HABILITATION À DIRIGER DES RECHERCHES

Discipline : Mathématiques

présentée par

Gregorio Baldi

Lieu de Hodge et applications : quelques résultats et conjectures

Soutenue le 26 mai 2025 devant le jury composé de :

Yves André	Sorbonne Université	Examinateur
Benjamin Bakker	University Illinois Chicago	Rapporteur
Jean-Benoît Bost	Université Paris-Saclay	Rapporteur
François Charles	DMA - École normale supérieure	Examinateur
David Fisher	Rice University	Rapporteur
Claire Voisin	Sorbonne Université	Examinatrice

Contents

Abstract

> Time is like a river of passing events, a rushing torrent. For as soon as a thing has appeared, it is swept away, and another comes in its place, and this too will be swept away.
>
> M. Aurelius, Meditations (Book IV)

Algebraic Geometry studies algebraic varieties over complex numbers–the solution sets of algebraic equations with complex coefficients. A key idea, promoted by Noether, is that understanding a variety begins with examining its lower-dimensional subvarieties. After the introduction of topological methods, Hodge, followed by Deligne and Griffiths, enhanced the cohomology groups of these varieties with a $\mathbb{C}^*$-action, inspired by partial differential equations and harmonic integrals. This $\mathbb{C}^*$-action, conjecturally, is crucial for understanding subvarieties and hence the essential features of any variety.

Over the years this framework, closely linked to Grothendieck's motives, has led to several breakthroughs. When a variety evolves within a family (e.g., hypersurfaces of fixed degree), the first step in understanding the $\mathbb{C}^*$-action is to study those with topological cycles invariant under $\mathbb{C}^*$. This leads to the Hodge locus. Recent work, including mine, suggests that the behaviour of the Hodge locus in variations of Hodge structures is governed by a simple codimension computation, described as the "typical-vs-atypical" dichotomy. The atypical part connects with insights from Arithmetic Geometry, model theory, Galois theory of foliations, and transcendence theory. A significant outcome, demonstrated by Klingler, Ullmo, and myself, shows that the Hodge locus of the moduli space of hypersurfaces with sufficiently high dimension and degree is a finite union of algebraic subvarieties, and isolated points having lower-than-expected codimension. In this document I survey the main ideas behind such works and various applications.

Beyond the standard Hodge theoretic setting (which includes the well-known and studied case of the Noether-Lefschetz locus), the two most notable cases of study will be:

- representations of complex hyperbolic lattices,
- orbit closures inside strata of abelian differentials.

In particular, the *Zilber-Pink philosophy* for variations of mixed Hodge structures will be used as a unifying principle to several finiteness results scattered in the literature. In the rest of the text I will present my main results on the topic. I take this opportunity to simplify some parts of the theory developed in those papers–e.g. there's a first special case of the geometric ZP in [23], a more general one (along the same lines) in [21] and the most general one in the later [24]. In particular I tried to tell a simple story, with the sketch of the main arguments, in a historically reversed order.

Abstract (version francaise)

La géométrie algébrique étudie les variétés algébriques sur les nombres complexes–les ensembles de solutions d'équations algébriques à coefficients complexes. Une idée clé, promue par Noether, est que la compréhension d'une variété commence par l'examen de ses sous-variétés de dimension inférieure. Après l'introduction des méthodes topologiques, Hodge, suivi par Deligne et Griffiths, ont enrichi les groupes de cohomologie de ces variétés avec une action de $\mathbb{C}^*$ inspirée par les équations différentielles partielles et les intégrales harmoniques. Cette action de $\mathbb{C}^*$ est, conjecturalement, cruciale pour comprendre les sous-variétés et donc les caractéristiques essentielles de toute variété.

Au fil des ans, ce cadre, étroitement lié aux motifs de Grothendieck, a conduit à plusieurs percées. Lorsqu'une variété évolue au sein d'une famille (par exemple, les hypersurfaces de degré fixe), la première étape pour comprendre l'action de $\mathbb{C}^*$ est d'étudier celles ayant des cycles

topologiques invariants sous $\mathbb{C}^*$. Cela conduit au lieu de Hodge. Des travaux récents, y compris les miens, suggèrent que le comportement du lieu de Hodge dans les variations de structures de Hodge est gouverné par un simple calcul de codimension, décrit comme la dichotomie « typique vs atypique ». La partie atypique est liée à des idées issues de la géométrie arithmétique, de la théorie des modèles, de la théorie de Galois des feuilletages, et de la théorie de la transcendance. Un résultat significatif, démontré par Klingler, Ullmo et moi-même, montre que le lieu de Hodge de l'espace des modules des hypersurfaces de dimension et de degré suffisamment élevés est une union finie de sous-variétés algébriques, avec des points isolés de codimension inférieure à celle attendue. Dans ce document, je présente les idées principales derrière ces travaux ainsi que diverses applications.

Au-delà du cadre classique de la théorie de Hodge (qui inclut le cas bien connu et étudié du lieu de Noether-Lefschetz), les deux cas d'étude les plus notables sont :

- les représentations des réseaux hyperboliques complexes,
- les clôtures d'orbites dans les strates de différentielles abéliennes.

En particulier, la *philosophie de Zilber-Pink* pour les variations de structures de Hodge mixtes sera utilisée comme principe unificateur pour plusieurs résultats de finitude dispersés dans la littérature. Dans la suite du texte, je présenterai mes principaux résultats sur ce sujet. J'en profite pour simplifier certaines parties de la théorie développée dans ces articles–par exemple, un premier cas particulier de la ZP géométrique se trouve dans [23], un cas plus général (dans la même lignée) dans [21], et le cas le plus général dans le plus récent [24]. J'ai essayé de raconter une histoire simple, avec un aperçu des principaux arguments, dans un ordre historiquement inversé.

Selected List of the author's publications

(A) **Effective atypical intersections and applications to orbit closures**.
With D. Urbanik.
arXiv:2406.16628. See [24].

(B) **Rich representations and superrigidity**.
With N. Miller, M. Stover, and E. Ullmo.
Ergodic Theory Dynam. Systems. Published online 2025:1-24. doi:10.1017/etds.2024.136. See [22].

(C) **Non-density of the exceptional components of the Noether-Lefschetz locus**.
With B. Klingler and E. Ullmo.
IMRN, Volume 2024, Issue 21, November 2024, Pages 13642-13650. See [20].

(D) **On the distribution of the Hodge locus**.
With B. Klingler and E. Ullmo.
Invent. Math., 235(2):441-487, 2024. See [21].

(E) **On the geometric Zilber-Pink theorem and the Lawrence-Venkatesh method**.
With B. Klingler and E. Ullmo.
Expo. Math. 41 (2023), no.3, 718-722. See [19].

(F) **Special subvarieties of non-arithmetic ball quotients and Hodge theory**.
With E. Ullmo.
Ann. of Math., 197(1):159-220, (2023). See [23].

Finally we point out two recent surveys of the author [17, 18] which, among other things, cover also some related work.

Acknowledgements. First of all, I'd like to thank all my collaborators: B. Klingler, N. Miller, M. Stover, E. Ullmo, and D. Urbanik. A special thanks to Emmanuel: we started together this beautiful trip (at the beginning of 2019), and I've learned a lot from you. I am also grateful to the member of the jury and the referees of the manuscript. A special thank to J.-B. Bost for his detailed comments that helped improving the text.

The list of amazing people I've met in this journey is certainly too long to fit here, but I've enjoined many conversations over the years. I'd like to thank the many friends I've encountered in this path, in particular E. Ambrosi and my colleagues at the IMJ. Finally thank you Giada for always being with me.

Part 1. Preliminaries

1. Introduction

1.1. **Hodge Theory.** Compact Kähler manifolds are peculiar in that they are equipped with several compatible structures: complex, Riemannian, and symplectic. This rich interplay makes them fundamental objects in mathematics. Classical Hodge theory [71] enhances the classical topological invariants of the underlying space of a Kähler manifold X by incorporating a $\mathbb{C}^*$-action. Building on de Rham's work, the key insight is that for any given Riemannian metric on X, every cohomology class has a canonical **harmonic** representative. Since complex differential forms on can be uniquely expressed as sums of forms of type (p, q), the fact that the (p, q) components of a harmonic form remain harmonic gives rise to the celebrated Hodge decomposition. This decomposition splits the cohomology of X with complex coefficients into a direct sum of complex vector spaces:

$$H^n(X, \mathbb{C}) = \bigoplus_{p+q=n} H^{p,q}(X) \qquad\qquad (\textbf{HoDec}).$$

This decomposition is elegantly encoded by a $\mathbb{C}^*$-action, where $H^{p,q}$ is the subspace on which $z \in \mathbb{C}^*$ acts as $z^p \overline{z}^q$. The symmetry under complex conjugation (namely $\overline{H^{p,q}} = H^{q,p}$) reflects the fact that $\mathbb{C}^*$ is viewed as a real algebraic torus.

Classical Algebraic Geometry, on the other hand, studies algebraic varieties over complex numbers–the solution sets of algebraic equations with complex coefficients. The Kodaira embedding theorem characterizes smooth complex projective varieties among all compact Kähler manifolds and underpins the concept of a polarized Hodge structure, which grants favorable properties to the period domains of such objects that play a central role in Griffiths' approach [66, 61] to Hodge theory. Another key distinction between algebraic and general Kähler manifolds is that while general Kähler manifolds may not admit complex submanifolds of positive dimension, algebraic varieties possess them in abundance. For instance, an algebraic subvariety $Z \subset X$ of codimension j (or more generally, an algebraic cycle) gives rise to an **algebraic** cohomology class $[Z] \in H^{2k}(X, \mathbb{Q})$, whose image in $H^{2k}(X, \mathbb{C})$ lies in $H^{k,k}(X)$ relatively to (HoDec). A rational cohomology class lying in some $H^{k,k}(X)$ is called a **Hodge class** (equivalently those are classes fixed by the action of $\mathbb{C}^*$). This suggests the tempting belief that the Hodge structure of a compact complex variety is governed by the geometry of its subvarieties, or more precisely, its Chow group. A special case of this belief is indeed one of the Clay millennium problems:

Conjecture 1.1 (Hodge)**.** *For every smooth projective complex algebraic variety X, a class $\lambda \in H^*(X, \mathbb{Q})$ is algebraic if and only if it is Hodge.*

This conjecture remains a central open problem in mathematics, lying at the intersection of Complex Analytic and Algebraic Geometry. Roughly speaking, its validity asserts that there are many algebraic cycles.

1.2. **Cycles and Hodge classes in families: the Hodge locus.** Recent years have seen significant activity in understanding how algebraic cycles vary within families, following a path laid in the previous centuries by figures such as M. Noether, Picard, Poincaré, Lefschetz, Weil, Griffiths, Grothendieck, and Deligne, among others. For a smooth projective family of algebraic varieties $f : \mathcal{X} \to S$, the Hodge structures associated with each fiber $\mathcal{X}_s := f^{-1}(s)$ enrich the local system $R^* f_* \mathbb{Z}_{\text{prim}}$ of primitive cohomology with the structure of $\mathbb{V}$, a **polarized variation of Hodge structures** (VHS). Since Hodge theory is at its core a transcendental phenomenon, understanding how algebraic cycles and Hodge classes vary with $s \in S(\mathbb{C})$ is a challenging problem. This has led to the study of the Hodge locus, a subset of the base $S(\mathbb{C})$ where the dimension (over $\mathbb{Q}$) of the cohomology group $H^{k,k}(\mathcal{X}_s, \mathbb{Q})$, or possibly some tensor construction thereof, is larger than expected.

In 1995, Cattani-Deligne-Kaplan [33] achieved a breakthrough in Hodge theory by proving that the Hodge locus on a nonsingular projective variety S parametrizing a polarized VHS $\mathbb{V}$ is a countable union of **algebraic** subvarieties. More recently, this result has been established using methods from o-minimal geometry [13]. The Cattani-Deligne-Kaplan result is regarded by many as the most compelling evidence for the Hodge conjecture's validity, alongside Deligne's proof that Hodge classes on abelian varieties are absolute and various special cases such as the Lefschetz (1,1) theorem, and results by Schoen and Markman on Hodge-Weil 4-folds (very recently culminated in [85]). The Hodge locus also plays a central role in conjectures like André-Oort and Zilber-Pink, which will be discussed in more detail later.

A simple yet rich example, still far from being fully understood, involves hypersurfaces. A hypersurface X of degree d and dimension n is defined as the zero set in $\mathbb{CP}^n+1$ of a homogeneous polynomial F of degree d in $n+2$ variable. It is smooth if the partial derivatives of F do not vanish simultaneously on $\mathbb{C}^{n+2} - \{0\}$.

For fixed n and d, the parameter space of hypersurfaces in $\mathbb{CP}^{n+1}$ of degree d is simply the complex vector space of homogeneous polynomials of degree d, denoted $V = \mathbb{C}[X_0, \dots, X_{n+1}]_d$. However, the parameter space of **smooth** hypersurfaces $U_{n,d} = V - \Delta$ is more intricate. The relevant VHS enriches the (primitive) cohomology $H^n(X, \mathbb{Z})$, and the Hodge locus remains far from understood. In the case $n = 2$ and d, this locus strictly contains the **Noether-Lefschetz locus**, which can be defined purely using algebraic geometry:

$$\mathrm{NL}_d := \{[X] \in U_{2,d} : \mathrm{Pic}(\mathbb{CP}^3) \to \mathrm{Pic}(X) \text{ is not an isomorphism}\}.$$

For a surface X outside NL_d, every curve on X has the pleasant and useful property that it is the complete intersection of X with another surface in $\mathbb{CP}^3$. This object has been the subject of many beautiful studies by Griffiths, Green, Voisin, Ciliberto, Harris, and Miranda and recently by myself with Klingler and Ullmo [20].

The most pressing challenges in the field revolve around two fundamental questions:

Question 1.2. What can we say about about Hodge loci? What do Hodge loci reveal to us?

1.3. **The typical-vs-atypical dichotomy.** In the case of arbitrary VHS, one of the most representative results is one I obtained in collaboration with B. Klingler and E. Ullmo (see also the ICM report [76] for more details). Let $(S, \mathbb{V})$ be a pair where S is a smooth quasi-projective complex variety and $\mathbb{V}$ an integral and polarised VHS on S, and denote by $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ the Hodge locus (more precise definitions will come later). We have:

Theorem 1.3 ([21, Sec. 2])**.** *The Hodge locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ *can be decomposed into its* ***typical*** *and* ***atypical*** *part, behaving as follows:*

(1) If the typical Hodge locus is non-empty, then the Hodge locus is dense in S.
(2) The geometric Zilber-Pink (ZP) conjecture holds true. That is, there are only finitely many maximal atypical special subvarieties of positive *period dimension.*

If the ***level*** *of $\mathbb{V}$ is ≥ 3 (cf. Section 5.1), then the positive-dimensional Hodge locus is a strict algebraic subvariety of S. (Equivalently, there are only finitely many* maximal *special subvarieties of positive period dimension.)*

From the period-domain perspective, the Hodge locus is simply the **intersection** between sub-period domains and the image along the period map of S in $\Gamma\backslash D$. Informally, a special subvariety is considered **atypical** if its codimension is less than expected, based on a simple dimension count within the period domain $\Gamma\backslash D$ naturally attached to $\mathbb{V}$. Otherwise, it is **typical**.

The concept of level refines the weight of a Hodge structure. These notions, though simple, are pervasive. They have even found interesting applications in string theory, as demonstrated by the work of Grimm and van de Heisteeg [67], and I believe such applications are on the cusp of substantial growth.

Conjecturally, the **zero-dimensional** atypical Hodge locus should consist of finitely many maximal components. This part of the Zilber-Pink conjecture (referred to as the arithmetic part of ZP) remains elusive and is the richest in arithmetic properties. Only for variations of Hodge structures of level 1 has the Pila-Zannier strategy [95] managed to reduce the problem to a *Large*

Galois Orbits hypothesis [38]. This case corresponds largely to that of Shimura varieties, where the paradigm of atypical intersections was first developed by Zilber [114], Bombieri-Masser-Zannier [27], and Pink [97]. One of the most spectacular recent results in this area is the proof of the André-Oort conjecture [94], which is the culmination of years of work by many researchers.

The key tools underpinning this theorem are o-minimality (a branch of model theory) and **functional transcendence** (FT)—in the level 1 case, see the ICM reports of Pila [92] and Tsimerman [105]. These concepts, along with the Zilber-Pink philosophy, have dramatically expanded our understanding of the Hodge locus, though many questions remain open, and new challenges have emerged. A notable early case of this theorem was explored in my earlier work [23], which dealt with the case where S is a ball quotient, and $\mathbb{V}$ is a VHS designed to control the totally geodesic subvarieties. More recently, I introduced, in collaboration with D. Urbanik [24], a new and more general framework for proving related results. FT has quickly evolved to higher levels of generality, and today, it is best understood in the context of foliated principal bundles, as recently introduced in [26].

In essence, FT asserts the following: Let $P \to S$ be an algebraic principal G-bundle (G a suitable algebraic group) with a principal and equivariant connection ∇. If a subvariety of P intersects non-transversally with an analytic leaf $\mathcal{L} \subset P$, then the projection to S of such an intersection is contained within a **strict weakly-special subvariety** (and thus, it is not Zariski dense). This viewpoint is adopted in my work with Urbanik to tackle more subtle cases of subvarieties of automorphic bundles over period domains, not predicted by Theorem 1.3. This is usually referred to as the **Ax-Schanuel theorem**.

Much work has already been done on atypical intersections and their connection with FT, yet many new avenues are opening. . .

2. History and preliminaries

The focus of this text is the *Hodge locus of a variation of Hodge structures*, a crucial object in Algebraic and Arithmetic Geometry, closely linked to but broader than the Hodge conjecture. I will provide a historical perspective on the Hodge locus, tracing its evolution through the Zilber-Pink viewpoint (borrowed from Number Theory) and its emerging connections to abelian differentials and non-arithmetic complex hyperbolic lattices.

To set the stage, I will first review the historical background and common ground of the three proposed research axes, introducing key concepts such as the Mumford-Tate group and the Zilber-Pink conjecture.

The beginnings: 1882-1920. Max Noether states his famous theorem on curves: if $d \geq 4$, there is a countable union of closed subvarieties of the parameter space $U_{2,d}$ of smooth degree d surfaces outside of which the restriction map $\mathrm{Pic}(\mathbb{P}^3) \to \mathrm{Pic}(X)$ is an isomorphism. (For $d = 2$ or $d = 3$, there is no analogous statement; smooth complex surfaces of these degrees have Picard groups of rank 2 and 7, respectively.)

In the 1920s, Lefschetz proves this using topological methods, marking the debut of the Noether-Lefschetz locus $\mathrm{NL}_d \subset U_{2,d}$ (i.e., the locus of surfaces for which the restriction map is **not** an isomorphism) in the study of surfaces. Hodge theory enters the first proofs of Noether's theorem through Lefschetz's theorem on $(1,1)$-classes[1]: given a compact Kähler manifold X, the first Chern class $c_1 : \mathrm{Pic}(X) \to H^2(X, \mathbb{C})$ surjects onto $H^{1,1} \cap H^2(X, \mathbb{Z}) \subset H^2(X, \mathbb{C})$. Lefschetz's original proof worked on projective surfaces and used normal functions, introduced by Poincaré. A posteriori, this translates the Noether-Lefschetz problem from a cycle-theoretic task into an analytic one. It is also the only case of the Hodge conjecture proved for all Kähler manifolds.

Hodge theory: 1950-1970. Hodge [71] introduced the Hodge decomposition for every compact Kähler manifold. In 1968, Griffiths [66, 61] initiated his study of **variation of Hodge structures**, including the ones enriching the local system associated to a smooth projective

[1]We refer also the the work of Kodaira-Spencer [78] for the first modern proof and remark that, even after the Hodge decomposition is established delicate algebraization issues have to be considered.

family of algebraic varieties (e.g. the primitive part of $H^2(X, \mathbb{Z})$ for the degree d surfaces). Deligne [41] vastly generalized Hodge's results, showing that the cohomology of any complex algebraic variety (not necessarily compact) is functorially endowed with a **mixed** $\mathbb{Z}$-Hodge structure.

Griffiths School: 1983 and Early 90s. Griffiths and his students—Carlson, Green, and Harris—[32] introduced new ideas from infinitesimal variations of Hodge structures. Green and Voisin proved the *Explicit Noether-Lefschetz Theorem*: each irreducible component Y of NL_d satisfies

$$d-3 \leq \operatorname{codim}_{U_{2,d}} Y \leq h^{2,0} = \binom{d-1}{3}.$$

The upper bound on codimension follows easily from the Hodge decomposition: a class $\int_\lambda \in H^2(X, \mathbb{Q})$ is algebraic if and only if it has type $(1,1)$, equivalent to $\int_\lambda \omega = 0$ for all $\omega \in H^0(X, \Omega^2_X)$. Additionally, the family of smooth degree d surfaces containing a line forms a component of NL_d of codimension $d-3$. However, proving that this provides a lower bound is more subtle and depends on delicate algebraic considerations.

In a 1988 paper, Ciliberto, Harris, and Miranda introduced the following key definitions: a component Y of NL_d is **general** if it has codimension $h^{2,0}$, and **exceptional** otherwise (originally called *special*). This general-special dichotomy spurred many further results from numerous contributors. For example, if $d \geq 5$ (if $d = 4$, there are no exceptional components), it is known that the union of the general components of NL_d is Zariski-dense in $U_{2,d}$ [36], and that NL_d is even dense in $U_{2,d}$ for the analytic topology [108, Prop. 5.20]. See [37] for explicit exceptional components.

Arithmetic/Diophantine Geometry: 1969-2000s. The viewpoint presented here traces back to Serre's work on *open image theorems* for abelian varieties and the related work of Mumford [90] and Tate from 1969. They sought to predict the image of the Galois group acting on the ℓ-adic Tate module of an abelian variety A, introducing the notion of the Mumford-Tate group.

Definition 2.1. *The* ***Mumford-Tate group*** *of a* $\mathbb{Q}$*-Hodge structure* V*,* $\mathbf{MT}(V)$*, is the Tannakian group associated to the category* $\langle V \rangle^{\otimes} \subset \mathbb{Q}\mathrm{HS}$*. Equivalently, it is the Zariski closure in* $\mathrm{GL}(V)$ *(considered as a* $\mathbb{Q}$*-algebraic group) of the image of the* $\mathbb{C}^*$*-action defining the Hodge decomposition on* V*.*

(When the Hodge structure is polarized, as it will usually be the case in this text, the Mumford-Tate group can also be described as the stabilizer in $\mathrm{GL}(V)$ of the Hodge classes in arbitrary tensor operations.)

Deligne [41] proved that the image of the Galois group acting on Tate module is a subgroup of $\mathbf{MT} \otimes \mathbb{Q}_\ell$. Pohlmann, in '68, proved the *Mumford–Tate conjecture* for abelian varieties with Complex Multiplication (which asserts that, under the Artin comparison isomorphism, the ℓ-adic Galois monodromy agrees with the $\mathbb{Q}_\ell$-points of Mumford-Tate group). I recall now the most general definition of Complex Multiplication for Hodge structures:

Definition 2.2. *A Hodge structure* V *is* ***CM*** *if* $\mathbf{MT}(V)$ *is commutative.*

In 1989, André [1], and later Oort in 1995 [91], proposed the André-Oort conjecture (AO) for Shimura varieties, partly inspired by non-abelian analogues of the Manin-Mumford conjecture and the distribution of CM points on moduli spaces (see also [3]). Shimura varieties, a special class of varieties that include modular curves and, more generally, moduli spaces parametrizing principally polarized Abelian varieties of a given dimension (possibly with additional prescribed structures), were originally introduced by Shimura in the 1960s in his study of the theory of complex multiplication.

Theorem 2.3 (AO for Shimura varieties, [94, 104])**.** *Let* S *be a Shimura variety. An irreducible subvariety is a sub-Shimura variety if and only if it contains a Zariski-dense set of CM points.*

In 1999, Bombieri, Masser, and Zannier [113] began studying other types of atypical intersections. The first example was curves against algebraic subgroups of multiplicative groups. Independently, Zilber explored similar ideas in the setting of exponential sum equations and the Schanuel conjecture (2002). Pink, motivated by unifying the Mordell-Lang and AO conjectures, proposed his own version of the conjecture. A further important step came with the advent of the Pila-Zannier strategy for Manin-Mumford (see the recent book [93]).

String Theory: 1984-2000s. The interest of string theorists in Calabi-Yau 3-manifolds stems from their connection to conformal field theories (CFTs). Gukov and Vafa [69] posed a question regarding the existence of infinitely many Calabi-Yau manifolds with complex multiplication of a fixed dimension, motivated by the connection to rational conformal field theories (RCFT), introduced earlier by Friedan-Qiu-Shenker. Around the same time, Moore explored the arithmetic-string theory connection, particularly the role of attractor varieties in black hole constructions within IIB string theory. These investigations revealed connections to atypical intersections that are not necessarily of CM type. For more recent investigations, see also [30].

The last decade. Let $\mathbb{V}$ be a polarizable variation of Hodge structures, simply $\mathbb{Z}$-VHS from now on, on an irreducible smooth quasi-projective variety S. At this point, it's useful to formally introduce the **Hodge locus** of $(S, \mathbb{V})$ and the concept of **level**. I start by the Hodge locus:

$$\mathrm{HL}(S, \mathbb{V}^{\otimes}) := \{s \in S(\mathbb{C}) : \mathbf{MT}(\mathbb{V}_s) \subsetneq \mathbf{MT}(\mathbb{V})\}, \tag{2.0.1}$$

where $\mathbf{MT}(\mathbb{V})$ denotes the Mumford–Tate group at a very general point of S, fixed and denoted by 0. Informally, the Hodge locus parametrizes the closed points of S corresponding to Hodge structures with additional *Hodge symmetries*. Cattani, Deligne, and Kaplan [33] proved that the Hodge locus is a countable union of **algebraic** subvarieties of S, known as **special subvarieties** for $(S, \mathbb{V})$. According to Griffiths' theory, $\mathbb{V}$ induces a period map

$$\Psi : S \to \Gamma\backslash D, \quad s \mapsto [\mathbb{V}_s], \tag{2.0.2}$$

where D is a subset of the flag variety of chains of subspaces of fixed dimensions of $\mathbb{V}_0 \otimes \mathbb{C}$ (the extra conditions are imposed by the polarization). It is natural a homogeneous space under G, the real points of $\mathbf{MT}(\mathbb{V})$. In particular $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ can be interpreted as the preimage of Mumford-Tate subdomains of $\Gamma\backslash D$. Indeed to each subvariety $Y \subset S$, there is associated a sub-period domain $\Gamma_Y\backslash D_Y \subset \Gamma\backslash D$ such that $Y \subset \Psi^{-1}(\Gamma_Y\backslash D_Y)$ (the right hand side is in fact known to be an *algebraic* subvariety, as we will see later), and Y is special when it is a component of $\Psi^{-1}(\Gamma_Y\backslash D_Y)$.

Intuitively, the **level** measures the minimum number of Lie bracket iterations

$$[\mathfrak{g}^{-1}, [\mathfrak{g}^{-1}, \dots [\mathfrak{g}^{-1}, \mathfrak{g}^{-1}] \dots]$$

one needs before obtaining zero, where $\mathfrak{g}^{-1} \subset \mathrm{Lie}(G \otimes \mathbb{C})$ represents the Griffiths sub-space of the holomorphic tangent space of D at the fixed base point 0. More will be discussed in Section 5.1, but we remark here that:

- If D is Hermitian, i.e. $\Gamma\backslash D$ is a connected Shimura varieties, it is well known that $[\mathfrak{g}^{-1}, \mathfrak{g}^{-1}] = 0$, so the level is 1.
- The natural VHS $\mathbb{V}$ on $U_{2,d}$ has level 1 if $d = 4$, and 2 if $d \geq 5$. (There are no periods for $d \leq 3$).
- The natural VHS $\mathbb{V}$ on $U_{n,d}$ has level ≥ 3 when n *and* d *are big enough*, cf. Corollary 6.16 for the precise statement. ù

From a Hodge-theoretic perspective, Hodge structures of even weight can look more natural, as algebraic classes always lie in even dimensional cohomology. For instance, the primitive cohomology of a degree d surface ($d \geq 4$) is concentrated in H^2, and the **vectorial** Hodge locus, i.e. the locus where there are some extra Hodge vectors, is precisely the Noether-Lefschetz locus. The Hodge classes in tensorial constructions of such H^2 are geometrically more subtle and play a crucial role here. This is a first evidence of the importance of the so called **Tannakian**

viewpoint and is concretely reflected in the study of arbitrary Mumford-Tate subdomains of period domains.

Until recent progress, much less was known about levels ≥ 3. In 2017, Klingler [75] began exploring the Zilber-Pink conjecture for arbitrary variations of Hodge structures. A.J. de Jong had formulated similar conjectures in a personal note from 2004. In [21], Klingler, Ullmo, and I proposed a stronger conjecture, now widely regarded as the correct version, along with another conjecture describing the typical Hodge locus. This also marked the beginning of the idea that 'atypical intersections govern typical ones."

Definition 2.4. *A special subvariety* $Y = \Psi^{-1}(\Gamma_Y \backslash D_Y)^0$ *is **atypical** if* $\operatorname{codim}_{\Psi(S)}(\Psi(Y)) < \operatorname{codim}_D(D_Y)$*, and **typical** otherwise. The union of the atypical (resp. typical) special subvarieties is denoted by* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}}$ *(resp.* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$*).*

(Possibly after replacing the base by a finite covering, Griffiths proved that the period map is a proper holomorphic map. In fact, recently a stronger conjecture of Griffiths on the quasi-projectivity of images of period maps has been established [10]. Therefore, it makes sense to speak of the dimensions of $\Psi(S)$ and $\Psi(Y)$.)

Let $(S, \mathbb{V})$ be a pair where S is a smooth quasi-projective complex variety and $\mathbb{V}$ an integral and polarised VHS on S with associated period map $\Psi : S \to \Gamma \backslash D$.

Conjecture 2.5 (Zilber–Pink Conjecture for VHS)**.** *There are only finitely many maximal atypical subvarieties for* $(S, \mathbb{V})$*. In particular the atypical Hodge locus is not Zariski-dense.*

Conjecture 2.6 (Density of the Typical Hodge Locus)**.** *The following are equivalent:*
- *(1)* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ *is not empty;*
- *(2)* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ *is dense in* $S(\mathbb{C})$*;*
- *(3) There exists a sub-period domain* $D' \subset D$ *such that* $\dim \Psi(S) - \operatorname{codim}_D D' \geq 0$.

This was made possible thanks to several developments regarding functional transcendence, culminating in the so-called **Ax-Schanuel** (AS) type theorems: Blázquez-Sanz, Casale, Freitag, and Nagloo [26] (in the very general setting of principal G-bundles), Bakker-Tsimerman [15, 16] (in the period domain first, and then in the period torsor). The case of level 1 was obtained in [87]. A representative statement to keep in mind is the following:

Theorem 2.7 (Special case of AS in the period domain, Bakker-Tsimerman)**.** *Let* $Y \subset D$ *be an algebraic subvariety of codimension* $\geq \dim S$ *and denote by* Y_Γ *its image in* $\Gamma \backslash D$*. Each component of* $\Psi^{-1}(Y_\Gamma)$ *lies in a **strict weakly special subvariety**[2] of* S*.*

Even more recently, a new avenue opened with my work with D. Urbanik, where I've extended the setting of Zilber-Pink to the notable case of $\mathbb{Z}$-variations of **mixed** HS [24] and found applications of our theory to Teichmüller problems. For simplicity I will mainly discuss the pure case, but this generalization is crucial for several applications and it is highly non-trivial.

The future is bright. It is worth mentioning that, in its debut, functional transcendence was tightly related with **o-minimality**. O-minimality has been applied with success to make progress on questions in Hodge theory (Griffiths conjecture, definable period maps [10]), and has recently had its own explosion of results guided by Binyamini (sharply o-minimal sets, the resolution of Wilkie's conjecture, see e.g. [25]). Another remarkable example is the Linear Shafarevich Conjecture in the quasi-projective case, by Bakker, Brunebarbe, and Tsimerman [12], which further uses the definable setting and among other things, the Ax-Schanuel theorem for abelian varieties.

The applications of the Zilber–Pink conjecture viewpoint have quickly permeated various areas and captured the attention of many mathematicians. To mention a few examples:

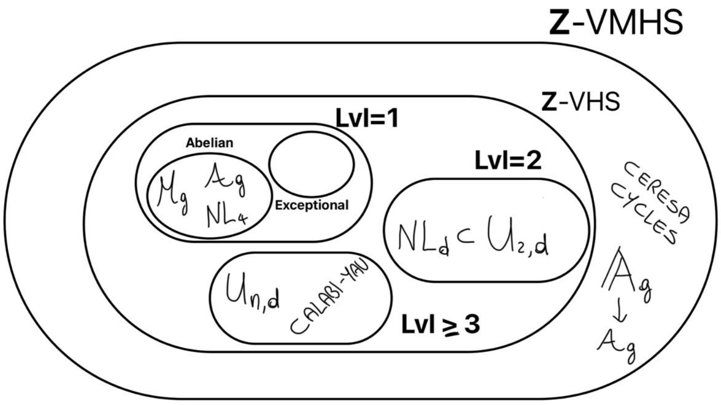
- Study of the torsion locus of the Ceresa normal function by Gao and Zhang [59], as well as related work by Kerr-Tayou [73] and Hain [70], has offered further advancements.

[2]They are generalization of the special subvarieties and defined as follows: maximal subvarieties with given algebraic monodromy group. They will be discussed in detail in Section 3.2.

- Description of the maximal compact subvarieties of Siegel modular varieties [68] — an area that has seen incremental progress over many years.
- A Diophantine direction, driven by the Lawrence and Venkatesh method [82], with later contributions by Lawrence-Sawin; Javanpeykar-Krämer-Lehn-Maculan; and Klingler, Ullmo and myself [19].
- A Conjecture of Matsushita on Lagrangian fibrations of hyperkähler manifolds [8, 110].
- The work of Dimitrov–Gao–Habegger and Kühne on the uniform Mordell-Lang conjecture [57].
- Lam and Tripathy [80, Thm. 1.1.3.] provide, conditional on the Zilber-Pink conjecture, a family of counterexamples to the Attractor Conjecture in all suitably high, odd dimensions.
- Finally, in [54, Thm. 5.3.1.], the authors use a special case of Zilber-Pink recently established by Richard-Yafaev [98] to construct infinitely many rank-two local systems of geometric origin which are not pullbacks of hypergeometric local systems.

We remark that Zilber-Pink offers a common foundation for considering different questions, as well as a guide for what to expect. However, the key ideas needed to progress are new each time and require different inputs from various theories (Algebraic and Arithmetic Geometry, Ergodic Theory, Model Theory, etc.).

To conclude this overview, the following diagram might serve as a road map for Hodge theory in families:



Part 2. Hodge theory in families

3. The Hodge locus

To describe the general framework we will work with, we first recall the concept of variations of **mixed** Hodge structures refining the above definitions.

3.1. **Preliminary definitions.** A mixed $\mathbb{Z}$-Hodge structure ($\mathbb{Z}$MHS) is a triple $(V, W_\bullet, F^\bullet)$ consisting of a (torsion-free) finitely generated $\mathbb{Z}$-module V, a finite ascending filtration $W_\bullet$ of $V_\mathbb{Q} := V \otimes_\mathbb{Z} \mathbb{Q}$ (called the *weight filtration*) and a finite decreasing filtration $F^\bullet$ of $V_\mathbb{C}$ (called the *Hodge filtration*) such that for each $n \in \mathbb{Z}$,

$$(\mathrm{Gr}_n^W V_\mathbb{Q}, \mathrm{Gr}_n^W F^\bullet)$$

is a pure $\mathbb{Q}$-Hodge structure of weight n. A *graded polarization* of a mixed $\mathbb{Z}$-Hodge structure is the datum of a polarization on the pure $\mathbb{Q}$-Hodge structure $\mathrm{Gr}^W V = \bigoplus_n \mathrm{Gr}_n^W V$. Analogously one defines $\mathbb{Q}$MHS.

The following is [96, 1.4, 1.5] and explains under which conditions a homomorphism $h : \mathbb{S}_\mathbb{C} \to \mathbf{GL}(V_\mathbb{C})$ gives a MHS (where $\mathbb{S}$ denotes the Deligne torus: the Weil restriction from $\mathbb{C}$ to $\mathbb{R}$ of $\mathbb{C}^*$):

Lemma 3.1. *Let V be a finite dimensional $\mathbb{Q}$-vector space. A morphism $h : \mathbb{S}_\mathbb{C} \to \mathbf{GL}(V_\mathbb{C})$ defines a $\mathbb{Q}$MHS on V if and only if there exists a connected $\mathbb{Q}$-algebraic subgroup $\mathbf{G} \subset \mathbf{GL}(V)$ such that h factors through $\mathbf{G}_\mathbb{C}$ and satisfies the following conditions (where we write W_{-1} for the unipotent radical of $\mathbf{G}$):*

- *The composition morphism $\mathbb{S}_\mathbb{C} \to \mathbf{G}_\mathbb{C} \to (\mathbf{G}/W_{-1})_\mathbb{C}$ is defined over the reals;*
- *Precomposing $\mathbb{S} \to (\mathbf{G}/W_{-1})_\mathbb{R}$ with the weight homomorphism $\mathbb{G}_{m,\mathbb{R}} \to \mathbb{S}$ we obtain a cocharacter of the center of $(\mathbf{G}/W_{-1})_\mathbb{R}$ defined over $\mathbb{Q}$;*
- *The weight filtration on $\mathrm{Lie}\,\mathbf{G}$ defined by $\mathrm{Ad} \circ h$ satisfies $W_0(\mathrm{Lie}\,\mathbf{G}) = \mathrm{Lie}\,\mathbf{G}$ and $W_{-1}(\mathrm{Lie}\,\mathbf{G}) = \mathrm{Lie}\,W_{-1}$.*

Definition 3.2. *A* mixed Hodge datum *is a pair $(\mathbf{G}, D)$ where $\mathbf{G}$ is a connected linear algebraic $\mathbb{Q}$-group and D is a $\mathbf{G}(\mathbb{R})W_{-1}(\mathbb{C})$-conjugacy class of homomorphisms $\mathbb{S} \to \mathbf{G}(\mathbb{C})$, satisfying the conditions of Lemma 3.1.*

Definition 3.3. *Let S be a smooth quasi-projective complex variety and $\mathcal{O}_S$ its sheaf of holomorphic functions, and $R \subset \mathbb{C}$ a ring. A* variation of mixed R-Hodge structures *(RVMHS) over S is a triple $(\mathbb{V}, \mathbf{W}_\bullet, F^\bullet)$, where:*

(1) $\mathbb{V}$ is a locally constant R_S-module on S,
(2) $\mathbf{W}_\bullet$ is a finite increasing filtration (called the weight filtration) of the K-local system $\mathbb{V}_K$ by K-local sub-systems,
(3) $F^\bullet$ is a finite descending filtration (called the Hodge filtration) of the holomorphic vector bundle $\mathcal{V} := \mathbb{V} \otimes_{R_S} \mathcal{O}_S$ by holomorphic subbundles,

such that

(a) for each $s \in S$, the triple $(\mathbb{V}_s, (\mathbf{W}_\bullet)_s, F_s^\bullet)$ is a mixed R-Hodge structure.
(b) the flat connection $\nabla : \mathcal{V} \longrightarrow \mathcal{V} \otimes \Omega_S^1$ whose sheaf of horizontal sections is $\mathbb{V}_\mathbb{C}$ satisfies the Griffiths' transversality condition

$$\nabla F^\bullet \subset \Omega_S^1 \otimes F^{\bullet-1}.$$

A graded polarization Ψ for $(\mathbb{V}, \mathbf{W}_\bullet, F^\bullet)$ is a sequence

$$\Psi_k : \mathrm{Gr}_k^W(\mathbb{V}_K) \times \mathrm{Gr}_k^W(\mathbb{V}_K) \longrightarrow K(-k)_S$$

of ∇-flat bilinear forms inducing graded polarisations $\Psi_{k,s}$ on the mixed R-Hodge structure $(\mathbb{V}_s, (\mathbf{W}_\bullet)_s, F_s^\bullet)$ for all $s \in S$.

An important admissibility condition for variations of mixed Hodge structures was introduced by Kashiwara in [72] (see also, for example, [58, Def. 3.3]). From now on we denote by $\mathbb{Z}$VMHS the category of graded-polarizable and admissible VMHS over S. When $\mathbb{Z}$ is replaced by $\mathbb{Q}$, this is a Tannakian category, as discussed for example in [75, §2.4], and references therein. (The fact that the category of admissible $\mathbb{Q}$VMHS is abelian due to Kashiwara, see [72, §4.5 and Prop. 5.2.6] for the relevant statements.)

Definition 3.4. *For every $s \in S$, the* Mumford-Tate group $\mathbf{G}_s(\mathbb{V})$ *of the Hodge structure $\mathbb{V}_s$ is the Tannakian group of $\langle \mathbb{V}_s \rangle^{\otimes}$ of $\mathbb{Q}MHS$ (here the exact faithful $\mathbb{Q}$-linear tensor functor* $\mathrm{For} : \mathbb{Q}MHS \to \mathrm{Vec}_{\mathbb{Q}}$ *is simply the forgetful functor).*

It is a connected $\mathbb{Q}$-algebraic group, and

- reductive if $\mathbb{V}_s$ (or $\mathbb{V}$) is pure; in general
- it is an extension of $\mathbf{G}_s(\mathrm{Gr}^W V)$ by a unipotent group (here W denotes the weight filtration on $\mathbb{V}$).

We note that Gr^W can be viewed as an exact functor [41, Them. 2.3.5(iv)] from $\mathbb{Q}$VMHS to $\mathbb{Q}$VHS. We have also a notion of *Mumford–Tate group* of the $\mathbb{Q}$VMHS $\mathbb{V}$, simply as the Tannakian group of $\mathbb{V}$. Here we pick the fiber functor associated to some base point $s \in S$, that is the functor

$$\omega_s : \mathbb{Q}\,\mathrm{VMHS} \to \mathbb{Q}\,\mathrm{MHS} \to \mathrm{Vec}_{\mathbb{Q}}, \qquad \mathbb{V} \mapsto \mathbb{V}_s \mapsto \mathrm{For}(\mathbb{V}_s).$$

We denote the Mumford–Tate group by $\mathbf{MT}(\mathbb{V}, \omega_s)$. Thanks to [42, II, Thm. 3.2], the resulting group, by the general theory, will not depend on such choice of a base point s (that is there's a *unique natural* isomorphism $\mathbf{MT}(\mathbb{V}, \omega_s) \cong \mathbf{MT}(\mathbb{V}, \omega_{s'})$). See also [2] for an equivalent point of view. By the general Tannakian formalism we have a canonical inclusion $\mathbf{MT}(\mathbb{V}_s) \subset \mathbf{MT}(\mathbb{V}, \omega_s)$, at every $s \in S$.

We define the following subset of $S^{\mathrm{an}} = S(\mathbb{C})$.

$$\mathrm{HL}(S, \mathbb{V}^{\otimes}) := \{s \in S^{\mathrm{an}} : \mathbf{MT}(\mathbb{V}_s) \neq \mathbf{MT}(\mathbb{V}, \omega_s)\}. \tag{3.1.1}$$

It is naturally a countable union of complex subspaces, and in fact Cattani, Deligne and Kaplan [33, Thm. 1.1] proved the following in the pure case, and Patrick, Pearlstein, and Schnell in the mixed case [29]. We also refer to [13, Thm. 1.6] (see also [14]) for an alternative proof, using o-minimality, and also to [9], for the mixed case.

Theorem 3.5. *Let S be a smooth connected complex quasi-projective algebraic variety and $\mathbb{V}$ be a $\mathbb{Z}$VMHS over S. Then* $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ *is a countable union of closed irreducible algebraic subvarieties of S: the strict maximal special subvarieties of S for $\mathbb{V}$.*

In general, the notion of *special subvariety* is defined as follows:

Definition 3.6. *Let $(S, \mathbb{V})$ be a $\mathbb{Z}$VMHS. A* special subvariety $Y \subset S$ *is a closed irreducible complex algebraic subvariety which is maximal for the Mumford-Tate group $\mathbf{G}_Y$ associated to $(Y, \mathbb{V}_{|Y})$.*

For us, from now on, all special subvarieties, so in particular all components of the Hodge locus, will be equipped with their *reduced* algebraic structure.

Then the datum of such a pair $(S, \mathbb{V})$ is the same datum as a *(mixed) period map*. Let $(\mathbf{G}, D)$ be a (connected, mixed) Hodge datum as in [75, Def. 3.3] associated to $(S, \mathbb{V})$, and $\Gamma\backslash D$ the associated mixed Mumford–Tate domain (called also Hodge variety by some authors). (Here and in the sequel, we always assume that Γ is a torsion free finite index subgroup of $\mathbf{G}(\mathbb{Z})$). We will be mainly concerned with period maps:

$$\Psi : S^{\mathrm{an}} \to \Gamma\backslash D,$$

that are locally liftable, holomorphic and horizontal maps (see [75, §3.5]).

Definition 3.7.

(1) A subvariety Z of S is said of positive period dimension for $\mathbb{V}$ *if $\Psi(Z^{\mathrm{an}})$ has positive dimension (i.e. is not a point).*

(2) *The* Hodge locus of positive period dimension $\mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{pos}}$ *is the union of the special subvarieties of S for $\mathbb{V}$ of positive period dimension.*

Remark 3.8. In the pure case the Mumford-Tate domain D decomposes as a product $D_1 \times \cdots \times D_k$, according to the decomposition of the adjoint group $\mathbf{G}^{\mathrm{ad}}$ into a product $\mathbf{G}_1 \times \cdots \times \mathbf{G}_k$ of simple factors (notice that some factors $\mathbf{G}_i$ may be $\mathbb{R}$-anisotropic). A subvariety Z of positive period dimension of S is said of positive period dimension *factorwise* if moreover the projection of $\Phi(Z^{\mathrm{an}})$ on each factor $\Gamma_i\backslash D_i$, $1 \leq i \leq r$, has positive dimension.

Definition 3.9. *A special subvariety $Y \subset S$, with Hodge datum $(\mathbf{G}_Y, D_Y)$ (associated to $\mathbb{V}_{|Y^{\mathrm{sm}}}$) is* atypical *if*

$$\mathrm{codim}_{\Gamma\backslash D}\,\Psi(Y^{\mathrm{an}}) < \mathrm{codim}_{\Gamma\backslash D}\,\Psi(S^{\mathrm{an}}) + \mathrm{codim}_{\Gamma\backslash D}\,\Gamma_Y\backslash D_Y. \tag{3.1.2}$$

Otherwise, it is said to be typical*.*

Here Γ_Y simply denotes $\Gamma \cap \mathbf{G}_Y(\mathbb{Q})$. Moreover the map $\Gamma_Y\backslash D_Y \to \Gamma\backslash D$ is quasi-finite, so the formal "codimension" makes sense.

Remark 3.10. One can rewrite the atypical condition appearing in Definition 3.9 in terms of dimension, rather than codimension:

$$\dim \Psi(S^{\mathrm{an}}) - \dim \Psi(Y^{\mathrm{an}}) < \dim D - \dim D_Y.$$

Definition 3.11. *The* atypical Hodge locus $\mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{atyp}} \subset \mathrm{HL}(S,\mathbb{V}^{\otimes})$ *(resp. the* typical Hodge locus $\mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{typ}} \subset \mathrm{HL}(S,\mathbb{V}^{\otimes})$*) is the union of the atypical (resp. strict typical) special subvarieties of S for $\mathbb{V}$.*

We expect the Hodge locus

$$\mathrm{HL}(S,\mathbb{V}^{\otimes}) = \mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{atyp}} \cup \mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{typ}}$$

to satisfy the Conjecture 2.5, Conjecture 2.6 (cf. [21] as well as [24, Sec. 4] for the mixed case). Formally there is no difference in the statements for the mixed and the pure case.

3.2. **Weakly special subvarieties in Hodge Theory.** Essentially from Theorem 3.5, we observe that there are countably many special subvarieties for $(S,\mathbb{V})$. It is also easy to observe that an intersection of special subvarieties is a union of special subvarieties. In this section we compare such notions and proprieties with the more general concept of *weakly special subvarieties*.

Much of what we will need is a consequence of the following [2, Thm. 1] (see also [41]). Let $\mathbb{V}$ be a $\mathbb{Z}$VMHS on S. For any closed point $s \in S$, let $\mathbf{G}_{S,s}$ be the Mumford-Tate group of $(S,\mathbb{V})$ at ω_s, and $\mathbf{H}_{S,s}$ the algebraic monodromy at s of the $\mathbb{Q}$-local system $\mathbb{V}_{\mathbb{Q}}$ (that is the connected component of the identity of the Zariski closure of the topological monodromy).

Theorem 3.12 (André-Deligne Monodromy Theorem)**.** *The monodromy group $\mathbf{H}_{S,s}$ is a normal subgroup of the derived subgroup of $\mathbf{G}_{S,s}$.*

We notice here that, in the general mixed case, the derived subgroup of $\mathbf{G}_{S,s}$ is not a semisimple group.

Definition 3.13. *A subset D' of the classifying space D is called* a weak Mumford-Tate subdomain *if there exists an element $t \in D'$ and a normal subgroup $\mathbf{N} \subset \mathbf{MT}(t)$ such that $D' = \mathbf{N}(\mathbb{R})^{+}\mathbf{N}^{u}(\mathbb{C}) \cdot t$, where $\mathbf{N}^{u}$ denotes the unipotent radical of $\mathbf{N}$.*

A *weakly special submanifold* of the Hodge manifold $\Gamma\backslash D$ is an irreducible component of the image in $\Gamma\backslash D$ of a weak Mumford-Tate subdomain. The datum $((\mathbf{H}, D_H), t)$ is called a *weak Hodge subdatum* of $(\mathbf{G}, D)$. Given $(S,\mathbb{V})$ with associated period map $\Psi : S \to \Gamma\backslash D$, every weakly special subvariety of $(S,\mathbb{V})$ can be described as an irreducible component of the preimage along Ψ of a weakly special submanifold of the Hodge variety $\Gamma\backslash D$; see [16, Cor. 2.9]. With this new description of weakly special subvarieties in terms of the period domain $\Gamma\backslash D$, we obtain the following:

Lemma 3.14. *A special subvariety Y for $(S, \mathbb{V})$ is weakly special. The property of being weakly special is closed under taking intersections and passing to irreducible components.*

(In particular it makes sense to consider special and weakly special closures: the special (resp. weakly special) closure of a variety $Y \subset S$ is the intersection of all special (resp. weakly special) subvarieties in which it is contained.)
The closure of special and weakly special subvarieties under intersection gives natural notions of special and weakly special closures: the special (resp. weakly special) closure of a variety $Y \subset S$ is the intersection of all special (resp. weakly special) subvarieties in which it is contained.

For our exposition we now follow [11, §3.5, and §6]. Let V be a finite dimensional real vector space equipped with an increasing filtration $W_\bullet$ and a collection of non-degenerate bilinear forms $q_k : \mathrm{Gr}_k^W \otimes \mathrm{Gr}_k^W \to \mathbb{R}$ that are $(-1)^k$-symmetric. Fix a partition of the dimension of V into non-negative integers $h^{p,q}$ satisfying Hodge symmetry. For any integer k we denote by $\check{\Omega}_k$ the complex projective algebraic variety parametrizing the $(q_k)_\mathbb{C}$-isotropic filtrations $\{F_k'^p\}$ on $\mathrm{Gr}_k^W V_\mathbb{C}$ with $\dim F_k'^p = \sum_{r \geq p} h^{r,k-r}$. The complex group $\mathrm{Aut}(q_k)$ acts transitively by algebraic automorphisms on $\check{\Omega}_k$.

Let $(V, W_\bullet, F^\bullet, q)$ be a mixed $\mathbb{Z}$-Hodge structure as above, polarized by q. We denote by D the corresponding *mixed period domain*: the set of decreasing filtrations $F'^\bullet$ of $V_\mathbb{C}$ such that $(V_\mathbb{R}, W_\bullet, F'^\bullet, q)$ is a real mixed Hodge structure graded-polarized by q and such that

$$\dim_\mathbb{C}(F'^p \mathrm{Gr}_{p+q}^W \cap \overline{F'}^q \mathrm{Gr}_{p+q}^W) = \dim F^p \mathrm{Gr}_{p+q}^W \cap \overline{F}^q \mathrm{Gr}_{p+q}^W = h^{p,q}.$$

By definition, D is a semi-algebraic (in the sense of real algebraic geometry) open subset of the smooth projective complex variety $\check{D}$ that parametrizes the decreasing filtrations of $V_\mathbb{C}$ by complex vector subspaces such that the filtration induced on the graded pieces $\mathrm{Gr}_k^W V_\mathbb{C}$ is inside $\check{\Omega}_k$ for each k. In particular $\check{D}$ maps to the product of the $\check{\Omega}_k$.

Given a $\mathbb{Z}$VMHS $\mathbb{V}$ on S, we have associated a mixed Hodge datum $(\mathbf{G}_S, D_S)$ as in Definition 3.2 (we always assume that $\mathbf{G}_S$ is the generic Mumford-Tate group of $\mathbb{V}$). In this case, $\mathbf{G}_S$ comes with a faithful linear representation $\rho : \mathbf{G}_S \to GL(V)$, and thanks to this representation, we obtain a morphism from D_S to the appropriate classifying space $\check{D}$ introduced above.

Notation. We denote by $\check{D}_S$ the Zariski closure of D_S in $\check{D}$ (it comes with a transitive action of $\mathbf{G}(\mathbb{C})$ and by $\check{D}_S^0 \subset \check{D}_S$ the monodromy orbit of some base point $h_0 \in D_S$, and write $D_S^0 = \check{D}_S^0 \cap D_S$. In particular to a $\mathbb{Z}$VMHS $\mathbb{V}$ on S we obtain the complete *Hodge-monodromy datum*:

$$(\mathbf{H}_S \subset \mathbf{G}_S, D_S^0 \subset D_S)$$

For geometric Zilber-Pink we will work with a finer notion of atypicality, that we now recall. This corresponds to what, in [21, Def. 5.3] is called *monodromically atypical*. See also Lem. 5.6, Ex 5.7, and Rmk. 5.7 in *op. cit.* for a more detailed discussion, in the pure case, of such a definition and comparisons with other variants. (The link between the two notions ultimately comes from Theorem 3.12.)

Definition 3.15. *A weakly special subvariety Y for $(S, \mathbb{V})$ is called* monodromically atypical *if*

$$\dim \Psi(S^{\mathrm{an}}) - \dim \Psi(Y^{\mathrm{an}}) < \dim \check{D}_S^0 - \dim \check{D}_Y^0. \tag{3.2.1}$$

Suppose Ψ is quasi-finite, so every point is a weakly special subvariety. Then unless Ψ is dominant, each point is in fact a monodromically atypical weakly special subvariety.

We conclude with an important result that will be applied many times, cf. [24, Prop. 4.18].

Proposition 3.16. *Let $(S, \mathbb{V})$ be a $\mathbb{Z}$VMHS. Then there exists a countable collection $\{h_i : C_i \to B_i\}_{i=1}^\infty$ of families of irreducible algebraic varieties of S, all of whose fibres are weakly special, and such that every weakly special subvarieties arises as a fibre in some such family. Moreover, the maps $C_i \to S$ may be assumed quasi-finite.*

3.3. **Main results: atypical.** Let $(S, \mathbb{V})$ be a $\mathbb{Z}$VMHS, with associated Hodge-monodromy datum $(\mathbf{H}_S \subset \mathbf{G}_S, D_S^0 \subset D_S)$. The following is usually referred to as the *Mixed geometric Zilber-Pink conjecture.*

Theorem 3.17 (Baldi-Urbanik [24, Thm. 7.1])**.** *There is a finite set $\Sigma = \Sigma_{(S,\mathbb{V})}$ of triples $(\mathbf{H}, D_H, \mathbf{N})$, where $(\mathbf{H}, D_H)$ is some sub-Hodge datum of the generic Hodge datum $(\mathbf{G}_S, D_S)$, $\mathbf{N}$ is a normal subgroup of $\mathbf{H}$ whose reductive part is semisimple, and such that the following property holds.*

For each monodromically atypical maximal (among all monodromically atypical subvarieties) $Y \subset S$ there is some $(\mathbf{H}, D_H, \mathbf{N}) \in \Sigma$ such that, up to the action of Γ, D_Y^0 is the image of $\mathbf{N}(\mathbb{R})^+\mathbf{N}(\mathbb{C})^u \cdot y$, for some $y \in D_H$.

In the pure case, it becomes:

Corollary 3.18 (Baldi-Klingler-Ullmo [21, Thm. 3.1])**.** *Let $(S, \mathbb{V})$ be a pure $\mathbb{Z}$VHS, with generic Hodge datum $(\mathbf{G}, D)$. Let Z be an irreducible component of the Zariski closure of* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{ws,w-atyp}}$ *in S. Then:*

(a) Either Z is a maximal monodromically atypical special subvariety;
(b) Or the adjoint Mumford-Tate group $\mathbf{G}_Z^{\text{ad}}$ decomposes as a non-trivial product $\mathbf{H}_Z^{\text{ad}} \times \mathbf{L}_Z$; Z contains a Zariski-dense set of fibers of $\Psi_{\mathbf{L}_Z}$ which are monodromically atypical weakly special subvarieties of S for Ψ, where (possibly up to an étale covering)

$$\Psi_{|Z^{\text{an}}} = (\Psi_{\mathbf{H}_Z}, \Psi_{\mathbf{L}_Z}) : Z^{\text{an}} \to \Gamma_{\mathbf{G}_Z} \backslash D_{G_Z} = \Gamma_{\mathbf{H}_Z} \backslash D_{H_Z} \times \Gamma_{\mathbf{L}_Z} \backslash D_{L_Z} \subset \Gamma \backslash D \ \ ;$$

and Z is Hodge generic in a special subvariety $\Psi^{-1}(\Gamma_{\mathbf{G}_Z} \backslash D_{G_Z})^0$ of S for Ψ which is monodromically typical (and therefore typical).

3.4. **Main results: typical.** We prove that $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ is algebraic in most cases. A simple measure of the complexity of $\mathbb{V}$ is its level: roughly, the length of the Hodge filtration on the holomorphic tangent space of D. See Definition 5.5 for the precise definition in terms of the algebraic monodromy group of $\mathbb{V}$. While special subvarieties usually abound for $\mathbb{Z}$VHSs of level one (e.g. families of abelian varieties or families of K3 surfaces) and for some $\mathbb{Z}$VHS of level two (e.g Green's famous example of the Noether-Lefschetz locus for degree d $(d > 3)$ surfaces in $\mathbb{P}^3_{\mathbb{C}}$, see [108, Proposition 5.20]), we show:

Theorem 3.19 (Baldi-Klingler-Ullmo [21, Thm. 1.5])**.** *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. If $\mathbb{V}$ is of level at least* 3 *then* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ *is a finite union of maximal atypical special subvarieties (hence is algebraic). In particular, if moreover $\mathbf{G}^{\text{ad}}$ is simple, then* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{pos}}$ *is algebraic in S.*

Theorem 3.20 (Baldi-Klingler-Ullmo [21, Thm. 3.9])**.** *If the typical Hodge locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{pos,typ}}$ *is nonempty then* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{pos,typ}}$ *is analytically dense in S.*

Hodge theory actually gives a simple combinatorial criterion to decide whether $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{typ}}$ is empty or not. Indeed see the recent works of Eterovic-Scanlon, Khelifa-Urbanik [49, 74]. They defined:

Definition 3.21. *A strict Hodge sub-datum $(\mathbf{H}, D_H) \subset (\mathbf{G}, D_G)$ is said* $\mathbb{V}$-admissible *if*

$$\dim \Psi(S^{\text{an}}) + \dim D_M \geq \dim D.$$

Theorem 3.22 (Eterovic-Scanlon, Khelifa-Urbanik)**.** *If $(\mathbf{H}, D_H) \subset (\mathbf{G}, D_G)$ is a $\mathbb{V}$-admissible Hodge sub-datum, then*

$$\mathrm{HL}(S, \mathbb{V}^{\otimes}, \mathbf{H}) := \{s \in S : \exists g \in \mathbf{G}(\mathbb{Q})^+, \mathbf{MT}(\mathbb{V}_s) \subset g\mathbf{H}g^{-1}\}$$

is dense in S^{an}.

We refer also to [103] for results regarding the *equidistribution* of the typical Hodge locus.

Definition 3.23. *Let $\mathbb{V}_1, \mathbb{V}_2 \in \mathbb{Z}VHS/S$. We say that $\mathbb{V}_1$ and $\mathbb{V}_2$ are* isogenous *if there is an equivalence of tensor categories $\langle \mathbb{V}_{1,\mathbb{Q}} \rangle^{\otimes} \cong \langle \mathbb{V}_{2,\mathbb{Q}} \rangle^{\otimes}$, where $\langle \mathbb{V}_{i,\mathbb{Q}} \rangle^{\otimes}$ denotes the smallest Tannakian subcategory of $\mathbb{Q}VHS$'s containing $\mathbb{V}_{i,\mathbb{Q}}$.*

Of course the Tannakian categories $\langle \mathbb{V}_{i,\mathbb{Q}} \rangle^{\otimes}$ appearing above are equivalent (as tensor categories) to the category of finite dimensional representations of their generic Mumford-Tate group. (The equivalence is realised by the functor of $\otimes$-automorphisms of the fiber functor). It can happen that two VHS have isomorphic Mumford-Tate groups, but the isomorphism does not induce an equivalence of tenor categories. cf. Def. 1.10 and Thm. 2.11 in the article of Deligne and Milne [42].

Remark 3.24. If two complex principally polarized abelian varieties A, B that are isogenous in the usual sense (either via a polarized or an unpolarized isogeny), then the Hodge structures $H^1(A,\mathbb{Z}), H^1(B,\mathbb{Z})$ are also isogenous in the sense of Definition 3.23 (here we take as base S the spectrum of $\mathbb{C}$). However the converse is in general not true, for example $H^1(A,\mathbb{Z})$ is isogenous to $H^1(A \times A,\mathbb{Z})$. However, for two principally polarized g-dimensional abelian varieties with Mumford–Tate group $\mathbf{GSp}_{2g}$ the two notions agree, since this is the case when the Mumford–Tate group is as big as possible.

Theorem 3.25 (Baldi-Miller-Stover-Ullmo [22, Thm. 1.18])**.** *Let S be a smooth quasi-projective variety and $\mathbb{V}_1, \mathbb{V}_2$ two pure polarized $\mathbb{Z}VHS$s on S. Assume that the generic Mumford–Tate groups of $\mathbb{V}_1$ and $\mathbb{V}_2$ are $\mathbb{Q}$-simple. If*

$$\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}} = \mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset,$$

then $\mathbb{V}_1$ is isogenous to $\mathbb{V}_2$. As a consequence, $\mathrm{HL}(S, \mathbb{V}_1^{\otimes}) = \mathrm{HL}(S, \mathbb{V}_2^{\otimes})$.

4. Atypical intersections–proofs

In this section we briefly explain the setting and the results of [24]. It will be the starting point for all later considerations.

4.1. **Riemann-Hilbert correspondence.** Given a smooth irreducible algebraic variety S and complex local system $\mathbb{V} \to S$, there is the *Riemann-Hilbert correspondence* functor

$$\begin{aligned} \tau : \mathrm{LocSys}(S^{\mathrm{an}}) &\to \mathrm{MIC}_{\mathrm{reg}}(S) \\ \mathbb{V} &\mapsto \tau(\mathbb{V}) = (\mathcal{H}, \nabla). \end{aligned}$$

Here $\mathrm{MIC}_{\mathrm{reg}}$ denotes the category of locally free modules $\mathcal{H}$ on S with regular-singular integrable connection ∇. The functor has the property that the space of ∇^{an}-flat sections of the locally free sheaf $\mathcal{H}^{\mathrm{an}}$ may be identified with $\mathbb{V}$ up to natural isomorphism. Moreover, the analytic vector bundle $\mathbb{V} \otimes_{\mathbb{C}} \mathcal{O}^{\mathrm{an}}$ with its natural flat connection has a unique algebraic structure $(\mathcal{H}, \nabla)$ with regular singularities. The functor τ is an equivalence of (Tannakian) categories [40].

Given a fixed such $(\mathcal{H}, \nabla)$, the tensor powers $\mathcal{H}^{a,b} := \mathcal{H}^{\otimes a} \otimes (\mathcal{H}^*)^{\otimes b}$ also carry natural algebraic connections, where we use $(-)^*$ to denote the dual vector bundle. Correspondingly, write $\mathbb{V}^{a,b} = \mathbb{V}^{\otimes a} \otimes (\mathbb{V}^*)^{\otimes b}$. Now let $\mathbf{H}_S$ be the algebraic monodromy group of $\mathbb{V}$. Using Chevalley's theorem on invariants, we may understand the group $\mathbf{H}_S$ as the identity component of the stabilizer of a one-dimensional subsystem $\mathbb{L} \subset \bigoplus_{a,b} \mathbb{V}^{a,b}$, and hence it admits an algebraic realization $\mathbf{H}_{S,s}$ in each fibre $\mathbf{GL}(\mathcal{H}_s)$ as the identity component of the stabilizer of $\tau(\mathbb{L})_s$.

Associated to this data we may construct a certain algebraic subbundle $P(s_0)$ of the principal left $\mathbf{GL}(\mathcal{H}_{s_0})$-bundle $\mathcal{P}er(\mathcal{H}) := \mathcal{I}so(\mathcal{H}, \mathcal{O}_S \otimes_{\mathbb{C}} \mathcal{H}_{s_0})$; here the fibres $\mathcal{P}er(\mathcal{H})_s$ above points $s \in S$ are identified with the set of isomorphisms $\mathrm{Iso}(\mathcal{H}_s, \mathcal{H}_{s_0})$. We write points of $\mathcal{P}er(\mathcal{H})$ as pairs (β, s), where $\beta : \mathcal{H}_s \to \mathcal{H}_{s_0}$. To construct $P(s_0)$ we define

$$P(s_0) := \{(\beta, s) \in \mathcal{P}er(\mathcal{H}) : \beta(\tau(\mathbb{L})_s) = \tau(\mathbb{L})_{s_0}\} \tag{4.1.1}$$

where $\mathbb{L} \subset \bigoplus_{a,b} \mathbb{V}^{a,b}$ is as above. This defines an algebraic subbundle of $\mathcal{P}er(\mathcal{H})$, and using flat frames we may observe that each $P(s_0)$ is naturally a left torsor for the Zariski closure $\mathbf{H}_S^{\mathrm{full}}$ of the monodromy representation at s_0.

4.2. **Principal bundles preliminaries and Ax-Schanuel after Blázquez-Sanz, Casale, Freitag, and Nagloo.** In this section we first introduce the formalism of H-bundles and principal connections. The reader may refer for example to [26, §2] for more details, although we will try to be as self contained as possible. Throughout we work over the complex numbers.

Let S be an (irreducible) smooth variety, and H a linear algebraic group. Let $\pi : P \to S$ be an H-*principal* bundle, that is H acts on P (on the right, with the action simply denoted by $\cdot$) and such that the action induces an isomorphism

$$P \times H \simeq P \times_S P, \ (p, g) \mapsto (p, p \cdot g).$$

In particular the fibers of π are principal homogeneous spaces, and the choice of a point $p \in P_s := \pi^{-1}(s)$ induces an isomorphism $H \cong P_s$, given by $g \mapsto p \cdot g$. We have the following exact sequence of S-bundles

$$0 \to \ker d\pi \to TP \to TS \times_S P \to 0, \tag{4.2.1}$$

where $\ker d\pi$ is the so called *vertical bundle*. A *connection* ∇ is simply a section of (4.2.1). We say that ∇ is H-*principal* if it is H-equivariant. We additionally say that ∇ is *flat* if it lifts to a morphism of Lie algebras: the equation $\nabla_{[v,w]} = [\nabla_v, \nabla_w]$ holds on the level of maps of $\mathbb{C}$-schemes. From now on all connections are H-principal and flat.

Since ∇ is flat, the space of rational ∇-horizontal vector field is a singular foliation on P. Moreover, if ∇ is an algebraic flat connection we obtain a regular foliation $\mathcal{F}$, cf. [26, §2.1 and 2.2].

Definition 4.1. *An* integral submanifold *of $\mathcal{F}$ is an m-dimensional immersed analytic submanifold $S \subset P$ (not necessarily embedded in P whose tangent space at each point is generated by the values of vector fields in $\mathcal{F}$ (i.e. the image of ∇). Maximal connected integral submanifolds are called* leaves*.*

Frobenius theorem ensures that through any regular point passes a unique leaf. Finally, we notice here that $\mathcal{F}$ has also *vertical* leaves included in the fibers of P at non regular points of ∇, but we will never consider such case. (In particular when speaking of leaves, we always mean *horizontal* leaves, cf. [26, end of §2.2].)

Example 4.2. The component P containing (id_{s_0}, s_0) of the bundle $P(s_0)$ introduced above is a principal $\mathbf{H}_S$-bundle that naturally comes with an algebraic flat connection. Consider the algebraic connection

$$\delta : \mathcal{H}om(\mathcal{H}, \mathcal{O}_S \otimes_{\mathbb{C}} \mathcal{H}_{s_0}) \to \Omega^1_S \otimes \mathcal{H}om(\mathcal{H}, \mathcal{O}_S \otimes_{\mathbb{C}} \mathcal{H}_{s_0}).$$

Using this map we construct a map of algebraic bundles

$$\nabla : TS \times_S \mathcal{P}er(\mathcal{H}) \to T\mathcal{P}er(\mathcal{H}).$$

This map sends a pair $(v, (\beta, s))$ to the differential operator $\nabla(v, (\beta, s))$ at (β, s) which computes the derivative of a function on $\mathcal{P}er(\mathcal{H})$ in the flat direction along v. By using δ to write out an explicit system of algebraic differential equations, one can see ∇ is an algebraic map. Then ∇ restricts to a map $TS \times_S P \to TP$, which gives the desired section of (4.2.1).

From now on we work with groups H that are a semidirect product of a unipotent group and a semisimple one (even if the more general case of *sparse* groups is considered in the literature).

The following is an immediate consequence of the main theorem of Blázquez-Sanz, Casale, Freitag, and Nagloo [26, Thm. 3.6].

Theorem 4.3. *Let S be a smooth algebraic variety and complex local system $\mathbb{V} \to S$. Let (P, ∇) the associated bundle of Example 4.2. Assume that the monodromy group $H = \mathbf{H}_S$ of $\mathbb{V}$ is a semidirect product of a semisimple and a unipotent group. Let V be a subvariety of P, $x \in V$, and let $\mathcal{L} \subset P$ be the leaf through x. Let U be an analytic irreducible component of $V \cap \mathcal{L}$. If*

$$\operatorname{codim}_P U < \operatorname{codim}_P V + \operatorname{codim}_P \mathcal{L},$$

then the projection of U in S is contained in a weakly-special subvariety.

4.3. **Hodge theory intermezzo.** We also recall the Ax-Schanuel Theorem for period maps. It is implied by the more general theorem Theorem 4.3, as we will explain in a moment. The original proof follows a series of papers, see [15, 35, 58, 16].

Let $(S, \mathbb{V})$ be a pair where S is a smooth quasi-projective complex variety and $\mathbb{V}$ a $\mathbb{Z}$VMHS on S with associated period map $\Psi : S \to \Gamma\backslash D$.

Theorem 4.4 (Ax-Schanuel in the period domain)**.** *Let $W \subset S \times \check{D}^0$ be an algebraic subvariety. Let U be an irreducible complex analytic component of $W \cap S \times_{\Gamma\backslash D^0} D^0$ such that*

$$\operatorname{codim}_{S\times\check{D}^0} U < \operatorname{codim}_{S\times\check{D}^0} W + \operatorname{codim}_{S\times\check{D}^0}(S \times_{\Gamma\backslash D^0} D^0) \ .$$

Then the projection of U to S is contained in a strict weakly special subvariety of $(S, \mathbb{V})$.

Notice that $S\times_{\Gamma\backslash D^0} D^0$ is simply the image of the graph of $\tilde{\Psi} : \widetilde{S} \to D^0$ under $\widetilde{S}\times D^0 \to S\times D^0$. Then the intersection $W \cap S \times_{\Gamma\backslash D^0} D^0$ can be identified with the intersection in $S \times D^0$ between W and the image of $\widetilde{S}$ in $S \times D^0$ along the map $(\pi, \tilde{\Psi})$.

4.3.1. *Relation with Period Torsors Associated to* $\mathbb{V}$. The relationship between Theorem 4.4 and Theorem 4.3 is ultimately explained by the torsor P in Example 4.2 associated to the underlying local system of $\mathbb{V}$. See indeed the discussion in [16]. Let $(\mathbf{H}_S, D_S^0)$ be the weak Hodge datum associated to $\mathbb{V}$. We may fix a basepoint $s_0 \in S$, and regard this data as being associated to the fibre $\mathbb{V}_{s_0}$. Then in particular, $\check{D}^0 = \check{D}_S^0$ is a variety parameterizing mixed Hodge data at s_0. Recall that P was constructed as subbundle of the total space of the vector bundle $\mathcal{H}om(\mathcal{H}, \mathcal{O}_S \otimes_{\mathbb{C}} \mathcal{H}_{s_0})$, where $\mathcal{H}$ is the algebraic bundle underlying $\mathbb{V} \otimes_{\mathbb{Z}} \mathcal{O}_{S^{\mathrm{an}}}$. Then we obtain a natural algebraic map $r : P \to \check{D}^0$. Concretely, this map acts as

$$[\eta \in \operatorname{Hom}(\mathcal{H}_s, \mathcal{H}_{s_0})] \mapsto \eta(W_{\bullet,s}, F_s^{\bullet}),$$

where we use that both the Hodge flag and the filtration by weight are given by algebraic subbundles of $\mathcal{H}$. This is a combination of the fact that the weight filtration on $\mathbb{V}$ is given by local subsystems to which we can apply the Riemann-Hilbert functor, together with the corresponding result for the Hodge filtration on the pure graded quotients, which is a consequence of work of Schmid [100]. To check that the map is well-defined (the image lands inside $\check{D}^0$ rather than some larger flag variety), it suffices by $\mathbf{H}_S(\mathbb{C})$-invariance to check this for a single leaf, where it follows from the fact that the associated period map lands inside $\check{D}^0$.

4.4. **A constructibility fact.** A useful fact which we will use repeatedly is the following:

Proposition 4.5. *Let (P, ∇) be an algebraic H-principal bundle with flat connection on an algebraic variety S, and let $f : \mathcal{Z} \to \mathcal{Y}$ be a family of subvarieties of P. Then the locus*

$$\mathcal{Y}(f, e) := \{(x, y) : \dim_x(\mathcal{L}_x \cap \mathcal{Z}_y) \geq e\}$$

is algebraically constructible.

Above, by $\dim_x$ we denote the local analytic dimension at the point x.

Proof. By stratifying the base $\mathcal{Y}$, we may assume $\mathcal{Y}$ is smooth.

Consider the bundle (Q, δ) obtained from (P, ∇) by pulling back along the natural projection $\pi : S \times \mathcal{Y} \to S$. From the definition of the fibre product one obtains an embedding $\mathcal{Z} \subset Q$. Then $\mathcal{Z} \subset Q$ is foliated by the intersections $\mathcal{L}_z \cap f^{-1}(y)$ where $\mathcal{L}_z \subset Q$ is an analytic leaf passing through $z \in \mathcal{Z}$. The locus

$$\mathcal{Z}(f, e) := \{z \in \mathcal{Z} : \dim_z \mathcal{L}_z \cap f^{-1}(y) \geq e\}$$

is closed analytic. Note that if $x = \pi(z)$ is the image of z in P, then the germ of the intersection $\mathcal{L}_z \cap f^{-1}(y)$ at z is identified with the germ of the intersection $\mathcal{L}_x \cap f^{-1}(y)$ at x, with $\mathcal{L}_x$ a leaf of P passing through x.

In fact we claim that $\mathcal{Z}(f,e)$ is also algebraic, and describe an explicit way to construct it. We may assume that $\mathcal{Z}$ is irreducible. We denote points of $\mathcal{Z}(f,e) \subset \mathcal{Z}$ by (x,y), in accordance with the embedding into the product $P \times \mathcal{Y}$. For each e, we define

$$\mathcal{Z}(f,e)^1 := \{(x,y) \in \mathcal{Z} : \dim T_x\mathcal{L}_x \cap T_x f^{-1}(y) = e\}.$$

Because ∇ is algebraic, so is the bundle $\bigcup_{\mathcal{L}\subset P} T\mathcal{L}$ above P, hence the locus $\mathcal{Z}(f,e)^1$ is defined by algebraically constructible conditions.

Now let e_f be the dimension of the smallest component of $\mathcal{L} \cap f^{-1}(y)$ as $\mathcal{L}$ and y vary over all possible choices. Then for $e \leq e_f$ we have $\mathcal{Z}(f,e) = \mathcal{Z}(f,e_f) = \mathcal{Z}$, so it suffices to show that $\mathcal{Z}(f,e)$ is algebraic for $e > e_f$. Now for $e = e_f + 1$ the locus $\mathcal{Z}(f,e)^1$ is not dense in $\mathcal{Z}$: the intersections of the fibres $f^{-1}(y)$ with the leaves $\mathcal{L}$ must be smooth at a generic point, so have tangent space dimension at most e_f at a generic point. Letting $\mathcal{Z}' \subset \mathcal{Z}$ denote the closure of $\mathcal{Z}(f,e)^1$, we may construct the family $f' : \mathcal{Z}' \to \mathcal{Y}$ and reduce to showing the algebraicity of $\mathcal{Z}(f',e)$ for all e. The proof is therefore complete by Noetherian induction.

The locus $\mathcal{Y}(f,e)$ is just the image of $\mathcal{Z}(f,e)$, hence is algebraically constructible. □

4.5. **Ax-Schanuel in families.** Give $h : C \to B$ a family of subvarieties of S, we define

$$\mathcal{Y}(f,e,h) := \left\{ (x,y) \in P \times \mathcal{Y} : \begin{array}{c} \text{there exists an embedded analytic germ} \\ (x,\mathcal{D}) \subset (x, \mathcal{Z}_y \cap \mathcal{L}_x) \text{ of dimension} \\ \text{at least } e \text{ mapping into a fibre of } h \end{array} \right\}$$

The following says that if one applies the Ax-Schanuel theorem to a family of subvarieties then the resulting intersections belong to a family of weakly special subvarieties.

Theorem 4.6. *Fix some family of subvarieties of P, $f : \mathcal{Z} \to \mathcal{Y}$, and an integer $e \geq 1$, such that each fibre of f has dimension $< \dim H + e$.*

Suppose that there exists a countable collection of families $\{h_i : C_i \to B_i\}_{i=1}^{\infty}$ of subvarieties of S, all of whose fibres are weakly special, and that all weakly special subvarieties of S are among the fibres of these families. Then the set $\mathcal{Y}(f,e)$ introduced in Proposition 4.5 is contained in $\bigcup_{i=1}^{m} \mathcal{Y}(f,e,h_i)$ for some index m after reordering the i's.

4.6. **Proof of the geometric Zilber-Pink (after Baldi-Urbanik).** Let S be a smooth quasi-projective variety, and $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS, with associated Hodge-monodromy datum $(\mathbf{H}_S \subset \mathbf{G}_S, D_S^0 \subset D_S)$. To ease the notation, we simply set $\check{D} = \check{D}_S^0$, and $\mathbf{G} = \mathbf{G}_S$. Let $\pi : P \to S$ be the period torsor associated to $(S, \mathbb{V})$ as in Section 4.3.1, with its algebraic evaluation map $r : P \to \check{D}$.

The proof of the following is quite standard.

Lemma 4.7. *Consider the set*

$$\mathcal{Q} := \{[\check{Q}] : \check{Q} \text{ is a weakly special subdomain of } \check{D}\}.$$

Then there exists a variety $\mathcal{Y}$, an embedding $\mathcal{Q} \hookrightarrow \mathcal{Y}(\mathbb{C})$, and a family $g : \mathcal{D} \to \mathcal{Y}$ of algebraic subvarieties of $\check{D}$ such that the fibre above the point $[\check{Q}]$ is the variety $\check{Q}$.

Thanks to the next lemma, in the course of the proof of the geometric Zilber-Pink, we will freely translate the group theoretic data $(\mathbf{H}, D_H, \mathbf{N})$ into families of weakly special subvarieties.

Lemma 4.8. *The theorem is equivalent to proving that monodromically atypical weakly special subvarieties of S belong to finitely many algebraic families, where each fibre of each family is a monodromically atypical weakly special variety.*

For simplicity, from now on, we drop the adjective "monodromically" in front of words typical and atypical (Definition 3.9 plays no role in this section, and we are always concerned with Definition 3.15).

Proof of Theorem 3.17. It suffices to prove the theorem after replacing S with a finite étale covering. Then standard techniques in Hodge theory can be used to reduce to the case where the period map Ψ is quasi-finite; see indeed [11] and the references therein. We will prove that, for each e, the e-dimensional atypical weakly special subvarieties of S belong to finitely many algebraic families as in Lemma 4.8. We start by parametrizing the weakly special subvarieties of D:

Proposition 4.9. *There exists a family $f : \mathcal{Z} \to \mathcal{Y}$ of subvarieties of $P \times S$ such that all weakly special subvarieties Y arise as follows: there is a point $y \in \mathcal{Y}(\mathbb{C})$ such that Y is a component of the projection to S of $\mathcal{Z}_y \cap \mathcal{L}$ for some leaf $\mathcal{L}$ above S.*

We let f be as in Proposition 4.9, and, for any integer j, write $f^{(<j)}$ for the subfamily defined by the property that the varieties $r(\mathcal{Z}_y)$ have dimension $< j$. (This is a priori a constructible condition on $\mathcal{Y}$, but one can make it algebraic by replacing $\mathcal{Y}$ with a finite union of strata.) We say a pair of integers (e, j) is in the *atypical range* if $j < \rho(e) := e + \dim \check{D}^0_S - \dim S$. For each e we consider

$$\mathcal{Y}(e) := \mathcal{Y}(f^{(<\rho(e))}, e) = \{(x, y) \in P \times \mathcal{Y} : \dim_x(\mathcal{L}_x \cap \mathcal{Z}_y) \geq e, \dim r(\mathcal{Z}_y) < \rho(e)\}. \tag{4.6.1}$$

This is the locus in $P \times \mathcal{Y}$ where e-dimensional atypical intersections associated to atypical weakly specials can appear, indeed, since the family f comes as pull-back from $\check{D}^0_S$, the condition $\dim_x(\mathcal{L}_x \cap \mathcal{Z}_y) \geq e$ is equivalent to $\dim_{r(x)} r(\mathcal{L}_x \cap \mathcal{Z}_y) \geq e$. Thanks to Proposition 4.5 and Proposition 4.9 this is a constructible set. In fact, we will be interested in two subsets of $\mathcal{Y}(e)$:

$$\mathcal{K}(e)' = \left\{(x, y) \in \mathcal{Y}(e) : \begin{array}{c} \text{there exists } (x, y') \in \mathcal{Y}(e') \\ \text{with } e' > e \text{ and } \mathcal{Z}_y \subsetneq \mathcal{Z}_{y'} \end{array}\right\} \tag{4.6.2}$$

$$\mathcal{K}(e) = \left\{(x, y) \in \mathcal{K}(e)' : \begin{array}{c} \text{there exists } (x, y') \in \mathcal{Y}(e) \\ \text{with } \mathcal{Z}_{y'} \subsetneq \mathcal{Z}_y \end{array}\right\} \tag{4.6.3}$$

These are once more constructible subsets, thanks to Proposition 4.5.

The following proposition is the key step to prove the geometric Zilber-Pink and crucially uses the Ax-Schanuel theorem.

Proposition 4.10. *For each $(x, y) \in \mathcal{Y}(e) \setminus \mathcal{K}(e)$, let U be a component at x of $\mathcal{Z}_y \cap \mathcal{L}_x$ of dimension $\geq e$. Then the projection of U to S surjects onto the germ of a strict weakly special subvariety Y of S (of dimension e).*

Because all the varieties in our family f arise by pulling back subvarieties of $\check{D}^0_S$, it will be convenient for the following arguments to refer to the Ax-Schanuel theorem for intersections with the graph of a period map in $S \times \check{D}^0_S$ (namely Theorem 4.4), although of course all the arguments could instead be rephrased in terms of intersections with leaves of P. We have a natural map $(\pi \times r) : P \to S \times \check{D}^0_S$, where $\pi : P \to S$ is the structure morphism.

Proof. Let $(x, y) \in \mathcal{Y}(e) \setminus \mathcal{K}(e)$ and a component $U \subset (\mathcal{Z}_y \cap \mathcal{L}_x)$ with $\dim_x U \geq e$. Since $\mathcal{Z}_y$ is a fibre of the family $f^{(<\rho(e))}$, the definition of $\rho(e)$ implies that we are in the atypical range, i.e.

$$\operatorname{codim}_{S \times \check{D}^0_S}(\pi \times r)(U) < \operatorname{codim}_{S \times \check{D}^0_S}(\pi \times r)(\mathcal{L}_x) + \operatorname{codim}_{S \times \check{D}^0_S} r(\mathcal{Z}_y). \tag{4.6.4}$$

Hence, setting $s = \pi(x)$, if $(Z, s) \subset (S, s)$ is a germ obtained as the image of U, then by the Ax-Schanuel theorem Theorem 4.4 the germ (Z, s) lies in a strict weakly special subvariety of $Y \subset S$ of S. Let Y be the smallest weakly special subvariety of S containing (Z, s), and let $(\mathbf{H}_Y, D^0_Y \subset \check{D}^0_Y)$ be its monodromy datum. We fix a Hodge-theoretic embedding of $(\mathbf{H}_Y, D^0_Y \subset \check{D}^0_Y)$ into $(\mathbf{H}_S, D^0_S \subset \check{D}^0_S)$ and after replacing x and $\mathcal{Z}_y$ by an $\mathbf{H}_S(\mathbb{C})$-translate we may assume that the orbit $\check{D}' := \mathbf{H}_Y(\mathbb{C}) \cdot r(x)$ agrees with $\check{D}^0_Y$. We want to show that $r(\mathcal{Z}_y) = \check{D}'$.

Lemma 4.11. *The image $r(\mathcal{Z}_y) \subset \check{D}^0_S$ lies in $\check{D}'$.*

Proof. Consider an irreducible component C' of $r(\mathcal{Z}_y) \cap \check{D}'$ containing the image of U, which must have dimension $\leq j := \dim r(\mathcal{Z}_y)$. To prove the lemma, it is enough to show that we must have $\dim C' = j$. By construction, C' is again the image of a fibre of f, so we can write $C' = r(\mathcal{Z}_{y'})$, for some $y' \in \mathcal{Y}$. Since $\mathcal{Z}_{y'}$ contains the projection of U particular $(x, y') \in \mathcal{Y}(e)$

and $\mathcal{Z}_{y'} \subseteq \mathcal{Z}_y$. The last inclusion is not an equality precisely when $\dim C' < j$. But points where this happens are all contained in $\mathcal{K}(e)$ as a consequence of the second condition (4.6.3) defining $\mathcal{K}(e)$. $\square$

Suppose now instead that we have a strict inclusion $r(\mathcal{Z}_y) \subsetneq \check{D}'$, which is equivalent to the fact that $\dim \check{D}' - j > 0$. A priori Y is either an atypical or a typical weakly special subvariety of $(S, \mathbb{V})$. We argue each case separately and in both cases we will derive the desired contradiction: for the former we contradict (4.6.2), and for the latter the minimality of Y (this details are omitted).

We have shown that $r(\mathcal{Z}_y) = \check{D}'$, which implies that U is the germ of the weakly special subvariety Y, concluding the proof of Proposition 4.10. $\square$

Thus, by Proposition 4.10, each germ U arising in the definition $\mathcal{Y}(e) \setminus \mathcal{K}(e)$ surjects onto the germ of a weakly special subvariety of S of dimension e. The next lemma shows that, in fact, the locus $\mathcal{Y}(e) \setminus \mathcal{K}(e)$ contains all germs of maximal atypical subvarieties.

Lemma 4.12. *Each maximal (among atypical) atypical weakly special Y of dimension e induces a point of $\mathcal{Y}(e) \setminus \mathcal{K}(e)$.*

Proof. As observed after (4.6.1), it is true by construction that each smooth germ of a d-dimensional atypical subvariety Y induces a point $(x, y) = (x_Y, y_Y) \in \mathcal{Y}(e)$, where $r(\mathcal{Z}_y) = \mathbf{H}_Y(\mathbb{C}) \cdot r(x) =: \check{D}'$ and $\dim_x(\mathcal{L}_x \cap \mathcal{Z}_y) = e$.

We need to check that not all such points are contained in $\mathcal{K}(e)$:

- If there exits $(x, y') \in \mathcal{Y}(e')$ with $e' > e$ and $\check{D}_Y \subsetneq r(\mathcal{Z}_{y'})$, we would like to conclude that Y is not maximal among atypical. We can assume that y is chosen in such a way that $\mathcal{Z}_{y'}$ is minimal among the fibres of $f^{(<\rho(e'))}$ containing $\check{D}_Y$. Consider W the Zariski closure in S of the projection of $\mathcal{L}_x \cap \mathcal{Z}_{y'}$. Since we are still in the atypical range for e', by Ax-Schanuel, W lies in some strict weakly special subvariety Y_W, and we must have $Y \subset W \subset Y_W$. We claim that the smallest such Y_W has to by atypical. Let $(\mathbf{H}_W, \check{D}_W)$ be its monodromy datum. Consider $r(\mathcal{Z}_{y'}) \cap \check{D}_W$, by the minimality assumption of $\mathcal{Z}_{y'}$ and the stability under intersections of the family, we must have $r(\mathcal{Z}_{y'}) \subset \check{D}_W$. If we have equality, we conclude that Y is atypical. Assume therefore that $r(\mathcal{Z}_{y'}) \subsetneq \check{D}_W$. If Y_W is typical (i.e. $\operatorname{codim}_S Y_W = \operatorname{codim}_{\check{D}_S^0} \check{D}_W$), then $Y \times r(\mathcal{Z}_{y'}) \subset Y_W \times \check{D}_W$ gives rise to an atypical intersection, and this contradicts the minimality of Y_W.
- If there exists $(x, y') \in \mathcal{Y}(e)$ with $r(\mathcal{Z}_{y'}) \subset \check{D}'$, then $\dim_x(\mathcal{L}_x \cap \mathcal{Z}_{y'}) = e$ and therefore the projection of $\mathcal{L}_x \cap \mathcal{Z}_{y'}$ to S is equal to Y. This implies that $\mathcal{Z}_{y'}$ is invariant under the monodromy of Y, and therefore that $r(\mathcal{Z}_{y'}) = \check{D}'$.

$\square$

Now we consider a countable collection $\{h_i : C_i \to B_i\}_{i=1}^{\infty}$ of families of weakly special subvarieties, each of whose fibres are atypical and e-dimensional. One can construct this by refining the above families so that all the fibres have dimension e, and monodromy data is small enough to be in the atypical range. What we proved so far gives $\mathcal{Y}(e) \setminus \mathcal{K}(e) = \bigcup_{i=1}^{\infty} \mathcal{Y}(f^{(<\rho(e))}, e, h_i)$, and thanks to Lemma 4.12 all weakly special subvarieties induce points of this set. Applying either the Ax-Schanuel in families, or directly the constructibly fact Proposition 4.5, one sees that in fact finitely many h_i suffice, which gives the result. $\square$

5. Typical intersections–proofs

5.1. **Level, after Baldi-Klingler-Ullmo** [21].

5.1.1. *Preliminaries.*

Definition 5.1. *Let $K = \mathbb{Q}$ or $\mathbb{R}$, or a totally real number field. If $K = \mathbb{Q}$ or $\mathbb{R}$, a K-*Hodge-Lie algebra *is a reductive (finite dimensional) Lie algebra $\mathfrak{g}$ over K endowed with a K-Hodge structure of weight zero*

$$\mathfrak{g}_{\mathbb{C}} = \bigoplus_{i\in\mathbb{Z}} \mathfrak{g}^i, \qquad \text{and} \quad \mathfrak{g}^{-i} = \overline{\mathfrak{g}^i} \quad \forall i \in \mathbb{Z} \ ,$$

such that the Lie bracket $[\cdot,\cdot] : \bigwedge^2 \mathfrak{g} \to \mathfrak{g}$ is a morphism of K-Hodge structures, and the negative of the Killing form $B_{\mathfrak{g}} : \mathfrak{g}^{\mathrm{ad}} \otimes \mathfrak{g}^{\mathrm{ad}} \to K$ is a polarisation of the K-Hodge substructure $\mathfrak{g}^{\mathrm{ad}}$ (where we identify the adjoint Lie algebra $\mathfrak{g}^{\mathrm{ad}}$ with the derived one $\mathfrak{g}^{\mathrm{der}} := [\mathfrak{g}, \mathfrak{g}]$).

*If K is a totally real number field a K-*Hodge-Lie algebra *is the datum of a reductive Lie algebra $\mathfrak{g}$ over K and of a $\mathbb{Q}$-Hodge-Lie algebra structure on $\mathrm{Res}_{K/\mathbb{Q}}\, \mathfrak{g}$.*

(What we denoted above by $\mathfrak{g}^i$ is, in the literature, often denoted by $\mathfrak{g}^{i,-i}$.)

Remark 5.2. Given $h : \mathbb{S} \to \mathbf{G}_{\mathbb{R}}$ a pure Hodge structure on a real algebraic group $\mathbf{G}_{\mathbb{R}}$ in the sense of [101, page 46], the adjoint action of h on the Lie algebra $\mathfrak{g}_{\mathbb{R}}$ endows $\mathfrak{g}_{\mathbb{R}}$ with the structure of a real Hodge-Lie algebra. Conversely one easily checks that a real Hodge-Lie algebra structure on $\mathfrak{g}_{\mathbb{R}}$ integrates into a Hodge structure on some connected real algebraic group $\mathbf{G}_{\mathbb{R}}$ with Lie algebra $\mathfrak{g}_{\mathbb{R}}$. Notice also that if a simple real Lie algebra $\mathfrak{g}_{\mathbb{R}}$ admits a Hodge-Lie structure its complexification $\mathfrak{g}_{\mathbb{C}}$ is still simple, see [101, 4.4.10].

5.1.2. *Definitions.*

Definition 5.3. *Given a simple $\mathbb{R}$-Hodge-Lie algebra $\mathfrak{g}_{\mathbb{R}}$, its* level *is the largest integer k such that $\mathfrak{g}^k \neq 0$. The* level *of a simple $\mathbb{Q}$-Hodge-Lie algebra $\mathfrak{g}$ is the* maximum *of the level of the irreducible factors of $\mathfrak{g}_{\mathbb{R}} := \mathfrak{g} \otimes_{\mathbb{Q}} \mathbb{R}$. The* level *of a semi-simple $\mathbb{Q}$-Hodge-Lie algebra is the* minimum *of the levels of its simple factors. If $K = \mathbb{Q}$ or $\mathbb{R}$, the* level *of a K-Hodge structure V is the level of its adjoint K-Hodge-Lie Mumford-Tate algebra $\mathfrak{g}^{\mathrm{ad}}$.*

Remark 5.4. Notice that our Definition 5.3 is *not* the standard one. Usually, the level of any pure real Hodge structure V is defined as the maximum of $k - l$ for $V^{k,l} \neq 0$. We believe that our definition, which takes into account only the adjoint Mumford-Tate Lie algebra of V, is more appropriate. For instance, the level of the weight 2 Hodge structure $H^2(S, \mathbb{Q})$, for S a $K3$-surface, is one for Definition 5.3, reflecting the fact that the motive of S is of abelian type, while it would be two with the usual definition.

It follows from Remark 5.2 that if $\mathbb{V}$ is a polarizable $\mathbb{Z}$VHS on a smooth connected quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and algebraic monodromy group $\mathbf{H}$, then any Hodge generic point $x \in \tilde{\Phi}(\widetilde{S^{\mathrm{an}}})$ defines a $\mathbb{Q}$-Hodge-Lie algebra structure $\mathfrak{g}^{\mathrm{ad}}_x$ on $\mathfrak{g}^{\mathrm{ad}}$, with $\mathbb{Q}$-Hodge-Lie subalgebra $\mathfrak{h}_x$. One immediately checks that the levels of these two structures are independent of the choice of the Hodge generic point x in D.

Definition 5.5. *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with algebraic monodromy group $\mathbf{H}$. The* level *of $\mathbb{V}$ is the level of the $\mathbb{Q}$-Hodge-Lie algebra $\mathfrak{h}_x$ for x any Hodge generic point of $\tilde{\Phi}(\widetilde{S^{\mathrm{an}}})$.*

5.2. **Proof of Theorem 3.19.**

The proof uses the fact that the $\mathbb{Q}$-Hodge-Lie algebra $\mathfrak{h}$ of the algebraic monodromy group $\mathbf{H}$ is generated in level 1, see below; and the analysis by Kostant [79] of root systems of Levi factors for complex semi-simple Lie algebras.

Definition 5.6. *Let $K = \mathbb{Q}$ or $\mathbb{R}$. A K-Hodge-Lie algebra $\mathfrak{g}$ is said to be* generated in level 1 *if the smallest K-Hodge-Lie subalgebra of $\mathfrak{g}$ whose complexification contains $\mathfrak{g}^{-1}$ (or equivalently $\mathfrak{g}^1$) coincides with $\mathfrak{g}$.*

Proposition 5.7. *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with algebraic monodromy group $\mathbf{H}$. Then the $\mathbb{Q}$-Hodge-Lie subalgebra structure $\mathfrak{h}_s$ defined by any point $s \in S$ on $\mathfrak{h}$ is generated in level 1.*

More precisely, the $\mathbb{R}$-Hodge-Lie subalgebra structure defined by any point $s \in S$ on the non-compact part $\mathfrak{h}_{\mathbb{R}}^{\mathrm{nc}}$ of $\mathfrak{h}_{\mathbb{R}} := \mathfrak{h} \otimes_{\mathbb{Q}} \mathbb{R}$ is generated in degree 1, and the compact part $\mathfrak{h}_{\mathbb{R}}^{\mathrm{c}}$ is of Hodge type $(0,0)$.

Proposition 5.8. *Suppose $\mathfrak{g}_{\mathbb{R}}$ is a simple $\mathbb{R}$-Hodge-Lie algebra generated in level 1, and of level at least 3. If $\mathfrak{g}'_{\mathbb{R}} \subset \mathfrak{g}_{\mathbb{R}}$ is an $\mathbb{R}$-Hodge-Lie subalgebra satisfying ${\mathfrak{g}'}^i = \mathfrak{g}^i$ for all $|i| \geq 2$ then $\mathfrak{g}' = \mathfrak{g}$.*

Assuming for the moment Proposition 5.7 and Proposition 5.8, let us finish the proof of Theorem 3.19. Let $Y \subset S$ be a monodromically typical weakly special subvariety for $\mathbb{V}$, with algebraic monodromy group $\mathbf{H}_Y$. In the following we fix a point $x \in D_{H_Y} \cap \widetilde{\Psi}(\widetilde{S^{\mathrm{an}}})$ and consider the $\mathbb{Q}$-Hodge-Lie algebra pair $\mathfrak{h}_Y \subset \mathfrak{h}$ defined by this point. The typicality condition writes:

$$\dim\left(\bigoplus_{i\geq 1} \mathfrak{h}_Y^{-i}\right) - \dim \Psi(Y^{\mathrm{an}}) = \dim\left(\bigoplus_{i\geq 1} \mathfrak{h}^{-i}\right) - \dim \Psi(S^{\mathrm{an}}) \ . \tag{5.2.1}$$

As Y is monodromically typical, we have

$$\dim \mathfrak{h}_Y^{-1} - \dim \Psi(Y^{\mathrm{an}}) = \dim \mathfrak{h}^{-1} - \dim \Psi(S^{\mathrm{an}}) \ . \tag{5.2.2}$$

We thus deduce from (5.2.1) and (5.2.2) that

$$\dim\left(\bigoplus_{i\geq 2} \mathfrak{h}_Y^{-i}\right) = \dim\left(\bigoplus_{i\geq 2} \mathfrak{h}^{-i}\right) \ ,$$

thus $\mathfrak{h}_Y^{-i} = \mathfrak{h}^{-i}$ as $\mathfrak{h}_Y^{-i} \subset \mathfrak{h}^{-i}$, for all $i \geq 2$. As $\mathfrak{h}_Y^{i} = \overline{\mathfrak{h}_Y^{-i}}$ and $\mathfrak{h}^{i} = \overline{\mathfrak{h}^{-i}}$ we finally deduce:

$$\mathfrak{h}_Y^{i} = \mathfrak{h}^{i} \quad \text{for all } \ |i| \geq 2, \tag{5.2.3}$$

The assumption that $\mathfrak{h}$ is of level at least 3 says that each simple $\mathbb{Q}$-factor of $\mathfrak{h}$ is at level at least 3. Moreover the equality (5.2.3) remains true if we replace $\mathfrak{h}$ by any of its $\mathbb{Q}$-simple factors and $\mathfrak{h}_Y$ by its intersection with this simple factor. To deduce that $\mathfrak{h}_Y = \mathfrak{h}$, we can thus without loss of generality assume that $\mathfrak{h}$ is $\mathbb{Q}$-simple.

By Proposition 5.7 it follows that one simple $\mathbb{R}$-factor $\mathfrak{l}_{\mathbb{R}}$ of $\mathfrak{h}_{\mathbb{R}}^{\mathrm{nc}}$ is generated in level 1 and of level at least 3. It then follows from Proposition 5.8 that the intersection $\mathfrak{h}_{Y,\mathbb{R}} \cap \mathfrak{l}_{\mathbb{R}}$ equals $\mathfrak{l}_{\mathbb{R}}$. As both $\mathfrak{h}_Y$ and $\mathfrak{h}$ are defined over $\mathbb{Q}$ and $\mathfrak{h}$ is the smallest $\mathbb{Q}$-algebra whose $\mathbb{R}$-extension contains $\mathfrak{l}_{\mathbb{R}}$, we conclude that $\mathfrak{h}_Y = \mathfrak{h}$. Thus $Y = S$, which finishes the proof . ☐

Proof of Proposition 5.7. Passing to a finite étale cover of S if necessary, we can and will assume without loss of generality that the monodromy of the local system $\mathbb{V}$ is contained in $\mathbf{H}(\mathbb{R})$. Let $\mathbb{V}_{\mathfrak{h}}$ be the $\mathbb{Q}$VHS associated to the adjoint representation of the algebraic monodromy group $\mathbf{H}$. For each $s \in S$, let $\mathfrak{h}_s$ be the fiber at s of $\mathbb{V}_{\mathfrak{h}}$. This is a semi-simple $\mathbb{Q}$-Hodge-Lie algebra. We write $\mathfrak{h}_{\mathbb{R},s} = \mathfrak{h}_{\mathbb{R},s}^{\mathrm{nc}} \oplus \mathfrak{h}_{\mathbb{R},s}^{\mathrm{c}}$ for the decomposition of the $\mathbb{R}$-Hodge-Lie-algebra $\mathfrak{h}_{\mathbb{R},s}$ into its non-compact and compact part.

As $\mathbf{H}$ has no $\mathbb{Q}$-anisotropic factor, $\mathfrak{h}_{\mathbb{R},s}$ is the smallest subalgebra of $\mathfrak{h}_{\mathbb{R},s}$ defined over $\mathbb{Q}$ and containing $\mathfrak{h}_{\mathbb{R},s}^{\mathrm{nc}}$. Proving that $\mathfrak{h}_{\mathbb{R},s}^{\mathrm{nc}}$ is generated in level 1 thus implies that $\mathfrak{h}_s$ is generated in level 1. Let us now turn to the proof of this last statement.

Let $\mathfrak{h}'_{\mathbb{C},s} \subset \mathfrak{h}_{\mathbb{C},s}$ be the complex Lie subalgebra generated by $\mathfrak{h}_s^{-1}$ and $\mathfrak{h}_s^{1}$. It is naturally defined over $\mathbb{R}$: $\mathfrak{h}'_{\mathbb{C},s} = \mathfrak{h}'_{\mathbb{R},s} \otimes_{\mathbb{R}} \mathbb{C}$ and $\mathfrak{h}'_{\mathbb{R},s} \subset \mathfrak{h}_{\mathbb{R},s}$ is a real Lie subalgebra. We claim that the collection $\left(\mathfrak{h}'_{\mathbb{R},s}\right)_{s\in S}$ defines a local subsystem of $\mathbb{V}_{\mathfrak{h}_{\mathbb{R}}} := \mathbb{V}_{\mathfrak{h}} \otimes_{\mathbb{Q}} \mathbb{R}$. To prove the claim, it is enough to show that the corresponding holomorphic subbundle of $\mathcal{V}_{\mathfrak{h}} := \mathbb{V}_{\mathfrak{h}_{\mathbb{C}}} \otimes_{\mathbb{C}} \mathcal{O}_S$ is stable under the flat connection ∇ defined by $\mathbb{V}_{\mathfrak{h}_{\mathbb{C}}}$. But a local holomorphic vector field X on S identifies under the period map with a local section u of $F^{-1}\mathcal{V}_{\mathfrak{h}}$ over $\Psi(S^{\mathrm{an}})$; and, under this identification, the derivation ∇_X at a point s is nothing else than $\mathrm{ad}\, u$, which preserves $\mathfrak{h}'_{\mathbb{C},s}$ at each point according to the very definition of the latter. Hence the claim.

It follows that $\mathfrak{h}'_{\mathbb{R},s}$ is an $\mathbf{H}_s(\mathbb{R})$-submodule of the Lie algebra $\mathfrak{h}_{\mathbb{R},s}$ under the adjoint action, i.e. an ideal of $\mathfrak{h}_{\mathbb{R},s}$. As $\mathfrak{h}_{\mathbb{R},s}$ is semi-simple we obtain a decomposition of real Lie algebras, hence of $\mathbb{R}$-Hodge-Lie algebras:

$$\mathfrak{h}_{\mathbb{R},s} = \mathfrak{h}'_{\mathbb{R},s} \oplus \mathfrak{l}_s \ ;$$

and of the associated flag domains $D_H = D_{H'} \times D_L$. The subvariety $\widetilde{\Psi}(\widetilde{S^{\mathrm{an}}}) \subset D_H$ is horizontal, thus tangent to $D_{H'}$ at every point. Hence there exists $d_L \in D_L$ such that $\widetilde{\Psi}(\widetilde{S^{\mathrm{an}}}) \subset D_{H'} \times \{d_L\}$. It follows that the monodromy $(\mathrm{Ad}\,\rho)(\pi_1(S,s))$ of the real local system $\mathbb{V}_{\mathfrak{h}_{\mathbb{R}}}$, hence also its Zariski-closure $\mathbf{H}_{\mathbb{R}}$, is contained in $\mathbf{H}' \times \mathbf{M}_L$, where $\mathbf{M}_L$ is the $\mathbb{R}$-anisotropic stabilizer of d_L in $\mathbf{L}$. Thus $\mathbf{H}_{\mathbb{R}} = \mathbf{H}'_{\mathbb{R}} \times \mathbf{M}_L$; $\mathbf{H}'_{\mathbb{R}}$ is the non-compact part of $\mathbf{H}_{\mathbb{R}}$, which is thus generated in degree 1; and $\mathbf{L} = \mathbf{M}_L$ is the compact part of $\mathbf{H}_{\mathbb{R}}$, of pure type $(0,0)$. Hence the result. □

Remark 5.9. Proposition 5.7 should be compared with [99, Prop. 3.4]. In *op. cit.* Robles obtains a weaker result for the more general situation of an horizontal subvariety Z of D_H not necessarily coming from an $\mathbb{R}$VHS on a quasi-projective base (her result is stated for $(\mathbf{H}, D_H) = (\mathbf{G}, D)$ but the proof adapts immediately to the general case; moreover, as indicated to us by Robles the $\mathbb{Q}$s appearing in [99, Proposition 3.4] have to be replaced by $\mathbb{R}$'s). In her case the group $\mathbf{H}'_{\mathbb{R}}$ is not necessarily a factor of $\mathbf{H}_{\mathbb{R}}$. Our stronger conclusion (and easier proof) comes from the ubiquitous use of Deligne's semisimplicity theorem.

Proof of Proposition 5.8. The proof follows from the results of Kostant in [79]. We will use his notation in our setting to help the reader. Thus let us write $\mathfrak{m} := \mathfrak{g}^0$, a Levi factor of the proper parabolic subalgebra $\mathfrak{q} := F^0\mathfrak{g}_{\mathbb{C}}$. Let $\mathfrak{t} := \mathrm{cent}\,\mathfrak{m}$ be the center of $\mathfrak{m}$ and let $\mathfrak{s} := [\mathfrak{m}, \mathfrak{m}]$. Thus $\mathfrak{m} = \mathfrak{t} \oplus \mathfrak{s}$. Let $\mathfrak{r}$ be the orthogonal complement for the Killing form of $\mathfrak{m}$ in $\mathfrak{g}$ so that $\mathfrak{g}_{\mathbb{C}} = \mathfrak{m} \oplus \mathfrak{r}$ and $[\mathfrak{m}, \mathfrak{r}] \subset \mathfrak{r}$. A nonzero element $\nu \in \mathfrak{t}^*$ is called a $\mathfrak{t}$-root if $\mathfrak{g}_\nu \neq 0$, where

$$\mathfrak{g}_\nu = \{z \in \mathfrak{g}_{\mathbb{C}}, \quad \mathrm{ad}\,x(z) = \nu(x)z, \ \forall\, x \in \mathfrak{t}\}.$$

We denote by $R \subset \mathfrak{t}^*$ the set of all $\mathfrak{t}$-roots. Following [79, Theorem 0.1], the root space $\mathfrak{g}_\nu$ is an irreducible $\mathrm{ad}\,\mathfrak{m}$-module for any $\nu \in R$, and any irreducible $\mathfrak{m}$-submodule of $\mathfrak{r}$ is of this form. Moreover if $\nu, \mu \in R$ and $\nu + \mu \in R$ then $[\mathfrak{g}_\nu, \mathfrak{g}_\mu] = \mathfrak{g}_{\nu+\mu}$, see indeed Theorem 2.3 in *op. cit.*.

Let $R^+ \subset R$ be the set of positive $\mathfrak{t}$-roots defined in [79, page 139]. Thus the unipotent radical $\mathfrak{n} = F^1\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{q} = F^0\mathfrak{g}_{\mathbb{C}}$ coincides with $\bigoplus_{\nu \in R^+} \mathfrak{g}_\nu$. Let $T \in \mathfrak{t}$ be the grading element defining the Hodge graduation $\mathfrak{g}_{\mathbb{C}} = \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}^i$, see [99, Section 2.2]. Thus $\mathfrak{g}_\nu \subset \mathfrak{g}^{\nu(T)}$ and $\nu \in R^+$ if and only if $\nu(T) > 0$.

A $\mathfrak{t}$-root in R^+ is called simple if it cannot be written as a sum of two elements of R^+. Let $R_{\mathrm{simple}} \subset R^+$ denote the set of simple roots. In [79, Theorem 2.7] Kostant proves that the elements of R_{simple} form a basis of $\mathfrak{t}^*$. If $R_{\mathrm{simple}} = \{\beta_1, \cdots, \beta_l\}$ (where $\dim \mathfrak{t} = l$) then moreover the $\mathfrak{g}_{\beta_i}$, $i \in I := \{1, \cdots, l\}$, generate $\mathfrak{n}$ under bracket.

As $\mathfrak{g}_{\mathbb{R}}$ is a simple real algebra, it follows from Remark 5.2 that $\mathfrak{g}_{\mathbb{C}}$ is a simple complex Lie algebra. Since $\mathfrak{g}_{\mathbb{R}}$ is assumed to be generated in level 1, it follows that for all $i \in I$, $\mathfrak{g}_{\beta_i} \subset \mathfrak{g}^1$ (otherwise β_i would not be a simple root); and $\mathfrak{g}^1 = \bigoplus_{i \in I} \mathfrak{g}_{\beta_i}$.

Let $\mathfrak{h}_{\mathbb{C}} \subset \mathfrak{g}_{\mathbb{C}}$ be the complex Lie subalgebra generated by $\bigoplus_{i \geq 2} \mathfrak{g}^i$ and $\bigoplus_{i \geq 2} \mathfrak{g}^{-i}$. It is naturally an $\mathfrak{m}$-submodule of $\mathfrak{g}_{\mathbb{C}}$, hence $\mathfrak{h}_{\mathbb{C}} \cap \mathfrak{g}^1$ decomposes as a direct sum $\bigoplus_{j \in J \subset I} \mathfrak{g}_{\beta_j}$. As $\mathfrak{g}_{\mathbb{R}}$ is generated in degree one and the level k of $\mathfrak{g}$ is at least 3, it follows from [79, Theorem 2.3] that the subspace $[\mathfrak{g}^{-2}, \mathfrak{g}^3] \subset \mathfrak{h}_{\mathbb{C}} \cap \mathfrak{g}^1$ is not 0, thus J is not empty. Notice that, for any $i \in I - J$ and $j \in J$, one has $[\mathfrak{g}_{\beta_i}, \mathfrak{g}_{\beta_j}] = 0$. Indeed the Killing form is $\mathrm{ad}(\mathfrak{m})$-invariant, and $\mathfrak{h}$ is closed under $\mathrm{ad}(\mathfrak{m})$; hence $\mathfrak{h}^\perp$ is closed under $\mathrm{ad}(\mathfrak{m})$, *a fortiori* under $\mathrm{ad}(\mathfrak{t})$, whence $(\mathfrak{h}^\perp)^1 = \bigoplus_{i \in I-J} \mathfrak{g}_{\beta_i}$. Since $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$ as a Lie algebra, the desired conclusion follows.

If $J \neq I$ this contradicts the fact that the restriction to $\mathfrak{t}$ of the highest root of $\mathfrak{g}_{\mathbb{C}}$ is of the form $\sum_{i \in I} n_{\beta_i}\beta_i$, with all $n_{\beta_i} > 0$ for all $i \in I$, see [79, (2.39) and (2.44)]. Thus $I = J$ and $\mathfrak{h}_{\mathbb{C}} = \mathfrak{g}_{\mathbb{C}}$. □

5.3. **Typical Locus – All or nothing.** In this section we prove Theorem 3.20. As usual, we let

$$\Psi : S^{\mathrm{an}} \to \Gamma\backslash D$$

be the period map associated to $\mathbb{V}$, where $(\mathbf{G}, D) = (\mathbf{G}_S, D_S)$ is the generic Hodge datum associated to $(S, \mathbb{V})$. The proof of the above result builds on a local computation (at some smooth point).

Let $\Gamma'\backslash D'$ be a period subdomain of $\Gamma\backslash D$ such that

$$0 \leq \dim((\Psi(S^{\mathrm{an}}) \cap \Gamma'\backslash D')^0) = \dim \Psi(S^{\mathrm{an}}) + \dim \Gamma'\backslash D' - \dim \Gamma\backslash D, \tag{5.3.1}$$

that is we have one typical intersection $Z = \Psi^{-1}(\Gamma'\backslash D')^0$. Assume for simplicity that the image of Z under the period map is smooth (otherwise one needs to stratify). Let $\mathbf{H} \subset \mathbf{G}$ be the generic Mumford–Tate of D', and, as usual, write $H = \mathbf{H}(\mathbb{R})^+ \subset G = \mathbf{G}(\mathbb{R})^+$ and $\mathfrak{h}$ for its Lie algebra, which is an $\mathbb{R}$-Hodge substructure of $\mathfrak{g} = \mathrm{Lie}(G)$. Finally we denote by $N_G(H)$ the normaliser of H in G. Let

$$\mathcal{C}_H := G/N_G(H) = \{H' = gHg^{-1} : g \in G\}$$

be the set of all subgroups of G that are conjugated to H (under G) with its natural structure of real-analytic manifold. Set

$$\Pi_H := \{(x, H') \in D \times \mathcal{C}_H : x(\mathbb{S}) \subset H'\} \subset D \times \mathcal{C}_H,$$

and let π_1 (resp. π_2) be the natural projection to D (resp. to $\mathcal{C}_H$). Notice that π_i are real-analytic G-equivariant maps (where G acts diagonally on Π_H).

Let $\bar{S}$ be the preimage of $\Psi(S^{\mathrm{an}})$ in D, along the natural projection map $D \to \Gamma\backslash D$, and $\tilde{S}$ be the preimage of $\overline{S}$ in Π_H, along π_1. By restricting π_2 we have a real-analytic map

$$f : \tilde{S} \to \mathcal{C}_H.$$

By a simple topological argument, as explained for example in [34, Proposition 1], to prove that $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is dense in S, it is enough to prove that f is generically a submersion (that is a submersion outside a nowhere-dense real analytic subset B of $\tilde{S}$).[3] As being submersive is an open condition (for the real analytic topology), it is enough to find a smooth point in $\tilde{S}$ at which f is submersive.

Let us analyse this condition. Let $y = (P, H') \in D \times \mathcal{C}_H$ be a point of $\tilde{S}$. The real tangent space of $\mathcal{C}_H$ at $f(y)$ is canonically isomorphic to

$$\mathfrak{g}/\operatorname{Lie}(N_G(H')) = \mathfrak{g}/\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}').$$

The image of df at $y = (\overline{P}, H')$ is equal to

$$(\mathfrak{m} + T_P(\overline{S})_{\mathbb{R}} + \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}'))/\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}') \tag{5.3.2}$$

where $\mathfrak{m}$ is the Lie algebra of M (from the identification $D = G/M$). Thus f is a submersion at $y = (P, H')$ if and only if (5.3.2) is equal to $\mathfrak{g}/\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')$. Now we work on the *complex* tangent spaces (by tensoring all our $\mathbb{R}$-Lie algebras with $\mathbb{C}$), and use the fact that $\mathfrak{h}, \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')$, and $\mathfrak{m}$ are naturally endowed with an $\mathbb{R}$-Hodge structure. Thus

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{g}^{-w} \oplus \cdots \oplus \mathfrak{g}^{-1} \oplus \mathfrak{g}^0 \oplus \mathfrak{g}^1 \oplus \cdots \oplus \mathfrak{g}^w,$$

for some $w \geq 1$. To ease the notation set

$$U := T_P(\overline{S})_{\mathbb{C}} \subset \mathfrak{g}^{-1}.$$

It follows that f is a submersion at $y = (P, H')$ if and only if the following two conditions are satisfied (compare with [34, Proposition 2]):

$$\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')^{-k} = \mathfrak{g}^{-k} \tag{5.3.3}$$

[3]Because of its simplicity, for the convenience of the reader, we recall here Chai's argument. Let $\Omega \subset S^{\mathrm{an}}$ be an non-empty open subset. Since f is open, the image of $\pi^{-1}(\Omega) - B$ in $\mathcal{C}_H$ is open and non-empty, hence meets in a dense set the $\mathbf{G}(\mathbb{Q})_+$-conjugates of H, as $\mathbf{G}(\mathbb{Q})_+$ is topologically dense in G.

for any $k = w, \dots, 2$ (recall that $\mathfrak{m}$ is pure of type $(0,0)$, so $\mathfrak{m}^{-k} = 0$ in this range), and

$$U + \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')^{-1} = \mathfrak{g}^{-1}. \tag{5.3.4}$$

Let us now exhibit a point $y = (P, H') \in \tilde{S}$ satisfying both (5.3.3) and (5.3.4). Let P be one of the lifts in D of a smooth Hodge-generic point of $\Psi(Z^{\mathrm{an}})$ and $H' = \mathbf{H}'(\mathbb{R})^+$ where $\mathbf{H}'$ is the Mumford-Tate group at that point. Then (5.3.1) implies the following decomposition of the holomorphic tangent bundle of D at P:

$$U + T_P(D')_{\mathbb{C}} = T_P(D)_{\mathbb{C}} = \mathfrak{g}^{-w} \oplus \cdots \oplus \mathfrak{g}^{-1}. \tag{5.3.5}$$

In particular we have $(\mathfrak{h}')^{-k} = \mathfrak{g}^{-k}$, for all $k > 1$, which implies that (5.3.3) holds true, since $(\mathfrak{h}')^{-k} \subset \mathfrak{h}'$. moreover, since $(\mathfrak{h}')^{-1} + U = \mathfrak{g}^{-1}$, also (5.3.4) is satisfied at y.

5.4. **Isogenies.** To better explain Theorem 3.25, we give a simple and more concrete corollary about two principally polarized families of abelian varieties $h_i : \mathcal{A}_i \to S$.

Corollary 5.10 (Baldi-Miller-Stover-Ullmo [22, Cor. 6.6])**.** *Fix $g \geq 2$. Let S be a smooth quasi projective variety and $h_i : \mathcal{A}_i \to S$ be two principally polarized abelian scheme of relative dimension g whose algebraic monodromy group is $\mathbf{Sp}_{2g}$. Set $\mathbb{V}_i := R^1 h_{i,*}\mathbb{Z}$ for the associated VHS. If*

$$\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}} = \mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset, \tag{5.4.1}$$

then $\mathcal{A}_1$ is isogenous to $\mathcal{A}_2$ as S-abelian schemes (and the isogeny preserves the polarizations).

Proof. We will use Remark 3.24 and the following fact: Given two principally polarized families of abelian varieties $h_i : \mathcal{A}_i \to S$, one has $\mathrm{Hom}_S(\mathcal{A}_1, \mathcal{A}_2) = \mathrm{Hom}(R_1 h_{1,*}\mathbb{Z}, R_1 h_{2,*}\mathbb{Z})$; see [41, Cor. 4.4.15] which needs the full monodromy assumption. Indeed Theorem 3.25 implies that the two $\mathbb{Z}$VHSs $\mathbb{V}_i$ for $i \in \{1,2\}$ are isogenous in the sense of Definition 3.23, and the above facts imply that this notion of isogeny translates precisely to the traditional one. ☐

Proof of Theorem 3.25. The main step in the proof is the following lemma for which we do not need to assume that the generic Mumford–Tate group of $\mathbb{V}_i$ is $\mathbb{Q}$-simple.

Lemma 5.11. *Let $f_i : S \to \Gamma_i \backslash D_i$ be the period map associated with $\mathbb{V}_i$, $i = 1, 2$, and let*

$$f_1 \times f_2 : S \longrightarrow \Gamma_1 \backslash D_1 \times \Gamma_2 \backslash D_2$$

be the period map associated with $\mathbb{V}_1 \oplus \mathbb{V}_2$. Then $\mathrm{HL}(S, (\mathbb{V}_1 \oplus \mathbb{V}_2)^{\otimes})_{\mathrm{pos,atyp}}$ is analytically dense in S.

Proof. As $\mathrm{HL}(S, \mathbb{V}_i^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset$, by Theorem 3.20 there exists sequences of special subvarieties $\{Z_{i,\ell}\}_{\ell \in \mathbb{N}}$ in $\Gamma_i \backslash D_i$ and a Zariski dense sequence $\{W_\ell\}_{\ell \in \mathbb{N}}$ of components of both $\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}}$ and $\mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}}$ such that

$$W_\ell = f_1^{-1}(Z_{1,\ell})^0 = f_2^{-1}(Z_{2,\ell})^0$$

for some components $f_i^{-1}(Z_{i,\ell})^0$ of $f_i^{-1}(Z_{i,\ell})$. The fact that W_ℓ are again typical intersections, guaranteed by Theorem 3.20, implies that W_ℓ is Hodge generic for $\mathbb{V}_i$ in $Z_{i,\ell}$. By passing to a subsequence, we assume that $Z_{i,\ell}$ has fixed dimension z_i. Similarly, we may assume also that the W_ℓ have dimension that don't depend on ℓ.

Since W_ℓ is realized in two ways as a typical intersection, we have

$$\operatorname{codim}_{\Gamma_i \backslash D_i} f_i(W_\ell) = \operatorname{codim}_{\Gamma_i \backslash D_i} Z_{i,\ell} + \operatorname{codim}_{\Gamma_i \backslash D_i}(f_i(S)).$$

for each $i \in \{1,2\}$. That is, if $w_i := \dim f_i(W_\ell)$, $d_i := \dim \Gamma_i \backslash D_i$, and $s_i := \dim f_i(S)$, then

$$d_i - w_i = d_i - z_i + d_i - s_i. \tag{5.4.2}$$

Note that we do not assume that our period maps are immersive. Notice that

$$W_\ell \subset (f_1 \times f_2)^{-1}(Z_{1,\ell} \times Z_{2,\ell}) \subsetneq S.$$

Let W'_ℓ be an irreducible component of $(f_1 \times f_2)^{-1}(Z_{1,\ell} \times Z_{2,\ell})$ containing W_ℓ. We claim that W'_ℓ is an atypical intersection of positive period dimension in the Hodge locus $\mathrm{HL}(S, (\mathbb{V}_1 \oplus \mathbb{V}_2)^{\otimes})$ for all ℓ. For this, we need to show that

$$\mathrm{codim}_{\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2}(f_1 \times f_2(W'_\ell)) < \mathrm{codim}_{\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2}(Z_{1,\ell} \times Z_{2,\ell}) + \mathrm{codim}_{\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2}(f_1 \times f_2(S)).$$

Summing Equation (5.4.2) for $i = 1, 2$ we have

$$d_1 + d_2 - w_1 - w_2 = d_1 + d_2 - z_1 - z_2 + d_1 + d_2 - s_1 - s_2. \tag{5.4.3}$$

This implies the requisite equation by noticing that

$$\max\{\dim(f_1(Y)), \dim(f_2(Y))\} \leq \dim(f_1 \times f_2(Y)) \leq \dim(f_1(Y)) + \dim(f_2(Y)).$$

for any Y in S. □

Lemma 5.12. *Each of the following holds:*

(1) If the monodromy of $\mathbb{V}_i$ is equal to the derived subgroup of its Mumford–Tate group $\mathbf{G}_i$, then $(f_1 \times f_2)(S)$ is not Hodge generic in $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$.

(2) If the Mumford–Tate group $\mathbf{G}_i$ of $\mathbb{V}_i$ is $\mathbb{Q}$-simple, then $\mathbf{G}_1^{\mathrm{ad}} \cong \mathbf{G}_2^{\mathrm{ad}}$, $D_1 \simeq D_2$. Denoting by $(\mathbf{G}, D)$ the corresponding Hodge datum, we have that $(f_1 \times f_2)(S)$ is Hodge generic in a modular correspondence

$$\Gamma\backslash D \subset \Gamma_1\backslash D \times \Gamma_2\backslash D.$$

Proof. By Lemma 5.11, S has a Zariski dense set of positive dimensional atypical components of the Hodge locus for the period map $f_1 \times f_2$. Therefore S satisfies conclusion (a) or (b) of the geometric Zilber-Pink conjecture stated above (namely Corollary 3.18). If conclusion (a) holds, then $(f_1 \times f_2)(S)$ is not Hodge generic in $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$, as desired. If the monodromy of $\mathbb{V}_i$ is equal to the derived subgroup of its Mumford–Tate group $\mathbf{G}_i$, then the projection of $f_i(S)$ on every factor of $\Gamma_i\backslash D_i$ is positive dimensional and Hodge generic. The same is true for the period map $f_1 \times f_2$. Therefore conclusion (b) is not satisfied. This finishes the proof of the first part of the lemma.

If the Mumford–Tate group $\mathbf{G}_i$ of $\mathbb{V}_i$ is $\mathbb{Q}$-simple, then the monodromy of $\mathbb{V}_i$ is equal to the derived subgroup of its Mumford–Tate group $\mathbf{G}_i$ and by the first part, $(f_1 \times f_2)(S)$ is not Hodge generic in $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$. Moreover in this situation, the only strict special subvarieties of $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$ whose projections on both factors are surjective are the modular correspondences. This forces $D_1 \simeq D_2$ and $\mathbf{G}_1^{\mathrm{ad}} \cong \mathbf{G}_2^{\mathrm{ad}}$. This finishes the proof of the second part of the lemma. □

Theorem 3.25 is now a consequence of the characterization of isogenous $\mathbb{Z}$VHSs. □

5.5. **Sketch of the proof of Theorem 3.22.** Denote by $\tilde{S}$ the universal cover of S, and $\tilde{s} \in \tilde{S}$ a Hodge generic point. Fix $(\mathbf{H}, D_H) \subset (\mathbf{G}, D_G)$ a $\mathbb{V}$-admissible Hodge sub-datum and $g \in \mathbf{G}(\mathbb{R})$ such that $\tilde{s} \in g \cdot D_H$. Consider $\mathcal{U} = \tilde{S} \cap g \cdot D_H \subset D_G$, which contains $\tilde{s}$. Since $\tilde{s}$ is Hodge generic, Ax-Schanuel (with the admissibility condition plugged in) implies that $\mathcal{U}$ has an analytic irreducible component of the expected dimension (at $\tilde{s}$). The same holds true for any g' sufficiently close to g, and we conclude by density of the rational points of $\mathbf{G}$ in the real ones.

Part 3. **Applications**

6. First applications

In this section we present some concrete instances of the general Zilber-Pink viewpoint to classical problems in Algebraic Geometry involving Jacobians, Neron-Severi group of degree d surfaces (and generalizations to higher dimensional hypersurfaces), and integral points of certain moduli spaces.

6.1. **Jacobians with given Mumford–Tate group, and a question of Serre.**

In [90] Mumford shows the existence of principally polarized abelian varieties X of dimension 4 having trivial endomorphism ring, that are not Hodge generic in $\mathcal{A}_4$ (they have an exceptional Hodge class in $H^4(X^2,\mathbb{Z})$). A question often attributed to Serre is to describe "as explicitly as possible" such abelian varieties *of Mumford's type.* The most satisfying way would be to write the equation of a smooth projective curve over $\overline{\mathbb{Q}}$ of genus 4, whose Jacobian is of Mumford's type. At least 20 years ago, Gross [45, Problem 1] asked the weaker question over $\mathbb{C}$.

Theorem 6.1 (Baldi-Klingler-Ullmo [21, Thm. 3.17])**.** *There exists a smooth projective curve $C/\overline{\mathbb{Q}}$ of genus 4 whose Jacobian has Mumford-Tate group isogenous to a $\mathbb{Q}$-form of the complex group $\mathbb{G}_{\mathrm{m}} \times \mathrm{SL}_2 \times \mathrm{SL}_2 \times \mathrm{SL}_2$.*

Remark 6.2 (Relation with Serre's open image theorem)**.** Let K be a number field (with fixed embeddings $K \subset \overline{\mathbb{Q}} \subset \mathbb{C}$), and let A/K be a g-dimensional principally polarised abelian variety. For any prime number ℓ, consider the natural Galois representation associated to the ℓ-adic Tate module $T_\ell(A)$ of A:

$$\rho = \rho_{A,\ell} : \mathrm{Gal}(\overline{K}/K) \to \mathbf{GSp}_{2g}(\mathbb{Z}_\ell).$$

Assume that $\mathrm{End}(A/\mathbb{C}) = \mathbb{Z}$ and $g \not\equiv 0 \mod 4$. Then Serre proved that ρ has open image in $\mathbf{GSp}_{2g}(\mathbb{Z}_\ell)$ and asked for explicit counterexamples in dimension 4. The above is the first known instance of Jacobian J of a smooth projective genus 4-curve such that $\mathrm{End}(J) = \mathbb{Z}$ and $\mathrm{Im}(\rho_{J,\ell})$ is not open in $\mathbf{GSp}_8(\mathbb{Z}_\ell)$.

Remark 6.3. Actually our proof shows the existence of infinitely many such curves. Finding explicit equations for such a curve remains an open problem.

Sketch of the proof. Let $\mathcal{M}_4$ be the moduli space of curves of genus 4, and $j : \mathcal{M}_4 \hookrightarrow \mathcal{A}_4$ be the Torelli morphism. Denote the image of j by

$$\mathcal{T}_4^0 = j(\mathcal{M}_4) \subset \mathcal{A}_4,$$

which will be referred to as the *open Torelli locus*, and by $\mathcal{T}_4$ its Zariski closure (the so called *Torelli locus*). It is well known that $\mathcal{T}_4$ is Hodge generic in $\mathcal{A}_4$ and that $1 = 10-9 = \mathrm{codim}_{\mathcal{A}_4}(\mathcal{T}_4)$.

Recall that $\mathcal{A}_4$ contains a special curve Y whose generic Mumford-Tate group $\mathbf{H}$ is isogenous to a $\mathbb{Q}$-form of $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3/\mathbb{C}$, as proven by Mumford [90]. From Theorem 3.22, we see that $\mathcal{T}_4$ cuts many Hecke translates Y_n of Y, that is: the union, varying n, of $\mathcal{T}_4 \cap Y_n$ is dense in $\mathcal{T}_4$ (there are actually many different ways to see this fact[4]). In particular, upon extracting a sub-sequence of Y_n, we have

$$Y_n \cap \mathcal{T}_4^0 \neq \emptyset, \tag{6.1.1}$$

since it is not possible that all intersections happen on $\mathcal{T}_4 - \mathcal{T}_4^0$.

Since the Y_n are one dimensional, for each $P \in Y_n$, there are only two possibilities:

- P is a special point, i.e. $\mathbf{MT}(P)$ is a torus (so P corresponds to a 4-fold with CM);
- $\mathbf{MT}(P) = \mathbf{MT}(Y_n)$, which is isogenous to some $\mathbb{Q}$-form of $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3$.

[4]The quickest is perhaps to notice that the Torelli locus in $\mathcal{A}_4$ is a subvariety of codimension one, which is in fact ample as a divisor by a result of Igusa.

Therefore, to conclude, we have to find a n and a non-special $P \in Y_n \cap \mathcal{T}_4^0$. Heading for a contradiction, suppose that, for all n, all points of $Y_n \cap \mathcal{T}_4^0$ are special. By density of the intersections, this means that $\mathcal{T}_4$ contains a dense set of special points. André–Oort now implies that $\mathcal{T}_4$ is special, which is the contradiction we were looking for. This shows the existence of a point of $\mathcal{T}_4^0(\mathbb{C})$, corresponding to a curve $C/\mathbb{C}$ of genus 4 whose Jacobian has Mumford-Tate group isogenous to a $\mathbb{Q}$-form of the $\mathbb{C}$-group $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3$. To conclude the proof of Theorem 6.1 we just have to observe that $\mathcal{M}_4, \mathcal{A}_4, j$ and all the Y_n can be defined over $\overline{\mathbb{Q}} \subset \mathbb{C}$ and therefore all intersections we considered during the proof are defined over $\overline{\mathbb{Q}}$. The result follows. $\square$

6.2. **Noether-Lefschetz locus and questions of Harris and Voisin.** We work over $\mathbb{C}$ and fix an integer $d \geq 4$. Let $U_d = \mathbb{P}H^0(\mathbb{P}^3, \mathcal{O}(d)) - \Delta$ be the scheme parametrizing smooth surfaces X of degree d in $\mathbb{P}^3$. Consider the so called *Noether-Lefschetz locus*:

$$\mathrm{NL}_d := \{[X] \in U_d : \mathrm{Pic}(\mathbb{P}^3) \to \mathrm{Pic}(X) \text{ is not an isomorphism}\}.$$

We refer to the notes [28] for a discussion more detailed than the one given in the introduction.

Remark 6.4. NL_d is a naturally a countable union of subsets $\{Y_i\}_i$ of U_d, such that each Y_i is given by the complex points of a subvariety of U_d. In particular we can speak of irreducible components of NL_d. Notice also that each irreducible component of NL_d has a natural schematic structure, which is often non-reduced. In what follows we will always consider these components with their reduced structure.

We study the relationship between the following theorem and the recent general results on the distribution of the Hodge locus discussed before [62, 63, 106].

Theorem 6.5 (Explicit Noether-Lefschetz theorem (Green-Voisin))**.** *Each irreducible component Y of NL_d is such that*

$$d - 3 \leq \operatorname{codim}_{U_d} Y \leq h^{2,0} = \binom{d-1}{3}.$$

In view of Theorem 6.5 the following definition is natural (it appeared explicitly[5] in [36, page 668], see also the previous works [32, 65]):

Definition 6.6. *A component Y of NL_d is said to be* general *if it has codimension $h^{2,0}$, and* exceptional *otherwise.*

It is known that the union of the general components of NL_d is Zariski-dense in U_d [36], and that NL_d is even dense in U_d for the analytic topology (an argument of Green, written for instance in [108, Prop. 5.20]). See also [37] for the construction of some explicit exceptional components. We refer also to the recent works [49, 74] (inspired by the Zilber-Pink paradigm proposed in [21]). In particular such results show, abstractly, the existence of general components in NL_d and the density of their union.

From now on, we assume $d \geq 5$ (quadrics and cubic surfaces are both rational and will have no periods; while for $d = 4$, there are no exceptional components). Harris conjectured in [64, page 301] that NL_d has only finitely many exceptional components. This conjecture was first disproved by Voisin in [107], for d big enough and divisible by 4. We refer also to [28, Example 3.10] for a brief overview of the ingenious construction of Voisin, which ultimately comes from NL_4.

After disproving Harris conjecture, Voisin asked the following [107, 0.5]:

Question 6.7 (Voisin)**.** Is the union of exceptional components of the Noether-Lefschetz locus Zariski dense in U_d?

The main result of this section provides a negative answer to Question 6.7 (as expected by Voisin).

[5]In the [36] the terminology *special* is used in place of *exceptional*. In order to avoid a possible source of confusion, since the word 'special' is at odds with its use in [21], we decided to rename it from the start.

Theorem 6.8 (Baldi-Klingler-Ullmo [20, Thm. 6, Thm. 15]). *The union of the exceptional (i.e. of codimension $< h^{2,0}$) components of* NL_d *is contained in a finite union of strict* special *subvarieties of* U_d *for* $\mathbb{V}$ *(in particular it is not Zariski-dense in* U_d*).*

In particular each irreducible component W of the Zariski closure of the union of the exceptional components is contained in a strict special subvariety of U_d of the form $\Psi^{-1}(\Gamma_W \backslash D_W)^0$, where $(\mathbf{G}_W, D_W)$ is its associated Hodge datum, strictly contained in the generic Hodge datum $(\mathbf{G}, D)$ of $(U_d, \mathbb{V})$. Since the components of NL_d are maximal among the components of the Hodge locus of $\mathbb{V}$ having an extra Hodge vector, either W is one of the exceptional components, or the Mumford-Tate group $\mathbf{G}_W$ has to be associated to some non-trivial Hodge class in some tensorial construction of $\mathbb{V}_{|W}$. We don't know whether such a Hodge tensor comes from an algebraic cycle in a suitable power of the generic surface X of W. This is what happens in [107]. We record here an application to a finiteness question of the reasoning just explained.

Corollary 6.9 (Baldi-Klingler-Ullmo [20, Cor. 16]). *Among the exceptional components of* NL_d*, only finitely many have algebraic monodromy isomorphic, over the real numbers, to the group* $H := \mathrm{SO}(2h^{2,0}, h^{1,1}_{\mathrm{prim}} - 1)$*.*

Proof of Corollary 6.9. If there were infinitely many exceptional components with monodromy group (abstractly) isomorphic to H, Theorem 6.8 would imply the existence of a sub-Mumford-Tate datum $(\mathbf{G}_W, D_W)$ strictly contained in the generic Hodge datum $(\mathbf{G}, D)$ with the property that $\mathbf{G}_{W,\mathbb{R}}$ strictly contains a subgroup isomorphic to H. The very last assertion contradicts the maximality of $\mathrm{SO}(2h^{2,0}, h^{1,1}_{\mathrm{prim}} - 1)$ in $G = \mathrm{SO}(2h^{2,0}, h^{1,1}_{\mathrm{prim}})$, established by Dynkin [44, Thm. 1.2] (the argument in *op. cit.* is over $\mathbb{C}$, but the statement over $\mathbb{R}$ follows immediately in this case). Therefore the components of NL_d under examination form a finite collection. ☐

We also briefly discuss a related question on the generic Picard rank of the components of NL_d.

Question 6.10. (Ciliberto-Harris-Miranda, [36, page 668]) What is the Picard group of a surface corresponding to a general point of a general component of NL_d?

The answer is expected to be $\mathbb{Z}^2$, but this is not known (see also [83] for related discussion and results). Let us call a component Y of NL_d *Picard generic* if the Picard group of a surface corresponding to a general point of Y is $\mathbb{Z}^2$, and *Picard exceptional* otherwise. The same arguments give also the following:

Theorem 6.11 (Baldi-Klingler-Ullmo [20, Thm. 8]). *The union of the Picard generic general components of* NL_d *is dense in* $U_d(\mathbb{C})$*, while the union of the Picard exceptional components is contained in a finite union of strict* special *subvarieties of* U_d *for* $\mathbb{V}$ *(in particular is not Zariski-dense in* U_d*).*

The main ingredient in the proof of Theorem 6.8 is the following:

Proposition 6.12. *Every irreducible component* Y *of* NL_d *is special of positive period dimension, and:*

(1) If Y *is exceptional (in the sense of Definition 6.6), then it is atypical (in the sense of the previous sections);*
(2) If Y *is general, then it is typical if and only if it it has the expected generic Mumford-Tate group (i.e. isogenous to* $\mathbb{G}_m \times \mathrm{SO}(2h^{2,0}, h^{1,1}_{\mathrm{prim}} - 1)$*.*

Remark 6.13. In particular a general component of NL_d with algebraic monodromy/generic Mumford-Tate group strictly contained in the expected one is atypical. Once more the union of such components will not be Zariski dense in U_d in virtue of the geometric Zilber-Pink Theorem 3.17. Therefore we observe that the (analytic) density of NL_d in U_d comes from the union of its components which are at the same time general and typical.

Proof. Recall that Y is special (of positive period dimension), of the form $\Psi^{-1}(\Gamma_H\backslash D_H)^0$. In particular one expects that

$$\operatorname{codim}_{U_d} Y \leq \operatorname{codim}_D D_H = h^{2,0}$$

and the equality to be the "general case". It might happen that $Y = \Psi^{-1}(\Gamma_{H_0}\backslash D_{H_0})^0$, for some $\Gamma_{H_0}\backslash D_{H_0}$ strictly contained in $\Gamma_H\backslash D_H$. Let $(\mathbf{G}_Y, D_Y)$ be the Hodge datum associated to Y (or, more precisely, to $\mathbb{V}$ restricted to a normalization of the reduced scheme associated to Y). Then, Y is typical if and only if

$$\operatorname{codim}_{U_d} Y = \operatorname{codim}_D D_Y$$

(and atypical otherwise).

By construction $(\mathbf{G}_Y, D_Y)$ is contained in $(\mathbf{H}, D_H)$, and therefore

$$\operatorname{codim}_D D_Y = \operatorname{codim}_D D_H + \operatorname{codim}_{D_H} D_Y.$$

Hence Y is typical if and only if

$$\operatorname{codim}_{U_d} Y = h^{2,0} + \operatorname{codim}_{D_H} D_Y.$$

From the bound in Theorem 6.5, we see that general agree with typical if the monodromy is the expected one, and exceptional implies atypical. □

To explain the full power of Conjecture 2.5 (compared to its geometric counterpart Theorem 3.17), we record here the following special conjectural case, which really involves the zero-dimensional components and is beyond our current understanding.

Conjecture 6.14. *Suppose that $\mathcal{Y} \to \mathbb{P}^1$ is a Lefschetz pencil of degree d surfaces in $\mathbb{P}^3$. If $d \geq 5$ the Picard number of $\mathcal{Y}_s$ is greater or equal to 2 for at most a finite number of values of $s \in \mathbb{P}^1(\mathbb{C})$.*

The proof of Theorem 6.8 and the previous arguments show that the above is a special case of the Zilber-Pink conjecture. In fact, this special case of ZP was noticed a long time ago by de Jong [39, Sec. 3.2.2.] and should be attributed to him (see also [21, Sec. 3.6.] for a history of the Zilber-Pink conjecture, and a comparison with de Jong's unpublished viewpoint).

Remark 6.15 (The Noether-Lefschetz locus for arbitrary threefolds)**.** To elucidate the role of the Zilber-Pink viewpoint, we conclude with a general result on the distribution of the Noether-Lefschetz locus for arbitrary (Y, L) where Y is a smooth projective threefold and L a very ample line bundle on Y. We consider $U_{Y,L} \subset \mathbb{P}H^0(L)$, the parameter space of smooth surfaces $X \subset Y$ such that $\mathcal{O}_Y(X) = L$ (i.e. smooth surfaces in the same equivalence class as L), and define

$$\mathrm{NL}_L := \{[X] \in U_{Y,L} : \mathrm{Pic}(Y) \to \mathrm{Pic}(X) \text{ is not a surjection}\}.$$

Similarly, the results described so far elucidate also the geometry of the more general NL_L.

6.3. **Higher dimensional smooth hypersurfaces.** In this section we offer a simple geometric illustration of Theorem 3.19:

Corollary 6.16 (Baldi-Klingler-Ullmo [21, Sec. 9.1])**.** *Let $\mathbb{P}_{\mathbb{C}}^{N(n,d)}$ be the projective space parametrising the hypersurfaces X of $\mathbb{P}_{\mathbb{C}}^{n+1}$ of degree d (where $N(n,d) = \binom{n+d+1}{d} - 1$). Let $U_{n,d} \subset \mathbb{P}_{\mathbb{C}}^{N(n,d)}$ be the Zariski-open subset parametrising the smooth hypersurfaces X and let $\mathbb{V} \to U_{n,d}$ be the $\mathbb{Z}$VHS corresponding to the primitive cohomology $H^n(X, \mathbb{Z})_{\mathrm{prim}}$.*

If $n = 3$ and $d \geq 5$; or $n = 4$ and $d \geq 6$; or $n = 5, 6, 8$ and $d \geq 4$; or $n = 7$ or ≥ 9 and $d \geq 3$, then the level of $\mathbb{V} \to U_{n,d}$ is at least 3, and therefore $\mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}} \subset U_{n,d}$ is algebraic.

6.4. **Some cases of the refined Bombieri-Lang conjecture.** Let $U_{n,d}$ be the Hilbert scheme of smooth hypersurfaces of degree d in $\mathbb{P}^{n+1}$, this is a smooth affine scheme over $\mathbb{Z}$. In a recent breakthrough, Lawrence and Venkatesh proved the following:

Theorem 6.17 ([82, Thm. 10.1, Prop. 10.2]). *There exist $n_0 \in \mathbb{N}_{\geq 3}$ and a function $d_0 : \mathbb{N} \to \mathbb{N}$ such that,*

$$\text{for every } n \geq n_0 \text{ and } d \geq d_0(n), \tag{6.4.1}$$

the set $U_{n,d}(\mathbb{Z}[S^{-1}])$ is not Zariski dense in $U_{n,d}$, for every finite set of primes S.

Consider $f_{n,d} : X_{n,d} \to U_{n,d}$ the universal family of smooth degree d hypersurfaces in $\mathbb{P}^{n+1}$. We denote by $\mathbb{V}$ the polarized $\mathbb{Z}$-variation of Hodge structure $(R^n f^{\mathrm{an}}_{n,d,\mathbb{C}_*}\mathbb{Z})_{\mathrm{prim}}$ on $U_{n,d,\mathbb{C}}$ and by $\Psi : U^{\mathrm{an}}_{n,d,\mathbb{C}} \to \Gamma\backslash D$ the associated period map. An irreducible algebraic subvariety $Y \subset U_{n,d,\mathbb{C}}$ is said to be *of positive period dimension* if $\Psi(Y^{\mathrm{an}}_{\mathbb{C}})$ has positive dimension. We prove the following reinforcement of Theorem 6.17:

Theorem 6.18 (Baldi-Klingler-Ullmo [19, Main Thm.]). *As long as* (6.4.1) *is satisfied, there exists a closed strict subscheme $E \subset U_{n,d}$ such that, for all finite set of primes S, we have*

$$\overline{U_{n,d}(\mathbb{Z}[S^{-1}])}_{\mathrm{pos}} \subset E,$$

where $\overline{U_{n,d}(\mathbb{Z}[S^{-1}])}_{\mathrm{pos}}$ denotes the union of the irreducible components of the Zariski closure of $U_{n,d}(\mathbb{Z}[S^{-1}])$ in $U_{n,d}$ of positive period dimension. That is: the Zariski closure of $U_{n,d}(\mathbb{Z}[S^{-1}]) - E(\mathbb{Z}[S^{-1}])$ has period dimension zero.

Remark 6.19. The complement of $U_{n,d}$ in $\mathbb{P}^{N(n,d)}$ is a hypersurface. Hence $U_{n,d}$ is an open affine subvariety, stable under the natural $\mathbf{PGL}(n+2)$-action on $\mathbb{P}^{N(n,d)}$. Let $\mathcal{M}_{n,d} := [\mathbf{PGL}(n+2)\backslash U_{n,d}]$ be the stack of smooth hypersurfaces in $\mathbb{P}^{n+1}$ of degree d. This is a finite type separated Deligne-Mumford algebraic stack over $\mathbb{Z}$ with affine coarse space. The period map $\Psi : U^{\mathrm{an}}_{n,d,\mathbb{C}} \to \Gamma\backslash D$ factorizes through $\mathcal{M}^{\mathrm{an}}_{n,d,\mathbb{C}}$. Notice moreover that the Torelli theorem assures that the period map $\mathcal{M}^{\mathrm{an}}_{n,d,\mathbb{C}} \to \Gamma\backslash D$ is quasi-finite for $(n,d) \neq (2,3)$. The moduli stack $\mathcal{M}^{\mathrm{an}}_{n,d,\mathbb{C}}$ is thus *Brody hyperbolic.* A famous conjecture of Bombieri-Lang thus predicts that $\mathcal{M}_{n,d}(\mathbb{Z}[S^{-1}])$ is finite. In particular E in Theorem 6.18 should be empty.

The following elucidation of the Lawrence-Venkatesh method for proving Theorem 6.17 will be crucial for us. Lawrence and Venkatesh actually prove that (quoting the third paragraph of [82, Section 1.1])–see also the last three lines of [82, Theorem 10.1].

Theorem 6.20 (Lawrence-Venkatesh). *The monodromy for the universal family of hypersurfaces must drop over each component of the Zariski closure of the integral points. I.e. for any S, there exists a closed subscheme V_S of $U_{n,d}$ (over $\mathbb{Z}[S^{-1}]$) whose irreducible components are of positive period dimension and not* monodromy generic, *such that $U_{n,d}(\mathbb{Z}[S^{-1}]) - V_S(\mathbb{Z}[S^{-1}])$ is finite.*

By the André-Deligne monodromy theorem Theorem 3.12 (see for example [21, Section 4 and 5]) and the fact that the $\mathbb{Z}$VHS $\mathbb{V}$ is irreducible, it follows that each V_S lies in the Hodge locus of positive period dimension $\mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}}$.

Remark 6.21. The Lawrence-Venkatesh method requires the choice of an auxiliary prime number p, and the choice of an identification between $\mathbb{C}$ and $\overline{\mathbb{Q}}_p$. Indeed, to prove that the $\mathbb{Z}[S^{-1}]$-points of $U_{n,d}$ are not Zariski dense, Lawrence and Venkatesh prove that some *p-adic period map* sending $x \in U_{n,d}(\mathbb{Z}[S^{-1}])$ to some p-adic representation of the absolute Galois group of $\mathbb{Q}_p$ has fibers that are not Zariski dense in $U_{n,d}$. This is done by working on a residue disk in $U_{n,d}(\mathbb{Q}_p)$ and the p-adic and complex period maps are then related by a study of the Gauss-Manin connection [82, Lemma 3.2]. What their proof actually shows, with respect to our fixed embedding $\overline{\mathbb{Q}} \subset \mathbb{C}$ is that for each S, there exists an automorphism ι_p of $\mathbb{C}$ such that V_S is contained in $\mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})^{\iota_p}_{\otimes} \subset U_{n,d,\mathbb{C}}$. What allows us to say that V_S lies in $\mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}}$ is the fact that $\mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}}$ is actually defined over some number field.

Proof. The proof is essentially a combination of Theorem 6.17 and our earlier results on the Hodge locus.

It follows from Corollary 6.16 that $\mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}}$ is a (closed, strict) algebraic subvariety of $U_{n,d}$ and, thanks to the elucidation of Theorem 6.17, we have

$$\bigcup_S \overline{U_{n,d}(\mathbb{Z}[S^{-1}])}_{\mathrm{pos}} = \bigcup_S V_S \subset \mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}},$$

where the union ranges over all finite set of primes S. It follows from Corollary 6.16 that

$$E' := \overline{\bigcup_S V_S}^{\mathrm{Zar}} \subset \mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}}.$$

We remark here that the above inclusion may happen to be strict. Therefore we obtained a closed $\overline{\mathbb{Q}}$-subvariety $E' \subset \mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}}$ containing all V_S (seen as $\overline{\mathbb{Q}}$-varieties). The Zariski closure E in $\mathbb{P}^N_{\mathbb{Z}}$ of E' enjoys the desired property. The proof of the Theorem is concluded. □

7. Complex hyperbolic lattices

In the next section we present the main results of [23]. A cornerstone in the representation theory of lattices in Lie groups is Margulis's superrigidity theorem [84, p. 2]. Indeed, the implications of superrigidity are profound, including arithmeticity of irreducible lattices in all noncompact semisimple Lie groups except $\mathrm{PO}(1, n)$ and $\mathrm{PU}(1, n)$, i.e., for all but real and complex hyperbolic lattices. Specifically, Margulis proved superrigidity of irreducible lattices in semisimple Lie groups of real rank at least two. Using the theory of harmonic maps, this was later extended to the rank one groups $\mathrm{Sp}(1, n)$ and $\mathrm{F}_4^{(-20)}$ by Corlette and Gromov–Schoen. While the harmonic maps method was extended to give a new proof in higher rank [88], it is open whether the methods of Margulis using Ergodic Theory can give an independent proof in the rank one cases. In this section we study lattices in $\mathrm{PU}(1, n)$, and we briefly discuss this setting.

Let $n \in \mathbb{Z}_{\geq 1}$ and consider the following hermitian form on $\mathbb{C}^{n+1}$

$$h(z) = |z_0|^2 + \cdots + |z_{n-1}|^2 - |z_n|^2$$

We define the n-dimensional complex ball:

$$\mathbb{B}^n = \mathbb{B}^n_{\mathbb{C}} := \{[z_0 : z_1 : \cdots : z_n] \in \mathbb{P}^n_{\mathbb{C}} : h(z) < 0\} \subset \mathbb{C}^n \subset \mathbb{P}^n_{\mathbb{C}}.$$

and its group of automorphisms:

$$\mathrm{PU}(1, n) := \{g \in \mathrm{PGL}_{n+1}(\mathbb{C}) : g \cdot \mathbb{B}^n = \mathbb{B}^n\}$$

We have: $\mathbb{B}^n \cong \mathrm{PU}(1, n)/\,\mathrm{U}(n)$ and our task is to understand the finite volume quotients

$$S_\Gamma := \Gamma \backslash \mathbb{B}^n.$$

7.1. Non–arithmetic complex hyperbolic lattices.

Regarding commensurability classes of non-arithmetic lattices in $\mathrm{PU}(1, n)$, at the time of writing this paper, we have:

- When $n = 2$, by the work of Deligne, Mostow and Deraux, Parker, Paupert, there are 22 known commensurability classes in $\mathrm{PU}(1, 2)$.
- When $n = 3$, by the work of Deligne, Mostow and Deraux, there are 2 known commensurability classes of non-arithmetic lattices in $\mathrm{PU}(1, 3)$. In both cases the trace field is $\mathbb{Q}(\sqrt{3})$ and the lattices are not cocompact.

For $n > 3$ non-arithmetic lattices in $\mathrm{PU}(1, n)$ are currently not known to exist. One of the biggest challenges in the study of complex hyperbolic lattices is to understand for each n how many commensurability classes of non-arithmetic lattices exist in $\mathrm{PU}(1, n)$.

7.2. **An arithmeticity criterion.** Consider the quotient by Γ of the symmetric space X associated to $G = \mathrm{PU}(1,n)$. In the complex hyperbolic case we obtain a ball quotient $S_\Gamma = \Gamma\backslash X$ which has a natural structure of a quasi-projective variety, as proven by Baily-Borel [7] in the arithmetic case, and by Mok [43] in general. By the commensurability criterion for arithmeticity of Margulis [84] we can decide whether Γ is arithmetic or not by looking at modular correspondences in the product $S_\Gamma \times S_\Gamma$. That is Γ is arithmetic if and only if S_Γ admits infinitely many totally geodesic correspondences. In this section, by totally geodesic subvarieties we always mean complex subvarieties of S_Γ, whose smooth locus is totally geodesic with respect to the canonical Kähler metric. Even if Γ is arithmetic, S_Γ may have no strict totally geodesic subvarieties. However, the existence of countably many Hecke correspondences implies that, if S_Γ contains one totally geodesic subvariety, then it contains countably many of such. We study totally geodesic subvarieties of S_Γ from multiple point of views, ultimately explaining how often and why they appear.

Theorem 7.1 (Baldi-Ullmo [23, Thm 1.2.1], and independently Bader-Fiser-Miller-Stover [6])**.** *If S_Γ contains infinitely many maximal complex totally geodesic subvarieties, then $\Gamma \subset G$ is arithmetic.*

Remark 7.2. The above statement is equivalent to the following one. Let $(Y_i)_{i\in\mathbb{N}}$ be a family of totally geodesic subvarieties of S_Γ. Then every irreducible component of the Zariski closure of the union $\bigcup_{i\in\mathbb{N}} Y_i$ is either one of the Y_i's, or it is a special and arithmetic subvariety of S_Γ.

Bader, Fisher, Miller and Stover [5, 6] proved the real and complex hyperbolic version of the theorem using some superrigidity theorems and results on equidistribution from homogeneous dynamics (see also earlier work of Margulis-Mohammadi [86]). This answers affirmatively a question of C. McMullen and A. Reid.

Remark 7.3. A similar strategy applied to $S_\Gamma \times S_\Gamma$ gives the a new proof of Margulis commensurator theorem for lattices in $\mathrm{PU}(1,n)$, $n > 1$ and some lattices in $\mathrm{PU}(1,1)$ (including triangle groups). See indeed [23, Thm 6.2.2].

7.3. **Basic idea behind Theorem 7.1.** Let $K \subset \mathbb{R}$ be the smallest field for which there is a K-form of G, which we denote by $\mathbf{G}$, and $\Gamma \subset \mathbf{G}(K)$. For simplicity, let us assume that K is a totally real number field and that $\Gamma \subset \mathbf{G}(\mathcal{O}_K)$, where ($\mathcal{O}_K$ denotes the ring of integers of K.

We can define the *arithmetic approximation* of the hermitian symmetric space S_Γ as follows. Since Γ is Zariski dense in G and has entries in K, it defines a K-form of G, which we denote by $\mathbf{G}$. We observe that $\Gamma \subset \mathbf{G}(\mathcal{O}_K)$ (up to finite index and conjugation). Now $\mathbf{G}(\mathcal{O}_K)$ is an arithmetic lattice in the real points $\widehat{\mathbf{G}} := \mathrm{Res}_{K/\mathbb{Q}}\,\mathbf{G}$. In more geometrical therms, we can realise it as the fundamental group of a symmetric space

$$AA(S_\Gamma) := \mathbf{G}(\mathcal{O}_K)\backslash\widehat{\mathbf{G}}(\mathbb{R})/\widehat{K},$$

where $\widehat{K}$ is some compact maximal subgroup. Given a totally geodesic submanifold $Y \subset S_\Gamma$ (let's say of dimension $m > 0$), we can perform the same construction. Indeed, by assumption, Y can be written as a sub-ball quotient $\Gamma_Y\backslash\mathbb{B}^m$, where Γ_Y is a lattice in $H = \mathrm{PU}(1,m)$, and find

$$AA(Y) = \mathbf{H}(\mathcal{O}_K)\backslash\widehat{\mathbf{H}}(\mathbb{R})/\widehat{K}_H.$$

We also obtain

$$AA(Y) \subset AA(S_\Gamma).$$

Moreover we have an immersive map (at least in the category of topological spaces)

$$S_\Gamma \to AA(S_\Gamma)$$

and every totally geodesic submanifold Y is closely related to the intersection in $AA(S_\Gamma)$ between S_Γ and $AA(Y)$. Intuitively we think that the higher the codimension of S_Γ in $AA(S_\Gamma)$, the more "non-arithmetic" Γ is. Indeed it's easy to observe that Γ is arithmetic iff such a codimension is zero. Now we can try compute the real dimensions of all objects involved. Naively, assuming that Γ is non-arithmetic, we would expect:

$$\dim S_\Gamma = 2n, \dim Y = m, \dim AA(S_\Gamma) \approx [K:\mathbb{Q}]\cdot 2n, \dim AA(Y) \approx [K:\mathbb{Q}]\cdot m.$$

If so, we have that

$$\operatorname{codim}_{AA(S_\Gamma)} Y < \operatorname{codim}_{AA(S_\Gamma)} S_\Gamma + \operatorname{codim}_{AA(S_\Gamma)} AA(Y);$$

indeed a direct computation shows that

$$[K:\mathbb{Q}]\cdot 2n - m < [K:\mathbb{Q}]\cdot 2n - 2n + [K:\mathbb{Q}]\cdot 2n - [K:\mathbb{Q}]\cdot m.$$

If such construction was performed *inside* algebraic geometry, such intersections would be quite odd: generically two curves in $\mathbb{P}^2_{\mathbb{C}}$ intersect in a bunch of points, and so on... It's tempting to deduce the result from our geometric Zilber-Pink Theorem 3.17.

The main steps needed to make the strategy sketched above work are the following:

(1) Realise S_Γ inside a Shimura variety/a period domain for $\mathbb{Z}$-VHS;
(2) Show that totally geodesic subvarieties are atypical intersections;

The main tools for each step are:

(1) Simpson's theory [101] and a result of Esnault-Groechenig [48] (see also [81]);
(2) Monodromy/Mumford–Tate computations;

7.4. **Step 1: constructing a $\mathbb{Z}$VHS.** By construction, S_Γ supports a real VHS, which we denote by $\mathbb{V}$. This is the so called *uniformizing VHS.*

Theorem 7.4 ([23, Thm 1.3.1])**.** *For every $\gamma \in \Gamma$, the trace of $\mathrm{Ad}(\gamma)$ lies in the ring of integers of a totally real number field K. As a consequence, up to conjugation by G, Γ lies in $\mathbf{G}(\mathcal{O}_K)$. Moreover $\mathbb{V}$ induces a $\mathbb{Z}$-variation of Hodge structure $\widehat{\mathbb{V}}$ on S_Γ.*

To construct a $\mathbb{Z}$-VHS we need a strong rigidity condition, namely *cohomological rigidity.* Building on a study initiated by Weil [111] in the cocompact case, Garland and Raghunathan proved the following.

Theorem 7.5 ([60, Thm. 1.10])**.** *Let G be a semisimple Lie group, not locally isomorphic to* SL_2*, nor to* $\mathrm{SL}_2(\mathbb{C})$*. For any lattice Γ in G, the first Eilenberg–MacLane cohomology group of Γ with respect to the adjoint representation is zero. In symbols:*

$$H^1(\Gamma, \mathrm{Ad}) = 0.$$

To see why such a vanishing is related to a rigidity result, observe that the the space of first-order deformations of $\rho : \Gamma \hookrightarrow G$ is naturally identified with $H^1(\Gamma, \mathrm{Ad})$, where

$$\mathrm{Ad} : \Gamma \xrightarrow{\rho} G \xrightarrow{\mathrm{Ad}} \mathrm{Aut}(\mathfrak{g}) \tag{7.4.1}$$

is the adjoint representation.

The following is proven by Esnault and Groechenig [48] and its proof relies on Drinfeld's theorem on the existence of ℓ-adic companions over a finite field. Theorem 7.4 is a consequence of such result.

Theorem 7.6 ([48, Thm. 1.1])**.** *Let S be a smooth connected quasi-projective complex variety. Then a complex local system $\mathbb{V}$ on S is integral, i.e. it comes as extension of scalars from a local system of projective $\mathcal{O}_L$-modules of finite type (for some number field $L \subset \mathbb{C}$), whenever it is:*

(1) Irreducible;
(2) Quasi-unipotent local monodromies around the components at infinity of a compactification with normal crossings divisor $i : S \hookrightarrow \overline{S}$;
(3) Cohomologically rigid, that is $\mathbb{H}^1(\overline{S}, i_{!*}\operatorname{End}^0(\mathbb{V}))$ vanishes;
(4) Of finite determinant.

Here $i_{!*}\operatorname{End}^0(\mathbb{V})$ denotes the intermediate extension seen as a perverse sheaf. See [48, Remark 2.4] for more details. Moreover $\mathbb{H}^1(\overline{S}, i_{!*}\operatorname{End}^0(\mathbb{V}))$ is the Zariski tangent space at the moduli point of $\mathbb{V}$ of the Betti moduli stack of complex local systems of given rank with prescribed determinant and prescribed local monodromies along the components of the normal crossing divisor $\overline{S} - i(S)$.

Briefly, to obtain the theorem discussed above from those two ingredients; we notice that infinitesimal rigidity implies cohomologically rigidity, moreover a computation with the explicit toroidal compactification of S_Γ shows that $\mathbb{V}$ has quasi-unipotent monodromy at ∞, and that twists by $\sigma : K \to \mathbb{R}$ preserve infinitesimal rigidity. The crucial fact, due to Simpson and Corlette is that any rigid local system naturally underlies a (complex) VHS. Therefore each $\mathbb{V}^\sigma$ is a VHS. Eventually $\bigoplus_\sigma \mathbb{V}^\sigma$ has a natural structure of $\mathbb{Z}$VHS.

Now we introduce the *Fundamental commutative diagram* (which generalises the theory of modular embeddings for triangle groups):

- $\widehat{\mathbb{V}} := \bigoplus_i \mathbb{V}^{\sigma_i}$, $\sigma_1 = \mathrm{id}, \dots, \sigma_r : K \to \mathbb{R}$;
- $\widehat{\mathbf{G}} :=$ Weil restriction from K to $\mathbb{Q}$ of $\mathbf{G}$.

Griffiths theory of period domains and period maps gives a commutative diagram in the complex analytic category:

$$\begin{tikzcd} X \arrow[r, hook, "\tilde{\Psi}"] \arrow[d, "\pi"] & D = D_{\widehat{G}} \arrow[d, "\pi"] \\ S_\Gamma{}^{\mathrm{an}} \arrow[r, "\Psi"] & \widehat{\mathbf{G}}(\mathbb{Z})\backslash D \end{tikzcd}$$

It may happen that $\widehat{\mathbf{G}}(\mathbb{Z})\backslash D$ is a Shimura variety. But in general $\widehat{\mathbf{G}}(\mathbb{Z})\backslash D$ is not algebraic, as we discussed before, and Ψ is only known to be holomorphic.

More facts about $D_{\widehat{G}}$ and $\tilde{\Psi}$ (essentially following reinterpreting in a geometric way the fact that Ψ is the period map associated to a VHS which is $\mathbb{Q}$ simple but splits over K):

- $D_{\widehat{G}}$ is a $\widehat{G} = \widehat{\mathbf{G}}(\mathbb{R}) = \prod G_{\sigma_i}$-orbit of one of the HSs constructed above, and the stabiliser is compact;
- All the G_{σ_i} are isomorphic over $\mathbb{C}$, so they are $\mathrm{PU}(p_{\sigma_i}, q_{\sigma_i})$, $p_{\sigma_i} + q_{\sigma_i} = n+1$;
- We can write $D_{\widehat{G}} = X \times X'$ where X' is homogeneous under $\prod_{i>1}^r G_{\sigma_i}$;
- $\tilde{\Psi}$ is holomorphic and Γ-equivariant:
$$\tilde{\Psi}(\gamma.x) = (\gamma.x, \sigma_2(\gamma).x_{\sigma_2}, \dots, \sigma_r(\gamma).x_{\sigma_r}),$$
where x_{σ_i} is the fibre of $\mathbb{V}^{\sigma_i}$ at x;
- $\tilde{\Psi}$ detects arithmeticity: Γ is arithmetic iff X' is a point, i.e. G_{σ_i} is compact for any $i > 1$ (Mostow-Vinberg).

7.5. **Step 2: Γ-special vs $\mathbb{Z}$-special.** Two ways for constructing *special* algebraic subvarieties of S_Γ:

Γ. Take a subgroup $H \subset G$, such that $\Gamma_H = \Gamma \cap H$ a lattice and $H.x \subset X$ a sub-Hermitian domain; then $\pi(H.x) \subset S_\Gamma$ is algebraic;

$\mathbb{Z}$. Take a $\mathbb{Q}$-subgroup $\mathbf{M} \subset \widehat{\mathbf{G}}$, which is the Mumford-Tate group of some element $\hat{x} \in D_{\widehat{G}}$; then $\Psi^{-1}(\mathbf{M}(\mathbb{Z})\backslash\mathbf{M}(\mathbb{R}).\hat{x})'$ is algebraic (Cattani-Deligne-Kaplan).

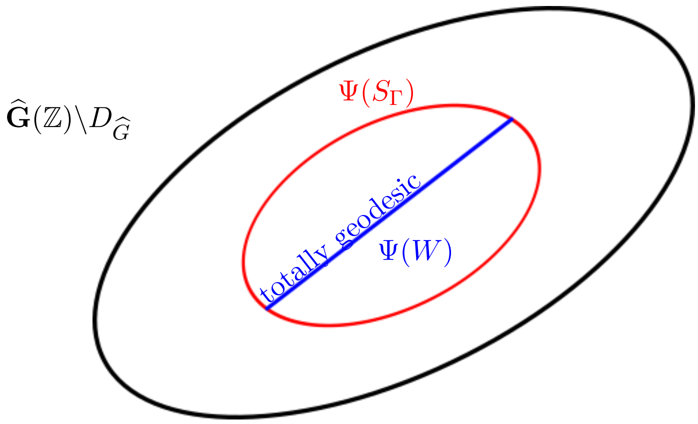
We call the former Γ-*special subvarieties* and the latter $\mathbb{Z}$-*special.* It's easy to see that the Γ-specials are also $\mathbb{Z}$-special. The idea is the following. We simply take $\mathbf{M}$ to be the Weil restriction from K to $\mathbb{Q}$ of the K-form of H given by Γ_H.

What about the converse implication? Can we take all $\mathbf{M}$'s to be Weil restriction? The answer is yes, but the proof is more complicated and we omit here the details.

Theorem 7.7 (Baldi-Ullmo [23, Cor 5.5.2])**.** *Let $W \subset S_\Gamma$ be an irreducible algebraic subvariety. The following are equivalent:*

(1) W is totally geodesic;
(2) W is bi-algebraic: some (equivalently any) analytic component of the preimage of W along $\pi : X \to S_\Gamma$ is algebraic;
(3) W is Γ-special;
(4) W is $\mathbb{Z}$-special;
(5) W is a component of $\Psi^{-1}(\pi(Y))$ for some algebraic subvariety Y of $D^\vee$.

The following picture can clarify what is going on:



where

$$\Psi(W) = \Psi(S_\Gamma) \cap \widehat{\mathbf{H}}(\mathbb{Z})\backslash D_{\widehat{H}}$$

for some $\widehat{\mathbf{H}}(\mathbb{Z})\backslash D_{\widehat{H}} \hookrightarrow \widehat{\mathbf{G}}(\mathbb{Z})\backslash D_{\widehat{G}}$. That's to say that the maximal totally geodesic subvarieties are components of the Hodge locus (of positive period dimension) for S_Γ and the $\mathbb{Z}$VHS we have constructed above. From here, we use the technology developed so far to control their distribution.

7.6. **End of the proof.** The idea is now that if Γ is not arithmetic, W is an atypical intersection, and our Theorem 3.17 then implies the main theorem.

Rather than proving in detail the fact that *non arithmetic implies atypical intersection*, let us discuss the simplest example, where we can compute everything directly:

- $\Gamma \subset G = \mathrm{PU}(1,2)$ non arithmetic, with trace field K of degree 2 over $\mathbb{Q}$. So $\widehat{G} = G \times G$.
- $W \subset S_\Gamma$ special subvariety (associated to $\mathbf{H} \subset \mathbf{G}/K$);
- Suppose that $\widehat{\mathbf{G}}(\mathbb{Z})\backslash D_{\widehat{G}}$ is a Shimura variety;
- Write $W = S_\Gamma \cap \widehat{\mathbf{H}}(\mathbb{Z})\backslash D_{\widehat{H}}$;
- $\operatorname{codim}_{\widehat{\mathbf{G}}(\mathbb{Z})\backslash D_{\widehat{G}}} S_\Gamma = 2$;
- $\operatorname{codim}_{\widehat{\mathbf{G}}(\mathbb{Z})\backslash D_{\widehat{G}}} \widehat{\mathbf{H}}(\mathbb{Z})\backslash D_{\widehat{H}} = 2$;
- $\operatorname{codim}_{\widehat{\mathbf{G}}(\mathbb{Z})\backslash D_{\widehat{G}}} W = 3$.

That is: totally geodesic subvarieties are atypical. This concludes the proof, by applying Theorem 3.17.

8. Rich representations and superrigidity

In this section we present the main results of [22], the proofs are essentially applications of Theorem 3.25, so we simply state the main results.

Let Γ be a lattice in a Lie group G. A representation ρ from Γ to a topological group H is said to *extend* if there is a continuous homomorphism $\widetilde{\rho} : G \to H$ so that ρ is the restriction of $\widetilde{\rho}$ to Γ under its lattice embedding in G. Then Γ is *superrigid* if every unbounded, Zariski dense representation of Γ to a connected, adjoint simple algebraic group H over a local field k of characteristic zero extends. It is known that there are lattices in $\mathrm{PO}(1,n)$ and $\mathrm{PU}(1,n)$ that are not superrigid. The study of representations of real and complex hyperbolic lattices is very difficult and rich in open questions, and the extent to which superrigidity fails has yet to be made precise. For example, one motivation is the first part of a question asked by David Fisher [53, Qtn. 3.15]:

Question 8.1 (D. Fisher)**.** Let $\Gamma \subset G$ be a lattice where $G = \mathrm{SO}(1,n)$ or $\mathrm{SU}(1,n)$. What conditions on a representation $\rho : \Gamma \to \mathrm{GL}_m(k)$ imply that ρ extends or almost extends? What conditions on Γ imply that Γ is arithmetic?

In particular, we build on the techniques from [23] to explore *geometric* conditions on a representation that allow one to use tools from dynamics and/or Hodge theory to prove new rigidity phenomena. The conditions we provide, which address Question 8.1 through the behavior of totally geodesic submanifolds under maps related to the representation, have also been considered in studying the infinite volume setting. There is some history of success along these lines. For just one family of examples, some classes of representations studied in the literature have *maximal* geometric behavior, where two distinct instances are:

- Maximal for the *Toledo invariant* (Koziarz–Maubon and many others)
- Maximal for their *limit set* (Besson, Courtois and Gallot and Shalom)

These considerations will lead us to study a class of representations generalizing those shown to be superrigid in [5, 6], namely those that are *geodesically rich*, we refer to [22, Sec. 1.2].

8.1. **Siu's immersion problem.**

An original motivation for studying the more general rigidity questions considered here is Siu's *immersion problem* [102, Prob. (b), p. 182]. Let $\mathbb{B}^n$ denote the unit ball in $\mathbb{C}^n$ with its Bergman metric. If Γ is a lattice in the holomorphic isometry group $\mathrm{PU}(1,n)$ of $\mathbb{B}^n$, then S_Γ will denote the *ball quotient* $\Gamma\backslash\mathbb{B}^n$.

Question 8.2 (Siu, 1985)**.** Is it true that every holomorphic embedding between compact ball quotients of dimension at least 2 must have totally geodesic image?

To our knowledge the only previous result is by Cao and Mok [31], which answered Question 8.2 in the affirmative for $f : \Gamma\backslash\mathbb{B}^n \to \Lambda\backslash\mathbb{B}^m$ when $n < m < 2n$, and uniform lattices. See [6, Thm. 1.7] and the surrounding references for more on Siu's analogous *submersive problem.* The following is our main contribution toward a positive answer to Siu's immersion problem.

Theorem 8.3 (Baldi-Miller-Stover-Ullmo [22, Thm. 1.3])**.** *Let $f : S_{\Gamma_1} \to S_{\Gamma_2}$ be an immersive holomorphic map between arithmetic ball quotients such that $f(S_{\Gamma_1})$ does not lie in any smaller totally geodesic subvariety of S_{Γ_2}. Suppose that S_{Γ_1} contains a proper totally geodesic subvariety of positive dimension. Then the following are equivalent:*

(1) The map f is covering.
(2) There is a generic sequence of positive-dimensional proper totally geodesic subvarieties $Y_\ell \subset S_{\Gamma_1}$ such that each $f(Y_\ell)$ lies is a proper totally geodesic subvariety $W_\ell \subset S_{\Gamma_2}$.

The positive-dimensional hypothesis on the subvarieties of S_{Γ_1} is to rule out the trivial counterexample of points. The arithmetic ball quotients with totally geodesic subvarieties of every possible codimension are those of *simplest type*, which are constructed using hermitian forms over number fields. In this case, Theorem 8.3 reduces Question 8.2 to checking a single dimension.

Corollary 8.4 (Baldi-Miller-Stover-Ullmo [22, Cor. 1.4])**.** *If Question 8.2 has a positive answer when S_{Γ_1} is a 2-dimensional arithmetic ball quotient of simplest type, then it has a positive answer for S_{Γ_1} an arithmetic ball quotient of simplest type of any dimension $n \geq 2$.*

Remark 8.5. By [31], the first open case is of a holomorphic embedding $f : S_{\Gamma_1} \to S_{\Gamma_2}$ with $\dim S_{\Gamma_1} = 2, \dim S_{\Gamma_2} = 4$. Here Theorem 8.3 implies that it suffices to prove that there is an infinite collection of distinct totally geodesic curves $C_i \subset S_{\Gamma_1}$ such that each $f(C_i)$ lies in some strict totally geodesic subvariety $Y_i \subset S_{\Gamma_2}$.

From a dynamical point of view, the dimension at least two hypothesis in the previous results is natural and critical in that it allows one to access tools from unipotent dynamics. From a Hodge theoretic point of view, one typically works with a complex subvariety. However, there is still something one can say for closed geodesics.

Theorem 8.6 (Baldi-Miller-Stover-Ullmo [22, Thm. 1.6])**.** *Let $S = \Gamma_n \backslash \mathbb{B}^n$ be a ball quotient with an immersive holomorphic map*

$$f : S \longrightarrow \Gamma_m \backslash \mathbb{B}^m$$

such that $f(S)$ does not lie in any smaller totally geodesic subvariety of $\Gamma_m \backslash \mathbb{B}^m$. Let $\{C_\ell\}$ be a generic sequence of distinct closed geodesics in S in the sense that they are not all contained in a finite union of maximal totally geodesic subvarieties of S. Assume that for each ℓ there exits a strict complex totally geodesic subvariety $Y_\ell \subset \Gamma_m \backslash \mathbb{B}^m$ such that $f(C_\ell) \subset Y_\ell$. Then f is a totally geodesic immersion.

We decided not to present the details of the proofs, but we point out that they follow essentially from Theorem 3.25, that we proved above.

9. Beyond the Hodge locus: orbit closures and abelian differentials

In this section we present the main results of [24].

9.1. **Recap of orbit closures.** In the last 15 years, moduli spaces of translation surfaces (X, ω) have attracted a lot of attention, thanks to the work of Eskin, Filip, McMullen, Mirzakhani, Mohammadi, Möller, Wright, and many others. A variety of tools have been used to try to grasp them, mostly from dynamics, analysis, and algebraic geometry. In particular, the work of the aforementioned authors gives striking results on so-called *orbit closures* $\mathcal{N} = \overline{\mathbf{GL}_2(\mathbb{R}) \cdot (X, \omega)}$. We push the Hodge-theoretic and algebro-geometric understanding of such loci $\mathcal{N}$ one step further. Using tools from differential geometry and differential algebra we will show that one can understand the distribution of orbit closures just by knowing whether they are *typical or atypical intersections.*

We start by recalling the main protagonists of the theory of moduli spaces of translation surfaces, mostly following the recent notes [52]. Let X be an irreducible smooth projective algebraic curve over $\mathbb{C}$, or simply a compact *Riemann surface.* A *translation surface* is a pair (X, ω), where X is a compact Riemann surface and ω is a non-zero global section of the cotangent bundle of X (also known as an *abelian differential*). If $g \geq 1$ is the genus of X, then ω has $2g-2$ zeros, counted with multiplicities. Let $\mathcal{M}_g$ denote the moduli space of genus g compact Riemann surfaces and $\Omega\mathcal{M}_g \to \mathcal{M}_g$ the bundle whose fibre over $[X] \in \mathcal{M}_g$ is just the space of all (non zero) holomorphic 1-forms on X. This is the total space of the so called *Hodge bundle* (with the zero section removed). It is a rank g (algebraic) vector bundle on $\mathcal{M}_g$, naturally equipped with a stratification by smooth (not necessarily irreducible) algebraic subvarieties $\Omega\mathcal{M}_g(\kappa)$, where $\kappa = (\kappa_0, \dots, \kappa_n)$ satisfies $\sum_i \kappa_i = 2g-2$ and denotes the multiplicities of zeros of ω. We will also write $n = \text{length}(\kappa) - 1$. (In fact, to avoid orbifold issues, we will later work with finite covers of such moduli spaces).

A holomorphic 1-form ω on X can be thought of as giving a collection of charts on X to $\mathbb{C}$, such that:

- the transition maps are translations,
- the charts can ramify at finitely many points corresponding to the zeros of ω,
- they are locally given by $z \mapsto \int_{z_0}^{z} \omega$.

Such charts induce the so-called *period coordinates* on $\Omega\mathcal{M}_g(\kappa)$. Fix a base point $(X_0, \omega_0) \in \Omega\mathcal{M}_g(\kappa)$, and look at the developing map/period coordinates, where $\widetilde{-}$ denotes the universal covering

(9.1.1) $$\text{Dev} : \widetilde{\Omega\mathcal{M}_g(\kappa)} \to H^1(X_0, Z(\omega_0); \mathbb{C}).$$

Concretely, the map sends (X',ω') to $\omega' \in H^1(X',Z(\omega');\mathbb{C})$ and then applies the flat transport to obtain an element of $H^1(X_0,Z(\omega_0);\mathbb{C})$. By a theorem of Veech the map Dev is a local biholomorphic.

We also obtain an action of $\mathrm{GL}_2(\mathbb{R})$ on $\Omega\mathcal{M}_g(\kappa)$, locally given by a diagonal action on a product of copies of $\mathbb{C} \cong \mathbb{R}^2$. Such an action is real analytic and outside the realm of algebraic geometry, but it turns out to be closely related to Hodge theory.

We will study the distribution of strict orbit closures $\mathcal{N} := \overline{\mathrm{GL}_2(\mathbb{R}) \cdot (X,\omega)} \subset \Omega\mathcal{M}_g(\kappa)$, where $\overline{(\cdot)}$ denotes the topological closure, also known in the literature as *affine invariant submanifolds*. (For terminology and notations we try to follow [52].)

We start by recalling a theorem of Eskin, Filip, and Wright (see §9.2 for definitions):

Theorem 9.1 ([46, Thm. 1.5])**.** *In each (irreducible component of each) stratum $\Omega\mathcal{M}_g(\kappa)$, all but finitely many orbit closures have rank 1 and degree at most 2. In each genus there is a finite union of orbit closures $\mathcal{N}$ of rank 2 and degree 1 such that all but finitely many of the orbit closures of rank 1 and degree 2 are a codimension 2 subvariety of one of these $\mathcal{N}$.*

Our goal is to present a new and effective proof of the above, inspired by the geometric ZP Theorem 3.17.

9.2. **Structure of orbit closures.** We recall a special case of a theorem of Eskin, Mirzakhani, and Mohammadi [47, Thm. 2.1] see also [52, Thm. 1]:

Theorem 9.2 (Linearity/topological rigidity)**.** *Each orbit closure $\mathcal{N} \subset \Omega\mathcal{M}_g(\kappa)$ is locally in period coordinates a* linear manifold*, i.e. any sufficiently small open $U \subset \mathcal{N}$ is such that* $\mathrm{Dev}(U)$ *is an open set inside a linear subspace.*

Let $\mathcal{N} \subset \Omega\mathcal{M}_g(\kappa)$ be an orbit closure, and denote by $T\mathcal{N}$ its tangent bundle. The linearity of orbit closures, as recalled in Theorem 9.2, allows one to realize $T\mathcal{N}$ as a local subsystem of H^1_{rel} whose fiber at (X,ω) is the first relative cohomology group of X with respect to $Z(\omega)$.

Remark 9.3. We pause for a moment to recall some basic facts about *relative cohomology*. Let (X,Z) be a pair of a Riemann surface and a finite set of points $Z = \{z_1,\ldots,z_n\} \subset X$ (that will usually be given by the zeros of some (non-zero) 1-form ω on X). We have a short exact sequence:

$$0 \to \tilde{H}^0(Z;\mathbb{Z}) \to H^1(X,Z;\mathbb{Z}) \to H^1(X;\mathbb{Z}) \to 0. \tag{9.2.1}$$

Above, $\tilde{H}^0(Z;\mathbb{Z})$ is the *reduced cohomology*: it denotes the group of all formal linear combinations of points in Z, modulo the element which takes each point in Z with coefficient 1 (it is the dual of the reduced homology group). We will always refer to the middle term as *relative cohomology* and the the last term as *absolute cohomology*. Note that the form ω naturally induces a class inside both $H^1(X,Z(\omega);\mathbb{Z})$ and $H^1(X;\mathbb{Z})$ by integrating over homological cycles and using homology-cohomology duality.

We have a version of (9.2.1) for flat bundles above $\Omega\mathcal{M}_g(\kappa)$ (see also [52, 3.1.13]):

$$0 \to W_0 \to H^1_{rel} \to H^1 \to 0. \tag{9.2.2}$$

The notation is justified by the fact that W_0 denotes the weight-zero local subsystem for the mixed VHS structure on H^1_{rel}.

As a consequence we have another short exact sequence of real bundles above $\mathcal{N}$:

$$0 \to W_0(T\mathcal{N}) \to T\mathcal{N} \to H^1(T\mathcal{N}) \to 0. \tag{9.2.3}$$

To explain the notation: $H^1(T\mathcal{N})$ is just the image of $T\mathcal{N} \subset H^1_{rel}$ in H^1 (the absolute cohomology), and $W_0(T\mathcal{N}) = T\mathcal{N} \cap W_0$.

Let $T \subset \mathbf{SL}_2(\mathbb{R})$ denote the upper triangular group.

Theorem 9.4 (Isolation property, [47, Thm. 2.3], see also [52, Thm. 4.1.8.]). *For every sequence of orbit closures $(\mathcal{N}_i)$ there are T-invariant measures μ_i, and after passing to a subsequence $(\mathcal{N}_i, \mu_i)$ there is another linear immersed submanifold $\mathcal{M}$, with finite T-invariant measure μ such that $\mathcal{N}_i \subset \mathcal{M}$ and $\mu_i \to \mu$.*

Note that the above result implies the same statement with T replaced by $\mathbf{SL}_2(\mathbb{R})$, which is how we will use it; see [52, Rem. 4.1.9 (ii)].

We can now recall the following theorem of Filip [51] (which builds on the earlier work [50], as well as earlier work of Wright [112] and Möller [89]), see also [52, Thm. 2 and Thm. 4.4.2]. We denote by $K = K_{\mathcal{N}} \subset \mathbb{R}$ the smallest field over which $\mathcal{N}$ is K-linear in the sense of Theorem 9.2. We have:

Theorem 9.5 (S. Filip). *Let $\mathcal{N} \subset \Omega\mathcal{M}_g(\kappa)$ be an orbit closure. There exists, up to isogeny, a factor $\mathcal{F} \subset \mathcal{J}$ of the relative Jacobian over $\mathcal{N}$, and a subgroup $\mathcal{S}$ of the free abelian group on the zeros of 1-forms, such that:*

- *(1) $\mathcal{F}$ admits real multiplication by K (i.e. $K = \mathrm{End}^0(\mathcal{F}_{\mathcal{N}})$), which is, in particular, a totally real number field;*
- *(2) The Abel-Jacobi map, possibly twisted by real multiplication, $AJ : \mathcal{S} \to \mathcal{F}$ assumes torsion values;*
- *(3) For each $(X, \omega) \in \mathcal{N}$, ω is an eigenform for the real multiplication on the fibre of $\mathcal{F}$ above X (see also* [52, §4.2.17]*).*

Furthermore, these conditions, together with a dimension bound, characterize $\mathcal{N}$.

(See [52, §4.5] for more details on the *twisted torsion* condition.)

Corollary 9.6. *Each orbit closure is algebraic, that is a Zariski closed subvariety of $\Omega\mathcal{M}_g$. In fact, orbit closures are defined over $\overline{\mathbb{Q}}$, and their Galois conjugates are again orbit closures.*

From Theorem 9.5, one may introduce *real multiplication, torsion and eigenform* invariants (see Section 9.4, for a geometric description). In particular, we write

- $r = r_{\mathcal{N}}$, the *(cylinder) rank* $\frac{1}{2}\dim H^1(T\mathcal{N})$ (which is in fact an integer, since $T\mathcal{N}$ is known to be symplectic thanks to [4]);
- $d = d_{\mathcal{N}}$, to simply denote the degree of the number field defining the linear equations of $\mathcal{N}$ (that is $[K : \mathbb{Q}]$);
- And finally $t = t_{\mathcal{N}}$ the *torsion corank* is $\dim W_0(T\mathcal{N})$.

With these notations, the dimension bound appearing at the end of Theorem 9.5 is just

$$\dim \mathcal{N} = 2r_{\mathcal{N}} + t_{\mathcal{N}}.$$

Remark 9.7. Combing Corollary 9.6 and Theorem 9.4, we observe that the Zariski closure of a collection of orbit closures is a finite union of orbit closures

9.3. **A new take on the Eskin-Filip-Wright finiteness.** We study orbit closures via various *period maps.* Denote by $\mathcal{A}_g$ the moduli space of principally polarized abelian varieties of dimension g, and by j the Torelli morphism $j : \mathcal{M}_g \to \mathcal{A}_g,\ X \mapsto (J(X), \theta_X)$ which associates to a genus g curve X its Jacobian $J(X)$ (naturally equipped with its principal polarization θ_X). For every n, let $\mathcal{A}_{g,n} \to \mathcal{A}_g$ denote the fibration whose fibre over a point $[A] \in \mathcal{A}_g$ is $\mathrm{Sym}^{[n]} A$ (the unordered n-tuples of points of A).

The varieties $\mathcal{A}_g$ and $\mathcal{A}_{g,n}$ can be interpreted as mixed period spaces (or mixed Shimura varieties). For $\mathcal{A}_g$ this is the standard Siegel space description, and a detailed construction of $\mathcal{A}_{g,1}$ as a mixed Shimura variety appears in [56, §2]. In general one has a natural bijection

$$\mathcal{A}_{g,n} \simeq \{(H, E) : H \in \mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g,\ \ E \in \mathrm{Ext}^1_{\mathrm{MHS}}(H, \mathbb{Z}^n))\},$$

where we view $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g \simeq \mathcal{A}_g$ as the moduli space for principally polarized pure weight one $\mathbb{Z}$-Hodge structures, and we view $\mathbb{Z}^n$ as a pure weight zero Hodge structure; see [52, 4.3.14]. One obtains an algebraic mixed period map $\varphi : \Omega\mathcal{M}_g(\kappa) \to \mathcal{A}_{g,n}$ which sends (X, ω) to the isomorphism class of the extension. The data of φ is equivalent to the data of the mixed $\mathbb{Z}$VHS

H^1_{rel} (at least when viewed as a map of orbifolds). Note that φ is quasi-finite, cf. the discussion at the end of [77, §3.4]. We will typically work with the factorization

$$\Omega\mathcal{M}_g(\kappa) \xrightarrow{\Omega\varphi} \Omega\mathcal{A}_{g,n} \to \mathcal{A}_{g,n},$$

of φ, where the first arrow is obtained by seeing the differential ω on the Jacobian of X. (Where $\Omega\mathcal{A}_{g,n}$ denotes the total space of the Hodge bundle above $\mathcal{A}_{g,n}$, without the zero section.)

Remark 9.8. With the language of (mixed) Shimura varieties one can equivalently describe $\mathcal{A}_{g,n}$ as follows. The classifying space for the extensions of a weight one $\mathbb{Z}$VHS of dimension 2g by $\mathbb{Z}(0)$ is the mixed Shimura variety $\mathbb{A}_g$. It is the universal principally polarized abelian variety of dimension g over $\mathcal{A}_g$. Finally

$$\mathcal{A}_{g,n} \cong \mathbb{A}_g \times_{\mathcal{A}_g} \mathbb{A}_g \times_{\mathcal{A}_g} \cdots \times_{\mathcal{A}_g} \mathbb{A}_g / S_n$$

(product of n factors modded out by the action of S_n, the symmetric groups on n elements). In fact, $\mathcal{A}_{g,n}$ is a *mixed Shimura variety of Kuga type*, see [56, §2] for more details on this. This means that, with the notation from Hodge theory, that $W_{-2} = 0$ and W_{-1} is the unipotent radical of its generic Mumford–Tate group, which is simply given by $\mathbb{G}_a^{2g\cdot n}$.

9.4. **Orbit closures as (a)typical intersections, inspired by Filip.** In this section we give a geometric reinterpretation of the conditions appearing in Theorem 9.5. Cf. [52] and implicitly [51].

We will think of an orbit closure $\mathcal{N}$ (with associated invariants r, d, t as in Section 9.2) as the reduced subvariety underlying an irreducible component of $(\Omega\varphi)^{-1}(E[\mathcal{N}])$, where $E[\mathcal{N}] \subset \Omega\mathcal{A}_{g,n}$ is an algebraic subvariety constructed using the conditions appearing in Theorem 9.5. In accordance with the structure of Theorem 9.5, the variety $E[\mathcal{N}]$ is constructed in three stages, depicted in the following diagram:

$$\begin{tikzcd} \mathcal{N} \arrow[r] \arrow[d, hook] & E[\mathcal{N}] \arrow[r, two heads] \arrow[d, hook] & M[\mathcal{N}] \arrow[r, two heads] \arrow[d, hook] & S[\mathcal{N}] \arrow[d, hook] \\ \Omega\mathcal{M}_g(\kappa) \arrow[r] & \Omega\mathcal{A}_{g,n} \arrow[r, two heads] & \mathcal{A}_{g,n} \arrow[r, two heads] & \mathcal{A}_g \end{tikzcd} . \tag{9.4.1}$$

We now give details and definitions regarding the objects appearing in the above diagram.

(1) Let $S[\mathcal{N}] \subset \mathcal{A}_g$ be the smallest weakly special subvariety of $\mathcal{A}_g$ containing the image of $\mathcal{N}$. It lies in the locus $R[\mathcal{N}]$ of g-dimensional principally polarized abelian varieties that have real multiplication *of the same type of* $\mathcal{N}$ (in the sense of Theorem 9.5). The latter is a special subvariety of $\mathcal{A}_g$ whose the generic Mumford-Tate group is isomorphic to $\left(\mathrm{Res}_{K/\mathbb{Q}} \mathrm{GSp}_{2r,K}\right) \times \mathrm{GSp}_{2(g-dr)}$.
(2) Let $M[\mathcal{N}] \subset \mathcal{A}_{g,n}$ be the smallest weakly special subvariety of $\mathcal{A}_{g,n}$ containing $\mathcal{N}$. It lies in $T[\mathcal{N}]$, the bundle over $R[\mathcal{N}]$ consisting of n-tuples of points satisfying the same torsion conditions as $\mathcal{N}$.
(3) Let $E[\mathcal{N}]$ be the the bundle of 1-forms over $M[\mathcal{N}]$ inside in the K-eigenspace containing the restriction of the section ω to $\mathcal{N}$.

The notation $S[\mathcal{N}]$ is chosen as we think of the variety $S[\mathcal{N}]$ as giving the *Shimura condition*. Similarly, $M[\mathcal{N}]$ gives the *mixed-Shimura condition*, and $E[\mathcal{N}]$ includes also the eigenform condition in addition to the previous two.

The conditions $R[\mathcal{N}]$ and $T[\mathcal{N}]$ are simply geometric reinterpretations[6] of the first two items in Theorem 9.5, whereas the conditions $S[\mathcal{N}]$ and $M[\mathcal{N}]$ may in principle be stronger. The condition $E[\mathcal{N}]$ corresponds to the third condition in Theorem 9.5. One can think of $M[\mathcal{N}]$ as a

[6]We stress here the fact that the first two conditions of Theorem 9.5 are written in terms of rational Hodge structures (as already noticed and used by Filip). At first sight, the *twist by real multiplication* appearing in the condition on the Abel-Jacobi variety could look like a statement about K-Hodge structures, but this becomes a $\mathbb{Q}$-condition after including Galois conjugate conditions.

variety that records all mixed Hodge-theoretic conditions present on $\mathcal{N}$, and $E[\mathcal{N}]$ as a variety that records, in addition to $M[\mathcal{N}]$, the third condition in Theorem 9.5.

Definition 9.9 (Relative (a)typicality — Intersection theoretic version)**.** *Let $\mathcal{M}$ be a orbit closure in some stratum $\Omega\mathcal{M}_g(\kappa)$ (possibly equal to component of the latter). Then an orbit closure $\mathcal{N} \subset \mathcal{M}$ is said* (intersection theoretically) atypical *(relative to $\mathcal{M}$) if*

$$\operatorname{codim}_{E[\mathcal{M}]} \mathcal{N} < \operatorname{codim}_{E[\mathcal{M}]}(E[\mathcal{N}]) + \operatorname{codim}_{E[\mathcal{M}]}(\mathcal{M}),$$

and (intersection theoretically) typical *(relative to $\mathcal{M}$) otherwise.*

The codimensions appearing in Definition 9.9 are formal, but one can make sense of them by taking the image of $\mathcal{N}$ in $\Omega\mathcal{A}_{g,n}$. For most of this section, we will just work with the intersection-theoretic notion of (a)typical.

We are now ready to prove our main result. That is to give a new proof, in our unifying setting, of the finiteness parts of Theorem 9.1. More precisely, with the vocabulary of Definition 9.9:

Theorem 9.10 (Baldi-Urbanik [24, Thm. 6.5])**.** *Let $\mathcal{M} \subset \Omega\mathcal{M}_g$ be an orbit closure (possibly equal to an irreducible component of some stratum $\Omega\mathcal{M}_g(\kappa)$). Then $\mathcal{M}$ contains at most finitely many (strict) orbit closures $\mathcal{N}$ that are:*

(1) Maximal, i.e. the only suborbit closures of $\mathcal{M}$ containing $\mathcal{N}$ are $\mathcal{N}$ and $\mathcal{M}$;
(2) Atypical (relative to $\mathcal{M}$) in the sense of Definition 9.9.

In particular in each (connected component of each) stratum $\Omega\mathcal{M}_g(\kappa)$, all but finitely many orbit closures have rank 1 and degree at most 2.

For simplicity we treat only the case $\mathcal{M} = \Omega\mathcal{M}_g(\kappa)$ some fixed orbit closure. We fix a base point $s_0 := (X, \omega) \in \mathcal{M}$. We write $(\mathbf{H}, D^0 \subset \check{D}^0) = (\mathbf{H}_{\mathcal{M}}, D^0_{\mathcal{M}} \subset \check{D}^0_{\mathcal{M}})$ for the associated weak Shimura (or Hodge) datum associated to $\mathcal{M}$. For example $\mathbf{H} = \mathrm{Sp}_{2g} \ltimes \mathbb{G}_{\mathrm{a}}^{2g \cdot n}$ if $\mathcal{M} = \Omega\mathcal{M}_g(\kappa)$, where $n = |\kappa| - 1$.

9.4.1. *Rephrasing the atypical condition in the period torsor.* We now let $(\mathcal{H}, \nabla)$ be the algebraic vector bundle with regular singular connection associated to $\mathbb{V} = H^1_{\mathrm{rel}}$, and consider the bundle P and map ν constructed in §4.3.1. The graded quotient of $\mathcal{H}$ which corresponds to H^1_{abs} we denote $\mathcal{H}_{\mathrm{abs}}$. Then ω is an algebraic section of $\mathcal{H}_{\mathrm{abs}}$, and we then obtain any algebraic map $r : P \to \mathcal{H}_{abs,s_0}$ given by

$$[\eta \in \mathrm{Hom}(\mathcal{H}_s, \mathcal{H}_{s_0})] \mapsto \eta(\omega_s).$$

Lemma 9.11. *The map r lands inside a unique eigenspace $E_{\mathcal{M}} = E_{\mathcal{M},s_0} \subset \mathcal{H}_{abs,s_0}$ for the field of real multiplication $K_{\mathcal{M}}$ associated to $\mathcal{M}$. The complex algebraic variety $(E_{\mathcal{M}} \setminus \{0\}) \times \check{D}^0$ consists of a single $\mathbf{H}(\mathbb{C})$-orbit, and the map $r \times \nu : P \to (E_{\mathcal{M}} \setminus \{0\}) \times \check{D}^0$ is $\mathbf{H}(\mathbb{C})$-equivariant and surjective.*

Proof. Let $\mathcal{L} \subset P$ be a leaf coming from the rational structure of the underlying local system $\mathbb{V}$. Then from the third part of Theorem 9.5, the image $r(\mathcal{L})$ lies in a $K_{\mathcal{M}}$-eigenspace $E_{\mathcal{M}} = E_{\mathcal{M},s_0}$. Because $\mathcal{L}$ is Zariski dense in P as a consequence of the previous discussion on leaves, we have $r^{-1}(\overline{r(\mathcal{L})}^{\mathrm{Zar}}) = P$ and hence $r(P) \subset E_{\mathcal{M}}$.

It is clear by construction that the map is $\mathbf{H}(\mathbb{C})$-invariant, so surjectivity will follow if we can show that in fact $(E_{\mathcal{M}} \setminus \{0\}) \times \check{D}^0$ is an orbit of $\mathbf{H}(\mathbb{C})$. This is easy in the case $\mathcal{M}$ equals a stratum and more delicate in general. $\square$

Given a suborbit closure $\mathcal{N} \subset \mathcal{M}$, we will write $Z_{\mathcal{N}} := E_{\mathcal{N}} \times \check{D}^0_{\mathcal{N}}$ for the analogous subvariety of $E_{\mathcal{M}} \times \check{D}^0$ associated to $\mathcal{N}$. It is well-defined up to the action of the monodromy group $\Gamma_{\mathcal{M}}$.

Corollary 9.12. *We have* $\dim(E_{\mathcal{N}} \times \check{D}^0_{\mathcal{N}}) = \dim E[\mathcal{N}]$. *If $\mathcal{N}$ is an intersection-theoretically atypical suborbit closure of $\mathcal{M}$, then we have the following inequality of formal codimensions:*

$$\operatorname{codim}_P \mathcal{N} < \operatorname{codim}_P (r \times \nu)^{-1}(E_{\mathcal{N}} \times \check{D}^0_{\mathcal{N}}) + \operatorname{codim}_P \mathcal{M}.$$

Proof. The first equality of dimensions is formal and can be seen by translating the abelian variety data into its Hodge-theoretic incarnation. For the second we use the fact that $P \to (E_{\mathcal{M}} \backslash \{0\}) \times \check{D}^0$ has constant fibre dimension to rewrite the inequality in Definition 9.9. The key fact is that

$$\begin{aligned}\operatorname{codim}_P (r \times \nu)^{-1}(E_{\mathcal{N}} \times \check{D}^0_{\mathcal{N}}) &= \dim[E_{\mathcal{M}} \times \check{D}^0_{\mathcal{M}}] - \dim[E_{\mathcal{N}} \times \check{D}^0_{\mathcal{N}}] \\ &= \dim E[\mathcal{M}] - \dim E[\mathcal{N}].\end{aligned}$$

□

9.4.2. *Over-parametrization.* In this section, we describe a family of subvarieties of P that (over) parametrizes all the data that can give rise to suborbit closures of $\mathcal{M}$, regardless whether such an orbit closure is typical or atypical.

Lemma 9.13. *There exists an algebraic family $f : \mathcal{Z} \to \mathcal{Y}$ of subvarieties of $Z_{\mathcal{M}} = E_{\mathcal{M}} \times \check{D}^0$ such that all subvarieties $Z_{\mathcal{N}} \subset Z_{\mathcal{M}}$ associated to suborbit closures $\mathcal{N} \subset \mathcal{M}$ arise as a fibre $f^{-1}(y)$ for some $y \in \mathcal{Y}$.*

Proof. It suffices to handle the two "factors" separately. In particular, since $E_{\mathcal{N}} \subset E_{\mathcal{M}}$ is an inclusion of linear subspaces, it is clear all possible choices for such a subspace are parameterized by a union of Grassmannian varieties. It thus suffices to show that all weakly special subdomains of $\check{D}^0$ belong to a common algebraic family, which is a consequence of Lemma 9.14 below. □

Let $(\mathbf{H}_{\mathcal{M}}, \check{D}^0_{\mathcal{M}})$ be the monodromy datum associated to $(\mathcal{M}, \mathbb{V}_{|\mathcal{M}})$ (where $\mathbb{V}$ is the standard $\mathbb{Z}$VMHS on some stratum containing $\mathcal{M}$).

Lemma 9.14. *There is an algebraic family (over a disconnected base) $f : \mathcal{Z} \to \mathcal{Y}$ of subvarieties of $\check{D}^0_{\mathcal{M}}$ such that, for every weakly special sub-datum $(\mathbf{H}_i, \check{D}_i)$, $\check{D}_i$ appears as a fibre of f.*

Proof. This is a consequence of [55, Lem. 12.3] and [56, Sec. 8.2]. In particular, it is enough to take as base $\mathcal{Y}$ the disjoint union of finitely many copies of $\mathcal{G} \times \check{D}^0_{\mathcal{M}}$, where the group $\mathcal{G}$ is defined right above [55, Lem. 12.3]. The algebraic family obtained parametrizes orbits at some point $x \in \check{D}^0_{\mathcal{M}}$ of some $\mathcal{G}(\mathbb{C})$-translate of a finite set of representatives of weakly special subdomains of $\check{D}^0_{\mathcal{M}}$. (In the pure case a stronger result can be found for example in [109, §4.2].) □

We note that we have already proved the same statement in a more general setting mixed Hodge theoretic setting (see especially Proposition 4.9)

In what follows we abusively write $f : \mathcal{Y} \to \mathcal{Z}$ for the pullback of the family in Lemma 9.13 to P along the map $r \times \nu$. We write $f^{(j)} : \mathcal{Y}^{(j)} \to \mathcal{Z}^{(j)}$ for the subfamily where the fibres have dimension j.

9.4.3. *Sketch of the proof of Theorem 9.10.* Let $(\mathcal{N}_i)_{i \in \mathbb{N}}$ be an infinite sequence of intersection theoretically atypical (relatively to $\mathcal{M}$) suborbit closures that don't lie in any other suborbit closure. The first input is as the beginning of [46, Proof of Thm. 1.5]: the Zariski closure of their union is a finite union of orbit closures (cf. Remark 9.7), so it is enough to show that they are not Zariski dense in $\mathcal{M}$. We may fix the dimension e of the suborbit closures $\mathcal{N}_i$ we consider; it suffices to prove the theorem for each e separately. Similarly, we may fix the dimension j of the inverse images $(r \times \nu)^{-1}(E_{\mathcal{N}_i} \times \check{D}^0_{\mathcal{N}_i})$.

Now let us apply Theorem 4.6 to our situation using the family $f^{(j)}$. The hypotheses are satisfied by Corollary 9.12, which gives the atypicality inequality with $e = j - \dim \mathbf{H}_{\mathcal{M}}$, and Proposition 3.16 which produces the desired weakly special families. We therefore obtain finitely many families $\{h_i : C_i \to B_i\}_{i=1}^m$ of strict weakly special subvarieties of S such that each $\mathcal{N}_i$ maps into a fibre of one of the h_i, and each map $C_i \to \mathcal{M}$ is quasi-finite. The proof is then completed by the following:

Lemma 9.15. *Let $h : C \to B$ be an algebraic family of subvarieties of $\mathcal{M}$ such that $\pi : C \to \mathcal{M}$ is quasi-finite. Then only finitely many fibres of h contain a maximal suborbit closure of $\mathcal{M}$.*

Proof. Let $I = (\mathcal{N}_i)_{i=1}^{\infty}$ be a sequence of orbit closures appearing in the fibres of h. Thanks to Theorem 9.4, after passing to a subsequence there is an orbit closure $\mathcal{T}$ containing the orbit closures associated to I such that the sequence of measures associated to I converges to $\mu_{\mathcal{T}}$. Fix an open analytic neighbourhood $O \subset \mathcal{T}$, chosen small enough so that $\pi^{-1}(O) = O_1 \sqcup \cdots \sqcup O_\ell$ is a disjoint union of analytic neighbourhoods of $\pi^{-1}(\mathcal{T})$ which map bijectively onto O. Then by equidistribution O intersects infinitely many orbit closures arising from the fibres of h, hence one of the O_i, say it is O_1, must satisfy the property that the fibres of h above $h(O_1)$ contain infinitely many orbit closures.

Applying Theorem 9.4 again to the subsequence of orbit closures in fibres lying above $h(O_1)$, and pulling back along $C \to \mathcal{M}$, we obtain an algebraic subvariety of C which projects to an algebraically constructible set contained in O_1. After taking O_1 small enough this implies that the projection consists of points, so in fact the orbit closures of interest lie in a single fibre. □

References

[1] Y. André. *G-functions and geometry*, volume E13 of *Aspects Math.* Wiesbaden etc.: Friedr. Vieweg &| Sohn, 1989.

[2] Y. André. Mumford-Tate groups of mixed Hodge structures and the theorem of the fixed part. *Compositio Math.*, 82(1):1–24, 1992.

[3] Y. André. Finiteness of couples of singular modular invariants on a plane non-modular algebraic curve. *J. Reine Angew. Math.*, 505:203–208, 1998.

[4] A. Avila, A. Eskin, and M. Möller. Symplectic and isometric $\mathrm{SL}(2,\mathbb{R})$-invariant subbundles of the Hodge bundle. *J. Reine Angew. Math.*, 732:1–20, 2017.

[5] U. Bader, D. Fisher, N. Miller, and M. Stover. Arithmeticity, superrigidity, and totally geodesic submanifolds. *Ann. of Math. (2)*, 193(3):837–861, 2021.

[6] U. Bader, D. Fisher, N. Miller, and M. Stover. Arithmeticity, superrigidity and totally geodesic submanifolds of complex hyperbolic manifolds. *Invent. Math.*, 233(1):169–222, 2023.

[7] W. L. Baily, Jr. and A. Borel. Compactification of arithmetic quotients of bounded symmetric domains. *Ann. of Math. (2)*, 84:442–528, 1966.

[8] B. Bakker. A short proof of a conjecture of Matsushita. *arXiv e-prints*, page arXiv:2209.00604, Sept. 2022.

[9] B. Bakker, Y. Brunebarbe, B. Klingler, and J. Tsimerman. Definability of mixed period maps. *J. Eur. Math. Soc. (JEMS)*, 26(6):2191–2209, 2024.

[10] B. Bakker, Y. Brunebarbe, and J. Tsimerman. o-minimal GAGA and a conjecture of Griffiths. *Invent. Math.*, 232(1):163–228, 2023.

[11] B. Bakker, Y. Brunebarbe, and J. Tsimerman. Quasi-projectivity of images of mixed period maps. *J. Reine Angew. Math.*, 804:197–219, 2023.

[12] B. Bakker, Y. Brunebarbe, and J. Tsimerman. The linear Shafarevich conjecture for quasiprojective varieties and algebraicity of Shafarevich morphisms. *arXiv e-prints*, page arXiv:2408.16441, Aug. 2024.

[13] B. Bakker, B. Klingler, and J. Tsimerman. Tame topology of arithmetic quotients and algebraicity of Hodge loci. *J. Am. Math. Soc.*, 33(4):917–939, 2020.

[14] B. Bakker, B. Klingler, and J. Tsimerman. Erratum to: “Tame topology of arithmetic quotients and algebraicity of Hodge loci”. *J. Am. Math. Soc.*, 36(4):1305–1308, 2023.

[15] B. Bakker and J. Tsimerman. The Ax–Schanuel conjecture for variations of Hodge structures. *Invent. Math.*, 217(1):77–94, 2019.

[16] B. Bakker and J. Tsimerman. Functional Transcendence of Periods and the Geometric André–Grothendieck Period Conjecture. *arXiv e-prints (Forum Math. Sigma, to appear)*, page arXiv:2208.05182, Aug. 2022.

[17] G. Baldi. Hodge theory and o-minimality at CIRM. *arXiv e-prints*, page arXiv:2502.03071, Feb. 2025.

[18] G. Baldi. What makes an algebraic curve special? *arXiv e-prints*, page arXiv:2502.06366, Feb. 2025.

[19] G. Baldi, B. Klingler, and E. Ullmo. On the geometric Zilber-Pink theorem and the Lawrence-Venkatesh method. *Expo. Math.*, 41(3):718–722, 2023.

[20] G. Baldi, B. Klingler, and E. Ullmo. Non-density of the exceptional components of the Noether–Lefschetz locus. *International Mathematics Research Notices*, 2024(21):13642–13650, November 2024.

[21] G. Baldi, B. Klingler, and E. Ullmo. On the distribution of the Hodge locus. *Invent. Math.*, 235(2):441–487, 2024.

[22] G. Baldi, N. Miller, M. Stover, and E. Ullmo. Rich representations and superrigidity. *arXiv e-prints (Ergodic Theory Dynam. Systems, published online)*, page arXiv:2402.03601, Feb. 2024.

[23] G. Baldi and E. Ullmo. Special subvarieties of non-arithmetic ball quotients and Hodge theory. *Annals of Mathematics*, 197(1):159 – 220, 2023.

[24] G. Baldi and D. Urbanik. Effective atypical intersections and applications to orbit closures. *arXiv e-prints*, page arXiv:2406.16628, June 2024.

[25] G. Binyamini. Log-Noetherian functions. *arXiv e-prints*, page arXiv:2405.16963, May 2024.

[26] D. Blázquez-Sanz, G. Casale, J. Freitag, and J. Nagloo. A differential approach to the Ax-Schanuel, I. *arXiv e-prints*, page arXiv:2102.03384, Feb. 2021.

[27] E. Bombieri, D. Masser, and U. Zannier. Intersecting a curve with algebraic subgroups of multiplicative groups. *Int. Math. Res. Not.*, 1999(20):1119–1140, 1999.

[28] J. Brevik and S. Nollet. Developments in Noether-Lefschetz theory. In *Hodge theory, complex geometry, and representation theory. NSF-CBMS regional conference in mathematics. Hodge theory, complex geometry, and representation theory, Fort Worth, TX, USA, June 18, 2012*, pages 21–50. Providence, RI: American Mathematical Society (AMS), 2014.

[29] P. Brosnan, G. Pearlstein, and C. Schnell. The locus of Hodge classes in an admissible variation of mixed Hodge structure. *C. R., Math., Acad. Sci. Paris*, 348(11-12):657–660, 2010.

[30] P. Candelas, X. de la Ossa, M. Elmi, and D. van Straten. A one parameter family of Calabi-Yau manifolds with attractor points of rank two. *J. High Energy Phys.*, 2020(10):74, 2020. Id/No 202.

[31] H.-D. Cao and N. Mok. Holomorphic immersions between compact hyperbolic space forms. *Invent. Math.*, 100(1):49–61, 1990.

[32] J. Carlson, M. Green, P. Griffiths, and J. Harris. Infinitesimal variations of Hodge structure. I. *Compos. Math.*, 50:109–205, 1983.
[33] E. Cattani, P. Deligne, and A. Kaplan. On the locus of Hodge classes. *J. Amer. Math. Soc.*, 8(2):483–506, 1995.
[34] C.-L. Chai. Density of members with extra Hodge cycles in a family of Hodge structures. *Asian J. Math.*, 2(3):405–418, 1998.
[35] K. C. T. Chiu. Ax-Schanuel for variations of mixed Hodge structures. *Math. Ann.*, 391(2):1681–1710, 2025.
[36] C. Ciliberto, J. Harris, and R. Miranda. General components of the Noether-Lefschetz locus and their density in the space of all surfaces. *Math. Ann.*, 282(4):667–680, 1988.
[37] C. Ciliberto and A. F. Lopez. On the existence of components of the Noether-Lefschetz locus with given codimension. *Manuscr. Math.*, 73(4):341–357, 1991.
[38] C. Daw and J. Ren. Applications of the hyperbolic Ax-Schanuel conjecture. *Compos. Math.*, 154(9):1843–1888, 2018.
[39] A. J. de Jong. Beyond the André-Oort conjecture. *6 pages note, communicated to the second author in June 2021*, 2004.
[40] P. Deligne. *Equations différentielles à points singuliers réguliers*, volume 163 of *Lect. Notes Math.* Springer, Cham, 1970.
[41] P. Deligne. Théorie de Hodge. II. *Inst. Hautes Études Sci. Publ. Math.*, 40:5–57, 1971.
[42] P. Deligne, J. S. Milne, A. Ogus, and K. yen Shih. *Hodge cycles, motives, and Shimura varieties*, volume 900. Springer, Cham, 1982.
[43] P. Deligne and G. D. Mostow. Monodromy of hypergeometric functions and nonlattice integral monodromy. *Inst. Hautes Études Sci. Publ. Math.*, (63):5–89, 1986.
[44] E. B. Dynkin. Maximal subgroups of the classical groups. *Transl., Ser. 2, Am. Math. Soc.*, 6:245–378, 1957.
[45] S. J. Edixhoven, B. J. J. Moonen, and F. Oort. Open problems in algebraic geometry. *Bull. Sci. Math.*, 125(1):1–22, 2001.
[46] A. Eskin, S. Filip, and A. Wright. The algebraic hull of the Kontsevich-Zorich cocycle. *Ann. Math. (2)*, 188(1):281–313, 2018.
[47] A. Eskin, M. Mirzakhani, and A. Mohammadi. Isolation, equidistribution, and orbit closures for the $\mathrm{SL}(2,\mathbb{R})$ action on moduli space. *Ann. Math. (2)*, 182(2):673–721, 2015.
[48] H. Esnault and M. Groechenig. Cohomologically rigid local systems and integrality. *Selecta Math. (N.S.)*, 24(5):4279–4292, 2018.
[49] S. Eterović and T. Scanlon. Likely intersections. *arXiv e-prints*, page arXiv:2211.10592, Nov. 2022.
[50] S. Filip. Semisimplicity and rigidity of the Kontsevich-Zorich cocycle. *Invent. Math.*, 205(3):617–670, 2016.
[51] S. Filip. Splitting mixed Hodge structures over affine invariant manifolds. *Ann. Math. (2)*, 183(2):681–713, 2016.
[52] S. Filip. Translation surfaces: dynamics and Hodge theory. *EMS Surv. Math. Sci.*, 11(1):63–151, 2024.
[53] D. Fisher. Superrigidity, arithmeticity, normal subgroups: results, ramifications, and directions. In *Dynamics, geometry, number theory. The impact of Margulis on modern mathematics*, pages 9–46. Chicago, IL: The University of Chicago Press, 2022.
[54] J. Fresán, Y. H. J. Lam, and Y. Qin. On Siegel's problem and Dwork's conjecture for G-functions. *arXiv e-prints*, page arXiv:2502.02147, Feb. 2025.
[55] Z. Gao. Towards the André-Oort conjecture for mixed Shimura varieties: the Ax-Lindemann theorem and lower bounds for Galois orbits of special points. *J. Reine Angew. Math.*, 732:85–146, 2017.
[56] Z. Gao. Mixed Ax-Schanuel for the universal abelian varieties and some applications. *Compos. Math.*, 156(11):2263–2297, 2020.
[57] Z. Gao. Recent developments of the Uniform Mordell-Lang Conjecture. *arXiv e-prints*, page arXiv:2104.03431, Apr. 2021.
[58] Z. Gao and B. Klingler. The Ax-Schanuel conjecture for variations of mixed Hodge structures. *Math. Ann.*, 388(4):3847–3895, 2024.
[59] Z. Gao and S.-W. Zhang. Heights and periods of algebraic cycles in families. *arXiv e-prints*, page arXiv:2407.01304, July 2024.
[60] H. Garland and M. S. Raghunathan. Fundamental domains for lattices in (R-)rank 1 semisimple Lie groups. *Ann. of Math. (2)*, 92:279–326, 1970.
[61] M. Green, P. Griffiths, and M. Kerr. *Mumford-Tate groups and domains*, volume 183 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2012. Their geometry and arithmetic.
[62] M. L. Green. Koszul cohomology and the geometry of projective varieties. II. *J. Differ. Geom.*, 20:279–289, 1984.
[63] M. L. Green. A new proof of the explicit Noether-Lefschetz theorem. *J. Differ. Geom.*, 27(1):155–159, 1988.
[64] M. L. Green. Components of maximal dimension in the Noether-Lefschetz locus. *J. Differ. Geom.*, 29(2):295–302, 1989.
[65] P. Griffiths and J. Harris. On the Noether-Lefschetz theorem and some remarks on codimension-two cycles. *Math. Ann.*, 271:31–51, 1985.

[66] P. A. Griffiths. Periods of integrals on algebraic manifolds. III: Some global differential-geometric properties of the period mapping. *Publ. Math., Inst. Hautes Étud. Sci.*, 38:125–180, 1970.
[67] T. W. Grimm and D. van de Heisteeg. Exact flux vacua, symmetries, and the structure of the landscape. *J. High Energy Phys.*, 2025(1):90, 2025. Id/No 5.
[68] S. Grushevsky, G. Mondello, R. Salvati Manni, and J. Tsimerman. Compact Subvarieties of the Moduli Space of Complex Abelian Varieties. *arXiv e-prints*, page arXiv:2404.06009, Apr. 2024.
[69] S. Gukov and C. Vafa. Rational conformal field theories and complex multiplication. *Commun. Math. Phys.*, 246(1):181–210, 2004.
[70] R. Hain. The Rank of the Normal Functions of the Ceresa and Gross–Schoen Cycles. *arXiv e-prints*, page arXiv:2408.07809, Aug. 2024.
[71] W. V. D. Hodge. The topological invariants of algebraic varieties. Proc. Intern. Congr. Math. (Cambridge, Mass., Aug. 30-Sept. 6, 1950) 1, 182-192 (1952)., 1952.
[72] M. Kashiwara. A study of variation of mixed Hodge structure. *Publ. Res. Inst. Math. Sci.*, 22:991–1024, 1986.
[73] M. Kerr and S. Tayou. On the torsion locus of the Ceresa normal function. *arXiv e-prints*, page arXiv:2406.19366, June 2024.
[74] N. Khelifa and D. Urbanik. Existence and density of typical Hodge loci. *Math. Ann.*, 391(1):819–842, 2025.
[75] B. Klingler. Hodge loci and atypical intersections: conjectures, 2017. To appear in the book Motives and Complex Multiplication.
[76] B. Klingler. Hodge theory, between algebraicity and transcendence. In *International congress of mathematicians 2022, ICM 2022, Helsinki, Finland, virtual, July 6–14, 2022. Volume 3. Sections 1–4*, pages 2250–2284. Berlin: European Mathematical Society (EMS), 2023.
[77] B. Klingler and L. A. Lerer. Abelian differentials and their periods: the bi-algebraic point of view. *arXiv e-prints. To appear in J. Eur. Math. Soc.*, page arXiv:2202.06031, Feb. 2022.
[78] K. Kodaira and D. C. Spencer. On a theorem of Lefschetz and the lemma of Enriques-Severi-Zariski. *Proc. Natl. Acad. Sci. USA*, 39:1273–1278, 1953.
[79] B. Kostant. Root systems for Levi factors and Borel-de Siebenthal theory. In *Symmetry and spaces. In Honor of Gerry Schwarz on the occasion of his 60th birthday*, pages 129–152. Basel: Birkhäuser, 2010.
[80] Y. H. J. Lam and A. Tripathy. Attractors are not algebraic. *arXiv e-prints*, page arXiv:2009.12650, Sept. 2020.
[81] A. Langer and C. Simpson. Rank 3 rigid representations of projective fundamental groups. *Compos. Math.*, 154(7):1534–1570, 2018.
[82] B. Lawrence and A. Venkatesh. Diophantine problems and p-adic period mappings. *Invent. Math.*, 221(3):893–999, 2020.
[83] A. F. Lopez. *Noether-Lefschetz theory and the Picard group of projective surfaces*, volume 438 of *Mem. Am. Math. Soc.* Providence, RI: American Mathematical Society (AMS), 1991.
[84] G. A. Margulis. *Discrete subgroups of semisimple Lie groups*, volume 17 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)].* Springer-Verlag, Berlin, 1991.
[85] E. Markman. Cycles on abelian 2n-folds of Weil type from secant sheaves on abelian n-folds. *arXiv e-prints*, page arXiv:2502.03415, Feb. 2025.
[86] A. Mohammadi and G. Margulis. Arithmeticity of hyperbolic 3-manifolds containing infinitely many totally geodesic surfaces. *Ergodic Theory Dyn. Syst.*, 42(3):1188–1219, 2022.
[87] N. Mok, J. Pila, and J. Tsimerman. Ax-Schanuel for Shimura varieties. *Ann. of Math. (2)*, 189(3):945–978, 2019.
[88] N. Mok, Y. T. Siu, and S.-K. Yeung. Geometric superrigidity. *Invent. Math.*, 113(1):57–83, 1993.
[89] M. Möller. Variations of Hodge structures of a Teichmüller curve. *J. Am. Math. Soc.*, 19(2):327–344, 2006.
[90] D. Mumford. A note of Shimura's paper: Discontinuous groups and Abelian varieties. *Math. Ann.*, 181:345–351, 1969.
[91] F. Oort. Canonical liftings and dense sets of CM-points. In *Arithmetic geometry. Proceedings of a symposium, Cortona, Arezzo, Italy, October 16–21, 1994*, pages 228–234. Cambridge: Cambridge University Press, 1997.
[92] J. Pila. O-minimality and Diophantine geometry. In *Proceedings of the International Congress of Mathematicians (ICM 2014), Seoul, Korea, August 13–21, 2014. Vol. I: Plenary lectures and ceremonies*, pages 547–572. Seoul: KM Kyung Moon Sa, 2014.
[93] J. Pila. *Point-counting and the Zilber-Pink conjecture*, volume 228 of *Camb. Tracts Math.* Cambridge: Cambridge University Press, 2022.
[94] J. Pila, A. N. Shankar, J. Tsimerman, H. Esnault, and M. Groechenig. Canonical Heights on Shimura Varieties and the André-Oort Conjecture. *arXiv e-prints*, page arXiv:2109.08788, Sept. 2021.
[95] J. Pila and U. Zannier. Rational points in periodic analytic sets and the Manin-Mumford conjecture. *Atti Accad. Naz. Lincei, Cl. Sci. Fis. Mat. Nat., IX. Ser., Rend. Lincei, Mat. Appl.*, 19(2):149–162, 2008.
[96] R. Pink. *Arithmetical compactification of mixed Shimura varieties*, volume 209 of *Bonn. Math. Schr.* Bonn: Univ. Bonn, Math.-Naturwiss. Fak., 1989.

[97] R. Pink. A combination of the conjectures of Mordell-Lang and André-Oort. In *Geometric methods in algebra and number theory*, pages 251–282. Basel: Birkhäuser, 2005.
[98] R. Richard and A. Yafaev. Generalised André-Pink-Zannier Conjecture for Shimura varieties of abelian type. *arXiv e-prints. To appear in Publ. Math. IHES*, page arXiv:2111.11216, Nov. 2021.
[99] C. Robles. Schubert varieties as variations of Hodge structure. *Sel. Math., New Ser.*, 20(3):719–768, 2014.
[100] W. Schmid. Variation of Hodge structure: The singularities of the period mapping. *Invent. Math.*, 22:211–319, 1973.
[101] C. T. Simpson. Higgs bundles and local systems. *Inst. Hautes Études Sci. Publ. Math.*, 75:5–95, 1992.
[102] Y.-T. Siu. Some recent results in complex manifold theory related to vanishing theorems for the semipositive case. Arbeitstag. Bonn 1984, Proc. Meet. Max-Planck-Inst. Math., Bonn 1984, Lect. Notes Math. 1111, 169-192., 1985.
[103] S. Tayou and N. Tholozan. Equidistribution of Hodge loci. II. *Compos. Math.*, 159(1):1–52, 2023.
[104] J. Tsimerman. The André-Oort conjecture for $\mathcal{A}_g$. *Ann. of Math. (2)*, 187(2):379–390, 2018.
[105] J. Tsimerman. Functional transcendence and arithmetic applications. In *Proceedings of the international congress of mathematicians, ICM 2018, Rio de Janeiro, Brazil, August 1–9, 2018. Volume II. Invited lectures*, pages 435–454. Hackensack, NJ: World Scientific; Rio de Janeiro: Sociedade Brasileira de Matemática (SBM), 2018.
[106] C. Voisin. Une précision concernant le théorème de Noether. (A precision concerning the theorem of Noether). *Math. Ann.*, 280(4):605–611, 1988.
[107] C. Voisin. Contrexemple à une conjecture de J. Harris. (A counterexample to a conjecture of J. Harris). *C. R. Acad. Sci., Paris, Sér. I*, 313(10):685–687, 1991.
[108] C. Voisin. *Théorie de Hodge et géométrie algébrique complexe*, volume 10 of *Cours Spéc. (Paris)*. Paris: Société Mathématique de France, 2002.
[109] C. Voisin. Hodge loci. In *Handbook of moduli. Volume III*, pages 507–546. Somerville, MA: International Press; Beijing: Higher Education Press, 2013.
[110] C. Voisin. Torsion points of sections of Lagrangian torus fibrations and the Chow ring of hyper-Kähler manifolds. In *Geometry of moduli. Proceedings of the Abel symposium, Svinøya Rorbuer, Svolvær, Norway, August 7–11, 2017*, pages 295–326. Cham: Springer, 2018.
[111] A. Weil. Remarks on the cohomology of groups. *Ann. of Math. (2)*, 80:149–157, 1964.
[112] A. Wright. The field of definition of affine invariant submanifolds of the moduli space of abelian differentials. *Geom. Topol.*, 18(3):1323–1341, 2014.
[113] U. Zannier. *Some problems of unlikely intersections in arithmetic and geometry. With appendixes by David Masser*, volume 181 of *Ann. Math. Stud.* Princeton, NJ: Princeton University Press, 2012.
[114] B. Zilber. Exponential sums equations and the Schanuel conjecture. *J. Lond. Math. Soc., II. Ser.*, 65(1):27–44, 2002.

CNRS, IMJ-PRG, Sorbonne Université, 4 place Jussieu, 75005 Paris, France
E-mail address, G. Baldi: `baldi@imj-prg.fr`