Hodge loci and atypical intersections: conjectures

Bruno Klingler

Abstract. We present a conjecture on the geometry of the Hodge locus of a (graded polarizable, admissible) variation of mixed Hodge structures over a complex smooth quasi-projective base, generalizing to this context the Zilber–Pink conjecture for mixed Shimura varieties (in particular the André–Oort conjecture).

Contents

1. Introduction: Hodge locus, atypical locus and main conjecture

1.1. Hodge locus

The general context of this paper is the study of the following geometric problem. Let k be an algebraically closed field and let $f\colon X \to S$ be a smooth morphism of quasi-projective k-varieties. Can we describe the locus of closed points $s \in S$ where the motive $[\mathcal{X}_s]$ of the fiber $\mathcal{X}_s$ is "simpler" than the motive of the fiber at a very general point? Here "simpler" means that the fiber $\mathcal{X}_s$ and its powers contain more algebraic cycles than the very general fiber and its powers. If a Tannakian formalism of k-motives were available, this would be equivalent to saying that the motivic Galois group $\mathrm{GMot}(\mathcal{X}_s)$ is smaller than the motivic Galois group of the very general fiber.

We restrict ourselves to $k = \mathbb{C}$. From now on all algebraic varieties are over $\mathbb{C}$. Following a common abuse of notation we will still denote by S

the complex analytic space S^{an} associated with an algebraic variety S; the meaning will be clear from the context. By a point of S we always understand a closed point.

We consider the Hodge incarnation of our problem. Let $\mathbb{V} \to S$ be a variation of mixed Hodge structures (VMHS) over a smooth quasi-projective variety S. In this introduction we will consider $\mathbb{Q}$VMHS (we will restrict ourselves to $\mathbb{Z}$VMHS when monodromy arguments are involved, as it simplifies the exposition). The weight filtration on $\mathbb{V}$ is denoted by $W_\bullet$ and the Hodge filtration on $\mathbb{V} \otimes_{\mathbb{Q}} \mathcal{O}_S$ by $F^\bullet$. In this paper all such variations are assumed to be *graded-polarizable and admissible.* A typical example of such a gadget is $\mathbb{V} = R^m f_* \mathbb{Q}$ for $f \colon \mathcal{X} \to S$ smooth algebraic, locally trivial for the usual topology. Precise definitions of Hodge theory are recalled in Section 2. The non-expert reader will think of the case of a (polarizable) variation of pure Hodge structures (corresponding geometrically to the case where f is moreover proper).

Replacing algebraic cycles by Hodge classes and motivated by the Hodge conjecture, one wants to understand the Hodge locus $\mathrm{HL}(S, \mathbb{V}) \subset S$, namely the subset of S of points s for which exceptional Hodge tensors for $\mathbb{V}_s$ occur.

The Tannakian formalism available for Hodge structures is particularly useful for describing $\mathrm{HL}(S, \mathbb{V})$. We recall it here, as it will be crucial for the statement of our main conjecture. For every $s \in S$, the Mumford–Tate group $\mathbf{P}_s(\mathbb{V})$ of the Hodge structure $\mathbb{V}_s$ is the Tannakian group of the Tannakian category $\langle \mathbb{V}_s^{\otimes} \rangle$ of mixed $\mathbb{Q}$-Hodge structures tensorially generated by $\mathbb{V}_s$ and its dual $\mathbb{V}_s^\vee$. Equivalently, the group $\mathbf{P}_s(\mathbb{V})$ is the stabiliser of the Hodge tensors for $\mathbb{V}_s$, *i.e.* the Hodge classes in the rational Hodge structures tensorially generated by $\mathbb{V}_s$ and its dual. This is a connected $\mathbb{Q}$-algebraic group, which is reductive if $\mathbb{V}_s$ is pure, and an extension of the reductive group $\mathbf{P}_s(\mathrm{Gr}^W_\bullet \mathbb{V})$ by a unipotent group in general (where W denotes the weight filtration on $\mathbb{V}$). A point $s \in S$ is said to be Hodge generic for $\mathbb{V}$ if $\mathbf{P}_s(\mathbb{V})$ is maximal. If S is irreducible, two Hodge generic points of S have the same Mumford–Tate group, called the generic Mumford–Tate group $\mathbf{P}_S(\mathbb{V})$ of $(S, \mathbb{V})$. The Hodge locus $\mathrm{HL}(S, \mathbb{V})$ is the subset of points of S which are not Hodge generic.

A fundamental result of Cattani–Deligne–Kaplan [8] (in the pure case, extended to the mixed case in [6]) states that $\mathrm{HL}(S, \mathbb{V})$ is a countable union of closed irreducible algebraic subvarieties of S. A *special subvariety* of $(S, \mathbb{V})$ is by definition an irreducible subvariety of S maximal among the irreducible subvarieties with a fixed generic Mumford–Tate group. Special subvarieties of dimension zero are called *special points* of $(S, \mathbb{V})$. A special point s whose Mumford–Tate group $\mathbf{P}_s(\mathbb{V})$ is a torus is called a point with complex multiplication (*CM-point*) for $(S, \mathbb{V})$.

In a nutshell, we would like to address the following vaguely stated

Question 1.1. Given a smooth quasi projective variety S, any variation of mixed Hodge structures $\mathbb{V} \to S$ produces naturally a countable collection of irreducible subvarieties of S: its special subvarieties. Can one describe the distribution of the special subvarieties strictly contained in S, in particular of the CM-points?

1.2. Zariski-closure of the Hodge locus

A first precise version of Question 1.1 would be to describe the Zariski-closure of $\mathrm{HL}(S, \mathbb{V})$, in particular to answer the following

Question 1.2. *Are there any geometric constraints on the Zariski-closure of* $\mathrm{HL}(S, \mathbb{V})$*? Can one describe the couples* $(S, \mathbb{V})$ *such that* $\mathrm{HL}(S, \mathbb{V})$ *is Zariski-dense in* S*?*

Particular cases of this problem have been classically considered by complex algebraic geometers, essentially when $\mathbb{V}$ is pure of weight 1 or 2, and using infinitesimal methods which lead to density results even for the Archimedean topology:

Example 1.3 ($\mathbb{V}$ pure of weight 1). When $\mathbb{V}$ is pure of weight 1 (and hence we are essentially considering families of Abelian varieties), Question 1.2 has been raised in a particular case in [11] and [21]. A typical result is the following. Let $S \subset \mathcal{A}_g$ be a subvariety of codimension at most g of the moduli space $\mathcal{A}_g$ of principally polarized Abelian varieties of dimension g. Then the set S_k of points $s \in S$ such that the corresponding Abelian variety A_s admits an Abelian subvariety of dimension k is dense (for the Archimedean topology) in S for any integer k between 1 and $g-1$. Let $\mathbb{V}$ be the VHS restriction to S of the Hodge incarnation $R^1 f_* \mathbb{Q}$ of the universal Abelian variety $f \colon \mathfrak{A}_g \to \mathcal{A}_g$ over $\mathcal{A}_g$. As $S_k \subset \mathrm{HL}(S, \mathbb{V})$ it follows that the set $\mathrm{HL}(S, \mathbb{V})$ is dense in S. More generally let $(\mathbf{G}, X)$ be a pure Shimura datum, $\mathrm{Sh}_K(\mathbf{G}, X)$ an associated Shimura variety, $\mathbf{H} \subset \mathbf{G}$ a $\mathbb{Q}$-reductive subgroup and $S \subset \mathrm{Sh}_K(\mathbf{G}, X)$ an algebraic subvariety. Let $\mathbb{V} \to S$ be the restriction to S of the variation of pure Hodge structures on $\mathrm{Sh}_K(\mathbf{G}, X)$ associated with any faithful algebraic representation of $\mathbf{G}$. Denote by $\mathrm{HL}(S, \mathbb{V}, \mathbf{H}) \subset \mathrm{HL}(S, \mathbb{V})$ the subset of points $s \in S$ whose Mumford–Tate group $\mathbf{P}_s(\mathbb{V})$ is $\mathbf{G}(\mathbb{Q})$-conjugated to $\mathbf{H}$. In [9] Chai defines an invariant $c(\mathbf{G}, X, \mathbf{H}) \in \mathbb{N}$, whose value is g in the example above, which has the property that $\mathrm{HL}(S, \mathbb{V}, \mathbf{H})$ is dense in S as soon as S has codimension at most $c(\mathbf{G}, X, \mathbf{H})$ in $\mathrm{Sh}_K(\mathbf{G}, X)$.

Example 1.4 ($\mathbb{V}$ pure of weight 2: Noether–Lefschetz locus). Consider the open subvariety $S \subset \mathbf{P}H^0(\mathbf{P}^3_{\mathbb{C}}, \mathcal{O}(d))$ parametrising smooth surfaces of degree d in $\mathbf{P}^3_{\mathbb{C}}$. For $d > 3$, the classical Noether theorem states that the very general surface $Y \in S$ has Picard group $\mathbb{Z}$: every curve on Y is a complete intersection of Y with another surface in $\mathbf{P}^3_{\mathbb{C}}$. The countable union $\mathrm{NL}(S)$ of algebraic subvarieties of S corresponding to surfaces with bigger Picard group is called the Noether–Lefchetz locus of S. Let $\mathbb{V} \to S$ be the VHS $R^2 f_* \mathbb{Q}$, where $f \colon \mathcal{Y} \to S$ denotes the universal family. Clearly $\mathrm{NL}(S) \subset \mathrm{HL}(S, \mathbb{V})$. In [10] Ciliberto, Harris and Miranda proved that $\mathrm{NL}(S)$ is Zariski-dense in S. Green (see [35, Prop. 5.20]) proved the stronger result that $\mathrm{NL}(S)$ is dense in S for the Archimedean topology. In particular $\mathrm{HL}(S, \mathbb{V})$ is dense in S.

Examples 1.3 and 1.4 indicate that special subvarieties for $(S, \mathbb{V})$ are quite common in general. Even if they have a Hodge theoretic significance, they are not special enough to force any global shape for the Zariski-closure of $\mathrm{HL}(S, \mathbb{V})$. Hence we cannot expect a naive answer to the naive Question 1.2.

1.3. Atypical locus

In this note we define a natural subset $S_{\mathrm{atyp}}(\mathbb{V}) \subset \mathrm{HL}(S, \mathbb{V})$: the *atypical locus* of $(S, \mathbb{V})$. While the Zariski-closure of $\mathrm{HL}(S, \mathbb{V})$ can be wild, we conjecture (Main Conjecture 1.9) that the structure of $S_{\mathrm{atyp}}(\mathbb{V})$ is simple. This generalizes the Zilber–Pink Conjecture for Shimura varieties to any (graded-polarized, admissible) $\mathbb{Q}$VMHS over any smooth quasi-projective base.

1.3.1. Hodge codimension. The crucial notion for defining the atypical locus $S_{\mathrm{atyp}}(\mathbb{V}) \subset S$ is the notion of Hodge codimension:

Definition 1.5 (Hodge codimension). Let S be an irreducible quasi-projective variety and $\mathbb{V} \to S^{\mathrm{sm}}$ a variation of mixed Hodge structures on the smooth locus S^{sm} of S. Let $\mathbf{P}_S$ be the generic Mumford–Tate group of $(S^{\mathrm{sm}}, \mathbb{V})$ and $\mathfrak{p}_S$ its Lie algebra (endowed with its canonical mixed $\mathbb{Q}$-Hodge structure, of weight ≤ 0). Define the Hodge codimension of S with respect to $\mathbb{V}$ as

$$\text{H-cd}(S, \mathbb{V}) := \dim_{\mathbb{C}}(\mathrm{Gr}_F^{-1}(\mathfrak{p}_S \otimes_{\mathbb{Q}} \mathbb{C})) - \mathrm{rk}\,\mathrm{Im}\overline{\nabla},$$

where $\overline{\nabla} \colon TS^{\mathrm{sm}} \to \mathrm{Gr}_F^{-1}(W_0\mathrm{End}\,\mathbb{V} \otimes_{\mathbb{Q}} \mathcal{O}_S)$ is the Kodaira–Spencer map of $(S, \mathbb{V})$ (see Section 2.5).

Remark 1.6. It follows from Definition 1.5 that if $Y \subset Y' \subset S$ is a pair of irreducible subvarieties and if $\mathbf{P}_Y = \mathbf{P}'_Y$ then $\text{H-cd}(Y', \mathbb{V}_{|Y'^{\mathrm{sm}}}) \leq \text{H-cd}(Y, \mathbb{V}_{|Y^{\mathrm{sm}}})$.

1.3.2. Atypical subvarieties.

Definition 1.7 (Atypical subvariety). Let S be an irreducible smooth quasi-projective variety and $\mathbb{V} \to S$ a variation of mixed Hodge structures on S. An irreducible subvariety $Y \subset S$ is said to be atypical for $(S, \mathbb{V})$ if

$$\text{H-cd}(Y, \mathbb{V}_{|Y^{\mathrm{sm}}}) < \text{H-cd}(S, \mathbb{V}). \tag{1.1}$$

We denote by $S_{\mathrm{atyp}}(\mathbb{V}) \subset S$ the subset of S given by the union of all atypical subvarieties for $(S, \mathbb{V})$.

One easily checks that $S_{\mathrm{atyp}}(\mathbb{V})$ is contained in $\mathrm{HL}(S, \mathbb{V})$ (Remark 4.4).

1.4. Optimal subvarieties

Let us introduce a notion cunningly different from atypicality:

Definition 1.8 (Optimal subvariety). Let S be an irreducible smooth quasi-projective variety and $\mathbb{V} \to S$ a variation of mixed Hodge structures on S. An irreducible subvariety $Y \subset S$ is said to be optimal for $(S, \mathbb{V})$ if, for any irreducible subvariety $Y \subsetneq Y' \subset S$ containing Y strictly, the following inequality holds:

$$\text{H-cd}(Y, \mathbb{V}_{|Y^{\mathrm{sm}}}) < \text{H-cd}(Y', \mathbb{V}_{|Y'^{\mathrm{sm}}}). \tag{1.2}$$

Notice that if $Y \subset S$ is optimal for $(S, \mathbb{V})$ then Y is atypical for $(S, \mathbb{V})$ and that, conversely, any irreducible $Y \subset S$ which is atypical for $(S, \mathbb{V})$ *and maximal for this property* is optimal for $(S, \mathbb{V})$.

1.5. Statements of the main conjecture

We have seen in Section 1.2 that the Zariski-closure of the Hodge locus $\mathrm{HL}(S, \mathbb{V})$ can be complicated. The main object of this text is to present the following conjectures (shown in Section 5 to be equivalent), which predict that the subset $S_{\mathrm{atyp}}(\mathbb{V})$ of $\mathrm{HL}(S, \mathbb{V})$ on the contrary has a simple geometry (or equivalently that optimal subvarieties are rare):

Conjecture 1.9 (Main conjecture, form 1). *For any irreducible smooth quasi-projective variety S endowed with a variation of mixed Hodge structures $\mathbb{V} \to S$, the subset $S_{\mathrm{atyp}}(\mathbb{V})$ is a finite union of special subvarieties of S.*

Conjecture 1.10 (Main conjecture, form 2). *For any irreducible smooth quasi-projective variety S endowed with a variation of mixed Hodge structures $\mathbb{V} \to S$, the subset $S_{\mathrm{atyp}}(\mathbb{V})$ is a strict algebraic subset of S.*

Conjecture 1.11 (Main conjecture, form 3). *For any irreducible smooth quasi-projective variety S endowed with a variation of mixed Hodge structures $\mathbb{V} \to S$, the subset $S_{\mathrm{atyp}}(\mathbb{V})$ is not Zariski-dense in S.*

Conjecture 1.12 (Main conjecture, form 4). *Any irreducible smooth quasi-projective variety S endowed with a variation of mixed Hodge structures $\mathbb{V} \to S$ contains only finitely many irreducible subvarieties optimal for $(S, \mathbb{V})$.*

1.6. Organization of the paper

This note is organized as follows. Section 2 provides a recollection on mixed Hodge theory for the non-expert. Section 3 defines the notion of mixed Hodge varieties (a generalization in the complex analytic category of mixed Shimura varieties) and the corresponding period maps. Although most of the material in this section reorganizes classical results, our treatment is resolutely "group-oriented": Deligne's formalism of Shimura data (and its generalisation to Hodge data) seems to offer an unrivaled functorial setting for Hodge theory. Section 4 explains Conjecture 1.9 in terms of the geometry of period maps and atypical intersections in the sense of [36]. In Section 5 we prove the equivalence of Conjectures 1.9, 1.10, 1.11 and 1.12, and explain the relation between these conjectures and more classical statements like the Zilber–Pink conjecture for Shimura varieties (and its particular case the André–Oort conjecture). Section 6 details the simplest example of Conjecture 1.9 outside the world of Shimura varieties: the case of Calabi–Yau 3-folds. Section 7 describes the relation between Conjecture 1.9 and a functional transcendence statement of Ax–Schanuel type for period maps (Conjecture 7.5).

2. Mixed Hodge theory

2.1. Deligne torus

The Deligne torus is the restriction of scalars $\mathbf{S} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbf{G}_m$. So $\mathbf{S}_{\mathbb{C}}$ is canonically isomorphic to $\mathbf{G}_{\mathrm{m},\mathbb{C}} \times \mathbf{G}_{\mathrm{m},\mathbb{C}}$ but the action of complex conjugation on $\mathbf{S}_{\mathbb{C}}$ is given by the usual one twisted by the exchange of the two factors. In particular $\mathbf{S}(\mathbb{R}) = \mathbb{C}^* \subset \mathbf{S}(\mathbb{C}) = \mathbb{C}^* \times \mathbb{C}^*$ consists of the points of the form $(z, \overline{z})$.

Let $w\colon \mathbf{G}_{\mathrm{m},\mathbb{R}} \to \mathbf{S}$ be the cocharacter whose value on real points is given by $\mathbb{R}^* \subset \mathbb{C}^*$. We define the cocharacter $\mu\colon \mathbf{G}_{\mathrm{m},\mathbb{C}} \to \mathbf{S}_{\mathbb{C}}$ to be the unique cocharacter such that $\overline{z} \circ \mu$ is trivial and $z \circ \mu = \mathrm{Id} \in \mathrm{End}\,(\mathbf{G}_{\mathrm{m},\mathbb{C}})$, where z and $\overline{z}$ are the two characters of $\mathbf{S}$ generating its character group such that the induced maps on points $\mathbb{C}^* = \mathbf{S}(\mathbb{R}) \subset \mathbf{S}(\mathbb{C}) \to \mathbf{G}_m(\mathbb{C}) = \mathbb{C}^*$ are the identity, resp. complex conjugation. On $\mathbb{C}$-points, identifying $\mathbf{S}(\mathbb{C}) = \mathbb{C}^* \times \mathbb{C}^*$, we have $\mu\colon \mathbb{C}^* \to \mathbb{C}^* \times \mathbb{C}^*$ given by $w \mapsto (w, 1)$.

2.2. Pure Hodge structures

We denote by R one of the rings $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. Given an R-module V we write $V_{\mathbb{R}} := V \otimes_R \mathbb{R}$ and $V_{\mathbb{C}} := V \otimes_R \mathbb{C}$.

A pure R-Hodge structure (resp. of weight $n \in \mathbb{Z}$) is a Noetherian R-module V together with a morphism of algebraic groups $\varphi\colon \mathbf{S} \to \mathbf{GL}(V_{\mathbb{R}})$ (resp. such that $\varphi \circ w$ is given by $\mathbb{C}^* = \mathbf{S}(\mathbb{R}) \ni z \mapsto z^{-n} \cdot \mathrm{Id}_V$). Notice that if R is a field, any pure R-Hodge structure is a direct sum of pure R-Hodge structures of fixed weight.

Equivalently, a pure R-Hodge structure of weight $n \in \mathbb{Z}$ is a Noetherian R-module V together with a bigrading $V_{\mathbb{C}} = \bigoplus_{p+q=n} V^{p,q}$ satisfying $\overline{V^{p,q}} = V^{q,p}$, or a decreasing filtration called the Hodge filtration F^p on $V_{\mathbb{C}}$ such that $F^p \oplus \overline{F^{n-p+1}} \xrightarrow{\simeq} V_{\mathbb{C}}$. The equivalence between φ, the Hodge filtration and the bigrading is as follows: the subspace $V^{p,q}$ of $V_{\mathbb{C}}$ is the eigenspace of $\mathbf{S}(\mathbb{C}) = \mathbb{C}^* \times \mathbb{C}^*$ associated with the character $(z, w) \mapsto z^{-p}w^{-q}$, $F^p = \bigoplus_{p' \geq p} V^{p',n-p'}$ and $V^{p,q} = F^p \cap \overline{F^q}$.

For example there exists a unique R-Hodge structure of weight $-2n$ on $V = (2\pi i)^n R$ called the Tate-Hodge structure and denoted $R(n)$.

A morphism $f\colon (V, \varphi) \to (V', \varphi')$ of pure R-Hodge structures is a morphism $f\colon V \to V'$ of R-modules such that $f_{\mathbb{R}}\colon V_{\mathbb{R}} \to V'_{\mathbb{R}}$ commutes with the action of $\mathbf{S}$.

We denote by $\mathcal{H}_R$ the category of pure R-Hodge structures and, for every $n \in \mathbb{Z}$, by $\mathcal{H}_R^n$ the full subcategory of $\mathcal{H}_R$ of pure R-Hodge structures of weight n.

If (V, φ) is a pure R-Hodge structure of weight n, a polarisation for (V, φ) is a morphism of R-Hodge structures $Q\colon V^{\otimes 2} \to R(-n)$ such that $(2\pi i)^n Q(x, \varphi(i)y)$ is a positive-definite bilinear form on $V_{\mathbb{R}}$.

2.3. Mixed Hodge structures

We denote by K the field $R \otimes_{\mathbb{Z}} \mathbb{Q}$.

A mixed R-Hodge structure (RMHS) is a triple $(V, W_\bullet, F^\bullet)$ consisting of a Noetherian R-module V, a finite ascending filtration $W_\bullet$ of $V_K := V \otimes_R K$ (called the weight filtration) and a finite decreasing filtration $F^\bullet$ of $V_{\mathbb{C}}$ (called the Hodge filtration) such that for each $n \in \mathbb{Z}$ the couple $(\mathrm{Gr}_n^W V_K, \mathrm{Gr}_n^W(F^\bullet))$ is a pure K-Hodge structure of weight n.

A pure R-Hodge structure V of weight $n \in \mathbb{Z}$ is then a special case of a mixed R-Hodge structure by defining the weight filtration as $W_{n'}V_K = V_K$ for $n' \geq n$ and $W_{n'}V_K = 0$ for $n' < n$. The notions of weight greater or smaller than $n \in \mathbb{Z}$ are defined in the obvious way.

We say that an RMHS is of type $\varepsilon \subset \mathbb{Z} \times \mathbb{Z}$ if the Hodge numbers $h^{p,q} := \dim_{\mathbb{C}}(\mathrm{Gr}_{p+q}^W V_K)^{p,q}$ are zero for $(p,q) \notin \varepsilon$ and non-zero for $(p,q) \in \varepsilon$.

A morphism $f\colon (V, W_\bullet, F^\bullet) \to (V', W'_\bullet, F'^\bullet)$ of RMHS is a morphism $f\colon V \to V'$ of R-modules respecting the weight and the Hodge filtration.

A graded polarization of a mixed R-Hodge structure is the datum of a polarization on the pure K-Hodge structure $\mathrm{Gr}^W V_K$.

Following [12, Th. 2.3.5], the category $\mathcal{MH}_R$ of mixed R-Hodge structures is Abelian (where the kernels and cokernels of morphisms are endowed with the induced filtrations); the functors $\mathrm{Gr}_n^W : \mathcal{MH}_R \to \mathcal{H}_K^n$ and $\mathrm{Gr}_F^p\colon V \mapsto \mathrm{Gr}_F^p(V_{\mathbb{C}})$ are exact. Moreover every morphism $f\colon (V, W_\bullet, F^\bullet) \to (V', W'_\bullet, F'^\bullet)$ of RMHS is strictly compatible with the weight and the Hodge filtrations (meaning that the inclusions $f(W_n) \subset W'_n$ and $f(F^p) \subset F'^p$ satisfy $f(W_n) = f(V_K) \cap W'_n$ and $f(F^p) = f(V_{\mathbb{C}}) \cap F'^p$).

We wish to extend our group-theoretic description from pure Hodge structures to mixed ones. Let $(V, W_\bullet, F^\bullet)$ be a mixed $\mathbb{R}$-Hodge structure. A splitting of $(V, W_\bullet, F^\bullet)$ is a bigrading $V_{\mathbb{C}} = \bigoplus_{p,q} V^{p,q}$ such that the equalities $W_n V = \sum_{p+q \leq n} V^{p,q}$ and $F^p = \sum_{r \geq p} V^{r,s}$ hold. Deligne proved that any mixed $\mathbb{R}$-Hodge structure admits a unique preferred splitting:

Proposition 2.1 (Deligne). *Let $(V, W_\bullet, F^\bullet)$ be a mixed $\mathbb{R}$-Hodge structure. It admits a unique splitting $(V^{p,q})_{p,q}$ satisfying:*

$$V^{p,q} = \overline{V^{q,p}} \mod \bigoplus_{r<p, s<q} V^{r,s}. \tag{2.1}$$

This splitting is functorial.

Remark 2.2. A mixed $\mathbb{R}$-Hodge structure $(V, W_\bullet, F^\bullet)$ such that Deligne's splitting satisfies moreover $V^{p,q} = \overline{V^{q,p}}$ is said to be *split over* $\mathbb{R}$.

Given a mixed R-Hodge structure $(V, W_\bullet, F^\bullet)$, Deligne's splitting on $V_{\mathbb{R}}$ defines a unique morphism $\varphi_{\mathbb{C}}\colon \mathbf{S}_{\mathbb{C}} \to \mathbf{GL}(V_{\mathbb{C}})$ such that $V^{p,q}$ is the eigenspace for the character $(z, w) \mapsto z^{-p} w^{-q}$ of $\mathbf{S}_{\mathbb{C}}$. If $(V, W_\bullet, F^\bullet)$ splits over $\mathbb{R}$ the morphism $\varphi_{\mathbb{C}}$ is the complexification of a morphism $\varphi\colon \mathbf{S} \to \mathbf{GL}(V_{\mathbb{R}})$. In particular we recover our initial definition of a pure R-Hodge structure.

Conversely:

Proposition 2.3. [28, 1.4 and 1.5] *Let V be a Noetherian R-module and let $\varphi_{\mathbb{C}} \colon \mathbf{S}_{\mathbb{C}} \to \mathbf{GL}(V_{\mathbb{C}})$ be a group morphism. It defines an RMHS on V if and only if, denoting by $\mathbf{P}$ the K-algebraic group Zariski closure over K of $\varphi_{\mathbb{C}}(\mathbf{S}_{\mathbb{C}})$, by $\mathbf{W}$ the unipotent radical of $\mathbf{P}$ and by $\pi : \mathbf{P} \to \mathbf{G} := \mathbf{P}/\mathbf{W}$ the reductive quotient of $\mathbf{P}$, the following conditions are satisfied:*

(1) *The composite $\mathbf{S}_{\mathbb{C}} \xrightarrow{\varphi_{\mathbb{C}}} \mathbf{P}_{\mathbb{C}} \xrightarrow{\pi} \mathbf{G}_{\mathbb{C}}$ is defined over $\mathbb{R}$.*

(2) *The composite $\mathbf{G}_{m,\mathbb{R}} \xrightarrow{w} \mathbf{S} \xrightarrow{\pi\circ\varphi} \mathbf{G}_{\mathbb{R}}$ is defined over K.*

(3) *The weight filtration of the mixed R-Hodge structure on $\mathfrak{p}$ defined by $\mathrm{Ad}_{\mathbf{P}} \circ \varphi_{\mathbb{C}}$ satisfies $W_{-1}(\mathfrak{p}) = \mathfrak{w}$, where $\mathfrak{w}$ denotes the Lie algebra of the K-group $\mathbf{W}$.*

Remark 2.4. Notice that, by (2), $\mathrm{Ad}_{\mathbf{P}} \circ \varphi_{\mathbb{C}}$ endows $\mathfrak{g}$ with a pure K-Hodge structure of weight 0, and hence, by (3), $\mathfrak{p}$ with a mixed K-Hodge structure of weight ≤ 0. In particular $W_0\mathfrak{p} = \mathfrak{p}$.

2.4. Mumford–Tate groups

The category $\mathcal{MH}_K$ of mixed K-Hodge structures is a K-linear tensor category which is rigid and has an obvious exact faithful K-linear tensor functor $\omega \colon (V_K, W_\bullet, F^\bullet) \mapsto V_K$.

For any RMHS $(V, W_\bullet, F^\bullet)$ we denote by $\langle V \rangle$ the Tannakian subcategory of $\mathcal{MH}_K$ generated by $(V_K, W_\bullet, F^\bullet)$ and by ω_V the restriction of the tensor functor ω to $\langle V \rangle$; in other words $\langle V \rangle$ is the smallest full subcategory containing $(V_K, W_\bullet, F^\bullet)$ and the trivial KMHS and stable under $\oplus$, $\otimes$, and taking subquotients. Then the functor $\mathrm{Aut}^{\otimes}(\omega_V)$ is representable by some closed K-algebraic subgroup $\mathbf{P}(V)$ of $\mathbf{GL}(V_K)$, called the Mumford–Tate group of $\mathbb{V}$, and ω_V defines an equivalence of categories $\langle V \rangle \simeq \mathbf{Rep}_K\mathbf{P}(V)$ ([13, II, 2.11]).

There are various equivalent definitions for $\mathbf{P}(V)$, in particular in terms of Hodge tensors. Recall that a Hodge class for V is a vector in $F^0V_{\mathbb{C}} \cap W_0V_K$. For integers $m, n \geq 0$, let $T^{m,n}V_K$ denote the mixed K-Hodge structure $V^{\otimes m} \otimes \mathrm{Hom}(V, R)^{\otimes n} \otimes_R K$. A Hodge tensor for V is a Hodge class in some $T^{m,n}V_K$.

Lemma 2.5 ([1, Lemma 2])**.** *Let $(V, W_\bullet, F^\bullet)$ be a mixed R-Hodge structure. Then:*

(a) *Any tensor fixed by $\mathbf{P}(V)$ in some $T^{m,n}V_K$ is a Hodge tensor. Conversely $\mathbf{P}(V)$ is the stabilizer in $\mathbf{GL}(V_K)$ of the Hodge tensors for V.*

(b) *$\mathbf{P}(V)$ is the K-Zariski-closure of the image of φ in $\mathbf{GL}(V_K)$ (and hence it is connected), and if moreover V is pure polarizable then $\mathbf{P}(V)$ is reductive.*

(c) *The group $\mathbf{P}(V)$ preserves $W_\bullet$ and is an extension of $\mathbf{P}(\mathrm{Gr}^W V_K)$ by a unipotent subgroup; in particular if V is graded-polarizable then the group $\mathbf{P}(\mathrm{Gr}^W V_K)$ is the quotient of $\mathbf{P}(V)$ by its unipotent radical.*

2.5. Variations of mixed Hodge structures

Hodge theory as recalled above can be considered as the particular case over a point of Hodge theory over an arbitrary base in the category of complex manifolds.

Let S be a complex manifold and $\mathcal{O}_S$ its sheaf of holomorphic functions. A variation of mixed R-Hodge structures (RVMHS) over S is a triple $(\mathbb{V}, W_\bullet, F^\bullet)$, where:

(1) $\mathbb{V}$ is a locally constant R_S-module on S,
(2) $W_\bullet$ 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 $(\mathbb{V}_s, (W_\bullet)_s, F^\bullet_s)$ is a mixed R-Hodge structure,
(b) the flat connection $\nabla\colon \mathcal{V} \to \mathcal{V} \otimes \Omega^1_S$ whose sheaf of horizontal sections is $\mathbb{V}_\mathbb{C}$ satisfies the Griffiths' transversality condition

$$\nabla F^\bullet \subset \Omega^1_S \otimes F^{\bullet-1}. \tag{2.2}$$

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

$$\Psi_k\colon \mathrm{Gr}^W_k(\mathbb{V}_K) \times \mathrm{Gr}^W_k(\mathbb{V}_K) \longrightarrow K(-k)_S$$

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

Variations of mixed Hodge structure over a smooth quasi-projective base S can have a pathological behaviour at infinity. The *admissibility* condition was defined by Kashiwara [22] to remedy this problem. As its statement is technical and we won't need it in this expository note (but it will be crucial in the proofs!), we content ourselves to refer to [22], [32] and [7, Def. 7.2]. Notice that every geometric variation of mixed Hodge structures is admissible. *From now on any $\mathbb{Z}$VMHS is graded-polarizable and admissible.*

Let $(\mathbb{V}, W_\bullet, F^\bullet)$ be an RVMHS over S. An important infinitesimal invariant associated with it is the Kodaira–Spencer map

$$\overline{\nabla}\colon TS^{\mathrm{sm}} \longrightarrow \mathrm{Gr}^{-1}_F(W_0\mathrm{End}\,\mathbb{V} \otimes_\mathbb{Q} \mathcal{O}_S), \tag{2.3}$$

which is nothing but the flat connection ∇ considered at first order. Indeed the maps $\nabla\colon F^p \to F^{p-1} \otimes_{\mathcal{O}_S} \Omega^1_S$ and $\nabla\colon F^{p-1} \to F^{p-2} \otimes_{\mathcal{O}_S} \Omega^1_S$ induce an $\mathcal{O}_S$-linear morphism $\overline{\nabla}\colon \mathrm{Gr}^p_F\mathcal{V} \to \mathrm{Gr}^{p-1}\mathcal{V} \otimes_{\mathcal{O}_S} \Omega^1_S$. The alternative writing (2.3) follows from the fact that ∇ respects the weight filtration.

2.6. Hodge locus and special subvarieties

Let $R = \mathbb{Z}$ or $\mathbb{Q}$. With the notation of the previous section, for s in S let $\mathbf{P}_s \subset \mathbf{GL}(\mathbb{V}_{\mathbb{Q},s})$ denote the Mumford–Tate group of the fiber $\mathbb{V}_s$. By [1, Lemma 4] (following Deligne in the pure case) the group $\mathbf{P}_s$ is locally constant on the complement $S^0 := S \setminus HL(S, \mathbb{V})$ of a countable union $\mathrm{HL}(S, \mathbb{V}) \subset S$ of proper irreducible analytic subvarieties of S. The locus $\mathrm{HL}(S, \mathbb{V})$ is called the Hodge locus of $(S, \mathbb{V})$. Let $\pi_S \colon \tilde{S} \to S$ denote a universal covering and choose a flat trivialization $\pi_S^* \mathbb{V}_{\mathbb{Q}} \simeq \tilde{S} \times V$. The choice of a point $\tilde{s} \in \tilde{S}$ such that $\pi_S(\tilde{s}) = s$ gives an identification $\mathbb{V}_{\mathbb{Q},s} \simeq V$, and hence an injective homomorphism $i_{\tilde{s}} \colon \mathbf{P}_s \hookrightarrow \mathbf{GL}(V)$. For $s \in S^0$ (such a point of S is called Hodge generic) the image $\mathbf{P}_S(\mathbb{V}) := \mathrm{Im}(i_{\tilde{s}}) \subset \mathbf{GL}(V)$ neither depends on w nor on the choice of $\tilde{s}$ (it is called the generic Mumford–Tate group of (S, V)); for all $s \in HL(S, \mathbb{V})$ and $\tilde{s} \in \tilde{S}$ above s, the image of $i_{\tilde{s}}$ is a proper subgroup of $\mathbf{P}_S(\mathbb{V})$.

When S is quasi-projective and $\mathbb{V}$ is of geometric origin, the Hodge conjecture implies that $\mathrm{HL}(S, \mathbb{V})$ is in fact a countable union of closed irreducible algebraic subvarieties of S. The following result is fundamental in the study of Hodge loci:

Theorem 2.6 ([8] **in the pure case,** [6] **in the mixed one).** *Suppose that S is quasi-projective. Then the Hodge locus* $\mathrm{HL}(S, \mathbb{V})$ *is a countable union of closed irreducible* algebraic *subvarieties of S.*

The irreducible components of the Zariski-closure of the strata of the Hodge locus $\mathrm{HL}(S, \mathbb{V})$ where $\mathbf{P}_s$ is locally constant are called *special subvarieties* of $(S, \mathbb{V})$. Special subvarieties of dimension zero are called *special points* of $(S, \mathbb{V})$. Special points whose Mumford–Tate group is a torus are called *CM points* (where CM stands for Complex Multiplication).

3. Mixed Hodge varieties

In this section we define (connected) mixed Hodge varieties, a generalization of (connected) mixed Shimura varieties defined by Deligne [12] and Pink [28]. All the ingredients are due to Pink [28]. We refer to the recent monograph [16] for more details in the pure case.

3.1. Mixed Hodge data

In this section $R = K = \mathbb{Q}$. We want to parametrize mixed $\mathbb{Q}$-Hodge structures, with fixed weight filtration and given Hodge numbers, by a homogeneous space described in terms of homomorphisms $\varphi_{\mathbb{C}} \colon \mathbf{S}_{\mathbb{C}} \to \mathbf{P}_{\mathbb{C}}$ and carrying a canonical complex structure. The following result of Pink will guide our definition.

Proposition 3.1. [28, 1.7] *Let $(\mathbf{P}, X_{\mathbf{P}})$ be a pair of a connected linear algebraic group $\mathbf{P}$ over $\mathbb{Q}$ and a $\mathbf{P}(\mathbb{R})\mathbf{W}(\mathbb{C})$-conjugacy class $X_{\mathbf{P}}$ in* $\mathrm{Hom}(\mathbf{S}_{\mathbb{C}}, \mathbf{P}_{\mathbb{C}})$ *(where $\mathbf{W}$ denotes the unipotent radical of $\mathbf{P}$). Assume that for one $\varphi \in X_{\mathbf{P}}$ (and then for any) the conditions* (1), (2) *and* (3) *of Proposition 2.3 are satisfied.*

Let $\rho \colon \mathbf{P} \to \mathbf{GL}(M)$ be any faithful (finite dimensional, algebraic) representation of $\mathbf{P}$. Then:

(a) *The image* $\mathcal{D}_{\mathbf{P},X_\mathbf{P}}$ *of the obvious map*

$$\psi\colon X_\mathbf{P} \longrightarrow \{\text{mixed } \mathbb{Q}\text{-Hodge structures on M}\}$$

admits a unique structure of complex manifold such that the Hodge filtration on $M_\mathbb{C}$ *depends analytically on* $\psi(\varphi)$*. This structure is invariant under* $\mathbf{P}(\mathbb{R})\mathbf{W}(\mathbb{C})$ *and* $\mathbf{W}(\mathbb{C})$ *acts analytically on* $\mathcal{D}_{\mathbf{P},X_\mathbf{P}}$.

(b) *The complex manifold* $\mathcal{D}_{\mathbf{P},X_\mathbf{P}}$ *is independent of the choice of the faithful representation* M.

Proposition 3.1 is proved by noticing that, as the weight filtration is constant on M when φ varies in $X_\mathbf{P}$, the Hodge filtration on M gives an injective map of $\mathcal{D}_{\mathbf{P},X_\mathbf{P}}$ into the flag manifold $\hat{\mathcal{D}}_{\mathbf{P},X_\mathbf{P}} := \mathbf{P}(\mathbb{C})/\exp(F^0\mathfrak{p}_\mathbb{C})$, which is easily shown to be an open embedding.

Remark 3.2. It is important to notice that:

- the surjective map $\psi\colon X_\mathbf{P} \to \mathcal{D}_{\mathbf{P},X_\mathbf{P}}$ is not a bijection in general. This is the case if and only if $F^0\mathfrak{w}_\mathbb{C}$ is trivial (see [28, 1.8 (b)]);
- the group $\mathbf{P}(\mathbb{R})$ does not in general act transitively on $\mathcal{D}_{\mathbf{P},X_\mathbf{P}}$ for the action of $\mathbf{P}(\mathbb{R})$ on $\mathcal{D}_{\mathbf{P},X_\mathbf{P}}$ defined in Proposition 3.1. It acts transitively on "the set of real points of $\mathcal{D}_{\mathbf{P},X_\mathbf{P}}$", namely the points for which the corresponding mixed Hodge structure on (any) M is split over $\mathbb{R}$.

Definition 3.3. A pair $(\mathbf{P}, X_\mathbf{P})$ as in Proposition 3.1 is called a *mixed Hodge datum*. We call $\mathcal{D} := \mathcal{D}_{\mathbf{P},X_\mathbf{P}}$ the *mixed Mumford–Tate domain* associated with the mixed Hodge datum $(\mathbf{P}, X_\mathbf{P})$.

Definition 3.4. A connected mixed Hodge datum is a triple $(\mathbf{P}, X_\mathbf{P}, \mathcal{D}^+)$, where $(\mathbf{P}, X_\mathbf{P})$ is a mixed Hodge datum and $\mathcal{D}^+$ is a connected component of $\mathcal{D} := \mathcal{D}_{\mathbf{P},X_\mathbf{P}}$.

Definition 3.5. A morphism of mixed Hodge data $(\mathbf{P}, X_\mathbf{P}) \to (\mathbf{P}', X_{\mathbf{P}'})$ is a homomorphism $\rho\colon \mathbf{P} \to \mathbf{P}'$ of $\mathbb{Q}$-algebraic groups which induces maps $X_\mathbf{P} \to X_{\mathbf{P}'}$. It naturally induces a holomorphic map $\mathcal{D} \to \mathcal{D}'$. A morphism of connected mixed Hodge data $(\mathbf{P}, X_\mathbf{P}, \mathcal{D}^+) \to (\mathbf{P}', X_{\mathbf{P}'}, \mathcal{D}'^+)$ is a morphism of mixed Hodge data $(\mathbf{P}, X_\mathbf{P}, \mathcal{D}) \to (\mathbf{P}', X_{\mathbf{P}'}, \mathcal{D}')$ whose last component maps $\mathcal{D}^+$ to $\mathcal{D}'^+$.

3.2. Mixed Hodge datum of Shimura type

Let $(\mathbf{P}, X_\mathbf{P})$ be a mixed Hodge datum and M a representation of $\mathbf{P}$. The representation M defines a $\mathbf{P}(\mathbb{R})\mathbf{W}(\mathbb{C})$-equivariant local system $\mathbb{M}$ on $\mathcal{D}$, supporting a family of mixed Hodge structures: a flat weight filtration $W_\bullet$ on $\mathbb{M}$ and a holomorphic Hodge filtration $F^\bullet$ on $\mathbb{M} \otimes_\mathbb{Q} \mathcal{O}_\mathcal{D}$ such that, for each $s \in \mathcal{D}$, the triple $(\mathbb{M}_s, (W_\bullet)_s, F^\bullet_s)$ is a mixed $\mathbb{Q}$-Hodge structure. In most cases however, the Griffiths transversality condition (2.2) is not satisfied:

Proposition 3.6. [28, 1.10] *Let* $(\mathbf{P}, X_\mathbf{P})$ *be a mixed Hodge datum and* M *a representation of* $\mathbf{P}$*. The triple* $(\mathbb{M}, W_\bullet, F^\bullet)$ *is a variation of mixed* $\mathbb{Q}$*-Hodge structures on* $\mathcal{D}$ *if and only if the Hodge structure on* $\mathfrak{p}$ *is of type*

$$\{(-1,1),(0,0),(1,-1)\},\{(-1,0),(0,-1)\},\{(-1,-1)\}.$$

Definition 3.7. A mixed Hodge datum $(\mathbf{P}, X_{\mathbf{P}})$ satisfying the condition of Proposition 3.6 is said to be of Shimura type.

Let us now relate mixed Hodge data of Shimura type with mixed Shimura varieties in the sense of [28]. Let $(\mathbf{P}, X_{\mathbf{P}})$ be a mixed Hodge datum of Shimura type. Let $\mathbf{U} \subset \mathbf{W}$ be the unique connected subgroup such that $\mathrm{Lie}\,\mathbf{U} = W_{-2}\mathrm{Lie}\,\mathbf{W}$. Following [28, 1.15] we define $\mathcal{X} \subset X_{\mathbf{P}}$ as the subset of the $\varphi_{\mathbb{C}}$'s defined by the following condition stronger than Proposition 2.3 (1):

(1') $\mathbf{S}_{\mathbb{C}} \xrightarrow{\varphi_{\mathbb{C}}} \mathbf{P}_{\mathbb{C}} \xrightarrow{\pi'} (\mathbf{P}/\mathbf{U})_{\mathbb{C}}$ is defined over $\mathbb{R}$.

In [28, 1.16] Pink proves that the restriction $\psi_{|\mathcal{X}} \colon \mathcal{X} \to \mathcal{D}$ of $\psi \colon X_{\mathbf{P}} \to \mathcal{D}$ is a bijection. The triple $(\mathbf{P}, \mathcal{D}, \psi_{|\mathcal{X}}^{-1})$ is then a mixed Shimura datum in the sense of [28, 2.1]. Conversely if $(\mathbf{P}, \mathcal{X}, h)$ is a mixed Shimura datum as in loc. cit. let us define $X_{\mathbf{P}}$ as the $\mathbf{P}(\mathbb{R})\mathbf{W}(\mathbb{C})$-conjugacy class of any point of $h(\mathcal{X})$. Then $(\mathbf{P}, X_{\mathbf{P}})$ is a mixed Hodge datum.

3.3. Connected mixed Hodge varieties

3.3.1. Definitions. We refer to [5] and [25, p. 33, 34, 42] for an introduction to arithmetic groups, congruence subgroups and neat subgroups.

Let $(\mathbf{P}, X_{\mathbf{P}}, \mathcal{D}^+)$ be a connected mixed Hodge datum. The stabilizer $\mathbf{P}(\mathbb{R})^+$ of $\mathcal{D}^+$ is easily seen to be open in $\mathbf{P}(\mathbb{R})$. Set $\mathbf{P}(\mathbb{Q})^+ := \mathbf{P}(\mathbb{Q}) \cap \mathbf{P}(\mathbb{R})^+$. Then any congruence subgroup $\Gamma \subset \mathbf{P}(\mathbb{Q})^+$ acts properly discontinuously on $\mathcal{D}^+$, so that $\Gamma\backslash\mathcal{D}^+$ is a complex analytic space with at most finite quotient singularities. Every sufficiently small finite index congruence subgroup of Γ acts freely on $\mathcal{D}^+$. Replacing Γ by such a subgroup if necessary, the quotient $\Gamma\backslash\mathcal{D}^+$ is thus a complex manifold and the map $\mathcal{D}^+ \longrightarrow \Gamma\backslash\mathcal{D}^+$ is unramified.

Definition 3.8 (connected mixed Hodge variety). Let $(\mathbf{P}, X_{\mathbf{P}}, \mathcal{D}^+)$ be a connected mixed Hodge datum and $\Gamma \subset \mathbf{P}(\mathbb{Q})^+$ a congruence subgroup. The complex analytic variety $\mathrm{Hod}^0_\Gamma(\mathbf{P}, X_{\mathbf{P}}, \mathcal{D}^+) := \Gamma\backslash\mathcal{D}^+$ is called the connected mixed Hodge variety associated with $(\mathbf{P}, X_{\mathbf{P}}, \mathcal{D}^+)$ and Γ. It is called a connected (pure) Hodge variety if $\mathbf{P}$ is reductive. The class of an element $x \in \mathcal{D}^+$ is denoted $[x] \in \Gamma\backslash\mathcal{D}^+$.

As in [28, 3.4] one obtains:

Lemma 3.9. *Let $\rho \colon (\mathbf{P}, X_{\mathbf{P}}, \mathcal{D}^+) \to (\mathbf{P}', X_{\mathbf{P}'}, \mathcal{D}'^+)$ be a morphism of connected mixed Hodge data and $\Gamma \subset \mathbf{P}(\mathbb{Q})^+$ and $\Gamma' \subset \mathbf{P}'(\mathbb{Q})^+$ congruence lattices such that $\rho(\Gamma) \subset \Gamma'$. Then the map*

$$[\rho] \colon \Gamma\backslash\mathcal{D}^+ \longrightarrow \Gamma'\backslash\mathcal{D}'^+$$

mapping $[x]$ to $[\rho \circ x]$ is well-defined and holomorphic. Such a map is called a Hodge morphism of connected mixed Hodge varieties.

Definition 3.10 (mixed Hodge variety). Let $(\mathbf{P}, X_{\mathbf{P}})$ be a mixed Hodge datum and let $K \subset \mathbf{P}(\mathbb{A}_{\mathrm{f}})$ be a compact open subgroup of the finite adelic points of $\mathbf{P}$. The mixed Hodge variety $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ is the complex analytic space

$\mathbf{P}(\mathbb{Q})\backslash(\mathcal{D}\times\mathbf{P}(\mathbb{A}_{\mathrm{f}})/K)$. Here $\mathbf{P}(\mathbb{Q})$ acts diagonally on $\mathcal{D}$ and $\mathbf{P}(\mathbb{A}_{\mathrm{f}})$, the group K acts on $\mathbf{P}(\mathbb{A}_{\mathrm{f}})$ on the right, and $\mathbf{P}(\mathbb{A}_{\mathrm{f}})$ is endowed with the adelic topology.

As in [25, Lemma 5.11, 5.12, 5.13] one shows:

Lemma 3.11. *Fix $\mathcal{D}^+$ a connected component of $\mathcal{D}$. Let $\mathcal{C}$ be a (finite) set of representatives for the finite double coset space $\mathbf{P}(\mathbb{Q})^+\backslash\mathbf{P}(\mathbb{A}_{\mathrm{f}})/K$. Then we have a homeomorphism $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}}) = \coprod_{g\in\mathcal{C}}\Gamma_g\backslash\mathcal{D}^+$, where Γ_g is the congruence subgroup $gKg^{-1}\cap\mathbf{P}(\mathbb{Q})^+$ of $\mathbf{P}(\mathbb{Q})^+$.*

In particular the mixed Hodge variety $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ is a finite union of connected mixed Hodge varieties.

When $(\mathbf{P}, X_{\mathbf{P}})$ is a mixed Hodge datum of Shimura type and $(\mathbf{P}, \mathcal{X}, h)$ is the associated mixed Shimura datum as in Section 3.2, the mixed Hodge variety $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ coincides with the mixed Shimura variety $M^K(\mathbf{P}, \mathcal{X})(\mathbb{C})$ defined in [28, 3.1]. We will say that $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ is of Shimura type and denote it by $\mathrm{Sh}_K(\mathbf{P}, X_{\mathbf{P}})$.

Let $\mathcal{D}^+$ be a connected component of $\mathcal{D}$ and $\Gamma\subset\mathbf{P}(\mathbb{Q})^+$ a torsion-free congruence subgroup. If M is a representation of $\mathbf{P}$, the family $(\mathbb{M}, W_\bullet, F^\bullet)$ of mixed $\mathbb{Q}$-Hodge structures on $\mathcal{D}^+$ descends to the connected mixed Hodge variety $\Gamma\backslash\mathcal{D}^+$. This defines a variation of mixed $\mathbb{Q}$-Hodge structures (*i.e.* satisfies Griffiths' transversality) if and only if $\Gamma\backslash\mathcal{D}^+$ is a connected mixed Shimura variety.

3.3.2. Algebraicity. In general the connected mixed Hodge variety $\Gamma\backslash\mathcal{D}^+$ and the mixed Hodge variety $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ are complex analytic varieties which do not admit any algebraic structure. If $(\mathbf{P}, X_{\mathbf{P}})$ is a pure Hodge datum one can show ([18]) that the pure Hodge variety $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ admits an algebraic structure if and only if it fibers holomorphically or antiholomorphically over a (connected) Shimura variety $\mathrm{Sh}_K(\mathbf{P}', X_{\mathbf{P}'})$. A similar result should hold in the mixed case.

3.4. Special subvarieties of Hodge varieties

Let $(\mathbf{P}, X_{\mathbf{P}}, \mathcal{D}^+)$ be a connected mixed Hodge datum and $Y := \Gamma\backslash\mathcal{D}^+$ an associated connected mixed Hodge variety. Although a representation M of $\mathbf{P}$ does not in general define a variation of mixed $\mathbb{Q}$-Hodge structures on Y, we can still define a notion of special subvariety of Y in purely group theoretical terms.

Definition 3.12. The image of any Hodge morphism $T\to Y$ between connected mixed Hodge varieties is called a special subvariety of Y. It is said to be of Shimura type if T is a connected mixed Shimura variety.

Definition 3.13. The Hodge locus $\mathrm{HL}(Y)$ of the connected mixed Hodge variety Y is the union of all special subvarieties of Y.

3.5. Period maps

Let $(\mathbf{P}, X_{\mathbf{P}})$ be a mixed Hodge datum, $\mathcal{D}$ the associated mixed Mumford–Tate domain, $K \subset \mathbf{P}(\mathbb{A}_f)$ a neat compact open subgroup and M an algebraic representation of $\mathbf{P}$. In general, Griffiths' transversality for $(\mathbb{M}, W_\bullet, F^\bullet)$ is recovered by restricting ourselves to *horizontal* subvarieties of $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$, which are defined as follows. The tangent bundle $T\mathcal{D}$ is naturally equivariant under $\mathbf{P}(\mathbb{R})\mathbf{W}(\mathbb{C})$, associated to the representation $F^{-\infty}\mathfrak{p}/F^0\mathfrak{p}$. The Hodge filtration on $\mathfrak{p}$ defined by the mixed Hodge structure $\varphi_{\mathbb{C}} \colon \mathbf{S}_{\mathbb{C}} \to \mathbf{P}_{\mathbb{C}}$ thus defines a natural $\mathbf{P}(\mathbb{R})\mathbf{W}(\mathbb{C})$-equivariant filtration $F^\bullet$ on the holomorphic tangent bundle $T\mathcal{D}$.

Definition 3.14. The horizontal tangent bundle $T_h\mathcal{D}$ is defined to be the holomorphic subbundle $F^{-1}T\mathcal{D}$, associated to the representation $\mathrm{Gr}_F^{-1}\mathfrak{p}$. By equivariance it descends to a holomorphic subbundle

$$T_h\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}}) \subset T\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}}),$$

called the horizontal tangent bundle of the mixed Hodge variety $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$. One denotes by $T_h\mathrm{Hod}^0_\Gamma(\mathbf{P}, X_{\mathbf{P}})$ its restriction to the connected mixed Hodge variety $\mathrm{Hod}^0_\Gamma(\mathbf{P}, X_{\mathbf{P}}) = \Gamma\backslash\mathcal{D}^+$.

Remark 3.15. The equality $T_h\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}}) = T\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ holds if and only if $(\mathbf{P}, X_{\mathbf{P}})$ is of Shimura type, *i.e.* if $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ is a mixed Shimura variety.

Definition 3.16. Let S be a complex manifold. A holomorphic map $\Phi \colon S \to \mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ is said to be horizontal if $\Phi_*(TS) \subset T_h\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$.

Definition 3.17. Let S be a complex manifold. A holomorphic map $f \colon S \to \mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ is said to be locally liftable horizontal if, for each point $s \in S$, there exists a neighbourhood U_s of s in S and a commutative diagram

$$\begin{tikzcd} U_s \arrow[r, "\tilde{f}"] \arrow[dr, "f"'] & \mathcal{D} \arrow[d] \\ & \mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}}) \end{tikzcd}$$

such that $\tilde{f}$ is horizontal.

Definition 3.18. Let $(\mathbf{P}, X_{\mathbf{P}})$ be a mixed Hodge datum, $K \subset \mathbf{P}(\mathbb{A}_f)$ a compact open subgroup and $\mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ the associated Hodge variety. Let S be a complex manifold. A map

$$\Phi \colon S \longrightarrow \mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$$

is called a period map if it is holomorphic, locally liftable and horizontal.

If $\Phi \colon S \to \mathrm{Hod}_K(\mathbf{P}, X_{\mathbf{P}})$ is a period map the pullback $\Phi^*(\mathbb{M}, W_\bullet, F^\bullet)$ is a $\mathbb{Q}$VMHS on S. Conversely suppose $\mathbb{V} \to S$ is a $\mathbb{Z}$VMHS over a connected complex manifold S. Fix a Hodge generic base point $s \in S$, let $\mathbf{P}_S :=$

$\mathbf{P}_S(\mathbb{V}) = \mathbf{P}(\mathbb{V}_s)$ be the generic Mumford–Tate group of $\mathbb{V}$, $\mathbf{W}_S$ its unipotent radical, $\mathfrak{p}_S$ the Lie algebra of $\mathbf{P}_S$ and X_S the $\mathbf{P}_S(\mathbb{R})\mathbf{W}_S(\mathbb{C})$-conjugacy class of $\varphi_s \colon \mathbf{S}_\mathbb{C} \to \mathbf{GL}(\mathbb{V}_{s,\mathbb{C}})$. The pair $(\mathbf{P}_S, X_S)$ is a mixed Hodge datum. Let $\mathcal{D}_S^+$ be the connected component of $\mathcal{D}_S := \mathcal{D}_{\mathbf{P}_S, X_S}$ containing the image $\psi(\varphi_s)$. The triple $(\mathbf{P}_S, X_S, \mathcal{D}_S^+)$ is a connected mixed Hodge datum. Let $\Gamma := \mathbf{P}(\mathbb{Q}) \cap \mathbf{GL}(\mathbb{V}_{s,\mathbb{Z}})$. We define the connected mixed Hodge variety $\mathrm{Hod}^0(S, \mathbb{V})$ as $\mathrm{Hod}^0_\Gamma(\mathbf{P}_S, X_S, \mathcal{D}_S^+)$.

The Hodge filtration for $\mathbb{V} \to S$ defines a period map

$$\Phi_S \colon S \longrightarrow \mathrm{Hod}^0(S, \mathbb{V}). \tag{3.1}$$

Moreover the Kodaira–Spencer map $\overline{\nabla} \colon TS^0 \to \mathrm{Gr}_F^{-1} W_0 \operatorname{End} \mathbb{V}$ is naturally interpreted as the composed morphism of fiber bundles

$$\begin{tikzcd} TS \arrow[r, "d\Phi_S"] \arrow[rr, bend right, "\overline{\nabla}"'] & \Phi_S^* \, T_h \mathrm{Hod}^0(S, \mathbb{V}) \arrow[r, hook] & \mathrm{Gr}_F^{-1} W_0 \operatorname{End} \mathbb{V} \, . \end{tikzcd}$$

Remark 3.19. By a classical theorem of Griffiths [19, Theor. 9.6] one can, enlarging S if necessary, assume that $\Phi_S(S)$ is a closed complex analytic subvariety of $\mathrm{Hod}^0(S, \mathbb{V})$. A long-standing conjecture of Griffiths states that the closed complex analytic horizontal subvariety $\Phi_S(S)$ should admit a canonical structure of quasi-projective algebraic variety. In the pure case we refer to the work of Sommese [31] for partial results and to [24] and [17] which both announce Griffiths's conjecture (I was not able to understand their proofs). *Note added in proof: since this paper was submitted, Griffiths's conjecture was proven in* [2].

3.6. Special subvarieties and period maps

The following proposition follows from the definitions.

Proposition 3.20. *Let $\mathbb{V} = (\mathbb{V}_\mathbb{Z}, W_\bullet, F^\bullet)$ be a $\mathbb{Z}$VMHS over a smooth quasi-projective variety S with associated period map*

$$\Phi_S \colon S \longrightarrow \mathrm{Hod}^0(S, \mathbb{V}).$$

Let Z be an irreducible subvariety of S. The following conditions are equivalent:

(1) *Z is a special subvariety for $(S, \mathbb{V})$.*

(2) *Z is an irreducible component of the Φ_S-preimage of a special subvariety of $\mathrm{Hod}^0(S, \mathbb{V})$.*

Thus:

$$\mathrm{HL}(S, \mathbb{V}) = \Phi_S^{-1}\left(\Phi_S(S) \cap \mathrm{HL}(\mathrm{Hod}^0(S, \mathbb{V}))\right).$$

4. Atypicality and optimality in terms of period maps

Let $\mathbb{V} = (\mathbb{V}_{\mathbb{Z}}, W_\bullet, F^\bullet)$ be a $\mathbb{Z}$VMHS over a smooth quasi-projective variety S with associated period map

$$\Phi_S \colon S \longrightarrow \mathrm{Hod}^0(S, \mathbb{V}).$$

4.1. Hodge codimension and period maps

Let us clarify geometrically Definition 1.5 of Hodge codimension in terms of Φ_S and intersection theory.

It follows from the description of $\mathcal{D}_S$ that

$$\dim_{\mathbb{C}} \mathrm{Gr}_F^{-1} \mathfrak{p}_S = \mathrm{rk}\, T_h \mathrm{Hod}^0(S, \mathbb{V})$$

and

$$\mathrm{rk}\, \mathrm{Im} \overline{\nabla} = \dim_{\mathbb{C}} \Phi_S(S).$$

Hence:

Corollary 4.1. *The Hodge codimension* $\text{H-cd}(S, \mathbb{V})$ *as introduced in Definition 1.5 is equal to the codimension of the tangent space at a Hodge generic point of* $\Phi_S(S)$ *in the corresponding horizontal tangent space of* $\mathrm{Hod}^0(S, \mathbb{V})$*:*

$$\text{H-cd}(S, \mathbb{V}) = \mathrm{rk}\, T_h \mathrm{Hod}^0(S, \mathbb{V}) - \dim_{\mathbb{C}} \Phi_S(S).$$

In other words, the Hodge codimension $\text{H-cd}(S, \mathbb{V})$ of $(S, \mathbb{V})$ is the natural codimension of $\Phi_S(S)$ in $\mathrm{Hod}^0(S, \mathbb{V})$ once the Griffiths' transversality condition is taken into account.

Remark 4.2. In the simple case where S is a closed algebraic subvariety of a connected mixed Shimura variety $\mathrm{Sh}_K^0(\mathbf{P}, X)$, and $\mathbb{V}$ is the restriction to S of a variation of mixed Hodge structures on $\mathrm{Sh}_K^0(\mathbf{P}, X)$ associated to an algebraic representation of $\mathbf{P}$, then $\text{H-cd}(S, \mathbb{V})$ coincides with what Pink in [29] calls the *defect* of S in Sh: the codimension of S in its *special closure* (*i.e.* the smallest special subvariety of Sh containing S).

4.2. Atypicality and period maps

It follows that the atypicality condition (1.2) is better understood in terms of the period map $\Phi_S \colon S \to \mathrm{Hod}^0(S, \mathbb{V})$ and intersection theory. Let $Y \subset S$ be an irreducible algebraic subvariety. Let $\mathbf{P}_S$ and $\mathbf{P}_Y$ be the generic Mumford–Tate groups of $\mathbb{V}$ and $\mathbb{V}_{Y^{\mathrm{sm}}}$ respectively, with Lie algebras $\mathfrak{p}_S$ and $\mathfrak{p}_Y$ and Mumford–Tate domain $\mathcal{D}_S$ and $\mathcal{D}_Y$. The restriction $\Phi_{S|Y}$ of the period map Φ_S to Y factorises uniquely as

$$\begin{tikzcd} Y \arrow[r, "\Phi_Y"] \arrow[rr, bend right, "\Phi_{S|Y}"'] & \mathrm{Hod}^0(Y, \mathbb{V}_{Y^{\mathrm{sm}}}) \arrow[r, hook, "i_{Y,S}"] & \mathrm{Hod}^0(S, \mathbb{V}) \end{tikzcd} .$$

We thus obtain the following definition, equivalent to Definition 1.7:

Definition 4.3. Let S be an irreducible smooth quasi-projective variety and $\mathbb{V} \to S$ a $\mathbb{Q}$VMHS on S. A subvariety $Y \subset S$ is said to be atypical for $(S, \mathbb{V})$ if it is irreducible and

$$\dim_{\mathbb{C}} \Phi_S(S) - \dim_{\mathbb{C}} \Phi_S(Y) < \operatorname{rk} T_h \mathrm{Hod}^0(S, \mathbb{V}) - \operatorname{rk} T_h \mathrm{Hod}^0(Y, \mathbb{V}_{Y^{\mathrm{sm}}}); \quad (4.1)$$

i.e., an atypical subvariety of $(S, \mathbb{V})$ is an irreducible subvariety $Y \subset S$ such that the subvariety $\Phi_S(Y)$ of $\Phi_S(S)$ has an excess intersection with the Hodge locus $\mathrm{HL}(\mathrm{Hod}^0(S, \mathbb{V})) \subset \mathrm{Hod}^0(S, \mathbb{V})$.

Remark 4.4. If $Y \subset S$ is Hodge generic in S the atypicality condition (4.1) is equivalent to $\dim_{\mathbb{C}} \Phi_S(Y) > \dim_{\mathbb{C}} \Phi_S(S)$, and hence is never satisfied. Thus a subvariety $Y \subset S$ atypical for $\mathbb{V}$ is always contained in the Hodge locus $\mathrm{HL}(S, \mathbb{V})$.

Remark 4.5. It follows from Remark 1.6 and Definition 1.7 that if $Y \subset Y' \subset S$ if a pair of irreducible subvarieties, if $\mathbf{P}_Y = \mathbf{P}_{Y'}$ and if Y is atypical for $(S, \mathbb{V})$ then Y' is atypical for $(S, \mathbb{V})$.

Remark 4.6. If follows immediately from Definition 4.3 that any $Y \subset S$ atypical for $(S, \mathbb{V})$ is contained in a unique subvariety of S atypical for $(S, \mathbb{V})$, having $\mathbf{P}_Y$ as generic Mumford–Tate group, and maximal for these properties: the irreducible component of $\Phi_S^{-1}(\Phi_S(S) \cap \mathrm{Hod}^0(Y, \mathbb{V}_{Y^{\mathrm{sm}}}))$ containing Y. In particular maximal atypical subvarieties of S are special subvarieties of S.

Remark 4.7. As the set of special subvarieties for $(\mathbf{P}_S, X_S)$ is countable, the set of maximal atypical subvarieties for $(S, \mathbb{V})$ is countable. In particular it follows from Remark 4.6 that $S_{\mathrm{atyp}}(\mathbb{V}) \subset \mathrm{HL}(S, \mathbb{V}) \subset S$ is also a countable union of irreducible algebraic subvarieties of S.

Remark 4.8. If $f \colon S \to S'$ is a surjective morphism of smooth quasi-projective varieties and the $\mathbb{Z}$VMHS $\mathbb{V}$ on S is the pull-back $f^*\mathbb{V}'$ of some $\mathbb{Z}$VMHS $\mathbb{V}'$ on S' then $S_{\mathrm{atyp}}(\mathbb{V}) = f^{-1} S'_{\mathrm{atyp}}(\mathbb{V}')$.

4.3. Optimality and period maps

Similarly, the optimality condition is better understood in terms of period maps: an irreducible subvariety $Y \subset S$ is optimal for $(S, \mathbb{V})$ if it is not strictly contained in a better approximation to a special subvariety of $\mathrm{Hod}^0(S, \mathbb{V})$.

4.4. Examples

Example 4.9. Let S be a Shimura variety and let $\mathbb{V}$ be a standard variation of Hodge structure on S. Then $Y \subset S$ is atypical if and only if $\dim_{\mathbb{C}} Y > \operatorname{rk} T_h \mathrm{Hod}^0(Y, \mathbb{V}_{|Y^{\mathrm{sm}}})$, which is impossible. Hence $S_{\mathrm{atyp}} = \emptyset$.

Example 4.10 (Atypical special subvarieties of Shimura type with dominant period map).

Lemma 4.11. *Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth quasi-projective variety S. Let $Y \subset S$ be a special subvariety for $(S, \mathbb{V})$, of Shimura type and* *such that the period map $\Phi_Y : Y \to \mathrm{Sh}^0(Y)$ is dominant. Then Y is not atypical for $(S, \mathbb{V})$ if and only if $(\mathbf{P}_S, \mathcal{D}_S)$ is of Shimura type, the period map $\Phi_S \colon S \to \mathrm{Sh}^0(S, \mathbb{V})$ is dominant, and $\mathbb{V}$ is the restriction to S of a standard variation of mixed Hodge structures on $\mathrm{Sh}^0(S, \mathbb{V})$.*

Proof. Let $\Phi_S \colon S \to \mathrm{Hod}^0(S, \mathbb{V})$ be the period map for $(S, \mathbb{V})$. As Y is special of Shimura type with dominant period map, $\mathrm{rk}\, T_h \mathrm{Hod}^0(Y, \mathbb{V}_{Y^{\mathrm{sm}}}) = \dim_{\mathbb{C}} \Phi_S(Y)$ and the atypicality condition (4.1) for Y reads as:

$$\dim_{\mathbb{C}} \Phi_S(S) < \mathrm{rk}\, T_h \mathrm{Hod}^0(S, \mathbb{V}). \tag{4.2}$$

The variety $\Phi_S(S)$ is an horizontal subvariety of $\mathrm{Hod}^0(S, \mathbb{V})$ and hence the inequality $\dim_{\mathbb{C}} \Phi_S(S) \leq \mathrm{rk}\, T_h \mathrm{Hod}^0(S, \mathbb{V})$ always holds. It follows that the special subvariety Y is not atypical for $(S, \mathbb{V})$ if and only if $\dim_{\mathbb{C}} \Phi_S(S) = \mathrm{rk}\, T_h \mathrm{Hod}^0(S, \mathbb{V})$. This implies that the horizontal distribution $T_h \mathrm{Hod}^0(S, \mathbb{V})$ of $T\mathrm{Hod}^0(S, \mathbb{V})$ is integrable.

Suppose first that $\mathbb{V}$ is pure. In that case standard Lie theoretic considerations imply that $\mathcal{D}_S^+$ is a Hermitian symmetric domain. The existence of one algebraic leaf (namely S) for the foliation on $\mathrm{Hod}^0(S, \mathbb{V})$ defined by $T_h \mathrm{Hod}^0(S, \mathbb{V})$ implies that $T_h \mathrm{Hod}^0(S, \mathbb{V}) = T\mathrm{Hod}^0(S, \mathbb{V})$ (*i.e.* $\mathbb{V}$ is of unconstrained type in the terminology of [16, p. 12]). Thus $\mathbb{V}$ comes from a $\mathbb{Z}$VMHS on $\mathrm{Hod}^0(S, \mathbb{V})$ and hence $(\mathbf{P}_S, X_S)$ is of Shimura type by Proposition 3.6. The same conclusion holds in the mixed case by a classical dévissage to the pure case.

The facts that $\Phi_S \colon S \to \mathrm{Sh}^0(S, \mathbb{V})$ is dominant and $\mathbb{V}$ is the restriction to S of a standard variation of mixed Hodge structures on $\mathrm{Sh}^0(S, \mathbb{V})$ are now clear. □

Notice the following special case of Lemma 4.11:

Lemma 4.12. *Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth quasi-projective variety S. Let $x \in S$ be a CM-point. Then x is not atypical for $(S, \mathbb{V})$ if and only if $(\mathbf{P}_S, X_S)$ is of Shimura type and the period map $\Phi_S \colon S \to \mathrm{Sh}^0(S, \mathbb{V})$ is dominant.*

5. On the main conjecture and its corollaries

Proposition 5.1. *Conjectures 1.9, 1.10, 1.11 and 1.12 are equivalent.*

Proof. Obviously Conjecture 1.9 implies Conjecture 1.10, which in turns implies Conjecture 1.11. Let us show that Conjecture 1.11 implies Conjecture 1.9. So let S be an irreducible smooth quasi-projective variety and $\mathbb{V} \to S$ a $\mathbb{Z}$VMHS on S. Let $Y \subset S$ be an irreducible component of the Zariski-closure of $S_{\mathrm{atyp}}(\mathbb{V})$. It follows from Remark 4.6 that Y is the Zariski-closure of the union $\bigcup_{i \in I} X_i$, where I is a countable set and the X_i are pairwise distinct maximal atypical special subvarieties for $(S, \mathbb{V})$. There exists an index $i \in I$

such that the variety X_i is not atypical for $(Y, \mathbb{V}_{|Y})$: otherwise $Y_{\mathrm{atyp}}(\mathbb{V}_{|Y})$ would be Zariski-dense in Y, contradicting Conjecture 1.11. Then

$$\text{H-cd}(Y, \mathbb{V}_{|Y^{\mathrm{sm}}}) = \text{H-cd}(X_i, \mathbb{V}_{|X_i^{\mathrm{sm}}}) < \text{H-cd}(S, \mathbb{V}), \tag{5.1}$$

and hence Y is atypical in S. By maximality of the X_i's, it follows that the set I contains only one element $\{i\}$. Thus $Y = X_i$ is a special subvariety of $(S, \mathbb{V})$ and the result.

It remains to show that Conjectures 1.9 and 1.12 are equivalent. Assume that Conjecture 1.12 holds true. As maximal atypical subvarieties for $(S, \mathbb{V})$ are optimal for $(S, \mathbb{V})$ Conjecture 1.9 follows. Conversely assume that Conjecture 1.9 holds true and let us prove Conjecture 1.12 by induction on $s = \dim_{\mathbb{C}} S$. For $s = 1$ the variety S is a curve, optimal subvarieties of S are points and coincide with (maximal) atypical subvarieties of S, hence the result in this case. Suppose Conjecture 1.12 holds true for varieties of dimension at most $s-1$ and suppose S is of dimension s. Let $Z_1, \dots, Z_n$ be the finite collection of maximal atypical subvarieties for $(S, \mathbb{V})$. Let $(Y_i)_{i \in I}$ be a set of pairwise distinct optimal subvarieties for $(S, \mathbb{V})$. For each $i \in I$ the optimal subvariety $Y_i \subset S$ is atypical, and hence there exists $\phi(i) \in \{1, \dots, n\}$ such that $Y_i \subset Z_{\phi(i)}$. Moreover by the very definition of optimality either $Y_i = Z_{\phi(i)}$ or Y_i is optimal for $(Z_{\phi(i)}, \mathbb{V}_{|Z^{\mathrm{sm}}_{\phi(i)}})$. As $\dim_{\mathbb{C}} Z_{\phi(i)} < s$ it follows by the induction hypothesis that the set I is finite. Hence the result. □

5.1. Conjecture 1.9 as a generalized Zilber–Pink conjecture

Notice that the restriction of Conjecture 1.9 to the class of pairs $(S, \mathbb{V})$ where S is a subvariety of a Shimura variety $\mathrm{Sh}^0(S, \mathbb{V})$ and $\mathbb{V}$ is the restriction to S of a standard $\mathbb{Z}$VMHS on $\mathrm{Sh}^0(S, \mathbb{V})$ is the Zilber–Pink conjecture as stated for example by Pila ([27, Conj. 2.3]) (this conjecture was formulated by Zilber [38] in the case of multiplicative groups, it is a stronger version of the original Pink conjecture [29, Conj. 1.1]).

5.2. Conjecture 1.9 and CM-points: a generalized André–Oort conjecture

Conjecture 1.9 immediately implies:

Conjecture 5.2 (André–Oort conjecture for $\mathbb{Z}$VHS, version 1). *Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth irreducible quasi-projective variety S. Suppose that the union $\mathbf{SL}(S, \mathbb{V})$ of special subvarieties for $(S, \mathbb{V})$, which are of Shimura type with dominant period map, is Zariski-dense in S. Then the Hodge datum $(\mathbf{P}_S, X_S)$ is of Shimura type and the period map $\Phi_S \colon S \to \mathrm{Sh}^0(S, \mathbb{V})$ is dominant.*

Let us show how to deduce Conjecture 5.2 from our main Conjecture 1.9. Let $\Phi_S \colon S \to \mathrm{Hod}^0(S, \mathbb{V})$ be the period map for $\mathbb{V}$. By assumption the special subvarieties of Shimura type and dominant period map are Zariski-dense in S. It thus follows from Conjecture 1.11 and Noetherian induction that there exists (a Zariski-dense set of) not atypical special subvarieties of Shimura type with dominant period map in S. The conclusion follows from Lemma 4.11.

Notice that Conjecture 5.2 immediately implies, and, by density of CM-points in Shimura varieties, is in fact equivalent to, the following:

Conjecture 5.3 (André–Oort conjecture for $\mathbb{Z}$VHS, version 2). *Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth irreducible quasi-projective variety S. Suppose that the union of CM-points for $(S, \mathbb{V})$ is Zariski-dense in S. Then $(\mathbf{P}_S, X_S)$ is of Shimura type and the period map $\Phi_S \colon S \to \mathrm{Sh}^0(S, \mathbb{V})$ is dominant.*

Remark 5.4. When S is moreover defined over $\overline{\mathbb{Q}}$, Conjecture 5.3 was originally stated (in a slightly different way) by Green–Griffiths–Kerr: see [16, Conj. VIII. B.1 p. 275].

Notice that Conjecture 5.2 is equivalent to the union of the following two conjectures:

Conjecture 5.5 (classical André–Oort). *Let $Y \subset \mathrm{Sh}^0_K(\mathbf{P}, X)$ be a closed irreducible algebraic subvariety of a connected mixed Shimura variety. If Y contains a Zariski-dense set of CM-points then Y is a special subvariety of $\mathrm{Sh}_K(\mathbf{P}, X)$.*

Conjecture 5.6. *Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth irreducible quasi-projective variety S. Suppose that the set of CM-points for $(S, \mathbb{V})$ is Zariski-dense in S. Then $(\mathbf{P}_S, X_S)$ is of Shimura type.*

Many works have been devoted to the André–Oort conjecture 5.2, culminating to its proof when $\mathrm{Sh}^0_K(\mathbf{P}, X)$ is of Abelian type. We refer to [23] for a detailed analysis of the André–Oort conjecture and references to the related works. On the other hand Conjecture 5.6 is completely open.

The proof of the classical André–Oort conjecture 5.5 relies on two different ingredients: on the one hand a precise analysis of the arithmetic of Galois orbits of CM-points (lower bound and heights), on the other hand a geometric analysis of the distribution in $\mathrm{Sh}_K(\mathbf{P}, \mathcal{D})$ of *positive dimensional special subvarieties*. If $(S, \mathbb{V})$ is not of Shimura type, the arithmetic of CM-points for $(S, \mathbb{V})$ seems difficult to understand. On the other hand it seems worth focusing on positive dimensional special subvarieties, namely the following geometric part of Conjecture 5.2:

Conjecture 5.7. *Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth irreducible quasi-projective variety S. Suppose that the union $\mathbf{SL}^{>0}(S, \mathbb{V})$ of* positive dimensional *special subvarieties for $(S, \mathbb{V})$, which are of Shimura type with dominant period map, is Zariski-dense in S. Then $(\mathbf{P}_S, X_S)$ is of Shimura type and the period map $\Phi_S \colon S \to \mathrm{Sh}^0(S, \mathbb{V})$ is dominant.*

One could even ask the following question, more general than Conjecture 5.2:

Question 5.8. Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth irreducible quasi-projective variety S. Suppose that the union of special subvarieties for $(S, \mathbb{V})$ which are of Shimura type (but not necessarily with dominant period maps) is Zariski-dense in S. Is it true that necessarily $(\mathbf{P}_S, X_S)$ is of Shimura type?

6. Example: Calabi–Yau 3-folds

Let us describe the first non-trivial incarnation of Conjecture 5.2 outside of the world of Shimura varieties. Let X be a smooth projective Calabi–Yau threefold (*i.e.* the canonical bundle K_X is trivial and X has trivial fundamental group). Let $H_{\mathbb{Z}} := H^3(X, \mathbb{Z})$ with its natural polarized weight three $\mathbb{Z}$-Hodge structure $H_{\mathbb{C}} = H^{3,0} \oplus H^{2,1} \oplus H^{1,2} \oplus H^{0,3}$. One can endow $H_{\mathbb{Z}}$ with two different weight one Hodge structures:

- the Weil Hodge structure H^1_W for which $H^{1,0}_W := H^{0,3} \oplus H^{2,1}$;
- the Griffiths Hodge structure H^1_G for which $H^{1,0}_G := H^{3,0} \oplus H^{2,1}$.

These two weight one Hodge structures define two complex structures on the torus $H_{\mathbb{R}}/H_{\mathbb{Z}}$, the Weil intermediate Jacobian $J(X)_W := H_{\mathbb{Z}} \backslash H_{\mathbb{C}}/H^1_W$ which is an Abelian variety but does not vary holomorphically with X, and the Griffiths intermediate Jacobian $J(X)_G := H_{\mathbb{Z}} \backslash H_{\mathbb{C}}/H^1_G$ which is a mere complex torus but varies holomorphically with X. In [4] Borcea proves that $H_{\mathbb{Z}}$ has CM if and only if both the Hodge structures H^1_W and H^1_G have CM and their Mumford–Tate tori mutually commute in $\mathbf{GL}(H_{\mathbb{Q}})$. Let $S = \mathrm{Def}(X)$ be the family of Calabi–Yau threefolds deformation space of X and let $\mathbb{V}$ be the corresponding polarized weight 3 variation of $\mathbb{Z}$-Hodge structure on S^{sm} with fiber $H_{\mathbb{Z}}$ at X. One can choose X so that S contains infinitely many CM-points. To the best of my knowledge, in all examples (see [34], [37], [30]) the irreducible subvarieties of S containing a Zariski-dense set of CM-points and maximal for these properties are of Shimura type, as predicted by Conjecture 5.2. On the other hand it is not clear to me that there are only finitely many such subvarieties as predicted by Conjecture 1.9. Notice that a weaker version of Conjecture 5.2 in this case (and more generally for Calabi–Yau n-folds) already appears in [20].

7. Functional transcendence

One main tool for attacking Conjecture 1.9 or 5.2 consists in establishing functional transcendence statements for the period map Φ_S. It is adapted from the Pila–Zannier strategy for proving the André–Oort conjecture 5.5, hopefully using o-minimal techniques. We refer once more to [23] for a description of this strategy in the case of the André–Oort conjecture and focus here on the expected statements.

7.1. Weakly special subvarieties

These functional transcendence statements detect *weakly special subvarieties*, a generalisation of special subvarieties.

Definition 7.1. Consider any Hodge morphisms $R \xleftarrow{\pi} T \xrightarrow{i} Y$ between connected mixed Hodge varieties and any point $r \in R$. Then any irreducible component of $i(\pi^{-1}(r))$ is called a weakly special subvariety of Y.

Let $\mathbb{V} \to S$ be a $\mathbb{Z}$VMHS over a smooth quasi-projective base S with associated period map $\Phi_S\colon S \to \mathrm{Hod}^0(S, \mathbb{V})$. Any irreducible component of $\Phi_S^{-1}(\Phi_S(S) \cap Z)$, where Z is a weakly special subvariety of the connected mixed Hodge variety $\mathrm{Hod}^0(S, \mathbb{V})$ is called a weekly special subvariety for $(S, \mathbb{V})$.

In particular (taking R a point) special subvarieties of $(S, \mathbb{V})$ are weakly special for $(S, \mathbb{V})$.

7.2. Bi-algebraic geometry

The format for the functional transcendence statements we are interested in is the notion of bi-algebraic structure:

Definition 7.2. A bi-algebraic structure on a connected algebraic variety S is a pair

$$(D\colon \tilde{S} \longrightarrow X,\ h\colon \pi_1(S) \longrightarrow \mathrm{Aut}(X)),$$

where $\pi\colon \tilde{S} \to S$ denotes the universal cover, X is an algebraic variety, $\mathrm{Aut}(X)$ its group of algebraic automorphisms, $h\colon \pi_1(S) \to \mathrm{Aut}(X)$ is a group morphism (called the holonomy representation) and D is an h-equivariant holomorphic map (called the developing map).

Definition 7.3. Let S be a connected algebraic variety endowed with a bi-algebraic structure (D, h).

(i) An irreducible analytic subvariety $Y \subset \tilde{S}$ is said to be an irreducible algebraic subvariety of $\tilde{S}$ if Y is an analytic irreducible component of $D^{-1}(\overline{D(Y)}^{\mathrm{Zar}})$, where $\overline{D(Y)}^{\mathrm{Zar}}$ denotes the Zariski-closure of $D(Y)$ in X.

(ii) An irreducible algebraic subvariety $Y \subset \tilde{S}$, resp. $W \subset S$, is said to be *bi-algebraic* if $\pi(Y)$ is an algebraic subvariety of S, resp. any (equivalently one) analytic irreducible component of $\pi^{-1}(W)$ is an irreducible algebraic subvariety of $\tilde{S}$.

If $(S, \mathbb{V})$ is a $\mathbb{Z}$VMHS with lifted period map $\tilde{\Phi}_S\colon \tilde{S} \to \mathcal{D}_S^+$, we denote by

$$\hat{\Phi}_S\colon \tilde{S} \longrightarrow \hat{\mathcal{D}}_S$$

the composite $j_S \circ \tilde{\Phi}_S$ where $j_S\colon \mathcal{D}_S \hookrightarrow \hat{\mathcal{D}}_S$ denotes the open embedding of the Mumford–Tate domain $\mathcal{D}_S$ in its compact dual (see Section 3.1). The pair $(\hat{\Phi}_S\colon \tilde{S} \to \hat{\mathcal{D}}_S, \rho_S\colon \pi_1(S) \to \mathbf{P}_S(\mathbb{C}) \longrightarrow \mathrm{Aut}(\hat{\mathcal{D}}_S)(\mathbb{C}))$ defines a natural bi-algebraic structure on S, called the bi-algebraic structure of $(S, \mathbb{V})$. The relation between the bi-algebraic structure of $(S, \mathbb{V})$ and Hodge theory in the pure case is given by the following proposition, whose proof will be provided in a sequel to this note. It is proven by Ullmo–Yafaev [33] in the case where S is a Shimura variety, and in some special case by Friedman and Laza [14]:

Proposition 7.4. *Let $(S, \mathbb{V})$ be a $\mathbb{Z}$VHS. The bi-algebraic subvarieties of (S, V) are the weakly special subvarieties of $(S, \mathbb{V})$.*

A similar result should hold in the mixed case (see [15, Cor. 8.3] for the case of $\mathbb{Z}$VMHS of Shimura type).

7.3. Ax–Schanuel for variation of mixed $\mathbb{Z}$-Hodge structures

The main functional transcendence conjecture in this setting is:

Conjecture 7.5 (Ax–Schanuel for $\mathbb{Z}$VMHS). *Let $(S, \mathbb{V})$ be a $\mathbb{Z}$VMHS. Let $U \subset \tilde{S} \times S$ be an algebraic subvariety and let W be an irreducible component of $U \cap \Delta$ (where Δ denotes the graph of $\pi\colon \tilde{S} \to S$). Then $\mathcal{D}_U W \geq \dim \overline{W}^{ws}$, where $\overline{W}^{ws}$ denotes the smallest weakly special special subvariety of S containing $\pi(W)$.*

When applied to a subvariety $U \subset \tilde{S} \times S$ of the form $Y \times \overline{\pi(Y)}^{\mathrm{Zar}}$ for $Y \subset \tilde{S}$ algebraic, Conjecture 7.5 specializes to the following:

Conjecture 7.6 (Ax-Lindemann for $\mathbb{Z}$VMHS). *Let $(S, \mathbb{V})$ be a $\mathbb{Z}$VMHS. Let $Y \subset \tilde{S}$ be an algebraic subvariety. Then $\overline{\pi(Y)}^{\mathrm{Zar}}$ is a bi-algebraic subvariety of $(S, \mathbb{V})$, i.e. weakly special for $(S, \mathbb{V})$.*

Establishing Conjecture 7.6 is a crucial step in establishing Conjecture 5.7.

Note added in proof: since this paper was submitted, Conjecture 7.5 has been proven in [3]*, following the strategy of* [26] *for Shimura varieties.*

Acknowledgments

I would like to thank Emmanuel Ullmo for his comments and his suggestion to add the interpretation of our main conjecture in terms of optimal subvarieties (see Conjecture 1.12). I thank the referees for their careful reading, in particular for forcing me to clarify the (subtle) relation between my definition of mixed Hodge datum and Pink's definition of mixed Shimura datum (see Section 3.2).

References

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

[2] B. Bakker, Y. Brunebarbe, and J. Tsimerman. o-minimal GAGA and a conjecture of Griffiths. `http://arxiv.org/abs/1811.12230.pdf`.

[3] B. Bakker and J. Tsimerman. The Ax-Schanuel conjecture for variations of Hodge structures. *Invent. Math.*, 217(1):77–94, 2017.

[4] C. Borcea. Calabi-Yau threefolds and complex multiplication. In *Essays on mirror manifolds*, pages 489–502. Int. Press, Hong Kong, 1992.

[5] A. Borel. *Introduction aux groupes arithmétiques.* Publications de l'Institut de Mathématique de l'Université de Strasbourg, XV. Actualités Scientifiques et Industrielles, No. 1341. Hermann, Paris, 1969.

[6] 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.

[7] J.-L. Brylinski and S. Zucker. An overview of recent advances in Hodge theory. In *Several complex variables, VI*, volume 69 of *Encyclopaedia Math. Sci.*, pages 39–142. Springer, Berlin, 1990.

[8] E. Cattani, P. Deligne, and A. Kaplan. On the locus of Hodge classes. *J. Amer. Math. Soc.*, 8(2):483–506, 1995.

[9] C.-L. Chai. Density of members with extra Hodge cycles in a family of Hodge structures. *Asian J. Math.*, 2(3):405–418, 1998.

[10] 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.

[11] E. Colombo and G. P. Pirola. Some density results for curves with nonsimple Jacobians. *Math. Ann.*, 288(1):161–178, 1990.

[12] P. Deligne. Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*, Proc. Sympos. Pure Math., XXXIII, pages 247–289. Amer. Math. Soc., Providence, R.I., 1979.

[13] P. Deligne, J. S. Milne, A. Ogus, and K-y. Shih. *Hodge cycles, motives, and Shimura varieties*, volume 900 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1982.

[14] R. Friedman and R. Laza. On some Hermitian variations of Hodge structure of Calabi-Yau type with real multiplication. *Michigan Math. J.*, 63(1):83–99, 2014.

[15] Z. Gao. Towards the Andre-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.

[16] 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.

[17] M. Green, P. Griffiths, R. Laza, and C. Robles. Series of lectures at KIAS. May 2017.

[18] P. Griffiths, C. Robles, and D. Toledo. Quotients of non-classical flag domains are not algebraic. *Algebr. Geom.*, 1(1):1–13, 2014.

[19] P. A. Griffiths. Periods of integrals on algebraic manifolds. III. Some global differential-geometric properties of the period mapping. *Inst. Hautes Études Sci. Publ. Math.*, 38:125–180, 1970.

[20] S. Gukov and C. Vafa. Rational conformal field theories and complex multiplication. *Comm. Math. Phys.*, 246(1):181–210, 2004.

[21] E. Izadi. Density and completeness of subvarieties of moduli spaces of curves or abelian varieties. *Math. Ann.*, 310(2):221–233, 1998.

[22] M. Kashiwara. A study of variation of mixed Hodge structure. *Publ. Res. Inst. Math. Sci.*, 22(5):991–1024, 1986.

[23] B. Klingler, E. Ullmo, and A. Yafaev. Bi-algebraic geometry and the André-Oort conjecture. In *Algebraic geometry: Salt Lake City 2015*, volume 97 of *Proc. Sympos. Pure Math.*, pages 319–359. Amer. Math. Soc., Providence, RI, 2018.

[24] K. Liu and Y. Shen. Algebraicity of the image of period map. `https://arxiv.org/pdf/1608.06029.pdf`.

[25] J. S. Milne. Introduction to Shimura varieties. In *Harmonic analysis, the trace formula, and Shimura varieties*, volume 4 of *Clay Math. Proc.*, pages 265–378. Amer. Math. Soc., Providence, RI, 2005.

[26] N. Mok, J. Pila, and J. Tsimerman. Ax-Schanuel for Shimura varieties. *Ann. of Math.*, 189(3):945–978, 2019.

[27] J. Pila. O-minimality and Diophantine geometry. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. 1*, pages 547–572. Kyung Moon Sa, Seoul, 2014.

[28] R. Pink. Compactification of mixed Shimura varieties. Dissertation, Bonner Math. Schriften 209.

[29] R. Pink. A combination of the conjectures of Manin–Mumford, Mordell–Lang and André–Oort. In *Geometric methods in algebra and number theory*, volume 235 of *Progress Math.*, pages 251–282. Birkhäuser Boston, Boston, MA, 2005.

[30] J. C. Rohde. *Cyclic coverings, Calabi-Yau manifolds and complex multiplication*, volume 1975 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.

[31] A. J. Sommese. On the rationality of the period mapping. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 5(4):683–717, 1978.

[32] J. Steenbrink and S. Zucker. Variation of mixed Hodge structure. I. *Invent. Math.*, 80(3):489–542, 1985.

[33] E. Ullmo and A. Yafaev. A characterization of special subvarieties. *Mathematika*, 57(2):263–273, 2011.

[34] E. Viehweg and K. Zuo. Complex multiplication, Griffiths-Yukawa couplings, and rigidity for families of hypersurfaces. *J. Algebraic Geom.*, 14(3):481–528, 2005.

[35] C. Voisin. *Théorie de Hodge et géométrie algébrique complexe*, volume 10 of *Cours Spécialisés*. Société Mathématique de France, Paris, 2002.

[36] U. Zannier. *Some problems of unlikely intersections in arithmetic and geometry*, volume 181 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2012. With appendixes by David Masser.

[37] Y. Zhang. Some results on families of Calabi-Yau varieties. In *Mirror symmetry. V*, volume 38 of *AMS/IP Stud. Adv. Math.*, pages 361–378. Amer. Math. Soc., Providence, RI, 2006.

[38] B. Zilber. Exponential sums equations and the Schanuel conjecture. *J. London Math. Soc. (2)*, 65(1):27–44, 2002.

Bruno Klingler
Institut für Mathematik,
Humboldt Universität zu Berlin,
Unter den Linden 6,
10099 Berlin
GERMANY

e-mail: bruno.klingler@hu-berlin.de