbb{C}$. The computations above have shown that this map provides an isomorphism

$$\mathrm{H}^0\left(A, \Omega^1_{A/\mathbb{C}}\right) \to \mathrm{H}^0\left(C_0, \Omega^1_{C_0/\mathbb{C}}\right) \oplus \mathrm{H}^0\left(A_1, \Omega^1_{A_1/\mathbb{C}}\right) \oplus \mathrm{H}^0\left(A_2, \Omega^1_{A_2/\mathbb{C}}\right).$$

(We take a moment to remark that the right-hand side is even a decomposition of $\mathrm{H}^0\left(A, \Omega^1_{A/\mathbb{C}}\right)$ into μ_3-eigenspaces!) This corresponds to an isomorphism on one piece of the Hodge structure, which we note upgrades to an isomorphism of Hodge structures because the relevant Hodge structures are concentrated in $(0,1)$ and $(1,0)$, which are complex conjugates. We conclude that A is isogenous to $C_0 \times A_1 \times A_2$ by Theorem 2.40.

3. We compute some signatures. For our notation, we let $F_0 := \mathbb{Q}(\zeta_3)$ have the embeddings $\{\tau_1, \tau_2\}$, where $\tau_\bullet \in \operatorname{Gal}(F_0/\mathbb{Q})$ sends $\zeta_3 \mapsto \zeta_3^\bullet$; similarly, we let $F_1 = F_2 := \mathbb{Q}(\zeta_9)$ have the embeddings $\{\sigma_1, \sigma_2, \sigma_4, \sigma_5, \sigma_7, \sigma_8\}$ where $\sigma_\bullet \in \operatorname{Gal}(\mathbb{Q}(\zeta_9)/\mathbb{Q})$ sends $\zeta_9 \mapsto \zeta_9^\bullet$. Here are our signatures.

   - On C_0, we see that H^{10} is spanned by dx/y^2, so with μ_3 acting on y, we get the signature $\Phi_0(\tau_1) = 1$ and $\Phi_0(\tau_2) = 0$.
   - On C_1, we see that H^{10} has basis given by $\left\{x^2\,dx/y^8, x^2\,dx/y^7, x\,dx/y^5\right\}$. Thus, with μ_9 acting on y, we get the signature
   $$\Phi_1(\sigma_i) = \begin{cases} 0 & \text{if } i \in \{5,7,8\}, \\ 1 & \text{if } i \in \{1,2,4\}. \end{cases}$$
   One can check that Φ_1 satisfies the check of Remark 2.57, proving that A_1 is simple.
   - On C_2, we see that H^{10} has basis given by $\left\{x\,dx/y^8, dx/y^7, dx/y^4\right\}$. Thus, with μ_9 acting on y, we get the signature
   $$\Phi_2(\sigma_i) = \begin{cases} 0 & \text{if } i \in \{4,7,8\}, \\ 1 & \text{if } i \in \{1,2,5\}. \end{cases}$$
   One can check that Φ_2 satisfies the check of Remark 2.57, proving that A_1 is simple.

The above computation allows us to conclude that we have decomposed A into simple abelian varieties with complex multiplication.

4. We compute $\mathrm{MT}(A)_{\mathbb{C}}$. Because A has complex multiplication, we see that $\mathrm{MT}(A)$ is a torus by Proposition 2.53 embedded in T_F, where $F := F_0 \times F_1 \times F_2$. As such, we may use Proposition 2.86 and the surrounding discussion following the proof to compute equations cutting out $\mathrm{MT}(A) \subseteq \mathrm{T}_F$. In particular, set $L := \mathbb{Q}(\zeta_9)$, which we note is a Galois extension of $\mathbb{Q}$ containing $F_0F_1F_2$. Then we note that $\mathrm{H}^1_{\mathrm{B}}(A, L) = \mathrm{H}^1_{\mathrm{B}}(C_0, L) \oplus \mathrm{H}^1_{\mathrm{B}}(A_1, L) \oplus \mathrm{H}^1_{\mathrm{B}}(A_2, L)$ can be given a basis

$$\{u_1, u_2, v_1, v_2, v_4, v_5, v_7, v_8, w_1, w_2, w_4, w_5, w_7, w_8\},$$

where the subscript partially indicates the F-eigenvalue. Then we see that $(\mathrm{T}_F)_L \subseteq \mathrm{GL}\left(\mathrm{H}^1_{\mathrm{B}}(A, L)\right)$ embeds as

$$\{\operatorname{diag}(\mu_1, \mu_2, \lambda_1, \lambda_2, \lambda_4, \lambda_5, \lambda_7, \lambda_8, \kappa_1, \kappa_2, \kappa_4, \kappa_5, \kappa_7, \kappa_8) : \mu_\bullet, \lambda_\bullet, \kappa_\bullet \in \mathbb{G}_{m,L}\}.$$

The discussion following Proposition 2.86 explains that equations cutting out $Z(\mathrm{MT}(A))^\circ_L \subseteq (\mathrm{T}_F)_L$ can be viewed as elements of the kernel of the map

$$\mathrm{X}^*\left((\mathrm{N}_{\Phi_0^*}, \mathrm{N}_{\Phi_1^*}, \mathrm{N}_{\Phi_2^*})\right) : \mathrm{X}^*(\mathrm{T}_F) \to \mathrm{X}^*(T_L).$$

Using the established bases for these lattices, we see that our map can be written as the matrix

$$\begin{array}{c|cc|cccccc|cccccc} & \mu_1 & \mu_2 & \lambda_1 & \lambda_2 & \lambda_4 & \lambda_5 & \lambda_7 & \lambda_8 & \kappa_1 & \kappa_2 & \kappa_4 & \kappa_5 & \kappa_7 & \kappa_8 \\ \sigma_1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ \sigma_2 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ \sigma_4 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ \sigma_5 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ \sigma_7 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ \sigma_8 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}.$$

Then one can compute a basis of the kernel of the matrix, which tells us that $\mathrm{MT}(A)_L \subseteq (\mathrm{T}_F)_L$ is cut out by the following equations. To begin, it turns out that $(A_1)_L$ and $(A_2)_L$ are isogenous, which we can see from the six equations

$$\begin{aligned} \lambda_1 &= \kappa_5, \\ \lambda_2 &= \kappa_1, \\ \lambda_4 &= \kappa_2, \\ \lambda_5 &= \kappa_7, \\ \lambda_7 &= \kappa_8, \\ \lambda_8 &= \kappa_4. \end{aligned}$$

(Namely, these equations imply an isomorphism of $\mathrm{MT}(A)$-representations $\mathrm{H}^1_{\mathrm{B}}(A_1, L) \cong \mathrm{H}^1_{\mathrm{B}}(A_2, L)$ and hence an isomorphism of Hodge structures, which gives the isogeny by Theorem 2.40.) Then there are the equations given by the polarization (via Lemma 2.65)

$$\begin{aligned} \mu_1\mu_2 &= \kappa_1\kappa_8, \\ \kappa_1\kappa_8 &= \kappa_2\kappa_7, \\ \kappa_1\kappa_8 &= \kappa_4\kappa_5. \end{aligned}$$

Lastly, there is the exceptional equation

$$\mu_1\kappa_7 = \kappa_5\kappa_8.$$

In total, we find that $\mathrm{MT}(A)_L$ is a torus isomorphic to $\mathbb{G}^4_{m,L}$ via the cocharacters $(\kappa_1, \kappa_2, \kappa_4, \kappa_8)$.

5. We use the previous step to compute $G^1_\ell(A_K)$ when ℓ splits completely in $K \coloneqq K_A^{\mathrm{conn}}$. Recall that we know the Mumford–Tate conjecture that $G_\ell(A)^\circ = \mathrm{MT}(A)_{G_\ell}$ by Remark 2.146. Thus, we choose ℓ to split completely in K_A^{conn} so that $L \subseteq \mathbb{Q}_\ell$, allowing us to engage in the diagonalization of the previous step. Now, to compute $G^1_\ell(A_K)$ from $G_\ell(A_K)$, we simply need to add in the equation that the multipler is 1. This reveals that $G^1_\ell(A_{K_A^{\mathrm{conn}}}) \subseteq (\mathrm{T}_F)_{\mathbb{Q}_\ell}$ is cut out by the following equations. As before, we have the six equations

$$\begin{aligned} \lambda_1 &= \kappa_5, \\ \lambda_2 &= \kappa_1, \\ \lambda_4 &= \kappa_2, \\ \lambda_5 &= \kappa_7, \\ \lambda_7 &= \kappa_8, \\ \lambda_8 &= \kappa_4 \end{aligned}$$

given by the isogeny $(A_1)_L \sim (A_2)_L$, and we have the equations given by the polarization

$$\begin{aligned}\mu_1\mu_2 &= 1,\\ \kappa_1\kappa_8 &= 1,\\ \kappa_2\kappa_7 &= 1,\\ \kappa_4\kappa_5 &= 1.\end{aligned}$$

Lastly, there is still the exceptional equation

$$\mu_1\kappa_7 = \kappa_5\kappa_8.$$

In total, we find that $G_\ell^1(A)$ is a torus isomorphic to $\mathbb{G}_{m,L}^3$ via the cocharacters $(\kappa_1, \kappa_2, \kappa_4)$. In total, we see $G_\ell^1(A_K)^\circ \subseteq \mathrm{GL}_{14}$ is

$$\left\{\operatorname{diag}\left(\frac{\kappa_2}{\kappa_1\kappa_4}, \frac{\kappa_1\kappa_4}{\kappa_2}, \kappa_4^{-1}, \kappa_1, \kappa_2, \kappa_2^{-1}, \kappa_1^{-1}, \kappa_4, \kappa_1, \kappa_2, \kappa_4, \kappa_4^{-1}, \kappa_2^{-1}, \kappa_1^{-1}\right) : \kappa_\bullet \in \mathbb{G}_{m,\mathbb{Q}_\ell}\right\}.$$

6. At last, we compute $\mathrm{ST}(A_K)$ where K contains K_A^{conn}. By Theorem 3.23, we see that ST does not depend on the choice of ℓ, so we may as well choose ℓ to split completely in K_A^{conn}. Then we may simply base-change the result of the previous step to $\mathbb{C}$, and then we may take maximal compact subgroups to see ST is

$$\left\{\operatorname{diag}\left(\frac{\kappa_2}{\kappa_1\kappa_4}, \frac{\kappa_1\kappa_4}{\kappa_2}, \kappa_4^{-1}, \kappa_1, \kappa_2, \kappa_2^{-1}, \kappa_1^{-1}, \kappa_4, \kappa_1, \kappa_2, \kappa_4, \kappa_4^{-1}, \kappa_2^{-1}, \kappa_1^{-1}\right) : \kappa_\bullet \in \mathrm{U}_1\right\}.$$

Once again, we remark that the product of maximal compact subgroups continues to be maximal compact. ■

Proposition 3.33. Define A to be the Jacobian of the proper curve C with affine chart $y^9 = x\left(x^2+1\right)$. Then $\mathrm{MT}(A)_{\mathbb{C}}$ is a torus isomorphic to $\mathbb{G}_{m,\mathbb{C}}^4$. We use this to compute $\mathrm{ST}(A_K)$ where K contains K_A^{conn}.

Proof. This argument is essentially the same as Proposition 3.32, so we will be a bit briefer.

1. Once again, we see that C has genus 7, so A has dimension 7, and we have a basis of holomorphic differentials given by
$$\left\{\frac{dx}{y^4}, \frac{dx}{y^5}, \frac{dx}{y^6}, \frac{dx}{y^7}, \frac{dx}{y^8}, \frac{x\,dx}{y^7}, \frac{x\,dx}{y^8}\right\}.$$
This time around, we see that μ_{18} acts on C by $\zeta_{18}\cdot(x,y) = (-x, -\zeta_9 y)$.

2. We decompose A into pieces.

   - As usual, C_0 projects onto $y^3 = x\left(x^2+1\right)$ by $(x,y) \mapsto \left(x, y^3\right)$. (This is the quotient of C by μ_3.) The Riemann–Hurwitz formula yields that C_0 is an elliptic curve with complex multiplication by μ_3 acting on the y-coordinate. We see that C_0 has a basis of holomorphic differentials given by dx/y^2, which pulls back to dx/y^6 ln C.
   - Now, C projects onto the proper curve C_1 with affine chart $y^9 = x^5(x+1)$ by $(x,y) \mapsto \left(x^2, xy\right)$, so A has $A_1 \coloneqq \operatorname{Jac} C_1$ as a factor. (This is the quotient of C by μ_2.) The Riemann–Hurwitz formula implies that C_1 has genus 3, and then we can compute that it has a basis of holomorphic differentials given by $\left\{x^4\,dx/y^8, x^3\,dx/y^7, x^2\,dx/y^5\right\}$, which pulls back to $\left\{x\,dx/y^8, dx/y^7, dx/y^5\right\}$ on C (up to scalar).

     Note that C_1 has an action by μ_9 acting on the y-coordinate, so $\mathbb{Q}(\zeta_9) \subseteq \mathrm{End}_{\mathbb{C}}(A_1)_{\mathbb{Q}}$. We will check in the next step that A_1 is simple by computing its signature and applying Remark 2.57.

We can see on the level of differentials that the induced map $C_0 \times A_1 \to A$ is injective, so we let A_2 be the cokernel. In terms of Hodge structures, we can see from the computation that

$$\mathrm{H}^1_{\mathrm{B}}(A, \mathbb{Q}) = \mathrm{H}^1_{\mathrm{B}}(C_0, \mathbb{Q}) \oplus \mathrm{H}^1_{\mathrm{B}}(A_1, \mathbb{Q}) \oplus \mathrm{H}^1_{\mathrm{B}}(A_2, \mathbb{Q})$$

is a decomposition of μ_{18}-representations because the left two spaces on the right-hand side are stable under the μ_{18}-action. We conclude that $\mathbb{Q}(\zeta_9) \subseteq \mathrm{End}_{\mathbb{C}}(A_2)_{\mathbb{Q}}$ as well.

3. We compute some signatures. As before, we let $F_0 \coloneqq \mathbb{Q}(\zeta_3)$ have $\{\tau_1, \tau_2\} = \mathrm{Gal}(\mathbb{Q}(F_0)/\mathbb{Q})$ where $\tau_\bullet \colon \zeta_3 \mapsto \zeta_3^\bullet$, and we let $F_1 = F_2 \coloneqq \mathbb{Q}(\zeta_9)$ have $\{\sigma_1, \dots, \sigma_8\} = \mathrm{Gal}(\mathbb{Q}(\zeta_9)/\mathbb{Q})$ has $\sigma_\bullet \colon \zeta_9 \mapsto \zeta_9^\bullet$.

   - On C_0, we look at the μ_9-eigenbasis of H^{10} to conclude that our signature has $\Phi_0(\tau_1) = 1$ and $\Phi_0(\tau_2)$.
   - On C_1, we look at the μ_9-eigenbasis of H^{10} to conclude that our signature is

$$\Phi_1(\sigma_i) = \begin{cases} 0 & \text{if } i \in \{5, 7, 8\}, \\ 1 & \text{if } i \in \{1, 2, 4\}. \end{cases}$$

     One can check that Φ_1 satisfies the check of Remark 2.57, proving that A_1 is simple.
   - On A_2, we take the remaining differentials from A to find that our signature is

$$\Phi_2(\sigma_i) = \begin{cases} 0 & \text{if } i \in \{4, 7, 8\}, \\ 1 & \text{if } i \in \{1, 2, 5\}. \end{cases}$$

     Again, one checks that Φ_2 satisfies the check of Remark 2.57

4. At this point, we recognize that our signatures are the same as in Proposition 3.32 up to swapping Φ_1 and Φ_2. Thus, up to some reordering of letters, the exact same computation goes through. Let's provide the result.

   To be explicit, we give $\mathrm{H}^1_{\mathrm{B}}(A, L) = \mathrm{H}^1_{\mathrm{B}}(C_0, L) \oplus \mathrm{H}^1_{\mathrm{B}}(A_1, L) \oplus \mathrm{H}^1_{\mathrm{B}}(A_2, L)$ a basis

$$\{u_1, u_2, v_1, v_2, v_4, v_5, v_7, v_8, w_1, w_2, w_4, w_5, w_7, w_8\},$$

   where the subscript partially indicates the F-eigenvalue, where $F \coloneqq F_0 \times F_1 \times F_2$. Then we set $L \coloneqq \mathbb{Q}(\zeta_9)$, and we see $(\mathrm{T}_F)_L \subseteq \mathrm{GL}\left(\mathrm{H}^1_{\mathrm{B}}(A, L)\right)$ embeds as

$$\{\mathrm{diag}(\mu_1, \mu_2, \lambda_1, \lambda_2, \lambda_4, \lambda_5, \lambda_7, \lambda_8, \kappa_1, \kappa_2, \kappa_4, \kappa_5, \kappa_7, \kappa_8) : \mu_\bullet, \kappa_\bullet, \lambda_\bullet \in \mathbb{G}_{m,L}\}.$$

   With this choice of lettering, the equations that end up cutting out $\mathrm{MT}(A)_L \subseteq (\mathrm{T}_F)_L$ are exactly the same, so $\mathrm{MT}(A)_L \cong \mathbb{G}^4_{m,L}$ via the cocharacters $(\kappa_1, \kappa_2, \kappa_4, \kappa_8)$.

   One is now able to compute $G^1_\ell(A)$ in the case where ℓ splits completely in $K \coloneqq K_A^{\mathrm{conn}}$. One finds the exact same equations via the same computation, so we find $G^1_\ell(A_K) \subseteq \mathrm{GL}_{14}$ is given by

$$\left\{\mathrm{diag}\left(\frac{\kappa_2}{\kappa_1\kappa_4}, \frac{\kappa_1\kappa_4}{\kappa_2}, \kappa_4^{-1}, \kappa_1, \kappa_2, \kappa_2^{-1}, \kappa_1^{-1}, \kappa_4, \kappa_1, \kappa_2, \kappa_4, \kappa_4^{-1}, \kappa_2^{-1}, \kappa_1^{-1}\right) : \kappa_\bullet \in \mathbb{G}_{m,\mathbb{Q}_\ell}\right\}.$$

   Base-changing to $\mathbb{C}$ and taking a maximal compact subgroup, we find $\mathrm{ST}(A_K)$ is

$$\left\{\mathrm{diag}\left(\frac{\kappa_2}{\kappa_1\kappa_4}, \frac{\kappa_1\kappa_4}{\kappa_2}, \kappa_4^{-1}, \kappa_1, \kappa_2, \kappa_2^{-1}, \kappa_1^{-1}, \kappa_4, \kappa_1, \kappa_2, \kappa_4, \kappa_4^{-1}, \kappa_2^{-1}, \kappa_1^{-1}\right) : \kappa_\bullet \in \mathrm{U}_1\right\},$$

   as required. ∎

3.1.4 Moment Statistics

In this subsection, we explain how to numerically verify the Sato–Tate conjecture (Conjecture 3.19). Fix an abelian variety A of dimension g defined over a number field K, and choose a prime ℓ and embedding $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$; for example, this allows us to define the usual ℓ-adic representation $\rho_\ell\colon \operatorname{Gal}(\overline{K}/K) \to \operatorname{GL}\left(\mathrm{H}^1_{\text{ét}}(A_{\overline{K}}, \mathbb{Q}_\ell)\right)$.

The main idea is that the map sending $g \in \operatorname{ST}(A)$ to the characteristic polynomial of $g \in \operatorname{GL}_{2g}(\mathbb{C})$ is well-defined up to conjugacy classes, so it defines a (continuous) map $\operatorname{Conj}(\operatorname{ST}(A)) \to \mathbb{C}^{2g+1}$, where $\mathbb{C}^{2g+1}$ simply lists out the coefficients of the characteristic polynomial. In this way, we can push the Haar measure on $\operatorname{ST}(A)$ all the way to $\mathbb{C}^{2g+1}$ to compute what the distribution of the characteristic polynomial will be.

Of course, in practice, it may be difficult to compute the characteristic polynomial of

$$\left[\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right] \in \operatorname{Conj}(\operatorname{ST}(A))$$

for some prime $\mathfrak{p}$ of K such that A has good reduction at $\mathfrak{p}$. For our application, we will only be interested in superelliptic curves, for which this can be computed in SageMath [Aru+19]. To help out the computation a bit more, we make two quick remarks.

Remark 3.34. Let $P(T)$ be the characteristic polynomial of $\mathrm{Frob}_\mathfrak{p}$ acting on $\mathrm{H}^1_{\text{ét}}(A_{\overline{\mathbb{F}_\mathfrak{p}}}, \mathbb{Q}_\ell)$. Then we remark that $P(1)$ has a geometric interpretation as $\#A(\mathbb{F}_\mathfrak{p})$.

Remark 3.35. It suffices to only consider primes $\mathfrak{p}$ which are totally split in K because such primes have density 1. This is helpful because primes that split $\mathfrak{p}$ completely have residue fields isomorphic to $\mathbb{F}_p$ where $p \in \mathbb{Z}$ is the prime sitting below $\mathfrak{p}$, so we are frequently able to reduce the computation to something only involving integral coefficients.

As before, let's begin with some elliptic curve examples. Here, we note that the characteristic polynomial of $\frac{1}{\mathrm{N}(\mathfrak{p})}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})$ will have degree 2, with leading coefficient 1, and the condition on the multiplier (from Proposition 3.15) implies that the constant coefficient is 1. Thus, we see that the only interesting coefficient of the characteristic polynomial is given by the trace.

Lemma 3.36. The map $\operatorname{tr}\colon \operatorname{Conj}(\mathrm{SU}_2) \to [-2,2]$ is a homeomorphism, and the pushforward of the normalized Haar measure of SU_2 onto $\operatorname{Conj}(\mathrm{SU}_2) = [-2,2]$ is given by the semicircle measure $\frac{1}{2\pi}\sqrt{4-t^2}\,dt$.

Proof. We show the claims separately.

1. We show that $\operatorname{tr}\colon \operatorname{Conj}(\mathrm{SU}_2) \to [-2,2]$ is a well-defined homeomorphism. Note that $\operatorname{tr}\colon \operatorname{Conj}(\mathrm{SU}_2) \to \mathbb{C}$ is continuous, and all spaces in sight are compact and Hausdorff, so it is enough to check that tr is a bijection.

   A priori, tr is only defined as a map $\operatorname{tr}\colon \operatorname{Conj}(\mathrm{SU}_2) \to \mathbb{C}$. To begin, we note that any element of SU_2 is diagonalizable by a unitary matrix, and the corresponding diagonal matrix must then look like $\operatorname{diag}(\lambda, \overline{\lambda})$ where $|\lambda|^2 = 1$. By writing $\lambda = e^{i\theta}$, we see that the trace of this element is $2\cos\theta$, so we see that $\operatorname{tr}\colon \operatorname{Conj}(\mathrm{SU}_2) \to [-2,2]$ is a well-defined surjection.

   It remains to check that tr is injective. Because each conjugacy class is represented by a diagonal matrix, it is enough to check that $g_1 \coloneqq \operatorname{diag}(\lambda_1, \overline{\lambda}_1)$ and $g_2 \coloneqq \operatorname{diag}(\lambda_2, \overline{\lambda}_2)$ have $\operatorname{tr} g_1 = \operatorname{tr} g_2$ only if g_1 and g_2 are conjugate. Well, write $\lambda_\bullet = e^{i\theta_\bullet}$, and then we see that

   $$2\cos\theta_1 = 2\cos\theta_2,$$

   which implies that $\{\pm\theta_1\} = \{\pm\theta_2\}$, so $\{\lambda_1, \overline{\lambda}_1\} = \{\lambda_2, \overline{\lambda}_2\}$. We now do casework: if $\lambda_1 = \lambda_2$, then we see that $g_1 = g_2$ on the nose; otherwise, $\lambda_1 = \overline{\lambda}_2$, and we see that

   $$\begin{bmatrix} & 1 \\ -1 & \end{bmatrix}\begin{bmatrix} \lambda_1 & \\ & \overline{\lambda}_1 \end{bmatrix}\begin{bmatrix} & -1 \\ 1 & \end{bmatrix} = \begin{bmatrix} \lambda_2 & \\ & \overline{\lambda}_2 \end{bmatrix},$$

so g_1 is conjugate to g_2.

2. We now compute the required measures. A linear algebra argument with the condition $gg^\dagger = 1_2$ shows that any element of SU_2 can be written uniquely in the form

$$\begin{bmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{bmatrix}$$

where $|\alpha|^2 + |\beta|^2 = 1$. In this way, we see that SU_2 is isomorphic to the unit group of the quaternions $\mathbb{H}$, so SU_2 is diffeomorphic to S^3 and inherits a Haar measure by pullback. Explicitly, one finds that SU_2 inherits an action on S^3 by rotations, so the Lebesgue measure on S^3 is invariant under the group. Note that we have yet to normalize the Haar mesure on SU_2.

We would now like to compute the volume of SU_2 with given trace t. Writing $\alpha = a+bi$ and $\beta = c+di$, we see that we are forcing $a = \frac{1}{2}t$, which then requires the remaining coordinates to live in a sphere of radius $\sqrt{1-\frac{1}{4}t^2}$. Thus, we see that our normalized Haar measure is

$$\frac{\sqrt{1-\frac{1}{4}t^2}\,dt}{\int_{-2}^{2}\sqrt{1-\frac{1}{4}t^2}\,dt}.$$

A quick substitution with $t = 2\cos\theta$ in the bottom integral reveals that it equals π, whereupon we find that the desired measure is $\frac{1}{2\pi}\sqrt{4-t^2}\,dt$ after some rearranging. ∎

Remark 3.37. In the sequel, it is occasionally more convenient to identify $\mathrm{Con}(\mathrm{SU}_2)$ with the collection of diagonal matrices $\mathrm{diag}\left(e^{i\theta}, e^{-i\theta}\right)$ where $\theta \in [0,\pi)$. Then we see that the trace is $2\cos\theta$, so we produce a measure of $\frac{2}{\pi}\sin^2\theta\,d\theta$ on $[0,\pi)$.

Example 3.38 (no complex multiplication)**.** We continue with the elliptic curve $E\colon y^2 = x^3 + x + 1$ over $\mathbb{Q}$ studied in Example 3.25. Then we recall that $\mathrm{ST}(E) = \mathrm{SU}_2$, so we may use the computation of Lemma 3.36 to see that the Sato–Tate conjecture (Conjecture 3.17) implies that the values

$$\left\{\mathrm{tr}\,\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right\}_{\mathfrak{p}\text{ prime}}$$

equidistribute according to the semicircle measure $\frac{1}{2\pi}\sqrt{4-t^2}\,dt$ on $[-2,2]$.

Example 3.39 (complex multiplication)**.** We continue with the elliptic curve $E\colon y^2 = x^3 + 1$ over $\mathbb{Q}(\zeta_3)$ studied in Example 3.26. Then we recall that $\mathrm{ST}(E) \cong \mathrm{U}_1$ embedded as $z \mapsto \mathrm{diag}(z,\overline{z})$. We may write U_1 as $\mathrm{U}_1 = \left\{e^{i\theta} : \theta \in [0,2\pi)\right\}$, so we can equip this group with the normalized Haar measure $\frac{1}{2\pi}\,d\theta$. (The map $e^{i\theta} \mapsto \theta$ is a homeomorphism away from a set of measure 0.) Noting the trace of $\mathrm{diag}\left(e^{i\theta}, e^{-i\theta}\right)$ is $2\cos\theta$, we see the Sato–Tate conjecture (Conjecture 3.17) implies that the values

$$\left\{\mathrm{tr}\,\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right\}_{\mathfrak{p}\text{ prime}}$$

equidistribute according to the measure $\frac{1}{\pi}\cdot\frac{1}{\sqrt{4-t^2}}\,dt$ on $[-2,2]$.

Example 3.40 (potential complex multiplication)**.** We continue with the elliptic curve $E\colon y^2 = x^3 + 1$ over $\mathbb{Q}(\zeta_3)$ studied in Example 3.27. Then we recall that $\mathrm{ST}(E) \cong \mathrm{U}_1 \rtimes S_2$, where $\mathrm{U}_1 \subseteq \mathrm{GL}_{2,\mathbb{C}}$ is embedded as $z \mapsto \mathrm{diag}(z, \overline{z})$, and $S_2 = \{1, w\}$ acts by switching the coordinates. Again, we give $\mathrm{U}_1 = \{e^{i\theta} : \theta \in [0, 2\pi)\}$ the normalized Haar measure $\frac{1}{2\pi}\,d\theta$, so $\mathrm{U}_1 \rtimes S_2$ gets the normalized Haar measure $\frac{1}{4\pi}\,d\theta$. For $u = \mathrm{diag}\left(e^{i\theta}, e^{-i\theta}\right) \in U$, we note that the trace of $(u, 1) \in \mathrm{U}_1 \rtimes S_2$ is simply $2\cos\theta$ while the trace of $(u, w) \in \mathrm{U}_1 \rtimes S_2$ vanishes. Thus, we see the Sato–Tate conjecture (Conjecture 3.17) implies that the values

$$\left\{ \mathrm{tr}\, \frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}} \iota\rho_\ell(\mathrm{Frob}_\mathfrak{p}) \right\}_{\mathfrak{p}\text{ prime}}$$

equidistribute according to the measure $\frac{1}{2\pi} \cdot \frac{1}{\sqrt{4-t^2}}\,dt + \frac{1}{2}\delta_0\,dt$ on $[-2, 2]$. Here, δ_0 refers to the δ-distribution concentrated at 0.

We now return to the Jacobian of (the normalization of the proper curve with affine chart) $y^9 = x(x-1)(x-\lambda)$. It will be helpful to take products of Haar measures in the sequel. The following result is an easier form of [DE14, Proposition 1.5.6].

Lemma 3.41. Fix a locally compact topological group G. Choose closed subgroups $H, K \subseteq G$ such that $G = HK$ and $K \subseteq C_G(H)$. Letting dh and dg be left Haar measures on H and K, respectively, we find that $dk\,dh$ is a left Haar measure on G.

Proof. We are tasked with showing that the integral

$$\int_H \int_K f(hk)\,dk\,dh$$

is left-invariant for G. It is left-invariant for H with no content, so it suffices to show the same for K. This follows after some manipulation because K commutes with H. ∎

Remark 3.42. In fact, [DE14, Proposition 1.5.6] shows something much stronger: one can replace the strong group-theoretic condition that $K \subseteq C_G(H)$ with merely that K is compact. In fact, a careful reading of the proof there reveals that we may even replace the condition that K is compact with merely having $H \cap K$ compact and $\Delta_G|_K = 1$, where Δ_G is the modular function on G.

Here is our application.

Proposition 3.43. Let A be the Jacobian of the normalization of the proper curve with affine chart $y^9 = x(x-1)(x-\lambda)$, where λ lives in a number field. Suppose that A does not have complex multiplication. We compute a Haar measure on $\mathrm{ST}(A_K)$ whenever K contains K_A^{conn}.

Proof. The Sato–Tate computation of Proposition 3.29 (combined with the conjugacy class computation of Lemma 3.36) reveals that an element of $\mathrm{Conj}(\mathrm{ST}(A))$ can be written as

$$\mathrm{diag}\left(\begin{bmatrix} e^{i\alpha_0} & \\ & e^{-i\alpha_0} \end{bmatrix}, \begin{bmatrix} e^{i\alpha_1+i\theta_1} & \\ & e^{i\alpha_1-i\theta_1} \end{bmatrix}, \begin{bmatrix} e^{i\alpha_1+i\alpha_4+i\theta_2} & \\ & e^{i\alpha_1+i\alpha_4-i\theta_2} \end{bmatrix}, \begin{bmatrix} e^{i\alpha_4+i\theta_4} & \\ & e^{i\alpha_4-i\theta_4} \end{bmatrix}, \right.$$
$$\left. \begin{bmatrix} e^{-i\alpha_4+i\theta_4} & \\ & e^{-i\alpha_4-i\theta_4} \end{bmatrix}, \begin{bmatrix} e^{-i\alpha_1-i\alpha_4+i\theta_2} & \\ & e^{-i\alpha_1-i\alpha_4-i\theta_2} \end{bmatrix}, \begin{bmatrix} e^{-i\alpha_1+i\theta_1} & \\ & e^{-i\alpha_1-i\theta_1} \end{bmatrix} \right)$$

where $\alpha_\bullet \in [0, 2\pi)$ and $\theta_\bullet \in [0, \pi)$. Technically, the map $(\alpha_\bullet, \theta_\bullet)\colon [0, 2\pi)^4 \times [0, \pi)^3 \to \mathrm{Conj}(\mathrm{ST}(A))$ is the finite-to-one because $Z(\mathrm{ST}(A))^\circ \cap \mathrm{ST}(A)^{\mathrm{der}}$ is finite, but this will make no effect on our computations as long as we normalize to have total volume 1 and only integrate against genuine functions on $\mathrm{Conj}(\mathrm{ST}(A))$.

Anyway, we see that the trace is given by

$$2\cos\alpha_0 + 2\cos(\alpha_1+\theta_1) + 2\cos(\alpha_1-\theta_1) + 2\cos(\alpha_1+\alpha_4+\theta_2) + 2\cos(\alpha_1+\alpha_4-\theta_2) + 2\cos(\alpha_4+\theta_4) + 2\cos(\alpha_4-\theta_4).$$

We finish by remarking that Lemma 3.41 gives our Haar measure as

$$\frac{1}{(2\pi)^3}\,d\alpha_0\,d\alpha_1\,d\alpha_4 \cdot \frac{1}{\pi^2}\left(2\sin^2\theta_1 \cdot 2\sin^2\theta_2 \cdot 2\sin^2\theta_4\right) d\theta_1\,d\theta_2\,d\theta_4,$$

which is what we wanted. (Note we used Remark 3.37 for the Haar measure on SU_2.) ■

Proposition 3.44. Let A be the Jacobian of the normalization of the proper curve with affine chart $y^9 = x^3 - 1$. Suppose that A does not have complex multiplication. We compute a Haar measure on $\mathrm{ST}(A_K)$ whenever K contains K_A^{conn}.

Proof. The Sato–Tate computation of Proposition 3.32 reveals that an element of $\mathrm{Conj}(\mathrm{ST}(A))$ can be written as

$$\mathrm{diag}\left(e^{i\alpha_2-i\alpha_1-\alpha_4}, e^{i\alpha_1+i\alpha_4-i\alpha_2}, e^{-i\alpha_4}, e^{i\alpha_1}, e^{i\alpha_2}, e^{-\alpha_2}, e^{-\alpha_1}, e^{i\alpha_4}, e^{i\alpha_1}, e^{i\alpha_2}, e^{i\alpha_4}, e^{-i\alpha_4}, e^{-i\alpha_2}, e^{-i\alpha_1}\right)$$

where $\alpha_\bullet \in [0, 2\pi)$. For example, we see that the trace is given by

$$2\cos\cos(\alpha_1-\alpha_2+\alpha_4) + 4\cos\alpha_1 + 4\cos\alpha_2 + 4\cos\alpha_4$$

We finish by remarking that Lemma 3.41 gives our Haar measure as

$$\frac{1}{(2\pi)^3}\,d\alpha_1\,d\alpha_2\,d\alpha_4,$$

which is what we wanted. ■

Remark 3.45. As remarked at the end of the proof of Proposition 3.33, we can run the exact same computation with working the curve given by $y^9 = x\left(x^2+1\right)$ because the resulting Sato–Tate group is the same up to reordering the basis.

Remark 3.46. For the previous examples, there are more interesting coefficients in the characteristic polynomial than merely the trace. Hoever, they are rather lengthy to write down, so we have chosen not to.

It still remains to explain how we numerically verify the Sato–Tate conjecture. The idea is that we can try to compute

$$\mathrm{tr}\,\frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}}\iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})$$

for various primes $\mathfrak{p}$ and then compare it with what is expected from

$$\int_{\mathrm{Conj}(\mathrm{ST}(A))} \mathrm{tr}\,g\,dg,$$

where dg refers to the pushforward of the Haar measure from $\mathrm{Conj}(\mathrm{ST}(A))$. One usually expects the above integral to vanish, so one can either look at other coefficients of the characteristic polynomial or at powers of $\mathrm{tr}\,g$. In the sequel, we will compute with only powers of $\mathrm{tr}\,g$ for simplicity, but we do remark that one can typically recover the other coefficients via a combination of Vieta's formulae and Newton's sums.

As usual, let's begin with elliptic curves. Here, explicit formulae are possible.

Example 3.47 (no complex multiplication)**.** We continue with the elliptic curve $E\colon y^2 = x^3 + x + 1$ over $\mathbb{Q}$ studied in Examples 3.25 and 3.38. Fix some integer $m \geq 0$. Using the given Haar measure (from Remark 3.37), we find that one expects the average of $\left\{\left(\operatorname{tr} \frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}} \iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right)^m\right\}_{\mathfrak{p} \text{ prime}}$ to be

$$\int_0^\pi (2\cos\theta)^m \frac{2}{\pi} \sin^2\theta \, d\theta = \begin{cases} \frac{1}{m/2+1}\binom{m}{m/2} & \text{if } m \text{ is even,} \\ 0 & \text{if } m \text{ is odd,} \end{cases}$$

where the last equality is verified by expanding $2\cos\theta = e^{i\theta} + e^{-i\theta}$ and $4\sin^2\theta = 2 - e^{2i\theta} - e^{-2i\theta}$.

Example 3.48 (complex multiplication)**.** We continue with the elliptic curve $E\colon y^2 = x^3 + 1$ over $\mathbb{Q}(\zeta_3)$ studied in Examples 3.26 and 3.39. Fix some integer $m \geq 0$. Using the given Haar measure, we find that one expects the average of $\left\{\left(\operatorname{tr} \frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}} \iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right)^m\right\}_{\mathfrak{p} \text{ prime}}$ to be

$$\int_0^{2\pi} (2\cos\theta)^m \frac{1}{2\pi} \, d\theta = \begin{cases} \binom{m}{m/2} & \text{if } m \text{ is even,} \\ 0 & \text{if } m \text{ is odd,} \end{cases}$$

where the last equality is verified by expanding $2\cos\theta = e^{i\theta} + e^{-i\theta}$.

Example 3.49 (complex multiplication)**.** We continue with the elliptic curve $E\colon y^2 = x^3 + 1$ over $\mathbb{Q}$ studied in Examples 3.27 and 3.40. Fix some integer $m \geq 0$. Using the given Haar measure, we find that one expects the average of $\left\{\left(\operatorname{tr} \frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}} \iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right)^m\right\}_{\mathfrak{p} \text{ prime}}$ to be

$$\int_0^{2\pi} (2\cos\theta)^m \frac{1}{4\pi} \, d\theta = \begin{cases} \frac{1}{2}\binom{m}{m/2} & \text{if } m \text{ is even,} \\ 0 & \text{if } m \text{ is odd,} \end{cases}$$

where the last equality is verified by expanding $2\cos\theta = e^{i\theta} + e^{-i\theta}$.

We now return to $y^9 = x(x-1)(x-\lambda)$. Here, we do not attempt to give explicit formulae, but we list the first few expected values, which were computed using SageMath.

Example 3.50. Let A be the Jacobian of the normalization of the proper curve with affine chart $y^9 = x(x-1)(x-10)$. SageMath can verify that A does not have complex multiplication. For $m \in \{0, 1, \ldots, 6\}$, we use Proposition 3.43 to find that we expect the aveage of $\left(\operatorname{tr} \frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}} \iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right)^m$ as $\mathfrak{p}$ varies over primes K (for K containing K_A^{conn}) to be as follows.

m	0	1	2	3	4	5	6
expected	1	0	8	0	186	0	7160
actual	1.0	0.0	7.8	0.2	180	16	6400

Here, the "actual" amounts have been rounded to two significant digits, and they were computed by averaging over primes $p < 216289$ which were $1 \pmod 9$; the condition $p \equiv 1 \pmod 9$ corresponds to splitting completely in $\mathbb{Q}(\zeta_9)$ (see Remark 3.35). These "actual" amounts suggest that $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$, a fact which we will verify in the next chapter.

Example 3.51. Let A be the Jacobian of the normalization of the proper curve with affine chart $y^9 = x^3 - 1$, where λ lives in a number field. For $m \in \{0, 1, \ldots, 6\}$, we use Proposition 3.44 to find that we expect the aveage of $\left(\operatorname{tr} \frac{1}{\sqrt{\mathrm{N}(\mathfrak{p})}} \iota\rho_\ell(\mathrm{Frob}_\mathfrak{p})\right)^m$ as $\mathfrak{p}$ varies over primes K (for K containing K_A^{conn}) to be as follows.

m	0	1	2	3	4	5	6
expected	1	0	26	0	2118	0	239300
actual	1.0	0.0	25	6.0	2000	890	220000

Here, the "actual" amounts have been rounded to two significant digits, and they were computed by averaging over primes $p < 100000$ which were $1 \pmod 9$; the condition $p \equiv 1 \pmod 9$ corresponds to splitting completely in $\mathbb{Q}(\zeta_9)$ (see Remark 3.35). These "actual" amounts suggest that $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$, a fact which we will verify in the next chapter.

Remark 3.52. If one runs the same computation as in the previous example with $y^9 = x\left(x^2 + 1\right)$, one should further restrict primes past $p \equiv 1 \pmod 9$ in order to see the correct moment statistics. This is because now $K_A^{\mathrm{conn}} \neq \mathbb{Q}(\zeta_9)$.

3.2 The Utility of L-Functions

In this section, we will explain how L-functions are used in analytic number theory. Before delving into the main content of this section, we give a rough indication of what an L-function is, though we will wait to explain why we care. One generally expects an L-function to have a Dirichlet series

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

which converges in some region $\{s \in \mathbb{C} : \operatorname{Re} s > \sigma\}$, where σ is a real number. (We may call σ the "abscissa" of convergence.) In this situation, one may find that σ to is a pole of $L(s)$ (though not always), but we usually expect $L(s)$ to admit a meromorphic continuation beyond $\{s \in \mathbb{C} : \operatorname{Re} s > \sigma\}$.

Another important feature is that L-functions frequently come with "Euler products" that look like

$$L(s) = \prod_p L_p(s),$$

where the "Euler factor" $L_p(s)$ is a rational function in p^{-s}. We will be mostly interested in non-vanishing and holomorphy of our L-functions, and these properties tend to be insensitive to adjusting finitely many Euler factors. Thus, we pick up the following notation.

Notation 3.53. Given two infinite products $\prod_p a_p$ and $\prod_p b_p$, we write

$$\prod_p a_p \doteq \prod_p b_p$$

if and only if the two products are equal up to a finite number of nonzero terms.

3.2.1 The Prime Number Theorem

To prove an equidistribution result, one needs to end up proving some natural density results. For a natural density result, one needs to be able to count a total in order to estimate the denominator. Thus, for Conjecture 3.17, we will need to count the number of primes. As such, in this subsection, we will pick up some

tools from analytic number theory, and then we will prove the prime number theorem as an application. Our exposition is very standard; for example, all arguments are results can be found in [Mur08, Chapter 3].

Formally, the prime number theorem states that

$$\sum_{p \le x} 1 \sim \frac{x}{\log x}.$$

Now, even though we are interested in counting primes, it is easier to prove a result of the form

$$\sum_{p \le x} \log p \sim x$$

because the right-hand side is simpler (roughly speaking). Quickly, we give names to our "prime-counting" functions of interest.

Definition 3.54. For $x > 0$, define $\pi(x)$ as the number of primes $p \le x$, and define

$$\psi(x) \coloneqq \sum_{\substack{p \text{ prime}, k > 0 \\ p^k \le x}} \log p.$$

For brevity, we let $\Lambda(n)$ be $\log p$ if n is a power of a prime p and 0 otherwise; then $\psi(x) = \sum_{n \le x} \Lambda(n)$.

It is easier to estimate ψ than π, but their estimates can be shown to be equivalent. To explain this, we use Abel summation.

Proposition 3.55 (Abel summation)**.** Choose a sequence of complex numbers $\{b_n\}_{n \ge 1}$, and set $B(x) \coloneqq \sum_{n \le x} b_n$. For any continuously differentiable $f \colon [0, \infty) \to \mathbb{C}$, we have

$$\sum_{a \le n \le x} b_n f(n) = B(x) f(x) - \int_1^x B(t) f'(t)\, dt.$$

Proof. The main idea is to write $b_n = B(n) - B(n-1)$, so telescoping shows

$$\sum_{n \le x} b_n f(n) = B(\lfloor x \rfloor) f(\lfloor x \rfloor) - \sum_{n < \lfloor x \rfloor} B(n)(f(n+1) - f(n)).$$

Now, $f(n+1) - f(n) = \int_n^{n+1} f'(t)\, dt$, so the sum collapses into the integral

$$\sum_{n \le x} b_n f(n) = B(\lfloor x \rfloor) f(\lfloor x \rfloor) - \int_1^{\lfloor x \rfloor} B(t) f'(t)\, dt.$$

It remains to move from $\lfloor x \rfloor$ to x, for which we note that $B(t) = B(\lfloor x \rfloor)$ for $t \in [\lfloor x \rfloor, x]$, so

$$B(x) f(x) - B(\lfloor x \rfloor) f(\lfloor x \rfloor) = \int_{\lfloor x \rfloor}^x B(t) f'(t)\, dt,$$

thereby completing the proof. ∎

Corollary 3.56. If $\psi(x) \sim x$, then $\pi(x) \sim x / \log x$.

Proof. Given $\psi(x) \sim x$, we begin by claiming $\sum_{p \le x} \log p \sim x$. Indeed,

$$\left| \psi(x) - \sum_{p \le x} \log p \right| = \sum_{\substack{p \text{ prime}, k > 1 \\ p^k \le x}} \log p.$$

We bound this sum unintelligently: it is

$$\sum_{k=2}^{\log_2 x} \sum_{p \le x^{1/k}} \log p \le (\log_2 x)(\sqrt{x} \log x),$$

which is $o(x)$, and the claim follows.

We now show $\pi(x) \sim x/\log x$. This requires Abel summation in the form of Proposition 3.55. Indeed, we see $\pi(x)$ equals

$$\sum_{n \le x} 1_{\text{is prime}}(n) \log n \cdot \frac{1}{\log n} = \frac{1}{\log x} \sum_{p \le x} \log p + \int_2^x \left(\sum_{p \le x} \log p \right) \frac{1}{t(\log t)^2} \, dt.$$

Thus, it remains to show that the integral is $o(x/\log x)$. Well, $\sum_{p \le x} \log p \sim x$, so it is enough to show that the intgral $\int_2^x (\log t)^{-2} \, dt$ is $o(x/\log x)$. Well, for x large, we see that

$$\int_2^{\sqrt{x}} \frac{1}{(\log t)^2} \, dt + \int_{\sqrt{x}}^x \frac{1}{(\log t)^2} \, dt \le \sqrt{x} + \frac{x}{(\log \sqrt{x})^2},$$

which is manifestly $o(x/\log x)$. ∎

Remark 3.57. In fact, one can reverse the application of Proposition 3.55 to show the reverse implication, but we will not need this.

We will spend the rest of our time trying to show that $\psi(x) \sim x$. We will use a weak form of the Weiner–Ikehara theorem to prove this from some analytic properties of the Riemann zeta function. As such, we spend some time working towards the Weiner–Ikehara theorem. Our approach follows [New80] and uses the following Tauberian theorem.

Theorem 3.58 (Newman)**.** Let $f \colon [0, \infty) \to \mathbb{C}$ be a bounded and piecewise continuous function, and let $F(s) \coloneqq \int_{\mathbb{R}^+} f(t) e^{-st} \, dt$ denote the Laplace transform. Suppose that $F(s)$ admits an analytic continuation to the half-plane $\{s \in \mathbb{C} : \operatorname{Re} s \ge 0\}$. Then the integral

$$\int_{\mathbb{R}^+} f(t) \, dt$$

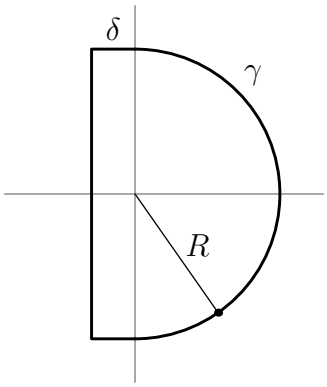
converges and equals $F(0)$.

Proof. In order to estimate with convergent integrals, for any $T > 0$, we define the function $F_T \colon \mathbb{C} \to \mathbb{C}$

$$F_T(s) \coloneqq \int_0^T f(t) e^{-st} \, dt.$$

We quickly remark that F is analytic on $\{s \in \mathbb{C} : \operatorname{Re} s > 0\}$ for free because boundedness of f implies that the integral converges in this region; similarly, we note that F_T is automatically entire for any $T > 0$.

Our goal is to show that $F_T(0) \to F(0)$ as $T \to \infty$. We will estimate $|F(0) - F_T(0)|$ via some clever contour integration. Fix some $R > 0$, which will eventually tend to ∞. Then we note that compactness of the interval $\{bi : -R \le b \le R\}$ implies that there is $\delta > 0$ such that the analytic continuation of F extends to an open set containing the box $\{a + bi : a \ge -\delta, -R \le b \le R\}$. We now let γ denote the following contour, oriented counterclockwise.



We also let γ_+ and γ_- denote the parts in the right-half and left-half planes, respectively. Now, the main trick is to note that

$$F(0) - F_T(0) = \frac{1}{2\pi i} \int_\gamma \frac{F(s) - F_T(s)}{s} \cdot e^{sT} \left(1 + \frac{s^2}{R^2}\right) ds$$

by the Cauchy integral formula. (The magic will come from the strange factor $e^{sT}\left(1 + s^2/R^2\right)$.) We now estimate this integral as (in order) $T \to \infty$, $\delta \to 0$, and $R \to \infty$.

- We estimate the integral on γ_+. This can be done directly. On one hand, expanding out the integral reveals

  $$|F(s) - F_T(s)| \leq \|f\|_\infty \cdot \frac{e^{-T \operatorname{Re} s}}{\operatorname{Re} s}.$$

  On the other hand, we note s/R is on the unit circle, so

  $$\left|\frac{e^{st}}{s}\left(1 + \frac{s^2}{R^2}\right)\right| \leq \frac{e^{T \operatorname{Re} s}}{R} \cdot 2 \operatorname{Re}(s/R).$$

  Combining estimates, we bound our integral by

  $$|F(0) - F_T(0)| \leq \frac{\|f\|_\infty}{R},$$

  which vanishes as $R \to \infty$, as required.

- To estimate the integral on γ_-, we split the integral into a sum of integrals of F and F_T separately. In this point, we bound the integral of F_T. Here, F_T is entire, we may replace the contour γ_- with a semicircle of radius R in the left-half plane. Proceeding as in the above point, we note that

  $$|F_T(s)| \leq \|f\|_\infty \cdot \frac{e^{-T \operatorname{Re} s}}{-\operatorname{Re} s}$$

  by expanding out the integral, and then estimating $e^{st}\left(1 + s^2/R^2\right)$ as before yields

  $$\left|\frac{1}{2\pi i} \int_{\gamma_-} \frac{F_T(s)}{s} \cdot e^{sT} \left(1 + \frac{s^2}{R^2}\right) ds\right| \leq \frac{\|f\|_\infty}{R},$$

  which again vanishes as $R \to \infty$.

- It remains to bound the integral of F over γ_-. This will require some care. We will split the estimates into the horizontal and vertical pieces. Throughout, R and δ remain fixed, and we will only send $T \to \infty$; in particular, F is bounded in the region of interest, so we may ignore its contribution.

  – On the horizontal pieces, for $\delta > 0$ small enough, we may still find that our integrand is on the order of $e^{T \operatorname{Re} s} \cdot 2 \operatorname{Re} s$, as in the first point. However, we note that the function $x \mapsto xe^{-x}$ on $\mathbb{R}^+$ achieves its maximum at $(1, 1/e)$, so with $\operatorname{Re} s < 0$, we see that our integrand is bounded by e^{-1}/T. To complete our estimate, we send $T \to \infty$.

– On the vertical piece, for $\delta > 0$ small enough, we note

$$\left| \frac{e^{sT}}{s} \left(1 + \frac{s^2}{R^2}\right) \right| \leq \frac{3e^{-\delta T}}{\delta}.$$

Sending $T \to \infty$ causes this piece to vanish. ■

We are now ready to prove our weakened Weiner–Ikehara theorem. We follow [Vat15, Theorem 2].

Theorem 3.59 (Weiner–Ikehara)**.** Choose a sequence of nonnegative real numbers $\{b_n\}_{n \geq 1}$, and set $L(s) \coloneqq \sum_{n \geq 1} b_n n^{-s}$ and $B(x) \coloneqq \sum_{n \leq x} b_n$. Suppose the following.

(i) The series $L(s)$ converges absolutely for $\operatorname{Re} s > 1$.

(ii) The function $L(s)$ admits a meromorphic continuation to $\operatorname{Re} s = 1$ and has no poles except possibly a simple pole at $s = 1$ with residue c.

(iii) We have $B(x) = O(x)$.

Then $B(x) = cx + o(x)$.

Proof. There are two steps.

1. By Proposition 3.55, we see that
$$L(s) = s \int_1^\infty B(t) t^{-s-1} \, dt$$
holds for $\operatorname{Re} s > 1$. Now, the idea is to apply Theorem 3.58 to the integral
$$\int_0^\infty \frac{B(e^t) - ce^t}{e^t} e^{-st} \, dt = \frac{L(s+1)}{s+1} - \frac{c}{s},$$
where the equality follows from the previous one after the substitutions $s \mapsto s+1$ and $t \mapsto e^t$. Notably, we are allowed to apply Theorem 3.58 because one already knows that the function $e^{-t} B(e^t) - c$ is bounded by (iii), and the right-hand side provides the required analytic continuation. Thus, we are told that
$$\int_0^\infty \frac{B(e^t) - ce^t}{e^t} \, dt = \int_1^\infty \frac{B(t) - ct}{t^2} \, dt$$
converges.

2. We are now ready to conclude. We must show that $B(x)/x \to 1$ as $x \to \infty$. Suppose for the sake of contradiction this is not the case; then either $\limsup_{x \to \infty} B(x)/x > c$ or $\liminf_{x \to \infty} B(x)/x < c$. We handle the case $\limsup_{x \to \infty} B(x)/x > c$ because the other case is similar. In this case, there is $\varepsilon > 0$ and an infinite sequence $\{x_i\}_{i \geq 1}$ tending to infinity such that $B(x_i)/x_i > c(1+\varepsilon)$ for all $i \geq 1$. For any such x_i, we see that
$$\int_{x_i}^{(1+\varepsilon)x_i} \frac{B(t) - ct}{t^2} \, dt \geq \int_{x_i}^{(1+\varepsilon)x_i} \frac{(c+\varepsilon)x_i - ct}{t^2} \, dt.$$
Upon a change of variables, we see this integral equals $\int_1^{1+\varepsilon} (c(1+\varepsilon) - ct) t^{-2} \, dt$, which is some nonzero constant not depending on x_i. Because we can let the x_i tend to infinity, we conclude that the integral $\int_1^\infty (B(t) - ct) t^{-2} \, dt$ cannot converge! This is our required contradiction. ■

Remark 3.60. Note that L is by definition real on the real axis (when the series converges), which implies that the residue c must be real because the residue equals the limit of $sL(s)$ as $s \to 1^+$.

Remark 3.61. The hypothesis (c) in the statement of Theorem 3.59 is not necessary, but one requires a somewhat more technical proof.

We will now apply Theorem 3.59 to show $\psi(x) \sim x$. Because (iii) of Theorem 3.59 is unrelated to the other two conditions, we handle it first. The argument is combinatorial.

Lemma 3.62 (Chebychev)**.** We have $\psi(x) = O(x)$.

Proof. Arguing as in Corollary 3.56, it is enough to show that $\sum_{p \le x} \log p = O(x)$. We proceed in steps.

1. For any $n \ge 0$, we claim that $\sum_{n<p\le 2n} \log p < 2n \log 2$. The idea is to consider $\binom{2n}{n}$. By expanding out its prime factorization, we note that $\binom{2n}{n}$ has each prime factor p in the range $n < p \le 2n$, so $\log \binom{2n}{n} \ge \sum_{n<p\le 2n} \log p$. On the other hand, the binomial theorem requires $\binom{2n}{n} < (1+1)^{2n}$, so $\log \binom{2n}{n} < 2n \log 2$, as required.
2. For any $\nu \ge 0$, we claim that $\sum_{p \le 2^\nu} \log p < 2^{\nu+1} \log 2$. Indeed, this sum is
$$\sum_{k=0}^{\nu-1} \left(\sum_{2^k < p \le 2^{k+1}} \log p \right) \le \sum_{k=0}^{\nu-1} 2^{k+1} \log 2$$
by the previous step, from which the claim follows.
3. We complete the proof. For any $x > 1$, we may find $\nu \ge 0$ such that $2^\nu \le x < 2^{\nu+1}$. Then $\sum_{p \le x} \log p$ is bounded by $2^{\nu+2} \log 2$ by the previous step, but this is in turn bounded by $4x \log 2$, so we are done. ∎

For (i) of Theorem 3.59, we must explain the relevance of the Riemann ζ-function to our argument.

Definition 3.63 (Riemann ζ-function)**.** We define the *Riemann ζ-function* by
$$\zeta(s) \coloneqq \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Lemma 3.64. For s such that $\operatorname{Re} s > 1$, we have $\zeta(s) \ne 0$ and
$$\frac{\zeta'}{\zeta}(s) = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$
This series also converges absolutely for $\operatorname{Re} s > 1$.

Proof. Unique prime factorization produces the Euler product
$$\zeta(s) = \prod_p \frac{1}{1-p^{-s}},$$
so taking logarithms implies
$$\log \zeta(s) = \sum_p -\log\left(1-p^{-s}\right).$$
Using the Taylor expansion of $\log(1-x)$, we find that
$$\log \zeta(s) = \sum_p \sum_{k=1}^{\infty} \frac{1}{kp^{ks}}.$$

The claimed equality would now follow by taking the derivative with respect to s, but of course, we must know that $\log \zeta(s)$ is an analytic function to be able to do this. Well, we will actually show that the right-hand side is absolutely convergent, which we note then implies $\zeta(s) \neq 0$. To check absolute convergence, we may rearrange our sum, so we sum over k. The $k = 1$ term is bounded by $\zeta(\operatorname{Re} s)$, which is finite. For the remaining terms, we bound our sum in magnitude by

$$\sum_{n=1}^{\infty} \sum_{k=2}^{\infty} \frac{1}{n^k} = \sum_{n=1}^{\infty} \frac{1/n^2}{1 - 1/n}.$$

The summand is $\frac{1}{n(n-1)}$, so the entire sum converges. ■

Thus, the function $L(s)$ arising from trying to show $\psi(x) \sim x$ is simply $\zeta'(s)/\zeta(s)$. It remains to show the required facts about meromorphic continuation for (ii). We begin by showing that ζ continues.

Lemma 3.65. The function $\zeta(s)$ admits a meromorphic continuation to $\operatorname{Re} s > 0$ with no poles except a simple pole at $s = 1$ with residue 1.

Proof. We use Abel summation. By Proposition 3.55, we see that

$$\zeta(s) = s \int_1^{\infty} \lfloor t \rfloor \, t^{-s-1} \, dt.$$

Now, we write $\lfloor t \rfloor = t - \{t\}$ to see

$$\zeta(s) = \frac{s}{s-1} - s \int_1^{\infty} \{t\} t^{-s-1} \, dt.$$

The listed claims will follow once we show that the reminaing integral $I(s)$ is analytic on $\operatorname{Re} s > 0$. Well, we see

$$|I(s)| \leq \int_1^{\infty} \frac{1}{t^{\operatorname{Re} s + 1}} \, dt = \frac{1}{\operatorname{Re} s},$$

so the integral is always finite, so the integral is analytic because the integrand is.[3] ■

Thus, the check (ii) of Theorem 3.59 amounts to the following non-vanishing result.

Proposition 3.66. If $s \in \mathbb{C}$ has $\operatorname{Re} s = 1$, then $\zeta(s) \neq 0$.

Proof. The following proof is tricky. We proceed in steps, following [Mur08, Section 3.2].

1. For $\sigma > 1$ and $t \in \mathbb{R}$, we claim that

$$\operatorname{Re} \log \zeta(\sigma + it) \stackrel{?}{=} \sum_{n=2}^{\infty} \frac{\Lambda(n) \cos(t \log n)}{n^{\sigma} \log n}.$$

Well, the argument of Lemma 3.64 (equivalently, integrating the statement) shows that

$$\log \zeta(s) = \sum_{p} \sum_{k=1}^{\infty} \frac{1}{k p^{ks}},$$

[3] This point techincally requires some care because one needs to apply some kind of dominated convergence theorem as in [Mat01]. Our proof actually shows that $I(s)$ is analytic on any region $\{s : \operatorname{Re} s > \sigma\}$ for any $\sigma > 0$, from which $I(s)$ being analytic on $\{s : \operatorname{Re} s > 0\}$ follows by taking unions.

where $s = \sigma + it$. This sum absolutely converges (as shown in Lemma 3.64), so we may view it as a sum over prime-powers $n = p^k$, in which case we see that the summand is $\Lambda(n)n^{-s}/\log n$. Thus, we see that

$$\log \zeta(s) = \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n^\sigma \log n} \cdot n^{-it}.$$

We conclude by noting that $\operatorname{Re} n^{-it} = \cos(t \log n)$.

2. For $\sigma > 1$ and $t \in \mathbb{R}$, we claim that

$$\left|\zeta(\sigma)^3 \zeta(\sigma + it)^4 \zeta(\sigma + 2it)\right| \stackrel{?}{\geq} 1.$$

Well, by taking logarithms, it is enough to show that

$$3 \operatorname{Re} \log \zeta(\sigma) + 4 \operatorname{Re} \log \zeta(\sigma + it) + \operatorname{Re} \log \zeta(\sigma + 2it) \stackrel{?}{\geq} 0.$$

By the previous step, we see that it is enough to check that

$$3 + 4\cos\theta + \cos 2\theta \geq 0$$

for any $\theta \in \mathbb{R}$. This amounts to mimizing the function $4\cos\theta + \cos 2\theta$; taking the derivative reveals that minima will occur when $\sin x = 0$ or $\cos x = 1$, so x is a multiple of π. Thus, we complete this step by noting that the above inequality holds when x is a multiple of π.

3. We conclude the proof. Fix some nonzero real number t, and we would like to show that $\zeta(1 + it) \neq 0$. Well, suppose for the sake of contradiction that $\zeta(1 + it) = 0$. Then

$$\lim_{\sigma \to 1^+} \zeta(\sigma)^3 \zeta(\sigma + it)^4 \zeta(\sigma + 2it) = 0$$

because the order of the zero at $\sigma = 0$ is at least $-3 + 4 + 0 > 0$. This contradicts the previous step, so we are done. ■

We are now ready to prove the Prime number theorem.

Theorem 3.67 (Prime number)**.** We have $\pi(x) \sim x/\log x$.

Proof. It only remains to synthesize the discussion from this subsection. By Corollary 3.56, it is enough to show $\psi(x) \sim x$. For this, we will use Theorem 3.59 applied to the sequence $\{\Lambda(n)\}_{n \geq 1}$, for which Lemma 3.64 explains makes the Dirichlet series equal to $-\zeta'(s)/\zeta(s)$. It remains to check the three conditions in Theorem 3.59.

(i) The absolute convergence of $-\zeta'(s)/\zeta(s)$ follows because $\Lambda(n) = O(n^\varepsilon)$ for any $\varepsilon > 0$, so the series converges absolutely and uniformly on compacts on any region $\{s \in \mathbb{C} : \operatorname{Re} s > \varepsilon\}$ for any $\varepsilon > 0$.

(ii) Because $\zeta(s)$ is nonzero on $\{s : \operatorname{Re} s = 1\}$, we conclude that $-\zeta'(s)/\zeta(s)$ admits a meromorphic continuation to this line. We already know that we are defined everywhere except at $s = 1$, and we see that having ζ have a simple pole with reside 1 at $s = 1$ implies the same for $-\zeta'(s)/\zeta(s)$ by expanding out a Taylor series at $s = 1$.

(iii) Lastly, we see $\psi(x) = O(x)$ by Lemma 3.62. ■

3.2.2 The Prime Ideal Theorem

In the sequel, we will want to count not just rational primes but also for number fields, so we want to extend Theorem 3.67 to number fields. The method remains the same, though we will not give a complete proof now because showing the required meromorphic continuation is harder. Our exposition loosely follows [RV99, Sections 7.4 and 7.7], which in turn follows [Hei10].

Here are our prime-counting functions.

Definition 3.68. Fix a number field K. For $x > 0$, define $\pi_K(x)$ as the number of prime ideals $\mathfrak{p}$ with $\mathrm{N}\,\mathfrak{p} \le x$. Now, define Λ_K as a function on the ideals of $\mathcal{O}_K$ by

$$\Lambda_K(I) \coloneqq \begin{cases} \log \mathrm{N}\,\mathfrak{p} & \text{if } I = \mathfrak{p}^k \text{ for } k \ge 1, \\ 0 & \text{else,} \end{cases}$$

and we set $\psi_K(x) \coloneqq \sum_{\mathrm{N}(I) \le x} \Lambda_K(I)$.

This time around, the relevant L-function for the Weiner–Ikehara theorem is as follows.

Definition 3.69 (Dedekind zeta function)**.** Fix a number field K. Then we define the *Dedekind ζ-function* as

$$\zeta_K(s) \coloneqq \sum_{I \subseteq \mathcal{O}_K} \frac{1}{\mathrm{N}(I)^s}.$$

Remark 3.70. As in Lemma 3.64, we note that one has an Euler product

$$\zeta_K(s) = \prod_{\mathfrak{p} \text{ prime}} \frac{1}{1 - \mathrm{N}\,\mathfrak{p}^{-s}}.$$

It will later be convenient to "twist" our Dedekind zeta function slightly.

Definition 3.71 (Hecke L-function)**.** Fix a number field K and a continuous character $\chi\colon K^\times \backslash \mathbb{A}_K^\times \to \mathbb{C}^\times$. Factoring $\chi = \prod_v \chi_v$ as a product over places of K, we define the *Hecke L-function* as $L(\chi) \coloneqq \prod_{\mathfrak{p}} (1 - \chi_{\mathfrak{p}}(\mathfrak{p}))^{-1}$, where

$$\chi_{\mathfrak{p}}(\mathfrak{p}) \coloneqq \begin{cases} \chi_{\mathfrak{p}}(\varpi_{\mathfrak{p}}) & \text{if } \chi_{\mathfrak{p}}|_{\mathcal{O}_{\mathfrak{p}}^\times} = 1, \\ 0 & \text{else,} \end{cases}$$

where $\varpi_{\mathfrak{p}} \in \mathfrak{p}$ is a uniformizer. We may call the former case "unramified" and the latter case "ramified." If χ is a unitary character (i.e., $\operatorname{im} \chi \subseteq S^1$), then we may also write $L(s, \chi) \coloneqq L\left(\chi |\cdot|^s\right)$.

Remark 3.72. By expanding out $(1 - \chi_{\mathfrak{p}}(\mathfrak{p}))^{-1} = \sum_{k=0}^{\infty} \chi_{\mathfrak{p}}(\mathfrak{p})^k$, we see that one can recover a "Dirichlet series" expansion for $L(s, \chi)$ in the form

$$L(s, \chi) = \sum_{I \subseteq \mathcal{O}_K} \frac{\chi(I)}{\mathrm{N}(I)^s}$$

for suitably defined $\chi(I)$ (depending on its prime factorization).

Example 3.73. If χ is the trivial character, then we recover the Dedekind ζ-function.

Example 3.74. Take $K = \mathbb{Q}$. Given a character $\chi\colon (\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^\times$, we abuse notation and lift χ to a character $K^\times\backslash\mathbb{A}_K^\times \to \mathbb{C}^\times$ via the composition

$$\mathbb{Q}^\times\backslash\mathbb{A}_\mathbb{Q}^\times = \mathbb{R}^+ \times \prod_p \mathbb{Z}_p^\times \twoheadrightarrow \prod_p \mathbb{Z}_p^\times \twoheadrightarrow (\mathbb{Z}/n\mathbb{Z})^\times \xrightarrow{\chi} \mathbb{C}^\times.$$

Upon expanding out the Euler product, we find

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

We begin by stating part of the required analytic check.

Lemma 3.75. Fix a number field K and a continuous unitary character $\chi\colon K^\times\backslash\mathbb{A}_K^\times \to \mathbb{C}^\times$. Then $L(s,\chi)$ converges absolutely and is nonzero for s such that $\operatorname{Re} s > 1$.

Proof. We will instead show that

$$\log L(s,\chi) = \sum_{\mathfrak{p}} -\log\left(1 - \chi_\mathfrak{p}(\mathfrak{p})\,\mathrm{N}\,\mathfrak{p}^{-s}\right)$$

absolutely converges when $\operatorname{Re} s > 1$. (Some formal business involving Euler products can then show that the Dirichlet series described in Remark 3.72 also converges absolutely.) Using the Taylor expansion, our sum is

$$\log L(s,\chi) = \sum_{\mathfrak{p}} \sum_{k\geq 1} \frac{\chi_\mathfrak{p}(\mathfrak{p})^k}{k\,\mathrm{N}\,\mathfrak{p}^{ks}}.$$

Now, to check absolute convergence, we see that may replace $\left|\chi_\mathfrak{p}(\mathfrak{p})^k\right| \in \{0,1\}$ with 1, essentially reducing to the case where $\chi = 1$.

We now find a way to reduce to the case for $K = \mathbb{Q}$, where the result follows from the argument of Lemma 3.64. Well, for each prime $\mathfrak{p}$, we see that $\mathrm{N}\,\mathfrak{p} \geq p$ where p is the prime lying over $\mathfrak{p}$. Further, there are at most $[K:\mathbb{Q}]$ primes of K sitting above p, so we see that

$$|\log L(s,\chi)| \leq [K:\mathbb{Q}] \sum_p \sum_{k\geq 1} \frac{1}{kp^{k\operatorname{Re} s}}.$$

We now have reduced to the situation in Lemma 3.64, so we are done. ∎

Remark 3.76. Term-by-term differentiation shows that the Dirichlet series defining $-L'(s,\chi)/L(s,\chi)$ continues to be absolutely convergent for s satisfying $\operatorname{Re} s > 1$. In particular, we see that

$$-\frac{\zeta_K'(s)}{\zeta_K(s)} = \sum_{I\subseteq\mathcal{O}_K} \frac{\Lambda_K(I)}{\mathrm{N}(I)^s}.$$

This time around, one lacks the integration trick done in Lemma 3.65. The proof is significantly more involved, so we merely state the result we need.

Theorem 3.77 (Hecke)**.** Fix a number field K and a continuous unitary character $\chi\colon K^\times\backslash\mathbb{A}_K^\times \to \mathbb{C}^\times$. Then $L(s,\chi)$ admits a meromorphic continuation to $\{s : \operatorname{Re} s > 0\}$. Further, $L(s,\chi)$ has no poles except a simple pole when $\chi\,|\cdot|^s$ is trivial on all unramified primes.

Proof. It is possible to prove an analytic continuation to $\{s : \operatorname{Re} s = 1\}$ "combinatorially," essentially by counting ideals of bounded norm; for example, see [ME05, Section 11.2]. However, the best proofs of this result go through Tate's thesis [Tat10]. See also [RV99, Theorem 7-19]. ■

Because it is more within reach (and closer in flavor to the results we are interested in), we will prove the needed non-vanishing result.

Proposition 3.78. Fix a number field K and a continuous unitary character $\chi\colon K^\times\backslash\mathbb{A}_K^\times \to \mathbb{C}^\times$. If $s \in \mathbb{C}$ has $\operatorname{Re} s = 1$, then $L(s,\chi) \neq 0$.

Proof. Note that $L(s+it,\chi) = L\left(s, \chi|\cdot|^{it}\right)$, so we may twist χ in order to assume that $s = 1$. Now, if χ is trivial on the finite adeles $(\mathbb{A}_K^\infty)^\times$, then Theorem 3.77 explains that there is a pole, so there is nothing to do. We now admit two lengthy cases. There are two lengthy cases.

- Suppose that $\chi_{\mathfrak{p}}^2$ is nontrivial on some unramified prime $\mathfrak{p}$. In this case, we may proceed as in Proposition 3.66: for $\sigma > 1$, an expansion as in Lemma 3.75 finds that

$$\operatorname{Re}\log L(\sigma,\chi) = \sum_{\mathfrak{p}}\sum_{k\geq 1}\frac{\cos(k\theta_{\mathfrak{p}})}{\operatorname{N}\mathfrak{p}^{k\sigma}},$$

  where $\theta_{\mathfrak{p}} \in \mathbb{R}$ is chosen so that $\chi_{\mathfrak{p}}(\mathfrak{p}) = e^{i\theta_{\mathfrak{p}}}$. But now the trigonmetric identity $3 + 4\cos\theta + \cos 2\theta$ proven in Proposition 3.66 verifies that

$$\left|L(\sigma,1)^3 L(\sigma,\chi)^4 L(\sigma,\chi^2)\right| \gg 1,$$

  where the implied constant comes from replacing the Euler product for $L(\sigma,1)$ with one with the correct Euler factors at ramified primes. We now send $\sigma \to 0^+$ and see that having $L(\sigma) = 0$ would force the entire quantity to vanish by pole-counting: we have a zero of order at least $-3+4+0 > 0$, where notably, the hypothesis implies that there is no pole at $L(1,\chi^2)$.

- Now suppose that χ^2 is trivial on all unramified primes. The idea is to consider the Dirichlet series $L(s) \coloneqq \zeta_K(s)L(s,\chi)$, for which one can use the Dirichlet convolution to find equals

$$L(s) = \sum_{I\subseteq\mathcal{O}_K}\left(\sum_{I=AB}\chi(B)\right)\frac{1}{\operatorname{N}(I)^s},$$

  for suitably defined $\chi(I)$. We are going to appeal to some somewhat difficult fact about Dirichlet series. To this end, we want some input from the coefficient $b(I) \coloneqq \sum_{I=AB}\chi(B)$. Multiplicativity reveals that

$$b(I) = \prod_{\mathfrak{p}}\left(1 + \chi(\mathfrak{p}) + \cdots + \chi(\mathfrak{p}^{\nu_{\mathfrak{p}}(I)})\right),$$

  and χ outputs to $\{-1,0,1\}$, so we see that $b(I)$ is a nonnegative integer always. Furthermore, $b(I)$ is nonzero when I is a square (because each factor is nonzero), so we see that $L(s) \geq \zeta_K(2s)$. For example, the pole at $s = 1$ for $\zeta_K(2s)$ then implies that $L(s)$'s abscissa of holomorphy cannot go past $\{s : \operatorname{Re} s = 1/2\}$.

  Now, because $L(s)$ has all nonnegative coefficients, its abscissa of holomorphy agrees with its abscissa of absolute convergence [RV99, Lemmas 7-29]. Thus, if $L(s,\chi)$ has a zero at $s = 1$, then $L(s)$ will succeed at being holomorphic at $s = 1$, so the abscissa of holomorphy for $L(s)$ goes all the way to $\operatorname{Re} s = 0$ by Theorem 3.77. This contradicts the previous paragraph. ■

At long last, we are ready to apply Theorem 3.59.

Theorem 3.79 (Prime ideal)**.** We have

$$\{\mathfrak{p} : \mathrm{N}\,\mathfrak{p} \le x\} \sim \frac{x}{\log x}.$$

Proof. Arguing as in Corollary 3.56, it is enough to check that the function

$$\psi_K(x) \coloneqq \sum_{\substack{\mathfrak{p} \text{ prime}, k \ge 1 \\ \mathrm{N}\,\mathfrak{p}^k \le x}} \log \mathrm{N}\,\mathfrak{p}$$

satisfies $\psi_K(x) \sim x$, for which we use Theorem 3.59. Now, Remark 3.76 explains that $-\zeta_K'/\zeta_K$ is the relevant Dirichlet series. We are now ready to run the checks of Theorem 3.59.

(i) The absolute convergence of $-\zeta'(s)/\zeta(s)$ on $\{s : \mathrm{Re}\, s > 1\}$ follows from Lemma 3.75.

(ii) Note that $\zeta_K(s)$ continues to $\{s : \mathrm{Re}\, s = 1\}$ by Theorem 3.77, and it is nonvanishing by combining Lemma 3.75 with Proposition 3.78. Thus, we achieve the continuation of $-\zeta_K'(s)/\zeta_K(s)$, and we can compute that the residue of its simple pole at $s = 1$ is 1.

(iii) To check $\psi_K(x) = O(x)$, we claim that

$$\sum_{\mathrm{N}(I) \le x} \Lambda_K(I) \overset{?}{\le} \sum_{\substack{p \text{ prime}, k \ge 1 \\ p^k \le x}} [K : \mathbb{Q}] \log p^k.$$

Indeed, it is enough to only consider I of the form $\mathfrak{p}^k$; summing over the primes p below $\mathfrak{p}$, we may upper-bound $\Lambda_K(I)$ by $\log p^k$ and then maximize the number of terms in the sum by noting that there are at most $[K : \mathbb{Q}]$ primes $\mathfrak{p}$ above p and bounding $\mathrm{N}\,\mathfrak{p}^k \le p^k$.

Now, arguing as in Corollary 3.56, one finds that $\psi_K(x) = O(x)$ now follows from $\psi(x) = O(x)$. Roughly speaking, the size of this sum is dominated by the $k = 1$ term (what is left is $O\left(\sqrt{x}(\log x)^2\right)$), and the $k = 1$ term is a constant multiple of $\psi(x)$, so we are done. ∎

3.2.3 Equidistribution

In this subsection, we will prove a few facts about equidistribution, following [Fit15, Section 2] and [Ser98, Appendix to Chapter I]. Although the term already appeals in the statement of our conjecture (Conjecture 3.17), we go ahead and provide a suitable definition. We will assume some measure theory throughout, though we remark that our measures μ will all be Radon on compact Hausdorff spaces X, so they may be thought of as continuous linear functionals on $C(X)$ by the Reisz representation theorem [Fol99, Theorem 7.2], where $C(X)$ denotes the space of complex continuous functions on X.

Definition 3.80 (equidistributed)**.** Fix a compact Hausdorff space X with a probability Radon measure μ (namely, $\mu(X) = 1$). Then a sequence $\{x_n\}_{n \ge 1}$ is *equidistriubted* with respect to μ if and only if any $f \in C(X)$ has

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(x_i) = \int_X f \, d\mu.$$

Remark 3.81. One may want to upgrade this definition from $f \in C(X)$ to $f \in L^1(X)$ or similar, but this is somewhat tricky: functions in $L^1(X)$ are well-defined up to a measure-zero subset, and it is frequently the case that countable subsets of X are measure zero. Concretely, with $X = [0, 1]$, we find that no countable sequence will equidistribute by testing against the function f which indicates this sequence!

The definition has been chosen to be quite strong, but this makes it difficult to check. As such, we pick up the following lemma.

Lemma 3.82. Fix a compact Hausdorff space X with a probability Radon measure μ. The following are equivalent for a sequence $\{x_n\}_{n\geq 1}$.

(i) The sequence $\{x_n\}_{n\geq 1}$ equidistributes.

(ii) Suppose that $F \subseteq C(X)$ is a subset of functions such that linear combinations of functions in F forms a dense subspace of $C(X)$. Then for any $f \in F$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x_i) = \int_X f \, d\mu.$$

Proof. Of course (i) implies (ii) because $F \subseteq C(X)$. For the reverse inclusion, let $V \subseteq C(X)$ denote the subset for which the conclusion holds. We know $F \subseteq V$, and we would like to show that $V = L^1(X)$. Certainly V is a subspace, and by the hypothesis of F, we see that V is a dense subspace of $L^1(X)$. Thus, for any $f \in C(X)$, we fix some $\varepsilon > 0$, and we may find $g_\varepsilon \in V$ such that $\|f - g_\varepsilon\|_\infty < \varepsilon$. Then

$$\begin{aligned}\left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x_i) - \int_X f \, d\mu \right| &\leq \left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g_\varepsilon(x_i) - \int_X g_\varepsilon \, d\mu \right| \\ &\quad + \left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} (f - g_\varepsilon)(x_i) \right| + \left| \int_X (f - g_\varepsilon) \, d\mu \right|.\end{aligned}$$

The rightmost term vanishes because $g \in V$, and the remaining terms are bounded by 2ε, which goes to 0 we send $\varepsilon \to 0^+$. ∎

In the sequel, we will be interested in the case where $X = \mathrm{Conj}(G)$ where G is some compact Hausdorff topological group; here X is given the quotient topology induced by the canonical projection $G \twoheadrightarrow X$. We quickly note that X is certainly compact, and X is Hausdorff because G is normal (and conjugacy classes are closed because they are images of certain continuous maps $G \to G$). We now note that Fourier analysis can detect equidistribution.

Lemma 3.83. Fix a compact Hausdorff topological group G with probability Haar measure μ, and set $X \coloneqq \mathrm{Conj}(X)$. The following are equivalent for a sequence $\{x_n\}_{n\geq 1}$ of X.

(i) The sequence $\{x_n\}_{n\geq 1}$ equidistributes.

(ii) For any nontrivial finite-dimensional complex irreducible continuous representation ρ, one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \operatorname{tr} \rho(x_i) = 0.$$

Proof. Quickly, we note that (ii) has $\operatorname{tr} \rho(x_i)$ well-defined because the character of a representation is well-defined up to conjugacy. Now (i) implies (ii) is immediate because $(\operatorname{tr} \circ \rho)\colon X \to \mathbb{C}$ is a continuous function.

For (ii) implies (i), we use Lemma 3.82. By (ii) above, we see that the conclusion of (ii) in Lemma 3.82 holds for each of the nontrivial irreducible characters $\operatorname{tr} \circ \rho$ of G because

$$\int_G \rho \, d\mu = 0$$

by the nontriviality of ρ. Additionally, we note that the conclusion of (ii) in Lemma 3.82 also holds for the trivial character because then everything in sight is 1. Thus, it remains to check that irreducible characters

of G form a dense subset of $C(X)$. In fact, characters are dense in $L^2(G)$ by (a corollary to) the Peter–Weyl theorem [Fol16, Proposition 5.23], so we are done. ■

Remark 3.84. Of course, one may replace the application of the Peter–Weyl theorem when it is easier to prove. For example, if G is a finite abelian group, then the relevant Fourier analysis is much easier to prove.

Example 3.85. Consider the compact abelian group $G = \mathbb{R}/\mathbb{Z}$ so that $G = X$. We claim that the sequence $\{n\alpha\}_{n\geq 0}$ equidistributes in G for any irrational $\alpha \in \mathbb{R}$.

Quickly, we note that the representations of G are one-dimensional because G is abelian. Further, we claim they all take the form $t \mapsto e^{2\pi imt}$ for $m \in \mathbb{Z}$: indeed, any character of G must lift to a character $\mathbb{R} \to \mathbb{C}^\times$, but it must land in S^1 because G is compact, so our character further lifts to a homomomorphism $\mathbb{R} \to \mathbb{R}$. Continuous homomomorphisms $\mathbb{R} \to \mathbb{R}$ are just scalars, so the claim follows upon ensuring that the induced map $\mathbb{R} \to S^1$ has $\mathbb{Z}$ in its kernel.

To conclude the proof, it is now enough to compute that any nonzero m makes

$$\sum_{n=0}^{N} e^{2\pi imn\alpha} = \frac{e^{2\pi im(N+1)\alpha} - 1}{e^{2\pi im\alpha} - 1},$$

which is $O_m(1)$ and hence $o_m(N)$.

As in the example, we remark that the condition (ii) may also be read as

$$\sum_{n=1}^{N} \operatorname{tr} \rho(x_i) = o(N),$$

so it is the sort of thing that one may hope to prove using the Weiner–Ikehara theorem (Theorem 3.59). We explain the application as follows.

Proposition 3.86 (Serre)**.** Fix a compact Hausdorff topological group G with probability Haar measure μ, and set $X := \operatorname{Conj}(X)$. Further, fix a number field K, and order the set of finite places $\mathfrak{p}$ by norm (breaking ties arbitrarily), and let $\{x_\mathfrak{p}\}_\mathfrak{p}$ be a sequence in X. Now, for each finite-dimensional complex continuous representation ρ of G, define the L-function

$$L(s, \rho) := \prod_{\mathfrak{p}} \frac{1}{\det\left(1 - \rho(x_\mathfrak{p}) \operatorname{N}\mathfrak{p}^{-s}\right)},$$

and suppose that $L(s, \rho)$ admits a non-vanishing holomorphic analytic continuation to the line $\{s : \operatorname{Re} s = 1\}$ for each nontrivial irreducible ρ. Then the sequence $\{x_\mathfrak{p}\}_\mathfrak{p}$ equidistributes in G.

Proof. We apply Lemma 3.83. Using Theorem 3.79 to count the number of prime ideals $\mathfrak{p}$ of norm less than some bound, we see that we need

$$\sum_{\operatorname{N}\mathfrak{p} \leq x} \operatorname{tr} \rho(x_\mathfrak{p}) \stackrel{?}{=} o_\rho\left(\frac{x}{\log x}\right)$$

for all nontrivial irreducible complex representations ρ of G. We go ahead and fix such a representation ρ; set $d := \dim \rho$ for brevity. We proceed in steps.

1. As in Theorem 3.67, the idea is to apply the Weiner–Ikehara theorem to the logarithmic derivative $L(s, \rho)$. The correct "twisted" prime-counting function is a little involved, so we postpone its computation for a moment. Instead, let's go ahead and compute $-L'(s, \rho)/L(s, \rho)$. For each $\mathfrak{p}$, let $\{\lambda_{\mathfrak{p}1}, \dots, \lambda_{\mathfrak{p}d}\}$

denote the eigenvalues of $\rho(x_{\mathfrak{p}})$ (counted with multiplicity), so we see that

$$\log\left(\det\left(1-\rho(x_{\mathfrak{p}})\,\mathrm{N}\,\mathfrak{p}^{-s}\right)\right)=\sum_{i=1}^{d}\log\left(1-\lambda_{\mathfrak{p}i}\,\mathrm{N}\,\mathfrak{p}^{-s}\right),$$

so taking the logarithmic derivative as in Lemma 3.64 of $L(s,\rho)$ yields

$$-\frac{L'(s,\rho)}{L(s,\rho)}=\sum_{\mathfrak{p}}\sum_{k\geq 1}\sum_{i=1}^{d}\frac{\lambda_{\mathfrak{p}i}^{k}\log\mathrm{N}\,\mathfrak{p}}{\mathrm{N}\,\mathfrak{p}^{ks}}.$$

Thus, we see that the correct weights are given by

$$\Lambda_{\rho}(I):=\begin{cases}\sum_{i=1}^{d}\lambda_{\mathfrak{p}i}^{k}\log\mathrm{N}\,\mathfrak{p} & \text{if } I=\mathfrak{p}^{k} \text{ and } k\geq 1,\\ 0 & \text{else.}\end{cases}$$

In particular, $-L'(s,\rho)/L(s,\rho)=\sum_{I\subseteq\mathcal{O}_K}\Lambda_{\rho}(I)/\,\mathrm{N}(I)^{s}$; this sum is purely formal, in the sense that one side makes sense as soon as the other does. Do note that $\Lambda_{\rho}(\mathfrak{p})=\operatorname{tr}\rho(x_{\mathfrak{p}})\log\mathrm{N}\,\mathfrak{p}$ for each prime $\mathfrak{p}$. Also, note

2. We now see that arguing as in Corollary 3.56 shows that it will be enough to check that

$$\sum_{\substack{I\subseteq\mathcal{O}_K\\ \mathrm{N}(I)\leq x}}\Lambda_{\rho}(I)\stackrel{?}{=}o_{\rho}(x).$$

This is somewhat involved, so we will provide some detail. Well, we group this sum as

$$\sum_{\substack{\mathfrak{p}\text{ prime},k>1\\ \mathrm{N}\,p^{k}\leq x}}\Lambda_{\rho}\left(\mathfrak{p}^{k}\right).$$

We now have two observations.

- We note we may discard the terms with $k\geq 2$. Because G is compact, the eigenvalues $\lambda_{\mathfrak{p}i}$ are all roots of unity, so the sum $\sum_{i=1}^{d}\lambda_{\mathfrak{p}i}^{k}$ is $O_{\rho}(1)$, so we may ignore its contribution. Now, for each prime $\mathfrak{p}$, we may rudely bound $\Lambda_{\rho}\left(\mathfrak{p}^{k}\right)$ as $\log p^{[K:\mathbb{Q}]k}$, where p is the prime under $\mathfrak{p}$. On the other hand, the number of primes $\mathfrak{p}$ with $\mathrm{N}\,\mathfrak{p}$ can be naïvely bounded by $[K:\mathbb{Q}]$, so we will do so. Thus, we see that our contribution totals to

$$[K:\mathbb{Q}]^{2}\sum_{k=2}^{\log_2 x}\sum_{p\leq x^{1/k}}\log p\leq[K:\mathbb{Q}]^{2}(\log_2 x)(\sqrt{x}\log x)$$

as in Corollary 3.56. We conclude that our hypothesis is equivalent to

$$\sum_{\mathrm{N}\,\mathfrak{p}\leq x}\Lambda_{\rho}(\mathfrak{p})\stackrel{?}{=}o_{\rho}(x).$$

- We now use Abel summation in the form of Proposition 3.55 to see

$$\sum_{\mathrm{N}\,\mathfrak{p}\leq x}\operatorname{tr}\rho(x_{\mathfrak{p}})=\frac{1}{\log x}\sum_{\mathrm{N}\,\mathfrak{p}\leq x}\Lambda_{\rho}(\mathfrak{p})+\int_{2}^{x}\left(\sum_{\mathrm{N}\,\mathfrak{p}\leq x}\Lambda_{\rho}(x)\right)\frac{1}{t(\log t)^{2}}\,dt$$

(Technically, we should stratify the sum over terms of given norm before applying Abel summation.) The left term in the right-hand side is now $o(x/\log x)$ by the hypothesis, and the right term is $o(x/\log x)$ as argued in Corollary 3.56.

3. We are now ready to complete the proof using Theorem 3.59. Here are our checks.

   (i) It is enough to check that
   $$\sum_{\mathrm{N}(I)=n} \Lambda_\rho(I) = O_\rho(n^\varepsilon)$$
   for each $\varepsilon > 0$. We may assume that I is a prime-power $\mathfrak{p}^k$. As in the previous step, we see that the contribution from $\sum_{i=1}^d \lambda_{\mathfrak{p}i}^k \log \mathrm{N}\,\mathfrak{p}$ is $O_\rho(1)$ is $O_\rho(1)$, so it has no effect. We now argue as in Theorem 3.79: the number of I with $I = \mathfrak{p}^k$ is bounded by $[K : \mathbb{Q}]$, and they only contribute $\log \mathrm{N}\,\mathfrak{p} = O(\log n)$. The result follows.

   (ii) This follows immediately from the hypothesis on $L(s, \rho)$.

   (iii) From Theorem 3.79, we already know that
   $$\sum_{\mathrm{N}(I)\le x} \Lambda_K(x) = O(x).$$
   The previous step explains that the contribution $\sum_{i=1}^d \lambda_{\mathfrak{p}i}^k \log \mathrm{N}\,\mathfrak{p}$ is $O_\rho(1)$, so we conclude. ∎

Remark 3.87. Essentially the same proof as in (i) of step 3 above shows that $\log L(s, \rho)$ converges absolutely in the region $\{s : \mathrm{Re}\, s > 1\}$, so $L(s, \rho)$ converges absolutely and is nonzero. Indeed, one merely needs to re-weight the Dirichlet series coefficient $\sum_{\mathrm{N}(I)=n} \Lambda_\rho(I)$ to undo the derivative, effectively removing a $\log n$ factor.

3.2.4 The Chebotarev Density Theorem

We now give a standard application of Proposition 3.86, to the Chebotarev desnity theorem. For any Galois extension L/K of number fields, our goal is to show that the Frobenius conjugacy classes $\mathrm{Frob}_\mathfrak{p}$ equidistribute in $\mathrm{Conj}(\mathrm{Gal}(L/K))$. In light of Proposition 3.86, we see that we are interested in the following L-functions.

Definition 3.88 (Artin L-function)**.** For a number field K, let $\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(V)$ be a finite-dimensional complex representation. Then we define the *Artin L-function*
$$L(s, \rho) \coloneqq \prod_\mathfrak{p} \frac{1}{\det\left(1 - \rho(\mathrm{Frob}_\mathfrak{p})\,\mathrm{N}\,\mathfrak{p}^{-s} | V^{I_\mathfrak{p}}\right)},$$
where $I_\mathfrak{p} \subseteq \mathrm{Gal}(\overline{K}/K)$ denotes the inertia subgroup of $\mathfrak{p}$.

Remark 3.89. Let us explain this factor. Formally, one should fix a prime $\mathfrak{P}$ of $\overline{K}$ living above $\mathfrak{p}$ (alternatively, one could choose a compatible system of primes for every subfield of $\overline{K}$), and then $I_\mathfrak{p}$ and $\mathrm{Frob}_\mathfrak{p}$ mean $I_\mathfrak{P}$ and $\mathrm{Frob}_\mathfrak{P}$, respectively. Let's check that this definition is independent of the choice of $\mathfrak{P}$: any other prime living above $\mathfrak{p}$ looks like $g\mathfrak{P}$ for some $g \in \mathrm{Gal}(\overline{K}/K)$. Then $I_{g\mathfrak{P}} = gI_\mathfrak{P}g^{-1}$ and $\mathrm{Frob}_{g\mathfrak{P}} = g\mathrm{Frob}_\mathfrak{P}g^{-1}$. Thus, we see that $v \mapsto \rho(g)v$ sends $V^{I_\mathfrak{P}} \to V^{I_{g\mathfrak{P}}}$ and sends the action of $\mathrm{Frob}_\mathfrak{P}$ to the action of $\mathrm{Frob}_{g\mathfrak{P}}$. As such, the characteristic polynomials must be equal.

Example 3.90. Taking ρ to be the trivial representation, we find that $L(s, 1)$ and $\zeta_K(s)$ are equal to a finite number of Euler factors. Recall that we may write $L(s, 1) = \zeta_K(s)$.

Before going any further, we state some helpful facts about Artin L-functions.

Lemma 3.91. For a number field K, let $\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(V)$ be a finite-dimensional complex representation. Then $\ker\rho$ is open, and ρ has finite image.

Proof. We have two steps.

1. We show the following "no small subgroups" result: we claim that there is an open neighborhood $U \subseteq \mathrm{GL}(V)$ of the identity which does not contain any nontrivial subgroup. Indeed, recall that $\exp\colon \mathfrak{gl}(V) \to \mathrm{GL}(V)$ is a local diffeomorphism; say that it is a diffeomorphism on some bounded open subset $U_1 \subseteq \mathfrak{gl}(V)$, and then set $U \coloneqq \exp\left(\frac{1}{2}U_1\right)$.

   Now, suppose for the sake of contradiction that U contains a subgroup $H \subseteq \mathrm{GL}(V)$. Then note that any $\exp(x) \in H$ for $x \in \frac{1}{2}U_1$ must have $\exp(2x)$ in H and hence in U, so $2x \in \frac{1}{2}U_1$ as well; this shows that $\frac{1}{2}U_1$ is unbounded, which is a contradiction.

2. We complete the proof. Choose an open neighborhood $U \subseteq \mathrm{GL}(V)$ of the identity as in the previous step. Then $\rho^{-1}(U) \subseteq \mathrm{Gal}(\overline{K}/K)$ is an open subset, but the profinite topology of $\mathrm{Gal}(\overline{K}/K)$ promises that this open subset contains an open subgroup $H \subseteq \mathrm{Gal}(\overline{K}/K)$. Then $\rho(H) \subseteq U$ is a subgroup, which must be trivial, so we conclude that $H \subseteq \ker\rho$. Now, H is a subgroup of finite index, so we conclude the same is true for $\ker\rho$. ∎

Lemma 3.92 (additive)**.** For a number field K, let ρ_1 and ρ_2 be finite-dimensional complex representations of $\mathrm{Gal}(\overline{K}/K)$. Then $L(s, \rho_1 \oplus \rho_2) = L(s, \rho_1)L(s, \rho_2)$.

Proof. This follows because, for any $g \in G$, the characteristic polynomial of $(\rho_1 \oplus \rho_2)(g)$ is the product of the characteristic polynomials of $\rho_1(g)$ and $\rho_2(g)$. ∎

Lemma 3.93 (induction)**.** Fix a finite extension L/K of number fields. Given a finite-dimensional complex representation $\rho\colon \mathrm{Gal}(\overline{K}/L) \to \mathrm{GL}(V)$, we have

$$L(s, \rho) = L\left(s, \mathrm{Ind}_{\mathrm{Gal}(\overline{K}/L)}^{\mathrm{Gal}(\overline{K}/K)} \rho\right).$$

Proof. We follow [Neu99, Proposition VII.10.4(iv)]. Once again, we equate Euler factors over a prime $\mathfrak{p}$ of K. For psychological reasons, we come down to finite extensions. By Lemma 3.91, we may find a finite Galois extension M of K extending L such that $\mathrm{Gal}(M/L)$ is in the kernel of ρ. Then we may replace all instances of $\overline{K}$ with M without changing the value of the L-function; for example, ρ is certainly well-defined throughout.

Now, for brevity, set $G \coloneqq \mathrm{Gal}(M/K)$ and $H \coloneqq \mathrm{Gal}(M/L)$, and let $\widetilde{\rho} \coloneqq \mathrm{Ind}_H^G \rho$ denote the induction; further, set $V_\rho \coloneqq V$ and $V_{\widetilde{\rho}} \coloneqq \mathrm{Ind}_H^G V$ for clarity. We want to show that

$$\frac{1}{\det\left(1 - \widetilde{\rho}(\mathrm{Frob}_\mathfrak{p})\,\mathrm{N}\,\mathfrak{p}^{-s} | V_{\widetilde{\rho}}^{I_\mathfrak{p}}\right)} \overset{?}{=} \prod_{\mathfrak{q}|\mathfrak{p}} \frac{1}{\det\left(1 - \rho(\mathrm{Frob}_\mathfrak{q})\,\mathrm{N}\,\mathfrak{q}^{-s} | V_\rho^{I_\mathfrak{q}}\right)},$$

where $\mathfrak{q}$ varies over primes of L lying over $\mathfrak{p}$. Note that these inertia subgroups can now be brought down to automorphisms of M. We now proceed in steps.

1. We begin with a special case. Suppose that there is a single prime $\mathfrak{P}$ of M above $\mathfrak{p}$, and set $\mathfrak{q} \coloneqq \mathfrak{P} \cap L$. In this case, $G = D_\mathfrak{P}$, and we would like to show that

   $$\det\left(1 - \widetilde{\rho}(\mathrm{Frob}_\mathfrak{P})T | V_{\widetilde{\rho}}^{I_\mathfrak{P}}\right) = \det\left(1 - \rho(\mathrm{Frob}_\mathfrak{P})^{[G:H]} T^{[G:H]} | V_\rho^{I_\mathfrak{P}}\right),$$

   where T is a formal variable replacing $\mathrm{N}\,\mathfrak{p}^{-s}$. (Note that $\mathrm{N}\,\mathfrak{q} = \mathrm{N}\,\mathfrak{p}^{[G:H]}$.) Note that we may as well replace L/K with M/L, effectively allowing us to assume that H is trivial.

For psychological reasons, we explain how to reduce to the case where $I_{\mathfrak{P}}$ is trivial by adjusting the representations. On one hand, we'd like to replace V_ρ with $V_\rho^{H\cap I_{\mathfrak{P}}}$. On the other hand, note that $\left(\operatorname{Ind}_H^G \rho\right)^{I_{\mathfrak{P}}}$ can be descended to $\operatorname{Ind}_{H/(H\cap I_{\mathfrak{P}})}^{G/I_{\mathfrak{P}}} V_\rho^{H\cap I_{\mathfrak{P}}}$: a function $f\colon G\to V$ succeeds at being invariant under $I_{\mathfrak{P}}$ if and only if it descends to a function on $G/I_{\mathfrak{P}}$, and we see that we must restrict outputs to $V^{H\cap I_{\mathfrak{P}}}$ because $h\in H\cap I_{\mathfrak{P}}$ will have $f(x)=f(xh)=\rho(h)f(x)$. Thus, by taking the quotient by $I_{\mathfrak{P}}$ everywhere, we may assume that it is trivial.

Now, $D_{\mathfrak{P}}$ has become cyclic of order $f\coloneqq f(\mathfrak{P}/\mathfrak{p})$ generated by the Frobenius, so we will be able to compute $\operatorname{Ind}_1^G \rho$ relatively easily. Indeed, view $V_{\widetilde{\rho}}$ as the vector space $V^{\oplus f}$ by sending the function f to the f-tuple $\left(f(\operatorname{Frob}_{\mathfrak{P}}^i)\right)_{i=0}^{f-1}$. Then we see that $\operatorname{Frob}_{\mathfrak{P}}$ acts on $V^{\oplus f}$ by

$$(v_0,v_1,\dots,v_{f-2},v_{f-1})=\left(\rho(\operatorname{Frob}_{\mathfrak{P}})^f v_{f-1},v_0,\dots,v_{f-3},v_{f-2}\right).$$

We may now compute the determinant of $1-\widetilde{\rho}(\operatorname{Frob}_{\mathfrak{P}})T$ by commuting the determinant of the matrix

$$\begin{bmatrix} 1 & -T & 0 & \cdots & 0 & 0\\ 0 & 1 & -T & \cdots & 0 & 0\\ 0 & 0 & 1 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & -T\\ -\rho(\operatorname{Frob}_{\mathfrak{P}})^f T & 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

which we see is $\det\left(1-\rho(\operatorname{Frob}_{\mathfrak{P}})^f T^f\right)$ after some row-reduction.

2. We now return to the general case for the remainder of the proof. All decomposition, inertial, and Frobenius elements will be taken over $\mathfrak{p}$ unless otherwise specified. We begin by computing the action of $\widetilde{\rho}$. The idea is to use Mackey theory. Indeed, for fixed prime $\mathfrak{P}$ of M above $\mathfrak{p}$, we are only interested in the action of $D_{\mathfrak{P}}$ on $\operatorname{Ind}_H^G \rho$, so we note there is an isomorphism

$$\operatorname{Res}_{D_{\mathfrak{P}}}^G \operatorname{Ind}_H^G \rho \overset{?}{\cong} \bigoplus_{g\in H\backslash G/D_{\mathfrak{P}}} \operatorname{Ind}_{D_{\mathfrak{P}}\cap gHg^{-1}}^{D_{\mathfrak{P}}} \rho_g,$$

where $\rho_g(d)=\rho\left(gdg^{-1}\right)$. Let's quickly explain this. There is a forward map sending a function $f\colon G\to V$ to the tuple of functions $(f_g)_g$ where $f_g\colon D_{\mathfrak{P}}\to V$ is defined by $f_g(x)\coloneqq f(xg)$. There is also a backward map sending the tuple $(f_g)_g$ to the function $f\colon G\to V$ given by $f(hgd)\coloneqq\rho(h)f_g(d)$. These maps are G-invariant and can be checked to be $D_{\mathfrak{P}}$-invariant, so we have our isomorphism.

Thus, we see that

$$\det\left(1-\widetilde{\rho}(\operatorname{Frob}_{\mathfrak{P}})\operatorname{N}\mathfrak{p}^{-s}|V_{\widetilde{\rho}}^{I_{\mathfrak{P}}}\right)=\prod_{g\in H\backslash G/D_{\mathfrak{P}}}\det\left(1-\operatorname{Frob}_{\mathfrak{P}}\operatorname{N}\mathfrak{p}^{-s}|(\operatorname{Ind}_{D_{\mathfrak{P}}\cap gHg^{-1}}^{D_{\mathfrak{P}}}\rho_g)^{I_{\mathfrak{P}}}\right).$$

Undoing conjugation by g, we can rewrite this as

$$\det\left(1-\widetilde{\rho}(\operatorname{Frob}_{\mathfrak{P}})\operatorname{N}\mathfrak{p}^{-s}|V_{\widetilde{\rho}}^{I_{\mathfrak{P}}}\right)=\prod_{g\in H\backslash G/D_{\mathfrak{P}}}\det\left(1-\operatorname{Frob}_{g\mathfrak{P}}\operatorname{N}\mathfrak{p}^{-s}|(\operatorname{Ind}_{D_{g\mathfrak{P}}\cap H}^{D_{g\mathfrak{P}}}\rho)^{I_{g\mathfrak{P}}}\right).$$

3. We translate the product using some group theory. For this, we need to enumerate the primes of L above $\mathfrak{p}$. Note $\operatorname{Gal}(M/K)$ acts transitively on the set of primes of M above $\mathfrak{p}$, so $g\mapsto g\mathfrak{P}$ defines a bijection from $G/D_{\mathfrak{P}}$ to this set of primes. Then restricting to L, we see that $g\mapsto(g\mathfrak{P}\cap L)$ is a surjective map from $G/D_{\mathfrak{P}}$ to the set of primes in L above $\mathfrak{p}$; this map descends to $H\backslash G/D_{\mathfrak{P}}$, where we claim that it actually defines a bijection. Indeed, $(g\mathfrak{P}\cap L)=(g'\mathfrak{P}\cap L)$ implies that $g\mathfrak{P}$ and $g'\mathfrak{P}$ are both primes of M sitting above the same prime of L, so there is $h\in\operatorname{Gal}(M/L)$ such tthat $hg\mathfrak{P}=g'\mathfrak{P}$, which implies $hgD_{\mathfrak{P}}=g'D_{\mathfrak{P}}$.

Thus, we see that

$$\prod_{\mathfrak{q}|\mathfrak{p}} \det\left(1-\rho(\mathrm{Frob}_{\mathfrak{q}})\,\mathrm{N}\,\mathfrak{q}^{-s}|V_\rho^{I_{\mathfrak{q}}}\right) = \prod_{g\in H\backslash G/D_{\mathfrak{P}}} \det\left(1-\rho(\mathrm{Frob}_{g\mathfrak{P}})^{f_g}\,\mathrm{N}\,\mathfrak{p}^{-f_g s}|V_\rho^{I_{g\mathfrak{P}}}\right),$$

where $f_g = f(g\mathfrak{P}/(g\mathfrak{P}\cap L)) = [D_{g\mathfrak{P}} : D_{g\mathfrak{P}}\cap H]$ is the required inertial degree.

4. We are now ready to complete the proof. In light of the previous two steps, we would like to show that any $\mathfrak{P}'$ of M above $\mathfrak{p}$ has

$$\det\left(1-\mathrm{Frob}_{\mathfrak{P}'}T|(\mathrm{Ind}_{D_{\mathfrak{P}'}\cap H}^{D_{\mathfrak{P}'}}\rho)^{I_{\mathfrak{P}'}}\right) = \det\left(1-\rho(\mathrm{Frob}_{\mathfrak{P}'})^{[D_{\mathfrak{P}'}:D_{g\mathfrak{P}'}\cap H]}T^{[D_{\mathfrak{P}'}:D_{g\mathfrak{P}'}\cap H]}|V_\rho^{I_{\mathfrak{P}'}}\right),$$

where T is a formal variable replacing $\mathrm{N}\,\mathfrak{p}^{-s}$. Now, we note that we may define $K' \coloneqq M^{D_{\mathfrak{P}'}}$ and $L' \coloneqq M^{D_{\mathfrak{P}'}\cap H}$, whereupon we see that $\mathfrak{P}'$ is the only prime above of M the prime $\mathfrak{p}' \coloneqq \mathfrak{P}'\cap K'$ in K'. The above equality then follows from the special case in the first step applied to the extension L'/K'. ■

Remark 3.94. Given subgroups $D\subseteq H\subseteq G$ and a representation ρ of H, the above proof used the fact that

$$\mathrm{Res}_D^G\,\mathrm{Ind}_H^G\,\rho \cong \bigoplus_{\eta\in H\backslash G/D} \mathrm{Ind}_{D\cap\eta^{-1}H\eta}^{D}\,\mathrm{Ind}_{D\cap\eta^{-1}H\eta}^{D}\,\rho_\eta,$$

where $\rho_\eta(d) \coloneqq \rho(\eta d\eta^{-1})$. This fact is remarkably useful.

Example 3.95. Let L/K be a Galois extension of number fields with Galois group G. Then $\mathrm{Ind}_{\mathrm{Gal}(\overline{K}/L)}^{\mathrm{Gal}(\overline{K}/K)}1$ is the regular representation of G, so by decomposing the regular representation into irreducible representations and using Lemmas 3.92 and 3.93, we find

$$\zeta_L(s) = \prod_{\rho\in\mathrm{IrRep}(G)} L(s,\rho)^{\dim\rho},$$

where $\mathrm{IrRep}(G)$ refers to the set of irreducible representations of G.

In light of Proposition 3.86, we need to show that nontrivial irreducible ρ give $L(s,\rho)$ a non-vanishing holomorphic continuation to the line $\{s : \mathrm{Re}\,s = 1\}$. The rough idea is to use the Brauer induction theorem to reduce to the abelian case, and then the abelian case can be turned over to Hecke L-functions by class field theory.

Thus, we begin with the abelian case. As promised, this is essentially class field theory.

Proposition 3.96. Fix a number field K, and let $\rho\colon \mathrm{Gal}(\overline{K}/K)\to\mathbb{C}^\times$ be a continuous character. Then there is a continuous unitary character $\chi\colon K^\times\backslash\mathbb{A}_K^\times\to\mathbb{C}^\times$ such that

$$L(s,\rho) = L(s,\chi)$$

for s such that $\mathrm{Re}\,s>1$.

Proof. The main point is that global class field theory in the form of [Mil20a, Theorem 5.3] provides an isomorphism

$$\widehat{K^\times\backslash\mathbb{A}_K^\times} \cong \mathrm{Gal}(\overline{K}/K)^{\mathrm{ab}}.$$

With this in mind as a guide, we construct the character χ. Because the target of ρ is abelian, we see that ρ factors through $\mathrm{Gal}(\overline{K}/K)^{\mathrm{ab}} = \mathrm{Gal}(K^{\mathrm{ab}}/K)$. For convenience, we recall from Lemma 3.91 that ρ descends

to a representation of $\mathrm{Gal}(L/K)$ for some minimal finite Galois extension L/K, and once again, we find that $\mathrm{Gal}(L/K)$ is abelian. Thus, we may define χ as the composite

$$K^\times\backslash\mathbb{A}_K^\times \twoheadrightarrow K^\times \mathrm{N}_{L/K}(\mathbb{A}_L^\times)\backslash\mathbb{A}_K^\times \cong \mathrm{Gal}(L/K) \xrightarrow{\rho} \mathbb{C}^\times,$$

where the isomorphism is given by global class field theory [Mil20a, Theorem 5.3]; explicitly, on finite primes $\mathfrak{p}$ of K unramified in L, it is trivial on $\mathcal{O}_\mathfrak{p}^\times \subseteq \mathbb{A}_K^\times$ and sends a uniformizer $\varpi_\mathfrak{p} \in \mathcal{O}_\mathfrak{p}$ to $\mathrm{Frob}_\mathfrak{p} \in \mathrm{Gal}(L/K)$. By construction, χ is a continuous character, and it is unitary because ρ must output to S^1 by the compactness of $\mathrm{Gal}(\overline{K}/K)$.

We now compare the Euler factors of $L(s,\rho)$ and $L(s,\chi)$ at a prime $\mathfrak{p}$ of K. There are two cases.

- Suppose that $\mathfrak{p}$ is a prime unramified in L/K. Then we see that

  $$\det\left(1-\rho(\mathrm{Frob}_\mathfrak{p})\,\mathrm{N}\,\mathfrak{p}^{-s} \mid \mathbb{C}\right) = 1-\chi_\mathfrak{p}(\varpi_\mathfrak{p})\,\mathrm{N}\,\mathfrak{p}^{-s}$$

  by construction of χ (and properties of the global class field theory map), so we are done.

- Suppose that $\mathfrak{p}$ is a prime ramified in L/K. On one hand, ρ is nontrivial on $I_\mathfrak{p} \subseteq \mathrm{Gal}(L/K)$, so $\mathbb{C}^{I_\mathfrak{p}}$ must be zero-dimensional, so the Euler factor of $L(s,\rho)$ is 1. On the other hand, we note that $\mathrm{N}_{L/K}(\mathbb{A}_L^\times)$ does not contain $\mathcal{O}_\mathfrak{p}^\times$ by a computation of norm subgroups, so χ is nontrivial on $\mathcal{O}_\mathfrak{p}^\times$ by tracking through the global class field theory isomorphism, so the Euler factor of $L(s,\chi)$ is also 1. ■

Remark 3.97. Because a Dirichlet series is uniquely determined by its coefficients, we see that the character χ is uniquely determined by ρ. However, this is not a bijection: the disagreement between the topologies of $K^\times\backslash\mathbb{A}_K^\times$ and $\mathrm{Gal}(\overline{K}/K)^{\mathrm{ab}}$ means that there are many more continuous unitary characters $K^\times\backslash\mathbb{A}_K^\times \to \mathbb{C}^\times$.

Corollary 3.98. Fix a number field K, and let $\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathbb{C}^\times$ be a nontrivial continuous character. Then $L(s,\rho)$ admits a nonvanishing holomorphic continuation to $\{s : \mathrm{Re}\, s = 1\}$.

Proof. Note $L(s,\rho)$ is already holomorphic and nonvanishing on $\{s : \mathrm{Re}\, s > 0\}$ by Remark 3.87. Now, construct the continuous unitary character $\chi\colon K^\times\backslash\mathbb{A}_K^\times \to \mathbb{C}^\times$ as in Proposition 3.96 so that $L(s,\chi) = L(s,\rho)$. The nonvanishing now follows from Proposition 3.78. Lastly, the continuation follows from Theorem 3.77 as soon as we check that $\chi|\cdot|^{1+it}$ is never trivial on all unramified primes. This follows by the nontriviality of ρ, which requires there to be an unramified prime $\mathfrak{p}$ where $\rho(\mathrm{Frob}_\mathfrak{p}) \neq 1$; this corresponds to the needed fact about χ. ■

We are now in a position to prove equidistribution of Frobenius elements in $\mathrm{Gal}(K^{\mathrm{ab}}/K)$, from which one can prove the general case by a clever reduction argument. However, we will be honest to our discussion of equidistribution and prove nonvanishing holomorphic continuation to $\{s : \mathrm{Re}\, s = 1\}$ for $L(s,\rho)$ for all nontrivial irreducible continuous representations ρ.

The idea is to write ρ as a "linear combination" of inductions of characters. Then the result will follow from the abelian case combined with our properties about L-functions. One almost achieves the full holomorphic nonvanishing as well, but it would be technically possible to see the trivial character in our linear combination, thus possibly introducing a pole or zero.

Of course, it does not a priori make sense to talk about linear combination of representations, so we must pass to their linearization: virtual characters. Thus, we will want to define the Artin L-function of a class function. To motivate, use Lemma 3.91 to descend ρ to some representation $\mathrm{Gal}(L/K) \to \mathbb{C}^\times$. For some $g \in \mathrm{Gal}(L/K)$, we let $\lambda_1,\dots,\lambda_d$ denote the eigenvalues of g (with algebraic multiplicities), so we see that

$$\log\left(\det(1-\rho(g)T)^{-1}\right) = \sum_{i=1}^{d} -\log(1-\lambda_i T).$$

Now, expanding out the Taylor series reveals that

$$\frac{1}{\det(1-\rho(g)T)} = \exp\left(\sum_{k=1}^{\infty} \frac{\operatorname{tr}\rho\left(g^k\right)T^k}{k}\right).$$

We are now ready to make the following definition.

Definition 3.99 (Artin L-function)**.** Fix a Galois extension L/K of number fields with Galois group G. For a class function $\chi\colon G \to \mathbb{C}$, we define the *Artin L-function* as

$$L(s,\chi) \coloneqq \prod_{\mathfrak{p}\text{ unr.}} \exp\left(\sum_{k=1}^{\infty} \frac{\chi(\operatorname{Frob}_{\mathfrak{p}}^k)}{k\,\mathrm{N}\,\mathfrak{p}^{-s}}\right),$$

where the product is taken over primes of K unramified in L.

Example 3.100. The discussion preceding the definition shows that $L(s,\rho) \doteq L(s,\operatorname{tr}\circ\rho)$ for any finite-dimensional complex representation $\rho\colon \operatorname{Gal}(L/K) \to \mathrm{GL}(V)$.

Remark 3.101. A notable defect of this definition is that we have not defined our Euler factors at ramified primes. This will cause us to use some $\doteq$s in the sequel; this is no issue because finitely many Euler factors will not change holomorphy or nonvanishing.

Here are the standard properties of these L-functions, which are carried over from our previous discussion.

Lemma 3.102. Fix a Galois extension M/K of number fields with Galois group G.

(a) If $\chi\colon G \to \mathbb{C}$ is a class function, then $L(s,\chi)$ converges absolutely to a nonvanishing holomorphic function in the region $\{s : \operatorname{Re} s > 1\}$.

(b) Additive: if $\chi_1,\chi_2\colon G \to \mathbb{C}$ are class functions, then $L(s,\chi_1+\chi_2) = L(s,\chi_1)L(s,\chi_2)$.

(c) Inflation: let L/K be a Galois subextension such that $\operatorname{Gal}(M/L) = H$. If $\chi\colon G/H \to \mathbb{C}$ is a class function, then $L(s,\chi) \doteq L(s,\widetilde{\chi})$, where $\widetilde{\chi}\colon G \to \mathbb{C}$ is the induced class function.

Proof. Here, (a) follows as in Remark 3.87 by noting that the series expansion for $\log L(s,\chi)$ absolutely converges to a finite value; notably, G is finite, so χ is bounded, so it does not meaningfully contribute. Continuing, (b) follows by a direct expansion of the Euler product, and (c) follows because the Euler factors are exactly the same for any prime $\mathfrak{p}$ of K unramified in M (and hence also unramified in K). ∎

The suitable analogue of Lemma 3.93 on induction remains true, but we will not need it in the full generality of complex class functions. However, we do need to know how to induct character.

Notation 3.103. Fix a subgroup H of a finite group G. Given a class function $\chi\colon H \to \mathbb{C}$, define the induced class function

$$\operatorname{Ind}_H^G \chi(g) \coloneqq \frac{1}{|H|} \sum_{\substack{\eta\in G \\ \eta g \eta^{-1}\in H}} \chi\left(\eta g \eta^{-1}\right).$$

Lemma 3.104. Fix a subgroup H of a finite group G. Given a finite-dimensional representation $\rho\colon H \to \mathrm{GL}(V)$, then we check that $\operatorname{tr}\circ\operatorname{Ind}_H^G\rho = \operatorname{Ind}_H^G(\operatorname{tr}\circ\rho)$.

Proof. We will use many of the same tricks appearing in Lemma 3.93. Fix some $g \in G$, and we would like to check the result at g. We proceed in steps.

1. We begin with the special case where G is cyclic and generated by g. If $H = G$, there is nothing to do. Otherwise, if $H \neq G$, then $\operatorname{Ind}_H^G(\operatorname{tr}\circ\rho)(g)$ is an empty, so we must show $\operatorname{tr}\operatorname{Ind}_H^G \rho(g)$ vanishes. Well, view elements $\operatorname{Ind}_H^G V$ as sequences of vectors $\{v_{Hg'}\}$ indexed by $H\backslash G$, and then we see that $\operatorname{Ind}_G^H \rho(g)$ acts by a (generalized) permutation matrix which is a sum of nontrivial cycles of length $[G : H]$. Thus, this operator has no trace.

2. We now show the general case. By Remark 3.94, we see that

$$\operatorname{Res}_{\langle g\rangle}^G \operatorname{Res}_H^G \rho \cong \bigoplus_{\eta\in H\backslash G/\langle g\rangle} \operatorname{Ind}_{\langle g\rangle\cap\eta^{-1}H\eta}^{\langle g\rangle} \rho_\eta,$$

   where $\rho_\eta(g') \coloneqq \rho\left(\eta g'\eta^{-1}\right)$. Thus, we see that

$$\operatorname{tr}\operatorname{Ind}_H^G \rho(g) = \sum_{\eta\in H\backslash G/\langle g\rangle} \operatorname{tr}\operatorname{Ind}_{\langle g\rangle\cap\eta^{-1}H\eta}^{\langle g\rangle} \rho_\eta\left(g\right).$$

   Now, by the previous case, we see that terms vanish as long as $g \notin \eta^{-1}H\eta$; on the other hand, if $g \in \eta^{-1}H\eta$, then we get a contribution of $\operatorname{tr}\rho\left(\eta g\eta^{-1}\right)$, so we see

$$\operatorname{tr}\operatorname{Ind}_H^G \rho(g) = \sum_{\substack{\eta\in H\backslash G/\langle g\rangle \\ \eta g\eta^{-1}\in H}} \operatorname{tr}\rho\left(\eta g\eta^{-1}\right).$$

   The result now follows by replacing the sum over $H\backslash G/\langle g\rangle$ with a sum over G. ∎

In order to allow us to stop talking about L-functions as quickly as possible, let's go ahead and explicate the inductive approach to meromorphic continuation via Brauer's theorem. We begin with the following non-standard definition.

Definition 3.105 (Brauer)**.** Fix a finite group G. Then G is *Brauer* if and only if, for any finite-dimensional complex irreducible representation ρ, there is a sequence of pairs $\{(a_i, H_i, \psi_i)\}_{i=1}^n$ where $a_i \in \mathbb{Z}$ and $H_i \subseteq G$ is a subgroup and $\psi_i\colon H_i \to \mathbb{C}^\times$ is a representation such that

$$\operatorname{tr}\circ\rho = \sum_{i=1}^n a_i \operatorname{Ind}_{H_i}^G \psi_i$$

as virtual character. A representation of the form $\operatorname{Ind}_{H_i}^G \psi_i$ is said to be *monomial*.

Lemma 3.106. Fix a Galois extension L/K of number fields with Galois group G. Suppose that G is Brauer. For any finite-dimensional complex representaion ρ of G, the function $L(s,\rho)$ admits a meromorphic continuation to $\{s : \operatorname{Re} s = 1\}$ with no poles or zeroes except possibly a pole or zero at $s = 1$. Further, the order of the pole at $s = 1$ is $\langle \operatorname{tr}\circ\rho, 1\rangle$.

Proof. By the additivity of Lemma 3.92, we may assume that ρ is irreducible. By Example 3.100, it is enough to check the result for $L(s, \operatorname{tr}\circ\rho)$. Because G is Brauer, we receive an expansion $\operatorname{tr}\circ\rho = \sum_{i=1}^n a_i \operatorname{Ind}_{H_i}^G \psi_i$ of $\operatorname{tr}\circ\rho$ into a $\mathbb{Z}$-linear combination of inductions of characters, which implies that

$$L(s, \operatorname{tr}\circ\rho) = \prod_{i=1}^n L\left(s, \operatorname{Ind}_{H_i}^G \psi_i\right)^{a_i}$$

by Lemma 3.102. By Example 3.100, we may now think of each $L\left(s, \operatorname{Ind}_{H_i}^G \psi_i\right)$ as an Artin L-function of a representation (up to finitely many Euler factors), so Lemma 3.93 tells us that this L-function is $L(s, \psi_i)$.

The meromorphic continuation now essentially follows from Corollary 3.98, which tells us each nontrivial ψ_i grants a nonvanishing holomorphic continuation of $L(s,\psi_i)$ to $\{s : \operatorname{Re} s \geq 1\}$. Note the same is true for trivial ψ_i except at the point $s = 1$ where we find a pole in $L(s,\psi_i)$ because this is a Dedekind ζ-function by Example 3.90; see Theorem 3.77 and proposition 3.78. Taking the appropriate product of these contributions proves the statement.

It remains to prove the last sentence. This will require a trick. On one hand, by the discussion in the previous paragraph, we see that the order of the pole is

$$\sum_{\substack{1\leq i\leq n\\ \psi_i = 1_{H_i}}} a_i.$$

On the other hand, we see $\langle \operatorname{tr}\circ\rho, 1\rangle$ equals

$$\sum_{i=1}^{n} a_i \left\langle \operatorname{Ind}_{H_i}^{G} \psi_i, 1_G \right\rangle = \sum_{i=1}^{n} a_i \langle \psi_i, 1_{H_i}\rangle$$

by Frobenius reciprocity. The last sentence now follows. ■

Thus, we will achieve our nonvanishing holomorphi continuation as soon as we check that all finite groups are Brauer; we will complete the nonvanishing later by a careful analysis of $s = 1$.

Our current goal is to prove Brauer's theorem that all finite groups are Brauer; our exposition follows [Ser77, Chapter 10]. We begin by creating a large supply of Brauer groups.

Lemma 3.107. Let G be a finite nilpotent group. Then G is Brauer.

Proof. We induct on $|G|$. For our base case, we note that if G is already abelian (for example, $|G| = 1$), then there is nothing to do because all irreducible representations are already one-dimensional.

Thus, for our induction, we may assume that G is nonabelian, and we fix some complex irreducible representation $\rho\colon G \to \mathrm{GL}(V)$ of G. Because taking induction commutes with taking quotients, we may replace G with $G/\ker\rho$, effectively allowing us to assume that ρ is injective. We will show directly that ρ can be induced from a character, which will complete the proof; we proceed in steps.

1. We claim that there is an abelian normal subgroup $N \subseteq G$ strictly containing $Z(G)$. This follows quickly because G is nilpotent: because G is nonabelian and nilpotent, we see that $G/Z(G)$ is nontrivial and has nontrivial center, so we let $N \subseteq G$ be the pre-image of the center. Then N strictly contains $Z(G)$ and is normal because it is the pre-image of a normal subgroup along a surjective homomorphism.

2. We now decompose $\operatorname{Res}_N^G \rho$ into irreducibles as

$$\operatorname{Res}_N^G \rho = \bigoplus_{\psi\in\operatorname{Hom}(N,\mathbb{C}^\times)} V^\psi,$$

   where $V^\psi \subseteq V$ denotes the ψ-eigenvectors of V. (The sum is over the characters of N; this decomposition exists because N is abelian.) Now, because $N \subseteq G$ is normal, we know that each of the spaces $\rho(g)V^\psi \subseteq V$ continues to be N-invariant and in fact will be N-isotypic. Thus, we see that G acts on the collection $\{V^\psi\}_\psi$, and it must act transitively because the span of the G-orbit of some V^ψ will be a G-subrepresentation of the irreducible representation ρ.

3. We claim that $\operatorname{Res}_N^G \rho$ is not isotypic. This is by the construction of N: this would imply that N acts by scalars on V, thereby implying that $\rho(N)$ commutes with $\rho(G)$, thereby giving $N \subseteq Z(G)$ because ρ is faithful. This contradicts the construction of N as strictly containing $Z(G)$.

4. We now complete the proof. Choose some $\psi_0 \in \operatorname{Hom}(N,\mathbb{C}^\times)$, and let $G_0 \subseteq G$ be the stabilizer of the action given in the second step. Then ρ restricts to a representation $\rho_0\colon G_0 \to \mathrm{GL}(V_0)$ where $V_0 \coloneqq V^{\psi_0}$. Now, we claim that $\rho = \operatorname{Ind}_{G_0}^G \rho_0$, which will complete the proof because G_0 is a strictly smaller nilpotent group than G. Well, for the isomorphism, view $\operatorname{Ind}_{G_0}^G \rho_0$ as $\mathbb{C}[G]\otimes_{\mathbb{C}[G_0]} \rho_0$, and then define the map $\operatorname{Ind}_{G_0}^G \rho_0 \to \rho$ by sending $g\otimes v_0$ to gv_0. ■

Thus, to show that any group G is Brauer, one may simply show that any virtual character for an irreducible complex representation is a $\mathbb{Z}$-linear combination of ones induced from nilpotent subgroups. Now that we are working with virtual characters than one-dimensional ones, we pick up the following notation.

Definition 3.108 (virtual character)**.** Fix a finite group G. Then we let $R(G)$ denote the free $\mathbb{Z}$-module of class functions $G \to \mathbb{C}$ generated by the virtual characters $\operatorname{tr} \circ \rho$ as ρ varies over finite-dimensional complex representations of G. One frequently calls $R(G)$ the *ring of virtual characters*.

Remark 3.109. By taking tensor products of representations, we see that $R(G)$ is a subring of the set of functions $G \to \mathbb{C}$. By induction and restriction of representations, we see that Res^G_H and Ind^G_H induce ring homomorphisms $R(G) \to R(H)$ and $R(H) \to R(G)$, respectively.

Thus, to check that a group G is Brauer, it will be enough to show that the map

$$\operatorname{Ind}\colon \bigoplus_{\substack{H \subseteq G \\ H \text{ Brauer}}} R(H) \to R(G)$$

is surjective, for any element of one of the $R(H)$s can be expanded into a sum of virtual characters induced from linear characters. For example, we will eventually show that one can restrict this direct sum to nilpotent subgroups.

Remark 3.110. While we're here, we remark that one can check the surjectivity of this map after tensoring with any free $\mathbb{Z}$-module because this essentially takes both sides to a finite power. In particular, in the sequel, we will frequently work with $R(G)_{\mathbb{Z}[\zeta_n]}$ where $n = |G|$, which is conveneint because the functions in $R(G)$ output to the ring $\mathbb{Z}[\zeta_n]$. (Indeed, for any $g \in G$ and representation ρ, because $g^n = 1$, the eigenvalues of g are all nth roots of unity, so $\operatorname{tr} \rho(g) \in \mathbb{Z}[\zeta_n]$.)

Most of our work in eventually proving that all finite groups are Brauer will come from a construction of many virtual characters. We begin with a couple preliminary lemmas.

Lemma 3.111. Fix a finite group G of order n, and choose a class function $f\colon G \to \mathbb{Z}[\zeta_n]$. Then nf is in the image of the map

$$\operatorname{Ind}\colon \bigoplus_{\substack{H \subseteq G \\ H \text{ cyclic}}} R(H)_{\mathbb{Z}[\zeta_n]} \to R(G)_{\mathbb{Z}[\zeta_n]}.$$

Proof. The proof has two steps.

1. We show that n is in the image of the given map. For each cyclic subgroup $H \subseteq G$, define $\theta_H\colon H \to \mathbb{Z}$ as $|H|$ times the indicator function of generating H. Then we claim that

$$n \overset{?}{=} \sum_{\substack{H \subseteq G \\ H \text{ cyclic}}} \operatorname{Ind}^G_H \theta_H.$$

   Well, for any $g \in G$, we begin by computing $\operatorname{Ind}^G_H \theta_H(g)$ as

$$\frac{1}{|H|} \sum_{\substack{\eta \in G \\ \eta g \eta^{-1} \in H}} \theta_H\left(\eta g \eta^{-1}\right) = \left|\{\eta \in G : \eta g \eta^{-1} \text{ generates } H\}\right|.$$

   Now, upon summing over all H, we see that each $\eta g \eta^{-1}$ surely generates exactly one cyclic subgroup H, so the claim follows.

2. We complete the proof. By the previous step, we see that nf equals

$$\left(\sum_{\substack{H \subseteq G \\ H \text{ cyclic}}} \operatorname{Ind}_H^G \theta_H \right) f = \sum_{\substack{H \subseteq G \\ H \text{ cyclic}}} \operatorname{Ind}_H^G (\theta_H f|_H),$$

so we will be done as soon as we check that $\theta_H f|_H \in R(H)_{\mathbb{Z}[\zeta_n]}$. Well, because H is cyclic, orthogonality of characters permits to merely check that $\langle \theta_H f|_H, \psi \rangle \in \mathbb{Z}[\zeta_n]$ for any character $\psi \colon H \to \mathbb{C}^\times$; however, this follows by a direct expansion of the inner product because ψ outputs to $\mathbb{Z}[\zeta_n]$ and $f|_H$ outputs to $|H|\,\mathbb{Z}[\zeta_n]$. ■

For the next lemma, we need a piece of notation.

Notation 3.112. Fix a finite group G of order n, and fix a prime p. Choose $g \in G$, whose order we write as $\operatorname{ord}(g) = mp^\nu$ where $p \nmid m$. Then we may find integers x and y such that $xm + yp^\nu = 1$. Now, for any $g \in G$, we define $g_p \coloneqq g^{xm}$ (which has prime-power order) and $g_p' \coloneqq g^{yp^\nu}$ (which has order coprime to p) so that $g = g_p g_p'$.

Lemma 3.113. Fix a finite group G of order n. For any class function $f \colon G \to \mathbb{Z}[\zeta_n]$ in $R(G)_{\mathbb{Z}[\zeta_n]}$, we have

$$f(g) \equiv f(g_p') \pmod{p\mathbb{Z}[\zeta_n]}$$

for any $g \in G$ and prime p.

Proof. Because we are only interested in the values of f on powers of g, we may as well work with $f|_{\langle g \rangle}$. Now, because $\langle g \rangle$, we may write $f|_{\langle g \rangle}$ as $\mathbb{Z}[\zeta_n]$-linear combination of linear characters $\langle g \rangle \to \mathbb{C}^\times$. Notably, the conclusion is $\mathbb{Z}[\zeta_n]$-linear in f, so we may as well assume that $f|_{\langle g \rangle}$ is a linear character $\langle g \rangle \to \mathbb{C}^\times$.

Now, recall that g_p' is g^{yp^ν} where the order of g equals mp^ν for $p \nmid m$ and $x, y \in \mathbb{Z}$ satisfy $xm + yp^\nu = 1$. Thus, $f(g)$ will be an mp^νth root of unity, so it will be enough to check that

$$\zeta \equiv \zeta^{yp^\nu} \pmod{p\zeta[\zeta]},$$

where ζ is a primitive mp^νth root of unity. Well, by the Frobenius automorphism, it is enough to check that sufficiently large pth powers of both sides are equal. For this, note that $p^\nu \equiv (1+y)p^\nu \pmod{mp^\nu}$, so it is enough to take p^νth powers. ■

Now is as good a time as any to begin our main argument.

Theorem 3.114 (Brauer)**.** Let G be a finite group. Then G is Brauer.

Proof. Let n be the order of n. The idea is to show that the map

$$\operatorname{Ind}_{\mathbb{Z}[\zeta_n]} \colon \bigoplus_{\substack{H \subseteq G \\ H \text{ nilpotent}}} R(H)_{\mathbb{Z}[\zeta_n]} \to R(G)_{\mathbb{Z}[\zeta_n]}$$

is surjective, which completes the proof. Namely, one can undo the base-change by $\mathbb{Z}[\zeta_n]$ because $\mathbb{Z}[\zeta_n]$ is a free $\mathbb{Z}$-algebra of finite rank; then one merely notes that any complex irreducible representation ρ of G can be written as a linear combination of inductions $R(H)$ for nilpotent subgroups $H \subseteq G$, but anything in $R(H)$ is induced from a linear character by Lemma 3.107. In fact, we will find that we may merely consider H which are the product of a cyclic group and a p-group.

As always, we proceed in steps.

1. To ground ourselves, we note that it is enough to check that 1 is in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$. Indeed, it is then enough to check that the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$ is an ideal, for which it is enough to check that $\mathrm{Ind}_H^G R(H) \subseteq R(G)$ is an ideal for any subgroup $H \subseteq G$. Well, for any $f_H \in R(H)$ and $f_G \in R(G)$, we see that $f_G \mathrm{Ind}_H^G f_H = \mathrm{Ind}_H^G(f_G|_H f_H)$.

   Our proof will eventually build a small supply of constant functions in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$, which will produce the constant function 1 by taking suitable linear combinations.

2. We proceed with the key construction. Fix a prime p, and choose $x \in G$ of order coprime to p. Then we claim that there is $f\colon G \to \mathbb{Z}$ in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$ such that $f(x) \not\equiv 0 \pmod p$ while $f(y) = 0$ of y has order coprime to p and is not conjugate to x.

   In fact, we will induce f directly from the subgroup $H = \langle x\rangle \times P$, where P is a Sylow p-subgroup of the centralizer $C(x) \subseteq G$. (Note that H is nilpotent because it is the product of nilpotent groups. Also, H is in fact a subgroup because $\langle x\rangle$ has cardinality coprime to p, thus making the induced map $\langle x\rangle \times P \to G$ an injective homomorphism.) We now define $f_H\colon H \to G$ by

$$f_H\left(x^i y\right) \coloneqq \begin{cases} |\langle x\rangle| & \text{if } x^i = x, \\ 0 & \text{else,} \end{cases}$$

   for any $x^i \in \langle x\rangle$ and $y \in P$. We quickly check that $f_H \in R(H)$: note that $f_H = f_H \circ \mathrm{pr}_{\langle x\rangle}$, so it is enough to check that $f_H|_{\langle x\rangle}$, which follows from Lemma 3.111

   It remains to check that $f \coloneqq \mathrm{Ind}_H^G f_H$ satisfies the required conditions. For example, of course f is in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$ and defines a function $G \to \mathbb{Z}$. Before doing anything else, we remark on the condition $\eta g \eta^{-1} \in H$ for $g, \eta \in G$. Namely, if g has order coprime to p, then $\eta g \eta^{-1} \in H$ continues to have order coprime to p; thus, by writing it out as $x^i y$ for $y \in P$, we find that we must have $\eta g \eta^{-1} \in \langle x\rangle$. In particular, to have $f_H\left(\eta g \eta^{-1}\right) \neq 0$, we must have $\eta g \eta^{-1} = x$ on the nose!

   - Using the previous paragraph, we compute $f(x)$ as

$$\frac{1}{|\langle x\rangle| \cdot |P|} \sum_{\substack{\eta \in G \\ \eta x \eta^{-1} = x}} |\langle x\rangle|.$$

     This sum is now $|C(x)| / |P|$, which is coprime to p because $P \subseteq C(x)$ is a Sylow p-subgroup.
   - Again using the paragraph preceding our checks, we see that any g of order coprime to p must have g conjugate to x in order for the sum $\mathrm{Ind}_H^G f_H(g)$ to have any nonzero terms. We conclude that $f(g) = 0$ when g has order coprime to p but is not conjugate to x.

3. Fix a prime p. Then we claim that there is $f\colon G \to \mathbb{Z}$ in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$ such that $p \nmid f(x)$ for all $x \in G$. Quickly, we note that Lemma 3.113 allows us to merely check the conclusion for $x \in G$ of order coprime to p.

   Now, let $X \subseteq G$ be a set of representatives of the conjugacy classes of the elements of G with order coprime to p. Then for each $x \in X$, we construct $f_x\colon G \to \mathbb{Z}$ in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$ via the previous step. Then we define

$$f \coloneqq \sum_{x \in X} f_x.$$

   We now check that f works. Certainly f is in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$. Further, the construction of the f_xs means that any $y \in G$ of order coprime to p will produce a nonzero contribution $\pmod p$ at exactly one summand (namely, the $x \in X$ conjugate to y).

4. We complete the proof. For each prime p, we factor $n = mp^\nu$ where $p \nmid m$; then we claim that m is in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$. By letting p vary over the prime factors of n, this allows us to conclude that 1 is in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$ by taking $\mathbb{Z}$-linear combinations, thereby completing the proof.

   Now fix a prime p. The previous step provides $f\colon G \to \mathbb{Z}$ such that $f(x) \not\equiv 0 \pmod p$ for all $x \in G$. By replacing f with a suitably large power (which we may do because the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$ is an ideal), we

may achieve that $f \equiv 1 \pmod{p^\nu}$ for all $x \in G$. Then $mf - m$ is a class function $G \to \mathbb{Z}$ with values divisible by n, so Lemma 3.111 tells us it is in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$. We are now allowed to conclude m is in the image of $\mathrm{Ind}_{\mathbb{Z}[\zeta_n]}$. ■

At long last, we may prove the Chebotarev density theorem.

Theorem 3.115 (Chebotarev density)**.** Fix a number field K. For each prime $\mathfrak{p}$ of K, choose a prime $\mathfrak{P}$ of $\overline{K}$ above $\mathfrak{p}$, and let $x_\mathfrak{p}$ be the conjugacy class of $\mathrm{Frob}_\mathfrak{P}$ in $\mathrm{Gal}(\overline{K}/K)$. Then the sequence $\{x_\mathfrak{p}\}_\mathfrak{p}$ equidistributes in $\mathrm{Conj}(\mathrm{Gal}(\overline{K}/K))$.

Proof. By Proposition 3.86, it is enough to check that the L-functions $L(s, \rho)$ have nonvanishing holomorphic continuation to $\{s : \mathrm{Re}\, s \geq 1\}$ for each nontrivial complex irreducible representation ρ of $\mathrm{Gal}(\overline{K}/K)$. Well, fix some such ρ. By Lemma 3.91, we can find a finite Galois extension L of K with Galois group G such that ρ descends to $\mathrm{Gal}(L/K)$. Now, G is Brauer by Theorem 3.114, so Lemma 3.106 provides a meromorphic continuation to $\{s : \mathrm{Re}\, s \geq 1\}$ which is holomorphic and nonvanishing for $s \neq 1$. Furthermore, $\langle \mathrm{tr} \circ \rho, 1\rangle = 0$ because ρ is nontrivial and irreducible, so holomorphy and nonvanishing follows. ■

3.2.5 Abelian Varieties with Complex Multiplication

For this subsection, we will let A be an abelian variety of dimension g defined over a number field K with complex multiplication by an order $\mathcal{O}$ of a CM number field E. We will prove the Sato–Tate conjecture for A.

Lemma 3.116. Fix an abelian variety A over a number field K satisfying Conjecture 3.19. For any representation r of $\mathrm{ST}(A)$, there is an ℓ and an algebraic extension $\widetilde{r}$ to $G_\ell(A)$ and an integer $w \in \mathbb{Z}$ such that $\widetilde{r}(t) = t^w$ (for scalars t) and

$$L(s - w/2, \widetilde{r} \circ \rho_{A,\ell}) \doteq \prod_\mathfrak{p} \frac{1}{\det\left(1 - r\left(\iota\rho_{A,\ell}(\mathrm{Frob}_\mathfrak{p})\, \mathrm{N}\,\mathfrak{p}^{-1/2}\right) \mathrm{N}\,\mathfrak{p}^{-s}\right)}.$$

Proof. This is [Joh17, p. 6315]. Note r is a continuous representation of the compact Lie group $\mathrm{ST}(A)$, so it upgrades to an algebraic representation of $G_\ell^1(A)_\iota$ and hence of $G_\ell^1(A)$. (This algebraic representation will descend to $\mathbb{Q}_\ell$ for some ℓ.) Now, $G_\ell(A) = \mathbb{G}_{m,\mathbb{Q}_\ell} \cdot G_\ell^1(A)$, so extending r is a matter of making sure $\widetilde{r}$ is well-defined on $\mathbb{G}_{m,\mathbb{Q}_\ell} \cap G_\ell^1(A)$. However, this is some finite subgroup of $\mathbb{G}_{m,\mathbb{Q}_\ell}$, say μ_n, so we merely need to select $w \in \mathbb{Z}$ so that $r(\zeta_n) = \zeta_n^w$. The final equality now follows by a direct expansion. ■

Theorem 3.117. Let A be an abelian variety over a number field K with complex multiplication by a CM algebra E. Then the Sato–Tate conjecture is true for A.

Proof. We proceed in steps.

1. By Proposition 3.86, it is enough to check that

$$\prod_\mathfrak{p} \frac{1}{\det\left(1 - r\left(\iota\rho_{A,\ell}(\mathrm{Frob}_\mathfrak{p})\, \mathrm{N}\,\mathfrak{p}^{-1/2}\right) \mathrm{N}\,\mathfrak{p}^{-s}\right)}$$

admits a non-vanishing holomorphic analytic continuation to the region $\{s : \mathrm{Re}\, s \geq 1\}$ for all nontrivial irreducible representation r of $\mathrm{ST}(A)$. Now, Lemma 3.116 allows us to extend r to an algebraic representation $\widetilde{r}$ of $G_\ell(A)$ such that there is an integer $w \in \mathbb{Z}$ for which $\widetilde{r}(t) = t^w$ on scalars. Then we are tasked with checking that $L(s - w/2, \widetilde{r} \circ \rho_{A,\ell})$ admits a non-vanishing holomorphic analytic continuation to the region $\{s : \mathrm{Re}\, s \geq 1 + w/2\}$.

2. We remark that what makes this case achievable is that Remark 2.130 explains that $G_\ell(A) \subseteq \mathrm{T}_E$ is just a torus, so $\widetilde{r}$ is just a character $G_\ell(A) \to \mathbb{G}_{m,\mathbb{Q}_\ell}$.[4] Thus, we are interested in finding a Hecke character which understands $\rho_{A,\ell}$. This is necessarily a little tricky because $\rho_{A,\ell}$ is ℓ-adic, but Hecke characters are archimedean.

   We will use the Fundamental theorem of complex multiplication, following [Con04, Theorem 3.7]. In particular, we would like to apply Proposition 2.143. Quickly, let's reduce to the case where K is a CM extension of E. Observe that almost all primes of K are totally split over $\mathbb{Q}$, so the Sato–Tate conjecture for a subfield of K will imply it for K: indeed, that density-1 subset of totally split primes does not change Frobenius upon restriction to the smaller field. But now, there is an abelian variety isogenous to A defined over the reflex field of E, so we may take K to be that CM field. (Note isogenies induce isomorphisms on the level of the Galois representation.)

   We may now apply the proof of Proposition 2.143: Theorem 2.138 provides a continuous character $\lambda\colon \mathbb{A}_{K,f}^\times \to E^\times$, and then we know there is a suitable reflex norm $\mathrm{N}_{\Phi^*}\colon \mathrm{T}_K \to \mathrm{T}_E$ such that

$$\rho_A(\mathrm{Art}_K\, s_f) = \lambda(s_f)\,\mathrm{N}_{\Phi^*}(s_f)^{-1}$$

   for any $s_f \in \mathbb{A}_{K,f}^\times$. This is currently valued in $\mathbb{A}_{E,f}^\times$, so to value it in archimedean places, we define $\chi\colon \mathbb{A}_K^\times \to \mathbb{A}_{E,\infty}^\times$ by

$$\chi(s) \coloneqq \lambda(s_f)\,\mathrm{N}_{\Phi^*}(s_\infty)^{-1}.$$

   Note χ is continuous by construction. Further, for $t \in K^\times$, we already knew that $\mathrm{Art}_K\, t = \mathrm{id}$, so $\lambda(t) = \mathrm{N}_{\Phi^*}(t)$, so we continue to have $K^\times \subseteq \ker\chi$. The point is that we have built a Hecke character χ which keeps track of our Galois information in λ.

3. With χ in hand, we may complete the proof. Importantly, Proposition 2.143 explains that χ actually factors through $G_\ell(A)(\mathbb{A}_E)$, so $\widetilde{r}\circ\chi$ is a well-defined Hecke character $K^\times\backslash\mathbb{A}_K^\times \to \mathbb{C}^\times$. Because $\rho_A(\mathrm{Frob}_\mathfrak{p})$ will be given by $\lambda(\mathfrak{p})$ for almost all $\mathfrak{p}$, we further see that this Hecke character $\widetilde{r}\circ\chi$ has

$$L(s - w/2, \widetilde{r}\circ\rho_{A,\ell}) \doteq L\left(s, |\cdot|^{-w/2}\,(\widetilde{r}\circ\chi)\right).$$

   Quickly, let's check that the right-hand Hecke character is unitary: it's enough to show that the image is bounded, which is clearer on the left-hand side because the elements $\iota\rho_{A,\ell}(\mathrm{Frob}_\mathfrak{p})$ can be placed (up to conjugacy) in the compact group $\mathrm{ST}(A)$.

   Thus, Theorem 3.77 tells us that we get a nonvanishing meromorphic continuation, and we have a pole at a certain value of $s = 1 + it$ if and only if $|\cdot|^{1-w/2+it}\,(\widetilde{r}\circ\chi) = |\cdot|$, which is equivalent to

$$\widetilde{r}\circ\chi \overset{?}{=} |\cdot|^{w/2+it}.$$

   Note that $(\widetilde{r}\circ\chi)$ defines an algebraic map $\mathrm{T}_K(\mathbb{A}_{\mathbb{Q},\infty}) \to \mathrm{T}_E(\mathbb{A}_{\mathbb{Q},\infty})$, so it can only ever be integral powers of $|\cdot|$, meaning we only have to worry about $t = 0$. But $\widetilde{r}\circ\chi = |\cdot|^{w/2}$ will go back and imply that $r\left(\iota\rho_{A,\ell}(\mathrm{Frob}_\mathfrak{p})\,\mathrm{N}\,\mathfrak{p}^{-1/2}\right)$ is trivial for almost all $\mathfrak{p}$, which implies that r is trivial because Frobenius elements are dense in $\mathrm{Gal}(\overline{K}/K)$ by Theorem 3.115. This completes the proof. ∎

Remark 3.118. In fact, [Joh17, Proposition 16] proves the Sato–Tate conjecture for abelian varieties with potential complex multiplication.

[4] We may choose ℓ favorably because the conjugacy class of $\rho_{A,\ell}(\mathrm{Frob}_\mathfrak{p})$ does not depend on ℓ: the Fundamental theorem of complex multiplication (in the form Proposition 2.141 explains that the Frobenius action must be given by a uniform scalar in $E^\times$.

Chapter 4
The Fermat Curve

Usually mathematicians have to shoot somebody to get this much publicity.

—Thomas R. Nicely

In ths chapter, we will study the Galois representation attached to the projective $\mathbb{Q}$-curve $X_N^1 \subseteq \mathbb{P}^1_{\mathbb{Q}}$ cut out by the equation

$$X_N \colon X^N + Y^N + Z^N = 0,$$

where $N \geq 3$ is some nonnegative integer. For the rest of this chapter, we will fix N and thus denote this curve by $X \subseteq \mathbb{P}^1_{\mathbb{Q}}$.

4.1 Homology and Cohomology

The exposition of this section follows [Ots16, Sections 2 and 3]. We will spend this section setting up some notation and proving basic facts about how these objects relate to each other.

4.1.1 The Group Action

Throughout, it will be helpful to note that the finite alegbraic $\mathbb{Q}$-group

$$G_N \coloneqq \frac{\mu_N \times \mu_N \times \mu_N}{\Delta \mu_N}$$

acts on X_N; here, $\Delta\mu_N \subseteq \mu_N \times \mu_N \times \mu_N$ refers to the diagonally embedded copy of μ_N. As with X_N, we will denote this group by G for the rest of the chapter, and we will let $\zeta \coloneqq \zeta_N$ be a primitive Nth root of unity.

Notably, the action map $G \times X \to X$ is defined over $\mathbb{Q}$ even though $G(\mathbb{Q})$ is trivial. For brevity, we will denote elements of G by $g_{[r:s:t]} \coloneqq [\zeta^r : \zeta^s : \zeta^t]$. We will also have occasion to study the character group $\widehat{G} \coloneqq \widehat{G}_N$, which we identify with

$$\widehat{G}_N = \left\{(a,b,c) \in (\mathbb{Z}/N\mathbb{Z})^3 : a+b+c = 0\right\}.$$

Explicitly, given a triple (a,b,c), we let $\alpha_{(a,b,c)}$ denote the corresponding character, which sends $g_{[r:s:t]} \mapsto \zeta^{ra+bs+tc}$.

In the sequel, we will have many vector spaces induced by X via (co)homology, which therefore have a G-action by functoriality. With this in mind, we make the following definition.

Definition 4.1. Given a $\mathbb{Q}(\zeta)$-vector space H with a G-action, we define

$$\mathrm{H}_\alpha \coloneqq \{v \in \mathrm{H} : g \cdot v = \alpha(g)v\}$$

to be the α-eigenspace for each $\alpha \in \widehat{G}$.

One inconvenience of this definition is that the vector spaces H of interest are frequently defined over $\mathbb{Q}$, but H_α is not.

Thus, we note that some $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on $\widehat{G}$ as follows: say $\tau(\zeta) = \zeta^u$ for some $u \in (\mathbb{Z}/N\mathbb{Z})^\times$, and then

$$(\tau\alpha)([\zeta^r : \zeta^s : \zeta^t]) = \alpha\left([\zeta^{u^{-1}r}, \zeta^{u^{-1}s} : \zeta^{u^{-1}t}]\right),$$

so we see that $\tau\alpha = u^{-1}\alpha$, where the multiplication $u^{-1}\alpha$ is understood to happen where α is a triple in $(\mathbb{Z}/N\mathbb{Z})^3$. With this in mind, given $\alpha \in \widehat{G}$, we let $[\alpha] \subseteq \widehat{G}$ be the collection of characters of the form $u\alpha$ as $u \in (\mathbb{Z}/N\mathbb{Z})^\times$ varies; for example, $-\alpha \in [\alpha]$. The point of this discussion is that we are able to build a decomposition

$$\mathbb{Q}[G] \cong \prod_{[\alpha] \in G/(\mathbb{Z}/N\mathbb{Z})^\times} \mathbb{Q}([\alpha]),$$

where $\mathbb{Q}([\alpha])$ is the image of the map $\mathbb{Q}[G] \to \mathbb{C}$ given by the characters in $[\alpha]$. We are now ready to make the following definition.

Definition 4.2. Given some $\mathbb{Q}$-vector space H with a G-action, we are now ready to define

$$\mathrm{H}_{[\alpha]} \coloneqq \left\{ v \in \mathrm{H} : v \otimes 1 \in \bigoplus_{\beta \in [\alpha]} (\mathrm{H} \otimes_{\mathbb{Q}} \overline{\mathbb{Q}})_\alpha \right\}.$$

The discussion of the Galois action of the previous paragraph implies that $\mathrm{H}_{[\alpha]}$ is a generalized eigenspace of the G-action on H. In particular, we find that $\mathrm{H}_{[\alpha]} \otimes \overline{\mathbb{Q}} = \bigoplus_{\beta \in [\alpha]} \mathrm{H}_\beta$, so $\mathrm{H} = \bigoplus_{[\alpha]} \mathrm{H}_{[\alpha]}$.

4.1.2 Differential Forms

In this subsection, we will define a few differential forms. A reasonable reference for this subsection is [Lan11, Section 1.7]. A computation with the Riemann–Hurwitz formula shows that the genus of X is $\frac{(N-1)(N-2)}{2}$, so we know that there are many holomorphic differential forms. On the other hand, we know that the space of differential forms lives in $\mathrm{H}^1_{\mathrm{dR}}(X(\mathbb{C}), \mathbb{C})$, which is equipped with a G-action. Anyway, we are now ready to define our differential form.

Definition 4.3. Fix notation as above. For $a \in \mathbb{Z}/N\mathbb{Z}$, let $[a]$ be a representative in $\{0, 1, \ldots, N-1\}$. For any $\alpha_{(a,b,c)} \in \widehat{G}$, we define the differential form

$$\omega_{\alpha_{(a,b,c)}} \coloneqq x^{[a]} y^{[b]-N} \frac{dx}{x}$$

in the affine patch $x^N + y^N + 1 = 0$ of X. In the sequel, we may also denote this differential form by $\omega_{(a,b,c)}$.

Remark 4.4. Because $x^N + y^N + 1 = 0$ implies $x^{N-1}\,dx = -y^{N-1}\,dy$, we also see that

$$\omega_{(a,b,c)} = -x^{[a]-N}y^{[b]}\,\frac{dy}{y}.$$

Further, we can pass to the affine patch $1 + v^N + u^N = 0$ of X by substituting $(x, y) = (1/u, v/u)$, for which we note $d(1/u)/(1/u) = -du/u$ so that

$$\omega_{(a,b,c)} = -u^{N-[a]-[b]}v^{[b]-N}\,\frac{du}{u}.$$

Remark 4.5. Following [Col87, Section VI], we remark that it will be numerically convenient to work with a rational multiple of the $\omega_\bullet$s for some computations in the sequel. Namlely, we define $\nu_\alpha := K(\alpha)\omega_\alpha$ when $\alpha = (a, b, c)$ has nonzero entries, where

$$K(a,b,c) := \begin{cases} \frac{N-[a]-[b]}{N} & \text{if } [a]+[b] > N, \\ 1 & \text{if } [a]+[b] < N. \end{cases}$$

From Remark 4.4, we see that $\omega_{(a,b,c)}$ always succeeds at being meromorphic with poles only at points of the form $[X : Y : 0]$, and it is closed (i.e., has vanishing residues) if and only if $0 \notin \{a, b, c\}$. Further, we see that $\omega_{(a,b,c)}$ succeeds at being holomorphic provided that we also have $[a] + [b] < N$, which we note is equivalent to $[a] + [b] + [c] = N$.

We have now provided $\frac{(N-1)(N-2)}{2}$ holomorphic differentials of X, so we would like to check that we have actually found a basis of $\mathrm{H}^0(X(\mathbb{C}), \Omega^1_{X/\mathbb{C}})$. Well, these differential forms are nonzero by construction,[1] and they are linearly independent because they are all eigenvectors for the G-action.

Lemma 4.6. Fix notation as above. For each $\alpha \in \widehat{G}$, the differential form ω_α is an eigenvector for the G-action with eigenvalue α.

Proof. Say $\alpha = \alpha_{(a,b,c)}$ for some $a, b, c \in \mathbb{Z}/N\mathbb{Z}$. Then for any $g_{[r:s:0]} \in G$, we note

$$\begin{aligned}(g_{[r:s:0]})^*\omega_{(a,b,c)} &= (\zeta^r x)^{[a]}(\zeta^s y)^{[b]-N}\,\frac{d(\zeta^r x)}{(\zeta^r x)} \\ &= \zeta^{ar+bs}\cdot x^{[a]}y^{[b]-N}\,\frac{dx}{x} \\ &= \alpha_{(a,b,c)}(g_{[r:s:0]})\omega_{(a,b,c)}.\end{aligned}$$

The reason to $g_{[r:s:0]}$ in the above computation is that we need the G-action to stay in the affine patch of points of the form $[X : Y : 1]$. ∎

Remark 4.7. Thus, we see that our differential forms must be linearly independent because they are eigenvectors with different eigenvalues. As such, we have constructed eigenbases of $\mathrm{H}^1_{\mathrm{dR}}(X(\mathbb{C}), \mathbb{C})$ and $\mathrm{H}^0(X(\mathbb{C}), \Omega^1_{X/\mathbb{C}})$.

While we're here, we compute the Poincaré pairing of our basis of differential forms.

[1] Later, Remark 4.13 will give another way to prove this via periods.

Lemma 4.8. Fix notation as above. Choose $\alpha, \alpha' \in \widehat{G}$ such that $\alpha = (a, b, c)$ and $\alpha' = (a', b', c')$ have nonzero entries. Then the Poincaré pairing

$$P\colon \mathrm{H}^1_{\mathrm{dR}}(X(\mathbb{C}), \mathbb{C}) \times \mathrm{H}^1_{\mathrm{dR}}(X(\mathbb{C}), \mathbb{C}) \to \mathbb{C}$$

given by $(\omega, \eta) \mapsto \frac{1}{2\pi i} \int_X (\omega \wedge \eta)$ sends $(\omega_\alpha, \omega_{\alpha'})$ to

$$P(\omega_\alpha, \omega_{\alpha'}) = \begin{cases} 0 & \text{if } \alpha \neq -\alpha', \\ (-1)^N \frac{N}{N-[a]-[b]} & \text{if } \alpha = -\alpha'. \end{cases}$$

Proof. We use the Poincaré residue, which implies that

$$P(\omega, \eta) = \sum_{x \in X(\mathbb{C})} \mathrm{Res}_x \left(\eta \int \omega \right),$$

where the sum is over poles, and $\int \omega$ refers to any choice of local primitive for ω in the neighborhood of x. To use this, we note that the computation of Remark 4.4 implies that ω_α and $\omega_{\alpha'}$ can only have poles at the points $[1 : -\zeta^s : 0]$ for some $s \in \mathbb{Z}/N\mathbb{Z}$, and in this neighborhood, we may write

$$\omega_\alpha = -u^{N-[a]-[b]} v^{[b]-N} \frac{du}{u}$$

and similarly for $\omega_{\alpha'}$. In particular, we see that

$$-\frac{1}{N-[a]-[b]} u^{N-[a]-[b]} v^{[b]-N}$$

makes a reasonable primitive for ω_α, so the Poincaré residue yields

$$P(\omega_\alpha, \omega_{\alpha'}) = \sum_{s \in \mathbb{Z}/N\mathbb{Z}} \mathrm{Res}_{(-\zeta^s, 0)} \left(-\frac{1}{N-[a]-[b]} u^{N-[a]-[b]} v^{[b]-N} \cdot -u^{N-[a']-[b']} v^{[b']-N} \frac{du}{u} \right).$$

Now, if $\alpha \neq \alpha'$, then we see that we are computing the residues of some monomial times du/u, but the power of u in the monomial is nonzero, so the residues all vanish. Lastly, we need to discuss what happens with $\alpha = -\alpha'$, where we see

$$\begin{aligned} P(\omega_\alpha, \omega_{-\alpha}) &= \sum_{s \in \mathbb{Z}/N\mathbb{Z}} \mathrm{Res}_{(-\zeta^s, 0)} \left(-\frac{1}{N-[a]-[b]} u^{N-[a]-[b]} v^{[b]-N} \cdot -u^{N-[-a]-[-b]} v^{[-b]-N} \frac{du}{u} \right) \\ &= \sum_{s \in \mathbb{Z}/N\mathbb{Z}} \mathrm{Res}_{(-\zeta^s, 0)} \left(-\frac{1}{N-[a]-[b]} u^{N-[a]-[b]} v^{[b]-N} \cdot u^{[a]+[b]-N} v^{-[b]} \frac{du}{u} \right) \\ &= \frac{1}{N-[a]-[b]} \sum_{s \in \mathbb{Z}/N\mathbb{Z}} \mathrm{Res}_{(-\zeta^s, 0)} \left(v^{-N} \frac{du}{u} \right) \\ &= \frac{1}{N-[a]-[b]} \sum_{s \in \mathbb{Z}/N\mathbb{Z}} (-\zeta^s)^{-N} \\ &= (-1)^N \frac{N}{N-[a]-[b]}, \end{aligned}$$

as desired. ∎

Remark 4.9. Following Remark 4.5, we see that $\alpha \in \widehat{G}$ with nonzero entries will have

$$P(\nu_\alpha, \nu_{-\alpha}) = (-1)^N$$

because exactly one of $K(\alpha)$ or $K(-\alpha)$ will absorb the given rational constant. This is essentially the reason for working with the $\nu_\bullet$s instead of the $\omega_\bullet$s.

4.1.3 Some Group Elements

In this subsection, we define a few elements of $\mathbb{Q}[G]$ which we will then use in the next subsection. We begin with the three elements

$$t := \sum_{g \in G} g, \qquad v := \sum_{s \in \mathbb{Z}/N\mathbb{Z}} g_{[0:s:0]}, \qquad \text{and} \qquad h := \sum_{r \in \mathbb{Z}/N\mathbb{Z}} g_{[r:0:0]}.$$

We take a moment to note that these elements satisfy the relations $tg = gt = t$ for any $g \in G$, and $t = hv = vh$, and $v^2 = Nv$ and $h^2 = Nh$. In the sequel, we will get a lot of mileage out of the idempotent

$$p := \frac{1}{N^2} \sum_{r,s \in \mathbb{Z}/N\mathbb{Z}} (1 - g_{[r:0:0]})(1 - g_{[0:s:0]}).$$

Let's check that p is idempotent.

Lemma 4.10. Fix notation as above.

(a) Then p is idempotent.

(b) For any $r, s \in \mathbb{Z}/N\mathbb{Z}$, we have $(1 - g_{[r:0:0]})(1 - g_{[0:s:0]})p = (1 - g_{[r:0:0]})(1 - g_{[0:s:0]})$.

Proof. Both claims hinge upon the fact that a direct expansion of $(1 - g_{[r:0:0]})(1 - g_{[0:s:0]})$ implies

$$p = \frac{1}{N^2} \left(N^2 - Nh - Nv + t\right).$$

We now show the claims separately.

(a) This is a direct computation: write

$$\begin{aligned} p^2 &= \frac{1}{N^4} \left(N^2 - Nh - Nv + t\right) \left(N^2 - Nh - Nv + t\right) \\ &= \frac{1}{N^4} \left(N^4 + N^2h^2 + N^2v^2 + t^2 - 2N^3h - 2N^3v + 2N^2t + N^2hv - 2Nht - 2Nvt\right) \\ &= \frac{1}{N^4} \left(N^4 + N^3h + N^3v + N^2t - 2N^3h - 2N^3v + 2N^2t + N^2t - 2N^2t - 2N^2t\right) \\ &= \frac{1}{N^4} \left(N^4 - N^3h - N^3v + N^2t\right) \\ &= p. \end{aligned}$$

(b) We will compute as in (a): note $h(1 - g_{[r:0:0]}) = 0$ and $v(1 - g_{[0:s:0]}) = 0$, so

$$\begin{aligned} (1 - g_{[r:0:0]})(1 - g_{[0:s:0]})p &= (1 - g_{[r:0:0]})(1 - g_{[0:s:0]}) \cdot \frac{1}{N^2} \left(N^2 - Nh - Nv + hv\right) \\ &= (1 - g_{[r:0:0]})(1 - g_{[0:s:0]}) \cdot \frac{N^2}{N^2} + 0 + 0 + 0 \\ &= (1 - g_{[r:0:0]})(1 - g_{[0:s:0]}), \end{aligned}$$

as required. ∎

4.1.4 Homology

In this subsection, we will study $\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{Q})$. By the end, we will define a 1-cycle $\gamma := \gamma_N$ such that $\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{Q}) = \mathbb{Q}[G]\cdot[\gamma]$. Morally, this means that we can understand our homology by focusing on this one cycle.

To begin, we need some path in $X(\mathbb{C})$, so we define $\delta\colon [0,1]\to X(\mathbb{C})$ by

$$\delta(t) := \left[t^{1/N} : (1-t)^{1/N} : \zeta_{2N}^{-1}\right].$$

Notably, $\delta(0) = [0:1:\zeta_{2N}^{-1}]$ and $\delta(1) = [1:0:\zeta_{2N}^{-1}]$, so $g = [\zeta^r : \zeta^s : 1]$ has $g_*\delta(0) = [0:\zeta^s:\zeta_{2N}^{-1}]$ and $g_*\delta(1) = [\zeta^r:0:\zeta_{2N}^{-1}]$. The point of this computation is that we see

$$(1 - g_{[r:0:0]} - g_{[0:s:0]} + g_{[r:s:0]})_*\delta \in \mathrm{Z}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{Q}).$$

We are now ready to define γ.

Definition 4.11. Fix notation (and in particular δ) as above. Then we define

$$\gamma := \frac{1}{N^2}\sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} (1-g_{[r:0:0]})(1-g_{[0:s:0]})_*\delta.$$

Note $\gamma = p_*\delta$.

The above computation shows that $\gamma\in \mathrm{Z}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{Q})$. We will want to know to its periods later. Note that the following result is essentially a special case of [Del18, Lemma 7.12].

Lemma 4.12. Fix notation as above. Suppose $(a,b,c)\in(\mathbb{Z}/N\mathbb{Z})^3$ has no nonzero entries. Then

$$\int_\gamma \omega_{(a,b,c)} = \zeta_{2N}^{[a]+[b]-N}\frac{\Gamma\left(\frac{[a]}{N}\right)\Gamma\left(\frac{[b]}{N}\right)}{\Gamma\left(\frac{[a]}{N}+\frac{[b]}{N}\right)}.$$

Proof. This is a direct computation. Denote the integral by $P(\gamma,\omega_{(a,b,c)})$. By adjunction, $\int_{p_*\delta}\omega_{(a,b,c)} = \int_\delta p^*\omega_{(a,b,c)}$. This allows us to compute

$$\begin{aligned}
P(\gamma,\omega_{(a,b,c)}) &= \frac{1}{N^2}\int_\delta \sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} (1-g_{[r:0:0]})(1-g_{[0:s:0]})^*\omega_{(a,b,c)}\\
&= \frac{1}{N^2}\int_\delta \sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} \left(1-\zeta^{ar}\right)\left(1-\zeta^{bs}\right)\omega_{(a,b,c)}\\
&= \left(\frac{1}{N^2}\sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} \left(1-\zeta^{ar}\right)\left(1-\zeta^{bs}\right)\right)\int_\delta \omega_{(a,b,c)}\\
&= \left(\frac{1}{N^2}\sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} \left(1-\zeta^{ar}\right)\left(1-\zeta^{bs}\right)\right)\zeta_{2N}^{[a]+[b]-N}\int_0^1 t^{[a]/N}(1-t)^{[b]/N-1}\frac{dt}{t}.
\end{aligned}$$

The last integral (famously) equals the Beta function, and it evaluates to $\Gamma\left(\frac{[a]}{N}\right)\Gamma\left(\frac{[b]}{N}\right)\Gamma\left(\frac{[a]+[b]}{N}\right)^{-1}$. We take a moment to check that

$$\sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} \left(1-\zeta^{ar}\right)\left(1-\zeta^{bs}\right) \stackrel{?}{=} N^2.$$

Well, $\left(1-\zeta^{ar}\right)\left(1-\zeta^{bs}\right) = 1-\zeta^{ar}-\zeta^{bs}+\zeta^{ar+bs}$, and because $a,b\neq 0$, we see that summing over r and s causes the terms not equal to 1 to vanish. Thus, we are left with N^2. ∎

Remark 4.13. Because the right-hand side is nonzero, Lemma 4.12 implies that the differential forms $\omega_{(a,b,c)}$ are nonzero.

Remark 4.14. Following Remark 4.5, it will be helpful to also compute $\int_\gamma \nu_{(a,b,c)}$. We claim that

$$\int_\gamma \nu_{(a,b,c)} \stackrel{?}{=} (-1)^{\lfloor([a]+[b])/N\rfloor}\zeta_{2N}^{[a]+[b]-N}\Gamma\left(\frac{[a]}{N}\right)\Gamma\left(\frac{[b]}{N}\right)\Gamma\left(\frac{[a+b]}{N}\right)^{-1}.$$

We have two cases. If $[a]+[b] < N$, then $\nu_{(a,b,c)} = \omega_{(a,b,c)}$, so this is immediate from Lemma 4.12. Otherwise, if $[a]+[b] > N$, then $\nu_{(a,b,c)} = \frac{N-[a]-[b]}{N}\omega_{(a,b,c)}$, so this follows from Lemma 4.12 as soon as we compute

$$\frac{N-[a]-[b]}{N}\Gamma\left(\frac{[a]+[b]}{N}\right)^{-1} \stackrel{?}{=} -\Gamma\left(\frac{[a+b]}{N}\right)^{-1}.$$

This follows because $\Gamma\left(\frac{[a]+[b]}{N}\right) = \frac{[a]+[b]-N}{N}\Gamma\left(\frac{[a+b]}{N}\right)$.

We are now ready to show that $\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{Q}) = \mathbb{Q}[G]\cdot[\gamma]$.

Lemma 4.15. Fix notation as above. Then $\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{Q}) = \mathbb{Q}[G]\cdot[\gamma]$.

Proof. It is enough to show that $\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{C}) = \mathbb{C}[G]\cdot[\gamma]$. Note that there is a canonical pairing

$$\begin{array}{ccc}\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{C}) \times \mathrm{H}^1_{\mathrm{dR}}(X(\mathbb{C}),\mathbb{C}) & \to & \mathbb{C} \\ (c,\omega) & & \mapsto \int_c \omega\end{array}$$

which is perfect by the de Rham theorem. We already have a basis $\{\omega_{(a,b,c)}\}_{a,b,c\neq 0}$ of $\mathrm{H}^1_{\mathrm{dR}}(X(\mathbb{C}),\mathbb{C})$, so we will find a dual basis for $\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{C})$ inside $\mathbb{C}[G]\cdot[\gamma]$. Well, for $g\in G$ and $\alpha\in\widehat{G}$, we see

$$\int_{g^*\gamma}\omega_\alpha = \int_\gamma g^*\omega_\alpha$$

equals $\alpha(g)P(\gamma,\omega_\alpha)$, where $P(\gamma,\omega_\alpha) \coloneqq \int_\gamma \omega_\alpha$ is the (nonzero!) period computed in Lemma 4.12. With this in mind, we define

$$c_\alpha \coloneqq \frac{1}{N^2 P(\gamma,\omega_\alpha)}\sum_{g\in G}\alpha(g)^{-1}g^*[\gamma]$$

for each $\alpha = \alpha_{(a,b,c)}$ with $a,b,c\neq 0$. Then we see that $\int_{c_\alpha}\omega_\beta = 1_{\alpha=\beta}$ by the orthogonality relations, so $\{c_\alpha\}$ is a dual basis of $\mathrm{H}_1^{\mathrm{B}}(X(\mathbb{C}),\mathbb{C})$, and it lives in $\mathbb{C}[G]\cdot[\gamma]$ by its construction. ∎

4.2 Galois Action: the Étale Site

We now use the notation set up in the previous section to write out the Galois action on the space of some absolute Hodge cycles attached to X. Roughly speaking, we will be interested in computing ℓ-adic monodromy groups of (quotients of) X, which requires us to have some understanding of the Galois representation

$$\rho\colon\ \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}\left(\mathrm{H}^1_{\text{ét}}(X_{\overline{\mathbb{Q}}},\mathbb{Q}_\ell)\right).$$

In particular, we recall from section 2.4.3 that it will really suffice to be able to compute the Galois action on cetain Tate classes living in

$$\mathrm{H}^1_{\text{ét}}(X_{\overline{\mathbb{Q}}},\mathbb{Q}_\ell)^{\otimes p}\otimes \mathrm{H}^1_{\text{ét}}(X_{\overline{\mathbb{Q}}},\mathbb{Q}_\ell)^{\vee\otimes p} \cong \mathrm{H}^1_{\text{ét}}(X_{\overline{\mathbb{Q}}},\mathbb{Q}_\ell)^{\otimes 2p}(p),$$

for some nonnegative index $p \geq 0$, which is the main point of this section. In particular, the Künneth theorem tells us that we will be interested in the cohomology group $\mathrm{H}^{2p}_{\text{ét}}(X^{2p}_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)(p)$.

Roughly speaking, the outline will be to pass to absolute Hodge cycles. Indeed, by the Mumford–Tate conjecture, one is able to correspond Tate classes to Hodge classes, and Hodge classes are known to be absolutely Hodge, and our construction of absolute Hodge cycles makes it clear how they should specialize to a Tate class. In this way, we find that we can attempt to compute Galois action on Tate classes by instead computing Galois action on absolute Hodge classes. This is useful because absolute Hodge cycles have a de Rham component, so we can run our computations on the de Rham component, which is the only place where we can hope to have a basis.

Throughout this section, p is a nonnegative index. We take a moment to note that the action of G on X upgrades into an action of G^{2p} on X^{2p}. Our exposition closely follows [GGL24, Subsection 8.5]. As in section 4.1.1, we will identify $\widehat{G}^{2p}$ with some subset of tuples in $(\mathbb{Z}/N\mathbb{Z})^{6p}$. And for a vector space H defined over $\mathbb{Q}(\zeta)$ (respectively, $\mathbb{Q}$) and character $\alpha \in \widehat{G}^{2p}$, we define H_α (respectively, $\mathrm{H}_{[\alpha]}$) as the corresponding α-eigenspace (respectively, $[\alpha]$-generalized eigenspace). Then given a vector $v \in \mathrm{H}$, we may also write v_α for the component in H_α.

In the sequel, we will find utility out of the following two subsets of $\widehat{G}^{2p}$.

Definition 4.16. Fix notation as above.

- We define the subset $\mathfrak{A}^{2p}$ to be equal to the subset of $\alpha \in \widehat{G}^{2p}$ having nonzero entries as a tuple in $(\mathbb{Z}/N\mathbb{Z})^{6p}$.
- We define the subset $\mathfrak{B}^{2p}$ to be equal to the subset of $\alpha \in \mathfrak{A}^{2p}$ such that $\alpha = (a_1, \ldots, a_{6p})$ satisfies

$$\frac{1}{N}\sum_{i=1}^{6p}[ua_i] = 3p$$

for all $u \in (\mathbb{Z}/N\mathbb{Z})^\times$.

Morally, the characters in $\mathfrak{A}_{2p}$ correspond to basis vectors of $\mathrm{H}^1_{\text{ét}}(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)^{\otimes p}$, and the characters in $\mathcal{B}_{2p}$ correspond to Hodge classes (see Proposition 4.23).

4.2.1 Hodge Cycles on X^{2p}

To understand the geometry of X, we will only be interested in tensor powers of $\mathrm{H}^1(X)$ (for a choice of cohomology theory H), which by the Künneth formula embed as

$$\mathrm{H}^1(X)^{\otimes 2p} \subseteq \mathrm{H}^{2p}\left(X^{2p}\right).$$

When H is de Rham cohomology H_{dR}, we thus see we are interested in when the image of an element in $\mathrm{H}^1_{\mathrm{dR}}(X)^{\otimes p}$ succeeds at being a Hodge cycle. Well, note that the action of G on $\mathrm{H}^1_{\mathrm{dR}}(X, \mathbb{C})$ extends to an action of G^{2p} on $\mathrm{H}^1_{\mathrm{dR}}\left(X, \mathbb{C}\right)^{\otimes 2p}$. This action diagonalizes with one-dimensional eigenspaces by extending Remark 4.7. We will use properties of the diagonalization to read off when we have an element of bidegree (p,p) in $\mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p}, \mathbb{C}\right)$.

Following [Del18, Proposition 7.6], it will be useful to have the following definition.

Definition 4.17 (weight)**.** Given a function $f\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, we define its *weight map* as the function $\langle f\rangle\colon (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}$ defined by

$$\langle f\rangle(u) \coloneqq \frac{1}{N}\sum_{a\in\mathbb{Z}/N\mathbb{Z}} f(ua)[a]$$

For $p \geq 0$, we note that we may identify $\widehat{G}^{2p}$ with a tuple in $(\mathbb{Z}/m\mathbb{Z})^{2p}$, and then we define the *weight* $\langle\alpha\rangle$ of a character $\alpha \in \widehat{G}^{2p}$ as $\langle 1_\alpha\rangle(1)$, where $1_\alpha\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}$ is the multiplicity of an element in $\mathbb{Z}/N\mathbb{Z}$ in the tuple α.

Remark 4.18. The point of this definition is as follows: given $\alpha \in \widehat{G}$ with $\alpha = (a, b, c)$ having nonzero entries, we note that ω_α has two possible cases.

- If $[a] + [b] + [c] = N$ so that $\langle \alpha \rangle = 1$, then $\omega_{(a,b,c)}$ is holomorphic so that $\omega_\alpha \in \mathrm{H}^{10}(X)$.
- If $[a] + [b] + [c] = 2N$ so that $\langle \alpha \rangle = 2$, then ω_α is not holomorphic so that $\omega_\alpha \in \mathrm{H}^{01}(X)$.

In all cases, we find $\omega_\alpha \in \mathrm{H}^{2-\langle \alpha \rangle, \langle \alpha \rangle - 1}(X)$.

Remark 4.19. If f is instead a function $f \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{C}$, we may similarly define the weight by the formula

$$\langle f \rangle(u) \coloneqq \sum_{a \in \frac{1}{N}\mathbb{Z}/\mathbb{Z}} f(ua)\langle a \rangle,$$

where $\langle a \rangle$ now refers to the element in $[0, 1)$ in the class of a.

Remark 4.20. Suppose that $\alpha \in \mathfrak{A}^{2p}$ has 1_α of constant weight. Then we claim that $\langle \alpha \rangle = 3p$. Indeed, we must have

$$\frac{1}{N}\sum_{i=1}^{6p}[a_i] = \frac{1}{N}\sum_{i=1}^{6p}[-a_i],$$

but $[-a_i] = N - a_i$ then forces the sum to equal $3p$.

We now upgrade Remark 4.18 to $\mathrm{H}^1_{\mathrm{dR}}(X, \mathbb{C})^{\otimes 2p}$.

Notation 4.21. Choose $\alpha \in \widehat{G}^{2p}$ as $\alpha = (\alpha_1, \ldots, \alpha_{2p})$ having nonzero entries. Then we set

$$\omega_\alpha \coloneqq \omega_{\alpha_1} \otimes \cdots \otimes \omega_{\alpha_{2p}}.$$

We define ν_α similarly.

Lemma 4.22. Choose $\alpha \in \widehat{G}^{2p}$ as $\alpha = (\alpha_1, \ldots, \alpha_{2p})$ having nonzero entries (i.e., $\alpha \in \mathfrak{A}_{2p}$). Then ω_α embedded in $\mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p}, \mathbb{C}\right)$ is of bidegree $(4p - \langle \alpha \rangle, \langle \alpha \rangle - 2p)$.

Proof. Because the Künneth isomorphism upgrades to an isomorphism of Hodge structures, it is enough to note that $\omega_{\alpha_i} \in \mathrm{H}^{2-\langle \alpha_i \rangle, \langle \alpha_i \rangle - 1}$ (see Remark 4.18) implies ω_α has bidegree

$$\left(4p - \sum_{i=1}^{2p}\langle \alpha_i \rangle, \sum_{i=1}^{2p}\langle \alpha_i \rangle - 2p\right).$$

The lemma follows because weight is additive. ∎

Proposition 4.23. Choose $\alpha \in \mathfrak{A}^{2p}$. Then $\mathrm{H}^{2p}_{\mathrm{B}}\left(X^{2p}\right)_{[\alpha]}$ is one-dimensional over $\mathbb{Q}([\alpha])$, and the following are equivalent.

(a) $\mathrm{H}^{2p}_{\mathrm{B}}\left(X^{2p}\right)_{[\alpha]}(p)$ consists entirely of Hodge classes.

(b) We have $\langle u\alpha \rangle = 3p$ for all $u \in (\mathbb{Z}/N\mathbb{Z})^\times$.

Proof. Expand $\alpha = (\alpha_1, \dots, \alpha_{2p})$. We begin by embedding

$$\mathrm{H}^{2p}_{\mathrm{B}}\left(X^{2p}, \mathbb{Q}\right)_{[\alpha]} \otimes_{\mathbb{Q}} \mathbb{C} = \bigoplus_{u \in (\mathbb{Z}/N\mathbb{Z})^\times} \mathrm{H}^{2p}_{\mathrm{B}}\left(X^{2p}, \mathbb{C}\right)_{u\alpha}$$

into

$$\mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p}, \mathbb{C}\right) = \bigoplus_{\substack{q_1, \dots, q_{2p} \\ q_1 + \cdots + q_{2p} = 2p}} \mathrm{H}^{q_1}_{\mathrm{dR}}(X, \mathbb{C}) \otimes \cdots \otimes \mathrm{H}^{q_{2p}}_{\mathrm{dR}}(X, \mathbb{C}),$$

where this last equality holds by the Künneth isomorphism. Quickly, we reduce to the case where $q_1 = \cdots = q_{2p} = 1$: for each $u \in (\mathbb{Z}/N\mathbb{Z})^\times$, we note that $u\alpha$ has nonzero entries. On the other hand, the G-action on $\mathrm{H}^0(X) = \mathbb{C}$ is always trivial, so we note that if any of the $q_\bullet$s are not equal to 1, then one of them must equal 0, meaning that

$$\left(\mathrm{H}^{q_1}_{\mathrm{dR}}(X, \mathbb{C}) \otimes \cdots \otimes \mathrm{H}^{q_{2p}}_{\mathrm{dR}}(X, \mathbb{C})\right)_{u\alpha} = \mathrm{H}^{q_1}_{\mathrm{dR}}(X, \mathbb{C})_{u\alpha_1} \otimes \cdots \otimes \mathrm{H}^{q_{2p}}_{\mathrm{dR}}(X, \mathbb{C})_{u\alpha_{2p}}$$

is the zero vector space. Thus, we see that

$$\mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p}, \mathbb{C}\right)_{[\alpha]} = \bigoplus_{u \in (\mathbb{Z}/N\mathbb{Z})^\times} \left(\mathrm{H}^1_{\mathrm{dR}}(X, \mathbb{C})^{\otimes 2p}\right)_{u\alpha}.$$

The comparison isomorphism now implies that $\mathrm{H}^{2p}_{\mathrm{B}}\left(X^{2p}, \mathbb{Q}\right)_{[\alpha]}$ has dimension $[\mathbb{Q}([\alpha]) : \mathbb{Q}]$ over $\mathbb{Q}$ and thus one dimension over $\mathbb{Q}([\alpha])$.

It remains to show that (a) and (b) are equivalent. Well, the $\mathbb{Q}$-vector space $\mathrm{H}^{2p}_{\mathrm{B}}\left(X^{2p}, \mathbb{Q}\right)_{[\alpha]}(p)$ will consist of Hodge classes if and only if $\left(\mathrm{H}^1_{\mathrm{dR}}(X, \mathbb{C})^{\otimes 2p}\right)_{u\alpha}$ is of bidegree (p, p), which is equivalent to $\langle u\alpha \rangle = 3p$ by Lemma 4.22. ∎

4.2.2 An Absolute Hodge Cycle

Thus far, we have access to classes ω_α, and we know how to compute their periods against the Betti cycle γ. We will be able to compute the Galois action on γ because it already comes from a Betti cycle, but we then need to know how to translate this into a Galois action on the ω_α; importantly, note that ω_αs have no obvious Galois action, and indeed, they cannot because they may not even be defined over a number field. To do this, we need a way to put γ and the ω_α on the same footing; following [GGL24, Section 8.5], we use absolute Hodge classes.

For example, the machinery of cohomology tells us how to take γ and then apply some cycle class maps to produce an absolute Hodge class. Let's be more explicit: we may pass the class $\gamma^{2p} \otimes (2\pi i)^{-p}$ through the maps

$$\mathrm{H}^{\mathrm{B}}_{2p}\left(X^{2p}, \mathbb{C}\right)(-p) \cong \mathrm{H}^{2p}_{\mathrm{B}}\left(X^{2p}, \mathbb{C}\right)(p) \subseteq \mathrm{H}^{2p}_{\mathbb{A}}(X)(p),$$

where the last map is the cycle class map. In order to ensure that we output an absolute Hodge cycle, we apply Proposition 4.23: we see that the generalized eigenspace for $[\alpha]$ contains all Hodge classes if and only if $\alpha \in \mathfrak{B}^{2p}$, we simply define $\gamma^{2p}_{[\alpha]} \in \mathrm{H}^{\mathrm{B}}_{2p}\left(X^{2p}, \mathbb{Q}\right)$ to be the projection to the $[\alpha]$-component, and we now know that its image $\gamma^{2p}_{[\alpha],\mathrm{AH}}$ is a Hodge class, hence an absolute Hodge class by Theorem 2.45.

> **Remark 4.24.** We remark that this last paragraph actually argues that the projection
>
> $$C^p_{\mathrm{AH}}(X)_{[\alpha]} \twoheadrightarrow \mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p}, \mathbb{C}\right)(p)_{[\alpha]}$$
>
> is an isomorphism for any $\alpha \in \mathfrak{B}^{2p}$. In particular, both spaces are 1-dimensional vector spaces over $\mathbb{Q}([\alpha])$.

Perhaps we should check that $\gamma^{2p}_{[\alpha],\mathrm{AH}}$ is nonzero. Roughly speaking, we expect this to hold by the period computations of Lemma 4.12.

Proposition 4.25. Choose $\alpha \in \mathfrak{B}^{2p}$. Then

$$\pi_\infty\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\right) = \sum_{\substack{\beta\in[\alpha]\\ \beta=(a_1,b_1,c_1,\ldots,a_{2p},b_{2p},c_{2p})}} \left((2\pi i)^{-p}\prod_{i=1}^{2p}\frac{N-[a_i]-[b_i]}{N}\int_\gamma \omega_{(-a_i,-b_i,-c_i)}\right)\omega_\beta.$$

Proof. We know that the ω_β form an eigenbasis of $\mathrm{H}_{\mathrm{dR}}^{2p}\left(X^{2p},\mathbb{C}\right)(p)_{[\alpha]}$ by restricting Remark 4.7 to the $[\alpha]$-generalized eigenspace. Thus, we know that $\pi_\infty(\gamma_{[\alpha],\mathrm{AH}})$ is certainly a linear combination of the ω_βs, so we write

$$\pi_\infty\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\right) = \sum_{\beta\in[\alpha]} z_\beta\omega_\beta,$$

and it remains to compute the coefficients z_β. For this, we use the computation of the Poincaré pairing computation from Lemma 4.8 (iterated $2p$ times), whereupon we see that

$$P\left(\pi_\infty\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\right),\omega_{-\beta}\right) = z_\beta\prod_{i=1}^{2p}(-1)^N\frac{N}{N-[a_i]-[b_i]},$$

where $\beta=(a_1,b_1,c_1,\ldots,a_{2p},b_{2p},c_{2p})$. Thus, to get the correct answer for z_β, we would like to show that

$$P\left(\pi_\infty\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\right),\omega_{-\beta}\right) \stackrel{?}{=} (2\pi i)^{-p}\int_\gamma \omega_{-\beta}.$$

(Note that the sign has disappeared because $(-1)^{N\cdot 2p}=1$.) To compute this Poincaré pairing, we would like to remember that $\gamma_{[\alpha],\mathrm{AH}}$ comes from a Betti class. As such, we remark that the composite

$$\mathrm{H}_{2p}^{\mathrm{B}}\left(X^{2p},\mathbb{C}\right)(-p)\cong\mathrm{H}_{\mathrm{B}}^{2p}\left(X^{2p},\mathbb{C}\right)(p)\subseteq\mathrm{H}_{\mathbb{A}}^{2p}\left(X^{2p}\right)(p)\twoheadrightarrow\mathrm{H}_{\mathrm{dR}}^{2p}\left(X^{2p},\mathbb{C}\right)$$

is just the usual cycle class map from Betti to de Rham cohomology. Thus, we see that the Poincaré pairing with $\gamma_{[\alpha],\mathrm{AH}}^{2p}$ may be computed as the integration pairing

$$P\left(\pi_\infty\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\right),\omega_{-\beta}\right) = (2\pi i)^{-p}\int_{\gamma_{[\alpha]}^{2p}}\omega_{-\beta}.$$

To complete the proof, we note that we may pass from integrating over $\gamma_{[\alpha]}^{2p}$ to γ^{2p} because the adjunctive property of the integration pairing allows us to pass the projection onto the $[\alpha]$-component to $-\beta$, but $\omega_{-\beta}$ already lives in the $[\alpha]$-generalized eigenspace. ∎

Thus, we see that $\gamma_{[\alpha],\mathrm{AH}}$ is nonzero because we have found nonzero coefficients: the integrals are nonzero by Lemma 4.12. While we're here, we translate this into a statement with $\nu_\bullet$s.

Corollary 4.26. Choose $\alpha \in \mathfrak{B}^{2p}$. Then

$$\pi_\infty\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\right) = \sum_{\substack{\beta\in[\alpha]\\ \beta=(a_1,b_1,c_1,\ldots,a_{2p},b_{2p},c_{2p})}} \left((2\pi i)^{-p}\prod_{i=1}^{2p}\int_\gamma \nu_{(-a_i,-b_i,-c_i)}\right)\nu_\beta.$$

Proof. The same proof as in Proposition 4.25 applies when combined with Remark 4.9. ∎

Remark 4.27. Following Remark 4.9, it will be computationally helpful to rewrite our formula in terms of the $\nu_\bullet$s because this will make the mysterious rational constant disappear.

In order to hide these integrals for now, we introduce the following notation.

Notation 4.28. For $\alpha \in \mathfrak{B}^{2p}$ such that $\alpha = (\alpha_1, \ldots, \alpha_{2p})$, we define

$$\operatorname{Per}\left(\gamma^{2p}, \nu_\alpha\right) \coloneqq (2\pi i)^{-p} \prod_{i=1}^{2p} \int_\gamma \nu_{\alpha_i}.$$

Note that this number is algebraic by Proposition 4.23 because it is the integral of a differential against an absolute Hodge class. (See the end of the proof of Proposition 4.25.)

Remark 4.29. In order to compute these integrals, we note Remark 4.14 grants the product of the integrals equals

$$\prod_{i=1}^{2p} (-1)^{\lfloor ([a_i]+[b_i])/N \rfloor} \zeta_{2N}^{[a_i]+[b_i]-N} \Gamma\left(\frac{[a_i]}{N}\right) \Gamma\left(\frac{[b_i]}{N}\right) \Gamma\left(\frac{[-c_i]}{N}\right)^{-1}.$$

We quickly note that $\left\lfloor \frac{[a_i]+[b_i]}{N} \right\rfloor \in \{0,1\}$ equals 0 exactly p times and equals 1 exactly p times because $\alpha \in \mathfrak{B}^{2p}$; additionally, $\zeta_{2N}^{-N\cdot 2p} = 1$, so that power vanishes. Thus, our period equals

$$(-2\pi i)^{-p} \prod_{i=1}^{2p} \zeta_{2N}^{[a_i]+[b_i]} \frac{\Gamma\left(\frac{[a_i]}{N}\right) \Gamma\left(\frac{[b_i]}{N}\right)}{\Gamma\left(\frac{[-c_i]}{N}\right)}.$$

We will also want to express the $\nu_\bullet$s in terms of γ.

Corollary 4.30. Choose $\alpha \in \mathfrak{B}^{2p}$. For any $\beta \in [\alpha]$, we have

$$\nu_\beta = \frac{1}{\#G^{2p}(\overline{\mathbb{Q}}) \operatorname{Per}\left(\gamma^{2p}, \nu_{-\alpha}\right)} \sum_{g \in G^{2p}(\overline{\mathbb{Q}})} \beta(g)^{-1} \cdot \pi_\infty\left(g^* \gamma^{2p}_{[\alpha],\mathrm{AH}}\right).$$

Proof. By the orthogonality of characters applied to Corollary 4.26, we find that

$$\frac{1}{\#G^{2p}(\overline{\mathbb{Q}})} \sum_{g \in G^{2p}(\overline{\mathbb{Q}})} \beta(g)^{-1} \cdot \pi_\infty\left(g^* \gamma^{2p}_{[\alpha],\mathrm{AH}}\right) = \operatorname{Per}\left(\gamma^{2p}, \nu_{-\alpha}\right) \nu_{\beta,\mathrm{AH}},$$

so the result follows. ∎

4.2.3 Computation of the Galois Action

In this subsection, we compute the Galois action on our absolute Hodge cycles. To ground ourselves, we begin by noting that we are expecting a permutation matrix.

Lemma 4.31. Choose $\alpha \in \mathfrak{A}^{2p}$ and a prime ℓ such that $\ell \equiv 1 \pmod{N}$. Given $\sigma \in \operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ such that $\sigma(\zeta_N) = \zeta_N^u$ for some $u \in (\mathbb{Z}/N\mathbb{Z})^\times$, we find that σ maps

$$\mathrm{H}^{2p}_{\text{ét}}\left(X^{2p}_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell\right)_\alpha \to \mathrm{H}^{2p}_{\text{ét}}\left(X^{2p}_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell\right)_{u^{-1}\alpha}.$$

Proof. Choose $v \in \mathrm{H}^{2p}_{\text{ét}}\left(X^{2p}_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell\right)_\alpha$. Then for any $g \in G^{2p}(\mathbb{Q}_\ell)$, we find that

$$\sigma(g \cdot v) = \sigma(g) \cdot \sigma(v)$$

because the action of G^{2p} is defined over $\mathbb{Q}$ and hence Galois-invariant. Rearranging, we see that

$$\begin{aligned} g \cdot \sigma(v) &= \sigma\left(\sigma^{-1}(g)\right) \cdot \sigma(v) \\ &= \sigma\left(\sigma^{-1}(g) \cdot v\right) \\ &= \sigma\left(\alpha\left(\sigma^{-1}(g)\right) \cdot v\right) \\ &= \alpha\left(\sigma^{-1}(g)\right) \sigma(v), \end{aligned}$$

where the last equality holds because the Galois action is $\mathbb{Q}_\ell$-linear. A direct computation then shows $\alpha\left(\sigma^{-1}(g)\right) = \sigma^{-1}(\alpha(g))$ and then $\sigma^{-1}(\alpha(g)) = \left(u^{-1}\alpha\right)(g)$. ∎

We now move towards the computation of the Galois action on absolute Hodge classes. This requires a warning. Our computation will be able to succeed by using de Rham classes as representatives for absolute Hodge classes. However, de Rham classes have no Galois action: only absolute Hodge classes have Galois action (through the ℓ-adic components). The key to keeping track of the differences between these elements is to keep track of our base-changes. In particular, for any prime ℓ, we may specify an embedding $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$ and note the "comparison" isomorphisms

$$\begin{aligned} \mathrm{H}^{2p}_{\mathrm{dR}}(X, \mathbb{C})(p)_{[\alpha]} &= C^p_{\mathrm{AH}}\left(X^{2p}_{\overline{\mathbb{Q}}}\right)_{[\alpha]} \otimes_{\mathbb{Q}} \mathbb{C} \\ &= C^p_{\mathrm{AH}}\left(X^{2p}_{\overline{\mathbb{Q}}}\right)_{[\alpha]} \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \otimes_\iota \mathbb{C} \\ &= \mathrm{H}^{2p}_{\text{ét}}\left(X^{2p}_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell\right)(p)_{[\alpha]} \otimes_\iota \mathbb{C}, \end{aligned}$$

where the last isomorphism is given by the Betti-to-étale comparison isomorphism. (We remark that these identifications are all G^{2p}-invariant.) For example, in the sequel, we may write bizzarre things such as

$$\gamma^{2p}_{[\alpha],\mathrm{AH}} \otimes 1 \in C^p_{\mathrm{AH}}\left(X^{2p}_{\overline{\mathbb{Q}}}\right) \otimes_{\mathbb{Q}} \overline{\mathbb{Q}} \qquad \text{or} \qquad \nu_\alpha \otimes 1 \in \mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p}, \mathbb{Q}\right) \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}$$

and then pretend that these elements live in the same vector space.

As promised in the previous section, we are able to compute the Galois action on γ. Explicitly, this amounts to the following.

Lemma 4.32. Choose $\alpha \in \mathfrak{B}^{2p}$.

(a) There is a function $\lambda\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{Q}([\alpha])^\times$ such that

$$\sigma\left(\gamma^{2p}_{[\alpha],\mathrm{AH}}\right) = \lambda(\sigma)\gamma^{2p}_{[\alpha],\mathrm{AH}}.$$

(b) For any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and $g \in G^{2p}(\overline{\mathbb{Q}})$, we have

$$\sigma\left(g^*\gamma^{2p}_{[\alpha],\mathrm{AH}}\right) = \lambda(\sigma) \cdot \sigma(g)^*\gamma^{2p}_{[\alpha],\mathrm{AH}}.$$

(c) For any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, we compute $\iota_\alpha(\lambda(\sigma)) \in \mathbb{Q}(\zeta_{2N})$ as

$$\iota_\alpha(\lambda(\sigma)) = \frac{\sigma\left(\mathrm{Per}\left(\gamma^{2p}, \nu_{-\alpha}\right)\right)}{\mathrm{Per}\left(\gamma^{2p}, \nu_{-\alpha}\right)}.$$

Proof. Here, (a) follows because $C^p_{\mathrm{AH}}\left(X^{2p}\right)_{[\alpha]}$ is a one-dimensional vector space over $\mathbb{Q}([\alpha])$ which is stable under the Galois action (because its Betti component is defined over $\mathbb{Q}$); thus, we see that $\gamma^{2p}_{[\alpha],\mathrm{AH}}$ is a basis vector of this space, so (a) follows. Continuing, (b) follows because the action of G^{2p} on X^{2p} is defined over

$\mathbb{Q}$, implying that

$$\begin{aligned}\sigma\left(g^*\gamma_{[\alpha],\mathrm{AH}}^{2p}\right) &= \sigma(g)^*\sigma\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\right)\\ &= \sigma(g)^*\left(\lambda(\sigma)\gamma_{[\alpha],\mathrm{AH}}^{2p}\right)\\ &= \lambda(\sigma)\cdot\sigma(g)^*\gamma_{[\alpha],\mathrm{AH}}^{2p},\end{aligned}$$

where the last equality holds because $\sigma(g)^*$ is linear.

Lastly, (c) will require a computation. We will work in the de Rham component; the idea is to project onto the α-component. Working in $C_{\mathrm{AH}}^p\left(X^{2p}\right)\otimes_{\mathbb{Q}}\mathbb{C}$, one has the equalities

$$\left(\lambda(\sigma)\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right) = \left(\sigma\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right).$$

We now project onto the α-eigenspace; because the G^{2p}-action is defined over $\mathbb{Q}$, the projection commutes with the Galois action, leaving us with

$$\left(\lambda(\sigma)\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right)_\alpha = \sigma\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right)_\alpha.$$

On one hand, by definition of ι_α, we see that the left-hand side will equal $\iota_\alpha(\lambda(\sigma))\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right)$; then projecting onto the de Rham component leaves us with

$$\pi_\infty\left(\left(\lambda(\sigma)\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right)_\alpha\right) = \lambda(\sigma)\,\mathrm{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)\nu_\alpha$$

by Corollary 4.26. On the other hand, for the right-hand side, we will want to project onto the de Rham component first (which commutes with Galois action by our identifications). To complete the proof, we now run computations in $\mathrm{H}_{\mathrm{dR}}^{2p}\left(X^{2p},\mathbb{C}\right) = \mathrm{H}_{\mathrm{dR}}^{2p}\left(X^{2p},\mathbb{Q}\right)\otimes_{\mathbb{Q}}\mathbb{C}$, for which we use Corollary 4.26 to see

$$\begin{aligned}\pi_\infty\left(\sigma\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right)_\alpha\right) &= \sigma\left(\pi_\infty\left(\gamma_{[\alpha],\mathrm{AH}}^{2p}\otimes 1\right)_\alpha\right)\\ &= \sigma\left(\nu_\alpha\otimes\mathrm{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)\right)\\ &= \sigma\left(\mathrm{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)\right)\nu_\alpha,\end{aligned}$$

where the last equality holds because the Galois action on $\mathrm{H}_{\mathrm{dR}}^{2p}\left(X^{2p},\mathbb{Q}\right)$ is trivial. Comparing the previous two computations completes the proof. ∎

We are now ready for our main theorem.

Theorem 4.33. Choose $\alpha\in\mathfrak{B}^{2p}$. For any $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ such that $\sigma(\zeta_N)=\zeta_N^u$ for $u\in(\mathbb{Z}/N\mathbb{Z})^\times$, we have

$$\sigma(\nu_\alpha\otimes 1) = \nu_{u^{-1}\alpha}\otimes\frac{\sigma\left(\mathrm{Per}(\gamma^{2p},\nu_{-u^{-1}\alpha})\right)}{\mathrm{Per}(\gamma^{2p},\nu_{-\alpha})},$$

where this Galois action takes place in $\mathrm{H}_{\mathrm{dR}}^{2p}\left(X^{2p},\mathbb{Q}\right)(p)_{[\alpha]}\otimes_{\mathbb{Q}}\overline{\mathbb{Q}} = C_{\mathrm{AH}}^p\left(X^{2p}\right)_{[\alpha]}\otimes_{\mathbb{Q}}\overline{\mathbb{Q}}$.

Proof. We combine the computed Galois action in Lemma 4.32 with the change-of-basis results Corollaries 4.26 and 4.30. To begin, Corollary 4.30 lets us write

$$\begin{aligned}\sigma(\nu_\alpha\otimes 1) &= \sigma\left(\frac{1}{\#G^{2p}(\overline{\mathbb{Q}})}\sum_{g\in G^{2p}(\overline{\mathbb{Q}})} g^*\gamma_{\mathrm{AH}}^{2p}\otimes\frac{1}{\alpha(g)\,\mathrm{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)}\right)\\ &= \frac{1}{\#G^{2p}(\overline{\mathbb{Q}})}\sum_{g\in G^{2p}(\overline{\mathbb{Q}})}\sigma\left(g^*\gamma_{\mathrm{AH}}^{2p}\otimes\frac{1}{\alpha(g)\,\mathrm{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)}\right)\\ &= \frac{1}{\#G^{2p}(\overline{\mathbb{Q}})}\sum_{g\in G^{2p}(\overline{\mathbb{Q}})}\sigma\left(g^*\gamma_{\mathrm{AH}}^{2p}\right)\otimes\frac{1}{\alpha(g)\,\mathrm{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)},\end{aligned}$$

where the last equality takes place in $C^p_{\mathrm{AH}}\left(X^{2p}\right) \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}$ so that the Galois action is happening in the left component. Continuing, Lemma 4.32 tells us that

$$\sigma\left(g^*\gamma^{2p}_{[\alpha],\mathrm{AH}}\right) = \sigma(g)\lambda(\sigma)\cdot\sigma(g)^*\gamma^{2p}_{[\alpha],\mathrm{AH}},$$

so

$$\sigma(\nu_\alpha \otimes 1) = \frac{1}{\#G^{2p}(\overline{\mathbb{Q}})}\sum_{g\in G^{2p}(\overline{\mathbb{Q}})} \lambda(\sigma)\cdot\sigma(g)^*\gamma^{2p}_{[\alpha],\mathrm{AH}} \otimes \frac{1}{\alpha(g)\operatorname{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)}.$$

(We will wait to evaluate $\lambda(\sigma)$ until the end because a trick is required to move it through the tensor product.) We now go back to the basis of $\nu_\bullet$s via Corollary 4.26, writing

$$\sigma(\nu_\alpha \otimes 1) = \frac{1}{\#G^{2p}(\overline{\mathbb{Q}})}\sum_{\substack{g\in G^{2p}(\overline{\mathbb{Q}})\\ \beta\in[\alpha]}} \lambda(\sigma)\cdot\sigma(g)^*\nu_\beta \otimes \frac{\operatorname{Per}(\gamma^{2p},\nu_{-\beta})}{\alpha(g)\operatorname{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)}.$$

Now, $\sigma(g)^*\nu_\beta \otimes 1 = \nu_\beta \otimes \beta(\sigma(g))$, where the equality is now taking place in $\mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p},\mathbb{Q}\right)\otimes_{\mathbb{Q}}\overline{\mathbb{Q}}$. Continuing, we see $\beta(\sigma(g)) = \sigma(\beta(g)) = \beta(g)^u$ because evaluating a character is Galois-invariant. Rearranging the sums, we now see that we can isolate the sum

$$\frac{1}{\#G^{2p}(\overline{\mathbb{Q}})}\sum_{g\in G^{2p}(\overline{\mathbb{Q}})}\frac{(u\beta)(g)}{\alpha(g)},$$

which orthogonality of characters tells us is the indicator for $\beta = u^{-1}\alpha$. Thus, we are left with

$$\sigma(\nu_\alpha\otimes 1) = \lambda(\sigma)\nu_{u^{-1}\alpha}\otimes\frac{\operatorname{Per}(\gamma^{2p},\nu_{-u^{-1}\alpha})}{\operatorname{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)}.$$

It remains to move $\lambda(\sigma)$ through the tensor product. Note that this is not totally trivial because the tensor product only lets us move rational numbers through. Anyway, it is enough to check the required equality in the de Rham component, allowing us to use the proof of Lemma 4.32 to note

$$\lambda(\sigma)\nu_{u^{-1}\alpha}\otimes\operatorname{Per}\left(\gamma^{2p},\nu_{-u^{-1}\alpha}\right) = \nu_{u^{-1}\alpha}\otimes\sigma\left(\operatorname{Per}\left(\gamma^{2p},\nu_{-u^{-1}\alpha}\right)\right),$$

from which the required result follows after some rearranging. ■

Remark 4.34. Because the G^{2p}-action commutes with the Galois action, it is not difficult to directly check that an α-eigenvector should go to a $u^{-1}\alpha$-eigenvector.

Remark 4.35. As a sanity check, it is not hard to see that Theorem 4.33 actually defines a group representation.

Remark 4.36. Following Remark 2.164, one can use Theorem 4.33 allows ons to compute the connected monodromy field K_A^{conn} of the Jacobian A of any quotient C of the Fermat curve X_N. Indeed, Remark 2.164 explains that this is essentially a matter of computing enough the field of definition of enough Tate classes (used to cut out the torus $G^\circ_\ell(A)$). In particular, we already know that $\mathbb{Q}(\zeta_N)\subseteq K_A^{\mathrm{conn}}$ (because of the endomorphisms), and then Theorem 4.33 explains that $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_N))$ fixes a Tate cycle ν_α if and only if it fixes the period $\operatorname{Per}\left(\gamma^{2p},\nu_{-\alpha}\right)$.

Let's see an example.

Corollary 4.37. Choose $\alpha := (a,b,c) \in \mathfrak{A}^1$, and set $\alpha' := (a',b',c')$ to be $-\alpha$. Then $(\alpha,\alpha') \in \mathfrak{B}^2$, and for any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ such that $\sigma(\zeta_N) = \zeta_N^u$ for $u \in (\mathbb{Z}/N\mathbb{Z})^\times$, we have

$$\sigma(\nu_{(\alpha,\alpha')} \otimes 1) = \nu_{u^{-1}(\alpha,\alpha')} \otimes (-1)^{\langle u^{-1}\alpha\rangle - \langle\alpha\rangle}.$$

In particular, σ fixes $\nu_{(\alpha,\alpha')} \otimes 1$ if and only if $u-1$ is divisible by $N/\gcd(a,b,c,N)$.

Proof. To see that $(\alpha,\alpha') \in \mathfrak{B}^2$, we note that any $u \in (\mathbb{Z}/N\mathbb{Z})^\times$ still has $u\alpha = -u\alpha'$, so $\{\langle u\alpha\rangle, \langle -u\alpha\rangle\} = \{1,2\}$.

Looking at Theorem 4.33, we see the main part of proof will be computing our periods. The main point is that the reflection formula for Γ (recalled later in Proposition 4.40) reassures us that

$$\Gamma\left(\frac{[a]}{N}\right)\Gamma\left(\frac{[-a]}{N}\right) = \frac{\pi}{\sin\frac{a\pi}{N}}.$$

We now combine this with the computation in Remark 4.29 to achieve

$$\mathrm{Per}\left(\gamma^{2p}, \nu_{-(\alpha,\alpha')}\right) = -(2\pi i)^{-1} \cdot \zeta_{2N}^{[-a]+[-b]+[a]+[b]} \cdot \frac{\pi}{\sin\frac{[a]\pi}{N}} \cdot \frac{\pi}{\sin\frac{[b]\pi}{N}} \cdot \frac{\sin\frac{[c]\pi}{N}}{\pi}.$$

Note that $[a]+[-a] = N$, so the power of ζ_{2N} disappears. Continuing, we expand $\sin z = \frac{1}{2i}\left(z + z^{-1}\right)$, which yields

$$\mathrm{Per}\left(\gamma^{2p}, \nu_{-(\alpha,\alpha')}\right) = -\frac{\left(\zeta_{2N}^c - \zeta_{2N}^{-c}\right)}{\left(\zeta_{2N}^a - \zeta_{2N}^{-a}\right)\left(\zeta_{2N}^b - \zeta_{2N}^{-b}\right)}.$$

Continuing, we factor $\zeta_{2N}^c/\zeta_{2N}^{-a-b} = \zeta_{2N}^{N\langle\alpha\rangle} = (-1)^{\langle\alpha\rangle}$, leaving us with

$$\mathrm{Per}\left(\gamma^{2p}, \nu_{-(\alpha,\alpha')}\right) = -(-1)^{\langle\alpha\rangle} \cdot \frac{\left(1 - \zeta_N^{-c}\right)}{\left(\zeta_N^a - 1\right)\left(\zeta_N^b - 1\right)}.$$

We now plug into Theorem 4.33 to reveal

$$\sigma(\nu_{(\alpha,\alpha')} \otimes 1) = \nu_{u^{-1}(\alpha,\alpha')} \otimes \frac{\sigma\left((-1)^{\langle u^{-1}\alpha\rangle} \cdot \frac{\left(1-\zeta_N^{-u^{-1}c}\right)}{\left(\zeta_N^{u^{-1}a}-1\right)\left(\zeta_N^{u^{-1}b}-1\right)}\right)}{(-1)^{\langle\alpha\rangle} \cdot \frac{\left(1-\zeta_N^{-c}\right)}{\left(\zeta_N^a-1\right)\left(\zeta_N^b-1\right)}},$$

which rearranges into the desired expression because $\sigma\left(\zeta_N^{u^{-1}}\right) = \zeta_N$.

It now remains the last sentence. Well, we see that σ fixes $\nu_{(\alpha,\alpha')}$ if and only if $u^{-1}\alpha = \alpha$, which is equivalent to $u\alpha = \alpha$. By taking $\mathbb{Z}$-linear combinations, it is equivalent to asking for $(u-1)\gcd(a,b,c) \equiv 0 \pmod N$, from which the claim follows. ∎

4.2.4 Some Examples

We begin with the superelliptic curve $C\colon y^9 = x^3 - 1$.

Proposition 4.38. Define A to be the Jacobian of the proper curve C with affine chart $y^9 = x^3 - 1$. Then we show $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$, and we compute $\mathrm{ST}(A)$.

Proof. We will freely use the computation executed in Proposition 3.32. Thoughout, $A := \mathrm{Jac}\, C$, and we recall that we have a decomposition $A = C_0 \times A_1 \times A_2$ (over $\mathbb{Q}$) into geometrically simple abelian varieties. We proceed in steps.

1. Even though this is not a Fermat curve, it is a quotient of the Fermat curve X_N with $N \coloneqq 9$: this is witnessed by the quotient map from the affine patch $x^9+y^9+1=0$ to C given by $\psi(x,y) \coloneqq \left(-x^3, y\right)$. Thus, we will be able to use the Galois-invariant embedding $\psi\colon \mathrm{H}^1_{\text{ét}}(C_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell) \hookrightarrow \mathrm{H}^1_{\text{ét}}(X_{N,\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)$ to use Theorem 4.33 by restricting to the Galois submodule. To make this explicit, we recall that we have a basis
$$\left\{\frac{dx}{y^4}, \frac{dx}{y^5}, \frac{dx}{y^6}, \frac{dx}{y^7}, \frac{dx}{y^8}, \frac{x\,dx}{y^7}, \frac{x\,dx}{y^8}\right\}$$
of $\mathrm{H}^{10}(C)$, we see that we can pass this basis through ψ^* to see that $\mathrm{H}^{10}(C) \subseteq \mathrm{H}^{10}(X)$ has basis
$$\left\{\nu_{351}, \nu_{342}, \nu_{333}, \nu_{324}, \nu_{315}, \nu_{621}, \nu_{612}\right\}.$$
Combining with the conjugate differentials yeilds a full basis of $\mathrm{H}^1_{\mathrm{dR}}(C,\mathbb{Q}) \subseteq \mathrm{H}^1_{\mathrm{dR}}(X,\mathbb{Q})$.

2. We now explain how to pass the étale site. By Conjecture 3.19, which is known in this case by Theorem 3.23, we may choose any ℓ, so we choose ℓ so that $\mathbb{Q}_\ell$ contains any algebraic numbers we will need in the sequel (most notably, we want ζ_N and our periods). For each $p \geq 0$, we recall that any $\alpha \in \mathfrak{B}^{2p}$ produces idenitifications
$$\mathrm{H}^{2p}_{\mathrm{dR}}\left(X^{2p},\mathbb{Q}\right)_{[\alpha]} \otimes_{\mathbb{Q}} \mathbb{C} = C^p_{\mathrm{AH}}\left(X^{2p}\right)_{[\alpha]} \otimes_{\mathbb{Q}} \mathbb{C} \hookrightarrow \mathrm{H}^{2p}_{\text{ét}}\left(X^{2p}, \mathbb{Q}_\ell\right)(p)_{[\alpha]} \otimes_\iota \mathbb{C},$$
where $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$ is some fixed embedding. In this way, we see that we are allowed to treat an expression like $\nu_{351} \otimes 1$ as an element of $\mathrm{H}^{2p}_{\text{ét}}\left(X^{2p}, \mathbb{Q}_\ell\right) \otimes_\iota \mathbb{C}$; for carefully chosen ℓ, a Galois descent argument is even able to reassure us that the basis vectors $\nu_\alpha \otimes 1$ produces from the previous step can be found in $\mathrm{H}^{2p}_{\text{ét}}\left(X^{2p}, \mathbb{Q}_\ell\right)(p)_{[\alpha]}$.
Thus, in the notation of Proposition 3.32, we see that ψ^* pulls the basis vectors $\{u_1 \otimes 1, v_1 \otimes 1, v_2 \otimes 1, v_4 \otimes 1, w_1 \otimes 1, w_2 \otimes 1, w_5 \otimes 1\}$ to
$$\{\nu_{333} \otimes 1, \quad \nu_{315} \otimes 1, \nu_{621} \otimes 1, \nu_{342} \otimes 1, \quad \nu_{612} \otimes 1, \nu_{324} \otimes 1, \nu_{351} \otimes 1\},$$
and one can recover ψ^* on the rest of the basis by taking conjugates.

3. We are now ready to begin executing Proposition 2.159; for this, Remark 2.160 informs us that we need to build a space of W' of Tate classes cutting out $G_\ell(A)^\circ \subseteq \mathrm{GL}_{14,\mathbb{Q}_\ell}$. We begin by adding W_1, made up of the endomorphisms, which ensures (for example) that $G_\ell(A)^\circ$ is diagonal. Then Proposition 3.32 computed that we also have the "polarization equations"
$$\begin{aligned}\mu_1\mu_2 &= \kappa_1\kappa_8,\\ \kappa_1\kappa_8 &= \kappa_2\kappa_7,\\ \kappa_1\kappa_8 &= \kappa_4\kappa_5,\end{aligned}$$
and the exceptional equation
$$\mu_1\kappa_7 = \kappa_5\kappa_8.$$
We remark that the polarization equations translate into a Tate class like $\nu_{(\alpha,-\alpha,\beta,-\beta)} \otimes 1$ understood as an element in $\mathrm{H}^4_{\text{ét}}(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)(2) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$, but this Tate class actually already come from a class in W_1 (see Corollary 4.37), so we may safely ignore it. Thus, we only have to translate the exceptional equation into the tensor
$$\nu_{333,675,648,612} \otimes 1 \in \mathrm{H}^4_{\text{ét}}(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)(2) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$$
and its Galois orbit.

4. We claim that $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_N)$. By Remark 2.160, it is enough to know that $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ is the largest subgroup of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixing W'. We already know that our endomorphisms, except the isogeny $(A_1)_{\overline{\mathbb{Q}}} \cong (A_2)_{\overline{\mathbb{Q}}}$, are defined over $\mathbb{Q}(\zeta_N)$ (see also Corollary 4.37 for these equations and the polarization). The isogeny corresponds to equations $\kappa_u = \lambda_{2u}$ for each $u \in (\mathbb{Z}/9\mathbb{Z})^\times$, which means that we would like to check that
$$\mathrm{Per}\left(\gamma^{2p}, \nu_{u(612,378)}\right)$$

is in $\mathbb{Q}(\zeta_N)$. Well, by Remark 4.14, this element is

$$(-2\pi i)^{-1}\zeta_{2N}^{u(6+2+3+7)}\cdot\frac{\Gamma\left(\frac{[6u]}{9}\right)\Gamma\left(\frac{[2u]}{9}\right)}{\Gamma\left(\frac{[8u]}{9}\right)}\cdot\frac{\Gamma\left(\frac{[3u]}{9}\right)\Gamma\left(\frac{[7u]}{9}\right)}{\Gamma\left(\frac{[u]}{9}\right)}.$$

A quick application of the reflection formula as in Corollary 4.37 shows this is in $\mathbb{Q}(\zeta_N)$.

It remains to check that $\sigma\in\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ fixes the Galois orbit of $\nu_{333,675,648,612}\otimes 1$. Well, looking at Theorem 4.33, it is enough to check that σ fixes

$$\mathrm{Per}\left(\gamma^4,\nu_{u(333,675,648,612)}\right)$$

for any $u\in(\mathbb{Z}/N\mathbb{Z})^\times$. Well, by Remark 4.14, we see this equals

$$(-2\pi i)^{-2}\zeta_{2N}^{u(3+3+6+7+6+4+6+1)}\cdot\frac{\Gamma\left(\frac{[3u]}{9}\right)\Gamma\left(\frac{[3u]}{9}\right)}{\Gamma\left(\frac{[6u]}{9}\right)}\cdot\frac{\Gamma\left(\frac{[6u]}{9}\right)\Gamma\left(\frac{[7u]}{9}\right)}{\Gamma\left(\frac{[4u]}{9}\right)}\cdot\frac{\Gamma\left(\frac{[6u]}{9}\right)\Gamma\left(\frac{[4u]}{9}\right)}{\Gamma\left(\frac{[u]}{9}\right)}\cdot\frac{\Gamma\left(\frac{[6u]}{9}\right)\Gamma\left(\frac{[u]}{9}\right)}{\Gamma\left(\frac{[7u]}{9}\right)}.$$

After the dust settles, we are left with

$$(-2\pi i)^{-2}\cdot\Gamma\left(\frac{3}{9}\right)^2\Gamma\left(\frac{6}{9}\right)^2.$$

Now, the reflection formula yields $\Gamma\left(\frac{3}{9}\right)\Gamma\left(\frac{6}{9}\right)=\frac{\pi}{\sin\frac{\pi}{3}}$, so we see that this period lives in $\mathbb{Q}(\zeta_N)$ and hence is fixed by σ; in fact, it is rational!

5. Choose $\sigma\in\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ to satisfy $\sigma(\zeta_N)=\zeta_N^u$. We compute the action of σ on W'. For example, the previous step actually shows that σ fixes the Galois orbit of $\nu_{333,675,648,612}\otimes 1$, so it remains to compute the action on W_1. Note that G acts on the $\mathbb{C}$-vector space, so the action can be diagonalized. Given some character $(\alpha,\beta)\in\mathfrak{A}^2$, we note that $(W_1)_{(\alpha,\beta)}$ is at most one-dimensional spanned by $\nu_{(\alpha,\beta)}\otimes 1$, and this element being a Tate class is equivalent to $\mathrm{H}^2_{\mathrm{B}}(X,\mathbb{Q})(1)_{[\alpha]}$ has Hodge cycles by the Mumford–Tate conjecture (known in this case by Remark 2.146), which is equivalent to $(\alpha,\beta)\in\mathfrak{B}^2$ by Proposition 4.23. With the aide of a computer, we can enumerate all such (α,β), and we see that they come in two forms.

   - We could have $\alpha=(a,b,c)$ and $\beta=-\alpha$. In this case, Corollary 4.37 explains that

   $$\sigma(\nu_{(\alpha,\beta)}\otimes 1)=\nu_{u^{-1}(\alpha,\beta)}\otimes(-1)^{\langle u^{-1}\alpha\rangle-\langle\alpha\rangle}.$$

   - We could have $\alpha=(a,b,c)$ and $\beta=(-a,-c,-b)$. As in Corollary 4.37, the main point is to compute our periods. Well, by Remark 4.14, we find

   $$\mathrm{Per}\left(\gamma^{2p},\nu_{-(\alpha,\beta)}\right)=-(2\pi i)^{-1}\zeta_{2N}^{[a]+[-a]+[-b]+[c]}\cdot\frac{\Gamma\left(\frac{[-a]}{N}\right)\Gamma\left(\frac{[-b]}{N}\right)}{\Gamma\left(\frac{[c]}{N}\right)}\cdot\frac{\Gamma\left(\frac{[a]}{N}\right)\Gamma\left(\frac{[c]}{N}\right)}{\Gamma\left(\frac{[-b]}{N}\right)},$$

   which after an appliation of the reflection formula gives

   $$\begin{aligned}\mathrm{Per}\left(\gamma^{2p},\nu_{-(\alpha,\beta)}\right)&=(2\pi i)^{-1}\zeta_{2N}^{[-b]+[c]}\cdot\frac{\pi}{\sin\frac{a\pi}{N}}\\&=\frac{\zeta_{2N}^{[-b]+[c]}}{\zeta_{2N}^{a}-\zeta_{2N}^{-a}}\\&=\frac{\zeta_{2N}^{[-b]+[c]+[a]}}{\zeta_N^a-1}.\end{aligned}$$

It will be convenient to write this entirely in terms of ζ_N, so we note that N being odd forces $\zeta_{2N} = -\zeta_N^{(N+1)/2}$, so this equals $(-1)^{[a]+[-b]+[c]}\zeta_N^{(a-b+c)(N+1)/2}$. The purpose of this rewrite is that all ζ_Ns will go away in the computation

$$\sigma(\nu_{(\alpha,\beta)} \otimes 1) = \nu_{u^{-1}(\alpha,\beta)} \otimes \frac{\sigma\left(\mathrm{Per}\left(\gamma^{2p}, \nu_{-u^{-1}(\alpha,\beta)}\right)\right)}{\mathrm{Per}\left(\gamma^{2p}, \nu_{-(\alpha,\beta)}\right)}$$

because $\sigma\left(\zeta_N^{u^{-1}}\right) = 1$, so we are left with

$$\sigma(\nu_{(\alpha,\beta)} \otimes 1) = \nu_{u^{-1}(\alpha,\beta)} \otimes (-1)^{[u^{-1}a]+[-u^{-1}b]+[u^{-1}c]+[a]+[-b]+[c]}.$$

6. Now choose $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ to satisfy $\sigma(\zeta_N) = \zeta_N^5$, which we note is a generator. We now compute

$$\{g \in \mathrm{GL}_{14,\mathbb{Q}_\ell} : g|_{W'} = \sigma|_{W'}\}.$$

For this, we recall from Proposition 2.159 that we are looking at the component of $G_\ell(A)$ containing the image of σ. In particular, we know that σ is a permutation matrix sending $(\nu_\alpha \otimes 1) \mapsto (\nu_{2\alpha} \otimes 1)$ (up to scalar), so we need g to also be a permutation matrix also sending $(\nu_\alpha \otimes 1) \mapsto (\nu_{2\alpha} \otimes 1)$ (again up to scalar). Well, for each available α, we will compute relations among scalars $\{\lambda_\alpha\}$ defined to satisfy $g(\nu_\alpha \otimes 1) = (\nu_{2\alpha} \otimes \lambda_\alpha)$. Because $G_\ell(A)^\circ$ is a torus of rank 4, we are expecting to be able to write all $\lambda_\bullet$s in terms of four of them.

With this in mind, we use the previous step as follows to produce the required relations. For brevity, let λ be the multiplier of g with respect to the pairing induced by the polarization; this multiplier becomes the action of g on $\mathbb{Q}_\ell(1)$.

- We need g to satisfy
$$g(\nu_{(\alpha,-\alpha)} \otimes 1) = \nu_{2(\alpha,-\alpha)} \otimes (-1)^{\langle 2\alpha\rangle - \langle\alpha\rangle},$$
so $\lambda_\alpha\lambda_{-\alpha} = (-1)^{\langle 2\alpha\rangle - \langle\alpha\rangle}\lambda$.
- For available (a,b,c), we need g to satisfy
$$g(\nu_{(a,b,c,-a,-c,-b)}) = \nu_{(2a,2b,2c,-2a,-2c,-2b)} \otimes (-1)^{[2a]+[-2b]+[2c]+[a]+[-b]+[c]},$$
so $\lambda_{(a,b,c)}\lambda_{(-a,-c,-b)} = \lambda(-1)^{[2a]+[-2b]+[2c]+[a]+[-b]+[c]}$. For convenience, we note that $\pmod 2$ computations have
$$[2a] + [-2b] + [2c] + [a] + [-b] + [c] \equiv [2a] + [2b] + [2c] + [a] + [b] + [c] \equiv \langle 2\alpha\rangle - \langle\alpha\rangle,$$
so we are seeing the same sign as before.
- We need g to fix $\nu_{u(333,675,648,612)}$, so $\lambda_{u(333)}\lambda_{u(675)}\lambda_{u(648)}\lambda_{u(612)} = \lambda^2$.

The above points tell us that we can determine g uniquely by choosing $(\kappa_1, \kappa_2, \kappa_4) = (\lambda_{612}, \lambda_{324}, \lambda_{648})$ and λ. Explicitly, we get the matrix

$$\left[\begin{array}{ccc}
-\kappa_1\kappa_4/\kappa_2 & & \\
\lambda\kappa_2/\kappa_1\kappa_4 & & \\
\hline
 & \lambda/\kappa_2 & \\
 & \lambda/\kappa_4 & \\
 & \kappa_1 & \\
 & & \lambda/\kappa_1 \\
 & & \kappa_4 \\
 & -\kappa_2 & \\
\hline
 & & -\lambda/\kappa_4 \\
 & & \kappa_1 \\
 & & \kappa_2 \\
 & & \lambda/\kappa_2 \\
 & & \lambda/\kappa_1 \\
 & & \kappa_4
\end{array}\right]$$

as representing g.

Thus, upon enforcing the multiplier to equal 1 and base-changing to $\mathbb{C}$, we see that $\mathrm{ST}(A)$ is generated by $\mathrm{ST}(A)^\circ$ (computed in Proposition 3.32) and the matrix

$$\begin{bmatrix} 1 & -1 & & & & & & & \\ & & & 1 & & & & & \\ & & 1 & & & & & & \\ & & & & & 1 & & & \\ & & & & & & 1 & & \\ & & & -1 & & & & & \\ & & & & & & & -1 & \\ & & & & & & 1 & & \\ & & & & & 1 & & & \\ & & & & & & & 1 & \\ & & & & & & & & 1 \end{bmatrix}.$$

This completes our computation. ■

We now use the above computation to compute the Sato–Tate group of some generic superelliptic curves.

Theorem 4.39. For given $\lambda \in \mathbb{Q}(\zeta_9) \setminus \{0,1\}$, define A to be the Jacobian of the proper curve $\widetilde{C}$ with affine chart $y^9 = x(x-1)(x-\lambda)$. Suppose that A does not have complex multiplication. Then we show $K_A^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$, and we compute $\mathrm{ST}(A)$.

Proof. As usual, we proceed in steps. Throughout, we freely use the computation of Proposition 3.29.

0. Quickly, we note that we may pass from $y^9 = x(x-1)(x-\lambda)$ (for $\lambda \notin \{0,1\}$) to $y^9 = \left(x^2+x+1\right)(x-\lambda)$ (for $\lambda \notin \{\zeta_3, \overline{\zeta}_3\}$). Indeed, consider the isomorphism $f\colon \mathbb{P}^1 \to \mathbb{P}^1$ defined over $\mathbb{Q}(\zeta_9)$ by fixing ∞ and sending $0 \mapsto \zeta_3$ and $1 \mapsto \overline{\zeta}_3$. Then the curves $y^9 = x(x-1)(x-\lambda)$ and $y^9 = \left(x^2+x+1\right)(x-f(\lambda))$ are isomorphic by an isomorphism of the "ground" $\mathbb{P}^1$. Thus, the connected monodromy field over $\mathbb{Q}(\zeta_9)$ of both curves must be the same. Because K_A^{conn} for both curves must contain $\mathbb{Q}(\zeta_9)$ anyway (there are endomorphisms whose field of definition is $\mathbb{Q}(\zeta_9)$ already), we see that this movement must be harmless!

1. We lift our situation to an abelian scheme. Let S be $\mathbb{A}^1_{\mathbb{Q}} \setminus \{\zeta_3, \overline{\zeta}_3\}$, and we let $\mathcal{C} \to S$ be the curve cut out by the equation $y^9 = \left(x^2+x+1\right)(x-\lambda)$ as λ varies over S; then we can normalize and complete C to produce a family of smooth projective curves $\widetilde{\mathcal{C}} \to S$. Then $\mathcal{A} \coloneqq \mathrm{Pic}^0 \mathcal{C}/S$ is an abelian scheme over S. In particular, for each $\lambda \in \mathbb{Q} \setminus \{\zeta_3, \overline{\zeta}_3\}$, we can specialize to $\lambda \in S$ to produce $A_\lambda \coloneqq \mathcal{A}_\lambda$ as the Jacobian of the curve $\widetilde{C}_\lambda \coloneqq \widetilde{\mathcal{C}}_\lambda$.

   While we're here, we set up a family of Galois representations. In order to avoid any difficult étale cohomology, we will do this cheaply using the Tate module. For each $n \geq 1$, we have a finite flat group scheme $\mathcal{A}[n] \to S$, so each $\lambda \in S(\mathbb{Q})$ gets a natural Galois-invariant pullback square as follows.

$$\begin{array}{ccc} A_\lambda[n] & \longrightarrow & \mathcal{A}[n] \\ \downarrow & & \downarrow \\ \lambda & \longrightarrow & S \end{array}$$

   Taking limits over n, we get Galois-invariant inclusions $V_\ell A \to V_\ell \mathcal{A}$, where $V_\ell \mathcal{A}$ can be interpreted as a sheaf with stalks given by $V_\ell A$. The moral of the story is that we will be able to use a special point in S in order to compute the Galois action for generic $\lambda \in S$.

2. As before, we will use Proposition 2.159 in order to compute $G_\ell(A_\lambda)$ when A_λ does not have complex multiplication. Thus, Remark 2.160 asks us to find a space W' of Tate classes cutting out $G_\ell(A)^\circ$. We may as well work with $\mathrm{MT}(A)$ by the Mumford–Tate conjecture, which is known in our case by Proposition 2.152. As before, we go ahead and add in W_1 to account for the endomorphisms of A. We also add the class of the polarization to W'. Thus, our Tate classes so far cut out $\mathrm{L}(A)$. The computation

of Proposition 3.29 tells us that $\mathrm{MT}(A)$ is $\mathrm{L}(A)$ cut out by one additional equation from the center, given by

$$\lambda_1\lambda_4\lambda_7 = \lambda_2\lambda_5\lambda_8.$$

Writing $g \in \mathrm{MT}(A) \subseteq \mathrm{L}(A)$ as $\mathrm{diag}(g_0, g_1, g_2, g_4, g_5, g_7, g_8)$ as in Proposition 3.29, we see that the above equation corresponds to the equation

$$\det g_1 g_4 g_7 = \det g_2 g_5 g_8,$$

which we see corresponds to the exceptional Tate class

$$(v_1 \wedge v_1') \otimes (v_4 \wedge v_4') \otimes (v_7 \wedge v_7') \otimes 1 \in \mathrm{H}^6_{\text{ét}}(A_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)(3),$$

where $v \wedge v' = \frac{1}{2}(v \otimes v' - v' \otimes v)$. Explicitly, the computation of Proposition 3.29 tells us that $g \in G_\ell(A)$ acts on $(v_1 \wedge v_1') \otimes (v_4 \wedge v_4') \otimes (v_7 \wedge v_7')$ by some power of the multiplier, which is then cancelled out some by the Tate twist.

3. We claim that $K_{A_\lambda}^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$ for generic λ. Our endomorphisms come from the automorphisms of the curve, which are all defined over $\mathbb{Q}(\zeta_9)$. Additionally, the polarization is certainly defined over $\mathbb{Q}(\zeta_9)$.

   It remains to handle the Galois orbit of the exceptional class given in the previous step. By the discussion at the end of the first step, it is enough to compute the Galois action at a single λ where this Tate class can be found. Well, we take $\lambda = 1$ so that we can appeal to the computations of Proposition 4.38. To explicate our basis, we will take $\{v_1, \dots, v_8\} = \{v_1, \dots, v_8\}$ and $\{v_1', \dots, v_8'\} = \{w_1, \dots, w_8\}$. Unravelling the Tate class, we see that it is a linear combination of the Tate classes given by permuting the triples in the subscript of the Tate class

   $$\nu_{315,612,342,648,378,675}.$$

   (We also need to consider $\nu_{2(315,612,342,648,378,675)}$ for the full Galois orbit, but the computation is essentially the same.) We would like to check that this Tate class is defined over $\mathbb{Q}(\zeta_9)$. Well, by Theorem 4.33, it is enough to check that the period

   $$\mathrm{Per}\left(\gamma^6, \nu_{315,612,342,648,378,675}\right)$$

   lives in $\mathbb{Q}(\zeta_9)$. After expanding the Γs, we are eventually left with some power of ζ_9 multiplied by

   $$\left((2\pi i)^{-1}\Gamma\left(\frac{3}{9}\right)\Gamma\left(\frac{6}{9}\right)\right)^3,$$

   which we see is in $\mathbb{Q}(\zeta_9)$.

4. We compute $G_\ell(A_\lambda)$ for generic λ. Above we computed that the Tate classes cutting out $G_\ell(A_\lambda)$ for generic λ are a strict subset of those needed for $\lambda = 1$, so one finds that $G_\ell(A_1) \subseteq G_\ell(\lambda)$ for generic λ. In particular, Proposition 4.38 tells us that $G_\ell(A_\lambda)$ must contain

   $$\begin{bmatrix} & -1 & & & & & & \\ 1 & & & & & & & \\ & & & & 1 & & & \\ & & & & & -1 & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & & 1 & \\ & & & & & & & 1 \\ & & & & & & & 1 \\ & & & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \end{bmatrix},$$

   where we have reordered the basis. However, having $K_{A_\lambda}^{\mathrm{conn}} = \mathbb{Q}(\zeta_9)$ implies by Proposition 2.159 that $[G_\ell(A_\lambda) : G_\ell(A_\lambda)^\circ] = 9$, and we can see that the group generated by $G_\ell(A)^\circ$ and the above matrix also has $G_\ell(A_\lambda)^\circ$ as an index-9 subgroup. Thus, we conclude that $G_\ell(A_\lambda)$ is generated by $G_\ell(A)^\circ$ (computed in Proposition 3.29) and the above matrix.

5. We conclude that $\mathrm{ST}(A)$ equals is generated by $\mathrm{ST}(A)^\circ$ (computed in Proposition 3.29) and the matrix given in the previous step. This completes the computation. ■

4.3 Calculations of the Periods

Our calculation of the Galois action on absolute Hodge cycles above (Theorem 4.33) found that the main difficulty reduces to a computation of the periods $\mathrm{Per}\left(\gamma^{2p}, \nu_\alpha\right)$. In general, it is not an easy problem to compute the periods of a variety, even an abelian variety with complex multiplication. However, we have already put in a lot of work into being able to do this: Remark 4.29 explains that $\alpha \in \mathfrak{B}^{2p}$ will have

$$\mathrm{Per}\left(\gamma^{2p}, \nu_\alpha\right) = (2\pi i)^{-p} \prod_{i=1}^{2p} \zeta_{2N}^{[a_i]+[b_i]} \frac{\Gamma\left(\frac{[a_i]}{N}\right)\Gamma\left(\frac{[b_i]}{N}\right)}{\Gamma\left(\frac{[-c_i]}{N}\right)}.$$

It remains to compute these ratios, which comes down to being able to do arithmetic with products of Γ-functions. This is the primary goal of this section.

4.3.1 Properties of Γ

To set ourselves up for the remaining subsections, we will now prove all needed properties of the Γ-function

$$\Gamma(s) \coloneqq \int_{\mathbb{R}^+} t^s e^{-t} \frac{dt}{t}$$

from scratch. We will be rather streamlined. Our end goal is to prove the following proposition.

Proposition 4.40. The function $\Gamma(s)$ admits a meromorphic continuation to $\mathbb{C}$ with only simple poles at the nonpositive integers. Further, it satisfies the following properties.

(a) Translation: $\Gamma(s+1) = s\Gamma(s)$.

(b) Reflection: $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$.

(c) Multiplication: for any positive integer d,

$$\Gamma(s)\Gamma\left(s+\frac{1}{d}\right)\Gamma\left(s+\frac{2}{d}\right)\cdots\Gamma\left(s+\frac{d-1}{d}\right) = (2\pi)^{(d-1)/2} d^{1/2-ds}\Gamma(ds).$$

Among (a)–(c), only (a) admits a quick proof.

Proof of Proposition 4.40(a). Assuming that the integral form is well-defined, we find that the result holds by integration by parts.

$$\begin{aligned}
\Gamma(s+1) &= \int_{\mathbb{R}^+} t^{s+1} e^{-t} \frac{dt}{t} \\
&= -\int_{t\in\mathbb{R}^+} t^s \, d\left(e^{-t}\right) \\
&= -t^s e^{-t}\Big|_{t=0}^{t=\infty} + s\int_{\mathbb{R}^+} t^s e^{-t}\frac{dt}{t} \\
&= s\Gamma(s),
\end{aligned}$$

as required. ∎

Example 4.41. A direct integral computation shows that $\Gamma(1) = 1$, so we note that we may read the integration by parts above backwards to see that we have shown that the integral defining $\Gamma(n)$ converges and equals $(n-1)!$ for any positive integer n.

Now that we have some idea how to bound the integral defining Γ, we are able to prove the meromorphic continuation.

Proof of meromorphic continuation of Proposition 4.40. We have two steps.

1. We claim that the integral converges absolutely and uniformly on compacts in the region $\{s : \operatorname{Re} s > 0\}$, which will prove that Γ is holomorphic there. Here, we may bound the integral absolutely by
$$\int_{\mathbb{R}^+} \left|t^s e^{-t}\right| \, dt \le \int_0^1 t^{\operatorname{Re} s - 1} \, dt + \int_1^\infty t^{\lceil \operatorname{Re} s - 1 \rceil} e^{-t} \, dt.$$
The left integral equals $\frac{1}{\operatorname{Re} s - 1}$, so it converges absolutely on compacts. The right integral is bounded by $\Gamma(\lceil \operatorname{Re} s - 1 \rceil)$, which we know by Example 4.41 to converge.

2. We complete the meromorphic continuation. The equation $\Gamma(s+1) = s\Gamma(s)$ allows us to inductively holomorphically continue $\Gamma(s)$ to the region $\mathbb{C} \setminus \{0, -1, -2, \dots\}$. This equation written as $\Gamma(s) = \frac{1}{s}\Gamma(s+1)$ also explains that Γ admits a simple pole at $s = 0$, which can then be inductively continued to produce simple poles on the nonpositive integers. ■

Example 4.42. We compute $\Gamma(1/2)$. The proof above shows that the integral converges, so we would like to compute $\int_{\mathbb{R}^+} t^{-1/2} e^{-t} \, dt$. Taking $u = \sqrt{t}$, we see that $2\,du = t^{-1/2}\,dt$, so
$$\Gamma(1/2) = \int_{\mathbb{R}} e^{-u^2} \, du.$$
The technique of squaring the integral and passing to polar coordinates shows that the integral equals $\sqrt{\pi}$.

We now turn to the reflection formula.

Proof of Proposition 4.40(b). We will have to do some work. The following slick argument is taken from David Speyer, who credits Paul Monsky [Spe]. We will show that the function $f(s) := \Gamma(s)\Gamma(1-s)\sin \pi s$ is constant. Note that this will complete the proof because we can compute the constant is π by writing
$$f(1/2) = \Gamma(1/2)^2 \sin \frac{\pi}{2}$$
and using Example 4.42. We now proceed in steps. The idea is that the ambient 1-periodicity of f means that we only have worry about bounds on $f(x+iy)$ as $|y| \to \infty$.

1. We claim that there is a holomorphic function $g \colon \mathbb{C}^\times \to \mathbb{C}$ such that $f(s) = g\left(e^{2\pi i s}\right)$. To begin, note that $\Gamma(s)$ has simple poles at the nonpositive integers, so $\Gamma(1-s)$ has simple poles at the positive integers, so $f(s)$ is entire. Furthermore, we claim that $f(s+1) = f(s)$. By analytic continuation, it is enough to check this away from the real axis. Because the function $\sin \pi s$ satisfies $\sin \pi(s+1) = -\sin \pi s$, it is enough to compute
$$\Gamma(s+1)\Gamma(1-(s+1)) = s\Gamma(s) \cdot \frac{1}{-s}\Gamma(1-s).$$
We now turn towards defining g. The function $s \mapsto e^{2\pi i s}$ is an entire surjection $\mathbb{C}^\times \to \mathbb{C}^\times$ with non-vanishing derivative everywhere, so one can at least locally invert it. Thus, we may use a local inverse suitably composed with f to define g locally. This local definition of g however extends to a definition on all $\mathbb{C}^\times$ because $f(s+1) = f(s)$.

2. We provide some bounds on the function g. We begin with some bounds on Γ: if $x \in [0,1]$ and $|y| \geq 1$, then

$$\begin{aligned}|\Gamma(x+iy)| &= \left|\frac{1}{x+iy}\right| \cdot |\Gamma((x+1)+iy)| \\ &\leq \Gamma(x+1) \\ &\leq \max_{x\in[1,2]} \Gamma(x),\end{aligned}$$

which is absolutely bounded by some constant C. Moving to f, we see

$$|f(x+iy)| \leq C^2 e^{-\pi y}.$$

Lastly, moving to g, we see that $\left|g\left(e^{2\pi i(x+iy)}\right)\right| \leq C^2 e^{-\pi y}$. We evaluate this in two extreme cases: sending $y \to \infty$ tells us that $|g(q)| \leq C^2 |q|^{1/2}$ as $|w| \to \infty$; on the other hand, sending $y \to -\infty$ tells us that $|g(q)| \leq C^2 |q|^{-1/2}$ as $q \to 0$.

3. We complete the proof. Our goal is to show that f is constant, so it is enough to show that g is constant. It is enough to show that $g(s)$ and $g(1/s)$ both extend to holomorphic functions at $s = 0$ because this will imply that g extendeds to a bounded holomorphic function, which is constant.

   It is therefore enough to show the following lemma in complex analysis: suppose $g\colon B(0,1) \setminus \{0\} \to \mathbb{C}$ is a holomorphic function such that $|g(q)| \leq |q|^{-1/2}$ as $q \to 0$. Then we want to show that g extends to a holomorphic function at 0. Well, the function $g_1(q) := qg(q)$ continues to be holomorphic on $B(0,1) \setminus \{0\}$, but now we see that it has a removable singularity at 0 with $qg(q) \to 0$ as $q \to 0$, so g_1 admits a holomorphic continuation to $B(0,1)$ by taking $g_1(q) = 0$. We may now divide out by the zero to define $g(q)$ at $q = 0$. ∎

We now turn to the multiplication formula. This will be harder still. We will require two lemmas.

Lemma 4.43 (Stirling's approximation)**.** As $s \to \infty$, we have

$$\Gamma(s+1) \sim \left(\frac{s}{e}\right)^s \sqrt{2\pi s}.$$

Proof. The following argument is taken from [Con, Section 3]. In order to make the asymptotic terms appear, we set $x := \frac{t-s}{\sqrt{s}}$ so that

$$\begin{aligned}\Gamma(s+1) &= \int_{\mathbb{R}^+} t^s e^{-t}\, dt \\ &= \int_{-\sqrt{s}}^{\infty} (\sqrt{s}x + s)^s e^{-(\sqrt{s}x+s)} \sqrt{s}\, dx \\ &= \left(\frac{s}{e}\right)^s \sqrt{s} \underbrace{\int_{-\sqrt{s}}^{\infty} \left(1 + \frac{x}{\sqrt{s}}\right)^s e^{-\sqrt{s}x}\, dx}_{I(\sqrt{s}) :=}\end{aligned}$$

It remains to check that $I(s) \to \sqrt{2\pi}$ as $s \to \infty$. This will be done using the Dominated convergence theorem. Define $f_s\colon \mathbb{R} \to \mathbb{R}$ by $f_s(x) := \left(1 + \frac{x}{s}\right)^{s^2} e^{-sx}$ so that $I(s) = \int_{\mathbb{R}} f_s(x)\, dx$. (Here, f_s is defined to be 0 on $(-\infty, -s]$.) We have two steps.

1. We claim that $f_s(x) \to e^{-x^2/2}$ as $s \to \infty$. It is enough to check equality after taking $\log$s, so we would like to show that

$$\lim_{s\to\infty} \left(s^2 \log\left(1 + \frac{x}{s}\right) - sx\right) \stackrel{?}{=} -\frac{x^2}{2}.$$

Now, $\log\left(1+\frac{x}{s}\right) = \sum_{k\geq 1} \frac{1}{k}\left(\frac{x}{s}\right)^k$, so the Monotone convergence theorem (used for s large) gives

$$\begin{aligned}\lim_{s\to\infty}\left(s^2\log\left(1+\frac{x}{s}\right)-sx\right) &= \lim_{s\to\infty}\left(s^2\sum_{k\geq 1}\frac{1}{k}\left(\frac{x}{s}\right)^k - sx\right)\\ &= \underbrace{0}_{k=1} - \underbrace{\frac{x^2}{2}}_{k=2} + \sum_{k\geq 3}\lim_{s\to\infty}\frac{1}{k}\left(\frac{x}{s}\right)^k,\end{aligned}$$

which evaluates to $-x^2/2$, as needed.

2. We now apply the Dominated convergence theorem to see that $I(s) \to \int_{\mathbb{R}} e^{-x^2/2}\,dx$, where the integral equals $\sqrt{2\pi}$ as remarked in Example 4.42. In light of the previous step, it remains to find a dominating function for the f_ss. We will do this based on sign.

   - For $x \leq 0$, we claim that $f_s(x) \leq e^{-x^2/2}$. If $s \leq -x$, then $f_s(x) = 0$, so there is nothing to do; otherwise, we take $s > -x$. After taking logarithms, we see that we would like to check that the function
   $$s^2\log\left(1+\frac{x}{s}\right) - sx + \frac{x^2}{2}$$
   is nonpositive for $x \leq 0$. This function vanishes at $x = 0$, so it is enough to check that it is increasing, for which we note its derivative (with respect to x) is
   $$\frac{s^2}{1+\frac{x}{s}}\cdot\frac{1}{s} - s + x = \frac{x^2}{s+x},$$
   which is nonnegative for $s \geq -x$.
   - For $x \geq 0$ (and $s \geq 1$), we claim that $f_s(x) \leq f_1(x)$. After taking logarithms, we see that we would like to show that
   $$(\log(1+x) - x) - \left(s^2\log\left(1+\frac{x}{s}\right) - sx\right)$$
   is nonnegative for $x \geq 0$. This function vanishes at $x = 0$, so it is enough to check that it is increasing, for which we note its derivative (with respect to x) is
   $$\left(\frac{1}{1+x} - 1\right) - \left(\frac{s^2}{1+\frac{x}{s}}\cdot\frac{1}{s} - s\right) = \frac{x^2(s-1)}{(1+x)(s+x)},$$
   which is nonnegative for $s \geq 1$.

   Thus, we see that our dominating function may be taken to be $e^{-x^2/2}$ in the negative region and $f_1(x)$ in the positive region. ■

Lemma 4.44 (Euler form)**.** If $s > 0$, then

$$\Gamma(s) = \lim_{n\to\infty}\frac{n!n^s}{s(s+1)\cdots(s+n)}.$$

Proof. We evaluate the limit directly, using Lemma 4.43. Manipulating directly, we see the limit is

$$\Gamma(s)\lim_{n\to\infty}\frac{\Gamma(n+1)n^s}{\Gamma(s+n+1)}.$$

We now see the desired $\Gamma(s)$ term, so we want to show that the remaining limit equals 0. By Lemma 4.43 and taking logarithms, we see that we would like to show that

$$\lim_{n\to\infty}(x\log x - x + s\log x - (s+x)\log(s+x) + s + x) \stackrel{?}{=} 0.$$

After some simplification, this limit is seen to equal the limit of $s - n\log\left(1+\frac{s}{n}\right)$, which can be evaluated to 0 by expanding out the power series for $\log(1+x)$. ■

Remark 4.45. The right-hand side in fact defines a holomorphic function on $\mathbb{C} \setminus \{0, -1, -2, \ldots\}$, so the given equality extends to this region by analytic continuation. This prior claim can be checked by verifying that the right-hand side converges uniformly on compact sets in the region $\{s : \operatorname{Re} s > 0\}$ and also satisfies the equation $\Gamma(s+1) = s\Gamma(s)$.

Proof of Proposition 4.40(c). The following argument is taken from [Var]. By analytic continuation, it is enough to check the identity when s is real and positive. We simply expand out the right-hand side using the Euler form (Lemma 4.44) and Stirling's approximation (Lemma 4.43). To avoid off-by-one errors, we note that

$$\begin{aligned}\Gamma(s) &= \lim_{n\to\infty} \frac{n!n^s}{s(s+1)\cdots(s+n)} \\ &= \lim_{n\to\infty} \frac{n!n^{s-1}}{s(s+1)\cdots(s+n-1)} \\ &= \lim_{n\to\infty} \frac{\sqrt{2\pi}e^{-n}n^{n+s-1/2}}{s(s+1)\cdots(s+n-1)}.\end{aligned}$$

Namely, the denominator now has precisely n terms. Now,

$$\begin{aligned}\prod_{k=0}^{d-1} \Gamma\left(s+\frac{k}{d}\right) &= \lim_{n\to\infty} \prod_{k=0}^{d-1} \frac{\sqrt{2\pi}e^{-n}n^{n+s+k/d-1/2}}{\left(s+\frac{k}{d}\right)\left(s+\frac{k+d}{d}\right)\cdots\left(s+\frac{k+(n-1)d}{d}\right)} \\ &= \lim_{n\to\infty} \frac{(\sqrt{2\pi})^d e^{-nd}n^{nd+ds+(0+\cdots+(d-1))/d-d/2}}{s\left(s+\frac{1}{d}\right)\cdots\left(s+\frac{nd-1}{d}\right)} \\ &= \lim_{n\to\infty} \frac{(\sqrt{2\pi})^d e^{-nd}n^{nd+ds-1/2}d^{nd}}{ds\,(ds+1)\cdots(ds+nd-1)}.\end{aligned}$$

We would like the Euler form (Lemma 4.44) for $\Gamma(ds)$ to come out of this limit, and this will be done by substituting $nd \to \infty$ into the limit for the coordinate $n \to \infty$. With this in mind, we move the strange factors from the right-hand side of the desired equality in Proposition 4.40 to the left-hand side, writing our limit as

$$(2\pi)^{-(d-1)/2}d^{ds-1/2}\prod_{k=0}^{d-1}\Gamma\left(s+\frac{k}{d}\right) = \lim_{n\to\infty}\frac{\sqrt{2\pi}e^{-nd}(nd)^{nd+ds-1/2}}{ds\,(ds+1)\cdots(ds+nd-1)},$$

which is indeed the Euler form for $\Gamma(ds)$. ∎

4.3.2 Unrefined Algebraicity

It will be worthwhile to give ourselves some language to describe the sorts of products we want to evalaute. A priori, we are basically computing a product which looks like

$$\prod_{i\in\mathbb{Z}} \Gamma\left(\frac{i}{N}\right)^{a_i},$$

where $\{a_i\}_{i\in\mathbb{Z}}$ is a sequence of integers arranged so that the above product is finite. (Namely, i/N should never be in $\mathbb{Z}_{\leq 0}$ if $a_i > 0$, and only finitely many of the a_i should fail to vanish.) However, by using the fact that $\Gamma(s+1) = s\Gamma(s)$, we may slide all factors of the product to $(0,1)$, meaning that we want to compute a product which looks like

$$\prod_{i=1}^{N-1} \Gamma\left(\frac{i}{N}\right)^{f(i/N)},$$

where $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ is some function.

Notation 4.46. For any function $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$, we define

$$\Gamma(f) \coloneqq \prod_{i=1}^{N-1} \Gamma\left(\frac{i}{N}\right)^{f(i/N)}.$$

Dually, for any element $a \in \mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$, we may write $a = \sum_{i=0}^{N-1} a_i \cdot \overline{i/N}$, and we define $\Gamma(a)$ according to the function $i/N \mapsto a_i$. Note that $\Gamma(a)$ does not admit a value if a is nonzero at $0/N$.

Remark 4.47. Because we are only interested in computing the periods $\operatorname{Per}\left(\gamma^{2p}, \nu_\alpha\right)$ where $\alpha \in \mathfrak{B}^{2p}$, we may restrict our view to functions $a\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ such that the weight $\langle a\rangle\colon (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{Z}$ is constant.

Now, Proposition 4.40 gives us two further properties about products of Γs we may use. By suitably translating, we are able to compute products which look like

$$\Gamma\left(\frac{a}{N}\right)\Gamma\left(\frac{N-a}{N}\right) \qquad \text{and} \qquad \Gamma\left(\frac{da}{N}\right)^{-1}\prod_{k=0}^{d-1}\Gamma\left(\frac{a}{N}+\frac{k}{d}\right),$$

$a, b \in \{1, \dots, N\}$, and we require $d \mid N$ and $N \nmid da$ in the second product. Here is some notation to keep track of this.

Notation 4.48. For a positive divisor d of N and $a \in \mathbb{Z}/N\mathbb{Z}$, we define the function $\varepsilon_{d,a}\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ to be the characteristic function of the set

$$\left\{\frac{N-da}{N}\right\} \cup \left\{\frac{a}{N}+\frac{k}{d} : k \in \{0, \dots, d-1\}\right\}.$$

Similarly, for any $a \in (\mathbb{Z}/N\mathbb{Z}) \setminus \{0\}$, we define $s_a\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ to be the characteristic function of $\left\{\frac{a}{N}, \frac{-a}{N}\right\}$.

Remark 4.49. Abusing notation slightly, we may identify $\varepsilon_{d,a}$ and s_a with the corresponding elements in $\mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$.

The moral is that we can compute $\Gamma(\varepsilon_{d,a})$ and $\Gamma(s_a)$, so we would like to see which functions $\frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ can be written as linear combinations of $\varepsilon_{d,a}$s and s_as. Recalling that we are only interested in functions of constant weight, we pick up the following results in this direction.

Lemma 4.50. For any positive divisor $d \mid N$ and $a \in \frac{1}{N}\mathbb{Z}/\mathbb{Z}$, the weight functions $\langle\varepsilon_{d,a}\rangle$ and $\langle s_a\rangle$ are constant.

Proof. This is [Del79, Example, p. 343]. Note $s_a = \varepsilon_{1,a}$ when $a \neq 0$, so we are reduced to the considering $\varepsilon_{d,a}$s. Before doing any computation, we note that we will write $[q]$ to be the representative in $[0,1)$ of an element $q \in \mathbb{Q}/\mathbb{Z}$. We now proceed in steps.

1. For any $u \in (\mathbb{Z}/N\mathbb{Z})^\times$, we find v such that $uv \equiv 1 \pmod N$ and compute

$$\begin{aligned}\langle\varepsilon_{d,a}\rangle(u) &= \frac{1}{N}\sum_{b\in(\mathbb{Z}/N\mathbb{Z})} 1_{\varepsilon_{d,a}}\left(\frac{ub}{N}\right)[b] \\ &= \left[-\frac{dva}{N}\right] + \sum_{k=0}^{d-1}\left[\frac{va}{N}+\frac{vk}{d}\right] \\ &= \langle\varepsilon_{d,va}\rangle(1).\end{aligned}$$

Thus, we see that we would like to show that $\langle \varepsilon_{d,a} \rangle(1) = \langle \varepsilon_{d,ua} \rangle(1)$ for any $u \in (\mathbb{Z}/N\mathbb{Z})^\times$; for example, there is nothing to show in the case where $a = 0$.

2. Setting $e \coloneqq N/d$, we note that $\varepsilon_{d,a} = \varepsilon_{d,a+e}$ pointwise. Thus, we are reduced to the case where $a \in [0, e)$ by shifting a appropriately.

3. Now, for any $q \in \mathbb{R}/\mathbb{Z} \setminus \frac{1}{d}\mathbb{Z}/\mathbb{Z}$, we define $\varepsilon_{d,q}$ as the indicator of the set $\{[-dq]\} \cup \{q + k/d : k \in \{0, 1 \ldots, d-1\}\}$. We claim that $\langle \varepsilon_{d,q} \rangle(1)$ does not depend on q, which will complete the proof in the cases where $a/N = q$ by the first step. As in the second step, we note that $\varepsilon_{d,q}$ only depends on the class of q in $\mathbb{Q}/\frac{1}{d}\mathbb{Z}$, so we may assume that $q \in (0, 1/d)$. Now, as in the first step, we compute

$$\begin{aligned} \langle \varepsilon_{d,q} \rangle(1) &= [-dq] + \sum_{k=0}^{d-1} \left[q + \frac{k}{d} \right] \\ &\overset{*}{=} (1 - dq) + \sum_{k=0}^{d-1} \left(q + \frac{k}{d} \right) \\ &= 1 + \sum_{k=0}^{d-1} \frac{k}{d}, \end{aligned}$$

which is independent of q. Here, the key equality $\overset{*}{=}$ holds notably because $q \in (0, 1/d)$. ■

Remark 4.51. In fact, the above proof shows that $\langle \varepsilon_{d,a} \rangle$ is $\frac{d+1}{2}$ when $N \nmid da$.

One also has a partial converse.

Proposition 4.52 (Koblitz–Ogus)**.** Let $f \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Q}$ be a function of constant weight such that $f(0) = 0$. Then f is a $\mathbb{Q}$-linear combination of the functions

$$\{\varepsilon_{d,a} : d \mid N, d \text{ is prime}, N \nmid da\} \sqcup \{s_a : N \mid a\}.$$

Proof. This is [Del79, Proposition, p. 344]. Approximately speaking, the idea is that we want to decompose f into a sum over some cosets, which is a job for Fourier analysis. Before doing anything, we set up some notation. Let E be the given set of $\varepsilon_{d,a}$s. Note that the given statement is one about some functions E in a vector space spanning the full space, which can be checked by extending scalars, so we go ahead and extend scalars to $\mathbb{C}$.

Now, for a given function $f \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{C}$ and a divisor $d \mid N$, we define $f_d \colon (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}$ by $f_d(u) \coloneqq f(u/d)$. For example, because $f(0) = 0$, we see that $f_1 = 0$. Continuing, for each function $f \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{C}$, we define $d(f)$ to be the smallest divisor of N such that $f_{d(f)}$ is nonzero, setting $d(f) = N$ if $f = 0$. Lastly, for convenience, we also define $I_d \subseteq (\mathbb{Z}/N\mathbb{Z})^\times$ (for $d \mid N$) to be the subgroup of elements $u \in (\mathbb{Z}/N\mathbb{Z})^\times$ such that $u \equiv 1 \pmod{d}$. Note that there is a short exact sequence

$$1 \to I_d \subseteq (\mathbb{Z}/N\mathbb{Z})^\times \twoheadrightarrow (\mathbb{Z}/d\mathbb{Z})^\times \to 1.$$

We now proceed in steps.

1. The general approach is to induct on $d(f)$. In particular, if $d(f) = 1$, then $f = 0$, so there is nothing to do. Thus, we may fix a divisor $d \mid N$ bigger than 1, and we would like to show that any (fixed) f of constant weight with $d(f) = d$ lives in $\operatorname{span}_{\mathbb{C}} E$, assuming this is true for any f' with $d(f') < d(f)$. As such, our goal is to find $g \in \operatorname{span}_{\mathbb{C}} E$ such that $d(f - g) < d(f)$.

   We need to do something to get ourselves off the ground, so we go ahead and specify some kinds of functions f with $d(f) = d$ for which we are already able to conclude.

(a) Suppose that f_d factors through $(\mathbb{Z}/d\mathbb{Z})^\times/\{\pm 1\}$. Then $f_d(-a) = f_d(a)$ for each a, so we may define

$$f' \coloneqq f - \sum_{a \in (\mathbb{Z}/d\mathbb{Z})^\times/\{\pm 1\}} f_d(a) s_{a/d}.$$

By construction, $f'_d = 0$ while $f'_e = f_e$ for any other divisor $e \mid N$, so $d(f') < d(f)$.

(b) Suppose that f_d factors through $\left(\mathbb{Z}/\frac{d}{p}\mathbb{Z}\right)^\times$ for some prime factor $p \mid d$. Let $I_{d/p,d}$ be the kernel of the projection $(\mathbb{Z}/d\mathbb{Z})^\times \twoheadrightarrow \left(\mathbb{Z}/\frac{d}{p}\mathbb{Z}\right)^\times$ so that f_d is invariant under $I_{d/p,d}$. Now, for each $a \in (\mathbb{Z}/d\mathbb{Z})^\times$, we note that

$$\frac{a}{d} I_{d/p,d} = \left(\left\{ \frac{d-pa}{d} \right\} \cup \left\{ \frac{a}{d} + \frac{k}{p} : k \in \{0, \dots, p-1\} \right\} \right) \cap \frac{1}{d}(\mathbb{Z}/d\mathbb{Z})^\times$$

because both sides are simply the elements of the form $\frac{b}{d}$ where $b \in (\mathbb{Z}/d\mathbb{Z})^\times$ has $a \equiv b \pmod{\frac{d}{p}}$. Thus, as in (a), we may subtract out suitable multiples of $\varepsilon_{p,a}$s from f to cause f_d to vanish while not changing f_e for any $e > d$, thereby making $d(f)$ smaller.

In the remaining steps, we will show that any f_d is a linear combination of functions of the type described in (a) and (b), which completes the induction and thus the proof.

2. The aforementioned goal will be achieved via Fourier analysis. Discrete Fourier analysis allows one to write f_d as a linear combination of characters $\chi\colon (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}^\times$, writing

$$f = \sum_{\chi\colon (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}^\times} \langle f, \chi_d \rangle \chi.$$

Because we want to show f_d is a linear combination of functions which factor through $(\mathbb{Z}/d\mathbb{Z})^\times/\{\pm 1\}$ or $\left(\mathbb{Z}/\frac{d}{p}\mathbb{Z}\right)^\times$, we may as well show that f_d is a linear combination of even and imprimitive characters. Taking the contraposition, we must show $\langle f_d, \chi_d \rangle = 0$ for any odd primitive character $\chi_d\colon (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}^\times$.

3. Forget the context of the previous step for a sentence. Continuing with the Fourier analysis, we will show in the next step that any function $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{C}$ and any character $\widetilde{\chi}\colon (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ has

$$\langle \langle f \rangle, \widetilde{\chi} \rangle = \sum_{\substack{d \mid N \\ \widetilde{\chi}|_{I_d} = 1}} -L(0, \chi_d) \, |I_d| \, \langle f_d, \chi_d \rangle, \tag{4.1}$$

where $\chi_d\colon (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}^\times$ is the character induced from $\widetilde{\chi}$. Let's explain how this completes the proof, returning to the context of the previous step.

We apply (4.1) to our f and some character $\widetilde{\chi}\colon (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ induced from a chosen odd primitive character $\chi_d\colon (\mathbb{Z}/d\mathbb{Z})^\times$; we want to show that $\langle f_d, \chi_d \rangle = 0$. Let's look at both sides of (4.1).

- Because $\langle f \rangle$ is constant and $\widetilde{\chi}$ is nontrivial, the left-hand side $\langle \langle f \rangle, \widetilde{\chi} \rangle$ vanishes.
- On the other hand, the right-hand side sees contributions only from divisors $e \mid N$ for which $I_e \subseteq \ker \widetilde{\chi}$. But then the image of I_e in $(\mathbb{Z}/d\mathbb{Z})^\times$ will be contained in $\ker \chi_d$, which forces $I_e \subseteq I_d$ because $\ker \chi_d$ is trivial (because χ_d is primitive). Thus, our sum only consider divisors $e \mid d$, but because $d(f) = d$, we see that $f_e = 0$ whenever $e < d$. In total, our right-hand side features only the term $-L(0, \chi_d) \, |I_d| \, \langle f_d, \chi_d \rangle$.

The above two points combine to imply $-L(0, \chi_d) \, |I_d| \, \langle f_d, \chi_d \rangle = 0$, so $\langle f_d, \chi_d \rangle = 0$ because χ_d being odd and primitive implies $L(0, \chi_d) \neq 0$. (Namely, $L(0, \chi_d) \neq 0$ by combining the functional equation for this Dirichlet L-function with the non-vanishing result Proposition 3.78.)

4. It remains to check the equality (4.1). This is a direct computation. Expanding everything out, we see

$$\langle\langle f\rangle,\widetilde{\chi}\rangle=\frac{1}{N}\sum_{\substack{u\in(\mathbb{Z}/N\mathbb{Z})^\times\\ a\in\mathbb{Z}/N\mathbb{Z}}}\frac{\langle a\rangle}{N}f\left(\frac{au}{N}\right)\overline{\chi}(u).$$

In order to make f_ds appear, we stratify the sum over a, writing

$$\langle\langle f\rangle,\widetilde{\chi}\rangle=\sum_{d|N}\frac{1}{d}\sum_{\substack{u\in(\mathbb{Z}/N\mathbb{Z})^\times\\ v\in(\mathbb{Z}/d\mathbb{Z})^\times}}\langle v\rangle f_d\left(uv\right)\overline{\chi}(u).$$

Eventually, the sum over v will turn into a term like $\langle f_d,\chi_d\rangle$, so we need to get rid of the sum over u. Let $U_d'\subseteq(\mathbb{Z}/N\mathbb{Z})^\times$ be a set of coset representatives for $(\mathbb{Z}/N\mathbb{Z})^\times/I_d$ so that $(\mathbb{Z}/N\mathbb{Z})^\times=U_d'I_d$. Then the internal sum over u looks like

$$\sum_{\substack{u'\in U_d'\\ u\in I_d}}f_d(uu'v)\overline{\chi}(uu').$$

Note $f_d(uu'v)=f_d(u'v)$, so we may sum $\overline{\chi}$ over just u alone. If $I_d\not\subseteq\ker\chi$, then this sum over u vanishes; otherwise, the sum over u is $|I_d|$, so the total sum is

$$\sum_{u'\in U_d'}f_d(u'v)\overline{\chi}(u')\,|I_d|=\chi_d(v)\,|I_d|\,\langle f_d,\chi_d\rangle.$$

Plugging this back in, we see

$$\langle\langle f\rangle,\widetilde{\chi}\rangle=\sum_{d|N}\left(\frac{1}{d}\sum_{v\in(\mathbb{Z}/d\mathbb{Z})^\times}\langle v\rangle\chi_d(v)\right)|I_d|\,\langle f_d,\chi_d\rangle.$$

The claim now follows by [Was12, Proposition 4.1, Theorem 4.2]. ∎

Corollary 4.53. Let $f\colon\frac{1}{N}\mathbb{Z}/\mathbb{Z}\to\mathbb{Z}$ be a function of constant weight w. Then

$$\pi^{-w}\Gamma(f)\in\overline{\mathbb{Q}}.$$

Proof. By adding or subtracting 1_0s (which have weight 0), we may assume that $f(0)=0$. The hypothesis and conclusion are $\mathbb{Q}$-linear in f (note that fractional powers are permitted in an algebraicity question), so Proposition 4.52 tells us that it is enough to check the result for f being one of the $\varepsilon_{d,a}$s in the statement; recall from Remark 4.51 that $\langle\varepsilon_{d,a}\rangle=\frac{d+1}{2}$.

In fact, for any divisor $d\mid N$ and choice of $a\in\mathbb{Z}/N\mathbb{Z}$ with $N\nmid da$, we claim that $\pi^{-w}\Gamma(\varepsilon_{d,a})\in\overline{\mathbb{Q}}^\times$, where $w=\frac{d+1}{2}$ is the weight. Indeed, by combining the reflection and multiplication formulae (Proposition 4.40), we see that $\Gamma(\varepsilon_{d,a})$ is

$$\Gamma\left(\frac{N-da}{N}\right)\prod_{k=0}^{d-1}\Gamma\left(\frac{a}{N}+\frac{k}{d}\right)\equiv\pi^{(d+1)/2}\pmod{\overline{\mathbb{Q}}^\times},$$

so the result follows. ∎

4.3.3 The Universal Distribution

This section follows [Kub79b]. We are now permitted to make the following definition.

Definition 4.54 (distribution)**.** A *distribution relation* is an element of $\mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$ of the form

$$\overline{a} - \sum_{\substack{b \in \frac{1}{N}\mathbb{Z}/\mathbb{Z} \\ db = a}} \overline{b},$$

where $d \mid N$ is a positive divisor. A *distrubtion* is a function $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to A$ to an abelian group A whose natural extension to $\mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$ vanishes on all distribution relations. A distribution is *odd* if and only if it also satisfies $f(-a/N) = -f(a/N)$ for all a.

Example 4.55 (universal)**.** Let U_N be the abelian group given by taking the quotient of $\mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$ by the subgroup generated by the distribution relations. Then there is a natural inclusion $i\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to U_N$, which we see is a distribution by construction. In fact, we see that every distribution $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to A$ factors uniquely through i, so i is initial in the category of distributions.

Example 4.56. By Proposition 4.40, the function $\frac{1}{\sqrt{2\pi}}\Gamma\colon \mathbb{Q}/\mathbb{Z} \to \mathbb{C}^\times/\overline{\mathbb{Q}}^\times$ is an odd distribution. Namely, this function descends to $\mathbb{Q}/\mathbb{Z}$ by the translation property, it is a distribution by the multiplication formula, and it is odd by the reflection formula. The Lang–Rohrlich conjecture asserts that $\frac{1}{\sqrt{2\pi}}\Gamma$ is a universal odd distribution; we refer to [And04, Lemma 24.6.1.1] for some related conjectures.

Example 4.56 explains why we are discussing distributions in this section: products of Γs can be tracked through as satisfying these distribution relations. We also remark that integer-valued functions of constant weight 0 live a new life here.

Lemma 4.57. Let $D_N^- \subseteq \mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$ be the $\mathbb{Z}$-module generated by the distribution relations and the elements $\overline{a} + \overline{-a}$ and $\overline{0}$. After identifying $\mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$ with functions $\frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$, we see D_N^- is generated by the elements $\overline{0}$ and $\varepsilon_{d,a}$ where $d \mid N$ is a divisor and $a \in (\mathbb{Z}/N\mathbb{Z})^\times$.

Proof. For nonzero a, note that $\varepsilon_{1,a}$ is simply the generator $\overline{a/N} + \overline{-a/N}$, and $\varepsilon_{d,a}$ produces the distribution relation

$$-\varepsilon_{d,a} + \varepsilon_{1,da} = \overline{\frac{da}{N}} - \sum_{k=0}^{d-1} \overline{\frac{a}{N} + \frac{k}{d}}.$$

Thus, up to adding or subtracting some $\varepsilon_{1,\bullet}$, we see that the distribution relations are in bijection with the $\varepsilon_{d,a}$s, so these elements generate the same subgroup of $\mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right]$. ∎

Remark 4.58. It is not hard to see that one may inductively write $\varepsilon_{d,a}$s as a $\mathbb{Z}$-linear combination of $\varepsilon_{p,a}$s where p is a prime. (For that matter, one can inductively write distribution relations in terms of ones where the divisor $d \mid N$ is prime.) The point is that we really only have to consider $\varepsilon_{p,a}$s (with p prime) and $\varepsilon_{1,a}$s in Lemma 4.57.

The goal of the present subsection is to show the following structure result [Kub79b, Theorem 1.8].

Theorem 4.59 (Kubert)**.** Let $i\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to U_N$ be an initial distribution. Then U_N is a free abelian group of rank $\varphi(N)$.

Proof from Propositions 4.61 and 4.63. We will go ahead and outline the argument, referring forward to results we will prove in the sequel. There are two main steps.

1. In Proposition 4.61, we show that any distribution f has $\langle \operatorname{im} f \rangle$ admitting a generating set of $\varphi(N)$ elements.
2. In Proposition 4.63, we exhibit a distribution r with $\dim_{\mathbb{Q}} \langle \operatorname{im} r \rangle_{\mathbb{Q}} = \varphi(N)$.

Let's quickly explain why these two implications allow us to conclude the proof. By the first step, we see that there is a surjection $\mathbb{Z}^{\varphi(N)} \twoheadrightarrow U_N$ of abelian groups, and we will be done as soon as we know that this map is an isomorphism. Well, because i is an initial distribution, we see that the distribution r factors through i, meaning that there is an induced surjection

$$\mathbb{Z}^{\varphi(N)} \twoheadrightarrow U_N \twoheadrightarrow \langle \operatorname{im} r \rangle.$$

However, this composite must become an isomorphism after tensoring with $\mathbb{Q}$ (for dimension reasons) by the second step, so the composite must in particular be injective. We conclude that the map $\mathbb{Z}^{\varphi(N)} \twoheadrightarrow U_N$ is an isomorphism. ∎

It remains to provide the proofs of Propositions 4.61 and 4.63. Before going further, we need some notation.

Notation 4.60. By the Chinese remainder theorem, summation provides an isomorphism

$$\sum_{p \mid N} \frac{1}{p^{\nu_p(N)}} \mathbb{Z}/\mathbb{Z} \to \frac{1}{N}\mathbb{Z}/\mathbb{Z}.$$

For any $s \in \frac{1}{N}\mathbb{Z}/\mathbb{Z}$ and $p \mid N$, we define $s_p \in \frac{1}{p^{\nu_p(N)}}\mathbb{Z}/\mathbb{Z}$ to be the corresponding p-component. Similarly, if we have $\frac{a}{N} \in \frac{1}{N}\mathbb{Z}/\mathbb{Z}$, we let $\frac{a_p}{p^{\nu_p(N)}}$ be the p-component.

Because it is faster, we now proceed with Proposition 4.63.

Proposition 4.61. Let $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to A$ be a distribution. Then $\langle \operatorname{im} f \rangle$ admits a generating set of $\varphi(N)$ elements.

Proof. This result is [Kub79b, Proposition 1.8], though we follow the isomorphic proof given in [Was12, Proposition 12.10]. The idea is to use the distribution relations to minimize the number of generators. There are two steps.

1. We claim that the collection

$$S_N \coloneqq \left\{ f\left(\frac{a}{N}\right) : a_p = 0 \text{ or } \gcd(a_p, p) = 1 \right\}$$

generates $\langle \operatorname{im} f \rangle$. We proceed by induction on the number of primes factors of N, where the statement has little content if $N = 1$.

Now, for a given N, choose some $a/N \in \frac{1}{N}\mathbb{Z}/\mathbb{Z}$, and we want to show that $f(a/N) \in \langle S_N \rangle$. Quickly, if $a_p = 0$ for some prime $p \mid N$, then in fact $a/N \in \frac{1}{d}\mathbb{Z}/\mathbb{Z}$ for some divisor $d \mid N$ with strictly fewer prime factors, so $f(a/N) \in \langle S_d \rangle$ by the induction.

Thus, we may assume that $a_p \neq 0$ for all $p \mid N$. In this case, we hope to use a distribution relation to find $f(a/N)$ in $\langle S_N \rangle$. In particular, note that we can write $a = dx$ where $d \mid N$ and $\gcd(x, N) = 1$: indeed, simply write $\frac{a}{N}$ in reduced terms as $\frac{x}{e}$, and then $a = \frac{N}{e} \cdot x$ is a suitable expansion. (In particular, e is the order of a, so $p \mid e$ for all primes e, so $\gcd(x, e) = 1$ implies $\gcd(x, N) = 1$.) Thus, $f(a/N)$ equals

$$f\left(d \cdot \frac{x}{N}\right) = \sum_{k=0}^{d-1} f\left(\frac{x}{N} + \frac{k}{d}\right),$$

and now every term in the right-hand side lives in S_N.

2. We claim that the collection

$$T_N \coloneqq \left\{ f\left(\frac{a}{N}\right) : a_p = 0, \text{ or } a_p \neq 1 \text{ and } \gcd(a_p, p) = 1 \right\}$$

generates $\langle \operatorname{im} f \rangle$. Once again, we proceed by induction on the number of prime factors of N, where the statement has little content if $N = 1$. Note that the previous step tells us that it is enough to check that $S_N \subseteq \langle T_N \rangle$.

As such, we go ahead and pick up some $f(a/N) \in S_N$, and to show that $f(a/N) \in \langle T_N \rangle$. As in the prior step, we note that having $a_p = 0$ for any prime p implies that $f(a/N) \in S_d$ for some $d \mid N$ with fewer prime factors, yielding $f(a/N) \in \langle T_N \rangle$ by the induction. Thus, we may assume that $a_p \neq 0$ for all p.

We will induct on the number $\omega(a/N)$ of primes p such that $a_p = 1$. Of course, if $\omega(a/N) = 0$, then $a/N \in T_N$ already, so there is nothing to do. Otherwise, suppose that our a/N has at least one prime $q \mid N$ with $a_q = 1$. We now use the distribution relations twice: set

$$\frac{b}{M} \coloneqq \sum_{\substack{p \mid N \\ p \neq q}} \frac{a_p}{p^{\nu_p(N)}},$$

and then we note that we have two equalities

$$f\left(q^{\nu_q(N)} \cdot \frac{b}{M}\right) = \sum_{k=0}^{q^{\nu_q(N)}-1} f\left(\frac{b}{M} + \frac{k}{q^{\nu_q(N)}}\right),$$

$$f\left(q^{\nu_q(N)-1} \cdot \frac{b}{M}\right) = \sum_{k=0}^{q^{\nu_q(N)-1}-1} f\left(\frac{b}{M} + \frac{qk}{q^{\nu_q(N)}}\right).$$

Both left-hand sides are in $\langle T_N \rangle$ by the induction on the number of prime factors. Now, subtracting these two equations produces the relation

$$\sum_{k \in (\mathbb{Z}/q^{\nu_q(N)})^\times} f\left(\frac{b}{M} + \frac{k}{q^{\nu_q(N)}}\right) \in \langle T_N \rangle.$$

Note $\frac{a}{N} = \frac{b}{M} + \frac{1}{q^{\nu_q(N)}}$ is the first term in this sum while the other terms in the sum have strictly smaller ω (because the q-component is not equal to 1), so we are done by the induction.

Note that the second step completes the proof because $\#S_N$ equals

$$\prod_{p \mid N} \# \left(\{0\} \cup \left(\mathbb{Z}/p^{\nu_p(N)}\mathbb{Z}\right)^\times \setminus \{1\} \right),$$

which is simply $\#(\mathbb{Z}/N\mathbb{Z})^\times = \varphi(N)$ by the Chinese remainder theorem. ∎

Remark 4.62. The proof of Proposition 4.61 actually gives explicit generators of $\operatorname{im} f$. One can unwind this (and the proof of Theorem 4.59) to give explicit generators of U_N defined in Example 4.55.

We now turn to the construction for Proposition 4.63.

Proposition 4.63. There exists a distribution $r \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to A$ such that $\dim_{\mathbb{Q}} \langle \operatorname{im} r \rangle_{\mathbb{Q}} = \varphi(N)$.

Proof. To understand why this is difficult, we note that we are basically trying to compute the dimension of the vector space $U_{N,\mathbb{Q}}$, where U_N is the rather horrendous abelian group constructed Example 4.55. Technically, Proposition 4.61 tells us what should be a basis, but this vector space has so many relations that

it is difficult to determine if these elements are actually linearly independent. The usual proof (in [Was12, Chapter 12] or [Kub79b, Section 3]) uses cyclotomy theory and some facts about character sums, reducing the task to a non-vanishing of some special value. These topics are moderately tangential to this thesis, so we will not discuss them. Instead, we will follow [Kub79b, Section 4] and provide a direct combinatorial construction.

Our target space will be $A_N \coloneqq \mathbb{Q}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$, and we note that $(\mathbb{Z}/N\mathbb{Z})^\times$ has a natural permutation action on A_N. Throughout, ord denotes the additive order of a group element. We require two elements of A_N.

- For $s \in \frac{1}{N}\mathbb{Z}/\mathbb{Z}$, we define
$$X_N(s) \coloneqq \sum_{\substack{x \in (\mathbb{Z}/N\mathbb{Z})^\times \\ x \cdot N/\operatorname{ord} s \equiv Ns}} \overline{x}.$$
For example, if $\operatorname{ord} s = N$, then $X_N(s) = \{Ns\}$. In general, if $s = a/d$ where $d = \operatorname{ord} s$ so that $\gcd(a,d) = 1$, then the xs take the form $(a + kd)$.

- For prime divisors $p \mid N$, we define
$$Y_N(p) \coloneqq \sum_{\substack{y \in (\mathbb{Z}/N\mathbb{Z})^\times \\ py \equiv 1 \pmod{N/p^{\nu_p(N)}}}} \overline{y}.$$
Notably, the value of $y \in (\mathbb{Z}/N\mathbb{Z})^\times$ only has freedom in the p-component, so $Y_N(p)$ has $\varphi\left(p^{\nu_p(N)}\right)$ elements.

Because $X_N(s)$ and $Y_N(p)$ are basically subsets of $(\mathbb{Z}/N\mathbb{Z})^\times$, we may write $\#X_N(s)$ or $\#Y_N(p)$ to mean the number of their elements. We are now ready to define $r_N \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Q}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$ by

$$r_N(s) \coloneqq \frac{X_N(s)}{\varphi(N)} \prod_{p \mid \operatorname{ord} s} \left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right).$$

It remains to run many checks on r_N. They are all some explicit combinatorial manipulations.

1. For $c \in (\mathbb{Z}/N\mathbb{Z})^\times$, we check that $r_N(cs) = cr_N(s)$. Note that $cX_N(s) = X_N(cs)$ because both contain the x such that $cxN/\operatorname{ord} s \equiv Ns$. It now suffices to check that
$$c\left(\left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right) \cdot X\right) \stackrel{?}{=} \left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right) \cdot cX$$
for any prime divisor $p \mid N$ and $X \in \mathbb{Q}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$. Well, it is enough to check this claim for $X \in (\mathbb{Z}/N\mathbb{Z})^\times$, whereupon doing some rearragnement shows that it is enough to check that $c(Y_N(p)X) = Y_N(p)(cX)$, which is true by definition of the $(\mathbb{Z}/N\mathbb{Z})^\times$-action on A_N.

2. For any divisor $M \mid N$, we claim that the diagram
$$\begin{array}{ccc} \frac{1}{M}\mathbb{Z}/\mathbb{Z} & \xrightarrow{r_M} & \mathbb{Q}\left[(\mathbb{Z}/M\mathbb{Z})^\times\right] \\ \cap\mathrel{|} & & \downarrow i \\ \frac{1}{N}\mathbb{Z}/\mathbb{Z} & \xrightarrow{r_N} & \mathbb{Q}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right] \end{array}$$
commutes, where i is given by $i(\overline{y}) = \frac{\varphi(M)}{\varphi(N)} \sum_{x \equiv y \pmod M} \overline{x}$ for any $y \in (\mathbb{Z}/M\mathbb{Z})^\times$. Note that i is injective and $\mathbb{Q}$-linear by construction, but it is not a ring map because it does not map $1 \mapsto 1$. However, the leading constant is chosen to make i multiplicative: for $\overline{y}_1, \overline{y}_2 \in (\mathbb{Z}/M\mathbb{Z})^\times$, we see $i(\overline{y}_1)i(\overline{y}_2)$ equals
$$\left(\frac{\varphi(M)}{\varphi(N)}\right)^2 \sum_{\substack{x_1 \equiv y_1 \pmod M \\ x_2 \equiv y_2 \pmod M}} \overline{x}_1\overline{x}_2 = \left(\frac{\varphi(M)}{\varphi(N)}\right)^2 \sum_{\substack{x \equiv y_1 y_2 \pmod M \\ x' \equiv 1 \pmod M}} \overline{x},$$

where we have substituted $(x, x') = (x_1x_2, x_1/x_2)$. We conclude $i(\overline{y}_1)i(\overline{y}_2) = i(\overline{y}_1\overline{y}_2)$, an equation which extends $\mathbb{Q}$-linearly to all A_N.

The main computation will be to compute $i(r_M(s))$ for $s \in \frac{1}{M}\mathbb{Z}/\mathbb{Z}$. Using the multiplicativity of the previous paragraph, we see

$$i(r_M(s)) = \frac{i(X_M(s))}{\varphi(M)} \prod_{p|\operatorname{ord} s} i\left(1 - \frac{Y_M(p)}{\#Y_M(p)}\right).$$

We see that we have to compute $i(X_M(s))$ and $i(Y_M(p))$.

- Note $\frac{\varphi(N)}{\varphi(M)} i(X_M(s)) = X_N(s)$: some $x \in (\mathbb{Z}/N\mathbb{Z})^\times$ finds itself in $\frac{\varphi(N)}{\varphi(M)} i(X_M(s))$ if and only if $x \cdot M/\operatorname{ord} s \equiv Ms$, which is equivalent to $x \cdot N/\operatorname{ord} s \equiv Ns$.
- We claim $\frac{i(Y_M(p))}{\#Y_M(p)} = i(1)\frac{Y_N(p)}{\#Y_N(p)}$. Note that the reduction map $Y_N(p) \to Y_M(p)$ is surjective: any y with $py \equiv 1 \pmod{M/p^{M/\nu_p(M)}}$ may be lifted to a multiplicative inverse of $p \pmod{N/p^{N/\mu_p(N)}}$. We thus see that the support of $\frac{\varphi(N)}{\varphi(M)} i(1) Y_M(p)$ agrees with the support of $\frac{\varphi(N)}{\varphi(M)} i(Y_M(p))$; however, each element in $\frac{\varphi(N)}{\varphi(M)} i(Y_M(p))$ is overcounted by a factor of $\varphi\left(p^{\nu_p(N)}\right)/\varphi\left(p^{\nu_p(M)}\right)$ because we already had freedom in the p-component. Adjusting for this completes the claim.

We now see

$$i(r_M(s)) = \frac{i(X_M(s))}{\varphi(M)} \prod_{p|\operatorname{ord} s} i(1)\left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right).$$

To get rid of the factor of $i(1)$, we note that $i(X_M(s))i(1) = i(X_M(s))$ by the multiplicativity. Lastly, we may substitute $\frac{i(X_M(s))}{\varphi(M)} = \frac{X_N(s)}{\varphi(N)}$, writing

$$i(r_M(s)) = \frac{X_N(s)}{\varphi(N)} \prod_{p|\operatorname{ord} s} \left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right),$$

which is indeed $r_N(s)$.

3. We claim that r_N is a distribution. Namely, for any divisor $d \mid N$ and $s \in \frac{1}{N}\mathbb{Z}/\mathbb{Z}$, we must check hat

$$r_N(ds) \stackrel{?}{=} \sum_{k=0}^{d-1} r_N\left(s + \frac{k}{d}\right).$$

We begin with a few reductions. By adjusting s by some k/d, we may assume that $\operatorname{ord} s$ is divisible by d. By inductively applying the distribution relations, we may assume that d is prime. Lastly, because i defined in the previous step is injective, we can pass from r_N to $r_{\operatorname{ord} s}$, allowing us to assume that $\operatorname{ord} s = N$. We now have two cases for the prime divisor d of N.

- Suppose that $d^2 \mid N$. In this case, all primes dividing $\operatorname{ord} s = N$ continue to divide $\operatorname{ord} ds = N/d$. Additionally, $s + \frac{k}{d}$ always has order N, so

$$\sum_{k=0}^{d-1} r_N\left(s + \frac{k}{d}\right) = \left(\frac{1}{\varphi(N)} \sum_{k=0}^{d-1} X_N\left(s + \frac{k}{d}\right)\right) \prod_{p|N}\left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right).$$

Because $s + \frac{k}{d}$ has order N, we see $X_N\left(s + \frac{k}{d}\right) = \overline{N(s + \frac{k}{d})}$. On the other hand, $X_N(ds)$ consists of the x for which $dx \equiv d(Ns)$, which is equivalent to having $x = N(s + \frac{k}{d})$. We conclude

$$\sum_{k=0}^{d-1} r_N\left(s + \frac{k}{d}\right) = \frac{X_N(ds)}{\varphi(N)} \prod_{p|N}\left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right),$$

which is $r_N(ds)$.

- Suppose $d \mid N$ while $d^2 \nmid N$. The same computation essentially goes through except for two caveats: $\operatorname{ord} ds = N/d$ has one fewer prime factor, and $s + \frac{k}{d}$ need not have order N. In particular, the p-component of $s + \frac{k}{d}$ is the same as the same p-component of s, so the order of $s + \frac{k}{d}$ is either N or N/d. Further, and we see that it will be N/d only when $s + \frac{k}{d}$ has d-component equal to 0 for exactly when k; say that $t = s + \frac{k_0}{d}$ is this value of k. Then

$$\sum_{k=0}^{d-1} r_N\left(s + \frac{k}{d}\right) = \frac{1}{\varphi(N)}\left(X_N(t) + \sum_{k=1}^{d-1} X_N\left(t + \frac{k}{d}\right)\left(1 - \frac{Y_N(d)}{\#Y_N(d)}\right)\right) \prod_{p \mid N/d}\left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right).$$

  Comparing this to $r_N(ds) = r_N(dt)$, we see that we have left to show

$$X_N(dt) \stackrel{?}{=} X_N(t) + \sum_{k=1}^{d-1} X_N\left(t + \frac{k}{d}\right)\left(1 - \frac{Y_N(d)}{\#Y_N(d)}\right),$$

  which is equivalent to

$$X_N(dt) + \frac{1}{\#Y_N(d)} \sum_{k=1}^{d-1} X_N\left(t + \frac{k}{d}\right) Y_N(d) \stackrel{?}{=} X_N(t) + \sum_{k=1}^{d-1} X_N\left(t + \frac{k}{d}\right).$$

  We now must compute the various X_Ns.

  - Each $t + \frac{k}{d}$ has order N by construction of t, so $X_N\left(t + \frac{k}{d}\right) = \overline{N(t + \frac{k}{d})}$. As such, multiplying by $Y_N(d)$ will leave us with $x \in (\mathbb{Z}/N\mathbb{Z})^\times$ such that $x \equiv \frac{N}{d}(t + \frac{k}{d}) \pmod{N/d}$, which is equivalent to $x \equiv \frac{N}{d} t \pmod{N/d}$; in particular, the sum on the left-hand side counts all these elements $\#Y_N(d) = (d-1)$ times. On the other hand, $X_N(t)$ consists of the x for which $dx \equiv Nt$, which is equivalent to $x \equiv \frac{N}{d} t \pmod{N/d}$, so

$$\frac{1}{\#Y_N(d)} \sum_{k=1}^{d-1} X_N\left(t + \frac{k}{d}\right) = X_N(t).$$

  - Similarly, dt has order N/d, so $X_N(dt)$ consists of the $x \in (\mathbb{Z}/N\mathbb{Z})^\times$ such that $dx \equiv Ndt$. Well, this is equivalent to having $x \equiv \overline{N(t + \frac{k}{d})}$, so

$$X_N(dt) = \sum_{k=1}^{d-1} X_N\left(t + \frac{k}{d}\right).$$

  Combining the above two points completes the computation.

4. We begin computing $\langle \operatorname{im} r_N \rangle_{\mathbb{Q}}$. For each prime $p \mid N$, define the fractional ideal

$$U_p \coloneqq X_N\left(\frac{p^{\nu_p(N)}}{N}\right) \mathbb{Z}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right] + \left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right) \mathbb{Z}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right].$$

   We claim that $\langle \operatorname{im} r_N \rangle$ equals $\prod_{p \mid N} U_p$. Because r_N respects the $(\mathbb{Z}/N\mathbb{Z})^\times$-action, it is enough to check that $\langle \operatorname{im} r_N \rangle$ is given by generators of this ideal. Well, a generic generator of $\prod_{p \mid N} U_p$ looks like

$$\prod_{p \nmid M} X_N\left(\frac{p^{\nu_p(N)}}{N}\right) \prod_{p \mid M}\left(1 - \frac{Y_N(p)}{\#Y_N(p)}\right),$$

   where M is some divisor of N; in fact, we may as well assume $\nu_p(M) \in \{0, \nu_p(N)\}$ for all primes p. We claim that the above element is $\varphi(N) r_N(1/M)$; this claim completes this step. To show the claim, we note the right product is already seen in $r_N(1/M)$. It thus remains to show that

$$\prod_{p \nmid M} X_N\left(\frac{p^{\nu_p(N)}}{N}\right) \stackrel{?}{=} X_N\left(\frac{1}{M}\right).$$

Indeed, the left-hand side is made of products $\prod_{p\nmid M} x_p$ where $x_p \cdot p^{\nu_p(N)} \equiv p^{\nu_p(N)} \pmod N$, which is equivalent to a condition on $x_p \equiv 1 \pmod{N/p^{\nu_p(N)}}$. By the Chinese remainder theorem, such products are in bijection with xs such that $x \equiv 1 \pmod{N/M}$, which is $X_N(1/M)$.

5. We claim that $(U_p)_{\mathbb{Q}} = \mathbb{Q}[(\mathbb{Z}/N\mathbb{Z})^\times]$. Because $(\#Y_N(p) - Y_N(p)) \in U_p$, it is enough to check that $Y_N(p) \in U_p$. In fact, we claim that $Y_N(p)$ is a multiple of $X_N\left(p^{\nu_p(N)}/N\right)$, which will complete the proof. Well, $Y_N(p)$ has x such that $px \equiv 1 \pmod{N/p^{\nu_p(N)}}$, and $X_N\left(p^{\nu_p(N)}/N\right)$ has x such that $x \equiv 1 \pmod{N/p^{\nu_p(N)}}$ as discussed in the previous step. Thus, we see
$$p' X_N\left(\frac{p^{\nu_p(N)}}{N}\right) = Y_N(p),$$
where $p' \in (\mathbb{Z}/N\mathbb{Z})^\times$ is chosen so that $p' \equiv p \pmod{N/p^{\nu_p(N)}}$, and the claim follows.

Thus, we have checked that r_N is a distribution, and the last two steps check that $\langle \operatorname{im} r_N\rangle_{\mathbb{Q}} = A_N$, so $\dim\langle \operatorname{im} r_N\rangle_{\mathbb{Q}} = \varphi(N)$ follows. ∎

4.3.4 Cohomology of the Universal Distribution

Let $i\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to U_N$ be the initial distribution of Example 4.55, and further let U_N^- be the quotient of U_N by the elements $\langle \overline{a} + \overline{-a}\rangle_{a\in\frac{1}{N}\mathbb{Z}/\mathbb{Z}}$. The quotient U_N^- is of interest to us because Γ factors through U_N^- by combining the reflection formula (Proposition 4.40) with Example 4.56.

We are now ready to state the main result of this subsection.

Theorem 4.64. Let $i\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to U_N$ be the initial distribution of Example 4.55, and further let U_N^- be the quotient of U_N by the elements $\langle \overline{a} + \overline{-a}\rangle_{a\in\frac{1}{N}\mathbb{Z}/\mathbb{Z}}$.

(a) The torsion subgroup $U_{N,\mathrm{tors}}^-$ is 2-torsion.

(b) If N is odd or divisible by 4, then $\dim_{\mathbb{F}_2} U_{N,\mathrm{tors}}^- = 2^{\omega(N)-1}$, where $\omega(N)$ is the number of distinct prime factors of N.

Proof from Propositions 4.65 and 4.67. As in the previous subsection, we go ahead and outline the argument, referring forward to results we will prove in the sequel. There are two steps: in Proposition 4.65, we show that $U_{N,\mathrm{tors}}^-$ is isomorphic to the cohomology group $\mathrm{H}^2(\langle\pm1\rangle, U_N)$, thereby proving (a). The dimension computation for this cohomology group is carried out in Proposition 4.67. ∎

We now turn our attention to the proofs of Proposition 4.65 and Proposition 4.67.

Proposition 4.65. Let $i\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to U_N$ be the initial distribution of Example 4.55. Further, let U_N^- be the quotient by the elements $\langle \overline{a} + \overline{-a}\rangle_{a\in\frac{1}{N}\mathbb{Z}/\mathbb{Z}}$. Then the torsion subgroup of U_N^- is isomorphic to
$$\mathrm{H}^2(\langle\pm1\rangle, U_N).$$

Proof. This is an application of Theorem 4.59. We follow [GGL24, Proposition 6.3.3]. Note that the action of $\langle\pm1\rangle \subseteq (\mathbb{Z}/N\mathbb{Z})^\times$ on $\frac{1}{N}\mathbb{Z}/\mathbb{Z}$ extends to U_N. We will actually show that $U_{N,\mathrm{tors}}^-$ is isomorphic to the Tate cohomology group
$$\mathrm{H}_{\mathrm{T}}^0(\langle\pm1\rangle, U_N) = \frac{U_N^{\langle\pm1\rangle}}{\mathrm{N}_{\langle\pm1\rangle}(U_N)},$$
which is enough because the group cohomology of a cyclic group is 2-periodic. We have two inclusions.

- On one hand, the denominator of $\mathrm{H}_{\mathrm{T}}^0(\langle\pm1\rangle, U_N)$ is basically modding out by the elements $\overline{a} + \overline{-a}$. Thus, we have an inclusion $\mathrm{H}_{\mathrm{T}}^0(\langle\pm1\rangle, U_N) \subseteq U_N^-$, so $\mathrm{H}_{\mathrm{T}}^0(\langle\pm1\rangle, U_N) \subseteq U_{N,\mathrm{tors}}^-$ because Tate cohomology groups are torsion.

- On the other hand, choose some $f \in U_{N,\mathrm{tors}}^-$, and we would like to check that $f \in U_N^{\langle \pm 1 \rangle}$. Well, we are given that there is some $D > 0$ such that Df vanishes in U_N^-, so $Df = (\overline{1} + \overline{-1})g$ (in U_N) for some $g \in U_N$. However, this implies that $(\overline{1} - \overline{-1})Df = 0$ in U_N, which requires $(\overline{1} - \overline{-1})f = 0$ because U_N is torsion-free by Theorem 4.59! We conclude that $f \in U_N^{\langle \pm 1 \rangle}$. ■

Before proceeding with the long proof of Proposition 4.67, we pick up a group-theoretic lemma.

Lemma 4.66. Fix finite abelian groups G and H. If M is a free $\mathbb{Z}[G \times H]$-module, then M^H and M/M^H are both free $\mathbb{Z}[G]$-modules.

Proof. Because M is a module over $G \times H$, we see that M^H is still a G-module. Quickly, note that M is a sum of $\mathbb{Z}[G \times H]$s, so because taking $(\cdot)^H$ and the quotient are both additive functors, it suffices to check the result for $M = \mathbb{Z}[G \times H]$. We now show that M^H and M/M^H are free independently.

- We show that M^H is a free $\mathbb{Z}[G]$-module. Indeed, some element $\sum_{(g,h)} a_{(g,h)}(g,h)$ is H-invariant if and only if $a_{(g,h)} = a_{(g,h')}$ always, in which case we see that

$$\sum_{(g,h) \in G \times H} a_{(g,h)}(g,h) = \sum_{g \in G} \left(a_{(g,1)}(g,1) \sum_{h \in H} (1,h) \right).$$

  Thus, we see that the map $\mathbb{Z}[G] \to \mathbb{Z}[G \times H]^H$ given by multiplying by $\sum_h (1,h)$ is an isomorphism.

- We show that M/M^H is a free $\mathbb{Z}[G]$-module. Quickly, observe that $\mathbb{Z}[G \times H]$ is free over $\mathbb{Z}[G]$ with a basis given by $\{(1,h)\}_{h \in H}$, so we may apply a linear transformation to see that $\mathbb{Z}[G \times H]$ is free over $\mathbb{Z}[G]$ with basis instead given by

$$\left\{ \sum_h (1,h) \right\} \sqcup \{(1,h)\}_{h \neq 1}.$$

  The first element is a basis of $\mathbb{Z}[G \times H]^H$ over $\mathbb{Z}[G]$ by the previous point, so we see that the quotient is free over $\mathbb{Z}[G]$ with basis given by the remaining entries. ■

Proposition 4.67 (Kubert)**.** Fix a positive integer N which is odd or divisible by 4, and let $i \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to U_N$ be the initial distribution of Example 4.55. Then

$$\dim_{\mathbb{F}_2} \mathrm{H}_{\mathrm{T}}^{\bullet}(\langle \pm 1 \rangle, U_N) = 2^{\omega(N)-1},$$

where $\omega(N)$ is the number of distinct prime factors of N.

Proof. Our argument follows [Kub79a, Section 2]. We continue with the set-up of Proposition 4.63, but we drop all the subscript Ns because we will work with fixed N throughout. Thus, we may also set $\nu_p \coloneqq \nu_p(N)$ for each prime p. In particular, by the universal property (and as outlined in Theorem 4.59), we see that U_N is isomorphic to the image of induced map $r \colon \mathbb{Z}\left[\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right] \to \mathbb{Q}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$.

We will need a few other pieces of notation. For bookkeeping reasons, we say that a divisor $M \mid N$ is admissible if and only if $\nu_p(M) \in \{0, \nu_p\}$ for all primes p; roughly speaking, M keeps track of a subset of primes dividing N. For example, for each admissible divisor $M \mid N$, we define

$$U(M) \coloneqq \prod_{p \mid M} U_p,$$

where U_p is the ideal defined at the end of the proof of Proposition 4.63; for example, $U(1) = \mathbb{Z}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$ and $U(N) = \operatorname{im} r$. In short, $U(M)$s will allow us to make certain inductive arguments.

Continuing, for each admissible divisor $M \mid N$, we define the subgroup $C(M) \subseteq (\mathbb{Z}/N\mathbb{Z})^\times$ by

$$C(M) \coloneqq \{a \in (\mathbb{Z}/N\mathbb{Z})^\times : a \equiv 1 \pmod{N/M}\}.$$

Thus, $C(M) \cong \prod_{p|M} (\mathbb{Z}/p^{\nu_p}\mathbb{Z})^\times$ is isomorphic to $(\mathbb{Z}/M\mathbb{Z})^\times$. For example, $C(1) = (\mathbb{Z}/N\mathbb{Z})^\times$ and $C(N) = \{1\}$. We also remark that the sum $Y(p)$ is fixed by $C(p^{\nu_p})$ by contruction (in fact, the set admits a transitive action), and a quick expansion of the definitions reveals that $X\left(p^{\nu_p}/N\right) = C(p^{\nu_p})$. As usual, we may identify $C(p^{\nu_p})$ with an element of $\mathbb{Z}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$ given by $\sum_{a \in C(p^{\nu_p})} \overline{a}$.

In the end, we will show that

$$\dim_{\mathbb{F}_2} \mathrm{H}^\bullet_{\mathrm{T}}\left(\langle \pm 1 \rangle, U(M)^{C(N/M)}\right) \stackrel{?}{=} 2^{\omega(M)-1}$$

for any admissible divisor $M \mid N$ bigger than 1, via an induction; taking $M = N$ then produces the desired result. Our proof now proceeds in many steps. We remark that our first few steps are picking up some technica tools used later.

1. Let $\varepsilon_p \in \mathbb{Q}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$ be the idempotent $\frac{1}{\#C(p^{\nu_p})} \sum_{a \in C(p^{\nu_p})} \overline{a}$. Then we claim that $x \in \mathbb{Q}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$ is fixed by $C(p^{\nu_p})$ if and only if $(1-\varepsilon_p)x = 0$. This is some abstract group theory. In one direction, if x is fixed by $C(p^{\nu_p})$, then
$$\frac{1}{\#C(p)} \sum_{a \in C(p^{\nu_p})} ax = \frac{1}{\#C(p^{\nu_p})} \sum_{a \in C(p^{\nu_p})} x$$
is simply x. In the other direction, if $(1-\varepsilon_p)x = 0$, then $x = \varepsilon_p x$; however, $a\varepsilon_p = \varepsilon_p$ for all $a \in C(p^{\nu_p})$ by a rearrangement of the terms in ε_p, so we see that $\varepsilon_p x$ is certainly fixed by $C(p^{\nu_p})$.

2. For an admissible divisor $M \mid N$ and prime $p \mid (N/M)$, we claim that $(1-\varepsilon_p)U\left(Mp^{\nu_p}\right) = (1-\varepsilon_p)U_M$ and
$$U\left(Mp^{\nu_p}\right)^{C(p^{\nu_p})} \stackrel{?}{=} C(p^{\nu_p}) \cdot U(M) + \left(1 - \frac{Y(p)}{\#Y(p)}\right) U(M)^{C(p^{\nu_p})},$$
where $C(p)$ refers to the element $\sum_{a \in C(p)} \overline{a}$ by abuse of notation.

   For this, we note $U\left(Mp^{\nu_p}\right) = U_p U(M)$ by definition, so
$$U\left(Mp^{\nu_p}\right) = C\left(p^{\nu_p}\right) U(M) + \left(1 - \frac{Y(p)}{\#Y(p)}\right) U(M).$$
   The first and the second claimed equalities are linked by the previous step, which tells us that we are interested in the kernel of $(1-\varepsilon_p)$. As such, let's look at how $(1-\varepsilon_p)$ behaves on each term.

   - Certainly $(1-\varepsilon_p)$ vanishes on $C\left(p^{\nu_p}\right)$, so the left term above lives in the kernel.
   - Similarly, $Y(p)$ is fixed by $C\left(p^{\nu_p}\right)$, so it is in the kernel of $(1-\varepsilon_p)$, from which one sees $(1-\varepsilon_p)\left(1 - \frac{Y(p)}{\#Y(p)}\right) = (1-\varepsilon_p)$. For example, we see that multiplying by $(1-\varepsilon_p)$ kills the coefficient $\left(1 - \frac{Y(p)}{\#Y(p)}\right)$. Additionally, the kernel of $(1-\varepsilon_p)$ will simply be the kernel of $(1-\varepsilon_p)$ acting on $U(M)$.

   Combining these two points completes the proof.

3. Suppose that M and M' are admissible divisors of N such that $MM' \mid N$. Then we claim that $U(M)$ is free as a $C(M')$-module; further, if $MM' \neq N$, we claim that $U(M)$ is free as a $\pm C(M')$-module.

   For this, we induct on M. If $M = 1$, then $U(M)$ is free over all subgroups of $(\mathbb{Z}/N\mathbb{Z})^\times$, so there is nothing to do. Thus, we focus on the inductive step, so suppose that the statement is true for M, and we would like to show it for Mp^{ν_p} for some prime $p \mid (N/M)$. Then the previous step provides short exact sequences as follows.
$$\begin{array}{ccccccccc}
0 & \longrightarrow & U(M)^{C(p^{\nu_p})} & \longrightarrow & U(M) & \longrightarrow & (1-\varepsilon_p)U(M) & \longrightarrow & 0 \\
 & & & & & & \| & & \\
0 & \longrightarrow & U(Mp^{\nu_p})^{C(p^{\nu_p})} & \longrightarrow & U(Mp^{\nu_p}) & \longrightarrow & (1-\varepsilon_p)U(Mp^{\nu_p}) & \longrightarrow & 0
\end{array}$$

The main claim is that $U(M)^{C(p^{\nu_p})} = U(Mp^{\nu_p})^{C(p^{\nu_p})}$. Let's quickly explain why this claim will complete this step. Fix an admissible divisor $M' \mid (N/Mp^{\nu_p})$, and then there are two things to show.

- We would like to show that $U(Mp^{\nu_p})$ is free over $C(M')$. By the inductive hypothesis, we know that $U(M)$ is free over $C(M')$ and $C(p^{\nu_p})$ and even over the product of the two groups. Thus, Lemma 4.66 tells us that $U(M)^{C(p^{\nu_p})} = U(Mp^{\nu_p})^{C(p^{\nu_p})}$ and $(1-\varepsilon_p)U(M) = (1-\varepsilon_p)U(Mp^{\nu_p})$ are both free over $C(M')$. Because the right term of the bottom short exact sequence is free, we conclude that the bottom short exact sequence thus splits, forcing $U(Mp^{\nu_p})$ to be a sum of free $C(M')$-modules and hence free.
- Suppose $MM'p^{\nu_p} \neq N$. Then we would like to show that $U(Mp^{\nu_p})$ is free over $\pm C(M')$. In this case, $U(M)$ is free over $\pm C(p^{\nu_p})C(M')$ by the induction, but $\pm C(M') \cap C(p^{\nu_p})$ is trivial: any element a in the intersection has $a \equiv 1 \pmod{N/p^{\nu_p}}$ and $a \equiv \pm 1 \pmod{N/M'}$, but the $+1$ is forced by having $N/(M'p^{\nu_p})$ be bigger than 2 by the hypotheses on N. Thus, $U(M)$ is actually free over $\pm C(M') \times C(p^{\nu_p})$, and now the argument can proceed as in the previous step.

It remains to prove the main claim $U(M)^{C(p^{\nu_p})} = U(Mp^{\nu_p})^{C(p^{\nu_p})}$. Well, the previous step grants

$$U\left(Mp^{\nu_p}\right)^{C(p^{\nu_p})} = C(p^{\nu_p}) \cdot U(M) + \left(1 - \frac{Y(p)}{\#Y(p)}\right) U(M)^{C(p^{\nu_p})}.$$

By the inductive hypothesis, $U(M)$ is free over $C(p^{\nu_p})$, so $U(M)^{C(p^{\nu_p})} = C(p^{\nu_p}) \cdot U(M)$. Thus, it is enough to show that $\left(1 - \frac{Y(p)}{\#Y(p)}\right) U(M)^{C(p^{\nu_p})} \subseteq U(M)^{C(p^{\nu_p})}$. Well, $Y(p)$ is stable under $C(p^{\nu_p})$, so we can express $\left(1 - \frac{Y(p)}{\#Y(p)}\right) \cdot C(p^{\nu_p})$ as $(1-y) \cdot C(p^{\nu_p})$ for some $y \in Y(p)$, and the result follows because $U(M)$ is a fractional ideal for $\mathbb{Z}\left[(\mathbb{Z}/N\mathbb{Z})^\times\right]$.

4. As a last tool, we show that the induced action of $(\mathbb{Z}/N\mathbb{Z})^\times$ on $\mathrm{H}^\bullet\left(\langle \pm 1\rangle, U(M)^{C(N/M)}\right)$ is trivial for any admissible divisor $M \mid N$. In fact, it's enough to check that the action of $C(p^{\nu_p})$ is trivial because these subgroups generate $(\mathbb{Z}/N\mathbb{Z})^\times$. Now, note that $U(M)^{C(N/M)}$ automatically has trivial action by $C(p^{\nu_p})$ if $p \nmid M$, so we now focus on the case $p \mid M$.

   Well, for a given $a \in C(p^{\nu_p})$, we would like to show that the action of a is trivial, for which it is enough to show that the action of $(1-a)$ is zero. Well, we claim that multiplication by $(1-a)$ factors through $\mathrm{H}^\bullet_{\mathrm{T}}\left(\langle \pm 1\rangle, U(M/p^{\nu_p})^{C(N/M)}\right)$, which we note vanishes because $U(M/p^{\nu_p})^{C(N/M)}$ is free over $\langle \pm 1\rangle$ by the previous step!

   Now, to show that multiplication by $(1-a)$ factors as claimed, it is enough by functoriality to show that multiplication by $(1-a)$ on $U(M)$ factors through $U(M/p^{\nu_p})$. Well, we see

$$U(M) = C\left(p^{\nu_p}\right) U(M/p^{\nu_p}) + \left(1 - \frac{Y(p)}{\#Y(p)}\right) U(M/p^{\nu_p}).$$

   Note $(1-a)C\left(p^{\nu_p}\right) = 0$, and $(1-a)\left(1 - \frac{Y(p)}{\#Y(p)}\right) = (1-a)$ because $Y(p)$ is fixed by a. Thus, $(1-a)U(M) \subseteq U(M/p^{\nu_p})$, as needed.

5. If M and Mp^{ν_p} are admissible divisors of N, we claim that there is a short exact sequence

$$0 \to \mathrm{H}^\bullet_{\mathrm{T}}\left(\langle \pm 1\rangle, U(M)^{C(N/M)}\right) \to \mathrm{H}^\bullet_{\mathrm{T}}\left(\langle \pm 1\rangle, U(Mp^{\nu_p})^{C(N/Mp^{\nu_p})}\right) \to \mathrm{H}^{\bullet+1}_{\mathrm{T}}\left(\langle \pm 1\rangle, U(M)^{C(N/M)}\right) \to 0.$$

   Note that we will have to do something nontrivial (beyond immediately applying a long exact sequence) because the middle group has a different invariant subgroup C acting on it. Our extra input will come from the morphism

$$\begin{array}{ccccccccc}
0 & \longrightarrow & U(M)^{C(p^{\nu_p})} & \longrightarrow & U(M) & \xrightarrow{(1-\varepsilon_p)} & (1-\varepsilon_p)U(M) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle (1-\frac{Y(p)}{\#Y(p)})} & & \downarrow{\scriptstyle (1-\frac{Y(p)}{\#Y(p)})} & & \downarrow & & \\
0 & \longrightarrow & U(Mp^{\nu_p})^{C(p^{\nu_p})} & \longrightarrow & U(Mp^{\nu_p}) & \xrightarrow{(1-\varepsilon_p)} & (1-\varepsilon_p)U(Mp^{\nu_p}) & \longrightarrow & 0
\end{array}$$

of exact sequences discussed in the third step; note that the left arrow is well-defined by the second step, and right arrow is then induced by the diagram. Before continuing, we make a few simplifiactions to this diagram.

- In the third step, we showed that $U(M)^{C(p^{\nu_p})} = U(Mp^{\nu_p})^{C(p^{\nu_p})}$.
- Recall that $Y(p)$ is fixed by $C(p^{\nu_p})$, so $(1-\varepsilon_p)\left(1-\frac{Y(p)}{\#Y(p)}\right) = (1-\varepsilon_p)$, thereby implying that the right arrow is simply the identity. We no longer care about the exact content of this right-hand term, so we denote it by $K := (1-\varepsilon_p)U(M) = (1-\varepsilon_p)U(Mp^{\nu_p})$.
- Using the fact that the set $Y(p)$ has a transitive action by $C(p^{\nu_p})$, we see that multiplying an element of $U(M)^{C(p^{\nu_p})}$ by $Y(p)$ is the same as multiplying it by any other element. Thus, we go ahead and fix some element $y_p \in Y(p)$, and we see that the left arrow is simply multiplication by $(1-y_p)$.

Our diagram now looks like the following.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U(M)^{C(p^{\nu_p})} & \longrightarrow & U(M) & \longrightarrow & K & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle (1-y_p)} & & \downarrow{\scriptstyle \left(1-\frac{Y(p)}{\#Y(p)}\right)} & & \| & & \\ 0 & \longrightarrow & U(M)^{C(p^{\nu_p})} & \longrightarrow & U(Mp^{\nu_p}) & \longrightarrow & K & \longrightarrow & 0 \end{array}$$

We now take $C(N/Mp^{\nu_p})$-invariants and $\langle\pm1\rangle$-cohomology to recover the result. Taking $C(N/Mp^{\nu_p})$-invariants keeps the exactness because $U(M)^{C(p^{\nu_p})}$ is free over $C(N/Mp^{\nu_p})$ by using Lemma 4.66 and the result in step 3. Thus, our diagram looks like the following.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U(M)^{C(N/M)} & \longrightarrow & U(M)^{C(N/Mp^{\nu_p})} & \longrightarrow & K' & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle (1-y_p)} & & \downarrow{\scriptstyle \left(1-\frac{Y(p)}{\#Y(p)}\right)} & & \| & & \\ 0 & \longrightarrow & U(M)^{C(N/M)} & \longrightarrow & U(Mp^{\nu_p})^{C(N/Mp^{\nu_p})} & \longrightarrow & K' & \longrightarrow & 0 \end{array}$$

Here, K' is the induced quotient, which we continue to not care about. We now take $\langle\pm1\rangle$-cohomology. For brevity, we will set $\mathrm{H}^\bullet_{\mathrm{T}}(M') := \mathrm{H}^\bullet_{\mathrm{T}}\left(\langle\pm1\rangle, U(M')^{C(N/M')}\right)$ for any admissible divisor $M' \mid N$.

$$\begin{array}{ccccccccccc} \mathrm{H}^{\bullet-1}_{\mathrm{T}}(\langle\pm1\rangle, K') & \longrightarrow & \mathrm{H}^\bullet_{\mathrm{T}}(M) & \longrightarrow & 0 & \longrightarrow & \mathrm{H}^\bullet_{\mathrm{T}}(\langle\pm1\rangle, K') & \longrightarrow & \mathrm{H}^{\bullet+1}_{\mathrm{T}}(M) & \longrightarrow & 0 \\ \| & & \downarrow{\scriptstyle (1-y_p)} & & \downarrow & & \| & & \downarrow{\scriptstyle (1-y_p)} & & \\ \mathrm{H}^{\bullet-1}_{\mathrm{T}}(\langle\pm1\rangle, K') & \longrightarrow & \mathrm{H}^\bullet_{\mathrm{T}}(M) & \longrightarrow & \mathrm{H}^\bullet_{\mathrm{T}}(Mp^{\nu_p}) & \longrightarrow & \mathrm{H}^\bullet_{\mathrm{T}}(\langle\pm1\rangle, K') & \longrightarrow & \mathrm{H}^{\bullet+1}_{\mathrm{T}}(M) & & \end{array}$$

Here, the 0s arise because $U(M)^{C(N/Mp^{\nu_p})}$ is free over $\langle\pm1\rangle$ by the third step. We now make a few simplications.

- The 0s in the top row imply that $\mathrm{H}^\bullet_{\mathrm{T}}(\langle\pm1\rangle, K') \to \mathrm{H}^{\bullet+1}_{\mathrm{T}}(M)$ is an isomorphism.
- By the previous step, we know that the $(1-y_p)$ arrows are the 0 map. Thus, the commutativity of the diagram implies that the arrows $\mathrm{H}^{\bullet-1}_{\mathrm{T}}(\langle\pm1\rangle, K') \to \mathrm{H}^\bullet_{\mathrm{T}}(M)$ and $\mathrm{H}^\bullet_{\mathrm{T}}(\langle\pm1\rangle, K') \to \mathrm{H}^{\bullet+1}_{\mathrm{T}}(M)$ in the bottom row are both the zero map.

The above two observations turns the bottom row into

$$0 \to \mathrm{H}^\bullet_{\mathrm{T}}(M) \to \mathrm{H}^\bullet_{\mathrm{T}}(Mp^{\nu_p}) \to \mathrm{H}^{\bullet+1}_{\mathrm{T}}(M) \to 0.$$

6. We now complete the proof by induction. We want to compute $\dim_{\mathbb{F}_2} \mathrm{H}^\bullet_{\mathrm{T}}\left(\langle\pm1\rangle, U(M)^{C(N/M)}\right)$ for any admissible divisor $M \mid N$. The previous step grants an "inductive step" that

$$\dim_{\mathbb{F}_2} \mathrm{H}^\bullet_{\mathrm{T}}\left(\langle\pm1\rangle, U(Mp^{\nu_p})^{C(N/Mp^{\nu_p})}\right) = \sum_{i\in\{0,1\}} \dim_{\mathbb{F}_2} \mathrm{H}^i_{\mathrm{T}}\left(\langle\pm1\rangle, U(M)^{C(N/M)}\right)$$

whenever M and Mp^{ν_p} is an admissible divisor of N. For example, we find that

$$\dim_{\mathbb{F}_2} \mathrm{H}_{\mathrm{T}}^{\bullet}\left(\langle\pm 1\rangle, U(p^{\nu_p})^{C(N/p^{\nu_p})}\right) = \sum_{i\in\{0,1\}} \dim_{\mathbb{F}_2} \mathrm{H}_{\mathrm{T}}^{i}\left(\langle\pm 1\rangle, U(1)^{C(N)}\right)$$

by taking $M = 1$. But now this dimension is independent of the cohomological index, so we inductively see that

$$\dim_{\mathbb{F}_2} \mathrm{H}_{\mathrm{T}}^{\bullet}\left(\langle\pm 1\rangle, U(M)^{C(N/M)}\right) = 2^{\omega(M)-1} \sum_{i\in\{0,1\}} \dim_{\mathbb{F}_2} \mathrm{H}_{\mathrm{T}}^{i}\left(\langle\pm 1\rangle, U(1)^{C(N)}\right)$$

for any admissible divisor $M \mid N$ such that $M > 1$.

The proof will be over as soon as we check

$$\sum_{i\in\{0,1\}} \dim_{\mathbb{F}_2} \mathrm{H}_{\mathrm{T}}^{i}\left(\langle\pm 1\rangle, U(1)^{C(N)}\right) = 1.$$

Well, note $U(1) = \mathbb{Z}[(\mathbb{Z}/N\mathbb{Z})^{\times}]$, so the $C(N)$-fixed points are given by $C(N) \cdot \mathbb{Z}[(\mathbb{Z}/N\mathbb{Z})^{\times}] = \mathbb{Z}C(N)$. This has trivial action by $\langle\pm 1\rangle$, so we are computing the Tate cohomology of the trivial $\langle\pm 1\rangle$-module $\mathbb{Z}$. Well, one has

$$\begin{cases} \mathrm{H}_{\mathrm{T}}^{0}(\langle\pm 1\rangle, \mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}, \\ \mathrm{H}_{\mathrm{T}}^{-1}(\langle\pm 1\rangle, \mathbb{Z}) = 0, \end{cases}$$

so we see that the sum of the $\mathbb{F}_2$-dimensions is in fact 1. ∎

Remark 4.68. Choose admissible divisors $M \mid M'$. The fifth step of the argument shows that there is an inclusion $U(M)^{C(N/M)} \subseteq U(M')^{C(N/M')}$ which then induces an inclusion

$$\mathrm{H}_{\mathrm{T}}^{\bullet}\left(\langle\pm 1\rangle, U(M)^{C(N/M)}\right) \to \mathrm{H}_{\mathrm{T}}^{\bullet}\left(\langle\pm 1\rangle, U(M')^{C(N/M')}\right)$$

on (Tate) cohomology. As seen in the fifth step of the argument, these inclusions explain "half" of the elements of a given $\mathrm{H}_{\mathrm{T}}^{\bullet}\left(U(M')^{C(N/M')}\right)$ by taking $M \mid M'$ to be M'/p^{ν_p} for some prime $p \mid M'$. The "other half" arises from a quotient and is thus harder to describe.

Example 4.69. Let's exhibit a nontrivial element in $\mathrm{H}_{\mathrm{T}}^{0}(\langle\pm 1\rangle, U_N)$. Remark 4.68 explains that there is an inclusion $\mathbb{Z}[(\mathbb{Z}/N\mathbb{Z})^{\times}]^{(\mathbb{Z}/N\mathbb{Z})^{\times}} \subseteq U_N$ which induces an inclusion

$$\mathrm{H}_{\mathrm{T}}^{0}\left(\langle\pm 1\rangle, \mathbb{Z}[(\mathbb{Z}/N\mathbb{Z})^{\times}]^{(\mathbb{Z}/N\mathbb{Z})^{\times}}\right) \subseteq \mathrm{H}_{\mathrm{T}}^{0}(\langle\pm 1\rangle, U_N).$$

Now, $\mathbb{Z}[(\mathbb{Z}/N\mathbb{Z})^{\times}]^{(\mathbb{Z}/N\mathbb{Z})^{\times}} \subseteq U_N$ is isomorphic to $\mathbb{Z}$ generated by $\sum_{a\in(\mathbb{Z}/N\mathbb{Z})^{\times}} \overline{a}$. Because this module has the trivial $\langle\pm 1\rangle$-action, we see that this generating element $\sum_{a\in(\mathbb{Z}/N\mathbb{Z})^{\times}} \overline{a}$ provides a nontrivial class in $\mathrm{H}_{\mathrm{T}}^{0}(\langle\pm 1\rangle, U_N)$.

4.3.5 Refined Algebraicity

The previous subsections (and in particular Theorem 4.64) allows us to upgrade Proposition 4.52.

Lemma 4.70. Let $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ be a function of constant weight. Then $2f$ is a $\mathbb{Z}$-linear combination of the functions 1_0 and $\varepsilon_{d,a}$ where $N \nmid da$.

Proof. This is [GGL24, Proposition 6.3.6]. By adding or subtracting 1_0s (which have weight 0), we may assume that $f(0) = 0$. By Proposition 4.52, we know that there is some denominator $D > 0$ such that Df is a $\mathbb{Z}$-linear combination of the functions 1_0 and $\varepsilon_{d,a}$ where $N \nmid da$, and we can see that there are no 1_0s because $f(0) = 0$. Thus, Lemma 4.57 tells us that Df (up to 1_0) vanishes in the group U_N^- described in Theorem 4.64. This group is actually 2-torsion by Theorem 4.64, so we conclude that $2f$ vanishes in U_N^-. Another application of Lemma 4.57 tells us that $2f$ is a $\mathbb{Z}$-linear combination of the $\varepsilon_{d,a}$s. ∎

Remark 4.71. In fact, once we know that $2f$ is a $\mathbb{Z}$-linear combination of 1_0 and the $\varepsilon_{d,a}$s, one can use some linear algebra to explicitly find this linear combination. We take a moment to note that Remark 4.58 tells us that we are allowed to only use $\varepsilon_{1,a}$s and $\varepsilon_{p,a}$s where $p \mid N$ is prime.

Here is the appliation to products of Γ.

Lemma 4.72. Let K_N be the extension of $\mathbb{Q}(\zeta_{2N}, i)$ generated by the elements $\pi^{-w}\Gamma(f)$, where the function $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ is of constant weight w and a $\mathbb{Z}$-linear combination of the $\varepsilon_{d,a}$s. Then

$$K_N = \mathbb{Q}(i, \zeta_{2N})\left(\{p^{p/N} : \text{prime } p \mid N\}\right).$$

Proof. It is enough to handle f which are equal to some $\varepsilon_{d,a}$. One can inductively write $\varepsilon_{d,a}$ as a sum of $\varepsilon_{1,\bullet}$s and $\varepsilon_{p,\bullet}$s, so we can just handle those. By the reflection formula (Proposition 4.40), $\Gamma(\varepsilon_{1,a})$ is in $\mathbb{Q}(\zeta_{2N}, i)$, so we don't have to worry about these elements.

Continuing, by the multipliation formula (Proposition 4.40), we see

$$\frac{\Gamma(\varepsilon_{p,a})}{\Gamma(\varepsilon_{1,pa})} = (2\pi)^{(p-1)/2} p^{1/2 - pa/N}.$$

We now have two cases on the parity of p.

- If p is odd, then these elements show $p^{1/2-pa/N} \in K_N$. However, $p^{1/2} \in \mathbb{Q}(i, \zeta_{2N})$ already, so we are only generating $p^{p/N} \in K_N$.
- Similarly, if $p = 2$, then these elements show $2^{1/2} \cdot 2^{1/2-2/N} \in K_N$. Thus, we are again only generating $2^{2/N} \in K_N$. ∎

Proposition 4.73. Let L_N be the extension of

$$K_N = \mathbb{Q}(i, \zeta_{2N})\left(\{p^{p/N} : \text{prime } p \mid N\}\right)$$

generated by the elements $\pi^{-w}\Gamma(f)$, where $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ is a function of constant weight w. Then the extension L_N/K_N is multiquadratic. If N is odd or divisible by 4, the degree is bounded by

$$\log_2[L_N : K_N] \le 2^{\omega(N)-1} - 1.$$

Proof. We proceed in steps, showing the various claims separately.

1. To check that this extension is multiquadratic, we will actually check that $(\pi^{-w}\Gamma(f))^2$ is in K_N for each f; note that $\pi^{-w}\Gamma(f)$ is already algebraic by Corollary 4.53. Now, by Lemma 4.70, we may write $2f$ as a $\mathbb{Z}$-linear combination of $\varepsilon_{d,a}$s, so $\Gamma(f)^2$ can be written as a product of $\Gamma(\varepsilon_{d,a})$s. But up to a power π, Lemma 4.72 assures us that $\Gamma(\varepsilon_{d,a})$ is in K_N.

2. It remains to bound $[L_N : K_N]$ when N is odd. By the previous step and Kummer theory [Lan02, Theorem VI.8.1], we would like to show that the 2-subgroup $\Gamma_N \subseteq K_N^\times / K_N^{\times 2}$ generated by the elements $(\pi^{-w}\Gamma(f))^2$ has
$$\dim_{\mathbb{F}_2} \Gamma_N \overset{?}{\leq} 2^{\omega(N)-1} - 1.$$
This bound will come from Theorem 4.64. To be formal, let $\varphi\colon \operatorname{Mor}_{\mathrm{cw}}(\frac{1}{N}\mathbb{Z}/\mathbb{Z}, \mathbb{Z}) \to K_N^\times$ be the homomorphism taking functions $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ of constant weight w to $(\pi^{-w}\Gamma(f))^2 \in K_N^\times$. By construction, we see that this homomorphism sends elements of the form $\varepsilon_{d,a}$ to $K_N^{\times 2}$, as discussed in Lemma 4.72. Thus, Lemma 4.57 tells us that φ descends to a homomorphism
$$\overline{\varphi}\colon U_{N,\mathrm{tors}}^- \to \frac{K_N^\times}{K_N^{\times 2}},$$
and Γ_N is the image of this map. Theorem 4.64 explains that the domain $U_{N,\mathrm{tors}}^-$ is already a 2-torsion group and has $\mathbb{F}_2$-dimension bounded by $2^{\omega(N)-1}$, so we will be done if we can lower the dimension any further.

3. We complete the proof by showing that $\overline{\varphi}$ has a nontrivial kernel. Indeed, consider the constant function $f_1 \equiv 1$. We have two checks.

   - On one hand, we claim that $f_1 \in \ker \varphi$. Then
   $$f_1 = 1_0 + 1_{2|N} 1_{1/2} \sum_{a=1}^{\lfloor (N-1)/2 \rfloor} \varepsilon_{1,a},$$
   where 1_0 and $1_{1/2}$ are indicators. Certainly 1_0 and $1_{1/2}$ are in $\ker \overline{\varphi}$ because $\Gamma(1_0) = \Gamma(1_{1/2}) = 1$ (see Example 4.42), and the $\varepsilon_{1,a}$s are in $\ker \overline{\varphi}$ as already noted. We conclude $f_1 \in \ker \varphi$.
   - On the other hand, we claim that f_1 is a nontrivial element of $U_{N,\mathrm{tors}}^-$. This is a little tricky. Under the isomorphism $U_{N,\mathrm{tors}}^- \cong \mathrm{H}_{\mathrm{T}}^0(\langle \pm 1 \rangle, U_N)$ of Proposition 4.65, f_1 corresponds to the (Tate) cohomology class
   $$\sum_{a \in (\mathbb{Z}/N\mathbb{Z})} \overline{a} \in \mathrm{H}_{\mathrm{T}}^0(\langle \pm 1 \rangle, U_N).$$
   However, this element is nontrivial by Example 4.69!

   We conclude that $\ker \overline{\varphi}$ is nontrivial, so $\Gamma_N = \operatorname{im} \overline{\varphi}$ satisfies
   $$\dim_{\mathbb{F}_2} \operatorname{im} \overline{\varphi} < \dim_{\mathbb{F}_2} U_{N,\mathrm{tors}}^-,$$
   so we are done by Theorem 4.64. ∎

Remark 4.74. The first step of the proof has the pleasant consequence of providing an explicit algorithm to compute the algebraic numbers $\pi^{-w}\Gamma(f)$, as discussed in [GGL24, Theorem 6.3.9]. Indeed, it suffices to compute the square $\pi^{-2w}\Gamma(2f)$. Now, Remark 4.71 says that we can use linear algebra to write $2f$ as a $\mathbb{Z}$-linear combination of some $\varepsilon_{d,a}$s, and then we can compute $\Gamma(\varepsilon_{d,a})$ using the reflection and multiplication formulae of Proposition 4.40 (as explained in Corollary 4.53).

Remark 4.75. Whether equality is achieved in Proposition 4.73 is an interesting question. It seems to be true in small examples; see Remark 4.78.

We now apply our theory to periods of the Fermat curve. To begin, we note that periods of Fermat curves can handle fairly general functions of constant weight.

Lemma 4.76. Let $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ be a function of constant weight w such that $f(0) = 0$. Then there is some index $p \geq 0$ and $\alpha \in \mathfrak{B}^{2p}$ and a list $\{a_i\}_{i=1}^n \subseteq (\mathbb{Z}/N\mathbb{Z})$ such that

$$f = 1_\alpha - \sum_{i=1}^n \varepsilon_{1,a_i}.$$

Proof. We will actually show that there is a list $\{a_i\}_{i=1}^n$ such that $f + \sum_{i=1}^n \varepsilon_{1,a_i}$ equals 1_α for some $\alpha \in \mathfrak{B}^{2p}$. In fact, it is enough to get $\alpha \in \mathfrak{A}^{2p}$: we already know that $f + \sum_{i=1}^n \varepsilon_{1,a_i}$ has constant weight by Lemma 4.50, so the weight will correctly be $3p$ as soon as this is some suitably $\alpha \in \mathfrak{A}^{2p}$ by Remark 4.20. As a last reduction, we note that we may assume $\operatorname{im} f \subseteq \mathbb{Z}_{\geq 0}$ by adding in suitable $\varepsilon_{1,\bullet}$s.

We now induct on $\|f\|_1 = \sum_{i=0}^n f(i/N)$. Here are some small cases.

- If $\|f\|_1 = 0$, then $f = 0$, and we can take $p = 0$ and α to be empty.
- It is not possible for f to be supported on a single nonzero entry because such a function cannot have constant weight.
- Suppose $\|f\|_1 = 2$. Because f should not be supported at a single point, we have $f = \overline{a/N} + \overline{b/N}$ for some $a, b \in \mathbb{Z}/N\mathbb{Z}$. We claim that $f = \varepsilon_{1,a}$. Well, f needs to have constant weight, so

  $$[a] + [b] = [-a] + [-b].$$

  Thus, $[a] + [b] = N$, so $b = -a$, as required.

We now proceed with the induction. Suppose that $\|f\|_1 > 2$. Because f is nonzero, f is supported on at least two points, which we name a/N and b/N where $a, b \in (\mathbb{Z}/N\mathbb{Z})$. We have two cases.

- Suppose that $b = -a$. Then $f - \varepsilon_{1,a}$ continues to have nonnegative image and constant weight, but $\|f - \varepsilon_{1,a}\|_1 < \|f\|_1$, so we may apply the inductive hypothesis to $f - \varepsilon_{1,a}$ to conclude the proof.
- Suppose that $b \neq -a$. Then there is a nonzero $c \in (\mathbb{Z}/N\mathbb{Z})$ such that $a + b + c = 0$, and we define $\alpha := (a, b, c)$ to be in $\mathfrak{A}^1$. We now see that

  $$f - 1_\alpha + \varepsilon_{1,c}$$

  has nonnegative image and constant weight, but $\|f - 1_\alpha + \varepsilon_{1,c}\|_1 < \|f\|_1$. We now again conclude by applying the inductive hypothesis. ∎

Theorem 4.77. Let K_A^{conn} be the connected monodromy field of the Jacobian A of the Fermat curve X_N, and define the field

$$K_N = \mathbb{Q}(i, \zeta_{2N})\left(\{p^{p/N} : \text{prime } p \mid N\}\right).$$

(a) We have $K_N \subseteq K_A^{\mathrm{conn}}(i, \zeta_{2N})$.

(b) The extension $K_A^{\mathrm{conn}}(i, \zeta_{2N})/K_N$ is multiquadratic.

(c) If N is odd or divisible by 4, then

$$\log_2[K_A^{\mathrm{conn}}(i, \zeta_{2N}) : K_N] \leq 2^{\omega(N)-1} - 1,$$

where $\omega(N)$ is the number of distinct prime factors of N.

Proof. As explaind in Remark 4.36, K_A^{conn} is the extension of $\mathbb{Q}(\zeta_N)$ which contains the periods

$$\operatorname{Per}\left(\gamma^{2p}, \nu_\alpha\right) = (2\pi i)^{-p} \prod_{i=1}^{2p} \zeta_{2N}^{[a_i]+[b_i]} \frac{\Gamma\left(\frac{[a_i]}{N}\right)\Gamma\left(\frac{[b_i]}{N}\right)}{\Gamma\left(\frac{[-c_i]}{N}\right)},$$

where $\alpha \in \mathfrak{B}^{2p}$ varies. By the reflection formula (Proposition 4.40), this period is in $\pi^{-\langle\alpha\rangle}\Gamma(1_\alpha)\mathbb{Q}(i,\zeta_{2N})$. Now, Lemma 4.76 explains that any function $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ of constant weight can be transformed into some 1_α for $\alpha \in \mathfrak{B}^{2p}$ at merely the cost of some 1_0s and $\varepsilon_{1,a}$s, so $K_A^{\mathrm{conn}}(i,\zeta_{2N})$ is actually generated by $\pi^{-w}\Gamma(f)$, where f may now vary over all functions $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ of some constant weight w. Part (a) now follows from Lemma 4.72, and parts (b) and (c) now follow from Proposition 4.73. ■

Remark 4.78. Let A be the Jacobian of the curve $y^2 = x^N - 1$, which is a quotient of the Fermat curve X_N. In [GGL24, Theorem 7.1.1], it is shown that K_A^{conn} is multiquadratic over merely $\mathbb{Q}(\zeta_N)$ via some algebro-geometric arguments. If N is odd, then Theorem 4.77 shows that

$$\log_2[K_A^{\mathrm{conn}}(i) : \mathbb{Q}(i,\zeta_N)] \le 2^{\omega(N)-1} - 1.$$

In particular, note that the $p^{p/N}$s define odd-degree cyclic extensions of $\mathbb{Q}(i,\zeta_N)$ and hence cannot live in the multiquadratic extension K_A^{conn} of $\mathbb{Q}(\zeta_N)$. The above bound agrees with the table in [GGL24, Example 6.4.10]; in fact, that table suggests that equality may hold without the added is!

Let's see an example computation.

Proposition 4.79. Define A to be the Jacobian of the proper curve C with affine chart $y^9 = x\left(x^2+1\right)$. Then we show $K_A^{\mathrm{conn}} = \mathbb{Q}\left(\zeta_N, 2^{1/3}, 2^{2/9}\cdot 3^{1/6}\right)$.

Proof. This computation follows the one in Proposition 4.38. We will freely use the computation executed in Proposition 3.33. Thoughout, $A \coloneqq \operatorname{Jac} C$, and we recall that we have a decomposition $A = C_0 \times A_1 \times A_2$ (over $\mathbb{Q}$) into geometrically simple abelian varieties. We proceed in steps.

1. Set $N \coloneqq 18$, and we note that there is a quotient map $X_N \to C$ from the affine patch $x^{18}+y^{18}+1 = 0$ to C given by $\psi(x,y) \coloneqq \left(x^9, xy^2\right)$. Thus, we will be able to use the Galois-invariant embedding $\psi\colon \mathrm{H}^1_{\text{ét}}(C_{\overline{\mathbb{Q}}},\mathbb{Q}_\ell) \hookrightarrow \mathrm{H}^1_{\text{ét}}(X_{N,\overline{\mathbb{Q}}},\mathbb{Q}_\ell)$ to use Theorem 4.33 by restricting to the Galois submodule. To make this explicit, we recall that we have a basis

$$\left\{\frac{dx}{y^4}, \frac{dx}{y^5}, \frac{dx}{y^6}, \frac{dx}{y^7}, \frac{dx}{y^8}, \frac{x\,dx}{y^7}, \frac{x\,dx}{y^8}\right\}$$

of $\mathrm{H}^{10}(C)$, we see that we can pass this basis through ψ^* to see that $\mathrm{H}^{10}(C) \subseteq \mathrm{H}^{10}(X)$ has basis

$$\left\{\nu_{5,10,3}, \nu_{4,8,6}, \nu_{3,6,9}, \nu_{2,4,12}, \nu_{1,2,15}, \nu_{11,4,3}, \nu_{10,2,6}\right\}.$$

Combining with the conjugate differentials yeilds a full basis of $\mathrm{H}^1_{\mathrm{dR}}(C,\mathbb{Q}) \subseteq \mathrm{H}^1_{\mathrm{dR}}(X,\mathbb{Q})$.

2. We pass to the étale site in exactly the same way as in Proposition 4.38. In the notation of Proposition 3.33, we see that ψ^* pulls the basis $\{u_1\otimes 1, v_1\otimes 1, v_2\otimes 1, v_4\otimes 1, w_1\otimes 1, w_2\otimes 1, w_5\otimes 1\}$ to

$$\{\nu_{3,6,9}\otimes 1, \quad \nu_{10,2,6}\otimes 1, \nu_{2,4,12}\otimes 1, \nu_{4,8,6}\otimes 1, \quad \nu_{1,2,15}\otimes 1, \nu_{11,4,3}\otimes 1, \nu_{5,10,3}\otimes 1\}.$$

3. We are now ready to begin executing Proposition 2.159; for this, Remark 2.160 informs us that we need to build a space of W' of Tate classes cutting out $G_\ell(A)^\circ \subseteq \mathrm{GL}_{14,\mathbb{Q}_\ell}$. We begin by adding W_1, made up of the endomorphisms, which ensures (for example) that $G_\ell(A)^\circ$ is diagonal. Then Proposition 3.33 computed that we also have the "polarization equations"

$$\begin{aligned}\mu_1\mu_2 &= \kappa_1\kappa_8,\\ \kappa_1\kappa_8 &= \kappa_2\kappa_7,\\ \kappa_1\kappa_8 &= \kappa_4\kappa_5,\end{aligned}$$

and the exceptional equation

$$\mu_1\kappa_7 = \kappa_5\kappa_8.$$

We remark that the polarization equations translate into a Tate class like $\nu_{(\alpha,-\alpha,\beta,-\beta)} \otimes 1$ understood as an element in $\mathrm{H}^4_{\text{ét}}(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)(2) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$, but this Tate class actually already come from a class in W_1 (see Corollary 4.37), so we may safely ignore it. Thus, we only have to translate the exceptional equation into the tensor

$$\nu_{(3,6,9),(7,14,15),(13,8,15),(1,2,15)} \otimes 1 \in \mathrm{H}^4_{\text{ét}}(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_\ell)(2) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$$

and its Galois orbit.

4. Arguing as in Remark 4.36, we know that the periods of the Tate classes given in the previous step generate K_A^{conn}, so it remains to compute these periods. We already know that our endomorphisms, except for the isogeny $(A_1)_{\overline{\mathbb{Q}}} \cong (A_2)_{\overline{\mathbb{Q}}}$, are defined over $\mathbb{Q}(\zeta_N)$ (see also Corollary 4.37). We now handle the remaining cycles.

   - The isogeny $A_1 \cong A_2$ corresponds to equations $\kappa_u = \lambda_{2u}$ for each $u \in (\mathbb{Z}/18\mathbb{Z})^\times$, which means that we would like to compute

     $$\mathrm{Per}\left(\gamma^{2p}, \nu_{u(1,2,15),u(16,14,6)}\right).$$

     Well, by Remark 4.14, this element is

     $$(-2\pi i)^{-1}\zeta_{2N}^{u(1+2+16+14)} \cdot \frac{\Gamma\left(\frac{[u]}{18}\right)\Gamma\left(\frac{[2u]}{18}\right)}{\Gamma\left(\frac{[3u]}{18}\right)} \cdot \frac{\Gamma\left(\frac{[16u]}{18}\right)\Gamma\left(\frac{[14u]}{18}\right)}{\Gamma\left(\frac{[12u]}{18}\right)}.$$

     One can check that the term on the left is in $\pi^{-1}\mathbb{Q}(\zeta_N)$, so it remains to handle the product of Γs. We handle the case where $u = 1$ because the others turn out to be essentially Galois conjugates. (Indeed, Theorem 4.33 explains that the remaining us belong to the same Galois orbit.) Using the algorithm suggested in Remark 4.74, one finds that this product equals

     $$\Gamma(-\varepsilon_{1,8} - \varepsilon_{2,3} - \varepsilon_{2,4} - \varepsilon_{2,6} - \varepsilon_{2,7} + \varepsilon_{3,1} + \varepsilon_{3,2} + 2\varepsilon_{3,4}),$$

     which evaluates to

     $$\left(-\zeta_N^5 + \zeta_N^2 + \zeta_N + 1\right) \cdot \left(2^{22} \cdot 3^3\right)^{1/18},$$

     up to a (correct) power of π.

   - It remains to compute

     $$\mathrm{Per}\left(\gamma^4, \nu_{(3,6,9),(7,14,15),(13,8,15),(1,2,15)} \otimes 1\right).$$

     Well, by Remark 4.14, we see this equals

     $$(-2\pi i)^{-2}\zeta_{2N}^{(3+6+7+14+13+8+1+2)} \cdot \frac{\Gamma\left(\frac{3}{18}\right)\Gamma\left(\frac{6}{18}\right)}{\Gamma\left(\frac{9}{18}\right)} \cdot \frac{\Gamma\left(\frac{7}{18}\right)\Gamma\left(\frac{14}{18}\right)}{\Gamma\left(\frac{3}{18}\right)} \cdot \frac{\Gamma\left(\frac{13}{18}\right)\Gamma\left(\frac{8}{18}\right)}{\Gamma\left(\frac{3}{18}\right)} \cdot \frac{\Gamma\left(\frac{1}{18}\right)\Gamma\left(\frac{2}{18}\right)}{\Gamma\left(\frac{3}{18}\right)}.$$

     As above, the term on the left belongs to $\pi^{-2}\mathbb{Q}(\zeta_N)$, so it remains to handle the product of Γs. Once again using the algorithm suggested in Remark 4.74, one finds that this product equals

     $$\Gamma\left(\varepsilon_{1,7} - \varepsilon_{1,8} + \frac{1}{2}\varepsilon_{1,9} + \varepsilon_{2,5} - \varepsilon_{2,7} + \varepsilon_{2,8} + \varepsilon_{3,1} - \varepsilon_{3,3} + \varepsilon_{3,4} - \varepsilon_{3,5}\right),$$

     which evaluates to $4 \cdot 2^{6/18}$ up to a (correct) power of π.

   Altogether, we can combine these two calculations to show $K_A^{\mathrm{conn}} = \mathbb{Q}\left(\zeta_N, 2^{2/9} \cdot 3^{1/6}\right)$ because this field already contains $2^{1/3}$. ∎

Remark 4.80. Perhaps it is notable that the exceptional Hodge class is defined over a smaller field than the endomorphisms!

4.4 Galois Action: the Crystalline Site

In this section, we use techinques from crystalline cohomology in order to compute the Galois action of the (geometric) Frobenius. Our exposition follows [Col87, Section IV]. For convenience, we will take p to be an odd prime not dividing N; this is enough to characterize the Galois action by the Chebotarev density theorem. Our main reference [Col87, Section IV] does permit $p = 2$, though one needs to work harder to handle $p \mid N$ (see [Col90]).

4.4.1 Morita's p-Adic Γ-Function

In (approximately) the same way that the Galois action on absolute Hodge classes produces periods which are products of Γs, the Galois action on the crystalline site will produce products of the p-adic Γ_p-function. Because we are taking $p \nmid N$, it will turn out to be enough to work with Morita's p-adic Γ_p-function. The purpose of the present section is to define Γ_p and give some of its basic properties. Our exposition follows [Rob00, Section 7.1].

In short, the function $\Gamma_p \colon \mathbb{Z}_p \to \mathbb{Z}_p$ is intended to be a continuous extension of the factorial function. However, the factorial function introduces many powers of p, which would cause Γ_p to vanish on large integers, which is undesirable, so it makes more sense to try to extend

$$n \mapsto \prod_{\substack{1 \le k < n \\ p \nmid k}} k.$$

However, this function fails to have a continuous extension: one has to add a sign.

Lemma 4.81. Fix an odd prime p. For any integers n and ν, where $\nu \ge 0$, we have

$$\prod_{\substack{n \le k < n + p^\nu \\ p \nmid k}} k \equiv -1 \pmod{p^\nu}.$$

Proof. If $\nu = 0$, there is nothing to do, so we assume $\nu \ge 1$. Taken $\pmod{p^\nu}$, this product is simply

$$\prod_{k \in (\mathbb{Z}/p^\nu\mathbb{Z})^\times} k.$$

Now, for each $k \notin \{\pm 1\}$, we have $k \not\equiv k^{-1}$, so we can pair k adn k^{-1} off in the product to cancel them out. Thus, the product is $1 \cdot -1 \equiv -1 \pmod{p^\nu}$. ∎

Proposition 4.82. Fix an odd prime p. The function $\Gamma_p \colon \mathbb{N} \to \mathbb{Z}_p^\times$ given by

$$\Gamma_p(n) \coloneqq (-1)^n \prod_{\substack{1 \le k < n \\ p \nmid k}} k$$

satisfies $|\Gamma_p(n) - \Gamma_p(m)|_p \le |n - m|_p$.

Proof. Suppose $|n - m|_p = p^{-\nu}$ so that $n - m = ap^\nu$ where $p \nmid a$. By symmetry, we may assume that $a \ge 0$. Then

$$\Gamma_p(n) - \Gamma_p(m) = \Gamma_p(m) \left((-1)^{n-m} \prod_{m \le k < m + ap^\nu} k - 1 \right).$$

Note $|\Gamma_p(m)|_p = 1$ by construction, so we may focus on the right-hand term. We must show that it is $0 \pmod{p^\nu}$. Well, Lemma 4.81 yields

$$(-1)^{n-m} \prod_{m \le k < m+ap^\nu} k \equiv (-1)^{n-m+a} \pmod{p^\nu},$$

so it is enough to note $n - m + a \equiv n - m + ap^\nu \equiv 0 \pmod 2$. (Recall p is odd!) ■

Definition 4.83 (Morita's Γ_p)**.** Fix an odd prime p. We define *Morita's* $\Gamma_p \colon \mathbb{Z}_p \to \mathbb{Z}_p^\times$ by continuously extending the map

$$\Gamma_p(n) \coloneqq (-1)^n \prod_{\substack{1 \le k < n \\ p \nmid k}} k.$$

In other words, $\Gamma_p(n+1) = (-1)^{n+1} n! / \left(p^{\lfloor n/p \rfloor} \lfloor n/p \rfloor !\right)$.

Remark 4.84. Let's explain why this continuous extension exists. Proposition 4.82 explains that $\Gamma_p|_{\mathbb{N}}$ is Lipshitz continuous (for the p-adic topology), so it is uniformly continuous, so it has a unique extension to $\mathbb{N} \subseteq \mathbb{Z}_p$ because $\mathbb{N}$ is dense in $\mathbb{Z}_p$.

Because Γ_p was constructed via a continuous extension from $\mathbb{N}$, we will only be able to prove facts about Γ_p using continuous extensions. Here is our analogue of Proposition 4.40.

Proposition 4.85. Fix an odd prime p. Define the auxiliary functions $h_p \colon \mathbb{Z}_p \to \mathbb{Z}_p$ and $R \colon \mathbb{Z}_p \to \{1, \dots, p\}$ by

$$h_p(x) \coloneqq \begin{cases} -x & \text{if } x \in \mathbb{Z}_p^\times, \\ -1 & \text{if } x \in p\mathbb{Z}_p, \end{cases} \qquad \text{and} \qquad R(x) \equiv x \pmod p.$$

(a) Translation: for $x \in \mathbb{Z}_p$, we have $\Gamma_p(x+1) = h_p(x)\Gamma_p(x)$.

(b) Reflection: for $x \in \mathbb{Z}_p$, we have $\Gamma_p(x)\Gamma_p(1-x) = (-1)^{R(x)}$.

(c) Multiplication: for any positive integer d with $p \nmid d$,

$$\Gamma_p(x)\Gamma_p\left(x + \frac{1}{d}\right) \cdots \Gamma_p\left(x + \frac{d-1}{d}\right) = \varepsilon_d d^{1-R(dx)} \left(d^{p-1}\right)^{\frac{R(dx)-dx}{p}} \Gamma_p(dx),$$

where $\varepsilon_d \coloneqq \Gamma_p(0)\Gamma_p(1/d) \cdots \Gamma_p((d-1)/d)$.

Remark 4.86. To understand this result, it is useful to note that every function we have written down is continuous. For example, h_p is continuous because it is the disjoint union of continuous functions, and R is continuous because it factors through $\mathbb{Z}_p/p\mathbb{Z}_p$. Lastly, the exponent map $x \mapsto \left(d^{p-1}\right)^{(R(dx)-dx)/p}$ is continuous: in fact, for $\alpha \in (1+p\mathbb{Z}_p)$, the exponent map $\alpha^\bullet \colon \mathbb{N} \to (1+p\mathbb{Z}_p)$ is Lipschitz continuous and hence admits a unique continuous extension to $\mathbb{Z}_p$. After factoring, it is enough to check $\beta^{p^\nu} - 1 \equiv 0 \pmod{p^\nu}$ for $\beta \in (1+p\mathbb{Z}_p)$, which we can show by induction on ν: there is nothing to do for $\nu = 0$, and the inductive step follows because

$$\beta^{p^{\nu+1}} - 1 = \underbrace{\left(\beta^{p^\nu} - 1\right)}_{\equiv 0 \pmod{p^\nu}} \underbrace{\left(1 + \beta^{p^\nu} + \beta^{2p^\nu} + \cdots + \beta^{(p-1)p^\nu}\right)}_{\equiv 0 \pmod p}.$$

Proof. This is in [Rob00, Sections 7.1.2–7.1.3]. Quickly, we note that all functions in sight are continuous by Remark 4.86. We are going to show each part individually and have a rather easier time than in Proposition 4.40, which we emphasize is because the density of $\mathbb{N} \subseteq \mathbb{Z}_p$ allows us to prove these results by induction (and combinatorics).

(a) We would like to show that $\Gamma_p(x+1)/\Gamma_p(x) = h_p(x)$. Both sides are continuous functions, so it is enough to check the result on the dense subset $\mathbb{N} \subseteq \mathbb{Z}_p$. Well, for $n \in \mathbb{N}$, we write

$$\frac{\Gamma_p(n+1)}{\Gamma_p(n)} = \frac{(-1)^{n+1}}{(-1)^n} \cdot \frac{\displaystyle\prod_{\substack{1 \le k \le n \\ p \nmid k}} k}{\displaystyle\prod_{\substack{1 \le k < n \\ p \nmid k}} k}.$$

The left factor provides a sign -1. As for the right factor, we get n if $p \nmid n$; otherwise if $p \mid n$, then the products are equal. In total, we see that the quotient is $h_p(n)$, as desired.

(b) Because both sides are continuous functions on $\mathbb{Z}_p$, it suffices to verify the equality for $n \in \mathbb{N}$, for which we use induction. Our base case is $n = 0$, where we see $\Gamma_p(0)\Gamma_p(1) = 1 \cdot -1$ equals $(-1)^p$ because p is odd. For the induction, we calculate

$$\begin{aligned}
\frac{\Gamma_p(x+1)\Gamma_p(1-(x+1))}{\Gamma_p(x)\Gamma_p(1-x)} &= \frac{h_p(x)\Gamma_p(x)\Gamma_p(-x)}{h_p(-x)\Gamma_p(x)\Gamma_p(-x)} \\
&= \frac{h_p(x)}{h_p(-x)} \\
&= \begin{cases} -1 & \text{if } x \in \mathbb{Z}_p^\times, \\ +1 & \text{if } x \in p\mathbb{Z}_p, \end{cases}
\end{aligned}$$

where the last equality follows from a quick piecewise computation of h_p. For the induction, it remains to check that $h_p(x)/h_p(-x) = (-1)^{R(x+1)-R(x)}$. Well, we may quickly compute

$$R(x+1) - R(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Z}_p^\times, \\ 1-p & \text{if } x \in p\mathbb{Z}_p, \end{cases}$$

by construction of $R(x)$, from which the result follows.

(c) For brevity, we define

$$G(x) \coloneqq \frac{1}{\Gamma_p(dx)} \prod_{k=0}^{d-1} \Gamma_p\left(x + \frac{k}{d}\right),$$

and we would like to show

$$G(x) \overset{?}{=} \varepsilon_d d^{1-R(dx)} \left(d^{p-1}\right)^{\frac{R(dx)-dx}{p}}.$$

Both sides are continuous functions, so we may check this on the dense subset $\frac{1}{d}\mathbb{N} \subseteq \mathbb{Z}_p$, which we will do by induction on $n \in \mathbb{N}$. At $n = 0$, the equation reads $G(0) = \varepsilon_d$, which is true by definition of ε_d. For the induction, we compute the ratio of consecutive terms on the left-hand side as

$$\begin{aligned}
\frac{G\left(x+\frac{1}{d}\right)}{G(x)} &= \frac{\Gamma_p(dx)}{\Gamma_p(dx+1)} \cdot \frac{\Gamma_p\left(x+\frac{1}{d}\right)\cdots\Gamma_p\left(x+\frac{d-1}{d}\right)\Gamma_p(x+1)}{\Gamma_p(x)\Gamma_p\left(x+\frac{1}{d}\right)\cdots\Gamma_p\left(x+\frac{d-1}{d}\right)} \\
&= \frac{h_p(x)}{h_p(dx)} \\
&= \begin{cases} d^{-1} & \text{if } x \in \mathbb{Z}_p^\times, \\ 1 & \text{if } x \in p\mathbb{Z}_p, \end{cases}
\end{aligned}$$

where the last equality follows by a quick piecewise computation of h_p.

It remains to compute the ratio of consecutive terms on the right-hand side. The ε_m will cancel out, so we are just asking to have the correct power of d, which we see is

$$\left(1 - R(dx+1) + (p-1)\cdot\frac{R(dx+1) - dx - 1}{p}\right) - \left(1 - R(dx) + (p-1)\cdot\frac{R(dx) - dx}{p}\right),$$

which simplifies to

$$-\frac{1}{p}(R(dx+1) - R(dx)) - \frac{p-1}{p} = \begin{cases} -1 & \text{if } x \in \mathbb{Z}_p^\times, \\ 0 & \text{if } x \in p\mathbb{Z}_p, \end{cases}$$

where we used the computation of $R(x+1) - R(x)$ in (b). The result follows. ∎

Now, Proposition 4.85 provides us with a distribution, as in Example 4.56.

Corollary 4.87. Fix an odd prime p. For any positive integer N with $p \nmid N$, the function $\Gamma_p \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}_p^\times/(\overline{\mathbb{Q}}^\times \cap \mathbb{Z}_p^\times)$ is an odd distribution.

Proof. This follows from Proposition 4.85. More precisely, the translation property allows our function to descend to the quotient $\frac{1}{N}\mathbb{Z}/\mathbb{Z}$, and the multiplication formula implies that we actually have a distribution. (One needs to combine the reflection formula with Corollary 4.91 to see why Γ_p outputs to $\overline{\mathbb{Q}}^\times$ on distribution relations.) Lastly, the reflection formula implies that our distribution is odd. ∎

Having a distribution gives allows us to compute certain products of Γ_p, in the same way as Γ.

Notation 4.88. Fix an odd prime p such that $p \nmid N$. For any function $f \colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$, we define

$$\Gamma_p(f) \coloneqq \prod_{i=1}^{N-1} \Gamma_p\left(\frac{i}{N}\right)^{f(i/N)}.$$

Remark 4.89. The same algorithm suggested in Remark 4.74 allows one to compute $\Gamma_p(f)$ for any f of constant weight. Once again, it suffices to compute the square $\Gamma_p(2f)$, but then some linear algebra discussed in Remark 4.71 allows one to express $2f$ as a $\mathbb{Z}$-linear combination of $\varepsilon_{d,a}$s, and $\Gamma_p(\varepsilon_{d,a})$ can be computed using the reflection and multiplication formulae.

4.4.2 An Encounter with Quadratic Reciprocity

In the multiplication formula of Proposition 4.85, one may complain that the value of ε_d is non-explicit, but the reflection formula implies that it is not too mysterious.

Example 4.90. Fix an odd prime p. By (b), we see that $\Gamma_p(1/2)^2 = (-1)^{R(1/2)} = (-1)^{(p+1)/2}$. In particular, if $p \equiv 1 \pmod 4$, then $\frac{p+1}{2}$ is odd, so

$$\Gamma_p\left(\frac{1}{2}\right)^2 = -1.$$

Thus, we have a somewhat explicit construction of $\sqrt{-1} \in \mathbb{Z}_p^\times$; in particular, it is the $\sqrt{-1}$ such that $\Gamma_p(1/2) \equiv \Gamma_p((p+1)/2) \equiv -((p-1)/2)! \pmod p$.

Corollary 4.91. Fix an odd prime p. For any positive integer d with $d \nmid N$, set

$$\varepsilon_d \coloneqq \prod_{k=0}^{d-1} \Gamma_p\left(\frac{k}{d}\right).$$

(a) If d is odd, then $\varepsilon_d \in \{\pm 1\}$ so that $\varepsilon_d^2 = +1$.

(b) If d is even, then $\varepsilon_d \in \{\pm \Gamma_p(1/2)\}$ so that $\varepsilon_d^2 = (-1)^{(p+1)/2}$.

Proof. Note that $\Gamma_p(0) = 1$, so we may safely ignore the $k = 0$ term. The point is to use the reflection formula of Proposition 4.85: if d is odd, we see

$$\varepsilon_d = \prod_{k=0}^{\lfloor (d-1)/2 \rfloor} \Gamma_p\left(\frac{k}{d}\right)\Gamma_p\left(1-\frac{k}{d}\right).$$

If d is even, the same equality holds with an added $\Gamma_p(1/2)$ on the right-hand side to account for the middle term $k = d/2$. Now, the reflection formula implies that each factor in the product is in $\{\pm 1\}$, so we see $\varepsilon_d \in \{\pm 1\}$ in the odd case and $\varepsilon_d \in \{\pm \Gamma_p(1/2)\}$ in the even case. The result then follows in the odd case, and the result follows in the even case by Example 4.90. ∎

Example 4.92. Fix an odd prime p with $3 \nmid p$, and we will compute $\Gamma_p(0)\Gamma_p(1/3)\Gamma_p(2/3)$. By the reflection formula (as in Corollary 4.91), this is $(-1)^{R(1/3)}$. Now, choosing $a \in \{1, 2\}$ such that $ap \equiv -1 \pmod 3$, we see that $R(1/3) = (ap+1)/3$, which is $a + 1 \pmod 2$, so we conclude

$$\varepsilon_3 = \begin{cases} -1 & \text{if } p \equiv 1 \pmod 3, \\ +1 & \text{if } p \equiv 2 \pmod 3. \end{cases}$$

An interesting feature of Example 4.92 is that

$$\varepsilon_3 = \left(\frac{-3}{p}\right),$$

where the right-hand side is the Legendre symbol. (This can be checked directly using quadratic reciprocity.) Some careful bookkeeping will show that this is true in general; amusingly, our bookkeeping will also contain a proof of quadratic reciprocity, though the proof presented here is almost certainly original to the author. Our argument follows [Coh07b, Theorem 6.4.14].

Definition 4.93 (half-system)**.** Fix an odd positive integer m. Then a *half-system* $H \subseteq \mathbb{Z}/m\mathbb{Z}$ is a subset for which there is a disjoint union

$$\mathbb{Z}/m\mathbb{Z} = H \sqcup -H \sqcup \{0\}.$$

Proposition 4.94. Fix an odd positive integer m and a half-system H. For each integer a not divisible by m, define

$$J_H(a, m) \coloneqq \prod_{i \in H} (-1)^{1_{\notin H}(ai)}.$$

Then $J_H(a, m) = \left(\frac{a}{m}\right)$, where the right-hand side is the Jacobi symbol.

Proof. This is basically an upgraded version of Gauss's lemma. We show this result in steps, following [Coh07a, Subsection 2.2.3].

For brevity, for $i \in (\mathbb{Z}/m\mathbb{Z}) \setminus \{0\}$, we define $\sigma_H(i)$ to be the unique element of $\{\pm i\} \cap H$, and we define $\varepsilon_H(i) = (-1)^{1_{\notin H}(i)}$ to be the sign so that $i = \varepsilon_H(i)\sigma_H(i)$. While we're here, we remark that $i \mapsto \sigma_H(ai)$ restricts to a bijection $H \to H$: it's enough to check that this map is injective, for which we note that having $\sigma_H(ai) = \sigma_H(aj)$ implies $ai \equiv \pm aj$, which implies $i \equiv \pm j$, which implies $i = j$ because $i, j \in H$ already.

1. We show that $J_H(a, m) = \left(\frac{a}{m}\right)$ when m equals a prime p. It is enough to check this equality $\pmod p$, for which we may use Euler's criterion to realize the left-hand side as $a^{(p-1)/2} \pmod p$. Drawing inspiration from a classical calculation of $a^{p-1} \equiv 1 \pmod p$, we note that
$$a^{(p-1)/2} \prod_{i \in H} i = \prod_{i \in H} ai.$$
Now, recall $ai \equiv \varepsilon_H(ai)\sigma_H(ai) \pmod p$, so
$$a^{(p-1)/2} \prod_{i \in H} i = \underbrace{\left(\prod_{i \in H} \varepsilon_H(ai)\right)}_{J_H(a,p)} \left(\prod_{i \in H} \sigma_H(ai)\right) \pmod p.$$
Because $i \mapsto \sigma_H(ai)$ is a bijection $H \to H$ (discussed above), the result follows.

2. We claim that J_H is independent of H. Well, let H' be another half-system. Choose a permutation π of $\{1, \dots, m-1\}$ so that $\pi(H) = H'$. Because H and H' are both half-systems, we note that π may as well be a product of transpositions of the form $(j, m-j)$. As such, we go ahead and let $\eta \colon H \to \{\pm 1\}$ be the sign so that $\pi(i) = \eta(i)i$ for each $i \in H$. Now, the main claim is that
$$\varepsilon_H(ai) \stackrel{?}{=} \eta(i)\eta(\sigma_H(ai))\varepsilon_{H'}(a\pi(i))$$
for each $i \in H$. Observe that this step will be concluded as soon as we take the product of the above equality over all i. The proof of the above equality is a direct calculation. On one hand, note $ai = \varepsilon_H(ai)\sigma_H(ai)$. On the other hand, ai is
$$\eta(i)a\pi(i) = \eta(i)\varepsilon_{H'}(a\pi(i))\sigma_{H'}(a\pi(i)).$$
Comparing the previous two sentences, it remains to show that
$$\sigma_{H'}(a\pi(i)) \stackrel{?}{=} \eta(\sigma_H(ai))\sigma_H(ai).$$
Note that both sides are $\pm ai \pmod m$, so it is enough to note that both sides are in H' by construction.
In the sequel, we may now abbreviate J_H to J because H does not matter.

3. We show that the function $J(a, m)$ is completely multiplicative in m. This will complete the proof by the definition of the Jacobi symbol. As such, choose odd positive integers m_1 and m_2 which are coprime to a, and we want to show that $J(a, m_1)J(a, m_2) = J(a, m_1m_2)$; set $m \coloneqq m_1m_2$ for brevity.
We will choose a special kind of half-system for this calculation. Choose half-systems $H_1 \subseteq \mathbb{Z}/m_1\mathbb{Z}$ and $H_2 \subseteq \mathbb{Z}/m_2\mathbb{Z}$, and then we construct the half-system
$$H \coloneqq \{i : i \pmod{m_1} \in H_1\} \cup m_1H_2.$$
Quickly, let's check that $H \subseteq \mathbb{Z}/m\mathbb{Z}$ is a half-system: given a nonzero $a \in \mathbb{Z}/m\mathbb{Z}$, we need to check that $\{\pm a\} \cap H$ is a singleton. We have two cases: if $a \pmod{m_1}$ is nonzero, then exactly one of $\pm a \pmod{m_1}$ is in H_1; otherwise, we may write $a = m_1b$ for some unique $b \in \mathbb{Z}/m_2\mathbb{Z}$, so exactly one of $\pm b$ is in H_2, completing the check.
It remains to compute $J(a, m)$, for which we write
$$\prod_{i \in H} \varepsilon_H(ai) = \left(\prod_{\substack{i \in H_1 \\ q \in (\mathbb{Z}/m_2\mathbb{Z})}} \varepsilon_H(a(i + qm_1))\right)\left(\prod_{b \in H_2} \varepsilon_H(am_1i)\right).$$
We now handle each product individually.

- For the left-hand product, we note that each term $\varepsilon_H(ai + aqm_1)$ has $ai + aqm_1 \equiv ai \not\equiv 0 \pmod{m_1}$, so $ai + aqm_1 \in H$ if and only if $ai \pmod{m_1} \in H_1$. Thus, our product is
$$\prod_{\substack{i\in H_1\\ q\in(\mathbb{Z}/m_2\mathbb{Z})}} \varepsilon_{H_1}(ai) = J(a,m_1)^{m_2}.$$
Because m_2 is odd, and $J(a,m_1) \in \{\pm 1\}$, this simplifies to $J(a,m_1)$.
- For the right-hand product, we note that $am_1 i$ is divisible by m_1, so $am_1 i \in H$ if and only if $ai \in H_2$. Thus, our product is
$$\prod_{b\in H_2} \varepsilon_{H_2}(ai) = J(a,m_2).$$

Combining the above two points reveals that $J(a,m) = J(a,m_1)J(a,m_2)$. ∎

Remark 4.95. Typical proofs of quadratic reciprocity are able to get away with using only the first step of Proposition 4.94 because they are allowed to work with only primes in the "denominators" of the Jacobi symbol. However, our application will desire general positive odd numbers, which has required the digression to half-systems in order to make possible the calculation in the third step of the proof.

Example 4.96. Consider the half-system $H = \{1,2,\dots,(m-1)/2\}$. Then we claim that $\varepsilon_H(ai) = (-1)^{\lfloor 2ai/m\rfloor}$, from which Proposition 4.94 will imply that
$$\left(\frac{a}{m}\right) = (-1)^{\sum_{i=1}^{(m-1)/2} \left\lfloor \frac{2ai}{m}\right\rfloor}.$$
Well, to show the claim, we write $ai = mq + r$ for some $r \in \{0,\dots,m-1\}$. Then $\lfloor \frac{2ai}{m}\rfloor \equiv \lfloor \frac{2r}{m}\rfloor \pmod 2$, which indicates if $r > m/2$. The claim follows.

Example 4.97. We use the half-system $H = \{1,2,\dots,(m-1)/2\}$ to show that $\left(\frac{2}{m}\right) = (-1)^{(m-1)(m+1)/8}$ (via Proposition 4.94). Indeed, for $i \in H$ (i.e., $i < m/2$), we see that $2i \notin H$ (i.e., $2i > m/2$) if and only if $i > m/4$. Thus, we are trying to count the number of integers between $m/4$ and $m/2$, which is just $\lfloor m/2\rfloor - \lfloor m/4\rfloor$. Writing $m = 8q + r$ for some $r \in [0,7]$, we see that $\lfloor m/2\rfloor - \lfloor m/4\rfloor$ depends only on r, and a direct calculation shows that it is even if and only if $r \in \{\pm 1 \pmod 8\}$. The claim follows.

Lemma 4.98. Fix odd and coprime positive integers m and n. Then
$$(-1)^{\sum_{i=1}^{(m-1)/2} R_n\left(\frac{i}{m}\right)} = \left(\frac{-n}{m}\right),$$
where $R_n\colon (\mathbb{Z}/n\mathbb{Z}) \to \{1,\dots,n\}$ chooses a representative of $\mathbb{Z}/n\mathbb{Z}$.

Proof. The idea is to make the left-hand sum look like the sum in Example 4.96. To this end, choose positive integers a and b such that $am = bn + 1$, which is possible because $\gcd(m,n) = 1$. Furthermore, by possibly replacing (a,b) with $(a+n, b+m)$, we may assume that b is even and a is odd. Our goal is to turn the remainders in the statement of the lemma into the quotients found in Example 4.96.

We are interested in the sum
$$\sum_{i=1}^{(m-1)/2} R_n\left(\frac{i}{m}\right) = \sum_{i=1}^{(m-1)/2} R_n\left(ai\right).$$

Now, the right-hand summands $R_n(ai)$ are usually $ai - n\left\lfloor \frac{ai}{n} \right\rfloor$, except when $n \mid ai$, in which case we just get an extra n. Of course, $n \mid ai$ happens exactly when $n \mid i$, for a total of $\lfloor m/(2n) \rfloor$ times. Thus, the right-hand sum is

$$\sum_{i=1}^{(m-1)/2} R_n(ai) = n\left\lfloor \frac{m}{2n} \right\rfloor + \sum_{i=1}^{(m-1)/2} ai - \sum_{i=1}^{(m-1)/2} \left\lfloor \frac{ai}{n} \right\rfloor.$$

Taken $\pmod 2$, this simplifies to

$$\frac{(m-1)(m+1)}{8} + \left\lfloor \frac{m}{2n} \right\rfloor + \sum_{i=1}^{(m-1)/2} \left\lfloor \frac{ai}{n} \right\rfloor.$$

(We have used the fact that a is odd.) Of course, really it is the last sum which is most interesting, for which we use a trick: $am = bn + 1$ implies that $bi/m = ai/n - i/mn$, and i/mn is less than $1/n$, so

$$\left\lfloor \frac{ai}{n} \right\rfloor = \left\lfloor \frac{bi}{m} \right\rfloor - 1_{n \mid i}.$$

Thus, we currently have

$$\frac{(m-1)(m+1)}{8} + \sum_{i=1}^{(m-1)/2} \left\lfloor \frac{bi}{m} \right\rfloor.$$

At last, we now put this in the exponent of (-1), which Examples 4.96 and 4.97 explains yields

$$\left(\frac{2}{m}\right)\left(\frac{b/2}{m}\right) = \left(\frac{b}{m}\right),$$

where we have used the fact that b is even. We are now done because $b \equiv -n^{-1} \pmod m$. ∎

Proposition 4.99. Fix an odd prime p. For any positive odd integer d coprime to p, we have

$$\prod_{i=0}^{d-1} \Gamma_p\left(\frac{i}{d}\right) = \left(\frac{-p}{d}\right).$$

Proof. We refine the argument of Corollary 4.91, following [Coh07b, Theorem 11.6.14]. By the reflection formula of Proposition 4.85, we see that our product is

$$\prod_{i=1}^{(d-1)/2} \Gamma_p\left(\frac{i}{d}\right)\Gamma_p\left(1 - \frac{i}{d}\right) = (-1)^{\sum_{i=1}^{(d-1)/2} R_p(i/d)}.$$

The result now follows from Lemma 4.98. ∎

While we're here, we note that we can use Lemma 4.98 to prove quadratic reciprocity.

Remark 4.100. One expects that one can use Lemma 4.98 to prove quadratic reciprocity because quadratic reciprocity approximately asserts that the Jacobi symbol $\left(\frac{n}{m}\right)$ is well-defined up to the class $m \pmod 4 n$, and Lemma 4.98 more or less only cares about information of $\pmod n$ due to the R_n.

Lemma 4.101. Let $O \subseteq \mathbb{N}$ be the set of positive odd integers. Suppose the function $J\colon O \times O \to \{\pm 1\}$ satisfies the following properties.

- $J(1,1) = 1$.
- $J(m+2n, n) = J(m,n)$.
- $J(m, 2m+n) = (-1)^{(m-1)/2} J(m,n)$.
- If $0 < m < 2n$, then $J(2n-m, n) = (-1)^{(n-1)/2} J(m,n)$.
- If $0 < n < 2m$, then $J(m, 2m-n) = (-1)^{(m-1)/2} J(m,n)$.

Then for any coprime $m, n \in O$, we have

$$J(m,n)J(n,m) = (-1)^{\frac{m-1}{2}\cdot\frac{n-1}{2}}.$$

Proof. We will show this by induction on $m+n$, where the base case of $m+n=2$ has $(m,n)=(1,1)$ and so $J(1,1)=1$. Note that the claim is symmetric in (m,n), so we may assume that $m<n$.

We now proceed with the induction. Given some (m,n), we have the following cases.

- If $n > 2m$, then we reduce the claim from (m,n) to $(m, n-2m)$. Well, we write

  $$J(m,n)J(n,m) = (-1)^{(m-1)/2} J(m, n-2m)\cdot J(n-2m, m).$$

  By the induction, this equals $(-1)^{(m-1)/2}\cdot(-1)^{(m-1)(n-2m-1)/4}$, so we are done upon noticing

  $$\frac{m-1}{2} + \frac{(m-1)(n-2m-1)}{4} \equiv \frac{(m-1)(n-1)}{4} \pmod 2,$$

  which we see by expanding $(m-1)(n-2m-1)/4 = (m-1)(n-1)/4 - m(m-1)/2$.

- If $n < 2m$, then we reduce the claim from (m,n) to $(m, 2m-n)$, where induction then applies $m + (2m-n) < m+n$ because $m<n$. Well, we write

  $$J(m,n)J(n,m) = (-1)^{(m-1)/2} J(m,2m-n)\cdot(-1)^{(m-1)/2} J(2m-n, m).$$

  By the induction, this equals $(-1)^{(m-1)(2m-n-1)/2}$, so we are done upon noticing

  $$\frac{(m-1)(2m-n-1)}{4} \equiv \frac{(m-1)(n-1)}{4} \pmod 2,$$

  which we see by expanding $(m-1)(2m-n-1)/4 = -(m-1)(n-1)/4 + (m-1)(2m-2)/4$. ∎

Theorem 4.102 (Quadratic reciprocity)**.** For any coprime odd positive integers m and n, we have

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{m-1}{2}\cdot\frac{n-1}{2}}.$$

Proof. We show that the function $J(m,n) := \left(\frac{m}{n}\right)$ satisfies the conditions of Lemma 4.101. Let's quickly deal with the easy ones: note $\left(\frac{1}{1}\right) = 1$ and $\left(\frac{m+2n}{n}\right) = \left(\frac{m}{n}\right)$ are properties of the Jacobi symbol. Continuing, Euler's criterion shows $\left(\frac{2n-m}{n}\right)$ equals

$$\left(\frac{-m}{n}\right) = (-1)^{(n-1)/2}\left(\frac{m}{n}\right).$$

There are two more checks, for which we will use Lemma 4.98. Quickly, we let $r_m\colon (\mathbb{Z}/m\mathbb{Z}) \to \{0,\dots,n-1\}$ denote the function which chooses a representative, and we note that Lemma 4.98 implies

$$(-1)^{\sum_{i=1}^{(n-1)/2} r_m\left(-\frac{i}{n}\right)} = \left(\frac{m}{n}\right)$$

because $r_m(-x) = m - R_m(x)$. We now run our last two checks.

- We show $\left(\frac{m}{2m+n}\right) = (-1)^{(m-1)/2}\left(\frac{m}{n}\right)$. As discussed above, Lemma 4.98 shows that it is enough to check
$$\sum_{i=1}^{(2m+n-1)/2} r_m\left(-\frac{i}{2m+n}\right) \stackrel{?}{\equiv} \frac{m-1}{2} + \sum_{i=1}^{(n-1)/2} r_m\left(-\frac{i}{n}\right) \pmod 2.$$
Well, note $r_m(-i/(2m+n)) = r_n(-i/n)$, so the real problem is the extra terms. Namely, we see
$$\sum_{i=1}^{(2m+n-1)/2} r_m\left(-\frac{i}{2m+n}\right) = \sum_{i=1}^{(n-1)/2} r_m\left(-\frac{i}{n}\right) + \sum_{i=(n+1)/2}^{(n-1)/2+m} r_m\left(-\frac{i}{n}\right).$$
This latter sum is a simply a sum over $i \in (\mathbb{Z}/m\mathbb{Z})$, so it comes out to $m(m-1)/2 \equiv (m-1)/2 \pmod 2$. The claim follows.
- If $0 < n < 2m$, we show $\left(\frac{m}{2m-n}\right) = (-1)^{(m-1)/2}\left(\frac{m}{n}\right)$. As discussed above, Lemma 4.98 shows that it is enough to check
$$\sum_{i=1}^{(2m-n-1)/2} r_m\left(-\frac{i}{2m-n}\right) + \sum_{i=1}^{(n-1)/2} r_m\left(-\frac{i}{n}\right) \stackrel{?}{\equiv} \frac{m-1}{2} \pmod 2.$$
The point is to replace i with $m-i$ in the left sum, which makes the summands into $r_m(i/(2m-n)) = r_m(-i/n)$. Thus, the two sums glue into
$$\sum_{i=1}^{m-1} r_m\left(-\frac{i}{n}\right) = \frac{m(m-1)}{2},$$
which is again $\equiv (m-1)/2 \pmod 2$. ∎

Remark 4.103. The usual proof of quadratic reciprocity via Gauss's lemma uses the identity
$$\sum_{i=1}^{(m-1)/2} \left\lfloor \frac{2ni}{m} \right\rfloor + \sum_{j=1}^{(n-1)/2} \left\lfloor \frac{2mi}{n} \right\rfloor = \frac{(m-1)(n-1)}{4},$$
which admits a geometric proof. We take a moment to note that this identity can be used to produce a similar proof of quadratic reciprocity as above because it directly shows
$$\left(\frac{m}{n}\right) = (-1)^{\frac{(m-1)(n-1)}{4} + \sum_{j=1}^{(n-1)/2} \left\lfloor \frac{2mi}{n} \right\rfloor},$$
from which the intuition of Remark 4.100 applies.

4.4.3 Computation of the Galois Action

In this section, we will use crystalline cohomology in order to compute the action of a geometric Frobenius. Our tool will be the following computational result of Coleman.

Theorem 4.104. Fix a smooth connected projective curve X over a finite unramified extension K of $\mathbb{Q}_p$. Let F be the induced geometric Frobenius on X. Fix a uniformizing parameter T at a point $x \in X$, and choose differentials ω and ν of the second kind on X which are regular on the residue class of x. Suppose ν is holomorphic on X and that $F^*\omega = \lambda\nu$ for some $\lambda \in K$. Given expansions

$$\omega = \sum_{n\geq 0} c_\omega(n) T^n \frac{dT}{T} \qquad \text{and} \qquad \nu = \sum_{n\geq 0} c_\nu(n) T^n \frac{dT}{T},$$

we have

$$\lambda = \lim_{i\to\infty} \frac{p \operatorname{Frob}_p c_\omega(n_i)}{c_\nu(pn_i)},$$

for some sequence $\{n_i\}_i$ of integers such that $|n_i|_p \to 0$.

We refer to [Col90, Theorem 17] for the proof, which uses properties of crystalline cohomology that we are too lazy to introduce here.

For our application where X is the Fermat curve, we note that $F^*\omega_\alpha$ lives in the span of $\omega_{p\alpha}$ by Theorem 4.33. (Notably, F is the absolute Frobenius on X, which can be seen to be the action by a geometric Frobenius at p, which satisfies $\zeta_N \mapsto \zeta_N^{p^{-1}}$. For example, one can check this on the level of the Tate module and then note that étale cohomology is dual.) Working out the relevant limit produces the following result.

Theorem 4.105. Fix an odd prime p not dividing N.

Proof. We use Theorem 4.104. In particular, the equation $x^N + y^N + 1 = 0$ allows us to write $\omega_{(a,b,c)} = -\zeta_{2N}^b x^a \left(x^N+1\right)^{b/N-1} \frac{dx}{x}$ at $(0, \zeta_{2N})$, so we achieve the expansion

$$\begin{aligned}\omega_{(a,b,c)} &= -\zeta_{2N}^b x^a \left(x^N+1\right)^{b/N-1} \frac{dx}{x} \\ &= -\zeta_{2N}^b \sum_{n\geq 0} \binom{\frac{b}{N}-1}{n} x^{a+Nn} \frac{dx}{x}.\end{aligned}$$

Now, we choose λ so that $\lambda \operatorname{Frob}_p^* \omega_{(a,b,c)} = \omega_{p^{-1}(a,b,c)}$, and our goal is to compute λ. ∎

4.4.4 Comparison of the Galois Actions

Theorem 4.106. Let $f\colon \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \mathbb{Z}$ be a function of constant weight w, and choose an odd prime $p \nmid N$. Fix a prime $\mathfrak{P}$ of $\overline{\mathbb{Q}}$ lying over p. Then

$$\iota_{\mathfrak{P}} \left(\frac{\operatorname{Frob}_{\mathfrak{P}} \pi^{-w} \Gamma(f)}{\pi^{-w}\Gamma(pf)} \right) = (-1)^w \Gamma_p(-pf),$$

where $\iota_{\mathfrak{P}}\colon \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$ is the induced embedding.

Let's go ahead and check some easy cases.

Example 4.107. Suppose f is 1_0. Then both sides of Theorem 4.106 are 1.

Example 4.108. We check Theorem 4.106 for $f = \varepsilon_{1,a}$, where $a \in \mathbb{Z}/N\mathbb{Z}$ is nonzero.

Proof. We compute the left-hand and right-hand sides separately.

- By the reflection formula for Γ (Proposition 4.40), the left-hand side equals
$$\frac{\mathrm{Frob}_{\mathfrak{P}}\left(\frac{1}{2\pi i}\cdot\frac{\pi}{\sin\frac{\pi a}{N}}\right)}{\frac{1}{2\pi i}\cdot\frac{\pi}{\sin\frac{\pi[pa]}{N}}} = \frac{\zeta_{2N}^{[pa]}-\zeta_{2N}^{-[pa]}}{\zeta_{2N}^{pa}-\zeta_{2N}^{-pa}}.$$
Writing $[pa] = pa - kN$ for $k \coloneqq \lfloor pa/N \rfloor$, we see that this quotient equals $\zeta_{2N}^{kN} = (-1)^k$.
- The reflection formla for Γ_p (Proposition 4.85) shows the right-hand side equals
$$(-1)\Gamma_p\left(\frac{[-pa]}{N}\right)\Gamma_p\left(1-\frac{[-pa]}{N}\right) = -(-1)^{R_p([-pa]/N)}.$$
Now, $[pa] = pa - kN$, so $[-pa] = -pa + kN + N$, so $[-pa]/N \equiv k+1 \pmod p$. Thus, the right-hand side also equals $(-1)^k$. ∎

Example 4.109. Fix a divisor $d \mid N$ and $a \in \mathbb{Z}/N\mathbb{Z}$. We check that Theorem 4.106 for $f = \varepsilon_{d,a}$ is equivalent to quadratic reciprocity for the positive odd integers d and p.

Proof. It will be enough to check the result for $\varepsilon_{d,a} - \varepsilon_{1,da}$, which will ease the application of the multiplication formula. We compute the left-hand and right-hand sides more or less separately.

Before starting these calculations, we note that we may as well choose a so that $a/N < 1/d$ (because our answer merely depends on the class $a \pmod{N/d}$). Further, we choose an integer b so that $0 < b/N < 1/d$ and $b/N \equiv pa/N \pmod{\frac{1}{d}\mathbb{Z}}$; such a b exists and is unique because it is equivalent to asking for $0 < b < N/d$ to satisfy $b \equiv pa \pmod{N/d}$. The point of choosing such a b is for $\frac{db}{N} = \left\{\frac{dpa}{N}\right\}$. As such, we also set $k \coloneqq \lfloor dpa/N \rfloor$ for brevity so that $db/N = dpa/N - k$.

- By the multiplication formula for Γ (Proposition 4.40), the left-hand side
$$\frac{\mathrm{Frob}_{\mathfrak{P}}\left((2\pi i)^{-(d-1)/2}\frac{1}{\Gamma(da/N)}\prod_{i=0}^{d-1}\Gamma\left(\frac{a}{N}+\frac{i}{d}\right)\right)}{(2\pi i)^{-(d-1)/2}\frac{1}{\Gamma(db/N)}\prod_{i=0}^{d-1}\Gamma\left(\frac{b}{N}+\frac{i}{d}\right)}$$
equals
$$\frac{\mathrm{Frob}_{\mathfrak{P}}\left(i^{-(d-1)/2}\cdot d^{1/2}\cdot d^{-da/N}\right)}{i^{-(d-1)/2}\cdot d^{1/2}\cdot d^{-db/N}}.$$
Now, for $a \in \mathbb{Z}_p^\times$, we see $\mathrm{Frob}_{\mathfrak{P}}(\sqrt{a}) = \pm\sqrt{a}$, where the $+$ sign is used if and only if $\sqrt{a} \in \mathbb{Z}_p^\times$; thus, $\mathrm{Frob}_{\mathfrak{P}}(\sqrt{a})/\sqrt{a} = \left(\frac{a}{p}\right)$. In total, this quantity collapses to
$$(-1)^{(d-1)(p-1)/4}\left(\frac{d}{p}\right)\cdot d^{-k}\cdot\frac{d^{pda/N}}{\mathrm{Frob}_{\mathfrak{P}}\, d^{da/N}}.$$
- By the multiplication formula for Γ_p (Proposition 4.85), the right-hand side
$$(-1)^{(d-1)/2}\cdot\frac{1}{\Gamma_p\left(1-\frac{db}{N}\right)}\prod_{i=0}^{d-1}\Gamma_p\left(\left(\frac{1}{d}-\frac{b}{N}\right)+\frac{i}{d}\right)$$

(note $\frac{1}{d} - \frac{b}{N} \equiv -\frac{pa}{N} \pmod{\frac{1}{d}\mathbb{Z}}$) equals

$$\left(\frac{p}{d}\right) d^{1-R(1-db/N)} \left(d^{p-1}\right)^{-\frac{1-db/N-R(1-db/N)}{p}},$$

where we have silently applied Proposition 4.99. Note that $1-db/N = 1+k-pda/N$, so $R(1-db/N) = R(1+k)$. Further, $0 < a/N < 1/d$ implies that $0 < dpa/N < p$, so $k < p$, so $R(1+k) = 1+k$. Thus, the above simplifies to

$$\left(\frac{p}{d}\right) d^{-k} \left(d^{p-1}\right)^{da/N}.$$

It remains to check that

$$(-1)^{(d-1)(p-1)/4} \left(\frac{d}{p}\right) \cdot d^{-k} \cdot \frac{d^{pda/N}}{\mathrm{Frob}_{\mathfrak{P}}\, d^{da/N}} \stackrel{?}{=} \left(\frac{p}{d}\right) d^{-k} \left(d^{p-1}\right)^{da/N}.$$

For example, we may cancel out the d^{-k}s. On one hand, taking the Nth power makes the power of d equal to $d^{(p-1)da}$ on both sides, so the equaliy implies an equality of the remaining signs, which we see to imply quadratic reciprocity.

On the other hand, given quadratic reciprocity, we may go ahead and cancel out those signs. Then we see that the Nth powers are equal, so these two elements may only differ by an element of $\mu_N \subseteq \overline{\mathbb{F}}_p^{\times} \hookrightarrow \mathbb{Q}_p^{\mathrm{unr}\times}$. In order to check correctness of this root of unity, it is enough to check $\pmod{\mathfrak{P}}$, where Frobenius is actually taking a pth power, from which the equality follows: indeed, both sides are $1 \pmod{\mathfrak{P}}$. ■

Remark 4.110. Combining Examples 4.107 to 4.109, we see that Theorem 4.106 has been checked "by hand" for any f which is a $\mathbb{Z}$-linear combination of the $\varepsilon_{d,a}$s (and 1_0). (We are implicitly using a multiplicativity of the formula.) However, any function f of constant weight has $2f$ equal to a $\mathbb{Z}$-linear combination of the $\varepsilon_{d,a}$s by Lemma 4.70, so we see that Theorem 4.106 will be true after squaring both sides (effectively doubling f). In other words, we have checked Theorem 4.106 "up to sign."

BIBLIOGRAPHY

[Wei49] André Weil. "Numbers of solutions of equations in finite fields". In: *Bull. Amer. Math. Soc.* 55 (1949), pp. 497–508. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1949-09219-4. URL: https://doi-org.libproxy.mit.edu/10.1090/S0002-9904-1949-09219-4.

[ST61] Gorō Shimura and Yutaka Taniyama. *Complex multiplication of Abelian varieties and its applications to number theory / by Goro Shimura and ... Yutaka Taniyama.* eng. Publications of the Mathematical Society of Japan ; 6. Tokyo: Mathematical Society of Japan, 1961.

[Tat66] John Tate. "Endomorphisms of abelian varieties over finite fields". In: *Inventiones mathematicae* 2.2 (Apr. 1966), pp. 134–144. ISSN: 1432-1297. DOI: 10.1007/BF01404549. URL: https://doi.org/10.1007/BF01404549.

[Poh68] Henry Pohlmann. "Algebraic Cycles on Abelian Varieties of Complex Multiplication Type". In: *Annals of Mathematics* 88.2 (1968), pp. 161–180. ISSN: 0003486X, 19398980. URL: http://www.jstor.org/stable/1970570 (visited on 12/04/2024).

[ST68] Jean-Pierre Serre and John Tate. "Good Reduction of Abelian Varieties". In: *Annals of Mathematics* 88.3 (1968), pp. 492–517. ISSN: 0003486X, 19398980. URL: http://www.jstor.org/stable/1970722 (visited on 03/30/2025).

[Sen73] Shankar Sen. "Lie Algebras of Galois Groups Arising from Hodge-Tate Modules". In: *Annals of Mathematics* 97.1 (1973), pp. 160–170. ISSN: 0003486X, 19398980. URL: http://www.jstor.org/stable/1970879 (visited on 12/16/2024).

[Del74] Pierre Deligne. "La conjecture de Weil. I". In: *Inst. Hautes Études Sci. Publ. Math.* 43 (1974), pp. 273–307. ISSN: 0073-8301. URL: http://www.numdam.org/item?id=PMIHES_1974__43__273_0.

[Mum74] David Mumford. *Abelian varieties*. eng. 2nd ed. / with appendices by C.P. Ramanujam and Yuri Manin. Published for the Tata Institute of Fundamental Research, Bombay [by] Oxford University Press, 1974. ISBN: 9780195605280.

[Rib76] Kenneth A. Ribet. "Galois Action on Division Points of Abelian Varieties with Real Multiplications". In: *American Journal of Mathematics* 98.3 (1976), pp. 751–804. URL: http://www.jstor.org/stable/2373815 (visited on 11/29/2024).

[Del77] P. Deligne. *Étale Cohomology: Starting Points*. 1977. URL: https://www.jmilne.org/math/Documents/DeligneArcata.pdf.

[Har77] Robin Hartshorne. *Algebraic Geometry*. Graduate Texts in Mathematics, No. 52. New York: Springer-Verlag, 1977.

[Ser77] Jean-Pierre Serre. *Linear representations of finite groups*. Graduate Texts in Mathematics, Vol. 42. Translated from the second French edition by Leonard L. Scott. Springer-Verlag, New York-Heidelberg, 1977, pp. x+170. ISBN: 0-387-90190-6.

[Del79] P. Deligne. "Valeurs de fonctions L et périodes d'intégrales". In: *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*. Proc. Sympos. Pure Math., XXXIII. With an appendix by N. Koblitz and A. Ogus. Amer. Math. Soc., Providence, RI, 1979, pp. 313–346.

[Kub79a] Daniel S. Kubert. "The $\mathbf{Z}/2\mathbf{Z}$ cohomology of the universal ordinary distribution". In: *Bull. Soc. Math. France* 107.2 (1979), pp. 203–224. ISSN: 0037-9484. URL: http://www.numdam.org/item?id=BSMF_1979__107__203_0.

[Kub79b] Daniel S. Kubert. "The universal ordinary distribution". In: *Bull. Soc. Math. France* 107.2 (1979), pp. 179–202. ISSN: 0037-9484. URL: http://www.numdam.org/item?id=BSMF_1979__107__179_0.

[Del80] Pierre Deligne. "La conjecture de Weil. II". In: *Inst. Hautes Études Sci. Publ. Math.* 52 (1980), pp. 137–252. ISSN: 0073-8301. URL: http://www.numdam.org/item?id=PMIHES_1980__52__137_0.

[New80] D. J. Newman. "Simple analytic proof of the prime number theorem". In: *Amer. Math. Monthly* 87.9 (1980), pp. 693–696. ISSN: 0002-9890. DOI: 10.2307/2321853. URL: https://doi-org.libproxy.berkeley.edu/10.2307/2321853.

[Zar83] Yu.G. Zarhin. In: *Journal für die reine und angewandte Mathematik* 1983.341 (1983), pp. 193–220. DOI: doi:10.1515/crll.1983.341.193. URL: https://doi.org/10.1515/crll.1983.341.193.

[Mur84] V. Kumar Murty. "Exceptional hodge classes on certain abelian varieties". In: *Mathematische Annalen* 268.2 (June 1984), pp. 197–206. ISSN: 1432-1807. DOI: 10.1007/BF01456085. URL: https://doi.org/10.1007/BF01456085.

[Fal86] Gerd Faltings. "Finiteness Theorems for Abelian Varieties over Number Fields". In: *Arithmetic Geometry*. Ed. by Gary Cornell and Joseph H. Silverman. New York, NY: Springer New York, 1986, pp. 9–26. ISBN: 978-1-4613-8655-1. DOI: 10.1007/978-1-4613-8655-1_2. URL: https://doi.org/10.1007/978-1-4613-8655-1_2.

[Ros86] Michael Rosen. "Abelian Varieties over $\mathbb{C}$". In: *Arithmetic Geometry*. Ed. by Gary Cornell and Joseph H. Silverman. New York, NY: Springer New York, 1986, pp. 79–101. ISBN: 978-1-4613-8655-1. DOI: 10.1007/978-1-4613-8655-1_4. URL: https://doi.org/10.1007/978-1-4613-8655-1_4.

[Col87] Robert F. Coleman. "The Gross-Koblitz formula". In: *Galois representations and arithmetic algebraic geometry (Kyoto, 1985/Tokyo, 1986)*. Vol. 12. Adv. Stud. Pure Math. North-Holland, Amsterdam, 1987, pp. 21–52. DOI: 10.2969/aspm/01210021. URL: https://doi-org.libproxy.berkeley.edu/10.2969/aspm/01210021.

[Kat88] Nicholas M. Katz. *Gauss sums, Kloosterman sums, and monodromy groups*. Vol. 116. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 1988, pp. x+246. ISBN: 0-691-08432-7; 0-691-08433-5. DOI: 10.1515/9781400882120. URL: https://doi-org.libproxy.mit.edu/10.1515/9781400882120.

[BLR90] Siegfried Bosch, Werner Lütkebohmert, and Michel Raynaud. *Néron models*. Vol. 21. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, 1990, pp. x+325. ISBN: 3-540-50587-3. DOI: 10.1007/978-3-642-51438-8. URL: https://doi.org/10.1007/978-3-642-51438-8.

[Col90] Robert F. Coleman. "On the Frobenius matrices of fermat curves". In: *p-adic Analysis: Proceedings of the International Conference held in Trento, Italy, May 29–June 2, 1989*. Ed. by Francesco Baldassarri, Siegfried Bosch, and Bernard Dwork. Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, pp. 173–193. ISBN: 978-3-540-46906-3. DOI: 10.1007/BFb0091138. URL: https://doi.org/10.1007/BFb0091138.

[Ich91] Takashi Ichikawa. "Algebraic groups associated with abelian varieties". en. In: *Mathematische Annalen* 289.1 (Mar. 1991), pp. 133–142. ISSN: 1432-1807. DOI: 10.1007/BF01446564. URL: https://doi.org/10.1007/BF01446564 (visited on 10/18/2024).

[Jan92] Uwe Jannsen. "Motives, numerical equivalence, and semi-simplicity". In: *Invent. Math.* 107.3 (1992), pp. 447–452. ISSN: 0020-9910. DOI: 10.1007/BF01231898. URL: https://doi-org.libproxy.mit.edu/10.1007/BF01231898.

[GH94] Phillip Griffiths and Joseph Harris. *Principles of algebraic geometry*. Wiley Classics Library. Reprint of the 1978 original. John Wiley & Sons, Inc., New York, 1994, pp. xiv+813. ISBN: 0-471-05059-8. DOI: 10.1002/9781118032527. URL: https://doi-org.libproxy.mit.edu/10.1002/9781118032527.

[Ill94] Luc Illusie. "Crystalline cohomology". In: *Motives (Seattle, WA, 1991)*. Vol. 55. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, 1994, pp. 43–70. DOI: 10.1090/pspum/055.1/1265522. URL: https://doi-org.libproxy.mit.edu/10.1090/pspum/055.1/1265522.

[Yan94] H. Yanai. "On Degenerate CM-Types". In: *Journal of Number Theory* 49.3 (1994), pp. 295–303. ISSN: 0022-314X. DOI: https://doi.org/10.1006/jnth.1994.1095. URL: https://www.sciencedirect.com/science/article/pii/S0022314X8471095X.

[LP95] M. Larsen and R. Pink. "Abelian varieties, ℓ-adic representations, and ℓ-independence". In: *Mathematische Annalen* 302.1 (May 1995), pp. 561–579. ISSN: 1432-1807. DOI: 10.1007/BF01444508. URL: https://doi.org/10.1007/BF01444508.

[MZ95] B. J. J. Moonen and Yu. G. Zarhin. "Hodge classes and Tate classes on simple abelian fourfolds". In: *Duke Math. J.* 77.3 (1995), pp. 553–581. ISSN: 0012-7094. DOI: 10.1215/S0012-7094-95-07717-5. URL: https://doi-org.libproxy.berkeley.edu/10.1215/S0012-7094-95-07717-5.

[Ful98] William Fulton. *Intersection theory*. Second. Vol. 2. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 1998, pp. xiv+470. ISBN: 3-540-62046-X; 0-387-98549-2. DOI: 10.1007/978-1-4612-1700-8. URL: https://doi-org.libproxy.mit.edu/10.1007/978-1-4612-1700-8.

[Ser98] Jean-Pierre Serre. *Abelian l-adic representations and elliptic curves*. Vol. 7. Research Notes in Mathematics. With the collaboration of Willem Kuyk and John Labute, Revised reprint of the 1968 original. A K Peters, Ltd., Wellesley, MA, 1998, p. 199. ISBN: 1-56881-077-6.

[Fol99] Gerald B. Folland. *Real analysis*. Second. Pure and Applied Mathematics (New York). Modern techniques and their applications, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. xvi+386. ISBN: 0-471-31716-0.

[Mil99] J. S. Milne. "Lefschetz classes on abelian varieties". In: *Duke Math. J.* 96.3 (1999), pp. 639–675. ISSN: 0012-7094. DOI: 10.1215/S0012-7094-99-09620-5. URL: https://doi-org.libproxy.berkeley.edu/10.1215/S0012-7094-99-09620-5.

[Moo99] Ben Moonen. *Notes on Mumford–Tate Groups*. 1999. URL: https://www.math.ru.nl/~bmoonen/Lecturenotes/CEBnotesMT.pdf (visited on 10/18/2024).

[Neu99] Jürgen Neukirch. *Algebraic number theory*. Vol. 322. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder. Springer-Verlag, Berlin, 1999, pp. xviii+571. ISBN: 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0. URL: https://doi-org.libproxy.berkeley.edu/10.1007/978-3-662-03983-0.

[RV99] Dinakar Ramakrishnan and Robert J. Valenza. *Fourier analysis on number fields*. Vol. 186. Graduate Texts in Mathematics. Springer-Verlag, New York, 1999, pp. xxii+350. ISBN: 0-387-98436-4. DOI: 10.1007/978-1-4757-3085-2. URL: https://doi-org.libproxy.berkeley.edu/10.1007/978-1-4757-3085-2.

[Rob00] Alain M. Robert. *A course in p-adic analysis*. Vol. 198. Graduate Texts in Mathematics. Springer-Verlag, New York, 2000, pp. xvi+437. ISBN: 0-387-98669-3. DOI: 10.1007/978-1-4757-3254-2. URL: https://doi-org.libproxy.mit.edu/10.1007/978-1-4757-3254-2.

[Hat01] Allen Hatcher. *Algebraic Topology*. Cambridge, 2001.

[Mat01] Lutz Mattner. "Complex differentiation under the integral". In: *Nieuw Arch. Wiskd. (5)* 2.1 (2001), pp. 32–35. ISSN: 0028-9825.

[Lan02] Serge Lang. *Algebra*. third. Vol. 211. Graduate Texts in Mathematics. Springer-Verlag, New York, 2002, pp. xvi+914. ISBN: 0-387-95385-X. DOI: 10.1007/978-1-4613-0041-0. URL: https://doi-org.libproxy.mit.edu/10.1007/978-1-4613-0041-0.

[And04] Yves André. *Une introduction aux motifs (motifs purs, motifs mixtes, périodes)*. Vol. 17. Panoramas et Synthèses [Panoramas and Syntheses]. Société Mathématique de France, Paris, 2004, pp. xii+261. ISBN: 2-85629-164-3.

[Con04] Brian Conrad. *The Main Theorem of Complex Multiplication*. 2004. URL: https://math.stanford.edu/~conrad/vigregroup/vigre04/mainthm.pdf.

[Rib04] Kenneth Ribet. *Review of Abelian ℓ-adic Representations and Elliptic Curves*. 2004. URL: https://math.berkeley.edu/~ribet/Articles/mg.pdf.

[ME05] M. Ram Murty and Jody Esmonde. *Problems in algebraic number theory*. Second. Vol. 190. Graduate Texts in Mathematics. Springer-Verlag, New York, 2005, pp. xvi+352. ISBN: 0-387-22182-4.

[Coh07a] Henri Cohen. *Number theory. Vol. I. Tools and Diophantine equations*. Vol. 239. Graduate Texts in Mathematics. Springer, New York, 2007, pp. xxiv+650. ISBN: 978-0-387-49922-2.

[Coh07b] Henri Cohen. *Number theory. Vol. II. Analytic and modern tools*. Vol. 240. Graduate Texts in Mathematics. Springer, New York, 2007, pp. xxiv+596. ISBN: 978-0-387-49893-5.

[Tay07] Richard Taylor. *Richard Taylor, The Sato–Tate conjecture*. Youtube. 2007. URL: https://www.youtube.com/watch?v=mpbWQbk18_g#t=20m15s.

[Vas07] Adrian Vasiu. *Some cases of the Mumford–Tate conjecture and Shimura varieties*. arXiv:math/0212066. Dec. 2007. DOI: 10.48550/arXiv.math/0212066. URL: http://arxiv.org/abs/math/0212066 (visited on 10/24/2024).

[Mil08] James S. Milne. *Abelian Varieties (v2.00)*. Available at www.jmilne.org/math/. 2008.

[Mur08] M. Ram Murty. *Problems in analytic number theory*. Second. Vol. 206. Graduate Texts in Mathematics. Readings in Mathematics. Springer, New York, 2008, pp. xxii+502. ISBN: 978-0-387-72349-5.

[NSW08] Jürgen Neukirch, Alexander Schmidt, and Kay Wingberg. *Cohomology of Number Fields*. 2nd ed. Grundlehren der mathematischen Wissenschaften. Springer Berlin, Heidelberg, 2008.

[Hei10] H. Heilbronn. "Zeta Functions and L-Functions". In: *Algebraic Number Theory: Proceedings of an Instructional Conference*. Ed. by J. W. S. Cassels and A. Fröhlich. 2nd ed. London Mathematical Society, 2010.

[Moo10] Ben Moonen. "Special subvarieties arising from families of cyclic covers of the projective line". In: *Doc. Math.* 15 (2010), pp. 793–819. ISSN: 1431-0635.

[Tat10] John T. Tate. "Fourier Analysis in Number Fields and Hecke's Zeta-Functions". In: *Algebraic Number Theory: Proceedings of an Instructional Conference*. Ed. by J. W. S. Cassels and A. Fröhlich. 2nd ed. London Mathematical Society, 2010.

[Bar+11] Tom Barnet-Lamb et al. "A family of Calabi-Yau varieties and potential automorphy II". In: *Publ. Res. Inst. Math. Sci.* 47.1 (2011), pp. 29–98. ISSN: 0034-5318. DOI: 10.2977/PRIMS/31. URL: https://doi.org/10.2977/PRIMS/31.

[Lan11] Serge Lang. *Complex Multiplication*. 1st ed. Springer New York, NY, 2011.

[DM12] P. Deligne and J.S. Milne. "Tannakian Categories". In: *Hodge Cycles, Motives, and Shimura Varieties*. Vol. 900. Lecture Notes in Mathematics. Springer Berlin, Heidelberg, 2012.

[Was12] Lawrence C. Washington. *Introduction to Cyclotomic Fields*. 2nd ed. Graduate Texts in Mathematics. Springer New York, NY, 2012.

[Lom13] Davide Lombardo. *Mumford-Tate groups and Hodge classes on Abelian varieties of low dimension*. 2013. URL: https://core.ac.uk/download/pdf/18603931.pdf.

[Mil13] James S. Milne. *Lectures on Etale Cohomology (v2.21)*. Available at www.jmilne.org/math/. 2013.

[Bar+14] Thomas Barnet-Lamb et al. "Potential automorphy and change of weight". In: *Ann. of Math. (2)* 179.2 (2014), pp. 501–609. ISSN: 0003-486X. DOI: 10.4007/annals.2014.179.2.3. URL: https://doi.org/10.4007/annals.2014.179.2.3.

[DE14] Anton Deitmar and Siegfried Echterhoff. *Principles of harmonic analysis*. Second. Universitext. Springer, Cham, 2014, pp. xiv+332. ISBN: 978-3-319-05791-0; 978-3-319-05792-7. DOI: 10.1007/978-3-319-05792-7. URL: https://doi.org/10.1007/978-3-319-05792-7.

[BK15] Grzegorz Banaszak and Kiran S. Kedlaya. "An Algebraic Sato-Tate Group and Sato-Tate Conjecture". In: *Indiana University Mathematics Journal* 64.1 (2015). Publisher: Indiana University Mathematics Department, pp. 245–274. ISSN: 0022-2518. URL: https://www.jstor.org/stable/26315458 (visited on 10/18/2024).

[Fit15] Francesc Fité. "Equidistribution, L-functions, and Sato-Tate groups". In: *Trends in number theory*. Vol. 649. Contemp. Math. Amer. Math. Soc., Providence, RI, 2015, pp. 63–88. DOI: 10.1090/conm/649/13020. URL: https://doi-org.libproxy.berkeley.edu/10.1090/conm/649/13020.

[Ton15] Brian Conrad (notes by Tony Feng). *Abelian Varieties*. 2015. URL: https://virtualmath1.stanford.edu/~conrad/249CS15Page/handouts/abvarnotes.pdf.

[Vat15] Akshaa Vatwani. "A simple proof of the Wiener–Ikehara Tauberian theorem". In: *Math. Student* 84.3-4 (2015), pp. 127–134.

[Yu15] Chia-Fu Yu. "A NOTE ON THE MUMFORD-TATE CONJECTURE FOR CM ABELIAN VARIETIES". In: *Taiwanese Journal of Mathematics* 19.4 (2015), pp. 1073–1084. ISSN: 10275487, 22246851. URL: http://www.jstor.org/stable/taiwjmath.19.4.1073 (visited on 11/30/2024).

[Fol16] Gerald B. Folland. *A course in abstract harmonic analysis*. Second. Textbooks in Mathematics. CRC Press, Boca Raton, FL, 2016, xiii+305 pp.+loose errata. ISBN: 978-1-4987-2713-6.

[Lom16] Davide Lombardo. "On the ℓ-adic Galois representations attached to nonsimple abelian varieties". In: *Ann. Inst. Fourier (Grenoble)* 66.3 (2016), pp. 1217–1245. ISSN: 0373-0956. DOI: 10.5802/aif.3035. URL: https://doi-org.libproxy.berkeley.edu/10.5802/aif.3035.

[Ots16] Noriyuki Otsubo. "Homology of the Fermat tower and universal measures for Jacobi sums". In: *Canad. Math. Bull.* 59.3 (2016), pp. 624–640. ISSN: 0008-4395. DOI: 10.4153/CMB-2016-012-0. URL: https://doi.org/10.4153/CMB-2016-012-0.

[Joh17] Christian Johansson. "On the Sato-Tate conjecture for non-generic abelian surfaces". In: *Trans. Amer. Math. Soc.* 369.9 (2017). With an appendix by Francesc Fité, pp. 6303–6325. ISSN: 0002-9947. DOI: 10.1090/tran/6847. URL: https://doi.org/10.1090/tran/6847.

[Mil17] J. S. Milne. *Algebraic Groups: The Theory of Group Schemes of Finite Type over a Field*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2017. DOI: 10.1017/9781316711736.

[Com18] Johan Commelin. *The Mumford–Tate conjecture for products of abelian varieties*. arXiv:1804.06840. Apr. 2018. URL: http://arxiv.org/abs/1804.06840 (visited on 10/18/2024).

[Del18] P. Deligne. "Hodge Cycles on Abelian Varieties". In: *Hodge Cycles, Motives, and Shimura Varieties*. Vol. 900. Lecture Notes in Mathematics. Springer Berlin, Heidelberg, 2018.

[Aru+19] Vishal Arul et al. "Computing zeta functions of cyclic covers in large characteristic". In: *Proceedings of the Thirteenth Algorithmic Number Theory Symposium*. Vol. 2. Open Book Ser. Math. Sci. Publ., Berkeley, CA, 2019, pp. 37–53.

[Sut19] Andrew Sutherland. "Sato-Tate distributions". en. In: *Contemporary Mathematics*. Ed. by Alina Bucur and David Zureick-Brown. Vol. 740. American Mathematical Society, 2019, pp. 197–248. ISBN: 978-1-4704-3784-8 978-1-4704-5629-0. DOI: 10.1090/conm/740/14904. URL: http://www.ams.org/conm/740 (visited on 10/18/2024).

[Mil20a] J.S. Milne. *Class Field Theory (v4.03)*. Available at www.jmilne.org/math/. 2020.

[Mil20b] James S. Milne. *Abelian Varieties (v0.10)*. Available at www.jmilne.org/math/. 2020.

[Ked21] Kiran S. Kedlaya. *Notes on Class Field Theory*. 2021. URL: https://kskedlaya.org/papers/cft-ptx.pdf.

[CC22] Victoria Cantoral-Farfán and Johan Commelin. "The Mumford-Tate conjecture implies the algebraic Sato-Tate conjecture of Banaszak and Kedlaya". In: *Indiana Univ. Math. J.* 71.6 (2022), pp. 2595–2603. ISSN: 0022-2518.

[SP] The Stacks project authors. *The Stacks project*. https://stacks.math.columbia.edu. 2022.

[Kre23] Miles Kretschmer. *Valedictorian Speech*. 2023. URL: https://www.youtube.com/live/UOK02uuq3G4?si=qMWVp9BIwQwWvZX7&t=1609.

[Vak23] Ravi Vakil. *The Rising Sea: Foundations of Algebraic Geometry*. 2023. URL: https://math.stanford.edu/~vakil/216blog/FOAGjul3123public.pdf.

[GGL24] Andrea Gallese, Heidi Goodson, and Davide Lombardo. *Monodromy groups and exceptional Hodge classes*. arXiv:2405.20394. July 2024. DOI: 10.48550/arXiv.2405.20394. URL: http://arxiv.org/abs/2405.20394 (visited on 10/18/2024).

[Con] Keith Conrad. *Stirling's Formula*. URL: https://kconrad.math.uconn.edu/blurbs/analysis/stirling.pdf.

[EGM] Bas Edixhoven, Gerard van der Geer, and Ben Moonen. *Abelian Varieties*. URL: http://van-der-geer.nl/~gerard/AV.pdf.

[Moo] Ben Moonen. *AN INTRODUCTION TO MUMFORD-TATE GROUPS*. en. URL: https://www.math.ru.nl/~bmoonen/Lecturenotes/MTGps.pdf.

[Spe] David E Speyer. *One-line proof of the Euler's reflection formula*. MathOverflow. URL:https://mathoverflow.net/q (version: 2017-08-08). eprint: https://mathoverflow.net/q/76447. URL: https://mathoverflow.net/q/76447.

[Var] Random Variable. *Ahlfors "Prove the formula of Gauss"*. Mathematics Stack Exchange. URL:https://math.stackex (version: 2017-01-15). eprint: https://math.stackexchange.com/q/753251. URL: https://math.stackexchange.com/q/753251.

List of Definitions