Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1207-1229

Monotonicity in half-spaces of positive solutions to $-\Delta_p u = f(u)$ in the case $p > 2$

ALBERTO FARINA, LUIGI MONTORO AND BERARDINO SCIUNZI

Abstract. We consider weak distributional solutions to the equation $-\Delta_p u = f(u)$ in half-spaces under zero Dirichlet boundary condition. We assume that the nonlinearity is positive and superlinear at zero. For $p > 2$ (the case $1 < p \leq 2$ is already known) we prove that any positive solution is strictly monotone increasing in the direction orthogonal to the boundary of the half-space. As a consequence we deduce some Liouville-type theorems for the Lane-Emden-type equation. Furthermore any nonnegative solution turns out to be $C^{2,\alpha}$ smooth.

Mathematics Subject Classification (2010): 35B05 (primary); 35B65, 35J70 (secondary).

1. Introduction

We consider the problem

$$\begin{cases} -\Delta_p u = f(u) & \text{in } \mathbb{R}^N_+ \\ u(x', y) \geqslant 0 & \text{in } \mathbb{R}^N_+ \\ u(x', 0) = 0 & \text{on } \partial\mathbb{R}^N_+, \end{cases} \tag{1.1}$$

where $N \geq 2$ and $f(\cdot)$ satisfies:

(h_f) the nonlinearity f is *positive*, *i.e.*, $f(t) > 0$ for $t > 0$, it is locally Lipschitz continuous in $\mathbb{R}^+ \cup \{0\}$, and

$$\lim_{t \to 0^+} \frac{f(t)}{t^{p-1}} = f_0 \in \mathbb{R}^+ \cup \{0\}.$$

In the following we denote a generic point in $\mathbb{R}^N$ by (x', y) with $x' = (x_1, x_2, \ldots, x_{N-1})$ and $y = x_N$, we assume with no loss of generality that $\mathbb{R}^N_+ = \{y > 0\}$. Furthermore, according to the regularity results in [18,32,41] (see also the recent

Received November 19, 2015; accepted June 7, 2016.
Published online December 2017.

developments in [31,40]), we assume that $u \in C^{1,\alpha}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ and that is fulfills the equation in the weak distributional sense. Actually, in our case the regularity up to the boundary does not follow directly from [32] and an argument by reflection is needed. This is quite standard and will be described later on in this paper.

By the strong maximum principle (see [42]), it follows that any nonnegative nontrivial solution is actually (strictly) positive. In this case we study the monotonicity of the solution in the direction orthogonal to the boundary of the half-space.

The main tool is the Alexandrov-Serrin moving plane method that dates back to [1,39]. It is well known that the moving plane procedure allows one to prove monotonicity and symmetry properties of the solutions to general PDE's. In the case of bounded domains and in the semilinear case $p = 2$, this study was started in the celebrated papers [5,27]. In the case of unbounded domains the main examples, arising from many applications, are provided by the whole space $\mathbb{R}^N$ and by the half-space $\mathbb{R}^N_+$. We refer to [7,27,28] for the case of the whole space, where radial symmetry of the solutions is expected. In this paper we will address the case when the domain is a half-space. We refer the reader to [2–4,10,16,17,19,25,35] for previous results concerning monotonicity of the solutions in half-spaces, in the nondegenerate case.

The case of p-Laplace equations is really harder to study. In fact the p-Laplacian is a nonlinear operator and, as a consequence, comparison principles are not equivalent to maximum principles. The degenerate nature of the operator also causes lack of regularity of the solutions. Furthermore, in the case $p > 2$ that we are considering, the use of weighted Sobolev spaces is naturally associated to the study of qualitative properties of the solutions. This issue is more delicate in unbounded domains. We cannot describe here in full detail, this fact, that will be clarified in the body the paper. Let us only say that the use of weighted Sobolev spaces is necessary in the case $p > 2$, and it requires in turn the use of a weighted Poincaré type inequality with weight $\rho = |\nabla u|^{p-2}$ (see [13]). The latter involves constants that may blow up when the solution approaches zero, and thus may happen also for positive solutions in unbounded domains. Once again the lack of compactness plays an important role.

The first results in bounded domains and in the case $1 < p < 2$ were obtained in [12]. The case $p > 2$ requires the above-mentioned use of weighted Sobolev spaces and was solved in [13], for positive nonlinearities ($f(t) > 0$ for $t > 0$). In the case of the whole space, we refer the reader to the recent results in [11,38,43].

The first results concerning the p-Laplace operator and problems in half-spaces have been obtained in [15] in dimension two. The same techniques have been also exploited in the fully nonlinear case in [8]. In higher dimension, the first results have been obtained in the singular case $1 < p < 2$ in [21,23] (see also [26]) where positive locally Lipschitz continuous nonlinearities are considered. A partial answer in the more difficult degenerate case $p > 2$ was obtained in [22], where power-like nonlinearities are considered under the restriction $2 < p < 3$. Here, considering a larger class of nonlinearities, namely positive nonlinearities that are superlinear at zero, we remove the condition $2 < p < 3$ and prove the following:

Theorem 1.1. *Let* $p > 2$ *and let* $u \in C^{1,\alpha}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ *be a positive solution to* (1.1) *with* $|\nabla u| \in L^\infty(\mathbb{R}^N_+)$. *Then, under the assumption* (h_f), *it follows that*

$$\frac{\partial u}{\partial y} > 0 \quad in \quad \mathbb{R}^N_+.$$

As a consequence $u \in C^{2,\alpha'}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ *for some* $0 < \alpha' < 1$.

Our monotonicity result holds in particular for Lane-Emden type equations, namely in the case $f(u) = u^q$ with $q \geq p-1$. Note that, the case $q \leq p-1$, or more generally the case when, for some $t_0 > 0$, it holds

$$f(t) \geq c\, t^{p-1} \qquad \text{for } t \in [0, t_0],$$

is already contained in [22, Theorem 3]. Furthermore Theorem 1.1 is proved without *a-priori* assumptions on the behavior of the solution, that is, at infinity the solution may decay at zero in some regions, while it can be far from zero in some other regions. It is crucial that, in our result, only local regularity of the solution is required. Indeed assuming that the solution has summability properties at infinity, *i.e.*, assuming that the solution belongs to some Sobolev space, the monotonicity result is somehow easier to deduce and it generally leads to the nonexistence of such solutions (we refer to [33], see also [44]). Finally it is worth emphasizing that we prove the first step of the moving plane procedure in a very general setting. Indeed, in Theorem 3.1, we prove that any positive solution is monotone increasing near the boundary for any $1 < p < \infty$ only assuming that the nonlinearity f is continuous in $\mathbb{R}^+ \cup \{0\}$ and for some $T > 0$, it holds that $|f(t)| \leq \bar{k}\, t^{p-1}$ for $t \in [0, T]$ and some $\bar{k} = \bar{k}(T) > 0$.

The technique developed to prove Theorem 1.1 also allows us to deduce a monotonicity result for solutions to equations involving a different class of nonlinearities. We have the following

Theorem 1.2. *Let* $p > 2$ *and let* $u \in C^{1,\alpha}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+}) \cap W^{1,\infty}(\mathbb{R}^N_+)$ *be a positive solution to* (1.1). *Suppose that* $f(\cdot)$ *is locally Lipschitz continuous in* $\mathbb{R}^+ \cup \{0\}$ *and that there exists* $t_0 > 0$ *such that*

$$f(s) > 0 \quad for \quad 0 < t < t_0, \qquad f(s) < 0 \quad for \quad t > t_0.$$

Assume furthermore that

$$\lim_{t \to 0^+} \frac{f(t)}{t^{p-1}} = f_0 \in \mathbb{R}^+ \cup \{0\}, \qquad \lim_{t \to t_0} \frac{f(t)}{(t_0 - t)|t_0 - t|^{p-2}} = f^0 \in \mathbb{R}^+ \cup \{0\}. \quad (1.2)$$

Then

$$\frac{\partial u}{\partial y} > 0 \quad in \quad \mathbb{R}^N_+.$$

As a consequence, $u \in C^{2,\alpha'}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ *for some* $0 < \alpha' < 1$.

Theorem 1.2 is mainly a corollary of Theorem 1.1 and it extends to the degenerate case $p > 2$ earlier results in [23] (see Theorem 1.3 there and see also [21, Theorem 1.8]). It applies, for instance, to solutions of

$$-\Delta_p u = u\big(1-u^2\big)\big|1-u^2\big|^q,$$

where $q \geq p-2$. When $p = 2$ and $q = 0$, the above equation reduces to

$$-\Delta u = u\big(1-u^2\big),$$

which is the celebrated Allen-Cahn equation arising in a famous conjecture of De Giorgi.

The monotonicity of the solution implies in particular stability its, see [9,24]. This allows us to deduce some Liouville-type theorems. Following [9,20], we set

$$q_c(N,p) = \frac{[(p-1)N-p]^2 + p^2(p-2) - p^2(p-1)N + 2p^2\sqrt{(p-1)(N-1)}}{(N-p)[(p-1)N-p(p+3)]}.$$

We refer to [9,20] and the references therein for more details and we only note here that the exponent $q_c(N,p)$ is larger than the classical critical Sobolev exponent. Once that, we know that by Theorem 1.1, the solutions are monotone and therefore stable, the same proof of [22, Theorem 4] provides the following Liouville-type result:

Theorem 1.3. *Let $p > 2$ and let $u \in C^{1,\alpha}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ be a non-negative weak solution of* (1.1) *in $\mathbb{R}^N_+$ with $|\nabla u| \in L^\infty(\mathbb{R}^N_+)$ and*

$$f(u) = u^q.$$

Assume that

$$\begin{cases} (p-1) < q < \infty & \text{if} \quad N \leqslant \frac{p(p+3)}{p-1} \\ (p-1) < q < q_c(N,p) & \text{if} \quad N > \frac{p(p+3)}{p-1}. \end{cases}$$

Then $u = 0$. If moreover we assume that u is bounded, then it follows that $u = 0$ assuming only that

$$\begin{cases} (p-1) < q < \infty & \text{if} \quad (N-1) \leqslant \frac{p(p+3)}{p-1} \\ (p-1) < q < q_c((N-1),p) & \text{if} \quad (N-1) > \frac{p(p+3)}{p-1}. \end{cases}$$

The paper is organized as follows. In Section 2 we recall some known results for the reader's covenience. In Section 3 we prove some preliminary results and then we prove Theorems 1.1 and 1.2.

2. Preliminaries

We start by stating some notation and preliminary results. Generic fixed and numerical constants will be denoted by C (with subscript in some case) and they will be allowed to vary within a single line or formula.

For $0 \leq \alpha < \beta$, we define the strip $\Sigma_{(\alpha,\beta)}$ as

$$\Sigma_{(\alpha,\beta)} := \mathbb{R}^{N-1} \times (\alpha, \beta) \tag{2.1}$$

and we will denote by Σ_β the strip

$$\Sigma_\beta := \mathbb{R}^{N-1} \times (0, \beta).$$

Then we define the cylinder

$$\mathcal{C}_{(\alpha,\beta)}(R) = \mathcal{C}(R) := \Sigma_{(\alpha,\beta)} \cap \left\{B'(0, R) \times \mathbb{R}\right\}, \tag{2.2}$$

where $B'(0, R)$ is the ball in $\mathbb{R}^{N-1}$ of radius R and center at zero. Given $\lambda \in \mathbb{R}$ we will define $u_\lambda(x)$ by

$$u_\lambda(x) = u_\lambda(x', y) := u(x', 2\lambda - y) \quad \text{in } \Sigma_{2\lambda}. \tag{2.3}$$

Finally we use the notation

$$u^+ := \max\{u, 0\}.$$

In the sequel of the paper we will often use the strong maximum principle. We refer to [42] (see also [34]) and we recall here the statement.

Theorem 2.1 (Strong maximum principle and Hopf's lemma). *Let Ω be a domain in $\mathbb{R}^N$ and suppose that $u \in C^1(\Omega)$, $u \geqslant 0$ in Ω, weakly solves*

$$-\Delta_p u + cu^q = g \geqslant 0 \quad \text{in} \quad \Omega,$$

with $1 < p < \infty$, $q \geqslant p-1$, $c \geqslant 0$ and $g \in L^\infty_{\mathrm{loc}}(\Omega)$. If $u \not\equiv 0$ then $u > 0$ in Ω. Moreover for any point $x_0 \in \partial\Omega$ where the interior sphere condition is satisfied, and such that $u \in C^1(\Omega \cup \{x_0\})$ and $u(x_0) = 0$ we have that $\frac{\partial u}{\partial s} > 0$ for any inward directional derivative, that is, if y approaches x_0 in a ball $B \subseteq \Omega$ that has x_0 on its boundary, then $\lim_{y\to x_0} \frac{u(y)-u(x_0)}{|y-x_0|} > 0$.

Let us recall that the linearized operator $L_u(v, \varphi)$ for a fixed solution u of $-\Delta_p(u) = f(u)$ is well defined for every v, $\varphi \in H^{1,2}_\rho(\Omega)$ with $\rho \equiv |\nabla u|^{p-2}$, by

$$L_u(v, \varphi) \equiv \int_\Omega \left[|\nabla u|^{p-2}(\nabla v, \nabla\varphi) + (p-2)|\nabla u|^{p-4}(\nabla u, \nabla v)(\nabla u, \nabla\varphi) - f'(u)v\varphi\right].$$

We refer [13] for more details and in particular for the definition of the weighted Sobolev spaces involved. Let us only recall here that the space $H^{1,2}_\rho(\Omega)$ can be defined as the space of functions v such that $\|v\|_{H^{1,2}_\rho(\Omega)}$ is bounded and

$$\|v\|_{H^{1,2}_\rho(\Omega)} := \|v\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\Omega,\rho)}.$$

This is the same space obtained performing the completion of smooth functions under the norm above. The space $H^{1,2}_{0,\rho}(\Omega)$ is obtained taking the closure of $C^\infty_c(\Omega)$ under the same norm and $\|\nabla v\|_{L^2(\Omega,\rho)}$ is an equivalent norm in $H^{1,2}_{0,\rho}(\Omega)$.

Moreover, $v \in H^{1,2}_\rho(\Omega)$ is a weak solution of the linearized equation if

$$L_u(v,\varphi) = 0$$

for any $\varphi \in H^{1,2}_{0,\rho}(\Omega)$. By [13] we have that $u_{x_i} \in H^{1,2}_\rho(\Omega)$ for $i = 1, \dots, N$, and $L_u(u_{x_i}, \varphi)$ is well defined for every $\varphi \in H^{1,2}_{0,\rho}(\Omega)$, with

$$L_u(u_{x_i}, \varphi) = 0 \qquad \forall \varphi \in H^{1,2}_{0,\rho}(\Omega).$$

In other words, the derivatives of u are weak solutions of the linearized equation. Consequently by the strong maximum principle for the linearized operator (see [14]) we have the following

Theorem 2.2. *Let $u \in C^1(\overline{\Omega})$ be a weak solution of $-\Delta_p(u) = f(u)$ in a bounded smooth domain Ω of $\mathbb{R}^N$ with $\frac{2N+2}{N+2} < p < \infty$, f positive ($f(s) > 0$ for $s > 0$) and locally Lipschitz continuous. Then, for any $i \in \{1, \dots, N\}$ and any domain $\Omega' \subset \Omega$ with $u_{x_i} \geqslant 0$ in Ω', we have that either $u_{x_i} \equiv 0$ in Ω' or $u_{x_i} > 0$ in Ω'.*

We now state the weighted Poincaré-type inequality proved in [13] that will be useful in the sequel.

Theorem 2.3 (Weighted Poincaré-type inequality). *Let $w \in H^{1,2}_\rho(\Omega)$ be such that*

$$|w(x)| \leq \hat{C} \int_\Omega \frac{|\nabla w(y)|}{|x-y|^{N-1}} dy, \tag{2.4}$$

with Ω a bounded domain and $\hat{C}$ a positive constant. Let ρ be a weight function such that

$$\int_\Omega \frac{1}{\rho^\tau |x-y|^\gamma} dy \leq C^* \qquad \text{for any} \quad x \in \Omega, \tag{2.5}$$

with $\max\{(p-2), 0\} \leqslant \tau < p-1$, $\gamma < N-2$ ($\gamma = 0$ *if* $N = 2$). *Then*

$$\int_\Omega w^2 \leq C_p \int_\Omega \rho |\nabla w|^2, \tag{2.6}$$

where $C_p = C_p(d, C^*)$, with $d =$ diam (Ω). Moreover, $C_p \to 0$ if $d \to 0$.

We remark that, for the sake of simplicity and for the reader's convenience, we make explicit the dependence of C_p on the parameters, which will play a crucial role, in the sequel and that we need to control. The other parameters involved are fixed in our application and we refer the reader to Theorem 8 and to [22, Corollary 2 in Section 5] (see also [13]).

We will use the weighted Poincaré-type inequality with $\rho = |\nabla u|^{p-2}$. The next proposition gives some sufficient conditions for (2.5).

Proposition 2.4. *Let* $1 < p < \infty$ *and* $u \in C^{1,\alpha}(\Omega)$ *be a weak solution to*

$$-\Delta_p u = h(x) \quad \text{in } \Omega,$$

with $h \in W^{1,\infty}(\Omega)$. *Let* $\Omega' \subset\subset \Omega$ *and* $0 < \delta < \operatorname{dist}(\Omega', \partial\Omega)$ *and assume that* $h > 0$ *in* $\overline{\Omega'_\delta}$, *where*

$$\Omega'_\delta = \{x \in \Omega : d\big(x, \Omega'\big) < \delta\} \Subset \Omega.$$

Let us fix β_1, β_2 *such that*

$$\inf_{x \in \Omega'_\delta} h(x) \geq \beta_1 > 0 \quad \text{and} \quad \delta \geq \beta_2 > 0.$$

Then there exits a positive constant $C^* = C^*(\beta_1, \beta_2)$ *such that*

$$\int_{\Omega'} \frac{1}{|\nabla u|^\tau} \frac{1}{|x - y|^\gamma} \leqslant C^*,$$

with $\max\{(p-2)\,,\,0\} \leqslant \tau < p - 1$.

Remark 2.5. The proof of Proposition 2.4 follows from [13] (see also [36, 37]), however we refer to [22, Proposition 1 in Section 4] for the version stated here. Let us also point out that, as above, we prefer to omit the dependence of the constant C^* on other parameters that are fixed and therefore not relevant in our application.

Later we will frequently exploit the classical Harnack inequality for p-Laplace equations. We refer to [34, Theorem 7.2.1] and the references therein. At some point, as it will be clear later, it will be crucial the use of a boundary-type Harnack inequality. We thus state here an adapted version of the more general and deep result of M.F. Bidaut-Véron, R. Borghol and L. Véron, see [6, Theorem 2.8].

Theorem 2.6 (Boundary Harnack inequality). *Let* $R_0 > 0$ *define the cylinder* $\mathcal{C}_{(0,L)}(2R_0)$ *as in* (2.2) *and let* u *be such that*

$$-\Delta_p u = c(x) u^{p-1} \qquad \text{in} \quad \mathcal{C}_{(0,L)}(2R_0),$$

with u *vanishing on* $\mathcal{C}_{(0,L)}(2R_0) \cap \{y = 0\}$ *and with* $\|c(x)\|_{L^\infty(\mathcal{C}_{(0,L)}(2R_0))} \leq C_0$. *Then*

$$\frac{1}{C}\frac{u(z_2)}{\rho(z_2)} \leq \frac{u(z_1)}{\rho(z_1)} \leq C\frac{u(z_2)}{\rho(z_2)}, \quad \forall z_1, z_2 \in B_{R_0} \cap \mathcal{C}_{(0,L)}(2R_0) : 0 < \frac{|z_2|}{2} \leq |z_1| \leq 2|z_2|\,,$$

where $C = C(p, N, C_0)$ *and* $\rho(\cdot)$ *is the distance function to* $\partial\mathbb{R}^N_+$.

Finally, we state a lemma that will be useful in the proof of Proposition 3.3 below, see [21, Lemma 2.1].

Lemma 2.7. *Let* $\theta > 0$ *and* $\nu > 0$ *such that* $\theta < 2^{-\nu}$. *Let*

$$\mathcal{L} : (1, +\infty) \to \mathbb{R}$$

be a non-negative and non-decreasing function such that

$$\begin{cases} \mathcal{L}(R) \leq \theta \mathcal{L}(2R) & \forall R > 1 \\ \mathcal{L}(R) \leq C R^{\nu} & \forall R > 1\,. \end{cases}$$

Then $\mathcal{L}(R) = 0$.

3. Proof of Theorem 1.1

We will give the proof of Theorem 1.1 at the end of this section. Let us begin by showing that any positive solution to (1.1) is increasing in the y-direction near the boundary $\partial\mathbb{R}^N_+$. We prove such a result for problems involving a more general class of nonlinearities and for any $1 < p < \infty$. We have the following:

Theorem 3.1. *Let* $1 < p < \infty$ *and let* $u \in C^{1,\alpha}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ *be a positive weak solution to* (1.1) *with* $|\nabla u| \in L^{\infty}(\mathbb{R}^N_+)$. *Assume that the nonlinearity* f *is continuous in* $\mathbb{R}^+ \cup \{0\}$ *and, for some* $T > 0$, *it holds that*

$$|f(t)| \leq \bar{k}\, t^{p-1} \qquad \text{for} \quad t \in [0, T]$$

for some $\bar{k} = \bar{k}(T) > 0$. *Then it follows that there exists* $\lambda > 0$ *such that*

$$\frac{\partial u}{\partial y}\big(x', y\big) > 0 \quad \text{in } \ \Sigma_\lambda. \tag{3.1}$$

In particular the result holds true under the condition (h_f).

Proof. We argue by contradiction and we assume that there exists a sequence of points $P_n = (x'_n, y_n)$ such that

$$\frac{\partial u}{\partial y}\big(x'_n, y_n\big) \leq 0 \quad \text{and} \quad y_n \underset{n\to+\infty}{\longrightarrow} 0. \tag{3.2}$$

We consider the sequence $\hat{x}_n$ defined by $\hat{x}_n = \big(x'_n, 1\big)$. We set $\alpha_n = u\big(x'_n, 1\big)$, and

$$w_n\big(x', y\big) = \frac{u\big(x' + x'_n, y\big)}{\alpha_n}. \tag{3.3}$$

We remark that $w_n(0,1)=1$ and we have

$$\begin{aligned}-\Delta_p w_n(x) &= \frac{1}{\alpha_n^{p-1}} f\big(u(x'+x_n',y)\big)\\ &= \frac{1}{\alpha_n^{p-1}} \frac{f\big(u(x'+x_n',y)\big)}{u^{p-1}\big(x'+x_n',y\big)} u^{p-1}(x'+x_n',y) = c_n(x) w_n^{p-1}(x),\end{aligned} \tag{3.4}$$

for

$$c_n(x) = \frac{f\big(u(x'+x_n',y)\big)}{u^{p-1}\big(x'+x_n',y\big)}\,. \tag{3.5}$$

Since for any $L>0$ we have that $u \in L^\infty(\Sigma_{(L)})$ (by the Dirichlet condition and because $|\nabla u|$ is bounded in $\mathbb{R}^N_+$), by the assumption on the nonlinearity f, we obtain that

$$\|c_n(x)\|_{L^\infty(\Sigma_L)} \le \|c_n(x)\|_{L^\infty(\Sigma_{2L})} \le C_0(L). \tag{3.6}$$

Now we consider real numbers L, R and R_0 satisfying

$$0 < 2R_0 < 1 < R < L. \tag{3.7}$$

We claim that:

$$\|w_n\|_{L^\infty(\mathcal{C}_{(0,L)}(R))} \le C(L,R,R_0)\,.$$

Since $w_n(0,1)=1$, by the classical Harnack inequality, see [34, Theorem 7.2.1], we have that

$$\|w_n\|_{L^\infty(\mathcal{C}_{(0,L)}(R)\cap\{y\ge \frac{R_0}{4}\})} \le C_H^i(L,R,R_0)\,. \tag{3.8}$$

Now we apply Theorem 2.6 to deduce that

$$\|w_n\|_{L^\infty(\mathcal{C}_{(0,L)}(R)\cap\{y\le \frac{R_0}{4}\})} \le C_H^b(L,R,R_0). \tag{3.9}$$

To this end, let $\tilde{P}=(\tilde{x}',\tilde{y})$ be such that $\tilde{x}' \in B_R'(0)$ and $0<\tilde{y}<\frac{R_0}{4}$ and consider a corresponding point $\check{Q}=\big(\check{x}',0\big)$ such that

$$\check{x}' \in B'(0,R) \qquad \text{and} \qquad \tilde{P} \in \partial B_{R_0}(\check{Q})\,.$$

Recalling the choice $2R_0<R<L$, it is easy to check that such a point exists (and in general is not unique), see Figure 3.1.

By [6] (see Theorem 2.6) and recalling (3.6), we infer that

$$\frac{w_n(\tilde{P})}{\tilde{y}} \le C\,\frac{w_n(\check{x}',R_0)}{R_0},$$

and, recalling also that $w_n(x,0)=0$, we deduce that

$$\|w_n\|_{L^\infty(\mathcal{C}_{(0,L)}(R)\cap\{y\le \frac{R_0}{4}\}))} \le \frac{C}{4}\cdot C_H^i(L,R,R_0),$$

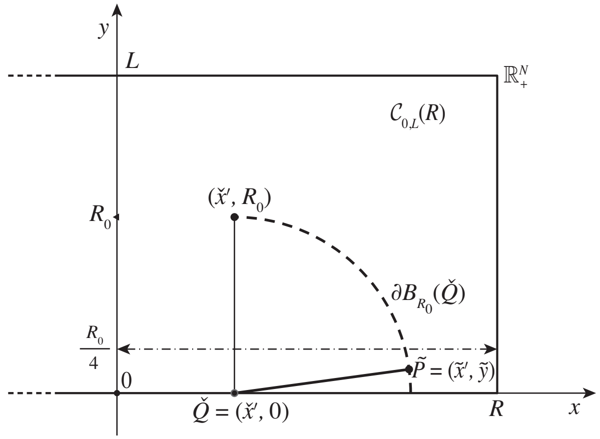

Figure 3.1.

that is (3.9) holds, with $C^b_H(L, R, R_0) = C \cdot C^i_H(L, R, R_0)$. Finally using (3.8) and (3.9) it follows that

$$\|w_n\|_{L^\infty(\mathcal{C}_{(0,L)}(R))} \leq C(L, R, R_0).$$

Now consider u (and consequently $u(x' + x'_n, y)$ in (3.3)), defined on the entire space $\mathbb{R}^N$ by odd reflection. That is

$$u\big(x', y\big) = -u\big(x', -y\big) \quad \text{in} \quad \{y < 0\},$$

and consequently

$$f(t) = -f(-t) \quad \text{if} \quad t < 0.$$

In this case we will refer to the cylinder

$$\mathcal{C}_{(-L,L)}(R) = B'_R(0) \times (-L, L).$$

By standard regularity theory, see, *e.g.*, [41, Theorem 1], since $\|w_n\|_{L^\infty(\mathcal{C}_{(-L,L)}(R))} \leq C(L, R, R_0)$, we have that

$$\|w_n\|_{C^{1,\alpha}_{\mathrm{loc}}(\mathcal{C}_{(-L,L)}(R))} \leqslant C(L, R, R_0)$$

for some $0 < \alpha < 1$. This allows us to use the Ascoli-Arzelà theorem to get

$$w_n \overset{C^{1,\alpha'}_{\mathrm{loc}}(\mathcal{C}_{(-L,L)}(R))}{\longrightarrow} w_0$$

up to subsequences, for $\alpha' < \alpha$. Furthermore, thanks to (3.6), we infer that

$$c_n(\cdot) \to c_0(\cdot) \tag{3.10}$$

weakly star in $L^{\infty}(\mathcal{C}_{(-L,L)}(R))$ up to subsequences. This and the fact that $w_0 \in C^{1,\alpha'}(\mathcal{C}_{(-L,L)}(R))$ allows us to deduce easily that

$$\begin{cases} -\Delta_p w_0 = c_0(x)\, w_0^{p-1} & \text{in } \mathcal{C}_{(0,L)}(R) \\ w_0(x', y) \geqslant 0 & \text{in } \mathcal{C}_{(0,L)}(R) \\ w_0(x', 0) = 0 & \text{on } \partial\mathcal{C}_{(0,L)}(R) \cap \partial\mathbb{R}^N_+ . \end{cases}$$

By the strong maximum principle, and recalling that $w_n(0, 1) = 1$ for all $n \in \mathbb{N}$, we deduce that $w_0 > 0$ in $\mathcal{C}_{(0,L)}(R)$ and, by Hopf's lemma, we infer that

$$\frac{\partial w_0}{\partial y}(0, 0) > 0 .$$

We conclude the proof noticing that a contradiction occurs since by (3.2) we should have that $\frac{\partial w_0}{\partial y}(0, 0) \leq 0$. □

Corollary 3.2. *Under the hypotheses of Theorem* 3.1, *there exists* $\lambda > 0$ *such that, for all* $0 < \theta \leq \frac{\lambda}{2}$, *it holds that*

$$u \leq u_\theta \quad \text{in } \Sigma_\theta .$$

Proof. Given λ from Theorem 3.1, using (3.1), it is sufficient to recall the definition of u_θ in (2.3). □

We now prove a technical result we are going to use in the sequel to prove our main result; we may refer to it as a weak comparison principle in narrow domains. We define the projection $\mathcal{P}$ as

$$\begin{aligned} \mathcal{P} : \mathbb{R}^N &\longrightarrow \mathbb{R}^{N-1} \\ (x', y) &\longrightarrow x' . \end{aligned}$$

In the proof of the next proposition, we will use the following inequalities: $\forall \eta, \eta' \in \mathbb{R}^N$ with $|\eta| + |\eta'| > 0$ there exists positive constants $\dot{C}, \check{C}$ depending on p such that

$$\begin{aligned} \big[|\eta|^{p-2}\eta - |\eta'|^{p-2}\eta'\big]\big[\eta - \eta'\big] &\geq \dot{C}\big(|\eta| + |\eta'|\big)^{p-2}\big|\eta - \eta'\big|^2, \\ \big||\eta|^{p-2}\eta - \big|\eta'\big|^{p-2}\eta'\big| &\leq \check{C}\big(|\eta| + \big|\eta'\big|\big)^{p-2}\big|\eta - \eta'\big|. \end{aligned} \tag{3.11}$$

Proposition 3.3. *Let* $p > 2$ *and let* $u \in C^{1,\alpha}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ *be a positive weak solution to* (1.1) *with* $|\nabla u| \in L^{\infty}(\mathbb{R}^N_+)$. *For* $0 \leq \alpha < \beta \leq \lambda$, *let* $\Sigma_{(\alpha,\beta)}$ *be the strip defined in* (2.1) *and assume that*

$$u \leq u_\lambda \qquad \text{on} \quad \partial\Sigma_{(\alpha,\beta)} . \tag{3.12}$$

Assume furthermore that, setting

$$\mathcal{I}^+_{(\lambda)} = \Big\{(x', \lambda) \,:\, x' \in \mathcal{P}\big(\operatorname{supp}(u - u_\lambda)^+\big)\Big\},$$

it holds that

$$u(x) \geq \gamma > 0 \quad on \quad \mathcal{I}^+_{(\lambda)}. \tag{3.13}$$

Then, for fixed $\Lambda > 0$ *such that* $\Lambda \geq 2\lambda + 1$, *there exists* $h_0 = h_0(f, p, \gamma, N, \|\nabla u\|_{L^\infty(\Sigma_\Lambda)})$ *such that if* $\beta - \alpha \leq h_0$ *we have* $u \leq u_\lambda \quad in \quad \Sigma_{(\alpha,\beta)}$.

Proof. Recalling that $u_\lambda(x', y) = u(x', 2\lambda - y)$, we remark that $(u - u_\lambda)^+ \in L^\infty(\Sigma_{(\alpha,\beta)})$ since we assumed $|\nabla u|$ is bounded. Let us now define

$$\Psi = (u - u_\lambda)^+ \varphi_R^2,$$

where $\varphi_R(x', y) = \varphi_R(x') \in C^\infty_c(\mathbb{R}^{N-1})$, $\varphi_R \geq 0$ such that

$$\begin{cases} \varphi_R \equiv 1 & \text{in } B'(0, R) \subset \mathbb{R}^{N-1} \\ \varphi_R \equiv 0 & \text{in } \mathbb{R}^{N-1} \setminus B'(0, 2R) \\ |\nabla \varphi_R| \leq \frac{C}{R} & \text{in } B'(0, 2R) \setminus B'(0, R) \subset \mathbb{R}^{N-1}, \end{cases} \tag{3.14}$$

where $B'(0, R)$ denotes the ball in $\mathbb{R}^{N-1}$ with center 0 and radius $R > 0$. From now on, for the sake of simplicity, we set $\varphi_R(x', y) := \varphi(x', y)$. By (3.14) and since $u \leq u_\lambda$ on $\partial\Sigma_{(\lambda,\beta)}$ (see (3.12)), it follows that $\Psi \in W^{1,p}_0(\mathcal{C}_{(\alpha,\beta)}(2R))$. Since u is a solution to problem (1.1), then it follows that u, u_λ are solutions to

$$\begin{cases} -\Delta_p u = f(u) & \text{in } \Sigma_{(\alpha,\beta)} \\ -\Delta_p u_\lambda = f(u_\lambda) & \text{in } \Sigma_{(\alpha,\beta)} \\ u \leq u_\lambda & \text{on } \partial\Sigma_{(\alpha,\beta)}. \end{cases} \tag{3.15}$$

Then using Ψ as a test function in both equations of problem (3.15) and substracting we get

$$\begin{aligned} &\int_{\mathcal{C}(2R)} \big(|\nabla u|^{p-2}\nabla u - |\nabla u_\lambda|^{p-2}\nabla u_\lambda, \nabla(u - u_\lambda)^+\big)\varphi^2 \\ &\quad + \int_{\mathcal{C}(2R)} \big(|\nabla u|^{p-2}\nabla u - |\nabla u_\lambda|^{p-2}\nabla u_\lambda, \nabla\varphi^2\big)(u - u_\lambda)^+ \\ &= \int_{\mathcal{C}(2R)} \big(f(u) - f(u_\lambda)\big)(u - u_\lambda)^+\varphi^2, \end{aligned} \tag{3.16}$$

where $\mathcal{C}(\cdot)$ denotes the cylinder defined in (2.2). By (3.11) and the fact that $p \geq 2$, from (3.16) we deduce that

$$
\begin{aligned}
&\dot{C} \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} |\nabla (u - u_\lambda)^+|^2 \varphi^2 \\
&\leq \int_{\mathcal{C}(2R)} \big(|\nabla u|^{p-2}\nabla u - |\nabla u_\lambda|^{p-2}\nabla u_\lambda, \nabla (u - u_\lambda)^+\big)\varphi^2 \\
&= - \int_{\mathcal{C}(2R)} \big(|\nabla u|^{p-2}\nabla u - |\nabla u_\lambda|^{p-2}\nabla u_\lambda, \nabla \varphi^2\big)(u - u_\lambda)^+ \\
&\quad + \int_{\mathcal{C}(2R)} \big(f(u) - f(u_\lambda)\big)(u - u_\lambda)^+ \varphi^2 \\
&\leq \int_{\mathcal{C}(2R)} \Big|\big(|\nabla u|^{p-2}\nabla u - |\nabla u_\lambda|^{p-2}\nabla u_\lambda, \nabla \varphi^2\big)\Big| (u - u_\lambda)^+ \\
&\quad + \int_{\mathcal{C}(2R)} \big(f(u) - f(u_\lambda)\big)(u - u_\lambda)^+ \varphi^2 \\
&\leq \check{C} \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} |\nabla (u - u_\lambda)^+| |\nabla \varphi^2| (u - u_\lambda)^+ \\
&\quad + \int_{\mathcal{C}(2R)} \big(f(u) - f(u_\lambda)\big)(u - u_\lambda)^+ \varphi^2,
\end{aligned}
\tag{3.17}
$$

where in the last line we used the Schwarz inequality and the second of (3.11). Setting

$$
I_1 := \check{C} \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u - u_\lambda)^+\big| \big|\nabla \varphi^2\big| (u - u_\lambda)^+ \tag{3.18}
$$

and

$$
I_2 := \int_{\mathcal{C}(2R)} \big(f(u) - f(u_\lambda)\big)(u - u_\lambda)^+ \varphi^2, \tag{3.19}
$$

equation (3.17) becomes

$$
\dot{C} \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u - u_\lambda)^+\big|^2 \varphi^2 \leq I_1 + I_2. \tag{3.20}
$$

In order to estimate the terms I_1 and I_2 in (3.20) we will use the weighted Poincaré-type inequality (2.6) (see [13]) and a covering argument that goes back to [22]. Let us consider the hypercubes Q_i of $\mathbb{R}^N$ defined by

$$
Q_i = Q_i' \times [\alpha, \beta],
$$

where $Q_i' \subset \mathbb{R}^{N-1}$ are hypercubes of $\mathbb{R}^{N-1}$ with edge $\beta - \alpha$ and such that

$$
\bigcup_i Q_i' = \mathbb{R}^{N-1}.
$$

Moreover we assume that $Q_i \cap Q_j = \emptyset$ for $i \neq j$ and

$$\bigcup_{i=1}^{\overline{N}} \overline{Q_i} \supset \mathcal{C}(2R). \tag{3.21}$$

It follows that each set Q_i has diameter

$$\operatorname{diam}(Q_i) = d_Q = \sqrt{N}(\beta - \alpha), \qquad i = 1, \ldots, \overline{N}. \tag{3.22}$$

The covering in (3.21) will allow us to use in each Q_i the weighted Poincaré-type inequality and to take advantage of the constant C_p in Theorem 2.3, that turns out not to be depending on the index i of (3.21). We will recollect the estimates later.

Let us define

$$w(x) = \begin{cases} \Big(u - u_\lambda\Big)^+(x', y) & \text{if } (x', y) \in \overline{Q}_i \\ -\Big(u - u_\lambda\Big)^+(x', 2\beta - y) & \text{if } (x', y) \in \overline{Q}_i^r, \end{cases} \tag{3.23}$$

where $(x', y) \in \overline{Q}_i^r$ if and only if $(x', 2\beta - y) \in \overline{Q}_i$. We claim that

$$\int_{Q_i} w^2 \leq C_p(Q_i) \int_{Q_i} (|\nabla u| + |\nabla u_\lambda|)^{p-2} |\nabla w|^2 \tag{3.24}$$

where $C_p(Q_i)$ is given by Theorem 2.3 and it goes to zero if the diameter of Q_i does. Actually, since $p \geq 2$, we will deduce (3.24) from

$$\int_{Q_i} w^2 \; \leq \; C_p(Q_i) \; \int_{Q_i} |\nabla u_\lambda|^{p-2} |\nabla w|^2 \,. \tag{3.25}$$

The fact that Theorem 2.3 can be applied to deduce (3.25) is somewhat technical and we describe the procedure below.

We have $\int_{Q_i \cup Q_i^r} w(x)dx = 0$ and therefore, see [29, Lemma 7.14, Lemma 7.16], it follows that

$$w(x) = \hat{C} \int_{Q_i \cup Q_i^r} \frac{(x_i - z_i) D_i w(z)}{|x - z|^N} dz \qquad \text{a.e. } x \in Q_i \cup Q_i^r,$$

where $\hat{C} = \hat{C}(d_Q, N)$, is a positive constant. Then for almost every $x \in Q_i$ we have

$$\begin{aligned} |w(x)| &\leq \hat{C} \int_{Q_i \cup Q_i^r} \frac{|\nabla w(z)|}{|x - z|^{N-1}} dz \\ &= \hat{C} \int_{Q_i} \frac{|\nabla w(z)|}{|x - z|^{N-1}} dz + \hat{C} \int_{Q_i^r} \frac{|\nabla w(z)|}{|x - z|^{N-1}} dz \\ &\leq 2\hat{C} \int_{Q_i} \frac{|\nabla w(z)|}{|x - z|^{N-1}} dz \,, \end{aligned}$$

where in the last inequality we used the following standard changing of variables

$$(z^t)' = z' \quad \text{and} \quad z^t_N = 2\beta - z_N,$$

the fact that for $x \in Q_i$ it holds that $(|x - z|)\big|_{z\in Q_i} \le (|x - z^t|)\big|_{z\in Q_i}$, and that by (3.23) it holds that $|\nabla w(z)| = |\nabla w(z^t)|$.

Hence (2.4) holds and, in order to prove (3.25), we need to show that (2.5) holds with

$$\rho = |\nabla u_\lambda|^{p-2}.$$

Note now that if w vanishes identically in Q_i, then there is nothing to prove. Otherwise it is easy to see that from our assumptions (see (3.13)) and the classical Harnack inequality, it follows that there exists $\bar\gamma > 0$ such that

$$u \ge \bar\gamma > 0 \qquad \text{in} \quad \tilde Q_i' \times [\lambda/2\,,\, 4\lambda], \tag{3.26}$$

where

$$\tilde Q_i' := \left\{x \in \mathbb{R}^{N-1} \;:\; \operatorname{dist}(x, Q_i') < 1\right\}.$$

Let us consider $Q_i^{\mathcal{R}_\lambda}$ obtained by the reflection of Q_i with respect to the hyperplane $T_\lambda = \{(x', y) \in \mathbb{R}^N \;:\; y = \lambda\}$. Since $Q_i^{\mathcal{R}_\lambda}$ is bounded away from the boundary $\mathbb{R}^N$, namely

$$\operatorname{dist}\left(Q_i^{\mathcal{R}_\lambda}, \{y = 0\}\right) \ge \lambda > 0$$

because of (3.26), then Proposition 2.4 applies with

$$\beta_1 = \min_{t\in[\bar\gamma, \|u\|_{L^\infty(\Sigma_\Lambda)}]} f(t) \quad \text{and} \quad \beta_2 = \lambda,$$

and we obtain that

$$\int_{Q_i^{\mathcal{R}_\lambda}} \frac{1}{|\nabla u|^{p-2}} \frac{1}{|x - y|^\gamma}\, dy \le C_1^*(\beta_1, \beta_2) \qquad \text{for any} \quad x \in Q_i^{\mathcal{R}_\lambda}.$$

By symmetry we deduce that

$$\int_{Q_i} \frac{1}{|\nabla u_\lambda|^{p-2}} \frac{1}{|x - y|^\gamma}\, dy \le C_1^*(\beta_1, \beta_2) \qquad \text{for any} \quad x \in Q_i,$$

so that we can exploit Theorem 2.3 to obtain (3.25) and consequently (3.24).

Let us now estimate the right-hand side of (3.20). Recalling (3.18) we get

$$\begin{aligned}
I_1 &= 2\check{C}\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{p-2}|\nabla(u-u_\lambda)^+|\varphi|\nabla\varphi|(u-u_\lambda)^+\\
&= 2\check{C}\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{\frac{p-2}{2}}|\nabla(u-u_\lambda)^+|\varphi(|\nabla u|+|\nabla u_\lambda|)^{\frac{p-2}{2}}|\nabla\varphi|(u-u_\lambda)^+\\
&\le \delta'\check{C}\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{p-2}|\nabla(u-u_\lambda)^+|^2\varphi^2\\
&\quad+\frac{\check{C}}{\delta'}\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{p-2}|\nabla\varphi|^2[(u-u_\lambda)^+]^2,
\end{aligned}$$

where in the last inequality we used the weighted Young inequality, with δ' to be chosen later. Hence

$$I_1 \le I_1^a + I_1^b, \tag{3.27}$$

where

$$\begin{aligned}
I_1^a &= \delta'\check{C}\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{p-2}\big|\nabla(u-u_\lambda)^+\big|^2\varphi^2,\\
I_1^b &= \frac{\check{C}}{\delta'}\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{p-2}|\nabla\varphi|^2\big[(u-u_\lambda)^+\big]^2.
\end{aligned}\tag{3.28}$$

Using the covering in (3.21), the properties of the cut-off function in (3.14) and the fact that $|\nabla u|$ and $|\nabla u_\lambda|$ are bounded, by (3.24) we deduce that

$$\begin{aligned}
I_1^b &\le \sum_{i=1}^{\overline{N}}\frac{C}{\delta' R^2}\int_{\mathcal{C}(2R)\cap Q_i}\big[(u-u_\lambda)^+\big]^2\\
&\le \max_i C_P(Q_i)\sum_{i=1}^{\overline{N}}\frac{C}{\delta' R^2}\int_{\mathcal{C}(2R)\cap Q_i}(|\nabla u|+|\nabla u_\lambda|)^{p-2}\big|\nabla(u-u_\lambda)^+\big|^2\\
&\le C_P^*\frac{C}{\delta' R^2}\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{p-2}\big|\nabla(u-u_\lambda)^+\big|^2
\end{aligned}\tag{3.29}$$

where $C_P^* = \max_i C_P(Q_i)$ and $C = C(p, \|\nabla u\|_{L^\infty(\Sigma_\Lambda)})$.

Now we estimate the term I_2 in (3.20). Since f is locally Lipschitz continuous because of (3.19), arguing as in (3.29), we get that

$$\begin{aligned}
I_2 &\le \int_{\mathcal{C}(2R)}\frac{f(u)-f(u_\lambda)}{u-u_\lambda}\big[(u-u_\lambda)^+\big]^2\\
&\le C_P^*\cdot C\int_{\mathcal{C}(2R)}(|\nabla u|+|\nabla u_\lambda|)^{p-2}\big|\nabla(u-u_\lambda)^+\big|^2,
\end{aligned}$$

where C_P^* is as in (3.29) and $C = C(f, \lambda, \|\nabla u\|_{L^\infty(\Sigma_\Lambda)})$. The constant C will depend on the Lipschitz constant of f in the interval $\big[0, \max\big\{\|u\|_{L^\infty(\Sigma_\Lambda)}, \|u_\lambda\|_{L^\infty(\Sigma_\Lambda)}]\big\}\big]$. By (3.20), (3.27), (3.28) and (3.29), up to redefining the constants, we obtain

$$
\begin{aligned}
&C \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u-u_\lambda)^+\big|^2 \varphi^2 \\
&\leq \delta' \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u-v)^+\big|^2 \\
&\quad + \frac{C_P^*}{R} \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u-u_\lambda)^+\big|^2 \\
&\quad + C_P^* \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u-u_\lambda)^+\big|^2 .
\end{aligned}
\tag{3.30}
$$

Let us choose δ' small in (3.30) such that $C - \delta' > C/2$ and fix $R > 1$. Then we obtain

$$
\begin{aligned}
&\int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u-u_\lambda)^+\big|^2 \varphi^2 \\
&\leq 4 \frac{C_P^*}{C} \int_{\mathcal{C}(2R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u-u_\lambda)^+\big|^2 .
\end{aligned}
\tag{3.31}
$$

To conclude we set now

$$
\mathcal{L}(R) := \int_{\mathcal{C}(R)} (|\nabla u| + |\nabla u_\lambda|)^{p-2} \big|\nabla (u-u_\lambda)^+\big|^2 . \tag{3.32}
$$

We can fix $h_0 = h_0(f, p, \gamma, \lambda, N, \|\nabla u\|_{L^\infty(\Sigma_\Lambda)})$ positive, such that if

$$
\beta - \alpha \leq h_0,
$$

(recall that $C_P^* \to 0$ in this case since $\operatorname{diam}(Q_i) \to 0$, see (3.22)) then

$$
\theta := 4 \frac{C_P^*}{C} < 2^{-N}.
$$

Then, by (3.31) and (3.32), we have

$$
\begin{cases}
\mathcal{L}(R) \leq \theta \mathcal{L}(2R) & \forall R > 1 \\
\mathcal{L}(R) \leq C R^N & \forall R > 1.
\end{cases}
$$

From Lemma 2.7 with $\nu = N$ and $\theta < 2^{-N}$, we get

$$
\mathcal{L}(R) \equiv 0
$$

and consequently that $(u - u_\lambda)^+ \equiv 0$. □

The proof of our main result will follow by the moving-plane procedure, strongly based on Proposition 3.3. As it will be clear later, it will be needed to substitute λ by $\lambda + \varepsilon$ in order to proceed further from the maximal position. To do this we need to be very accurate in the estimate of the constants involved, namely we need to control the role of h_0 in Proposition 3.3. This is the reason why we introduced the larger strip Σ_Λ, that allows us to control $|\nabla u|$. But we still need to control the dependence of h_0 on γ (see (3.13)). Equivalently we need a uniform control (with respect to ε) on the infimum of u far from the boundary, and in the set where u is greater than u_λ. This motivates the following:

Lemma 3.4. *Let $\lambda > 0$ and let u be a solution to* (1.1), *with $|\nabla u| \in L^\infty(\mathbb{R}^N_+)$ and u_λ defined as in* (2.3). *Assume here that (h_f) is fulfilled with $f_0 = 0$ and define*

$$\mathcal{I}^+_{(\lambda,\varepsilon)} = \Big\{(x', \lambda) \,:\, x' \in \mathcal{P}\big(\operatorname{supp}(u - u_{\lambda+\varepsilon})^+\big)\Big\}.$$

Then there exist $\varepsilon_0 > 0$ and $\gamma > 0$ such that

$$u(x) \geq \gamma \quad on \quad \mathcal{I}^+_{(\lambda,\varepsilon)},$$

for all $0 \leq \varepsilon \leq \varepsilon_0$.

Proof. By contradiction, given $\varepsilon_0 > 0$ and $\gamma > 0$, we find $0 \leq \varepsilon \leq \varepsilon_0$ and a point $Q_\varepsilon = (x'_\varepsilon, \lambda)$ with $Q_\varepsilon \in \mathcal{I}^+_{(\lambda,\varepsilon)}$ such that

$$u\big(x'_\varepsilon, \lambda\big) \leq \gamma.$$

It is convenient to consider $\varepsilon_0 = \gamma = 1/n$ and the corresponding $\varepsilon = \varepsilon_n \leq \varepsilon_0$ defined by contradiction as above, that obviously approaches zero as n tends to infinity. Also we use the notation $Q_{\varepsilon_n} \in \mathcal{I}^+_{(\lambda,\varepsilon_n)}$. On a corresponding sequence $P_n = (x'_n, y_n)$ we have that

$$u\big(x'_n, y_n\big) \geq u_{\lambda+\varepsilon_n}\big(x'_n, y_n\big) \quad \text{with } \big(x'_n, y_n\big) \in \Sigma_{\lambda+\varepsilon_n}, \tag{3.33}$$

where the existence of the sequence (x'_n, y_n) follows by the fact that $Q_{\varepsilon_n} \in \mathcal{I}^+_{(\lambda,\varepsilon_n)}$ and (up to subsequences) $y_n \to y_0 \in [0, \lambda]$. Moreover $\lim_{n\to+\infty} u\big(x'_n, \lambda\big) \to 0$. Let us set

$$w_n\big(x', y\big) = \frac{u\big(x' + x'_n, y\big)}{\alpha_n} \tag{3.34}$$

and $\alpha_n = u\big(x'_n, \lambda\big)$, with $\lim\limits_{n\to+\infty} \alpha_n = 0$. We remark that $w_n(0, \lambda) = 1$. Then we have

$$-\Delta_p w_n(x) = c_n(x) w_n^{p-1}(x), \tag{3.35}$$

for

$$c_n(x) = \frac{f\big(u(x' + x'_n, y)\big)}{u^{p-1}\big(x' + x'_n, y\big)}. \tag{3.36}$$

Since for any $L > 0$ we have that $u \in L^\infty(\Sigma_{(L)})$ (by the Dirichlet condition and because $|\nabla u|$ is bounded in $\mathbb{R}^N_+$), by (h_f) we obtain that

$$\|c_n(x)\|_{L^\infty(\Sigma_L)} \leq C(L). \tag{3.37}$$

For $L > \lambda$ we consider the cylinder $\mathcal{C}_{(0,L)}(R)$ and, arguing as in the proof of Theorem 3.1 (see the first claim there), we deduce that

$$\|w_n\|_{L^\infty(\mathcal{C}_{(0,L)}(R))} \leq C(L)\,.$$

Now, as in the proof of Theorem 3.1, we consider u defined on the entire space $\mathbb{R}^N$ by odd reflection and, by standard regularity theory (see [18,41]), we deduce that

$$\|w_n\|_{C^{1,\alpha}_{\mathrm{loc}}(\mathcal{C}_{(-L,L)}(R))} \leqslant C(L)$$

for some $0 < \alpha < 1$. This allows us to use the Ascoli-Arzelà theorem and get

$$w_n \overset{C^{1,\alpha'}_{\mathrm{loc}}(\mathcal{C}_{(-L,L)}(R))}{\longrightarrow} w_{L,R}$$

up to subsequences, for $\alpha' < \alpha$. Replacing L by $L + n$ ($n \in \mathbb{N}$), and R by $R + n$, we can repeat the argument above and then perform a standard diagonal process to define w in the entire space $\mathbb{R}^N$ in such a way that w is locally the limit of subsequences of w_n. It turns out that, by construction, setting

$$w_+(x) = w(x) \cdot \chi_{\overline{\mathbb{R}^N_+}}$$

we have that

$$\begin{cases} -\Delta_p w_+ = 0 & \text{in } \mathbb{R}^N_+ \\ w_+(x', y) \geqslant 0 & \text{in } \mathbb{R}^N_+ \\ w_+(x', 0) = 0 & \text{on } \partial\mathbb{R}^N_+\,. \end{cases}$$

This is a simple computation where in (3.35) we need to use the fact that $c_n(x) \to 0$ as $n \to +\infty$ uniformly on compact sets. This follows considering that w_n is uniformly bounded on compact sets and then, by (3.34), it follows that $u(x + x'_n, y) \to 0$ as $n \to +\infty$. By (3.36) and recalling that

$$\lim_{t \to 0} \frac{f(t)}{t^{p-1}} = 0,$$

finally it follows that $c_n(x) \to 0$ on compact sets.

By the strong maximum principle we have now that $w_+ > 0$, in view of the fact that (by uniform convergence of w_n) $w_+(0, \lambda) = 1$. By [30, Theorem 3.1], it follows that w_+ must be affine linear, *i.e.*, $w_+(x', y) = ky$, for some $k > 0$ by the

Dirichlet condition. If $y_0 \in [0, \lambda)$, by (3.33) and by the uniform convergence of $w_n \to w_+$, we would have

$$w_+(0, y_0) \geq (w_+)_\lambda(0, y_0).$$

This is a contradiction since $w_+(x', y) = ky$ for some $k > 0$.

Therefore let us assume that $y_n \to \lambda$ and note that, by the mean value theorem, at some point ξ_n lying on the segment from $(0, y_n)$ to $(0, 2(\lambda + \varepsilon_n) - y_n)$, it should hold that

$$\frac{\partial w_n}{\partial y}(0, \xi_n) \leq 0\,.$$

Since $w_n \to w_+$ in $C^{1,\alpha}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ we would have that

$$\frac{\partial w_+}{\partial y}(0, \lambda) \leq 0\,.$$

Again this is a contradiction since $w_+(x', y) = ky$, for some $k > 0$, and the result is proved. $\square$

The results proved above allow us to conclude the proof of our main result.

Proof of Theorem 1.1. We consider here the case when (h_f) is fulfilled with $f_0 = 0$ since in the simpler case $f_0 > 0$ the result follows directly by [22, Theorem 3]. Thanks to Corollary 3.2 we have that the set

$$\Lambda \equiv \{t > 0 \;:\; u \leqslant u_\alpha \quad \text{in} \quad \Sigma_\alpha \qquad \forall \alpha \leqslant t\}$$

is not empty. To conclude the proof, if we set

$$\bar{\lambda} = \sup \Lambda,$$

which now is well defined, we have to show that $\bar{\lambda} = +\infty$. By contradiction assume that $\bar{\lambda} < +\infty$ and set

$$W^+_\varepsilon := \left(u - u_{\bar{\lambda}+\varepsilon}\right)^+ \chi_{\Sigma_{\bar{\lambda}+\varepsilon}}.$$

We point out that given $0 < \delta < \bar{\lambda}/2$, there exists ε_0 such that for all $0 < \varepsilon \leq \varepsilon_0$ it follows that

$$\operatorname{supp} W^+_\varepsilon \subset \Sigma_\delta \,\cup\, \Sigma_{(\bar{\lambda}-\delta, \bar{\lambda}+\varepsilon)}.$$

This follows by an analysis of the limiting profile at infinity. We do not add the details since the proof is exactly the one in [21, Proposition 4.1]. For δ and ε_0 sufficiently small Proposition 3.3 applies in Σ_δ and in $\Sigma_{(\bar{\lambda}-\delta,\bar{\lambda}+\varepsilon)}$ with $\lambda = \bar{\lambda}+\varepsilon$ and $\Lambda = 2\bar{\lambda} + 1$. It is crucial here the fact that, thanks to Lemma 3.4, the parameter h_0 in the statement of Proposition 3.3 can be chosen independently of ε since there γ does not depend on ε. Then we conclude that $W^+_\varepsilon \equiv 0$. This is a contradiction with the definition of $\bar{\lambda}$, so we have proved that $\bar{\lambda} = \infty$. This implies the monotonicity of u in the half-space, that is $\frac{\partial u}{\partial y}(x) \geqslant 0$ in $\mathbb{R}^N_+$. By Theorem 2.2, since u is not trivial, it follows

$$\frac{\partial u}{\partial y}(x) > 0 \quad \text{in} \quad \mathbb{R}^N_+.$$

Finally, to prove that $u \in C^{2,\alpha'}_{\mathrm{loc}}(\overline{\mathbb{R}^N_+})$ just note that from the fact that $\frac{\partial u}{\partial y} > 0$ we deduce that the set of critical points $\{\nabla u = 0\}$ is empty and consequently the equation is no more degenerate. The $C^{2,\alpha'}$ regularity follows therefore by standard regularity results, see [29]. □

Proof of Theorem 1.2. By [21, Theorem 1.7] it follows that $0 < u \leq t_0$. Thanks to the behaviour of the nonlinearity near t_0 (see (1.2)), the strong maximum principle applies and implies that actually $0 < u < t_0$ in the half space. Arguing now as in the proof of [23, Theorem 1.3] it follows that u is strictly bounded away from t_0 in Σ_λ for any $\lambda > 0$. Now the monotonicity of the solution follows by our Theorem 1.1 (in the case $f_0 > 0$ the result follows also directly by [22, Theorem 3]). Note in fact that the condition (h_f) is satisfied in the range of values that the solutions takes in any strip and this is sufficient in order to run over again the moving plane procedure. □

References

[1] A. D. ALEXANDROV, *A characteristic property of the spheres*, Ann. Mat. Pura Appl. **58** (1962), 303–354.

[2] H. BERESTYCKI, L. CAFFARELLI and L. NIRENBERG, *Inequalities for second order elliptic equations with applications to unbounded domains*, Duke Math. J. **81** (1996), 467–494.

[3] H. BERESTYCKI, L. CAFFARELLI AND L. NIRENBERG, *Further qualitative properties for elliptic equations in unbounded domains*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), **25** (1997), 69–94.

[4] H. BERESTYCKI, L. CAFFARELLI and L. NIRENBERG, *Monotonicity for elliptic equations in an unbounded Lipschitz domain*, Comm. Pure Appl. Math. **50** (1997), 1089–1111.

[5] H. BERESTYCKI and L. NIRENBERG, *On the method of moving planes and the sliding method*, Bol. Soc. Bras. Mat., Nova Sér. **22** (1991), 1–37.

[6] M. F. BIDAUT-VÉRON, R. BORGHOL and L. VÉRON, *Boundary Harnack inequality and a priori estimates of singular solutions of quasilinear elliptic equations*, Calc. Var. Partial Differential Equations **27** (2006), 159–177.

[7] L. CAFFARELLI, B. GIDAS and J. SPRUCK, *Asymptotic symmetry and local behavior of semilinear elliptic equations with critical sobolev growth*, Comm. Pure Appl. Math. **42** (1989), 271–297.

[8] F. CHARRO, L. MONTORO and B. SCIUNZI, *Monotonicity of solutions of fully nonlinear uniformly elliptic equations in the half-plane*, J. Differential Equations **251** (2011), 1562–1579.

[9] L. DAMASCELLI, A. FARINA, B. SCIUNZI and E. VALDINOCI, *Liouville results for m-Laplace equations of Lane-Emden-Fowler type*, Ann. Inst. H. Poincaré Anal. Non Linéaire **26** (2009), 1099–1119.

[10] L. DAMASCELLI and F. GLADIALI, *Some nonexistence results for positive solutions of elliptic equations in unbounded domains*, Rev. Mat. Iberoam. **20** (2004), 67–86.

[11] L. DAMASCELLI, S. MERCHÁN, L. MONTORO and B. SCIUNZI, *Radial symmetry and applications for a problem involving the $-\Delta_p(\cdot)$ operator and critical nonlinearity in $\mathbb{R}^N$*, Adv. Math. **256** (2014), 313–335.

[12] L. DAMASCELLI and F. PACELLA, *Monotonicity and symmetry of solutions of p-Laplace equations, $1 < p < 2$, via the moving plane method*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **26** (1998), 689–707.

[13] L. DAMASCELLI and B. SCIUNZI, *Regularity, monotonicity and symmetry of positive solutions of m-Laplace equations*, J. Differential Equations **206** (2004), 483–515.

[14] L. DAMASCELLI and B. SCIUNZI, *Harnack inequalities, maximum and comparison principles, and regularity of positive solutions of m-Laplace equations*, Calc. Var. Partial Differential Equations **25** (2006), 139–159.

[15] L. DAMASCELLI and B. SCIUNZI, *Monotonicity of the solutions of some quasilinear elliptic equations in the half-plane, and applications*, Differential Integral Equations **23** (2010), 419–434.

[16] E. N. DANCER, *Some notes on the method of moving planes*, Bull. Aust. Math. Soc. **46** (1992), 425–434.

[17] E. N. DANCER, *Some remarks on half space problems*, Discrete Contin. Dyn. Syst. **25** (2009), 83–88.

[18] E. DI BENEDETTO, *$C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations*, Nonlinear Anal. **7** (1983), 827–850.

[19] A. FARINA, *Rigidity and one-dimensional symmetry for semilinear elliptic equations in the whole of $\mathbb{R}^N$ and in half spaces*, Adv. Math. Sci. Appl. **13** (2003), 65–82.

[20] A. FARINA, *On the classification of solutions of the Lane-Emden equation on unbounded domains of $\mathbb{R}^N$*, J. Math. Pures Appl. (9) **87** (2007), 537–561.

[21] A. FARINA, L. MONTORO and B. SCIUNZI, *Monotonicity and one-dimensional symmetry for solutions of $-\Delta_p u = f(u)$ in half-spaces*, Calc. Var. Partial Differential Equations **43** (2012), 123–145.

[22] A. FARINA, L. MONTORO and B. SCIUNZI, *Monotonicity of solutions of quasilinear degenerate elliptic equations in half-spaces*, Math. Ann. **357** (2013), 855–893.

[23] A. FARINA, L. MONTORO, G. RIEY and B. SCIUNZI, *Monotonicity of solutions to quasilinear problems with a first-order term in half-spaces*, Ann. Inst. H. Poincaré Anal. Non Linéaire **32** (2015), 1–22.

[24] A. FARINA, B. SCIUNZI and E. VALDINOCI, *Bernstein and De Giorgi type problems: new results via a geometric approach*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **7** (2008), 741–791.

[25] A. FARINA and E. VALDINOCI, *Flattening results for Elliptic PDEs in unbounded domains with applications to Overdetermined Problems*, Arch. Ration. Mech. Anal. **195** (2010).

[26] E. GALAKHOV, *A comparison principle for quasilinear operators in unbounded domains*, Nonlinear Anal. **70** (2009), 4190–4194.

[27] B. GIDAS, W. M. NI and L. NIRENBERG, *Symmetry and related properties via the maximum principle*, Comm. Math. Phys. **68** (1979), 209–243.

[28] B. GIDAS, W. M. NI and L. NIRENBERG, *Symmetry of positive solutions of nonlinear elliptic equations in $\mathbb{R}^N$*, Adv. Math. Suppl. Stud. **7A** (1981), 369–403.

[29] D. GILBARG and N. S. TRUDINGER, "Elliptic Partial Differential Equations of Second Order", Reprint of the 1998 Edition, Springer.

[30] T. KILPELÄINEN, H. SHAHGHOLIAN and X. ZHONG, *Growth estimates through scaling for quasilinear partial differential equations*, Ann. Acad. Sci. Fenn. Math. **32** (2007), 595–599.

[31] T. KUUSI and G. MINGIONE, *Universal potential estimates*, J. Funct. Anal. **262**, (2012), 4205–4269.

[32] G.M. LIEBERMAN, *Boundary regularity for solutions of degenerate elliptic equations*, Nonlinear Anal. **12** (1988), 1203–1219.

[33] C. MERCURI and M. WILLEM, *A global compacteness result for the p-Laplacian involving critical nonlinearities*, Discrete Contin. Dyn. Syst. **28** (2010), 469–493.
[34] P. PUCCI and J. SERRIN, "The Maximum Principle", Birkhäuser, Boston, 2007.
[35] A. QUAAS and B. SIRAKOV, *Existence results for nonproper elliptic equations involving the Pucci operator*, Comm. Partial Differential Equations **31** (2006), 987–1003.
[36] B. SCIUNZI, *Some results on the qualitative properties of positive solutions of quasilinear elliptic equations*, NoDEA Nonlinear Differential Equations Appl. **14** (2007), 315–334.
[37] B. SCIUNZI, *Regularity and comparison principles for p-Laplace equations with vanishing source term*, Commun. Contemp. Math. **16** (2014), 1450013, 20.
[38] B. SCIUNZI, *Classification of positive $\mathcal{D}^{1,p}(\mathbb{R}^N)$-solutions to the critical p-Laplace equation in $\mathbb{R}^N$*, Adv. Math. **291** (2016), 12–23.
[39] J. SERRIN, *A symmetry problem in potential theory*, Arch. Ration. Mech. Anal. **43** (1971), 304–318.
[40] E. V. TEIXEIRA, *Regularity for quasilinear equations on degenerate singular sets*, Math. Ann. **358** (2014), 241–256.
[41] P. TOLKSDORF, *Regularity for a more general class of quasilinear elliptic equations*, J. Differential Equations **51** (1984), 126–150.
[42] J. L. VAZQUEZ, *A strong maximum principle for some quasilinear elliptic equations*, Appl. Math. Optim. **12** (1984), 191–202.
[43] J. VÉTOIS, *A priori estimates and application to the symmetry of solutions for critical p-Laplace equations*, J. Differential Equations **260** (2016), 149–161.
[44] H. H. ZOU, *A priori estimates and existence for quasi-linear elliptic equations*, Calc. Var. Partial Differential Equations **33** (2008), 417–437.

Université de Picardie Jules Verne
LAMFA, CNRS UMR 7352
Amiens, France
alberto.farina@u-picardie.fr

Dipartimento di Matematica e Informatica
Università della Calabria
Ponte Pietro Bucci, 31B
I-87036 Arcavacata di Rende, Cosenza, Italia
montoro@mat.unical.it
sciunzi@mat.unical.it

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1231-1254

Classification of Kähler homogeneous manifolds of non-compact dimension two

RU AHMADI AND BRUCE GILLIGAN

Abstract. Suppose G is a connected complex Lie group and H is a closed complex subgroup such that $X := G/H$ is Kähler and the codimension of the top non-vanishing homology group of X with coefficients in $\mathbb{Z}_2$ is equal to two. We show that such an X has the structure of a holomorphic fiber bundle whose fiber and base are constructed from certain "basic building blocks", *i.e.*, $\mathbb{C}$, $\mathbb{C}^*$, Cousin groups, and flag manifolds.

Mathematics Subject Classification (2010): 32M10 (primary); 32Q15 (secondary).

1. Introduction

In this paper we consider complex homogeneous manifolds of the form G/H, where G is a connected complex Lie group and H is a closed complex subgroup of G. The existence of complex analytic objects on such a G/H, like non-constant holomorphic functions, plurisubharmonic functions and analytic hypersurfaces, is related to when G/H could be Kähler. So the first question one might consider concerns the existence of Kähler structures and we restrict ourselves to that question here. The structure of compact Kähler homogeneous manifolds is now classical [32] and [13] and the structure in the case of G-invariant metrics is also known [16]. Our investigations here concern non-compact complex homogeneous manifolds having a Kähler metric that is not necessarily G-invariant.

Some results are known under restrictions on the type of group G that is acting. The base of the holomorphic reduction of any complex solvmanifold is always Stein [28], where the proof uses some fundamental ideas in [31]. For G a solvable complex Lie group and G/H Kähler the fiber of the holomorphic reduction of G/H is a Cousin group, see [37] and the holomorphic reduction of a finite covering of G/H is a principal Cousin group bundle, see [20]. If G is semisimple, then G/H

This work was partially supported by an NSERC Discovery Grant.

Received October 6, 2015; accepted in revised form June 12, 2016.
Published online December 2017.

admits a Kähler structure if and only if H is algebraic [11]. For G reductive there is the characterization that G/H is Kähler if and only if $S \cdot H$ is closed in G and $S \cap H$ is an algebraic subgroup of S, a maximal semisimple subgroup of G, see [23, Theorem 5.1]. There is also a result if G is the direct product of its radical and a maximal semisimple subgroup under some additional assumptions on the isotropy subgroup and on the structure of G/H [36].

One way to proceed is to impose some topological restraints on $X := G/H$. In [18] we classified Kähler homogeneous manifolds X having more than one end by showing that X is either the product of a Cousin group of hypersurface type and a flag manifold or X admits a homogeneous fibration as a $\mathbb{C}^*$-bundle over the product of a compact complex torus and a flag manifold. Now in the setting of proper actions of Lie groups Abels introduced the notion of non-compact dimension, see [2] and [3, Section 2]. We do not wish to assume that our Lie group actions are necessarily proper ones, so we take a dual approach and define the non-compact dimension d_X of a connected smooth manifold X to be the codimension of the top non-vanishing homology group of X with coefficients in $\mathbb{Z}_2$, see Section 2. Our goal in this paper is to classify Kähler homogeneous manifolds G/H with $d_{G/H} = 2$. All such spaces are holomorphic fiber bundles where the fibers and the bases of the bundles involved consist of Cousin groups, flag manifolds, $\mathbb{C}$, and $\mathbb{C}^*$. We now present the statement of our main result, where T denotes a compact complex torus, C a Cousin group, and Q a flag manifold. Throughout the rest of this paper, if G is a mixed group, *i.e.*, is neither solvable nor semisimple, then S denotes a maximal semisimple subgroup of G. In particular, if G is simply connected, one has its Levi-Malcev decomposition $G = S \ltimes R$, where R is the radical of G.

Theorem 1.1 (Main theorem). *Suppose $X := G/H$ with $d_X = 2$, where G is a connected complex Lie group and H is a closed complex subgroup of G. Then X is Kähler if and only if X is one of the following:*

Case I. H **discrete:** *A finite covering of X is biholomorphic to a product $C \times A$, with C a Cousin group, A a Stein Abelian Lie group and $d_C + d_A = 2$.*

Case II: H **is not discrete:**

(1) *Suppose $\mathcal{O}(X) = \mathbb{C}$ and let $G/H \to G/N$ be its normalizer fibration;*
  - (a) *X is a $(\mathbb{C}^*)^k$-bundle over $C \times Q$ with $d_C + k = d_X = 2$;*
  - (b) *X is $T \times G/N$ with $\mathcal{O}(G/N) = \mathbb{C}$ and G/N fibers as a $\mathbb{C}$-bundle over a flag manifold; there are two subcases depending on whether S acts transitively on G/N or not;*

(2) *Suppose $\mathcal{O}(X) \neq \mathbb{C}$ and let $G/H \to G/J$ be its holomorphic reduction;*
  - (a) *$d_{G/J} = 2$ and G/J is Stein. Then*
    - (i) *$G/J = \mathbb{C}$ or*
    - (ii) *G/J is the 2-dimensional affine quadric, and in both of these cases $X = T \times Q \times G/J$ or*
    - (iii) *G/J is the complement of a quadric curve in $\mathbb{P}_2$, and X or a two-to-one covering of X is a product $T \times Q \times G/J$ or*
    - (iv) *$G/J = (\mathbb{C}^*)^2$ and a finite covering of X is $T \times Q \times G/J$;*

(b) $d_{G/J} = 2$ *and* G/J *is not Stein. Then a finite covering of* X *is biholomorphic to* $T \times Y$ *with* Y *a flag manifold bundle over the holomorphic reduction* G/J, *a* $\mathbb{C}^*$*-bundle over an affine cone minus its vertex;*

(c) $d_{G/J} = 1$ *and* G/J *is Stein. Then a finite covering of* X *is biholomorphic to* $C \times Q \times \mathbb{C}^*$, *where* $d_C = 1$*;*

(d) $d_{G/J} = 1$ *and* G/J *is not Stein. Then a finite covering of* X *is a* $\mathbb{C}^*$*-bundle over* $T \times \widetilde{Y}$, *where* $\widetilde{Y}$ *is the universal covering of* Y *which is a flag manifold bundle over the holomorphic reduction* G/J, *an affine cone minus its vertex. Moreover,* $d_{J/H} = 1$ *and* $\mathcal{O}(J/H) = \mathbb{C}$.

The paper is organized as follows. In section two we gather a number of technical tools. In particular, we note that Proposition 2.11 deals with the setting where the fiber of the normalizer fibration is a Cousin group and its base is a flag manifold. It is essential for Case II (1) (a) in the Main theorem and can be used to simplify the proof when $d_X = 1$ given in [18], see Remark 2.12. Section three is devoted to the case when the isotropy subgroup is discrete. Sections four and five deal with general isotropy and contain the proof of the classification when there are no non-constant holomorphic functions and when there are non-constant holomorphic functions, respectively. In section six we note that the manifolds listed in the classification are indeed Kähler. In the last section we present some examples.

ACKNOWLEDGEMENTS. We thank Prof. A. T. Huckleberry and the referees for their comments that led to significant improvements. We also thank the Arbeitsgruppe Transformationsgruppen at the Ruhr-Universität Bochum for their kind hospitality during a sojourn at which time part of this paper was written.

For H discrete this classification was presented in the first author's dissertation [4].

2. Technical tools

The purpose of this section is to collect a number of definitions and basic tools that are needed in the following.

2.1. Basic notions

Definition 2.1. A *Cousin group* is a complex Lie group G with $\mathcal{O}(G) = \mathbb{C}$. The terminology *toroidal group* is also found in the literature. Every Cousin group is Abelian and is the quotient of $\mathbb{C}^n$ by a discrete subgroup having rank $n+k$ for some k with $1 \leq k \leq n$. For details, we refer the reader to [1].

Definition 2.2. A *flag manifold* (the terminology *homogeneous rational manifold* is also in common usage) is a homogeneous manifold of the form S/P, where S is a connected semisimple complex Lie group and P is a parabolic subgroup of S. One source concerning the structure of flag manifolds is [8, Section 3.1].

Definition 2.3. For X a connected (real) smooth manifold we define

$$d_X := \dim_{\mathbb{R}} X - \min\{r \mid H_k(X, \mathbb{Z}_2) = 0 \quad \forall \quad k > r\},$$

i.e., d_X is the codimension of the top non-vanishing homology group of X with coefficients in $\mathbb{Z}_2$. We call d_X the *non-compact dimension* of X.

Remark 2.4. For later purposes we note that $d_X = 0$ if and only if X is compact.

Proposition 2.5. *Suppose X is a connected Stein manifold. Then* $\dim_{\mathbb{C}} X \leq d_X$.

Proof. For X Stein one has $H_k(X, \mathbb{Z}_2) = 0$ for all $k > \dim_{\mathbb{C}} X$ by [39]. □

2.2. Fibration methods

Throughout the paragraph we make use of a number of fibrations that are now classical.

(1) **Normalizer fibration:** Given G/H let $N = N_G(H^0)$ be the normalizer in G of the connected component of the identity H^0 of H. Since H normalizes H^0, we have $H \subset N$ and the *normalizer fibration* is given by $G/H \to G/N$;
(2) **Holomorphic reduction:** Given G/H we set $J := \{g \in G \mid f(gH) = f(eH)$ for all $f \in \mathcal{O}(G/H)\}$. Then J is a closed complex subgroup of G containing H and we call the fibration $p : G/H \to G/J$ the *holomorphic reduction* of G/H. By construction G/J is holomorphically separable and $\mathcal{O}(G/H) \cong p^*(\mathcal{O}(G/J))$.

Suppose a manifold X admits a locally trivial fiber bundle $X \xrightarrow{F} B$ with F and B connected smooth manifolds. One would then like to know how d_F and d_B are related to d_X whenever possible. The following result was proved in [7, Section 2] using spectral sequences.

Lemma 2.6 (The fibration lemma). *Suppose $X \xrightarrow{F} B$ is a locally trivial fiber bundle with X, F, B smooth manifolds. Then*

(1) *if the bundle is orientable* (e.g., *if* $\pi_1(B) = 0$), *then* $d_X = d_F + d_B$;
(2) *if B has the homotopy type of a q-dimensional CW complex, then* $d_X \geq d_F + (\dim B - q)$;
(3) *if B is homotopy equivalent to a compact manifold, then* $d_X \geq d_F + d_B$.

Remark 2.7. If B is homogeneous, then one knows that B is homotopy equivalent to a compact manifold if:

(1) the isotropy subgroup of B has finitely many connected components [35]; *e.g.*, in an algebraic setting;
(2) if B is a solvmanifold [34]; indeed, every solvmanifold is a vector bundle over a compact solvmanifold [9].

2.3. Special case of a question of Akhiezer

Later we will need a result that is based on [7, Lemma 8]. Since that lemma was stated in a way suitable for its particular application in [7], we reformulate it in a form suitable for the present context.

Lemma 2.8 ([7, Lemma 8]). *Let G be a connected, simply connected complex Lie group with Levi-Malcev decomposition $G = S \ltimes R$ with $\dim_{\mathbb{C}} R = 2$ and Γ a discrete subgroup of G such that $X = G/\Gamma$ is Kähler. Then Γ is contained in a subgroup of G of the form $A \ltimes R$, where A is a proper algebraic subgroup of S.*

This has the following consequence which we use later.

Theorem 2.9. *Suppose G is a connected, simply connected, complex Lie group with Levi-Malcev decomposition $G = S \ltimes R$ with $\dim_{\mathbb{C}} R = 2$. Let Γ be a discrete subgroup of G such that $X = G/\Gamma$ is Kähler, Γ is not contained in a proper parabolic subgroup of G and $\mathcal{O}(G/\Gamma) \simeq \mathbb{C}$. Then $S = \{e\}$,* i.e., *G is solvable.*

Proof. By Lemma 2.8 the subgroup Γ is contained in a proper subgroup of G of the form $A \ltimes R$, where A is a proper algebraic subgroup of S. Since $R \cdot \Gamma$ is closed in G, *e.g.*, see [19], there are fibrations

$$G/\Gamma \longrightarrow G/R \cdot \Gamma \longrightarrow S/A,$$

where $G/R \cdot \Gamma = S/\Lambda$ with $\Lambda := S \cap R \cdot \Gamma$. If A is reductive, then S/A is Stein and we get non-constant holomorphic functions on X as pullbacks using the above fibrations. But this contradicts the assumption that $\mathcal{O}(X) \simeq \mathbb{C}$. If A is not reductive then [29, Theorem 30.1] applies and A is contained in a proper parabolic subgroup of S. But this implies Γ is also contained in a proper parabolic subgroup of G, thus contradicting the assumption that this is not the case. $\square$

2.4. The algebraic setting revisited

Throughout this paragraph we repeatedly use two results of Akhiezer concerning the invariant d_X in the setting where $X = G/H$ and G is a connected linear algebraic group over $\mathbb{C}$ and H is an algebraic subgroup of G. For the convenience of the reader we now state these here.

Theorem 2.10 ($d = 1$ in [5]; $d = 2$ in [6]). *Suppose G is a connected linear algebraic group over $\mathbb{C}$, H is an algebraic subgroup of G and $X := G/H$.*

(1) $d_X = 1 \Longrightarrow$ *H is contained in a parabolic subgroup P of G with $P/H = \mathbb{C}^*$;*
(2) $d_X = 2 \Longrightarrow$ *H is contained in a parabolic subgroup P of G with P/H being:*
   (a) $\mathbb{C}$;
   (b) *the affine quadric Q_2;*
   (c) *the complement of a quadric curve in $\mathbb{P}_2$;*
   (d) $(\mathbb{C}^*)^2$.

2.5. Cousin group bundles over flag manifolds

In this section we prove a result concerning the structure of Kähler homogeneous manifolds whose normalizer fibrations are Cousin group bundles over flag manifolds, without assumptions on the invariant d. We show that one can reduce the problem to the case where a complex reductive group is acting transitively and employ some now classical details about the structure of parabolic subgroups, see [8] or [17]. A crucial point occurs in diagram (2.1) below, where the right vertical arrow is a holomorphic fiber bundle and the left vertical one is algebraic as a consequence of [23, Theorem 5.1].

Proposition 2.11. *Suppose $X := G/H$ is a Kähler homogeneous manifold whose normalizer fibration $G/H \to G/N$ has fiber N/H a Cousin group and base $Q := G/N$ a flag manifold. Then there exists a closed complex subgroup I of N containing H such that the fibration $G/H \to G/I$ realizes X as a $(\mathbb{C}^*)^k$-bundle over a product $G/I = Q \times C$, where C is a Cousin group with $d_C = d_X - k$.*

Proof. Our first task is to show that there is a reductive complex Lie group acting holomorphically and transitively on X. Write $N/H = \mathbb{C}^q/\Gamma$ and note that there exists a subgroup $\widehat{\Gamma} < \Gamma$ such that $\widehat{N/H} := \mathbb{C}^q/\widehat{\Gamma}$ is isomorphic to $(\mathbb{C}^*)^q$ and is a covering group of N/H, see [1, Section 1.1]. In particular, the reductive complex Lie group $\widehat{G} := S \times \widehat{N/H}$ acts transitively on X. We drop the hats from now on and assume, by considering a finite covering, if necessary, that $G = S \times Z$ is a reductive complex Lie group, where $Z \cong (\mathbb{C}^*)^q$ is the center of G and S is a maximal semisimple subgroup.

Since X is Kähler, the S-orbit $S \cdot H/H = S/S \cap H$ is closed in X and $S \cap H$ is an algebraic subgroup of S [23, Theorem 5.1]. Consider the induced fibration on the left hand side of the following diagram

$$
\begin{tikzcd}
S/S \cap H \arrow[r, hook] \arrow[d, "F"'] & G/H = X \arrow[d, "N/H"] \\
S/P \arrow[r, phantom, "="] & G/N
\end{tikzcd}
\tag{2.1}
$$

where $F := P/S \cap H$ is the induced fiber. The bundle $G/H \to G/N$ is defined by a representation $\rho : N \longrightarrow \mathrm{Aut}^0(N/H) \cong N/H$ with the group $\rho(N)$ lying in the connected component of the identity of the automorphism group of N/H since N is connected. Since N/H is Abelian, ρ factors through the canonical projection from N to $N/N' \cong P/P' \times Z$. Note that $P/P' \cong (\mathbb{C}^*)^p$, see [8, Proposition 8, Section 3.1]. Since $P/P' \times Z \cong (\mathbb{C}^*)^{p+q}$ is reductive, the factorized homomorphism is algebraic and the image $\rho(N)$ is a closed subgroup of N/H that is isomorphic to an algebraic torus $(\mathbb{C}^*)^k$ given as the quotient of $P/P' \times Z$ by an algebraic subgroup. Let $\sigma : N \to N/H^0 \to (N/H^0)/(H/H^0) = N/H$ be the composition of the quotient homomorphisms. The subgroup $I := \sigma^{-1} \circ \rho(N)$ is a closed, complex subgroup of N, and therefore of G, that contains H. Thus one has the fibration $G/H \to G/I$ whose typical fiber F is biholomorphic to $(\mathbb{C}^*)^k$. We claim that the bundle $G/I \to G/N$ is holomorphically trivial. This follows from the fact that the N-action on the neutral fiber of the bundle $G/I \to G/N$ is trivial. Otherwise, the dimension of the N-orbit in N/H would be bigger than k, as we assumed above, and this would give a contradiction. Finally, since N/H is a Cousin group, $C := N/I$ is also a Cousin group and the statement about the topological invariant follows because $d_{N/I} = d_{G/I}$, since S/P is compact and simply connected, and $d_{G/I} = d_X - k$. □

Remark 2.12. The case $d_X = 1$ is treated in [18, Proposition 5], where X is assumed to have more than one end. For X Kähler this is equivalent to $d_X = 1$.

3. The discrete case

Throughout this section we assume that $X = G/\Gamma$ is Kähler with $d_X = 2$, where G is a connected, simply connected, complex Lie group and Γ is a discrete subgroup of G. We first show that G is solvable. Then we prove that a finite covering of such an X is biholomorphic to a product $C \times A$, where C is a Cousin group and A is a holomorphically separable complex Abelian Lie group.

3.1. The reduction to solvable groups

We first handle the case when the Kähler homogeneous manifold has no nonconstant holomorphic functions.

Lemma 3.1. *Assume Γ is a discrete subgroup of a connected, simply connected complex Lie group G that is not contained in a proper parabolic subgroup of G, with $X := G/\Gamma$ Kähler, $\mathcal{O}(X) = \mathbb{C}$, and $d_X \leq 2$. Then G is solvable.*

Proof. Assume $G = S \ltimes R$ is a Levi decomposition. Since the R-orbits are closed, we have a fibration

$$G/\Gamma \longrightarrow G/R \cdot \Gamma = S/\Lambda,$$

where $\Lambda := S \cap R \cdot \Gamma$ is Zariski dense and discrete in S, see [19]. Now if $\mathcal{O}(R \cdot \Gamma/\Gamma) = \mathbb{C}$, then the result was proved in [19]. Otherwise, let

$$R \cdot \Gamma/\Gamma \longrightarrow R \cdot \Gamma/J =: Y$$

be the holomorphic reduction. Then Y is holomorphically separable and since R acts transitively on Y, it follows that Y is Stein [28]. One has $2 = d_X \geq d_Y \geq \dim_{\mathbb{C}} Y$. Further we claim that J^0 is a normal subgroup of G. In order to see this, note first that $\mathcal{O}(G/N_G(J^0)) = \mathbb{C}$ because one has the fibration $G/\Gamma \to G/N_G(J^0)$. If $N_G(J^0) \neq G$, then it follows from [27, Corollary 6] that $N_G(J^0)$ is contained in a proper parabolic subgroup of G. However, this implies that Γ is also contained in the same proper parabolic subgroup, which contradicts our assumptions. As a consequence, the quotient group $\widehat{R} := R/J^0$ has complex dimension one or two. If $\dim \widehat{R} = 1$, then $\widehat{G} := G/J^0$ is a product $S \times \widehat{R}$ and this implies $S = \{e\}$ by [36]. If $\dim \widehat{R} = 2$, then $\widehat{G}$ is either a product, see [36] again, or it is a non-trivial semidirect product. In the latter case the result follows from Theorem 2.9. □

In the next Proposition we reduce ourselves to the case when the maximal semisimple factor is $SL(2,\mathbb{C})$. We first prove a technical lemma in that setting.

Lemma 3.2. *Suppose G/Γ is Kähler and $d_{G/\Gamma} \leq 2$, where Γ is a discrete subgroup of a connected, complex Lie group of the form $G = SL(2,\mathbb{C}) \ltimes R$ with R the radical of G. Then Γ is not contained in a proper parabolic subgroup of G.*

Proof. Assume the contrary, *i.e.*, that Γ is contained in a proper parabolic subgroup and let P be a maximal such subgroup of G. Note that P is isomorphic to $B \ltimes R$, where B is a Borel subgroup of $SL(2,\mathbb{C})$. Let $P/\Gamma \to P/J$ be the holomorphic reduction. Then P/Γ is a Cousin group [37] and P/J is Stein [28]. Note that $J \neq P$, since otherwise P would be Abelian, giving a contradiction. The Fibration lemma and Proposition 2.5 imply $\dim_{\mathbb{C}} P/J = 1$ or 2. So P/J is biholomorphic to $\mathbb{C}$, $\mathbb{C}^*$, $\mathbb{C}^* \times \mathbb{C}^*$, or the complex Klein bottle.

In the first two cases P/J is equivariantly embeddable in $\mathbb{P}_1$ and by [30] it follows that G/J is Kähler. In the latter two cases the fiber J/Γ is compact by the Fibration lemma and we can push down the Kähler metric on X to obtain a Kähler metric on G/J, see [12]. In particular, the S-orbit $S/(S \cap J)$ in G/J is Kähler and so its isotropy $S \cap J$ is algebraic [11]. Now consider the diagram

$$\begin{array}{ccccc} G/\Gamma & \xrightarrow{F} & G/J & \xrightarrow{Y} & G/P \\ & & & & \| \\ \cup & & \cup & & \mathbb{P}_1 \\ & & & & \| \\ S/S\cap\Gamma & \xrightarrow{F_S} & S/S\cap J & \xrightarrow{Z} & S/B. \end{array}$$

Note that since $Y := P/J$ is noncompact and $d_{G/\Gamma} = 2$, it follows from the Fibration lemma that either $d_F = 1$ or F is compact. Since F is an Abelian Lie group, it is clear that $d_{F_S} \leq d_F$.

We list below, up to isomorphism, the algebraic subgroups of B and in each case we derive a contradiction.

(1) $\dim_{\mathbb{C}} S \cap J = 2$. Then $S \cap J = B$. This yields the contradiction $d_{S/S\cap\Gamma} \leq d_{F_S} + d_{S/B} = 1 + 0 = 1 < 3 = d_{S/S\cap\Gamma}$, since $S \cap \Gamma$ is finite;
(2) $\dim_{\mathbb{C}} S \cap J = 1$. There are two possible cases.
    (a) If $S \cap J = \mathbb{C}^*$, then $S/S \cap J$ is an affine quadric or the complement of a quadric curve in $\mathbb{P}_2$ and thus $Z = \mathbb{C}$. So $P/J \neq \mathbb{C}^*$ and it is either $\mathbb{C}$ or $(\mathbb{C}^*)^2$, *i.e.*, $d_{P/J} = 2$. Then the Fibration lemma implies that F is compact and, since the fiber F_S is closed in F, it must also be compact. But this forces $S \cap \Gamma$ to be an infinite subgroup of $S \cap J$ which is a contradiction;

(b) If $S \cap J = \mathbb{C}$, then $S/S \cap J$ is a finite quotient of $\mathbb{C}^2 \setminus \{(0,0)\}$ and so $Z = \mathbb{C}^*$. Now $P/J = \mathbb{C}$, $(\mathbb{C}^*)^2$ or $\mathbb{C}^*$. In the first two instances F would be compact and we get the same contradiction as in (a). In the last case $d_F = 1$ by the Fibration lemma and F_S is either compact or $\mathbb{C}^*$. Again $S \cap \Gamma$ is infinite with the same contradiction as in (a);

(3) $\dim_{\mathbb{C}} S \cap J = 0$. Here $S \cap J$ is finite, since it is an algebraic subgroup of B. Then $\dim S/S \cap J = 3$ and we see that $\dim G/J = 3$, since we know $\dim G/P = 1$ and $\dim P/J \le 2$. Then $P/J = (\mathbb{C}^*)^2$ and, since the fiber $S/S \cap J$ is both open and closed in G/J, it follows that $S/S \cap J = G/J$ and $d_{S/S\cap J} = 2$. But F is compact and thus so is F_S and we get the contradiction that $d_{S/S\cap\Gamma} = 2 < 3 = d_{S/S\cap\Gamma}$.

As a consequence, Γ is not contained in a proper parabolic subgroup of G. □

Proposition 3.3. *Suppose G/Γ is Kähler with $d_{G/\Gamma} \le 2$. Then G is solvable.*

Proof. First note that G cannot be semisimple. If that were so, then Γ would be algebraic, hence finite and thus G/Γ would be Stein. But then $2 = d_{G/\Gamma} \ge \dim_{\mathbb{C}} G/\Gamma = \dim_{\mathbb{C}} G$ which is a contradiction, since necessarily $\dim_{\mathbb{C}} G \ge 3$ for any complex semisimple Lie group G.

So assume $G = S \ltimes R$ is mixed. The proof is by induction on the dimension of G. Now if a proper parabolic subgroup of G contains Γ, then a maximal one does too, it is solvable by induction and thus it has the special form $B \ltimes R$, where B is isomorphic to a Borel subgroup of $S = SL(2,\mathbb{C})$. But this is impossible because of Lemma 3.2.

Lemma 3.1 handles the case $\mathcal{O}(G/\Gamma) = \mathbb{C}$. So we assume $\mathcal{O}(G/\Gamma) \ne \mathbb{C}$ with holomorphic reduction $G/\Gamma \to G/J$. The Main Result in [7] gives the following possibilities for the base G/J:

(1) $\mathbb{C}$;
(2) affine quadric Q_2;
(3) $\mathbb{P}_2 \setminus Q$, where Q is quadric curve;
(4) an affine cone minus its vertex;
(5) $\mathbb{C}^*$-bundle over an affine cone minus its vertex.

In case (1) the bundle is holomorphically trivial, its compact fiber being a torus, and the group that is acting effectively is solvable. In cases (2) and (3) we have fibrations $G/\Gamma \to G/J \to G/P = \mathbb{P}_1$ and so Γ is contained in a proper parabolic subgroup of G, contradicting what was shown in the previous paragraph.

In order to handle cases (4) and (5) we recall that an affine cone minus its vertex fibers equivariantly as a $\mathbb{C}^*$-bundle over a flag manifold. Thus we get fibrations

$$G/\Gamma \longrightarrow G/J \longrightarrow G/P.$$

Note that it cannot be the case that $G \neq P$, since then Γ would be contained in a proper parabolic subgroup, a possibility that has been ruled out. So $G = P$ and there is no flag manifold involved in this setting. Thus G/J (or a 2-1 covering) is biholomorphic to $\mathbb{C}^*$ or $(\mathbb{C}^*)^2$. In the second case the fiber J/Γ is compact and thus a torus, so G is solvable. If $G/J = \mathbb{C}^*$, then J/Γ is Kähler with $\dim J < \dim G$ and $d_{J/\Gamma} = 1$ by the Fibration lemma. By induction J is solvable and so G is solvable too, because $G/J = \mathbb{C}^*$. $\square$

3.2. A product decomposition

In order to prove the classification we need the following splitting result.

Proposition 3.4. *Suppose G is a connected, simply connected solvable complex Lie group that contains a discrete subgroup Γ such that G/Γ is Kähler and has holomorphic reduction $G/\Gamma \to G/J$ with base $(\mathbb{C}^*)^2$ and fiber a torus. Then a finite covering of G/Γ is biholomorphic to a product.*

Proof. First assume that J^0 is normal in G and let $\alpha : G \to G/J^0$ be the quotient homomorphism with differential $d\alpha : \mathfrak{g} \to \mathfrak{g}/\mathfrak{j}$. Then G/J^0 is a two dimensional complex Lie group that is either Abelian or solvable. In the Abelian case $G_0 := \alpha^{-1}(S^1 \times S^1)$ is a subgroup of G that has compact orbits in X, since these orbits fiber as torus bundles over $S^1 \times S^1$ in the base. The result now follows from [22, Theorem 6.14].

Next assume that J^0 is normal and G/J^0 is isomorphic to the two dimensional Borel group B with Lie algebra $\mathfrak{b}$. Let $\mathfrak{n}_{\mathfrak{b}}$ denote the nilradical of $\mathfrak{b}$. Then $\mathfrak{n} := d\alpha^{-1}(\mathfrak{n}_{\mathfrak{b}})$ is the nilradical of G. Let N denote the corresponding connected Lie subgroup of G. Now choose $\gamma_N \in \Gamma_N := N \cap \Gamma$ such that $\alpha(\gamma_N) \neq eJ^0$. There exists $x \in \mathfrak{n}$ such that $\exp(x) = \gamma_N$. Let U be the connected Lie group corresponding to $\langle \gamma_N \rangle_{\mathbb{C}}$. Since Γ centralizes J^0 (see [20, Theorem 1]), it follows that $\mathfrak{n} = \mathfrak{u} \oplus \mathfrak{j}$ and $N = U \times J^0$ is Abelian. Set $\Gamma_U := \Gamma \cap U$ and $\Gamma_J := \Gamma \cap J^0$. Then $N/\Gamma_N = U/\Gamma_U \times J^0/\Gamma_J$.

Since $\Gamma/\Gamma_N = \mathbb{Z}$, we may choose $\gamma \in \Gamma$ such that γ projects to a generator of Γ/Γ_N. Also set $A := \exp(\langle w \rangle_{\mathbb{C}})$ for fixed $w \in \mathfrak{g} \setminus \mathfrak{n}$. Since A is complementary to N, we have $G = A \ltimes N$. Now there exist $\gamma_A \in A$ and $\gamma_N \in N$ such that $\gamma = \gamma_A \cdot \gamma_N$. Both γ and γ_N centralize J^0 and thus γ_A does too. Also $\gamma_A = \exp(h)$ for some $h = sw$ with $s \in \mathbb{C}$. Therefore,

$$[h, \mathfrak{j}] = 0. \tag{3.1}$$

Since $\mathfrak{a} + \mathfrak{u}$ is isomorphic to $\mathfrak{b} = \mathfrak{g}/\mathfrak{j}$ as a vector space, there exists $e \in \mathfrak{u}$ such that

$$[d\alpha(h), d\alpha(e)] = 2d\alpha(e).$$

Let $\{e_1, \ldots, e_{n-2}\}$ be a basis for $\mathfrak{j}$. There exist structure constants a_i such that

$$[h, e] = 2e + \sum_{i=1}^{n-2} a_i e_i \ ,$$

and the remaining structure constants are all 0 by (3.1). Note that, conversely, any choice of the structure constants a_i determines a solvable Lie algebra $\mathfrak{g}$ and the corresponding connected simply-connected complex Lie group $G = A \ltimes N$.

We now compute the action of $\gamma_A \in A$ on N by conjugation. In order to do this note that the restriction $\mathrm{ad}_h : \mathfrak{n} \to \mathfrak{n}$ of ad_h to $\mathfrak{n}$ is expressed by the matrix

$$M \ := \ [\mathrm{ad}_h] = \begin{pmatrix} 2 & 0 \ldots 0 \\ a_1 & 0 \ldots 0 \\ \vdots & \vdots \ddots \vdots \\ a_{n-2} & 0 \ldots 0 \end{pmatrix}.$$

So the action of A on N is through the one parameter group of linear transformations $t \mapsto e^{tM}$ for $t \in \mathbb{C}$. For $k \geq 1$

$$(tM)^k = \frac{1}{2}(2t)^k M,$$

and it follows that

$$e^{tM} = \frac{1}{2}\left(e^{2t} - 1\right) M + \mathrm{Id}.$$

Since $\gamma_A \in \Gamma$ and Γ is a subgroup of J, the element $\alpha(\gamma_A)$, where $\alpha : G \to G/J^0$ is the quotient homomorphism defined above, acts trivially on the base $Y = G/J$. So $t = \pi i k$ where $k \in \mathbb{Z}$. Hence γ_A acts trivially on U. Also γ_N acts trivially on N, since N is Abelian. Thus γ acts trivially on N and as a consequence, although G is a non-Abelian solvable group the manifold $X = G/\Gamma$ is just the quotient of $\mathbb{C}^n$ by a discrete additive subgroup of rank $2n-2$. Its holomorphic reduction is the original torus bundle which, since we are now dealing with the Abelian case, is trivial.

Now assume J^0 is not normal in G, set $N := N_G(J^0)$, and let $G/J \xrightarrow{N/J} G/N$ be the normalizer fibration. Since the base G/N of the normalizer fibration is an orbit in some projective space, G/N is holomorphically separable and thus Stein [28]. Since we also have $2 \geq d_{G/N} \geq \dim_{\mathbb{C}} G/N$ we see that $G/N \cong \mathbb{C}, \mathbb{C}^*$ or $\mathbb{C}^* \times \mathbb{C}^*$. We claim that we must have $G/N = \mathbb{C}^*$, *i.e.*, we have to eliminate the other two possiblities. Assume $G/N \cong \mathbb{C}$. Since $d_X \leq 2$ the Fibration lemma implies $d_{N/J} = 0$, *i.e.*, N/J is biholomorphic to a torus T. Thus $G/J = T \times \mathbb{C}$. However, $G/J = \mathbb{C}^* \times \mathbb{C}^*$ giving a contradiction. Now assume $G/N \cong \mathbb{C}^* \times \mathbb{C}^*$. By Chevalley's theorem [15] the commutator group G' acts algebraically. Hence the G'-orbits are closed and one gets the commutator fibration $G/N \xrightarrow{\mathbb{C}} G/G' \cdot N$. Since G is solvable, it follows that G' is unipotent and the G'-orbits are cells, *i.e.*,

$G' \cdot N/N \cong \mathbb{C}$. By the Fibration lemma the base of the commutator fibration is a torus. But it is proved in [26] that the base of a commutator fibration is always Stein which is a contradiction. This proves the claim that $G/N \cong \mathbb{C}^*$ and by the Fibration lemma $d_{N/J} = 1$ and hence $N/J = \mathbb{C}^*$.

Since G is simply connected, G admits a Hochschild-Mostow hull [25], *i.e.*, there exists a solvable linear algebraic group

$$G_a = (\mathbb{C}^*)^k \ltimes G$$

that contains G as a Zariski dense, topologically closed, normal complex subgroup. By passing to a subgroup of finite index we may, without loss of generality, assume the Zariski closure $G_a(\Gamma)$ of Γ in G_a is connected. Then $G_a(\Gamma) \supseteq J^0$ and $G_a(\Gamma)$ is nilpotent [20]. Let $\pi : \widehat{G_a}(\Gamma) \to G_a(\Gamma)$ be the universal covering and set $\widehat{\Gamma} := \pi^{-1}(\Gamma)$. Since $\widehat{G_a}(\Gamma)$ is a simply connected, nilpotent, complex Lie group, the exponential map from the Lie algebra $\mathfrak{g}_a(\Gamma)$ to $\widehat{G_a}(\Gamma)$ is bijective. For any subset of $\widehat{G_a}(\Gamma)$ and, in particular for the subgroup $\widehat{\Gamma}$, the smallest closed, connected, complex (respectively real) subgroup $\widehat{G_1}$ (respectively $\widehat{G_0}$) of $\widehat{G_a}(\Gamma)$ containing $\widehat{\Gamma}$ is well-defined. By construction $\widehat{G_0}/\widehat{\Gamma}$ is compact, see [38, Theorem 2.1]. Set $G_1 := \pi(\widehat{G_1})$ and $G_0 := \pi(\widehat{G_0})$ and consider the CRS manifold (G_1, G_0, Γ). Note that the homogeneous CR-manifold G_0/Γ fibers as a T-bundle over $S^1 \times S^1$. In order to understand the complex structure on the base $S^1 \times S^1$ of this fibration consider the diagram

$$\begin{array}{ccccc}
\widehat{G}_0/\widehat{\Gamma} & \subset & \widehat{G}_1/\widehat{\Gamma} & \subseteq & \widehat{G}_a(\Gamma)/\widehat{\Gamma} \\
\| & & \| & & \| \\
G_0/\Gamma & \subset & G_1/\Gamma & \subseteq & G_a(\Gamma)/\Gamma \\
T\big\downarrow & & T\big\downarrow & & \big\downarrow T \\
\end{array}$$

$$S^1 \times S^1 = G_0/G_0 \cap (J^0 \cdot \Gamma) \subset G_1/J^0 \cdot \Gamma \subseteq G_a/J^0 \cdot \Gamma\,.$$

As observed in [20, Theorem 1], the manifold $G_a/J^0 \cdot \Gamma$ is a holomorphically separable solvmanifold and thus is Stein [28]. So $G_1/J^0 \cdot \Gamma$ is also Stein and thus $G_0/G_0 \cap (J^0 \cdot \Gamma)$ is totally real in $G_1/J^0 \cdot \Gamma$. The corresponding complex orbit $G_1/J^0 \cdot \Gamma$ is then biholomorphic to $\mathbb{C}^* \times \mathbb{C}^*$. It now follows by [20, Theorem 6.14] that a finite covering of G_1/Γ splits as a product of a torus with $\mathbb{C}^* \times \mathbb{C}^*$ and, in particular, that a subgroup of finite index in Γ is Abelian.

Now set $A := \{ \exp t\xi \mid t \in \mathbb{C} \}$, where $\xi \in \mathfrak{g} \setminus \mathfrak{n}$ and $\mathfrak{n}$ is the Lie algebra of N^0. Then $G = A \ltimes N^0$ and any $\gamma \in \Gamma$ can be written as $\gamma = \gamma_A.\gamma_N$ with $\gamma_A \in A$ and $\gamma_N \in N$. The fiber $G/\Gamma \to G/N$ is the N^0-orbit of the neutral point and Γ acts on it by conjugation. Since N/Γ is Kähler and has two ends, it follows by [18, Proposition 1] that a finite covering of N/Γ is biholomorphic to a product of the torus and $\mathbb{C}^*$. (By abuse of language we still call the subgroup having finite index Γ.) Now the bundle $G/\Gamma \to G/N$ is associated to the bundle

$$\mathbb{C} = G/N^0 \longrightarrow G/N = \mathbb{C}^*$$

and thus $G/\Gamma = \mathbb{C} \times_\rho (T \times \mathbb{C}^*)$, where $\rho : N/N^0 \to \mathrm{Aut}(T \times \mathbb{C}^*)$ is the adjoint representation. Since Γ is Abelian, this implies that γ_A acts trivially on $\Gamma_N := \Gamma \cap N^0$. Now suppose J has complex dimension k. Then γ_A is acting as a linear map on $N^0 = \mathbb{C} \ltimes J^0 = \mathbb{C}^{k+1}$ and commutes with the additive subgroup Γ_N that has rank $2k + 1$ and spans N^0 as a linear space. This implies γ_A that acts trivially on N^0 and, as a consequence, the triviality of a finite covering of the bundle, as required. □

3.3. The classification for discrete isotropy

In the following we classify Kähler G/Γ when Γ is discrete and $d_X \leq 2$. Note that $d_X = 0$ means X is compact and this is the now classical result of Borel-Remmert [13]; the case $d_X = 1$ corresponds to X having more than one end and was handled in [18].

Theorem 3.5 ([4]). *Let G be a connected simply connected complex Lie group and Γ a discrete subgroup of G such that $X := G/\Gamma$ is Kähler and $d_X \leq 2$. Then G is solvable and a finite covering of X is biholomorphic to a product $C \times A$, where C is a Cousin group and A is $\{e\}, \mathbb{C}^*, \mathbb{C}$, or $(\mathbb{C}^*)^2$. Moreover, $d_X = d_C + d_A$.*

Proof. By Proposition 3.3 G is solvable. If $\mathcal{O}(X) \cong \mathbb{C}$, then X is a Cousin group [37]. Otherwise, $\mathcal{O}(X) \neq \mathbb{C}$ and let

$$G/\Gamma \xrightarrow{J/\Gamma} G/J$$

be the holomorphic reduction. Its base G/J is Stein [28], its fiber J/Γ is biholomorphic to a Cousin group [37], and a finite covering of the bundle is principal [20]. Since G/J is Stein, by Proposition 2.5 one has

$$\dim_{\mathbb{C}} G/J \leq d_{G/J} \leq d_X \leq 2.$$

If $d_X = 1$, then $d_{G/J} = 1$ and G/J is biholomorphic to $\mathbb{C}^*$. A finite covering of this bundle is principal, with the connected Cousin group as structure group, and so is holomorphically trivial [24]. If $d_X = 2$, the Fibration lemma implies $G/J \cong \mathbb{C}$, $\mathbb{C}^*$, $\mathbb{C}^* \times \mathbb{C}^*$ or a complex Klein bottle [7]. The case of $\mathbb{C}^*$ is handled as above and a torus bundle over $\mathbb{C}$ is trivial by Grauert's Oka Principle [24]. Finally, since a Klein bottle is a 2-1 cover of $\mathbb{C}^* \times \mathbb{C}^*$, it suffices to consider the case $\mathbb{C}^* \times \mathbb{C}^*$. That case is handled by Proposition 3.4. □

Remark 3.6. This theorem proves the classification in Case I in the Main theorem. *i.e.*, if the isotropy is discrete. One should note that any complex manifold X that has a finite covering biholomorphic to $C \times A$, where C is a Cousin group and A a Stein Abelian Lie group, is Kähler.

4. The classification when $\mathcal{O}(X) = \mathbb{C}$

Proof. Let $\pi : G/H \to G/N$ be the normalizer fibration and recall that its base G/N is equivariantly embedded in some complex projective space $\mathbb{P}_q$. Let $\overline{G}$ denote the algebraic closure of the image of G in $PGL(q+1, \mathbb{C})$ and G' be the commutator group of G. Chevalley showed that $G' = \overline{G}'$ (see [15, Theorem 13, page 173] or [14, Corollary II.7.9]) and, as a consequence, G' is acting as an algebraic group on G/N. This fact and the fact that G' is normal in G imply the existence of the fibration $G/N \to G/N \cdot G'$. Now the base $G/N \cdot G'$ of the commutator fibration is an Abelian affine algebraic group that is Stein [26] and thus, because of the assumption $\mathcal{O}(G/H) = \mathbb{C}$, we also have $\mathcal{O}(G/N) = \mathbb{C}$ and the base $G/N \cdot G'$ must be a point. Otherwise, one could pullback non-constant holomorphic functions to G/N in order to obtain a contradiction. Since G' acts on G/N as an algebraic group of transformations and $d_{G/N} \le d_X = 2$, there is a parabolic subgroup P of G' containing $N \cap G'$ (see [6] or Theorem 2.10) and we now consider the fibrations

$$G/H \longrightarrow G/N = G'/N \cap G' \longrightarrow G'/P.$$

Our strategy in the remainder of the proof is to use the Fibration lemma 2.6 applied to each of the above fibrations and the information we know on the fiber $N/H = (N/H^0)/(H/H^0)$ of the normalizer fibration. Note that H/H^0 is a discrete subgroup of the complex Lie group N/H^0. Since $2 = d_{G/H} \ge d_{N/H}$ and N/H is Kähler whenever G/H is, Theorem 3.5 applies and a finite covering of N/H is biholomorphic to a product $C \times A$, where C is a Cousin group and A is a Stein Abelian Lie group with $d_C + d_A = d_{N/H}$. In particular, $A = \mathbb{C}^p \times (\mathbb{C}^*)^q$ by the classification of complex Abelian Lie groups, see [33, Theorem 3.2], and $d_A = 2p + q$. In addition, we have $2 = d_{G/H} \ge d_{G/N}$ and we look at the cases $d_{G/N} = 0$ (*i.e.*, G/N compact, see Remark 2.4), $d_{G/N} = 1$, and $d_{G/N} = 2$.

First we assume G/N is compact and thus a flag manifold, *i.e.*, $N \cap G' = P$ is a parabolic subgroup of G', and suppose $\mathcal{O}(N/H) = \mathbb{C}$. The fact that N/H is a Cousin group follows from the observations in the previous paragraph. The structure in this case is given in Proposition 2.11: X fibers as a $(\mathbb{C}^*)^k$-bundle over a product $Q \times C$, where Q a flag manifold and C is a Cousin group with $d_C + k = 2$.

Next suppose G/N compact and $\mathcal{O}(N/H) \ne \mathbb{C}$ with holomorphic reduction $N/H \to N/I$. Recall that I is a closed complex subgroup of G containing H and thus we get an intermediate fibration $G/H \to G/I$. In each case below we will show that $\mathcal{O}(G/I) \ne \mathbb{C}$ and this will contradict the assumption that $\mathcal{O}(G/H) = \mathbb{C}$. From what we noted above there are three possibilities for a finite covering of N/H:

(i) $N/I = \mathbb{C}^*$ and $I/H =: C$ is a Cousin group of hypersurface type;
(ii) $N/I = (\mathbb{C}^*)^2$ and $I/H = T$ is a torus;
(iii) $N/I = \mathbb{C}$ and $I/H = T$ is a torus.

In case (i) the space G/I fibers as a $\mathbb{C}^*$-bundle over the flag manifold G/N. Either this bundle is non-trivial and G/I is an affine cone minus its vertex or the bundle is trivial, and so a product $\mathbb{C}^* \times G/N$. In either case one has $\mathcal{O}(G/I) \neq \mathbb{C}$, the desired contradiction.

In case (ii) the space G/I fibers over the flag manifold G/N with typical fiber $N/I = \mathbb{C}^* \times \mathbb{C}^*$. We consider the possibilities for the S-orbits in G/I. First suppose that S acts transitively on $G/I = S/S \cap I$. Then $S \cap I$ is algebraic, since G/I is Kähler, [11]. By the main result in [6] the bundle $S/S \cap I \to S/S \cap P$ is principal. Let $J := \mathbb{C}^*$ be a subgroup of the structure group, *e.g.*, $J := \mathbb{C}^* \times \{e\}$, and consider the right J-action on $S/S \cap I$. This action equivariantly fibers $S/S \cap I$ as a $\mathbb{C}^*$-bundle over a $\mathbb{C}^*$-bundle over $S/S \cap P$. We then proceed as in the previous paragraph. If the S-orbits have complex codimension one in G/I, then we have again a $\mathbb{C}^*$-bundle over a $\mathbb{C}^*$-bundle over G/N, where the latter space still has non-constant holomorphic functions.

Finally, if any S-orbit has complex codimension two in G/I, then it is a section of this bundle because the flag manifold G/N is simply connected. Indeed, we claim that all S-orbits are sections in this setting. As a consequence, G/I splits as a product $\mathbb{C}^* \times \mathbb{C}^* \times G/N$ and again $\mathcal{O}(G/I) \neq \mathbb{C}$.

In (iii) the group action on the fiber $\mathbb{C}$ is by affine transformations [6] and the S-action on the space G/I is transitive. This gives us the following diagram

$$\begin{array}{ccccc} G/I & = & S/S \cap I & = & G/I \\ & & \downarrow & & \downarrow \\ G/N & = & S/S \cap P & = & G/N\,. \end{array}$$

By [10, Proposition 1 in Section 5.2] the group $S \cap I$ contains a maximal torus of $S \cap P$. This implies $S \cap I$ is normal in $S \cap P$ if and only if $S \cap I$ coincides with $S \cap P$ and this is not the case in our setting, since $\dim_{\mathbb{C}} N/I = 1 > 0$. Now the S-orbits are transversal to this one dimensional fiber and thus are coverings of the flag manifold G/N. Again, the fact that G/N is simply connected implies that these orbits are sections and the bundle $G/I \to G/N$ is a product. Once more we have $\mathcal{O}(G/I) \neq \mathbb{C}$.

Now suppose $d_{G/N} = 1$. As noted above, G' acts algebraically and transitively on G/N. It then follows that G/N is an affine cone minus its vertex by [5] or as recalled in Theorem 2.10, and consequently $\mathcal{O}(G/N) \neq \mathbb{C}$ contradicting the assumption that $\mathcal{O}(G/H) = \mathbb{C}$.

Suppose $d_{G/N} = 2$ and a finite covering of $P/N \cap G'$ is biholomorphic to $(\mathbb{C}^*)^2$. An argument analogous to the one given above in (ii) now yields a contradiction.

Finally, suppose $d_{G/N} = 2$ and $P/N \cap G' = \mathbb{C}$. There are two possibilities depending on whether S is transitive on G/N or not, and we first suppose S acts transitively on G/N. By the Fibration lemma N/H is compact, and thus a compact complex torus since G/H is Kähler. Since N/H is a Cousin group, as in the proof of Proposition 2.11 we may assume that G is a reductive complex Lie group. Thus the S-orbits are closed in G/H [23, Theorem 5.1] and one has the following diagram

$$\begin{array}{ccc} S/S\cap H & \hookrightarrow & G/H \\ F\Big\downarrow & & \Big\downarrow N/H \\ S/I & = & G/N\,. \end{array}$$

So F is a closed subgroup of N/H and it is compact. However, $F = I/S \cap H$ is the quotient of algebraic groups. This is only possible if $I/S \cap H$ is finite. Since we have the fibration $S/I \to S/P$ with $P/I = \mathbb{C}$ and S/P a flag manifold, we see that S/I is simply connected. Every S-orbit in X is a holomorphic section of the bundle $G/H \to G/N$ and $X = T \times S/I$ is a product. This is Case II (1) (b) in the Main theorem when S is transitive on G/N.

Otherwise, S does not act transitively on G/N. The radical $R_{G'}$ of G' is a unipotent group acting algebraically on G/N yielding a fibration

$$G/N \xrightarrow{F} G/N \cdot R_{G'}$$

where $F = \mathbb{C}^p$ with $p > 0$. The Fibration lemma and the assumption $d_X = 2$ imply that N/H is compact, thus a torus, $F = \mathbb{C}$ and $Z := G/N \cdot R_{G'}$ is compact and thus a flag manifold. Now $G/N \neq \mathbb{C} \times Z$ because one would then have $\mathcal{O}(G/N) \neq \mathbb{C}$, contradicting $\mathcal{O}(X) = \mathbb{C}$. So G/N is a non-trivial line bundle over Z and there are two S-orbits in G/N, a compact one Y_1 which is the zero section of the line bundle and is biholomorphic to Z and an open one Y_2. The latter holds, since the existence of another closed orbit would imply the triviality of the $\mathbb{C}^*$-bundle $G/N \setminus Y_1$ over Z. We write X as a disjoint union $X_1 \cup X_2$ with $X_i := \pi^{-1}(Y_i)$ for $i = 1, 2$. Then X_1 is a Kähler torus bundle over Z and it is trivial by [13]. A finite covering of X_2 is also trivial since X_2 is Kähler and satisfies $d_{X_2} = 1$ [18, Main theorem, case (b)]. Note that the S-orbits in X_1 (respectively X_2) are holomorphic sections of the torus bundle lying over the corresponding S-orbit Y_1 (respectively Y_2).

Let $x_2 \in X_2$ and $M_2 := S.x_2$. Since X is Kähler, the boundary of M_2 consists of S-orbits of strictly smaller dimension [23, Theorem 3.6], and for dimension reasons, these necessarily lie in X_1. Let $M_1 \subset \overline{M_2}$ be such an S-orbit in X_1 and let $p \in M_1$. As observed in the previous paragraph, M_1 is biholomorphic to Y_1 which is a flag manifold. Therefore, $M_1 = K \cdot p = K/L$, where K is a maximal compact subgroup of S and L is the corresponding isotropy subgroup at the point p and is compact. The L-action at the L-fixed point p can be linearized. This means that there exist an L-invariant open neighbourhood U of p in X, an open neighbourhood V of 0 in $T_p(X)$, a linear map $\Phi : T_p(X) \to T_p(X)$, and a biholomorphic map α of U onto V with $\alpha(p) = 0$ such that $\alpha^{-1} \circ \Phi \circ \alpha$ gives the L-action on U. In this setup Φ leaves invariant the complex vector subspaces $T_p(K/L)$ and $T_p(\pi^{-1}(\pi(p)))$ and thus also a complementary complex vector subspace W of $T_p(X)$. So we have the following decomposition:

$$T_p(X) = T_p\left(\pi^{-1}(\pi(p))\right) \oplus W \oplus T_p(K/L).$$

Let (t, w, v) be the corresponding coordinates in $T_p(X)$. Set $\widehat{M}_1 := S \cdot \widehat{x}_1$ for some $\widehat{x}_1 \in U \cap X_1$ with $\alpha(\widehat{M}_1 \cap U) \cap T_p(\pi^{-1}(\pi(p))) = \{(t_0\, w_0, v_0)\}$, where $t_0 \neq 0$. Since L is a subgroup of S and the S-orbits are transversal to the complex torus $\pi^{-1}(\pi(p))$, it follows that Φ acts as the identity on the subspace $T_p(\pi^{-1}(\pi(p)))$. So every point of $\Phi \circ \alpha(\widehat{M}_1 \cap U)$ has t-coordinate equal to $t_0 \neq 0$. As a consequence, $\widehat{M}_1$ does not intersect $\overline{M}_2$ inside the set U. Now for any other point of M_1 its isotropy subgroup for the K-action is a conjugate of the group L and the argument just given applied to that conjugate of the group L at that fixed point shows that M_2 is the unique S-orbit that contains M_1 in its closure and also that $\overline{M_2} = M_2 \cup M_1$ is a complex submanifold of X that is a holomorphic section of the bundle $\pi : X = G/H \to G/N$. This bundle is thus trivial and X is biholomorphic to $T \times G/N$. This is Case II (1) (b) in the Main theorem when S is not transitive on G/N.

This completes the classification if $\mathcal{O}(X) = \mathbb{C}$. $\square$

5. The classification when $\mathcal{O}(X) \neq \mathbb{C}$

We first prove a generalization of Proposition 3.4 for arbitrary isotropy.

Proposition 5.1. *Let G be a connected, simply connected, solvable complex Lie group, H a closed complex subgroup of G with G/H Kähler, $G/H \to G/J$ its holomorphic reduction with fiber $T = J/H$ a compact complex torus and base $G/J = (\mathbb{C}^*)^2$. Then a finite covering of G/H is biholomorphic to $T \times (\mathbb{C}^*)^2$.*

Proof. If H^0 is normal in G, then this is Proposition 3.4. Otherwise, let $N := N_G(H^0)$ and consider $G/H \to G/N$. Since G/N is an orbit in some projective space, G/N is holomorphically separable and the map $G/H \to G/N$ factors through the holomorphic reduction, *i.e.*, $J \subset N$. We first assume that $J = N$ and consider $\widehat{N} := N_G(J^0)$. Then the argument given in the third paragraph of the proof of Proposition 3.4 shows that $G/\widehat{N} = \mathbb{C}^*$ and $\widehat{N}/J = \mathbb{C}^*$. But then (a finite covering of) $\widehat{N}/H$ is isomorphic to $\mathbb{C}^* \times T$, see [18, Proposition1], implying that H^0 is normal in $\widehat{N}$. This gives the contradiction that $\widehat{N} = N$ while $\dim \widehat{N} > \dim J = \dim N$.

So we are reduced to the case where $J \neq N$ and, after going to a finite covering if necessary, $G/N = \mathbb{C}^*$ and $N/H \cong \mathbb{C}^* \times T$ is an Abelian complex Lie group, since N/H is Kähler with two ends, see [18, Proposition 1]. We have the diagram

$$
\begin{tikzcd}
X = G/H \arrow[rr, "T"] \arrow[dr] & & G/J = \mathbb{C}^* \times \mathbb{C}^* \arrow[dl] \\
& G/N = \mathbb{C}^*. &
\end{tikzcd}
$$

Since the top line is the holomorphic reduction and X is Kähler, a finite covering of this bundle is a T-principal bundle, see [20, Theorem 1]. Choose $\xi \in \mathfrak{g} \setminus \mathfrak{n}$ and set $A := \exp\langle\xi\rangle_{\mathbb{C}}$ and $B := N/H \cong \mathbb{C}^* \times T$. Then the group $\widehat{G} := A \ltimes B$ acts holomorphically and transitively on X, where A acts from the left and B acts from the right on the principal $\mathbb{C}^* \times T$-bundle $G/H \to G/N$. For dimension reasons the isotropy of this action is discrete and the result now follows by Proposition 3.4. $\square$

Proof. Let $G/H \to G/J$ be the holomorphic reduction. By the Fibration lemma one has $d_{G/J} \leq 2$ and we first consider the case when $d_{G/J} = 2$. In [7] there is a list of the possibilities for G/J which was also given above in the proof of Proposition 3.3. We now employ that list to determine the structure of X.

Suppose $G/J = \mathbb{C}$. By the Fibration lemma J/H is compact, Kähler, and so biholomorphic to the product of a torus T and a flag manifold Q. Thus $X = T \times Q \times \mathbb{C}$ by [24]; this is case (2) (a) (i) in the Main theorem.

Suppose G/J is an affine quadric. By the Fibration lemma we again have $J/H = T \times Q$. Then X is biholomorphic to a product, since G/J is Stein and it is simply connected [24]; this is case (2) (a) (ii) in the Main theorem.

If G/J is the complement of the quadric curve in $\mathbb{P}_2$, then a two-to-one covering of G/J is the affine quadric and the pullback of X to that space is again a product, as in the previous case; this is case (2) (a) (iii) in the Main theorem.

Suppose the base of the holomorphic reduction of G/H is $G/J = (\mathbb{C}^*)^2$. Since G/H is Kähler, every fiber of the fibration $G/H \to G/J$ is Kähler, and it is compact by the Fibration lemma and thus biholomorphic to $T \times Q$ [13], where T is a compact, complex torus and Q is a flag manifold. The S-orbits in the base G/N of the normalizer fibration $G/H \to G/N$ are compact and biholomorphic to the S-orbits in X. This follows from the fact that the fibers of the fibration of any flag manifold have to be flag submanifolds. But, since the isotropy of the fiber of the normalizer fibration is discrete and no positive dimensional flag manifold is parallelizable, the fibers of the induced fibration of the S-orbits by the normalizer fibration must be discrete, *i.e.*, the S-orbits in X cover the S-orbits in G/N and the latter is simply connected, so the covering is one-to-one. Next we consider the commutator fibration $G/N \to G/G' \cdot N$ of the base G/N of the normalizer fibration. Since $G/G' \cdot N$ is an Abelian, Stein Lie group [26], it follows that $G/G' \cdot N$ is isomorphic to $\{e\}$, $\mathbb{C}$, $\mathbb{C}^*$ or $(\mathbb{C}^*)^2$ by the Fibration lemma and Proposition 2.5. Note that $G/G' \cdot N \neq \mathbb{C}$, since, otherwise, the space X would be biholomorphic to a product $T \times Q \times \mathbb{C}$ by the Oka principle [24], with the base of its holomorphic reduction being $\mathbb{C}$, and this would contradict the assumption that this base is $(\mathbb{C}^*)^2$. Let $\sigma : G \to G/G' \cdot N$ be the quotient homomorphism and set $G_0 := \sigma^{-1}(K)$, where K is the maximal compact subgroup of $G/G' \cdot N$. Note that G_0 contains G' and thus also S. It then follows that the G_0-orbits in G/N are compact, homogeneous CR-manifolds that are products $Q \times (S^1)^k$ by [21, Proposition 4.4], where $k := \dim_{\mathbb{R}} K$ is equal to 0, 1, or 2. As a consequence, $G/N = Q \times (\mathbb{C}^*)^k$. Since S is acting trivially on the fiber N/H of the normalizer fibration and on the second factor in the last product decomposition, the composition of the projection map of the normalizer fibration and the projection of G/N onto Q is the fibration $G/H \to G/R \cdot H$ by the orbits of the radical R of G, *i.e.*, the R-orbits in G/H are closed. Since the base $G/R \cdot H = Q$ of the radical fibration is simply connected and the S-orbits in G/H cover this base, we see that the radical fibration has holomorphic sections and hence $G/H = Q \times Z$, where $Z := R \cdot H/H = R/R \cap H$ is a complex solvmanifold. Then Z fibers as a Kähler T-bundle over $(\mathbb{C}^*)^2$ and a finite covering of Z is a product by Proposition 5.1. Putting this together one sees that a finite covering of X is a product $T \times Q \times (\mathbb{C}^*)^2$. This gives Case II (2) (a) (iv) in the Main theorem.

Suppose G/J is a $\mathbb{C}^*$-bundle over an affine cone minus its vertex. Here $d_{G/J} = 2$ and G/J is not Stein. By the Fibration lemma J/H is compact and J/H inherits a Kähler structure from X. If we set $N := N_J(H^0)$, then the normalizer fibration $J/H \to J/N$ is a product with $N/H = T$ and $J/N = Q$ [13]. Since the fiber of the fibration $G/H \to G/N$ is compact, G/N is Kähler [12] and $d_{G/N} = 2$ by the Fibration lemma. First assume that $G/N = S/I$, where I is an algebraic subgroup of S [11]. Then the principal T-bundle $G/H \to S/I$ is of the form $G/H = S \times_\rho T$, where the representation $\rho : I \to \mathrm{Aut}^0(T) \cong T$ factors through I/I'. As in the proof of Proposition 2.11, we may assume that G is reductive and that the image $\rho(I)$ is, on the one hand, an algebraic subtorus $(\mathbb{C}^*)^k$ of the algebraic manifold $S/S \cap H$ which is closed in G/H [23, Theorem 5.1] and, on the other hand, a closed subgroup of the compact complex torus T. Hence $\rho(I) = \{e\}$ and, as a consequence, the bundle $G/H \to S/I$ is trivial. Example 7.4 shows that the Q-bundle $S/I \to S/S \cap J$ is not necessarily trivial. The setting where S is not transitive on G/N occurs if G/J is a product of $\mathbb{C}^*$ with an affine cone minus its vertex. As in the last part of the previous paragraph, one again has closed R-orbits and the radical fibration $G/H \to G/R \cdot H$. A finite covering splits as a product with the typical radical orbit $R/R \cap H$ being a Kähler T-bundle over $\mathbb{C}^*$ that has a finite covering that splits as a product [22, Theorem 6.14]. These considerations yield the possibilities in Case II (2) (b) in the Main theorem.

Suppose next that $d_{G/J} = 1$. By the Fibration lemma $d_{J/H} = 1$ and either $\mathcal{O}(J/H) = \mathbb{C}$ or $\mathcal{O}(J/H) \neq \mathbb{C}$. We first assume the former and show below that the latter gives a contradiction and thus it does not occur. Since J/H is Kähler, the classification given in [18, Proposition 5] applies and the normalizer fibration $J/H \to J/N$, where $N := N_J(H^0)$, realizes J/H as a Cousin bundle over a flag manifold.

The first case occurs if G/J is Stein. By Proposition 2.5 we have $\dim_{\mathbb{C}} G/J = 1$ and thus $G/J = \mathbb{C}^*$. Since S acts trivially on the Cousin group J/H and on G/J, the radical orbits are closed and one has the fibration $G/H \to G/R \cdot H$. Its base is biholomorphic to Q and the S-orbits in G/H are holomorphic sections. So this bundle is holomorphically trivial. Thus X is biholomorphic to $Q \times Z$, where Z is a hypersurface Cousin group bundle over $\mathbb{C}^*$. A finite covering of this splits as a product by [22, Theorem 6.14]. This is Case II (2) (c) in the Main theorem.

The other possibility is that G/J is not Stein and then G/J is an affine cone minus its vertex. By Proposition 2.11 there is a closed complex subgroup I of N containing H with $I/H = \mathbb{C}^*$ and $J/I = N/I \times J/N$, where $N/I =: T$ is a torus and $J/N =: Q$ is a flag manifold. Consider the T-bundle $G/I \to G/N$, set $Y := G/N$, and observe that Y is a Q-bundle over G/J. As such, Y has a finite fundamental group. This follows from the exact homotopy sequence of the bundle

$G/N \to G/J$ and the facts that a flag manifold is connected and simply connected and an affine cone minus its vertex has finite fundamental group. Consider the finite covering $\pi : \widetilde{Y} \to Y$, where $\widetilde{Y}$ is the universal covering of Y and let $\pi^*(G/I)$ be the pullback of G/I via the map π. The S-orbits in $\pi^*(G/I)$ are sections and the torus bundle splits as a product $\pi^*(G/I) = T \times \widetilde{Y}$. Thus a finite covering of X fibers as a $\mathbb{C}^*$-bundle over $T \times \widetilde{Y}$; this is Case II (2) (d) in the Main theorem. Example 7.3 shows that Y itself need not be a product.

Finally we assume that $\mathcal{O}(J/H) \neq \mathbb{C}$ when $d_{G/J} = 1$. Let $J/H \to J/I$ be its holomorphic reduction. Since $\mathcal{O}(J/H) \neq \mathbb{C}$, one has $\dim J/I > 0$ and thus $\dim J > \dim I$ and in all cases we will produce the contradiction that I and J have the same dimension. By [18, Proposition 3] a finite covering of J/H is biholomorphic to $I/H \times J/I$, where $I/H = T$ is a torus and $Z := J/I$ is an affine cone minus its vertex. Since the fibration $G/H \to G/I$ has T as fiber and T is compact, there is a Kähler structure on G/I, see [12] and by the Fibration lemma we have $d_{G/I} = 2$. First assume that a solvable subgroup of G acts transitively on G/I. Then $G/I \to G/J$ is a $\mathbb{C}^*$-bundle over $\mathbb{C}^*$ and G/I is Stein, *e.g.*, see [7, page 904]. Thus G/I is the base of the holomorphic reduction of G/H and this gives the contradiction that $I = J$. Next suppose that a maximal semisimple subgroup S of G acts transitively on G/I. Since $S/S \cap I$ is Kähler, $S \cap I$ is algebraic by [11]. Thus there exists a parabolic subgroup P of S containing $S \cap I$ such that $P/I \cap S$ is isomorphic to $(\mathbb{C}^*)^2$, see [6] or Theorem 2.10. Then G/I or a two-to-one covering of G/I is a homogeneous algebraic principal $\mathbb{C}^*$-bundle over an affine cone minus its vertex and is quasi-affine and thus holomorphically separable, see [7, Proposition 2]. So G/I is the base of the holomorphic reduction of G/H. But this again gives the contradiction $I = J$. The remaining case occurs when G is a mixed group, *i.e.*, G is neither solvable nor semisimple. First suppose $G/J = \mathbb{C}^*$ and J/I is an affine cone minus its vertex. Let $N := N_G(I^0)$ and consider the normalizer fibration $G/I \to G/N$ which we first assume to be a covering. As in the case when $G/J = (\mathbb{C}^*)^2$ handled above, by using the commutator fibration of G/N and [21, Proposition 4.4] we see that G/N is a product and so is holomorphically separable. Therefore, $\dim I = \dim N = \dim J$ gives the desired contradiction. In all the other cases we get a $\mathbb{C}^*$-bundle over an affine cone minus its vertex with codimension one S-orbits. Since an affine cone minus its vertex has a finite fundamental group, by passing to a finite covering we find that the S-orbits are holomorphic sections of the bundle $G/I \to G/J$, and G/I is a product and thus holomorphically separable. So we again get the contradiction that $I = J$. Thus the case $\mathcal{O}(J/H) \neq \mathbb{C}$ does not occur if $d_{G/J} = 1$.

This completes the classification when $\mathcal{O}(X) \neq \mathbb{C}$. □

6. Proof of the converse

The only component manifolds in the Main theorem that are not immediately recognizable as Kähler are the Y's that occur in II (2) (b) and II (2) (d). Since these are flag manifold bundles over holomorphically separable bases, these manifolds are Kähler by [30]. The proof of the converse follows from the following observations:

(1) The product of Kähler manifolds is a Kähler manifold;
(2) A connected, complex manifold X that is a finite, unramified covering of a complex maniold Y is Kähler if and only if Y is Kähler.

7. Examples

We now give non-trivial examples that can occur in the classification.

Example 7.1. The manifolds that occur in Proposition 2.11 need not be biholomorphic to a product of an S-orbit times an orbit of the center. For $k = d_X = 1$, let $\chi : B \to \mathbb{C}^*$ be a non-trivial character, where B is a Borel subgroup of $S := SL(2, \mathbb{C})$. Let C be a non-compact 2-dimensional Cousin group. Then C fibers as a $\mathbb{C}^*$-bundle over an elliptic curve T and let B act on C via the character χ. Set $X := S \times_B C$. Then X fibers as a principal C-bundle over S/B and is Kähler, but neither this bundle nor the corresponding $\mathbb{C}^*$-bundle is trivial.

Example 7.2. Let $S := SL(3, \mathbb{C})$ and

$$I := \left\{ \begin{pmatrix} * & 0 & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix} \right\} \subset B \subset P := \left\{ \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix} \right\},$$

where B is the Borel subgroup of S consisting of upper triangular matrices. Then $S/I \to S/B$ is an affine $\mathbb{C}$-bundle over the flag manifold S/B. Now consider the fibration $S/I \to S/P$. Its fiber is $P/I = \mathbb{P}_2 \setminus \{\text{point}\}$ and all holomorphic functions on S/I are constant along the fibers by Hartogs' principle and so must come from the base $S/P = \mathbb{P}_2$. But the latter is compact and so $\mathcal{O}(S/P) = \mathbb{C}$ and, as a consequence, we see that $\mathcal{O}(S/I) = \mathbb{C}$. Thus S/I is an example of a space that can be the base of the normalizer fibration as in the Main theorem II (1) (b) when S is transitive on that base.

Example 7.3. Consider the subgroups of $S := SL(5, \mathbb{C})$ defined by

$$I := \left\{ \begin{pmatrix} 1 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{pmatrix} \right\} \subset P := \left\{ \begin{pmatrix} * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{pmatrix} \right\}$$

and

$$J := P' = \left\{ \begin{pmatrix} 1 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{pmatrix} \right\}.$$

Then $J/I = \mathbb{P}_1$, $P/J = \mathbb{C}^*$, and $S/P = Q$ is a flag manifold that can be fibered as a Gr(2, 4)-bundle over $\mathbb{P}_4$. We have the fibrations

$$S/I \xrightarrow{\mathbb{P}_1} S/J \xrightarrow{\mathbb{C}^*} S/P = Q.$$

Note that S/J is holomorphically separable due to the fact that it can be equivariantly embedded as an affine cone minus its vertex in some projective space and since J/H is compact, S/J is the base of the holomorphic reduction of S/I. Since the fibration of S/I is not trivial, the spaces Y that occur in the Main theorem II (2) (d) need not split as the products of flag manifolds and affine cones minus their vertices.

Example 7.4. Using the groups defined in the previous example set $\widehat{S} := S \times S$ and $\widehat{I} := I \times I$. Then $Y := \widehat{S}/\widehat{I} = S/I \times S/I$ is an example that can occur in the Main theorem II (2) (b). Such a Y fibers as a non-trivial flag manifold over a $(\mathbb{C}^*)^2$-bundle over a flag manifold.

References

[1] Y. ABE and K. KOPFERMANN, "Toroidal Groups. Line Bundles, Cohomology and Quasi-Abelian Varieties", Lecture Notes in Mathematics, Vol. 1759, Springer-Verlag, Berlin, 2001.

[2] H. ABELS, *Proper transformation groups*, In: "Transformation Groups", Proc. Conf., Univ. Newcastle upon Tyne, Newcastle upon Tyne, 1976, London Math. Soc. Lecture Note Series, n. 26, Cambridge Univ. Press, Cambridge, 1977, 237–248.

[3] H. ABELS, *Some topological aspects of proper group actions; noncompact dimension of groups*, J. London Math. Soc. (2) **25** (1982), 525–538.

[4] S. AHMADI, "A Classification of Homogeneous Kähler Manifolds with Discrete Isotropy and Top Non-vanishing Homology in Codimension Two", Ph.D. Dissertation, University of Regina, 2013.

[5] D. N. AKHIEZER, *Dense orbits with two endpoints*, (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **41** (1977), 308–324; translation in Math. USSR-Izv. **11** (1977) (1978), 293–307.

[6] D. N. AKHIEZER, *Complex n-dimensional homogeneous spaces homotopically equivalent to* $(2n-2)$*-dimensional compact manifolds*, Selecta Math. Sov. **3** (1983/84), 286–290.

[7] D. N. AKHIEZER and B. GILLIGAN, *On complex homogeneous spaces with top homology in codimension two*, Canad. J. Math. **46** (1994), 897–919.

[8] D. N. AKHIEZER, "Lie Group Actions in Complex Analysis", Aspects of Mathematics, Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, 1995.

[9] L. AUSLANDER and R. TOLIMIERI, *Splitting theorems and the structure of solvmanifolds*, Ann. of Math. (2) **92** (1970), 164–173.

[10] H. AZAD, A. HUCKLEBERRY and W. RICHTHOFER, *Homogeneous CR-manifolds*, J. Reine Angew. Math. **358** (1985), 125–154.

[11] F. BERTELOOT and K. OELJEKLAUS, *Invariant plurisubharmonic functions and hypersurfaces on semi-simple complex Lie groups*, Math. Ann. **281** (1988), 513–530.

[12] A. BLANCHARD, *Sur les variétés analytiques complexes*, Ann. Sci. École Norm. Sup. (3) **73** (1956), 157–202.

[13] A. BOREL and R. REMMERT, *Über kompakte homogene Kählersche Mannigfaltigkeiten*, Math. Ann. **145** (1961/1962), 429–439.

[14] A. BOREL, "Linear Algebraic Groups", Second edition, Graduate Texts in Mathematics, 126, Springer-Verlag, New York, 1991, xii+288.
[15] C. CHEVALLEY, "Théorie des Groupes de Lie II. Groupes Algébriques", Hermann, Paris, 1951.
[16] J. DORFMEISTER and K. NAKAJIMA, *The fundamental conjecture for homogeneous Kähler manifolds*, Acta Math. **161** (1988), 23–70.
[17] G. FELS, A. HUCKLEBERRY and J. WOLF, "Cycle Spaces of flag Domains. A Complex Geometric Viewpoint", Progress in Mathematics, Vol. 245, Birkhäuser Boston, Inc., Boston, MA, 2006.
[18] B. GILLIGAN, K. OELJEKLAUS and W. RICHTHOFER, *Homogeneous complex manifolds with more than one end*, Canad. J. Math. **41** (1989), 163–177.
[19] B. GILLIGAN, *Invariant analytic hypersurfaces in complex Lie groups*, Bull. Aust. Math. Soc. **70** (2004), 343–349.
[20] B. GILLIGAN and K. OELJEKLAUS, *Two remarks on Kähler homogeneous manifolds*, Ann. Fac. Sci. Toulouse Math. (6) **17** (2008), 73–80.
[21] B. GILLIGAN and A. T. HUCKLEBERRY, *Fibrations and globalizations of compact homogeneous CR-manifolds*, (Russian) Izv. Ross. Akad. Nauk Ser. Mat. **73** (2009), 67–126; translation in Izv. Math. **73** (2009), 501–553.
[22] B. GILLIGAN and K. OELJEKLAUS, *Compact CR-solvmanifolds as Kähler obstructions*, Math. Z. **269** (2011), 179–191.
[23] B. GILLIGAN, C. MIEBACH and K. OELJEKLAUS, *Homogeneous Kähler and Hamiltonian manifolds*, Math. Ann. **349** (2011), 889–901.
[24] H. GRAUERT, *Analytische Faserungen über holomorph-vollständigen Räumen*, Math. Ann. **135** (1958), 263–273.
[25] G. HOCHSCHILD and G. D. MOSTOW, *On the algebra of representative functions of an analytic group II*, Amer. J. Math. **86** (1964), 869–887.
[26] A. T. HUCKLEBERRY and E. OELJEKLAUS, *Homogeneous Spaces from a Complex Analytic Viewpoint*, In: "Manifolds and Lie Groups" (papers in honor of Y. Matsushima), Progress in Math., Vol. 14, Birkhäuser, Boston, 1981, 159–186.
[27] A. T. HUCKLEBERRY and E. OELJEKLAUS, "Classification Theorems for Almost Homogeneous Spaces", Institut Élie Cartan, 9, Université de Nancy, Institut Élie Cartan, Nancy, 1984.
[28] A. T. HUCKLEBERRY and E. OELJEKLAUS, *On holomorphically separable complex solvmanifolds*, Ann. Inst. Fourier (Grenoble) **36** (1986), 57–65.
[29] J. E. HUMPHREYS, "Linear Algebraic Groups", Graduate Texts in Mathematics, n. 21, Springer-Verlag, New York-Heidelberg, 1975.
[30] K. KODAIRA, *On Kähler varieties of restricted type* (*an intrinsic characterization of algebraic varieties*), Ann. of Math. (2) **60** (1954), 28–48.
[31] J.-J. LOEB, *Action d'une forme réelle d'un groupe de Lie complexe sur les fonctions plurisousharmoniques*, Ann. Inst. Fourier (Grenoble) **35** (1985), 59–97.
[32] Y. MATSUSHIMA, *Sur les espaces homogènes kählériens d'un groupe de Lie réductif*, Nagoya Math. J. **11** (1957), 53–60.
[33] A. MORIMOTO, *On the classification of noncompact complex Abelian Lie groups*, Trans. Amer. Math. Soc. **123** (1966), 200–228.
[34] G. D. MOSTOW, *Factor spaces of solvable groups*, Ann. of Math. (2) **60** (1954), 1–27.
[35] G. D. MOSTOW, *On covariant fiberings of Klein spaces I*, Amer. J. Math. **77** (1955) 247–278. *Covariant fiberings of Klein spaces II*, Amer. J. Math. **84** (1962), 466–474.
[36] K. OELJEKLAUS and W. RICHTHOFER, *Recent results on homogeneous complex manifolds*, In: "Complex Analysis, III" (College Park, Md., 1985-86), 78–119, Lecture Notes in Math., Vol. 1277, Springer, Berlin, 1987.
[37] K. OELJEKLAUS and W. RICHTHOFER, *On the structure of complex solvmanifolds*, J. Differential Geom. **27** (1988), 399–421.

[38] M. S. RAGHUNATHAN, "Discrete Subgroups of Lie Groups", Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68, Springer-Verlag, New York-Heidelberg, 1972.
[39] J.-P. SERRE, *Quelques problèmes globaux relatifs aux variétés de Stein*, In: "Coll. sur les Fonctions de Plusieurs Variables", Bruxelles, 1953, 57–68.

Department of Mathematics and Statistics
University of Regina
Regina, Canada S4S 0A2
seyed.ruhala.ahmadi@gmail.com
gilligan@math.uregina.com

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1255-1299

Volume and self-intersection of differences of two nef classes

DAN POPOVICI

Abstract. Let $\{\alpha\}$ and $\{\beta\}$ be nef cohomology classes of bidegree (1, 1) on a compact n-dimensional Kähler manifold X such that the difference of intersection numbers $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\}$ is positive. We solve in a number of special but rather inclusive cases the quantitative part of Demailly's Transcendental Morse Inequalities Conjecture for this context predicting the lower bound $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\}$ for the volume of the difference class $\{\alpha - \beta\}$. We completely solved the qualitative part in an earlier work. We also give general lower bounds for the volume of $\{\alpha - \beta\}$ and show that the self-intersection number $\{\alpha - \beta\}^n$ is always bounded below by $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\}$. We also describe and estimate the relative psef and nef thresholds of $\{\alpha\}$ with respect to $\{\beta\}$ and relate them to the volume of $\{\alpha - \beta\}$. Finally, broadening the scope beyond the Kähler realm, we propose a conjecture relating the balanced and the Gauduchon cones of $\partial\bar\partial$-manifolds which, if proved to hold, would imply the existence of a balanced metric on any $\partial\bar\partial$-manifold.

Mathematics Subject Classification (2010): 32J27 (primary); 32U40, 32Q25, 32J25, 53C55 (secondary).

1. Introduction

Let X be a compact Kähler manifold with $\dim_{\mathbb{C}} X = n$ and let $\{\alpha\}, \{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ be nef Bott-Chern cohomology classes such that

$$\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0. \tag{1.1}$$

A (possibly transcendental) class $\{\alpha\} \in H^{1,1}_{BC}(X, \mathbb{R})$ being nef means (*cf*. [11, Definition 1.3]) that for some (hence all) fixed Hermitian metric ω on X and for every $\varepsilon > 0$, there exists a C^∞ form $\alpha_\varepsilon \in \{\alpha\}$ such that $\alpha_\varepsilon \geq -\varepsilon\,\omega$.

We have proved in [20, Theorem 1.1] that the class $\{\alpha - \beta\}$ is big (*i.e.* contains a Kähler current T). This solved the *qualitative part* of Demailly's *Transcendental Morse Inequalities Conjecture* for differences of two nef classes (*cf*. [6, Conjecture 10.1, (ii)]) on compact Kähler (and even more general) manifolds. This special

Received March 16, 2016; accepted in revised form June 12, 2016.
Published online December 2017.

form of the conjecture was originally motivated by attempts at extending to transcendental classes and to compact Kähler (not necessarily projective) manifolds the cone duality theorem of Boucksom, Demailly, Paun and Peternell [6, Theorem 2.2.] that plays a major role in the theory of classification of projective manifolds. Recall that T being a *Kähler current* means that T is a d-closed positive $(1,\ 1)$-current with the property that for some (hence all) fixed Hermitian metric ω on X, there exists $\varepsilon > 0$ such that $T \geq \varepsilon\,\omega$ on X. Nefness and bigness are quite different positivity properties for real (possibly transcendental) $(1,\ 1)$-classes and the by-now standard definitions just recalled extend classical algebraic definitions for integral classes on projective manifolds.

In this paper we give a partial answer to the *quantitative part* of Demailly's *Transcendental Morse Inequalities Conjecture* for differences of two nef classes:

Conjecture 1.1 ([6, Conjecture 10.1, (*ii*)]). Let $\{\alpha\}, \{\beta\} \in H^{1,\,1}_{BC}(X,\ \mathbb{R})$ be nef classes satisfying condition (1.1) on a compact Kähler manifold X with $\dim_{\mathbb{C}} X = n$. Then

$$\mathrm{Vol}(\{\alpha - \beta\}) \geq \{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}. \tag{1.2}$$

This is stated for arbitrary (*i.e.* possibly non-Kähler) compact complex manifolds in [6], but the volume is currently only known to be meaningful when X is of *class* $\mathcal{C}$, a case reducible to the Kähler case by modifications. Thus, we may assume without loss of generality that X is Kähler.

Recall that the volume is a way of gauging the "amount" of positivity of a class $\{\gamma\} \in H^{1,\,1}_{BC}(X,\ \mathbb{R})$ when X is Kähler (or merely of *class* $\mathcal{C}$) and was introduced in [5, Definition 1.3] as

$$\mathrm{Vol}(\{\gamma\}) := \sup_{T\in\{\gamma\},\, T\geq 0} \int_X T^n_{ac} \tag{1.3}$$

if $\{\gamma\}$ is *pseudo-effective* (*psef*), *i.e.* if $\{\gamma\}$ contains a *positive* $(1,\ 1)$-current $T \geq 0$, where T_{ac} denotes the absolutely continuous part of T in the Lebesgue decomposition of its coefficients (which are complex measures when $T \geq 0$). If the class $\{\gamma\}$ is not psef, then its volume is set to be zero. It was proved in [5, Theorem 1.2] that this volume (which is always a *finite* non-negative quantity thanks to the Kähler, or more generally *class* $\mathcal{C}$, assumption on X) coincides with the standard volume of a holomorphic line bundle L if the class $\{\gamma\}$ is integral (*i.e.* the first Chern class of some L). Moreover, the class $\{\gamma\}$ is big (*i.e.* contains a Kähler current) if and only if its volume is *positive*, by [5, Theorem 4.7].

Thus, under the assumptions of Conjecture 1.1, the main result in [20] ensures that $\mathrm{Vol}(\{\alpha - \beta\}) > 0$. In other words, $\{\alpha - \beta\}$ is positive in the *big* sense. The special case when $\{\beta\} = 0$ had been proved in [14, Theorem 2.12] and had served there as the main ingredient in the proof of the numerical characterisation of the Kähler cone. (In particular, the proof of the more general statement in [20] reproves in a much simpler way the main technical result in [14].) The thrust of Conjecture 1.1 is to estimate from below the "amount" of positivity of the class $\{\alpha - \beta\}$.

A first group of results that we obtain in the present paper can be summed up in the following positive answer to Conjecture 1.1 under an extra assumption. Recall that for *nef* classes $\{\gamma\}$, the volume equals the top self-intersection $\{\gamma\}^n$ (*cf*. [5, Theorem 4.1]), but for arbitrary classes, any order may occur between these two quantities.

Theorem 1.2. *Let X be a compact Kähler manifold with $\dim_{\mathbb{C}} X = n$ and let $\{\alpha\}$, $\{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ be nef classes such that $\{\alpha\}^n - n\{\alpha\}^{n-1}.\{\beta\} > 0$. Suppose, moreover, that*

$$\mathrm{Vol}\,(\{\alpha - \beta\}) \geq \{\alpha - \beta\}^n. \tag{1.4}$$

Then $\mathrm{Vol}(\{\alpha - \beta\}) \geq \{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}$.

Although there are examples when the volume of $\{\alpha - \beta\}$ is strictly less than the top self-intersection, the assumption (1.4), that we hope to be able to remove in future work, is satisfied in quite a number of cases, *e.g.*, when the class $\{\alpha - \beta\}$ is nef (treated in Section 2).

Actually, we prove in full generality in Section 5 the analogue of Conjecture 1.1 for the top self-intersection number $\{\alpha - \beta\}^n$ in place of the volume of $\{\alpha - \beta\}$.

Theorem 1.3. *Let X be a compact Kähler manifold with $\dim_{\mathbb{C}} X = n$ and let $\{\alpha\}, \{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ be nef classes such that $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\} > 0$. Then*

$$\{\alpha - \beta\}^n \geq \{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}.$$

Theorem 1.2 follows immediately from Theorem 1.3. Since the nef cone is the closure of the Kähler cone, we may assume without loss of generality that the classes $\{\alpha\}$ and $\{\beta\}$ are actually Kähler. As for the volume of $\{\alpha - \beta\}$ in the general case (*i.e.* without assumption (1.4)), we prove a lower bound that is weaker than the expected lower bound (1.2) in a way that depends explicitly on how far the class $\{\alpha - \beta\}$ is from being nef. The *nefness defect* of $\{\alpha - \beta\}$ is defined explicitly and investigated in relation to the volume in Subsections 4.2, 4.3 and 4.4. We call it the *nef threshold* (a term that is already present in the literature) of $\{\alpha\}$ with respect to $\{\beta\}$ and discuss it together with the analogous *psef threshold* of $\{\alpha\}$ with respect to $\{\beta\}$ in Section 4. In Section 4.4, we prove the following general lower bound for the volume of $\{\alpha - \beta\}$.

Theorem 1.4. *Let X be a compact Kähler manifold with $\dim_{\mathbb{C}} X = n$ and let $\{\alpha\}$, $\{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ be Kähler classes such that $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\} > 0$. Let $s_0 := N^{(\beta)}(\alpha) > 0$ be the nef threshold of $\{\alpha\}$ with respect to $\{\beta\}$. Then:*

(i) *if $s_0 \geq 1$, the class $\{\alpha - \beta\}$ is nef and the optimal volume estimate* (1.2) *holds;*
(ii) *if $s_0 < 1$, the class $\{\alpha - \beta\}$ is not nef and the next volume estimate holds:*

$$\mathrm{Vol}(\{\alpha - \beta\}) \geq \Big(\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}\Big)\left(\frac{\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}}{\{\alpha\}^n - ns_0\,\{\alpha\}^{n-1}.\,\{\beta\}}\right)^{n-1}. \tag{1.5}$$

A very special case of this result was also observed independently in [23] using the technique introduced in [20].

Taking our cue from the estimates we obtain in Section 3 for the supremum of $t \geq 0$ such that the class $\{\alpha\} - t\,\{\beta\}$ is psef in the setting of Conjecture 1.1, we define the psef and nef thresholds of $\{\alpha\}$ with respect to $\{\beta\}$ as the functions $P^{(\beta)},\ N^{(\beta)} : H^{1,1}_{BC}(X,\,\mathbb{R}) \to \mathbb{R}$,

(i) $P^{(\beta)}(\alpha) := \inf\int_X \alpha\wedge\gamma^{n-1}$;
(ii) $N^{(\beta)}(\alpha) := \inf\int_Y \alpha\wedge\omega^{n-p-1}$;

where in (i) the infimum is taken over all the Gauduchon metrics γ on X normalised by

$$[\beta]_{BC}.\big[\gamma^{n-1}\big]_A = \int_X \beta\wedge\gamma^{n-1} = 1,$$

while in (ii) the infimum is taken over all $p = 0, 1, \dots, n-1$, over all the irreducible analytic subsets $Y \subset X$ such that codim $Y = p$ and over all Kähler classes $\{\omega\}$ normalised by $\int_Y \beta\wedge\omega^{n-p-1} = 1$. The class $\{\beta\}$ is supposed to be big in the case of $P^{(\beta)}$ and Kähler in the case of $N^{(\beta)}$. (The subscripts BC and A will stand throughout for the Bott-Chern, respectively Aeppli cohomologies.) In Subsection 4, we prove the following formulae that justify the terminology and make it match existing notions in the literature:

$$P^{(\beta)}(\alpha) = \sup\{t\in\mathbb{R}\ /\ \text{the class } \{\alpha\} - t\,\{\beta\} \text{ is psef}\},$$
$$N^{(\beta)}(\alpha) = \sup\{s\in\mathbb{R}\ /\ \text{the class } \{\alpha\} - s\{\beta\} \text{ is nef}\}.$$

The psef/nef thresholds of $\{\alpha\}$ with respect to $\{\beta\}$ turn out to gauge quite effectively the amount of positivity that the class $\{\alpha\}$ has in the "direction" of the class $\{\beta\}$. We study their various properties in Section 4, estimate them in terms of intersection numbers as

$$\frac{\{\alpha\}^n}{n\,\{\alpha\}^{n-1}.\{\beta\}} \leq P^{(\beta)}(\alpha) \leq \frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}},$$

and by similar, more involved inequalities for $N^{(\beta)}(\alpha)$, and relate them to the volume of $\{\alpha - \beta\}$ as

$$\mathrm{Vol}(\{\alpha-\beta\}) \geq \left(1 - \frac{1}{P^{(\beta)}(\alpha)}\right)^n \{\alpha\}^n,$$

whenever the classes $\{\alpha\}$ and $\{\beta\}$ are Kähler.

Using these thresholds, we prove Conjecture 1.1 in yet another special case: when the psef and the nef thresholds of $\{\alpha\}$ with respect to $\{\beta\}$ are sufficiently *close* to each other (*cf.* Proposition 4.12). Of course, we always have: $N^{(\beta)}(\alpha) \leq P^{(\beta)}(\alpha)$.

As in our earlier work [20] and as in [25] that preceded it, we will repeatedly make use of two ingredients. The first one is *Lamari's positivity criterion*.

Lemma 1.5 ([18, Lemme 3.3]). *Let $\{\alpha\} \in H^{1,1}_{BC}(X, \mathbb{R})$ be any real Bott-Chern cohomology class on an n-dimensional compact complex manifold X. The following two statements are equivalent.*

(i) *There exists a* $(1, 1)$*-current T in $\{\alpha\}$ such that $T \geq 0$ on X* (i.e. $\{\alpha\}$ *is psef);*
(ii) $\int_X \alpha \wedge \gamma^{n-1} \geq 0$ *for all Gauduchon metrics γ on X.*

In fact, Lamari's result holds more generally for any (*i.e.* not necessarily d-closed) C^∞ real $(1, 1)$-form α on X, but we will not use this here. The second ingredient that we will often use is Yau's solution of the Calabi conjecture.

Theorem 1.6 ([26]). *Let X be a compact complex n-dimensional manifold endowed with a Kähler metric ω. Let $dV > 0$ be any C^∞ positive volume form on X such that $\int_X \omega^n = \int_X dV$. Then, there exists a unique Kähler metric $\widetilde{\omega}$ in the Kähler class $\{\omega\}$ such that $\widetilde{\omega}^n = dV$.*

There is a non-Kähler analogue of Yau's theorem by Tosatti and Weinkove [24] that will not be used in this work. Moreover, most of the techniques that follow are still meaningful or can be extended to the non-Kähler context. This is part of the reason why we believe that a future development of the matters dealt with in this paper may be possible in the more general setting of $\partial\bar\partial$-manifolds. The conjecture we propose in Section 6 is an apt illustration of this idea.

We will make repeated use of the technique introduced in [20] based on the Cauchy-Schwarz inequality for estimating from below certain integrals of traces of Kähler metrics introduced in [20]. Moreover, there are mainly two new techniques that we introduce in the current paper: (i) the observation, proof and use of certain pointwise inequalities involving products of positive smooth forms (*cf.* Appendix) reminiscent of the Hovanskii-Teissier inequalities and generalising [20, Lemma 3.1]; (ii) a technique for constructing what we call *approximate fixed points* for Monge-Ampère equations when we allow the right-hand side to vary (*cf.* proof of Proposition 5.1) whose rough idea originates in and was suggested by discussions the author had several years ago in a completely different context with different equations and for very different purposes with J.-P. Demailly to whom we are very grateful.

2. Special case of Conjecture 1.1 when $\{\alpha - \beta\}$ is nef

We start by noticing the following elementary inequality.

Lemma 2.1. *Let $\alpha > 0$ and $\beta \geq 0$ be C^∞ $(1, 1)$-forms on a complex manifold X with $\dim_{\mathbb{C}} X = n$ such that $\alpha - \beta \geq 0$. Then:*

$$(\alpha - \beta)^n \geq \alpha^n - n\,\alpha^{n-1} \wedge \beta \text{ at every point in } X. \tag{2.1}$$

If $d\alpha = d\beta = 0$ and if X is compact, then taking integrals we get:

$$\mathrm{Vol}(\{\alpha-\beta\}) = \int_X (\alpha-\beta)^n \geq \int_X \alpha^n - n \int_X \alpha^{n-1} \wedge \beta = \{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\}. \tag{2.2}$$

Proof. Let $x_0 \in X$ be an arbitrary point and let $z_1, \dots, z_n$ be local holomorphic coordinates centred at x_0 such that at x_0 we have:

$$\alpha = \sum_{j=1}^{n} i dz_j \wedge d\bar{z}_j \quad \text{and} \quad \beta = \sum_{j=1}^{n} \beta_j \, i dz_j \wedge d\bar{z}_j.$$

Then $\alpha - \beta = \sum_{j=1}^{n}(1-\beta_j)\, i dz_j \wedge d\bar{z}_j$ at x_0, while $\beta_j \geq 0$ and $1 - \beta_j \geq 0$ at x_0 for all j. Thus inequality (2.1) at x_0 translates to

$$(1-\beta_1)\cdots(1-\beta_n) \geq 1-(\beta_1+\cdots+\beta_n) \qquad \text{for all} \quad \beta_1, \dots, \beta_n \in [0,\ 1]. \tag{2.3}$$

This elementary inequality is easily proved by induction on $n \geq 1$. Indeed, (2.3) is an identity for $n = 1$, while if (2.3) has been proved for n, then we have:

$$\begin{aligned}(1-\beta_1)\cdots(1-\beta_n)(1-\beta_{n+1}) &\overset{(i)}{\geq} (1-(\beta_1+\cdots+\beta_n))\ (1-\beta_{n+1})\\ &= 1-(\beta_1+\cdots+\beta_n+\beta_{n+1})+\beta_{n+1}(\beta_1+\cdots+\beta_n)\\ &\geq 1-(\beta_1+\cdots+\beta_n+\beta_{n+1}),\end{aligned}$$

since $\beta_j \geq 0$ for all j. (We used $1 - \beta_{n+1} \geq 0$ to get (i) from the induction hypothesis.) Thus (2.3) is proved and (2.1) follows from it.

Now, if α and β are d-closed, they define Bott-Chern cohomology classes. Since $\alpha - \beta$ is a semi-positive C^∞ (1, 1)-form, its Bott-Chern class is nef (and even a bit more), hence its volume equals $\int_X (\alpha - \beta)^n$ by [5, Theorem 4.1] if X is compact. (Note that X is compact Kähler since α is a Kähler metric under the present assumptions.) The remaining part of (2.2) follows at once from (2.1) by integration. □

An immediate consequence of Lemma 2.1 is the desired volume lower bound (1.2) in the special case when the class $\{\alpha - \beta\}$ is assumed to be nef. Note, however, that $\{\alpha - \beta\}$ need not be nef in general even a posteriori in the setting of Conjecture 1.1.

Proposition 2.2. *Let X be a compact Kähler manifold with* $\dim_{\mathbb{C}} X = n$ *and let* $\{\alpha\}, \{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ *be nef Bott-Chern cohomology classes such that the class* $\{\alpha - \beta\}$ *is nef. Then*

$$\mathrm{Vol}(\{\alpha - \beta\}) \geq \{\alpha\}^n - n\, \{\alpha\}^{n-1}.\{\beta\}. \tag{2.4}$$

Proof. It suffices to prove inequality (2.4) in the case when the classes $\{\alpha\}$, $\{\beta\}$ and $\{\alpha - \beta\}$ are all Kähler. (Otherwise, we can add $2\varepsilon\{\omega\}$ to $\{\alpha\}$ and $\varepsilon\{\omega\}$ to $\{\beta\}$ for a fixed Kähler class $\{\omega\}$ and let $\varepsilon \downarrow 0$ in the end. The volume function is known to be continuous by [5, Corollary 4.11].) If we define the form α as the sum of any Kähler metric in the class $\{\alpha - \beta\}$ with any Kähler metric β in the class $\{\beta\}$, the forms α, β and $\alpha - \beta$ obtained in this way satisfy the hypotheses of Lemma 2.1, hence also the elementary inequality (2.1) and its consequence (2.2). □

Recall that the class $\{\alpha - \beta\}$ is big under the assumptions of Conjecture 1.1 by the main result in [20]. However, big positivity is quite different in nature to nef positivity. The general (*i.e.* possibly non-nef) case is discussed in the next sections.

3. Applications of Monge-Ampère equations

In this section we rewrite in a more effective way and observe certain consequences of the arguments in [20, Section 3].

Lemma 3.1. *Let X be any compact complex manifold with* $\dim_{\mathbb{C}} X = n$. *With any C^∞ $(1,\,1)$-forms $\alpha, \beta > 0$ and any Gauduchon metric γ, we associate the C^∞ $(1,\,1)$-form $\widetilde{\alpha} = \alpha + i\partial\bar\partial u > 0$ defined as the unique normalised solution* (*whose existence is guaranteed by the Tosatti-Weinkove theorem in* [24]) *of the Monge-Ampère equation:*

$$(\alpha + i\partial\bar\partial u)^n = c\,\beta\wedge\gamma^{n-1} \quad \textit{such that} \quad \sup_X u = 0, \tag{3.1}$$

where $c > 0$ is the unique constant for which the above equation admits a solution $u : X \to \mathbb{R}$. (*Of course, a posteriori, $c = (\int_X(\alpha + i\partial\bar\partial u)^n)/(\int_X \beta\wedge\gamma^{n-1})$, while if $d\alpha = 0$ then $c = \int_X \alpha^n = \{\alpha\}^n > 0$.*)

Then the following inequality holds:

$$\left(\int_X \widetilde{\alpha}\wedge\gamma^{n-1}\right)\cdot\left(\int_X \widetilde{\alpha}^{n-1}\wedge\beta\right) \geq \frac{1}{n}\left(\int_X \widetilde{\alpha}^n\right)\left(\int_X \beta\wedge\gamma^{n-1}\right). \tag{3.2}$$

Proof. Let us define $\det_\gamma \widetilde{\alpha}$ by requiring $\widetilde{\alpha}^n = (\det_\gamma \widetilde{\alpha})\,\gamma^n$ on X. Since $\beta\wedge\gamma^{n-1} = (1/n)\,(\Lambda_\gamma\beta)\,\gamma^n$, the Monge-Ampére equation (3.1) translates to

$$\det_\gamma \widetilde{\alpha} = \frac{c}{n}\,\Lambda_\gamma\beta. \tag{3.3}$$

Hence, we get the following identities and inequalities:

$$\begin{aligned}
&\left(\int_X \widetilde{\alpha}\wedge\gamma^{n-1}\right)\left(\int_X \widetilde{\alpha}^{n-1}\wedge\beta\right) = \left(\int_X \frac{1}{n}\,(\Lambda_\gamma\widetilde{\alpha})\,\gamma^n\right)\left(\int_X \frac{1}{n}\,(\Lambda_{\widetilde{\alpha}}\beta)\,(\det_\gamma\widetilde{\alpha})\,\gamma^n\right)\\
&\overset{(a)}{\geq} \frac{1}{n^2}\left(\int_X [(\Lambda_\gamma\widetilde{\alpha})\,(\Lambda_{\widetilde{\alpha}}\beta)]^{\frac{1}{2}}\,(\det_\gamma\widetilde{\alpha})^{\frac{1}{2}}\,\gamma^n\right)^2 \overset{(b)}{\geq} \frac{1}{n^2}\left(\int_X (\Lambda_\gamma\beta)^{\frac{1}{2}}\,(\det_\gamma\widetilde{\alpha})^{\frac{1}{2}}\,\gamma^n\right)^2\\
&\overset{(c)}{=} \frac{1}{n^2}\left(\sqrt{\frac{c}{n}}\int_X (\Lambda_\gamma\beta)\,\gamma^n\right)^2 = \frac{1}{n^2}\left(\sqrt{\frac{c}{n}}\,n\int_X \beta\wedge\gamma^{n-1}\right)^2 = \frac{c}{n}\left(\int_X \beta\wedge\gamma^{n-1}\right)^2,
\end{aligned}$$

which prove (3.2) since $c = (\int_X \widetilde{\alpha}^n)/(\int_X \beta\wedge\gamma^{n-1}) > 0$, where (a) is the Cauchy-Schwarz inequality, (b) follows from the inequality $(\Lambda_\gamma\widetilde{\alpha})\,(\Lambda_{\widetilde{\alpha}}\beta) \geq \Lambda_\gamma\beta$ proved in [20, Lemma 3.1], while (c) follows from (3.3). □

Corollary 3.2. *Let X be a compact Kähler manifold with* $\dim_{\mathbb{C}} X = n$. *Then, for every Kähler metrics α, β and every Gauduchon metric γ on X, the following inequality holds:*

$$\left(\int_X \alpha\wedge\gamma^{n-1}\right)\cdot\left(\int_X \alpha^{n-1}\wedge\beta\right) \geq \frac{1}{n}\left(\int_X \alpha^n\right)\left(\int_X \beta\wedge\gamma^{n-1}\right). \tag{3.4}$$

Proof. It is clear that (3.4) follows immediately from (3.2) since the assumption $d\alpha = d\beta = 0$ ensures that $\int_X \widetilde{\alpha}\wedge\gamma^{n-1} = \int_X \alpha\wedge\gamma^{n-1}$, $\int_X \widetilde{\alpha}^{n-1}\wedge\beta = \int_X \alpha^{n-1}\wedge\beta$ and $\int_X \widetilde{\alpha}^n = \int_X \alpha^n$. $\square$

Remark 3.3. Under the hypotheses of Corollary 3.2, for any γ satisfying the inequality $(\Lambda_\gamma\widetilde{\alpha})\,(\Lambda_{\widetilde{\alpha}}\beta) \geq n\Lambda_\gamma\beta$ (an improved version of [20, Lemma 3.1] which need not hold in general, but holds for some special choices of γ – *cf.* proof of Lemma 7.2), the lower bound on the right-hand side of (3.4) improves to $(\int_X \alpha^n)\,(\int_X \beta\wedge\gamma^{n-1})$. If this improved lower bound held for all Gauduchon metrics γ, Conjecture 1.1 would follow immediately (see Theorem 3.5 below).

We first notice a consequence of Corollary 3.2 for nef classes.

Corollary 3.4. *If* $\{\alpha\}$ *and* $\{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ *are nef classes on a compact Kähler manifold* X *with* $\dim_{\mathbb{C}} X = n$ *such that* $\{\alpha\}^n - n\{\alpha\}^{n-1}.\{\beta\} > 0$, *then* $\{\alpha\}^n > 0$ *and, unless* $\{\beta\}=0$, *the following non-orthogonality property holds:* $\{\alpha\}^{n-1}.\{\beta\} > 0$.

Proof. The nef hypothesis on $\{\alpha\}$ and $\{\beta\}$ ensures that $\{\alpha\}^{n-1}.\{\beta\} \geq 0$, hence $\{\alpha\}^n > 0$ since $\{\alpha\}^n > n\,\{\alpha\}^{n-1}.\{\beta\}$ by assumption. For the rest of the proof, we reason by contradiction: suppose that $\{\alpha\}^{n-1}.\{\beta\} = 0$ and that $\{\beta\} \neq 0$. By the nef hypothesis on $\{\alpha\}$ and $\{\beta\}$, for every $\varepsilon > 0$, there exist C^∞ forms $\alpha_\varepsilon \in \{\alpha\}$, $\beta_\varepsilon \in \{\beta\}$ such that $\alpha_\varepsilon + \varepsilon\,\omega > 0$ and $\beta_\varepsilon + \varepsilon\,\omega > 0$ for some arbitrary fixed Kähler metric ω on X. Applying (3.4) to the Kähler metrics $\alpha_\varepsilon + \varepsilon\,\omega$ and $\beta_\varepsilon + \varepsilon\,\omega$ in place of α and β and letting $\varepsilon \downarrow 0$, we get $\int_X \beta\wedge\gamma^{n-1} = 0$ for every Gauduchon metric γ on X. (Note that $\int_X \alpha_\varepsilon\wedge\gamma^{n-1} = \int_X \alpha\wedge\gamma^{n-1}$, $\int_X \beta_\varepsilon\wedge\gamma^{n-1} = \int_X \beta\wedge\gamma^{n-1}$ and $\int_X \alpha_\varepsilon^{n-1}\wedge\beta_\varepsilon = \{\alpha\}^{n-1}.\{\beta\} = 0$.) If we fix a d-closed positive current $T \geq 0$ in the class $\{\beta\}$ (such a current exists since the nef class $\{\beta\}$ is, in particular, pseudo-effective), this means that $\int_X T\wedge\gamma^{n-1} = 0$ for every Gauduchon metric γ on X. Consequently, $T = 0$, hence $\{\beta\} = \{T\} = 0$, a contradiction. $\square$

An immediate consequence of Corollary 3.2 is the following result in which the volume lower bound (3.6) falls short of the expected inequality (1.2). However, (3.6) solves the qualitative part of [6, Conjecture 10.1, (ii)] already solved in [20], while (3.5) gives moreover an effective estimate of the largest $t > 0$ for which the class $\{\alpha - t\beta\}$ remains pseudo-effective. This estimate will prompt the discussion of the psef and nef thresholds in the next section.

Theorem 3.5. *Let* X *be a compact Kähler manifold with* $\dim_{\mathbb{C}} X = n$ *and let* $\alpha, \beta > 0$ *be Kähler metrics such that* $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0$.

Then, for every $t \in [0, +\infty)$, *there exists a real* $(1, 1)$*-current* $T_t \in \{\alpha - t\beta\}$ *such that*

$$T_t \geq \left(1 - nt\,\frac{\{\alpha\}^{n-1}.\{\beta\}}{\{\alpha\}^n}\right)\alpha \quad \text{on} \quad X. \tag{3.5}$$

In particular, T_t *is a Kähler current for all* $0 \leq t < \frac{\{\alpha\}^n}{n\,\{\alpha\}^{n-1}.\{\beta\}}$, *so taking* $t = 1$ *(which is allowed by the assumption* $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0$*) we get that the class*

$\{\alpha-\beta\}$ *contains a Kähler current. Moreover, its volume satisfies:*

$$\begin{aligned}\mathrm{Vol}(\{\alpha-\beta\}) &\geq \big(\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}\big)\left(\frac{\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}}{\{\alpha\}^n}\right)^{n-1} \\ &\geq \{\alpha\}^n - n^2\,\{\alpha\}^{n-1}.\,\{\beta\}.\end{aligned} \tag{3.6}$$

Proof. Thanks to Lamari's positivity criterion (Lemma 1.5), the existence of a current $T_t \in \{\alpha - t\beta\}$ satisfying (3.5) is equivalent to

$$\int_X \left(\alpha - t\,\beta - \alpha + nt\,\frac{\{\alpha\}^{n-1}.\,\{\beta\}}{\{\alpha\}^n}\,\alpha\right)\wedge\gamma^{n-1} \geq 0$$

for every Gauduchon metric γ on X. This, in turn, is equivalent to

$$nt\,\frac{\{\alpha\}^{n-1}.\{\beta\}}{\{\alpha\}^n}\int_X \alpha\wedge\gamma^{n-1} \geq t\int_X \beta\wedge\gamma^{n-1} \quad \text{for every Gauduchon metric} \quad \gamma.$$

The last inequality is nothing but (3.4) which was proved in Corollary 3.2. This completes the proof of the existence of a current $T_t \in \{\alpha - t\beta\}$ satisfying (3.5).

Now, (3.5) implies that the absolutely continuous part T_{ac} of $T := T_1 \in \{\alpha-\beta\}$ has the same lower bound as T. Moreover, if $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\} > 0$, then

$$\mathrm{Vol}(\{\alpha-\beta\}) \geq \int_X T_{ac}^n \geq \left(1 - n\frac{\{\alpha\}^{n-1}.\{\beta\}}{\{\alpha\}^n}\right)^n \int_X \alpha^n \overset{(i)}{\geq} \left(1 - n^2\frac{\{\alpha\}^{n-1}.\,\{\beta\}}{\{\alpha\}^n}\right)\{\alpha\}^n,$$

which proves the claim (3.6). To obtain (i), we have used the elementary inequality $(1-\lambda)^n \geq 1 - n\lambda$ which holds for every $\lambda \in [0,\ 1]$. □

The above proof shows that a current $T_t \in \{\alpha - t\beta\}$ satisfying (3.5) exists even if we do not assume $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\} > 0$, although this information will be of use only under this assumption. Note that the non-orthogonality property $\{\alpha\}^{n-1}.\,\{\beta\} > 0$ ensured by Corollary 3.4 constitutes the obstruction to the volume lower estimate (3.6) being optimal (*i.e.* coinciding with the expected estimate (1.2)). We now point out an alternative way of inferring the same suboptimal volume lower bound (3.6) from the proof of Theorem 3.5.

Alternative wording of the proof of the volume lower estimate (3.6). By Lamari's positivity criterion (Lemma 1.5), the existence of a current T in the class $\{\alpha-\beta\}$ such that $T \geq \delta\alpha$ for some constant $\delta > 0$ (which must be such that $\delta < 1$) is equivalent to

$$\int_X ((1-\delta)\,\alpha - \beta)\wedge\gamma^{n-1} \geq 0, \quad \textit{i.e.}\text{ to} \quad \int_X \left((\alpha - \frac{1}{1-\delta}\,\beta\right)\wedge\gamma^{n-1} \geq 0,$$

for all Gauduchon metrics γ on X. Applying again Lamari's positivity criterion, this is still equivalent to the class $\{\alpha\} - (1/(1-\delta))\,\{\beta\}$ being pseudo-effective. Inequality (3.5) shows that the largest δ we can choose with this property is larger than or equal to

$$\delta_0 = 1 - n\,\frac{\{\alpha\}^{n-1}.\{\beta\}}{\{\alpha\}^n}, \quad \text{which gives} \quad 1-\delta_0 = n\,\frac{\{\alpha\}^{n-1}.\{\beta\}}{\{\alpha\}^n}. \tag{3.7}$$

On the other hand, we can write:

$$\begin{aligned} \mathrm{Vol}(\{\alpha-\beta\}) &= \mathrm{Vol}\left((1-t)\,\{\alpha\} + t\left(\{\alpha\} - \frac{1}{t}\,\{\beta\}\right)\right) \\ &\overset{(a)}{\geq} \mathrm{Vol}((1-t)\,\{\alpha\}) = (1-t)^n\,\{\alpha\}^n, \end{aligned} \tag{3.8}$$

where inequality (a) holds for every $t \in [0,\ 1]$ for which the class $\{\alpha\} - (1/t)\,\{\beta\}$ is *pseudo-effective*. By (3.7), $t := 1-\delta_0 = n\,\{\alpha\}^{n-1}.\{\beta\}/\{\alpha\}^n$ satisfies this property. With this choice of t, inequality (3.8) translates to the first inequality in (3.6).

Corollary 3.6. *Let* $\{\alpha\}, \{\beta\} \in H^{1,1}_{BC}(X,\ \mathbb{R})$ *be nef classes on a compact Kähler manifold* X *with* $\dim_{\mathbb{C}} X = n$ *such that* $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0$. *If* $\{\beta\} = 0$, *then* $\{\alpha\}$ *is big, while if* $\{\beta\} \neq 0$, *then* $\{\alpha\} - t\,\{\beta\}$ *is big for all* $0 \leq t < \frac{\{\alpha\}^n}{n\,\{\alpha\}^{n-1}.\,\{\beta\}}$. *Moreover, the volume lower bound* (3.6) *holds.*

The case when $\{\beta\} = 0$ is the key Theorem 2.12 in [14]. So, in particular, our method produces a much quicker proof of this fundamental result of [14]. The case when $\{\beta\} \neq 0$ is new, although the case $t = 1$ and the method of proof are those of [20]. Notice that the quantity $\{\alpha\}^n/n\,\{\alpha\}^{n-1}.\,\{\beta\} > 0$ is well defined when $\{\beta\} \neq 0$ by Corollary 3.4.

Proof. We fix an arbitrary Kähler metric ω on X and a constant $t \geq 0$ that will be specified shortly. The nefness assumption on $\{\alpha\}, \{\beta\}$ means that for every $\varepsilon > 0$, smooth forms $\alpha \in \{\alpha\}$ and $\beta \in \{\beta\}$ depending on ε can be found such that $\alpha_\varepsilon := \alpha + \varepsilon\,\omega$ and $\beta_\varepsilon := \beta + \frac{\varepsilon}{t}\,\omega$ are Kähler metrics. Notice that the class $\{\alpha_\varepsilon - t\beta_\varepsilon\} = \{\alpha - t\beta\}$ is independent of ε. On the other hand, the quantities $\{\alpha_\varepsilon\}^n = \{\alpha\}^n + \sum_{k=1}^n \varepsilon^k \binom{n}{k}\,\{\alpha\}^{n-k}.\{\omega\}^k$ and $\{\alpha_\varepsilon\}^{n-1}.\,\{\beta_\varepsilon\} = (\{\alpha\}^{n-1} + \sum_{l=1}^{n-1} \varepsilon^l \binom{n-1}{l}\,\{\alpha\}^{n-1-l}.\{\omega\}^l).(\{\beta\} + \frac{\varepsilon}{t}\,\{\omega\})$ converge to $\{\alpha\}^n$ and respectively $\{\alpha\}^{n-1}.\{\beta\}$ when $\varepsilon \to 0$. Thus, $\{\alpha_\varepsilon\}^n - n\,\{\alpha_\varepsilon\}^{n-1}.\{\beta_\varepsilon\} > 0$ if $\varepsilon > 0$ is small enough. Applying Theorem 3.5 to the Kähler metrics α_ε and β_ε, we infer that the class $\{\alpha_\varepsilon - t\beta_\varepsilon\} = \{\alpha - t\beta\}$ is big whenever $0 \leq t < \{\alpha_\varepsilon\}^n/n\,\{\alpha_\varepsilon\}^{n-1}.\,\{\beta_\varepsilon\}$. In particular, if $\{\beta\} = 0$, this means that the class $\{\alpha\}$ is big (since we can fix $\varepsilon > 0$ and choose $t = 0$). Meanwhile, if $\{\beta\} \neq 0$ and if we choose $t < \{\alpha\}^n/n\,\{\alpha\}^{n-1}.\{\beta\}$, then $t < \{\alpha_\varepsilon\}^n/n\,\{\alpha_\varepsilon\}^{n-1}.\{\beta_\varepsilon\}$ for all $\varepsilon > 0$ small enough and we conclude that $\{\alpha - t\beta\}$ is big. The volume lower bound (3.6) holds for $\{\alpha_\varepsilon\}$ and $\{\beta_\varepsilon\}$ for $t = 1$ and all sufficiently small $\varepsilon > 0$, so letting $\varepsilon \to 0$ and using the continuity of the volume, we get it for $\{\alpha\}$ and $\{\beta\}$. $\square$

4. Trace and volume of (1, 1)-cohomology classes

The implicit discussion of the relative positivity thresholds of a cohomology class with respect to another in Theorem 3.5 and in Corollary 3.6 prompts a further investigation of their relationships with the volume that we undertake to study in this section.

4.1. The psef threshold

Let X be a compact complex manifold in Fujiki's *class* $\mathcal{C}$, $n := \dim_{\mathbb{C}} X$.

Definition 4.1. For every big Bott-Chern class $\{\beta\} = [\beta]_{BC} \in H^{1,1}_{BC}(X, \mathbb{R})$, we define the β-directed trace (or the psef threshold in the β-direction) to be the function:

$$P^{(\beta)} \ : \ H^{1,1}_{BC}(X, \mathbb{R}) \to \mathbb{R}, \quad P^{(\beta)}(\alpha) := \inf \int_X \alpha \wedge \gamma^{n-1}, \tag{4.1}$$

for all Bott-Chern classes $\{\alpha\} = [\alpha]_{BC} \in H^{1,1}_{BC}(X, \mathbb{R})$, where the infimum is taken over all the Gauduchon metrics γ on X normalised such that

$$[\beta]_{BC}.[\gamma^{n-1}]_A = \int_X \beta \wedge \gamma^{n-1} = 1. \tag{4.2}$$

All the integrals involved in the above definition are clearly independent of the representatives α, β of the Bott-Chern classes $[\alpha]_{BC}$, $[\beta]_{BC}$ and of the representative γ^{n-1} of the Aeppli-Gauduchon class $[\gamma^{n-1}]_A \in H^{n-1,n-1}_A(X, \mathbb{R})$. Thus the infimum is taken over the subset S_β of the Gauduchon cone $\mathcal{G}_X$ consisting of classes $[\gamma^{n-1}]_A$ normalised by $[\beta]_{BC}.[\gamma^{n-1}]_A = 1$. The bigness assumption on $[\beta]_{BC}$ has been imposed to ensure that $[\beta]_{BC}.[\gamma^{n-1}]_A > 0$, so that $[\gamma^{n-1}]_A$ can be normalised with respect to $[\beta]_{BC}$ as in (4.2), for *every* class $[\gamma^{n-1}]_A \in \mathcal{G}_X$.

This definition is motivated in part by the next observation which is an immediate consequence of Lamari's positivity criterion: the β-directed trace $P^{(\beta)}$ coincides with the slope function introduced for big classes $\{\alpha\}$ in [8, Definition 3.7] and thus gauges the positivity of real (1, 1)-classes $\{\alpha\}$ with respect to a reference big class $\{\beta\}$. The quantity on the right-hand side of (4.3) below (*i.e.* the slope) may well be called the psef threshold of $\{\alpha\}$ in the $\{\beta\}$-direction (a term already used in the literature).

Proposition 4.2. *Suppose that* $\{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ *is a fixed big class. Then*

$$\begin{aligned} P^{(\beta)}(\alpha) &= \sup\{t \in \mathbb{R} / \text{the class } \{\alpha\} - t\{\beta\} \text{ is psef}\} \\ &= \sup\{t \in \mathbb{R} / \exists T \in \{\alpha\} \text{ current, } \exists\, \widetilde{\beta} \in \{\beta\} C^\infty\text{-form so that} T \geq t\widetilde{\beta}\} \\ &= \sup\{t \in \mathbb{R} / \forall \widetilde{\beta} \in \{\beta\} C^\infty\text{-form, } \exists\, T \in \{\alpha\} \text{ current so that } T \geq t\widetilde{\beta}\}, \end{aligned} \tag{4.3}$$

for every class $\{\alpha\} \in H^{1,1}_{BC}(X, \mathbb{R})$. *In particular, the set* $\{t \in \mathbb{R}$ / *the class* $\{\alpha\} - t\{\beta\}$ *is psef*$\}$ *equals the interval* $(-\infty, \ P^{(\beta)}(\alpha)]$.

Proof. Let $A_\alpha^\beta := \{t \in \mathbb{R} \ / \text{ the class } \{\alpha\} - t\,\{\beta\} \text{ is psef}\}$ and let $t_\alpha^\beta := \sup\ A_\alpha^\beta$. By Lamari's positivity criterion, the class $\{\alpha\} - t\,\{\beta\}$ is psef if and only if

$$\int_X \alpha\wedge\gamma^{n-1} \geq t\int_X \beta\wedge\gamma^{n-1} \quad \text{for all } [\gamma^{n-1}]_A \in \mathcal{G}_X \iff \int_X \alpha\wedge\gamma^{n-1} \geq t$$

for all $[\gamma^{n-1}]_A \in \mathcal{G}_X$ normalised such that $\int_X \beta\wedge\gamma^{n-1} = 1$. This proves the inequality $P^{(\beta)}(\alpha) \geq t_\alpha^\beta$. To prove that equality holds, we reason by contradiction. Suppose that $P^{(\beta)}(\alpha) > t_\alpha^\beta$. Pick any t_1 such that $P^{(\beta)}(\alpha) > t_1 > t_\alpha^\beta$. Then $\int_X \alpha\wedge\gamma^{n-1} > t_1$ for all Gauduchon metrics γ such that $[\beta]_{BC}.[\gamma^{n-1}]_A = 1$. This is equivalent to $\int_X \alpha\wedge\gamma^{n-1} > t_1\ \int_X \beta\wedge\gamma^{n-1}$ for all Gauduchon metrics γ, which thanks to Lamari's positivity criterion implies $\exists\, T \in \{\alpha\} - t_1\,\{\beta\}$ such that $T \geq 0$, *i.e.* the class $\{\alpha\} - t_1\,\{\beta\}$ is psef.

Thus $t_1 \in A_\alpha^\beta$, contradicting the choice $t_1 > t_\alpha^\beta = \sup\ A_\alpha^\beta$. □

An immediate consequence is the next statement showing that the β-directed trace (*i.e.* the psef threshold) gauges the positivity of real Bott-Chern $(1,\ 1)$-classes much as the volume does.

Corollary 4.3. *Suppose that $\{\beta\} \in H_{BC}^{1,1}(X,\ \mathbb{R})$ is a fixed big class. For any class $\{\alpha\} \in H_{BC}^{1,1}(X,\ \mathbb{R})$, the following equivalences hold:*

(i) $\{\alpha\}$ *is psef* $\iff P^{(\beta)}(\alpha) \geq 0$;
(ii) $\{\alpha\}$ *is big* $\iff P^{(\beta)}(\alpha) > 0$.

Proof. (i) follows at once from (4.3) and so does (ii) after we (trivially) notice that the class $\{\alpha\}$ is big if and only if there exists $\varepsilon > 0$ such that $\{\alpha\} - \varepsilon\{\beta\}$ is psef. Indeed, this is a consequence of the fixed class $\{\beta\}$ being supposed big. □

Next, we observe some easy but useful properties of the β-directed trace.

Proposition 4.4. *Suppose that $\{\beta\} \in H_{BC}^{1,1}(X,\ \mathbb{R})$ is a fixed big class.*

(i) *For all classes $\{\alpha_1\}, \{\alpha_2\} \in H_{BC}^{1,1}(X,\ \mathbb{R})$, we have*

$$P^{(\beta)}(\alpha_1+\alpha_2) \geq P^{(\beta)}(\alpha_1) + P^{(\beta)}(\alpha_2). \tag{4.4}$$

In particular, $P^{(\beta)}(\alpha_1) \geq P^{(\beta)}(\alpha_2)$ whenever $\{\alpha_1\} \geq_{psef} \{\alpha_2\}$ (in the sense that $\{\alpha_1 - \alpha_2\}$ is psef);

(ii) *For any class $\{\alpha\} \in H_{BC}^{1,1}(X,\ \mathbb{R})$ and any $t \in \mathbb{R}$, we have*

$$P^{(\beta)}(t\,\alpha) = t\,P^{(\beta)}(\alpha) \quad \textit{and, if } t > 0, \quad P^{(t\,\beta)}(\alpha) = \frac{1}{t}\,P^{(\beta)}(\alpha); \tag{4.5}$$

(iii) *For every big class $\{\alpha\} \in H_{BC}^{1,1}(X,\ \mathbb{R})$, we have*

$$P^{(\alpha)}(\alpha) = 1. \tag{4.6}$$

Proof. Let $\{\alpha_1\}, \{\alpha_2\} \in H^{1,1}_{BC}(X, \mathbb{R})$. Since $\int_X(\alpha_1+\alpha_2)\wedge\gamma^{n-1} = \int_X \alpha_1\wedge\gamma^{n-1} + \int_X \alpha_2 \wedge \gamma^{n-1}$ for every $[\gamma^{n-1}]_A \in H^{n-1,\,n-1}_A(X, \mathbb{R})$, we get

$$\inf \int_X (\alpha_1 + \alpha_2) \wedge \gamma^{n-1} \geq \inf \int_X \alpha_1 \wedge \gamma^{n-1} + \inf \int_X \alpha_1 \wedge \gamma^{n-1},$$

where the infima are taken over all $[\gamma^{n-1}] \in S_\beta$. This proves (i).

(ii) Follows immediately from $\int_X t\alpha \wedge \gamma^{n-1} = t \int_X \alpha \wedge \gamma^{n-1}$ and from the fact that $[\gamma^{n-1}]_A$ is $(t\beta)$-normalised if and only if $t[\gamma^{n-1}]_A$ is β-normalised.

(iii) Follows from $\int_X \alpha\wedge\gamma^{n-1}{=}1$ for all $[\gamma^{n-1}]_A$ such that $[\alpha]_{BC}.[\gamma^{n-1}]_A{=}1$. □

The next observation deals with the variation of $P^{(\beta)}$ when $\{\beta\}$ varies. As usual, an inequality $\{\alpha\} \geq_{psef} \{\beta\}$ between real $(1,\ 1)$-classes will mean that the difference class $\{\alpha - \beta\}$ is psef.

Proposition 4.5. *Let $\{\beta_1\}, \{\beta_2\} \in H^{1,1}_{BC}(X, \mathbb{R})$ be big classes.*

(i) *If $\{\beta_1\} \geq_{psef} C\,\{\beta_2\}$ for some constant $C > 0$, then*

$$P^{(\beta_1)} \leq \frac{1}{C}\, P^{(\beta_2)} \quad \textit{on the psef cone} \quad \mathcal{E}_X \subset H^{1,1}_{BC}(X, \mathbb{R}); \tag{4.7}$$

(ii) *The following inequality holds:*

$$P^{(\beta_2)}(\beta_1)\, P^{(\beta_1)} \leq P^{(\beta_2)} \quad \textit{on the psef cone} \quad \mathcal{E}_X \subset H^{1,1}_{BC}(X, \mathbb{R}). \tag{4.8}$$

Proof. If $\{\beta_1 - C\,\beta_2\}$ is psef, then $\int_X(\beta_1 - C\,\beta_2)\wedge\gamma^{n-1} \geq 0$, *i.e.* $[\beta_1]_{BC}.[\gamma^{n-1}]_A \geq C\,[\beta_2]_{BC}.[\gamma^{n-1}]_A$, for all classes $[\gamma^{n-1}]_A \in \mathcal{G}_X$. It follows that, for every psef class $\{\alpha\} \in H^{1,1}_{BC}(X, \mathbb{R})$, we have:

$$\int_X \alpha \wedge \frac{\gamma^{n-1}}{\int_X \beta_1 \wedge \gamma^{n-1}} \leq \frac{1}{C} \int_X \alpha \wedge \frac{\gamma^{n-1}}{\int_X \beta_2 \wedge \gamma^{n-1}} \quad \text{for all} \quad [\gamma^{n-1}]_A \in \mathcal{G}_X.$$

Taking infima over all $[\gamma^{n-1}]_A \in \mathcal{G}_X$, we get (4.7). On the other hand, it follows from (4.3) that

$$\{\beta_1\} \geq_{psef} P^{(\beta_2)}(\beta_1)\,\{\beta_2\},$$

which in turn implies (4.8) thanks to (4.7) applied with $C = P^{(\beta_2)}(\beta_1)$. □

4.2. The nef threshold

We now observe that the discussion of the psef threshold in Subsection 4.1 can be run analogously in the nef context using the following important result of [14, Corollary 0.4].

Theorem 4.6 (Demailly-Paun 2004). *Let X be a compact Kähler manifold, $\dim_{\mathbb{C}} X = n$. Then the dual of the nef cone $\overline{\mathcal{K}_X} \subset H^{1,1}(X, \mathbb{R})$ under the Serre duality is the closed convex cone $\mathcal{N}_X \subset H^{n-1,n-1}(X, \mathbb{R})$ generated by classes of currents of the shape $[Y] \wedge \omega^{n-p-1}$, where Y runs over the irreducible analytic subsets of X of any codimension $p = 0, 1, \ldots, n-1$ and $\{\omega\}$ runs over the Kähler classes of X.*

Let $\{\alpha\}, \{\beta\} \in H^{1,1}(X, \mathbb{R})$ be arbitrary classes on a compact Kähler n-fold X. By Theorem 4.6, for any $s \in \mathbb{R}$, the class $\{\alpha - s\beta\}$ is nef if and only if

$$\int_Y \alpha \wedge \omega^{n-p-1} \geq s \int_Y \beta \wedge \omega^{n-p-1}, \quad p = 0, 1, \ldots, n-1, \ \mathrm{codim}_X Y = p, \{\omega\} \in \mathcal{K}_X.$$

(As usual, $\mathcal{K}_X$ denotes the Kähler cone of X.) This immediately implies the following statement.

Proposition 4.7. *Let $\{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ be any Kähler class on a compact Kähler n-fold X. The nef threshold of any $\{\alpha\} \in H^{1,1}_{BC}(X, \mathbb{R})$ in the $\{\beta\}$-direction, defined by the first identity below, also satisfies the second identity:*

$$N^{(\beta)}(\alpha) := \inf \int_Y \alpha \wedge \omega^{n-p-1} = \sup\{s \in \mathbb{R} \ / \ \textit{the class } \{\alpha\} - s\{\beta\} \textit{ is nef}\}, \quad (4.9)$$

where the infimum is taken over all $p = 0, 1, \ldots, n-1$, over all the irreducible analytic subsets $Y \subset X$ such that codim $Y = p$ *and over all Kähler classes $\{\omega\}$ normalised such that $\int_Y \beta \wedge \omega^{n-p-1} = 1$. In particular, the set $\{s \in \mathbb{R} \ / \ \text{the class } \{\alpha\} - s\{\beta\} \text{ is nef}\}$ equals the interval $(-\infty, N^{(\beta)}(\alpha)]$.*

Thus, we obtain a function $N^{(\beta)} : H^{1,1}_{BC}(X, \mathbb{R}) \to \mathbb{R}$. It is clear that

$$N^{(\beta)}(\alpha) \leq P^{(\beta)}(\alpha) \quad \textit{for all} \quad \{\alpha\} \in H^{1,1}_{BC}(X, \mathbb{R}) \quad (4.10)$$

thanks to the supremum characterisations of the two thresholds and to the well-known implication "nef $\Longrightarrow$ psef".

It is precisely in order to ensure that $\int_Y \beta \wedge \omega^{n-p-1} > 0$, hence that $\{\omega\}$ can be normalised as stated, for any Kähler class $\{\omega\}$ and any $Y \subset X$ that we assumed $\{\beta\}$ to be Kähler. The two-fold characterisations of the nef and the psef thresholds yield at once the following consequence.

Observation 4.8. Suppose that no analytic subset $Y \subset X$ exists except in codimensions 0 and n. Then $N^{(\beta)}(\alpha) = P^{(\beta)}(\alpha)$ for all Kähler classes $\{\alpha\}, \{\beta\}$.

Proof. If $Y = X$ is the only analytic subset of X of codimension $p < n$, then $N^{(\beta)}(\alpha) = \inf \int_X \alpha \wedge \omega^{n-1}$ where the infimum is taken over all the Kähler classes $\{\omega\}$, *i.e.* over all the Aeppli-Gauduchon classes $[\omega^{n-1}]_A$ representable by the $(n-1)^{st}$ power of a Kähler metric, normalised such that $\int_X \beta \wedge \omega^{n-1} = 1$. Since these classes form a subset of all the Aeppli-Gauduchon classes $[\gamma^{n-1}]_A$ normalised by $\int_X \beta \wedge \gamma^{n-1} = 1$, we get $N^{(\beta)}(\alpha) \geq P^{(\beta)}(\alpha)$. However, the reverse inequality always holds, hence equality holds. □

An immediate consequence of Proposition 4.7 is the following analogue of Corollary 4.3 for the nef/Kähler context.

Corollary 4.9. *Suppose that $\{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ is a fixed Kähler class. For any class $\{\alpha\} \in H^{1,1}_{BC}(X, \mathbb{R})$, the following equivalences hold:*

(i) $\{\alpha\}$ *is nef* $\iff N^{(\beta)}(\alpha) \geq 0$;
(ii) $\{\alpha\}$ *is Kähler* $\iff N^{(\beta)}(\alpha) > 0$.

In particular, if no analytic subset $Y \subset X$ exists except in codimensions 0 *and n, then the following* (*actually known, see* [11]) *equivalences hold:*

(a) $\{\alpha\}$ *is nef* $\iff$ $\{\alpha\}$*is psef;*
(b) $\{\alpha\}$ *is Kähler* $\iff$ $\{\alpha\}$ *is big.*

Proof. (i) follows at once from (4.9) and so does (ii) after we (trivially) notice that the class $\{\alpha\}$ is Kähler if and only if there exists $\varepsilon > 0$ such that $\{\alpha\} - \varepsilon\{\beta\}$ is nef. Indeed, this is a consequence of the fixed class $\{\beta\}$ being supposed Kähler and of the Kähler cone being the interior of the nef cone. □

We immediately get analogues of Propositios 4.4 and 4.5 for $N^{(\beta)}(\alpha)$ in place of $P^{(\beta)}(\alpha)$ and for the order relation $\geq_{nef}$ in place of $\geq_{psef}$, where $\{\alpha\} \geq_{nef} \{\beta\}$ means that the class $\{\alpha - \beta\}$ is nef.

4.3. Relations of the psef/nef threshold to the volume

We now relate the β-directed trace of a Kähler class $\{\alpha\}$ to the volume of $\{\alpha - \beta\}$.

Proposition 4.10.

(i) *For any Kähler classes $\{\alpha\}, \{\beta\}$ on a compact Kähler n-fold X, we have:*

$$\frac{\{\alpha\}^n}{n\,\{\alpha\}^{n-1}.\{\beta\}} \overset{(a)}{\leq} P^{(\beta)}(\alpha) \overset{(b)}{\leq} \frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}}. \tag{4.11}$$

In fact, it suffices to suppose that $\{\beta\}$ is big in the inequality (b). In particular, if $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0$, then $P^{(\beta)}(\alpha) > 1$ (*hence we find again that $\{\alpha - \beta\}$ is big in this case*)*;*

(ii) *For any Kähler classes* $\{\alpha\}, \{\beta\}$ *such that* $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0$, *we have:*

$$\mathrm{Vol}(\{\alpha-\beta\}) \geq \left(1 - \frac{1}{P^{(\beta)}(\alpha)}\right)^n \{\alpha\}^n. \tag{4.12}$$

Note that the combination of (4.12) *and part (a) of* (4.11) *is the volume lower bound* (3.6).

Proof. (i) Inequality (b) is trivial: it suffices to choose $[\gamma^{n-1}]_A = t\,[\alpha^{n-1}]_A$ for the constant $t > 0$ satisfying the β-normalisation condition $[\beta]_{BC}.t\,[\alpha^{n-1}]_A = 1$, *i.e.* $t = 1/\{\alpha\}^{n-1}.\{\beta\}$, and to use the definition of $P^{(\beta)}(\alpha)$ as an infimum.

Inequality (a) follows from Corollary 3.2 by taking the infimum over all the Gauduchon metrics γ normalised by $\int_X \beta\wedge\gamma^{n-1} = 1$ in (3.4).

(ii) We saw in the second proof of the lower estimate (3.6) that (3.8) holds for every $t \in [0,\ 1]$ such that $\{\alpha\} - (1/t)\,\{\beta\}$ is psef. Now, (4.3) shows that the infimum of all these t is $1/P^{(\beta)}(\alpha)$. Thus (3.8) holds for $t = 1/P^{(\beta)}(\alpha)$, yielding (4.12). □

A similar link between the volume and the nef threshold is given in the next result by considering Monge-Ampère equations on analytic subsets $Y \subset X$.

Proposition 4.11. *For every Kähler classes* $\{\alpha\}, \{\beta\}$ *on a compact Kähler n-fold* X, *we have:*

$$\begin{aligned}\inf_{\substack{p=0,1,\dots,n-1,\\ Y\subset X,\ \mathrm{codim}\ Y=p}} \frac{\mathrm{Vol}_Y(\alpha)}{(n-p)\,\{\alpha\}^{n-p-1}.\{\beta\}.\{[Y]\}} &\overset{(a)}{\leq} N^{(\beta)}(\alpha)\\ \overset{(b)}{\leq} \inf_{\substack{p=0,1,\dots,n-1,\\ Y\subset X,\ \mathrm{codim}\ Y=p}} \frac{\mathrm{Vol}_Y(\alpha)}{\{\alpha\}^{n-p-1}.\{\beta\}.\{[Y]\}}&\end{aligned} \tag{4.13}$$

where the infima are taken over the analytic subsets $Y \subset X$. *We have set* $\mathrm{Vol}_Y(\alpha) := \int_Y \alpha^{n-p} = \int_X \alpha^{n-p}\wedge[Y]$ *and* $\{\alpha\}^{n-p-1}.\{\beta\}.\{[Y]\} := \int_Y \alpha^{n-p-1}\wedge\beta = \int_X \alpha^{n-p-1}\wedge \beta\wedge[Y]$ (*both quantities depending only on p and the classes* $\{\alpha\}, \{\beta\}, \{[Y]\}$).

Proof. Pick any Kähler metrics $\alpha \in \{\alpha\}$ and $\beta \in \{\beta\}$. Let $Y \subset X$ be any analytic subset of arbitrary codimension $p \in \{0, 1, \dots, n-1\}$ and let ω be any Kähler metric on X normalised such that $\int_Y \beta\wedge\omega^{n-p-1} = 1$. We can solve the following Monge-Ampère equation:

$$\widetilde{\alpha}_Y^{n-p} = \mathrm{Vol}_Y(\alpha)\,\beta\wedge\omega^{n-p-1} \quad \text{on} \quad Y, \tag{4.14}$$

in the sense that there exists a d-closed (weakly) positive $(1,\ 1)$-current $\widetilde{\alpha}_Y$ on Y (*cf.* [10, Definition 1.2]) lying in the restricted class $\{\alpha\}_{|Y}$ such that $\widetilde{\alpha}_Y$ is C^∞ on the regular part Y_{reg} of Y. We defer to the end of the proof the explanation of how this follows from results in the literature. We adopt the standard point of view (see [10, Section 1]) according to which C^∞ forms on a singular variety Y are defined locally as restrictions to Y_{reg} of C^∞ forms on an open subset of some $\mathbb{C}^N$

into which Y locally embeds. In what follows, the exterior powers and products involving $\widetilde{\alpha}_Y$ are to be understood on Y_{reg} even when we write Y.

If we define $\det_\omega \widetilde{\alpha}_Y$ by requiring $\widetilde{\alpha}_Y^{n-p} = (\det_\omega \widetilde{\alpha}_Y)\,\omega^{n-p}$ on Y, then (4.14) translates to the identity:

$$\det_\omega \widetilde{\alpha}_Y = \frac{\mathrm{Vol}_Y(\alpha)}{n-p}\,\Lambda_\omega \beta_{|Y} \quad \text{on} \quad Y. \tag{4.15}$$

Thus, the argument in the proof of Lemma 3.1 can be rerun on Y as follows:

$$\begin{aligned}
&\left(\int_Y \widetilde{\alpha}_Y \wedge \omega^{n-p-1}\right)\left(\int_Y \widetilde{\alpha}_Y^{n-p-1} \wedge \beta\right) \\
&= \frac{1}{(n-p)^2}\left(\int_Y (\Lambda_\omega \widetilde{\alpha}_Y)\,\omega^{n-p}\right)\left(\int_Y (\Lambda_{\widetilde{\alpha}_Y}\beta_{|Y})\,(\det_\omega \widetilde{\alpha}_Y)\,\omega^{n-p}\right) \\
&\overset{(a)}{\geq} \frac{1}{(n-p)^2}\left(\int_Y [(\Lambda_\omega \widetilde{\alpha}_Y)\,(\Lambda_{\widetilde{\alpha}_Y}\beta_{|Y})]^{\frac{1}{2}}\,(\det_\omega \widetilde{\alpha}_Y)^{\frac{1}{2}}\,\omega^{n-p}\right)^2 \\
&\overset{(b)}{\geq} \frac{1}{(n-p)^2}\left(\int_Y (\Lambda_\omega \beta_{|Y})^{\frac{1}{2}}\left(\frac{\mathrm{Vol}_Y(\alpha)}{n-p}\right)^{\frac{1}{2}}(\Lambda_\omega \beta_{|Y})^{\frac{1}{2}}\,\omega^{n-p}\right)^2 \\
&= \frac{\mathrm{Vol}_Y(\alpha)}{n-p}\left(\frac{1}{(n-p)}\int_Y (\Lambda_\omega \beta_{|Y})\,\omega^{n-p}\right)^2 = \frac{\mathrm{Vol}_Y(\alpha)}{n-p}\left(\int_Y \beta \wedge \omega^{n-p-1}\right)^2 \\
&\overset{(c)}{=} \frac{\mathrm{Vol}_Y(\alpha)}{n-p},
\end{aligned}$$

where (a) is an application of the Cauchy-Schwarz inequality, (b) follows from the pointwise inequality $(\Lambda_\omega \widetilde{\alpha}_Y)\,(\Lambda_{\widetilde{\alpha}_Y}\beta) \geq \Lambda_\omega \beta$ (*cf.* [20, Lemma 3.1]) and from (4.15), while (c) follows from the normalisation $\int_X \beta \wedge \omega^{n-p-1} = 1$.

Thus, since $\int_Y \widetilde{\alpha}_Y \wedge \omega^{n-p-1} = \int_Y \alpha \wedge \omega^{n-p-1}$ and $\int_Y \widetilde{\alpha}_Y^{n-p-1} \wedge \beta = \int_Y \alpha^{n-p-1} \wedge \beta$, we get:

$$\int_Y \alpha \wedge \omega^{n-p-1} \geq \frac{\mathrm{Vol}_Y(\alpha)}{(n-p)\,\{\alpha\}^{n-p-1}.\{\beta\}.\{[Y]\}}$$

for every analytic subset $Y \subset X$ and every Kähler metric ω normalised by $\int_Y \beta \wedge \omega^{n-p-1} = 1$. This proves inequality (a) in (4.13).

The proof of inequality (b) in (4.13) follows immediately by choosing the Kähler metric ω to be proportional to α, *i.e.* $\omega = t\alpha$ for the constant $t = t_Y > 0$ determined by the normalisation condition $\int_Y \beta \wedge \omega^{n-p-1} = 1$ once $Y \subset X$ has been chosen. Indeed, for every $p = 0, 1, \ldots, n-1$ and every analytic subset $Y \subset X$, we immediately get:

$$\inf_\omega \int_Y \alpha \wedge \omega^{n-p-1} \leq \int_Y \alpha \wedge (t\alpha)^{n-p-1} = \frac{\int_Y \alpha^{n-p}}{\int_Y \beta \wedge \alpha^{n-p-1}}$$

which implies part (b) of (4.13) after taking the infimum over p and Y.

It remains to explain how the solution of equation (4.14) is obtained. If Y is smooth, Yau's classical theorem in [26] ensures the existence and uniqueness of a Kähler metric $\widetilde{\alpha}_Y$ in $\{\alpha\}_{|Y}$ which solves (4.14). If Y is singular, we choose a desingularisation $\widetilde{Y}$ of Y that is a finite sequence of blow-ups with smooth centres in X:

$$\mu : \widetilde{Y} \longrightarrow Y, \quad \text{which is the restriction of} \quad \mu : \widetilde{X} \longrightarrow X.$$

Thus, $\mu : \widetilde{X} \setminus \mu^{-1}(Z) \longrightarrow X \setminus Z$ is a biholomorphism above the complement of the analytic subset $Z := Y_{sing}$ and $\widetilde{X}$ is a compact Kähler manifold, hence so is the submanifold $\widetilde{Y}$. Moreover, $\mu^\star(\beta \wedge \omega^{n-p-1})$ is a C^∞ semi-positive $(n-p,\, n-p)$-form on $\widetilde{X}$ that is strictly positive on $\widetilde{X} \setminus \mu^{-1}(Z)$. Clearly, $\mu^\star\{\alpha\} = \{\mu^\star\alpha\}$ is a semi-positive (hence also nef) big class on $\widetilde{X}$ and

$$\mathrm{Vol}_{\widetilde{Y}}\left(\mu^\star\{\alpha\}\right) = \int_{\widetilde{X}} \left(\mu^\star\alpha\right)^{n-p} \wedge [\widetilde{Y}] = \int_X \alpha^{n-p} \wedge [Y] = \mathrm{Vol}_Y(\alpha) > 0.$$

We consider the following Monge-Ampère equation on the (smooth) compact Kähler manifold $\widetilde{Y}$:

$$\widetilde{\alpha}_{\widetilde{Y}}^{n-p} = \mathrm{Vol}_{\widetilde{Y}}\left(\mu^\star\{\alpha\}\right) \mu^\star\left(\beta \wedge \omega^{n-p-1}\right) \quad \text{on} \quad \widetilde{Y}. \tag{4.16}$$

If the class $\mu^\star\{\alpha\}$ were Kähler, Yau's Theorem 3 in [26] on solutions of the Monge-Ampère equation with a degenerate (*i.e.*, semi-positive) smooth right-hand side would yield a unique d-closed $(1,\, 1)$-current $\widetilde{\alpha}_{\widetilde{Y}} \in \mu^\star\{\alpha\}_{|\widetilde{Y}}$ solving equation (4.16) such that $\widetilde{\alpha}_{\widetilde{Y}} \geq 0$ on $\widetilde{Y}$, $\widetilde{\alpha}_{\widetilde{Y}}$ is C^∞ on $\widetilde{Y} \setminus \mu^{-1}(Z)$ and $\widetilde{\alpha}_{\widetilde{Y}}$ has locally bounded coefficients on $\widetilde{Y}$. In our more general case where the class $\mu^\star\{\alpha\}$ is only semi-positive and big, Theorems A, B, C in [7] yield a unique d-closed $(1,\, 1)$-current $\widetilde{\alpha}_{\widetilde{Y}} \in \mu^\star\{\alpha\}_{|\widetilde{Y}}$ such that $\widetilde{\alpha}_{\widetilde{Y}} \geq 0$ on $\widetilde{Y}$ and

$$\left\langle \widetilde{\alpha}_{\widetilde{Y}}^{n-p} \right\rangle = \mathrm{Vol}_{\widetilde{Y}}\left(\mu^\star\{\alpha\}\right) \mu^\star\left(\beta \wedge \omega^{n-p-1}\right) \quad \text{on} \quad \widetilde{Y},$$

where $\langle \quad \rangle$ stands for the non-pluripolar product introduced in [7]. Moreover, $\widetilde{\alpha}_{\widetilde{Y}}$ is C^∞ on the ample locus of the class $\mu^\star\{\alpha\}_{|\widetilde{Y}}$ (*cf.* [7, Theorem C]), which in our case coincides with $\widetilde{Y} \setminus \mu^{-1}(Z)$, and $\widetilde{\alpha}_{\widetilde{Y}}$ has minimal singularities (*cf.* [7, Theorem B]) among the positive currents in the class $\mu^\star\{\alpha\}_{|\widetilde{Y}}$. Since this class contains C^∞ semi-positive forms (*e.g.*, $(\mu^\star\alpha)_{|\widetilde{Y}}$), its currents with minimal singularties have locally bounded potentials. Thus, $\widetilde{\alpha}_{\widetilde{Y}}$ has locally bounded (and even continuous) potentials, so $\langle \widetilde{\alpha}_{\widetilde{Y}}^{n-p} \rangle$ equals the exterior power $\widetilde{\alpha}_{\widetilde{Y}}^{n-p}$ in the sense of Bedford and Taylor [3]. In particular, $[\widetilde{\alpha}_{\widetilde{Y}}^k]_{BC} = [\mu^\star\alpha^k]_{BC}$ for all k, so $\int_{\widetilde{Y}} \widetilde{\alpha}_{\widetilde{Y}}^{n-p-1} \wedge \mu^\star\beta = \int_{\widetilde{Y}} \mu^\star\alpha^{n-p-1} \wedge \mu^\star\beta$. It remains to set

$$\widetilde{\alpha}_Y := \mu_\star \widetilde{\alpha}_{\widetilde{Y}}.$$

We thus get a d-closed positive $(1,\, 1)$-current $\widetilde{\alpha}_Y \in \{\alpha\}_{|Y}$ whose restriction to $Y_{reg} = Y \setminus Z$ is C^∞ and which solves the Monge-Ampère equation (4.14). $\square$

We can now relate both the psef and the nef thresholds $P^{(\beta)}(\alpha)$, $N^{(\beta)}(\alpha)$ to the volume of $\{\alpha-\beta\}$. The next result confirms Conjecture 1.1 in the case when these thresholds are sufficiently *close* to each other.

Proposition 4.12. *Let X be a compact Kähler manifold,* $\dim_{\mathbb{C}} X = n$, *and let* $\{\alpha\}, \{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ *be Kähler classes such that*

$$\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0. \tag{4.17}$$

If either of the following two conditions is satisfied:

$$(i)\quad N^{(\beta)}(\alpha) \geq 1 \qquad or \qquad (ii)\quad N^{(\beta)}(\alpha) \geq \frac{\frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}} - P^{(\beta)}(\alpha)}{n-1}, \tag{4.18}$$

then

$$\mathrm{Vol}(\{\alpha-\beta\}) \geq \{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\}. \tag{4.19}$$

Note that $P^{(\beta)}(\alpha) \geq \frac{\frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}} - P^{(\beta)}(\alpha)}{n-1}$ thanks to inequality (a) in (4.11). Since $P^{(\beta)}(\alpha) \geq N^{(\beta)}(\alpha)$ (*cf.* (4.10)), this shows that condition (ii) requires $N^{(\beta)}(\alpha)$ to be "close" to $P^{(\beta)}(\alpha)$. In particular, (ii) holds if $N^{(\beta)}(\alpha)$ and $P^{(\beta)}(\alpha)$ coincide.

Proof of Proposition 4.12. If $N^{(\beta)}(\alpha) \geq 1$, then the class $\{\alpha-\beta\}$ is nef (*cf.* Proposition 4.7), so (4.19) follows from Proposition 2.2 in this case.

Let us now suppose that $N^{(\beta)}(\alpha) < 1$ and that condition (ii) is satisfied. We set $s_0 := N^{(\beta)}(\alpha)$ and $t_0 := P^{(\beta)}(\alpha)$, so $s_0 < 1 < t_0$ (where the last inequality follows from $\{\alpha-\beta\}$ being big – the main result in [20]). We have

$$\{\alpha-\beta\} = \frac{t_0-1}{t_0-s_0}\,\{\alpha - s_0\beta\} + \frac{1-s_0}{t_0-s_0}\,\{\alpha - t_0\beta\}. \tag{4.20}$$

Since the class $(1-s_0)/(t_0-s_0)\cdot\{\alpha - t_0\beta\}$ is psef, we get the first inequality below:

$$\begin{aligned} \mathrm{Vol}(\{\alpha-\beta\}) &\geq \left(\frac{t_0-1}{t_0-s_0}\right)^n \mathrm{Vol}(\{\alpha - s_0\,\beta\}) \\ &\geq \left(1 - \frac{1-s_0}{t_0-s_0}\right)^n \left(\{\alpha\}^n - ns_0\{\alpha\}^{n-1}.\{\beta\}\right), \end{aligned} \tag{4.21}$$

where the second inequality follows from Proposition 2.2 since the class $\{\alpha - s_0\,\beta\}$ is nef. Let

$$f : [0,\ 1] \to [0,\ +\infty), \quad f(s) := \left(1 - \frac{1-s}{t_0-s}\right)^n \left(\{\alpha\}^n - ns\{\alpha\}^{n-1}.\{\beta\}\right). \tag{4.22}$$

Thus $f(1) = \{\alpha\}^n - n\{\alpha\}^{n-1}.\{\beta\}$ and (4.21) translates to $\mathrm{Vol}(\{\alpha-\beta\}) \geq f(s_0)$.

We will now show that f is non-increasing on the interval $[\frac{R-t_0}{n-1},\ 1]$, where we set:

$$R:=\frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}} \quad \text{or equivalently} \quad R=\sup\big\{r>0/\{\alpha\}^n-r\{\alpha\}^{n-1}.\{\beta\}>0\big\}. \tag{4.23}$$

Assumption (4.17) means that $R>n$. Deriving f, we get:

$$\begin{aligned} f'(s) &= -n\,\frac{(t_0-1)^n}{(t_0-s)^n}\,\{\alpha\}^{n-1}.\{\beta\}+n\frac{(t_0-1)^{n-1}}{(t_0-s)^{n-1}}\frac{t_0-1}{(t_0-s)^2}\Big(\{\alpha\}^n-ns\{\alpha\}^{n-1}.\{\beta\}\Big)\\ &= n\,\frac{(t_0-1)^n}{(t_0-s)^{n+1}}\,\Big(\{\alpha\}^n-((n-1)s+t_0)\,\{\alpha\}^{n-1}.\{\beta\}\Big), \qquad s\in[0,\ 1]. \end{aligned}$$

Since $t_0-1>0$ and $t_0-s>0$, the definition of R implies that $f'(s)\le 0$ for all s such that $(n-1)s+t_0\ge R$, *i.e.* for all $s\ge\frac{R-t_0}{n-1}$.

Recall that we are working under the assumption $s_0\in[\frac{R-t_0}{n-1},\ 1)$, so from f being non-increasing on $[\frac{R-t_0}{n-1},\ 1]$ we infer that $f(s_0)\ge f(1)=\{\alpha\}^n-n\{\alpha\}^{n-1}.\{\beta\}$. Since $\mathrm{Vol}(\{\alpha-\beta\})\ge f(s_0)$ by (4.21), we get (4.19). □

4.4. Nef/psef thresholds and volume revisited

We now prove Theorem 1.4. In so doing, we use a different method for obtaining a lower bound for the volume of $\{\alpha-\beta\}$ that takes into account the "angles" between $\{\alpha-s_0\,\beta\}$ and $\{\alpha-t\,\beta\}$ when t varies in a subinterval of $[1,\ t_0)$.

We start with a useful observation in linear algebra generalising inequality (2.1).

Lemma 4.13. *Let $\alpha>0$ and $\beta\ge 0$ be C^∞ $(1,\ 1)$-forms on an arbitrary complex manifold X with $\dim_{\mathbb{C}} X=n$ such that $\alpha-\beta\ge 0$. Then, for every $k\in\{0,1,\dots,n\}$, the following inequality holds:*

$$(\alpha-\beta)^{n-k}\wedge\alpha^k\ge\alpha^n-(n-k)\,\alpha^{n-1}\wedge\beta. \tag{4.24}$$

Proof. Let $x_0\in X$ be any point and $z_1,\dots,z_n$ local holomorphic coordinates about x_0 such that

$$\alpha=\sum_{j=1}^n i\,dz_j\wedge d\bar z_j \quad \text{and} \quad \beta=\sum_{j=1}^n\beta_j\,i\,dz_j\wedge d\bar z_j, \quad \text{hence}$$

$$\alpha-\beta=\sum_{j=1}^n(1-\beta_j)\,i\,dz_j\wedge d\bar z_j \quad \text{at} \quad x_0.$$

Thus $\beta_j\in[0,\ 1]$ for all $j=1,\dots,n$ by our assumptions and inequality (4.24) at x_0 translates to

$$\frac{k!\,(n-k)!}{n!}\sum_{1\le j_1<\dots<j_{n-k}\le n}\big(1-\beta_{j_1}\big)\dots\big(1-\beta_{j_{n-k}}\big)\ge 1-\frac{n-k}{n}\sum_{l=1}^n\beta_l,$$

which, in turn, translates to the following inequality after we set $\gamma_j := 1 - \beta_j \in [0,\ 1]$:

$$\frac{k!\,(n-k)!}{n!}\left(\sum_{1\leq j_1<\cdots<j_k\leq n}\frac{1}{\gamma_{j_1}\ldots\gamma_{j_k}}\right)\gamma_1\ldots\gamma_n \geq \frac{n-k}{n}\sum_{l=1}^{n}\gamma_l+k+1-n. \quad (4.25)$$

Note that the left hand side of (4.25) is meaningful even if some γ_j vanishes because it reappears in $\gamma_1\ldots\gamma_n$. We will prove inequality (4.25) by induction on $n \geq 1$ (where $k \in \{1, \ldots, n\}$ is fixed arbitrarily).

If $n = 1$, (4.25) reads $1 \geq 1$. Although it is not required by the induction procedure, we now prove (4.25) for $n = 3$ and $k = 1$ since this case will be used further down, *i.e.* we prove

$$\frac{1}{3}\,(\gamma_1\gamma_2+\gamma_2\gamma_3+\gamma_3\gamma_1) \geq \frac{2}{3}\,(\gamma_1+\gamma_2+\gamma_3)-1 \quad \text{for all} \quad \gamma_1, \gamma_2, \gamma_3 \in [0,\ 1]. \quad (4.26)$$

It is clear that (4.26) is equivalent to $(\gamma_1 - 1)\,(\gamma_2 - 1) + (\gamma_2 - 1)\,(\gamma_3 - 1) + (\gamma_3 - 1)\,(\gamma_1 - 1) \geq 0$ which clearly holds since $\gamma_j - 1 \leq 0$ for all j.

Now we perform the induction step. Suppose that we have proved (4.25) for all $1 \leq m \leq n$. Proving (4.25) for $n + 1$ amounts to proving the following inequality:

$$\begin{aligned} A_{k,\,n+1} &:= \frac{k!\,(n+1-k)!}{(n+1)!}\sum_{1\leq j_1<\cdots<j_k\leq n+1}\frac{\gamma_1\ldots\gamma_{n+1}}{\gamma_{j_1}\ldots\gamma_{j_k}} \\ &\geq \frac{n+1-k}{n+1}\sum_{l=1}^{n+1}\gamma_l+k-n. \end{aligned} \quad (4.27)$$

The left-hand term $A_{k,\,n+1}$ of (4.27) can be re-written as

$$\frac{k!\,(n+1-k)!}{(n+1)!}\,\frac{1}{n+1-k} \cdot\left(\gamma_1\sum_{\substack{\neq 1\\ 1\leq r_1<\cdots<r_{n-k}\leq n+1}}\gamma_{r_1}\ldots\gamma_{r_{n-k}}+\cdots+\gamma_{n+1}\sum_{\substack{\neq n+1\\ 1\leq r_1<\cdots<r_{n-k}\leq n+1}}\gamma_{r_1}\ldots\gamma_{r_{n-k}}\right),$$

where the meaning of the notation is that the sum whose coefficient is γ_s runs over all the ordered indices $r_1 < \cdots < r_{n-k}$ selected from the set $\{1, \ldots, n+1\} \setminus \{s\}$. Now, using inequality (4.25) for n (the induction hypothesis), for every $s \in \{1, \ldots, n+1\}$ we get:

$$\sum_{\substack{\neq s\\ 1\leq r_1<\cdots<r_{n-k}\leq n+1}}\gamma_{r_1}\ldots\gamma_{r_{n-k}} \geq \frac{n!}{k!(n-k)!}\left(\frac{n-k}{n}\sum_{l\in\{1,\ldots,n+1\}\setminus\{s\}}\gamma_l+k+1-n\right).$$

Plugging these inequalities into the last (re-written) expression for $A_{k,\,n+1}$, we get:

$$A_{k,\,n+1} \geq \frac{n-k}{n(n+1)} \left(\sum_{l\in\{1,\dots,n+1\}\setminus\{1\}} \gamma_1\,\gamma_l + \cdots + \sum_{l\in\{1,\dots,n+1\}\setminus\{n+1\}} \gamma_{n+1}\,\gamma_l \right) + \frac{k+1-n}{n+1}\,(\gamma_1 + \cdots + \gamma_{n+1}),$$

hence

$$\begin{aligned}
A_{k,\,n+1} &\geq \frac{2(n-k)}{n(n+1)} \sum_{1\leq j<k\leq n+1} \gamma_j\,\gamma_k + \frac{k+1-n}{n+1} \sum_{l=1}^{n+1} \gamma_l \\
&= \frac{2(n-k)}{n(n+1)} \frac{\sum_{1\leq j<k<l\leq n+1} (\gamma_j\,\gamma_k + \gamma_k\,\gamma_l + \gamma_l\,\gamma_j)}{n-1} + \frac{k+1-n}{n+1} \sum_{l=1}^{n+1} \gamma_l \\
&\overset{(a)}{\geq} \frac{2(n-k)}{(n-1)n(n+1)} \left(2 \sum_{1\leq j<k<l\leq n+1} (\gamma_j + \gamma_k + \gamma_l) - 3\binom{n+1}{3} \right) \\
&\quad + \frac{k+1-n}{n+1} \sum_{l=1}^{n+1} \gamma_l \\
&= \frac{2(n-k)}{(n-1)n(n+1)} \left(2\binom{n}{2} \sum_{l=1}^{n+1} \gamma_l - 3\binom{n+1}{3} \right) + \frac{k+1-n}{n+1} \sum_{l=1}^{n+1} \gamma_l \\
&= \frac{1}{n+1} \left(\frac{4(n-k)}{n(n-1)} \frac{n(n-1)}{2} + k + 1 - n \right) \sum_{l=1}^{n+1} \gamma_l \\
&\quad - \frac{2(n-k)}{(n-1)n(n+1)} 3 \frac{(n-1)n(n+1)}{2\cdot 3} = \frac{n-k+1}{n+1} \sum_{l=1}^{n+1} \gamma_l - (n-k),
\end{aligned}$$

where inequality (a) above follows from (4.26) applied to each sum $\gamma_j\gamma_k + \gamma_k\gamma_l + \gamma_l\gamma_j$. Thus we have got precisely the inequality (4.27) that we set out to prove. The proof of Lemma 4.13 is complete. □

Now suppose we are in the setting of Conjecture 1.1. We keep the notation of Subsection 4.3. Recall that $s_0 := N^{(\beta)}(\alpha)$ and $t_0 := P^{(\beta)}(\alpha)$. We assume that $s_0 < 1$ (since Conjecture 1.1 has been proved in the case when $s_0 \geq 1$).

We express the class $\{\alpha - \beta\}$ as a convex combination of the nef class $\{\alpha - s_0\,\beta\}$ and the big class $\{\alpha - t\,\beta\}$ for every $t \in [1,\, t_0)$ (*cf.* Theorem 3.5) in the following more flexible version of (4.20):

$$\{\alpha - \beta\} = \frac{t-1}{t-s_0}\,\{\alpha - s_0\beta\} + \frac{1-s_0}{t-s_0}\,\{\alpha - t\beta\}, \quad t \in [1,\, t_0). \tag{4.28}$$

We know from Theorem 3.5 that for every $t < \frac{R}{n}$ (*cf.* notation (4.23)) there exists a Kähler current T_t in the class $\{\alpha - t\beta\}$ such that $T_t \geq (1 - \frac{n}{R}\,t)\,\alpha$. Thus we get the following Kähler current in the class $\{\alpha - \beta\}$:

$$\begin{aligned} S_t &:= \frac{t-1}{t-s_0}\,(\alpha - s_0\beta) + \frac{1-s_0}{t-s_0}\,T_t \\ &\geq \frac{t-1}{t-s_0}\,(\alpha - s_0\beta) + \frac{1-s_0}{t-s_0}\,\left(1 - \frac{n}{R}\,t\right)\,\alpha, \quad t \in \left[1,\ \frac{R}{n}\right], \end{aligned} \tag{4.29}$$

since the class $\{\alpha - s_0\,\beta\}$ being nef allows us to assume without loss of generality that $\alpha - s_0\beta \geq 0$ (after possibly adding $\varepsilon\omega$ and letting $\varepsilon \downarrow 0$ in the end). Since the right-hand side of (4.29) is smooth, it also provides a lower bound for the absolutely continuous part of S_t, so we get the following lower bound for the volume for all $t \in [1,\ \frac{R}{n}]$:

$$\begin{aligned} &\mathrm{Vol}(\{\alpha - \beta\}) \geq \int_X S^n_{t,\,ac} \\ &\geq \frac{1}{(t-s_0)^n}\sum_{k=0}^{n}\binom{n}{k}(t-1)^{n-k}(1-s_0)^k\left(1-\frac{nt}{R}\right)^k\int_X(\alpha - s_0\beta)^{n-k}\wedge\alpha^k. \end{aligned} \tag{4.30}$$

Since the class $\{\alpha - s_0\beta\}$ is nef, using Lemma 4.13, we get the following:

Lemma 4.14. *Let X be a compact Kähler manifold,* $\dim_{\mathbb{C}} X = n$*, and let* $\{\alpha\}, \{\beta\} \in H^{1,1}_{BC}(X,\ \mathbb{R})$ *be Kähler classes such that* $\{\alpha\}^n - n\{\alpha\}^{n-1}.\{\beta\} > 0$. *Suppose that* $s_0 := N^{(\beta)}(\alpha) < 1$. *Then the following estimate holds:*

$$\begin{aligned} \mathrm{Vol}(\{\alpha - \beta\}) \geq \left(\frac{A\,t - s_0}{t - s_0}\right)^n \left(\{\alpha\}^n - \frac{s_0(t-1)}{A\,t - s_0}\,n\,\{\alpha\}^{n-1}.\,\{\beta\}\right), \\ \textit{for all}\quad t \in \left[1,\ \frac{R}{n}\right], \end{aligned} \tag{4.31}$$

where we denote $R := \{\alpha\}^n/\{\alpha\}^{n-1}.\,\{\beta\} > n$ *and* $A := 1 - \frac{n}{R}\,(1 - s_0) \in (s_0,\ 1)$.

Proof. From $A - s_0 = (1 - s_0)\,(1 - \frac{n}{R}) \in (0,\ 1)$ (because $s_0 \in (0,\ 1)$ and $1 - \frac{n}{R} \in (0,\ 1)$), we infer that $A > s_0$. That $A < 1$, is obvious.

Without loss of generality, we may assume that $\alpha - s_0\beta \geq 0$, so (4.24) applies to α and $s_0\beta$ and from (4.30) we get:

$$\begin{aligned} &\mathrm{Vol}(\{\alpha - \beta\}) \\ &\geq \frac{1}{(t-s_0)^n}\sum_{k=0}^{n}\binom{n}{k}(t-1)^{n-k}(1-s_0)^k\left(1-\frac{nt}{R}\right)^k\Big(\{\alpha\}^n - (n-k)s_0\{\alpha\}^{n-1}.\{\beta\}\Big) \\ &= \frac{1}{(t-s_0)^n}\left[t - 1 + (1-s_0)\left(1 - \frac{nt}{R}\right)\right]^n\{\alpha\}^n \\ &\quad - \frac{t-1}{(t-s_0)^n}\left[t - 1 + (1-s_0)\left(1 - \frac{nt}{R}\right)\right]^{n-1} n s_0\,\{\alpha\}^{n-1}.\,\{\beta\}, \end{aligned}$$

which proves (4.31) since $t - 1 + (1 - s_0)\,(1 - \frac{nt}{R}) = At - s_0$. $\square$

Thus, it becomes necessary to study the variation of the following function:

$$g:\left[1,\,\frac{R}{n}\right]\to\mathbb{R},\quad g(t):=\left(\frac{At-s_0}{t-s_0}\right)^n\left(\{\alpha\}^n-\frac{s_0(t-1)}{At-s_0}n\{\alpha\}^{n-1}.\{\beta\}\right),\quad (4.32)$$

since $\mathrm{Vol}(\{\alpha-\beta\})\geq g(t)$ for all $t\in[1,\,\frac{R}{n}]$. From $A-s_0=(1-s_0)\,(1-\frac{n}{R})\in(0,\,1)$, we get:

$$\begin{aligned} g(1) &= \left(1-\frac{n}{R}\right)^n\,\{\alpha\}^n,\\ \text{while}\quad g\left(\frac{R}{n}\right) &= \left(\frac{R-n}{R-ns_0}\right)^n\,\left(\{\alpha\}^n-ns_0\,\{\alpha\}^{n-1}.\{\beta\}\right). \end{aligned} \tag{4.33}$$

We see that $g(1)$ is precisely the lower bound obtained for the volume of $\{\alpha-\beta\}$ in (3.6), so this lower bound will be improved if $g(t)>g(1)$ for some $t\in(1,\,R/n]$.

Variation of g. Since $[(t-1)/(At-s_0)]'=(A-s_0)/(At-s_0)^2$ and $[(At-s_0)/(t-s_0)]'=(1-A)\,s_0/(t-s_0)^2$, for the derivative of $g(t)$ we get:

$$\begin{aligned} g'(t) &= n(1-A)\,s_0\,\frac{(At-s_0)^{n-1}}{(t-s_0)^{n+1}}\\ &\quad\times\left(\{\alpha\}^n-\frac{(ns_0-ns_0A+A-s_0)t-ns_0(1-A)-s_0\,(A-s_0)}{(1-A)\,(At-s_0)}\,\{\alpha\}^{n-1}.\{\beta\}\right). \end{aligned}$$

Now, $At-s_0>0$ for all $t\in[1,\,R/n]$ since $At-s_0\geq A-s_0=(1-s_0)(1-\frac{n}{R})>0$. Since $t\geq 1>s_0$, from the definition (4.23) of R, we get the equivalences:

$$\begin{aligned} g'(t)\geq 0 \iff & [ns_0\,(1-A)+A-s_0]\,t-ns_0\,(1-A)-s_0\,(A-s_0)\\ &\leq(1-A)(At-s_0)R \iff -\left[RA^2-(ns_0-1+R)\,A+(n-1)\,s_0\right]t\\ &+s_0\left[A-s_0+(n-R)\,(1-A)\right]\geq 0. \end{aligned} \tag{4.34}$$

- *Sign of* $RA^2-(ns_0-1+R)\,A+(n-1)\,s_0$. The discriminant of this 2^{nd} degree polynomial in A is

$$\Delta_R=R^2-2\,((n-2)\,s_0+1)\;R+(ns_0-1)^2. \tag{4.35}$$

The discriminant of Δ_R (viewed as a polynomial in R) is

$$\Delta'=16(n-1)\,s_0(1-s_0)>0\quad\text{since}\quad s_0\in(0,\,1). \tag{4.36}$$

Thus, the Δ_R vanishes at $R_1=(n-2)s_0+1-2\sqrt{(n-1)s_0(1-s_0)}$ and $R_2=(n-2)s_0+1+2\sqrt{(n-1)s_0(1-s_0)}$.

Lemma 4.15. *With our usual notation $R:=\{\alpha\}^n/\{\alpha\}^{n-1}.\{\beta\}$, we have: $R_1<R_2\leq n<R$.*

Proof. Only the inequality $R_2 \leq n$ needs a proof. It is equivalent to

$$\begin{aligned}&(n-2)s_0 + 2\sqrt{(n-1)s_0(1-s_0)} \leq n-1\\ \iff\ &2\sqrt{(n-1)s_0(1-s_0)} \leq (n-1)(1-s_0)+s_0\\ \iff\ &(\sqrt{(n-1)\,(1-s_0)} - \sqrt{s_0})^2 \geq 0,\end{aligned}$$

which clearly holds. □

The upshot is that $\Delta_R > 0$, so $RA^2 - (ns_0 - 1 + R)\,A + (n-1)\,s_0$ vanishes at $A_1 = (ns_0 - 1 + R - \sqrt{\Delta_R})/2R$ and $A_2 = (ns_0 - 1 + R + \sqrt{\Delta_R})/2R$.

Lemma 4.16. *With our notation* $A := 1 - \frac{n}{R}\,(1-s_0) \in [0,\,1)$, *we have:* $A_1 < A < A_2$.

Proof. The inequality $A_1 < A$ is equivalent to

$$\frac{ns_0 - 1 + R - \sqrt{\Delta_R}}{2R} < \frac{R - n + ns_0}{R} \iff n(2-s_0) - 1 - R < \sqrt{\Delta_R}. \quad (4.37)$$

If $n(2-s_0) - 1 - R \leq 0$, (4.37) is obvious. If $n(2-s_0) - 1 - R > 0$, inequality (4.37) is equivalent to

$$\begin{aligned}&R^2 + [n(2-s_0)-1]^2 - 2[n(2-s_0)-1]R < R^2 - 2[(n-2)s_0+1]R + (ns_0-1)^2\\ &\iff [n(2-s_0) - ns_0]\,[n(2-s_0) + ns_0 + 2]\\ &< 2[n(2-s_0) - (n-2)s_0 - 2]\,Rn(n-1)(1-s_0)\\ &< (1-s_0)(n-1)R \iff n < R,\end{aligned}$$

where the last inequality holds thanks to our assumption $\{\alpha\}^n - n\{\alpha\}^{n-1}.\{\beta\} > 0$.

The inequality $A < A_2$ is equivalent to

$$\frac{R - n + ns_0}{R} < \frac{ns_0 - 1 + R + \sqrt{\Delta_R}}{2R} \iff R + 1 - (2-s_0)n < \sqrt{\Delta_R}. \quad (4.38)$$

If $R + 1 - (2-s_0)n \leq 0$, (4.38) is obvious. If $R + 1 - (2-s_0)n > 0$, inequality (4.38) is equivalent to

$$\begin{aligned}&R^2 + [1-(2-s_0)n]^2 + 2[1-(2-s_0)n]R < R^2 - 2\,[(n-2)s_0+1]\,R + (ns_0-1)^2\\ &\iff 2[2 - (2-s_0)n + (n-2)s_0]\,R < [ns_0 - (2-s_0)n]\,[ns_0 - 2 + (2-s_0)n]\\ &\iff (n-1)\,(s_0-1)\,R < n(n-1)(s_0-1) \iff R > n \quad \text{since} \quad s_0 - 1 < 0,\end{aligned}$$

where the last inequality holds thanks to our assumption $\{\alpha\}^n – n\{\alpha\}^{n-1}.\{\beta\} > 0$. □

The obvious corollary of Lemma 4.16 is the following inequality:

$$RA^2 - (ns_0 - 1 + R)\,A + (n-1)\,s_0 < 0. \quad (4.39)$$

- *Monotonicity of* $g\ :\ [1,\ \frac{R}{n}] \to \mathbb{R}$. Picking up where we left off in (4.34), we get the equivalence:

$$g'(t) \geq 0 \iff t \geq s_0\,\frac{A - s_0 + (n-R)\,(1-A)}{RA^2 - (ns_0 - 1 + R)\,A + (n-1)\,s_0}. \quad (4.40)$$

Lemma 4.17. *The following inequalities hold:*

$$
\begin{aligned}
&\text{(a)} \quad \mathrm{A - s_0 + (n - R)\,(1 - A) < 0;} \\
&\text{(b)} \quad 1 > \mathrm{s_0}\, \frac{\mathrm{A - s_0 + (n - R)\,(1 - A)}}{\mathrm{RA^2 - (ns_0 - 1 + R)\,A + (n - 1)\,s_0}}.
\end{aligned}
\tag{4.41}
$$

Proof. (a) We have:

$$
\begin{aligned}
A - s_0 + (n - R)\,(1 - A) &= (1 - s_0)\,(1 - \frac{n}{R}) + \frac{n}{R}\,(1 - s_0)\,(n - R) \\
&= \frac{(1 - s_0)(n - R)(n - 1)}{R}
\end{aligned}
$$

and the last expression is negative since $n - R < 0$ while $1 - s_0 > 0$ and $n - 1 > 0$.

(b) Thanks to (4.39), inequality (b) in (4.41) is equivalent to

$$
\begin{aligned}
&RA^2 - (ns_0 - 1 + R)A + (n - 1)s_0 < s_0[A - s_0 + (n - R)(1 - A)] \\
\iff\ & RA^2 - (Rs_0 + R + s_0 - 1)\,A + s_0\,(s_0 + R - 1) < 0.
\end{aligned}
\tag{4.42}
$$

The discriminant of the left hand side in (4.42), viewed as a 2^{nd} degree polynomial in A, is $\Delta'' = (R - 1)^2(1 - s_0)^2$, so the left hand side of (4.42) vanishes at

$$
A_3 = \frac{R(s_0 + 1) + s_0 - 1 - (R - 1)(1 - s_0)}{2R} = s_0
$$

$$
\text{and} \quad A_4 = 1 - \frac{1 - s_0}{R}, \quad \text{where clearly} \quad A_3 < A_4.
$$

Thus, inequality (4.42) is equivalent to $s_0 < A < 1 - \frac{1-s_0}{R}$. We have seen in Lemma 4.14 that $A > s_0$. On the other hand, proving $A < 1 - \frac{1-s_0}{R}$ amounts to proving

$$
1 - \frac{n}{R}\,(1 - s_0) < 1 - \frac{1 - s_0}{R} \iff 1 < n \quad (\text{since} \quad 1 - s_0 > 0 \quad \text{and} \quad R > 0).
$$

The last inequality being obvious, the proof of (b) in (4.41) is complete. □

Conclusion 4.18. Inequality (4.40) holds strictly for every $t \geq 1$ thanks to part (b) of (4.41). So, in particular, $g'(t) > 0$ for all $t \in [1, \frac{R}{n}]$, *i.e.* the function $g : [1, \frac{R}{n}] \to \mathbb{R}$ is increasing.

Since $\mathrm{Vol}(\{\alpha - \beta\}) \geq g(t)$ for all $t \in [1, \frac{R}{n}]$ (*cf.* Lemma 4.14), the best lower bound for $\mathrm{Vol}(\{\alpha - \beta\})$ that we get through this method in the case when $s_0 := N^{(\beta)}(\alpha) < 1$ is

$$
\begin{aligned}
\mathrm{Vol}(\{\alpha - \beta\}) &\geq g\left(\frac{R}{n}\right) \\
&= (\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\})\,\left(\frac{\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\,\{\beta\}}{\{\alpha\}^n - ns_0\,\{\alpha\}^{n-1}.\,\{\beta\}}\right)^{n-1}.
\end{aligned}
\tag{4.43}
$$

This proves Theorem 1.4. Note that this lower bound for the volume improves on the lower bound $g(1)$ (*cf.* (4.33)) obtained in (3.6).

5. Intersection numbers

In this section we prove Theorem 1.3. We start by deriving analogues in bidegree $(p,\, p)$ with $p \geq 2$ of the inequalities established in Section 3. We will use the standard notion of positivity for $(q,\, q)$-forms whose definition is recalled at the beginning of the Appendix before Lemma 7.1.

Proposition 5.1. *Let X be a compact Kähler manifold with $\dim_{\mathbb{C}} X = n$ and let α, β be Kähler metrics on X. Then, for every $t \in [0, +\infty)$, every $p \in \{1, \ldots, n\}$ and every C^∞ positive $(n-p,\, n-p)$-form $\Omega^{n-p,\,n-p} \geq 0$ on X such that $\partial\bar\partial\Omega^{n-p,\,n-p} = 0$, we have:*

$$\int_X \big(\alpha^p - tp\,\alpha^{p-1}\wedge\beta\big)\wedge\Omega^{n-p,\,n-p} \geq \Big(1 - t\,\frac{n}{R}\Big)\int_X \alpha^p\wedge\Omega^{n-p,\,n-p}, \quad (5.1)$$

where, as usual, we let $R := \frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}}$. We also have:

$$\int_X \big(\alpha^p - t^p\beta^p\big)\wedge\Omega^{n-p,\,n-p} \geq \left(1 - t^p\,\frac{\binom{n}{p}}{R_p}\right)\int_X \alpha^p\wedge\Omega^{n-p,\,n-p}, \qquad (5.2)$$

where we let $R_p := \frac{\{\alpha\}^n}{\{\alpha\}^{n-p}.\{\beta\}^p}$.

Proof. We may and will assume without loss of generality that $\Omega^{n-p,\,n-p}$ is strictly positive. Inequality (5.1) is equivalent to

$$t\,\frac{n}{R}\int_X \alpha^p\wedge\Omega^{n-p,\,n-p} \geq tp\int_X \alpha^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p},$$

which, in turn, after the simplification of $t \geq 0$ and the unravelling of R, is equivalent to

$$\frac{n}{p}\bigg(\int_X \alpha^p\wedge\Omega^{n-p,\,n-p}\bigg)\cdot\bigg(\int_X \alpha^{n-1}\wedge\beta\bigg) \geq \{\alpha\}^n\int_X \alpha^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}. \quad (5.3)$$

This inequality can be proved using the method in the proof of Lemma 3.1, the pointwise inequality (7.5) proved in the Appendix and an *approximate fixed point technique* that we now describe. Here are the details.

Approximate fixed point technique

We consider the following Monge-Ampère equation whose unique C^∞ solution in the Kähler class $\{\alpha\}$ is denoted by $\widetilde{\alpha} := \alpha + i\partial\bar\partial\varphi > 0$:

$$\widetilde{\alpha}^n = \frac{\{\alpha\}^n}{\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,\,n-p}]_A}\,\alpha^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}. \tag{5.4}$$

By $\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,\,n-p}]_A$ we mean the positive real number $\int_X \alpha^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}$ which clearly depends only on the Bott-Chern classes $\{\alpha\}, \{\beta\} \in H^{1,\,1}(X,\,\mathbb{R})$ and on the Aeppli class $[\Omega^{n-p,\,n-p}]_A \in H_A^{n-p,\,n-p}(X,\,\mathbb{R})$.

We will vary the form α on the right-hand side of (5.4) in its Kähler class $\{\alpha\}$. Let $\mathcal{E}_\alpha := \{T \in \{\alpha\}\ /\ T \geq 0\}$ be the set of d-closed positive $(1,\,1)$-currents in the Kähler class $\{\alpha\}$. Thus $\mathcal{E}_\alpha$ is a *compact* convex subset of the locally convex space $\mathcal{D}'^{1,\,1}(X,\,\mathbb{R})$ endowed with the weak topology of currents. (The compactness is a consequence of the existence of Gauduchon metrics and holds for any psef class $\{\alpha\}$ even if X is not Kähler.) Fix an arbitrary Kähler metric ω in $\{\alpha\}$. For every $\varepsilon > 0$, we associate with equation (5.4) the map:

$$R_\varepsilon : \mathcal{E}_\alpha \to \mathcal{E}_\alpha, \quad R_\varepsilon(T) = \alpha_{T,\,\varepsilon}, \tag{5.5}$$

defined in three steps as follows. Let $T \in \mathcal{E}_\alpha$ be arbitrary.

(i) By the Blocki-Kolodziej version [4] for *Kähler* classes of Demailly's regularisation-of-currents theorem [11, Theorem 1.1], there exist C^∞ d-closed $(1,\,1)$-forms $\omega_\varepsilon \in \{\alpha\} = \{T\}$ for $\varepsilon > 0$ such that $\omega_\varepsilon \geq -\varepsilon\omega$ and $\omega_\varepsilon \to T$ in the weak topology of currents as $\varepsilon \to 0$. (The Kähler assumption on the class $\{\alpha\}$ crucially ensures that the possible negative part of ω_ε does not exceed $\varepsilon\omega$, see [4].)

Note that for every sequence of currents $T_j \in \mathcal{E}_\alpha$ converging weakly to a current $T \in \mathcal{E}_\alpha$ and for every fixed $\varepsilon > 0$, the sequence of C^∞ forms $(\omega_{j,\,\varepsilon})_j$ (obtained by applying to each T_j the Blocki-Kolodziej regularisation procedure just described producing a family $\omega_{j,\,\varepsilon} \to T_j$ as $\varepsilon \to 0$) converges in the C^∞ topology to the C^∞ form ω_ε (obtained by applying to T the Blocki-Kolodziej regularisation procedure producing a family $\omega_\varepsilon \to T$ as $\varepsilon \to 0$). In other words, for every fixed $\varepsilon > 0$, the map $\mathcal{E}_\alpha \ni T \mapsto \omega_\varepsilon \in C^\infty_{1,\,1}(X,\,\mathbb{C})$ is *continuous* if $\mathcal{E}_\alpha$ has been equipped with the weak topology of currents and the space of smooth $(1,\,1)$-forms has been given the C^∞ topology.

To see this, it suffices to work locally with currents $T_j = i\partial\bar\partial\psi_j \geq 0$ and $T = i\partial\bar\partial\psi \geq 0$ for which the psh potentials have the property that $\psi_j \longrightarrow \psi$ in the L^1 topology as $j \to +\infty$, and to show that for every fixed $\varepsilon > 0$ we have $i\partial\bar\partial\psi_j \star \rho_\varepsilon \longrightarrow i\partial\bar\partial\psi \star \rho_\varepsilon$ in the C^0 topology as $j \to +\infty$. (The convergence in the C^∞ topology follows from this by taking derivatives.) Indeed, currents are regularised in [4] by convolution of their local potentials with regularising kernels ρ_ε. Since $i\partial\bar\partial\psi_j \star \rho_\varepsilon = \psi_j \star i\partial\bar\partial\rho_\varepsilon$ and $i\partial\bar\partial\psi \star \rho_\varepsilon = \psi \star i\partial\bar\partial\rho_\varepsilon$, we have to ensure, for every fixed $\varepsilon > 0$, that

$$\int_{U'} (\psi_j - \psi)(y),\, u_\varepsilon(x-y) \underset{j\to+\infty}{\longrightarrow} 0 \ \text{ locally uniformly with respect to } x \in U' \Subset U$$

for every C^∞ function u_ε (which is an arbitrary coefficient of $i\partial\bar\partial\rho_\varepsilon$ in this case) defined on the open subset $U \subset X$ on which we work. This is clear from the L^1 convergence $\psi_j \longrightarrow \psi$ on U.

(ii) Set $u_{T,\,\varepsilon} := (1-\varepsilon)\,\omega_\varepsilon + \varepsilon\omega$. Thus $u_{T,\,\varepsilon}$ is a Kähler metric in the class $\{\alpha\}$ since it is C^∞ and $u_{T,\,\varepsilon} \geq -(1-\varepsilon)\,\varepsilon\,\omega + \varepsilon\,\omega = \varepsilon^2\,\omega > 0$. Moreover, $u_{T,\,\varepsilon} \to T$ in the weak topology of currents as $\varepsilon \to 0$.

(iii) Solve equation (5.4) with right-hand term defined by $u_{T,\,\varepsilon}$ instead of α:

$$\alpha_{T,\,\varepsilon}^n = \frac{\{\alpha\}^n}{\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,\,n-p}]_A}\, u_{T,\,\varepsilon}^{p-1} \wedge \beta \wedge \Omega^{n-p,\,n-p}. \tag{5.6}$$

This means that we denote by $\alpha_{T,\,\varepsilon}$ the unique Kähler metric in the Kähler class $\{\alpha\}$ solving equation (5.6) whose existence is ensured by Yau's theorem [26]. We put $R_\varepsilon(T) := \alpha_{T,\,\varepsilon}$. Thus, in particular, the image of R_ε consists of (smooth) Kähler metrics in $\{\alpha\}$.
Now, R_ε is a continuous self-map of the compact convex subset $\mathcal{E}_\alpha$ of the locally convex space $\mathcal{D}'^{1,\,1}(X,\,\mathbb{R})$, so by the Schauder fixed point theorem, there exists a current $T_\varepsilon \in \mathcal{E}_\alpha$ such that $T_\varepsilon = R_\varepsilon(T_\varepsilon) = \alpha_{T_\varepsilon,\,\varepsilon}$. Since $\alpha_{T_\varepsilon,\,\varepsilon} := \widetilde{\alpha}_\varepsilon$ is C^∞, by construction, the fixed-point current T_ε must be a C^∞ form, so $T_\varepsilon = \widetilde{\alpha}_\varepsilon$ and $\omega_\varepsilon \geq \widetilde{\alpha}_\varepsilon - \delta_\varepsilon\omega$ for some $\delta_\varepsilon \downarrow 0$ when $\varepsilon \to 0$. (The last statement follows from the fact that ω_ε converges in the C^∞ topology to T if T is C^∞ – see the explanations under (3) below.) Hence $u_{T_\varepsilon,\,\varepsilon} = (1-\varepsilon)\,\omega_\varepsilon + \varepsilon\omega \geq (1-\varepsilon)\,\widetilde{\alpha}_\varepsilon + [\varepsilon - (1-\varepsilon)\,\delta_\varepsilon]\,\omega$. We put $\eta_\varepsilon := \varepsilon - (1-\varepsilon)\,\delta_\varepsilon$, so $\eta_\varepsilon \to 0$ when $\varepsilon \to 0$.

To conclude, for every $\varepsilon > 0$, we have got a Kähler metric $\widetilde{\alpha}_\varepsilon$ in the Kähler class $\{\alpha\}$ such that

$$\begin{aligned}\widetilde{\alpha}_\varepsilon^n &= \frac{\{\alpha\}^n}{\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,n-p}]_A}[(1-\varepsilon)\omega_\varepsilon + \varepsilon\omega]^{p-1} \wedge \beta \wedge \Omega^{n-p,\,n-p}\\ &\geq (1-\varepsilon)^{p-1}\frac{\{\alpha\}^n}{\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,n-p}]_A}\widetilde{\alpha}_\varepsilon^{p-1}\wedge\beta\wedge\Omega^{n-p,n-p} - O(|\eta_\varepsilon|),\end{aligned} \tag{5.7}$$

where ω is an arbitrary, fixed Kähler metric in the class $\{\alpha\}$ and $O(|\eta_\varepsilon|)$ is a quantity that converges to zero as $\varepsilon \to 0$. The Kähler metric $\widetilde{\alpha}_\varepsilon$ can be viewed as an *approximate fixed point* in the class $\{\alpha\}$ of equation (5.4).

Explanations. Here are a few additional comments on the choice of a *continuous* regularising operator R_ε for every $\varepsilon > 0$. We are indebted to J.-P. Demailly and to A. Zeriahi for many of the ensuing remarks that were left out of the first version of this paper.

(1) The existence of a *continuous* regularising operator is an easy consequence of the regularisation theorem (whatever version of it may be used, be it Demailly's regularisation of currents [11, Theorem 1.1] or the Blocki-Kolodziej one [4] or any other one) applied to *finitely many* currents. The argument for this statement,

which has been very kindly and effectively explained to the author by J.-P. Demailly, makes use of the compactness and convexity of $\mathcal{E}_\alpha$ in $\mathcal{D}'^{1,1}(X, \mathbb{R})$. Indeed, the cone of positive currents has a compact and metrisable, hence countable, base. For this reason, there are several different topologies that induce the same topology on this cone (= the weak topology of currents). If we fix a smooth representative α of the class $\{\alpha\}$ (and α can be chosen to be a Kähler metric in this case, but this is irrelevant here), any current $T \in \mathcal{E}_\alpha$ can be written as $T = \alpha + i\partial\bar\partial\varphi \geq 0$ for a global quasi-psh (hence L^2 and indeed L^p for every $p \in [1, +\infty)$) potential φ that is unique up to a constant. We can equip the space of potentials $\{\varphi \ / \ i\partial\bar\partial\varphi \geq -\alpha\}$ with the topology induced by the L^2 Hilbert space topology, which is separable, hence has a countable orthonormal base. This topology induces on $\mathcal{E}_\alpha$ the weak topology of currents.

Now, by compactness of $\mathcal{E}_\alpha$, for every ε, there is a *finite* covering of $\mathcal{E}_\alpha$ by open balls of radius ε. Let $T_1, \dots, T_{N_\varepsilon} \in \mathcal{E}_\alpha$ be the centres of these balls and let $\mathcal{E}_{\alpha,\,\varepsilon}$ be the convex polyhedron generated by $T_1, \dots, T_{N_\varepsilon}$. We can take $\varepsilon = 1/m$ for $m \in \mathbb{N}^\star$ and by convexity of $\mathcal{E}_\alpha$ we get

$$\mathcal{E}_\alpha = \bigcup_{m=1}^{+\infty} \mathcal{E}_{\alpha,\,\frac{1}{m}}.$$

Thus, it suffices to regularise the finitely many currents $T_1, \dots, T_{N_\varepsilon}$ and to extend the regularisation to all the currents $T \in \mathcal{E}_{\alpha,\,\varepsilon}$ by mere convex combinations. This clearly produces a *continuous* regularising operator.

(2) The main result of [4], namely that in a Kähler class positive currents can be regularised with only an $O(\varepsilon)$ loss of positivity (so, ultimately, with no loss at all, as explained above – hence the Kähler metrics in a given Kähler class are dense in the positive currents of that class) can also be obtained as an easy consequence of Demailly's regularisation theorem [11]. The argument for this statement, which was very kindly explained to the author by A. Zeriahi, proceeds by first regularising by a mere cut-off operation. Indeed, let $T = \alpha + i\partial\bar\partial\varphi \geq 0$ be an arbitrary positive current in the Kähler class $\{\alpha\}$, where $\alpha > 0$ is a Kähler metric in this class. For every $\varepsilon > 0$, put $T_\varepsilon := \alpha + i\partial\bar\partial \max(\varphi, -\frac{1}{\varepsilon}) \geq 0$. The current T_ε is still positive since the maximum of any two α-psh functions is still α-psh when α is a Kähler metric ([17, Proposition 2.3, (4)]). We have $\max(\varphi, -\frac{1}{\varepsilon}) \downarrow \varphi$ pointwise and $T_\varepsilon \to T$ weakly as $\varepsilon \to 0$. Moreover, the currents T_ε have *bounded* potentials, so we can apply Demailly's regularisation theorem [11] to each of them to write T_ε as the weak limit of a sequence of C^∞ forms $T_{\varepsilon,\,\delta} \in \{\alpha\}$ as $\delta \to 0$. Since all the Lelong numbers of T_ε vanish (because the potential is bounded), Demailly's theorem [11] ensures that only a loss of positivity of $O(\delta)$ is introduced by the regularisation process. Taking the diagonal sequence with $\varepsilon = \delta$, we get an approximation of the original current T by C^∞ forms in its class with only an $O(\varepsilon)$ loss of positivity.

The interest in the Blocki-Kolodziej regularisation procedure [4] lies in its giving a much simpler proof of the existence of a good regularisation of currents (which is by no means unique) for the special case of a Kähler class than Demailly's proof of the general case.

(3) The Blocki-Kolodziej regularisation [4] proceeds by convolution of the local potentials of the current T with regularising kernels ρ_ε. This method produces a *continuous* regularising operator R_ε for every ε. Moreover, if T is a C^∞ form in the class $\{\alpha\}$, the C^∞ forms T_ε obtained by regularising T converge to T in the C^∞ topology as $\varepsilon \to 0$. This is because locally, if ψ is a C^∞ function defined on an open subset $\Omega \subset \mathbb{C}^n$ containing the origin, then all the derivatives of the convolutions $\rho_\varepsilon \star \psi$ converge uniformly on all the compact subsets $K \subset \Omega$ to the corresponding derivatives of ψ and the (standard) patching procedure used in [4] does not destroy this property. On the other hand, Yau's theorem [26] gives uniform estimates in all the C^l norms of the solution of the Monge-Ampère equation in terms of the right-hand side term of this equation. Putting these facts together, we get that the regularising operator R_ε obtained by regularisation followed by an application of Yau's theorem is indeed *continuous* in the weak topology of currents and, moreover, $R_\varepsilon(T)$ converges in the C^∞ topology to T whenever $T \in \mathcal{E}_\alpha$ is C^∞.

Use of the approximate fixed point

Let us fix any smooth volume form $dV > 0$ on X. The left hand side term in (5.3) reads:

$$
\begin{aligned}
&\frac{n}{p}\left(\int_X \alpha^p \wedge \Omega^{n-p,\,n-p}\right)\cdot\left(\int_X \alpha^{n-1}\wedge\beta\right)\\
&= \frac{n}{p}\left(\int_X \frac{\widetilde{\alpha}_\varepsilon^p \wedge \Omega^{n-p,\,n-p}}{dV}\,dV\right)\cdot\left(\int_X \frac{\widetilde{\alpha}_\varepsilon^{n-1}\wedge\beta}{\widetilde{\alpha}_\varepsilon^n}\,\frac{\widetilde{\alpha}_\varepsilon^n}{dV}\,dV\right)\\
&\overset{(a)}{\geq} \left[\int_X \left(\frac{n}{p}\,\frac{\widetilde{\alpha}_\varepsilon^p \wedge \Omega^{n-p,\,n-p}}{dV}\,\frac{\widetilde{\alpha}_\varepsilon^{n-1}\wedge\beta}{\widetilde{\alpha}_\varepsilon^n}\right)^{\frac12}\left(\frac{\widetilde{\alpha}_\varepsilon^n}{dV}\right)^{\frac12} dV\right]^2\\
&\overset{(b)}{\geq} \left[\int_X \left(\frac{\widetilde{\alpha}_\varepsilon^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}}{dV}\right)^{\frac12}\left(\frac{\widetilde{\alpha}_\varepsilon^n}{dV}\right)^{\frac12} dV\right]^2\\
&\overset{(c)}{\geq} (1-\varepsilon)^{p-1}\,\frac{\{\alpha\}^n}{\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,\,n-p}]_A}\\
&\quad\cdot\left[\int_X \left(\frac{\widetilde{\alpha}_\varepsilon^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}}{dV}\right)^{\frac12}\left(\frac{\widetilde{\alpha}_\varepsilon^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}}{dV}\right)^{\frac12} dV\right]^2 - O(|\eta_\varepsilon|)\\
&\overset{(d)}{=} (1-\varepsilon)^{p-1}\,\frac{\{\alpha\}^n}{\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,\,n-p}]_A}\left[\int_X \alpha^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}\right]^2 - O(|\eta_\varepsilon|)\\
&= (1-\varepsilon)^{p-1}\,\{\alpha\}^n\int_X \alpha^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p} - O(|\eta_\varepsilon|),
\end{aligned}
$$

for every $\varepsilon > 0$. Letting $\varepsilon \to 0$, we get the desired inequality (5.3) since $\eta_\varepsilon \to 0$. Inequality (a) was an application of the Cauchy-Schwarz inequality, (b) was an application of the pointwise inequality (7.5) of Lemma 7.1 in the Appendix, (c) followed from (5.7), while identity (d) followed from $\widetilde{\alpha}_\varepsilon$ belonging to the class $\{\alpha\}$. The proof of (5.1), which is equivalent to (5.3), is complete.

The proof of (5.2) runs along the same lines. Indeed, (5.2) is equivalent to

$$t^p\,\frac{\binom{n}{p}}{R_p}\int_X \alpha^p\wedge\Omega^{n-p,\,n-p}\geq t^p\int_X \beta^p\wedge\Omega^{n-p,\,n-p},$$

which, in turn, after the simplification of $t^p\geq 0$ and the unravelling of R_p, is equivalent to

$$\binom{n}{p}\left(\int_X \alpha^p\wedge\Omega^{n-p,\,n-p}\right)\cdot\left(\int_X \alpha^{n-p}\wedge\beta^p\right)\geq\{\alpha\}^n\int_X \beta^p\wedge\Omega^{n-p,\,n-p}. \quad (5.8)$$

The proof of (5.8) is almost identical to that of (5.3) spelt out above except for the replacement of equation (5.4) with the following Monge-Ampère equation:

$$\widetilde{\alpha}^n=\frac{\{\alpha\}^n}{\{\beta\}^p.[\Omega^{n-p,\,n-p}]_A}\,\beta^p\wedge\Omega^{n-p,\,n-p}, \quad (5.9)$$

and for the replacement of the pointwise inequality (7.5) with (7.4). Note that, since α does not feature on the right-hand side of equation (5.9), the *approximate fixed point technique* is no longer necessary in this case. It suffices to work with the unique Kähler-metric solution $\widetilde{\alpha}$ of (5.9). □

Remark 5.2. If an exact (rather than an approximate) fixed point for equation (5.4) had been sought, we would have needed to consider the following equation in which the Kähler-metric solution $\widetilde{\alpha}\in\{\alpha\}$ features on both sides:

$$\alpha^n=\frac{\{\alpha\}^n}{\{\alpha\}^{p-1}.\{\beta\}.[\Omega^{n-p,\,n-p}]_A}\,\widetilde{\alpha}^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}.$$

Equations of this type, going back to Donaldson's J-flow and to work by Chen, admit a solution under a certain assumption on the class $\{\alpha\}$. See [15] and the references therein for details. Our *approximate fixed point technique* does not require any particular assumption on $\{\alpha\}$.

We can now prove the main result of this section which subsumes Theorem 1.3.

Theorem 5.3. *Let X be a compact Kähler manifold with* $\dim_{\mathbb{C}} X=n$ *and let* $\{\alpha\},\{\beta\}$ *be Kähler classes such that* $\{\alpha\}^n-n\,\{\alpha\}^{n-1}.\{\beta\}>0$. *Then, for every* $k\in\{1,2,\dots,n\}$ *and every smooth positive* $(n-k,\,n-k)$*-form* $\Omega^{n-k,\,n-k}\geq 0$ *such that* $\partial\bar\partial\Omega^{n-k,\,n-k}=0$*, the following inequalities hold:*

$$\begin{aligned}\{\alpha^k-\beta^k\}.\big[\Omega^{n-k,\,n-k}\big]_A &\overset{(I_k)}{\geq}\{\alpha-\beta\}^k.\big[\Omega^{n-k,\,n-k}\big]_A\\ &\overset{(II_k)}{\geq}\big\{\alpha^k-k\,\alpha^{k-1}\wedge\beta\big\}.\big[\Omega^{n-k,\,n-k}\big]_A\\ &\overset{(III_k)}{\geq}\left(1-\frac{n}{R}\right)\{\alpha\}^k.\big[\Omega^{n-k,\,n-k}\big]_A\geq 0,\end{aligned} \quad (5.10)$$

where, as usual, $R := \frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}}$. *(Thus* $R > n$ *by assumption.) In particular,* (II_n) *and* (III_n) *read:*

$$\{\alpha - \beta\}^n \geq \{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} = \left(1 - \frac{n}{R}\right)\{\alpha\}^n > 0. \tag{5.11}$$

Proof. We may and will assume without loss of generality that $\Omega^{n-k,\,n-k}$ is strictly positive.
Inequality (III_k) is nothing but (5.1) for $t = 1$ and $p = k$.
We will now prove (II_k) by induction on $k \in \{1, \dots, n\}$. Let us fix Kähler metrics α, β in the classes $\{\alpha\}$, respectively $\{\beta\}$. For $k = 1$, (II_1) is obviously an identity. Now, proving (II_k) for an arbitrary k amounts to proving that the quantity

$$S_k := \int_X \Big((\alpha - \beta)^k - \alpha^k + k\,\alpha^{k-1} \wedge \beta\Big) \wedge \Omega^{n-k,\,n-k} \tag{5.12}$$

is non-negative. To this end, we first prove the identity:

$$S_k = \sum_{l=1}^{k-1} l \int_X (\alpha - \beta)^{k-l-1} \wedge \alpha^{l-1} \wedge \beta^2 \wedge \Omega^{n-k,\,n-k}, \quad k = 1, \dots, n. \tag{5.13}$$

This follows immediately by writing the next pointwise identities:

$$\begin{aligned}
&(\alpha - \beta)^k - \alpha^k + k\,\alpha^{k-1} \wedge \beta = -\beta \wedge \sum_{l=1}^{k} (\alpha - \beta)^{k-l} \wedge \alpha^{l-1} + k\,\alpha^{k-1} \wedge \beta \\
&= \sum_{l=1}^{k-1} \alpha^{l-1} \wedge \beta \wedge \Big(\alpha^{k-l} - (\alpha - \beta)^{k-l}\Big) = \sum_{l=1}^{k-1} \alpha^{l-1} \wedge \beta^2 \wedge \sum_{r=0}^{k-l-1} \alpha^{k-l-1-r} \wedge (\alpha - \beta)^r \\
&= \sum_{l=1}^{k-1} \sum_{r=0}^{k-l-1} \alpha^{k-r-2} \wedge \beta^2 \wedge (\alpha - \beta)^r \\
&= \sum_{r=0}^{k-2} \alpha^{k-r-2} \wedge \beta^2 \wedge (\alpha - \beta)^r + \dots \\
&\quad + \sum_{r=0}^{k-l-1} \alpha^{k-r-2} \wedge \beta^2 \wedge (\alpha - \beta)^r + \dots + \alpha^{k-2} \wedge \beta^2 \\
&= \beta^2 \wedge (\alpha - \beta)^{k-2} + 2\,\alpha \wedge \beta^2 \wedge (\alpha - \beta)^{k-3} + \dots \\
&\quad + l\,\alpha^{l-1} \wedge \beta^2 \wedge (\alpha - \beta)^{k-l-1} + \dots + (k-1)\,\alpha^{k-2} \wedge \beta^2 \\
&= \sum_{l=1}^{k-1} l\,(\alpha - \beta)^{k-l-1} \wedge \alpha^{l-1} \wedge \beta^2.
\end{aligned}$$

This clearly proves (5.13).

Now we can run the induction on $k \in \{1, \dots, n\}$ to prove (II_k). Suppose that $(II_1), \dots, (II_{k-1})$ have been proved. Combining them with (III_k) that was proved in (5.1) for all $k \in \{1, \dots, n\}$, we deduce that the classes $\{\alpha - \beta\}^{k-r}$ are positive in the following sense:

$$\{\alpha - \beta\}^{k-r}.\big[\Omega^{n-k+r,\, n-k+r}\big]_A \geq 0$$

for all $r \in \{1, \dots, k\}$ and for all C^∞ strictly positive $(n-k+r,\ n-k+r)$-forms $\Omega^{n-k+r,\, n-k+r} > 0$ such that $\partial\bar\partial\Omega^{n-k+r,\, n-k+r} = 0$.

Choosing forms of the shape $\Omega^{n-k+r,\, n-k+r} := \alpha^{r-2} \wedge \beta^2 \wedge \Omega^{n-k,\, n-k}$ with $\Omega^{n-k,\, n-k} > 0$ of bidegree $(n-k,\ n-k)$ satisfying $\partial\bar\partial\Omega^{n-k,\, n-k} = 0$, we get:

$$\{\alpha - \beta\}^{k-r}.\{\alpha\}^{r-2}.\{\beta\}^2.\big[\Omega^{n-k,\, n-k}\big]_A \geq 0, \quad r \in \{2, \dots, k\}.$$

Setting $r := l + 1$, this translates to

$$\int_X (\alpha - \beta)^{k-l-1} \wedge \alpha^{l-1} \wedge \beta^2 \wedge \Omega^{n-k,\, n-k} \geq 0, \quad l \in \{1, \dots, k-1\},$$

which means precisely that all the terms in the sum expressing S_k in (5.13) are non-negative. Hence, $S_k \geq 0$, which proves (II_k) (see 5.12).

Let us now prove (I_k) as a consequence of (II_k) and (III_k). For every $k \in \{1, \dots, n\}$, the following pointwise identities are obvious:

$$\begin{aligned}
\alpha^k - \beta^k - (\alpha - \beta)^k &= \beta \wedge \sum_{l=0}^{k-1} \alpha^{k-l-1} \wedge (\alpha - \beta)^l - \beta^k \\
&= \beta \wedge \left(\alpha^{k-1} - \beta^{k-1} + \sum_{l=1}^{k-1} \alpha^{k-l-1} \wedge (\alpha - \beta)^l\right) \\
&= \beta \wedge \left((\alpha - \beta) \wedge \sum_{r=0}^{k-2} \alpha^{k-r-2} \wedge \beta^r + \sum_{l=1}^{k-1} \alpha^{k-l-1} \wedge (\alpha - \beta)^l\right).
\end{aligned}$$

Hence, for every smooth $(n-k,\ n-k)$-form $\Omega^{n-k,\, n-k} \geq 0$ such that $\partial\bar\partial\Omega^{n-k,\, n-k} = 0$, we have:

$$\begin{aligned}
&\Big(\big\{\alpha^k - \beta^k\big\} - \{\alpha - \beta\}^k\Big).\big[\Omega^{n-k,\, n-k}\big]_A \\
&= \sum_{r=0}^{k-2} \int_X (\alpha - \beta) \wedge \alpha^{k-r-2} \wedge \beta^{r+1} \wedge \Omega^{n-k,\, n-k} \\
&\quad + \sum_{l=1}^{k-1} \int_X (\alpha - \beta)^l \wedge \alpha^{k-l-1} \wedge \beta \wedge \Omega^{n-k,\, n-k} \\
&= \sum_{r=0}^{k-2} \{\alpha - \beta\}.\Big[\Omega_r^{n-1,\, n-1}\Big]_A + \sum_{l=1}^{k-1} \{\alpha - \beta\}^l.\Big[\Gamma_l^{n-l,\, n-l}\Big]_A \\
&\geq 0,
\end{aligned}$$

where we have put $\Omega_r^{n-1,\,n-1} := \alpha^{k-r-2} \wedge \beta^{r+1} \wedge \Omega^{n-k,\,n-k}$ and $\Gamma_l^{n-l,\,n-l} := \alpha^{k-l-1} \wedge \beta \wedge \Omega^{n-k,\,n-k}$. It is clear that $\Omega_r^{n-1,\,n-1}$ and $\Gamma_l^{n-l,\,n-l}$ are positive $\partial\bar\partial$-closed forms of bidegree $(n-1,\, n-1)$, respectively $(n-l,\, n-l)$, so the last inequality follows from the combination of (II_k) and (III_k). Thus (I_k) is proved. $\square$

We immediately get the following consequence of Theorem 5.3 which is the analogue of Theorem 3.5 in bidegree $(k,\, k)$ for an arbitrary k.

Corollary 5.4. *Let X be a compact Kähler manifold with $\dim_{\mathbb{C}} X = n$ and let $\alpha, \beta > 0$ be Kähler metrics such that $\{\alpha\}^n - n\,\{\alpha\}^{n-1}.\{\beta\} > 0$. Then, for every $k \in \{1, 2, \dots, n\}$, there exist closed positive $(k,\, k)$-currents $U_k \in \{\alpha^k - \beta^k\}$ and $S_k \in \{(\alpha - \beta)^k\}$ such that*

$$U_k \geq \left(1 - \frac{n}{R}\right) \alpha^k \quad \textit{and} \quad S_k \geq \left(1 - \frac{n}{R}\right) \alpha^k \tag{5.14}$$

on X, where, as usual, we let $R := \frac{\{\alpha\}^n}{\{\alpha\}^{n-1}.\{\beta\}}$. (So $R > n$ by assumption.)

Proof. This follows immediately from Theorem 5.3 by using the analogue of Lamari's positivity criterion [18, Lemme 3.3] in bidegree $(k,\, k)$ for every k. $\square$

6. A conjecture in the non-Kähler context

Let X be a compact complex manifold with $\dim_{\mathbb{C}} X = n$. It is standard that if X is of *class $\mathcal{C}$*, then X is both *balanced* (*i.e.* it admits a balanced metric: a Hermitian metric ω such that $d\omega^{n-1} = 0$) by [1, Corollary 4.5] and a $\partial\bar\partial$-manifold (*i.e.* the $\partial\bar\partial$-lemma holds on X). On the other hand, there are a great deal of examples of balanced manifolds that are not $\partial\bar\partial$-manifolds (*e.g.*, the Iwasawa manifold), but it is still an open problem to find out whether or not every $\partial\bar\partial$-manifold admits a balanced metric. To the author's knowledge, all the examples of $\partial\bar\partial$-manifolds known so far are also balanced. We now briefly indicate how a generalised version of Demailly's *Transcendental Morse Inequalities Conjecture* for a difference of two nef classes might answer a stronger version of this question. The main idea is borrowed from Toma's work [22] in the projective setting and was also exploited in [9] in the Kähler setting.

It is standard that the canonical linear map induced in cohomology by the identity:

$$I_{n-1} : H_{BC}^{n-1,\,n-1}(X,\, \mathbb{C}) \to H_A^{n-1,\,n-1}(X,\, \mathbb{C}), \quad [\Omega]_{BC} \mapsto [\Omega]_A, \tag{6.1}$$

is well defined on every X, but it is neither injective, nor surjective in general. Moreover, the balanced cone of X consisting of Bott-Chern cohomology classes of bidegree $(n-1,\ n-1)$ representable by balanced metrics ω^{n-1}:

$$\mathcal{B}_X = \big\{[\omega^{n-1}]_{BC}\,/\,\omega > 0, \quad C^\infty\ (1,\ 1)\text{-form such that}\quad d\omega^{n-1} = 0 \quad \text{on} \quad X\big\} \subset H^{n-1,\,n-1}_{BC}(X,\ \mathbb{R}),$$

maps under I_{n-1} to a subset of the Gauduchon cone of X (introduced in [19]) consisting of Aeppli cohomology classes of bidegree $(n-1,\ n-1)$ representable by Gauduchon metrics ω^{n-1}:

$$\mathcal{G}_X = \big\{[\omega^{n-1}]_{A}\,/\,\omega > 0, \quad C^\infty\ (1,\ 1)\text{-form such that}\quad \partial\bar\partial\omega^{n-1} = 0 \quad \text{on} \quad X\big\} \subset H^{n-1,\,n-1}_{A}(X,\ \mathbb{R}).$$

Clearly, the inclusion $I_{n-1}(\mathcal{B}_X) \subset \mathcal{G}_X$ is strict in general. So is the inclusion $I_{n-1}(\overline{\mathcal{B}_X}) \subset \overline{\mathcal{G}}_X$ involving the closures of these two open convex cones.

Now, if X is a $\partial\bar\partial$-manifold, I_{n-1} is an isomorphism of the vector spaces $H^{n-1,\,n-1}_{BC}(X,\ \mathbb{C})$ and $H^{n-1,\,n-1}_{A}(X,\ \mathbb{C})$, as is well known. It is tempting to make the following:

Conjecture 6.1. If X is a compact $\partial\bar\partial$-manifold of dimension n, then $I_{n-1}(\overline{\mathcal{B}_X}) = \overline{\mathcal{G}}_X$.

If proved to hold, this conjecture would imply that every $\partial\bar\partial$-manifold is actually balanced since the Gauduchon cone is never empty (due to the existence of Gauduchon metrics by [16]), so the balanced cone would also have to be non-empty in this case. Moreover, a positive answer to this conjecture would have far-reaching implications for a possible future non-Kähler mirror symmetry theory since it would remove the ambiguity of choice between the balanced and the Gauduchon cones on $\partial\bar\partial$-manifolds. These two cones would be canonically equivalent on $\partial\bar\partial$-manifolds in this event.

One piece of evidence supporting Conjecture 6.1 is that it holds on every *class* $\mathcal{C}$ manifold X if the whole of Demailly's *Transcendental Morse Inequalities Conjecture* for a difference of two nef classes is confirmed when X is Kähler. This is the gist of the observations made in [22] and in [9] alluded to above. Indeed, if X is of *class* $\mathcal{C}$, we may assume without loss of generality that X is actually compact Kähler. As proved in [6], a complete positive answer to Conjecture 1.1 would imply that the pseudo-effective cone $\mathcal{E}_X \subset H^{1,\,1}(X, \mathbb{R})$ of classes of d-closed positive $(1,\ 1)$-currents T is the dual of the cone $\mathcal{M}_X \subset H^{n-1,\,n-1}(X, \mathbb{R})$ of *movable classes* (*i.e.* the closure of the cone generated by classes of currents of the shape $\mu_\star(\widetilde\omega_1 \wedge \cdots \wedge \widetilde\omega_{n-1})$, where $\mu : \widetilde X \to X$ is any modification of compact Kähler manifolds and the $\widetilde\omega_j$ are any Kähler metrics on $\widetilde X$ – see [6, Definition 1.3]). Since on $\partial\bar\partial$-manifolds (hence, in particular, on compact Kähler ones) the Bott-Chern, Dolbeault and Aeppli cohomologies are canonically equivalent, it is irrelevant in which of these cohomologies the groups $H^{1,\,1}(X, \mathbb{R})$ and $H^{n-1,\,n-1}(X, \mathbb{R})$ are considered.

The closure $\overline{\mathcal{G}}_X \subset H^{n-1,\,n-1}(X, \mathbb{R})$ of the Gauduchon cone is dual to the pseudo-effective cone $\mathcal{E}_X \subset H^{1,\,1}(X, \mathbb{R})$ by Lamari's positivity criterion (Lemma

1.5), while the same kind of argument (*i.e.* duality and Hahn-Banach) going back to Sullivan shows that the closure $\overline{\mathcal{B}}_X \subset H^{n-1,\,n-1}(X, \mathbb{R})$ of the balanced cone is dual to the cone

$$\mathcal{S}_X = \big\{[T]_A \,/\, T \geq 0,\ T \text{ is a } (1,\ 1)\text{-current such that } \quad \partial\bar\partial T = 0 \quad \text{on} \quad X\big\} \subset H_A^{1,\,1}(X,\ \mathbb{R}).$$

Note that $\mathcal{S}_X$ is closed if X admits a balanced metric ω^{n-1} (against which the masses of positive $\partial\bar\partial$-closed $(1,\ 1)$-currents T can be considered), hence so is it when X is Kähler. Thus, by duality, the identity $I_{n-1}(\overline{\mathcal{B}_X}) = \overline{\mathcal{G}}_X$ is equivalent to $I_1(\mathcal{E}_X) = \mathcal{S}_X$, where I_1 is the canonical linear map induced in cohomology by the identity:

$$I_1 : H_{BC}^{1,\,1}(X,\ \mathbb{C}) \to H_A^{1,\,1}(X,\ \mathbb{C}), \quad [\gamma]_{BC} \mapsto [\gamma]_A. \tag{6.2}$$

In general, I_1 is neither injective, nor surjective, but it is an isomorphism when X is a $\partial\bar\partial$-manifold.

With these facts understood, the identity $I_1(\mathcal{E}_X) = \mathcal{S}_X$ can be proved when X is Kähler (provided that Conjecture 1.1 can be solved in the affirmative) as explained in [9, Proposition 2.5] by an argument generalising to transcendental classes an earlier argument from [22] that we now recall for the reader's convenience.

The inclusion $I_1(\mathcal{E}_X) \subset \mathcal{S}_X$ is obvious. To prove the reverse inclusion, let $[T]_A \in \mathcal{S}_X$, *i.e.* $T \geq 0$ is a $(1,\ 1)$-current such that $\partial\bar\partial T = 0$. Since I_1 is an isomorphism, there exists a unique class $[\gamma]_{BC} \in H_{BC}^{1,1}(X,\ \mathbb{R})$ such that $I_1([\gamma]_{BC}) = [T]_A$. This means that $[\gamma]_A = [T]_A$. We will show that $[\gamma]_{BC} \in \mathcal{E}_X$. If the [6] conjecture (predicated on Conjecture 1.1) predicting that $\mathcal{E}_X$ is dual to $\mathcal{M}_X$ is confirmed, showing that $[\gamma]_{BC} \in \mathcal{E}_X$ amounts to showing that

$$[\gamma]_{BC} \cdot [\mu_\star(\widetilde\omega_1 \wedge \cdots \wedge \widetilde\omega_{n-1})]_A \geq 0 \tag{6.3}$$

for all modifications $\mu : \widetilde{X} \to X$ and all Kähler metrics $\widetilde\omega_j$ on $\widetilde{X}$. On the other hand, Alessandrini and Bassanelli proved in [2, Theorem 5.6] the existence and uniqueness of the inverse image under proper modifications $\mu : \widetilde{X} \to X$ of arbitrary complex manifolds of any positive $\partial\bar\partial$-closed $(1,\ 1)$-current $T \geq 0$ in such a way that the Aeppli cohomology class $[T]_A$ is preserved:

$$\exists!\ (1,\ 1)\text{-current } \mu^\star T \geq 0 \text{ on } \widetilde{X} \quad \text{such that} \quad \partial\bar\partial\,(\mu^\star T) = 0,$$
$$[\mu^\star T]_A = \mu^\star([T]_A) \quad \text{and} \quad \mu_\star(\mu^\star T) = T.$$

(Note that the inverse image $\mu^\star([T]_A)$ of any Aeppli class is trivially well defined by taking smooth representatives of the class and pulling them back. Indeed, $\partial\bar\partial$-closedness is preserved, while pullbacks of Aeppli-cohomologous smooth forms are trivially seen to be Aeppli-cohomologous.) Using this key ingredient from [2], we get:

$$
\begin{aligned}
&[\gamma]_{BC} \, . \, [\mu_\star(\widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_{n-1})]_A \\
&= \int_X \gamma \wedge \mu_\star(\widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_{n-1}) = \int_{\widetilde{X}} (\mu^\star\gamma) \wedge (\widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_{n-1}) \\
&= [\mu^\star\gamma]_A \, . \, [\widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_{n-1}]_{BC} = \int_{\widetilde{X}} (\mu^\star T) \wedge (\widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_{n-1}) \geq 0,
\end{aligned}
$$

which proves (6.3). Note that γ and $\mu^\star\gamma$ have no sign, so the key point has been the replacement in the integral over $\widetilde{X}$ of $\mu^\star\gamma$ by $\mu^\star T \geq 0$ which was made possible by $\widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_{n-1}$ being d-closed (so we could switch the roles of the Bott-Chern and the Aeppli cohomologies) and by the identity $[\mu^\star\gamma]_A = [\mu^\star T]_A$ following from $[\gamma]_A = [T]_A$ (see above) and from $[\mu^\star T]_A = \mu^\star([T]_A)$.

The techniques employed in this paper do not seem to be using the full force of the Kähler assumption on X and many of the arguments are valid in a more general context. This is part of the justification for proposing Conjecture 6.1.

7. Appendix: Hovanskii-Teissier-type inequalities

In this section we prove the pointwise inequalities for Hermitian metrics that were used in earlier sections. They generalise the inequality in [20, Lemma 3.1].

For the sake of enhanced flexibility, we shall deal with positive $(q,\, q)$-forms that are not necessarily the q^{th} power of a positive $(1,\, 1)$-form. Given any $q \in \{0, \ldots, n\}$ and any C^∞ real $(q,\, q)$-form $\Omega^{q,\, q}$ on X, we make use of the standard notion of (weak) positivity (see, *e.g.*, [13, III.1.1]): $\Omega^{q,\, q}$ is said to be *positive* (respectively *strictly positive*) if for any $(1,\, 0)$-forms $\alpha_1, \ldots \alpha_{n-q}$, the $(n,\, n)$-form $\Omega^{q,\, q} \wedge i\alpha_1 \wedge \overline{\alpha}_1 \wedge \cdots \wedge i\alpha_{n-q} \wedge \overline{\alpha}_{n-q}$ is non-negative (respectively positive). We write $\Omega^{q,\, q} \geq 0$ (respectively $\Omega^{q,\, q} > 0$) in this case. If, in local holomorphic coordinates $z_1, \ldots, z_n$, we write

$$
\frac{\Omega^{q,\, q}}{q!} = \sum_{|L|=|R|=q} \Omega_{L\bar{R}} \, i dz_L \wedge d\bar{z}_R, \tag{7.1}
$$

then it is clear by considering $\Omega^{q,\, q} \wedge idz_{s_1} \wedge d\bar{z}_{s_1} \wedge \cdots \wedge idz_{s_{n-q}} \wedge d\bar{z}_{s_{n-q}}$ that

$$
\Omega^{q,\, q} \geq 0 \quad \text{implies} \quad \Omega_{L\bar{L}} \geq 0 \quad \text{for all} \quad L \quad \text{with} \quad |L| = q. \tag{7.2}
$$

(We have used the usual notation: L and R stand for ordered multi-indices $L = (1 \leq l_1 < \cdots < l_q \leq n)$, respectively $R = (1 \leq r_1 < \cdots < r_q \leq n)$ of length q and $idz_L \wedge d\bar{z}_R := idz_{l_1} \wedge d\bar{z}_{r_1} \wedge \cdots \wedge idz_{l_q} \wedge d\bar{z}_{r_q}$.)

In the special case when $\Omega^{q,\, q} = \gamma^q$ for some positive definite smooth $(1,\, 1)$-form (= Hermitian metric) γ on X, if we write

$$
\gamma = \sum_{j,k=1}^{n} \gamma_{j\bar{k}} \, idz_j \wedge d\bar{z}_k, \tag{7.3}
$$

then Sylvester's criterion ensures that $M_{L\bar{L}}(\gamma) > 0$ for all multi-indices $L \subset \{1, \dots, n\}$ of any length $l \in \{1, \dots, n\}$. (For any multi-indices $L, K \subset \{1, \dots, n\}$ of equal lengths, $M_{K\bar{L}}(\gamma)$ denotes the minor of the matrix $(\gamma_{j\bar{k}})_{j,k}$ corresponding to the rows with index in K and the columns with index in L.) Clearly, $M_{L\bar{L}}(\gamma) = \Omega_{L\bar{L}}$ for all L with $|L| = q$.

Lemma 7.1. *Let α, β be arbitrary Hermitian metrics on a complex manifold X with $\dim_{\mathbb{C}} X = n$.*

The following pointwise inequalities hold for every $p \in \{1, \dots, n\}$ and for every smooth form $\Omega^{n-p,\,n-p} \geq 0$ of bidegree $(n-p,\ n-p)$ on X:

$$\binom{n}{p} \frac{\alpha^{n-p} \wedge \beta^p}{\alpha^n} \cdot \frac{\alpha^p \wedge \Omega^{n-p,\,n-p}}{\alpha^n} \geq \frac{\beta^p \wedge \Omega^{n-p,\,n-p}}{\alpha^n} \tag{7.4}$$

and

$$\frac{n}{p} \frac{\alpha^{n-1} \wedge \beta}{\alpha^n} \cdot \frac{\alpha^p \wedge \Omega^{n-p,\,n-p}}{\alpha^n} \geq \frac{\alpha^{p-1} \wedge \beta \wedge \Omega^{n-p,\,n-p}}{\alpha^n}. \tag{7.5}$$

Proof. Let us first prove (7.4). The special case when $p = 1$ and $\Omega^{n-1,\,n-1} = \gamma^{n-1}$ for some $(1,1)$-form $\gamma > 0$ was proved in [20, Lemma 3.1]. We fix any point $x \in X$ and choose local coordinates $z_1, \dots, z_n$ about x such that

$$\alpha(x) = \sum_{j=1}^{n} i dz_j \wedge d\bar{z}_j \quad \text{and} \quad \beta(x) = \sum_{j=1}^{n} \beta_j \, i dz_j \wedge d\bar{z}_j. \tag{7.6}$$

Thus $\beta_j > 0$ for all j. At x we get: $\frac{\beta^p}{p!} = \sum_{j_1 < \dots < j_p} \beta_{j_1} \dots \beta_{j_p} \bigwedge_{l \in \{j_1, \dots, j_p\}} (i dz_l \wedge d\bar{z}_l)$, hence

$$\frac{\alpha^{n-p} \wedge \beta^p}{\alpha^n} = \frac{1}{\binom{n}{p}} \sum_{j_1 < \dots < j_p} \beta_{j_1} \dots \beta_{j_p} = \frac{\beta_1 \dots \beta_n}{\binom{n}{p}} \left(\sum_{|K| = n-p} \frac{1}{\beta_K} \right) \quad \text{at } x, \tag{7.7}$$

where $\beta_K := \beta_{k_1} \dots \beta_{k_{n-p}}$ whenever $K = (1 \leq k_1 < \dots k_{n-p} \leq n)$. On the other hand, using (7.1) with $q = n - p$, we get at x:

$$\frac{\alpha^p \wedge \Omega^{n-p,\,n-p}}{\alpha^n} = \frac{1}{\binom{n}{p}} \sum_{|L| = n-p} \Omega_{L\bar{L}}$$

$$\text{and} \quad \frac{\beta^p \wedge \Omega^{n-p,\,n-p}}{\alpha^n} = \frac{\beta_1 \dots \beta_n}{\binom{n}{p}} \sum_{|L| = n-p} \frac{\Omega_{L\bar{L}}}{\beta_L}.$$

Thus, inequality (7.4) at x is equivalent to:

$$\left(\sum_{|L| = n-p} \Omega_{L\bar{L}} \right) \frac{\beta_1 \dots \beta_n}{\binom{n}{p}} \left(\sum_{|K| = n-p} \frac{1}{\beta_K} \right) \geq \frac{\beta_1 \dots \beta_n}{\binom{n}{p}} \sum_{|L| = n-p} \frac{\Omega_{L\bar{L}}}{\beta_L},$$

which clearly holds since $\Omega_{L\bar{L}} \geq 0$ and $\beta_K > 0$ for all multi-indices K, L.

Let us now prove (7.5). With the above notation, we have at x:

$$\frac{\alpha^{n-1}\wedge\beta}{\alpha^n} = \frac{1}{n}\sum_{j=1}^{n}\beta_j \quad \text{and} \quad \frac{\alpha^{p-1}}{(p-1)!} = \sum_{|J|=p-1} idz_J\wedge d\bar{z}_J,$$

and the second identity yields at x:

$$\frac{\alpha^{p-1}\wedge\beta}{(p-1)!} = \sum_{|J|=p-1}\sum_{j\in\{1,\dots,n\}\setminus J}\beta_j\, idz_j\wedge d\bar{z}_j\wedge idz_J\wedge d\bar{z}_J,$$

which, in turn, implies the following identity at x:

$$\frac{\alpha^{p-1}\wedge\beta\wedge\Omega^{n-p,\,n-p}}{(p-1)!\,(n-p)!} = \sum_{|J|=p-1}\sum_{j\in\{1,\dots,n\}\setminus J}\beta_j\,\Omega_{L_{jJ}\bar{L}_{jJ}}\,\frac{\alpha^n}{n!},$$

where we have set $\Omega_{L_{jJ}\bar{L}_{jJ}} := \Omega_{L\bar{L}}$ with $L := \{1,\dots,n\}\setminus(\{j\}\cup J)$ ordered increasingly. Thus, $\{j\}$, J and L form a partition of $\{1,\dots,n\}$, so any two of them uniquely determine the third.

Consequently, inequality (7.5) at x translates to

$$\begin{aligned}&\frac{n}{p}\,\frac{1}{\binom{n}{p}\,n}\left(\sum_{|L|=n-p}\Omega_{L\bar{L}}\right)\left(\sum_{j=1}^{n}\beta_j\right)\\ &\geq \frac{(p-1)!\,(n-p)!}{n!}\sum_{|J|=p-1}\sum_{j\in\{1,\dots,n\}\setminus J}\beta_j\,\Omega_{L_{jJ}\bar{L}_{jJ}},\end{aligned}$$

which is clear since $\frac{n}{p}\,\frac{1}{\binom{n}{p}n} = \frac{(p-1)!\,(n-p)!}{n!}$, $\Omega_{L\bar{L}}\geq 0$ for every L, $\beta_j>0$ for every j and the terms in the double sum on the right-hand side of the above inequality are precisely all the products of the shape $\Omega_{L\bar{L}}\,\beta_j$ with $j\notin L$, so they form a subset of the terms on the left hand side □

Note that inequalities (7.4) and (7.5) of Lemma 7.1 allow a kind of "simplification" of α^n between the numerators and the denominators. For possible future use, we notice a simultaneous reinforcement of inequalities (7.4) and (7.5) that has not been used in this paper. For this reason and since the proof of the general case involves rather lengthy calculations, we will only prove a special case.

Lemma 7.2. *Let α, β be arbitrary Hermitian metrics on a complex manifold X with* $\dim_{\mathbb{C}} X = n$. *Let $p\in\{1,\dots,n\}$ be arbitrary.*

If $\Omega^{n-p,\,n-p}$ is proportional to $\alpha^k\wedge\beta^{n-p-k}$ for some $k\in\{0,\dots,n-p\}$, then the factor $\binom{n}{p}$ can be omitted from (7.4). *In other words, for all $p,k\in\{0,\dots,n\}$ such that $p+k\leq n$ we have:*

$$\frac{\alpha^{n-p}\wedge\beta^p}{\alpha^n}\cdot\frac{\alpha^{p+k}\wedge\beta^{n-p-k}}{\alpha^n} \geq \frac{\alpha^k\wedge\beta^{n-k}}{\alpha^n} \quad \textit{on} \quad X. \tag{7.8}$$

Proof. We will only prove here the case when $p+k=n-1$, *i.e.*

$$\frac{\alpha^{n-p}\wedge\beta^p}{\alpha^n}\cdot\frac{\beta\wedge\alpha^{n-1}}{\alpha^n}\geq\frac{\alpha^{n-p-1}\wedge\beta^{p+1}}{\alpha^n}\quad\text{on}\quad X, \tag{7.9}$$

which is equivalent to $(\Lambda_\gamma\alpha)\,(\Lambda_\alpha\beta)\geq n\,\Lambda_\gamma\beta$ when $\gamma^{n-1}=t\,\alpha^{n-p-1}\wedge\beta^p$ for some constant $t>0$. Notice that this last inequality improves by a factor n in the special case when $\gamma^{n-1}=t\,\alpha^{n-p-1}\wedge\beta^p$ the general lower bound proved in [20, Lemma 3.1].

We fix an arbitrary point $x\in X$ and choose local coordinates as in (7.6). Using identities analogous to those in the proof of Lemma 7.1, we see that (7.9) translates at x to

$$\frac{n-p}{n}\left(\sum_{j_1<\dots<j_p}\beta_{j_1}\dots\beta_{j_p}\right)\left(\sum_{l=1}^n\beta_l\right)\geq(p+1)\sum_{k_1<\dots<k_{p+1}}\beta_{k_1}\dots\beta_{k_{p+1}}. \tag{7.10}$$

Now, the left hand side of inequality (7.10) equals

$$\frac{n-p}{n}\left((p+1)\sum_{k_1<\dots<k_{p+1}}\beta_{k_1}\dots\beta_{k_{p+1}}+\sum_{j_1<\dots<j_p}\beta_{j_1}\dots\beta_{j_p}(\beta_{j_1}+\dots+\beta_{j_p})\right),$$

so (7.10) is equivalent to

$$(n-p)\sum_{j_1<\dots<j_p}\beta_{j_1}\dots\beta_{j_p}(\beta_{j_1}+\ldots+\beta_{j_p})\geq p(p+1)\sum_{k_1<\dots<k_{p+1}}\beta_{k_1}\dots\beta_{k_{p+1}}. \tag{7.11}$$

We will now prove (7.11). Let us fix an arbitrary ordered sequence $1\leq k_1<\dots<k_{p+1}\leq n$. For every $r,s\in\{k_1,\dots,k_{p+1}\}$ with $r<s$, we have:

$$2\beta_{k_1}\dots\beta_{k_{p+1}}=(2\beta_r\,\beta_s)\prod_{l\notin\{r,s\}}\beta_l\leq\beta_r^2\prod_{l\notin\{r,s\}}\beta_l+\beta_s^2\prod_{l\notin\{r,s\}}\beta_l, \tag{7.12}$$

where all the products above bear on the indices $l\in\{k_1,\dots,k_{p+1}\}\setminus\{r,s\}$. Note that $\beta_r^2\prod_{l\notin\{r,s\}}\beta_l$ is obtained from $\beta_{k_1}\dots\beta_{k_{p+1}}$ by omitting β_s and counting β_r twice. Summing up these inequalities over all the $\binom{p+1}{2}$ pairs of indices $r<s$ selected from $k_1,\dots,k_{p+1}$, we get

$$\begin{aligned}\binom{p+1}{2}2\beta_{k_1}\dots\beta_{k_{p+1}}\leq\beta_{k_2}\dots\beta_{k_{p+1}}&\big(\beta_{k_2}+\dots+\beta_{k_{p+1}}\big)\\&+\beta_{k_1}\beta_{k_3}\dots\beta_{k_{p+1}}\big(\beta_{k_1}+\beta_{k_3}+\dots+\beta_{k_{p+1}}\big)\\&+\dots\dots\dots\\&+\beta_{k_1}\dots\beta_{k_p}\big(\beta_{k_1}+\dots+\beta_{k_p}\big).\end{aligned} \tag{7.13}$$

Note that for every $s \in \{1, \dots, p+1\}$, β_{k_s} does not feature in the s^{th} line on the right-hand side of (7.13). Adding up these inequalities over all the ordered sequences $1 \leq k_1 < \cdots < k_{p+1} \leq n$, we get the desired inequality (7.11) because any ordered sequence $1 \leq j_1 < \cdots < j_p \leq n$ occurs inside exactly $(n-p)$ ordered sequences $1 \leq k_1 < \cdots < k_{p+1} \leq n$. Indeed, the extra index for $1 \leq k_1 < \cdots < k_{p+1} \leq n$ can be chosen arbitrarily in $\{1, \dots, n\} \setminus \{j_1, \dots, j_p\}$, so there are $(n-p)$ choices for it.

This completes the proof of (7.11), hence the proof of (7.8) when $p + k = n - 1$. □

Again for the record, we notice that an application of Lemma 7.2 is an inequality between intersection numbers of cohomology classes reminiscent of the Hovanskii-Teissier inequalities (*cf.*, *e.g.*, [12, Proposition 5.2]). It has an interest of its own.

Proposition 7.3. *Let X be a compact Kähler manifold with* $\dim_{\mathbb{C}} X = n$ *and let* $\{\alpha\}, \{\beta\} \in H^{1,1}_{BC}(X, \mathbb{R})$ *be nef Bott-Chern cohomology classes. Then*

$$\big(\{\alpha\}^{n-p}.\{\beta\}^p\big)\big(\{\alpha\}^{p+k}.\{\beta\}^{n-p-k}\big) \geq \big(\{\alpha\}^n\big)\big(\{\alpha\}^k.\{\beta\}^{n-k}\big) \tag{7.14}$$

for all $p, k \in \{0, \dots, n\}$ such that $p + k \leq n$.

By the density of the nef cone in the Kähler cone, we may assume without loss of generality that $\{\alpha\}$ and $\{\beta\}$ are Kähler classes in which we fix respective Kähler metrics α, β.

Proof 1 (*deduced from a known result*).[1] For every $j \in \{0, \dots n\}$, let

$$c_j := \log\big(\{\alpha\}^j.\{\beta\}^{n-j}\big).$$

It is a standard result that the function $j \mapsto c_j$ is *concave*. Now, $k \leq n - p \leq n$ and

$$\begin{aligned} n - p &= \frac{p}{n-k}\,k + \frac{n-k-p}{n-k}\,n, \\ \text{hence, by concavity,}\quad c_{n-p} &\geq \frac{p}{n-k}\,c_k + \frac{n-k-p}{n-k}\,c_n. \end{aligned} \tag{7.15}$$

Similarly, $k \leq p + k \leq n$ and

$$\begin{aligned} p + k &= \frac{n-p-k}{n-k}\,k + \frac{p}{n-k}\,n, \\ \text{hence, by concavity,}\quad c_{p+k} &\geq \frac{n-p-k}{n-k}\,c_k + \frac{p}{n-k}\,c_n. \end{aligned} \tag{7.16}$$

Taking the sum of (7.15) and (7.16), we get: $c_{n-p} + c_{p+k} \geq c_n + c_k$, which is nothing but (7.14). □

[1] This argument was kindly pointed out to the author by S. Boucksom.

Proof 2. It uses the pointwise inequality (7.8) via the technique introduced in [20] and the *approximate fixed point* technique introduced in the proof of Proposition 5.1. The arguments are essentially a repetition of some of those used above, so we will only indicate the main points.

First notice that the case when $k = 0$ is an immediate consequence of the Hovanskii-Teissier inequalities (*cf.* [12, Proposition 5.2]) which spell:

$$\{\alpha\}^{n-p}.\{\beta\}^{p} \geq \big(\{\alpha\}^n\big)^{\frac{n-p}{n}}\big(\{\beta\}^n\big)^{\frac{p}{n}} \quad \text{and} \quad \{\alpha\}^{p}.\{\beta\}^{n-p} \geq \big(\{\alpha\}^n\big)^{\frac{p}{n}}\big(\{\beta\}^n\big)^{\frac{n-p}{n}}.$$

Multiplying these two inequalities, we get (7.14) for $k = 0$.

For the general case of an arbitrary k, we consider the Monge-Ampère equation:

$$\begin{aligned} \widetilde{\alpha}^n &= \frac{\{\alpha\}^n}{\{\alpha\}^k.\{\beta\}^{n-k}}\ \alpha^k \wedge \beta^{n-k}, \\ \text{or equivalently} \quad \det_\beta \widetilde{\alpha} &= \frac{\{\alpha\}^n}{\{\alpha\}^k.\{\beta\}^{n-k}}\ \frac{\alpha^k \wedge \beta^{n-k}}{\beta^n}, \end{aligned} \tag{7.17}$$

for which the *approximate fixed point technique* introduced in the proof of Proposition 5.1 produces, for every $\varepsilon > 0$, a Kähler metric $\widetilde{\alpha}_\varepsilon$ in the Kähler class $\{\alpha\}$ (in which we have fixed beforehand a Kähler metric ω) such that

$$\widetilde{\alpha}_\varepsilon^n = \frac{\{\alpha\}^n}{\{\alpha\}^k.\{\beta\}^{n-k}}[(1-\varepsilon)\,\omega_\varepsilon + \varepsilon\omega]^k \wedge \beta^{n-k} \geq (1-\varepsilon)^k \frac{\{\alpha\}^n}{\{\alpha\}^k.\{\beta\}^{n-k}} \widetilde{\alpha}_\varepsilon^k \wedge \beta^{n-k} - O(|\eta_\varepsilon|),$$

for some constant $\eta_\varepsilon \to 0$ as $\varepsilon \to 0$. Hence

$$\det_\beta \widetilde{\alpha}_\varepsilon \geq (1-\varepsilon)^k\ \frac{\{\alpha\}^n}{\{\alpha\}^k.\{\beta\}^{n-k}}\ \frac{\widetilde{\alpha}_\varepsilon^k \wedge \beta^{n-k}}{\beta^n} - O(|\eta_\varepsilon|). \tag{7.18}$$

We can now rerun the argument used several times above. For every $\varepsilon > 0$, we have:

$$\begin{aligned} &(\{\alpha\}^{n-p}.\{\beta\}^p)\left(\{\alpha\}^{p+k}.\{\beta\}^{n-p-k}\right) \\ &= \left(\int_X \frac{\widetilde{\alpha}_\varepsilon^{n-p} \wedge \beta^p}{\beta^n}\ \beta^n\right)\left(\int_X \frac{\widetilde{\alpha}_\varepsilon^{p+k} \wedge \beta^{n-p-k}}{\widetilde{\alpha}_\varepsilon^n}\ (\det_\beta \widetilde{\alpha}_\varepsilon)\ \beta^n\right) \\ &\overset{(a)}{\geq} \left[\int_X \left(\frac{\widetilde{\alpha}_\varepsilon^{n-p} \wedge \beta^p}{\beta^n}\ \frac{\widetilde{\alpha}_\varepsilon^{p+k} \wedge \beta^{n-p-k}}{\widetilde{\alpha}_\varepsilon^n}\right)^{\frac{1}{2}} (\det_\beta \widetilde{\alpha}_\varepsilon)^{\frac{1}{2}}\ \beta^n\right]^2 \\ &\overset{(b)}{\geq} \left[\int_X \left(\frac{\widetilde{\alpha}_\varepsilon^k \wedge \beta^{n-k}}{\beta^n}\right)^{\frac{1}{2}} (\det_\beta \widetilde{\alpha}_\varepsilon)^{\frac{1}{2}}\ \beta^n\right]^2 \\ &\overset{(c)}{\geq} (1-\varepsilon)^k\ \frac{\{\alpha\}^n}{\{\alpha\}^k.\{\beta\}^{n-k}}\ \left(\int_X \widetilde{\alpha}_\varepsilon^k \wedge \beta^{n-k}\right)^2 - O(|\eta_\varepsilon|) \\ &= (1-\varepsilon)^k\ \{\alpha\}^n\ \left(\{\alpha\}^k.\{\beta\}^{n-k}\right) - O(|\eta_\varepsilon|). \end{aligned}$$

As usual, (a) follows from the Cauchy-Schwarz inequality, (b) follows from the pointwise inequality (7.8), while (c) follows from the inequality (7.18). Letting $\varepsilon \to 0$, we get (7.14). $\square$

References

[1] L. ALESSANDRINI and G. BASSANELLI, *Metric properties of manifolds bimeromorphic to compact Kähler spaces*, J. Differential Geom. **37** (1993), 95–121.

[2] L. ALESSANDRINI and G. BASSANELLI, "The Class of Compact Balanced Manifolds Is Invariant under Modifications", Complex Analysis and Geometry, Trento, 1993, 117; Lecture Notes in Pure and Appl. Math., Vol. 173, Dekker, New York, 1996.

[3] E. BEDFORD and B. TAYLOR, *A new capacity for plurisubharmonic functions*, Acta Math. **149** (1982), 1–40.

[4] Z. BLOCKI and S. KOLODZIEJ, *On regularization of plurisubharmonic functions on manifolds*, Proc. Amer. Math. Soc. **135** (2007), 2089–2093.

[5] S. BOUCKSOM, *On the volume of a line bundle*, Internat. J. Math. **13** (2002), 1043–1063.

[6] S. BOUCKSOM, J.-P. DEMAILLY, M. PAUN and T. PETERNELL, *The pseudo-effective cone of a compact Kähler manifold and varieties of negative kodaira dimension*, J. Algebraic Geom. **22** (2013), 201–248.

[7] S. BOUCKSOM, P. EYSSIDIEUX, V. GUEDJ and A. ZERIAHI, *Monge-Ampère equations in big cohomology classes*, Acta Math. **205** (2010), 199–262.

[8] S. BOUCKSOM, C. FAVRE and M. JONSSON, *Differentiability of volumes of divisors and a problem of teissier*, J. Algebraic Geom. **18** (2009), 279–308.

[9] I. CHIOSE, R. RASDEACONU and I. SUVAINA, *Balanced metrics on uniruled manifolds*, arXiv e-print DG 1408.4769v1.

[10] J.-P. DEMAILLY, *Mesures de Monge-Ampère et caractérisation géométrique des variétés algébriques affines*, Mém. Soc. Math. France (N.S.) **19** (1985), 1–124.

[11] J.-P. DEMAILLY, *Regularization of closed positive currents and intersection theory*, J. Algebraic Geom. **1** (1992), 361–409.

[12] J.-P. DEMAILLY, *A numerical criterion for very ample line bundles*, J. Differential Geom. **37** (1993), 323–374.

[13] J.-P. DEMAILLY, "Complex Analytic and Algebraic Geometry", http://www-fourier.ujf-grenoble.fr/ demailly/books.html.

[14] J.-P. DEMAILLY and M. PAUN, *Numerical characterization of the Kähler cone of a compact Kähler manifold*, Ann. of Math. **159** (2004), 1247–1274.

[15] H. FANG, M. LAI and X. MA, *On a class of fully nonlinear flows in Kähler geometry*, J. Reine Angew. Math. **653** (2011), 189–220.

[16] P. GAUDUCHON, *Le théorème de l'excentricité nulle*, C.R. Acad. Sci. Paris, Sér. A **285** (1977), 387–390.

[17] V. GUEDJ and A. ZERIAHI, *Intrinsic capacities on compact Kähler manifolds*, J. Geom. Anal. **15** (2005), 607–639.

[18] A. LAMARI, *Courants kählériens et surfaces compactes*, Ann. Inst. Fourier Grenoble **49** (1999), 263–285.

[19] D. POPOVICI, *Aeppli cohomology classes associated with Gauduchon metrics on compact complex manifolds*, Bull. Soc. Math. France (3) **143** (2015), 1–37.

[20] D. POPOVICI, *Sufficient bigness criterion for differences of two nef classes*, Math. Ann. **364** (2016), 649–655.

[21] Y.-T. SIU, *Analyticity of sets associated to lelong numbers and the extension of closed positive currents*, Invent. Math. **27** (1974), 53–156.

[22] M. TOMA, *A note on the cone of mobile curves*, C. R. Math. Acad. Sci. Paris **348** (2010), 71–73.
[23] V. TOSATTI, *The Calabi-Yau theorem and Kähler currents*, Adv. Theor. Math. Phys. **20** (2016), 381–404.
[24] V. TOSATTI and B. WEINKOVE, *The complex Monge-Ampère equation on compact hermitian manifolds*, J. Amer. Math. Soc. **23** (2010), 1187–1195.
[25] J. XIAO, *Weak transcendental holomorphic Morse inequalities on compact Kähler manifolds*, Ann. Inst. Fourier (Grenoble) **65** (2015), 1367–1379.
[26] S. T. YAU, *On the Ricci curvature of a complex Kähler manifold and the complex Monge-Ampère equation*, Comm. Pure Appl. Math. **31** (1978), 339–411.

Institut de Mathématiques de Toulouse
Université Paul Sabatier
118 route de Narbonne
31062 Toulouse, France
popovici@math.univ-toulouse.fr

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1301-1355

A problem about Mahler functions

BORIS ADAMCZEWSKI AND JASON P. BELL

In memory of Alf van der Poorten

Abstract. Let K be a field of characteristic zero and k and l be two multiplicatively independent positive integers. We prove the following result that was conjectured by Loxton and van der Poorten during the Eighties: a power series $F(z) \in K[[z]]$ satisfies both a k- and a l-Mahler-type functional equation if and only if it is a rational function.

Mathematics Subject Classification (2010): 11J81 (primary); 11B85, 65Q20 (secondary).

1. Introduction

In a series of three papers [27–29] published in 1929 and 1930, Mahler initiated a totally new direction in transcendence theory. *Mahler's method*, a term coined much later by Loxton and van der Poorten, aims at proving transcendence and algebraic independence of values at algebraic points of locally analytic functions satisfying a certain type of functional equations. In its original form, it concerns equations of the form

$$F\left(z^k\right) = R(z, F(z)), \tag{1.1}$$

where $R(z, x)$ denotes a bivariate rational function with coefficients in a number field and $k \geq 2$ is an integer. For instance, using the fact that $F(z) = \sum_{n=0}^{\infty} z^{2^n}$ satisfies the basic functional equation

$$F\left(z^2\right) = F(z) - z,$$

Mahler was able to prove that $F(\alpha)$ is a transcendental number for every algebraic number α with $0 < |\alpha| < 1$. As observed by Mahler himself, his approach allows

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No 648132. The second author was supported by NSERC grant 31-611456.
Received June 16, 2016; accepted in revised form September 23, 2016.
Published online December 2017.

one to deal with functions of several variables and systems of functional equations as well. It also leads to algebraic independence results, transcendence measures, measures of algebraic independence, and so forth. Mahler's method was later developed by various authors, including Becker, Kubota, Loxton and van der Poorten, Masser, Nishioka, Töpfer, among others. For classical aspects of Mahler's theory, we refer the reader to the monograph of Ku. Nishioka [35] and the reference therein. However, a major deficiency of Mahler's method is that, contrary to Siegel E- and G-functions, there is not a single classical transcendental constant that is known to be the value at an algebraic point of an analytic function solution to a Mahler-type functional equation[1]. This may explain why it was somewhat neglected for almost fifty years.

At the beginning of the Eighties, Mahler's method really took on a new significance after Mendès France popularized the fact that some Mahler-type systems of functional equations naturally arise in the study of automata theory (see for instance [31]). Though already noticed in 1968 by Cobham [11], this connection remained relatively unknown at that time, probably because Cobham's work was never published in an academic journal. Cobham claimed that Mahler's method has the following nice consequence for the Hartmanis-Stearns problem about the computational complexity of algebraic irrational real numbers [21]: the expansion of an algebraic irrational number in an integer base cannot be generated by a finite automaton. His idea was to derive this result by applying Mahler's method to systems of functional equations of the form

$$\begin{pmatrix} F_1(z^k) \\ \vdots \\ F_n(z^k) \end{pmatrix} = A(z) \begin{pmatrix} F_1(z) \\ \vdots \\ F_n(z) \end{pmatrix} + B(z)\,, \tag{1.2}$$

where $A(z)$ is an $n \times n$ matrix and $B(z)$ is an n-dimensional vector, both having entries that are rational functions with algebraic coefficients. Though Cobham's conjecture is proved in [1] by means of a completely different approach, it still remained a challenging problem to complete the proof he envisaged. In this direction, a great deal of work has been done by Loxton and van der Poorten [25,26] and a particular attention was then paid to systems of functional equations as in (1.2) (see for instance [9,32,33,35,38]). Very recently, another proof of Cobham's conjecture using Mahler's method is finally obtained in [4,38], thus solving a long-standing problem in Mahler's method.

Let K be a field. We observe that a power series $F(z) \in K[[z]]$ is a component of a vector satisfying a system of functional equations of the form (1.2)[2] if and only

[1] A remarkable discovery of Denis (see [13]), which deserves to be better understood, is that Mahler's method can be also applied to prove transcendence and algebraic independence results involving periods of t-modules which are variants of the more classical periods of abelian varieties, in the framework of the arithmetic of function fields of positive characteristic. For a detailed discussion on this topic, we refer the reader to the recent survey by Pellarin [37], see also [36].

[2] We assume here that the entries of $A(z)$ and $B(z)$ are in $K(z)$.

if the family

$$1, F(z), F\left(z^k\right), F\left(z^{k^2}\right), \dots$$

is linearly dependent over the field $K(z)$, that is, if there exist a natural number n and polynomials $Q(z), P_0(z), \dots, P_n(z) \in K[z]$, not all of which are zero, such that

$$Q(z) + \sum_{i=0}^{n} P_i(z) F\left(z^{k^i}\right) = 0. \tag{1.3}$$

Following Loxton and van der Poorten [26], we use the following definition.

Definition 1.1. Let K be a field and $k \geq 2$ be an integer. A power series $F(z) \in K[[z]]$ is a *k-Mahler function*, or for short is *k-Mahler*, if it satisfies a functional equation of the form (1.3).

Beyond transcendence, Mahler's method and automata theory, it is worth mentioning that Mahler functions naturally occur as generating functions in various other topics such as combinatorics of partitions, numeration and the analysis of algorithms (see [15] and the references therein and also dozens of examples in [7,8] and [19, Chapter 14]). A specially intriguing appearance of Mahler functions is related to the study of Siegel G-functions and in particular of diagonals of rational functions[3]. Though no general result confirms this claim, one observes that many generating series associated with the p-adic valuation of the coefficients of G-functions with rational coefficients turn out to be p-Mahler functions.

As a simple illustration, we give the following example. Let us consider the algebraic function

$$\mathfrak{f}(z) := \frac{1}{(1-z)\sqrt{1-4z}} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{2k}{k} z^n .$$

Note that $\mathfrak{f}$ is a G-function which satisfies the following minimal differential equation:

$$\mathfrak{f}'(z) = \frac{(3-6z)}{(1-z)(1-4z)} \mathfrak{f}(z) .$$

Let us define the sequence

$$a(n) := \nu_3 \left(\sum_{k=0}^{n} \binom{2k}{k} \right) ,$$

where ν_3 denotes the 3-adic valuation. We claim that the function

$$\mathfrak{f}_1(z) := \sum_{n \geq 0} a(n) z^n \in \mathbb{Q}[[z]]$$

[3] See for instance [3] for a discussion of the links between diagonals of rational functions with algebraic coefficients and G-functions.

is a 3-Mahler function[4]. This actually comes from the following nice equality

$$\nu_3\left(\sum_{k=0}^{n}\binom{2k}{k}\right) = \nu_3\left(n^2\binom{2n}{n}\right), \tag{1.4}$$

independently proved by Allouche and Shallit in 1989 (unpublished) and by Zagier [45]. Indeed, setting $\mathfrak{f}_2(z) := \sum_{n\geq 0} a(3n)z^n$ and $\mathfrak{f}_3(z) := \sum_{n\geq 0} a(3n+1)z^n$, we infer from Equality (1.4) that

$$\begin{pmatrix} \mathfrak{f}_1(z^3) \\ \mathfrak{f}_2(z^3) \\ \mathfrak{f}_3(z^3) \end{pmatrix} = A(z)\begin{pmatrix} \mathfrak{f}_1(z) \\ \mathfrak{f}_2(z) \\ \mathfrak{f}_3(z) \end{pmatrix} + B(z),$$

with

$$A(z) := \frac{1}{z^3(1+z+z^2)}\begin{pmatrix} z(1+z+z^2) & -z^2 & -z \\ 0 & z^2(1+z) & -z^4 \\ 0 & -z^2 & z^2(1+z) \end{pmatrix}$$

and

$$B(z) := \frac{1}{z^3(1+z+z^2)}\begin{pmatrix} \dfrac{z(2z^2-1)}{z-1} \\ -\dfrac{z^4}{z-1} \\ \dfrac{z^2(1+z)}{z-1} \end{pmatrix}.$$

A simple computation then gives the relation

$$a_0(z) + a_1(z)\mathfrak{f}_1(z) + a_2(z)\mathfrak{f}_1\big(z^3\big) + a_3(z)\mathfrak{f}_1\big(z^9\big) + a_4(z)\mathfrak{f}_1\big(z^{27}\big) = 0,$$

where

$$\begin{aligned}
a_0(z) &:= z + 2z^2 - z^3 + z^4 + 3z^5 - z^7 + 3z^8 + z^9 - z^{11} + 3z^{12} - 2z^{14} \\
&\quad -z^{15} + 2z^{16} - 2z^{17} - 2z^{18} + 2z^{21}, \\
a_1(z) &:= -1 - z^4 - z^8 + z^9 + z^{13} + z^{17}, \\
a_2(z) &:= 1 + z + z^2 + z^3 + z^4 + z^5 + z^6 + z^7 + z^8 - z^{13} - z^{14} - z^{15} - z^{16} \\
&\quad -z^{17} - z^{18} - z^{19} - z^{20} - z^{21}, \\
a_3(z) &:= -z^3 - z^6 - z^7 - z^9 - z^{10} - z^{11} - z^{13} - z^{14} + z^{16} - z^{17} + z^{19} \\
&\quad +z^{20} + z^{22} + z^{23} + z^{24} + z^{26} + z^{27} + z^{30}, \\
a_4(z) &:= z^{21} - z^{48}.
\end{aligned}$$

[4] It would be interesting to know the set of primes p for which $\sum_{n\geq 0}\nu_p\left(\sum_{k=0}^{n}\binom{2k}{k}\right)z^n$ is a p-Mahler function.

Of course, one could produce similar examples associated with transcendental G-functions by considering the Hadamard product (denoted by $\odot$ below) of several well-chosen algebraic functions. For instance, the elliptic integral

$$\mathfrak{g}(z) := \frac{2}{\pi}\int_0^{\pi/2} \frac{d\theta}{\sqrt{1-16z\sin^2\theta}} = \frac{1}{\sqrt{1-4z}} \odot \frac{1}{\sqrt{1-4z}} = \sum_{n=0}^{\infty} \binom{2n}{n}^2 z^n$$

is a transcendental G-function which satisfies the following minimal differential equation

$$\left(z^2-16z^3\right)\mathfrak{g}''(z) + \left(z-32z^2\right)\mathfrak{g}'(z) - 4z\mathfrak{g}(z) = 0\,,$$

and it is not hard to see that, for every prime p,

$$\mathfrak{g}_p(z) := \sum_{n=0}^{\infty} \nu_p\left(\binom{2n}{n}^2\right) z^n$$

is a p-Mahler function. More precisely, one can show that $\mathfrak{g}_p$ satisfies a relation of the form

$$a_0(z) + a_1(z)\mathfrak{g}_p(z) + a_2(z)\mathfrak{g}_p(z^p) + a_3(z)\mathfrak{g}_p\left(z^{p^2}\right) + a_4(z)\mathfrak{g}_p\left(z^{p^3}\right) = 0\,,$$

where the $a_i(z)$ are polynomials of degree $\mathrm{O}(p^3)$ too long to be reproduced here.

Regarding (1.1), (1.2) or (1.3), it is tempting to ask about the significance of the integer parameter k. Already in 1976, van der Poorten [40] suggested that two solutions of Mahler-type functional equations associated with essentially distinct parameters should be completely different. For instance, one may naturally expect [40] (and it is now proved [34]) that the two functions

$$\sum_{n=0}^{\infty} z^{2^n} \quad\text{and}\quad \sum_{n=0}^{\infty} z^{3^n}$$

are algebraically independent over $\mathbb{C}(z)$. This idea was later formalized by Loxton and van der Poorten who made a general conjecture whose one-dimensional version can be stated as follows[5].

Conjecture 1.2 (Loxton and van der Poorten). Let k and l be two multiplicatively independent positive integers and L be a number field. Let $F(z) \in L[[z]]$ be a locally analytic function that is both k- and ℓ-Mahler. Then $F(z)$ is a rational function.

[5] Note that in fact this conjecture does not imply any statement concerning algebraic independence. It does, however, cover linear independence. Indeed, say that $F(z)$ and $G(z)$ are irrational power series such that F is 2-Mahler and G is 3-Mahler, then 1, F and G are linearly independent over $\mathbb{C}(z)$ (otherwise F is at once 2- and 3-Mahler, and thus rational).

We recall that two integers k and l larger than 1 are multiplicatively independent if there is no pair of positive integers (n, m) such that $k^n = \ell^m$, or equivalently, if $\log(k)/\log(\ell) \notin \mathbb{Q}$. Conjecture 1.2 first appeared in a 1987 paper of van der Poorten [41]. Since then it was explicitly studied in a number of different contexts including some papers of Loxton [24], Becker [9], Randé [42], Bell [10] and the monograph of Everest *et al.* [19]. Independently, Zannier also considered a similar question in [46].

In this paper, our aim is to prove the following result, which has been proven independently by Schäfke and Singer [39].

Theorem 1.3. *Let K be a field of characteristic zero and let k and l be two multiplicatively independent positive integers. Then a power series $F(z) \in K[[z]]$ is both k- and ℓ-Mahler if and only if it is a rational function.*

Let us make few comments.

- Taking K to be a number field in Theorem 1.3 gives Conjecture 1.2;
- If k and ℓ denote two multiplicatively dependent natural numbers, then it is easy to see that a power series is k-Mahler if and only if it is also ℓ-Mahler (see Remark 8.2);
- As explained in more detail in Section 2, one motivation for proving Theorem 1.3 is that it provides a far-reaching generalization of one fundamental result in the theory of sets of integers recognizable by finite automata: Cobham's theorem. Loxton and van der Poorten [24, 41] actually guessed that Conjecture 1.2 should be a consequence of some algebraic independence results for Mahler functions of several variables. In particular, they hoped to obtain a totally new proof of Cobham's theorem by using Mahler's method. Note, however, that our proof of Theorem 1.3 follows a totally different way and ultimately relies on Cobham's theorem, so we do not obtain an independent derivation of that result;
- Another important motivation for establishing Theorem 1.3 comes from the fact that this kind of statements, though highly natural and somewhat ubiquitous, are usually very difficult to prove. In particular, similar independence phenomena, involving two multiplicatively independent integers, are expected in various contexts but only very few results have been obtained up to now. As an illustration, we cite below three interesting open problems that rest on such a principle, all of them being widely open[6]. A long-standing question in dynamical systems is the so-called $\times 2\ \times 3$ problem addressed by Furstenberg [20]: prove that the only Borel measures on $[0, 1]$ that are simultaneously ergodic for $T_2(x) = 2x \pmod 1$ and $T_3(x) = 3x \pmod 1$ are the Lebesgue measure and measures supported by those orbits that are periodic for both actions T_2 and T_3. The following problem, sometimes attributed to Mahler, was suggested by Mendès France

[6] In all of these problems, the integers 2 and 3 may of course be replaced by any two multiplicatively independent integers larger than 1.

in [31] (see also [2]): given a binary sequence $(a_n)_{n\geq 0} \in \{0,1\}^{\mathbb{N}}$, prove that

$$\sum_{n=0}^{\infty} \frac{a_n}{2^n} \quad \text{and} \quad \sum_{n=0}^{\infty} \frac{a_n}{3^n}$$

are both algebraic numbers only if both are rational numbers. The third problem we mention appeared implicitly in work of Ramanujan (see [44]): prove that both 2^x and 3^x are integers only if x is a natural number. This is a particular instance of the four exponentials conjecture, a famous open problem in transcendence theory [43, Chapter 1, page 15].

The outline of the paper is as follows. In Section 2, we briefly discuss the connection between Theorem 1.3 and Cobham's theorem. In Section 3, we describe our strategy for proving Theorem 1.3. Then the remaining Sections 4-11 are devoted to the different steps of the proof of Theorem 1.3. Throughout this paper, k and l will denote integers larger than or equal to 2.

ACKNOWLEDGEMENTS. The authors would like to thank Michel Mendès France for his comments and encouragements, as well as the anonymous referee for his useful suggestions and careful reading. The first author is indebted to Éric Delaygue for his help with Maple. He is also most grateful to Macha and Vadim for inspiring discussions.

2. Connection with finite automata and Cobham's theorem

One motivation for proving Theorem 1.3 is that it provides a far-reaching generalization of a fundamental result in the theory of sets of integers recognizable by finite automata. The aim of this section is to briefly describe this connection. For more details and formal definitions on automatic sets and automatic sequences, we refer the reader to the book of Allouche and Shallit [6].

Let $k \geq 2$ be a natural number. A set $\mathcal{N} \subset \mathbb{N}$ is said to be *k-automatic* if there is a finite automaton (more formally a k-deterministic finite automaton) that accepts as input the expansion of n in base k and outputs 1 if $n \in \mathcal{N}$ and 0 otherwise. For example, the set of Thue-Morse integers $1, 2, 4, 7, 8, 11, 13, \ldots$, formed by the integers whose sum of binary digits is odd, is 2-automatic. The associated automaton is given in Figure 1 below. It has two states. This automaton successively reads the binary digits of n (starting, say, from the most significant digit and the initial state q_0) and thus ends the reading either in state q_0 or in state q_1. The initial state q_0 gives the output 0, while q_1 gives the output 1.

Another typical 2-automatic set of integers is given by the powers of 2: $1, 2, 4, 8, 16, \ldots$. Though these integers have very simple expansions in base 2, one can observe that this is not the case when writing them in base 3. One of the most important results in the theory of automatic sets formalizes this idea. It says

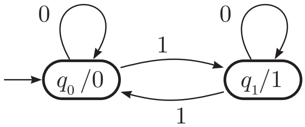

Figure 2.1. The finite-state automaton recognizing the set of Thue-Morse integers.

that only very well-behaved sets of integers can be automatic with respect to two multiplicatively independent numbers. Indeed, in 1969 Cobham [12] proved the following result.

Theorem 2.1 (Cobham). *Let k and ℓ be two multiplicatively independent integers. Then a set $\mathcal{N} \subseteq \mathbb{N}$ is both k- and ℓ-automatic if and only if it is the union of a finite set and a finite number of arithmetic progressions.*

The proof given by Cobham of his theorem is elementary but notoriously difficult, and it remains a challenging problem to find a more natural/conceptual proof (see for instance the comment in Eilenberg [17, page 118]). There are many interesting generalizations of this result. A very recent one is due to Durand [16] and we refer the reader to the introduction of [16] for a brief but complete discussion about such generalizations.

To conclude this section, let us briefly explain why Cobham's Theorem is a consequence of Theorem 1.3. Let us assume that $\mathcal{N} \subseteq \mathbb{N}$ is k-automatic. Set $F(x) := \sum_{n\in\mathcal{N}} x^n \in \mathbb{Z}[[x]]$. Then it is known that $F(x)$ is k-Mahler (see for instance [19, page 232]). In addition, let us assume that $\mathcal{N}$ is also ℓ-automatic where k and ℓ are multiplicatively independent. Then by Theorem 1.3, it follows that $F(x)$ is a rational function and thus the sequence of coefficients of $F(x)$ satisfies a linear recurrence. Since the coefficients of $F(x)$ take only two distinct values (0 and 1), we see that this linear recurrence is ultimately periodic. This exactly means that $\mathcal{N}$ is the union of a finite set and a finite number of arithmetic progressions, as claimed by Cobham's theorem.

3. Sketch of proof of Theorem 1.3

In this section we describe the main steps of the proof of Theorem 1.3.

Let R be a ring and $\mathfrak{P}$ be an ideal of R. If $F(x) = \sum_{n=0}^{\infty} f(n)x^n \in R[[x]]$, then we denote by $F_{\mathfrak{P}}(x)$ the reduction of $F(x)$ modulo $\mathfrak{P}$, that is

$$F_{\mathfrak{P}}(x) = \sum_{n=0}^{\infty} (f(n) \bmod \mathfrak{P})x^n \in (R/\mathfrak{P})[[x]] .$$

Let K be a field of characteristic zero and $F(x) \in K[[x]]$ be both k- and ℓ-Mahler.

Step 0. This is a preliminary step. In the introduction, we defined Mahler functions as those satisfying Equation (1.3) but it is not always convenient to work with this general form of equations. In Sections 4 and 6 we show that there is no loss of generality to work with some more restricted types of functional equations. Also in Section 8, we prove that one can assume without loss of generality some additional assumptions on k and ℓ; namely that there are primes p and q such that p divides k but does not divide ℓ and q divides ℓ but does not divide k.

Step 1. A first observation, proved in Section 5, is that the coefficients of the formal power series $F(x)$ only belong to some finitely generated $\mathbb{Z}$-algebra $R \subseteq K$. Then we prove the following useful local-global principle: $F(x)$ is a rational function if it has rational reduction modulo a sufficiently large set of maximal ideals of R. Using classical results of commutative algebra about Jacobson rings, we derive from our local-global principle that there is no loss of generality to assume that K is a number field and that R is a localization of the ring of integers of K formed by inverting a positive integer (that is, R is of the form $\mathcal{O}_K[1/N]$ for some positive integer N).

Comment. Our strategy consists now in applying again our local-global principle. Indeed, since R is of the form $\mathcal{O}_K[1/N]$, we have that the quotient ring $R/\mathfrak{P}$ is a finite field for every prime ideal $\mathfrak{P}$ of R. Our plan is thus to exploit the fact that $F_{\mathfrak{P}}(x)$ has coefficients in the finite set $R/\mathfrak{P}$ to prove that $F_{\mathfrak{P}}(x)$ is both a k- and an ℓ-automatic power series (see Section 7 for a definition), for some prime ideals $\mathfrak{P}$. If this is the case, then Cobham's theorem applies and we get that $F_{\mathfrak{P}}(x)$ is a rational function. The local-global principle actually implies that it is enough to prove that $F_{\mathfrak{P}}(x)$ is both k- and ℓ-automatic for infinitely many prime ideals $\mathfrak{P}$ of R.

Step 2. In Section 7, we underline the relation between k-Mahler, k-regular, and k-automatic power series. The latter two notions are defined in that section. In particular, we will use a result of Dumas [14] showing that every k-Mahler power series can be decomposed as

$$F(x) = G(x) \cdot \Pi(x),$$

where $G(x) \in R[[x]]$ is a k-regular power series and $\Pi(x) \in R[[x]]$ is the inverse of an infinite product of polynomials. Since $F(x)$ is also ℓ-Mahler, we also have a similar decomposition

$$F(x) = H(x) \cdot \Pi'(x),$$

where $H(x) \in R[[x]]$ is a ℓ-regular power series and $\Pi'(x) \in R[[x]]$ is the inverse of an infinite product of polynomials. Furthermore, the theory of regular power series implies that $G_{\mathfrak{P}}(x)$ is k-automatic and that $H_{\mathfrak{P}}(x)$ is ℓ-automatic for every prime ideal $\mathfrak{P}$ of R.

In Section 13 we will split both infinite products $\Pi(x)$ and $\Pi'(x)$ and get an expression of the form

$$F(x) = G(x) \cdot \Pi_1(x) \cdot \Pi_2(x) = H(x) \cdot \Pi'_1(x) \cdot \Pi'_2(x)$$

where $\Pi_1(x)$, $\Pi_2(x)$, $\Pi'_1(x)$, $\Pi'_2(x) \in R[[x]]$ are inverses of some other infinite products of polynomials.

Step 3. After proving preliminary results in Sections 9 and 10, we look at the singularities of Mahler functions at roots of unity in Section 11. We use asymptotic techniques to show that one can reduce to the case of considering Mahler equations whose singularities at roots of unity have a restricted form. This ensures, using some results of Section 7, that $\Pi_1(x)$ is k-automatic and that $\Pi'_1(x)$ is ℓ-automatic when reduced modulo every prime ideal $\mathfrak{P}$ of R.

Step 4. In our last step, we use Chebotarev's density theorem in order to ensure the existence of an infinite set $\mathcal{S}$ of prime ideals of R such that $\Pi_2(x)$ is k-automatic and $\Pi'_2(x)$ is ℓ-automatic when reduced modulo every ideal $\mathfrak{P} \in \mathcal{S}$.

Conclusion. Since the product of k-automatic power series is k-automatic, we infer from Steps 2, 3 and 4 that for every prime ideals $\mathfrak{P} \in \mathcal{S}$ the power series $F_{\mathfrak{P}}(x)$ is both k- and ℓ-automatic. By Cobham's theorem, $F_{\mathfrak{P}}(x)$ is rational for every such prime ideal. Then the local-global principle ensures that $F(x)$ is rational, as desired.

4. Preliminary reduction for the form of Mahler equations

In the introduction, we defined k-Mahler functions as power series satisfying a functional equation of the form given in (1.3). In the literature, they are sometimes defined as solutions of a more restricted type of functional equations. We recall here that these apparently stronger conditions on the functional equations actually lead to the same class of functions. In the sequel, it will thus be possible to work without loss of generality with these more restricted type of equations.

Lemma 4.1. *Let us assume that $F(x)$ satisfies a k-Mahler equation as in* (1.3). *Then there exist polynomials $P_0(x), \dots, P_n(x)$ in $K[x]$, with* $\gcd(P_0(x), \dots, P_n(x)) = 1$ *and $P_0(x)P_n(x) \neq 0$, and such that*

$$\sum_{i=0}^{n} P_i(x) F\left(x^{k^i}\right) = 0 . \tag{4.1}$$

Proof. Let us assume that $F(x)$ satisfies a k-Mahler equation as in (1.3). There thus exist some nonnegative integer n and polynomials $A(x)$, $A_0(x), \dots, A_n(x)$ in $K[x]$, with $A_n(x)$ nonzero, such that

$$\sum_{i=0}^{n} A_i(x) F\left(x^{k^i}\right) = A(x) .$$

We first show that we can assume that $A(x) = 0$. Indeed, let us assume that $A(x) \neq 0$. Applying the operator $x \mapsto x^k$ to this equation, we get that

$$\sum_{i=0}^{n} A_i\left(x^k\right) F\left(x^{k^{i+1}}\right) = A\left(x^k\right) .$$

Multiplying the first equation by $A(x^k)$ and the second by $A(x)$ and subtracting, we obtain the new equation

$$\sum_{i=0}^{n+1} B_i(x) F\left(x^{k^i}\right) = 0\,,$$

where $B_i(x) := A_i(x)A(x^k) - A_i(x^k)A(x)$ for every integer i, $1 \leq i \leq n$ and where $B_{n+1}(x) := A_n(x^k)A(x) \neq 0$. We can thus assume without loss of generality that $A(x) = 0$.

Now, among all such nontrivial relations of the form

$$\sum_{i=0}^{n} P_i(x) F\left(x^{k^i}\right) = 0\,, \tag{4.2}$$

we choose one with n minimal. Thus $P_n(x)$ is nonzero. We claim $P_0(x)$ is nonzero. Let us assume this is not the case. Pick the smallest integer j such that $P_j(x)$ is nonzero. By assumption, $j > 0$. Then there is some nonnegative integer a such that the coefficient of x^a in $P_j(x)$ is nonzero. Let b be the unique integer such that $a \equiv b \bmod k$ and $0 \leq b < k$. Let us define the operator Λ_b from $K[[x]]$ into itself by

$$\Lambda_b\left(\sum_{i=0}^{\infty} f(i)x^i\right) := \sum_{i=0}^{\infty} f(ki+b)x^i\,.$$

These types of operators are classically used for studying algebraic power series over fields of characteristic $p > 0$, where one takes $k = p$ (see for instance [6, Chapter 12] and the references therein). In this context, these operators are often referred to as Cartier operators. With this definition, every $F(x) \in K[[x]]$ has a unique decomposition as

$$F(x) = \sum_{b=0}^{k-1} x^b \Lambda_b(F)\big(x^k\big)\,,$$

which implies that

$$\Lambda_b\left(F(x)G\big(x^k\big)\right) = \Lambda_b\left(F(x)\right)G(x)$$

for every pair of power series $F(x), G(x) \in K[[x]]$. Applying Λ_b to Equation (4.2), we thus get that

$$0 = \Lambda_b\left(\sum_{i=j}^{n} P_i(x) F\left(x^{k^i}\right)\right) = \sum_{i=j-1}^{n-1} \Lambda_b\left(P_{i+1}(x)\right) F\left(x^{k^i}\right).$$

By construction, $\Lambda_b(P_j(x))$ is nonzero, which shows that this relation is nontrivial. This contradicts the minimality of n. It follows that $P_0(x)$ is nonzero.

Furthermore, if $\gcd(P_0(x), \dots, P_n(x)) = D(x) \neq 0$, it suffices to divide (4.2) by $D(x)$ to obtain an equation with the desired properties. This ends the proof. □

5. Reduction to the number field case

In this section we show that we may restrict our attention to the case where the base field K is replaced by a number field and more precisely by a localization of the ring of integers of that number field formed by inverting a single integer. This means a ring of the form $\mathcal{O}_K[1/N]$, where K denotes a number field, $\mathcal{O}_K$ the ring of integers of K, and N a positive integer.

Theorem 5.1. *Let us assume that the conclusion of Theorem* 1.3 *holds whenever the field K is replaced by a localization of the ring of integers of a number field of the form $\mathcal{O}_K[1/N]$. Then Theorem* 1.3 *is true.*

We first observe that the coefficients of a Mahler function in $K[[x]]$ actually belong to some finitely generated $\mathbb{Z}$-algebra $R \subseteq K$.

Lemma 5.2. *Let K be a field of characteristic zero, let $k \geq 2$ be an integer, and let $F(x) \in K[[x]]$ be a k-Mahler power series. Then there exists a finitely generated $\mathbb{Z}$-algebra $R \subseteq K$ such that $F(x) \in R[[x]]$.*

Proof. Let $F(x) := \sum_{n=0}^{\infty} f(n)x^n \in K[[x]]$ be a k-Mahler power series. We first infer from Lemma 4.1 that there exist a natural number n and polynomials $P_0(x), \dots, P_n(x) \in K[x]$ with $P_0(x)P_n(x) \neq 0$ such that

$$\sum_{i=0}^{n} P_i(x)F\left(x^{k^i}\right) = 0 .$$

Let d be a natural number that is strictly greater than the degrees of the polynomials $P_0(x), \dots, P_n(x)$. Let R denote the smallest $\mathbb{Z}$-algebra containing:

- The coefficients of $P_0(x), \dots, P_n(x)$;
- The coefficients $f(0), \dots, f(d)$;
- The multiplicative inverses of all nonzero coefficients of $P_0(x)$.

By definition, $R \subseteq K$ is a finitely generated $\mathbb{Z}$-algebra. We claim that $F(x) \in R[[x]]$. To see this, suppose that this is not the case. Let n_0 be the smallest nonnegative integer such that $f(n_0) \notin R$. By assumption, $n_0 > d$. Consider the equation

$$P_0(x)F(x) = -\sum_{i=1}^{n} P_i(x)F\left(x^{k^i}\right) . \tag{5.1}$$

Let j denote the order of $P_0(x)$ at $x = 0$ and let $c \neq 0$ denote the coefficient of x^j in $P_0(x)$. Then if we extract the coefficient of x^{n_0+j} in Equation (5.1), we see that $cf(n_0)$ can be expressed as an R-linear combination of $f(0), \dots, f(n_0 - 1)$. Hence $cf(n_0)$ belongs to R by the minimality of n_0. Since $c^{-1} \in R$ we see that $f(n_0) \in R$, a contradiction. This ends the proof. □

We now prove that the height of a rational function which satisfies a Mahler-type equation can be bounded by the maximum of the degrees of the polynomials defining the underlying equation.

Lemma 5.3. *Let K be a field, let n and d be natural numbers, and let $P_0(x), \ldots, P_n(x)$ be polynomials in $K[x]$ of degree at most d with $P_0(x)P_n(x) \neq 0$. Suppose that $F(x) \in K[[x]]$ satisfies the Mahler-type equation*

$$\sum_{i=0}^{n} P_i(x)F(x^{k^i}) = 0\,.$$

If $F(x)$ is rational, then there exist polynomials $A(x)$ and $B(x)$ of degree at most d with $B(0) = 1$ such that $F(x)$ is the power series expansion of $A(x)/B(x)$.

Proof. Without any loss of generality we can assume that $F(x)$ is not identically zero. If $F(x)$ is rational, then there exist two polynomials $A(x)$ and $B(x)$ in $K[x]$ with gcd 1 and with $B(0) = 1$ such that $F(x) = A(x)/B(x)$. Observe that

$$\sum_{i=0}^{n} P_i(x)A\left(x^{k^i}\right)/B\left(x^{k^i}\right) = 0\,.$$

Multiplying both sides of this equation by the product $B(x)B(x^k)\cdots B(x^{k^n})$, we see that $B(x^{k^n})$ divides

$$P_n(x)A\left(x^{k^n}\right)B(x)\cdots B\left(x^{k^{n-1}}\right).$$

Since $\gcd(A(x), B(x)) = 1$ and $A(x)$ is nonzero, we actually have that $B(x^{k^n})$ divides

$$P_n(x)B(x)\cdots B\left(x^{k^{n-1}}\right).$$

Let d_0 denote the degree of $B(x)$. Then we have

$$\begin{aligned} k^n d_0 &\leq \deg(P_n(x)) + \sum_{i=0}^{n-1} \deg\left(B\left(x^{k^i}\right)\right) \\ &\leq d + d_0\left(1 + k + \cdots + k^{n-1}\right) \\ &= d + d_0(k^n - 1)/(k-1)\,. \end{aligned}$$

Thus

$$d_0(k^{n+1} - 2k^n + 1)/(k-1) \leq d\,,$$

which implies $d_0 \le d$ since $(k^{n+1} - 2k^n + 1)/(k-1) \ge 1$ for every integer $k \ge 2$. A similar argument gives the same upper bound for the degree of $A(x)$. ☐

We derive from Lemma 5.3 a useful local-global principle for the rationality of Mahler functions with coefficients in a finitely generated $\mathbb{Z}$-algebra.

Lemma 5.4. *Let K be a field, let $k \ge 2$ be an integer, and let $R \subseteq K$ be a ring. Let us assume that $F(x) \in R[[x]]$ has the following properties.*

(i) *There exist a natural number n and polynomials $P_0(x), \dots, P_n(x) \in R[x]$ with $P_0(x)P_n(x) \ne 0$ such that*

$$\sum_{i=0}^{n} P_i(x) F\left(x^{k^i}\right) = 0\,;$$

(ii) *There exists a set $\mathcal{S}$ of maximal ideals of R such that $F(x)$ mod I is a rational power series in $(R/I)[[x]]$ for every $I \in \mathcal{S}$;*

(iii) *One has $\bigcap_{I \in \mathcal{S}} I = \{0\}$.*

Then $F(x)$ is a rational function.

Proof. Let d be a natural number greater than the degrees of all polynomials $P_0(x), \dots, P_n(x)$. By (ii), we have that for each maximal ideal I in $\mathcal{S}$, $F(x)$ mod I is a rational function. Thus by (i) and Lemma 5.3, we see that for each maximal ideal I in $\mathcal{S}$, there exist two polynomials $A_I(x)$ and $B_I(x) \in (R/I)[x]$ of degree at most d with $B_I(0) = 1$ and such that $F(x) \equiv A_I(x)/B_I(x) \mod I$. In particular, if $F(x) = \sum_{j \ge 0} f(j)x^j$, we see that the sequences in the set $\left\{(f(d+1+i+j) \mod I)_{j\ge 0} \mid i = 0, \dots, d\right\}$ are linearly dependent over R/I. Thus the determinant of each $(d+1) \times (d+1)$ submatrix of the infinite matrix

$$M := \begin{pmatrix} f(d+1) & f(d+2) & f(d+3) & \cdots \\ f(d+2) & f(d+3) & f(d+4) & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ f(2d+1) & f(2d+2) & f(2d+3) & \cdots \end{pmatrix}$$

lies in the maximal ideal I. Since this holds for every maximal ideal I in $\mathcal{S}$, we infer from (iii) that every $(d+1) \times (d+1)$ minor of M vanishes. It follows that M has rank at most d and thus the rows of M are linearly dependent over the field of fractions of R. In particular, there exist $c_0, \dots, c_d \in R$, not all zero, such that

$$\sum_{i=0}^{d} c_i f(d+1+i+j) = 0$$

for all $j \geq 0$. Letting $B(x) := c_d + c_{d-1}x + \cdots + c_0 x^d$, we see that $B(x)F(x)$ is a polynomial. Hence $F(x)$ is a rational function. This ends the proof. □

We are now ready to prove the main result of this section.

Proof of Theorem 5.1. Let K be a field of characteristic zero and let $F(x) \in K[[x]]$ be a power series that is both k- and ℓ-Mahler for some multiplicatively independent natural numbers k and ℓ. By Lemma 4.1, there are natural numbers n and m and polynomials $P_0(x), \ldots, P_n(x)$ and $Q_0(x), \ldots, Q_m(x)$ with

$$P_0(x)P_n(x)Q_0(x)Q_m(x) \neq 0$$

and such that

$$\sum_{i=0}^{n} P_i(x)F\left(x^{k^i}\right) = \sum_{j=0}^{m} Q_j(x)F\left(x^{\ell^j}\right) = 0\,. \tag{5.2}$$

Then by Lemma 5.2, there is a finitely generated $\mathbb{Z}$-algebra $R \subseteq K$ such that $F(x) \in R[[x]]$. By adjoining all the coefficients of $P_0(x), \ldots, P_n(x)$ and of $Q_0(x), \ldots, Q_m(x)$ to R, we can assume that $P_i(x)$ and $Q_j(x)$ are in $R[x]$ for $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$. By localizing at the multiplicatively closed set consisting of nonzero integers in R, we can assume that R is a finitely generated $\mathbb{Q}$-algebra.

Let $\mathcal{M} \subseteq \operatorname{Spec}(R)$ denote the collection of maximal ideals of R. Since R is a finitely generated $\mathbb{Q}$-algebra, R is a Jacobson ring and R/I is a finite extension of $\mathbb{Q}$ for every $I \in \mathcal{M}$ (see [18, Theorem 4.19, page 132]). Thus, for each maximal ideal I of R, the quotient field R/I is a number field. If we assume that the conclusion of Theorem 1.3 holds when the base field is a number field, then we get that $F(x)$ mod I is a rational function in $(R/I)[[x]]$ for it is clearly both k- and ℓ-Mahler[7]. Since R is a Jacobson ring that is also a domain, we have that $\bigcap_{I \in \mathcal{M}} I = \{0\}$ (*cf.* [18, page 132]). Then Lemma 5.4 implies that $F(x)$ is a rational function in $R[[x]]$. This shows it is sufficient to prove Theorem 1.3 in the case that K is a number field.

We can thus assume that $F(x) \in K[[x]]$ where K is a number field. Now, if we apply again Lemma 5.2, we see that there is a finitely generated $\mathbb{Z}$-algebra $R \subseteq K$ such that $F(x) \in R[[x]]$. Furthermore, every finitely generated $\mathbb{Z}$-subalgebra of a number field K has a generating set of the form $\{a_1/b, \ldots, a_t/b\}$, where b is a nonzero (rational) integer and $a_1, \ldots, a_t$ are algebraic integers in K. Thus R is a subalgebra of a localization of the ring of integers of a number field formed by inverting a single nonzero integer, that is $R \subseteq \mathcal{O}_K[1/b]$, where $\mathcal{O}_K$ denotes the ring of algebraic integers in K. Thus to establish Theorem 1.3 it is sufficient to prove the following result: let k and ℓ be two multiplicatively independent natural numbers,

[7] Note that since $P_0(0)Q_0(0) \neq 0$, we may assume that $P_0(0) = Q_0(0) = 1$ by multiplying the left side of (5.2) by $1/P_0(0)$ and the right side of (5.2) by $1/Q_0(0)$. This ensures that, for each functional equation, not all the coefficients vanish when reduced modulo a maximal ideal I of R. Hence $F(x)$ mod I is both k- and ℓ-Mahler.

let R be of the form $\mathcal{O}_K[1/b]$ where K is a number field, and let $F(x) \in R[[x]]$, then if $F(x)$ is both k- and ℓ-Mahler it is a rational function. This concludes the proof. □

6. Further reductions for the form of Mahler equations

In this section we refine the results of Section 4. We show that a power series satisfying a Mahler equation of the form given in (4.1) is also solution of a more restricted type of functional equations.

Lemma 6.1. *Let K be a field and $k \geq 2$ be an integer. Let us assume that $F(x) := \sum_{s\geq 0} f(s)x^s \in K[[x]]$ satisfies a k-Mahler equation of the form*

$$\sum_{i=0}^{n} P_i(x)F\left(x^{k^i}\right) = 0\,,$$

where $P_0(x), \dots, P_n(x) \in K[x]$, $\gcd(P_0(x), \dots, P_n(x)) = 1$ and $P_0(x)P_n(x) \neq 0$. Then there exists a natural number N such that, for every integer $a > N$ with $f(a) \neq 0$, $F(x)$ can be decomposed as

$$F(x) = T_a(x) + x^a F_0(x)\,,$$

where $T_a(x) \in K[x]$ is the Taylor approximation of $F(x)$ at $x = 0$ up to degree $a - 1$ and $F_0(x)$ has nonzero constant term and satisfies a k-Mahler equation

$$\sum_{i=0}^{n+1} Q_i(x)F_0\left(x^{k^i}\right) = 0$$

for some polynomials $Q_0, \dots, Q_{n+1} \in K[x]$ satisfying the following conditions.

(i) *It holds $Q_0(0) = 1$;*
(ii) *If $\alpha \neq 0$ and $P_0(\alpha) = 0$, then $Q_0(\alpha) = 0$;*
(iii) *If $\alpha \neq 0$, $P_0(\alpha) = 0$ and $\alpha^k = \alpha$, then $Q_j(\alpha) \neq 0$ for some $j \in \{1, \dots, n+1\}$.*

Proof. By assumption, we have that $F(x)$ satisfies a k-Mahler equation

$$\sum_{i=0}^{n} P_i(x)F\left(x^{k^i}\right) = 0\,,$$

where $P_0(x)P_n(x)$ is nonzero. Let N denote the order of vanishing of $P_0(x)$ at $x = 0$. Suppose that $a \geq N$ and $f(a) \neq 0$. Then we have that

$$F(x) = T_a(x) + x^a F_0(x)\,,$$

where $T_a(x)$ is the Taylor approximation of $F(x)$ up to degree $a-1$ and $F_0(x)$ is a power series with nonzero constant term. Then we have

$$\sum_{i=0}^{n} P_i(x)\left(T_a\left(x^{k^i}\right)+x^{k^i\cdot a}F_0\left(x^{k^i}\right)\right)=0\,,$$

which we can write as

$$\sum_{i=0}^{n} P_i(x)x^{k^i\cdot a}F_0\left(x^{k^i}\right)=C(x)\,, \tag{6.1}$$

where $C(x)$ denotes the polynomial

$$C(x):=-\sum_{i=0}^{n} P_i(x)T_a\left(x^{k^i}\right).$$

Set $S(x):=P_0(x)x^{-N}$. By definition of N, $S(x)$ is a polynomial with $S(0)\neq 0$. Then if we divide both sides of Equation (6.1) by x^{a+N}, we obtain that

$$S(x)F_0(x)+\sum_{i=1}^{n} P_i(x)x^{k^i a-a-N}F_0\left(x^{k^i}\right)=x^{-a-N}C(x)\,. \tag{6.2}$$

Observe that the left side is a power series with constant term $S(0)F_0(0)\neq 0$ and thus $C_0(x):=x^{-a-N}C(x)$ is a polynomial with $C_0(0)\neq 0$. Applying the operator $x\mapsto x^k$, we also obtain that

$$S\left(x^k\right)F_0\left(x^k\right)+\sum_{i=1}^{n} P_i\left(x^k\right)x^{k^{i+1}a-ka-kN}F_0\left(x^{k^{i+1}}\right)=C_0\left(x^k\right). \tag{6.3}$$

Multiplying (6.2) by $C_0(x^k)$ and (6.3) by $C_0(x)$ and then subtracting, we get that

$$\begin{aligned}&C_0\left(x^k\right)S(x)F_0(x)+\sum_{i=1}^{n} C_0\left(x^k\right)P_i(x)x^{k^i a-a-N}F_0\left(x^{k^i}\right)\\&-C_0(x)S\left(x^k\right)F_0\left(x^k\right)-\sum_{i=1}^{n} C_0(x)P_i\left(x^k\right)x^{k^{i+1}a-ka-kN}F_0\left(x^{k^{i+1}}\right)=0\,.\end{aligned}$$

Since $C_0(0)$ and $S(0)$ are nonzero, we see that $F_0(x)$ satisfies a non-trivial k-Mahler equation

$$\sum_{i=0}^{n+1} Q_i(x)F_0\left(x^{k^i}\right)=0\,,$$

where

$$Q_0(x):=\frac{C_0\left(x^k\right)S(x)}{\gcd\left(C_0(x),\,C_0\left(x^k\right)\right)}$$

and

$$Q_1(x) := \frac{C_0\big(x^k\big)P_1(x)x^{ka-a-N} - C_0(x)S\big(x^k\big)}{\gcd\Big(C_0(x), C_0\big(x^k\big)\Big)},$$

and, for $i \in \{2, \dots, n+1\}$,

$$Q_i(x) := \frac{x^{k^i a-ka-N}\Big(C_0\big(x^k\big)x^{(k-1)a}P_i(x) - C_0(x)P_{i-1}\big(x^k\big)\Big)}{\gcd\Big(C_0(x), C_0\big(x^k\big)\Big)},$$

with the convention that $P_{n+1}(x) := 0$. By construction, $Q_0(0) \neq 0$, which we may assume to be equal to 1 by multiplying our equation by $1/Q_0(0)$. Since $S(x)$ divides $Q_0(x)$, we have that if $P_0(\alpha) = 0$ for some nonzero α then $Q_0(\alpha) = 0$. Finally, suppose that $P_0(\alpha) = 0$ for some nonzero α such that $\alpha^k = \alpha$. We claim that $Q_i(\alpha)$ is nonzero for some $i \in \{1, \dots, n+1\}$. Note that since $\gcd(P_0(x), \dots, P_n(x)) = 1$, there is smallest positive integer j such that $P_j(\alpha)$ is nonzero. We claim that $Q_j(\alpha) \neq 0$. Indeed, otherwise α would be a root of $C_0(x)/\gcd(C_0(x), C_0(x^k))$, but this is impossible since $\alpha^k = \alpha$. This ends the proof. □

Corollary 6.2. *Let K be a field and let k and ℓ be multiplicatively independent natural numbers. Let $F(x) := \sum_{s\geq 0} f(s)x^s \in K[[x]]$ be a power series that is both k- and ℓ-Mahler and that is not a polynomial. Then there is a natural number a such that $F(x)$ can be decomposed as*

$$F(x) = T_a(x) + x^a F_0(x),$$

where $T_a(x)$ is the Taylor approximation of $F(x)$ up to degree $a-1$, $F_0(x)$ satisfies a k-Mahler equation as in Lemma 6.1, *and $F_0(x)$ also satisfies an ℓ-Mahler equation of the form*

$$\sum_{i=0}^{r} R_i(x)F_0\Big(x^{\ell^i}\Big) = 0$$

with $R_0(x), \dots, R_r(x) \in K[x]$ and $R_0(0) = 1$.

Proof. Applying Lemma 6.1 to $F(x)$, viewed as a k-Mahler function, we obtain the existence of a positive integer N_1 (which corresponds to N in Lemma 6.1) for which the conclusion of this lemma holds. Similarly, applying Lemma 6.1 to $F(x)$, viewed as a ℓ-Mahler function, we obtain the existence of a positive integer N_2. Now, we can choose $N_3 := \max(N_1, N_2)$ and pick $a > N_3$ such that $f(a) \neq 0$ to obtain the desired conclusion. □

7. Links with automatic and regular power series

The aim of this section is to emphasize the relation between k-Mahler, k-regular, and k-automatic power series. We gather some useful facts about automatic and regular power series that will turn out to be useful for proving Theorem 1.3. We also recall a result of Dumas [14] showing that every k-Mahler power series can be decomposed as the product of a k-regular power series of a special type and the inverse of an infinite product of polynomials. Such a decomposition will play a key role in the proof of Theorem 1.3.

7.1. Automatic and regular power series

We recall here basic facts about regular power series, which were introduced by Allouche and Shallit [7] (see also [8] and [6, Chapter 16]). They form a distinguished class of k-Mahler power series as well as a natural generalization of k-automatic power series.

A useful way to characterize k-automatic sequences, due to Eilenberg [17], is given in terms of the so-called k-kernel.

Definition 7.1. Let $k \geq 2$ be an integer and let $\mathbf{f} = (f(n))_{n\geq 0}$ be a sequence with values in a set E. The *k-kernel* of $\mathbf{f}$ is defined as the set

$$\left\{(f(k^a n + b))_{n\geq 0} \mid a \geq 0, b \in \{0, \dots, k^a - 1\}\right\} .$$

Theorem 7.2 (Eilenberg). *A sequence is k-automatic if and only if its k-kernel is finite.*

This characterization gives rise to the following natural generalization of automatic sequences introduced by Allouche and Shallit [7].

Definition 7.3. Let R be a commutative Noetherian ring and let $\mathbf{f} = (f(n))_{n\geq 0}$ be a R-valued sequence. Then $\mathbf{f}$ is said to be *k-regular* if the dimension of the R-module spanned by its k-kernel is finite.

In the sequel, we will say that a power series $F(x) \in R[[x]]$ is *k-regular* (respectively *k-automatic*) if its sequence of coefficients is k-regular (respectively k-automatic). Of course, with a subset $\mathcal{E}$ of $\mathbb{N}$, we can associate its characteristic sequence $\chi(n)$, taking values in $\{0, 1\}$, and thus a power series $F_{\mathcal{E}}(x) := \sum \chi(n)x^n \in \mathbb{Z}[[x]]$. When the set $\mathcal{E}$ is k-automatic, $F_{\mathcal{E}}(x)$ is a k-automatic power series. More generally, a power series $F(x) = \sum f(n)x^n$ with coefficients in a finite set S is k-automatic if and only if for every $s \in S$ the set $\{n \in \mathbb{N} \mid f(n) = s\}$ is k-automatic. In the following proposition, we collect some useful general facts about k-regular power series.

Proposition 7.4. *Let R be a commutative ring and $k \geq 2$ be an integer. Then the following properties hold.*

(i) *If $F(x) \in R[[x]]$ is k-regular and I is an ideal of R, then $F(x) \bmod I \in (R/I)[[x]]$ is k-regular;*
(ii) *If $F(x) \in R[[x]]$ is k-regular, then the coefficients of $F(x)$ take only finitely many distinct values if and only if $F(x)$ is k-automatic;*

(iii) *If $F(x) = \sum_{i\geq 0} f(i)x^i$ and $G(x) = \sum_{i\geq 0} g(i)x^i$ are two k-regular power series in $R[[x]]$, then the Cauchy product*

$$F(x)G(x) := \sum_{i=0}^{\infty}\left(\sum_{j=0}^{i}\binom{i}{j} f(j)g(i-j)\right)x^i$$

is k-regular.

Proof. The property (i) follows directly from the definition of a k-regular sequence, while (ii) and (iii) correspond respectively to Theorem 16.1.5 and Corollary 16.4.2 in [6]. $\square$

In Section 11, we will need to use that k-regular sequences with complex values do have strict restrictions on the growth of their absolute values, a fact evidenced by the following result.

Proposition 7.5. *Let $k \geq 2$ be a natural number and let $F(x) \in \mathbb{C}[[x]]$ be a k-regular power series. Then $F(x)$ is analytic in the open unit disk and there exist two positive real numbers C and m such that*

$$|F(x)| < C(1-|x|)^{-m},$$

for all $x \in B(0, 1)$.

Proof. Let $F(x) = \sum_{i=0}^{\infty} f(i)x^i \in \mathbb{C}[[x]]$ be a k-regular power series. Then there is some positive constant A and some integer $d > 0$ such that

$$|f(i)| \leq A(i+1)^d,$$

for every nonnegative integer i (see [6, Theorem 16.3.1]). This immediately gives that $F(x)$ is analytic in the open unit disk. Moreover, for $x \in B(0, 1)$,

$$|F(x)| \leq \sum_{i=0}^{\infty} A(i+1)^d|x|^i \leq \sum_{i=0}^{\infty} Ad!\binom{i+d}{d}|x|^i = Ad!(1-|x|)^{-d-1}.$$

The result follows. $\square$

7.2. Becker power series

Becker [9, Theorem 1] showed that a k-regular power series is necessarily k-Mahler. In addition to this, he proved [9, Theorem 2] the following partial converse (see Theorem 7.6 below). The general converse does not hold. For example, the power series in $\mathbb{Q}[[x]]$ defined by the k-Mahler equation

$$(1-x)F(x) = F\left(x^k\right)$$

and satisfying $F(0) = 1$ is not k-regular. This can easily be shown using Proposition 7.5.

Theorem 7.6 (Becker). *Let K be a field, let k be a natural number ≥ 2, and let $F(x) \in K[[x]]$ be a power series that satisfies a k-Mahler equation of the form*

$$F(x) = \sum_{i=1}^{n} P_i(x) F\left(x^{k^i}\right) \tag{7.1}$$

for some polynomials $P_1(x), \ldots, P_n(x) \in K[x]$. Then $F(x)$ is a k-regular power series.

Definition 7.7. In honour of Becker's result, a power series $F(x) \in K[[x]]$ that satisfies an equation of the form given in Equation (7.1) will be called a *k-Becker* power series.

Theorem 7.6 shows that the set of k-Becker power series is contained in the set of k-regular power series. However, the converse is not true. As an example, we provide the following result that will also be used in Section 13.

Proposition 7.8. *Let k be a natural number, and let $\omega \in \mathbb{C}$ be a root of unity with the property that if $j \geq 1$ then $\omega^{k^j} \neq \omega$. Then*

$$\left(\prod_{j=0}^{\infty}\left(1 - \omega x^{k^j}\right)\right)^{-1}$$

is k-regular but it is not k-Becker.

Proof. Since ω is a root of unity, the sequence $\omega, \omega^k, \omega^{k^2}, \ldots$ is eventually periodic and there is some smallest natural number N such that

$$\omega^{k^{2N}} = \omega^{k^N} .$$

Set $\beta := \omega^{k^N}$ and let us consider the polynomial

$$Q(x) = (1 - \beta x)\left(1 - \beta x^k\right) \cdots \left(1 - \beta x^{k^{N-1}}\right).$$

Then

$$\frac{Q\left(x^k\right)}{Q(x)} = \frac{1 - \beta x^{k^N}}{1 - \beta x} .$$

Since

$$1 - \beta x^{k^N} = 1 - \left(\beta x\right)^{k^N} ,$$

we see that $Q(x^k)/Q(x)$ is a polynomial.

Since

$$1 - (\beta x)^{k^N} = \frac{Q\left(x^k\right)}{Q(x)} \cdot (1 - \beta x) ,$$

we get that $(1-\omega x)$ divides the polynomial $Q(x^k)(1-\beta x)/Q(x)$. Furthermore, $(1-\omega x)$ cannot divide $(1-\beta x)$ since by assumption $\omega \neq \beta$. By Euclid's lemma, we thus obtain that

$$\frac{Q\left(x^k\right)}{Q(x)} = (1-\omega x)S(x)$$

for some polynomial $S(x)$.

Set

$$F(x) := \left(\prod_{j=0}^{\infty}\left(1-\omega x^{k^j}\right)\right)^{-1}$$

and $G(x) := Q(x)^{-1}F(x)$. Since $F(x)$ satisfies the k-Mahler recurrence

$$F\left(x^k\right) = (1-\omega x)F(x)\,,$$

we see that

$$G\left(x^k\right) = Q\left(x^k\right)^{-1}(1-\omega x)Q(x)G(x),$$

or equivalently,

$$G(x) = S(x)G\left(x^k\right).$$

Thus $G(x)$ is a k-Becker power series. By Proposition 7.4, $F(x)$ is k-regular as it is a product of a polynomial (which is k-regular) and a k-regular power series.

On the other hand, $F(x)$ cannot be a k-Becker power series. To see this, suppose that $F(x)$ satisfies an equation of the form

$$F(x) = \sum_{i=1}^{d} P_i(x)F\left(x^{k^i}\right).$$

Now, dividing both sides by $F(x^k)$, the right side becomes a polynomial in x, while the left side is $(1-\omega x)^{-1}$, a contradiction. The result follows. □

In Section 11, we will need the following basic result about k-Becker power series.

Lemma 7.9. *Let $k \geq 2$ and let us assume that $F(x) \in K[[x]]$ satisfies a k-Mahler equation of the form*

$$F(x) = \sum_{i=1}^{n} a_i F\left(x^{k^i}\right)$$

for some constants $a_1, \dots, a_n \in K$. Then $F(x)$ is constant.

Proof. Let us denote by $F(x) = \sum_{i\geq 0} f(i)x^i$ the power series expansion of $F(x)$. If $F(x)$ were non-constant, there would be some smallest positive integer i_0 such that $f(i_0) \neq 0$. Thus $F(x) = \lambda + x^{i_0} F_0(x)$ for some λ in K and some $F_0(x) \in K[[x]]$. But taking the coefficient of x^{i_0} in the right-hand side of the equation

$$F(x) = \sum_{i=1}^{n} a_i F\left(x^{k^i}\right),$$

we see that $f(i_0) = 0$, a contradiction. The result follows. □

Though there are some Mahler functions that are not Becker functions, the following result shows that every k-Mahler power series can be decomposed as the product of a k-Becker power series and the inverse of an infinite product of polynomials. This decomposition will turn out to be very useful to prove Theorem 1.3. This result appears as Theorem 31 in the Thèse de Doctorat of Dumas [14].

Proposition 7.10. *Let k be a natural number, let K be a field, and let $F(x) \in K[[x]]$ be a k-Mahler power series satisfying an equation of the form*

$$\sum_{i=0}^{n} P_i(x) F\left(x^{k^i}\right) = 0,$$

where $P_0(x), \ldots, P_n(x) \in K[x]$ and $P_0(0) = 1$. Then there is a k-Becker power series $G(x)$ such that

$$F(x) = \left(\prod_{i=0}^{\infty} P_0\left(x^{k^i}\right)\right)^{-1} G(x).$$

8. Conditions on k and ℓ

In this section K will denote an arbitrary field. We consider power series in $K[[x]]$ that are both k- and ℓ-Mahler with respect to two multiplicatively independent natural numbers k and ℓ. More specifically, we look at the set of natural numbers m for which such a power series is necessarily m-Mahler.

Proposition 8.1. *Let k and ℓ be two integers ≥ 2 and let $F(x) \in K[[x]]$ be a power series that is both k- and ℓ-Mahler. Let us assume that a and b are integers with the property that $m := k^a \ell^b$ is an integer and $m > 1$. Then $F(x)$ is also m-Mahler.*

Proof. Let V denote the $K(x)$-vector space spanned by all the power series that belong to the set $\left\{F(x^{k^a \ell^b}) \mid a, b \in \mathbb{N}\right\}$. Recall that by Lemma 4.1, we can assume that the corresponding Mahler equations are both homogeneous. Hence there exists some natural number N such that for every integer $n \geq N$ we have $F(x^{k^n}) = \sum_{i=0}^{N-1} P_{i,n}(x) F(x^{k^i})$ and $F(x^{\ell^n}) = \sum_{i=0}^{N-1} Q_{i,n}(x) F(x^{\ell^i})$ for some rational functions $P_{0,n}(x), \ldots, P_{N-1,n}(x), Q_{0,n}(x), \ldots, Q_{n,N-1}(x)$. Thus V is a $K(x)$-vector space of dimension at most N^2.

Suppose that a and b are integers such that $m := k^a \ell^b$ is an integer and $m > 1$. If a and b are nonnegative, then $F(x^{m^j}) \in V$ for every integer $j \geq 0$ and since the dimension of V is finite, we see that $F(x)$ is m-Mahler. Thus we may assume that at least one of a or b is negative. Since $m \geq 1$, at least one of a or b must also be positive. Without loss of generality, we may thus assume that $a > 0$ and $b < 0$.

We are now going to show that $F(x^{m^j}) \in V$ for every nonnegative integer j. To see this, we fix a nonnegative integer j. Then we observe that $m^j \ell^{-bj} = k^{ja}$ and thus $F(x^{m^j l^i})$ belongs to V for every integer $i \geq -bj$. Since $-bj \geq 0$, there exists a smallest nonnegative integer i_0 such that $F(x^{m^j \ell^i}) \in V$ for every integer $i \geq i_0$. If i_0 is zero, then we are done. We assume that i_0 is positive and look for a contradiction. By definition of i_0, we note that $F(x^{m^j \ell^{i_0 - 1}}) \notin V$. By assumption, $F(x)$ satisfies a ℓ-Mahler equation of the form

$$\sum_{i=0}^{N} P_i(x) F(x^{\ell^i}) = 0,$$

with $P_0(x), \ldots, P_N(x) \in K[x]$ and $P_0(x) \neq 0$. Applying the operator $x \mapsto x^{m^j \ell^{i_0 - 1}}$, we get that

$$P_0\left(x^{m^j \ell^{i_0-1}}\right) F\left(x^{m^j \ell^{i_0-1}}\right) = -\sum_{i=1}^{N} P_i\left(x^{m^j \ell^{i_0-1}}\right) F\left(x^{m^j \ell^{i_0-1+i}}\right).$$

By definition of i_0, the right side of this equation is in V, and so $F(x^{m^j \ell^{i_0-1}}) \in V$ since $P_0(x)$ is nonzero. This is a contradiction. It follows that $F(x^{m^j}) \in V$ for every nonnegative integer j.

Since V is a $K(x)$-vector space of dimension at most N^2, we see that $F(x), F(x^m), \ldots, F(x^{m^{N^2}})$ are linearly dependent over $K(x)$, which implies that $F(x)$ is m-Mahler. This ends the proof. □

Remark 8.2. Taking $k = \ell$ and $b = 0$ in Proposition 8.1, we see that if a power series $F(x)$ is k-Mahler then it is also k^a-Mahler for every $a \geq 1$. The converse is obvious. Consequently, if k and ℓ are multiplicatively dependent natural numbers, then $F(x)$ is k-Mahler if and only if it is ℓ-Mahler.

Corollary 8.3. *Let k and ℓ be two multiplicatively independent natural numbers and let $F(x) \in K[[x]]$ be a power series that is both k- and ℓ-Mahler. Then there exist two multiplicatively independent positive integers k' and ℓ' such that the following conditions hold.*

(i) *There is a prime number p that divides k' and does not divide ℓ';*
(ii) *There is a prime number q that divides ℓ' and does not divide k';*
(iii) *$F(x)$ is both k'- and ℓ'-Mahler.*

Proof. There exist prime numbers $p_1, \dots, p_m$ and nonnegative integers $a_1, \dots, a_m$, $b_1, \dots, b_m$ such that

$$k = \prod_{i=1}^{m} p_i^{a_i} \text{ and } \ell = \prod_{i=1}^{m} p_i^{b_i}.$$

Moreover, we can assume that, for each i, at least one of a_i or b_i is positive.

Note that if there are i and j such that $a_i = 0$ and $b_j = 0$, then we can take $k' := k$ and $\ell' := \ell$ and set $p := p_j$ and $q := p_i$ to obtain the desired result. Thus we can assume without loss of generality that $b_i > 0$ for $i \in \{1, \dots, m\}$. Then there is some $i_0 \in \{1, \dots, m\}$ such that $a_{i_0}/b_{i_0} \leq a_j/b_j$ for all $j \in \{1, \dots, m\}$. In particular, $c_j := a_j b_{i_0} - b_j a_{i_0}$ is a nonnegative integer for all $j \in \{1, \dots, m\}$. Hence

$$k' := k^{b_{i_0}} \ell^{-a_{i_0}} = \prod_{j=1}^{m} p_j^{c_j} \in \mathbb{N}.$$

Furthermore, p_{i_0} does not divide k' and since k and ℓ are multiplicatively independent, the c_i's are not all equal to zero.

Now we pick $i_1 \in \{1, \dots, m\}$ such that $c_{i_1}/b_{i_1} \geq c_j/b_j$ for all $j \in \{1, \dots, m\}$. Note that $c_{i_1} > 0$ since the c_i's are not all equal to zero. Set

$$\ell' := \ell^{c_{i_1}} (k')^{-b_{i_1}} = \prod_{j=1}^{m} p_j^{b_j c_{i_1} - b_{i_1} c_j} \in \mathbb{N}.$$

Since $c_{i_0} = 0, c_{i_1} > 0$ and the b_i's are positive, we get that p_{i_0} divides ℓ'. Moreover, p_{i_1} does not divide ℓ' while p_{i_1} divides k' for c_{i_1} is positive. In particular, k' and ℓ' are multiplicatively independent. Furthermore, Proposition 8.1 implies that $F(x)$ is both k'- and ℓ'-Mahler. Setting $q := p_{i_0}$ and $p = p_{i_1}$, we obtain that k' and ℓ' have all the desired properties. This concludes the proof. □

9. Asymptotic estimates for some infinite products

In this section, we study the behaviour around the unit circle of infinite products of the form

$$\left(\prod_{i=0}^{\infty} P(x^{k^i}) \right)^{-1},$$

where $P(x) \in \mathbb{C}[x]$ and $P(0) = 1$. We obtain some asymptotic estimates that will be necessary in Section 11.

We will prove that when α is a root of unity satisfying $\alpha^k = \alpha$ that is not a root of P, then this product is rather well-behaved when approaching α through certain well-chosen sets of points. Throughout Sections 9, 10, and 11, we make use of certain subsets of the unit circle having 1 as a limit point. We define these sets now.

Definition 9.1. Let $\varepsilon \in (0, 1)$ and let $\theta \in [-1, 1]$. Then we define

$$X_{\theta,\varepsilon} := \{\exp((-1 + i\theta)s) \mid s \in (0, \varepsilon)\}. \tag{9.1}$$

We take X_θ to be the set $\{0\} \cup \{\exp((-1 + i\theta)s \mid s \geq 0\}$.

We note that each X_θ is a compact subset of the closed unit disk. In fact, X_θ is homeomorphic to $\mathbb{R}_{\geq 0} \cup \{+\infty\}$.

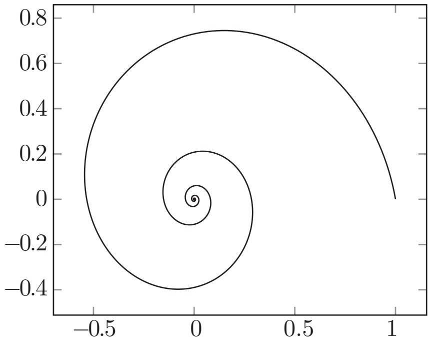

Figure 9.1. This picture of the full set X_θ, with $\theta = 5$, shows the spiral-like structure of the curve.

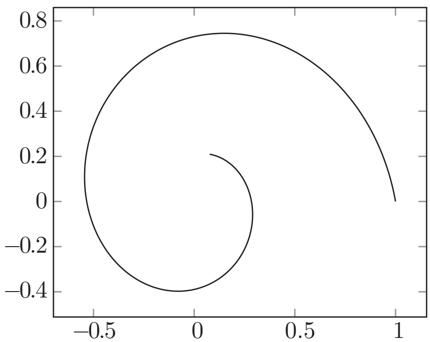

Figure 9.2. This picture shows the set $X_{\theta,\varepsilon}$, where we take $\theta = 5$ and $\epsilon = 1.5$.

Moreover, if $\theta \neq \theta'$, two sets of the form $X_{\theta,\varepsilon}$ and $X_{\theta',\varepsilon'}$ are always disjoint. This can be seen by noting that if $\exp((-1 + i\theta)s) = \exp((-1 + i\theta')s')$, with $\theta, \theta' \in [0, 1]$, then they have the same modulus and hence $s = s'$; next we must have that $\exp(i\theta s) = \exp(i\theta' s)$ and so $(\theta - \theta')s$ must be an integer multiple of 2π, which can only occur if $\theta = \theta'$ since $|\theta - \theta'| \leq 2$ and $0 < s < 1$. Finally, we remark that a set of the form $X_{\theta,\varepsilon}$ has the property that if $y \in X_{\theta,\varepsilon}$ and k is a positive integer then there is a unique point $z \in X_{\theta,\varepsilon}$ such that $z^k = y$.

Proposition 9.2. *Let $k \geq 2$ be a natural number, let α be root of unity that satisfies $\alpha^k = \alpha$, and let $P(x)$ be a nonzero polynomial with $P(0) = 1$ and $P(\alpha) \neq 0$. Then for all but countably many $\theta \in [-1, 1]$, there exist two positive real numbers A and $\varepsilon \in (0, 1)$, depending upon θ, such that*

$$|1-t|^A < \left| \left(\prod_{j=0}^{\infty} P\Big((t\alpha)^{k^j}\Big) \right)^{-1} \right| < |1-t|^{-A}$$

whenever $t \in X_{\theta,\epsilon}$.

In contrast, the following result shows that such infinite products behave differently when α is a root of P. In the case where $k = 2$, we point out that a different proof can be found in [5, Théorème 3]. Precise asymptotics for the coefficients of the power series expansion of this infinite product has also been studied by Mahler, de Bruijn, and Dumas and Flajolet (see [15] and the references therein). We give the following proof for the sake of completeness.

Lemma 9.3. *Let $k \geq 2$ be a natural number. Then if $\{t_n\}$ is a sequence of complex numbers with $|t_n| < 1$ for every n such that $t_n \to 1$ as $n \to \infty$ then*

$$\lim_{n\to\infty} \left| \prod_{j=0}^{\infty} \frac{1}{1-t_n^{k^j}} \right| \cdot |1-t_n|^A = \infty,$$

for every positive real number A.

Proof. By ignoring some initial terms of our sequence, we may assume that $|1 - t_n| \in (0, 1/k^9)$ for every n. Now let $t \in B(0, 1)$ be such that $|1 - t| \in (0, 1/k^9)$. Let $N \geq 2$ be the largest natural number such that $|1 - t| \in (0, k^{-(N+1)^2})$. Then

$$\begin{aligned}
\left| \prod_{j=0}^{\infty} \left(1-t^{k^j}\right)^{-1} \right| &\geq \left| \prod_{j=0}^{N} \left(1-t^{k^j}\right)^{-1} \right| \\
&= \left|(1-t)^{-(N+1)}\right| \left| \prod_{j=0}^{N} \left(1+t+\cdots+t^{k^j-1}\right)^{-1} \right| \\
&\geq \left|(1-t)^{-(N+1)}\right| \prod_{j=0}^{N} k^{-j} \\
&\geq |1-t|^{-(N+1)} k^{-(N+1)^2} \\
&> |1-t|^{-N}.
\end{aligned}$$

By definition of N, we obtain that $|1 - t| > k^{-(N+2)^2}$, which easily gives that

$$N > \sqrt{\frac{-\log|1-t|}{4\log k}}.$$

This ends the proof, for the right-hand side tends to infinity when t tends to 1. $\square$

We are now going to prove Proposition 9.2. We will need the following two auxiliary results.

Lemma 9.4. *Let $k \geq 2$ be a natural number. Then for $t \in (0,1)$, we have*

$$\sum_{i=1}^{\infty} t^i/i \;\geq\; (1-1/k)\sum_{i=0}^{\infty} t^{k^i}\,.$$

Proof. We have

$$\begin{aligned}
\sum_{i=1}^{\infty} t^i/i &= t + \sum_{i=0}^{\infty}\sum_{j=k^i+1}^{k^{i+1}} t^j/j \\
&\geq t + \sum_{i=0}^{\infty}\sum_{j=k^i+1}^{k^{i+1}} t^{k^{i+1}}/k^{i+1} \\
&= t + \sum_{i=0}^{\infty} t^{k^{i+1}}\left(k^{i+1}-k^i\right)/k^{i+1} \\
&= t + (1-1/k)\sum_{i=0}^{\infty} t^{k^{i+1}} \\
&\geq (1-1/k)\sum_{i=0}^{\infty} t^{k^i}\,,
\end{aligned}$$

which ends the proof. □

Lemma 9.5. *Let $k \geq 2$ be a natural number and let $\lambda \neq 1$ be a complex number. Then for all but countably many $\theta \in [-1,1]$, there exist two positive real numbers A and $\varepsilon \in (0,1)$, depending upon θ, such that*

$$|1-t|^A \;<\; \left|\prod_{j=0}^{\infty}\frac{1}{1-\lambda t^{k^j}}\right| \;<\; |1-t|^{-A}$$

whenever $t \in X_{\theta,\varepsilon}$.

Proof. We first prove the inequality on the right-hand side.

We note that for each $j \geq 0$ there are only finitely many complex solutions to the equation $1-\lambda t^{k^j}=0$, and thus there are at most countably many solutions as j ranges over all nonnegative integers. As already observed, for $\theta \neq \theta'$ with $\theta,\theta' \in [-1,1]$ and for $\varepsilon,\varepsilon' \in (0,1)$, one has $X_{\theta,\varepsilon}\cap X_{\theta',\varepsilon'}=\emptyset$. It follows that for all but countably many values of $\theta \in [-1,1]$ the equation $1-\lambda t^{k^j}=0$ has no solution on $X_{\theta,\varepsilon}$ whenever $\varepsilon \in (0,1)$. Moreover, since $\lambda \neq 1$, $t=1$ and $t=0$ are never a solution, this equation has no solution in X_θ. For the remainder of the proof we assume that $\theta \in [-1, 1]$ has this property.

Observe that X_θ is a compact set that is closed under the map $t \mapsto t^k$ and we have that $1 - \lambda t$ is nonzero for $t \in X_\theta$. By compactness, we see that there exist two positive real numbers ε_0 and c_0, $c_0 < 1$ and depending upon θ, such that

$$\inf\left\{\left|1 - \lambda t^{k^j}\right| : t \in X_{\theta,\varepsilon_0},\ j \geq 0\right\} > c_0 . \tag{9.2}$$

We fix $t \in X_{\theta,\varepsilon_0}$ and we let $N = N(t)$ to be the largest nonnegative integer such that $|t^{k^N}| \geq 1/2$. Then for $j \geq 1$ we have $\left|t^{k^{N+j}}\right| = \left|(t^{k^{N+1}})^{k^{j-1}}\right| < (1/2)^{k^{j-1}}$. Hence

$$\left|1 - \lambda t^{k^{N+j}}\right| \geq 1 - |\lambda|(1/2)^{k^{j-1}} .$$

Since the series $\sum_{j\geq 0}(1/2)^{k^{j-1}}$ converges, we get that the infinite product

$$\prod_{j=N(t)+1}^{\infty} \left|\frac{1}{1 - \lambda t^{k^j}}\right|$$

is uniformly bounded over X_{θ,ε_0} by some constant c_1. (We note that $\lambda \neq 1$ is fixed, $N = N(t)$ depends upon t, $t \in X_{\theta,\varepsilon_0}$, and it is necessary to begin the product at $N + 1$ in order to achieve uniformity in our bound.) Then

$$\begin{aligned}
\left|\prod_{j=0}^{\infty}\left(1 - \lambda t^{k^j}\right)^{-1}\right| &= \prod_{j=0}^{N}\left|1 - \lambda t^{k^j}\right|^{-1} \prod_{j=1}^{\infty}\left|1 - \lambda t^{k^{N+j}}\right|^{-1} \\
&\leq (1/c_0)^{N+1} c_1 \\
&= \left(k^{N+1}\right)^{-\log c_0/\log k} c_1 .
\end{aligned}$$

Furthermore, we have by assumption that $\left|t^{k^{N+1}}\right| < 1/2$ and thus $k^{N+1} < -\log 2/\log|t|$. This implies that

$$\left|\prod_{j=0}^{\infty}\left(1 - \lambda t^{k^j}\right)^{-1}\right| \leq c_1 \left(-\log 2/\log|t|\right)^{-\log c_0/\log k} .$$

Now we let t tend to 1 along X_{θ,ε_0}, that is we write $t = \exp((-1 + i\theta)s)$ with $s \in (0, \varepsilon_0)$ and we let s tend to 0. Then we have $|t| = \exp(-s)$ and so $\log|t| = -s$. Then when $t \to 1$ along the arc X_{θ,ε_0} we have that $|1 - t|/\log|t|$ tends to

$$\lim_{s\to 0} \frac{|1 - \exp((-1 + i\theta)s)|}{-s} = -|-1 + i\theta| \geq -\sqrt{2}$$

and hence there exists some positive real numbers $\varepsilon < \varepsilon_0$ and c_2 such that

$$c_1\,(-\log 2/\log|t|)^{-\log c_0/\log k} < c_2|1-t|^{\log c_0/\log k}\,,$$

whenever $t \in X_{\theta,\varepsilon}$. Since $c_0 < 1$, we obtain that there exists a positive real number A_1 such that

$$\left|\prod_{j=0}^{\infty}\left(1-\lambda t^{k^j}\right)^{-1}\right| < |1-t|^{-A_1}\,,$$

for all $t \in X_{\theta,\varepsilon}$. This gives the right-hand side bound in the statement of the lemma.

To get the left side, note that for all $t \in X_\theta$,

$$\left|\prod_{j=0}^{\infty}\frac{1}{1-\lambda t^{k^j}}\right| \geq \prod_{j=0}^{\infty}\left|1+|\lambda||t|^{k^j}\right|^{-1} \geq \prod_{j=0}^{\infty}\exp\left(-|\lambda|\cdot|t|^{k^j}\right).$$

By Lemma 9.4, we have

$$\prod_{j=0}^{\infty}\exp\left(-|\lambda|\cdot|t|^{k^j}\right) \geq \exp\left(-|\lambda|(1-1/k)^{-1}\sum_{i=1}^{\infty}|t|^i/i\right) = (1-|t|)^{|\lambda|k/(k-1)}\,.$$

We thus obtain that, for all $t \in X_\theta$,

$$\left|\prod_{j=0}^{\infty}\frac{1}{1-\lambda t^{k^j}}\right| \geq (1-|t|)^{A_2}\,,$$

where $A_2 := \lfloor|\lambda|k/(k-1)\rfloor + 1$. Now we note that, when $t \to 1$ along the arc X_{θ,ε_0}, we have $|1-t|/(1-|t|)$ tends to $|-1+i\theta| \in [1, \sqrt{2}]$, which can be seen by writing $t = \exp((-1+i\theta)s)$ and letting $s \to 0$ and taking limits. Since $\varepsilon < 1$, it follows that there is some positive constant $A_3 > A_2$ for which we have

$$\left|\prod_{j=0}^{\infty}\frac{1}{1-\lambda t^{k^j}}\right| > |1-t|^{A_3}\,,$$

whenever $t \in X_{\theta,\varepsilon}$.

Taking A to be equal to the maximum of A_1 and A_3, we get the desired result. □

Proof of Proposition 9.2. Let $\beta_1, \ldots, \beta_s$ denote the complex roots of P (considered with mutliplicities) so that we may factor $P(x)$ as $P(x) = (1-\beta_1^{-1}x)\cdots(1-\beta_s^{-1}x)$. We thus obtain

$$\left|\prod_{j=0}^{\infty}\frac{1}{P\left((t\alpha)^{k^j}\right)}\right| = \prod_{i=1}^{s}\left|\prod_{j=0}^{\infty}\frac{1}{1-\beta_i^{-1}\alpha t^{k^j}}\right|\,,$$

where $\beta_i^{-1}\alpha \neq 1$ for every $i \in \{1, \dots, s\}$. Then by Lemma 9.5, there are cocountable[8] subsets Y_i of $[-1, 1]$ such that for a given i and a given $\theta \in Y_i$, there exist a natural number A and a positive real number ε, $0 < \varepsilon < 1$, depending upon θ, such that

$$|1-t|^A < \left| \prod_{j=0}^{\infty} \left(1 - \beta_i^{-1}\alpha t\right)^{-1} \right| < |1-t|^{-A}$$

whenever $t \in X_{\theta,\varepsilon}$. Since the finite intersection of cocountable sets is cocountable, we see that taking $Y = Y_1 \cap \cdots \cap Y_s$, that whenever $\theta \in Y$ we have there exist natural numbers A_i and positive real numbers ε_i, $0 < \varepsilon_i < 1$, depending upon θ, such that

$$|1-t|^{A_i} < \left| \prod_{j=0}^{\infty} \left(1 - \beta_i^{-1}\alpha t\right)^{-1} \right| < |1-t|^{-A_i}$$

whenever $t \in X_{\theta,\varepsilon_i}$. Taking $\varepsilon := \min(\varepsilon_1, \dots, \varepsilon_s)$ and $A := \sum_{i=1}^{s} A_i$, we obtain the desired result. □

10. Asymptotic estimates for solutions of analytic Mahler-type systems

In this section we fix a non-trivial norm $\|\cdot\|$ on $\mathbb{C}^d$. We let $B(x,r)$ (respectively $\overline{B(x,r)}$) denote the open (respectively closed) ball of radius r centered at x. Our results will not depend on the choice of this norm. Throughout this section, we make use of the sets $X_{\theta,\varepsilon}$ and X_θ defined in Definition 9.1.

As defined in Section 7, a Becker function $F(x) \in \mathbb{C}[[x]]$ is an analytic function on the open unit disk satisfying a functional equation of the form:

$$F(x) = \sum_{i=1}^{n} P_i(x) F\left(x^{k^i}\right)$$

for some polynomials $P_1(x), \dots, P_n(x) \in \mathbb{C}[x]$. Of course, such an equation leads to a k-Mahler linear system

$$\begin{pmatrix} F(x) \\ \vdots \\ F(x^{k^{n-1}}) \end{pmatrix} = A(x) \begin{pmatrix} F(x^k) \\ \vdots \\ F(x^{k^n}) \end{pmatrix},$$

where $A(x)$ is an $n \times n$ matrix with polynomial entries. In what follows, we provide an asymptotic lower bound around certain points of the unit circle for solutions of similar systems but associated with more general matrices. Indeed, we consider matrices whose entries are only assumed to be analytic on $B(0,1)$ and continuous on $\overline{B(0,1)}$. This result will be used in Section 11.

[8] This means, of course, that the complement of Y_i in $[-1, 1]$ is a countable set.

Proposition 10.1. *Let d and k be two natural numbers, let α be a root of unity such that $\alpha^k = \alpha$ and let $A : \overline{B(0,1)} \to M_d(\mathbb{C})$ be a continuous matrix-valued function. Let us assume that $w(x) \in \mathbb{C}[[x]]^d$ satisfies the equation*

$$w(x) = A(x)w\left(x^k\right)$$

for all $x \in B(0,1)$. Let us also assume that the following properties hold.

(i) *The coordinates of $w(x)$ are analytic in $B(0,1)$ and continuous on $\overline{B(0,1)}$;*
(ii) *The matrix $A(\alpha)$ is not nilpotent;*
(iii) *There exist two positive real numbers ε and M such that $|\det(A(x))| > (1 - |x|)^M$ for every x with $1 - \varepsilon < |x| < 1$;*
(iv) *The set $\{w(x) \mid x \in B(0,1)\}$ is not contained in a proper vector subspace of $\mathbb{C}^d$.*

If ζ is a root of unity such that $\zeta^{k^j} = 1$ for some natural number j and $\theta \in [-1,1]$, then there exist a positive real number C and a subset $S \subseteq X_\theta$ that has 1 as a limit point such that

$$||w(t\alpha\zeta)|| > |1-t|^C$$

for all $t \in S$.

Before proving Proposition 10.1, we will need two auxiliary results.

Lemma 10.2. *Let d and k be two natural numbers, let α be a root of unity such that $\alpha^k = \alpha$, and let $A : \overline{B(0,1)} \to M_d(\mathbb{C})$ be a continuous matrix-valued function. Let us assume that $w(x) \in \mathbb{C}[[x]]^d$ satisfies the equation*

$$w(x) = A(x)w\left(x^k\right)$$

for all $x \in B(0,1)$. Let us also assume that the following properties hold.

(i) *The coordinates of $w(x)$ are analytic in $B(0,1)$ and continuous on $\overline{B(0,1)}$;*
(ii) *The matrix $A(\alpha)$ is not nilpotent;*
(iii) *The set $\{w(x) \mid x \in B(0,1)\}$ is not contained in a proper vector subspace of $\mathbb{C}^d$.*

Then if $\theta \in [-1,1]$, then there exist a positive real number C and a subset $S \subseteq X_\theta$ that has 1 as a limit point such that

$$||w(t\alpha)|| > |1-t|^C$$

for all $t \in S$.

Proof. Since $A(\alpha)$ is not nilpotent, there is some natural number e such that the kernel of $A(\alpha)^e$ and the kernel of $A(\alpha)^{e+1}$ are equal to a same proper subspace of $\mathbb{C}^d$, say W. Then there is a nonzero vector subspace V such that $A(\alpha)(V) \subseteq V$

and $V \oplus W = \mathbb{C}^d$. Moreover, by compactness, there is a positive real number c_0, $c_0 < 1$, such that

$$||A(\alpha)(w)|| \geq c_0 \tag{10.1}$$

whenever $w \in V$ is a vector of norm 1.

Since every vector x has a unique decomposition of the form $v \oplus w$ with v in V and w in W, we see that the map $\pi(x) := v$ gives a continuous linear projection map $\pi : \mathbb{C}^d \to V$ with the property that $u - \pi(u) \in W$ for all $u \in \mathbb{C}^d$. We infer from Inequality (10.1) that

$$||\pi(A(\alpha)(u))|| = \|A(\alpha)(\pi(u))\| \geq c_0||\pi(u)|| \tag{10.2}$$

for all $u \in \mathbb{C}^d$. Since A is continuous on $\overline{B(0,1)}$, Inequality (10.2) implies the existence of a positive constant $\varepsilon > 0$ such that

$$||\pi(A(x)(u))|| > (c_0/2)||\pi(u)||,$$

for all $u \in \mathbb{C}^d$ and all $x \in B(\alpha, \varepsilon) \cap \overline{B(0,1)}$. It follows by a simple induction that if $x_1, \ldots, x_m \in B(\alpha, \varepsilon) \cap \overline{B(0,1)}$ then

$$||\pi(A(x_1) \cdots A(x_m)(u))| \geq (c_0/2)^m ||\pi(u)||. \tag{10.3}$$

Let $\theta \in [-1, 1]$. We claim that there exists a complex number t_0 such that $t_0 \in X_\theta \cap B(1, \varepsilon)$ and $w(t_0\alpha) \notin W$. Otherwise, there would be a nonzero row vector u such that $u \cdot w(t\alpha) = 0$ for all $t \in X_\theta \cap B(1, \varepsilon)$. But $u \cdot w(x)$ is analytic in $B(0, 1)$ for $w(x)$ is and hence it would be identically zero on $B(0, 1)$ by the identity theorem since $X_\theta \cap B(1, \varepsilon)$ has accumulation points inside the open unit disk. This would contradict assumption (iii).

From now on, we fix a complex number t_0 with this property. For every $i \geq 1$, we then define t_i to be the unique element in X_θ such that $t_i^k = t_{i-1}$. Since $w(t_0\alpha) \notin W$, there exists a positive real number c_1 such that

$$||\pi(w(t_0\alpha))|| = c_1 > 0.$$

Furthermore, by construction, the sequence $t_0, t_1, t_2, \ldots$ belongs to $X_\theta \cap B(1, \varepsilon)$. We thus infer from (10.3) that

$$\begin{aligned} ||\pi(w(t_n\alpha))|| &= ||\pi(A(t_n\alpha)A(t_{n-1}\alpha)\cdots A(t_1\alpha)(w(t_0\alpha))|| \\ &\geq (c_0/2)^n ||\pi(w(t_0\alpha)|| \\ &\geq c_1(c_0/2)^n, \end{aligned}$$

for all $n \geq 1$. Furthermore, since the projection π is continuous, there is some positive real number c_2 such that $||\pi(u)|| < c_2||u||$ for all $u \in \mathbb{C}^d$. Thus

$$||w(t_n\alpha)|| \geq c_2^{-1}||\pi(w(t_n\alpha))|| \geq c_2^{-1}c_1(c_0/2)^n$$

for all $n \geq 1$.

On the other hand, we note that we have a map $\log : X_\theta \setminus \{0\} \to \mathbb{C}$ given by $\log(\exp((-1+i\theta)s) = (-1+i\theta)s$, and for each positive real number a, we have an a-th power map $X_\theta \to X_\theta$ given by $t \mapsto \exp(a\log(t))$. Since

$$\lim_{a\to 0^+} \frac{t_0^a - 1}{a} = \log(t_0)\,,$$

we have that $|t_0^a - 1|/a|t_0 - 1| \to |\log(t_0)|/|t_0 - 1|$, as $a \to 0^+$. Since t_0 is fixed, we let κ denote the quantity $|\log(t_0)|/|t_0 - 1|$.

Then there exists some $\varepsilon_0 \in (0,1)$ such that

$$|t_0^a - 1| < 2a\kappa|1 - t_0|$$

for $a \in (0, \varepsilon_0)$. Thus if n is large enough, say $n \geq n_0$, then $k^n > 1/\varepsilon_0$ and we have $|t_n - 1| = |(t_0)^{1/k^n} - 1| < 2\kappa|1 - t_0|/k^n$. Hence $k^n > 2\kappa|1 - t_0|/|1 - t_n|$. Then for $n \geq n_0$ we have

$$\begin{aligned}
||w(t_n\alpha)|| &> c_2^{-1}c_1(c_0/2)^n \\
&= c_2^{-1}c_1k^{n\log_k(c_0/2)} \\
&\geq \left(c_2^{-1}c_1\,(2\kappa|1 - t_0|)^{\log_k(c_0/2)}\right)|1 - t_n|^{-\log_k(c_0/2)}.
\end{aligned}$$

Thus if we take $C := -2\log_k(c_0/2) > 0$, the fact that t_n tends to 1 as n tends to infinity implies the existence of a positive integer $n_1 \geq n_0$ such that

$$||w(t_n\alpha)|| > |1 - t_n|^C\,,$$

for all $n \geq n_1$. Taking $S := \{t_n \mid n \geq n_1\}$, we obtain the desired result. □

Lemma 10.3. *Let $B : \overline{B(0,1)} \to M_d(\mathbb{C})$ be a continuous matrix-valued function whose entries are analytic inside the unit disk and continuous on the closed unit disk. Let us assume that there exist two positive real numbers ε and M such that $|\det(B(x))| > (1 - |x|)^M$ for every x such that $1 - \varepsilon < |x| < 1$. Then there exists a positive real number C such that for every column vector u of norm 1, we have*

$$||B(x)(u)|| \geq (1 - |x|)^C$$

for every x such that $1 - \varepsilon < |x| < 1$.

Proof. Our assumption implies that $B(x)$ is invertible for every x such that $1 - \varepsilon < |x| < 1$. Let $\Delta(x)$ denote the determinant of $B(x)$. Using the classical adjoint formula for the inverse of $B(x)$, we see that $B(x)^{-1}$ has entries $c_{i,j}(x)$ that have the property that they are expressible (up to sign) as the ratio of the determinant of a submatrix of $B(x)$ and $\Delta(x)$. Since the entries of $B(x)$ are continuous on $\overline{B(0,1)}$, each determinant of a submatrix of $B(x)$ is also continuous on $\overline{B(0,1)}$. By compactness, we see that there is a positive real number κ such that

$$|c_{i,j}(x)| \leq \kappa/|\Delta(x)| \leq \kappa(1 - |x|)^{-M}$$

for every $(i, j) \in \{1, \dots, d\}^2$ and every x such that $1 - \varepsilon < |x| < 1$. Thus there exists a positive real number C such that

$$\|B(x)^{-1}\| \leq (1 - |x|)^{-C}$$

for every x such that $1 - \varepsilon < |x| < 1$. It follows that if u is a vector of norm 1, then

$$\|B(x)(u)\| \geq (1 - |x|)^{C},$$

for every x such that $1 - \varepsilon < |x| < 1$. The result follows. □

Proof of Proposition 10.1. Let $\theta \in [-1, 1]$. Since $A(\alpha)$ is not nilpotent, we first infer from Lemma 10.2 that there exist a positive real number C_0 and a sequence $t_n \in X_\theta$, which tends to 1, such that $||w(t_n\alpha)|| > |1 - t_n|^{C_0}$ for every integer $n \geq 1$. Let $s_n \in X_\theta$ be such that $s_n^{k^j} = t_n$. Then

$$w(s_n\alpha\zeta) = A(s_n\alpha\zeta)A\Big(s_n^k\alpha\zeta^k\Big)\cdots A\Big(s_n^{k^{j-1}}\alpha\zeta^{k^{j-1}}\Big)(w(t_n\alpha)).$$

By assumption there exists a positive real number M such that $|\det(A(x))| > (1 - |x|)^M$ for every x with $1 - \varepsilon < |x| < 1$. Set

$$B(x) := A(x\alpha\zeta)A\Big(x^k\alpha\zeta^k\Big)\cdots A\Big(x^{k^{j-1}}\alpha\zeta^{k^{j-1}}\Big).$$

Then there is a positive real number C_1 such that if $(1 - \varepsilon)^{1/k^{j-1}} < |x| < 1$ then

$$\det(B(x)) > (1 - |x|)^M \cdots \Big(1 - |x|^{k^{j-1}}\Big)^M \geq (1 - |x|)^{jM}.$$

It follows from Lemma 10.3 that there exists a positive real number C_1 such that for n sufficiently large we have

$$\begin{aligned} ||w(s_n\alpha\zeta)|| = ||B(s_n)(w(t_n\alpha))|| &> (1 - |s_n|)^{C_1}||w(t_n\alpha)|| \\ &> (1 - |s_n|)^{C_1}|1 - t_n|^{C_0}. \end{aligned}$$

We have that $t_n = \exp((-1 + i\theta)u_n)$ where u_n is a sequence of positive numbers tending to 0. Taking limits, we then see that $|1 - t_n|/(1 - |s_n|) \to |-1 + i\theta| \cdot k^j$ and $|1 - s_n|/(1 - |s_n|) \to |-1 + i\theta|$ as $n \to \infty$. Hence there exists a positive real number C such that

$$||w(s_n\alpha\zeta)|| \geq |1 - s_n|^C$$

for all n sufficiently large. The result follows. □

11. Elimination of singularities at certain roots of unity

In this section we look at the singularities of k-Mahler functions at roots of unity of a certain form. Strictly speaking, we do not necessarily eliminate singularities, and so the section title is perhaps misleading. We do, however, show that one can reduce to the case of considering Mahler equations whose singularities at roots of unity have a restricted form.

Assumption-Notation 11.1. Throughout this section we make the following assumptions and use the following notation.

(a) We assume that k and l are integers, $k, l \geq 2$, for which: there exists a prime p such that $p|k$ and p does not divide ℓ, and there exists a prime q such that $q|\ell$ and q does not divide k. In particular, k and ℓ are two multiplicatively independent integers;

(b) We assume that $F(x)$ is a k-Mahler complex power series that satisfies an equation of the form

$$\sum_{i=0}^{a} A_i(x)F\left(x^{k^i}\right) = 0$$

with $A_0, \ldots, A_a \in \mathbb{C}[x]$ and $A_0(0) \neq 0$;

(c) We assume that $F(x)$ is an ℓ-Mahler complex power series that satisfies an equation of the form

$$\sum_{i=0}^{b} B_i(x)F\left(x^{\ell^i}\right) = 0$$

with $B_0, \ldots, B_b \in \mathbb{C}[x]$ and $B_0(0) \neq 0$.

In this section our aim is to prove the following result. It will be a key result for proving Theorem 1.3.

Theorem 11.2. *Let $F(x) \in \mathbb{C}[[x]]$ be a power series that satisfies Assumption-Notation* 11.1 *and that is not a polynomial. Then $F(x)$ satisfies a non-trivial k-Mahler equation of the form*

$$\sum_{i=0}^{c} P_i(x)F\left(x^{k^i}\right) = 0,$$

with the property that $P_0(0) = 1$ and $P_0(\alpha) \neq 0$ if α is a root of unity satisfying $\alpha^{k^j} = \alpha$ for some positive integer j.

Though this result is of a purely algebraic nature, our proof relies heavily on analytic methods. One may ask whether a purely algebraic proof exists.

Strategy of proof. Using Assumptions (b) and (c), Proposition 7.10 leads to two different expressions for F:

$$F(x) = \left(\prod_{j=0}^{\infty} A_0\left(x^{k^j}\right)\right)^{-1} H(x) \quad \text{and} \quad F(x) = \left(\prod_{j=0}^{\infty} B_0\left(x^{\ell^j}\right)\right)^{-1} G(x),$$

where H is k-Becker and G is ℓ-Becker. This gives:

$$\left(\prod_{j=0}^{\infty} A_0\left(x^{k^j}\right)\right)^{-1} = \left(\prod_{j=0}^{\infty} B_0\left(x^{\ell^j}\right)\right)^{-1} G(x)H(x)^{-1}.$$

We want to argue by contradiction assuming that A_0 has a root α satisfying $\alpha^{k^{i_0}} = \alpha$ for some positive integer i_0. The main idea is to use the asymptotics of Sections 7, 9, and 10 in order to show that the absolute values of the left-hand side and the right-hand side of the above Equality behave really differently in some neighbourhood of α, providing a contradiction. However, there are several technical difficulties and the proof will be divided into seven steps, as briefly described below.

In Step 1, we will first replace, for technical reasons, F by some function F_0 and the Equality above will be consequently replaced by

$$\left(\prod_{j=0}^{\infty} \widetilde{Q}_0\left(x^{k^{i_0 j}}\right)\right)^{-1} = \left(\prod_{j=0}^{\infty} R_0\left(x^{\ell^j}\right)\right)^{-1} G(x)H(x)^{-1}, \tag{11.1}$$

where $\widetilde{Q}_0$ is a polynomial satisfying $\widetilde{Q}_0(\alpha) = 0$ and $\alpha^{k^{i_0}} = \alpha$. Again for technical reasons, we will also have to replace the point α by $\alpha\zeta_0$, where ζ_0 is some well-chosen p^n-th root of unity (the choice of ζ_0 is made in Step 3). Here, p denotes the prime from Assumption (a).

At this point, one could use the results of Sections 7 and 9 to derive upper bounds showing that both $\left|\left(\prod_{j=0}^{\infty} R_0(x^{\ell^j})\right)^{-1}\right|$ and $|G(x)|$ do not grow too fast in some neighbourhood of the point $\alpha\zeta_0$. In contrast, it follows from Lemma 9.3 and Proposition 9.2 that $\left|\left(\prod_{j=0}^{\infty} \widetilde{Q}_0(x^{k^{i_0 j}})\right)^{-1}\right|$ becomes much bigger at certain well-chosen points near this point since $\widetilde{Q}_0(\alpha) = 0$ and $\alpha^{k^{i_0}} = \alpha$. This would be enough to derive a contradiction if we were able to obtain a lower bound for $|H(x)|$ around $\alpha\zeta_0$. Since H is a k-Becker function, it is easy to obtain a general upper bound (as we will do for G in Step 5), but we cannot obtain a suitable lower bound because the matrix associated with the underlying linear system of functional equations could be nilpotent.

In order to overcome this difficulty, we will replace H by the function $L(x) := H(x)\left(\prod_{j=0}^{\infty}(1 - \alpha^{-1}x^{k^{i_0 j}})^r\right)$, for some well-chosen rational parameter r. The choice of the parameter r will be given at Step 2. Once this last modification is made, one obtains, instead of Equality (11.1), an equality of the form:

$$\left|\prod_{j=0}^{\infty}\left(1-\alpha^{-1}x^{k^{i_0 j}}\right)^{-b}\right| = \left|\left(\prod_{j=0}^{\infty} R_0\left(x^{\ell^j}\right)\right)^{-1} G(x)\left(\prod_{j=0}^{\infty} S_0\left(x^{k^{i_0 j}}\right)\right) L(x)^{-1}\right|,$$

where S_0 is some polynomial and b is positive. It corresponds to Equality (11.8) in the proof.

In step 3, we will show that our choice of r allows to derive a suitable lower bound for $|L(x)|$ around $\alpha\zeta_0$ by applying Proposition 10.1. On the other hand, in Steps 4, 5, and 6, we will use the results of Sections 7 and 9 in order to provide suitable upper bounds for $\left|\left(\prod_{j=0}^{\infty} R_0(x^{\ell^j})\right)^{-1}\right|$, $|G(x)|$, and $\left|\prod_{j=0}^{\infty} S_0(x^{k^{i_0 j}})\right|$ around $\alpha\zeta_0$.

In step 7, we will finally gather all the bounds obtained in Steps 3, 4, 5, and 6 in order to deduce that, around $\alpha\zeta_0$, the right-hand side of Equality (11.8) is much smaller than the left-hand side should be according to Lemma 9.3. This will provide the desired contradiction.

With the preliminary results of Sections 6, 7, 9, and 10, we are now almost ready to prove Theorem 11.2. Before doing this, we give the following simple lemma. We recall that the Kronecker symbol $\delta_{i,j}$ is defined, as usual, by $\delta_{i,j}=1$ if $i=j$ and $\delta_{i,j}=0$ otherwise.

Lemma 11.3. *Let d be a natural number and let A be a $d \times d$ complex matrix whose (i, j)-entry is $\delta_{i,j+1}$ if $i \geq 2$. If there is an integer r such that the $(1, r)$-entry of of A is nonzero, then A is not nilpotent.*

Proof. Let $(a_1, \ldots, a_d)$ denote the first row of A. Then by the theory of companion matrices, A has characteristic polynomial $x^d - a_1x^{d-1} - a_2x^{d-2} - \cdots - a_d$. But if A is nilpotent, its characteristic polynomial must be x^d and hence the first row of A must be zero. □

Proof of Theorem 11.2. Consider the set I of all polynomials $P(x) \in \mathbb{C}[x]$ for which there exist positive integers a and b with $0 < a < b$ such that

$$P(x)F(x) \in \sum_{j=a}^{b} \mathbb{C}[x]F\left(x^{k^j}\right).$$

We note that I is an ideal of $\mathbb{C}[x]$. Let $P_0(x)$ be a generator for I. It follows from assumption (b) that $P_0(0) \neq 0$ and we can assume without loss of generality that $P_0(0) = 1$. Let us assume that α is a root of $P_0(x)$ with the property that $\alpha^{k^{i_0}} = \alpha$ for some positive integer i_0. We will obtain a contradiction from this assumption, which will prove the theorem.

Step 1 (*preliminaries*). Since $F(x)$ is k-Mahler, it is also k^{i_0}-Mahler and hence $F(x)$ satisfies a non-trivial polynomial equation of the form

$$\sum_{j=0}^{d} Q_j(x)F\left(x^{k^{i_0 j}}\right) = 0$$

with $Q_0, \ldots, Q_d$ polynomials and $Q_0(x)Q_d(x) \neq 0$. We pick such a nontrivial relation with Q_0 nonzero and the degree of Q_0 minimal. By assumption P_0 divides Q_0 and so α is a root is of $Q_0(x)$. The minimality of the degree of Q_0 also implies that $\gcd(Q_0(x), \ldots, Q_d(x)) = 1$. By Lemma 6.1, there exists some natural number N such that $F(x)$ can be decomposed as $F(x) = T(x) + x^N F_0(x)$, where $T(x)$ is a polynomial of degree $N - 1$ and $F_0(x)$ is a power series with nonzero constant term such that $F_0(x)$ satisfies a k^{i_0}-Mahler equation of the form

$$\sum_{j=0}^{e} \widetilde{Q}_j(x)F_0\left(x^{k^{i_0 j}}\right) = 0 \tag{11.2}$$

with $\widetilde{Q}_0(0) = 1$, $\widetilde{Q}_0(\alpha) = 0$ and $\widetilde{Q}_{j_0}(\alpha) \neq 0$ for some integer j_0, $0 < j_0 \leq e$. Moreover, by picking N sufficiently large, we may assume that $F_0(x)$ satisfies a nontrivial ℓ-Mahler equation

$$\sum_{j=0}^{f} R_j(x)F_0\left(x^{\ell^j}\right) = 0$$

for some polynomials $R_j(x)$ with $R_0(0) = 1$. Now, we infer from Proposition 7.10 that there is some ℓ-Becker power series $G(x)$ such that

$$F_0(x) = \left(\prod_{j=0}^{\infty} R_0\left(x^{\ell^j}\right)\right)^{-1} G(x), \tag{11.3}$$

and that there is some k-Becker power series $H(x)$ such that

$$F_0(x) = \left(\prod_{j=0}^{\infty} \widetilde{Q}_0\left(x^{k^{i_0 j}}\right)\right)^{-1} H(x)\,. \tag{11.4}$$

Step 2 (*Choice of the parameter* r). For $j = 0, \ldots, e$, we let c_j denote the order of vanishing of $\widetilde{Q}_j(x)$ at α, with the convention that $c_j = \infty$ if $\widetilde{Q}_j(x) = 0$. We note that by assumption $0 < c_0 < \infty$ and $c_{j_0} = 0 < c_0$. Let

$$b := \max\left\{\frac{c_0 - c_j}{j} \mid j = 1, \ldots, e\right\}\,. \tag{11.5}$$

Since at least one of $c_1, \dots, c_d$ is strictly less than c_0, we have that b is positive. Moreover, by definition there is some $j_1 \in \{1, \dots, e\}$ such that $c_{j_1} + bj_1 - c_0 = 0$. Then, for $j \in \{0, \dots, e\}$, we set

$$S_j(x) := \widetilde{Q}_j(x) \left(\prod_{n=0}^{j-1} \left(1 - \alpha^{-1} x^{k^{i_0 n}}\right)^b \right) \left(1 - \alpha^{-1} x\right)^{-c_0}. \tag{11.6}$$

Note that (11.5) implies that $S_0(x)$ is a polynomial in $\mathbb{C}[x]$ such that $S_0(0) = 1$ and $S_0(\alpha) \neq 0$.

Now, we set

$$L(x) := H(x) \left(\prod_{j=0}^{\infty} \left(1 - \alpha^{-1} x^{k^{i_0 j}}\right)^{b-c_0} \right). \tag{11.7}$$

In other words, we choose $r := b - c_0$. Then we infer from Equalities (11.3), (11.4), (11.6), and (11.7) that

$$\prod_{j=0}^{\infty} \left(1-\alpha^{-1} x^{k^{i_0 j}}\right)^{-b} = \left(\prod_{j=0}^{\infty} R_0\left(x^{\ell^j}\right) \right)^{-1} G(x) \left(\prod_{j=0}^{\infty} S_0\left(x^{k^{i_0 j}}\right) \right) L(x)^{-1}. \tag{11.8}$$

Step 3 (*Upper bound for* $|L(x)|^{-1}$). We first infer from (11.2) and (11.7) that the function L satisfies the following relation:

$$\sum_{n=0}^{e} \widetilde{Q}_n(x) \left(\prod_{j=n}^{\infty} S_0\left(x^{k^{i_0 j}}\right)^{-1} \right) \left(\prod_{j=n}^{\infty} \left(1 - \alpha^{-1} x^{k^{i_0 j}}\right)^{-b} \right) L\left(x^{k^{i_0 n}}\right) = 0,$$

which gives by (11.6) that

$$\begin{aligned} L(x) &= -\sum_{n=1}^{e} \left(\widetilde{Q}_n(x) \widetilde{Q}_0(x)^{-1} \prod_{j=0}^{n-1} S_0\left(x^{k^{i_0 j}}\right) \prod_{j=0}^{n-1} \left(1 - \alpha^{-1} x^{k^{i_0 j}}\right)^b \right) L\left(x^{k^{i_0 n}}\right) \\ &= -\sum_{n=1}^{e} \left(S_n(x) \prod_{j=1}^{n-1} S_0\left(x^{k^{i_0 j}}\right) \right) L\left(x^{k^{i_0 n}}\right). \end{aligned}$$

Let $A(x)$ denote the $e \times e$ matrix whose (i, j)-entry is $\delta_{i,j+1}$ if $i \geq 2$ and whose $(1, j)$-entry is

$$C_j(x) := -S_n(x) \prod_{j=1}^{n-1} S_0\left(x^{k^{i_0 j}}\right)$$

for $j = 1, \dots, e$. Then the previous computation gives us the following functional equation:

$$\left[L(x), L\left(x^{k^{i_0}}\right), \dots, L\left(x^{k^{i_0(e-1)}}\right)\right]^T = A(x)\left[L\left(x^{k^{i_0}}\right), \dots, L\left(x^{k^{i_0 e}}\right)\right]^T, \tag{11.9}$$

where ^{T} denotes the transpose.

In order to obtain the desired upper bound (namely, Inequality (11.11) that will be stated in the sequel), we are going to apply Proposition 10.1. We thus start by showing that the vector $[L(x), L(x^{k^{i_0}}), \ldots, L(x^{k^{i_0(e-1)}})]^T$ and the matrix $A(x)$ satisfy the assumptions (i), (ii), (iii), and (iv) of this proposition. We first note that $L(x)$ is not identically zero since $F(x)$ is not a polynomial. Furthermore, we assume that L is not a nonzero constant since otherwise the desired upper bound (11.11) would be immediately satisfied.

(i) By definition,

$$S_n(x) = \widetilde{Q}_n(x)\left(\prod_{j=0}^{n-1}\left(1-\alpha^{-1}x^{k^{i_0 j}}\right)^b\right)(1-\alpha^{-1}x)^{-c_0}.$$

Moreover, a simple computation gives that

$$\prod_{j=0}^{n-1}\left(1-\alpha^{-1}x^{k^{i_0 j}}\right)^b = \left(1-\alpha^{-1}x\right)^{bn} P_n(x)^b,$$

for some polynomial $P_n(x)$ that does not vanish at α. By definition of c_n, this shows that

$$S_n(x) = \left(1-\alpha^{-1}x\right)^{c_n+bn-c_0} P_n(x)^b R_n(x)\,, \tag{11.10}$$

where $P_n(x)$ and $R_n(x)$ are two polynomials that do not vanish at α. By the definition of b in (11.5), we have $c_n + bn - c_0 \geq 0$ for $n \in \{0, \ldots, e\}$, and thus $S_n(x)$ is analytic in the open unit disk and continuous on the closed unit disk. Since the finite product $\prod_{j=1}^{n-1} S_0(x^{k^{i_0 j}})$ is a polynomial, this shows that the entries of the matrix $A(x)$ are analytic on $B(0,1)$ and continuous on $\overline{B(0,1)}$;

(ii) As already observed, there is some integer j_1, $1 \leq j_1 \leq e$, such that $c_{j_1} + bj_1 - c_0 = 0$. Since $P_{j_1}(\alpha)R_{j_1}(\alpha) \neq 0$, Equation (11.10) implies that $S_{j_1}(\alpha) \neq 0$. On the other hand, we have that $\prod_{j=0}^{j_1-1} S_0(x^{k^{i_0 j}})$ does not vanish at α since $S_0(\alpha) \neq 0$ and $\alpha^{k^{i_0}} = \alpha$. We thus obtain that the $(1, j_1)$-entry of $A(\alpha)$ is nonzero. By Lemma 11.3, this implies that $A(\alpha)$ is not nilpotent;

(iii) By definition of the matrix A, we get that

$$\det A(x) = (-1)^e C_e(x) = (-1)^{e+1} S_e(x) \prod_{n=1}^{e-1} S_0\left(x^{k^{i_0 n}}\right).$$

By (11.10), we have that $S_e(x) = (1-\alpha^{-1}x)^{c_e+be-c_0} P_e(x)^b R_e(x)$, where $P_e(x)$ and $R_e(x)$ are polynomials. It follows that there exist two positive real numbers δ and M such that

$$|\det A(x)| > (1-|x|)^M$$

for every x such that $1-\delta < |x| < 1$;

(iv) We claim that

$$\left\{\left[L(x), L\left(x^{k^{i_0}}\right), \ldots, L\left(x^{k^{i_0(e-1)}}\right)\right]^T \mid x \in B(0,1)\right\}$$

cannot be contained in a proper subspace of $\mathbb{C}^e$. Indeed, if it were, then there would exist some nonzero row vector u such that

$$u\left[L(x), L\left(x^{k^{i_0}}\right), \ldots, L\left(x^{k^{i_0(e-1)}}\right)\right]^T = 0$$

for all $x \in B(0,1)$. But this would give that $L(x), \ldots, L(x^{k^{i_0(e-1)}})$ are linearly dependent over $\mathbb{C}$, and hence by Lemma 7.9, we would obtain that $L(x)$ is a constant function, a contradiction.

It follows from (i), (ii), (iii) and (iv) that we can apply Proposition 10.1 to the vector $[L(x), L(x^{k^{i_0}}), \ldots, L(x^{k^{i_0(e-1)}})]^T$. From now on, we fix a positive integer N_0 that will be assume to be large enough in step 4. Let μ be a primitive p^n-th root of unity with $n \geq N_0 + i_0(e-1)\nu_p(k)$. Here, $\nu_p(k)$ denotes the p-adic valuation of k and p is the prime number from assumption (a). By Proposition 10.1, for every $\theta \in [-1,1]$, there exist a positive integer M_0 and an infinite sequence $(t_\theta(n))_{n\geq 0} \in X_\theta \setminus \{1\}$ (denoted by $(t(n))_{n\geq 0}$ for short) which tends to 1 such that

$$\left\|\left[L(t(n)\alpha\mu), L\left(t(n)^{k^{i_0}}\alpha\mu^{k^{i_0}}\right), \ldots, L\left(t(n)^{k^{i_0(e-1)}}\alpha\mu^{k^{i_0(e-1)}}\right)\right]^T\right\| > |1-t(n)|^{M_0},$$

for every nonnegative integer n. By the pigeonhole principle, we can find an integer $n_0 \geq N_0$, a primitive p^{n_0}-th root of unity ζ_0, such that for every $\theta \in [-1,1]$ there exist a sequence $(s_\theta(n))_{n\geq 0}$ in $X_\theta \setminus \{1\}$ which tends to 1, and a positive integer A_1 (depending upon θ) satisfying

$$|L(s_\theta(n)\alpha\zeta_0)|^{-1} < |1-s_\theta(n)|^{-A_1} \tag{11.11}$$

for every positive integer n.

Remark 11.4. We fix the p^{n_0}-th root of unity ζ_0 once for all.

Step 4 (*Upper bound for* $|(\prod_{j\geq 0} R_0(x^{\ell^j}))^{-1}|$). From assumption (a), we get that if N_0 is large enough, then $R_0((\alpha\zeta_0)^{\ell^j}) \neq 0$ for every $j \geq 0$. Let n_1 and n_2, $n_1 < n_2$, be two positive integers such that

$$(\alpha\zeta_0)^{\ell^{n_1}} = (\alpha\zeta_0)^{\ell^{n_2}}. \tag{11.12}$$

Then for every $\theta \in [-1,1]$ and $t \in X_\theta \setminus \{0,1\}$ we have

$$\prod_{j=0}^{\infty} R_0\left((t\alpha\zeta_0)^{\ell^j}\right) = \prod_{j=0}^{n_1-1} R_0\left((t\alpha\zeta_0)^{\ell^j}\right) \prod_{i=n_1}^{n_2-1}\prod_{j=0}^{\infty} R_0\left(\left((t\alpha\zeta_0)^{\ell^i}\right)^{\ell^{j(n_2-n_1)}}\right).$$

Note that $\prod_{j=0}^{n_1-1} R_0(x^{\ell^j})$ is a polynomial that does not vanish at any point of the finite set $\left\{(\alpha\zeta_0)^{\ell^j}) \mid j \geq 0\right\}$. It follows that, for every $\theta \in [-1, 1]$, there exist two positive real numbers C_1 and ε_1 such that

$$\left| \left(\prod_{j=0}^{n_1-1} R_0\Big(t\alpha\zeta_0)^{\ell^j} \Big) \right)^{-1} \right| < C_1 ,$$

for all $t \in X_{\theta,\varepsilon_1}$. Furthermore, Equality (11.12) implies that for every integer i, $n_1 \leq i \leq n_2 - 1$, we have

$$\left((\alpha\zeta_0)^{\ell^i} \right)^{\ell^{j(n_2-n_1)}} = \left((\alpha\zeta_0)^{\ell^i} \right) .$$

Thus, for every integer i, $n_1 \leq i \leq n_2 - 1$, we can apply Proposition 9.2 to the infinite product

$$\left(\prod_{j=0}^{\infty} R_0 \left(\left((t\alpha\zeta_0)^{\ell^i} \right)^{\ell^{j(n_2-n_1)}} \right) \right)^{-1} .$$

This implies the existence of a cocountable subset Y_1 of $[-1, 1]$ such that for each $\theta \in Y_1$, there is a positive real number ε_2 and a positive integer A_2, both of which depend upon θ, such that

$$\left| \left(\prod_{j=0}^{\infty} R_0 \left((t\alpha\zeta_0)^{\ell^j} \right) \right)^{-1} \right| < |1 - t|^{-A_2} \tag{11.13}$$

for every $t \in X_{\theta,\varepsilon_2}$.

Step 5 (*Upper bound for* $|G(x)|$). Note first that, since $G(x)$ is a ℓ-Becker power series, Theorem 7.6 implies that $G(x)$ is ℓ-regular. By Proposition 7.5, there exist two positive real numbers C and m such that

$$|G(x)| < C(1 - |x|)^{-m} ,$$

for every complex number x in the open unit disk. This implies that there exist two positive real numbers ε_3 and A_3 such that

$$|G(x)| < (1 - |x|)^{-A_3} \tag{11.14}$$

for every complex number x with $1 - \varepsilon_3 < 1 - |x| < 1$.

Step 6 (*Upper bound for* $|\prod_{j\geq 0} S_0(x^{k^{i_0 j}})|$). First note that since $\alpha^{k^{i_0}} = \alpha$, $S_0(0) = 1$ and α is not a root of S_0, we can apply Proposition 9.2. We thus obtain the existence of a cocountable subset $Y_2 \subseteq [-1, 1]$ such that for each $\theta \in Y_2$, there is some positive real numbers δ_0 and a positive integer M_0 such that

$$\left|\prod_{j=0}^{\infty} S_0\Big((t\alpha)^{k^{i_0 j}}\Big)\right| < |1-t|^{M_0} \tag{11.15}$$

for every $t \in X_{\theta,\delta_0}$. Henceforth, we assume that we have selected $\theta \in Y_1 \cap Y_2$ and we assume that Equations (11.13) and (11.15) hold – this holds precisely when $t \in X_{\theta,\varepsilon_2} \cap X_{\theta,\delta_0} = X_{\theta,\min(\epsilon_2,\delta_0)}$.

We also note that $(\alpha\zeta_0)^{k^{i_0 j}} = \alpha$ for all $j \geq n_0$. This implies that

$$\prod_{j=0}^{\infty} S_0\Big((t\alpha\zeta_0)^{k^{i_0 j}}\Big) = R(t) \prod_{j=0}^{\infty} S_0\Big((t\alpha)^{k^{i_0 j}}\Big), \tag{11.16}$$

where

$$R(t) = \left(\prod_{j=0}^{n_0-1} S_0\Big((t\alpha\zeta_0)^{k^{i_0 j}}\Big)\right)\left(\prod_{j=0}^{n_0-1} S_0\Big((t\alpha)^{k^{i_0 j}}\Big)\right)^{-1}.$$

Since $\alpha^{k^{i_0 j}} = \alpha$ and $S_0(\alpha) \neq 0$, then, for every $\theta \in Y_2$, there are two positive real numbers δ_1 and C_2 such that

$$|R(t)| < C_2 \tag{11.17}$$

for every $t \in X_{\theta,\delta_1}$.

We thus infer from (11.15), (11.16), and (11.17) that for every $\theta \in Y_2$ there exist a positive real number ε_4 and a positive integer A_4, both of which depend upon θ, such that

$$\left|\prod_{j=0}^{\infty} S_0\Big((t\alpha\zeta_0)^{\ell^j}\Big)\right| < |1-t|^{-A_4} \tag{11.18}$$

for $t \in X_{\theta,\varepsilon_4}$.

Step 7 (*Conclusion*). Set

$$\Pi(x) := \left(\prod_{j=0}^{\infty} R_0\Big(x^{\ell^j}\Big)\right)^{-1} G(x) \left(\prod_{j=0}^{\infty} S_0\Big(x^{k^{i_0 j}}\Big)\right) L(x)^{-1}.$$

Let us fix a real number $\theta \in Y_1 \cap Y_2$. Collecting all the upper bounds obtained in (11.14), (11.13), (11.18), and (11.11), we obtain that

$$|\Pi(s_\theta(n))\alpha\zeta_0)| < |1 - s_\theta(n))|^{-(A_1+A_2+A_3+A_4)}$$

for every integer n large enough. We thus infer from Equality (11.8) that

$$\left|\prod_{j=0}^{\infty}\left(1-(s_\theta(n)\zeta_0)^{k^{i_0 j}}\right)^{-b}\right| = |\Pi(s_\theta(n)\alpha\zeta_0)| < |1-s_\theta(n)|^{-(A_1+A_2+A_3+A_4)}$$

for every integer n large enough. But this contradicts Lemma 9.3, since $\zeta_0^{k^j} = 1$ for all sufficiently large j. This concludes the proof. $\square$

12. Existence of good prime ideals

In this section we prove the following result.

Theorem 12.1. *Let R be a ring of the form $\mathcal{O}_K[1/M]$, where K denotes a number field and M denotes a positive integer. Let $P(x)$, $Q(x) \in R[x]$ be two polynomials with $P(0) = Q(0) = 1$ and such that none of the zeros of $P(x)Q(x)$ are roots of unity. Let k and l be two integers, $k, l \geq 2$, for which: there exists a prime p such that $p|k$ and p does not divide ℓ, and there exists a prime q such that $q|\ell$ and q does not divide k. Then there are infinitely many prime ideals $\mathfrak{P}$ in R such that*

$$\left(\prod_{i=0}^{\infty} P\left(x^{k^i}\right)\right)^{-1} \bmod \mathfrak{P} \quad \textit{and} \quad \left(\prod_{i=0}^{\infty} Q\left(x^{\ell^i}\right)\right)^{-1} \bmod \mathfrak{P}$$

are respectively k- and ℓ-automatic power series in $(R/\mathfrak{P})[[x]]$.

We do not know whether the conclusion to the statement of Theorem 12.1 holds if we allow P or Q to vanish at roots of unity, but we suspect that the statement is false in this setting.

Our proof is based on Chebotarev's density theorem for which we refer the reader for example to [22] and to the informative survey [23]. We first prove three auxiliary results.

Lemma 12.2. *Let K be a number field and let α be a nonzero element in K that is not a root of unity. Then for all sufficiently large natural numbers n the equation $\beta^n = \alpha$ has no solution $\beta \in K$.*

Proof. Let $\mathcal{O}_K$ be the ring of integers of K. Each nonzero prime ideal π of $\mathcal{O}_K$ gives rise to a rank one discrete valuation ν_π of the field K. Notice that if $\beta^n = \alpha$ then $\nu_\pi(\alpha) = n\nu_\pi(\beta)$. In particular, if there exists some prime π for which $\nu_\pi(\alpha)$ is nonzero then we see that, in the equation $\beta^n = \alpha$, n must divide $\nu_\pi(\alpha)$ and we get the result. We may write $\alpha = a/b$ with $a, b \in \mathcal{O}_K$, nonzero. Notice that since $\mathcal{O}_K$ is a Dedekind domain, the ideals (a) and (b) must factor into prime ideals. Now if (a) or (b) are different ideals, then there must be some nonzero prime ideal π of R for which the induced valuation of $\alpha = a/b$ is nonzero. The previous remark thus shows that we must have $(\alpha) = \alpha\mathcal{O}_K = \mathcal{O}_K$. We thus may assume without loss of generality that α is a unit in R. But if $\beta^n = \alpha$ then, since $\mathcal{O}_K$ is integrally closed, we must have $\beta \in \mathcal{O}_K$ and β must be a unit. By Dirichlet's unit theorem, the group of units of $\mathcal{O}_K$ is a finitely generated abelian group. Hence if $\beta^n = \alpha$ for infinitely many n, then α must be a torsion element of the units group. That is, α must be a root of unity, which ends the proof. $\square$

Lemma 12.3. *Let m be a natural number and let $d_1, \ldots, d_m$ be positive integers. Suppose that H is a subgroup of $\prod_{i=1}^m (\mathbb{Z}/d_i\mathbb{Z})$ with the property that there exist natural numbers $r_1, \ldots, r_m$ with*

$$1/r_1 + \cdots + 1/r_m < 1$$

such that for each $i \in \{1, \ldots, m\}$, there is an element $h_i \in H$ whose i-th coordinate has order r_i. Then there is an element $h \in H$ such that no coordinate of h is equal to zero.

Proof. For each $i \in \{1, \ldots, m\}$, we let

$$\pi_i : \prod_{i=1}^m (\mathbb{Z}/d_i\mathbb{Z}) \to \mathbb{Z}/d_i\mathbb{Z}$$

denote the projection onto the i-th coordinate. Given $(x_1, \ldots, x_m) \in \mathbb{Z}^m$ we have that $x_1h_1 + \cdots + x_mh_m \in H$. Observe that the density of integers y for which

$$\pi_i\left(\sum_{j \neq i} x_jh_j + yh_i\right) = 0$$

is equal to $1/r_i$. Since this holds for all $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m) \in \mathbb{Z}^{m-1}$, we see that the density of $(x_1, \ldots, x_m) \in \mathbb{Z}^m$ for which

$$\pi_i\left(\sum_{j=1}^m x_jh_j\right) = 0$$

is equal to $1/r_i$. Thus the density of $(x_1, \ldots, x_m) \in \mathbb{Z}^m$ for which

$$\pi_i\left(\sum_{j=1}^m x_jh_j\right) = 0$$

holds for some $i \in \{1, \ldots, m\}$ is at most

$$1/r_1 + \cdots + 1/r_m < 1\,.$$

In particular, we see that there is some $(x_1, \ldots, x_m) \in \mathbb{Z}^m$ such that the element $h := x_1h_1 + \cdots + x_mh_m \in H$ has no coordinate equal to zero. $\square$

Lemma 12.4. *Let $k \geq 2$ be an integer, let R be a ring of the form $\mathcal{O}_K[1/M]$, where K denotes a number field and M denotes a positive integer, let $\mathfrak{P}$ be a nonzero prime ideal of R, and let a be an element of R. Suppose that for some natural number n, the polynomial $1 - ax^{k^n} \bmod \mathfrak{P}$ has no roots in $R/\mathfrak{P}$. Then the infinite product*

$$\left(\prod_{j=0}^{\infty}\left(1 - ax^{k^j}\right)\right)^{-1} \bmod \mathfrak{P}$$

is a k-automatic power series in $(R/\mathfrak{P})[[x]]$.

Proof. Set $F(x) := \prod_{j=0}^{\infty}(1 - ax^{k^j})^{-1} \bmod \mathfrak{P}$. Without loss of generality we can assume that a does not belong to $\mathfrak{P}$. Let us first note that the sequence $a, a^k, a^{k^2}, \ldots$ is necessarily eventually periodic modulo $\mathfrak{P}$. However, it cannot be periodic, as otherwise the polynomial $1 - ax^{k^n}$ would have a root for every natural number n. Thus there exists a positive integer N such that

$$a \not\equiv a^{k^N} \equiv a^{k^{2N}} \mod \mathfrak{P}.$$

Set $b := a^{k^N}$ and let us consider the polynomial

$$Q(x) := (1 - bx)\big(1 - bx^k\big) \cdots \left(1 - bx^{k^{N-1}}\right).$$

Now arguing exactly as in the proof of Proposition 7.8, we see that there exists a polynomial $S(x) \in R[x]$ such that $G(x) := Q(x)^{-1}F(x)$ satisfies the equation

$$G(x) \equiv S(x)G\big(x^k\big) \mod \mathfrak{P}.$$

Thus Theorem 7.6 implies that $G(x) \bmod \mathfrak{P}$ is a k-regular power series in $(R/\mathfrak{P})[[x]]$. By Proposition 7.4, we see that $F(x) \bmod \mathfrak{P}$ is a k-regular power series since it is a product of a polynomial (which is k-regular) and a k-regular power series. Since the base field is finite, Proposition 7.4 gives that $F(x) \bmod \mathfrak{P}$ is actually a k-automatic power series. This ends the proof. □

Proof of Theorem 12.1. By assumption R is of the form $\mathcal{O}_K[1/M]$, where K denotes a number field and M denotes a positive integer. Let L be the Galois extension of K generated by all complex roots of the polynomial $P(x)Q(x)$. Thus there are $\alpha_1, \ldots, \alpha_d, \beta_1, \ldots, \beta_e \in L$ such that $P(x) = (1-\alpha_1 x)\cdots(1-\alpha_d x)$ and $Q(x) = (1 - \beta_1 x)\cdots(1 - \beta_e x)$. By assumption there is a prime p that divides k but does not divide ℓ and a prime q that divides ℓ but does not divide k. Let s be a natural number such that p^s and q^s are both larger than $d + e$. Since by assumption none of the roots of $P(x)Q(x)$ is a root of unity, Lemma 12.2 implies that, for $1 \leq i \leq d$ and $1 \leq j \leq e$, there are largest nonnegative integers n_i and m_j with the property that we can write $\alpha_i = \gamma_i^{p^{n_i}} u_i$ and $\beta_j = \delta_j^{q^{m_j}} v_j$ for some elements $\gamma_i, \delta_j \in L(e^{2\pi i/(p^s q^s)})$ and u_i, v_j roots of unity in $L(e^{2\pi i/(p^s q^s)})$.

Next let n denote a natural number that is strictly larger than the maximum of the n_i and the m_j for i and j such that $1 \le i \le d$ and $1 \le j \le e$. Set $E := L(e^{2\pi i/(p^n q^n)})$ and let F denote the Galois extension of E generated by all complex roots of the polynomial

$$\prod_{i=1}^{d}\prod_{j=1}^{e}\Big(x^{p^n}-\gamma_i\Big)\Big(x^{q^n}-\delta_j\Big).$$

For each i, $1 \le i \le d$, we pick a root $\gamma_{i,0}$ of $x^{p^n}-\gamma_i$, and for each j, $1 \le j \le e$, we pick a root $\delta_{j,0}$ of $x^{q^n}-\delta_j$.

Claim. We claim that for every integer i, $1 \le i \le d$, there is an automorphism σ_i in $\mathrm{Gal}(F/E)$ such that

$$\sigma_i(\gamma_{i,0}) = \gamma_{i,0}u\,,$$

with u a primitive p^r-th root of unity for some r greater than or equal to s. Similarly, for every integer j, $1 \le j \le e$, there is an automorphism τ_j in $\mathrm{Gal}(F/E)$ that such that

$$\tau_j(\delta_{j,0}) = \delta_{j,0}u'\,,$$

for some primitive $q^{r'}$-th root of unity u' with r' greater than or equal to s.

Proof of the claim. Note that

$$\left\{\frac{\sigma(\gamma_{i,0})}{\gamma_{i,0}} \;\mid\; \sigma \in \mathrm{Gal}(F/E)\right\}$$

forms a subgroup of the p^n-th roots of unity. To prove the claim we just have to prove that this group cannot be contained in the group of p^{s-1}-th roots of unity. Let us assume that this is the case. Then the product of the Galois conjugates of $\gamma_{i,0}$ must be $\tilde{\gamma}_i := \gamma_{i,0}^{p^t}v$ for some $t < s$ and some $p^{(s-1)}$-th root of unity v. Moreover, $\tilde{\gamma}_i$ lies in $L(e^{2\pi i/(p^n q^n)})$. Note that the Galois group of $L(e^{2\pi i/(p^n q^n)})$ over $L(e^{2\pi i/(p^s q^s)})$ has order dividing $\phi(p^n q^n)/\phi(p^s q^s) = p^{n-s}q^{n-s}$. Since all conjugates of $\tilde{\gamma}_i$ are equal to $\tilde{\gamma}_i$ times some root of unity, we see that the relative norm of $\tilde{\gamma}_i$ with respect to the subfield $L(e^{2\pi i/(p^s q^s)})$ is of the form $\tilde{\gamma}_i{}^d v'$ for some divisor d of $p^{n-s}q^{n-s}$ and some root of unity v'. Moreover,

$$\tilde{\gamma}_i{}^d v' \in L\Big(e^{2\pi i/(p^s q^s)}\Big).$$

Note that the gcd of d and p^{n-t} is equal to p^{n-s_0} for some integer $s_0 \ge s$. Since $\gamma_{i,0}^{p^n} = \tilde{\gamma}_i{}^{p^{n-t}} v^{-p^{n-t}} \in L(e^{2\pi i/(p^s q^s)})$, we see by expressing p^{n-s_0} as an integer linear combination of d and p^{n-t} that

$$\tilde{\gamma}_i{}^{p^{n-s_0}}\omega = \gamma_{i,0}^{p^{n-s_0+t}}\omega' \in L\Big(e^{2\pi i/(p^s q^s)}\Big)$$

for some roots of unity ω and ω' and some $s_0 \geq s$. But $s_0 - t \geq 1$ and so we see that α_i is equal to a root of unity times

$$\left(\gamma_{i,0}^{p^{n-s_0+t}} \omega'\right)^{p^{s_0-t+n_i}},$$

contradicting the maximality of n_i. This confirms the claim. $\square$

For an integer m, we let $\mathbb{U}_m$ denote the subgroup of $\mathbb{C}^*$ consisting of all m-th roots of unity. Note that we can define a group homomorhpism Φ from $\mathrm{Gal}(F/E)$ to $(\mathbb{U}_{p^n})^d \times (\mathbb{U}_{q^n})^e$ by

$$\Phi(\sigma) := (\sigma(\gamma_{1,0})/\gamma_{1,0}, \ldots, \sigma(\gamma_{d,0})/\gamma_{d,0}, \sigma(\delta_{1,0})/\delta_{1,0}, \ldots, \sigma(\delta_{e,0})/\delta_{e,0}).$$

We see that Φ is a group homomorphism since each $\sigma \in \mathrm{Gal}(F/E)$ fixes the p^n-th and q^n-th roots of unity. Set $H := \Phi(\mathrm{Gal}(F/E))$. The claim implies that the i-th coordinate in $(\mathbb{U}_{p^n})^d$ of $\Phi(\sigma_i)$ has order at least equal to p^s. Similarly, it also implies that the j-th coordinate in $(\mathbb{U}_{q^n})^e$ of $\Phi(\tau_j)$ has order at least equal to q^s. Since p^s and q^s are both greater than $d + e$, we have

$$d/p^s + e/q^s < 1.$$

Now, since $(\mathbb{U}_{p^n})^d \times (\mathbb{U}_{q^n})^e \cong (\mathbb{Z}/p^n\mathbb{Z})^d \times (\mathbb{Z}/q^n\mathbb{Z})^e$, we infer from Lemma 12.3 that there exists an element h in H such that every coordinate of h is different from the identity element. In other words, this means that there exists some element τ of $\mathrm{Gal}(F/E)$ that fixes no element in the set

$$\{\gamma_{i,0} \mid 1 \leq i \leq d\} \cup \{\delta_{j,0} \mid 1 \leq j \leq e\}.$$

Since by definition τ fixes all p^n-th and q^n-th roots of unity, we see more generally that no root of the polynomial

$$\prod_{i=1}^{d} \prod_{j=1}^{e} \left(x^{p^n} - \gamma_i\right)\left(x^{q^n} - \delta_j\right)$$

is fixed by τ. Since τ belongs to $\mathrm{Gal}(F/E)$, we can see τ as an element of $\mathrm{Gal}(F/K)$ that fixes all elements of E. We have thus produce an element τ of $\mathrm{Gal}(F/K)$ that fixes all roots of $P(x)Q(x)$ but that that does not fix any of the roots of the polynomial

$$\prod_{i=1}^{d} \prod_{j=1}^{e} \left(x^{p^n} - \gamma_i\right)\left(x^{q^n} - \delta_j\right).$$

It follows from Chebotarev's density theorem (see for instance the discussion in [23]) that there is an infinite set of nonzero prime ideals $\mathcal{S} \subseteq \mathrm{Spec}(R)$ such that if

$\mathfrak{P} \in \mathcal{S}$ then $P(x)Q(x) \bmod \mathfrak{P}$ factors into linear terms while the minimal polynomial of

$$\prod_{i=1}^{d}\prod_{j=1}^{e}\left(x^{p^n} - \gamma_i\right)\left(x^{q^n} - \delta_j\right)$$

over K has no root modulo $\mathfrak{P}$. In particular, there is a natural number N larger than n such that for all such prime ideals $\mathfrak{P}$, the polynomial $P(x)Q(x) \bmod \mathfrak{P}$ splits into linear factors, while the polynomial $P(x^{p^N})Q(x^{q^N}) \bmod \mathfrak{P}$ does not have any roots in $R/\mathfrak{P}$.

For such a prime ideal $\mathfrak{P}$, there thus exist $a_1, \dots, a_d, b_1, \dots, b_e$ in the finite field $R/\mathfrak{P}$ such that

$$P(x) \equiv (1 - a_1 x) \cdots (1 - a_d x) \mod \mathfrak{P}$$

and

$$Q(x) \equiv (1 - b_1 x) \cdots (1 - b_d x) \mod \mathfrak{P}\,.$$

Then

$$\left(\prod_{j=0}^{\infty} P\left(x^{k^j}\right)\right)^{-1} \equiv \prod_{i=1}^{d}\left(\prod_{j=0}^{\infty}\left(1 - a_i x^{k^j}\right)\right)^{-1} \mod \mathfrak{P}\,.$$

By Lemma 12.4 the right side is a product of k-automatic power series and hence, by Proposition 7.4, is k-automatic. Thus the infinite product

$$\left(\prod_{j=0}^{\infty} P\left(x^{k^j}\right)\right)^{-1} \mod \mathfrak{P}$$

is a k-automatic power series in $R/\mathfrak{P}[[x]]$. Similarly, we get that

$$\left(\prod_{j=0}^{\infty} Q\left(x^{\ell^j}\right)\right)^{-1} \equiv \prod_{i=1}^{e}\left(\prod_{j=0}^{\infty}\left(1 - b_i x^{\ell^j}\right)\right)^{-1} \mod \mathfrak{P}\,,$$

which implies that the infinite product

$$\left(\prod_{j=0}^{\infty} Q\left(x^{\ell^j}\right)\right)^{-1} \mod \mathfrak{P}$$

is a ℓ automatic power series in $R/\mathfrak{P}[[x]]$. This concludes the proof. $\square$

13. Proof of Theorem 1.3

We are now ready to prove our main result.

Proof of Theorem 1.3. Let K be a field of characteristic zero and let k and l be two multiplicatively independent positive integers.

We first note that if $F(x) \in K[[x]]$ is a rational function, then for every integer $m \geq 2$, it obviously satisfies a functional equation as in (1.3) with $n = 0$. Hence, $F(x)$ is m-Mahler, which gives a first implication.

To prove the converse implication, we fix $F(x) \in K[[x]]$ that is both k- and ℓ-Mahler and we aim at proving that $F(x)$ is a rational function. Of course, if $F(x)$ is a polynomial, there is nothing to prove. From now on, we thus assume that $F(x)$ is not a polynomial. By Corollary 8.3, we can assume that there are primes p and q such that p divides k while p does not divide ℓ and such that q divides ℓ while q does not divide k. By Theorem 5.1, we can assume that there is a ring R of the form $\mathcal{O}_K[1/M]$ (where K is a number field and M is a positive integer), such that $F(x) \in R[[x]]$ and satisfies the equations

$$\sum_{i=0}^{n} P_i(x)F(x^{k^i}) = 0$$

with $P_0, \ldots, P_d \in R[x]$ and

$$\sum_{i=0}^{m} Q_i(x)F\left(x^{\ell^i}\right) = 0$$

with $Q_0, \ldots, Q_e \in R[x]$. Without loss of generality, we can assume that all complex roots of $P_0(x)$ and $Q_0(x)$ belong to R (otherwise we could just enlarge R by adjoining these numbers). Furthermore, we can assume that $P_0(x)Q_0(x) \neq 0$. By Corollary 6.2, we can also assume that $P_0(0) = 1$ and that $Q_0(0) = 1$, for otherwise we could just replace $F(x)$ by the power series $F_0(x)$ given there. We choose a ring embedding of R in $\mathbb{C}$ and for the moment we regard $F(x)$ as a complex power series. By Theorem 11.2, we can assume that if α is a root of unity such that $\alpha^{k^j} = \alpha$ for some positive integer j, then $P_0(\alpha) \neq 0$. Similarly, we can assume that if β is a root of unity such that $\beta^{\ell^j} = \beta$ for some positive integer j, then $Q_0(\beta) \neq 0$.

By Proposition 7.10, we can write

$$F(x) = \left(\prod_{j=0}^{\infty} P_0\left(x^{k^j}\right)\right)^{-1} G(x)\,,$$

for some k-regular power series $G(x) \in R[[x]]$. Furthermore, we can decompose $P_0(x)$ as $P_0(x) = S_0(x)S_1(x)$, where $S_0(x)$ and $S_1(x)$ are two polynomials, the zeros of $S_0(x)$ are all roots of unity, none of the zeros of $S_1(x)$ is a root of unity, and $S_0(0) = S_1(0) = 1$. Since by assumption all roots of $P_0(x)$ lie in R, we get that both $S_0(x)$ and $S_1(x)$ belong to $R[x]$. By assumption if α is a root of $S_0(x)$ then for every positive integer j, one has $\alpha^{k^j} \neq \alpha$. Then, it follows from Proposition 7.8 that

$$\left(\prod_{j=0}^{\infty} S_0\left(x^{k^j}\right)\right)^{-1} \in R[[x]]$$

is a k-regular power series. Set $H(x) := \left(\prod_{j=0}^{\infty} S_0(x^{k^j})\right)^{-1} G(x)$. We infer from part (iii) of Proposition 7.4 that $H(x)$ is a k-regular power series. Moreover, one has

$$F(x) = \left(\prod_{j=0}^{\infty} S_1\left(x^{k^j}\right)\right)^{-1} H(x)\,. \tag{13.1}$$

Similarly, by Proposition 7.10, we can write

$$F(x) = \left(\prod_{j=0}^{\infty} Q_0\left(x^{k^j}\right)\right)^{-1} I(x)\,,$$

for some k-regular power series $I(x) \in R[[x]]$. As previously, we can decompose $Q_0(x)$ as $Q_0(x) = T_0(x)T_1(x)$, where $T_0(x)$ and $T_1(x)$ belong to $R[x]$, the zeros of $T_0(x)$ are all roots of unity, none of the zeros of $T_1(x)$ are roots of unity, and $T_0(0) = T_1(0) = 1$. By assumption if β is a root of $T_0(x)$ then for every positive integer j, one has $\beta^{\ell^j} \neq \beta$. Then it follows from Proposition 7.8 that

$$\left(\prod_{j=0}^{\infty} T_0\left(x^{\ell^j}\right)\right)^{-1} \in R[[x]]$$

is a ℓ-regular power series. Set $J := \prod_{j=0}^{\infty} T_0(x^{k^j})^{-1} I(x)$. Again, we see by Proposition 7.4 that $J(x)$ is ℓ-regular. Moreover, one has

$$F(x) = \left(\prod_{j=0}^{\infty} T_1\left(x^{k^j}\right)\right)^{-1} J(x)\,. \tag{13.2}$$

By Theorem 12.1, there is an infinite set of nonzero prime ideals $\mathcal{S}$ of R such that, for every prime ideal $\mathfrak{P}$ in $\mathcal{S}$,

$$\left(\prod_{j=0}^{\infty} S_1\left(x^{k^j}\right)\right)^{-1} \bmod \mathfrak{P}$$

is a k-automatic power series in $(R/\mathfrak{P})[[x]]$ and

$$\left(\prod_{j=0}^{\infty} T_1\left(x^{\ell^j}\right)\right)^{-1} \bmod \mathfrak{P}$$

is a ℓ-automatic power series in $(R/\mathfrak{P})[[x]]$. Then we infer from Equalities (13.1) and (13.2) that, for $\mathfrak{P} \in \mathcal{S}$, $F(x) \bmod \mathfrak{P}$ is k-regular for it is the product of two k-regular power series. Similarly, $F(x) \bmod \mathfrak{P}$ is a ℓ-regular power series.

We recall that since R is of the form $\mathcal{O}_K[1/M]$, it is a *Dedekind domain*; that is, it is a Noetherian normal domain of Krull dimension one. In particular, all nonzero prime ideals are maximal. Now since R is a finitely generated $\mathbb{Z}$-algebra and $\mathfrak{P}$ is a maximal ideal, the quotient ring $R/\mathfrak{P}$ is a finite field (see [18, Theorem 4.19, page 132]). By Proposition 7.4, this implies that $F(x) \bmod \mathfrak{P}$ is actually both k- and ℓ-automatic. By Cobham's theorem, we obtain that the sequence of coefficients of $F(x) \bmod \mathfrak{P}$ is eventually periodic and hence $F(x) \bmod \mathfrak{P}$ is a rational function.

Note that since $\mathcal{S}$ is infinite, the intersection of all ideals in $\mathcal{S}$ is the zero ideal (see [18, Lemma 4.16, page 130]). Moreover, $F(x) \bmod \mathfrak{P}$ is rational for every prime ideal $\mathfrak{P} \in \mathcal{S}$. Applying Lemma 5.4, we obtain that $F(x)$ is a rational function. This ends the proof. □

References

[1] B. ADAMCZEWSKI and Y. BUGEAUD, *On the complexity of algebraic numbers I. Expansions in integer bases*, Ann. of Math. **165** (2007), 547–565.

[2] B. ADAMCZEWSKI and Y. BUGEAUD, *Real and p-adic expansions involving symmetric patterns*, Int. Math. Res. Not. IMRN **2006** (2006), 17 pp.

[3] B. ADAMCZEWSKI and J. P. BELL, *Diagonalization and rationalization of algebraic Laurent series*, Ann. Sci. Éc. Norm. Supér. **46** (2013), 963–1004.

[4] B. ADAMCZEWSKI and C. FAVERJON *Méthode de Mahler: relations linéaires, transcendance et applications aux nombres automatiques*, Proc. Lond. Math. Soc. **115** (2017), 55–90.

[5] J.-P. ALLOUCHE, B. RANDÉ and L. THIMONIER, "Fonctions génératrices transcendantes à coefficients engendrés par automates", STACS 88 (Bordeaux 1988), Lecture Notes in Comput. Sci., Vol. 294, Springer, Berlin, 1988, 170–183.

[6] J.-P. ALLOUCHE and J. SHALLIT, "Automatic Sequences. Theory, Applications, Generalizations", Cambridge University Press, Cambridge, 2003.

[7] J.-P. ALLOUCHE and J. SHALLIT, *The ring of k-regular sequences*, Theoret. Comput. Sci. **98** (1992), 163–197.

[8] J.-P. ALLOUCHE and J. SHALLIT, *The ring of k-regular sequences, II*, Theoret. Comput. Sci. **307** (2003), 3–29.

[9] P.-G. BECKER, *k-Regular power series and Mahler-type functional equations*, J. Number Theory **49** (1994), 269–286.

[10] J. P. BELL, *A generalization of Cobham's theorem for regular sequences*, Sém. Lothar. Combin. 54A (2005/07), Art. B54Ap, 15 pp.

[11] A. COBHAM, *On the Hartmanis-Stearns problem for a class of tag machines*, In: "Conference Record of 1968 Ninth Annual Symposium on Switching and Automata Theory", Schenectady, New York, 1968, 51–60.

[12] A. COBHAM, *On the base-dependence of sets of numbers recognizable by finite automata*, Math. Systems Theory **3** (1969), 186–192.

[13] L. DENIS, *Indépendance algébrique de logarithmes en caractéristique p*, Bull. Aust. Math. Soc. **74** (2006), 461–470.

[14] PH. DUMAS, "Récurrences mahlériennes, suites automatiques, études asymptotiques", Thèse de Doctorat, Université Bordeaux I, Talence, 1993.

[15] PH. DUMAS and PH. FLAJOLET, *Asymptotique des récurrences mahlériennes: le cas cyclotomique*, J. Théor. Nombres Bordeaux **8** (1996), 1–30.
[16] F. DURAND, *Cobham's theorem for substitutions*, J. Eur. Math. Soc. (JEMS) **13** (2011), 1799–1814.
[17] S. EILENBERG, "Automata, Languages, and Machines", Vol. A, Academic Press, 1974.
[18] D. EISENBUD, "Commutative Algebra. With a View Toward Algebraic Geometry", Graduate Texts in Mathematics, Vol. 150, Springer-Verlag, New York, 1995.
[19] G. EVEREST, A. VAN DER POORTEN, I. SHPARLINSKI and T. WARD, "'Recurrence Sequences", Mathematical Surveys and Monographs, Vol. 104, American Mathematical Society, Providence, RI, 2003.
[20] H. FURSTENBERG, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1** (1967), 1–49.
[21] J. HARTMANIS and R. E. STEARNS, *On the computational complexity of algorithms*, Trans. Amer. Math. Soc. **117** (1965), 285–306.
[22] S. LANG, "Algebraic Number Theory", Graduate Texts in Mathematics, Vol. 110, Springer-Verlag, New York, 1994.
[23] H. W. LENSTRA and P. STEVENHAGEN, *Chebotarëv and his density theorem*, Math. Intelligencer **18** (1996), 26–37.
[24] J. H. LOXTON, *Automata and transcendence*, In: "New Advances in Transcendence Theory", Durham 1986, Cambridge University Press, 1988, 215–228.
[25] J. H. LOXTON and A. J. VAN DER POORTEN, *Arithmetic properties of the solutions of a class of functional equations*, J. Reine Angew. Math. **330** (1982), 159–172.
[26] J. H. LOXTON and A. J. VAN DER POORTEN, *Arithmetic properties of automata: regular sequences*, J. Reine Angew. Math. **392** (1988), 57–610.
[27] K. MAHLER, *Arithmetische Eigenschaften der Lösungen einer Klasse von Funktionalgleichungen*, Math. Ann. **101** (1929), 342–367.
[28] K. MAHLER, *Arithmetische Eigenschaften einer Klasse transzendental-transzendente Funktionen*, Math. Z. **32** (1930), 545–585.
[29] K. MAHLER, *Über das Verschwinden von Potenzreihen mehrerer Veränderlichen in speziellen Punktfolgen*, Math. Ann. **103** (1930), 573–587.
[30] K. MAHLER, *Some suggestions for further research*, Bull. Aust. Math. Soc. **29** (1984), 101–108.
[31] M. MENDÈS FRANCE, *Nombres algébriques et théorie des automates*, Enseign. Math. **26** (1980), 193–199.
[32] KU. NISHIOKA, *New approach in Mahler's method*, J. Reine Angew. Math. **407** (1990), 202–219.
[33] KU. NISHIOKA, *Algebraic independence measures of the values of Mahler's functions*, J. Reine Angew. Math. **420** (1991), 203–214.
[34] KU. NISHIOKA, *Algebraic independence by Mahler's method and S-unit equations*, Compos. Math. **92** (1994), 87–110.
[35] KU. NISHIOKA, "Mahler Functions and Transcendence", Lecture Notes in Math., Vol. 1631, Springer-Verlag, Berlin, 1997.
[36] F. PELLARIN, *Aspects de l'indépendance algébrique en caractéristique non nulle*, Astérisque 317 (2008) 205–242 (2008). Séminaire Bourbaki. Vol. 2006/2007, Exp. No. 973.
[37] F. PELLARIN, *An introduction to Mahler's method for transcendence and algebraic independence*, arXiv:1005.1216.
[38] P. PHILIPPON, *Groupes de Galois et nombres automatiques*, J. Lond. Math. Soc. (2) **92** (2015), 596–614.
[39] R. SCHÄFKE and M. SINGER, *Consistent systems of linear differential and difference equations*, J. Eur. Math. Soc. (JEMS), to appear.
[40] A. J. VAN DER POORTEN, *On the transcendence and algebraic independence of certain somewhat amusing numbers (results of Loxton and Van der Poorten)*, In: "Séminaire de Théorie des Nombres 1975–1976 (Univ. Bordeaux I, Talence)", Exp. n. 14, 13 pp.

[41] A. J. VAN DER POORTEN, *Remarks on automata, functional equations and transcendence*, In: “Séminaire de Théorie des Nombres de Bordeaux (1986-1987)”, Exp. n. 27, 11 pp.
[42] B. RANDÉ, *Récurrence* 2- *et* 3-*mahlériennes*, J. Théor. Nombres Bordeaux **5** (1993), 101–109.
[43] M. WALDSCHMIDT, “Diophantine Approximation on Linear Algebraic Groups”, Grundlehren der Mathematischen Wissenschaften, Vol. 326, Springer-Verlag, Berlin, 2000.
[44] M. WALDSCHMIDT, *Un demi-siècle de transcendance*, In: “Development of mathematics 1950-2000”, Birkhäuser, Basel, 2000, 1121–1186.
[45] D. ZAGIER, *Solution to advanced problem 6625*, Amer. Math. Monthly **99** (1992), 66–69.
[46] U. ZANNIER, *On a functional equation relating a Laurent series* $f(x)$ *to* $f(x^m)$, Aequationes Math. **55** (1998), 15–43.

Institut Camille Jordan
Université Claude Bernard Lyon 1
43 boulevard du 11 novembre 1918
69622 Villeurbanne Cedex, France
Boris.Adamczewski@math.cnrs.fr

Department of Pure Mathematics
University of Waterloo
Waterloo, ON, Canada
N2L 3G1
jpbell@uwaterloo.ca

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1357-1382

Multiplicative relations among singular moduli

JONATHAN PILA AND JACOB TSIMERMAN

Abstract. We consider some Diophantine problems of mixed modular-multiplicative type. In particular, we prove, for each $n \geq 1$, a finiteness result for n-tuples of singular moduli minimally satisfying a non-trivial multiplicative relation.

Mathematics Subject Classification (2010): 14G18 (primary); 03C64 (secondary).

1. Introduction

We consider some Diophantine problems of mixed modular-multiplicative type associated with the Zilber-Pink conjecture (ZP; see [4, 26, 35] and Section 2). Our results rely on the "modular Ax-Schanuel" theorem recently established by us [24].

Recall that a *singular modulus* is a complex number which is the j-invariant of an elliptic curve with complex multiplication; equivalently it is a number of the form $\sigma = j(\tau)$ where $j : \mathbb{H} \to \mathbb{C}$ is the elliptic modular function, $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ is the complex upper-half plane, and $\tau \in \mathbb{H}$ is a quadratic point (*i.e.* $[\mathbb{Q}(\tau) : \mathbb{Q}] = 2$).

Definition 1.1. An n-tuple $(\sigma_1, \ldots, \sigma_n)$ of distinct singular moduli will be called a *singular-dependent n-tuple* if the set $\{\sigma_1, \ldots, \sigma_n\}$ is multiplicatively dependent (*i.e.* $\prod \sigma_i^{a_i} = 1$ for some integers a_i not all zero), but no proper subset is multiplicatively dependent.

Theorem 1.2. *Let $n \geq 1$. There exist only finitely many singular-dependent n-tuples.*

The independence of proper subsets is clearly needed to avoid trivialities. The result is ineffective. Some examples (including a singular-dependent 5-tuple) can be found among the rational singular moduli (listed in [29, A.4]; see 6.3). Bilu-Masser-Zannier [3] show that there are no singular moduli with $\sigma_1\sigma_2 = 1$. This

JP was supported by an EPSRC grant entitled "O-minimality and Diophantine geometry", reference EP/J019232/1. JT was supported by an NSERC discovery grant.

Received February 3, 2015; accepted in revised form June 14, 2016.
Published online December 2017.

result is generalised by Bilu-Luca-Pizarro-Madariaga [2] to classify all solutions of $\sigma_1\sigma_2 \in \mathbb{Q}^\times$. Habegger [12] shows that only finitely many singular moduli are algebraic units.

In addition to the "modular Ax-Schanuel", we make use of isogeny estimates and other arithmetic ingredients, gathered in Section 6, and we require the following result showing that distinct rational "translates" of the j-function are multiplicatively independent modulo constants. To formulate it, recall that, for $g_1, g_2 \in \mathrm{GL}_2^+(\mathbb{Q})$, the functions $j(g_1 z)$, $j(g_2 z)$ are identically equal if and only if $[g_1] = [g_2]$ in $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2^+(\mathbb{Q})$; functions $f_1, \ldots, f_k : \mathbb{H} \to \mathbb{C}$ will be called *multiplicatively independent modulo constants* if there is no relation $\prod_{i=1}^k f_i^{n_i} = c$ where n_i are integers, not all zero, and $c \in \mathbb{C}$.

Theorem 1.3. *Let $g_1, \ldots, g_k \in \mathrm{GL}_2^+(\mathbb{Q})$. If the functions $j(g_1 z), \ldots, j(g_k z)$ are pairwise distinct then they are multiplicatively independent modulo constants.*

Theorem 1.3 is not predicted by ZP, nor would it follow from "Ax-Schanuel" for exp and j (see Section 3). But in view of Theorem 1.3, Theorem 1.2 is implied by ZP.

The Zilber-Pink setting is introduced in Section 2. After the proofs of 1.3 and 1.2 in Section 4 and Section 6, we discuss further ZP problems in the same setting in Section 7, Section 8, and Section 9, obtaining some partial results and some results conditional on certain "weakly bounded height conjectures" which we formulate in this setting. These are along the lines of a conjecture of Habegger [10] (see also [11, Appendix B]) in the modular setting, itself an analogue of the "Bounded Height Conjecture" for $(\mathbb{C}^\times)^n$ formulated by Bombieri-Masser-Zannier [4] and proved by Habegger [9].

ACKNOWLEDGEMENTS. We thank Gareth Jones for comments. We are grateful to the referee for corrections and careful comments as well as for raising the question addressed in Section 9.

2. The Zilber-Pink setting

We identify varieties and subvarieties with their sets of complex points (thus $Y(1)(\mathbb{C}) = \mathbb{C}$ and $\mathbb{G}_\mathrm{m}(\mathbb{C}) = \mathbb{C}^\times$). Varieties and subvarieties are assumed irreducible over $\mathbb{C}$.

For $m, n \in \mathbb{N} = \{0, 1, 2, \ldots\}$ set

$$X = X_{m,n} = Y(1)^m \times \mathbb{G}_\mathrm{m}^n.$$

Definition 2.1.

1. A *weakly special subvariety* of $Y(1)^m = X_{m,0} = \mathbb{C}^m$ is a subvariety of the following form. There is a "partition" $m_0, \ldots, m_k$ of $\{1, \ldots, m\}$, in which only

m_0 is permitted to be 0, but $k = 0$ is permitted such that $M = M_0 \times M_1 \times \ldots \times M_k$ where M_0 is a point in $\mathbb{C}^{m_0}$ (here $\mathbb{C}^{m_i}$ refers to the cartesian product of the coordinates contained in m_i, which is a subset of $\{1, \ldots, m\}$) and, for $i = 1, \ldots, k$, $M_i \subset \mathbb{C}^{m_i}$ is a modular curve;

2. A *special point* of $\mathbb{C}^m$ is a weakly special subvariety M of dimension zero (so $n_0 = \{1, \ldots, n\}$ and $M = M_0$) such that each coordinate of M is a singular modulus;
3. A *special subvariety* of $\mathbb{C}^m$ is a weakly special subvariety such that $m_0 = \emptyset$ or $M_0 \in \mathbb{C}^{m_0}$ is a special point. It is *strongly* special if $m_0 = \emptyset$;
4. A *weakly special subvariety* of $\mathbb{G}_{\mathrm{m}}^n = X_{0,n} = (\mathbb{C}^\times)^n$ is a coset of a subtorus, *i.e.* a subvariety defined by a finite system of equations $\prod x_i^{a_{ij}} = \xi_j$, $j = 1, \ldots, k$ where, for each j, $a_{ij} \in \mathbb{Z}$ are not all zero, $\xi_j \in \mathbb{C}^\times$ and the lattice generated by the exponent vectors $(a_{1j}, \ldots, a_{nj})$, $j = 1, \ldots, k$ is primitive;
5. A *special point* of $\mathbb{G}_{\mathrm{m}}^n$ is a torsion point;
6. A *special subvariety* of $\mathbb{G}_{\mathrm{m}}^n$ is a weakly special subvariety such that each ξ_j is a root of unity; *i.e.* it is a coset of a subtorus by a torsion point;
7. A *weakly special subvariety* of X is a product $M \times T$ where M, T are weakly special subvarieties of $Y(1)^m$, $\mathbb{G}_{\mathrm{m}}^n$, respectively, and likewise for a *special point* of X and *special subvariety* of X.

Definition 2.2. Let $W \subset X$ be a subvariety. A subvariety $A \subset W$ is called an *atypical component* (of W in X) if there is a special subvariety $T \subset X$ such that $A \subset W \cap T$ and

$$\dim A > \dim W + \dim T - \dim X.$$

The *atypical set* of W (in X) is the union of all atypical components (of W in X), and is denoted $\mathrm{Atyp}(W, X)$, or $\mathrm{Atyp}(W)$ if X is implicit from the context.

Variants of the following conjecture, in different settings, were made independently by Zilber [35], Bombieri-Masser-Zannier [4], and Pink [26].

Conjecture 2.3 (Zilber-Pink for X). Let $W \subset X$. Then $\mathrm{Atyp}(W)$ is a finite union of atypical components; equivalently, there are only finitely many maximal atypical components.

The full Zilber-Pink conjecture is the same statement about an arbitrary mixed Shimura variety (with its special subvarieties), and an algebraic subvariety $W \subset X$. In fact the above is the analogue of the statements in [4, 35] in the general setting considered by Pink, and is notionally stronger than the statement in [26]. For a general discussion of the conjecture see [34].

Definition 2.4. Let $A \subset X$ be a subvariety. We denote by $\langle A \rangle$ the smallest special subvariety containing A (which exists as it is just the intersection of all special subvarities containing A), and define the *defect* of A by

$$\delta(A) = \dim\langle A \rangle - \dim A.$$

Thus $A \subset W$ is atypical if $\delta(A) < \dim X - \dim W$, and W itself is atypical if $\langle W \rangle \neq X$.

Now in Conjecture 2.3 we only look for maximal atypical components, and we do not care if a larger atypical component contains a smaller but more atypical (*i.e.* smaller defect) one. But in fact the conjecture (taken over all special subvarieties of X) implies a formally stronger version (see [14, Proposition 2.4]).

Definition 2.5. A subvariety $W \subset V$ is called *optimal* for V if there is no strictly larger subvariety $W \subset W' \subset V$ with $\delta(W') \leq \delta(W)$.

Conjecture 2.6. Let $V \subset X$. Then V has only finitely many optimal subvarieties.

For a particular V and X, finding (or establishing the finiteness of) all optimal subvarieties could be more difficult than finding (or establishing the finiteness of) all maximal atypical subvarieties.

Now (as in [14]) we can repeat the same pattern of definitions with weakly special subvarieties instead of special ones. The smallest weakly special subvariety containing W we denote $\langle W\rangle_{\text{geo}}$, and we define the *geodesic defect* to be

$$\delta_{\text{geo}}(W) = \dim\langle W\rangle_{\text{geo}} - \dim W.$$

A subvariety $W \subset V$ is called *geodesic-optimal* if there is no strictly larger subvariety $W' \subset V$ with $\delta_{\text{geo}}(W') \leq \delta_{\text{geo}}(W)$. (This property is termed "cd-maximal" in the multiplicative setting in [27]). The following fact was established for modular, multiplicative and Abelian varieties separately in [14].

Proposition 2.7. *Let $V \subset X_{m,n}$. An optimal subvariety of V is geodesic-optimal.*

Proof. It is easy to adapt the proof of [14, Proposition 4.3] to show that $X_{m,n}$ has the "defect condition", and then the above follows from the formal properties of weakly special and special subvarieties, as in [14, Proposition 4.5]. □

Now we consider

$$V = V_n = \{(x_1, \dots, x_n, t_1, \dots, t_n) : x_i = t_i, i = 1, \dots, n\} \subset X_n = X_{n,n}.$$

We see that if a tuple $(\sigma_1, \dots, \sigma_n)$ of singular moduli satisfies a non-trivial multiplicative relation then the point

$$\Sigma = (\sigma_1, \dots, \sigma_n, \sigma_1, \dots, \sigma_n) \in V$$

lies in the intersection of V with a special subvariety of X of codimension $n + 1$. So such a point is an atypical component of V_n.

3. Mixed Ax-Schanuel

We now take again

$$X = X_{m,n} = Y(1)^m \times \mathbb{G}_{\text{m}}^n, \quad U = U_{m,n} = \mathbb{H}^m \times \mathbb{C}^n, \quad \text{and} \quad \pi : U \to X$$

given by

$$\pi(z_1, \dots, z_m, u_1, \dots, u_n) = \big(j(z_1), \dots, j(z_m), \exp(u_1), \dots, \exp(u_n)\big).$$

Definition 3.1.

1. An *algebraic subvariety* of U will mean a complex-analytically irreducible component of $Y \cap U$ where $Y \subset \mathbb{C}^m \times \mathbb{C}^n$ is an algebraic subvariety;
2. A *weakly special subvariety* of U is an irreducible component of $\pi^{-1}(W)$ where W is a weakly special subvariety of X. Likewise for *special subvariety* of U.

The following result leads to the analogue of the "Weak Complex Ax" (WCA; [14, Conjecture 5.10]) in this mixed modular-multiplicative setting. It is deduced from the same statement in the two extreme special cases: WCA for $Y(1)^n$, which is a consequence of the full modular Ax-Schanuel result established in [24], and WCA for $\mathbb{G}_{\mathrm{m}}^n$, which is a consequence of Ax-Schanuel [1].

Note that we could avoid talking about "algebraic subvarieties of U" by taking Y to be an algebraic subvariety of $\mathbb{C}^m \times \mathbb{C}^n$ and A to be a complex-analytically irreducible component of $Y \cap \pi^{-1}(V)$.

Theorem 3.2. *Let $V \subset X$ and $W \subset U$ be algebraic subvarieties and $A \subset W \cap \pi^{-1}(V)$ a complex-analytically irreducible component. Then*

$$\dim A = \dim V + \dim W - \dim X,$$

unless A is contained in a proper weakly special subvariety of U.

Proof. We suppose that A is not contained in a proper weakly special subvariety of U, and prove the dimension statement. We may suppose that A is Zariski-dense in W and that $\pi(A)$ is Zariski-dense in V.

Let V_0 be the image of V under the projection $X \to \mathbb{G}_{\mathrm{m}}^n$, and W_0 the image of W under the projection $p_0 : U \to \mathbb{C}^n$. Then the image A^* of A under p_0, being connected, is contained in some complex-analytically irreducible component $A_0 \subset W_0 \cap \exp^{-1}(V_0)$. Then A_0 is not contained in a proper weakly special subvariety of $\mathbb{C}^n$, otherwise A would be contained in a proper weakly special subvariety of U. So by Ax-Schanuel ([1]; see also [32]) we have

$$\dim A_0 \leq \dim W_0 + \dim V_0 - \dim \mathbb{C}^n.$$

Now we look at fibres in $\mathbb{H}^m$ and $\mathbb{C}^m$. We let $A_u, W_u \subset \mathbb{H}^m$, $V_t \subset \mathbb{C}^m$ be the fibres (of A, W, V respectively) over $u = (u_1, \dots, u_n) \in A_0$, $u \in W_0$, $t = (t_1, \dots, t_n) \in V_0$, respectively. Now A_0 must be Zariski-dense in W_0, else A could not be Zariski-dense in W, and similarly $\exp(A_0)$ must be Zariski-dense in V_0.

Since A is irreducible, the image A^* has constant dimension (see [16, V. 3.2, Corollary 2]) equal to the rank $\mathrm{rk}(p_0)$ of $p_0 : A \to A_0$, and $\dim A^* \leq \dim A_0$. Further we have [16, V. 3.3] that $\dim A = \mathrm{rk}(p_0) + \lambda(p_0) = \dim A^* + \lambda(p_0)$ where $\lambda(p_0)$ is the generic (*i.e.* minimal) fibre dimension of p_0.

The projection $W \to W_0$ has a generic fibre dimension away from a locus $W' \subset W$ of lower dimension, which does not contain A. So a generic fibre over A^* outside the image of W' is generic for A^* as well as W_0, and likewise for the corresponding fibre over V_0.

For $u \in A^*$, if A_u is not contained in a proper weakly special subvariety of $\mathbb{H}^m$, then by Ax-Schanuel for the j-function [24] we have,

$$\dim A_u \leq \dim W_u + \dim V_u - \dim \mathbb{H}^m.$$

If this holds generically, adding up the two last displays gives us the statement we want.

So we consider what happens when this fails generically. If the A_u were contained in a fixed proper weakly special, than A would be, which we have precluded. So the fibres must belong to a "moving family" of proper weakly specials. As elements of $\mathrm{GL}_2^+(\mathbb{Q})$ can't vary analytically, the only possibility is that some coordinates are constant on the fibres (though not constant on A).

Without loss of generality, we can suppose that these coordinates are $z_1, \ldots, z_k$. For $1 \leq \ell \leq k$, let V_ℓ be the image of V under the projection $X \to \mathbb{C}^\ell \times \mathbb{G}_{\mathrm{m}}^n$, and W_ℓ the image of W under the projection $p_\ell : U \to \mathbb{H}^\ell \times \mathbb{C}^n$. Then the image of A under p_ℓ, being connected, is contained in some complex-analytically irreducible component $A_\ell \subset W_\ell \cap \pi_{\ell,n}^{-1}(V_\ell)$. Note that this is consistent with the earlier definition of A_0, W_0, V_0.

Now we prove inductively that the dimension inequality holds at "level" ℓ, and once it holds at level k we are done. We assume that, for some $0 \leq h < k$:

(A) A_h is Zariski-dense in W_h and $\pi_h(A_h)$ is Zariski-dense in V_h;
(B) $\dim A_h \leq \dim W_h + \dim V_h - (n + h)$.

We know that these both hold for $h = 0$, and that (A) holds for all h.

Now z_{h+1} is constant on the fibres, so $\dim A_{h+1} = \dim A_h$. To show (B) we need only show that either $\dim W_{h+1} > \dim W_h$ or $\dim V_{h+1} > \dim V_h$.

Suppose that $\dim W_{h+1} = \dim W_h$. This means that, as functions on W, z_{h+1} is algebraic over $z_1, \ldots, z_h, u_1, \ldots, u_n$. But, as W is not contained in a proper weakly special subvariety, z_{h+1} is not constant on W nor does it satisfy any relation $z_{h+1} = g z_i$ where $1 \leq i \leq h$ and $g \in \mathrm{GL}_2^+(\mathbb{Q})$. But then, by the "Ax-Lindemann" result of [22] for the j-function, $j(z_{h+1})$ is algebraically independent of $j(z_1), \ldots, j(z_h), \exp(u_1), \ldots, \exp(u_n)$ as functions on W. Hence by the Zariski density these functions are independent as functions on A_{h+1}, and hence, by the Zariski-density of $\pi_{h+1}(A_{h+1})$ in V_{h+1}, we must have that $\dim V_{h+1} = \dim V_h + 1$. □

From this statement one may deduce, as explained in [23, above 5.7], the analogue of [14, Conjecture 5.10] (for j itself this follows from [24]).

Theorem 3.3. *Let $U' \subset U$ be a weakly special subvariety, and put $X' = \pi(U')$. Let $V \subset X'$ and $W \subset U'$ be subvarieties, and A a complex-analytically irreducible component of $W \cap \pi^{-1}(V)$. Then*

$$\dim A = \dim V + \dim W - \dim X'$$

unless A is contained in a proper weakly special subvariety of U'.

It is shown in [14] that Theorem 3.2 is equivalent by arguments using only the formal properties of the collection of weakly special subvarieties to the following version. We need the following definition from [14].

Definition 3.4. Fix a subvariety $V \subset X$.

1. A *component* with respect to V is a complex analytically irreducible component of $W \cap \pi^{-1}(V)$ for some algebraic subvariety $W \subset U$;
2. If A is a component with respect to V we define its *defect* to be $\partial(A) = \dim \mathrm{Zcl}(A) - \dim A$, where $\mathrm{Zcl}(A)$ denotes the Zariski closure of A;
3. A component A with respect to V is called *optimal* for V if there is no structly larger component B with respect to V with $\partial(B) \le \partial(A)$;
4. A component A with respect to V is called *geodesic* if it is a component of $W \cap \pi^{-1}(V)$ for some weakly special subvariety W.

Proposition 3.5. *Let $V \subset X$. An optimal component with respect to V is geodesic.*

Proof. The same as the proof that "Formulation A" implies "Formulation B" in [14]. (The proof of the reverse implication is also the same as given there.) □

4. Proof of Theorem 1.3

We start by recalling some background on trees and lattices associated to $\mathrm{GL}_2^+(\mathbb{Q})$. Let $T_{\mathbb{Q}} = \mathrm{PSL}_2(\mathbb{Z})\backslash \mathrm{PGL}_2^+(\mathbb{Q})$, where we assume their images are distinct. For a prime number p, $T_{\mathbb{Q}}$ maps into $T_p = \mathrm{PSL}_2(\mathbb{Z}_p)\backslash \mathrm{PGL}_2(\mathbb{Q}_p)$, and embeds into the product of the T_p over all p.

Now $T_{\mathbb{Q}}$ may be identified with the space of $\mathbb{Z}$-lattices in $\mathbb{Q}^2$ up to scaling, by sending g to the lattice spanned by $e_1 g, e_2 g$, where $e_1 = (1,0), e_2 = (0,1)$. Likewise, T_p may be identified with the space of $\mathbb{Z}_p$-lattices in $\mathbb{Q}_p^2$ up to scale. Moreover, T_p may be given the structure of a connected $(p+1)$-regular tree by saying that two lattices L, L' are adjacent if one can scale L' to be inside L with index p. There is a natural right action of $\mathrm{PGL}_2(\mathbb{Q}_p)$ on T_p: it acts on $\mathbb{Q}_p^2$ (up to scaling) in the natural way and thus on the lattices in it.

Since T_p is a tree there is a unique shortest path between any two nodes, and any path between those nodes traverses that path.

Our proof will study curves isogenous to the curve E_0 whose j-invariant is 0. These curves have CM by $\mathbb{Z}[\zeta]$, where $\zeta = \exp(2\pi i/3)$. A point $z \in \mathbb{H}$ with $j(z) = 0$ corresponds to the elliptic curve E_0 together with a basis v_1, v_2 for its integral homology $H_1(E_0, \mathbb{Z})$. For any sub-lattice $L \subset H_1(E_0, \mathbb{Q})$ we can define an elliptic curve E_L isogenous to E_0 which only depends on L up to scale. To do this, scale L until it contains $H_1(E_0, \mathbb{Z})$ and the quotient is cyclic. We can identify $Q_L = L/H_1(E_0, \mathbb{Z})$ with a subgroup of the torsion group of E_0 and take the quotient. Define $T'_{\mathbb{Q}}$ to be the space of lattices in $H_1(E_0, \mathbb{Q})$, up to scaling, and correspondingly T'_p the space of $\mathbb{Z}_p$-lattices in $H_1(E_0, \mathbb{Q}_p)$, up to scaling.

Now suppose that E_L is isomorphic to E_0. This implies that the quotient Q_L is the same as that of the kernel of an endomorphism x of E_0. If we identify $H_1(E, \mathbb{Z})$ with $\mathbb{Z}[\zeta]$, then the kernel of multiplication by x is $(x^{-1})/\mathbb{Z}[\zeta]$, where (m) denotes the fractional ideal generated by m. These correspond to elements of the fractional ideal group of $\mathbb{Z}_p[\zeta]$ (providing the endomorphisms giving the kernels) quotiented out by $\mathbb{Q}_p^\times$ (scaling). Explicitly we find the following.

1. If $p \equiv 1 \bmod 3$ then (p) has two disctinct primes above it, whose product is (p). Then $\mathbb{Z}_p[\zeta] = \mathbb{Z}_p \oplus \mathbb{Z}_p$ with ideal group $\mathbb{Z}^2$, which we quotient by the diagonal $\mathbb{Z}$. These nodes give a line in the tree: each such node being adjacent to two other such nodes for which the edges correspond to the two primes over (p);
2. If $p \equiv 2 \bmod 3$ then $\mathbb{Z}_p[\zeta] = \mathbb{Z}_{p^2}$, with ideal group $\mathbb{Z}$ which we quotient by $\mathbb{Z}$. Thus in this case there is just one node coming from curves isomorphic to E_0;
3. If $p = 3$ we get a ramified extension of $\mathbb{Z}_3$, which still has ideal group $\mathbb{Z}$ (generated by powers of the uniformiser) but now we quotient by $2\mathbb{Z}$ since 3 has valuation 2. We thus have two nodes coming from curves isomorphic to E_0, which are adjacent in the tree.

Note that in every case there is at least one node N' of T'_p adjacent to $H_1(E_0, \mathbb{Z})$ such that any lattice L for which the shortest path from $H_1(E_0, \mathbb{Z})$ to L goes through N' is not isomorphic to E_0.

Proposition 4.1. *Let $g_1, \dots, g_k \in \mathrm{GL}_2^+(\mathbb{Q})$ and suppose that the functions $j(g_i z)$ are distinct. Then there exists $z \in \mathbb{H}$ such that $j(g_i z) = 0$ for exactly one i.*

Proof. Suppose first that there exists a prime number p such that the images u_i of the g_i in T'_p are distinct. Without loss of generality we may assume that g_1, g_2 have images u_1, u_2 in T'_p whose distance is at least as large as that between the images of any distinct g_i, g_k. This implies there is a unique node N adjacent to u_1 such that the shortest path from u_1 to any other u_i goes through N. We may further suppose without loss of generality that $g_1 = 1$.

Fixing a basis v_1, v_2 for $H_1(E_0, \mathbb{Z})$ gives a map from T_p to T'_p, sending $\mathbb{Z}^2$ to $H_1(E_0, \mathbb{Z})$. By choosing v_1, v_2 appropriately we can send N to N'. It follows that the z with $j(z) = 0$ corresponding to this choice has $j(g_i z) \neq 0$ for all $i > 1$.

Now we give the proof without the simplifying assumption. While no single p may separate all the g_i, finitely many p do. Let $S = \{g_1, \dots, g_k\}$. Consider the image of S in T_2 and pick two nodes with maximal distance among images of pairs from S. Let u_2 be one of these "extremal" nodes, and let S_2 be the subset of S whose image in T_2 is u_2.

Now consider the image of S_2 in T_3, choose an extremal node u_3 and let S_3 be the subset of S_2 whose image in T_3 is u_3. After finitely many steps we arrive at a set S_p with only a single element. We may assume this element is g_1 and that $g_1 = 1$.

For each prime $q \leq p$ we let N_q be the unique node adjacent to u_q through which all paths from u_q to other images S_r go, where r is the prime preceding q (or $r = 0$ for $p = 2$).

Choose a basis v_1, v_2 of $H_1(E_0, \mathbb{Z})$ such that the induced map from T_q to T'_q takes N_q to N'_q for all $q \leq p$. The fact that this is possible amounts to the fact that $\mathrm{SL}_2(\mathbb{Z})$ subjects onto $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ for every n.

The claim now is that, for each $i > 1$, $j(g_i z) \neq 0$. To see this, let $q < p$ be the largest prime such that $g_i \in S_q$, and $q' \leq p$ the next prime after q. The above argument in the tree $T'_{q'}$ shows that $g_i z$ does not represent E_0. This proves the claim and the proposition follows. □

Proof of Theorem 1.3. Theorem 1.3 follows directly from Proposition 4.1 □

5. Arithmetic estimates

The proof of Theorem 1.2, and further results considered in the sequel, use some basic arithmetic estimates which are gathered here. Several of them were used for similar purposes in [13]. The absolute logarithmic Weil height of a non-zero algebraic number α is denoted $h(\alpha)$; the absolute Weil height is $H(\alpha) = \exp h(\alpha)$.

Constants $c_0, c_1, c_2 \ldots$ here and in the sequel are positive and absolute (though not necessarily effective), and have only the indicated dependencies.

Weak Lehmer inequality

A lower bound for the height by any fixed negative power of the degree suffices for our purposes. Loher has proved (see [17]): if $[K : \mathbb{Q}] = d \geq 2$ and $0 \neq \alpha \in K$ is not a root of unity then

$$h(\alpha) \geq \frac{1}{37} d^{-2} (\log d)^{-1}. \tag{5.1}$$

Singular moduli

For a singular modulus σ, we denote by R_σ the associated quadratic order and $D_\sigma = D(R_\sigma)$ its discriminant. Habegger [12, Lemma 1] shows that

$$h(\sigma) \geq c_1 \log |D_\sigma| - c_0, \tag{5.2}$$

based on results of Colmez and Nakkajima-Taguchi.

No singular modulus is a root of unity (we thank Gareth Jones for pointing this out: a singular modulus has a Galois conjugate which is real, but ± 1 are not singular moduli by inspecting the list of rational singular moduli, *e.g.*, in [29, A. 4]). This together with Kronecker's theorem imply, for a non-zero singular modulus σ,

$$h(\sigma) > c_2. \tag{5.3}$$

In the other direction [13, Lemma 4.3], for all $\epsilon > 0$,

$$h(\sigma) \leq c_3(\epsilon) |D_\sigma|^\epsilon. \tag{5.4}$$

Finally, we note that if τ is a pre-image of a singular modulus σ in the classical fundamental domain for the $\mathrm{SL}_2(\mathbb{Z})$ action then (see [22, 5.7])

$$H(\tau) \le 2D_\sigma. \tag{5.5}$$

Class numbers of imaginary quadratic fields

The class number of an imaginary quadratic order R will be denoted $\mathrm{Cl}(R)$. Recall that, for a singular modulus σ, $[\mathbb{Q}(\sigma) : \mathbb{Q}] = \mathrm{Cl}(R_\sigma)$. By Landau-Siegel, for every $\epsilon > 0$,

$$\mathrm{Cl}(R) \ge c_4(\epsilon)|D(R)|^{\frac{1}{2}-\epsilon}. \tag{5.6}$$

In the other direction,

$$\mathrm{Cl}(R) \le c_5(\epsilon)|D(R)|^{\frac{1}{2}+\epsilon} \tag{5.7}$$

with $c_5(\epsilon)$ explicit (see, *e.g.*, Paulin [20, Proposition 2.2], for a precise statement).

Faltings height of an elliptic curve

Let E be an elliptic curve defined over a number field. Let $h_{\mathrm{F}}(E)$ denote the semi-stable Faltings height of E, and j_E its j-invariant. Then ([31, 2.1]; see also [10])

$$\left| h(j_E) - \frac{1}{12} h_{\mathrm{F}}(E) \right| \le c_6 \log \max\{2, h(j_E)\} \tag{5.8}$$

with an absolute constant c_6.

Further, if E_1, E_2 are elliptic curves defined over a number field with a cyclic isogney of order N between them (*i.e.* $\Phi_N(j_{E_1}, j_{E_2}) = 0$) then ([28, 2.1.4]; see also [13, proof of Lemma 4.2])

$$|h_{\mathrm{F}}(E_1) - h_{\mathrm{F}}(E_2)| \le \frac{1}{2} \log N. \tag{5.9}$$

Isogeny estimate

Let K be a number field with $d = \max\{2, [K : \mathbb{Q}]\}$. Let E, E' be elliptic curves defined over K, with $h_{\mathrm{F}}(E)$ and $h_{\mathrm{F}}(E')$ their semi-stable Faltings heights. When E and E' are isogenous, the fundamental results of Masser and Wüstholz [19] give an estimate for the degree of as minimal isogeny between E, E' in terms of $[K : \mathbb{Q}]$ and the height of one of them. Gaudron and Rémond [8] prove the following explicit result improving that of Pellarin [21].

If E, E' are isogenous then there exists an isogeny $E \to E'$ of degree N satisfying

$$N \le 10^{13} d^2 \max\{h_{\mathrm{F}}(E), \log d, 1\}^2. \tag{5.10}$$

In particular there exists a cyclic isogeny with the same degree bound.

Estimate for the height of a multiplicative dependence

The following result, due to Yu (see [17]), allows us to get control of the height of a multiplicative relation on our singular moduli in terms of their height. It is thus a kind of "multiplicative isogeny estimate".

Let $\alpha_1, \dots, \alpha_n$ be multiplicatively dependent non-zero elements of a number field K of degree $d \geq 2$. Suppose that any proper subset of the α_i is multiplicatively independent. Then there exist rational integers $b_1, \dots, b_n$ with $\alpha_1^{b_1} \dots \alpha_n^{b_n} = 1$ and

$$|b_i| \leq c_7(n) d^n \log d\, h(\alpha_1) \dots h(\alpha_n)/h(\alpha_i), \quad i = 1, \dots, n. \tag{5.11}$$

6. Proof of Theorem 1.2

Fix n. Let $X = X_n = X_{n,n} = Y(1)^n \times \mathbb{G}_{\mathrm{m}}^n$, and let

$$V = V_n = \{(x_1, \dots, x_n, t_1, \dots, t_n) \in X : t_i = x_i, i = 1, \dots, n\}.$$

So $\dim V = \operatorname{codim} V = n$ and a singular-dependent n tuple $(x_1, \dots, x_n)$ gives rise to an atypical point $(x_1, \dots, x_n, x_1, \dots, x_n) \in V$.

Lemma 6.1. *A singular-dependent n-tuple may not be contained in an atypical component of V of positive dimension.*

Proof. A singular-dependent tuple can never be contained in a special subvariety of X defined by two (independent) multiplicative conditions, for between them we could eliminate one coordinate, contradicting the minimality.

Now a special subvariety of the form $M \times \mathbb{G}_{\mathrm{m}}^n$, where M is a special subvariety of $Y(1)^n$, can never intersect V atypically; neither can one of the form $Y(1)^n \times T$ where T is a special subvariety of $\mathbb{G}_{\mathrm{m}}^n$.

Let us consider a special subvariety of the form $M \times T$ where T is defined by one multiplicative condition. The intersection of $M \times T$ with V consists of those n-tuples of M which belong to T. This would typically have dimension $\dim M - 1$, and so to be atypical we must have $M \cap \mathbb{G}_{\mathrm{m}}^n \subset T$. Now Theorem 1.3 implies that M has two identically equal coordinates, but then cannot contain a singular-dependent tuple. □

Proof of Theorem 1.2. If $\sigma = j(\tau)$ is a singular modulus, so that $\tau \in \mathbb{H}$ is quadratic over $\mathbb{Q}$, we define its *complexity* $\Delta(\sigma)$ to be the absolute value of the discriminant of τ *i.e.* $\Delta(\sigma) = |D_\sigma| = |b^2 - 4ac|$ where $ax^2 + bx + c \in \mathbb{Z}[x]$ with $(a, b, c) = 1$ has τ as a root. For a tuple $(\sigma_1, \dots, \sigma_n)$ of singular moduli we define the complexity of σ to be $\Delta(\sigma) = \max(\Delta(\sigma_1), \dots, \Delta(\sigma_n))$.

Now suppose that V contains a point corresponding to a singular-dependent n-tuple of sufficiently large complexity, Δ. By Landau-Siegel (5.6) with $\epsilon = 1/4$, such a tuple has, for sufficiently large (though ineffective) Δ, at least $c_5 \Delta^{1/4}$ conjugates over $\mathbb{Q}$. Each is a singular-dependent n-tuple, and they give rise to distinct points in V.

Let F_j be the standard fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$, and $F_{\exp}$ the standard fundamental domain for the action of $2\pi i\mathbb{Z}$ (by translation) on $\mathbb{C}$.

We now consider the sets

$$Y = \left\{(z, u, r, s) \in F_j^n \times F_{\exp}^n \times \mathbb{R}^n \times \mathbb{R} : j(z) = \exp(u), r \cdot u = 2\pi i s\right\},$$

so that $(j(z), \exp(u)) \in V$ for $(z, u, r, s) \in Y$ and

$$Z = \left\{(z, r, s) \in F_j^n \times \mathbb{R}^n \times \mathbb{R} : \exists u, (z, u, r, s) \in Y\right\}.$$

Then Z is a definable set in the o-minimal structure $\mathbb{R}_{\mathrm{an}\,\exp}$.

A singular-dependent n-tuple $\sigma \in V$ has a pre-image

$$\tau = (z_1, \dots, z_n, u_1, \dots, u_n) \in F_j^n \times F_{\exp}^n,$$

and this gives rise to a point in Z, where the coordinates in $\mathbb{R}^{n+1}$ register the multiplicative dependence of the tuple, as follows. The F_j coordinates are the z_i, so they are quadratic points, and as recalled in (5.5) their absolute height is bounded by $2\Delta(\sigma_i)$. The point in $\mathbb{R}^{n+1}$ has integer coordinates $(b_1, \dots, b_n, b)$, not all zero, such that

$$\sum_{i=1}^{n} b_i u_i = 2\pi i b.$$

Now in view of the height estimate (5.4), and degree estimate (5.7) on the $j(z_i)$, (5.11) gives that the b_i in a multiplicative relation among the σ_i may be taken to be bounded in size by $c_8(n)\Delta^n$. With this bound on the $|b_i|$, and since the imaginary parts of the u_i are bounded by $2\pi i$, we get an upper bound on $|b|$. We find that the height of $(z_1, \dots, z_n, b_1, \dots, b_n, b)$ is bounded by $c_9(n)\Delta^n$.

In view of the Galois lower bound, a singular-dependent n-tuple of complexity Δ gives rise to at least

$$T^{\frac{1}{4n}} \text{ quadratic points on } Z \text{ with absolute height at most } T = c_{10}(n)\Delta^n.$$

For sufficiently large Δ, the Counting Theorem [25] applied to quadratic points on Z (considered in real coordinates) implies that it contains a semi-algebraic set of positive dimension. This implies (by the arguments used in [13,14]: the corresponding points (z, u) in $\mathbb{H}^n \times \mathbb{C}^n$ cannot be constant on all such semi-algebraic sets) that there is a complex algebraic $Y \subset U$ which intersects Z in a positive-dimensional component A which is atypical in dimension and contains singular-dependent n-tuples.

By the mixed Ax-Schanuel of Section 3 this implies that there is a positive-dimensional weakly special subvariety W containing Y containing a component B with $A \subset B$ and $\partial(B) \leq \partial(A)$. Moreover, it contains the special subvarieties that contain (some of) the singular-dependent points, so W is a special subvariety of positive dimension containing singular-dependent points of V, which we have seen is impossible.

So Δ is bounded, giving the finiteness. $\square$

Example 6.2. An example of a singular-dependent 5-tuple is (see [29, A. 4]):

$$\big(-2^{15}3^{3}5^{3}11^{3},\quad -2^{15},\quad 2^{3}3^{3}11^{3},\quad 2^{6}3^{3},\quad 2^{15}3^{1}5^{3}\big).$$

One also has a 3-tuple $(-2^{15}, -2^{15}3^{3}, 2^{6}3^{3})$ and 4-tuple $(2^{4}3^{3}5^{3}, -2^{15}3^{1}5^{3}, -3^{3}5^{3}, 2^{6}5^{3})$.

7. On the atypical set of V_n

The atypical set of V_n is the union of its proper optimal components (V_n itself is always optimal but never atypical). Since optimal components are geodesic-optimal Proposition 2.7, we will investigate the possibilities for these.

We observe that any geodesic-optimal components which dominate every coordinate can only come from an optimal strongly special subvariety. The finiteness of these, even if we cannot identify them, is guaranteed by o-minimality.

Definition 7.1. Complex numbers x, y will be called *Hecke equivalent* if $\Phi_N(x, y) = 0$ for some $N \geq 1$. *I.e.*, if the elliptic curves with j-invariants x and y are isogenous.

7.1. Geodesic-optimal components of dimension n

As already observed, V_n is not atypical since it dominates both $Y(1)^n$ and $\mathbb{G}_{\mathrm{m}}^n$. In other words, the defect of V_n is equal to its codimension.

7.2. Geodesic-optimal components of dimension $n-1$

Let $T \subset X$ be a geodesic subvariety of co-dimension 2. Can $T \cap V$ have dimension $n-1$? There are two equations defining T, each being one of four possible types: a single modular relation, a constant modular coordinate, a single multiplicative relation, a constant multiplicative coordinate.

Now if both equations are of modular (respectively multiplicative) type we never get an atypical component, because V dominates $Y(1)^n$ (respectively $\mathbb{G}_{\mathrm{m}}^n$). The same is true for any T which is defined purely by modular (respectively multiplicative) relations.

So we consider T defined by one condition of each type. Let us call T_1 the projection of T to the $Y(1)^n$ factor, which is a geodesic subvariety of codimension 1, and T_2 its projection to $\mathbb{G}_{\mathrm{m}}^n$. We get an atypical component of dimension $n-1$ if either $T_1 \cap \mathbb{G}_{\mathrm{m}}^n$ is contained in T_2, or if T_2 is contained in T_1 (*i.e.* when both are considered in the same copy of $(\mathbb{C}^\times)^n$).

If the modular condition is a modular relation (rather than a constant coordinate) then the first is excluded by Theorem 1.3, unless it is of the form $x_i = x_j$. If the multiplicative relation is not a fixed coordinate, the other inclusion is also impossible unless it is of the form $t_i = t_j$.

So we are reduced to considering constant coordinate conditions on both sides. This obviously leads to a component of dimension $n-1$ if the conditions coincide:

$x_i = \xi = t_i$. However such a component can only be atypical (*i.e.* arise from the intersection of V_n with a special subvariety of codimension (at most) 2 if ξ is both a singular modulus and a root of unity. But this never occurs, as remarked in Section 5.

This establishes ZP for V_1, which is the curve defined by $x_1 = t_1$ in $\mathbb{C} \times \mathbb{C}^\times$. And it shows that V_2 has no atypical subvarieties of positive dimension apart from the "diagonal" $x_1 = x_2$.

Proposition 7.2. *ZP holds for* V_2.

Proof. In view of the fact that the only atypical component of positive dimension is the "diagonal", which has defect zero, we are reduced to showing that V_2 has only finitely many optimal points, *i.e.* points which are atypical but not contained in the "diagonal". A point $(x_1, x_2, x_1, x_2) \in V_2$ is atypical if it lies on a special subvariety of codimension 3. There are then two cases: we have two independent modular conditions and one multiplicative, or two multiplicative and one modular relation.

The former case is exactly the question of singular-dependent 2 tuples, whose finiteness we have already established. The latter leads to the question of two (unequal) roots of unity which satisfy a modular relation. This is established in the following proposition, by a similar argument to that used in (5.2); and with this the proof is complete. □

We may observe that the optimal points of V_2 satisfy 3 special relations (never 4), so have "defect" 1.

Definition 7.3. A pair of distinct roots of unity is called a *modular pair* if they satisfy a modular relation.

Proposition 7.4. *There exist only finitely many modular pairs.*

Proof. Let (ζ_1, ζ_2) be such a point, where the order of ζ_i is M_i and $\Phi_L(\zeta_1, \zeta_2) = 0$. The point is that the order of the root of unity, and their bounded height, leads to a bound on the degree of the modular relation. Specifically, by (5.8), the semi-stable Faltings height of the corresponding elliptic curves E_1, E_2 with j-invariants ζ_1, ζ_2 are bounded, and so by the isogeny estimate (5.10) there is a modular relation $\Phi_N(\zeta_1, \zeta_2) = 0$ with $N \leq c_{11} \max\{M_1, M_2\}^5$. Thus such a point leads to a rational point on a suitable definable set whose height is bounded by a polynomial in the orders of the two roots, and if it is of sufficiently large complexity it forces the existence of a higher dimensional atypical intersection containing such points. But the only atypical set of dimension 1 is given by $x_1 = x_2, t_1 = t_2$. □

As modular relations always subsist between two numbers, there is no notion of "modular-multiplicative n-tuples" analogous to singular-dependent tuples. However, an immediate consequence of the above is that, for any n, there exists only finitely many n-tuples of distinct roots of unity which are pairwise Hecke equivalent (and none for sufficiently large n).

7.3. Geodesic-optimal components of dimension $n-2$

These arise from intersecting V_n with a geodesic subvariety T of codimension (at least) 3. We must have at least 1 relation of each type, and if they are all of "non-constant" type (no fixed coordinates) then we get finiteness by o-minimality.

If there is one constant condition, this immediately gives a second such condition of the other type, and then any additional non-constant condition (*i.e.* not forcing any further constant coordinates) will give a component of dimension $n-2$. However, no such component can be atypical.

Consider the case of 3 constant conditions. First the case of two fixed modular coordinates. This will give rise to an atypical intersection if the two fixed values are multiplicatively related. Next the case of two fixed multiplicative coordinates. This will give rise to an atypical component if the two fixed values are Hecke equivalent. The finiteness of such components follows from ZP for V_2, and they all have defect 2. Thus:

Proposition 7.5. *For $n \geq 1$, V_n has only finitely many maximal atypical components of dimension $n-2$.*

But for $n = 3$ we can in fact exclude "strongly atypical" altogether. Such a component has one of two shapes.

1. Two modular relations and one multiplicative relation. This would be atypical if the resulting modular curve satisfied the multiplicative relation, but this is impossible by Theorem 1.3;
2. Two multiplicative relations and one modular relation. This gives a "multiplicative curve", which can be parameterised as $(\zeta_1 t^{a_1}, \zeta_2 t^{a_2}, \zeta_3 t^{a_3})$, where ζ_i are roots of unity and a_i integers. As the $\Phi_N, N \geq 2$ are symmetric, two coordinates cannot satisfy a modular equation unless $a_i = a_j$ (so that $N = 1$ and $\Phi_1 = X - Y$) and $\zeta_i = \zeta_j$.

Proposition 7.6. *The positive dimensional atypical components of V_3 and their defects may be described as follows:*

1. *The intersection of V_3 with $x_i = x_j, i \neq j$ is a copy of V_2 contained in X_2 (hence of defect 2) and has some atypical points in it, which have defect 1. It contains also the subvariety with $x_1 = x_2 = x_3$, which has defect 0;*
2. *A singular-dependent 2-tuple $\sigma = (\sigma_1, \sigma_2)$ gives rise to an atypical component A_σ of dimension 1 and defect 2. (There may exist singular moduli which belong to two distinct such pairs σ, σ'. Then we get a point $(A_\sigma \cap A_{\sigma'})$ of defect 1);*
3. *A modular pair $\zeta = (\zeta_1, \zeta_2)$ gives rise to an atypical component B_ζ of dimension 1 and defect 2. (There may exist roots of unity belonging to two distinct modular pairs ζ, ζ'. Then we get a point $(B_\zeta \cap B_{\zeta'})$ of defect 1.)*

In particular, there are no positive dimensional "strongly atypical" components (i.e., *with no constant coordinates*).

Thus ZP for V_3 depends on the finiteness of its atypical points off all the above positive dimensional atypical components. This leads to some Diophantine questions which would establish ZP for V_3, which we study in the next section.

Remark 7.7. Note that X contains families of weakly special subvarieties which intersect V_n atypically, namely those defined be relations the form $x_i = x_j$ (and $t_i = t_j$) or $x_k = t_k = c_k \in \mathbb{C}^\times$ for various choices of $(i, j), i \neq j, k$. If m such conditions are imposed, the resulting weakly special subvariety has dimension $2n - 2m$ and intersects V_n in a component of dimension $n - m$, so has geodesic defect $n - m$.

Conjecture 7.8. The atypical geodesic components described in Remark 7.7 give all geodesic optimal subvarieties of V_n for any n; in particular, apart from components defined by "diagonal" equations $x_i = x_j$ there are no "strongly optimal" geodesic optimal components (*i.e.* with no constant coordinates).

8. Optimal points in V_3

The optimal points in $(x_1, x_2, x_3, x_1, x_2, x_3) \in V_3$ fall into two classes. Those which are atypical in satisfying at least 4 special conditions, but are not contained in atypical component of higher dimension. And those which are "more atypical", satisfying 5 special conditions (it is not possible to have 6: only a triple of singular moduli which were also roots of unity could achieve this), though lying in an atypical set of larger dimension but larger defect. Those lying on diagonals $x_i = x_j, i \neq j$ are easy to describe, we consider here those that do not.

Let us first consider points satisfying 5 special conditions. These also fall into two types: 3 modular, 2 multiplicative, or the other way around. If there are 3 modular conditions then each x_i is a singular modulus. The two multiplicative conditions mean either than one x_j is torsion, and the other two multiplicatively related, or the three are pairwise multiplicatively related. The former is impossible. Now only finitely many pairs of singular moduli have a multiplicative relation, so x_1, x_2 comes from a finite set, and x_3 comes also from a finite set. If there are three multiplicative relations then each x_i is torsion. Only finitely many pairs of (distinct) roots of unity satisfy isogenies, and we get finiteness (there are no "Hecke equivalences" involving three points!). All these points have defect 1.

Now we consider points $(x_1, x_2, x_3, x_1, x_2, x_3) \in V_3$, away from positive dimensional atypical subvarieties, satisfying 4 special conditions. The "generic" situation involves no singular moduli or roots of unity.

Problem 8.1. Prove that there exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers, which are not roots of unity and not singular moduli, such that they are pairwise Hecke equivalent, and also pairwise multiplicatively dependent.

The various arithmetic estimates seem insufficient to get a lower degree bound in terms of the "complexity": the degrees of the two isogenies and the heights of the two multiplicative relations. This seems to be a problem of a similar nature to that encountered in [13] dealing with curves which are not "asymmetric" (see [13, Section 1]).

There are three "non-generic" variations of which we can resolve two. The multiplicative relations may take the form that one coordinate is a root of unity, the other two coordinates being multiplicatively dependent. Similarly, the modular relations may take the form that one coordinate is singular, the other two Hecke equivalent. Or both. Note that if two coordinates are singular the point is not optimal, but lies on one of the atypical components in Propositions 7.6(3); if two coordinates are roots of unity, the point is on a component as in Proposition 7.6(2).

We consider the non-generic multiplicative condition first. Up to permutations we may assume the singular coordinate is x_1

Proposition 8.2. *There exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers such that:*

1. *x_3 is a root of unity, x_1, x_2 are multiplicatively dependent but not roots of unity;*
2. *The three points are pairwise Hecke equivalent, but are not singular moduli.*

Proof. Define the complexity Δ of such a triple to be the maximum of: the order M of the root of unity x_3 and the minimum degrees of isogenies N_1, N_2 between x_3 and x_1, x_2, respectively. By (5.8), the stable Faltings height of an elliptic curves whose j-invariant is a root of unity is absolutely bounded. Now by (5.9), $h(x_j) \ll (1 + \log \max\{N_j\})$, $j = 1, 2$, so by (5.10) the degrees $d_j = [\mathbb{Q}(x_3, x_j) : \mathbb{Q}] \gg N_j^{1/5}$. By (5.11) and (5.1) (to get a lower bound for $h(x_i)$) the height of a multiplicative relation between x_1, x_2 is bounded by some $c_{12}\Delta^{c_{13}}$. And $[\mathbb{Q}(x_3) : \mathbb{Q}] = \phi(M) \gg_\epsilon M^{1-\epsilon}$, where ϕ is the Euler ϕ-function. We may take $\epsilon = 1/2$ say.

Thus, a triple of complexity Δ gives rise to "many" (*i.e.* at least $c_{14}\Delta^{c_{15}}$) quadratic points on a certain definable set, and so all but finitely many such points lie on atypical components of positive dimension.

But no such triples lie on positive dimensional atypical components: by Proposition 7.6, such components have either two singular coordinates or two modular coordinates, so the conditions on our triples would then force all x_i to be singular, which is impossible (as then x_3 cannot be torsion) or all torsion, which leads to the same impossible requirement for x_1. □

Symmetrically, we have the case where the modular relations are of the non-generic form. We seem unable to establish finiteness here, so we pose it as a problem.

Problem 8.3. Prove that there exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers such that x_1 is singular, x_2, x_3 are Hecke equivalent, and the three are pairwise multiplicatively dependent.

Finally, we have the following.

Proposition 8.4. *There exist only finitely many triples x_1, x_2, x_3 of distinct non-zero algebraic numbers such that*

1. x_1 *is a singular modulus,* x_2, x_3 *are Hecke equivalent but are not singular moduli;*
2. x_3 *is a root of unity,* x_1, x_2 *are multiplicatively dependent but not roots of unity.*

Proof. Let D be the discriminant of x_1 (see Section 5), and M the (minimal) order of x_3. Take N minimal with $\Phi_N(x_2, x_3) = 0$, and B minimal for a non-trivial multiplicative relation $x_1^{b_1} x_2^{b_2} = 1$ with $B = \max\{b_1, b_2\}$. Set $\Delta = \max\{|D|, M, N\}$ to be the complexity of the tuple (x_1, x_2, x_3). Set $d = [\mathbb{Q}(x_1, x_2, x_3) : \mathbb{Q}]$.

Let E_ξ be the elliptic curve with j-invariant ξ. As in the proof of Proposition 8.2, $h_{\mathrm{F}}(E_{x_3})$ is bounded by some absolute c_{16}. Then, by the isogeny estimates (5.10), we have $N \leq c_{17}([\mathbb{Q}(x_2, x_3) : \mathbb{Q}])^5$. Also $M^{1-\epsilon} \ll_\epsilon \phi(M) = [\mathbb{Q}(x_3) : \mathbb{Q}]$, and $|D| \ll [\mathbb{Q}(x_1) : \mathbb{Q}]^4$ by (5.6).

Arguing as in [13], the height inequalities (5.8), (5.9) imply that $h(x_2)$ is bounded above by $c_{18}(1 + \log N)$. By the Weak Lehmer estimate (5.1) it is bounded below by $c_{19} d^{-3}$. Corresponding estimates for $h(x_1)$ are provided by (5.4) and (5.3). Therefore (5.11) ensures that

$$B \leq c_{20} d^3 D.$$

The rest of the proof is the same as the proof of Proposition 8.2. □

Thus Problems 8.1 and 8.3 imply (and are implied by) ZP for V_3. If one takes two complex numbers and three conditions, then either two "modular" conditions or two "multiplicative" special conditions will force the points to be special, and one can prove finiteness. However one can consider two complex numbers satisfying a special condition of each of three (or more) different types.

Let S be a Shimura curve corresponding to a quaternion algebra over $\mathbb{Q}$ (see, *e.g.*, Elkies [6]). There is a notion of *Hecke orbit* of a point on S (see, *e.g.*, [5]), an equivalence class of points under a certain equivalence relation. This relation is the existence of a "cyclic N-isogeny" between the corresponding parameterised objects for some N; see [6, Section 2.3, page 12]. If S has genus zero, there is an analogue $j_S : \mathbb{H} \to \mathbb{P}^1$ of the j-function (see [7,30]) which generates the function field of S, and we may speak of points in $\mathbb{C}$ being "Hecke equivalent (for S)" if they are in the same Hecke orbit.

Problems 8.5. Prove that there are only finitely many pairs of distinct non-zero algebraic numbers x_1, x_2 in each situation.

1. x_1, x_2 are Hecke equivalent (in the sense of 7.1), and multiplicatively dependent, and are also Hecke equivalent for some other Shimura curve;
2. x_1, x_2 are Hecke equivalent, and multiplicatively dependent, and the points with these x-coordinates are dependent in some specific elliptic curve;
3. As in the previous problems, but with more or different conditions: say the points are Hecke equivalent/dependent for 10 pairwise incommensurable Shimura curves.

Finally we state a "weakly bounded height conjecture" on the height of "just likely" intersections of mixed multiplicative-modular type under which Problems 8.1 and 8.3 are affirmed.

Definition 8.6. A *modular-dependent pair* is a point $(x, y) \in (\mathbb{C}^\times)^2$ such that there exists integers N, a, b, ℓ with $N \geq 2$, $\ell \geq 1$ and $\gcd(a, b) = 1$ such that

$$\Phi_N(x, y) = 0, \quad (x^a y^b)^\ell = 1.$$

The *complexity* $\Delta(x, y)$ of such a pair is the minimum of $\max(N, |a|, |b|, \ell)$ over all N, a, b, ℓ for which the above equations hold for x, y.

Conjecture 8.7. For $\epsilon > 0$ we have $h(x), h(y) \leq c_\epsilon \Delta(x, y)^\epsilon$ for all modular-dependent pairs (x, y).

Proposition 8.8. *Assume Conjecture* 8.7. *Then finiteness holds in Problems* 8.1 *and* 8.3.

Proof. Let (x, y) be a modular-dependent pair with complexity $\Delta = \Delta(x, y) = \max(|a|, |b|, \ell)$ for suitable a, b, ℓ. Constants denoted c are absolute but may vary at each occurrence.

Let E_x, E_y be elliptic curves with j-invariants x, y and semistable Faltings heights $h_{\mathrm{F}}(x) = h_{\mathrm{F}}(E_x)$ and $h_{\mathrm{F}}(y) = h_{\mathrm{F}}(E_y)$ respectively. Then E_x, E_y may both be defined over $\mathbb{Q}(x, y)$, and we set $d = [\mathbb{Q}(x, y) : \mathbb{Q}]$.

By the isogeny estimate (5.10), $N \leq cd^2 \max\{h_{\mathrm{F}}(x), \log d, 1\}^2$. Now $h_{\mathrm{F}}(x)$ and $h(x)$ differ by at most $c \log \max(2, h(x))$. So

$$N \leq cd^2 \max(1, \log d)^2 \big(1 + h(x) + c \log \max(2, h(x))\big)^2.$$

We have $d^2 \max(1, \log d)^2 \leq d^4$, and under Conjecture 8.7 (with $\epsilon = 1/20$ say) we have

$$N \leq cd^4 \Delta^{1/10}.$$

For the purposes of Proposition 8.1 and 8.3 we may assume that neither x nor y is a root of unity. By a Weak Lehmer inequality (5.1) we have $h(x) \geq cd^{-3}, \quad h(y) \geq cd^{-3}$. Since neither x, y is a root of unity, we find (5.11) that there exists a nontrivial multiplicative relation $x^\alpha y^\beta = 1$ with

$$|\alpha| \leq cd^3 h(y) \leq cd^3 \Delta^{1/10}, \quad |\beta| \leq cd^3 h(x) \leq cd^3 \Delta^{1/10}.$$

Again since x, y are not roots of unity, we have that (α, β) is a multiple of $(\ell a, \ell b)$. So we find that $|a|, |b|, c\ell \leq cd^3 \Delta^{1/10}$. Now $\Delta = \max(N, |a|, |b|, |\ell|)$ and so combining the various inequalities we find

$$\Delta \leq cd^7.$$

Now points x_1, x_2, x_3 as in Problem 8.1 give rise to rational points on some suitable definable set of height at most $\max\big(\Delta(x_1, x_2), \Delta(x_2, x_3), \Delta(x_1, x_3)\big)$. This lower estimate for the degree is then suitable to complete a finiteness proof for isolated points of this form by point-counting and o-minimality as in the proofs of Theorem 1.2, Propositions 8.2, and 8.4. The argument for Problem 8.3 is similar. □

9. On ZP for V_n

The referee asked us whether there is a natural generalization of the height-theoretic Conjecture 8.7 which would imply ZP for $V_n, n \geq 3$. We thank the referee for raising this question, to which we offer an affirmative answer here. As this Conjecture 9.5 is rather more speculative than the very special case in Conjecture 8.7 we have preferred to keep this section separate.

We continue to let $X_n = Y(1)^n \times \mathbb{G}_{\mathrm{m}}^n$ and $U_n = \mathbb{H}^n \times \mathbb{C}^n$, and F a standard fundamental domain for the action on U_n by $\mathrm{SL}_2(\mathbb{Z})^n \times \mathbb{Z}^n$ where $m \in \mathbb{Z}$ acts on $\mathbb{C}$ be translation by $2\pi i m$. Constants $c, c(n), c(n, \epsilon), \ldots$ depend only on the indicated quantities, but may differ at each occurence.

First, we will assume Conjecture 7.8. This seems to be necessary, in view of the following. We have seen that an optimal component with respect to $V_n \subset X_n$ is geodesic optimal. Now weakly special subvarieties of U_n are contained in larger definable families of "Mobius varieties" which are defined by some finite number of relations of the form $z_i = g_{ij} z_j, g_{ij} \in \mathrm{SL}_2(\mathbb{R})$ or of the form $z_k = c_k \in \mathbb{C}$ on the $\mathbb{H}^n$ variables and of the form $\sigma_{j=1}^n r_{ij} u_j = 0$ with $(r_1, \ldots, r_n) \in \mathbb{R}^n$ on the $\mathbb{C}^n$ variables (see Mobius varieties in [14, Section 6.2] and "linear varieties" in [22, 10.1]). Then the set of Mobius varieties which intersect $Z = \pi^{-1}(V_n) \cap F$ optimally among Mobius varieties gives the full set of weakly special varieties intersecting Z optimally. By o-minimality, the set of relations among non-constant coordinates is then finite, since the corresponding coefficients r_{ij} and group elements g_{ij} must in fact be rational (see [14, Proposition 6.6; 22, 10.2]).

We are thus led to consider, for example, the intersections of a fixed strongly special modular special subvariety M with a family of translates $\{aT : a \in A\}$ of a subtorus T, *i.e.* a family of weakly special multiplicative weakly special subvarieties. Here A can be taken to be a copy of some $(\mathbb{C}^\times)^m$. Since optimal components are geodesic-optimal, these components will correspond to those $a \in A$ for which $M \cap aT$ has atypical dimension, which give some subvarieties $A_i \subset A$. However, if the component is optimal, the corresponding special subvariety will in general be larger, and we will be led to consider atypical points in $A_i \subset (\mathbb{C}^\times)^m$, *i.e.* to some cases of ZP for the multiplicative group, which we do not know how to handle at present.

We will say that a point $C = (x_1, \ldots, x_n) \in (\mathbb{C}^\times)^n$ satisfies h special relations if the smallest special subvariety of X_n containing $(x_1, \ldots, x_n, x_1, \ldots, x_n)$ has codimension h. On the modular side, the relation of being in the same Hecke orbit divides the non-special coordinates into k equivalence classes. Such equivalence classes of non-special points we call *cliques*. Then we see that if C satisfies h special relations we have

$$h = n + m - k,$$

where m is the number of independent multiplicative relations satisfied by C, and k is the number of cliques. We set

$$\partial(C) = n - h = k - m.$$

Definition 9.1. A tuple $C = (c_1, \dots, c_n) \in (\mathbb{C}^\times)^n$ with pairwise distinct coordinates is called *n-optimal* if no proper subtuple C' has $\partial(C') \le \partial(C)$. *I.e.* removing any ℓ points from C loses at least $\ell + 1$ special relations.

Proposition 9.2. *Assuming Conjecture* 7.8*, ZP for all V_n is equivalent to the statement that, for all n, there are only finitely many n-optimal points.*

Proof. Assuming Conjecture 7.8, all optimal components are, up to permutations of coordinates, of the form $\{(W, W) \in V_n\}$ where $W \subset \mathbb{C}^n$ is of the form

$$\{(c_1, c_2, \dots, c_\ell, x_{\ell+1}, \dots, x_n) : x_i \in \mathbb{C}^*, x_i = x_j : (i, j) \in I\}$$

for some set I of pairs (i, j) with $\max(i, \ell) < j$ from $\{1, \dots, n\}$, where $c_1, \dots, c_\ell$ are distinct non-zero complex numbers.

If the tuple $C = (c_1, \dots, c_\ell)$ satisfies h special relations, we have

$$\dim W = n - \ell - |I|, \quad \dim\langle W\rangle = 2n - 2|I| - h,$$

whence

$$\delta(W) = 2n - 2|I| - h - (n - \ell - |I|) = n + (\ell - h) - |I| = n + \partial(C) - |I|.$$

Therefore, the component W is optimal just if C is ℓ-optimal. □

Suppose C is an n-optimal tuple, with (C, C) contained in some smallest special subvariety $T \subset X_n$. Then the component of $T \cap V_n$ containing (C, C) must be just the point $\{(C, C)\}$. Otherwise, the component is clearly not optimal. Thus, an n-optimal tuple is a tuple of algebraic numbers, and the degree $d(C) = [\mathbb{Q}(c_1, \dots, c_n) : \mathbb{Q}]$ is bounded in terms of the degrees of the equations defining T.

We now frame a "weakly bounded height conjecture" for certain "just likely" intersections that seems plausible and is sufficient to establish this finiteness (assuming Conjecture 7.8). Of course one only needs the conjecture to hold for optimal points, which must in fact be "unlikely".

Consider a point $C = (c_1, \dots, c_n) \in (\mathbb{C}^\times)^n$ with c_i distinct, together with a $\mathbb{Z}$-module Γ of exponents of multiplicative relations on C. That is, Γ is a $\mathbb{Z}$-submodule of the *relation group*

$$\Gamma(C) = \{(a_1, \dots, a_n) \in \mathbb{Z}^n : c_1^{a_1} \cdots c_n^{a_n} = 1\}.$$

Suppose C has k cliques and $\operatorname{rank}(\Gamma) = m$. Removing some points from C yields a tuple C', and it inherits a submodule Γ' of relations from Γ which are trivial on the points removed (*i.e.* Γ' is the submodule of exponent vectors for which the coordinates corresponding to $C - C'$ are zero). We call Γ' the *induced relations*.

Definition 9.3. A pair (C, Γ) consisting of a tuple $C \in (\mathbb{C}^\times)^n$ with pairwise distinct coordinates having k cliques and a $\mathbb{Z}$-module of $\Gamma \subset \mathbb{Z}^n$ of exponent vectors of multiplicative relations on C is called *grounded* if, for any subtuple C' formed by removing any κ cliques, where $0 < \kappa < k$, together with any number of special points, the induced relation module Γ' satisfies $\mathrm{rank}(\Gamma') < \mathrm{rank}(\Gamma) - \kappa$. *I.e.* losing κ cliques loses at least $\kappa + 1$ multiplicative relations.

Note that a grounded tuple can never contain a singleton clique, for omitting such a clique will lead to the loss of at most one multiplicative relation.

Definition 9.4. We define the *height* of a tuple $C = (c_1, \dots, c_n)$ to be

$$h(C) = \max(h(c_1), \dots, h(c_n)).$$

The *modular complexity* of a tuple is

$$\Delta_{\mathrm{mod}}(C) = \max\{N\}$$

over N such that there exists c_i, c_j (allowing $i = j$) with $\Phi_N(c_i, c_j) = 0$, and $N \geq 2$ minimal for this pair i, j. We define the *complexity* of a $\mathbb{Z}$-submodule of $\mathbb{Z}^n$ to be

$$\Delta(\Gamma) = \min\{T\}$$

over T such that there is a basis of Γ consisting of vectors with all entries of absolute value bounded by T. The *complexity* of a pair (C, Γ) is

$$\Delta(C, \Gamma) = \max\{\Delta_{\mathrm{mod}}(C), \Delta(\Gamma)\}.$$

Finally, the *complexity* of C is

$$\Delta(C) = \Delta(C, \Gamma(C)).$$

Conjecture 9.5. Let (C, Γ) be grounded, where $C \in (\mathbb{C}^\times)^n$, and suppose that $\mathrm{rank}(\Gamma)$ equals the number k of cliques of C. Then

$$h(C) \leq c(n, \epsilon)\Delta(C, \Gamma)^\epsilon$$

for any $\epsilon > 0$.

Note that such C is in the intersection of V_n with a special subvariety of dimension $n + k - \mathrm{rank}(\Gamma) = n$, hence is a "just likely" intersection, though this may not be the smallest special subvariety containing C.

It seems that one cannot hope to have a suitable weakly bounded height conjecture for tuples which are not grounded. For example, if one has a clique C' satisfying some relations Γ', then imposing just one additional relation $\gamma \in \mathbb{Z}^n$ on $C = (C', C'')$ for an additional clique C'' would allow the height of C'' to be roughly $\Delta(\mathbb{Z}\gamma)h(C')$. An interesting question seems to be whether one should expect in fact an upper bound of the form $\ll_n (\log \Delta(C, \Gamma))^{c(n)}$ in Conjecture 9.5.

We now gather some further arithmetic estimates. Various forms of the following result, which we do not need in the sharpest forms, appear in the literature; see [17]. The following is extracted from [18].

Proposition 9.6. *For an n-tuple C of degree $d = d(C)$ and height $h = h(C)$, there is a basis of $\Gamma(C)$ consisting of vectors of integers of size at most*

$$c(n)d^n \log(d+2)^{3n} \max(h, 1)^n.$$

Proof. This is a weakened form of the bound in [18, page 253] together with the estimates for the quantities there established on page 254. ☐

Lemma 9.7. *Suppose $C \in (\mathbb{C}^\times)^n$ is n-optimal. Then $(C, \Gamma(C))$ is grounded. Moreover, if C has k cliques, then there is a submodule Γ of relations on C with* $\mathrm{rank}(\Gamma) = k$ *such that (C, Γ) is grounded and $\Delta(\Gamma) \le c(n)\Delta(\Gamma(C))$.*

Proof. For the first assertion, if we remove κ cliques (and some special points) from C to form C' and lose only κ multiplicative relations then $\partial(C') \le \partial(C)$, and C was not n-optimal. So n-optimal is stronger than grounded.

For the second assertion, we show how to find a suitable submodule Γ of $\Gamma(C)$ of rank equal to k, the number of cliques.

We first show that there is a vector $v_1 \in \Gamma(C)$ with a height bound as in the assertion of the lemma which "involves" all special points and cliques, *i.e.* where the exponent is non-zero on every coordinate i where c_i is special, and for some coordinate in every clique.

Let B be a basis of $\Gamma(C)$ consisting of vectors of integers of size at most $\Delta(\Gamma(C))$. Since $(C, \Gamma(C))$ is n-optimal, such a vector w_j exists for each individual special coordinate, and for each individual clique; say there are J such vectors. Moreover, we can assume that each $w_j \in B$. We consider vectors of the form

$$w = \sum_{j=1}^{J} a_j w_j, \quad a_j \in \mathbb{Q}.$$

For each special coordinate or clique, the condition that w vanishes on that coordinate or clique gives a proper subspace of $\mathbb{Q}^n$. It therefore contains at most $c(n)T^{n-1}$ integer points in the box $[-T, T]^n$. We must avoid $J \le n$ such subspaces, so $T = c(n)$ suffices.

We now construct $v_2 \in \Gamma(C)$ such that, for every clique, v_2 does not vanish modulo v_1. For each clique individually the existence of such a vector w_j is assured since $(C, \Gamma(C))$ is grounded, and so we can take $w_j \in B$. A similar box argument produces v_2 (the number of subspaces to avoid is now at most the number of pairs of cliques), and we continue to produce $v_3, \ldots, v_k$, where v_k does not vanish modulo $\mathbb{Z}[v_1, \ldots, v_{k-1}]$ on any choice of $k-1$ cliques. ☐

Proposition 9.8. *Let σ be a singular modulus of discriminant D_σ. Then there exists $N \ge 2$ with $\Phi_N(\sigma, \sigma) = 0$ satisfying $N \le |D_\sigma|$. Conversely, if $\Phi_N(\sigma, \sigma) = 0$ where $N \ge 2$ then $|D_\sigma| \le cN^{20}$, with an explicit c.*

Proof. Let τ be a preimage of σ in the standard fundamental domain. Then τ satisfies a minimal quadratic equation over $\mathbb{Z}$ of the form $A\tau^2 + B\tau + C = 0$ which is reduced, meaning $|B| \le A \le C$ and $B \ge 0$ if $A = |B|$ or $A = C$. Thus $4AC = B^2 - D_\sigma \le AC - D_\sigma$ whence $3AC \le |D_\sigma|$. Now $g\tau = \tau$ for $g = \left(\begin{smallmatrix} -B & -C \\ A & 0 \end{smallmatrix}\right)$, which is primitive of determinant $N = AC \le |D_\sigma|$.

In the other direction, suppose $\Phi_N(\sigma, \sigma) = 0$. This means that $g\tau = h\tau$ for a matrix g of the form $g = \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$ with $0 < a, 0 \le b < d, ad = N$ (see [15, 5.1, page 52]), and $h \in \mathrm{SL}_2(\mathbb{Z})$. Now $|\mathrm{Re}(g\tau)| \le 2N$ and $|\mathrm{Im}(\tau)^{-1}| \le 2N$ so the matrix h has entries at most $c(2N)^9$ by [13, Lemma 5.1]. Thus τ is fixed by $h^{-1}g$, an integer matrix with entries bounded by $c(2N)^{10}$. This gives an integral quadratic polynomial satisfied by τ whose coefficients have size at most $2c(2N)^{10}$. The minimal equation for τ must divide this one, and so (with a new constant) $|D_\sigma| \le cN^{20}$ □

Theorem 9.9. *Assuming Conjectures* 7.8 *and* 9.5, *ZP holds for* $V_n \subset X_n$ *for all* n.

Remark 9.10. One might hope to prove at this juncture that Conjecture 9.5 implies finiteness of n-optimal tuples for each n (without assuming Conjecture 7.8). However, our proof will require Conjecture 7.8.

Proof. Suppose $C \in (\mathbb{C}^\times)^n$ is n-optimal. Then the point $(C, C) \in V_n$ is an optimal component. Thus C is grounded, and the relation group $\Gamma(C)$ has rank exceeding k, the number of cliques in C (because (C, C) must be an unlikely intersection).

By Lemma 9.7 we find a submodule Γ of relations on C with rank k and with $\Delta(\Gamma) \le c(n)\Gamma(C)$. By Conjecture 9.8 with $\epsilon = (20n)^{-1}$ we have

$$h(C) \le c(n)\Delta(C, \Gamma)^{1/(20n)}.$$

We now obtain a lower bound for $d(C)$ in terms of $\Delta(C)$. We start considering $\Delta_{\mathrm{mod}}(C)$. If x, y are distinct and $\Phi_N(x, y) = 0$ with N minimal then, as in the proof of Proposition 8.8,

$$N \le c(n)[\mathbb{Q}(x, y) : \mathbb{Q}]^4 \Delta(C, \Gamma)^{1/(10n)}.$$

If x is special, then we have $\Phi_N(x, x) = 0$ for some $N \le |D_x|$, while $d(x) \gg_\epsilon |D_x|^{1/2-\epsilon}$ (ineffectively) by Landau-Siegel. Thus again $N \le c[\mathbb{Q}(x) : \mathbb{Q}]^4$, and we find

$$\Delta_{\mathrm{mod}}(C) \le c(n)d(C)^4 \Delta(C, \Gamma)^{1/(10n)}$$

for some (ineffective if any c_i are special) positive $c(n)$.

Now by Proposition 9.6, there is a basis of $\Gamma(C)$ of height at most

$$\Delta(\Gamma(C)) \le c(n)d(C)^{4n} \Delta(C, \Gamma)^{1/10},$$

and since $\Delta(\Gamma) \le c(n)\Delta(C)$ we have that $\Delta(C) \le c(n)d(C)^{4n}\Delta(C)^{1/10}$. Hence

$$\Delta(C) \le c(n)d(C)^{5n}.$$

Now consider the uniformisation $\pi : F \to X$, which is definable in an o-minimal structure. As in the proof of 1.2, the point C and each of its conjugates gives rise to a rational point P of height $H(P) \leq c\Delta(C)$ on a suitable definable subset of a suitable power of $\mathrm{GL}_2(\mathbb{R})$. We follow the argument in the proof of Theorem 1.2 (which follows that in [13,14]). If Δ is sufficiently large then the Counting Theorem implies that the above-mentioned definable set contains positive-dimensional real semi-algebraic sets.

Since there are many conjugates of C giving rise to rational points, some positive-dimensional semi-algebraic set must give rise to a moving component of the given dimension and defect. Complexifying the real parameter, there is a larger component of V_n with the same defect. The mixed Ax-Schanuel implies there is a larger geodesic component with the same defect so that (in virtue of Conjecture 7.8) the point C was not n-optimal. This contradiction shows that the complexity of an optimal n-tuple is bounded. Then the degree $d(C)$ and the height $h(C)$ are bounded by some $c(n)$, and so there are only finitely many such C. □

References

[1] J. AX, *On Schanuel's conjectures*, Ann. of Math. **93** (1971), 252–268.

[2] Y. BILU, F. LUCA and A. PIZARRO-MADARIAGA, *Rational products of singular moduli*, J. Number Theory **168** (2016), 397–410.

[3] Y. BILU, D. MASSER and U. ZANNIER, *An effective "theorem of André" for CM-points on a plane curve*, Math. Proc. Cambridge Philos. Soc. **154** (2013), 145–152.

[4] E. BOMBIERI, D. MASSER and U. ZANNIER, *Anomalous subvarieties – structure theorems and applications*, Int. Math. Res. Not. IMRN **19** (2007).

[5] C. DAW and A. HARRIS, *Categoricity of modular and Shimura curves*, J. Inst. Math. Jussieu **16** (2017), 1075–1101.

[6] N. ELKIES, *Shimura curve computations*, In: "Algorithmic Number Theory", Third International Symposium, ANTS-III, Portland, oregon, 1998, J. Buhler (ed.), Lecture Notes in Comp. Sci., Vol. 1423, Springer, Berlin, 1998, 1–47.

[7] E. ERRTHUM, *Singular moduli of Shimura curves*, Canad. J. Math. **63** (2011), 826–861.

[8] É. GAUDRON and G. RÉMOND, *Théorème des périodes et degrés minimaux d'isogénies*, Comment. Math. Helv. **89** (2014), 343–403.

[9] P. HABEGGER, *On the bounded height conjecture*, Int. Math. Res. Not. IMRN **2009**, 860–886.

[10] P. HABEGGER, *Weakly bounded height on modular curves*, Acta Math. Vietnam. **35** (2010), 43–69.

[11] P. HABEGGER, *Effective height upper bounds on algebraic tori*, CIRM Luminy course notes, arXiv:1201.3255.

[12] P. HABEGGER, *Singular moduli that are algebraic units*, Algebra Number Theory **9** (2015), 1515–1524.

[13] P. HABEGGER and J. PILA, *Some unlikely intersections beyond André–Oort*, Compos. Math. **148** (2012), 1–27.

[14] P. HABEGGER and J. PILA, *O-minimality and certain atypical intersections*, Ann. Sci. École Norm. Sup. (4) **49** (2016), 813–858.

[15] S. LANG, "Elliptic Functions", Graduate Texts in Mathematics **112**, second edition, Springer, New York, 1987.

[16] S. LOJASIEWICZ, "Introduction to Complex Analytic Geometry", Birkhäuser, Basel, 1991.

[17] T. LOHER and D. W. MASSER, *Uniformly counting points of bounded height*, Acta Arith. **111** (2004), 277–297.
[18] D. MASSER, *Linear relations on algebraic groups*, In: "New Advances in Transcendence Theory", Baker (ed.), CUP, 1988, 248–262.
[19] D. MASSER and G. WÜSTHOLZ, *Estimating isogenies on elliptic curves*, Invent. Math. **100** (1990), 1–24.
[20] R. PAULIN, *An explicit André-Oort type result for $\mathbb{P}^1(\mathbb{C}) \times \mathbb{G}_{\mathrm{m}}(\mathbb{C})$ based on logarithmic forms*, Publ. Math. Debrecen **88** (2016), 21–33.
[21] F. PELLARIN, *Sur une majoration explicite pour un degré d'isogénie liant deux courbes elliptiques*, Acta Arith. **100** (2001), 203–243.
[22] J. PILA, *O-minimality and the André-Oort conjecture for $\mathbb{C}^n$*, Ann. of Math. **173** (2011), 1779–1840.
[23] J. PILA, *Functional transcendence via o-minimality, lecture notes, LMS-EPSRC minicourse*, 2013, In: "O-minimality and Diophantine Geometry", Jones and Wilkie (eds.), London Math. Soc. Lecture Note Series, Vol. 421, CUP, 2015.
[24] J. PILA and J. TSIMERMAN, *Ax-Schanuel for the j-function*, Duke Math. J. **165** (2016), 2587–2605.
[25] J. PILA and A. WILKIE, *The rational points of a definable set*, Duke Math. J. **133** (2006), 591–616.
[26] R. PINK, *A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang*, 2005 preprint, available from the author's webpage.
[27] B. POIZAT, *L'egalité au cube*, J. Symbolic Logic **66** (2001), 1647–1676.
[28] M. RAYNAUD, *Hauteurs et isogénies*, Seminaire sur les pinceaux arithmetiques: La conjecture de Mordell, Astérisque **127** (1985), 199–234.
[29] J.-P. SERRE, "Lectures on the Mordell-Weil Theorem", *Aspects of Mathematics* E15, Vieweg, Braunschweig, 1989.
[30] G. SHIMURA, *Construction of class fields and zeta functions of algebraic curves*, Ann. of Math. **85** (1967), 58–159.
[31] J. SILVERMAN, *Heights and elliptic curves*, In: "Arithmetic Geometry", Cornell and Silverman (eds), Springer, New York, 1986, 253–265.
[32] J. TSIMERMAN, *Ax-Schanuel and o-minimality*, In: "O-minimality and Diophantine Geometry", Jones and Wilkie (eds.), London Math. Soc. Lecture Note Series, Vol. 421, CUP, 2015.
[33] D. ZAGIER, *Elliptic modular functions and their applications*, In: "The 1-2-3 of Modular Forms", J. Bruinier, G. van der Geer, G. Harder, and D. Zagier (eds.), Universitext, Springer, Berlin, 2008, 1–103.
[34] U. ZANNIER, "Some Problems of Unlikely Intersections in Arithmetic and Geometry", with appendices by D. Masser, Annals of Mathematics Studies, Vol. 181, Princeton University Press, 2012.
[35] B. ZILBER, *Exponential sums equations and the Schanuel conjecture*, J. Lond. Math. Soc. (2) **65** (2002), 27–44.

Mathematical Institute
University of Oxford
Woodstock Road
Oxford OX2 6GG, UK
pila@maths.ox.ac.uk

Department of Mathematics
University of Toronto
215 Huron Street
Toronto HU1601B, Canada
jacobt@math.toronto.edu

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1383-1396

Divisorial Zariski decomposition and some properties of full mass currents

ELEONORA DI NEZZA, ENRICA FLORIS AND STEFANO TRAPANI

Abstract. Let α be a big class on a compact Kähler manifold. We prove that a decomposition $\alpha = \alpha_1 + \alpha_2$ into the sum of a modified nef class α_1 and a pseudoeffective class α_2 is the divisorial Zariski decomposition of α if and only if $\mathrm{vol}(\alpha) = \mathrm{vol}(\alpha_1)$. We deduce from this result some properties of full mass currents.

Mathematics Subject Classification (2010): 32J25 (primary); 32Q15, 32W20 (secondary).

Introduction

The study of the Zariski decomposition started with the work of Zariski [26] who defined it for an effective divisor in a smooth projective surface. Fujita extended the definition to the case of pseudoeffective divisors [13]. Due to the importance of the Zariski decomposition for surfaces, several generalizations to higher dimension exist (see [22] for a survey of these constructions). The divisorial Zariski decomposition for a cohomology class α on a Kähler manifold has been introduced by Boucksom in [7]. If α is the class of a big divisor on a projective manifold, the divisorial Zariski decomposition coincides with the σ-decomposition introduced by Nakayama [20]. The divisorial Zariski decomposition is a decomposition

$$\alpha = Z(\alpha) + \{N(\alpha)\}$$

into a "positive part", the Zariski projection $Z(\alpha)$, whose non-nef locus has codimension at least 2, and a "negative part" $\{N(\alpha)\}$ which is the class of an effective divisor and is rigid. The class $Z(\alpha)$ encodes some important information about α: $Z(\alpha)$ is big if and only if α is and $\mathrm{vol}(\alpha) = \mathrm{vol}(Z(\alpha))$.

The first named author is supported by a Marie Sklodowska Curie individual fellowship 660940–KRF–CY (MSCA-IF).

Received August 27, 2015; accepted in revised form July 5, 2016.
Published online December 2017.

In this note we give a criterion for a sum of two classes to be a divisorial Zariski decomposition. Our main result is:

Main Theorem. *Let X be a compact Kähler manifold of complex dimension n. Let α be a big class on X. Let $\alpha_1 \in H^{1,1}(X, \mathbb{R})$ be a modified nef class and $\alpha_2 \in H^{1,1}(X, \mathbb{R})$ be a pseudoeffective class. Then $\alpha = \alpha_1 + \alpha_2$ is the divisorial Zariski decomposition of α if and only if* $\mathrm{vol}(\alpha) = \mathrm{vol}(\alpha_1)$.

The relations between the Zariski decomposition of numerical classes of cycles on a projective variety and their volume have been largely studied recently in a series of papers [14, 15, 18]. The Main Theorem also goes in this direction: for instance, if X is projective and $\alpha = \{D\}$ is the class of a big divisor, we recover [15, Proposition 5.3] for cycles of codimension 1.

Our proof relies deeply on a result of existence and uniqueness of weak solutions of complex Monge-Ampère equations.

On the other hand the proof in [15] uses the differentiability of the volume function $f(t) = \mathrm{vol}(\alpha + t\{D\})$, which, at the moment, is known to be true only in the algebraic case. In Remark 2.3 we present a proof of the Main Theorem using the differentiability of the volume. As it is proved by Xiao [24, Proposition 1.1] , the differentiability of the volume is equivalent to the following quantitative version of a Demailly's conjecture [8, Conjecture 10.1], that states:

Let X be a compact Kähler manifold of complex dimension n, and let $\alpha, \beta \in H^{1,1}(X, \mathbb{R})$ be two nef classes. Then we have

$$\mathrm{vol}(\alpha - \beta) \geq \alpha^n - n\, \alpha^{n-1} \cdot \beta. \tag{0.1}$$

While this paper was being published, Witt Nyström [21] proved inequality (0.1) for projective manifolds. This, together with Remark 2.3, provides another proof of the Main theorem in the case where X is projective and α_2 is the class of an effective $\mathbb{R}$-divisor.

In the second part of this note we show that the Main Theorem is strictly related to the invariance of finite energy classes under bimeromorphic maps. More precisely, in Theorem 3.6 we show that finite energy classes are inviariant under a bimeromorphic map if and only if the volumes are preserved. This extends to any dimension, [12, Proposition 2.5], where a similar statement is proved in dimension 2 by the first named author using the Hodge index theorem.

We now give a brief outline of this note. Section 1 reviews background material on the divisorial Zariski decomposition and currents with full Monge-Ampère mass. In Section 2 we prove the Main Theorem and in Section 3 we give some applications to full mass currents. In particular we prove Theorem 3.6.

ACKNOWLEDGEMENTS. We would like to thank Sébastien Boucksom for several useful discussions on the subject and for communicating us the proof in Remark 2.3.

1. Preliminaries

Let (X, ω) be a compact Kähler manifold of complex dimension n and let $\alpha \in H^{1,1}(X, \mathbb{R})$ be a real $(1, 1)$-cohomology class. Recall that α is said to be *pseudo-effective* if it can be represented by a closed positive $(1, 1)$-current T; α is *nef* if and only if for any $\varepsilon > 0$ there exists a smooth form $\theta_\varepsilon \in \alpha$ such that $\theta_\varepsilon \geq -\varepsilon\omega$; α is *big* if and only if it can be represented by a *Kähler current*, *i.e.*, if and only if there exists a positive closed $(1, 1)$-current $T \in \alpha$ such that $T \geq \varepsilon\, \omega$ for some $\varepsilon > 0$ and α is a Kähler class if and only if it contains a Kähler form.

Given a smooth representative θ of the class α, it follows from $\partial\bar{\partial}$-lemma that any positive $(1, 1)$-current can be written as $T = \theta + dd^c\varphi$ where the global potential $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is a θ*-plurisubharmonic function* (θ-psh for short), *i.e.*, it is upper semicontinuous and $\theta + dd^c\varphi \geq 0$ in the sense of currents. Here d and d^c are real differential operators defined as

$$d := \partial + \bar{\partial}, \qquad d^c := \frac{i}{2\pi}\left(\bar{\partial} - \partial\right).$$

Let T be a closed positive $(1, 1)$-current. We denote by $\nu(T, x)$ its Lelong number at a point $x \in X$ defined as

$$\nu(T, x) = \nu(\varphi, x) := \sup\{\gamma \geq 0 \,:\, \varphi(z) \,\leq\, \gamma \log d(x, z) + C\},$$

where z is a coordinate in a coordinate neighborhood of x and d is a distance on it. The *Lelong number* of T along a prime divisor D is

$$\nu(T, D) := \inf\{\nu(T, x) \,:\, x \in D\}.$$

We refer the reader to [11] for a more extensive account on Lelong numbers.

There is a unique decomposition of T as a weakly convergent series

$$T = R + \sum_j \lambda_j [D_j],$$

where:

(i) $[D_j]$ is the current of integration over the prime divisor $D_j \subset X$;
(ii) $\lambda_j := \nu(T, D_j) \geq 0$;
(iii) R is a closed positive current with the property that $\operatorname{codim} E_c(R) \geq 2$ for every $c > 0$.

Recall that

$$E_c(R) := \{x \in X \,:\, \nu(R, x) \geq c\},$$

and that this is an analytic subset of X by a famous result due to Siu [23]. Such a decomposition is called the Siu decomposition of T.

Analytic and minimal singularities

A positive current $T = \theta + dd^c\varphi$ is said to have *analytic singularities* if there exists $c > 0$ such that locally on X,

$$\varphi = \frac{c}{2}\log\sum_{j=1}^{N}|f_j|^2 + u,$$

where u is smooth and $f_1, \dots, f_N$ are local holomorphic functions.

If T and T' are two closed positive currents on X, then T' is said to be *less singular* than T if their local potentials satisfy $\varphi \le \varphi' + O(1)$.

A positive current T is said to have *minimal singularities* (inside its cohomology class α) if it is less singular than any other positive current in α. Its θ-psh potentials φ will correspondingly be said to have minimal singularities.

Such θ-psh functions with minimal singularities always exist, one can consider for example

$$V_\theta := \sup\{\varphi\ \theta\text{-psh}, \varphi \le 0 \text{ on } X\}.$$

1.1. Big and modified nef classes

Definition 1.1. If α is a big class, we define its *ample locus* $\mathrm{Amp}(\alpha)$ as the set of points $x \in X$ such that there exists a Kähler current $T \in \alpha$ with analytic singularities and smooth in a neighbourhood of x.

The ample locus $\mathrm{Amp}(\alpha)$ is a Zariski open subset, and it is nonempty thanks to Demailly's regularization result (see [7]).

Observe that a current with minimal singularities $T_{\min} \in \alpha$ has locally bounded potential in $\mathrm{Amp}(\alpha)$.

Definition 1.2. Let α be a big class.

(1) Let $T \in \alpha$ be a positive $(1, 1)$-current, then we set

$$E_+(T) := \{x \in X \ : \ \nu(T, x) > 0\};$$

(2) We define the *non-Kähler locus* of α as

$$E_{nk}(\alpha) := \bigcap_T E_+(T)$$

ranging among all the Kähler currents in α.

By [7, Theorem 3.17(iii)] a class α is Kähler if and only if $E_{nk}(\alpha) = \emptyset$. Moreover by [7, Theorem 3.17(ii)] we have $E_{nk}(\alpha) = X \setminus \mathrm{Amp}(\alpha)$.

Definition 1.3. We say that α is *modified-nef* if and only if for every $\varepsilon > 0$ there exists a closed $(1, 1)$-current $T_\varepsilon \in \alpha$ with $T_\varepsilon \ge -\varepsilon\omega$ and $\nu(T_\varepsilon, D) = 0$ for any prime divisor D.

We recall now an alternative and useful definition of modified nef classes.

Proposition 1.4 ([7, Proposition 3.2]). *Let $\alpha \in H^{1,1}(X, \mathbb{R})$ be a pseudoeffective class. Then α is modified nef if and only if $\nu(\alpha, D) = 0$ for every prime divisor D.*

We refer to [7] for the defintion and properties of the *minimal multiplicity* $\nu(\alpha, D)$. We will be only interested in the case where α is big, and in this case the minimal multiplicity coincides with $\nu(T_{\min}, D)$, the Lelong number along D of a current in α with minimal singularities (*cf.* [7, Proposition 3.6(ii)]).

1.2. The divisorial Zariski decomposition

In this subsection we collect some basic results on the divisorial Zariski decomposition defined in [7]. They can all be found in [7] but we recall some statements frequently used in this note.

Let $\alpha \in H^{1,1}(X, \mathbb{R})$ be a pseudo-effective class. The *divisorial Zariski decomposition* of α is defined as follows:

Definition 1.5. The negative part of α is defined as $N(\alpha) := \sum \nu(\alpha, D)[D]$, where D are prime divisors. The Zariski projection of α is $Z(\alpha) := \alpha - \{N(\alpha)\}$. We call the decomposition $\alpha = Z(\alpha) + \{N(\alpha)\}$ the divisorial Zariski decomposition of α.

Properties. Let $\alpha = Z(\alpha) + \{N(\alpha)\}$ be the divisorial Zariski decomposition of α. Then

(1) The class $Z(\alpha)$ is modified nef [7, Proposition 3.8];
(2) $N(\alpha)$ is a divisor, *i.e.* there is a finite number of prime divisors D such that $\nu(\alpha, D) > 0$ [7, Proposition 3.12];
(3) The set of modified nef classes is a closed convex cone and it is the closure of the convex cone generated by the classes $\mu_\star\tilde{\alpha}$ where $\mu : \tilde{X} \to X$ is a modification and $\tilde{\alpha}$ is a Kähler class on $\tilde{X}$ [7, Proposition 2.3];
(4) The negative part $\{N(\alpha)\}$ is a *rigid* class, *i.e.* it contains only one positive current [7, Proposition 3.13];
(5) Let α be a modified nef and big class, $D_1, \ldots, D_k$ be prime divisors and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}^+$. Then [7, Proposition 3.18]

$$N\left(\alpha + \sum_i \lambda_i \{D_i\}\right) = \sum_i \lambda_i [D_i]$$

if and only if $D_j \subset E_{nk}(\alpha)$ for any j.

Proposition 1.6 ([7, Proposition 3.6(ii)]). *Let $\alpha \in H^{1,1}(X, \mathbb{R})$ be a big class and let $T_{\min} \in \alpha$ be a current with minimal singularities. Consider the Siu decomposition of $T_{\min}$,*

$$T_{\min} = R + \sum_j a_j [D_j]$$

where $a_j = \nu(T_{\min}, D_j)$. Then $\{R\} = Z(\alpha)$ and $\{\sum_j a_j D_j\} = \{N(\alpha)\}$. In particular, $\nu(\alpha, D) = \nu(T_{\min}, D)$ for any prime divisor D.

1.3. Volume of big classes

Fix $\alpha \in H^{1,1}_{big}(X, \mathbb{R})$. We introduce

Definition 1.7. Let $T_{\min}$ be a current with minimal singularities in α and let Ω a Zariski open set on which the potentials of $T_{\min}$ are locally bounded, then

$$\mathrm{vol}(\alpha) := \int_\Omega T_{\min}^n > 0 \tag{1.1}$$

is called the volume of α.

Note that the Monge-Ampère measure of $T_{\min}$ is well defined in Ω by [1] and that the volume is independent of the choice of $T_{\min}$ and Ω [4, Theorem 1.16].

Let $f : X \to Y$ be a birational modification between compact Kähler manifolds and let $\alpha_Y \in H^{1,1}(Y, \mathbb{R})$ be a big class. The volume is preserved by pull-backs,

$$\mathrm{vol}(f^\star \alpha_Y) = \mathrm{vol}(\alpha_Y)$$

(see [6]). On the other hand, it is not preserved by push-forwards. In general we have

$$\mathrm{vol}(f_\star \alpha_X) \geq \mathrm{vol}(\alpha_X)$$

(see Remark 3.4).

1.4. Full mass currents

Fix X a n-dimensional compact Kähler manifold, $\alpha \in H^{1,1}(X, \mathbb{R})$ be a big class and $\theta \in \alpha$ a smooth representative.

The non-pluripolar product

Given $T_1 := \theta_1 + dd^c\varphi_1, \dots, T_p := \theta_p + dd^c\varphi_p$ positive $(1, 1)$-currents, where θ_j are closed smooth $(1, 1)$-forms, following the construction of Bedford-Taylor [2] in the local setting, it has been shown in [4, Proposition 1.6] that the sequence of currents

$$\mathbf{1}_{\bigcap_j \{\varphi_j > V_{\theta_j} - k\}} \big(\theta_1 + dd^c \max(\varphi_1, V_{\theta_1} - k)\big) \wedge \dots \wedge \big(\theta_p + dd^c \max(\varphi_p, V_{\theta_p} - k)\big)$$

is non-decreasing in k and converges weakly to the so-called *non-pluripolar product*

$$\big\langle T_1 \wedge \dots \wedge T_p \big\rangle.$$

The resulting positive (p, p)-current does not charge pluripolar sets and it is *closed*. In the sequel we will focus on the particular case when $T_1 = \dots = T_p = T$ and $p = n$. We denote by $\langle T^n \rangle$ the *non-pluripolar measure* of T.

Let us stress that since the non-pluripolar product does not charge pluripolar sets,

$$\mathrm{vol}(\alpha) = \int_X \big\langle T_{\min}^n \big\rangle \tag{1.2}$$

whereas by [4, Proposition 1.20] for any positive (1, 1)-current $T \in \alpha$ we have

$$\mathrm{vol}(\alpha) \geq \int_X \langle T^n \rangle. \tag{1.3}$$

Definition 1.8. A closed positive (1, 1)-current T on X with cohomology class α is said to have *full Monge-Ampère mass* if

$$\int_X \langle T^n \rangle = \mathrm{vol}(\alpha).$$

We denote by $\mathcal{E}(X, \alpha)$ the set of such currents. Let φ be a θ-psh function such that $T = \theta + dd^c\varphi$. The *non-pluripolar Monge-Ampère measure* of φ is

$$\mathrm{MA}(\varphi) := \langle (\theta + dd^c\varphi)^n \rangle = \langle T^n \rangle.$$

We will say that φ has *full Monge-Ampère mass* if $\theta + dd^c\varphi$ has full Monge-Ampère mass. We denote by $\mathcal{E}(X, \theta)$ the set of corresponding functions.

2. Proof of the main theorem

Throughout this section X and Y will be compact Kähler manifolds of complex dimension n.

Theorem 2.1. *Let α be a big class on X. Let $\alpha_1 \in H^{1,1}(X, \mathbb{R})$ be a modified nef class and $\alpha_2 \in H^{1,1}(X, \mathbb{R})$ be a pseudoeffective class. Then $\alpha = \alpha_1 + \alpha_2$ is the divisorial Zariski decomposition of α if and only if* $\mathrm{vol}(\alpha) = \mathrm{vol}(\alpha_1)$.

Remark 2.2. In particular, Theorem 2.1 implies that the pseudoeffective class α_2 will be of the form $\alpha_2 = \sum_{j=1}^{N} \lambda_j \{D_j\}$ where D_j are prime divisors and $\lambda_j = \nu(\alpha, D_j) \geq 0$.

Proof of Theorem 2.1. If $\alpha = \alpha_1 + \alpha_2$ is the divisorial Zariski decomposition then by [7, Proposition 3.20] we have $\mathrm{vol}(\alpha) = \mathrm{vol}(\alpha_1)$.

Vice versa, assume that we have a decomposition as above with $\mathrm{vol}(\alpha) = \mathrm{vol}(\alpha_1) = V$. Let μ be a smooth volume form on X with total mass V and let $T_1 \in \mathcal{E}(X, \alpha_1)$ be the unique solution of the complex Monge-Ampère equation

$$\langle T_1^n \rangle = \mu.$$

Such T_1 exists and is unique by [4, Theorem 3.1]. Furtheremore, T_1 has minimal singularities in its cohomology class [4, Theorem 4.1]. Let τ be any positive closed (1, 1)-current in α_2 and set $T = T_1 + \tau$. By multilinearity of the non-pluripolar product [4, Proposition 1.4], we have $\langle T^n \rangle \geq \langle T_1^n \rangle$. By (1.2) and (1.3) we have

$$\int_X \langle T^n \rangle \leq \mathrm{vol}(\alpha) = \mathrm{vol}(\alpha_1) = \int_X \langle T_1^n \rangle.$$

Therefore $\langle T^n\rangle = \langle T_1^n\rangle = \mu$. Thus T is a solution of the Monge-Ampère equation $\langle T^n\rangle = \mu$ in the class α and by uniqueness, it follows that α_2 is rigid, *i.e.* there exists a unique positive closed $(1,1)$-current in α_2. Moreover, T has minimal singularities. Since $\mathrm{vol}(\alpha) = \int_X \langle T^n\rangle = \mu(X)$, by the multilinearity of the non-pluripolar product we get

$$\sum_{j=0}^{n-1} \binom{n}{j} \langle T_1^j \wedge \tau^{n-j}\rangle = 0.$$

Let $S \in \alpha_1$ be a Kähler current, *i.e.* $S \geq \varepsilon\omega$ for some $\varepsilon > 0$. Let Ω_1 be a non-empty Zariski open subset where S is smooth and let $\Omega = \mathrm{Amp}(\alpha) \neq \emptyset$. Since T has minimal singularities, then $T \in \alpha$ has locally bounded potential on Ω. In particular, the current τ has locally bounded potential in $\Omega_2 = \Omega \cap \Omega_1 = X \setminus \Sigma$. Then we have

$$0 \leq \varepsilon^{n-1} \int_{\Omega_2} \omega^{n-1} \wedge \tau \leq \int_{\Omega_2} S^{n-1} \wedge \tau \leq \int_{\Omega_2} T_1^{n-1} \wedge \tau = 0,$$

where the last inequality follows from [4, Proposition 1.20]. This implies that the current τ is supported on Σ.

By [11, Corollary 2.14], τ is of the form

$$\tau = \sum_{j=1}^{N} \lambda_j [D_j],$$

where D_j are irreducibile divisors and $\lambda_j \geq 0$. Moreover, observe that, since α_1 is modified nef and T_1 has minimal singularities, we have $\nu(T_1, D_j) = 0$ for any j by Proposition 1.4 hence $\lambda_j = \nu(T, D_j)$. In other words, $T = T_1 + \tau$ is the Siu decomposition of T. Since α is big and T has minimal singularities, by Proposition 1.6 we have $\nu(\alpha, D) = \nu(T, D)$, hence the conclusion. □

We would like to observe that in the algebraic case, for a projective manifold X, Theorem 2.1 can be proved using the differentiability of the volume [9].

We thank Sébastien Boucksom for the following remark:

Remark 2.3. Let $N^1(X)_{\mathbb{R}} \subset H^{1,1}(X, \mathbb{R})$ denote the real Néron-Severi space and $\alpha \in N^1(X)_{\mathbb{R}}$ be a big class. Assume $\alpha = \alpha_1 + \sum_{i=1}^N \lambda_i \{D_i\}$ with

(i) $\alpha_1 \in N^1(X)_{\mathbb{R}}$ a modified nef class such that $\mathrm{vol}(\alpha) = \mathrm{vol}(\alpha_1)$;
(ii) $\lambda_i \geq 0$;
(iii) D_i are prime divisors for any i.

Then $\alpha = \alpha_1 + \sum_{i=1}^N \lambda_i \{D_i\}$ is the divisorial Zariski decomposition of α. We claim that it is enough to prove that for any prime divisor $D \not\subset E_{nk}(\alpha)$,

$$\mathrm{vol}(\alpha_1 + tD) > \mathrm{vol}(\alpha_1) \quad \forall t > 0. \tag{2.1}$$

Indeed, to prove that $\alpha = \alpha_1 + \sum_{i=1}^{N} \lambda_i \{D_i\}$ is the divisorial Zariski decomposition of α, we have to check that $D_i \subset E_{nk}(\alpha_1)$ by Property 1.2(5). If $\lambda_i > 0$ and $D_i \not\subset E_{nk}(\alpha_1)$ then (2.1) yields

$$\operatorname{vol}(\alpha) \geq \operatorname{vol}(\alpha_1 + \lambda_i D_i) > \operatorname{vol}(\alpha_1) = \operatorname{vol}(\alpha),$$

hence a contradiction.

The inequality (2.1) easily follows from the differentiability of the volume. Indeed, by [9, Theorem A] we have

$$\frac{d}{dt}\Big|_{t=0} \operatorname{vol}(\alpha_1 + tD) = n\big\langle \alpha_1^{n-1} \big\rangle \cdot D$$

where $\langle \alpha_1^{n-1} \rangle$ denotes *the positive product* of α defined in [4, Definition 1.17]. Thanks to [9, Remark 4.2 and Theorem 4.9], we have $\langle \alpha_1^{n-1} \rangle \cdot D > 0$, hence $\operatorname{vol}(\alpha_1 + tD)$ is a continuous strictly increasing function for small $t > 0$, and so $\operatorname{vol}(\alpha_1 + tD) > \operatorname{vol}(\alpha_1)$.

Using the results in [21] by Witt Nyström and Boucksom, the above proof works when X is projective and $\alpha \in H^{1,1}(X, \mathbb{R})$.

3. Currents with full Monge-Ampère mass

In this section we state a few consequences of Theorem 2.1. The first result states that currents with full Monge-Ampère mass in α compute the coefficients of the divisorial Zariski decomposition of α.

Theorem 3.1. *Let α be a big class on X. If $T \in \mathcal{E}(X, \alpha)$ and $T_{\min} \in \alpha$ is a current with mininal singularities, then the set*

$$\big\{x \in X \,:\, \nu(T, x) > \nu(T_{\min}, x)\big\}$$

is contained in a countable union of analytic subsets of codimension ≥ 2 contained in $E_{nK}(\alpha)$. In particular, $\nu(T, D) = \nu(T_{\min}, D)$ for any irreducible divisor $D \subset X$.

Proof. If $T \in \mathcal{E}(X, \alpha)$ then $E_+(T) \subset E_{nk}(\alpha)$ because of [12, Proposition 1.9]. On the other hand if we write the Siu decomposition of T as

$$T = T_1 + \sum_{j \geq 1} \lambda_j \big[D_j\big],$$

where D_j are prime divisors and $\operatorname{codim} E_c(T_1) \geq 2$ for all $c > 0$, we have $D_j \subset X \setminus \operatorname{Amp}(\alpha)$. Hence there is a finite number of D_j such that $\lambda_j \neq 0$. In particular, $\nu(T_1, D_j) = 0$ for any j.

Set $\alpha_1 := \{T_1\}$ and note that, since α is big, α_1 is big. Moreover, α_1 is modified nef. Indeed, pick $T_{\min,1} \in \alpha_1$ a current with minimal singularities. Since $0 \leq \nu(T_{\min,1}, D_j) \leq \nu(T_1, D_j) = 0$, we have $\nu(T_{\min,1}, D) = 0$ for any D prime divisor. The claim then follows from Propositions 1.4 and 1.6.

Furthermore, the current $S = T_{\min,1} + \sum_{j=1}^{N} \lambda_j [D_j]$ is less singular than T, hence it has full Monge-Ampère mass [4, Corollary 2.3]. Therefore

$$\mathrm{vol}(\alpha) = \int_X \langle T^n \rangle = \int_X \langle S^n \rangle = \int_X \langle T^n_{\min,1} \rangle = \mathrm{vol}(\alpha_1).$$

We are now under the assumptions of Theorem 2.1, thus $\alpha = \alpha_1 + \sum_{j\geq 1} \lambda_j [D_j]$ is the divisorial Zariski decompostion of α and

$$\nu\big(T, D_j\big) = \lambda_j = \nu\big(\alpha, D_j\big) = \nu\big(T_{\min}, D_j\big),$$

where the last identity is Proposition 1.6.

Moreover,

$$B := \big\{x \in X \ : \ \nu(T, x) > \nu(T_{\min}, x)\big\} \subset \bigcup_{c\in\mathbb{Q}^+} E_c(T_1) \cup \bigcup_{j=1}^{N} \Sigma_j,$$

where $\Sigma_j := \{x \in D_j \ : \ \nu(T, x) > \lambda_j\}$. Indeed, if $x \in B$ is such that $x \in X \setminus \bigcup_{j=1}^{N} D_j$ then $\nu(T, x) = \nu(T_1, x) > \nu(T_{\min,1}, x) \geq 0$. If $x \in D_j$ for some j and $x \in B$ then $\nu(T, x) > \nu(T_{\min}, D_j) = \lambda_j$, that is $x \in \Sigma_j$. Finally, observe that by [23] both $E_c(T_1)$ and Σ_j are analytic subsets of codimension ≥ 2 for any $c > 0$ and j, respectively. □

Remark 3.2. In [19, Theorem 1.1 and Lemma 5.4] Lesieutre constructs an example of a big class α on a 4-dimensional manifold X whose non-nef locus $E_{nn}(\alpha)$ is an infinite countable union of irreducible curves and it is Zariski dense in a divisor $E \subset X$. Hence α is modified nef and Theorem 3.1 implies that if $T \in \mathcal{E}(X, \alpha)$ then the set $E_+(T) := \{x \in X \ : \ \nu(T, x) > 0\}$ contains $E_{nn}(\alpha)$ but it does not contain E. Therefore $E_+(T)$ is not a closed analytic subset. In particular, there does not exist a positive current with analytic singularities $T_+ \in \alpha$ that has full Monge-Ampère mass.

In [12], the first named author proved that finite energy classes (and in particular the energy class $\mathcal{E}$ defined in section 3) are in general not preserved by bimeromorphic maps (see [12, Example 1.7 and Proposition 2.3]). In order to circumvent this problem she introduced a natural condition.

Definition 3.3. Let $f : X \dashrightarrow Y$ be a bimeromorphic map and α be a big class on X. Let $\mathcal{T}_\alpha(X)$ denote the set of positive closed $(1,1)$-currents in α. We say that *Condition* (V) is satisfied if

$$f_\star\Big(\mathcal{T}_\alpha(X)\Big) = \mathcal{T}_{f_\star\alpha}(Y),$$

where $\mathcal{T}_{f_\star\alpha}(Y)$ is the set of positive currents in the image class $f_\star\alpha$.

Remark 3.4. Note that in general we have $f_\star\big(\mathcal{T}_\alpha(X)\big) \subseteq \mathcal{T}_{f_\star\alpha}(Y)$. This means in particular that the push-forward of a current with minimal singularities in α_X has not necessarily minimal singularities in $f_\star\alpha_X$, hence $\mathrm{vol}(f_\star\alpha_X) \geq \mathrm{vol}(\alpha_X)$.

The first named author showed in [12, Proposition 2.3] that Condition (V) implies that $f_\star\mathcal{E}(X, \alpha) = \mathcal{E}(Y, f_\star\alpha)$.

In the following we prove that Condition (V) is equivalent to the preservation of volumes.

Lemma 3.5. *Let $f : X \to Y$ be a birational morphism and let α be a big class on X. Let E_i, F_i be distinct prime divisors contained in the exceptional locus* $\mathrm{Exc}(f)$ *of f, then there exist $a_i, b_i \in \mathbb{R}^+$ such that*

$$\alpha = f^\star f_\star \alpha - \left[\sum_i a_i\{E_i\} - \sum_i b_i\{F_i\} \right]. \tag{3.1}$$

Moreover, Condition ($\mathtt{V}$) *is equivalent to:*

(i) $a_i \leq \nu(f^\star f_\star \alpha, E_i)$ *for any i;*
(ii) $-b_i \leq \nu(f^\star f_\star \alpha, F_i)$ *for any i.*

The statements in Lemma 3.5 are quite standard but we include a proof for the reader's convenience.

Proof. The identity (3.1) follows from the fact that for any $T \in \alpha$ positive $(1, 1)$-current, $T - f^\star f_\star T$ is supported on $\mathrm{Exc}(f)$ since f is a biholomorphism on $X \setminus \mathrm{Exc}(f)$. Therefore we conclude by [11, Corollary 2.14].

Assume Condition ($\mathtt{V}$) holds, that is, that any positive $(1, 1)$-current $S \in f_\star\alpha$ can be written as $S = f_\star T$ for some positive $(1, 1)$-current $T \in \alpha$. Since the cohomology classes of the excetional divisors of f are linearly independent, by (3.1) we have an identity of currents,

$$T + \sum_i a_i[E_i] = f^\star f_\star T + \sum_i b_i[F_i].$$

Thus, for any i we have $\nu(f^\star f_\star T, E_i) - a_i \geq 0$ and $\nu(f^\star f_\star T, F_i) + b_i \geq 0$. Hence (i) and (ii) since Condition ($\mathtt{V}$) holds in particular for currents with minimal singularities in $f_\star\alpha$.

Conversely, let $S \in f_\star\alpha$ be a positive $(1, 1)$-current. By the Siu decomposition the current

$$f^\star S - \sum_i \nu(f^\star S, E_i)[E_i] - \sum_i \nu(f^\star S, F_i)[F_i]$$

is positive. For any i, set $\lambda_i := \nu(f^\star S, E_i) - a_i$ and $\mu_i := \nu(f^\star S, F_i) + b_i$ and observe $\lambda_i, \mu_i \geq 0$ by (i) and (ii). Then

$$T := f^\star S - \sum_i \nu(f^\star S, E_i)[E_i] - \sum_i \nu(f^\star S, F_i)[F_i] + \sum_i \lambda_i[E_i] + \sum_i \mu_i[F_i]$$

is a positive $(1, 1)$-current in α and by construction we have $f_\star T = S$. $\square$

Theorem 3.6. *Let $f : X \dashrightarrow Y$ be a bimeromorphic map and let α be a big class on X. Then Condition ($\mathtt{V}$) holds if and only if* $\mathrm{vol}(\alpha) = \mathrm{vol}(f_\star\alpha)$.

Proof. Condition ($\mathtt{V}$) insures that there exists a positive current $T \in \alpha$ such that $f_\star T$ is a current with minimal singularities in $f_\star\alpha$. Then

$$\mathrm{vol}(\alpha) \geq \int_X \langle T^n \rangle = \int_Y \langle (f_\star T)^n \rangle = \mathrm{vol}(f_\star\alpha).$$

By Remark 3.4 we get $\mathrm{vol}(\alpha) = \mathrm{vol}(f_\star\alpha)$.

Let us now prove the converse implication. First, observe that, applying a resolution of singularities, a bimeromorphic map $f : X \dashrightarrow Y$ can be decomposed as $f = h^{-1} \circ g$,

$$\begin{tikzcd} & Z \arrow[dl, "h"'] \arrow[dr, "g"] & \\ X & & Y, \end{tikzcd}$$

where h, g are two birational morphisms and Z denotes a resolution of singularities for the graph of f. By the proof of [4, Proposition 1.12], for every birational morphism h we have $h^\star\big(\mathcal{T}_\alpha(X)\big) = \mathcal{T}_{h^\star\alpha}(Z)$, hence it suffices to prove the claim when f is a birational morphism.

Let E_i, F_i and a_i, b_i as in (3.1). By Lemma 3.5, Condition (V) is equivalent to:

(i) $a_i \leq \nu(f^\star f_\star\alpha, E_i)$ for any i;
(ii) $-b_i \leq \nu(f^\star f_\star\alpha, F_i)$ for any i.

Condition (ii) is satisfied since $\nu(f^\star f_\star\alpha, F_i) \geq 0$. Thus we are left to prove (i).

Consider $\beta := f^\star f_\star\alpha + \sum_i b_i\{F_i\}$. We notice that $f_\star\beta = f_\star\alpha$. Moreover, by Lemma 3.5, β satisfies Condition (V). Indeed, for any i we have $-b_i \leq \nu(f^\star f_\star\beta, F_i) = \nu(f^\star f_\star\alpha, F_i)$. By the first implication of this theorem, we get $\mathrm{vol}(\beta) = \mathrm{vol}(f_\star\beta) = \mathrm{vol}(f_\star\alpha)$.

Let $T_{\min} \in \alpha$ and $S_{\min} \in f_\star\alpha$ be currents with minimal singularities. Then $T_{\min} + \sum_i a_i[E_i]$ and $f^\star S_{\min} + \sum_i b_i[F_i]$ are both positive $(1, 1)$-currents in β with full Monge-Ampère mass. Indeed,

$$\int_X \left\langle \left(T_{\min} + \sum_i a_i[E_i]\right)^n \right\rangle = \int_X \langle T_{\min}^n \rangle = \mathrm{vol}(\alpha)$$
$$\int_X \left\langle \left(f^\star S_{\min} + \sum_i b_i[F_i]\right)^n \right\rangle = \int_Y \langle S_{\min}^n \rangle = \mathrm{vol}(f_\star\alpha),$$

and $\mathrm{vol}(\alpha) = \mathrm{vol}(f_\star\alpha) = \mathrm{vol}(\beta)$. By Theorem 3.1

$$a_j \leq \nu\left(T_{\min} + \sum_i a_i[E_i], E_j\right) = \nu\left(f^\star S_{\min} + \sum_i b_i[F_i], E_j\right) = \nu\big(f^\star S_{\min}, E_j\big)$$

for any prime divisor E_j, since the prime divisors F_i and E_j are distinct. By Proposition 1.6, $a_j \leq \nu(f^\star S_{\min}, E_j) = \nu(f^\star f_\star\alpha, E_j)$, hence the conclusion. $\square$

Theorem 3.7. *Let α be a big class and D be an irreducible divisor such that $D \cap \mathrm{Amp}(\alpha) \neq \emptyset$. Then*

$$\mathrm{vol}(\alpha + tD) > \mathrm{vol}(\alpha) \quad \forall t > 0.$$

Vice versa, if $D \cap \mathrm{Amp}(\alpha) = \emptyset$ then

$$\mathrm{vol}(\alpha + tD) = \mathrm{vol}(\alpha) \quad \forall t > 0.$$

Proof. We first reduce to the case α modified nef and big class. Let $\alpha = Z(\alpha) + \{N(\alpha)\}$ be the divisorial Zariski decomposition of α. By Lemma 3.8 $D \cap \mathrm{Amp}(\alpha) \neq \emptyset$ if and only if $D \cap \mathrm{Amp}(Z(\alpha)) \neq \emptyset$.

If the theorem is true for modified nef and big classes, we have

$$\mathrm{vol}(\alpha + tD) \geq \mathrm{vol}(Z(\alpha) + tD) > \mathrm{vol}(Z(\alpha)) = \mathrm{vol}(\alpha).$$

Thus we can assume that α is a modified nef and big class. Assume by contradiction that there exists t_0 such that $\mathrm{vol}(\alpha + t_0 D) = \mathrm{vol}(\alpha)$. It follows by Theorem 2.1 that $\beta = \alpha + t_0 D$ is the divisorial Zariski decomposition of β and so $D \subset E_{nk}(\alpha)$ Property 1.2(5). Since $E_{nk}(\alpha) = X \setminus \mathrm{Amp}(\alpha)$ [7, Proposition 3.17] we get a contradiction.

Vice versa, if $\alpha = Z(\alpha) + \{N(\alpha)\}$ is the divisorial Zariski decomposition of α and $D \cap \mathrm{Amp}(\alpha) = \emptyset$ (or equivalently $D \subset E_{nk}(Z(\alpha))$ by Lemma 3.8 below and [7, Theorem 3.17]) then by Property 1.2(5) we have that, for any $t > 0$, the divisorial Zariski decomposition of $\alpha + tD$ is

$$\alpha + tD = Z(\alpha) + (N(\alpha) + tD),$$

thus $\mathrm{vol}(\alpha + tD) = \mathrm{vol}(Z(\alpha)) = \mathrm{vol}(\alpha)$. ☐

Lemma 3.8. *Let $\alpha \in H^{1,1}_{\mathrm{big}}(X, \mathbb{R})$ and let $\alpha = Z(\alpha) + \{N(\alpha)\}$ be its divisorial Zariski decomposition. Then we have*

$$\mathrm{Amp}(\alpha) = \mathrm{Amp}(Z(\alpha)).$$

Proof. We first show the inclusion $\mathrm{Amp}(\alpha) \subset \mathrm{Amp}(Z(\alpha))$. Pick $x \in \mathrm{Amp}(\alpha)$. By definition there exists a Kähler current with analytic singularities $T \in \alpha$ which is smooth in a neighbourhood of x. Moreover $\nu(T_{\min}, x) = 0$ since $0 = \nu(T, x) \geq \nu(T_{\min}, x)$. Let $T = R + \sum_j a_j[D_j]$ be the Siu decomposition of T, then $x \notin \mathrm{supp}\, D_j$ for any j. The current $T - N(\alpha) \in Z(\alpha)$ has clearly analytic singularities, is smooth around x and it is also Kähler since $N(\alpha) \leq \sum_j a_j[D_j]$ by Proposition 1.6. Hence $x \in \mathrm{Amp}(Z(\alpha))$. Conversely, pick $x \in \mathrm{Amp}(Z(\alpha))$, then there exists a Kähler current with analytic singularities $T \in Z(\alpha)$ that is smooth in a neighbourhood of x (see Definition 1.1). It follows from Property 1.2(5) that $x \notin \mathrm{supp}\, N(\alpha)$. This implies that $T + N(\alpha) \in \alpha$ is a Kähler current with analytic singularites that is smooth in a neighbourhood of x. Hence $x \in \mathrm{Amp}(\alpha)$. ☐

References

[1] E. BEDFORD and B. A. TAYLOR, *A new capacity for plurisubharmonic functions*, Acta Math. **149** (1982), 1–40.

[2] E. BEDFORD and B. A. TAYLOR, *Fine topology, Silov boundary, and $(dd^c)^n$*, J. Funct. Anal. **72** (1987), 225–251.

[3] R. BERMAN, S. BOUCKSOM, V. GUEDJ and A. ZERIAHI, *A variational approach to complex Monge-Ampère equations*, Publ. Math. Inst. Hutes Études Sci. **117** (2013), 179–245.

[4] S. BOUCKSOM, P. EYSSIDIEUX, V. GUEDJ and A. ZERIAHI, *Monge-Ampère equations in big cohomology classes*, Acta Math. **205** (2010), 199–262.

[5] S. BOUCKSOM, "Cones positifs des variétés complexes compactes", PhD Thesis available at http://www-fourier.ujf-grenoble.fr/demailly/theses/boucksom.
[6] S. BOUCKSOM, *On the volume of a line bundle*, Internat. J. Math. **13** (2002), 1043–1063.
[7] S. BOUCKSOM, *Divisorial Zariski decompositions on compact complex manifolds*, Ann. Sci. Éc. Norm. Supér. (4) **37** (2004), 45–76.
[8] S. BOUCKSOM, J. P. DEMAILLY, M. PĂUN T. and PETERNELL, *The pseudo-effective cone of a compact Kähler manifold and varieties of negative Kodaira dimension*, J. Algebraic Geom. **22** (2013), 201–248.
[9] S. BOUCKSOM, C. FAVRE and M. JONSSON, *Differentiability of volumes of divisors and a problem of Teissier*, J. Algebraic Geom. **18** (2009), 279–308.
[10] J. P. DEMAILLY, *Regularization of closed positive currents and intersection theory*, J. Algebraic Geom. **1** (1992), 361–409.
[11] J. P. DEMAILLY, "Complex Analytic and Differential Geometry", Book available at http://www-fourier.ujf-grenoble.fr/ demailly/documents.html.
[12] E. DI NEZZA, *Stability of Monge-Ampère energy classes*, J. Geom. Anal. **25** (2015), 2565–2589.
[13] T. FUJITA, *On Zariski problem*, Proc. Japan Acad. **55** (1979), 106–110.
[14] M. FULGER, J. KOLLÁR and B. LEHMANN, *Volume and Hilbert function of $\mathbb{R}$-divisors*, Michigan Math. J. **65** (2016), 371–387.
[15] M. FULGER and B. LEHMANN, *Zariski decompositions of numerical cycle classes*, J. Algebraic Geom. **26** (2017), 43–106.
[16] V. GUEDJ and A. ZERIAHI, *Intrinsic capacities on compact Kähler manifolds*, J. Geom. Anal. **15** (2005), 607–639.
[17] V. GUEDJ and A. ZERIAHI, *The weighted Monge-Ampère energy of quasipsh functions*, J. Funct. Anal. **250** (2007), 442–82.
[18] B. LEHMANN, *Geometric characterizations of big cycles*, arXiv:1309.0880.
[19] J. LESIEUTRE, *The diminished base locus is not always closed*, Compositio Math. **150** (2014) 1729–1741.
[20] N. NAKAYAMA, "Zariski-Decomposition and Abundance", MSJ Memoirs, 14, Mathematical Society of Japan, Tokyo, 2004.
[21] D. WITT NYSTRÖM with an appendix by S. BOUCKSOM, *Duality between the pseudoeffective and the movable cone on a projective manifold*, arXiv:1602.03778.
[22] Y. PROKHOROV, *On Zariski decomposition problem*, Proc. Steklov Inst. Math. **240** (2003), 37–65.
[23] Y. T. SIU, *Analyticity of sets associated to Lelong numbers and the extension of meromorphic maps*, Bull. Amer. Math. Soc. **79** (1973), 1200–1205.
[24] J. XIAO, *Movable intersection and bigness criterion*, arXiv:1405.1582.
[25] J. WŁODARCZYK, *Toroidal varieties and the weak factorization theore*, Invent. Math. **154** (2003), 223–331.
[26] O. ZARISKI, *The theorem of Riemann-Roch for high multiples of an effective divisor on an algebraic surface*, Ann. of Math. (2) **76**, (1962), 560–615.

Department of Mathematics
Imperial College
South Kensington Campus
London SW7 2AZ, UK
e.di-nezza@imperial.ac.uk
e.floris@imperial.ac.uk

Dipartimento di Matematica
Università di Roma "Tor Vergata"
Via della Ricerca Scientifica, 1
00133 Roma, Italia
trapani@axp.mat.uniroma2.it

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1397-1416

On the Kodaira-Spencer map of Abelian schemes

YVES ANDRÉ

Abstract. Let A be an Abelian scheme over a smooth affine complex variety S, Ω_A the $\mathcal{O}_S$-module of 1-forms of the first kind on A, $\mathcal{D}_S\Omega_A$ the $\mathcal{D}_S$-module spanned by Ω_A in the first algebraic De Rham cohomology module, and $\theta_\partial : \Omega_A \to \mathcal{D}_S\Omega_A/\Omega_A$ the Kodaira-Spencer map attached to a tangent vector field ∂ on S. We compare the rank of $\mathcal{D}_S\Omega_A/\Omega_A$ to the maximal rank of θ_∂ when ∂ varies: we show that both ranks do not change when one passes to the "modular case", *i.e.* when one replaces S by the smallest weakly special subvariety of $\mathcal{A}_g$ containing the image of S (assuming, as one may up to isogeny, that A/S is principally polarized); we then analyse the "modular case" and deduce, for instance, that for any Abelian pencil of relative dimension g with Zariski-dense monodromy in Sp_{2g}, the derivative with respect to a parameter of a non zero Abelian integral of the first kind is never of the first kind.

Mathematics Subject Classification (2010): 14K20 (primary); 14G35, 32G20 (secondary).

This paper deals with Abelian integrals depending algebraically on parameters and their derivatives with respect to the parameters. Since the nineteenth century, it has been known that differentiation with respect to parameters does not preserve Abelian integrals of the first kind in general.

We study this phenomenon in the language of modern algebraic geometry, *i.e.* in terms of the algebraic De Rham cohomology $\mathcal{O}_S$-module $\mathcal{H}^1_{dR}(A/S)$ attached to an Abelian scheme A of relative dimension g over a smooth $\mathbb{C}$-scheme S, its submodule Ω_A of forms of the first kind on A, the Gauss-Manin connection ∇ and the associated Kodaira-Spencer map θ, *i.e.* the $\mathcal{O}_S$-linear map $T_S \otimes \Omega_A \stackrel{\theta}{\to} \mathcal{H}^1_{dR}(A/S)/\Omega_A$ induced by ∇.

We introduce and compare the following (generic) "ranks":

- $r = r(A/S) = \operatorname{rk} \mathcal{D}_S\Omega_A/\Omega_A$;
- $r' = r'(A/S) = \operatorname{rk} \theta$;
- $r'' = r''(A/S) = \max_\partial \operatorname{rk} \theta_\partial$;

where ∂ runs over local tangent vector fields on S (of course, $r'' = r'$ when S is a curve).

Received June 6, 2016; accepted in revised form August 2, 2016.
Published online December 2017.

One has $r'' \leq r' \leq r \leq g$ and these inequalities may be strict, even if there is no isotrivial factor (Paragraph 4.1.2). On the other hand, these ranks are insensitive to dominant base change, and depend only on the isogeny class of A/S (1.6). In particular, one may assume that A/S is principally polarized and (replacing S by an etale covering) admits a level $n \geq 3$ structure.

We prove that r and r'' are unchanged if one passes to the "modular case", *i.e.* if one replaces S by the smallest weakly special (= totally geodesic) subvariety of the moduli space $\mathcal{A}_{g,n}$ containing the image of S, and A by the universal Abelian scheme on S (3.1).

We prove that $r = r'$ in the "modular case", *i.e.* when S is a weakly special subvariety of $\mathcal{A}_{g,n}$ (3.2).

We then study the "PEM case", *i.e.* the case when the connected algebraic monodromy group is maximal with respect to the polarization and the endomorphisms, and emphasize the "restricted PEM case", *i.e.* where we moreover assume that if the center F of End $A \otimes \mathbb{Q}$ is a CM field, then Ω_A is a free $F \otimes_{\mathbb{Q}} \mathcal{O}_S$-module (4.1, 4.3); this includes, of course, the case when the algebraic monodromy group is Sp_{2g}.

Building on the previous results, we show that one has $r'' = r' = r = g$ in the restricted PEM case (4.4). If moreover S is a curve, we show that the derivative (with respect to a parameter) of a non zero Abelian integral of the first kind is never of the first kind (4.6).

Our methods are inspired by B. Moonen's paper [18]; we exploit the "bi-algebraic" properties of the Kodaira-Spencer map in the guise of a theorem of "logarithmic Ax-Schanuel type" for tangent vector bundles (2.2).

Since the problems under study occur in various parts of algebraic geometry and diophantine geometry, we have tried to make the results more accessible by including extended reminders: Section 1 about algebraic De Rham cohomology of Abelian schemes, Gauss-Manin connections and Kodaira-Spencer maps; Subsections 3.1 to 3.4 about weakly special subvarieties of connected Shimura varieties, relative period torsors, and automorphic bundles.

ACKNOWLEDGEMENTS. I thank D. Bertrand and B. Moonen for their careful reading.

1. Preliminaries

1.1. Invariant differential forms

Let S be a smooth connected scheme over a field k of characteristic zero.

Let $f : G \to S$ be a smooth commutative group scheme; we denote by $m : G \times_S G \to G$ the group law and by $e : S \to G$ the unit section. The *invariant differential* 1*-forms* on G are those satisfying $m^*\omega = p_1^*\omega + p_2^*\omega$ (where p_1, p_2

denote the projections); they form a locally free $\mathcal{O}_S$-module denoted by Ω_G, naturally isomorphic to $e^*\Omega^1_{G/S}$ and to $f_*\Omega^1_{G/S}$, and $\mathcal{O}_S$-dual to the Lie algebra $\operatorname{Lie} G$. One has $f^*\Omega_G \cong \Omega^1_{G/S}$. Moreover, invariant differential 1-forms are closed [19, 3.5] [9, 1.2.1].

Let us consider the special case when $G = A$ is an Abelian scheme of relative dimension g, or $G = A^\natural$ universal vectorial extension of A (Rosenlicht-Barsotti, *cf. e.g.*, [16, I]). Recall that $\operatorname{Ext}(A, \mathbb{G}_a) \cong R^1 f_*\mathcal{O}_A$ (using the fact that any rigidified $\mathbb{G}_a$-torsor over an S-Abelian scheme has a canonical S-group structure), so that $A^\natural$ is an extension of A by the vector group attached to the dual of $R^1 f_*\mathcal{O}_A$, which is a locally free $\mathcal{O}_S$-module of rank g. The projection $A^\natural \to A$ gives rise to an exact sequence of locally free $\mathcal{O}_S$-modules

$$0 \to \Omega_A \to \Omega_{A^\natural} \to R^1 f_*\mathcal{O}_A \to 0, \tag{1.1}$$

in a way compatible with base change $S' \to S$. On the other hand, if $A^t := Pic^0(A)$ denotes the dual Abelian scheme, Ω_{A^t} is naturally dual to $R^1 f_*\mathcal{O}_A$ (Cartier), and $\Omega_{A^{t\natural}}$ is naturally dual to $\Omega_{A^\natural}$ in such a way that the exact sequence (1.1) is dual to corresponding exact sequence for A^t [9, 1.1.1].

1.2. Algebraic De Rham cohomology

The *first algebraic De Rham cohomology* $\mathcal{O}_S$-module $\mathcal{H}^1_{dR}(G/S)$ is the hypercohomology sheaf $\mathbf{R}^1 f_*(\Omega^*_{G/S}, d)$. Assuming S affine, it can be computed *à la* Čech using an affine open cover $\mathcal{U}$ of G and taking as coboundary map on $C^p(\mathcal{U}, \Omega^q_{G/S})$ the sum of the Čech coboundary and $(-)^{p+1}$ times the exterior derivative d. In particular, since invariant differential forms are closed, there is a canonical $\mathcal{O}_S$-linear map $\Omega_G \to \mathcal{H}^1_{dR}(G/S)$.

If $G = A$ is an Abelian scheme, and $A^\natural$ its universal vectorial extension, it turns out that the canonical morphisms

$$\Omega_{A^\natural} \to \mathcal{H}^1_{dR}\big(A^\natural/S\big) \leftarrow \mathcal{H}^1_{dR}(A/S)$$

are isomorphisms [9, 1.2.2]. The exact sequence (1.1) thus gives rise to an exact sequence of locally free $\mathcal{O}_S$-modules

$$0 \to \Omega_A = f_*\Omega^1_{A/S} \to \mathcal{H}^1_{dR}(A/S) \to R^1 f_*\mathcal{O}_A = \Omega^\vee_{A^t} \to 0, \tag{1.2}$$

in a way compatible with base change $S' \to S$ and with duality $A \mapsto A^t$ (*cf.* also [14, 8.0]; $f_*\Omega^1_{A/S}$ and $R^1 f_*\mathcal{O}_A$ are the graded pieces gr^1 and gr^0 of the Hodge filtration of $\mathcal{H}^1_{dR}(A/S)$ respectively).

Any polarization of A endows the rank $2g$ vector bundle $\mathcal{H}^1_{dR}(A/S)$ with a symplectic form, for which Ω_A is a Lagrangian[1] subbundle, and the exact sequence (1.2) becomes autodual.

[1] *I.e.* isotropic of rank g.

When $S = \operatorname{Spec} k$, $\mathcal{H}^1_{dR}(A/S)$ can also be interpreted as the space of differentials of the second kind (*i.e.* closed rational 1-forms which are Zariski-locally sums of a regular 1-form and an exact rational form) modulo exact rational 1-forms. For any rational section τ of $A^\natural \to A$ and any $\eta \in \Omega_{A^\natural}$, $\tau^*\eta$ is of the second kind and depends on τ only up to the addition of an exact rational 1-form.

In the sequel, we abbreviate $\mathcal{H}^1_{dR}(A/S)$ by $\mathcal{H}$.

1.3. Gauss-Manin connection

Since the nineteenth century, it has been known that differentiating Abelian integrals with respect to parameters leads to linear differential equations, the prototype being the Gauss hypergeometric equation in the variable t satisfied by $\int_1^\infty z^{a-c}(1-z)^{c-b-1}(1-tz)^{-a}dz$. Manin gave an algebraic construction of this differential module (in terms of differentials of the second kind), later generalized by Katz-Oda and others to the construction of the Gauss-Manin connection on algebraic De Rham cohomology of any smooth morphism $X \to S$.

Let as before $A \xrightarrow{f} S$ be an Abelian scheme of relative dimension g over a smooth connected k-scheme S. If $k = \mathbb{C}$, the *Gauss-Manin connection* is determined by its analytification ∇^{an}, whose dual is the unique analytic connection on $(\mathcal{H}^\vee)^{an}$ which kills the period lattice

$$\ker \exp_A \cong \ker \exp_{A^\natural} \subset \left(\operatorname{Lie} A^\natural\right)^{an} = \left(\Omega^\vee_{A^\natural}\right)^{an} = (\mathcal{H}^\vee)^{an}. \tag{1.3}$$

The formation of $(\mathcal{H}, \nabla)$ is compatible with base change $S' \to S$ and with duality $A \mapsto A^t$. It is contravariant in A, and S-isogenies lead to isomorphisms between Gauss-Manin connections.

If S is affine and Ω_A and Ω_{A^t} are free, let us take a basis $\omega_1, \ldots, \omega_g$ of Ω_A and complete it into a basis $\omega_1, \ldots, \omega_g, \eta_1, \ldots, \eta_g$ of $\mathcal{H}$. Pairing with a basis $\gamma_1, \ldots, \gamma_{2g}$ of the period lattice on a universal covering $\tilde{S}$ of S^{an}, one gets a full solution matrix

$$Y = \begin{pmatrix} \Omega_2 & \mathrm{N}_2 \\ \Omega_1 & \mathrm{N}_1 \end{pmatrix} \in M_{2g}(\mathcal{O}(\tilde{S})) \tag{1.4}$$

for ∇ (with $(\Omega_1)_{ij} = \int_{\gamma_i} \omega_j$, etc...). This reflects into a family of differential equations[2]

$$\partial Y = Y \begin{pmatrix} R_\partial & S_\partial \\ T_\partial & U_\partial \end{pmatrix}, \tag{1.5}$$

where $R_\partial, S_\partial, T_\partial, U_\partial \in M_g(\mathcal{O}(S))$[3] depend $\mathcal{O}(S)$-linearly on the derivation $\partial \in \Gamma T_S$.

[2] We write the matrix of ∇_∂ on the right in order to let the monodromy act on the left on Y. This convention has many advantages. In particular, it is independent of the choice of $\gamma_1, \ldots, \gamma_{2g}$. Writing Y with the indices 2 above the indices 1 will be justified in Subsection 2.2.2 below.

[3] The fact that these matrices have entries in $\mathcal{O}(S)$ rather than $\mathcal{O}(S^{an})$ reflects the algebraic nature of the Gauss-Manin connection. Alternatively, it can be deduced from the next sentence.

It is well-known that the Gauss-Manin connection is regular at infinity (*cf.*, *e.g.*, [14, 14.1]), hence its $\mathcal{D}$-module theoretic properties are faithfully reflected by monodromy theoretic properties.

Remark 1.1. The Katz-Oda algebraic construction of ∇, in the case of $\mathcal{H}^1_{dR}(A/S)$, goes as follows [15, 1.4]. From the exact sequence

$$0 \to f^*\Omega^1_{S/k} \to \Omega^1_{A/k} \to \Omega^1_{A/S} \to 0, \tag{1.6}$$

passing to exterior powers, one gets the exact sequence of k-linear complexes of $\mathcal{O}_A$-modules

$$0 \to f^*\Omega^1_{S/k} \otimes \Omega^{*-1}_{A/S} \to \Omega^*_{A/k}/\left(f^*\Omega^2_{S/k} \otimes \Omega^{*-2}_{A/S}\right) \to \Omega^*_{A/S} \to 0. \tag{1.7}$$

Then ∇ is a coboundary map in the long exact sequence for $\mathbf{R}^* f_*$ applied to (1.7), that is

$$\mathbf{R}^1 f_*\Omega^*_{A/S} \xrightarrow{\nabla} \mathbf{R}^2 f_*\left(f^*\Omega^1_{S/k} \otimes \Omega^{*-1}_{A/S}\right) = \Omega^1_{S/k} \otimes \mathbf{R}^1 f_*\Omega^*_{A/S}, \tag{1.8}$$

and can be computed explicitly *à la* Čech, *cf.* [14, 3.4]. One checks that this map satisfies the Leibniz rule and the associated map

$$T_S = \left(\Omega^1_{S/k}\right)^\vee \xrightarrow{\partial \mapsto \nabla_\partial} \mathrm{End}_k\, \mathcal{H} \tag{1.9}$$

respects Lie brackets, so that ∇ corresponds to a $\mathcal{D}_S$-module structure on $\mathcal{H}$ (here $\mathcal{D}_S$ denotes the sheaf of rings of differential operators on S, which is generated by the tangent bundle T_S). In fact, it can also be interpreted as the first higher direct image of $\mathcal{O}_A$ in the $\mathcal{D}$-module setting (*cf. e.g.*, [6, 4] for an algebraic proof).

An alternative and more precise construction of ∇, which avoids homological algebra, consists in endowing $A^\natural$ with the structure of a commutative algebraic $\mathcal{D}$-group, which automatically provides a connection on (the dual of) its Lie algebra [3, 3.4, H5] [5, 6].

1.4. Kodaira-Spencer map

The Gauss-Manin connection does not preserve the subbundle $\Omega_A \subset \mathcal{H}$ in general. The composed map

$$\Omega_A \hookrightarrow \mathcal{H} \xrightarrow{\nabla} \Omega^1_S \otimes \mathcal{H} \twoheadrightarrow \Omega^1_S \otimes (\mathcal{H}/\Omega_A) = \Omega^1_S \otimes \Omega^\vee_{A^t} \tag{1.10}$$

is the *Kodaira-Spencer map* (or Higgs field). Like the Gauss-Manin connection, its formation commutes with base-change. Unlike the Gauss-Manin connection, it is an $\mathcal{O}_S$-linear map (also called the Higgs field of A/S [22]).

Remark 1.2. This map can be interpreted as a coboundary map in the long exact sequence for $R^* f_*$ applied to (1.6), and computed explicitly *à la* Čech, *cf.* [14, 3.4] [15, 1.3].

It can be rewritten as the map

$$\theta : \ T_S \otimes_{\mathcal{O}_S} \Omega_A \to \Omega^\vee_{A^t} = \mathrm{Lie}\, A^t. \tag{1.11}$$

If $\mathcal{D}_S^{\leq 1} \subset \mathcal{D}_S$ denotes the subsheaf of differential operators of order ≤ 1 on S, and $\mathcal{D}_S \Omega_A \subset \mathcal{H}$ the sub-$\mathcal{D}_S$-module generated by Ω_A in $\mathcal{H} = \mathcal{H}^1_{dR}(A/S)$. One has

$$\mathrm{Im}\, \theta = \mathcal{D}_S^{\leq 1} \Omega_A / \Omega_A \subset \mathcal{D}_S \Omega_A / \Omega_A \subset \mathcal{H}/\Omega_A = \mathrm{Lie}\, A^t. \tag{1.12}$$

The Kodaira-Spencer map can also be rewritten as the map

$$T_S \overset{\partial \mapsto \theta_\partial}{\to} \mathrm{Lie}\, A \otimes \mathrm{Lie}\, A^t, \tag{1.13}$$

which is invariant by duality $A \mapsto A^t$ [7, 9.1]; if A is polarized, it thus gives rise to a map

$$T_S \overset{\partial \mapsto \theta_\partial}{\to} S^2 \mathrm{Lie}\, A \cong \mathrm{Hom}_{\mathrm{sym}}\big(\Omega_A, \Omega_A^\vee\big). \tag{1.14}$$

In the situation and notation of the end of Subsection 1.3, the matrix of θ_∂ is T_∂ (which is a symmetric matrix if one chooses the basis $\omega_1, \dots, \eta_g$ to be symplectic).

Remarks 1.3.

i) Here is another interpretation of θ_∂ in terms of the universal vectorial extension $A^\natural$, assuming S affine [7, 9]: for any $\omega \in \Gamma \Omega_A$, pull-back the exact sequence of vector bundles associated to (1.6) by the morphism $\mathcal{O}_A \to \Omega^1_{A/S}$ corresponding to ω and get an extension of A by the vector group attached to Ω^1_S, so that the morphism from $A^\natural$ to this vectorial extension gives rise, at the level of invariant differential forms, to a morphism $\Omega_{A^t} \to \Omega^1_S$; thus to any ω and any $\partial \in \Gamma T_S = \mathrm{Hom}(\Omega^1_S, \mathcal{O}_S)$, one gets an element of $\Omega^\vee_{A^t}$, which is nothing but $\theta_\partial \cdot \omega$;

ii) The following equivalences are well-known:

 - A/S is isotrivial $\Leftrightarrow \theta = 0 \Leftrightarrow \mathcal{D}_S \Omega_A = \Omega_A \Leftrightarrow \nabla$ is isotrivial (*i.e.* has finite monodromy).

 Remembering that the Kodaira-Spencer map commutes to base-change, the only non trivial implications are: ∇ isotrivial $\Rightarrow \mathcal{D}_S \Omega_A = \Omega_A$, and $\theta = 0 \Rightarrow A/S$ isotrivial. The first implication comes from Deligne's "théorème de la partie fixe" [10, 4.1.2]. An elementary proof of the second one will be given below (Paragraph 2.1.1);

iii) In contrast to $\mathcal{D}_S \Omega_A$, $\mathcal{D}_S^{\leq 1} \Omega_A$ is not locally a direct factor of $\mathcal{H}$ in general: at some points $s \in S$ the rank of θ_s may drop (see however Theorem 3.2). In fact, the condition that the rank of θ_s is constant is very restrictive: for instance, if S is a proper curve, the condition that θ is everywhere an isomorphism is equivalent to the condition that the Arakelov inequality $\deg \Omega_A \leq \frac{g}{2}\deg \Omega^1_S$ is an equality, and implies that A/S is a modular family, parametrized by a Shimura curve [22].

1.4.1. Since the $\mathcal{O}_S$-module $\mathcal{H}/\mathcal{D}_S\Omega_A$ carries a $\mathcal{D}_S$-module structure, it is locally free [14, 8.8], hence $\mathcal{D}_S\Omega_A$ is locally a direct summand of $\mathcal{H}$. In fact, by Deligne's semisimplicity theorem [10, 4.2.6], $\mathcal{D}_S\Omega_A$ is even a direct factor of $\mathcal{H}$ (as a $\mathcal{D}_S$-module, hence as a vector bundle).

Lemma 1.4. *The formation of $\mathcal{D}_S\Omega_A$ commutes with dominant base change $S' \xrightarrow{\pi} S$ (with S' smooth connected).*

Proof. Since $\mathcal{H}$ commutes with base-change and $\mathcal{D}_S\Omega_A$ is locally a direct summand, it suffices to prove the statement after restricting S to a dense affine open subset. In particular, one may assume that π is a flat submersion, so that $T_{S'} \to \pi^*T_S$ and $\mathcal{D}_{S'} \to \pi^*\mathcal{D}_S$ are epimorphisms, and $\mathcal{D}_{S'}\Omega_{A_{S'}} = \pi^*\mathcal{D}_S\pi^*\Omega_A = \pi^*(\mathcal{D}_S\Omega_A)$. □

1.4.2. As in the introduction, let us define

$$r = r(A/S) := \operatorname{rk} \mathcal{D}_S\Omega_A/\Omega_A, \tag{1.15}$$

$$r' = r'(A/S) := \operatorname{rk} \theta = \operatorname{rk} \mathcal{D}_S^{\leq 1}\Omega_A/\Omega_A, \tag{1.16}$$

$$r'' = r''(A/S) := \max_{\partial} \operatorname{rk} \theta_\partial, \tag{1.17}$$

where ∂ runs over local tangent vector fields on S (and rk denotes a generic rank).

Lemma 1.5. *These are invariant by dominant base change $S' \xrightarrow{\pi} S$ (with S' smooth connected), and depend only on the isogeny class of A/S.*

Proof. For r, this follows from the previous lemma. Its proof also shows that $\mathcal{D}_S^{\leq 1}\Omega_A$ commutes with base change by flat submersions, which settles the case of r'. For r'', we may assume that S and S' are affine, that T_S is free and $T_{S'} = \pi^*T_S$, and pick a basis $\partial_1, \dots, \partial_d$ of tangent vector fields; the point is that $\max_{\lambda_i} \operatorname{rk} \sum \lambda_i \theta_{\partial_i}$ is the same when the λ_i's run in $\mathcal{O}(S)$ or in $\mathcal{O}(S')$ (consider the θ_{∂_i}'s as matrices and note that each minor determinant is a polynomial in the λ_i's).

The second assertion is clear since any isogeny induces an isomorphism at the level of $(\mathcal{H}, \nabla)$. □

Lemma 1.6.

(1) *$r'' = g$ holds if and only if there exists a local vector field ∂ such that $\theta_\partial.\omega \neq 0$ for every non-zero $\omega \in \Gamma\Omega_A$;*
(2) *$r' = g$ holds if and only if for every non-zero $\omega \in \Gamma\Omega_A$, there exists a local vector field ∂ such that $\theta_\partial.\omega \neq 0$.*

Proof. The first equivalence is immediate, while the second uses the symmetry of (1.13): assuming A polarized, and after restricting S to a dense open affine subset, one has $r' = g \Leftrightarrow \forall \omega \in \Gamma\Omega_A \setminus 0,\ \exists \eta \in \Gamma\Omega_A, \exists \partial \in \Gamma T_S,\ (\theta_\partial.\omega)\cdot\eta \neq 0$. Since $(\theta_\partial.\eta)\cdot\omega = (\theta_\partial.\omega)\cdot\eta$, one gets $\forall \omega \in \Gamma\Omega_A \setminus 0,\ \exists \eta \in \Gamma\Omega_A, \exists \partial \in \Gamma T_S,\ (\theta_\partial.\eta)\cdot\omega \neq 0$. □

2. Automorphic vector bundles and bi-algebraicity

2.1. Bi-algebraicity of the Kodaira-Spencer map

2.1.1. Let $\mathcal{A}_{g,n}$ be the moduli scheme of principally polarized Abelian varieties of dimension g with level n structure ($n \geq 3$), and let $\mathcal{X} \to \mathcal{A}_{g,n}$ be the universal Abelian scheme.

The universal covering of $\mathcal{A}_{g,n}^{an}$ is the Siegel upper half space $\mathfrak{H}_g$. We denote by $j_{g,n} : \mathfrak{H}_g \to \mathcal{A}_{g,n}^{an}$ the uniformizing map (for $g = n = 1$, this is the usual j-function). The pull-back of the dual of the period lattice $\ker \exp_{\mathcal{X}}$ on $\mathfrak{H}_g$ is a constant symplectic lattice Λ. On $\mathfrak{H}_g$, the Gauss-Manin connection of $\mathcal{X}/\mathcal{A}_{g,n}$ becomes a trivial connection with solution space $\Lambda_{\mathbb{C}}$.

On the other hand, $\mathfrak{H}_g$ is an (analytic) open subset of its "compact dual" $\mathfrak{H}_g^\vee$, which is the Grassmannian of Lagrangian subspaces $V \subset \Lambda_{\mathbb{C}}^\vee$ (*i.e.* isotropic subspaces of dimension g): the Lagrangian subspace V_τ corresponding to a point $\tau \in \mathfrak{H}_g$ is $\Omega_{\mathcal{X}_{j_{g,n}(\tau)}} \subset \mathcal{H}^1_{dR}(\mathcal{X}_{j_{g,n}(\tau)}) \cong \Lambda_{\mathbb{C}}^\vee$ (note that the latter isomorphism depends on τ, not only on $j_{g,n}(\tau)$). The Grassmannian $\mathfrak{H}_g^\vee$ is a homogeneous space for $Sp(\Lambda_{\mathbb{C}})$ (in block form $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ sends $\tau \in \mathfrak{H}_g$ to $(A\tau + B)(C\tau + D)^{-1}$). The vector bundle $j_{g,n}^*\mathrm{Lie}\mathcal{X}$ is the restriction to $\mathfrak{H}_g$ of the tautological vector bundle $\mathcal{L}$ on the Lagrangian Grassmannian $\mathfrak{H}_g^\vee$.

2.1.2. In this universal situation, the Kodaira-Spencer map (in the form of (1.14)) is an isomorphism

$$T_{\mathcal{A}_{g,n}} \xrightarrow{\sim} S^2\mathrm{Lie}\mathcal{X}, \tag{2.1}$$

and its pull-back to $\mathfrak{H}_g$ is the restriction of the canonical isomorphism

$$T_{\mathfrak{H}_g^\vee} \xrightarrow{\sim} S^2\mathcal{L} \tag{2.2}$$

cf. e.g., [8,12].

Any principally polarized Abelian scheme with level n structure A/S is isomorphic to the pull-back of $\mathcal{X}$ by a morphism $S \xrightarrow{\mu} \mathcal{A}_{g,n}$, and the Kodaira-Spencer map of A/S (in the form of (1.14)) is the pull-back by μ of the isomorphism (2.1) composed with $d\mu : T_S \to \mu^*T_{\mathcal{A}_{g,n}}$. In particular, the Kodaira-Spencer map θ of A/S vanishes if and only if the image of $S \to \mathcal{A}_{g,n}$ is a point, *i.e.* A/S is constant; moreover, if A/S is not constant, μ is generically finite, and ∂ is a non zero section of T_S, then θ_∂ is non zero.

2.2. Relative period torsor

2.2.1. The *bi-algebraicity* mentioned above refers to the pair of algebraic structures $\mathcal{A}_{g,n}$, $\mathfrak{H}_g^\vee$, which are transcendentally related via $\mathfrak{H}_g$ and $j_{g,n}$.

On the other hand, there is a purely algebraic relation between these two algebraic structures, through the *relative period torsor*. This is the $Sp(\Lambda_{\mathbb{C}})_{\mathcal{A}_{g,n}}$-torsor $\Pi_{g,n} \overset{\pi}{\to} \mathcal{A}_{g,n}$ of solutions of the Gauss-Manin connection ∇ of $\mathcal{X}$. More formally, this is the torsor of isomorphisms $\mathcal{H} \to \Lambda_{\mathbb{C}}^\vee \otimes \mathcal{O}_{\mathcal{A}_{g,n}}$ which respect the ∇-horizontal tensors[4]. Its generic fiber is the spectrum of the Picard-Vessiot algebra[5] attached to ∇, namely Spec $\mathbb{C}(\mathcal{A}_{g,n})[Y_{ij}]_{i,j=1,\dots,2g}$ (with the notation of Subsection 1.3).

2.2.2. The canonical horizontal isomorphism $\mathcal{H} \otimes_{\mathcal{O}_{\mathcal{A}_{g,n}}} \mathcal{O}_{\mathfrak{H}_g} \overset{\sim}{\to} \Lambda_{\mathbb{C}}^\vee \otimes_{\mathbb{C}} \mathcal{O}_{\mathfrak{H}_g}$ gives rise to an analytic map

$$k : \mathfrak{H}_g \to \Pi_{g,n} \tag{2.3}$$

with $\pi \circ k = j_{g,n}$. In local bases and with the notation of Subsection 1.3, k sends $\tau \in \mathfrak{H}_g$ to the point $Y(\tau) = \begin{pmatrix} \Omega_2(\tau) & \mathrm{N}_2(\tau) \\ \Omega_1(\tau) & \mathrm{N}_1(\tau) \end{pmatrix}$ of $\Pi_{j_{g,n}(\tau)}$. In particular, the image of k is Zariski-dense in $\Pi_{g,n}$.

On the other hand there is an algebraic $Sp(\Lambda_{\mathbb{C}})$-equivariant map

$$\rho : \Pi_{g,n} \to \mathfrak{H}_g^\vee, \tag{2.4}$$

which sends a point $p \in \Pi_{g,n}(\mathbb{C})$ viewed as an isomorphism $\mathcal{H}_{\pi(p)} \to \Lambda^\vee$ to the image of $\Omega_{\mathcal{X}_{\pi(p)}}$ in $\Lambda_{\mathbb{C}}^\vee$. In local bases and with the notation of Subsection 1.3, ρ sends $\begin{pmatrix} \Omega_2 & \mathrm{N}_2 \\ \Omega_1 & \mathrm{N}_1 \end{pmatrix}$ to $\tau = \Omega_2 \Omega_1^{-1}$; $\rho \circ k$ is the Borel embedding $\mathfrak{H}_g \hookrightarrow \mathfrak{H}_g^\vee$.

One thus has the following diagram

$$\mathfrak{H}_g \to \Pi_{g,n} \overset{(\pi,\rho)}{\twoheadrightarrow} \mathcal{A}_{g,n} \times \mathfrak{H}_g^\vee, \tag{2.5}$$

in which the first map has Zariski-dense image, and the second map (π, ρ) is surjective (of relative dimension $\frac{g(3g+1)}{2}$) since the restriction of ρ to any fiber of π is $Sp(\Lambda_{\mathbb{C}})$-equivariant and $\mathfrak{H}_g^\vee$ is homogeneous. It follows that the graph of $j_{g,n}$ is Zariski-dense[6] in $\mathfrak{H}_g^\vee \times \mathcal{A}_{g,n}$.

The function field of $\Pi_{g,n}$ is studied in detail in [4]: it is a differential field both for the derivations of $\mathcal{A}_{g,n}$ and for the derivations $\partial/\partial\tau_{ij}$ of $\mathfrak{H}_g^\vee$. Over $\mathbb{C}(\mathfrak{H}_g^\vee) = \mathbb{C}(\tau_{ij})_{i \le j \le g}$, it is generated by (iterated) derivatives with respect to the $\partial/\partial\tau_{ij}$'s of the modular functions (the field of modular functions being $\mathbb{C}(\mathcal{A}_{g,n})$).

[4] This is a reduction from GSp to Sp of the standard principal bundle considered in [17, III.3].

[5] In general, one has to adjoin the inverse of the wronskian together with the entries of a full solution matrix in order to build a Picard-Vessiot algebra, but here the wronskian is a rational function, since the monodromy is contained in Sp.

[6] This property does not extend to m-jets for $m \geq 3$.

2.3. Connected Shimura varieties and weakly special subvarieties

2.3.1. Let G be a reductive group over $\mathbb{Q}$, G^{ad} the quotient by the center, and $G^{ad}(\mathbb{R})^+$ the connected component of identity of the Lie group $G^{ad}(\mathbb{R})$.

Let X be a connected component of a conjugacy class $\mathfrak{X}$ of real-algebraic homomorphisms $\mathbb{C}^* \to G_{\mathbb{R}}$. For any rational representation $G \to GL(W)$, one then has a collection of real Hodge structures $(W_{\mathbb{R}}, h_x)_{x\in X}$ on $W_{\mathbb{R}}$ parametrized by X. If the weight is defined over $\mathbb{Q}$ (which is the case if $G = G^{ad}$ since the weight is 0 in this case), one even has a collection of rational Hodge structures $(W, h_x)_{x\in X}$.

In the sequel, we assume that $(G, \mathfrak{X})$ satisfies Deligne's axioms for a *Shimura datum*; these axioms ensure that X has a $G^{ad}(\mathbb{R})^+$-invariant metric, which makes X into a hermitian symmetric domain, and that the $(W_{\mathbb{R}}, h_x)$ (respectively (W, h_x)) come from variations of polarized Hodge structures on the analytic variety X (respectively if the weight is defined over $\mathbb{Q}$, for instance if $G = G^{ad}$); moreover, in the case of the adjoint representation on $\mathfrak{g} = \mathrm{Lie}\, G$, the variation of Hodge structures is of type $(-1, 1) + (0, 0) + (1, -1)$ (*cf. e.g.*, [17, II]).

2.3.2. Let Γ be a discrete subgroup of $G^{ad}(\mathbb{Q})^+$, quotient of a torsion-free congruence subgroup of $G(\mathbb{Q})$. Then $\Gamma\backslash X$ has a canonical structure of algebraic variety (Baily-Borel): the *connected Shimura variety* attached to (G, X, Γ). The variation of Hodge structures descends to it, with monodromy group Γ. The situation of Subsection 2.1 corresponds to the case $G = GSp_{2g}$, $X = \mathfrak{H}_g$, $\Gamma =$ the congruence subgroup of level $n \geq 3$ (*cf. e.g.*, [17, II]).

2.3.3. Let S be the connected Shimura variety attached to (G, X, Γ), and $j : X \to S$ the uniformizing map. An irreducible subvariety $S_1 \subset S$ is *weakly special* if there is a sub-Shimura datum $(H, \mathfrak{Y}) \to (G, \mathfrak{X})$, a decomposition $(H^{ad}, \mathfrak{Y}^{ad}) = (H_1, \mathfrak{Y}_1) \times (H_2, \mathfrak{Y}_2)$, and a point $y \in \mathfrak{Y}_2$ such that S_1 is the image of $Y_1 \times y$ in S (here Y_1 is a connected component of $\mathfrak{Y}_1$ contained in X) [21][7]; in particular, S_1 is isomorphic to the connected Shimura variety attached to $(H_1, Y_1, \Gamma^{ad} \cap H_1)$.

2.4. Automorphic vector bundles

2.4.1. Given a faithful rational representation W of G, the associated family of Hodge filtrations on $W_{\mathbb{C}}$ is parametrized by a certain flag variety $X^\vee$, the *compact dual* of X, which is a $G^{ad}_{\mathbb{C}}$-homogeneous space.

The isotropy group of a point $x \in X^\vee$ is a parabolic subgroup P_x, $K_x := P_x \cap G^{ad}(\mathbb{R})^+$ is a maximal compact subgroup, and there is a $G^{ad}(\mathbb{R})^+$-equivariant *Borel embedding*

$$X = G^{ad}(\mathbb{R})^+/K_x \overset{i}{\hookrightarrow} X^\vee = G^{ad}_{\mathbb{C}}/P_x. \tag{2.6}$$

[7] This is a *special* subvariety if y is a special point.

2.4.2. Associated to W, there is a variation of polarized Hodge structures on $S = \Gamma\backslash X$, hence an integrable connection ∇ with regular singularities at infinity on the underlying vector bundle $\mathcal{W}$. There is again a relative period torsor in this situation.

Assume for simplicity that $G = G^{ad}$. The monodromy group Γ is then Zariski-dense in G. The *relative period torsor*

$$\Pi \overset{\pi}{\to} S \tag{2.7}$$

is the G_S-torsor of isomorphisms $\mathcal{W} \to W_C \otimes \mathcal{O}_S$ which respects the ∇-horizontal tensors[8]. Its generic fiber is the Picard-Vessiot algebra attached to ∇.

The canonical horizontal isomorphism $\mathcal{W} \otimes_{\mathcal{O}_S} \mathcal{O}_X \overset{\sim}{\to} W_{\mathbb{C}} \otimes \mathcal{O}_X$ gives rise to an analytic map $k : X \to \Pi$ with $\pi \circ k = j$. There is an algebraic $G_{\mathbb{C}}$-equivariant map $\Pi \overset{\rho}{\to} X^\vee$ (which sends a point $p \in \Pi(\mathbb{C})$ viewed as an isomorphism $\mathcal{W}_{\pi(p)} \to W$ to the point of $X^\vee$ which parametrizes the image of the Hodge filtration of $\mathcal{W}_{\pi(p)}$); one has $\rho \circ k = i$.

One thus has the following factorization:

$$(j, i) : \; X \to \Pi \overset{(\pi,\rho)}{\twoheadrightarrow} S \times X^\vee \tag{2.8}$$

in which the first map has dense image, the second map (π, ρ) is surjective (since the restriction of ρ to any fiber of π is $G_{\mathbb{C}}$-equivariant with homogeneous target).

Since any faithful rational representation of G lies in the tannakian category generated by W and conversely, neither $X^\vee$ nor Π depend on the auxiliary W. On the other hand, π^* provides an equivalence between the category of vector bundles on S and the category of $G_{\mathbb{C}}$-vector bundles on Π [17, III.3.1].

2.4.3. A $G_{\mathbb{C}}$-equivariant vector bundle $\check{\mathcal{V}}$ on $X^\vee = G_{\mathbb{C}}/P_x$ is completely determined by its fiber at $x \in X$ together with the induced P_x-action (or else, the induced K_x-action). The quotient $\mathcal{V} := \Gamma\backslash i^*\check{\mathcal{V}}$ has a canonical structure of algebraic vector bundle on $S = \Gamma\backslash X$, the *automorphic vector bundle* attached to $\check{\mathcal{V}}$ [17, III.2.1, 3.6]. One has the equality of analytic vector bundles on X:

$$j^*\mathcal{V} = i^*\check{\mathcal{V}}. \tag{2.9}$$

There is also a purely algebraic relation between $\mathcal{V}$ and $\check{\mathcal{V}}$, through the relative period torsor [17, III.3.5]: one has the equality of algebraic $G_{\mathbb{C}}$-vector bundles on Π:

$$\pi^*\mathcal{V} = \rho^*\check{\mathcal{V}}. \tag{2.10}$$

[8] It coincides with the standard principal bundle considered in [17, III.3].

2.4.4. Any representation of $G_{\mathbb{C}}$ gives rise to a $G_{\mathbb{C}}$-equivariant vector bundle on $X^{\vee}$, hence to an automorphic vector bundle (which carries an integrable connection).

On the other hand, $T_{X^{\vee}}$ is a $G_{\mathbb{C}}$-equivariant vector bundle on $X^{\vee}$, and the corresponding automorphic vector bundle is nothing but T_S.

In the situation of Paragraph 2.1.2, the tangent bundle $T_{\mathfrak{H}_g^{\vee}}$ and its tautological bundle $\mathcal{L}$ are equivariant vector bundles, and the universal Kodaira-Spencer map (2.1) is an isomorphism of automorphic vector bundles on $\mathcal{A}_{g,n}$.

2.5. A theorem of logarithmic Ax-Schanuel type for tangent bundles

2.5.1. The theorem of logarithmic Ax-Schanuel type for connected Shimura varieties is the following [11, 2.3.1] (*cf.* also [21])[9]:

Theorem 2.1. *Let S be a connected Shimura variety ($S^{an} = \Gamma\backslash X$). Let $Z \subset S$ be an irreducible locally closed subset, and let $\tilde{Z}$ be an analytic component of the inverse image of Z in X.*

Then the image in S of the intersection with X of the Zariski closure of $\tilde{Z}$ in the compact dual $X^{\vee}$ is the smallest weakly special subvariety $S_1 \subset S$ containing Z.

Here is a sketch of proof. One can replace S by the smallest special subvariety containing Z. Fix a point $s \in Z(\mathbb{C})$ and a faithful rational representation of G, and consider the associated vector bundle $\mathcal{W}$ with integrable connection ∇ on S. Let $\hat{G}_1 \subset G$ be the Zariski closure of the monodromy group Γ_Z of $(\mathcal{W}_{|Z}, \nabla_{|Z})$ at s. Up to replacing Γ by a subgroup of finite index, $\hat{G}_1$ is connected and a normal subgroup of G (by [2, 5]). This gives rise to a weakly special subvariety $S_1 \subset S$ associated to a factor $G_1 = \hat{G}_1^{ad}$ of G^{ad}, which is in fact the smallest weakly special subvariety of S containing Z (*cf.* [18, 3.6], [21, 4.1] for details). On the other hand, since $\tilde{Z}$ is stable under Γ_Z, its Zariski closure in the $G_{\mathbb{C}}$-homogeneous space $X_1^{\vee}$ is stable under G_1, hence equal to $X_1^{\vee}$. □

Here is the analog for tangent vector bundles, assuming Z smooth:

Theorem 2.2. *In this situation, $T_{\tilde{Z}}$ is Zariski-dense in $T_{X_1^{\vee}}$.*

Proof. We may replace G by G_1 and S by S_1. Let $\overline{T_{\tilde{Z}}}$ be the Zariski closure of $T_{\tilde{Z}} = \tilde{Z} \times_Z T_Z$ in $T_{X^{\vee}}$. Let $(\mathcal{W}, \nabla)$ be as above, and let Π be the relative period torsor of S (we take over the notation (2.3) (2.5)). Since $(\mathcal{W}_{|Z}, \nabla_{|Z})$ has the same algebraic monodromy group as $(\mathcal{W}, \nabla)$, namely G, the generic fiber of the projection $\Pi_Z \overset{\pi_Z}{\to} Z$ is the spectrum of the Picard-Vessiot algebra attached

[9] Not to be confused with the (exponential) Ax-Schanuel theorem - a.k.a. hyperbolic Ax-Lindemann - for connected Shimura varieties, which concerns the maximal irreducible algebraic subvarieties of $X^{\vee}$ whose intersection with X is contained in $\tilde{Z}$, and which is a much deeper result (Pila-Tsimerman, Klingler-Ullmo-Yafaev).

to $(\mathcal{W}_{|Z}, \nabla_{|Z})$, and the image of $k_{|\tilde{Z}}(\tilde{Z})$ is Zariski-dense in Π_Z. It follows that $(k_{|\tilde{Z}} \times 1_{T_Z})(T_{\tilde{Z}}) = k_{|\tilde{Z}}(\tilde{Z}) \times_Z T_Z$ is Zariski-dense in $\Pi_Z \times_Z T_Z$.

In fact, $\nabla_{|Z}$ induces a connection on the torsor Π_Z, which amounts to a splitting of the natural exact sequence of $G_{\mathbb{C}}$-equivariant vector bundles on Π_Z:

$$T_{\Pi_Z/Z} \to T_{\Pi_Z} \overset{\leftharpoondown}{\to} \Pi_Z \times_Z T_Z,$$

and since $k_{|\tilde{Z}}(\tilde{Z})$ is horizontal, the Zariski-closure of $k_{|\tilde{Z}*}(T_{\tilde{Z}})$ in T_{Π_Z} is the $G_{\mathbb{C}}$-equivariant vector subbundle $\Pi_Z \times_Z T_Z$.

On the other hand, $T_{\tilde{Z}} \to T_{X^\vee}$ factors through the map $T_{\Pi_Z} \to T_{X^\vee}$ of $G_{\mathbb{C}}$-equivariant vector bundles induced by ρ, and one concludes that $\overline{T_{\tilde{Z}}}$ is a $G_{\mathbb{C}}$-equivariant vector subbundle of $T_{X^\vee}$. Hence $\overline{T_{\tilde{Z}}} = \breve{\mathcal{V}}$ for some automorphic vector subbundle $\mathcal{V} \subset T_S$.

It is known (see [13, VIII, 5]) that for any irreducible factor of X, the (real) representation of the corresponding factor of $\mathfrak{k}$ on the corresponding factor of $T_x X$ is irreducible, from which it follows that the automorphic vector subbundles of T_S are of the form $S \times_{S_1} T_{S_1}$ for some factor S_1 of the locally symmetric domain S. Since $\mathcal{V}$ contains T_Z and Z is not contained in any proper S_1, one concludes that $\mathcal{V} = T_S$ and $\overline{T_{\tilde{Z}}} = T_{X^\vee}$. □

Remark 2.3. In general, given an algebraic vector bundle $\mathcal{M}$ on an algebraic variety Y, the Zariski closure of an analytic subbundle over some Zariski-dense analytic subspace of Y is not necessarily an algebraic subbundle of $\mathcal{M}$: for instance, the Zariski closure in $T_{\mathbb{C}^2}$ of the tangent bundle of the graph in $\mathbb{C}^2$ of a Weierstrass $\wp$ function is a quadric bundle over $\mathbb{C}^2$, not a vector subbundle of $T_{\mathbb{C}^2}$ (a similar counterexample holds for the graph of the usual j-function and its bundle of jets of order ≤ 3, since j satisfies a rational non-linear differential equation of order 3).

On the other hand, Theorem 2.2 does not extend to arbitrary automorphic vector bundles, but one has the following easy consequence of Theorem 2.1:

Porism 2.4. In the same situation, let $\mathcal{V}$ be an automorphic vector bundle on S with corresponding vector bundle $\breve{\mathcal{V}}$ on $X^\vee$, and let $\mathcal{F}$ be a vector subbundle of the restriction of $\mathcal{V}$ to Z. Then Z and $\mathcal{F}$ are bi-algebraic if and only if Z is a weakly special subvariety and $\mathcal{F}$ is an automorphic vector bundle.

The assumption "Z is bi-algebraic" means that $\tilde{Z}$ is the intersection of X with an algebraic subvariety of $X^\vee$, and according to Theorem 2.1, this amounts to $Z = S_1$.

The assumption "$\mathcal{F}$ is bi-algebraic" means that its pull-back $\tilde{\mathcal{F}}$ in $\breve{\mathcal{V}}$ is an algebraic subvariety. Since $Z = S_1$, this amounts to say that this analytic subbundle of $\mathcal{V} \times_S X_1^\vee = \breve{\mathcal{V}} \times_{X^\vee} X_1^\vee$ is algebraic. It is in fact a $G_{1\mathbb{C}}$-vector subbundle, so that $\mathcal{F}$ is an automorphic vector bundle on $X_1^\vee$.

Using the relative period torsor, one can also prove the following stronger version of Theorem 2.1:

Scholium 2.5. in the setting of Theorem 2.1, the graph of $j_{|\tilde{Z}}$ is Zariski-dense in $X_1^\vee \times Z$.

Indeed, it follows from (2.8) (with S_1 in place of S) that the map $\Pi_{1|Z} \overset{(\rho,\pi_Z)}{\to} X_1^\vee \times Z$ is surjective. On the other hand, the image of $k_{|\tilde{Z}}(\tilde{Z})$ is Zariski-dense in $\Pi_{1|Z}$.

3. Transition to the modular case

We go back to the study of $r(A/S)$, $r'(A/S)$, $r''(A/S)$ for an Abelian scheme A/S. We may assume that the base field k is $\mathbb{C}$. According to Lemma 1.6, these ranks are invariant by dominant base change of S and by isogeny of A, hence one may assume that A/S admits a principal polarization and a Jacobi level n structure for some $n \geq 3$, and then replace S by the smooth locus $Z \subset \mathcal{A}_{g,n}$ of its image in the moduli space of principally polarized Abelian varieties of dimension g with level n structure, and A by the restriction $\mathcal{X}_Z$ of the universal Abelian scheme $\mathcal{X}$ on $\mathcal{A}_{g,n}$.

3.1. From Z to the smallest weakly special subvariety of $\mathcal{A}_{g,n}$ containing Z

Let us consider again the situation of Paragraph 2.5.1, with $S = \mathcal{A}_{g,n}$. Given a (locally closed) subvariety $Z \subset \mathcal{A}_{g,n}$, one constructs the smallest weakly special subvariety $S_1 \subset \mathcal{A}_{g,n}$ containing Z, taking $(\mathcal{W}, \nabla)$ equal to $\mathcal{H}^1_{dR}(\mathcal{X}/\mathcal{A}_{g,n})$ with its Gauss-Manin connection. By construction, $S_1(\mathbb{C}) = \Gamma_1\backslash X_1$ where X_1 is a hermitian symmetric domain attached to the adjoint group G_1 of the connected algebraic monodromy group of $\nabla_{|Z}$.

Theorem 3.1. *One has* $r(\mathcal{X}_Z/Z) = r(\mathcal{X}_{S_1}/S_1)$ *and* $r''(\mathcal{X}_Z/Z) = r''(\mathcal{X}_{S_1}/S_1)$.

Proof. Fix $s \in Z(\mathbb{C})$. By construction $\nabla_{|Z}$ and $\nabla_{|S_1}$ have the same connected algebraic monodromy group at s, namely $\hat{G}_1 \subset Sp_{2g}$ (up to replacing n by a multiple). It follows that $\mathcal{D}_Z\Omega_{\mathcal{X}_Z} = (\mathcal{D}_{S_1}\Omega_{\mathcal{X}_{S_1}})_{|Z}$, whence $r(\mathcal{X}_Z/Z) = r(\mathcal{X}_{S_1}/S_1)$.

On the other hand, the inequality $r''(\mathcal{X}_Z/Z) \leq r''(\mathcal{X}_{S_1}/S_1)$ is obvious. For any natural integer $h < g$, let Δ_h be the closed subset of $T_{\mathfrak{H}_g^\vee}$ corresponding to quadratic forms in $S^2\mathrm{Lie}\mathcal{X}$ of rank $\leq h$ (this is in fact a $Sp(\Lambda_{\mathbb{C}})$-subvariety; Δ_0 is the 0-section). Then $r''(\mathcal{X}_Z/Z)$ (respectively $r''(\mathcal{X}_{S_1}/S_1)$) is the greatest integer h such that $d\mu(\partial)$ in not contained in $\mu^*\Delta_{h-1}$. In order to prove the inequality $r''(\mathcal{X}_Z/Z) \geq r''(\mathcal{X}_{S_1}/S_1)$, it thus suffices to show that if $T_{\tilde{Z}}$ is not contained in Δ_h, neither is $T_{X_1^\vee}$, which follows from the fact that $T_{\tilde{Z}}$ is Zariski-dense in $T_{X_1^\vee}$ (2.2). □

3.2. Case of a weakly special subvariety of $\mathcal{A}_{g,n}$

We now assume that S is a weakly special subvariety of $\mathcal{A}_{g,n}$, with associated group $G = G^{ad}$, and that there is a finite covering $\hat{G}$ of G contained in $Sp(\Lambda_{\mathbb{Q}})$.

Theorem 3.2. *One has* $\mathrm{Im}\,\theta = \mathcal{D}_S\Omega_{\mathcal{X}_S}/\Omega_{\mathcal{X}_S}$, *hence* $r(\mathcal{X}_S/S) = r'(\mathcal{X}_S/S)$.

Proof. Fix an arbitrary point $x \in X$ and set $s = j(x) \in S$. Then $X = G(\mathbb{R})/K_x$, and $X^\vee = G_\mathbb{C}/P_x$ can also be written $\hat{G}_\mathbb{C}/\hat{P}_x \subset \mathfrak{H}_g^\vee$; $\hat{P}_x$ stabilizes the Lagrangian subspace $V_x := \Omega_{\mathcal{X}_s} \subset \Lambda_\mathbb{C}^\vee$. We write

$$\mathfrak{g} = \mathrm{Lie}\, G_\mathbb{C} = \mathrm{Lie}\, \hat{G}_\mathbb{C}, \quad \mathfrak{k}_\mathbb{C} = \mathrm{Lie}\, K_{x,\mathbb{C}}.$$

The Hodge decomposition of $\mathfrak{g}$ with respect to $ad \circ h_x$ takes the form $\mathfrak{u}^+ \oplus \mathfrak{k}_\mathbb{C} \oplus \mathfrak{u}^-$, where $\mathfrak{u}^+$, the Lie algebra of the unipotent radical of P_x, is of type $(-1,1)$, $\mathfrak{k}_\mathbb{C}$ is of type $(0,0)$, and $\mathfrak{u}^-$ of type $(1,-1)$. One has $\mathfrak{u}^+ \oplus \mathfrak{k}_\mathbb{C} = \mathrm{Lie}\, \hat{P}_x$, $\mathfrak{k}_\mathbb{C} \oplus \mathfrak{u}^-$ is the Lie algebra of an opposite parabolic group P_x^-, and $\mathfrak{k}_\mathbb{C}$ is the common (reductive) Levi factor (*cf.* also [18, 5]).

Looking at the Hodge type, one finds that

$$\left[\mathfrak{u}^+, \mathfrak{u}^+\right] = \left[\mathfrak{u}^-, \mathfrak{u}^-\right] = 0,\ \left[\mathfrak{k}_\mathbb{C}, \mathfrak{u}^+\right] \subset \mathfrak{u}^+,\ \left[\mathfrak{k}_\mathbb{C}, \mathfrak{u}^-\right] \subset \mathfrak{u}^-,\ \left[\mathfrak{u}^+, \mathfrak{u}^-\right] \subset \mathfrak{k}_\mathbb{C}. \quad (3.1)$$

By the Jacobi identity, it follows that

$$\left[\mathfrak{k}_\mathbb{C}, \left[\mathfrak{u}^+, \mathfrak{u}^-\right]\right] \subset \left[\mathfrak{u}^+, \mathfrak{u}^-\right], \quad (3.2)$$

i.e. $[\mathfrak{u}^+, \mathfrak{u}^-]$ is a Lie ideal of $\mathfrak{k}_\mathbb{C}$, hence a reductive Lie algebra.

We may identify $T_x X^\vee = T_s S$ with $\mathfrak{u}^-$. Note that $\Lambda_\mathbb{C}^\vee$ is a faithful representation of $\mathfrak{g}$ and that V_x is stable under $\mathrm{Lie}\, P_x = \mathfrak{u}^+ + \mathfrak{k}_\mathbb{C}$. Using the Hodge decomposition $V_x \oplus \bar{V}_x = \Lambda_\mathbb{C}^\vee \cong H^1_{dR}(\mathcal{X}_s)$, we can write the elements of $\mathfrak{g}$ as matrices in block form $\left(\begin{smallmatrix} R & S \\ T & \iota(R) \end{smallmatrix}\right)$, with $R \in \mathfrak{k}_\mathbb{C}$, $S \in \mathfrak{u}^+$, $T \in \mathfrak{u}^-$ and ι is the involution exchanging P_x and P_x^-. Identifying $\mathfrak{u}^+$ with $\left(\begin{smallmatrix} 0 & \mathfrak{u}^+ \\ 0 & 0 \end{smallmatrix}\right)$ (respectively $\mathfrak{u}^-$ with $\left(\begin{smallmatrix} 0 & 0 \\ \mathfrak{u}^- & 0 \end{smallmatrix}\right)$), one may write $[\mathfrak{u}^+, \mathfrak{u}^-] = \left(\begin{smallmatrix} \mathfrak{u}^+ \cdot \mathfrak{u}^- & 0 \\ 0 & \iota(\mathfrak{u}^+ \cdot \mathfrak{u}^-) \end{smallmatrix}\right)$. Therefore $\mathfrak{u}^+ \cdot \mathfrak{u}^-$ is a reductive Lie algebra acting on V_x. Accordingly, V_x decomposes as $V_0 \oplus V'$, where V_0 is the kernel of this action, and $(\mathfrak{u}^- \cdot \mathfrak{u})V' = V'$.

The identifications $\mathfrak{u}^- = T_s S$ and $V_x = (\Omega_{\mathcal{X}_S})_s$ lead to $V_x \oplus \mathfrak{u}^- V_x = (\mathcal{D}_S^{\leq 1} \Omega_{\mathcal{X}_S})_s$.

Claim. $V_x \oplus \mathfrak{u}^- V_x$ is the smallest $\mathfrak{g}$-submodule of $\Lambda_\mathbb{C}^\vee$ containing V_x. Therefore, it is the fiber at x of $\mathcal{D}_S \Omega_{\mathcal{X}_S}$.

The point is that $V_x + \mathfrak{u}^- V_x = V_x + \mathfrak{u}^- V'$ is stable under $\mathfrak{u}^-$, $\mathfrak{u}^-$ and $\mathfrak{k}_\mathbb{C}$, which follows from (3.1) and from the fact that V_x is stable under $\mathfrak{u}^+ + \mathfrak{k}_\mathbb{C}$. □

4. The case of maximal monodromy (subject to given polarization and endomorphisms)

4.1. Abelian schemes of PEM type

Definition 4.1. A principally polarized Abelian scheme A/S is *of PE-monodromy type – or PEM type* – if its geometric generic fibre is simple and the connected algebraic monodromy is maximal with respect to the polarization ψ and the endomorphisms.

In other words, the Zariski-closure of the monodromy group at $s \in S$ is the maximal algebraic subgroup of $Sp(H^1(A_s), \psi_s)$ which commutes with the action of End A/S (this condition is independent of $s \in S(\mathbb{C})$).

Let us make this more explicit. The endomorphism $\mathbb{Q}$-algebra $D := (\text{End}\, A/S) \otimes \mathbb{Q}$ is the same as the one of its generic fiber; since the latter is assumed to be geometrically simple, D is also the endomorphism $\mathbb{Q}$-algebra of the geometric generic fibre. According to Albert's classification, its falls into one of the following types:

I: a totally real field $F = D$;
II: a totally indefinite quaternion algebra D over a totally real field F;
III: a totally indefinite quaternion algebra D over a totally real field F;
IV: a division algebra D over a CM field F.

Let $E \supset F$ be a maximal subfield of D, which we can take to be a CM field except for type I, and let E^+ be a maximal totally real subfield. For any embedding $\lambda : E^+ \hookrightarrow \mathbb{R}$, let us order the embeddings $\lambda_1, \lambda_2 : E \hookrightarrow \mathbb{C}$ above λ if $E \neq E^+$ (and set $\lambda_1 = \lambda$ otherwise). We identify λ (respectively λ_1) with a homomorphism $E^+ \otimes \mathbb{C} \to \mathbb{C}$ (respectively $E \otimes \mathbb{C} \to \mathbb{C}$). Let us set

$$\mathcal{H}_\lambda = \mathcal{H} \otimes_{E^+ \otimes \mathbb{C}, \lambda} \mathbb{C} \ \text{ (respectively } \mathcal{H}_{\lambda_1} = \mathcal{H} \otimes_{E \otimes \mathbb{C}, \lambda_1} \mathbb{C}). \tag{4.1}$$

By functoriality of the Gauss-Manin connection, these are direct factors of $\mathcal{H}$ as $\mathcal{D}_S$-modules, and $\mathcal{H}_\lambda$ only depends (up to isomorphism) on the restriction $[\lambda]$ of λ to F^+.

Then the maximal possible connected complex monodromy group at an arbitrary point $s \in S(\mathbb{C})$ is of the form $\Pi_{[\lambda]}\, G_{[\lambda]}$ where $G_{[\lambda]}$ and its representation on $\mathcal{H}_{\lambda,s}$ are of the form

I: $Sp(\mathcal{H}_{\lambda,s}),\ St$;
II: $Sp(\mathcal{H}_{\lambda_1,s}),\ St \oplus St$;
III: $SO(\mathcal{H}_{\lambda_1,s}),\ St \oplus St$;
IV: $SL(\mathcal{H}_{\lambda_1,s}),\ St \oplus St^\vee$;

where St denotes the standard representation, and $St^\vee$ its dual. Moreover, for types I, II, III, $\mathcal{H}_{\lambda_1,s}$ is an even-dimensional space, *cf.*, *e.g.*, [1, 5].

Remark 4.2. If A/S is endowed with a level n structure, it is of PEM type if and only if the smallest weakly special subvariety of $\mathcal{A}_{g,n}$ containing the image of S is a special subvariety of PEL type in the sense of Shimura, *i.e.* the image in $\mathcal{A}_{g,n}$ of the moduli space for principally polarized Abelian varieties A such that $D \subset (\text{End}\, A) \otimes \mathbb{Q}$, equipped with level n structure [20] (for $S = \operatorname{Spec} k$, A is of PEM type if and only if A has complex multiplication).

One could also define the related (but weaker) notion of Abelian scheme A/S of PE Hodge type, on replacing the monodromy group by the Mumford-Tate group, *cf.*, *e.g.*, [1]. If A/S is endowed with a level n structure, it is of PE Hodge type if and only if the smallest *special subvariety* of $\mathcal{A}_{g,n}$ containing the image of S is a special subvariety of PEL type in the sense of Shimura.

4.1.1. Parallel to (4.1), one has a decomposition

$$\Omega_{A,\lambda} = \Omega_A \otimes_{E^+\otimes\mathbb{C},\lambda} \mathbb{C} \ \ \text{(respectively } \Omega_{A,\lambda_1} = \Omega_A \otimes_{E\otimes\mathbb{C},\lambda_1} \mathbb{C}). \tag{4.2}$$

The sequence (1.2) induces an exact sequence

$$0 \to \Omega_{A,\lambda_1} \to \mathcal{H}_{\lambda_1} \to \mathcal{H}_{\lambda_1}/\Omega_{A,\lambda_1} \to 0. \tag{4.3}$$

It turns out that for types I, II, III, $\mathcal{H}_{\lambda_1}/\Omega_{A,\lambda_1} \cong \Omega^\vee_{A,\lambda_1}$. This is not the case for type IV, and the pair

$$\big(r_{[\lambda]} = \dim \Omega_{A,\lambda_1,s},\ s_{[\lambda]} = \dim \mathcal{H}_{\lambda_1,s}/\Omega_{A,\lambda_1,s}\big) \tag{4.4}$$

is an interesting invariant called the *Shimura type* (for type IV, the PEL families depend not only on D, the polarization and the level structure, but also on these pairs, when $[\lambda]$ runs among the real embeddings of F^+). On the other hand, $\Omega_{A,\lambda_2} \cong \Omega_{A,\lambda_1}$ for types I, II, III, while $\Omega_{A,\lambda_2} \cong (\mathcal{H}_{\lambda_1}/\Omega_{A,\lambda_1})^\vee$ for type IV.

4.1.2. By functoriality, the Kodaira-Spencer map induces a map

$$\theta_{\partial,\lambda_1} :\ \Omega_{A,\lambda_1} \to \mathcal{H}_{\lambda_1}/\Omega_{A,\lambda_1}. \tag{4.5}$$

Therefore, $\operatorname{rk}\theta_{\partial,\lambda_1} \leq \min(r_{[\lambda]}, s_{[\lambda]})$. In particular, if for some $[\lambda]$, $r_{[\lambda]} \neq s_{[\lambda]}$, then $r' < g$.

Let us consider for example the Shimura family of PEL type of Abelian 3-folds with multiplication by an imaginary quadratic field E (type IV) and invariant $(r_{[\lambda]} = 1, s_{[\lambda]} = 2)$ (it is non empty by [20]). The base is a Shimura surface, and for this family one has $r'' = 2,\ r' = r = g = 3$. Let A/S be the restriction of this Abelian scheme to a general curve of this surface; then $r'' = r' = 2,\ r = g = 3$.

One gets examples with $r < g$ when $r_{[\lambda]} \cdot s_{[\lambda]} = 0$ for some $[\lambda]$.

4.2. Abelian schemes of restricted PEM type

Definition 4.3. A principally polarized Abelian scheme A/S is of *restricted PEM type* if it is of PEM type and for any (equivalently, for all) $s \in S(\mathbb{C})$, $(\Omega_A)_s$ is a free $E \otimes \mathbb{C}$-module.

In the latter condition, one could replace E by F. It is automatic for types I, II, III. For type IV, it amounts to the equality $r_\lambda = s_\lambda$ for every λ; in that case, $\Omega_{\lambda_1} \cong \Omega^\vee_{\lambda_2}$.

Theorem 4.4. *In the restricted PEM case, one has* $r'' = r' = r = g$.

Proof. Thanks to Lemma 1.6 and Theorem 3.1, we are reduced to prove that $r'' = g$ for a Shimura family of PEL type, provided $r_\lambda = s_\lambda$ for every λ in the type IV case. This amounts in turn to showing that there exists ∂ such that $\theta_{\partial,\lambda_1}$ has maximal rank, equal to the rank of $\mathcal{H}_{\lambda_1}$ which is twice the rank m of Ω_{A,λ_1}. Let $\mathfrak{g}$ be one of the Lie algebras $sp(2m),\ so(2m),\ sl(2m)$. In the notation of the proof of Theorem 3.2, The point is to show that $\mathfrak{u}^-$ contains an invertible element. But $\mathfrak{u}^+$ consists of lower left quadrants of elements of $\mathfrak{g}$ viewed as a $2m$-$2m$-matrices; and it is clear that the lower left quadrant of a general element of $\mathfrak{g}$ is an invertible m-m-matrix. □

Remarks 4.5.

i) One can be more precise and give an interpretation of the partial Kodaira-Spencer map at the level of $X^\vee$ as induced by isomorphisms

$$T_{G_{[\lambda]}/P_{[\lambda]}} \xrightarrow{\sim} S^2\mathcal{L}_{\lambda_1} \tag{4.6}$$

for type I and II (the lower left quadrant of an element of $\mathfrak{g}$ is symmetric),

$$T_{G_{[\lambda]}/P_{[\lambda]}} \xrightarrow{\sim} \mathcal{L}_{\lambda_1}^{\otimes 2} \tag{4.7}$$

for type III and IV (the lower left quadrant of an element of $\mathfrak{g}$ can be any m-m-matrix).

ii) Of course one has $r = g$ whenever $\mathcal{H}$ is an irreducible $\mathcal{D}_S$-module.

Claim. If $\mathrm{End}_S A = \mathbb{Z}$ and A/S is not isotrivial, then $\mathcal{H}$ is an irreducible $\mathcal{D}_S$-module.

Indeed, the conclusion can be reformulated as: the local system $R_1 f_*^{an}\mathbb{C}$ is irreducible. Since we know that it is semisimple [10, Section 4.2.6], this is also equivalent, by Schur's lemma, to $\mathrm{End}\, R_1 f_*^{an}\mathbb{C} = \mathbb{C}$ and also to $\mathrm{End}\, R_1 f_*^{an}\mathbb{Z} = \mathbb{Z}$. This equality then follows from the assumptions by the results of [10, Section 4.4]. More precisely, let Z be as in *loc. cit.* the center of $\mathrm{End}\, R_1 f_*^{an}\mathbb{Q}$; then Z is contained in $(\mathrm{End}_S A)\otimes\mathbb{Q}$ (*loc. cit.*, 4.4.7), hence equal to $\mathbb{Q}$, and by *loc. cit.* Proposition 4.4.11 (under conditions (a), (b), (c_1) or (c_2)), one deduces that $\mathrm{End}\, R_1 f_*^{an}\mathbb{Z} = \mathbb{Z}$.

It would be interesting to determine whether $r'' = g$ in this case, beyond the PEM case.

4.3. Differentiating Abelian integrals of the first kind with respect to a parameter

From the above results about differentiating differential forms of the first kind with respect to parameters, it is possible to draw results about differentiating their integrals.

An *Abelian integral of the first kind* on A is a $\mathbb{C}$-linear[10] combination of Abelian periods $\int_\gamma \omega$, with $\omega \in \Gamma\Omega_A$ and γ in the period lattice on a universal covering $\tilde{S}$ of S^{an}.

Theorem 4.6. *Assume that A is an Abelian scheme of restricted PEM type over an affine curve S. Let ∂ be a non-zero derivation of $\mathcal{O}(S)$. Then the derivative of a non zero Abelian integral of the first kind is never an Abelian integral of the first kind (on A).*

[10] Or $\mathcal{O}(S)$-linear, this amounts to the same.

Proof. Let us first treat the case when the monodromy of A/S is Zariski-dense in Sp_{2g} for clarity. We may assume that Ω_A is free. Then an Abelian integral of the first kind is an $\mathcal{O}(S)$-linear combination $\sum_{ij} \lambda_{ij} \int_{\gamma_i} \omega_j$ of entries of $\binom{\Omega_2}{\Omega_1}$. By (1.5), $\sum_{ij} \lambda_{ij} \partial \int_{\gamma_i} \omega_j = \sum_{ijk} \lambda_{ij} (\int_{\gamma_i} \omega_k (R_\partial)_{kj} + \int_{\gamma_i} \omega_k (T_\partial)_{kj})$, *i.e.* an $\mathcal{O}(S)$-linear combination of entries of $\binom{\Omega_2 R_\partial + \mathrm{N}_2 T_\partial}{\Omega_1 R_\partial + \mathrm{N}_1 T_\partial}$.

Since the monodromy of A/S is Zariski-dense in Sp_{2g}, $Y = \begin{pmatrix} \Omega_2 & \mathrm{N}_2 \\ \Omega_1 & \mathrm{N}_1 \end{pmatrix}$ is the generic point of a $Sp_{2g,\mathbb{C}(S)}$-torsor, by differential Galois theory in the fuchsian case (Picard-Vessiot-Schlesinger-Kolchin). Since there is no linear relations between the entries of a generic element of Sp_{2g}, there is no $\mathbb{C}(S)$-linear relations between the entries of $\begin{pmatrix} \Omega_2 & \mathrm{N}_2 \\ \Omega_1 & \mathrm{N}_1 \end{pmatrix}$, or else between the entries of $\begin{pmatrix} \Omega_2 & \mathrm{N}_2 T_\partial \\ \Omega_1 & \mathrm{N}_1 T_\partial \end{pmatrix}$ since T_∂ is invertible (4.4). One concludes that $\sum_{ij} \lambda_{ij} \partial \int_{\gamma_i} \omega_j = \sum \mu_{ij} \int_{\gamma_i} \omega_j$ with $\lambda_{ij}, \mu_{ij} \in \mathcal{O}(S)$ implies $\lambda_{ij} = \mu_{ij} = 0$.

The other cases are treated similarly, decomposing $\mathcal{H}$ into pieces of rank $2m$ indexed by λ as above, and replacing Sp_{2g} by Sp_{2m}, SO_{2m} or SL_{2m} according to the type. $\square$

References

[1] S. ABDULALI, *Abelian varieties and the general Hodge conjecture*, Compos. Math. **109** (1997), 341–355.

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

[3] D. BERTRAND and A. PILLAY, *A Lindemann-Weierstrass theorem for semi-Abelian varieties over function fields*, J. Amer. Math. Soc. **23** (2010), 491–533.

[4] D. BERTRAND and V. ZUDILIN, *On the transcendence degree of the differential field generated by Siegel modular forms*, J. Reine Angew. Math. **554** (2003), 47–68.

[5] J.-B. BOST, *Algebraization, transcendence, and D-group schemes*, Notre Dame J. Form. Log. **54** (2013), 377–434.

[6] M. CAILOTTO and L. FIOROT, *Algebraic connections vs. algebraic $\mathcal{D}$-modules: inverse and direct images*, Port. Math. **66** (2009), 303–320.

[7] C.-L. CHAI and G. FALTINGS, "Degenerations of Abelian Varieties", Ergeb. Math. Grenzgeb., Vol. 22, Springer, 1990.

[8] C. CILIBERTO and G. VAN DER GEER, *The moduli space of Abelian varieties and the singularities of the theta divisor*, Surv. Differ. Geom. **7** (2000), 61–81.

[9] R. COLEMAN, *Duality for the De Rham cohomology of an Abelian scheme*, Ann. Inst. Fourier Grenoble **48** (1998), 1379–1393.

[10] P. DELIGNE, *Théorie de Hodge II*, Inst. Hautes Études Sc. Publ. Math. **40** (1971), 5–57.

[11] Z. GAO, "The Mixed Ax-Lindemann Theorem and its Application to the Zilber-Pink Conjecture", PhD thesis Orsay-Leiden, 2014.

[12] G. VAN DER GEER, *The cohomology of the moduli space of Abelian varieties*, In: "Handbook of Moduli", G. Farkas and I. Morrison (eds.), Vol. I, Adv. Lectures in Math., Vol. 24, Higher Education Press, Beijing-Boston, 2012, 415–458.

[13] S. HELGASON, "Differential Geometry, Lie Groups, and Symmetric Spaces", Acad. Press, 1978.

[14] N. KATZ, *Nilpotent connections and the monodromy theorem. Application of a result of Turrittin*, Inst. Hautes Études Sc. Publ. Math. **39** (1970), 155–232.

[15] N. KATZ, *Algebraic solutions of differential equations (p-curvature and the Hodge filtration)*, Invent. Math. **18** (1972), 1–118.
[16] B. MAZUR and W. MESSING, "Universal Extensions and one Dimensional Crystalline Cohomology, Notes in Mathematics, Vol. 370, Springer, 1974.
[17] J. MILNE, *Canonical models of (mixed) Shimura varieties and automorphic vector bundles*, In: "Automorphic Forms, Shimura Varieties, and L-functions", Proc. Conf. Univ. Michigan, 1990, 283–414.
[18] B. MOONEN, *Linearity properties of Shimura varieties, I*, J. Algebraic Geom. **7** (1998), 539–567.
[19] B. MOONEN and G. VAN DER GEER, "Abelian Varieties", draft of a book.
[20] G. SHIMURA, *On an analytic family of polarized Abelian varieties and automorphic functions*, Ann. of Math. **78** (1963), 149–192.
[21] E. ULLMO and A. YAFAEV, *Algebraic flows on Shimura varieties*, preprint, 2015.
[22] E. VIEHWEG and K. ZUO, *A characterization of certain Shimura curves in the moduli stack of Abelian varieties*, J. Differential Geom. **66** (2004), 233–287.

Institut de Mathématiques de Jussieu
4 place Jussieu
75005 Paris, France
yves.andre@imj-prg.fr

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1417-1446

Polynomial semiconjugacies, decompositions of iterations, and invariant curves

FEDOR PAKOVICH

Abstract. We study the functional equation $A \circ X = X \circ B$, where A, B, and X are polynomials with complex coefficients. Using results of [13] about polynomials sharing preimages of compact sets in $\mathbb{C}$, we show that for given B its solutions may be described in terms of the filled-in Julia set of B. On this base, we prove a number of results describing a general structure of solutions. The results obtained imply in particular the result of Medvedev and Scanlon [10] about invariant curves of maps $F : \mathbb{C}^2 \to \mathbb{C}^2$ of the form $(x, y) \to (f(x), f(y))$, where f is a polynomial, and a version of the result of Zieve and Müller [22] about decompositions of iterations of a polynomial.

Mathematics Subject Classification (2010): 37F10 (primary); 14H99 (secondary).

1. Introduction

Let A and B be rational functions of degree at least two on the Riemann sphere. The functions A and B are called commuting if

$$A \circ B = B \circ A, \tag{1.1}$$

and conjugate if

$$A \circ X = X \circ B \tag{1.2}$$

for some rational function X of degree one.

If (1.2) is satisfied for some rational function X of degree at least two, the function B is called semiconjugate to A, and the function X is called a semiconjugacy from B to A. Unlike conjugation, semiconjugation is not an equivalency relation. We will use the notation $A \leq B$ if for given rational functions A and B there exists a non-constant rational function X such that (1.2) holds, and the notation $A \underset{X}{\leq} B$ if A,B, and X satisfy (1.2). The notation reflects the fact that the binary relation

Received October 30, 2015; accepted in revised form August 18, 2016.
Published online December 2017.

on the set of rational functions defined by equality (1.2) is a preorder. Indeed, if $A \underset{X}{\leq} B$ and $B \underset{Y}{\leq} C$ then $A \underset{X\circ Y}{\leq} C$.

Both equations (1.1) and (1.2) have "obvious" solutions. Namely, equation (1.1) has solutions of the form

$$A = R^{\circ m}, \quad B = R^{\circ n}, \tag{1.3}$$

where R is an arbitrary rational function and $m, n \geq 1$. Notice that such A and B have an iteration in common, that is

$$A^{\circ n} = B^{\circ m}. \tag{1.4}$$

In order to obtain solutions of equation (1.2) we can take arbitrary rational functions A_1, B_1 and set

$$F = A_1 \circ B_1, \quad G = B_1 \circ A_1.$$

Then the equality

$$(A_1 \circ B_1) \circ A_1 = A_1 \circ (B_1 \circ A_1) \tag{1.5}$$

implies that $F \underset{A_1}{\leq} G$. Similarly, $G \underset{B_1}{\leq} F$. Moreover, if now A_2, B_2 are rational functions such that the equality

$$G = A_2 \circ B_2 \tag{1.6}$$

holds, then the function $H = B_2 \circ A_2$ satisfies $G \underset{A_2}{\leq} H$ and $H \underset{B_2}{\leq} G$, implying that $F \underset{A_1\circ A_2}{\leq} H$ and $H \underset{B_2\circ B_1}{\leq} F$. This motivates the following definition of an equivalency relation on the set of rational functions: $F \sim G$ if there exist rational functions A_i, B_i, $1 \leq i \leq n$, such that

$$F = A_1 \circ B_1, \quad G = B_n \circ A_n,$$

and

$$B_i \circ A_i = A_{i+1} \circ B_{i+1}, \quad 1 \leq i \leq n-1.$$

Clearly, $F \sim G$ implies that $F \leq G$ and $G \leq F$. Notice that, since for any rational function X of degree one the equality

$$A = (A \circ X) \circ X^{-1}$$

implies that $A \sim X^{-1} \circ A \circ X$, any equivalence class is a collection of conjugacy classes.

Functional equation (1.1) was first studied by Fatou, Julia, and Ritt in the papers [5,8], and [21]. In all these papers it was assumed that the considered commuting functions A and B have no iterate in common. Fatou and Julia described solutions of (1.1) under the additional assumption that the Julia set of A or B does not coincide with the whole complex plane, and Ritt investigated the general case. Briefly, the Ritt theorem states that if rational functions A and B commute and no iterate of A is equal to an iterate of B, then, up to a conjugacy, A and B are either powers, or Chebyshev polynomials, or Lattès functions. Another proof of the Ritt theorem was given by Eremenko in [4]. Notice however that a description of commuting A and B with a common iterate is known only in the polynomial case. Thus, in a certain sense the classification of commuting rational functions is not yet completed. On the other hand, it was shown by Ritt [19,21] that in the polynomial case equality (1.1) implies that, up to the change

$$A \to \lambda \circ A \circ \lambda^{-1}, \quad B \to \lambda \circ B \circ \lambda^{-1},$$

where λ is a polynomial of degree one, either

$$A = z^n, \quad B = \varepsilon z^m,$$

where $\varepsilon^n = \varepsilon$, or

$$A = \pm T_n, \quad B = \pm T_m,$$

or

$$A = \varepsilon_1 R^{\circ m}, \quad B = \varepsilon_2 R^{\circ n},$$

where $R = zS(z^\ell)$ for some polynomial S and ε_1, ε_2 are l-th roots of unity. In fact, this conclusion remains true if instead of (1.1) one were to assume only that A and B share a completely invariant compact set in $\mathbb{C}$ (see [13]).

Equation (1.2) was investigated in the recent paper [17]. The main result of [17] states that if a rational function B is semiconjugate to a rational function A, then either $A \sim B$, or A and B are "minimal holomorphic self-maps" between orbifolds of non-negative Euler characteristic on the Riemann sphere. The latter class of functions is a natural extension of the class of Lattès functions and admits a neat characterization. However, as with the description of commuting rational functions, the description of solutions of (1.2) given in [17] is not completely satisfactory, since it gives no information about equivalent rational functions. In particular, the results of [17] do not provide any bounds on the number of conjugacy classes in an equivalence class of a rational function B or more generally on the number of conjugacy classes of A such that $A \leq B$. Another related problem is the following: is it true that if conditions $A \leq B$ and $B \leq A$ hold simultaneously, then $A \sim B$? Finally, it would be desirable to obtain some handy structural descriptions of the totality of X satisfying (1.2) for given A and B, and of the totality of A satisfying $A \leq B$ for given B.

In this paper we study equation (1.2) with emphasis on the above questions in the case where all the functions involved are *polynomials*. Notice that in distinction with the general case, for polynomials there exists quite a comprehensive theory of functional decompositions developed by Ritt [20]. Nevertheless, questions regarding polynomial decompositions may be highly non-trivial, and a number of recent papers are devoted to such questions arising from different branches of mathematics. Let us mention for example the paper [22] with applications to algebraic dynamics [6], or the paper [16] with applications to differential equations [18]. Another example is the recent paper [10] about invariant varieties for dynamical systems defined by coordinatwise actions of polynomials, a considerable part of which concerns properties of polynomial solutions of (1.2).

The main distinction between this paper and the above mentioned papers is the systematical use of ideas and results from the paper [13] which relates polynomials sharing preimages of compact sets in $\mathbb{C}$ with the functional equation

$$A \circ C = D \circ B.$$

In particular, the main result of [13] leads to a characterization of polynomial solutions of (1.2) in terms of filled-in Julia sets. Recall that for a polynomial B the filled-in Julia set $K(B)$ is defined as the set of points in $\mathbb{C}$ whose orbits under iterations of B are bounded. Since equality (1.2) implies the equalities

$$A^{\circ n} \circ X = X \circ B^{\circ n}, \quad n \geq 1,$$

it it easy to see that if X is a semiconjugacy from B to A, then the preimage $X^{-1}(K(A))$ coincides with $K(B)$. We show that this property is in fact characteristic.

Theorem 1.1. *Let A, B and X be polynomials of degree at least two such that $A \underset{X}{\leq} B$. Then*

$$X^{-1}(K(A)) = K(B). \tag{1.7}$$

Conversely, if equality (1.7) *holds and* $\deg A = \deg B$*, then there exists a polynomial of degree one μ such that*

$$(\mu \circ A) \circ X = X \circ B$$

and $\mu(K(A)) = K(A)$. More generally, if for given B and X the condition

$$X^{-1}(K) = K(B) \tag{1.8}$$

holds for some compact set K in $\mathbb{C}$, then there exists a polynomial A such that $A \underset{X}{\leq} B$ and $K(A) = K$.

For a fixed polynomial B of degree at least two denote by $\mathcal{E}(B)$ the set of polynomials X of degree at least two such that $A \underset{X}{\leq} B$ for some polynomial A. An immediate corollary of Theorem 1.1 is that a polynomial X is contained in $\mathcal{E}(B)$ if and only if $K(B)$ is a union of fibers of X. Another corollary is that if $A \underset{X}{\leq} B$, then for any decomposition $X = X_1 \circ X_2$ there exists a polynomial C such that

$$A \underset{X_1}{\leq} C, \quad C \underset{X_2}{\leq} B.$$

Notice that in particular this casts the problem of the description of decompositions of iterations of a polynomial, first considered in the paper [22], into the context of equation (1.2). Indeed, since $B \circ B^{\circ d} = B^{\circ d} \circ B$, the polynomial $B^{\circ d}$ is contained in $\mathcal{E}(B)$ and hence for any decomposition $B^{\circ d} = Y \circ X$ the equalities

$$B \circ Y = Y \circ A, \quad A \circ X = X \circ B$$

hold for some polynomial A.

The following statement is another corollary of the main result of [13].

Theorem 1.2. *For any $X_1, X_2 \in \mathcal{E}(B)$ there exists $X \in \mathcal{E}(B)$ such that* $\deg X = \mathrm{LCM}(\deg X_1, \deg X_2)$ *and*

$$X = U_1 \circ X_1 = U_2 \circ X_2$$

for some polynomials U_1, U_2. Furthermore, there exists $W \in \mathcal{E}(B)$ such that $\deg W = \mathrm{GCD}(\deg X_1, \deg X_2)$ *and*

$$X_1 = V_1 \circ W, \quad X_2 = V_2 \circ W$$

for some polynomials V_1, V_2.

For fixed polynomials A, B denote by $\mathcal{E}(A, B)$ the subset of $\mathcal{E}(B)$ (possibly empty) consisting of polynomials X such that $A \underset{X}{\leq} B$. In particular, the set $\mathcal{E}(B, B)$ consists of polynomials of degree at least two commuting with B. We will call a polynomial P *special* if it is conjugate to z^n or $\pm T_n$, or equivalently if there exists a Möbius transformation μ which maps $K(P)$ to $\mathbb{D}$ or $[-1, 1]$. The following result describes a general structure of $\mathcal{E}(A, B)$ for non-special A, B.

Theorem 1.3. *Let A and B be fixed non-special polynomials of degree at least two such that the set $\mathcal{E}(A, B)$ is non-empty, and let X_0 be an element of $\mathcal{E}(A, B)$ of minimal degree. Then a polynomial X belongs to $\mathcal{E}(A, B)$ if and only if $X = \widetilde{A} \circ X_0$ for some polynomial $\widetilde{A}$ commuting with A.*

Notice that in a sense this result is a generalization of the result of Ritt about commuting polynomials. Indeed, applying Theorem 1.3 for $B = A$ and $X = B$, we obtain that if A is non-special and $B \in \mathcal{E}(A, A)$, then $B = \widetilde{A} \circ R$, where R is a polynomial of minimal degree in $\mathcal{E}(A, A)$. Now we can apply Theorem 1.3 again to the polynomial $\widetilde{A}$ and so on, eventually obtaining the representation $B = \mu_1 \circ R^{\circ m_1}$, where μ_1 is a polynomial of degree one commuting with A. In particular, since $A \in \mathcal{E}(A, A)$, the equality $A = \mu_2 \circ R^{\circ m_2}$ holds for some polynomial μ_2 of degree one commuting with A.

Another corollary of Theorem 1.3 is the following result obtained by Medvedev and Scanlon in [10]: if $\mathcal{C} \subset \mathbb{C}^2$ is an irreducible algebraic curve invariant under the map $F : (x, y) \to (f(x), f(y))$, where f is a non-special polynomial, then there exists a polynomial p which commutes with f such that $\mathcal{C}$ has the form $z_1 = p(z_2)$ or $z_2 = p(z_1)$. More generally, we prove the following statement which supplements the results of [10] about algebraic curves invariant under the map $F : (x, y) \to (f(x), g(y))$, where f and g are non-special polynomials.

Theorem 1.4. *Let f and g be non-special polynomials of degree at least two and $\mathcal{C}$ a curve in $\mathbb{C}^2$. Then $\mathcal{C}$ is an irreducible (f, g)-invariant curve if and only if $\mathcal{C}$ has the form $u(x)-v(y)=0$, where u, v are polynomials of coprime degrees satisfying the equations*

$$t \circ u = u \circ f, \quad t \circ v = v \circ g \tag{1.9}$$

for some polynomial t.

Our next result describes the interrelations between the equivalence $\sim$, the preorder $\leq$, and decompositions of iterations.

Theorem 1.5. *Let A and B be polynomials of degree at least two. Then conditions $A \leq B$ and $B \leq A$ hold simultaneously if and only if $A \sim B$. Furthermore, $A \sim B$ if and only if there exist polynomials X, Y such that*

$$B \circ Y = Y \circ A, \quad A \circ X = X \circ B,$$

and $Y \circ X = B^{\circ d}$ for some $d \geq 0$.

For a fixed polynomial B of degree at least two denote by $\mathcal{F}(B)$ the set of polynomials A such that $A \leq B$. The following theorem gives a structural description of the set $\mathcal{F}(B)$.

Theorem 1.6. *Let B be a fixed non-special polynomial of degree $n \geq 2$. Then there exist $A \in \mathcal{F}(B)$ and a semiconjugacy X from B to A which are universal in the following sense: for any polynomial $C \in \mathcal{F}(B)$ there exist polynomials X_C, U_C such that $X = U_C \circ X_C$ and the diagram*

$$\begin{tikzcd} \mathbb{C} \arrow[r, "B"] \arrow[d, "X_C"] & \mathbb{C} \arrow[d, "X_C"] \\ \mathbb{C} \arrow[r, "C"] \arrow[d, "U_C"] & \mathbb{C} \arrow[d, "U_C"] \\ \mathbb{C} \arrow[r, "A"] & \mathbb{C} \end{tikzcd} \tag{1.10}$$

is commutative. Furthermore, the degree of X is bounded from above by a constant $c = c(n)$ which depends on n only.

We did not make special efforts to obtain an optimal estimation for $c(n)$, however our method of proof shows that

$$c(n) \leq (n-1)!n^{2\log_2 n+3}.$$

Thus, Theorem 1.6 gives an effective bound on the number of conjugacy classes of polynomials A such that $A \leq B$.

The paper is organized as follows. In the second section we give a very brief overview of the Ritt theory. In the third section we recall basic results of [13] and prove Theorem 1.1 and Theorem 1.2. We also prove the corollaries of Theorem 1.1 mentioned above. In the fourth section we first show that if $A \leq B$ and one of polynomials A or B is special, then the other one also is special (Theorem 4.4). Then we prove Theorem 1.3 and deduce from it the result of Ritt about commuting polynomials. We also apply Theorem 1.3 to the problem of description of curves in $\mathbb{C}^2$ invariant under maps $F : (x, y) \to (f(x), g(y))$, where f and g are polynomials, and prove Theorem 1.4. Finally, we prove Theorem 1.5.

In the fifth section we first show (Theorem 5.2) that if B is a non-special polynomial of degree n, and $X \in \mathcal{E}(B)$, then the degree l of any special compositional factor of X satisfies the inequality $l \leq 2n$. On this base we prove that if $X \in \mathcal{E}(B)$ is not a polynomial in B, then $\deg X$ is bounded from above by a constant which depends on n only. In turn, from this result we deduce Theorem 1.6. As another corollary of the boundedness of $\deg X$ we obtain the following result of Zieve and Müller [22]: if B is a non-special polynomial of degree $n \geq 2$, and X and Y are polynomials such that $Y \circ X = B^{\circ s}$ for some $s \geq 1$, then there exist polynomials $\widetilde{X}$, $\widetilde{Y}$ and $i, j \geq 0$ such that

$$Y = B^{\circ i} \circ \widetilde{Y}, \quad X = \widetilde{X} \circ B^{\circ j}, \quad \text{and} \quad \widetilde{Y} \circ \widetilde{X} = B^{\circ \widetilde{s}},$$

where $\widetilde{s}$ is bounded from above by a constant which depends on n only.

ACKNOWLEDGEMENTS. The author is grateful to the Max-Planck-Institut füer Mathematik for the hospitality and the support.

2. Overview of the Ritt theory

Let F be a polynomial with complex coefficients. The polynomial F is called *indecomposable* if the equality $F = F_2 \circ F_1$ implies that at least one of the polynomials F_1, F_2 is of degree one. Any representation of a polynomial F in the form $F = F_r \circ F_{r-1} \circ \cdots \circ F_1$, where $F_1, F_2, \ldots, F_r$ are polynomials, is called *a decomposition* of F. A decomposition is called *maximal* if all $F_1, F_2, \ldots, F_r$ are indecomposable and of degree greater than one. Two decompositions having an equal number of terms

$$F = F_r \circ F_{r-1} \circ \cdots \circ F_1 \quad \text{and} \quad F = G_r \circ G_{r-1} \circ \cdots \circ G_1$$

are called *equivalent* if either $r = 1$ and $F_1 = G_1$, or $r \geq 2$ and there exist polynomials μ_i, $1 \leq i \leq r-1$, of degree 1 such that

$$F_r = G_r \circ \mu_{r-1}, \quad F_i = \mu_i^{-1} \circ G_i \circ \mu_{i-1}, \quad 1 < i < r, \quad \text{and} \quad F_1 = \mu_1^{-1} \circ G_1.$$

The theory of polynomial decompositions established by Ritt can be summarized in the form of two theorems usually called the first and the second Ritt theorems (see [20]).

The first Ritt theorem states, roughly speaking, that any maximal decompositions of a polynomial may be obtained from any other by some iterative process involving the functional equation

$$A \circ C = D \circ B. \tag{2.1}$$

Theorem 2.1 ([20]). *Any two maximal decompositions $\mathcal{D}, \mathcal{E}$ of a polynomial P have an equal number of terms. Furthermore, there exists a chain of maximal decompositions $\mathcal{F}_i$, $1 \leq i \leq s$, of P such that $\mathcal{F}_1 = \mathcal{D}$, $\mathcal{F}_s \sim \mathcal{E}$, and $\mathcal{F}_{i+1}$ is obtained from $\mathcal{F}_i$ by a replacement of two successive polynomials $A \circ C$ in $\mathcal{F}_i$ by two other polynomials $D \circ B$ such that* (2.1) *holds.*

The second Ritt theorem in turn describes indecomposable polynomial solutions of (2.1). More precisely, it describes solutions satisfying the condition

$$\mathrm{GCD}(\deg A, \deg D) = 1, \quad \mathrm{GCD}(\deg C, \deg B) = 1, \tag{2.2}$$

which holds in particular if A, C, D, B are indecomposable (see Theorem 2.3 below).

Theorem 2.2 ([20]). *Let A, C, D, B be polynomials such that* (2.1) *and* (2.2) *hold. Then there exist polynomials $\sigma_1, \sigma_2, \mu, \nu$ of degree one such that, up to a possible replacement of A by D and of C by B, either*

$$A = \nu \circ z^s R^n(z) \circ \sigma_1^{-1}, \qquad C = \sigma_1 \circ z^n \circ \mu \tag{2.3}$$

$$D = \nu \circ z^n \circ \sigma_2^{-1}, \qquad B = \sigma_2 \circ z^s R(z^n) \circ \mu, \tag{2.4}$$

where R is a polynomial, $n \geq 1$, $s \geq 0$, and $\mathrm{GCD}(s, n) = 1$, *or*

$$A = \nu \circ T_m \circ \sigma_1^{-1}, \qquad C = \sigma_1 \circ T_n \circ \mu, \tag{2.5}$$

$$D = \nu \circ T_n \circ \sigma_2^{-1} \qquad B = \sigma_2 \circ T_m \circ \mu, \tag{2.6}$$

where T_n, T_m are the Chebyshev polynomials, $n, m \geq 1$, and $\mathrm{GCD}(n, m) = 1$.

Notice that the main difficulty in the practical use of Theorem 2.1 and Theorem 2.2 is the fact that classes of solutions appearing in Theorem 2.2 are not disjoint. Namely, any solution of the form (2.5), (2.6) with $n = 2$ can also be represented in the form (2.3), (2.4) (see, *e.g.*, [10,16,22] for further details).

The description of polynomial solutions of equation (2.1) in the general case in a certain sense reduces to the case where (2.2) holds by the following statement.

Theorem 2.3 ([3]). *Let* A, C, D, B *be polynomials such that* (2.1) *holds. Then there exist polynomials* U, V, $\widetilde{A}$, $\widetilde{C}$, $\widetilde{D}$, $\widetilde{B}$, *where*

$$\deg U = \mathrm{GCD}(\deg A, \deg D), \quad \deg V = \mathrm{GCD}(\deg C, \deg B),$$

such that

$$A = U \circ \widetilde{A}, \;\; D = U \circ \widetilde{D}, \;\; C = \widetilde{C} \circ V, \;\; B = \widetilde{B} \circ V,$$

and

$$\widetilde{A} \circ \widetilde{C} = \widetilde{D} \circ \widetilde{B}.$$

In particular, if $\deg C = \deg B$, *then there exists a polynomial* μ *of degree one such that*

$$A = D \circ \mu^{-1}, \quad C = \mu \circ B.$$

Theorem 2.2 implies the following description of polynomial solutions of equation (1.2) under the condition

$$\mathrm{GCD}(\deg X, \deg B) = 1 \tag{2.7}$$

(see [7]).

Theorem 2.4 ([7]). *Let* A, B, X *be polynomials such that* (1.2) *and* (2.7) *hold. Then there exist polynomials* μ, ν *of degree one such that either*

$$A = \nu \circ z^s R^n(z) \circ \nu^{-1}, \quad X = \nu \circ z^n \circ \mu, \quad D = \mu^{-1} \circ z^s R(z^n) \circ \mu,$$

where R *is a polynomial,* $n \geq 1$, $s \geq 0$, *and* $\mathrm{GCD}(s, n) = 1$, *or*

$$A = \nu \circ \pm T_m \circ \nu^{-1}, \quad X = \nu \circ T_n \circ \mu, \quad D = \mu^{-1} \circ \pm T_m \circ \mu,$$

where T_n, T_m *are the Chebyshev polynomials,* $n, m \geq 1$, *and* $\mathrm{GCD}(n, m) = 1$.

Notice, however, that Theorem 2.2, even combined with Theorem 2.3, provides very little information about solutions of (1.2) if (2.7) is not satisfied. A possible way to investigate the general case is to analyze somehow the totality of all decompositions of a polynomial P, basing on Theorem 2.1 and Theorem 2.2, and then to apply this analysis to (1.2) using the fact that we can pass from the decomposition $P = A \circ X$ to the decomposition $P = X \circ B$. This idea was used in [10]. A similar technique was used in [22], where it was applied to the study of decompositions of iterations of a polynomial. In this paper we use another method completely bypassing Theorem 2.1. Notice by the way that Theorem 2.1 does not hold for arbitrary rational functions (see, *e.g.*, [12]).

3. Semiconjugacies and Julia sets

3.1. Polynomials sharing preimages of compact sets

Let $f_1(z)$, $f_2(z)$ be non-constant complex polynomials and $K_1, K_2 \subset \mathbb{C}$ compact sets. In the paper [13] we investigated the following problem. Under what conditions on the collection $f_1(z), f_2(z), K_1, K_2$ do the preimages $f_1^{-1}(K_1)$ and $f_2^{-1}(K_2)$ coincide, that is,

$$f_1^{-1}(K_1) = f_2^{-1}(K_2) = K \tag{3.1}$$

for some compact set $K \subset \mathbb{C}$?

Using ideas from approximation theory, we relate equation (3.1) to the functional equation

$$g_1(f_1(z)) = g_2(f_2(z)), \tag{3.2}$$

where $f_1(z), f_2(z), g_1(z), g_2(z)$ are polynomials. It is easy to see that for any polynomial solution of (3.2) and any compact set $K_3 \subset \mathbb{C}$ we obtain a solution of (3.1) setting

$$K_1 = g_1^{-1}(K_3), \quad K_2 = g_2^{-1}(K_3). \tag{3.3}$$

Briefly, the main result of [13] states that, under a very mild condition on the cardinality of K, all solutions of (3.1) can be obtained in this way. Combined with Theorem 2.3 and Theorem 2.2 this leads to a very explicit description of solutions of (3.1).

Theorem 3.1 ([13]). *Let* $f_1(z)$, $f_2(z)$ *be polynomials,* $\deg f_1 = d_1$, $\deg f_2 = d_2$, $d_1 \leq d_2$, *and let* $K_1, K_2, K \subset \mathbb{C}$ *be compact sets such that* (3.1) *holds. Suppose that* $\mathrm{card}\{\mathrm{K}\} \geq \mathrm{LCM}(\mathrm{d}_1, \mathrm{d}_2)$. *Then, if* d_1 *divides* d_2, *there exists a polynomial* $g_1(z)$ *such that* $f_2(z) = g_1(f_1(z))$ *and* $K_1 = g_1^{-1}(K_2)$. *On the other hand, if* d_1 *does not divide* d_2, *then there exist polynomials* $g_1(z)$, $g_2(z)$, $\deg g_1 = d_2/d$, $\deg g_2 = d_1/d$, *where* $d = \mathrm{GCD}(d_1, d_2)$, *and a compact set* $K_3 \subset \mathbb{C}$ *such that* (3.2), (3.3) *hold. Furthermore, in this case there exist polynomials* $\widetilde{f}_1(z)$, $\widetilde{f}_2(z)$, $W(z)$, $\deg W(z) = d$, *such that*

$$f_1(z) = \widetilde{f}_1(W(z)), \qquad f_2(z) = \widetilde{f}_2(W(z)) \tag{3.4}$$

and there exist linear functions $\sigma_1(z)$, $\sigma_2(z)$ *such that either*

$$\begin{aligned} g_1(z) &= z^c R^{d_1/d}(z) \circ \sigma_1^{-1}, \qquad & \widetilde{f}_1(z) &= \sigma_1 \circ z^{d_1/d}, \\ g_2(z) &= z^{d_1/d} \circ \sigma_2^{-1}, \qquad & \widetilde{f}_2(z) &= \sigma_2 \circ z^c R\big(z^{d_1/d}\big), \end{aligned} \tag{3.5}$$

for some polynomial $R(z)$ *and* c *equal to the remainder after division of* d_2/d *by* d_1/d, *or*

$$\begin{aligned} g_1(z) &= T_{d_2/d}(z) \circ \sigma_1^{-1}, \qquad & \widetilde{f}_1(z) &= \sigma_1 \circ T_{d_1/d}(z), \\ g_2(z) &= T_{d_1/d}(z) \circ \sigma_2^{-1}, \qquad & \widetilde{f}_2(z) &= \sigma_2 \circ T_{d_2/d}(z), \end{aligned} \tag{3.6}$$

for the Chebyshev polynomials $T_{d_1/d}(z)$, $T_{d_2/d}(z)$.

Theorem 3.1 may be used to prove many other results (see [13] for details), the most notable of which is the following description of solutions of (3.1) in the case where $K_1 = K_2$, first obtained by T. Dinh [1,2] by methods of complex dynamics.

Theorem 3.2 ([2,13]). *Let $f_1(z)$, $f_2(z)$ be polynomials such that*

$$f_1^{-1}(T) = f_2^{-1}(T) = K \tag{3.7}$$

holds for some infinite compact sets T, $K \subset \mathbb{C}$. Then, if d_1 divides d_2, there exists a polynomial $g_1(z)$ such that $f_2(z) = g_1(f_1(z))$ and $g_1^{-1}(T) = T$. On the other hand, if d_1 does not divide d_2, then there exist polynomials $\widetilde{f_1}(z)$, $\widetilde{f_2}(z)$, $W(z)$, $\deg W(z) = d$, satisfying (3.4). *Furthermore, in this case one of the following conditions holds:*

1) *T is a union of concentric circles and*

$$\widetilde{f_1}(z) = \sigma \circ z^{d_1/d}, \qquad \widetilde{f_2}(z) = \sigma \circ \gamma z^{d_2/d} \tag{3.8}$$

for some linear function $\sigma(z)$ and $\gamma \in \mathbb{C}$;

2) *T is a segment and*

$$\widetilde{f_1}(z) = \sigma \circ \pm T_{d_1/d}(z), \qquad \widetilde{f_2}(z) = \sigma \circ \pm T_{d_2/d}(z), \tag{3.9}$$

for some linear function $\sigma(z)$ and the Chebyshev polynomials $T_{d_1/d}(z)$, $T_{d_2/d}(z)$.

3.2. Proofs of Theorem 1.1 and Theorem 1.2

Proof of Theorem 1.1. If $A \underset{X}{\leq} B$, then for any $n \geq 1$ the equality

$$A^{\circ n} \circ X = X \circ B^{\circ n}$$

holds. Therefore, if $z_1 = X(z_0)$, then the sequence $A^{\circ n}(z_1)$ is bounded if and only if the sequence $X \circ B^{\circ n}(z_0)$ is bounded. In turn, the last sequence is bounded if and only if the sequence $B^{\circ n}(z_0)$ is bounded. Thus, $A \underset{X}{\leq} B$ implies

$$X^{-1}(K(A)) = K(B). \tag{3.10}$$

Conversely, if (3.10) holds, then it follows from $B^{-1}(K(B)) = K(B)$ that

$$(X \circ B)^{-1}(K(A)) = K(B).$$

Thus,

$$X^{-1}(K(A)) = (X \circ B)^{-1}(K(A)).$$

Since $\deg X \mid \deg(X \circ B)$, applying to the latter equality Theorem 3.1 we conclude that

$$\widetilde{A} \circ X = X \circ B$$

for some polynomial $\widetilde{A}$. Furthermore, since we proved that for such $\widetilde{A}$ the equality $X^{-1}(K(\widetilde{A})) = K(B)$ holds, we see that $X^{-1}(K(\widetilde{A})) = X^{-1}(K(A))$, implying that $K(\widetilde{A}) = K(A)$. Finally, it follows from Theorem 3.1 applied to

$$A^{-1}(K) = \widetilde{A}^{-1}(K) = K,$$

where $K = K(\widetilde{A}) = K(A)$, that there exists a polynomial of degree one μ such that $\widetilde{A} = \mu \circ A$ and $\mu(K(A)) = K(A)$.

More generally, if

$$X^{-1}(K) = K(B) \tag{3.11}$$

for some compact set $K \subset \mathbb{C}$, then

$$X^{-1}(K) = (X \circ B)^{-1}(K),$$

implying by Theorem 3.1 that (1.2) holds for some polynomial A. Furthermore, since for such a polynomial A equality (3.10) holds, we conclude that $X^{-1}(K) = X^{-1}(K(A))$ and $K = K(A)$. □

Corollary 3.3. *Let B be a polynomial of degree at least two. Then a polynomial X is contained in $\mathcal{E}(B)$ if and only $K(B)$ is a union of fibers of X. In particular, if B_1 and B_2 are polynomials such that $K(B_1) = K(B_2)$, then $\mathcal{E}(B_1) = \mathcal{E}(B_2)$.*

Proof. Clearly, condition (3.11) implies that $K(B)$ is a union of fibers of X. Conversely, if $K(B)$ is a union of fibers of X, then

$$K(B) = X^{-1}\big(X(K(B))\big),$$

implying that (3.11) holds for the compact set $K = X(K(B))$. □

Corollary 3.4. *Let A, B, and X be polynomials such that $A \underset{X}{\leq} B$. Then for any decomposition $X = X_1 \circ X_2$ there exists a polynomial C such that*

$$A \underset{X_1}{\leq} C, \quad C \underset{X_2}{\leq} B.$$

Proof. By Theorem 1.1, $K(B) = X^{-1}(K(A))$. Since $X = X_1 \circ X_2$, this implies that $K(B) = X_2^{-1}(\widetilde{K})$, where $\widetilde{K} = X_1^{-1}(K(A))$. Therefore, by Theorem 1.1, there exists a polynomial C such that

$$C \circ X_2 = X_2 \circ B. \tag{3.12}$$

Now we have:

$$A \circ X_1 \circ X_2 = X_1 \circ X_2 \circ B = X_1 \circ C \circ X_2,$$

implying that $A \circ X_1 = X_1 \circ C$. □

Remark 3.5. Corollary 3.4 may be proved without using Theorem 1.1. Indeed, if $X = X_1 \circ X_2$, then it follows from

$$A \circ (X_1 \circ X_2) = X_1 \circ (X_2 \circ B),$$

by Theorem 2.3, that

$$X_1 \circ X_2 = U \circ \widetilde{W}, \quad X_2 \circ B = V \circ \widetilde{W}, \tag{3.13}$$

where

$$\deg \widetilde{W} = \mathrm{GCD}(\deg(X_1 \circ X_2), \deg(X_2 \circ B)).$$

Since $\deg X_2 \mid \deg \widetilde{W}$, Theorem 2.3 applied to the first equality in (3.13) implies that $\widetilde{W} = S \circ X_2$ for some polynomial S. Therefore,

$$X_2 \circ B = V \circ \widetilde{W} = V \circ S \circ X_2$$

and hence (3.12) holds for $C = V \circ S$.

Proof of Theorem 1.2. By Theorem 1.1, the condition $X_1, X_2 \in \mathcal{E}(B)$ implies that there exist $K_1, K_2 \subset \mathbb{C}$ such that

$$X_1^{-1}(K_1) = K(B), \quad X_2^{-1}(K_2) = K(B).$$

It now follows from Theorem 3.1 that there exist polynomials X, W, U_1, U_2, V_1, V_2 such that

$$\deg X = \mathrm{LCM}(\deg X_1, \deg X_2), \quad \deg W = \mathrm{GCD}(\deg X_1, \deg X_2),$$

and that equalities

$$X = U_1 \circ X_1 = U_2 \circ X_2$$

and

$$X_1 = V_1 \circ W, \quad X_2 = V_2 \circ W \tag{3.14}$$

hold. Furthermore, there exists $K_3 \subset \mathbb{C}$ such that

$$K_1 = U_1^{-1}(K_3), \quad K_2 = U_2^{-1}(K_3).$$

Therefore, $X^{-1}(K_3) = K(B)$, implying by Theorem 1.1 that $X \in \mathcal{E}(B)$. Finally, any of equalities (3.14) implies that $W \in \mathcal{E}(B)$ by Corollary 3.4. $\square$

4. Semiconjugacies between fixed A and B

4.1. Semiconjugacies between special polynomials

For a polynomial P and a finite set $K \subset \mathbb{C}$ denote by $P^{-1}_{\text{odd}}(K)$ the subset of $P^{-1}(K)$ consisting of points where the local multiplicity of P is odd. Notice that the chain rule implies that if $P = A \circ B$, then

$$P^{-1}_{\text{odd}}(K) = B^{-1}_{\text{odd}}\left(A^{-1}_{\text{odd}}(K)\right). \tag{4.1}$$

Lemma 4.1. *Let P be a polynomial of degree $n \geq 2$, and $K \subset \mathbb{C}$ a finite set containing at least two points. Assume that $P^{-1}_{\text{odd}}(K) = K$. Then K contains exactly two points, and P is conjugate to $\pm T_n$.*

Proof. Denote by e_z the multiplicity of P at $z \in \mathbb{C}$, and set $r = \text{card}(K)$. Since for any $y \in \mathbb{C}$ the set $P^{-1}(y)$ contains

$$n - \sum_{\substack{z \in \mathbb{C} \\ P(z)=y}} (e_z - 1)$$

points and

$$\sum_{z \in \mathbb{C}} (e_z - 1) = n - 1,$$

we have:

$$\text{card}\big(P^{-1}(K)\big) \geq rn - \sum_{z \in \mathbb{C}} (e_z - 1) = (r-1)n + 1 \tag{4.2}$$

(the minimum is attained if K contains all finite critical values of P). Therefore, if

$$\text{card}\left(P^{-1}_{\text{odd}}(K)\right) = \text{card}(K) = r,$$

then the set $P^{-1}(K)$ contains at least $(r-1)n + 1 - r$ points where the local multiplicity of P is greater than one, implying that

$$\sum_{z \in P^{-1}(K)} e_z \geq r + 2\left((r-1)n + 1 - r\right). \tag{4.3}$$

Since the sum in the left-hand side of (4.3) equals rn, this inequality implies that

$$(n-1)(r-2) \leq 0. \tag{4.4}$$

Thus, $r = 2$. Furthermore, since the equality in (4.4) is attained if and only if equality is attained in (4.3), we conclude that if $P^{-1}_{\text{odd}}(K) = K$, then $e_z = 2$ for each $z \in P^{-1}(K) \setminus K$, and the local multiplicity of P at each of the two points of K is equal to one.

Changing P to $\sigma^{-1} \circ P \circ \sigma$ for a convenient polynomial of degree one σ, we can assume that $K = \{-1, 1\}$. Then the condition on multiplicities of P implies that $P^2 - 1$ is divisible by $(P')^2$, and calculating the quotient we conclude that P satisfies the differential equation

$$n^2\bigl(1 - y^2\bigr) = \bigl(y'\bigr)^2\bigl(1 - z^2\bigr).$$

Since the general solution of the equation

$$\frac{y'}{\sqrt{1 - y^2}} = \pm\frac{n}{\sqrt{1 - z^2}}$$

is

$$\arccos y = \pm n \arccos z + c,$$

it follows now from $P(1) = \pm 1$ that

$$P = \pm\cos(n \arccos x) = \pm T_n(z). \qquad \square$$

Remark 4.2. Notice that the equality $T_n(-z) = (-1)^n T_n(z)$ implies that for even n the polynomials T_n and $-T_n$ are conjugate since $T_n = \alpha \circ (-T_n) \circ \alpha^{-1}$, where $\alpha(z) = -z$. For odd n however the polynomials T_n and $-T_n$ are not conjugate.

Lemma 4.3. *Let P be a polynomial and $a, b \in \mathbb{C}$. Then the set $P^{-1}_{\mathrm{odd}}\{a, b\}$ contains at least two points.*

Proof. It follows from the equality

$$2n = \sum_{\substack{z\in\mathbb{C} \\ P(z)=a}} e_z + \sum_{\substack{z\in\mathbb{C} \\ P(z)=b}} e_z$$

that the number

$$\sum_{z\in P^{-1}_{\mathrm{odd}}\{a,b\}} e_z$$

is even, implying that the number $\operatorname{card}(P^{-1}_{\mathrm{odd}}\{a, b\})$ also is even. On the other hand,

$$\operatorname{card}\left(P^{-1}_{\mathrm{odd}}\{a, b\}\right) \neq 0,$$

for otherwise $P^{-1}_{\mathrm{odd}}\{a, b\}$ contains at most $n/2 + n/2 = n$ points in contradiction with inequality (4.2). $\square$

Theorem 4.4. *Let A and B be polynomials of degree at least two such that $A \leq B$. Then A is conjugate to z^n if and only if B is conjugate to z^n. Similarly, A is conjugate to $\pm T_n$ if and only if B is conjugate to $\pm T_n$.*

Proof. Assume that B is conjugate to $\pm T_n$, and let X be a semiconjugacy from B to A. Changing B and X to $\sigma^{-1}\circ B\circ\sigma$ and $X\circ\sigma$, for a convenient polynomial σ of degree one, without loss of generality we can assume that $B=\pm T_n$. By Theorem 1.1, we have:

$$X^{-1}(K(A)) = K(B) = [-1,1]. \tag{4.5}$$

Set $m=\deg X$. Since

$$T_m^{-1}([-1,1]) = [-1,1], \tag{4.6}$$

equality (4.5) implies that

$$X^{-1}(K(A)) = T_m^{-1}([-1,1]).$$

It now follows from Theorem 3.1 that there exists a polynomial δ of degree one such that $X=\delta\circ T_m$. Therefore, changing A and X to $\delta^{-1}\circ A\circ\delta$ and $\sigma^{-1}\circ X$, we can assume that $X=T_m$. Thus, we have:

$$A\circ T_m = T_m\circ \pm T_n = (-1)^m T_n\circ T_m, \tag{4.7}$$

implying that $A=\pm T_n$.

Similarly, if $B=z^n$, then the equalities

$$X^{-1}(K(A)) = K(B) = \mathbb{D},$$

and $(z^m)^{-1}(\mathbb{D})=\mathbb{D}$ imply that $X=\delta\circ z^m$ for some polynomial δ of degree one, and arguing as above we conclude that A is conjugate to z^n.

Assume now that A is conjugate to $\pm T_n$. Without loss of generality we can assume that $A=\pm T_n$. Since $T_{n\,\mathrm{odd}}^{-1}\{-1,1\}=\{-1,1\}$, formula (4.1) implies that

$$(\pm T_n\circ X)^{-1}_{\mathrm{odd}}\{-1,1\} = X^{-1}_{\mathrm{odd}}\{-1,1\}.$$

It follows now from

$$\pm T_n\circ X = X\circ B \tag{4.8}$$

that

$$B^{-1}_{\mathrm{odd}}\left(X^{-1}_{\mathrm{odd}}\{-1,1\}\right) = X^{-1}_{\mathrm{odd}}\{-1,1\}. \tag{4.9}$$

Since by Lemma 4.3 the set $X^{-1}_{\mathrm{odd}}\{-1,1\}$ contains at least two points, this implies by Lemma 4.1 that the polynomial B is conjugate to $\pm T_n$.

Finally, if A is conjugate to z^n, we can assume that $A=z^n$, and considering zeroes of the left and the right parts of the equality

$$z^n\circ X = X\circ B,$$

we see that $B^{-1}(X^{-1}(0)) = X^{-1}(0)$. It follows now from inequality (4.2) that $X^{-1}(0)$ consists of a single point, implying easily that the polynomial B is conjugate to z^n. □

Remark 4.5. Since for even n the polynomials T_n and $-T_n$ are conjugate (see Remark 4.2), Theorem 4.4 implies that if B is conjugate to $\pm T_n$ for even n, then A and B are conjugate. On the other hand, if B is conjugate to $-T_n$ for odd n, then A is not necessarily conjugate to $-T_n$, but only to $\pm T_n$. Still, it follows from (4.7) that if B is conjugate to T_n, then A is conjugate to T_n.

Notice that Theorem 4.4 combined with Remark 4.5 implies the following corollary.

Corollary 4.6. *Let A and B be polynomials such that the conditions $A \leq B$ and $B \leq A$ hold simultaneously, and at least one of A and B is special. Then A and B are conjugate.*

4.2. Proof of Theorem 1.3

The following lemma is a well-known fact from the complex dynamics. For the reader's convenience we give a short proof based on Theorem 3.1.

Lemma 4.7. *Let A be a polynomial of degree n such that $K(A)$ is a union of circles with a common center. Then $K(A)$ is a disk, and A is conjugate to z^n. Similarly, if $K(A)$ is a segment, then A is conjugate to $\pm T_n$.*

Proof. Since for a polynomial A the complement to $K(A)$ in $\mathbb{CP}^1$ is connected (see, *e.g*., [11, Lemma 9.4]), if $K(A)$ is a union of circles with a common center, then $K(A)$ is a disk. Furthermore, changing if necessary A to a conjugate polynomial, we can assume that $K(A) = \mathbb{D}$. Thus, $A^{-1}(\mathbb{D}) = \mathbb{D}$. On the other hand, $(z^n)^{-1}(\mathbb{D}) = \mathbb{D}$, and applying to these equalities Theorem 3.1, we conclude that $A = \alpha z^n$, where $|\alpha| = 1$, implying that A is conjugate to z^n.

Similarly, if $K(A)$ is a segment, we can assume that $K(A) = [-1, 1]$, and to conclude in a similar way that A is conjugate to $\pm T_n$. $\square$

Proof of Theorem 1.3. Set $d_0 = \deg X_0$, and let $X \in \mathcal{E}(A, B)$ be a polynomial of degree d. By Theorem 1.1, we have:

$$X_0^{-1}(K(A)) = K(B), \quad X^{-1}(K(A)) = K(B).$$

Applying to these equalities Theorem 3.2 and taking into account that, by Lemma 4.7, $K(A)$ is neither a union of concentric circles nor a segment, we conclude that $X = \widetilde{A} \circ X_0$ for some polynomial $\widetilde{A}$. Substituting now this expression in (1.2) and using that $X_0 \in \mathcal{E}(A, B)$ we have:

$$A \circ \widetilde{A} \circ X_0 = \widetilde{A} \circ X_0 \circ B = \widetilde{A} \circ A \circ X_0,$$

implying that $A \circ \widetilde{A} = A \circ \widetilde{A}$.

Conversely, if A commutes with $\widetilde{A}$, then

$$A \circ (\widetilde{A} \circ X_0) = \widetilde{A} \circ A \circ X_0 = (\widetilde{A} \circ X_0) \circ B. \qquad \square$$

Theorem 1.3 implies in particular the following classification of commuting polynomials obtained by Ritt.

Theorem 4.8 ([21]). *Let A and B be commuting polynomials of degree at least two. Then, up to the change*

$$A \to \lambda \circ A \circ \lambda^{-1}, \quad B \to \lambda \circ B \circ \lambda^{-1}, \tag{4.10}$$

where λ is a polynomial of degree one, either

$$A = z^n, \quad B = \varepsilon z^m, \tag{4.11}$$

where $\varepsilon^n = \varepsilon$, or

$$A = \pm T_n, \quad B = \pm T_m, \tag{4.12}$$

or

$$A = \varepsilon_1 R^{\circ m}, \quad B = \varepsilon_2 R^{\circ n}, \tag{4.13}$$

where $R = zS(z^\ell)$ for some polynomial S, and ε_1, ε_2 are l-th roots of unity.

Proof. Assume first that A is conjugate to z^n. Without loss of generality we may assume that $A = z^n$. Applying Theorem 1.1 for $B = A$ and $X = B$, we have:

$$B^{-1}(K(A)) = K(A).$$

Since $K(A) = \mathbb{D}$, arguing as in Lemma 4.7 we conclude that $B = \varepsilon z^m$, and it follows from $A \circ B = B \circ A$ that $\varepsilon^n = \varepsilon$. If A is conjugate to $\pm T_n$, the proof is similar.

On the other hand, if A is non-special, then Theorem 1.3 implies that any $B \in \mathcal{E}(A, A)$ has the form $B = \widetilde{A} \circ R$, where R is a polynomial of the minimum possible degree in $\mathcal{E}(A, A)$. Now we can apply Theorem 1.3 again to the polynomial $\widetilde{A}$ and so on, obtaining eventually the representation $B = \mu_1 \circ R^{\circ m_1}$, where μ_1 is a polynomial of degree one commuting with A. In particular, since $A \in \mathcal{E}(A, A)$, the equality $A = \mu_2 \circ R^{\circ m_2}$ holds for some polynomial μ_2 of degree one commuting with A. Furthermore, since R commutes with $A = \mu_2 \circ R^{\circ m_2}$, the polynomial μ_2 commutes with R. This implies easily that, up to a conjugacy, $R = zS(z^\ell)$ for some polynomial S, and $\mu_2 = \varepsilon_2 z$ for some lth root of unity ε_2. Finally, since μ_1 commutes with the polynomial A, and $A = \mu_2 \circ R^{\circ m_2}$ has the form $z\widetilde{S}(z^\ell)$ for some polynomial $\widetilde{S}$, we conclude that $\mu_1 = \varepsilon_1 z$ for some lth root of unity ε_1. □

4.3. Semiconjugacies and invariant curves

It was shown in the recent paper [10] that the problem of describing semiconjugate polynomials is closely related to the problem of describing algebraic curves $\mathcal{C}$ in $\mathbb{C}^2$ invariant under maps of the form $F : (x, y) \to (f(x), g(y))$, where f, g are polynomials of degree at least two. Briefly, this relation may be summarized as follows (see [10, Proposition 2.34] for more details).

If $\mathcal{C}$ is an irreducible (f, g)-invariant curve, then its projective closure $\overline{\mathcal{C}}$ in $\mathbb{CP}^1 \times \mathbb{CP}^1$ is also (f, g)-invariant. Denote by $\bar{h}$ the restriction of F on $\overline{\mathcal{C}}$. Let $\widetilde{\mathcal{C}}$ be the desingularization of $\mathcal{C}$ and $\beta : \widetilde{\mathcal{C}} \to \overline{\mathcal{C}}$ a map biholomorphic off a finite set. Clearly, $\bar{h}$ lifts to a holomorphic map $h : \widetilde{\mathcal{C}} \to \widetilde{\mathcal{C}}$. Consider now the commutative diagram

$$\begin{array}{ccc} \widetilde{\mathcal{C}} & \xrightarrow{h} & \widetilde{\mathcal{C}} \\ \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\beta} \\ \overline{\mathcal{C}} & \xrightarrow{\bar{h}} & \overline{\mathcal{C}} \\ \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\alpha} \\ \mathbb{CP}^1 & \xrightarrow{f} & \mathbb{CP}^1, \end{array} \tag{4.14}$$

where $\alpha : \overline{\mathcal{C}} \to \mathbb{CP}^1$ is the projection map onto the first coordinate. Set $\pi = \alpha \circ \beta$. If π is a constant, then $\mathcal{C}$ is a line $z_1 = \xi$, where ξ is a fixed point of f, so assume that the degree of π is at least one. Observe that since $f^{-1}(\infty) = \infty$, the set $K = \pi^{-1}(\infty)$ and the map h satisfy the equality

$$h^{-1}(K) = K. \tag{4.15}$$

Since h is a holomorphic map between Riemann surfaces of the same genus and $\deg h = \deg f \geq 2$, it follows from the Riemann-Hurwitz formula that either $g(\widetilde{\mathcal{C}}) = 0$, or $g(\widetilde{\mathcal{C}}) = 1$ and h is unbranched. Since $\deg h \geq 2$, for unbranched h equality (4.15) is impossible. Therefore, $\widetilde{\mathcal{C}} = \mathbb{CP}^1$ and (4.15) implies easily that, up to the change $\alpha \circ h \circ \alpha^{-1}$, where α is a Möbius transformation, either $K = \infty$ and h is a polynomial, or $K = \{0, \infty\}$ and $h = z^{\pm \deg f}$. Thus,

$$f \circ \pi = \pi \circ h, \tag{4.16}$$

where either π and h are polynomials, or $h = z^{\pm \deg f}$ and π is a Laurent polynomial. The last case requires an additional investigation. The paper [10] refers (Fact 2.25) to a more general result of [9] (Theorem 10) implying that for a non-special polynomial f this possibility is excluded. Alternatively, one can use the results of [14] (*e.g.*, Theorem 6.4).

Considering in a similar way the projection onto the second coordinate, we obtain the equality

$$g \circ \rho = \rho \circ h. \tag{4.17}$$

Thus, for non-special f and g any irreducible (f,g)-invariant curve may be parametrized by some polynomials π, ρ satisfying a system given by equations (4.16), (4.17) for some polynomial h.

Notice that in a certain sense a description of (f, g)-invariant curves reduces to the case $f = g$ since the commutative diagram

$$\begin{array}{ccc} \mathbb{C}^2 & \xrightarrow{(h,h)} & \mathbb{C}^2 \\ \Big\downarrow{\scriptstyle(\pi,\rho)} & & \Big\downarrow{\scriptstyle(\pi,\rho)} \\ \mathbb{C}^2 & \xrightarrow{(f,g)} & \mathbb{C}^2 \end{array} \tag{4.18}$$

implies that any (f, g)-invariant curve is an image of an (h, h)-invariant curve under the map $(x, y) \to (\pi(x), \rho(y))$.

Theorem 1.3 allows to obtain easily the following description of (f, f)-invariant curves obtained in [10] (see Theorem 6.24 and the theorem on page 85).

Theorem 4.9. *Let f be a non-special polynomial of degree at least two, and $\mathcal{C}$ an irreducible (f, f)-invariant curve in $\mathbb{C}^2$. Then there exists a polynomial p which commutes with f such that $\mathcal{C}$ has either the form $z_1 = p(z_2)$ or $z_2 = p(z_1)$.*

Proof. If $\mathcal{C}$ is a line $z_1 = \xi$, then ξ is a fixed point of f, and the conclusion of the theorem holds for $p = \xi$. Similarly, the theorem holds if $\mathcal{C}$ is a line $z_2 = \xi$. Otherwise, as it was shown above, $\mathcal{C}$ may be parametrized by some non-constant polynomials π, ρ satisfying the system

$$f \circ \pi = \pi \circ h, \tag{4.19}$$
$$f \circ \rho = \rho \circ h \tag{4.20}$$

for some polynomial h. Furthermore, without loss of generality we may assume that there exists no polynomial w of degree greater than one such that

$$\pi = \widetilde{\pi} \circ w, \quad \rho = \widetilde{\rho} \circ w, \tag{4.21}$$

for some polynomials $\widetilde{\pi}$, $\widetilde{\rho}$. Indeed, if (4.21) holds, then applying Theorem 2.3 to the equality

$$(f \circ \widetilde{\pi}) \circ w = \widetilde{\pi} \circ (w \circ h),$$

we conclude that $w \circ h = \widetilde{h} \circ w$ for some polynomial $\widetilde{h}$, implying that we may change π to $\widetilde{\pi}$, ρ to $\widetilde{\rho}$, and h to $\widetilde{h}$.

Set $d = \mathrm{GCD}(\deg \rho, \deg \pi)$. Since f is not special, it follows from (4.19), (4.20) by Theorem 1.3 that if both ρ and π are of degree at least two, then $d > 1$, implying by Theorem 1.2 that (4.21) holds for some polynomials $\widetilde{\pi}$, $\widetilde{\rho}$ and w with $\deg w = d > 1$. Therefore, at least one of the polynomials ρ and τ is of degree one, say $\deg \rho = 1$. Then, $\mathcal{C}$ has the form $z_1 = p(z_2)$, where $p = \pi \circ \rho^{-1}$. Furthermore, equality (4.20) implies that $h = \rho^{-1} \circ f \circ \rho$, and substituting this expression into (4.19) we conclude that p commutes with f. $\square$

Proof of Theorem 1.4. For any polynomials of coprime degrees u and v the curve $\mathcal{C}_{u,v}: \ u(x) - v(y) = 0$ is irreducible (see [15, Proposition 3.1]). Furthermore, if (1.9) holds and (x_0, y_0) is a point on $\mathcal{C}_{u,v}$, then (1.9) yields the equality

$$u(f(x_0)) = t(u(x_0)) = t(v(y_0)) = v(g(y_0)),$$

implying that $(f(x_0), g(y_0))$ also is a point on $\mathcal{C}_{u,v}$.

Conversely, assume that $\mathcal{C}$ is an irreducible (f, g)-invariant curve which is not a line, and let π and ρ be polynomials parametrizing $\mathcal{C}$ and satisfying (4.16), (4.17) for some polynomial h. Then by Theorem 1.2, there exist polynomials u and v of coprime degrees such that

$$u \circ \pi = v \circ \rho.$$

Thus, any irreducible (f, g)-invariant curve $\mathcal{C}$ in $\mathbb{C}^2$ has the form $u(x) - v(y) = 0$ for some polynomials u, v of coprime degrees. Furthermore, since the polynomial

$$s = u \circ \pi = v \circ \rho$$

belongs to $\mathcal{E}(h)$ we have:

$$t \circ u \circ \pi = u \circ \pi \circ h = u \circ f \circ \pi,$$
$$t \circ v \circ \rho = v \circ \rho \circ h = v \circ g \circ \rho,$$

implying (1.9). □

A further analysis of system (1.9) using Proposition 5.4 and Proposition 5.5 proved below leads to a more precise description of (f, g)-invariant curves apparently equivalent to the one given by [10, Theorem 6.2]. Notice however that in [10] a more general case of *skew-invariant* curves and *skew-twists* between polynomials is considered, and the methods of our paper involving Julia sets seem not to be extendable to this more general situation.

4.4. Semiconjugacies between equivalent A and B

For a natural number $n > 1$ with a prime decomposition $n = p_1^{a_1} p_2^{a_2} \dots p_s^{a_s}$ set $\mathrm{rad}(n) = p_1 p_2 \dots p_s$. The following two theorems in totality provide a proof of Theorem 1.5.

Theorem 4.10. *Let A and B be polynomials of degree at least two. Then conditions $A \leq B$ and $B \leq A$ hold simultaneously if and only if $A \sim B$.*

Proof. The "if" part follows from the definition of $\sim$ (see the introduction). Furthermore, if at least one of A and B is special, then conditions $A \leq B$ and $B \leq A$ imply by Corollary 4.6 that A and B are conjugate and hence equivalent. So, we may assume that A and B are non-special.

Let Y and X be polynomials such that

$$B \underset{Y}{\leq} A, \quad A \underset{X}{\leq} B. \tag{4.22}$$

Set $n = \deg A = \deg B$. We can assume that $\deg X > 1$, $\deg Y > 1$ since otherwise A and B are conjugate and hence $A \sim B$. Since (4.22) implies that $Y \circ X$ commutes with B, Theorem 4.8 implies that

$$\operatorname{rad}(\deg X) \mid \operatorname{rad}(n). \tag{4.23}$$

In particular,

$$\operatorname{GCD}(\deg X, n) > 1. \tag{4.24}$$

Applying Theorem 2.3 to the equality

$$A \circ X = X \circ B, \tag{4.25}$$

we conclude that there exist polynomials $\widetilde{X}$, $\widetilde{B}$, and W such that

$$B = \widetilde{B} \circ W, \quad X = \widetilde{X} \circ W, \tag{4.26}$$

and $\deg W = \operatorname{GCD}(\deg X, n)$. Clearly, $B \sim W \circ \widetilde{B}$, and equalities (4.25) and (4.26) imply that

$$A \circ \widetilde{X} = \widetilde{X} \circ (W \circ \widetilde{B}). \tag{4.27}$$

Furthermore, $\deg \widetilde{X} < \deg X$, since $\deg W > 1$ by (4.24). If $\deg \widetilde{X} = 1$, then $A \sim W \circ \widetilde{B}$ since A and $W \circ \widetilde{B}$ are conjugate; hence,

$$A \sim W \circ \widetilde{B} \sim B,$$

and we are done. Otherwise, we can apply Theorem 2.3 in a similar way to equality (4.27) and so on. Since condition (4.23) ensures that the degrees of corresponding semiconjugacies decrease, we obtain in this way a finite chain of equivalences from B to A. □

Theorem 4.11. *Let A and B be polynomials of degree at least two. Then $A \sim B$ if and only if there exist polynomials X and Y such that*

$$B \circ Y = Y \circ A, \quad A \circ X = X \circ B, \tag{4.28}$$

and $Y \circ X = B^{\circ d}$ for some $d \geq 0$.

Proof. Taking into account Theorem 4.10, we only need to show that if equalities (4.28) hold, then they hold for some $\widetilde{X}$, $\widetilde{Y}$ such that $\widetilde{Y} \circ \widetilde{X} = B^{\circ d}$, $d \geq 0$. Since (4.28) implies that $Y \circ X$ commutes with B, it follows from Theorem 4.8 that either B is special, or, up to a conjugacy,

$$Y \circ X = \varepsilon_1 R^{\circ m_1}, \quad B = \varepsilon_2 R^{\circ m_2},$$

where $R = zS(z^n)$ for some polynomial S, and ε_1, ε_2 are nth roots of unity. In the first case, Corollary 4.6 implies that A and B are conjugate. Therefore, in this case

(4.28) holds for some Möbius transformations $\widetilde{Y}$ and $\widetilde{X}$ such that $\widetilde{Y} \circ \widetilde{X} = B^0$. In the second case set

$$\widetilde{X} = X \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)},$$

where $\varepsilon_3 = \varepsilon_2^{m_1}/\varepsilon_1$, and observe that the second of equalities (4.28) still holds for $\widetilde{X}$ since

$$\begin{aligned} A \circ \widetilde{X} &= A \circ X \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} = X \circ B \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} \\ &= X \circ \varepsilon_2 R^{\circ m_2} \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} = X \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} \circ \varepsilon_2 R^{\circ m_2} = \widetilde{X} \circ B. \end{aligned}$$

On the other hand, we have:

$$Y \circ \widetilde{X} = \varepsilon_1 R^{\circ m_1} \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} = \varepsilon_1 \varepsilon_3 R^{\circ m_2 m_1} = \varepsilon_2^{m_1} R^{\circ m_2 m_1} = B^{\circ m_1}. \qquad \square$$

5. Semiconjugacies for fixed B

5.1. Special factors of semiconjugacies

Lemma 5.1. *Let A and B be polynomials of degree $n \geq 2$ such that*

$$A \circ T_\ell = T_\ell \circ B, \quad l \geq 2. \tag{5.1}$$

Then $l \leq 2n$, unless $A = \pm T_n$ and $B = \pm T_n$. Similarly, if

$$A \circ z^\ell = z^\ell \circ B, \quad l \geq 2, \tag{5.2}$$

then $l \leq n$, unless $A = \alpha z^n$, $\alpha \in \mathbb{C}$, and $B = \beta z^n$, $\beta \in \mathbb{C}$.

Proof. If

$$n \leq \frac{l-1}{2}, \tag{5.3}$$

then the set

$$(T_\ell \circ B)^{-1}_{\text{odd}}\{-1, 1\} = B^{-1}_{\text{odd}}\{-1, 1\}$$

contains at most $l - 1$ points. Therefore, if equality (5.1) holds, then the set

$$(A \circ T_\ell)^{-1}_{\text{odd}}\{-1, 1\} \tag{5.4}$$

also contains at most $l - 1$ points. On the other hand, since -1 and 1 are the only finite critical values of T_n, if the set $A^{-1}_{\text{odd}}\{-1, 1\}$ contains at least one point distinct from ± 1, then set (5.4) contains at least l points. Since by Lemma 4.3 the set $A^{-1}_{\text{odd}}\{-1, 1\}$ contains at least two points, we conclude that if (5.3) holds, then

$$A^{-1}_{\text{odd}}\{-1, 1\} = \{-1, 1\}. \tag{5.5}$$

Therefore, by Lemma 4.1, $A = \pm T_n$, It follows now from (5.1) that

$$\pm T_{nl} = T_\ell \circ B,$$

implying that

$$T_\ell \circ B = \pm T_\ell \circ T_n,$$

and applying to the last equality Theorem 2.3 we see that

$$T_\ell = \pm T_\ell \circ \mu, \quad B = \mu^{-1} \circ T_n, \tag{5.6}$$

for some polynomial μ of degree one. Finally, it is easy to see, using for example the explicit formula

$$T_n = \frac{n}{2} \sum_{k=0}^{[n/2]} (-1)^k \frac{(n-k-1)!}{k!(n-2k)!} (2x)^{n-2k}, \tag{5.7}$$

that T_n has non-zero coefficients of its terms of degree n and $n-2$, and the coefficient equal zero for its term of degree $n-1$. Thus, the first of equalities (5.6) implies the equality $\mu = \pm x$.

Assume now that equality (5.2) holds and $n \leq l-1$. Then the polynomial in the right part of (5.2) has at most $l-1$ zeroes. On the other hand, since the unique finite critical value of z^ℓ is zero, it is easy to see that, unless

$$A = \alpha z^n, \quad \alpha \in \mathbb{C}, \tag{5.8}$$

the polynomial in the left part of (5.2) has at least l zeroes. Finally, (5.8) and (5.2) imply easily that $B = \beta z^n$, $\beta \in \mathbb{C}$. ☐

Theorem 5.2. *Let B be a non-special polynomial of degree $n \geq 2$, and X an element of $\mathcal{E}(B)$. Assume that $X = W_1 \circ z^\ell \circ W_2$ for some polynomials W_1, W_2 and $l \geq 1$. Then $l \leq n$. Similarly, if $X = W_1 \circ \pm T_\ell \circ W_2$, then $l \leq 2n$.*

Proof. If $X = W_1 \circ z^\ell \circ W_2$, then applying Corollary 3.4 twice we conclude that there exist polynomials $C_1$and C_2 such that the equalities

$$A \circ W_1 = W_1 \circ C_1, \quad C_1 \circ z^\ell = z^\ell \circ C_2, \quad C_2 \circ W_2 = W_2 \circ B \tag{5.9}$$

hold. Applying now Lemma 5.1 to the second equality in (5.9) we conclude that $l \leq n$, unless C_1 and C_2 are conjugate to z^n. On the other hand, in the last case the third equality in (5.9) implies by Theorem 4.4 that B is conjugate to z^n. If $X = W_1 \circ \pm T_\ell \circ W_2$, the proof is similar. ☐

Corollary 5.3. *Let B be a non-special polynomial of degree $n \geq 2$. Assume that $B^{\circ d} = W_1 \circ z^\ell \circ W_2$ for some polynomials W_1, W_2, and $l \geq 1$, $d \geq 1$. Then $l \leq n$. Similarly, if $B^{\circ d} = W_1 \circ \pm T_\ell \circ W_2$, then $l \leq 2n$.*

Proof. Direct consequence of Theorem 5.2, since $B^{\circ d}$ is a semiconjugacy from B to B. ☐

5.2. Proof of Theorem 1.6

For natural numbers n and m define $l = l(n, m)$ as the maximum number coprime with n which divides m. Thus,

$$m = lb, \tag{5.10}$$

where $\mathrm{rad}(b)|\mathrm{rad}(n)$ and $\mathrm{GCD}(n, l) = 1$. Define now $d = d(n, m)$ as the minimum number such that b in (5.10) satisfies $b \mid n^d$. The next proposition describes a general structure of elements of $\mathcal{E}(B)$ for non-special B.

Proposition 5.4. *Let B be a non-special polynomial of degree $n \geq 2$. Then any $X \in \mathcal{E}(B)$ has the form $X = \nu \circ z^{l(n,m)} \circ W$, where ν is a polynomial of degree one, and W is a compositional right factor of $B^{\circ d(n,m)}$. Furthermore, $l(n, m) < n$.*

Proof. Set $m = \deg X$, and let l, b, d be the numbers defined above. If A is a polynomial such that

$$A \circ X = X \circ B, \tag{5.11}$$

then the equality

$$A^{\circ d} \circ X = X \circ B^{\circ d}, \tag{5.12}$$

implies by Theorem 2.3 that

$$X = U \circ S, \quad B^{\circ d} = V \circ S, \tag{5.13}$$

for some polynomials U, V, S, where $\deg U = l$. Furthermore, equalities (5.11) and $X = U \circ S$ imply by Corollary 3.4 that

$$A \circ U = U \circ C \tag{5.14}$$

for some polynomial C. Since l is coprime with n, by Theorem 2.4 there exist polynomials μ, ν of degree one such that either

$$A = \nu \circ z^s R^\ell(z) \circ \nu^{-1}, \quad U = \nu \circ z^\ell \circ \mu, \quad C = \mu^{-1} \circ z^s R(z^\ell) \circ \mu,$$

where R is a polynomial, $n \geq 1$, $s \geq 0$, and $\mathrm{GCD}(s, l) = 1$, or

$$A = \nu \circ \pm T_n \circ \nu^{-1}, \quad U = \nu \circ T_\ell \circ \mu, \quad C = \mu^{-1} \circ \pm T_n \circ \mu,$$

where $\mathrm{GCD}(l, n) = 1$. In the last case however Theorem 4.4 applied to (5.11) implies that B is conjugate to T_n. Therefore, the first case must hold and hence $X = \nu \circ z^\ell \circ W$, where $W = \mu \circ S$ is a compositional right factor of $B^{\circ d}$. Moreover, since $n = rl + s$, where $r = \deg R$, the inequality $l < n$ holds whenever $r \neq 0$. On the other hand, if $r = 0$, then A is conjugate to z^n and hence B also is conjugate to z^n by Theorem 4.4. □

For a natural number $n > 1$ with a prime decomposition $n = p_1^{a_1} p_2^{a_2} \ldots p_s^{a_s}$ set $\mathrm{ord}_p(n) = a_i$, if $p = p_i$ for some i, $1 \leq i \leq s$, and $\mathrm{ord}_p\, n = 0$ otherwise.

Proposition 5.5. *If, under assumptions of Proposition* 5.4*, the polynomial X is not a polynomial in B, then $d(n,m) \leq 2\log_2 n + 3$.*

Proof. Set

$$a = n^d/b. \tag{5.15}$$

Clearly, for any prime p,

$$\mathrm{ord}_p(b) + \mathrm{ord}_p(a) = \mathrm{ord}_p(n)d,$$

implying that

$$\mathrm{ord}_p\, b = \mathrm{ord}_p(n)(d-1) + \mathrm{ord}_p(n) - \mathrm{ord}_p(a). \tag{5.16}$$

Observe that the definition of $d(n,m)$ implies that a is not divisible by n. Moreover, the number b is not divisible by n either, since otherwise equality (5.11) implies by Theorem 2.3 that X is a polynomial in B. Observe also that by Theorem 4.4 any polynomial A such that (5.11) holds is not special.

It follows from Theorem 2.3 applied to equality (5.12) that there exist polynomials N, F and Y, Z, where

$$\deg Z = l, \quad \deg Y = a,$$

such that

$$A^{\circ d} = N \circ Y, \quad X = N \circ Z,$$

and

$$Y \circ X = Z \circ B^{\circ d}. \tag{5.17}$$

Applying now Theorem 2.3 and Theorem 2.2 to the equality

$$Y \circ X = \left(Z \circ B^{d-i}\right) \circ B^i$$

for each i, $1 \leq i \leq d-1$, we obtain a collection of polynomials Y_i, X_i, W_i U_i, K_i, L_i, $1 \leq i \leq d-1$, such that

$$Y = U_i \circ Y_i, \quad Z \circ B^{\circ d-i} = U_i \circ K_i, \quad X = X_i \circ W_i, \quad B^{\circ i} = L_i \circ W_i, \tag{5.18}$$

and

$$Y_i \circ X_i = K_i \circ L_i. \tag{5.19}$$

Furthermore,

$$\deg Y_i = a_i, \quad \deg X_i = lb_i,$$

where

$$a_i = \frac{a}{\mathrm{GCD}\left(a, n^{d-i}\right)}, \quad b_i = \frac{b}{\mathrm{GCD}\left(b, n^i\right)}, \tag{5.20}$$

and there exist polynomials of degree one ν_i, σ_i, μ_i $1 \leq i \leq d-1$, such that either

$$Y_i = \nu_i \circ z^{a_i} \circ \sigma_i, \quad X_i = \sigma_i^{-1} \circ z^c R(z^{a_i}) \circ \mu_i, \tag{5.21}$$

where $R \in \mathbb{C}[z]$ and $\mathrm{GCD}(c, a_i) = 1$, or

$$Y_i = \nu_i \circ z^c R^{lb_i}(z) \circ \sigma_i, \quad X_i = \sigma_i^{-1} \circ z^{lb_i} \circ \mu_i, \tag{5.22}$$

where $R \in \mathbb{C}[z]$ and $\mathrm{GCD}(c, lb_i) = 1$, or

$$Y_i = \nu_i \circ T_{a_i} \circ \sigma_i, \quad X_i = \sigma_i^{-1} \circ T_{lb_i} \circ \mu_i, \tag{5.23}$$

where $\mathrm{GCD}(a_i, lb_i) = 1$.

Observe first that

$$a_i \geq 2^i, \quad b_i \geq 2^{d-i}. \tag{5.24}$$

Indeed, since $n \nmid a$, there exists $p \in \mathrm{rad}(n)$ such that $\mathrm{ord}_p(n) - \mathrm{ord}_p(a) > 0$. Thus, $\mathrm{ord}_p(b) > \mathrm{ord}_p(n^{d-1})$ by (5.15), and hence for any i, $1 \leq i \leq d-1$, the equality

$$\mathrm{ord}_p\left(\mathrm{GCD}(b, n^i)\right) = \mathrm{ord}_p(n)i$$

holds. It follows now from (5.20) and (5.16) that

$$\mathrm{ord}_p(b_i) = \mathrm{ord}_p(b) - \mathrm{ord}_p\left(\mathrm{GCD}(b, n^i)\right) = \mathrm{ord}_p(n)(d-1-i) + \mathrm{ord}_p(n) - \mathrm{ord}_p(a),$$

implying that

$$b_i \geq p^{\mathrm{ord}_p(n)(d-1-i)+\mathrm{ord}_p(n)-\mathrm{ord}_p(a)} \geq p^{\mathrm{ord}_p(n)(d-1-i)+1} \geq p^{(d-1-i)+1} = p^{d-i}.$$

Similarly, since $n \nmid b$, there exists $q \in \mathrm{rad}(n)$ such that $\mathrm{ord}_q(n) - \mathrm{ord}_q(b) > 0$ implying by (5.20) and (5.16) that that for any i, $1 \leq i \leq d-1$, the inequality $a_i \geq q^i$ holds. Since $p \geq 2$, $q \geq 2$, this proves (5.24).

In order to establish the required bound, observe that since

$$A^{\circ d} = N \circ U_i \circ Y_i,$$

it follows from Corollary 5.3 that if (5.21) or (5.23) holds, then $a_i \leq 2n$. On the other hand, since $X = X_i \circ W_i$, if (5.22) or (5.23) holds, then $b_i \leq lb_i \leq 2n$, by Theorem 5.2. Thus, for any i, $1 \leq i \leq d-1$, the inequality

$$\min\{a_i, b_i\} \leq 2n$$

holds. On the other hand, it follows from (5.24) that for $i_0 = \lfloor d/2 \rfloor$ the inequality

$$\min\{a_i, b_i\} \geq 2^{\lfloor d/2 \rfloor}$$

holds. Therefore, $2^{\lfloor d/2 \rfloor} \leq 2n$, implying that $2^{d/2} \leq 2\sqrt{2}n$. Thus, $d/2 \leq \log_2 n + 3/2$ and $d \leq 2\log_2 n + 3$. $\square$

Proof of Theorem 1.6. Observe first that if $X \in \mathcal{E}(B)$ is a semiconjugacy from B to A, then A is defined in a unique way since the equalities

$$A \circ X = X \circ B, \quad \widetilde{A} \circ X = X \circ B$$

imply $A \circ X = \widetilde{A} \circ X$, which in turn implies $A = \widetilde{A}$. In particular, this implies that for any $X_1, X_2 \in \mathcal{E}(B)$ such that $X_2 = \mu \circ X_1$ for some polynomial μ of degree one the corresponding polynomials $A_1, A_2 \in \mathcal{F}(B)$ are conjugate. Moreover, for any $A \in \mathcal{F}(B)$ there exists X such that

$$A \circ X = X \circ B \tag{5.25}$$

and X is not a polynomial in B, since (5.25) and $X = \widetilde{X} \circ B^{\circ s}$ imply that

$$A \circ \widetilde{X} = \widetilde{X} \circ B.$$

Finally, if $X_1, X_2 \in \mathcal{E}(B)$ and $\deg X_1 = \deg X_2$, then the corresponding polynomials in $A_1, A_2 \in \mathcal{F}(B)$ are conjugate, since Theorem 1.1 and Theorem 3.1 imply that there exists a polynomial μ of degree one such that $X_2 = \mu \circ X_1$.

Let X be an element of $\mathcal{E}(B)$ and $X = \nu \circ z^l \circ W$ its representation from Proposition 5.4. Then it follows from Proposition 5.5 that, unless X is a polynomial in B, the inequality $d \leq 2 \log_2 n + 3$ holds. Since, in addition, for the number l the inequality $l < n$ holds, this implies that up to the change $X \to \mu \circ X$, where μ is a polynomial of degree one, there exists at most a finite number of elements of $\mathcal{E}(B)$ which are not polynomials in B. Applying to these polynomials recursively Theorem 1.2 we obtain polynomials $X \in \mathcal{E}(B)$ and $A \in \mathcal{F}(B)$ which satisfy the conclusion of the theorem. □

Remark 5.6. Since the degree of the polynomial of X from Theorem 1.6 is equal to the least common multiple of degrees of all polynomials from $\mathcal{E}(B)$ which are not polynomials in B, it follows from Proposition 5.4 and Proposition 5.5 that $\deg X$ is bounded by the number $\psi(n) n^{2 \log_2 n + 3}$, where $\psi(n)$ denotes the least common multiple of all numbers less than n and coprime with n. In particular,

$$c(n) \leq (n-1)! n^{2 \log_2 n + 3}.$$

Corollary 5.7. *Let B be a polynomial of degree at least two. Then there exists at most a finite number of conjugacy classes of polynomials A such that $A \leq B$.*

Proof. If B is non-special, then the corollary follows from Theorem 1.6. For special B the corollary follows from Theorem 4.4. □

Corollary 5.8. *Each equivalence class of the relation $\sim$ contains at most a finite number of conjugacy classes.*

Proof. It follows from Corollary 5.7, since $A \sim B$ implies $A \leq B$. □

Corollary 5.9 ([22]). *Let B be a non-special polynomial of degree $n \geq 2$, and X and Y polynomials such that $Y \circ X = B^{\circ s}$ for some $s \geq 1$. Then there exist polynomials $\widetilde{X}$, $\widetilde{Y}$ and $i, j \geq 0$ such that*

$$Y = B^{\circ i} \circ \widetilde{Y}, \quad X = \widetilde{X} \circ B^{\circ j}, \quad \textit{and} \quad \widetilde{Y} \circ \widetilde{X} = B^{\circ \widetilde{s}},$$

where $\widetilde{s}$ is bounded from above by a constant which depends on n only.

Proof. Clearly, without loss of generality we may assume that X is not a polynomial in B. Since $B \circ B^{\circ d} = B^{\circ d} \circ B$, the polynomial $B^{\circ d}$ is contained in $\mathcal{E}(B)$ and hence X is contained in $\mathcal{E}(B)$ by Corollary 3.4. Furthermore, since $\operatorname{rad}(\deg X) \mid \operatorname{rad}(n)$, it follows from Proposition 5.4 and Proposition 5.5 that there exists a polynomial $\widetilde{Y}$ such that $\widetilde{Y} \circ X = B^{\circ(2\log_2 n+3)}$. Therefore, if $s > 2\log_2 n + 3$, then

$$B^{\circ s} = B^{\circ(s-2\log_2 n-3)} \circ B^{\circ(2\log_2 n+3)} = B^{\circ(s-2\log_2 n-3)} \circ \widetilde{Y} \circ X = Y \circ X,$$

implying that $Y = B^{\circ(s-2\log_2 n-3)} \circ \widetilde{Y}$. This proves the corollary, and shows that $\widetilde{s} \leq 2\log_2 n + 3$. □

Remark 5.10. The bound $\widetilde{s} \leq 2\log_2 n + 3$ in Corollary 5.9 is not optimal. It was shown in [22] that in fact $\widetilde{s} \leq \log_2(n + 2)$ and that this last bound cannot be improved. For more details we refer the reader to [22]. Notice however that for applications, similar to the ones given in [6], the actual form of the bound for $\widetilde{s}$ is not important.

References

[1] T. DINH, *Ensembles d'unicité pour les polynômes*, Ergodic Theory Dynam. Systems **22** (2002), 171–186.

[2] T. DINH, *Distribution des préimages et des points périodiques d'une correspondance polynomiale*, Bull. Soc. Math. France **133** (2005), 363–394.

[3] H. ENGSTROM, *Polynomial substitutions*, Amer. J. Math. **63** (1941), 249–255.

[4] A. EREMENKO, *Some functional equations connected with the iteration of rational functions* (Russian), Algebra i Analiz. **1** (1989), 102–116; translation in Leningrad Math. J. **1** (1990), 905–919.

[5] P. FATOU, *Sur l'iteration analytique et les substitutions permutables*, J. Math. Pures Appl. (9) **2** (1923), 343–384.

[6] D. GHIOCA, T. TUCKER and M. ZIEVE, *Linear relations between polynomial orbits*, Duke Math. J. **161** (2012), 1379–1410.

[7] H. INOU, *Extending local analytic conjugacies*, Trans. Amer. Math. Soc. **363** (2011), 331–343.

[8] G. JULIA, *Mémoire sur la permutabilité des fractions rationelles*, Ann. Sci. École Norm. Sup. **39** (1922), 131–215.

[9] A. MEDVEDEV, "Minimal sets in ACFA", Thesis (PhD) University of California, Berkeley. 2007. 96.

[10] A. MEDVEDEV and T. SCANLON, *Invariant varieties for polynomial dynamical systems*, Ann. of Math. **179** (2014), 81–177.

[11] J. MILNOR, "Dynamics in one Complex Variable", Princeton Annals in Mathematics, Vol. 160, Princeton, NJ: Princeton University Press, 2006.
[12] M. MUZYCHUK and F. PAKOVICH, *Jordan-Holder theorem for imprimitivity systems and maximal decompositions of rational functions*, Proc. Lond. Math. Soc. **102** (2011), 1–24.
[13] F. PAKOVICH, *On polynomials sharing preimages of compact sets, and related questions*, Geom. Funct. Anal. **18** (2008), 163–183.
[14] F. PAKOVICH, *Prime and composite Laurent polynomials*, Bull. Sci. Math. **133** (2009), 693–732.
[15] F. PAKOVICH, *Algebraic curves $P(x) - Q(y) = 0$ and functional equations*, Complex Var. Elliptic Equ. **56** (2011), 199–213.
[16] F. PAKOVICH, *Generalized "second Ritt theorem" and explicit solution of the polynomial moment problem*, Compos. Math. **149** (2013), 705–728.
[17] F. PAKOVICH, *On semiconjugate rational functions*, Geom. Funct. Anal. **26** (2016), 1217-1243.
[18] F. PAKOVICH, *Solution of the parametric Center Problem for Abel Equation*, J. Eur. Math. Soc. (JEMS) **19** (2017), 2343-2369.
[19] J. F. RITT, *On the iteration of rational functions*, Trans. Amer. Math. Soc. **21** (1920), 348–356.
[20] J. RITT, *Prime and composite polynomials*, Trans. Amer. Math. Soc. **23** (1922), 51–66.
[21] J. F. RITT, *Permutable rational functions*, Trans. Amer. Math. Soc. **25** (1923), 399–448.
[22] M. ZIEVE and P. MÜLLER, *On Ritt's polynomial decomposition theorem*, preprint, arXiv:0807.3578.

Department of Mathematics
Ben Gurion University of the Negev
P.O.B. 653 Beer Sheva
8410501 Israel
pakovich@math.bgu.ac.il

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1447-1459

Uniqueness of entire functions sharing a small function with linear differential polynomials

INDRAJIT LAHIRI AND SHUBHASHISH DAS

Abstract. We consider the situation when an entire function shares a small function with linear differential polynomials. Our result improves a result of H. Zhong.

Mathematics Subject Classification (2010): 30D35 (primary); 30D30 (secondary).

1. Introduction, definitions and results

Suppose that f is a meromorphic function in the complex plane $\mathbb{C}$. A meromorphic function $a = a(z)$, defined in $\mathbb{C}$, is called a small function of f if $T(r, a) = S(r, f)$, where $T(r, a)$ is Nevanlinna's characteristic function of a and $S(r, f)$ is any quantity satisfying $S(r, f) = o\{T(r, f)\}$ as $r \to \infty$ possibly outside a set of finite linear measure.

We denote by $E(a; f)$ the collection of the zeros of $f - a$, where a zero is counted according to its multiplicity. Also by $\overline{E}(a; f)$ and by $E_{1)}(a; f)$ we denote the collection of distinct zeros of $f - a$ and simple zeros of $f - a$ respectively.

Suppose that f and g are two meromorphic functions in $\mathbb{C}$ and $a = a(z)$ is a small function of f and g. We say that f and g share the small function a CM (counting multiplicities) or IM (ignoring multiplicities) if $E(a; f) = E(a; g)$ or $\overline{E}(a; f) = \overline{E}(a; g)$ respectively.

The investigation of uniqueness of an entire function sharing certain values with its derivatives was initiated by L. A. Rubel and C. C. Yang in 1977, see [6]. They proved the following result.

Theorem A ([6]). *Let f be a nonconstant entire function. If for two values a and b, $E(a; f) = E(a; f^{(1)})$ and $E(b; f) = E(b; f^{(1)})$, then $f \equiv f^{(1)}$.*

Let $f(z) = \exp(e^z) \int_0^z \exp(-e^t)(1 - e^t)dt$. Then $f^{(1)} - 1 = e^z(f - 1)$ and so $E(1; f) = E(1; f^{(1)})$. Clearly $f \not\equiv f^{(1)}$ and we see that the hypothesis of

The work of the second author is supported by CSIR fellowship.

Received November 6, 2015; accepted in revised form September 27, 2016.
Published online December 2017.

two-value sharing in Theorem A is essential. So it appeared to be an interesting problem to investigate the situation of a single value sharing by an entire function with its derivative. To this end, the first result came from G. Jank, E. Mues and L. Volkmann [3], which may be stated as follows.

Theorem B ([3]). *Let f be a nonconstant entire function. If for a nonzero constant a, $\overline{E}(a; f) = \overline{E}(a; f^{(1)})$ and $\overline{E}(a; f) \subset \overline{E}(a; f^{(2)})$, then $f \equiv f^{(1)}$.*

We easily note that the hypothesis of Theorem B is equivalent to the following: $E(a; f) = E(a; f^{(1)})$ and $\overline{E}(a; f) \subset \overline{E}(a; f^{(2)})$.

It is now a natural query whether the second order derivative can be replaced by a higher order one. H. Zhong [9] answered this query in the negative by means of the following example.

Example 1.1. Let $k(\geq 3)$ be a positive integer and $\omega(\neq 1)$ be a $(k-1)^{th}$ root of unity. If $g(z) = e^{\omega z} + \omega - 1$, then g, $g^{(1)}$ and $g^{(k)}$ share the value ω CM but neither $g \equiv g^{(1)}$ nor $g \equiv g^{(k)}$.

Accommodating the general order derivative, H. Zhong [9] proved the following result.

Theorem C ([9]). *Let f be a nonconstant entire function, $a(\neq 0)$ be a finite value and $n(\geq 1)$ be an integer. If $E(a; f) = E(a; f^{(1)})$ and $\overline{E}(a; f) \subset \overline{E}(a; f^{(n)}) \cap \overline{E}(a; f^{(n+1)})$, then $f \equiv f^{(n)}$.*

Suppose that f is a nonconstant entire function and $a_1, a_2, \ldots, a_n(\neq 0)$ are complex numbers.

Then
$$L = L(f) = a_1 f^{(1)} + a_2 f^{(2)} + \cdots + a_n f^{(n)} \tag{1.1}$$
is called a linear differential polynomial generated by f.

In 1999, P. Li [4] extended Theorem C to linear differential polynomials and proved the following result.

Theorem D ([4]). *Let f be a nonconstant entire function and L be defined by* (1.1)*. Suppose that a is a nonzero finite value. If $\overline{E}(a; f) = \overline{E}(a; f^{(1)})$ and $\overline{E}(a; f) \subset \overline{E}(a; L) \cap \overline{E}(a; L^{(1)})$, then $f \equiv f^{(1)} \equiv L$.*

In the present paper we extend Theorem C by considering shared small functions instead of shared values.

For two subsets A and B of $\mathbb{C}$, we denote by $A \Delta B$ the set $(A - B) \cup (B - A)$, which is called the symmetric difference of the sets A and B.

We refer the reader to the monograph [2] for standard definitions and notation of the value distribution theory.

Suppose that f is a meromorphic function and $a = a(z)$ is a small function of f. We denote by $n_{(2}(r, a; f)$ the number of multiple zeros of $f - a$ lying in $|z| \leq r$. The function
$$N_{(2}(r, a; f) = \int_0^r \frac{n_{(2}(t, a; f) - n_{(2}(0, a; f)}{t} dt + n_{(2}(0, a; f) \log r$$
is called the integrated counting function of multiple zeros of $f - a$.

Let $A \subset \mathbb{C}$. Then by $n_A(r, a; f)$ we denote the number of zeros of $f - a$ lying in $A \cap \{z : |z| \le r\}$. The function

$$N_A(r, a; f) = \int_0^r \frac{n_A(t, a; f) - n_A(0, a; f)}{t} dt + n_A(0, a; f) \log r$$

is called the integrated counting function of those zeros of $f - a$ that lie in A.

We now state the results of the present paper.

Theorem 1.2. *Let f be a nonconstant entire function and $a = a(z) (\not\equiv 0, \infty)$ be a small function of f such that $a^{(1)} \not\equiv a$. Suppose that $A = \overline{E}(a; f) \Delta \overline{E}(a; f^{(1)})$ and $B = \overline{E}(a; f) \backslash \{\overline{E}(a; L) \cap \overline{E}(a; L^{(1)})\}$, where L defined by* (1.1) *is nonconstant. Then $f \equiv L = \alpha e^z$, where $\alpha (\neq 0)$ is a constant, provided the following hold:*

(i) $N_{A \cup B}(r, a; f) + N_A(r, a; f^{(1)}) = S(r, f)$;
(ii) $E_{1)}(a; f) \subset \overline{E}(a; f^{(1)})$;
(iii) *each common zero of $f - a$ and $f^{(1)} - a$ has the same multiplicity.*

Putting $A = B = \emptyset$ we obtain the following corollary.

Corollary 1.3. *Let f be a nonconstant entire function and $a = a(z) (\not\equiv 0, \infty)$ be a small function of f such that $a^{(1)} \not\equiv a$. If $E(a; f) = E(a; f^{(1)})$ and $\overline{E}(a; f) \subset \overline{E}(a; L) \cap \overline{E}(a; L^{(1)})$, L being nonconstant, then $f \equiv L = \alpha e^z$, where $\alpha (\neq 0)$ is a constant and L is defined by* (1.1).

The following example shows that the hypothesis $a^{(1)} \not\equiv a$ is essential for Theorem 1.2 and Corollary 1.3.

Example 1.4. Let $f = e^z + \exp(e^z)$ and $a = e^z$. Then $a (\not\equiv 0, \infty)$ is a small function of f. Also $E(a; f) = E(a; f^{(1)}) = \emptyset$ and so $\overline{E}(a; f) \subset \overline{E}(a; L) \cap \overline{E}(a; L^{(1)})$. Clearly the conclusion of Theorem 1.2 and Corollary 1.3 does not hold.

We note that the function f of Example 1.4 is of infinite order. In the following theorem we see that the hypothesis "$a^{(1)} \not\equiv a$" can be removed from Corollary 1.3 if we consider an entire function of finite order.

Theorem 1.5. *Let f be a nonconstant entire function of finite order and $a = a(z) (\not\equiv 0, \infty)$ be a small function of f. If $E(a; f) = E(a; f^{(1)})$ and $\overline{E}(a; f) \subset \overline{E}(a; L) \cap \overline{E}(a; L^{(1)})$, then $f \equiv L = \alpha e^z$, where $\alpha (\neq 0)$ is a constant and L is defined by* (1.1).

Let f be a nonconstant meromorphic function in $\mathbb{C}$ and $a_1, a_2, \dots, a_l (\not\equiv 0)$ be small functions of f. A function of the form

$$\psi = \sum_{j=1}^{l} a_j (f)^{n_{0j}} \left(f^{(1)}\right)^{n_{1j}} \cdots \left(f^{(k)}\right)^{n_{kj}}$$

is called a differential polynomial generated by f, where $n_{ij} (i = 0, 1, \dots k;\ j = 1, 2, \dots l)$ and k are nonnegative integers.

The numbers $\gamma_\psi = \max_{1\le j\le l}\sum_{i=0}^{k} n_{ij}$ and $\Gamma_\psi = \max_{1\le j\le l}\sum_{i=0}^{k}(i+1)n_{ij}$ are respectively called the degree and weight of ψ.

ACKNOWLEDGEMENTS. The authors are thankful to the referee for valuable suggestions and observations towards the improvement of the paper.

2. Lemmas

In this section we present some necessary lemmas.

Lemma 2.1 ([1]; see also [7]). *Let f be a meromorphic function and k be a positive integer. Suppose that f is a solution of the following differential equation: $a_0w^{(k)} + a_1w^{(k-1)} + \cdots + a_kw = 0$, where $a_0(\neq 0), a_1, a_2, \ldots, a_k$ are constants. Then $T(r, f) = O(r)$. Furthermore, if f is transcendental, then $r = O(T(r, f))$.*

Lemma 2.2 ([1]). *Let f be a meromorphic function and n be a positive integer. If there exist meromorphic functions $a_0(\not\equiv 0), a_1, \ldots, a_n$ such that*

$$a_0f^n + a_1f^{n-1} + \cdots + a_{n-1}f + a_n \equiv 0,$$

then

$$m(r, f) \le nT(r, a_0) + \sum_{j=1}^{n} m(r, a_j) + (n-1)\log 2.$$

Lemma 2.3 ([5]; see also [8, page 28]). *Let f be a nonconstant meromorphic function. If*

$$R(f) = \frac{a_0f^p + a_1f^{p-1} + \cdots + a_p}{b_0f^q + b_1f^{q-1} + \cdots + b_q}$$

is an irreducible rational function in f with the coefficients being small functions of f and $a_0b_0 \not\equiv 0$, then

$$T(r, R(f)) = \max\{p, q\}T(r, f) + S(r, f).$$

Lemma 2.4. *Let $f, a_0, a_1, \ldots, a_p, b_0, b_1, \ldots, b_q$ be meromorphic functions. If*

$$R(f) = \frac{a_0f^p + a_1f^{p-1} + \cdots + a_p}{b_0f^q + b_1f^{q-1} + \cdots + b_q} \quad (a_0b_0 \not\equiv 0),$$

then

$$T(r, R(f)) = O\left(T(r, f) + \sum_{i=0}^{p} T(r, a_i) + \sum_{j=0}^{q} T(r, b_j)\right).$$

Proof. The lemma follows from the first fundamental theorem and the properties of the characteristic function. □

Lemma 2.5 ([2, page 68]). *Let f be a transcendental meromorphic function and $f^n P(z) = Q(z)$, where $P(z), Q(z)$ are differential polynomials generated by f and the degree of Q is at most n. Then $m(r, P) = S(r, f)$.*

Lemma 2.6 ([2, page 69]). *Let f be a nonconstant meromorphic function and*

$$g(z) = f^n(z) + P_{n-1}(z),$$

where $P_{n-1}(z)$ is a differential polynomial generated by f and of degree at most $n-1$.

If $N(r, \infty; f) + N(r, 0; g) = S(r, f)$, then $g(z) = h^n(z)$, where $h(z) = f(z) + \frac{a(z)}{n}$ and $h^{n-1}(z)a(z)$ is obtained by substituting $h(z)$ for $f(z)$, $h^{(1)}(z)$ for $f^{(1)}(z)$ etc. *in the terms of degree $n-1$ in $P_{n-1}(z)$.*

Let us note the special case, where $P_{n-1}(z) = a_0(z) f^{n-1} +$ terms of degree $n-2$ at most. Then $h^{n-1}(z)a(z) = a_0(z)h^{n-1}(z)$ and so $a(z) = a_0(z)$. Hence $g(z) = \left(f(z) + \frac{a_0(z)}{n}\right)^n$.

Lemma 2.7 ([2, page 47]). *Let f be a nonconstant meromorphic function and a_1, a_2, a_3 be distinct small functions of f. Then*

$$T(r, f) \leq \overline{N}(r, 0; f - a_1) + \overline{N}(r, 0; f - a_2) + \overline{N}(r, 0; f - a_3) + S(r, f).$$

We note that in Lemma 2.7 a_1, a_2, a_3 are allowed to be constants, and one of them may even be ∞.

3. Proofs of the theorems

Proof of Theorem 1.2. Let $\lambda = \frac{f^{(1)} - a}{f - a}$ and $g = f - a$. Then

$$g^{(1)} = \lambda g + a - a^{(1)} = \lambda_1 g + \mu_1, \tag{3.1}$$

where $\lambda_1 = \lambda$ and $\mu_1 = a - a^{(1)} = b$, say.

Differentiating (3.1) and using (3.1) repeatedly we get

$$g^{(k)} = \lambda_k g + \mu_k, \tag{3.2}$$

where $\lambda_{k+1} = \lambda_k^{(1)} + \lambda_1 \lambda_k$ and $\mu_{k+1} = \mu_k^{(1)} + \mu_1 \lambda_k$ for $k = 1, 2, \ldots$

We now divide the proof into two parts.

Part I

We prove that $T(r,\lambda)=S(r,f)$. If λ is constant, then obviously $T(r,\lambda)=S(r,f)$. So we suppose that λ is nonconstant. By the hypothesis (i), (ii) and (iii) we get

$$N(r,0;\lambda)+N(r,\infty;\lambda)\le N_A(r,0;f-a)+N_A\big(r,0;f^{(1)}-a\big)=S(r,f). \tag{3.3}$$

Putting $k=1$ in $\lambda_{k+1}=\lambda_k^{(1)}+\lambda_1\lambda_k$ we get $\lambda_2=\lambda^2+d_1\lambda$, where $d_1=\frac{\lambda^{(1)}}{\lambda}$. Again putting $k=2$ in $\lambda_{k+1}=\lambda_k^{(1)}+\lambda_1\lambda_k$ we have $\lambda_3=\lambda_2^{(1)}+\lambda_1\lambda_2=\lambda^3+3d_1\lambda^2+d_2\lambda$, where $d_2=d_1^2+d_1^{(1)}$. Similarly $\lambda_4=\lambda_3^{(1)}+\lambda_1\lambda_3=\lambda^4+6d_1\lambda^3+(6d_1^2+3d_1^{(1)}+d_2)\lambda^2+(d_2^{(1)}+d_1d_2)\lambda$. Therefore, in general, we get for $k\ge 2$

$$\lambda_k=\lambda^k+\sum_{j=1}^{k-1}\alpha_j\lambda^j, \tag{3.4}$$

where $T(r,\alpha_j)=O(\overline{N}(r,0;\lambda)+\overline{N}(r,\infty;\lambda))+S(r,\lambda)=S(r,f)$ for $j=1,2,\dots,k-1$.

Again putting $k=1$ in $\mu_{k+1}=\mu_k^{(1)}+\mu_1\lambda_k$ we get $\mu_2=\mu_1^{(1)}+\mu_1\lambda_1=b\lambda+b^{(1)}$. Also putting $k=2$ in $\mu_{k+1}=\mu_k^{(1)}+\mu_1\lambda_k$ we obtain by (3.4), $\mu_3=b\lambda^2+(b^{(1)}+bd_1+\alpha_1)\lambda+b^{(2)}$. Similarly $\mu_4=b\lambda^3+(2bd_1+b^{(1)}+b\alpha_2)\lambda^2+(b^{(2)}+2b^{(1)}d_1+bd^{(1)}+\alpha_1^{(1)}+bd_1^2+\alpha_1d_1+b\alpha_1)\lambda+b^{(3)}$. Therefore, in general, for $k\ge 2$

$$\mu_k=\sum_{j=1}^{k-1}\beta_j\lambda^j+b^{(k-1)}, \tag{3.5}$$

where $T(r,\beta_j)=O(\overline{N}(r,0;\lambda)+\overline{N}(r,\infty;\lambda))+S(r,\lambda)=S(r,f)$ for $j=1,2,\dots,k-1$ and $\beta_{k-1}=b$.

Let z_0 be a zero of $f-a$ and $f^{(1)}-a$ with multiplicity $q(\ge 2)$. Then z_0 is a zero of $f^{(1)}-a^{(1)}$ with multiplicity $q-1$. Hence z_0 is a zero of $b=a-a^{(1)}=(f^{(1)}-a^{(1)})-(f^{(1)}-a)$ with multiplicity $q-1$. Since $q\le 2(q-1)$, we have $N_{(2}(r,a;f)\le 2N(r,0;b)+N_A(r,a;f)=S(r,f)$.

We first suppose that either $n\ge 2$ or $n=1$ and $a_1\ne 1$. Let

$$\psi=\frac{(a-L(a))\big(f^{(1)}-a^{(1)}\big)-\big(a-a^{(1)}\big)(L-L(a))}{f-a}. \tag{3.6}$$

From (3.6) we get $N(r,\psi)\le N_{(2}(r,a;f)+N_{A\cup B}(r,a;f)+(n+1)N(r,\infty;a)=S(r,f)$ and so $T(r,\psi)=S(r,f)$ because $m(r,\psi)=S(r,f)$.

Using (3.2), (3.4) and (3.5) we get

$$\begin{aligned}L(g)&=a_1g^{(1)}+\sum_{k=2}^{n}a_kg^{(k)}\\&=a_1(\lambda g+b)+\sum_{k=2}^{n}a_k\left(\lambda^k+\sum_{j=1}^{k-1}\alpha_j\lambda^j\right)g+\sum_{k=2}^{n}a_k\left(\sum_{j=1}^{k-1}\beta_j\lambda^j+b^{(k-1)}\right).\end{aligned}$$

Therefore from (3.6) we get

$$\begin{aligned} 0 \equiv & \left\{\psi + a_1 b\lambda + \sum_{k=2}^{n} a_k b\left(\lambda^k + \sum_{j=1}^{k-1}\alpha_j\lambda^j\right) - \lambda(a - L(a))\right\} g \\ & + b\left\{ba_1 + \sum_{k=2}^{n} a_k\left(\sum_{j=1}^{k-1}\beta_j\lambda^j + b^{(k-1)}\right) - (a - L(a))\right\}. \end{aligned} \tag{3.7}$$

If $\psi + a_1b\lambda + \sum_{k=2}^{n} a_k b(\lambda^k + \sum_{j=1}^{k-1}\alpha_j\lambda^j) - \lambda(a - L(a)) \equiv 0$, then by Lemma 2.2 we get $m(r, \lambda) = S(r, f)$. Therefore by (3.3) we have $T(r, \lambda) = S(r, f)$.

Suppose that $\psi + a_1b\lambda + \sum_{k=2}^{n} a_k b(\lambda^k + \sum_{j=1}^{k-1}\alpha_j\lambda^j) - \lambda(a - L(a)) \not\equiv 0$. Then from (3.7) we get

$$g = -\frac{b\left\{ba_1 + \sum\limits_{k=2}^{n} a_k\left(\sum\limits_{j=1}^{k-1}\beta_j\lambda^j + b^{(k-1)}\right) - (a - L(a))\right\}}{\psi + a_1b\lambda + \sum\limits_{k=2}^{n} a_k b\left(\lambda^k + \sum\limits_{j=1}^{k-1}\alpha_j\lambda^j\right) - \lambda(a - L(a))}. \tag{3.8}$$

From (3.8) we get by Lemma 2.4, $T(r, g) = O(T(r, \lambda)) + S(r, f)$ and so $T(r, f) = O(T(r, \lambda)) + S(r, f)$. This implies that $S(r, f)$ is replaceable by $S(r, \lambda)$.

Also, from (3.8) we see that g is a rational function in λ, which can be made irreducible. We now put

$$g = \frac{P_s(\lambda)}{Q_{s+1}(\lambda)}, \tag{3.9}$$

where $P_s(\lambda)$ and $Q_{s+1}(\lambda)$ are relatively prime polynomials in λ of respective degrees s and $s + 1$. The coefficients of both the polynomials are small functions of λ. Without loss of generality we assume that $Q_{s+1}(\lambda)$ is a monic polynomial. We further note that the counting function of the common zeros of $P_s(\lambda)$ and $Q_{s+1}(\lambda)$, if any, is $S(r, \lambda)$, because $P_s(\lambda)$ and $Q_{s+1}(\lambda)$ are relatively prime and the coefficients are small functions of λ.

Since $N(r, \infty; g) = S(r, f) = S(r, \lambda)$, we see from (3.9) that $N(r, 0; Q_{s+1}(\lambda)) = S(r, \lambda)$. Also by (3.3) we know that $N(r, \infty; \lambda) = S(r, f) = S(r, \lambda)$. So by Lemma 2.6 we get

$$Q_{s+1}(\lambda) = \left(\lambda + \frac{c}{s+1}\right)^{s+1}, \tag{3.10}$$

where c is the coefficient of λ^s in $Q_{s+1}(\lambda)$.

If $c \not\equiv 0$, then by Lemma 2.7 we obtain

$$\begin{aligned} T(r, \lambda) &\le \overline{N}(r, 0; \lambda) + \overline{N}(r, \infty; \lambda) + \overline{N}\left(r, -\frac{c}{s+1}; \lambda\right) + S(r, \lambda) \\ &= \overline{N}(r, 0; Q_{s+1}(\lambda)) + S(r, \lambda) \\ &= S(r, \lambda), \end{aligned}$$

a contradiction. Therefore $c \equiv 0$ and we get from (3.9) and (3.10)

$$g = \frac{P_s(\lambda)}{\lambda^{s+1}}. \tag{3.11}$$

Differentiating (3.11) we obtain

$$g^{(1)} = d_1 \frac{\lambda P_s^{(1)}(\lambda) - (s+1)P_s(\lambda)}{\lambda^{s+1}},$$

where $d_1 = \frac{\lambda^{(1)}}{\lambda}$ and $T(r, d_1) = O(\overline{N}(r, 0; \lambda) + \overline{N}(r, \infty; \lambda)) + m(r, d_1) = S(r, f) + S(r, \lambda) = S(r, \lambda)$. So by Lemma 2.3 we have

$$T\left(r, g^{(1)}\right) = (s+1-p)T(r, \lambda) + S(r, \lambda), \tag{3.12}$$

for some integer $p, 0 \leq p \leq s$.

Again since $g^{(1)} = \lambda g + b$, where $b = a - a^{(1)} \not\equiv 0$, we get from (3.11)

$$g^{(1)} = \frac{P_s(\lambda)}{\lambda^s} + b$$

and so by Lemma 2.3 we have

$$T\left(r, g^{(1)}\right) = (s-p)T(r, \lambda) + S(r, \lambda), \tag{3.13}$$

where p is same as in (3.12). Now from (3.12) and (3.13) we get $T(r, \lambda) = S(r, \lambda)$, a contradiction.

Next we suppose that $n = 1$ and $a_1 = 1$. Let

$$\phi = \frac{\left(a - L^{(1)}(a)\right)(L - L(a)) - (a - L(a))\left(L^{(1)} - L^{(1)}(a)\right)}{f - a}.$$

Since in this case $L = f^{(1)}$, we get

$$\begin{aligned} \phi &= \frac{\left(a - a^{(2)}\right)\left(f^{(1)} - a^{(1)}\right) - \left(a - a^{(1)}\right)\left(f^{(2)} - a^{(2)}\right)}{f - a} \\ &= \frac{\left(a - a^{(2)}\right)g^{(1)} - bg^{(2)}}{g}. \end{aligned} \tag{3.14}$$

By the hypothesis we have $T(r, \phi) = S(r, f)$. Using (3.2), (3.4), (3.5) and (3.14) we get

$$\left\{b\lambda^2 + (\alpha_1 b - a + a^{(2)})\lambda + \phi\}g + b\{b^{(1)} + \beta_1 \lambda + a^{(2)} - a\right\} \equiv 0. \tag{3.15}$$

Following the similar argument of the preceding case and using (3.15) we can show that $m(r, \lambda) = S(r, f)$. So by (3.3) we have $T(r, \lambda) = S(r, f)$. This completes the proof of Part I.

Part II

First we verify that

$$T(r, f) \leq 3\overline{N}(r, 0; f - a) + S(r, f). \tag{3.16}$$

By the first fundamental theorem we get

$$\begin{aligned}
T(r, f) &= T(r, f - a) + S(r, f) \\
&= T\left(r, \frac{1}{f - a}\right) + S(r, f) \\
&= N\left(r, \frac{1}{f - a}\right) + m\left(r, \frac{1}{f - a}\right) + S(r, f) \\
&\leq N\left(r, \frac{1}{f - a}\right) + m\left(r, \frac{1}{f^{(1)} - a^{(1)}}\right) + S(r, f) \\
&= N\left(r, \frac{1}{f - a}\right) + T\left(r, f^{(1)}\right) - N\left(r, \frac{1}{f^{(1)} - a^{(1)}}\right) + S(r, f).
\end{aligned}$$

Now by Lemma 2.7 we get from above

$$\begin{aligned}
T(r, f) \leq N(r, 0; f - a) &+ \overline{N}\left(r, 0; f^{(1)} - a\right) + \overline{N}\left(r, 0; f^{(1)} - a^{(1)}\right) \\
&- N\left(r, 0; f^{(1)} - a^{(1)}\right) + S(r, f).
\end{aligned} \tag{3.17}$$

Let us denote by $N^{p}_{(k}(r, 0; F)$ the counting function of zeros of F with multiplicities not less than k and a zero of multiplicity $q(\geq k)$ is counted $q - p$ times, where $p \leq k$.

Now

$$\begin{aligned}
&N(r, 0; f - a) + \overline{N}\left(r, 0; f^{(1)} - a^{(1)}\right) - N\left(r, 0; f^{(1)} - a^{(1)}\right) \\
&= \overline{N}(r, 0; f - a) + N^{1}_{(2}(r, 0; f - a) - N^{1}_{(2}\left(r, 0; f^{(1)} - a^{(1)}\right) \\
&= \overline{N}(r, 0; f - a) + \overline{N}_{(2}(r, 0; f - a) + N^{2}_{(3}(r, 0; f - a) - N^{1}_{(2}\left(r, 0; f^{(1)} - a^{(1)}\right) \\
&\leq 2\overline{N}(r, 0; f - a) + N^{1}_{(2}\left(r, 0; f^{(1)} - a^{(1)}\right) - N^{1}_{(2}\left(r, 0; f^{(1)} - a^{(1)}\right) + S(r, f) \\
&= 2\overline{N}(r, 0; f - a) + S(r, f),
\end{aligned}$$

where $\overline{N}_{(2}(r, 0; f - a)$ is the integrated counting function of distinct multiple zeros of $f - a$.

Therefore from (3.17) we get

$$T(r, f) \leq 2\overline{N}(r, 0; f - a) + \overline{N}\left(r, 0; f^{(1)} - a\right) + S(r, f). \tag{3.18}$$

Since

$$\overline{N}(r,0;\, f^{(1)}-a) \le \overline{N}(r,0;\, f-a)+N_A(r,0;\, f^{(1)}-a) = \overline{N}(r,0;\, f-a)+S(r,\, f),$$

(3.16) is obtained from (3.18).

Since $T(r,\lambda) = S(r,\, f)$, we see that $T(r,\lambda_k) + T(r,\mu_k) = S(r,\, f)$ for $k = 1, 2, \dots$, where λ_k and μ_k are defined in (3.2). Now

$$\begin{aligned} L &= \sum_{k=1}^{n} a_k f^{(k)} = \sum_{k=1}^{n} a_k g^{(k)} + L(a) \\ &= \left(\sum_{k=1}^{n} a_k \lambda_k\right) g + \sum_{k=1}^{n} a_k \mu_k + L(a) = \xi g + \eta, \text{ say.} \end{aligned} \tag{3.19}$$

Clearly $T(r,\xi) + T(r,\eta) = S(r,\, f)$. Differentiating (3.19) we get

$$L^{(1)} = \xi^{(1)} g + \xi g^{(1)} + \eta^{(1)}. \tag{3.20}$$

Let $z_0 \notin A \cup B$, be a zero of $g = f - a$. Then from (3.19) and (3.20) we get $a(z_0) - \eta(z_0) = 0$ and $\xi(z_0)(a(z_0) - a^{(1)}(z_0)) + \eta^{(1)}(z_0) - a(z_0) = 0$.

If $a(z) - \eta(z) \not\equiv 0$, we get

$$\overline{N}(r,0;\, f-a) \le N_{A\cup B}(r,0;\, f-a) + N(r,0;\, a-\eta) + S(r,\, f) = S(r,\, f),$$

which contradicts (3.16). Therefore

$$a(z) \equiv \eta(z). \tag{3.21}$$

Again if $\xi(z)(a(z) - a^{(1)}(z)) + \eta^{(1)}(z) - a(z) \not\equiv 0$, we get

$$\begin{aligned} \overline{N}(r,0;\, f-a) \le N_{A\cup B}(r,0;\, f-a) &+ N\left(r,0;\, \xi\left(a-a^{(1)}\right) + \eta^{(1)} - a\right) \\ &+ S(r,\, f) = S(r,\, f), \end{aligned}$$

which contradicts (3.16). Therefore

$$\xi(z)\left(a(z) - a^{(1)}(z)\right) + \eta^{(1)}(z) - a(z) \equiv 0. \tag{3.22}$$

Since $a(z) \not\equiv a^{(1)}(z)$, from (3.21) and (3.22) we get $\xi(z) \equiv 1$. Hence from (3.19) and (3.21) we get $L \equiv g + a \equiv f$.

By actual calculation we see that $\lambda_2 = \lambda^2 + \lambda^{(1)}$ and $\lambda_3 = \lambda^3 + 3\lambda\lambda^{(1)} + \lambda^{(2)}$. We now verify, in general, that

$$\lambda_k = \lambda^k + P_{k-1}[\lambda], \tag{3.23}$$

where $P_{k-1}[\lambda]$ is a differential polynomial in λ with constant coefficients such that the degree $\gamma_{P_{k-1}} \leq k-1$ and the weight $\Gamma_{P_{k-1}} \leq k$. Also each term of $P_{k-1}[\lambda]$ contains some derivative of λ.

Let (3.23) be true. Then

$$\begin{aligned}\lambda_{k+1} &= \lambda_k^{(1)} + \lambda_1\lambda_k \\ &= \left(\lambda^k + P_{k-1}[\lambda]\right)^{(1)} + \lambda\left(\lambda^k + P_{k-1}[\lambda]\right) \\ &= \lambda^{k+1} + P_k[\lambda],\end{aligned}$$

where we note that differentiation does not increase the degree of a differential polynomial but increases its weight by 1. So (3.23) is verified by mathematical induction.

Since $\xi(z) \equiv 1$, by (3.19) and (3.23) we get

$$\sum_{k=1}^{n} a_k\lambda^k + \sum_{k=1}^{n} a_k P_{k-1}[\lambda] \equiv 1. \tag{3.24}$$

By the hypotheses (ii) and (iii) we see that λ has no simple pole. Let z_0 be a pole of λ with multiplicity $p(\geq 2)$. Then z_0 is a pole of $\sum_{k=1}^{n} a_k\lambda^k$ with multiplicity np and it is a pole of $\sum_{k=1}^{n} a_k P_{k-1}[\lambda]$ with multiplicity at most $(n-1)p+1$. Since $np > (n-1)p+1$, it follows that z_0 is a pole of the left hand side of (3.24) with multiplicity np, which is impossible. So λ is an entire function. If λ is transcendental, then by Lemma 2.5 we get from (3.24) that $T(r,\lambda) = S(r,\lambda)$, a contradiction. If λ is a polynomial of degree $d(\geq 1)$, then the left hand side of (3.24) is a polynomial of degree nd, which is also a contradiction. Therefore λ is a constant and so from (3.23) we get $\lambda_k = \lambda^k$ for $k = 1, 2, \ldots$. We suppose that $\lambda \neq 1$.

Since $L \equiv f$, we see by Lemma 2.1 that $T(r, f) = O(r)$ and so $T(r,a) = o(r)$, because a is a small function of f.

Since λ is a constant, by a simple calculation we get $\mu_k = \sum_{j=0}^{k-1} b^{(k-1-j)}\lambda^j$ for $k = 1, 2, \ldots$. Therefore from (3.19) we have

$$\eta = L(a) + \sum_{k=1}^{n} a_k\mu_k = L(a) + \sum_{k=1}^{n} a_k \left(\sum_{j=0}^{k-1} b^{(k-1-j)}\lambda^j\right). \tag{3.25}$$

From (3.21) and (3.25) we see that $a = a(z)$ is an entire function. Since $T(r,a) = o(r)$, by Lemma 2.1, (3.21) and (3.25) we observe that $a = a(z)$ is a polynomial.

Now from (3.1) we get

$$f^{(1)} = \lambda f + (1-\lambda)a = \lambda f + P_l, \tag{3.26}$$

where P_l is a polynomial of degree l.

Differentiating (3.26) $l+1$ times we get $f^{(l+2)} = \lambda f^{(l+1)}$ and so $f^{(l+1)} = \beta e^{\lambda z}$, where $\beta(\neq 0)$ is a constant. Now integrating $f^{(l+1)} = \beta e^{\lambda z}$, $l+1$ times we get

$$f = \frac{\beta}{\lambda^{l+1}} e^{\lambda z} + Q_t,$$

where Q_t is a polynomial of degree $t(\leq l)$.

Since $\xi(z) \equiv 1$ and $\lambda_k = \lambda^k$, we have $\sum_{k=1}^{n} a_k \lambda^k = 1$. Hence

$$L = \sum_{k=1}^{n} a_k f^{(k)} = \left(\sum_{k=1}^{n} a_k \lambda^k\right) \frac{\beta}{\lambda^{l+1}} e^{\lambda z} + \sum_{k=1}^{n} a_k Q_t^{(k)} = \frac{\beta}{\lambda^{l+1}} e^{\lambda z} + \sum_{k=1}^{n} a_k Q_t^{(k)}.$$

Since $f \equiv L$, we have $Q_t \equiv \sum_{k=1}^{n} a_k Q_t^{(k)}$ and this implies $Q_t \equiv 0$. Therefore $f = \frac{\beta}{\lambda^{l+1}} e^{\lambda z}$ and from (3.26) we get $\frac{\beta}{\lambda^{l}} e^{\lambda z} = \frac{\beta}{\lambda^{l}} e^{\lambda z} + (1-\lambda)a$, which is impossible as $\lambda \neq 1$ and $a \not\equiv 0$. Hence $\lambda = 1$ and so from (3.26) we obtain $f \equiv L = \alpha e^{z}$, where $\alpha(\neq 0)$ is a constant. This proves the theorem. □

Proof of Theorem 1.5. Let $a \equiv a^{(1)}$. Then $a = \beta e^{z}$, where $\beta(\neq 0)$ is a constant. Since $E(a; f) = E(a; f^{(1)})$ and f is of finite order, there exists a polynomial h such that $\frac{f^{(1)}-a}{f-a} = e^{h}$ and so $\frac{f^{(1)}-a^{(1)}}{f-a} = e^{h}$. Integrating we get $f = a + \gamma e^{\nu}$, where $\gamma(\neq 0)$ is a constant and $\nu^{(1)}(z) = e^{h(z)}$. Since f and so a are of finite order, we see that ν is a polynomial. Again $E(a; f) = E(a; f^{(1)}) = \emptyset$ and $f^{(1)} = a + \gamma \nu^{(1)} e^{\nu}$ imply that $\nu^{(1)}$ is a constant. So $\nu = cz + d$, where $c(\neq 0)$ and d are constants. Therefore $f = a + \gamma e^{cz+d}$ and this contradicts the fact that $a = \beta e^{z}$ is a small function of f. Hence $a \not\equiv a^{(1)}$ and the theorem follows from Corollary 1.3. This proves the theorem. □

References

[1] J. CHANG and M. FANG, *Entire functions that share a small function with their derivatives*, Complex Var. Theory Appl. **49** (2004), 871–895.

[2] W. K. HAYMAN, “Meromorphic Functions”, The Clarendon Press, Oxford, 1964.

[3] G. JANK, E. MUES and L. VOLKMANN, *Meromorphe Functionen, die mit ihrer ersten und zweiten Ableitung einen endlichen Wert teilen*, Complex Var. Theory Appl. **6** (1986), 51–71.

[4] P. LI, *Entire functions that share one value with their linear differential polynomials*, Kodai Math. J. **22** (1999), 446–457.

[5] A. MOHOŃKO, *The Nevanlinna characteristics of certain meromorphic functions*, Funktsional. Anal. i Prilozhen **14** (1971), 83–87.

[6] L. A. RUBEL and C. C. YANG, *Values shared by an entire function and its derivative*, In: “Complex Analysis, Kentucky, 1976”, Lecture Notes in Math., Vol. 599, Springer, 1977, 101–103.

[7] G. VALIRON, “Lectures on General Theory of Integral Functions”, Edouard Privat, Toulouse,1923.

[8] C. C. YANG and H. X. YI, "Uniqueness Theory of Meromorphic Functions", Science Press and Kluwer Academic Publishers, 2003.
[9] H. ZHONG, *Entire functions that share one value with their derivatives*, Kodai Math. J. **18** (1995), 250–259.

Department of Mathematics
University of Kalyani
West Bengal 741235, India
ilahiri@hotmail.com
dshubhashish.90@gmail.com

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1461-1489

Twisted cohomology of arrangements of lines and Milnor fibers

MARIO SALVETTI AND MATTEO SERVENTI

Abstract. Let $\mathcal{A}$ be an arrangement of affine lines in $\mathbb{C}^2$, with complement $\mathcal{M}(\mathcal{A})$. The (co)homology of $\mathcal{M}(\mathcal{A})$ with twisted coefficients is strictly related to the cohomology of the Milnor fibre associated to the conified arrangement, endowed with the geometric monodromy. Although several partial results are known, even the first Betti number of the Milnor fiber is not understood. We give here a vanishing conjecture for the first homology, which is of a different nature with respect to the known results. Let Γ be the graph of *double points* of $\mathcal{A}$: we conjecture that if Γ is connected then the geometric monodromy acts trivially on the first homology of the Milnor fiber (so that the first Betti number is combinatorially determined in this case). This conjecture depends only on the combinatorics of $\mathcal{A}$. We prove it in some cases with stronger hypothesis.

In the final parts, we introduce a new description in terms of the group given by the quotient of the commutator subgroup of $\pi_1(\mathcal{M}(\mathcal{A}))$ by the commutator of its *length-zero subgroup*. We use that to deduce some new interesting cases of a-monodromicity, including a proof of the conjecture under some extra conditions.

Mathematics Subject Classification (2010): 55N25 (primary); 57M05 (secondary).

1. Introduction

Let $\mathcal{A} := \{\ell_1, \ldots, \ell_n\}$ be an arrangement of affine lines in $\mathbb{C}^2$, with complement $\mathcal{M}(\mathcal{A})$. Let $\mathcal{L}$ be a rank-1 local system on $\mathcal{M}(\mathcal{A})$, which is defined by a unitary commutative ring R and an assignment of an invertible element $t_i \in R^*$ for each line $\ell_i \in \mathcal{A}$. Equivalently, $\mathcal{L}$ is defined by a module structure on R over the fundamental group of $\mathcal{M}(\mathcal{A})$ (such structure factorizes through the first homology of $\mathcal{M}(\mathcal{A})$). By "coning" $\mathcal{A}$ one obtains a three-dimensional central arrangement, with complement fibering over $\mathbb{C}^*$. The Milnor fiber F of such fibration is a surface of degree $n+1$, endowed with a natural monodromy automorphism of order $n+1$. It

Partially supported by INdAM and by Università di Pisa under the "PRA – Progetti di Ricerca di Ateneo" (Institutional Research Grants) – Project no. PRA_2016_67 "Geometria, Algebra e Combinatoria di Spazi di Moduli e Configurazioni".

Received August 7, 2015; accepted in revised form September 27, 2016.
Published online December 2017.

is well known that the trivial (co)homology of F with coefficients in a commutative ring A, as a module over the monodromy action, is obtained by the (co)homology of $\mathcal{M}(\mathcal{A})$ with coefficients in $R := A[t^{\pm 1}]$, where here the structure of R as a $\pi_1(\mathcal{M}(\mathcal{A}))$-module is given by taking all the t_i's equal to t and the monodromy action corresponds to t-multiplication. For reflection arrangements, relative to a Coxeter group $\mathbf{W}$, many computations were done, especially for the *orbit space* $\mathcal{M}_{\mathbf{W}}(\mathcal{A}) := \mathcal{M}(\mathcal{A})/\mathbf{W}$, which has an associated Milnor fiber $F_{\mathbf{W}} := F/\mathbf{W}$: in this case we know a complete answer for $R = \mathbb{Q}[t^{\pm 1}]$, for all groups of finite type (see [11,12,21]), and for some groups of affine type [6–8] (based on the techniques developed in [13,30]). For $R = \mathbb{Z}[t^{\pm 1}]$ a complete answer is known in case A_n (see [5]). Some results are known for (non quotiented) reflection arrangements (see [25,31]). A big amount of work in this case has been done on related questions, when $R = \mathbb{C}$, in that case the t_i's being non-zero complex numbers, trying to understand the jump-loci (in $(\mathbb{C}^*)^n$) of the cohomology (see for example [9,10,16,18,24,34]).

Some algebraic complexes computing the twisted cohomology of $\mathcal{M}(\mathcal{A})$ are known (see for example the above cited papers). In [22], the minimal cell structure of the complement which was constructed in [33] (see [15,28]) was used to find an algebraic complex which computes the twisted cohomology, in the case of real defined arrangements (see also [23]). The form of the boundary maps depends not only on the lattice of the intersections associated to $\mathcal{A}$ but also on its *oriented matroid:* for each singular point P of multiplicity m there are $m-1$ generators in dimension 2 whose boundary has non vanishing components along the lines contained in the "cone" of P and passing above P.

Many of the specific examples of arrangements with non-trivial cohomology (*i.e.*, having non-trivial monodromy) which are known are based on the theory of *nets* and *multinets* (see [19]): there are relatively few arrangements with non trivial monodromy in cohomology and some conjecture claim very strict restrictions for line arrangements (see [37]).

In this paper we state a vanishing conjecture of a very different nature, which is very easily stated and which involves only the lattice associated to the arrangement. Let Γ be the graph with vertex set $\mathcal{A}$ and edge set which is given by taking an edge (ℓ_i, ℓ_j) iff $\ell_i \cap \ell_j$ is a double point. Then our conjecture is as follows:

Conjecture 1.

$$\textit{Assume that } \Gamma \textit{ is connected; then } \mathcal{A} \textit{ has trivial monodromy.} \tag{1.1}$$

This conjecture is supported by several "experiments", since all computations we made confirm it. Also, all non-trivial monodromy examples we know have disconnected graph Γ. We give here a proof holding with further restrictions. Our method uses the algebraic complex given in [22] so our arrangements are real.

An arrangement with trivial monodromy will be called *a-monodromic*. We also introduce a notion of monodromic triviality over $\mathbb{Z}$. By using free differential calculus, we show that $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$ iff the fundamental group of the complement $\mathcal{M}(A)$ of the arrangement is commutative modulo the commutator subgroup of the length-zero subgroup of the free group F_n. As a consequence, we deduce that if $G := \pi_1(\mathcal{M}(\mathcal{A}))$ modulo its second derived group is commutative, then $\mathcal{A}$ has trivial monodromy over $\mathbb{Z}$.

In the final section we give an intrinsic characterization of the a-monodromicity. Let K be the kernel of the *length map* $G \to \mathbb{Z}$. We introduce the group $H := \frac{[G,G]}{[K,K]}$, and we show that such group exactly measures the "non-triviality" of the first homology of the Milnor fiber F, as well as its torsion. Any question about the first homology of F is actually a question about H. To our knowledge, H appears here for the first time (a preliminary partial version is appearing in [32]). We use this description to give some interesting new results about the a-monodromicity of the arrangement. First, we show that if G decomposes as a direct product of two groups, each of them containing an element of length 1, then $\mathcal{A}$ is a-monodromic (Theorem 8.11). This includes the case when G decomposes as a direct product of free groups. As a further interesting consequence, an arrangement which decomposes into two subarrangements which intersect each other transversally, is a-monodromic.

Also, we use this description to prove our conjecture under the hypothesis that we have a connected *admissible* graph of commutators (Theorem 8.13): essentially, this means to have enough double points $\ell_i \cap \ell_j$ which give as relation (mod $[K, K]$) the commutator of the fixed geometric generators β_i, β_j of G.

After having finished our paper, we learned about the paper [2] were the graph of double points is introduced and some partial results are shown, by very different methods.

2. Some recalls

We recall here some general constructions (see [36], also as a reference to most of the recent literature). Let M be a space with the homotopy type of a finite CW-complex with $H_1(M; \mathbb{Z})$ free Abelian of rank n, having basis $e_1, \dots, e_n$. Let $\underline{t} = (t_1, \dots, t_n) \in (\mathbb{C}^*)^n$ and denote by $\mathbb{C}_{\underline{t}}$ the Abelian rank one local system over M given by the representation

$$\phi : H_1(M; \mathbb{Z}) \longrightarrow \mathbb{C}^* = \mathrm{Aut}(\mathbb{C})$$

assigning t_i to e_i.

Definition 2.1. With this notation one calls

$$V(M) = \left\{\underline{t} \in (\mathbb{C}^*)^n : \dim_{\mathbb{C}} H_1(M; \mathbb{C}_{\underline{t}}) \geq 1\right\}$$

the *(first) characteristic variety* of M.

There are several other analogous definitions in all (co)homological dimensions, as well as refined definitions keeping into account the dimension actually reached by the local homology groups. For our purposes here we need to consider only the above definition.

The characteristic variety of a CW-complex M turns out to be an algebraic subvariety of the algebraic torus $(\mathbb{C}^*)^{b_1(M)}$ which depends only on the fundamental group $\pi_1(M)$ (see for example [10]).

Let now $\mathcal{A}$ be a complex hyperplane arrangement in $\mathbb{C}^n$. One knows that the complement $\mathcal{M}(A) = \mathbb{C}^n \setminus \bigcup_{H\in\mathcal{A}} H$ has the homotopy type of a finite CW-complex of dimension n. Moreover, in this case one knows by a general result (see [1]) that the characteristic variety of M is a finite union of torsion translated subtori of the algebraic torus $(\mathbb{C}^*)^{b_1(M)}$.

Now we need to briefly recall two standard constructions in arrangement theory (see [26] for details).

Let $\mathcal{A} = \{H_1, \dots, H_n\}$ be an affine hyperplane arrangement in $\mathbb{C}^n$ with coordinates $z_1, \dots, z_n$ and, for every $1 \le i \le n$, let α_i be a linear polynomial such that $H_i = \alpha_i^{-1}(0)$. The *cone* $c\mathcal{A}$ of $\mathcal{A}$ is a central arrangement in $\mathbb{C}^{n+1}$ with coordinates $z_0, \dots, z_n$ given by $\{\widetilde{H}_0, \widetilde{H}_1, \dots, \widetilde{H}_n\}$ where $\widetilde{H}_0$ is the coordinate hyperplane $z_0 = 0$ and, for every $1 \le i \le n$, $\widetilde{H}_i$ is the zero locus of the homogenization of α_i with respect to z_0.

Now let $\widetilde{\mathcal{A}} = \{\widetilde{H}_0, \dots, \widetilde{H}_n\}$ be a central arrangement in $\mathbb{C}^{n+1}$ and choose coordinates $z_0, \dots, z_n$ such that $\widetilde{H}_0 = \{z_0 = 0\}$; moreover, for every $1 \le i \le n$; let $\widetilde{\alpha}_i(z_0, \dots, z_n)$ be such that $\widetilde{H}_i = \widetilde{\alpha}_i^{-1}(0)$. The *deconing* of $\widetilde{\mathcal{A}}$ is the arrangement $d\widetilde{\mathcal{A}}$ in $\mathbb{C}^n$ given by $\{H_1, \dots, H_n\}$ where, if we set for every $1 \le i \le n$, $\alpha_i(z_1, \dots, z_1) = \widetilde{\alpha}_i(1, z_1, \dots, z_n)$, $H_i = \alpha_i^{-1}(0)$. One easily sees that $\mathcal{M}(c\mathcal{A}) = \mathcal{M}(\mathcal{A}) \times \mathbb{C}^*$ (and conversely $\mathcal{M}(\widetilde{\mathcal{A}}) = \mathcal{M}(d\widetilde{\mathcal{A}}) \times \mathbb{C}^*$).

The fundamental group $\pi_1(\mathcal{M}(\widetilde{\mathcal{A}})))$ is generated by elementary loops β_i, $i = 0, \dots, n$, around the hyperplanes and in the decomposition $\pi_1(\mathcal{M}(\mathcal{A})) \simeq \pi_1(\mathcal{M}(d\mathcal{A})) \times \mathbb{Z}$ the generator of $\mathbb{Z} = \pi_1(\mathbb{C}^*)$ corresponds to a loop going around all the hyperplanes. The generators can be ordered so that such a loop is represented by $\beta_0 \dots \beta_n$. Choosing $\widetilde{H}_0$ as the hyperplane at infinity in the deconing $\mathcal{A} = d\widetilde{\mathcal{A}}$, one has (see [10])

$$V(\widetilde{\mathcal{A}}) = \left\{\underline{t} \in (\mathbb{C}^*)^{n+1} : (t_1, \dots, t_n) \in V(d\mathcal{A}) \text{ and } t_0 \cdots t_n = 1\right\}.$$

It is still an open question whether the characteristic variety $V(\widetilde{\mathcal{A}})$ is combinatorially determined, that is, determined by the intersection lattice $L(\widetilde{\mathcal{A}})$. Actually, the question is partially solved: thanks to the above description we can write

$$V(\widetilde{\mathcal{A}}) = \check{V}(\widetilde{\mathcal{A}}) \cup T(\widetilde{\mathcal{A}}),$$

where $\check{V}(\widetilde{\mathcal{A}})$ is the union of all the components of $V(\widetilde{\mathcal{A}})$ passing through the unit element $\underline{1} = (1, 1, \dots, 1)$ and $T(\widetilde{\mathcal{A}})$ is the union of the translated tori of $V(\widetilde{\mathcal{A}})$.

The "homogeneous" part $\check{V}(\widetilde{\mathcal{A}})$ is combinatorially described through the *resonance variety*

$$\mathcal{R}^1(\widetilde{\mathcal{A}}) := \left\{a \in A^1 : H^1(A^\bullet, a \wedge \cdot) \neq 0\right\}$$

introduced in [18]. Here $A^\bullet$ is the Orlik-Solomon algebra over $\mathbb{C}$ of $\widetilde{\mathcal{A}}$. Denote by $\mathcal{V}(\widetilde{\mathcal{A}})$ the tangent cone of $V(\widetilde{\mathcal{A}})$ at $\underline{1}$; it turns out that $\mathcal{V}(\widetilde{\mathcal{A}}) \cong \mathcal{R}^1(\widetilde{\mathcal{A}})$. So, from $\mathcal{R}^1(\widetilde{\mathcal{A}})$ we can obtain the components of $V(\widetilde{\mathcal{A}})$ containing $\underline{1}$ by exponentiation.

It is also known (see [10,24]) that $\mathcal{R}^1(\widetilde{\mathcal{A}})$ is a subspace arrangement: $\mathcal{R}^1(\widetilde{\mathcal{A}}) = C_1 \cup \cdots \cup C_r$ with $\dim C_i \geq 2, C_i \cap C_j = 0$ for every $i \neq j$.

One makes a distinction between *local components* C_I of $R^1(\widetilde{\mathcal{A}})$, associated to a codimensional-2 flat I in the intersection lattice, which are contained in some coordinate hyperplanes; and *global components*, which are not contained in any coordinate hyperplane of A^1. Global components of dimension $k-1$ are known to correspond to (k,d)-*multinets* [19]. Let $\overline{\mathcal{A}}$ be the projectivization of $\widetilde{\mathcal{A}}$. A (k,d)-*multinet* on a multi-arrangement $(\overline{\mathcal{A}}, m)$, is a pair $(\mathcal{N}, \mathcal{X})$ where $\mathcal{N}$ is a partition of $\overline{A}$ into $k \geq 3$ classes $\overline{\mathcal{A}}_1, \ldots, \overline{\mathcal{A}}_k$ and $\mathcal{X}$ is a set of multiple points with multiplicity greater than or equal to 3 which satisfies a list of conditions. We just recall that $\mathcal{X}$ determines $\mathcal{N}$: construct a graph $\Gamma' = \Gamma'(\mathcal{X})$ with $\overline{\mathcal{A}}$ as vertex set and an edge from l to l' if and only if $l \cap l' \notin \mathcal{X}$. Then the connected components of Γ' are the blocks of the partition $\mathcal{N}$.

3. The Milnor fibre and a conjecture

Let $Q : \mathbb{C}^3 \to \mathbb{C}$ be a homogeneous polynomial (of degree $n+1$) which defines the arrangement $\widetilde{\mathcal{A}}$. Then Q gives a fibration

$$Q_{|\mathcal{M}(\tilde{\mathcal{A}})} : \mathcal{M}(\tilde{\mathcal{A}}) \to \mathbb{C}^* \tag{3.1}$$

with *Milnor fibre* $\mathbf{F} = Q^{-1}(1)$ and *geometric monodromy* $\pi_1(\mathbb{C}^*, 1) \to \mathrm{Aut}(F)$ induced by $x \to e^{\frac{2\pi i}{n+1}} \cdot x$ (see for example [35,38]).

Let A be any unitary commutative ring and

$$R := A\big[t, t^{-1}\big].$$

Consider the Abelian representation

$$\pi_1(\mathcal{M}(\widetilde{\mathcal{A}})) \to H_1(\mathcal{M}(\widetilde{\mathcal{A}}); \mathbb{Z}) \to \mathrm{Aut}(R) : \ \beta_j \to t\cdot$$

taking a generator β_j into t-multiplication. Let R_t be the ring R endowed with this $\pi_1(\mathcal{M}(\widetilde{\mathcal{A}}))$-module structure. Then the following it is well-known:

Proposition 3.1. *One has an R-module isomorphism*

$$H_*(\mathcal{M}(\widetilde{\mathcal{A}}), R_t) \cong H_*(F, A)$$

where t-multiplication on the left corresponds to the monodromy action on the right.

In particular for $R = \mathbb{Q}[t, t^{-1}]$, which is a PID, one has

$$H_*\left(\mathcal{M}(\widetilde{\mathcal{A}}), \mathbb{Q}[t^{\pm 1}]\right) \cong H_*(F, \mathbb{Q}).$$

Since the monodromy operator has order dividing $n+1$, $H_*(\mathcal{M}(\widetilde{\mathcal{A}}); R_t)$ decomposes into cyclic modules either isomorphic to R or to $\frac{R}{(\varphi_d)}$, where φ_d is a cyclotomic polynomial with $d|n+1$. It is another open problem to find a (possibly combinatorial) formula for the Betti numbers of F.

It derives from the spectral sequence associated to (3.1) that

$$n+1 = \dim(H_1(\mathcal{M}(\widetilde{\mathcal{A}}); \mathbb{Q})) = 1 + \dim \frac{H_1(F; \mathbb{Q})}{(\mu - 1)}$$

where on the right one has the coinvariants with respect to the monodromy action. Therefore

$$b_1(F) \geq n;$$

actually

$$b_1(F) = n \quad \Leftrightarrow \quad \mu = id.$$

Definition 3.2. An arrangement $\widetilde{\mathcal{A}}$ with trivial monodromy will be called *a-monodromic*.

Remark 3.3. The arrangement $\widetilde{\mathcal{A}}$ is a-monodromic if and only if

$$H_1(F; \mathbb{Q}) \cong \mathbb{Q}^n \left(\text{equivalently: } H_1(\mathcal{M}(\widetilde{\mathcal{A}}); R) \cong \left(\tfrac{R}{(t-1)}\right)^n\right)$$

Let $\mathcal{A} = d\widetilde{\mathcal{A}}$ be the affine part. In analogy with Definition 3.2 we say

Definition 3.4. The affine arrangement $\mathcal{A}$ is *a-monodromic* if

$$H_1(\mathcal{M}(\mathcal{A}); R) \cong \left(\frac{R}{(t-1)}\right)^{n-1}.$$

By the Kunneth formula one easily gets (with $R = \mathbb{Z}[t^{\pm 1}]$ or $R = \mathbb{Q}[t^{\pm 1}]$)

$$H_1(\mathcal{M}(\widetilde{\mathcal{A}}); R) \cong H_1(\mathcal{M}(\mathcal{A}); R) \otimes \frac{R}{(t^{n+1} - 1)} \oplus \frac{R}{(t-1)}. \tag{3.2}$$

It follows that if $\mathcal{A}$ has trivial monodromy then $\widetilde{\mathcal{A}}$ does. The converse is not true in general (see the example in Figure 6.7).

We can now state the conjecture presented in the introduction.

Conjecture 1. let Γ be the graph with vertex set $\mathcal{A}$ and edge-set all pairs (ℓ_i, ℓ_j) such that $\ell_i \cap \ell_j$ is a double point. Then if Γ is connected then $\mathcal{A}$ is a-monodromic.

Conjecture 2. let Γ be as before. Then if Γ is connected then $\widetilde{\mathcal{A}}$ is a-monodromic.

By formula (3.2) Conjecture 1 implies Conjecture 2.

A partial evidence of these conjecture is that the connectivity condition on the graph of double points gives strong restrictions on the characteristic variety, as we now show.

Remark 3.5. Let $\underline{t} = (t, \dots, t) \in (\mathbb{C}^*)^{n+1}$ give non-trivial monodromy for the arrangement $\widetilde{\mathcal{A}}$. Then $\underline{t} \in V(\widetilde{\mathcal{A}})$. Moreover, $\underline{t}$ can intersect $\check{V}(\widetilde{\mathcal{A}})$ only in some global component.

The next theorem shows how the connectivity of Γ is an obstruction to the existence of multinet structures.

Theorem 3.6. *If the above graph Γ is connected then the projectivized $\overline{\mathcal{A}}$ of $\widetilde{\mathcal{A}}$ does not support any multinet structure.*

Proof. Choose a set $\mathcal{X}$ of points of multiplicity greater than or equal to 3 and build $\Gamma'(\mathcal{X})$ as we said at the end of Section 2. This graph $\Gamma'(\mathcal{X})$ has $\overline{\mathcal{A}}$ as the set of vertices and the set of edges of Γ is contained in the set of edges of $\Gamma'(\mathcal{X})$. Since by hypothesis Γ is connected then $\Gamma'(\mathcal{X})$ has at most two connected components and so $\mathcal{X}$ cannot give a multinet structure an $\overline{\mathcal{A}}$. $\square$

Corollary 3.7. *If the graph Γ is connected, there is no global resonance component in $\mathcal{R}^1(\widetilde{\mathcal{A}})$.*

So, according to Remark 3.5, if Γ is connected then non trivial monodromy could appear only in the presence of some translated subtori in the characteristic variety.

4. Algebraic complexes

We shall prove the conjectures with extra assumptions on the arrangement. Our tool will be an algebraic complex which was obtained in [22], as a 2-dimensional refinement of that in [33], where the authors used the explicit construction of a *minimal cell complex* which models the complement. Since these complexes work for real defined arrangements, this will be our first restriction.

Of course, there are other algebraic complexes computing local system cohomology (see the references listed in the introduction). The one in [22] seemed to us particularly suitable to attack the present problem (even if we were not able to solve it in general).

First, the complex depends on a fixed and generic system of "polar coordinates". In the present situation, this just means to take an oriented affine real line ℓ which is transverse to the arrangement. We also assume (even if it is not strictly necessary) that ℓ is "far away" from $\mathcal{A}$, meaning that it does not intersect the closure of the bounded facets of the arrangement. This is clearly possible because the union of bounded chambers is a compact set (the arrangement is finite). The choice of ℓ induces a labelling on the lines $\{\ell_1, \dots, \ell_n\}$ in $\mathcal{A}$, where the indices of the lines agree with the ordering of the intersection points with ℓ, induced by the orientation of ℓ.

Let us choose a basepoint $O \in \ell$, coming before all the intersection points of ℓ with $\mathcal{A}$ (with respect to the orientation of ℓ). We recall the construction in [22] in the case of the Abelian local system defined before.

Let $\mathrm{Sing}(\mathcal{A})$ be the set of singular points of the arrangement. For any point $P \in \mathrm{Sing}(\mathcal{A})$, let $S(P) := \{\ell \in \mathcal{A} : P \in \ell\}$; so $m(P) = |S(P)|$ is the multiplicity of P.

Let $i_P,\ i^P$ be the minimum and maximum index of the lines in $S(P)$ (so $i_P < i^P$). We denote by $C(P)$ the subset of lines in $\mathcal{A}$ whose indices belong to the closed interval $[i_P, i^P]$. We also denote by

$$U(P) := \{\ell \in \mathcal{A} : \ \ell \text{ does not separate } P \text{ from the basepoint } O\}.$$

Let $(\mathcal{C}_*, \partial_*)$ be the 2-dimensional algebraic complex of free R-modules having one 0-dimensional basis element e^0, n 1-dimensional basis elements $e^1_j,\ j = 1, \dots, n$, (e^1_j corresponding to the line ℓ_j) and $\nu_2 = \sum_{P \in \mathrm{Sing}(\mathcal{A})} m(P) - 1$ 2-dimensional basis elements: to the singular point P of multiplicity $m(P)$ we associate generators $e^2_{P,h},\ h = 1, \dots, m(P) - 1$. The lines through P will be indicized as $\ell_{j_{P,1}}, \dots, \ell_{j_{P,m(P)}}$ (with growing indices).

As a dual statement to [22, Theorem 2], we obtain:

Theorem 4.1. *The local system homology $H_*(\mathcal{M}(\mathcal{A}); R)$ is computed by the complex $(\mathcal{C}_*, \partial_*)$ above, where*

$$\partial_1\big(e^1_j\big) = (t_j - 1)\, e^0$$

and

$$\begin{aligned}\partial_2(e^2_{P,h}) = &\sum_{\ell_j \in S(P)} \left(\prod_{\substack{i<j \text{ so that} \\ l_i \in U(P)}} t_i \right) \left(\prod_{i \in [j_{P,h+1} \to j)} t_i - \prod_{\substack{i<j \text{ so that} \\ l_i \in S(P)}} t_i \right) e^1_j \\ &+ \sum_{\ell_j \in C(P) \cap U(P)} \left(\prod_{\substack{i<j \text{ so that} \\ l_i \in U(P)}} t_i \right) \left(1 - \prod_{\substack{i \le j_{P,h}, i<j \\ l_i \in S(P)}} t_i \right) \left(\prod_{\substack{i \ge j_{P,h+1},\ i<j \\ l_i \in S(P)}} t_i - \prod_{\substack{i \ge j_{P,h+1} \\ l_i \in S(P)}} t_i \right) e^1_j ,\end{aligned} \tag{4.1}$$

where $[j_{P,h+1} \to j)$ is the set of indices of the lines in $S(P)$ which run from $j_{P,h+1}$ (included) to j (excluded) in the cyclic ordering of $1, \dots, n$.

By convention, a product over an empty set of indices equals 1.

When $R = A[t^{\pm 1}]$ and $t_i = t,\ i = 1, \ldots, n$, we obtain the local homology $H_*(\mathcal{M}(\mathcal{A}); R)$ by using an analogue algebraic complex, where all t_i's equal t in the formulas. In particular (4.1) becomes

$$\begin{aligned}\partial_2\left(e^2_{P,h}\right) = &\sum_{\ell_j \in S(P)} t^{\#\{\ell_i \in U(P): i<j\}}\left(t^{\#[j_{P,h+1}\to j)} - t^{\#\{\ell_i \in S(P):\ i<j\}}\right) e^1_j \\ &+ \sum_{\ell_j \in C(P)\cap U(P)} t^{\#\{\ell_i\in U(P): i<j\}+\#\{\ell_i \in S(P): i\geq j_{P,h+1}, i<j\}} \\ &\times \left(1 - t^{\#\{\ell_i \in S(P): i\leq j_{P,h}, i<j\}}\right)\left(1 - t^{\#\{\ell_i\in S(P):\ i\geq j_{P,h+1},\ i\geq j\}}\right) e^1_j.\end{aligned} \tag{4.2}$$

By separating in the first sum the case $j \geq j_{P,h+1}$ from the case $j \leq j_{P,h}$ we have:

$$\begin{aligned}\partial_2\left(e^2_{P,h}\right) = &\sum_{\substack{\ell_j \in S(P)\\ j\geq j_{P,h+1}}} t^{\#\{\ell_i \in U(P): i<j\}+\#\{\ell_i\in S(P): j_{P,h+1}\leq i<j\}}\left(1 - t^{\#\{\ell_i\in S(P): i\leq j_{P,h}\}}\right) e^1_j \\ &+ \sum_{\substack{\ell_j \in S(P)\\ j\leq j_{P,h}}} t^{\#\{\ell_i \in U(P): i<j\}+\#\{\ell_i\in S(P): i<j\}}\left(t^{\#\{\ell_i\in S(P): j_{P,h+1}\leq i\}} - 1\right) e^1_j \\ &+ \sum_{\ell_j \in C(P)\cap U(P)} t^{\#\{\ell_i\in U(P): i<j\}+\#\{\ell_i \in S(P): i\geq j_{P,h+1}, i<j\}} \\ &\times \left(1 - t^{\#\{\ell_i \in S(P): i\leq j_{P,h}, i<j\}}\right)\left(1 - t^{\#\{\ell_i\in S(P):\ i\geq j_{P,h+1},\ i\geq j\}}\right) e^1_j.\end{aligned} \tag{4.3}$$

In particular, let P be a double point. Then h takes only the value 1, and $j_{P,1}$, $j_{P,2}$ are the indices of the two lines passing through P. So formula (4.3) becomes

$$\begin{aligned}\partial_2\left(e^2_{P,1}\right) = &\, t^{\#\{\ell_i\in U(P): i<j_{P,2}\}}(1-t)\, e^1_{j_{P,2}} + t^{\#\{\ell_i \in U(P):\ i<j_{P,1}\}}\,(t-1)\ e^1_{j_{P,1}} \\ &+ \sum_{\ell_j\in C(P)\cap U(P)} t^{\#\{\ell_i\in U(P):\ i<j\}}\,(t-1)^2\ e^1_j.\end{aligned} \tag{4.4}$$

Since ∂_2 is divisible by $t-1$ we can rewrite (4.4) as

$$\partial_2\left(e^2_{P,1}\right) = (t-1)\ \tilde{\partial}_2\left(e^2_{P,1}\right) \tag{4.5}$$

where

$$\begin{aligned}\tilde{\partial}_2\left(e^2_{P,1}\right) = &\, t^{\#\{\ell_i\in U(P):\ i<j_{P,2}\}} e^1_{j_{P,2}} - t^{\#\{\ell_i\in U(P): i<j_{P,1}\}} e^1_{j_{P,1}} \\ &+ \sum_{\ell_j\in C(P)\cap U(P)} t^{\#\{\ell_i\in U(P):\ i<j\}(1-t)\ e^1_j}.\end{aligned} \tag{4.6}$$

5. A proof in particular cases

We give here a proof of Conjecture 1 with further hypothesis on $\mathcal{A}$.

Notice that the rank of ∂_2 is $n-1$ (the sum of all rows vanishes). Then the arrangement has no monodromy if and only if the only elementary divisor of ∂_2 is $\varphi_1 := t-1$, so ∂_2 diagonalizes to $\oplus_{i=1}^{n-1} \varphi_1$. This is equivalent to the reduced boundary $\tilde{\partial}_2$ having an invertible minor of order $n-1$.

Let Γ be the graph of double points. A choice of an admissible coordinate system gives a total ordering on the lines so it induces a labelling, varying between 1 and n, on the set of vertices $V\Gamma$ of Γ. Let T be a spanning tree of Γ (with induced labelling on VT).

Definition 5.1. We say that the induced labelling on $VT = V\Gamma$ is *very good* (with respect to the given coordinate system) if the sequence $n, \dots, 1$ is a *collapsing ordering* on T. In other words, the graph obtained by T by removing all vertices with label $\geq i$ and all edges having both vertices with label $\geq i$, is a tree, for all $i = n, \dots, 1$.

We say that the spanning tree T is *very good* if there exists an admissible coordinate system such that the induced labelling on VT is very good (see Figure 6.1).

Remark 5.2.

(1) A labelling over a spanning tree T gives a collapsing ordering if and only if for each vertex v, the number of adjacent vertices with lower label is ≤ 1. In this case, only the vertex labelled with 1 has no lower labelled adjacent vertices (by the connectness of T).
(2) Given a collapsing ordering over T, for each vertex v with label $i_v > 1$, let $\ell(v)$ be the edge which connects v with the unique adjacent vertex with lower label; by giving to $\ell(v)$ the label $i_v + \frac{1}{2}$, we obtain a discrete Morse function on the graph T (see [20]) with unique critical cell given by the vertex with label 1. The set of all pairs $(v, \ell(v))$ is the acyclic matching which is associated to this Morse function.

Let us indicate by Γ_0 the linear tree with n vertices: we consider Γ_0 as a CW-decomposition of the real segment $[1, n]$, with vertices $\{j\}$, $j = 1, \dots, n$, and edges the segments $[j, j+1]$, $j = 1, \dots, n-1$.

Definition 5.3. We say that a labelling induced by some coordinate system on the tree T is *good* if there exists a permutation $i_1, \dots, i_n$ of $1, \dots, n$ which gives a collapsing sequence both for T and for Γ_0. In other words, at each step we always remove either the maximum labelled vertex or the minimum, and this is a collapsing sequence for T.

We say that T is *good* if there exists an admissible coordinate system such that the induced labelling on VT is good (see Figure 6.2).

Notice that a very good labelling is a good labelling where at each step one removes the maximum vertex.

Consider some arrangement $\mathcal{A}$ with graph Γ and labels on the vertices which are induced by some coordinate system. Notice that changes of coordinates act on the labels by giving all possible *cyclic permutations*, which are generated by the transformation $i \to i+1 \bmod n$. So, given a labelled tree T, checking if T is very good (respectively good) consists in verifying if some cyclic permutation of the labels is very good (respectively good). This property depends not only on the "shape" of the tree, but also on how the lines are disposed in $\mathbb{R}^2$ (the associated oriented matroid). In fact, one can easily find arrangements where some "linear" tree is very good, and others where some linear tree is not good.

Definition 5.4. We say that an arrangement $\mathcal{A}$ is *very good* (respectively *good*) if Γ is connected and has a very good (respectively good) spanning tree.

It is not clear if this property is combinatorial, *i.e.*, if it depends only on the lattice. Of course, $\mathcal{A}$ very good implies $\mathcal{A}$ good.

Theorem 5.5. *Let $\mathcal{A}$ be a good arrangement. Then $\mathcal{A}$ is a-monodromic.*

Proof. We use induction on the number n of lines, the claim being trivial for $n = 1$. Take a suitable coordinate system as in Definition 5.4, such that the graph Γ has a spanning tree T with good labelling. Assume for example that at the first step we remove the last line, so the graph Γ' of the arrangement $\mathcal{A}' := \mathcal{A} \setminus \{\ell_n\}$ is connected and the spanning tree T' obtained by removing the vertex $\{\ell_n\}$ and the "leaf-edge" (ℓ_n, ℓ_j) (for some $j < n$) has a good labelling.

There are $n-1$ double points which correspond to the edges of T : only one of these is contained in ℓ_n, namely $\ell_n \cap \ell_j$ (see Remark 5.2). Let $\mathcal{D} := \{d_1, \dots, d_{n-1}\}$ be the set of such double points, with $d_{n-1} = \ell_n \cap \ell_j$. Let also $\mathcal{D}' := \{d_1, \dots, d_{n-2}\}$, which corresponds to the edges of T'. Let $(\mathcal{C}(\mathcal{D})_*, \partial_*)$ (respectively $(\mathcal{C}(\mathcal{D}')_*, \partial'_*)$) be the subcomplex of $\mathcal{C}(\mathcal{A})_*$ generated by the 2-cells which correspond to $\mathcal{D}$ (respectively $\mathcal{D}'$): then $\mathcal{C}(\mathcal{D})_2 = \oplus_{1 \le i \le n-1} Re_j$, and $\mathcal{C}(\mathcal{D}')_2 = \oplus_{1 \le i \le n-2} Re'_j$. Notice that, by the explicit formulas given in Section 4, the component of the boundary $\partial_2(e_j)$ along the 1-dimensional generator corresponding to ℓ_n equals $-\varphi_1$ for $j = n-1$, and vanishes for $j = 1, \dots, n-2$. Actually, the natural map taking e'_j into e_j, $j = 1, \dots, n-2$, identifies $\mathcal{C}(\mathcal{D}')_*$ with the sub complex of $\mathcal{C}(\mathcal{D})_*$ generated by the e_j's, $j = 1, \dots, n-2$,

$$\partial_2 = \left[\begin{array}{c|c} \partial'_2 & * \\ \hline 0 & -\varphi_1 \end{array}\right]. \tag{5.1}$$

Then by induction ∂'_2 diagonalizes to $\oplus_{j=1}^{n-2}\varphi_1$. Therefore ∂_2 diagonalizes to $\oplus_{j=1}^{n-1}\varphi_1$, which gives the thesis. If at the first step we remove the first line, the argument is similar, because $\partial_2(e_j)$ has no non-vanishing components along the generator corresponding to ℓ_1. □

Let us consider a different situation.

Definition 5.6. We say that a subset Σ of the set of singular points $\mathrm{Sing}(\mathcal{A})$ of the arrangement $\mathcal{A}$ is *conjugate-free* (with respect to a given admissible coordinate system) if $\forall P \in \Sigma$ the set $U(P) \cap C(P)$ is empty.

An arrangement $\mathcal{A}$ will be called *conjugate-free* if Γ is connected and contains a spanning tree T such that the set of points in $\mathrm{Sing}(\mathcal{A})$ that correspond to the edges ET of T is conjugate-free (see Figure 6.3).

Let Σ be conjugate-free: it follows from formula (4.3) that the boundary of all generators $e^2_{P,h}$, $P \in \Sigma$, can have non-vanishing components only along the lines which contain P.

Theorem 5.7. *Assume that $\mathcal{A}$ is conjugate-free. Then $\mathcal{A}$ is a-monodromic.*

Proof. The sub matrix of ∂_2 which corresponds to the double points ET is φ_1-times the incidence matrix of the tree T. Such matrix is the boundary matrix of the complex which computes the $\mathbb{Z}$-homology of T : it is a unimodular rank-$(n-1)$ integral matrix (see for example [3]). From this the result follows straightforwardly. □

We can have a mixed situation between Definitions 5.4 and 5.6 (see Figure 6.4).

Theorem 5.8. *Assume that Γ is connected and contains a spanning tree T which reduces, after a sequence of moves where we remove either the maximum or the minimum labelled vertex, to a subtree T' which is conjugate-free. Then $\mathcal{A}$ is a-monodromic.*

Proof. The thesis easily follows by induction on the number n of lines. In fact, either T is conjugate-free, and we use Theorem 5.7, or one of the subtrees $T \setminus \{\ell_n\}$, $T \setminus \{\ell_1\}$ satisfies again the hypothesis of the theorem. Assume that it is $T'' = T \setminus \{\ell_n\}$. Then the boundary map ∂_2 restricted to the 2-cells corresponding to ET'' has a shape similar to (5.1). Therefore by induction we conclude. □

Some examples are given in Section 6.

Remark 5.9. In all the theorems in this section, we have proven a stronger result: namely, the subcomplex spanned by the generators corresponding to the double points is a-monodromic.

6. Examples

In this section we give examples corresponding to the various definitions of Section 5. We include the computations of the local homology of the complements.

In Figure 6.1 we show an arrangement having a very good tree (Definition 5.1) and the associated sequence of contractions.

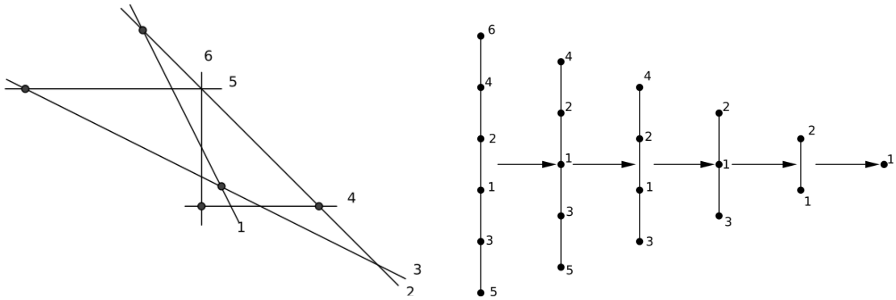

Figure 6.1.

In Figure 6.2 an arrangement with a good tree is given (Definition 5.3) together with its sequence of contractions.

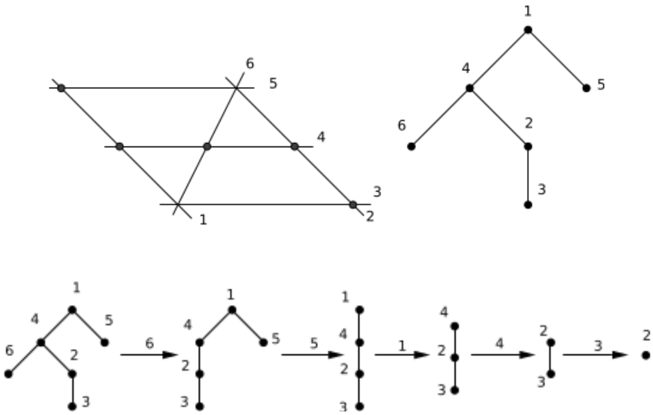

Figure 6.2.

An arrangement having a tree which is both conjugate-free (see Definition 5.6) and good is depicted in Figure 6.3

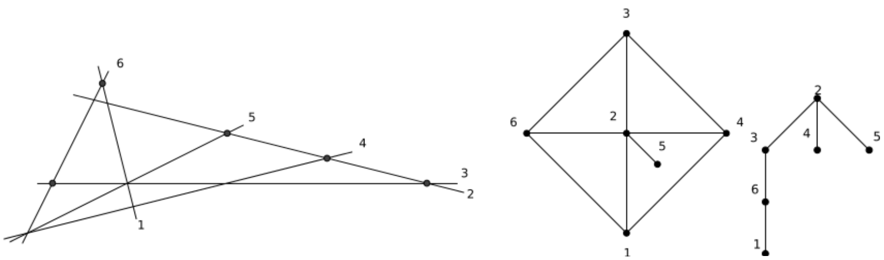

Figure 6.3.

In Figure 6.4 we show an arrangement with a tree which after 2 admissible contractions becomes conjugate free (see Theorem 5.8).

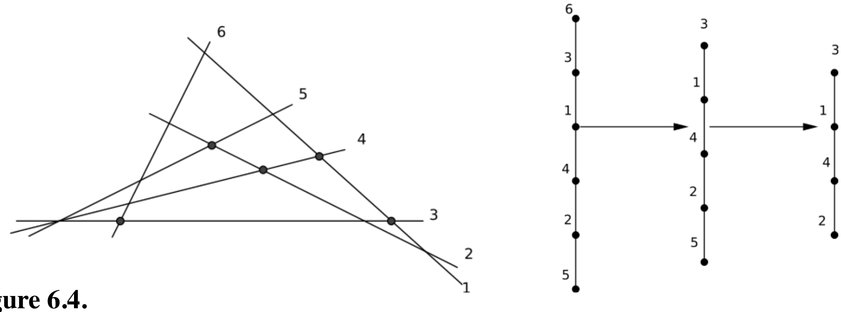

Figure 6.4.

Next we give some example of arrangements with non-trivial monodromy. Notice that the graph of double points is disconnected in these cases.

Notice also that in the first two examples one has non-trivial monodromy both for the given affine arrangement and its conifed arrangement in $\mathbb{C}^3$; in the last example, the given affine arrangement has non trivial monodromy while its conification is a-monodromic.

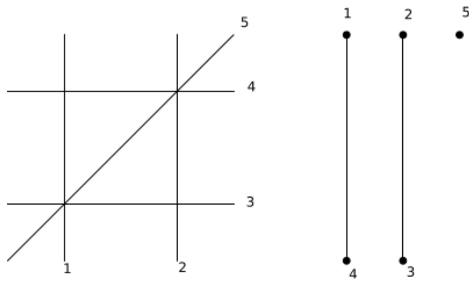

Figure 6.5. Deconed A3 arrangement

$$H_1(\mathcal{M}(\mathcal{A}), \mathbb{Q}[t^{\pm 1}]) \simeq \left(\frac{\mathbb{Q}[t^{\pm 1}]}{(t-1)}\right)^3 \oplus \frac{\mathbb{Q}[t^{\pm 1}]}{(t^3-1)}.$$

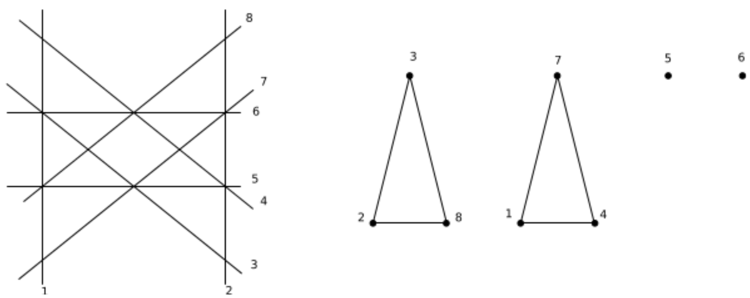

Figure 6.6. Deconed Pappus arrangement

$$H_1(\mathcal{M}(\mathcal{A}), \mathbb{Q}[t^{\pm 1}]) \simeq \left(\frac{\mathbb{Q}[t^{\pm 1}]}{(t-1)}\right)^6 \oplus \frac{\mathbb{Q}[t^{\pm 1}]}{(t^3-1)}.$$

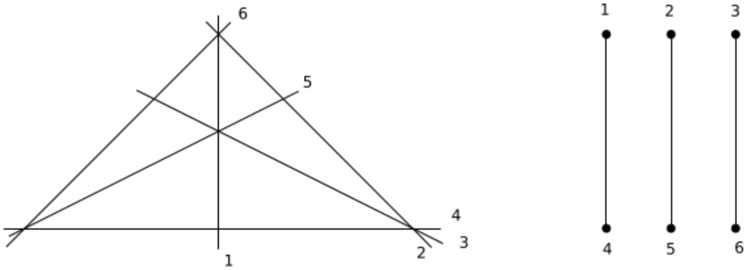

Figure 6.7. The "complete triangle" has non-trivial monodromy but its conification is a-monodromic

$$H_1(\mathcal{M}(\mathcal{A}), \mathbb{Q}[t^{\pm 1}]) \simeq \left(\frac{\mathbb{Q}[t^{\pm 1}]}{(t-1)}\right)^4 \oplus \frac{\mathbb{Q}[t^{\pm 1}]}{(t^3-1)}.$$

We focus here on the structure of the fundamental groups of the above examples, in particular in case of a-monodromic arrangements.

For arrangement in Figure 6.1: after taking line 5 to infinity we obtain an affine arrangement having only double points with two pairs of parallel lines, namely (the new) lines 2, 6 and 4, ∞. Therefore

$$\pi_1(\mathcal{M}(\mathcal{A})) = \mathbb{Z} \times \mathbb{Z} \times F_2 \times F_2.$$

We consider arrangement in Figure 6.2 and in Figure 6.5 together. The deconed $A3$ arrangement in Figure 6.5 is a well known $K(\pi, 1)$-arrangement: the fundamental group of the complement is the *pure braid group* P_4 in 4 strands. Notice that the projection onto the y coordinate fibers the complement over $\mathbb{C} \setminus \{2 \text{ pts}\}$ with fiber $\mathbb{C} \setminus \{3 \text{ pts}\}$. It is well known that this fibering is not trivial and we obtain a semi-direct product decomposition

$$\pi_1(\mathcal{M}(\mathcal{A})) = F_3 \rtimes F_2.$$

The same projection gives a fibering of the complement of the arrangement in Figure 6.2 over $\mathbb{C} \setminus \{3 \text{ pts}\}$ with fiber $\mathbb{C} \setminus \{3 \text{ pts}\}$. Notice that this is also a non-trivial fibering, so we have a semi-direct decomposition

$$\pi_1(\mathcal{M}(\mathcal{A})) = F_3 \rtimes F_3.$$

In particular, we have an a-monodromic arrangement such that the fundamental group of the complement is not a direct product of free groups.

In the arrangement of Figure 6.3 the line at infinity is transverse to the other lines. If we take line 5 at infinity we get an affine arrangement with only double points, with two pairs of parallel lines 1, 3 and 4, 6. Therefore we obtain a decomposition of $\pi_1(\mathcal{M}(\mathcal{A}))$ as in case of Figure 6.1.

The arrangement of Figure 6.4 has only one triple point. By taking line 5 to infinity we get an affine arrangement with only double points and one pair of parallel lines 3, 4. Therefore

$$\pi_1(\mathcal{M}(\mathcal{A})) = \mathbb{Z}^4 \times F_2.$$

The *complete triangle* in Figure 6.7 becomes, after taking any line at infinity, the affine arrangement $\mathcal{A}'$ which is obtained from the $A3$ deconed arrangement in Figure 6.5 by adding one more line $\bar{l}$ which is transverse to all the others. Therefore

$$\pi_1(\mathcal{M}(\mathcal{A})) = \mathbb{Z} \times (F_3 \rtimes F_2).$$

Remark 6.1. It turns out that the arrangement $\mathcal{A}'$ is a-monodromic. This is not a contradiction: in fact, one is considering two different local systems on $\mathcal{M}(\mathcal{A}') = \mathcal{M}(\mathcal{A})$. The a-monodromic one associates to an elementary loop around $\bar{l}$ the t-multiplication. This is different from the one obtained by exchanging one of the affine lines of the arrangement $\mathcal{A}$ in Figure 6.7 with the infinity line. In this case we should associate to an elementary loop around $\bar{l}$ the t^6-multiplication, and then apply formula (4.1).

7. Free calculus

In this section we reformulate our conjecture in terms of Fox calculus. Let $\mathcal{A} = \{l_1, \ldots, l_n\}$ be as above; if we denote by β_i an elementary loop around l_i we have that the fundamental group $\pi_1(\mathcal{M}(\mathcal{A}))$ is generated by $\beta_1, \ldots, \beta_n$ and a presentation of this group is given for example in [29]. Let $R = \mathbb{Q}[t^{\pm 1}]$ be as above with the given structure of $\pi_1(\mathcal{M}(\mathcal{A}))-$module.

We denote by F_n the free group generated by $\beta_1, \ldots, \beta_n$. Let $\varphi : F_n \to < t >$ be the group homomorphism defined by $\varphi(\beta_i) = t$ for every $1 \leq i \leq n$ where $< t >$ is the multiplicative subgroup of R generated by t. As in [4], if w is a word in the β_j's, we use the notation w^φ for $\varphi(w)$. Consider the algebraic complex which computes the local homology of $\mathcal{M}(\mathcal{A})$ introduced in Section 4. The following remark is crucial for the rest of this section: if $e^2_{P,j}$ is a two-dimensional generator corresponding to a two-cell which is attached along the word w in the β_j's, then $\left(\frac{\partial w}{\partial \beta_i}\right)^\varphi$ is the coefficient of e^1_i of the border of $e^2_{P,j}$. This easily follows from the combinatorial calculation of local system homology.

Let $l : F_n \longrightarrow \mathbb{Z}$ be the length function, given by

$$l(\beta_{i_1}^{\epsilon_1} \cdots \beta_{i_r}^{\epsilon_r}) = \sum_{k=1}^{r} \epsilon_k.$$

Then $\ker\varphi$ is the normal subgroup of F_n given by the words of lenght 0.

Each relation in the fundamental group $\pi_1(\mathcal{M}(\mathcal{A}))$ is a commutator (*cf.* [17, 29]), so it lies in $\ker\varphi$. So, in the sequel, we consider only words in $\ker\varphi$.

Remark 7.1. The arrangement $\mathcal{A}$ is a-monodromic if and only if (by definition) the $\mathbb{Q}[t^{\pm 1}]$-module generated by $\partial_2(e_j^2),\ j=1,\dots,\nu_2$, equals $(t-1)\mathrm{ker}\partial_1$. One has: $\mathrm{ker}\partial_1 = \{\sum_{j=1}^n\ x_j\ e_j^1\ :\ \sum_{j=1}^n\ x_j = 0\}$.
Let $R_j,\ j=1\dots,\nu_2$, be a complete set of relations in $\pi_1(\mathcal{M}(\mathcal{A}))$. We use now e_j^2 to indicate the two-dimensional generator corresponding to a two-cell which is attached along the word R_j. Then the boundary of e_j^2 is given by

$$\partial_2(e_j^2) = \sum_{i=1}^n \left(\frac{\partial R_j}{\partial \beta_i}\right)^{\varphi} e_i^1, \quad j=1,\dots,\nu_2. \tag{7.1}$$

Then by Remark 7.1 $\mathcal{A}$ is a-monodromic if and only if each element of the shape

$$P(t) := (1-t)\sum_{i=1}^n\ P_i(t)\ e_i^1, \quad \sum_{i=1}^n P_i(t) = 0, \quad (P_i(t)\in\mathbb{Q}[t^{\pm 1}], i=1\dots,n) \tag{7.2}$$

is a linear combination with coefficients in $\mathbb{Q}[t^{\pm 1}]$ of the elements in (7.1), *i.e.*:

$$P(t) = \sum_{j=1}^{\nu_2} Q_j(t)\partial_2(e_j^2), \quad Q_j(t)\in\mathbb{Q}[t^{\pm 1}]. \tag{7.3}$$

It is natural to wonder about solutions with coefficients in $\mathbb{Z}[t^{\pm 1}]$ instead of $\mathbb{Q}[t^{\pm 1}]$. We say that $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$ if there is a solution to (7.3) over $\mathbb{Z}[t^{\pm 1}]$ (when all the $P_i(t)$'s in (7.2) are in $\mathbb{Z}[t^{\pm 1}]$).

Theorem 7.2. *The arrangement $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$ if and only if $\pi_1(\mathcal{M}(\mathcal{A}))$ is commutative modulo* $[\mathrm{ker}\varphi, \mathrm{ker}\varphi]$. *More precisely, $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$ if and only if*

$$[F_n, F_n] = N[\mathrm{ker}\varphi, \mathrm{ker}\varphi],$$

where N is the normal subgroup generated by the relations R_j's .

Proof. A set of generators for $(t-1)\mathrm{ker}\partial_1$ as $\mathbb{Z}[t^{\pm 1}]$-modulo is given by all elements of the type

$$P_{rs} := (1-t)(e_r^1 - e_s^1)\ , \quad r\neq s.$$

Such an element can be re-written in the form (7.1) as

$$P_{rs} = \sum_{i=1}^n \left(\frac{\partial[\beta_r,\beta_s]}{\partial\beta_i}\right)^{\varphi} e_i^1,$$

where $[\beta_r,\beta_s] = \beta_r\beta_s\beta_r^{-1}\beta_s^{-1}$. Now there exists an expression like (7.3) for P_{rs}, with all $Q_j(t)\in\mathbb{Z}[t^{\pm 1}]$ if and only if

$$\left(\frac{\partial[\beta_r,\beta_s]}{\partial\beta_i}\right)^{\varphi} = \left(\frac{\partial\prod_{j=1}^{\nu_2}\ R_j^{Q_j(\beta_1)}}{\partial\beta_i}\right)^{\varphi}, \quad i=1,\dots,n. \tag{7.4}$$

Here $Q_j(\beta_1) \in \mathbb{Z}[F_n]$ is obtained by substituting t with β_1 (any word of length one would give the same here). Moreover, for R, w any words in kerφ we set $R^w := wRw^{-1}$, and for $a \in \mathbb{Z}$ we set $R^{aw} := R^w \dots R^w$ (a factors) if $a > 0$ and $(R^{-1})^w \dots (R^{-1})^w$ ($|a|$ factors) for $a < 0$. Also, we set $R^{aw+bu} := R^{aw}R^{bu}$. Then equalities (7.4) come from standard Fox calculus.

Then from Blanchfield theorem (see [4, Chapter 3]) it follows that

$$[F_n, F_n] \subset N[\ker\varphi, \ker\varphi].$$

The opposite inclusion follows because, as we said before Remark 7.1, for any arrangement one has $N \subset [F_n, F_n]$. □

Remark 7.3. The condition in Theorem 7.2 is equivalent to the equality

$$\frac{F_n}{N[\ker\varphi, \ker\varphi]} = \frac{F_n}{[F_n, F_n]} = H_1(\mathcal{M}(\mathcal{A}); \mathbb{Z}).$$

Since ker$\varphi \supset [F_n, F_n]$, so $[\ker\varphi, \ker\varphi] \supset [[F_n, F_n], [F_n, F_n]]$, the next corollary follows immediately from Theorem 7.2.

Corollary 7.4. *Assume that $\pi_1(\mathcal{M}(\mathcal{A}))/\pi_1(\mathcal{M}(\mathcal{A}))^{(2)}$ is Abelian, which is equivalent to the condition $\pi_1(\mathcal{M}(\mathcal{A}))^{(1)} = \pi_1(\mathcal{M}(\mathcal{A}))^{(2)}$, where $\pi_1(\mathcal{M}(\mathcal{A}))^{(i)}$ is the i-th element of the derived series of $\pi_1(\mathcal{M}(\mathcal{A}))$, for $i \geq 0$. Then $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$.*

The condition of Corollary 7.4 corresponds to the vanishing of the so-called *Alexander invariant* of $\pi_1(\mathcal{M}(\mathcal{A}))$.

As a subgroup of the free group F_n, the group kerφ is a free group We use the Reidemeister-Schreier method to write an explicit list of generators. Notice that for any fixed $1 \leq j \leq n$, the set $\{\beta_j^k : k \in \mathbb{Z}\}$ is a Schreier right coset representative system for $F_n/\ker\varphi$. Denote briefly by $s_{k,i}$ the element $s_{\beta_j^k, \beta_i} = \beta_j^k \beta_i \overline{(\beta_j^k \beta_i)}^{-1} = \beta_j^k \beta_i \beta_j^{-(k+1)}$. Then

$$\ker\varphi = \left\langle \{s_{k,i} : 1 \leq i \leq n,\ k \in \mathbb{Z}\}; s_{k,i} \right\rangle,$$

where $s_{k,i}$ is a relation if and only if $\beta_j^k \beta_i$ is freely equal to β_j^{k+1}; this happens if and only if $i = j$. So kerφ is the free group generated by $\{s_{k,i} : k \in \mathbb{Z}, 1 \leq i \leq n, i \neq j\}$. Its Abelianization

$$\text{ab}\,(\ker\varphi) = \ker\varphi/[\ker\varphi, \ker\varphi]$$

is the free Abelian group on the classes $\overline{s}_{k,i}$ of the generators $s_{k,i}$'s, $i \neq j$. Let

$$\text{ab} : \ker\varphi \longrightarrow \text{ab}\,(\ker\varphi)$$

be the Abelianization homomorphism. Now we define the automorphism σ of $\ker\varphi$ by

$$\sigma(s_{k,i}) = s_{k+1,i},$$

which passes to the quotient, so it defines an automorphism (call it again σ) of $\mathrm{ab}(\ker\varphi)$. Therefore we may view $\mathrm{ab}(\ker\varphi)$ as a finitely genereted free $\mathbb{Z}[\sigma^{\pm 1}]$-module, with basis $\overline{s}_{0,i}$ with $1 \le i \le n$ and $i \ne j$. In this language Theorem 7.2 translates as:

Theorem 7.5. *The arrangement $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$ if and only if the submodule* $(1-\sigma)\mathrm{ab}(\ker\varphi)$ *of* $\mathrm{ab}(\ker\varphi)$ *is generated by* $\mathrm{ab}(R_j)$, $j = 1, \dots, \nu_2$, *as* $\mathbb{Z}[\sigma^{\pm 1}]$*-module.*

Of course, one can give a conjecture holding over $\mathbb{Z}$.

Conjecture 3. Assume that Γ is connected; then $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$.
Conjecture 3 clearly implies Conjectures 1 and 2. Our experiments agree with this stronger conjecture.

We give explicit computations for the arrangements in Figures 6.1 and 6.5. The $\mathbb{Z}[\sigma^{\pm 1}]$-module $(1-\sigma)\mathrm{ab}(\ker\varphi)$ is generated by $\{(1-\sigma)\overline{s}_{0,i},\ , i \ne j\}$. We choose j as the last index in the natural ordering. All Abelianized relations are divisible by $(1-\sigma)$, so we just divide everything by $1-\sigma$ and verify that $\mathrm{ab}(\ker\varphi)$ is generated by $\mathrm{ab}(R_j)/(1-\sigma)$.

For the arrangement in Figure 6.1 we have to rewrite 13 relations coming from 11 double points and 1 triple point. After Abelianization we obtain:

(a) $\overline{s}_{0,2} - \overline{s}_{0,3}$;

(b) $\overline{s}_{0,2} - \overline{s}_{0,4}$;

(c) $\overline{s}_{0,3} - \overline{s}_{0,4}$;

(d) $\overline{s}_{0,1} - \overline{s}_{0,4}$;

(e) $\overline{s}_{0,1} - \overline{s}_{0,3}$;

(f) $\sigma\overline{s}_{0,2} + \overline{s}_{0,5}$;

(g) $(1+\sigma)\overline{s}_{0,2} - \sigma\overline{s}_{0,5}$;

(h) $\overline{s}_{0,1} + \sigma^{-1}(1-\sigma)\overline{s}_{0,2} - \sigma^{-2}(1+\sigma)\overline{s}_{0,5}$;

(i) $\overline{s}_{0,3} + (\sigma^{-1} - 1)\overline{s}_{0,5}$;

(j) $\overline{s}_{0,4} + (\sigma^{-1} - 1)\overline{s}_{0,5}$;

(k) $\overline{s}_{0,1} + (\sigma^{-1} - 1)\overline{s}_{0,2} - \sigma^{-1}\overline{s}_{0,5}$;

(l) $\overline{s}_{0,1} - \overline{s}_{0,2}$;

(m) $\overline{s}_{0,3} - \overline{s}_{0,5}$.

The generator $\overline{s}_{0,5}$ is obtained as $\sigma(\ (i) - (m)\ )$. From $\overline{s}_{0,5}$ we obtain in sequence all the other generators $\overline{s}_{0,3},\ \overline{s}_{0,1},\ \overline{s}_{0,4},\ \overline{s}_{0,2}$. According to Theorem 7.5 this gives the a-monodromicity of the arrangement in Figure 6.1.

For the arrangement $A3$ *deconed* in Figure 6.5 we have to rewrite two relations for each triple point and one relation for each double point. Their Abelianization is given by:

(a) $\overline{s}_{0,2} - \overline{s}_{0,3}$;

(b) $\sigma\overline{s}_{0,2} + \overline{s}_{0,4}$;

(c) $(\sigma + 1)\overline{s}_{0,2} - \sigma\overline{s}_{0,4}$;

(d) $\sigma\overline{s}_{0,1} + (1 - \sigma)\overline{s}_{0,2} + (\sigma^{-1})\overline{s}_{0,3} + (\sigma^{-2} - 1)\overline{s}_{0,4}$;

(e) $\overline{s}_{0,1} + (\sigma^{-1} - 1)\overline{s}_{0,2} - (\sigma^{-1})\overline{s}_{0,4}$;

(f) $(\sigma + 1)\overline{s}_{0,1} + (\sigma^{-1} - \sigma)\overline{s}_{0,2} - \overline{s}_{0,3} + (\sigma^{-2} - \sigma^{-1})\overline{s}_{0,4}$.

We perform the following base changes:

$(a') = (a)$;

$(b') = (b) - \sigma(a)$;

$(c') = (c) - (b) - (a)$;

$(d') = (d) - \sigma^{-2}(b) + \sigma^{-1}(a) - \sigma(e)$;

$(e') = (e) + \sigma^{-1}(b) - \sigma^{-1}(a)$;

$(f') = (f) - (\sigma^{-2} + \sigma^{-1} + 1)(b) + \sigma(a) + \sigma^{-1}(c) - (\sigma + 1)(e)$

and

$$\overline{s}'_{0,1} = \overline{s}_{0,1} + \sigma^{-1}\,\overline{s}_{0,3}; \qquad \overline{s}'_{0,2} = \overline{s}_{0,2} - \overline{s}_{0,3};$$
$$\overline{s}'_{0,3} = \overline{s}_{0,3}; \qquad \overline{s}'_{0,4} = \overline{s}_{0,4} + \sigma\,\overline{s}_{0,3}.$$

It is straighforward to verify, after these changes, that the submodule M generated by $\langle \mathrm{ab}(R_j) : \ j = 1, \dots, 6\rangle$ equals

$$\left\langle \overline{s}'_{0,1}, \overline{s}'_{0,2}, (1 + \sigma + \sigma^2)\overline{s}'_{0,3}, \overline{s}'_{0,4} \right\rangle.$$

So $M \subsetneq (1 - \sigma)\mathrm{ab}(\ker\varphi)$, in accordance with Theorem 7.5.

8. Further characterizations

In this section we give a more intrinsic picture. Let $\widetilde{\mathcal{A}} = \{H_0, H_1, \dots, H_n\}$ be the conified arrangement in $\mathbb{C}^3$. The fundamental group

$$\mathbf{G} = \pi_1(\mathcal{M}(\widetilde{\mathcal{A}})) \quad (= \pi_1(\mathcal{M}(\mathcal{A})) \times \mathbb{Z})$$

is generated by elementary loops $\beta_0, \dots, \beta_n$ around the hyperplanes. Let

$$\mathbf{F} = \mathbf{F_{n+1}}[\beta_0, \dots, \beta_n]$$

be the free group and $\mathbf{N}$ be the normal subgroup generated by the relations, so we have a presentation

$$1 \longrightarrow \mathbf{N} \longrightarrow \mathbf{F} \xrightarrow{\pi} \mathbf{G} \longrightarrow 1$$

The length map $\varphi : \mathbf{F} \to \langle t \rangle \cong \mathbb{Z}$ factors through π by a map

$$\psi : \mathbf{G} \to \mathbb{Z}.$$

Next, ψ factorizes through the Abelianization

$$\frac{\mathbf{G}}{[\mathbf{G},\mathbf{G}]} \cong H_1(\mathcal{M}(\widetilde{\mathcal{A}}); \mathbb{Z}) \cong \mathbb{Z}^{n+1} \cong \frac{\mathbf{F}}{[\mathbf{F},\mathbf{F}]}.$$

Let now $\mathbf{K} = \ker \psi$ so we have

$$1 \longrightarrow \mathbf{K} \longrightarrow \mathbf{G} \xrightarrow{\psi} \mathbb{Z} \longrightarrow 1, \tag{8.1}$$

and ψ factorizes through

$$\mathbf{G} \xrightarrow{\text{ab}} \frac{\mathbf{G}}{[\mathbf{G},\mathbf{G}]} \cong \mathbb{Z}^{n+1} \xrightarrow{\lambda} \mathbb{Z}.$$

We have the following commutative diagram:

$$\begin{tikzcd}
 & & 1 \arrow[d] & & \\
 & & \mathbf{N} \arrow[d] & & 1 \\
1 \arrow[r] & \ker(\varphi) \arrow[r] & \mathbf{F} \arrow[r, "\varphi"] \arrow[d, "\pi"'] & \mathbb{Z} \arrow[r] \arrow[ur] & 1 \\
1 \arrow[r] & {[\mathbf{G},\mathbf{G}]} \arrow[r] \arrow[d] & \mathbf{G} \arrow[r, "\text{ab}"] \arrow[ur, "\psi"] \arrow[d] & \frac{\mathbf{G}}{[\mathbf{G},\mathbf{G}]} \arrow[r] \arrow[u, "\lambda"'] & 1 \\
 & \mathbf{K} \arrow[ur] & 1 & & \\
1 \arrow[ur] & & & & .
\end{tikzcd} \tag{8.2}$$

Remark 8.1. One has $\ker(\lambda) = \frac{\mathbf{K}}{[\mathbf{G},\mathbf{G}]}$ so $\frac{\mathbf{K}}{[\mathbf{G},\mathbf{G}]} \cong \mathbb{Z}^n$.

Therefore diagram (8.2) extends to

$$
\begin{tikzcd}
 & & 1 \arrow[d] & & \\
 & & \mathbf{N} \arrow[d] & 1 & 1 \\
1 \arrow[r] & \ker(\varphi) \arrow[r] & \mathbf{F} \arrow[r, "\varphi"] \arrow[d, "\pi"'] & \mathbb{Z} \arrow[r] \arrow[u] \arrow[ur] & 1 \\
1 \arrow[r] & {[\mathbf{G},\mathbf{G}]} \arrow[r] \arrow[d] & \mathbf{G} \arrow[r, "\mathrm{ab}"] \arrow[ur, "\psi"] \arrow[d] & \frac{\mathbf{G}}{[\mathbf{G},\mathbf{G}]} \arrow[r] \arrow[u, "\lambda"'] & 1 \\
 & \mathbf{K} \arrow[ur] & 1 & \frac{\mathbf{K}}{[\mathbf{G},\mathbf{G}]} \arrow[u] & \\
1 \arrow[ur] & & & 1 \arrow[u] & .
\end{tikzcd}
\tag{8.3}
$$

Recall the $\mathbb{Z}[t^{\pm 1}]-$module isomorphism

$$H_1(\mathbf{G}; \mathbb{Z}[t^{\pm 1}]) \;\cong\; H_1(F; \mathbb{Z}) \tag{8.4}$$

where F is the Milnor fibre, and (by the Shapiro Lemma):

$$H_1(F; \mathbb{Z}) \cong H_1(\mathbf{K}; \mathbb{Z}) = \frac{\mathbf{K}}{[\mathbf{K}, \mathbf{K}]}. \tag{8.5}$$

Remark 8.2. There is an exact sequence

$$1 \longrightarrow \frac{[\mathbf{G}, \mathbf{G}]}{[\mathbf{K}, \mathbf{K}]} \longrightarrow \frac{\mathbf{K}}{[\mathbf{K}, \mathbf{K}]} \longrightarrow \frac{\mathbf{K}}{[\mathbf{G}, \mathbf{G}]} \cong \mathbb{Z}^n \longrightarrow 1. \tag{8.6}$$

From the definition before Theorem 7.2 one has

Lemma 8.3. *The arrangement $\widetilde{\mathcal{A}}$ is a-monodromic over $\mathbb{Z}$ if and only if*

$$H_1(F; \mathbb{Z}) \cong \mathbb{Z}^n.$$

It follows:

Theorem 8.4. *The arrangement $\widetilde{\mathcal{A}}$ is a-monodromic over $\mathbb{Z}$ if and only if*

$$\frac{[\mathbf{G},\mathbf{G}]}{[\mathbf{K},\mathbf{K}]} = 0. \tag{8.7}$$

Proof. It immediately follows from sequence 8.6 and from the property that a surjective endomorphism of a finitely generated free Abelian group is an isomorphism. □

Since $\mathbf{K} \supset [\mathbf{G},\mathbf{G}]$ it follows immediately (see Corollary 7.4) that:

Corollary 8.5. *If*

$$\mathbf{G}^{(1)} = [\mathbf{G},\mathbf{G}] = \mathbf{G}^{(2)} = [[\mathbf{G},\mathbf{G}],[\mathbf{G},\mathbf{G}]],$$

then the arrangement $\widetilde{\mathcal{A}}$ is a-monodromic.

We also have:

Corollary 8.6. *Let $\mathbf{G}$ have a central element of length 1. Then the arrangement $\widetilde{\mathcal{A}}$ is a-monodromic.*

Proof. Let $\gamma \in \mathbf{G}$ be a central element of length 1. From sequence (8.1) the group splits as a direct product

$$\mathbf{G} \cong \mathbf{K} \times \mathbb{Z}$$

where $\mathbb{Z} = <\gamma>$. Therefore clearly $[\mathbf{G},\mathbf{G}] = [\mathbf{K},\mathbf{K}]$. □

An example of the situation of the corollary is when one of the generators β_j commutes with all the others, *i.e.*, one hyperplane is transversal to the others. So, we find again in this way a well-known fact.

Consider again the exact sequence (8.6). Remind that the arrangement $\widetilde{\mathcal{A}}$ is a-monodromic (over $\mathbb{Q}$) if and only if $H_1(F;\mathbb{Q}) \cong \mathbb{Q}^n$. By tensoring sequence (8.6) by $\mathbb{Q}$ we obtain:

Theorem 8.7. *The arrangement $\widetilde{\mathcal{A}}$ is a-monodromic (over $\mathbb{Q}$) if and only if*

$$\frac{[\mathbf{G},\mathbf{G}]}{[\mathbf{K},\mathbf{K}]} \otimes \mathbb{Q} = 0.$$

Remark 8.8. All remarkable questions about the H_1 of the Milnor fibre F are actually questions about the group

$$\frac{[\mathbf{G},\mathbf{G}]}{[\mathbf{K},\mathbf{K}]}.$$

In particular:

(1) $H_1(F;\mathbb{Z})$ has torsion if and only if $\frac{[\mathbf{G},\mathbf{G}]}{[\mathbf{K},\mathbf{K}]}$ has torsion;

(2) $b_1(F) = n + rk\left(\frac{[\mathbf{G},\mathbf{G}]}{[\mathbf{K},\mathbf{K}]}\right)$.

(There are only complicated examples with torsion in higher homology of the Milnor fiber, recently found in [14].)

Corollary 8.9. *One has*

$$n \le b_1(F) \le n + rk\left(\frac{[\mathbf{G},\mathbf{G}]}{[[\mathbf{G},\mathbf{G}],[\mathbf{G},\mathbf{G}]]}\right) = n + rk\left(\frac{\mathbf{G}^{(1)}}{\mathbf{G}^{(2)}}\right).$$

Now we consider again the affine arrangement $\mathcal{A}$. Denoting by $\mathbf{G}' := \pi_1(\mathcal{M}(\mathcal{A}))$, we have

$$\mathbf{G} \cong \mathbf{G}' \times \mathbb{Z},$$

where the factor $\mathbb{Z}$ is generated by a loop around all the hyperplanes in $\widetilde{\mathcal{A}}$. As already said, it follows by the Kunneth formula that if $\mathcal{A}$ has trivial monodromy over $\mathbb{Z}$ (respectively $\mathbb{Q}$), so does $\widetilde{\mathcal{A}}$. Conversely, in Figure 6.7 we have an example where $\widetilde{\mathcal{A}}$ is a-monodromic but $\mathcal{A}$ has non-trivial monodromy.

The a-monodromicity of $\mathcal{A}$ (over $\mathbb{Z}$) is equivalent to

$$H_1(\mathcal{M}(\mathcal{A}); R) \cong \left(\frac{R}{(t-1)}\right)^{n-1}, \tag{8.8}$$

($R = \mathbb{Z}[q^{\pm 1}]$). By considering a sequence as in (8.1),

$$1 \longrightarrow \mathbf{K}' \longrightarrow \mathbf{G}' \xrightarrow{\psi} \mathbb{Z} \longrightarrow 1, \tag{8.9}$$

we can repeat the above arguments: in particular condition (8.8) is equivalent to

$$H_1(\mathbf{K}';\mathbb{Z}) = \frac{\mathbf{K}'}{[\mathbf{K}',\mathbf{K}']} = \mathbb{Z}^{n-1},$$

and we get an exact sequence like in (8.6) for $\mathbf{K}'$ and $\mathbf{G}'$. So we obtain:

Theorem 8.10. *The arrangement $\mathcal{A}$ is a-monodromic over $\mathbb{Z}$ (respectively over $\mathbb{Q}$) if and only if*

$$\frac{[\mathbf{G}',\mathbf{G}']}{[\mathbf{K}',\mathbf{K}']} = 0 \left(\textit{respectively } \frac{[\mathbf{G}',\mathbf{G}']}{[\mathbf{K}',\mathbf{K}']} \otimes \mathbb{Q} = 0\right).$$

By considering a presentation for $\mathbf{G}'$

$$1 \longrightarrow \mathbf{N}' \longrightarrow \mathbf{F}' \xrightarrow{\pi} \mathbf{G}' \longrightarrow 1$$

where $\mathbf{F}'$ is the group freely generated by $\beta_1, \ldots, \beta_n$, we have a diagram similar to (8.3) for $\mathbf{G}'$. From $\mathbf{N}' \subset [\mathbf{F}', \mathbf{F}'] \subset \ker\varphi$ we have isomorphisms

$$\frac{[\mathbf{G}', \mathbf{G}']}{[\mathbf{K}', \mathbf{K}']} \cong \frac{\pi^{-1}[\mathbf{G}', \mathbf{G}']}{\pi^{-1}[\mathbf{K}', \mathbf{K}']} \cong \frac{[\mathbf{F}', \mathbf{F}']}{\mathbf{N}'[\ker\varphi, \ker\varphi]},$$

which gives again Theorem 7.2.

Corollary 8.6 extends clearly to the affine case: therefore, if one line of $\mathcal{A}$ is in general position with respect to the others, then $\mathcal{A}$ is a-monodromic. This result has the following useful generalization, which has both a central and an affine versions. We give here the affine one.

Theorem 8.11. *Assume that the fundamental group G' decomposes as a direct product*

$$G' = A \times B$$

of two subgroups, each one having at least one element of length one. Then $\mathcal{A}$ is a-monodromic.

In particular, this applies to the case when G' decomposes as a direct product of free groups,

$$G' = F_{i_1} \times F_{i_2} \times \cdots \times F_{i_k}$$

where (at least) two of them have an element of length one.

Proof. First, observe that any commutator $[ab, a'b'] \in [G', G']$ equals $[a, a'][b, b']$. Therefore it is sufficient to show that $[A, A] \subset [K', K']$, and $[B, B] \subset [K', K']$. Let $a_0 \in A$, $b_0 \in B$ be elements of length one. Let $l = \psi(a)$, $l' = \psi(a')$ be the lengths of a and a' respectively. Then

$$[a, a'] = [ab_0^{-l}, a'b_0^{-l'}],$$

and the second commutator lies in $[K', K']$ by construction. This proves that $[A, A] \subset [K', K']$.

In the same way, by using a_0, we show that $[B, B] \subset [K', K']$. $\square$

Remark 8.12. This theorem includes the case when the arrangement is a disjoint union $\mathcal{A} = \mathcal{A}' \sqcup \mathcal{A}''$ of two subarrangements which intersect each other transversally. It is known that $\pi_1(\mathcal{M}(\mathcal{A}))$ is the direct product of $\pi_1(\mathcal{M}(\mathcal{A}'))$ with $\pi_1(\mathcal{M}(\mathcal{A}''))$ (see [27]) therefore by Theorem 8.11 the arrangement $\mathcal{A}$ is a-monodromic. This remark also seems new in the literature.

We can use this result (or even Corollary 8.6) to prove the a-monodromicity of those examples in Section 6 for which the fundamental group splits as a direcy product of free groups.

Another example is given by any affine arrangement having only double points: in this case $\mathcal{A} = \cup_{i=1}^{k} \mathcal{A}_i$ where the $\mathcal{A}_i$'s are sets of parallel lines. Then $\pi_1(\mathcal{A}) = \times_{i=1}^{k} F_{n_i}$ where F_{n_i} is the free group in $n_i = |A_i|$ generators. This gives an easy prove of the following known fact: if there exists a line in a projective arrangement $\mathcal{A}$ which contains all the points of multiplicity greater than 3, then $\mathcal{A}$ is a-monodromic.

To take care also of examples as that in Figure 6.2, where the fundamental group is not a direct product of free groups, let us introduce another class of graphs $\tilde{\Gamma}$ as follows. Let the affine arrangement $\mathcal{A}$ have n lines. Then:

(1) the vertex set of $\tilde{\Gamma}$ corresponds to the set of generators $\{\beta_i,\ i = 1, \dots, n\}$ of G';
(2) for each edge (β_i, β_j) of $\tilde{\Gamma}$, the commutator $[\beta_i, \beta_j]$ belongs to $[K', K']$;
(3) $\tilde{\Gamma}$ is connected.

We call a graph $\tilde{\Gamma}$ satisfying the previous conditions an *admissible* graph.

Theorem 8.13. *If $\mathcal{A}$ allows an admissible graph $\tilde{\Gamma}$ then $\mathcal{A}$ is a-monodromic.*

We need the following lemma.

Lemma 8.14. *Let $F_n = F[\beta_1, \dots, \beta_n]$ be the free group in the generators β_i's. Let φ be the length function (see Section 7) on F_n. Then for any sequence of indices $i_0, \dots, i_k$ one has*

$$[\beta_{i_0}, \beta_{i_1}][\beta_{i_1}, \beta_{i_2}] \dots [\beta_{i_{k-1}}, \beta_{i_k}][\beta_{i_k}, \beta_{i_0}] \in [\ker(\varphi), \ker(\varphi)]$$

for each "closed" product of commutators.

Proof of lemma. If $k \leq 2$ the result is trivial. If $k = 3$, a straighforward application of Blanchfield theorem ([4]) gives the result. For $k > 3$, we can write

$$\begin{aligned}&[\beta_{i_0}, \beta_{i_1}][\beta_{i_1}, \beta_{i_2}] \dots [\beta_{i_{k-1}}, \beta_{i_k}][\beta_{i_k}, \beta_{i_0}] \\ &\quad = ([\beta_{i_0}, \beta_{i_1}][\beta_{i_1}, \beta_{i_2}][\beta_{i_2}, \beta_{i_0}])([\beta_{i_0}, \beta_{i_2}] \dots [\beta_{i_{k-1}}, \beta_{i_k}][\beta_{i_k}, \beta_{i_0}]),\end{aligned}$$

and we conclude by induction on k. □

Remark 8.15. Clearly, Lemma 8.14 applied to the generators of G' gives that

$$[\beta_{i_0}, \beta_{i_1}][\beta_{i_1}, \beta_{i_2}] \dots [\beta_{i_{k-1}}, \beta_{i_k}][\beta_{i_k}, \beta_{i_0}] \in [K', K']$$

for each closed product of commutators.

Proof of Theorem 8.13. According to Theorem 8.10 what we have to prove is that any commutator $[\beta_i, \beta_j]$ belongs to $[K', K']$.

If i, j corresponds to an edge (β_i, β_j) of $\tilde{\Gamma}$, the result follows by definition. Otherwise, let $\beta_i = \beta_{i_0}, \beta_{i_1}, \dots, \beta_{i_k} = \beta_j$ be a path in $\tilde{\Gamma}$ connecting β_i with β_j. By definition, $[\beta_{i_j}, \beta_{i_{j+1}}] \in [K', K']$, $j = 0, \dots, k-1$, so $\prod_{j=0}^{k-1}[\beta_{i_j}, \beta_{i_{j+1}}] \in [K', K']$. By Lemma 8.14 and Remark 8.15

$$[\beta_{i_0}, \beta_{i_1}][\beta_{i_1}, \beta_{i_2}] \dots [\beta_{i_{k-1}}, \beta_{i_k}][\beta_{i_k}, \beta_{i_0}] \in [K', K'].$$

It follows that $[\beta_{i_0}, \beta_{i_k}] = [\beta_i, \beta_j] \in [K', K']$, which gives the thesis. □

We can use Theorem 8.13 to prove Conjecture 1 under further hypothesis.

Corollary 8.16. *Let $\mathcal{A}$ be an affine arrangement and let Γ be its associated graph of double points. Assume that Γ contains an admissible spanning tree $\tilde{\Gamma}$. Then $\mathcal{A}$ is a-monodromic.*

Of course, under the hypothesis of Corollary 8.16, the graph Γ is connected. Examples where Γ contains an admissible spanning tree are the conjugate-free arrangements in Definition 5.6. Here all commutators (corresponding to the edges of T) of the geometric generators are simply equal to 1 in the group G'. Therefore Theorem 8.13 is a generalization of Theorem 5.7.

Very little effort is needed to show that the whole graph Γ of double points in the arrangement of Figure 6.2 is admissible: therefore Corollary 8.16 applies to this case.

For the sake of completeness, we also mention that, for all the examples in Section 6 which have non trivial monodromy, all the quotient groups $[G', G']/[K', K']$ are free Abelian of rank 2. This fact is in accordance with the monodromy computations given in Section 6, since in all these cases one has φ_3-torsion. It also follows that, for such examples, the first homology group of the Milnor fiber has no torsion.

Remark 8.17. When the graph Γ of double points is not connected, then we can consider its decomposition into connected components $\Gamma = \sqcup_i \Gamma_i$. We have a corresponding decomposition $\mathcal{A} = \sqcup_i \mathcal{A}_i$ of the arrangement. By definition, every double point of $\mathcal{A}$ is a double point of exactly one of the $\mathcal{A}_i$'s, while each pair of lines in different $\mathcal{A}_i$'s either intersect in some point of multiplicity greater than 2, or are parallel (we are considering the affine case here). If our conjecture is true, then each $\mathcal{A}_i$ is a-monodromic. At the moment we are not able to speculate about how the monodromy of $\mathcal{A}$ is influenced by these data: apparently, the only knowledge of such decomposition gives little control on the multiplicities of the intersection points of different components, which can assume very different values. We are going to address these interesting problems in future work.

References

[1] D. ARAPURA, *Geometry of cohomology support loci for local systems I*, J. Algebraic Geom. **6** (1997), 563–597.

[2] P. BAILET, *On the monodromy of the Milnor fiber of hyperplane arrangements*, Canad. Math. Bull. **57** (2014), 697–707.

[3] N. BIGGS, "Algebraic Graph Theory", Cambridge University Press, 1974.

[4] J. BIRMAN, "Braids, Links and Mapping Class Groups", Princeton Univ. Press, 1975.

[5] F. CALLEGARO, *The homology of the Milnor fibre for classical braid groups*, Algebr. Geom. Topol. **6** (2006), 1903–1923.

[6] F. CALLEGARO, D. MORONI and M. SALVETTI, *Cohomology of Artin groups of type $\tilde{A}_n$, $\tilde{B}_n$ and applications*, In: "Proceedings of the Conference on Groups Homotopy and Configuration Spaces", N. Iwase (ed.), Geometry and Topology Publications, Coventry, 2008, 85–104.

[7] F. CALLEGARO, D. MORONI and M. SALVETTI, *Cohomology of affine Artin groups and applications*, Trans. Amer. Math. Soc. **360** (2008), 4169–4188.

[8] F. CALLEGARO, D. MORONI and M. SALVETTI, *The $K(\pi, 1)$ problem for the affine Artin group of type $\tilde{B}_n$ and its cohomology*, J. Eur. Math. Soc. (JEMS) **12** (2010), 1–22.

[9] D. COHEN and P. ORLIK, *Arrangements and local systems*, Math. Res. Lett. **7** (2000), 299–316.

[10] D. COHEN and A. SUCIU, *Characteristic varieties of arrangements*, Math. Proc. Cambridge Philos. Soc. **127** (1999), 33–53.

[11] C. DE CONCINI, C. PROCESI and M. SALVETTI, *Arithmetic properties of the cohomology of braid groups*, Topology **40** (2001), 739–751.

[12] C. DE CONCINI, C. PROCESI, M. SALVETTI and F. STUMBO, *Arithmetic properties of the cohomology of Artin groups*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28** (1999), 695–717.

[13] C. DE CONCINI and M. SALVETTI, *Cohomology of Artin groups*, Addendum: "The homotopy type of Artin groups", Math. Res. Lett. **3** (1996), 293–297.

[14] G. DENHAM and A. SUCIU, *Multinets, parallel connections, and Milnor fibrations of Arrangements*, Proc. Lond. Math. Soc. (3) **108** (2014), 1435–1470.

[15] A. DIMCA and S. PAPADIMA, *Hypersurface complements, Milnor fibers and higher homotopy groups of arrangements*, Ann. of Math. (2) **158** (2003), 473–507.

[16] A. DIMCA, S. PAPADIMA and A. SUCIU, *Topology and geometry of cohomology jump loci*, Duke Math. J. **148** (2009), 405–457.

[17] M. ELIYAHU, D. GARBER and M. TEICHER, *A conjugation-free geometric presentation of fundamental groups of arrangements*, Manuscripta Math. **133** (2010), 247–271.

[18] M. J. FALK, *Arrangements and cohomology*, Ann. Comb. **1** (1997), 135–157.

[19] M. J. FALK and S. YUZVINSKY, *Multinets, resonance varieties and pencils of plane curves*, Compos. Math. **143** (2007), 10069–1088.

[20] R. FORMAN, *Morse Theory for Cell Complexes*, Adv. Math. **134** (1998),90–145.

[21] E. V. FRENKEL, *Cohomology of the commutator subgroup of the braid group*, Funct. Anal. Appl. **22** (1988), 248–250.

[22] G. GAIFFI and M. SALVETTI *The Morse complex of a line arrangement*, J. Algebra **321** (2009), 316–337.

[23] G. GAIFFI, F. MORI and M. SALVETTI, *Minimal CW-complexes for Complement to Line Arrangements via Discrete Morse Theory*, In: "Topology of Alg Var. and Sing.", Amer. Math. Soc., Contemporary Math. Vol. 538, 2011, 293–308.

[24] A. LIBGOBER and S. YUZVINSKY, *Cohomology of the Orlik-Solomon algebras and local systems*, Compos. Math. **21** (2000), 337–361.

[25] A. MĂCINIC and S. PAPADIMA, *On the monodromy action on Milnor fibers of graphic arrangements*, Topology Appl. **156** (2009), 761–774.

[26] P. ORLIK and M. TERAO, "Arrangements of Hyperplanes", Grundl. der Math. Wiss., Vol. 300, Springer-Verlag, 1992.

[27] M. OKA and K. SAKAMOTO, *Product theorem of the fundamental group of a reducible curve*, J. Math. Soc. Japan **30** (1978), 599–602.

[28] R. RANDELL, *Morse theory, Milnor fibers and minimality of a complex hyperplane arrangement*, Proc. Amer. Math. Soc. **130** (2002), 2737–2743.

[29] M. SALVETTI, *Topology of the complement of real hyperplanes in* $\mathbb{C}^n$, Invent. Math. **88** (1987), 603–618.

[30] M. SALVETTI, *The homotopy type of Artin groups*, Math. Res. Lett. **1** (1994), 567–577.

[31] S. SETTEPANELLA, *Cohomology of pure braid groups of exceptional cases*, Topol. Appl. **156** (2009), 1008–1012.

[32] M. SALVETTI and M. SERVENTI, *Arrangements of lines and monodromy of associated Milnor fibers*, J. Knot Theory Ramif. **26** (2016).

[33] M. SALVETTI and S. SETTEPANELLA, *Combinatorial Morse theory and minimality of hyperplane arrangements*, Geom. Topol. **11** (2007), 1733–1766.

[34] A. SUCIU, *Translated tori in the characteristic varieties of complex hyperplane arrangements*, Topol. Appl. **118** (2002), 209–223.

[35] A. SUCIU, *Hyperplane arrangements and Milnor fibrations*, Ann. Fac. Sci. Toulouse Math. (6) **23** (2014), 417–481.

[36] A. SUCIU, *Around the tangent cone theorem*, In "Configuration Spaces: Geometry, Topology and Representation Theory", Springer INdAM Series, F. Callegaro, F. Cohen, C. De Concini, E. Feichtner, G. Gaiffi, M. Salvetti (eds.), 2016, 1–39.

[37] M. YOSHINAGA, *Resonant bands and local system cohomology groups for real line arrangements*, Vietnam J. Math. **42** (2014), 377–392.

[38] M. YOSHINAGA, *Milnor fibers of real line arrangements*, J. Singul. **7** (2013), 220–237.

Departimento di Matematica
Università di Pisa
Largo Pontecorvo, 5
56127 Pisa, Italia
salvetti@dm.unipi.it
serventi@mail.dm.unipi.it

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1491-1519

Lipschitz changes of variables between perturbations of log-concave measures

MARIA COLOMBO, ALESSIO FIGALLI AND YASH JHAVERI

Abstract. Extending a result of Caffarelli, we provide global Lipschitz changes of variables between compactly supported perturbations of log-concave measures. The result is based on a combination of ideas from optimal transportation theory and a new Pogorelov-type estimate. In the case of radially symmetric measures, Lipschitz changes of variables are obtained for a much broader class of perturbations.

Mathematics Subject Classification (2010): 49Q20 (primary); 35J96, 26D10 (secondary).

1. Introduction

In [4], Caffarelli built Lipschitz changes of variables between log-concave probability measures. More precisely, he showed that if $V, W \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ are convex functions with $D^2V(x) \le \Lambda_V \,\mathrm{Id}$ and $\lambda_W \,\mathrm{Id} \le D^2W(x)$ for a.e. $x \in \mathbb{R}^n$ with $0 < \Lambda_V, \lambda_W < \infty$, then there exists a Lipschitz map $T : \mathbb{R}^n \to \mathbb{R}^n$ such that $T_\#\big(e^{-V(x)}\,dx\big) = e^{-W(x)}\,dx$ [1] and

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \le \sqrt{\Lambda_V/\lambda_W}. \tag{1.1}$$

[1] Given two finite Borel measures μ and ν and a Borel map $T : \mathbb{R}^n \to \mathbb{R}^n$, recall that $T_\#\mu = \nu$ if

$$\int_{\mathbb{R}^n} \varphi(y)\,d\nu(y) = \int_{\mathbb{R}^n} \varphi(T(x))\,d\mu(x) \qquad \forall\,\varphi \text{ Borel and bounded.}$$

M. Colombo acknowledges the support of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM), of Dr. Max Rössler, of the Walter Haefner Foundation and of the ETH Zurich Foundation. A. Figalli has been partially supported by NSF Grant DMS-1262411 and NSF Grant DMS-1361122.

Received October 13, 2015; accepted September 27, 2016.
Published online December 2017.

The map T is obtained via optimal transportation. It is the unique solution of the Monge problem for quadratic cost

$$\min\left\{\int_{\mathbb{R}^n} |x - T(x)|^2 e^{-V(x)}\,dx : T_{\#}\big(e^{-V(x)}\,dx\big) = e^{-W(x)}\,dx\right\}$$

(see Section 2 for more details and [12] for a completely different construction of a Lipschitz change of variables in this setting). We note that a particularly important feature of Caffarelli's result is that the bound (1.1) is *independent* of the dimension n.

A consequence of Caffarelli's result is the possible deduction of certain functional inequalities (such as log-Sobolev or Poincaré-type inequalities) for log-concave measures from their corresponding Gaussian versions. For instance, denoting the standard Gaussian measure on $\mathbb{R}^n$ by γ_n, consider the Gaussian log-Sobolev inequality,

$$\int_{\mathbb{R}^n} f^2 \ln f\,d\gamma_n \le \int_{\mathbb{R}^n} |\nabla f|^2\,d\gamma_n + \left(\int_{\mathbb{R}^n} f^2\,d\gamma_n\right)\ln\left(\int_{\mathbb{R}^n} f^2\,d\gamma_n\right),$$

which holds for every function $f \in W^{1,2}(\mathbb{R}^n)$. For any measure ν such that there exists a Lipschitz change of variables between ν and the Gaussian measure, namely $\nu = T_{\#}\gamma_n$, we deduce, applying twice the change of variable formula, that

$$\begin{aligned}
&\int_{\mathbb{R}^n} f^2 \ln f\,d\nu = \int_{\mathbb{R}^n} f(T)^2 \ln f(T)\,d\gamma_n \\
&\le \int_{\mathbb{R}^n} |\nabla[f\circ T]|^2\,d\gamma_n + \left(\int_{\mathbb{R}^n} f(T)^2\,d\gamma_n\right)\ln\left(\int_{\mathbb{R}^n} f(T)^2\,d\gamma_n\right) \\
&\le \|\nabla T\|^2_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n} |\nabla f(T)|^2\,d\gamma_n + \left(\int_{\mathbb{R}^n} f(T)^2\,d\gamma_n\right)\ln\left(\int_{\mathbb{R}^n} f(T)^2\,d\gamma_n\right) \\
&= \|\nabla T\|^2_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n} |\nabla f|^2\,d\nu + \left(\int_{\mathbb{R}^n} f^2\,d\nu\right)\ln\left(\int_{\mathbb{R}^n} f^2\,d\nu\right).
\end{aligned}$$

Therefore, ν enjoys a log-Sobolev inequality with constant $\|\nabla T\|^2_{L^\infty(\mathbb{R}^n)}$.

Besides the natural consequences described in [4] and above, Caffarelli's Theorem has found numerous applications in various fields: indeed, it can be used to transfer isoperimetric inequalities, to obtain correlation inequalities, and more (see, for instance, [6, 7, 11, 13]). Some recent extensions and variations of Caffarelli's Theorem can be found in, for example, [14, 15, 17].

In this paper, we extend the result of Caffarelli by building Lipschitz changes of variables between perturbations of V and W that are not necessarily convex. Perturbations of log-concave measures (in particular, perturbations of Gaussian measures) appear, for instance, in quantum physics as a means to help understanding solutions to physical theories with nonlinear equations of motion. In cases where an explicit solution is unknown, perturbations of log-concave measures can be used to yield approximate solutions.

We let $\mathcal{P}(X)$ denote the space of probability measures on a metric space X. The main result of the paper is the following:

Theorem 1.1. *Let $V \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ be such that $e^{-V(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. Suppose that $V(0) = \inf_{\mathbb{R}^n} V$ and there exist constants $0 < \lambda, \Lambda < \infty$ for which $\lambda \,\mathrm{Id} \leq D^2V(x) \leq \Lambda \,\mathrm{Id}$ for a.e. $x \in \mathbb{R}^n$. Moreover, let $R > 0$, $q \in C^0_c(B_R)$, and $c_q \in \mathbb{R}$ be such that $e^{-V(x)+c_q-q(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. Assume that $-\lambda_q \,\mathrm{Id} \leq D^2q$ in the sense of distributions for some constant $\lambda_q \geq 0$. Then, there exists a constant $C = C(R, \lambda, \Lambda, \lambda_q) > 0$, independent of n, such that the optimal transport map T that takes $e^{-V(x)}\,dx$ to $e^{-V(x)+c_q-q(x)}\,dx$ satisfies*

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq C. \tag{1.2}$$

The crucial point here is that the estimate on the Lipschitz constant of the optimal transport map is independent of dimension, as it is in Caffarelli's results for log-concave measures.

In the case of spherically symmetric measures, we are able to weaken the assumptions on both the log-concave measure and its perturbation and still obtain a global Lipschitz change of variables. In particular, the Lipschitz constant is controlled only by the L^∞-norm of the positive and negative parts of the perturbation q, denoted by q^+ and q^-. In the following theorem, we first analyze the 1-dimensional problem:

Theorem 1.2. *Let $V : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ be a convex function and $q : \mathbb{R} \to \mathbb{R}$ be a bounded function such that $e^{-V(x)}\,dx$, $e^{-V(x)-q(x)}\,dx \in \mathcal{P}(\mathbb{R})$. Then, the optimal transport T that takes $e^{-V(x)}\,dx$ to $e^{-V(x)-q(x)}\,dx$ is Lipschitz and satisfies*

$$\|\log T'\|_{L^\infty(\mathbb{R})} \leq \|q^+\|_{L^\infty(\mathbb{R})} + \|q^-\|_{L^\infty(\mathbb{R})}. \tag{1.3}$$

We remark that while the map T in Theorem 1.2 is only unique up to sets of $e^{-V(x)}\,dx$-measure zero, arguing by approximation, we can find a particular transport T for which the estimate on $\log T'$ in (1.3) is satisfied almost everywhere in $\mathbb{R}$. Applying this 1-dimensional result to radially symmetric densities, we obtain the following:

Theorem 1.3. *Let $V : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a convex, radially symmetric function and $q : \mathbb{R}^n \to \mathbb{R}$ be a bounded, radially symmetric function such that $e^{-V(x)}\,dx$, $e^{-V(x)-q(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. Then, the optimal transport T that takes $e^{-V(x)}\,dx$ to $e^{-V(x)-q(x)}\,dx$ is Lipschitz and satisfies*

$$e^{-\|q^+\|_{L^\infty(\mathbb{R}^n)}-\|q^-\|_{L^\infty(\mathbb{R}^n)}}\,\mathrm{Id} \leq \nabla T(x) \leq e^{\|q^+\|_{L^\infty(\mathbb{R}^n)}+\|q^-\|_{L^\infty(\mathbb{R}^n)}}\,\mathrm{Id} \quad \text{for a.e. } x \in \mathbb{R}^n. \tag{1.4}$$

Note that the assumption $e^{-V(x)-q(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$ in Theorems 1.2 and 1.3, unlike in Theorem 1.1, is nonrestrictive. Since q is not required to be compactly supported, the normalization constant making $e^{-V(x)-q(x)}\,dx$ a probability measure if it were not already can simply be absorbed into q.

We further remark that the 1-dimensional estimate in Theorem 1.2 is false in higher dimensions when one does not assume that the densities are radially symmetric. More precisely, taking the reference measure $e^{-V(x)}\,dx$ to be the standard Gaussian measure, the estimate

$$\|D^2\phi - \mathrm{Id}\,\|_{L^\infty(\mathbb{R}^n)} \leq C\|q\|_{L^\infty(\mathbb{R}^n)} \tag{1.5}$$

cannot be true for $n > 1$ (see Remark 5.2 to understand the relationship between (1.3) and (1.5) for $n = 1$). This is manifest if we recall that the Monge-Ampère equation linearizes to the Poisson equation, which does not enjoy $C^{1,1}_{\mathrm{loc}}$ estimates for bounded right-hand side. In other words, given V and q to be chosen, letting ϕ_ε be the potential such that $\nabla\phi_\varepsilon$ takes $e^{-V(x)}\,dx$ to $e^{-V(x)-\varepsilon q(x)}\,dx$ (for simplicity, we omit the normalization constant that makes $e^{-V(x)-\varepsilon q(x)}\,dx$ a probability measure), and setting $\psi_\varepsilon(x) = (\phi_\varepsilon(x) - |x|^2/2)/\varepsilon$, we have that

$$\begin{aligned}\Delta\psi_\varepsilon + O(\varepsilon) = \frac{\log\det\nabla^2\phi_\varepsilon}{\varepsilon} &= \frac{-V + V(\nabla\phi_\varepsilon) + \varepsilon q(\nabla\phi_\varepsilon)}{\varepsilon}\\ &= \langle x, \nabla\psi_\varepsilon\rangle + q(\nabla\phi_\varepsilon) + O(\varepsilon)\end{aligned}$$

for every $\varepsilon > 0$. The estimate (1.5) implies that $\sup_{\varepsilon>0}\|D^2\psi_\varepsilon\|_{L^\infty(\mathbb{R}^n)} < \infty$ and, therefore, the existence of a $C^{1,1}_{\mathrm{loc}}$ solution to the Poisson equation with bounded right-hand side, an impossibility in higher dimensions.

Although this heuristic argument is convincing, the details of the proof are rather delicate, and we give them in the Appendix for completeness.

ACKNOWLEDGEMENTS. Y. Jhaveri would like to thank Pablo Stinga for helpful conversations. Part of this work was done while the authors were guests of the FIM at ETH Zürich in the Fall of 2014; the hospitality of the Institute is gratefully acknowledged.

2. Preliminaries

We begin with some preliminaries on optimal transportation and the Monge-Ampère equation, and we fix some notation.

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$. The Monge optimal transport problem for quadratic cost consists of finding the most efficient way to take μ to ν given that the transportation cost to move from a point x to a point y is $|x - y|^2$. Hence, one is led to minimize

$$\mathrm{cost}(T) := \int_{\mathbb{R}^n} |x - T(x)|^2\,d\mu(x)$$

among all maps T such that $T_{\#}\mu = \nu$. A relaxed formulation of Monge's problem, due to Kantorovich, is to minimize

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 \, d\pi(x, y)$$

among all transport plans π, namely the measures $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ whose marginals are μ and ν. By a classical theorem of Brenier [2], the existence and uniqueness of an optimal transport plan are guaranteed when μ is absolutely continuous and μ and ν have finite second moments. Additionally, the optimality of a transport plan π is equivalent to $\pi = (\mathrm{Id} \times \nabla\phi)_{\#}\mu$ where ϕ is a convex function, often called the potential associated to the optimal transport. As a consequence, it follows that in the Monge problem, unique optimal maps exist as gradients of convex functions.

Theorem 2.1. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ such that $\mu = f(x)\,dx$ and*

$$\int_{\mathbb{R}^n} |x|^2 \, d\mu(x) + \int_{\mathbb{R}^n} |y|^2 \, d\nu(y) < \infty.$$

Then, there exists a unique (up to sets of μ-measure zero) optimal transport T taking μ to ν. Moreover, there is a convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ such that $T = \nabla\phi$.

A direct consequence of Brenier's characterization of optimal transports as gradients of convex functions is that

$$\langle x - y, T(x) - T(y)\rangle \geq 0 \qquad \text{for a.e. } x, y \in \mathbb{R}^n, \tag{2.1}$$

which follows immediately from the monotonicity of gradients of convex functions.

Suppose now that $\mu = f(x)\,dx$ and $\nu = g(y)\,dy$, and let ϕ be a convex function such that $T = \nabla\phi$ for T the optimal transport that takes μ to ν. Assuming that $T = \nabla\phi$ is a smooth diffeomorphism, the standard change of variables formula implies that

$$f(x) = g(T(x)) \det \nabla T(x).$$

Hence, assuming that $g > 0$, we see that ϕ is a solution to the Monge-Ampère equation

$$\det D^2\phi = \frac{f}{g \circ \nabla\phi}.$$

This formal link between optimal transportation and Monge-Ampère (since, to deduce the above equation, we assumed that T was already smooth) is at the heart of the regularity of optimal transport maps (see, for instance, [8] for more details). In particular, Caffarelli showed the following in [3] (see also [9, Theorem 4.5.2]):

Theorem 2.2. *Let $X, Y \subset \mathbb{R}^n$ be bounded open sets, and $f : X \to \mathbb{R}^+$ and $g : Y \to \mathbb{R}^+$ be probability densities locally bounded away from zero and infinity. If Y is convex, then for any set $X' \subset\subset X$, the optimal transport $T = \nabla\phi : X \to Y$ between $f(x)\,dx$ and $g(y)\,dy$ is of class $C^{0,\alpha}(X')$ for some $\alpha > 0$. In addition, if $f \in C^{k,\beta}_{\mathrm{loc}}(X)$ and $g \in C^{k,\beta}_{\mathrm{loc}}(Y)$ for some $k \in \mathbb{N} \cup \{0\}$ and $\beta \in (0, 1)$, then $\phi \in C^{k+2,\beta}_{\mathrm{loc}}(X)$.*

As mentioned in [1], Caffarelli's regularity result on optimal transports can be extended to the case where f and g are defined on all of $\mathbb{R}^n$ and assumed to be locally bounded away from zero and infinity. Lastly, we note that optimal transport maps are stable under approximation (see [18]). In particular, let f_j and g_j be locally uniformly bounded probability densities such that $f_j \to f$ and $g_j \to g$ in L^1_{loc}. Then, the associated potentials $\phi_j \to \phi$ locally uniformly and $\nabla\phi_j \to \nabla\phi$ in measure.

We fix the following additional notation:

B_R	ball of radius R centered at the origin
$\mathcal{L}^n$	n-dimensional Lebesgue measure
$\mathcal{H}^d$	d-dimensional Hausdorff measure
$\mathbb{S}^{n-1}$	unit sphere in $\mathbb{R}^n$
ω_n	n-dimensional Lebesgue measure of $B_1 \subset \mathbb{R}^n$.

3. Lipschitz changes of variables between log-concave measures

We begin with two useful results of Caffarelli (see [4]). They provide some motivation, and we briefly recall their proofs both for completeness and because we shall need them later.

Lemma 3.1. *Let $\mu = f(x)\,dx$, $\nu = g(x)\,dx \in \mathcal{P}(\mathbb{R}^n)$ with finite second moments and $\nabla\phi = T$ be the optimal transport taking μ to ν. Assume that* $\log f \in L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ *and that g is bounded away from zero in the ball B_{j} for some* $\mathrm{j} > 0$ *and vanishes outside B_{j}. Then,*

$$T(x) \to \mathrm{j}\frac{x}{|x|} \qquad \textit{uniformly as } |x| \to \infty.$$

In particular, for any fixed $\varepsilon > 0$ and for all $\alpha \in \mathbb{S}^{n-1}$, the function $\phi(x+\varepsilon\alpha) + \phi(x-\varepsilon\alpha) - 2\phi(x) \to 0$ as $|x| \to \infty$.

Proof. We begin by noticing that, as a consequence of Theorem 2.2, T is continuous on $\mathbb{R}^n$ and, in particular, the map T is well defined at every point.

Let $x_0 \in \mathbb{R}^n$ and $\theta \in (0, \pi/4)$ be fixed, and consider the cone with vertex at $T(x_0)$ and pointing in the x_0-direction

$$\Gamma := \left\{ y \in \mathbb{R}^n : \angle(x_0, y - T(x_0)) \le \frac{\pi}{2} - \theta \right\}.$$

By (2.1) we see that

$$\angle(x - x_0, T(x) - T(x_0)) \le \frac{\pi}{2};$$

hence,

$$\angle(x - x_0, x_0) \le \angle(x - x_0, T(x) - T(x_0)) + \angle(x_0, T(x) - T(x_0)) \le \pi - \theta \qquad \forall\, x \text{ so that } T(x) \in \Gamma,$$

and so, up to a set of measure zero, the preimage of Γ under T is contained in the (concave) cone

$$\Omega := \{x \in \mathbb{R}^n : \angle(x_0, x - x_0) \leq \pi - \theta\}.$$

Moreover, since $T_\# \mu = \nu$,

$$\inf_{x \in B_{\mathrm{j}}} g(x)\, \mathcal{L}^n\big(\Gamma \cap B_{\mathrm{j}}\big) \leq \nu\big(\Gamma \cap B_{\mathrm{j}}\big) = \nu(\Gamma) \leq \mu(\Omega).$$

Let $B = B_{(|x_0| \tan\theta)/2}$, and notice that $\Omega \subseteq \mathbb{R}^n \setminus B$. This proves that $\mu(\Omega) \leq \mu(\mathbb{R}^n \setminus B)$.

Now, $\mu(\mathbb{R}^n \setminus B) \to 0$ as $|x_0| \to \infty$ since B covers $\mathbb{R}^n$ as $|x_0| \to \infty$. Recalling that g is bounded away from zero in B_{j}, we have that

$$\lim_{|x_0| \to \infty} \mathcal{L}^n\big(\Gamma \cap B_{\mathrm{j}}\big) = 0.$$

Letting $\theta \to 0$, we see that $T(x_0) \to \mathrm{j}\frac{x_0}{|x_0|}$. As the point x_0 was fixed arbitrarily, $\nabla\phi(x) = T(x) \to \mathrm{j}\frac{x}{|x|}$ uniformly as $|x| \to \infty$. Thus, ϕ behaves like the cone $\mathrm{j}|x|$ at infinity. In particular, for any fixed $\varepsilon > 0$ and for all $\alpha \in \mathbb{S}^{n-1}$, the function $\phi(x + \varepsilon\alpha) + \phi(x - \varepsilon\alpha) - 2\phi(x) \to 0$ as $|x| \to \infty$. □

Thanks to Lemma 3.1, [4,5], Caffarelli proved the following result.

Theorem 3.2. *Let V, $W \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ be such that $e^{-V(x)}\,dx$, $e^{-W(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. Suppose there exist constants $0 < \lambda_W$, $\Lambda_V < \infty$ such that $D^2V(x) \leq \Lambda_V\,\mathrm{Id}$ and $\lambda_W\,\mathrm{Id} \leq D^2W(x)$ for a.e. $x \in \mathbb{R}^n$. Then, the optimal transport T that takes $e^{-V(x)}\,dx$ to $e^{-W(x)}\,dx$ is globally Lipschitz and satisfies*

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq \sqrt{\Lambda_V/\lambda_W}. \tag{3.1}$$

Proof. By the stability of optimal transports, we may assume that W is equal to infinity outside the ball B_{j} for some fixed $\mathrm{j} > 0$. Indeed, define

$$W^{\mathrm{j}} := \begin{cases} W & \text{in } B_{\mathrm{j}} \\ \infty & \text{in } \mathbb{R}^n \setminus B_{\mathrm{j}} \end{cases}$$

and $c_{\mathrm{j}} \in (0, \infty)$ such that

$$\int_{\mathbb{R}^n} e^{c_{\mathrm{j}} - W^{\mathrm{j}}(x)}\,dx = 1.$$

Clearly, $e^{c_{\mathrm{j}} - W^{\mathrm{j}}} \to e^{-W}$ in $L^1(\mathbb{R}^n)$ as $\mathrm{j} \to \infty$. Hence, if we prove (3.1) for the optimal transport T^{j} that takes $e^{-V(x)}\,dx$ to $e^{c_{\mathrm{j}} - W^{\mathrm{j}}(x)}\,dx$, letting $\mathrm{j} \to \infty$ we obtain the same estimate for T.

Also, by Theorem 2.2, the convex potential $\phi : \mathbb{R}^n \to \mathbb{R}$ associated to the optimal transport T is of class C^3; therefore, ϕ satisfies the Monge-Ampère equation

$$\det D^2\phi(x) = \frac{e^{-V(x)}}{e^{-W(\nabla\phi(x))}},$$

or equivalently,

$$\log\det D^2\phi(x) = -V(x) + W(\nabla\phi(x)). \tag{3.2}$$

For fixed $\varepsilon > 0$, we define the incremental quotient of a function $f : \mathbb{R}^n \to \mathbb{R}$ at $(x, \alpha) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$ by

$$f^\varepsilon(x, \alpha) := f(x + \varepsilon\alpha) + f(x - \varepsilon\alpha) - 2f(x).$$

By the convexity of ϕ we see that $\phi^\varepsilon \geq 0$. Also, it follows by Lemma 3.1 that $\phi^\varepsilon \to 0$ as $|x| \to \infty$. Thus ϕ^ε attains a global maximum at some $(x_0, \alpha_0) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$. Up to a rotation, we assume that $\alpha_0 = \mathrm{e}_1$. Thus,

$$0 = \nabla\phi^\varepsilon(x_0, \mathrm{e}_1) = \nabla\phi(x_0 + \varepsilon\mathrm{e}_1) + \nabla\phi(x_0 - \varepsilon\mathrm{e}_1) - 2\nabla\phi(x_0). \tag{3.3}$$

Moreover, because e_1 is the maximal direction,

$$0 = \partial_\beta\phi^\varepsilon(x_0, \mathrm{e}_1) = \varepsilon\langle\nabla\phi(x_0 + \varepsilon\mathrm{e}_1) - \nabla\phi(x_0 - \varepsilon\mathrm{e}_1), \beta\rangle \qquad \forall\, \beta \perp \mathrm{e}_1.$$

Taking $\beta = \mathrm{e}_i$ for $i \neq 1$ and utilizing (3.3), we see that all the components but the first of $\nabla\phi(x_0 + \varepsilon\mathrm{e}_1)$, $\nabla\phi(x_0 - \varepsilon\mathrm{e}_1)$, and $\nabla\phi(x_0)$ are equal. Let $\delta := \langle\nabla\phi(x_0 + \varepsilon\mathrm{e}_1) - \nabla\phi(x_0 - \varepsilon\mathrm{e}_1), \mathrm{e}_1\rangle/2$, and observe that, by (3.3),

$$\begin{aligned}\langle\nabla\phi(x_0), \mathrm{e}_1\rangle \pm \delta &= \frac{1}{2}\langle\nabla\phi(x_0 + \varepsilon\mathrm{e}_1) + \nabla\phi(x_0 - \varepsilon\mathrm{e}_1), \mathrm{e}_1\rangle \\ &\quad \pm \frac{1}{2}\langle\nabla\phi(x_0 + \varepsilon\mathrm{e}_1) - \nabla\phi(x_0 - \varepsilon\mathrm{e}_1), \mathrm{e}_1\rangle = \langle\nabla\phi(x_0 \pm \varepsilon\mathrm{e}_1), \mathrm{e}_1\rangle.\end{aligned}$$

Hence, we conclude that

$$\nabla\phi(x_0 \pm \varepsilon\mathrm{e}_1) = \nabla\phi(x_0) \pm \delta\mathrm{e}_1. \tag{3.4}$$

Another consequence of ϕ^ε achieving a maximum at x_0 is

$$D^2\phi(x_0 + \varepsilon\mathrm{e}_1) + D^2\phi(x_0 - \varepsilon\mathrm{e}_1) - 2D^2\phi(x_0) \leq 0. \tag{3.5}$$

We recall that

$$\lim_{\varepsilon\to 0^+} \frac{\det(A + \varepsilon B) - \det(A)}{\varepsilon} = \det(A)\,\mathrm{tr}\,(A^{-1}B) \tag{3.6}$$

for all square matrices A and B with A invertible. Also, if we set $F(A) := \log\det A$, since F is concave on the space of positive semidefinite $n \times n$ matrices and recalling (3.6), we have

$$\nabla F\big(D^2\phi(x_0)\big) = \big(D^2\phi(x_0)\big)^{-1}$$

and

$$F\Big(D^2\phi(x_0 \pm \varepsilon e_1)\Big) \le F\Big(D^2\phi(x_0)\Big) + \Big\langle \Big(D^2\phi(x_0)\Big)^{-1}, D^2\phi(x_0 \pm \varepsilon e_1) - D^2\phi(x_0)\Big\rangle.$$

In particular, from (3.5) and the convexity of ϕ, we deduce that

$$F\Big(D^2\phi(x_0 + \varepsilon e_1)\Big) + F\Big(D^2\phi(x_0 - \varepsilon e_1)\Big) - 2F\Big(D^2\phi(x_0)\Big) \le 0.$$

Now, let us, for fixed $\varepsilon > 0$, consider the incremental quotient of (3.2) at (x_0, e_1). Using (3.4), we realize that

$$V^{\varepsilon}(x_0, e_1) \ge W^{\delta}(\nabla\phi(x_0), e_1). \tag{3.7}$$

Observe that

$$V^{\varepsilon}(x_0, e_1) = \int_0^{\varepsilon}\left(\int_{-t}^{t}\Big\langle D^2V(x_0 + se_1)e_1, e_1\Big\rangle\, ds\right)dt;$$

hence,

$$V^{\varepsilon}(x_0, e_1) \le \Lambda_V \varepsilon^2. \tag{3.8}$$

Furthermore, from (3.4), we similarly see that

$$\lambda_W \delta^2 \le W^{\delta}(\nabla\phi(x_0), e_1).$$

Combining this estimate with (3.8) and (3.7), we get

$$\varepsilon\sqrt{\Lambda_V/\lambda_W} \ge \delta. \tag{3.9}$$

Set $C := \sqrt{\Lambda_V/\lambda_W}$. Since

$$\phi^{\varepsilon}(x_0, e_1) = \int_0^{\varepsilon}\Big\langle \nabla\phi(x_0 + te_1) - \nabla\phi(x_0 - te_1), e_1\Big\rangle\, dt,$$

the convexity of ϕ, (3.4), and (3.9) give us that

$$\phi^{\varepsilon}(x_0, e_1) = 2\delta\varepsilon \le 2C\varepsilon^2,$$

and so

$$\|\nabla T\|_{L^{\infty}(\mathbb{R}^n)} = \left\|D^2\phi\right\|_{L^{\infty}(\mathbb{R}^n)} \le 2C.$$

Notice that this is the desired estimate up to a factor 2. We use a bootstrapping argument to remove this factor. Suppose that $0 \le \|D^2\phi\|_{L^{\infty}(\mathbb{R}^n)} \le a_0$ for some $a_0 > C$. For any $0 \le t \le \varepsilon$, by (3.4) and (3.9),

$$|\langle \nabla\phi(x_0 + te_1) - \nabla\phi(x_0 - te_1), e_1\rangle| \le \min\{2\varepsilon C, 2a_0 t\}.$$

Thus,

$$\phi^\varepsilon(x_0, e_1) \le \int_0^{\frac{\varepsilon C}{a_0}} 2a_0 t\,dt + \int_{\frac{\varepsilon C}{a_0}}^{\varepsilon} 2\varepsilon C\,dt = \varepsilon^2 \frac{(2Ca_0 - C^2)}{a_0}.$$

In other words, if $\|D^2\phi\|_{L^\infty(\mathbb{R}^n)} \le a_0$ with $a_0 > C$, then

$$\|D^2\phi\|_{L^\infty(\mathbb{R}^n)} \le \frac{(2Ca_0 - C^2)}{a_0}.$$

Starting with $a_0 = 2C$ and repeating the above procedure an infinite number of times, we prove (3.1) since C uniquely solves $(2Ca - C^2)/a = a$. $\square$

Remark 3.3. Notice that the above proof relies only on the local behavior of our densities e^{-V} and e^{-W}. In particular, the bounds on the Hessians of V and W are only used near the maximum point x_0 and its image $\nabla\phi(x_0)$, respectively. This simple observation will play an important role in the proof of Theorem 1.1.

Remark 3.4. The above result is not ideal. Indeed, if $V = W$, then $T = \mathrm{Id}$ and one would like to have the bound $\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \le 1$ instead of $\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \le \sqrt{\Lambda_V/\lambda_V}$.

4. Compactly supported perturbations: proof of Theorem 1.1

In the following lemma, we prove an upper bound on how far points travel under the transport map when the source measure is perturbed in a certain fixed ball B_P. We capture and quantify that our perturbations are compactly supported. Lemma 4.1 will be applied in the proof of Theorem 1.1 to the inverse transport.

Furthermore, given our convex function V, we consider, for $\mathrm{j} \in \mathbb{N}$,

$$V^{\mathrm{j}} := \begin{cases} V & \text{in } B_{\mathrm{j}} \\ \infty & \text{in } \mathbb{R}^n \setminus B_{\mathrm{j}}, \end{cases} \tag{4.1}$$

and we approximate $e^{-V(x)}\,dx$ with compactly supported measures $e^{c_{\mathrm{j}} - V^{\mathrm{j}}(x)}\,dx$. This approximation is in the spirit of Caffarelli's approximation in the proof of Theorem 3.2. It allows us to find maximum points of a suitable function and guarantees that they do not escape to infinity in the proof of Theorem 1.1. This approximation procedure is purely technical. Hence, on a first reading of Lemma 4.1, the reader may just take $\mathrm{j} = \infty$.

Lemma 4.1. *Let $V \in C^\infty(\mathbb{R}^n)$ be such that $\mu := e^{-V(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. Suppose that $V(0) = \inf_{\mathbb{R}^n} V$ and there exist constants $0 < \lambda, \Lambda < \infty$ such that $\lambda\,\mathrm{Id} \le D^2V(x) \le \Lambda\,\mathrm{Id}$ for all $x \in \mathbb{R}^n$. Moreover, let $P > 0$, $p \in C_c^\infty(B_P)$, and $c_p \in \mathbb{R}$ be such that $e^{-V(x)+c_p-p(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. Given $\mathrm{j} > P$, set V^{j} as in (4.1) and choose $c_{p,\mathrm{j}} \in (0,\infty)$ such that $\mu_{p,\mathrm{j}} := e^{c_{p,\mathrm{j}} - V^{\mathrm{j}}(x) + c_p - p(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. If T is* *the optimal transport map that takes $\mu_{p,\mathrm{j}}$ to μ, then there exist constants $P' = P'(P, \lambda, \Lambda, \|p\|_{L^\infty(\mathbb{R}^n)}) > 0$ and $\mathrm{j}' = \mathrm{j}'(n, V(0), P, \lambda, \Lambda, \|p\|_{L^\infty(\mathbb{R}^n)}) > P$ such that for all $\mathrm{j} \in [\mathrm{j}', \infty]$,*

$$T(B_P) \subseteq B_{P'}. \tag{4.2}$$

Even though this lemma is not independent of dimension as written (specifically, j' depends on n), the dimensional dependence does not affect the constant P' and disappears in the limit as $\mathrm{j} \to \infty$. Thus, we can indeed prove a global estimate on the optimal transport taking $e^{-V(x)}\,dx$ to $e^{-V(x)+c_q-q(x)}\,dx$ that is independent of dimension.

Lemma 4.1 is written under slightly different assumptions than Theorem 1.1. In particular, besides the obvious additional regularity assumptions on V and its perturbation, made only for simplicity, we have not required that the perturbation be semiconvex. That said, if we assume the the distributional Hessian of p is indeed bounded below by $-\lambda_p \,\mathrm{Id}$, then we can replace the dependence on $\|p\|_{L^\infty(\mathbb{R}^n)}$ with a dependence on λ_p, as explained in the following remark.

Remark 4.2. Let p be a function compactly supported in B_P that satisfies the semiconvexity condition $D^2 p \geq -\lambda_p \,\mathrm{Id}$ in the sense of distributions. Then, its L^∞-norm is controlled by a constant depending only on P and λ_p (in particular, it is independent of dimension):

$$\|p\|_{L^\infty(\mathbb{R}^n)} \leq 4\lambda_p P^2. \tag{4.3}$$

First, up to convolving p with a standard convolution kernel, we can assume that p is smooth. Then, we observe that every 1-dimensional restriction $f_\alpha(t) = p(t\alpha)$, for $t \in \mathbb{R}$ and $\alpha \in \mathbb{S}^{n-1}$, is compactly supported in $[-P, P]$ and has second derivative bounded below by $-\lambda_p$. This implies that

$$\|f'_\alpha\|_{L^\infty(\mathbb{R})} \leq 2\lambda_p P. \tag{4.4}$$

Indeed, suppose to the contrary that $f'_\alpha(t_0) > 2\lambda_p P$ for some $t_0 \in [-P, P]$. By integration, we would get

$$0 = f'_\alpha(P) \geq f'_\alpha(t_0) + \int_{t_0}^{P} f''_\alpha(\tau)\,d\tau > 2\lambda_p P + \lambda_p(P - t_0) > 0,$$

which is impossible. This proves (4.4), and (4.3) holds by integrating.

Before proceeding with the proof of Lemma 4.1, we recall a Talagrand-type transport inequality. Given $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^n)$, we denote the squared Wasserstein distance between μ_1 and μ_2 by $W_2^2(\mu_1, \mu_2)$ (see [18, Chapter 6] for the general definition), and we consider their relative entropy

$$\mathrm{Ent}(\mu_2|\mu_1) := \begin{cases} \displaystyle\int_{\mathbb{R}^n} \log\left(\frac{d\mu_2}{d\mu_1}\right) d\mu_2 & \text{if } \mu_2 \ll \mu_1 \\ \infty & \text{otherwise.} \end{cases}$$

Here, $d\mu_2/d\mu_1$ is the relative density of μ_2 with respect to μ_1. If $\mu_1 = e^{-V(x)}\,dx$ for some $V \in C^2(\mathbb{R}^n)$ such that $D^2V(x) \geq \lambda_V\,\mathrm{Id}$ for all $x \in \mathbb{R}^n$, we have that (see [6], applied in the particular case when μ_1 and μ_2 are probability measures)

$$W_2^2(\mu_1, \mu_2) \leq \frac{2}{\lambda_V}\mathrm{Ent}(\mu_2|\mu_1). \tag{4.5}$$

In our applications, $W_2^2(\mu_1, \mu_2)$ coincides with the cost of the optimal transport taking μ_2 to μ_1.

Proof of Lemma 4.1. Notice first that, as a consequence of Theorem 2.2, T is continuous.

Assume there exists a point $x_0 \in B_P$ with $T(x_0) \notin B_{10P}$ (otherwise, the statement is true with $P' = 10P$). We show that $T(x_0) \in B_{P'}$ for some $P' = P'(P, \lambda, \Lambda, \|p\|_{L^\infty(\mathbb{R}^n)}) > 0$ that will be chosen later. Let

$$\bar{x} := x_0 + 3P\,\frac{T(x_0) - x_0}{|T(x_0) - x_0|},$$

and define the constant C_0 and ball B by $C_0 P = |T(x_0) - x_0|$ and $B := B_P(\bar{x})$. Also, let $F : B \to \mathbb{R}^n$ be the projection of a point $y \in B$ onto the hyperplane through $T(x_0)$ and perpendicular to $y - x_0$. The map F is well-defined because $x_0 \notin B$ (see Figure 4.1). Let us assume that $\mathrm{j}' > 6P$, so that $B \subseteq B_{\mathrm{j}}$.

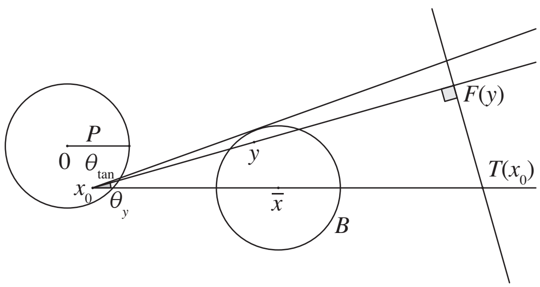

Figure 4.1. The optimal transport sends B far away.

By (2.1), we have that

$$\langle y - x_0, T(y) - T(x_0)\rangle \geq 0 \qquad \forall\, y \in B,$$

and as $F(y)$ is the closest point to y in the set $\{z \in \mathbb{R}^n : \langle y - x_0, z - T(x_0)\rangle \geq 0\}$,

$$|T(y) - x_0| \geq |F(y) - x_0| \qquad \forall\, y \in B$$

(see Figure 4.1). Given any $y \in B$, either x_0, y, and $\bar{x}$ determine a plane, call it Γ_y, within which x_0, $F(y)$, and $T(x_0)$ determine a right triangle, or x_0, y, and $\bar{x}$ are collinear. Thus,

$$|F(y) - x_0| = C_0 P \cos\theta_y$$

where θ_y is the angle between $F(y) - x_0$ and $T(x_0) - x_0$. Now, $\Gamma_y \cap \partial B$ is a circle of radius P centered at $\bar{x}$. Letting θ_{tan} be the angle between the line through x_0 and tangent to $\Gamma_y \cap \partial B$ and the line through $T(x_0)$ and x_0, we see that $\theta_y \leq \theta_{tan}$. (While there are two such tangent lines, the angles they determine with the line through $T(x_0)$ and x_0 are the same. Again, see Figure 4.1.) Moreover, $|x_0 - \bar{x}| = 3P$ and $\cos\theta_{tan} = 2\sqrt{2}/3$. Consequently,

$$|F(y) - x_0| \geq C_0 P \, \cos\theta_{\tan} \geq \frac{C_0 2\sqrt{2}P}{3}$$

and

$$|T(y) - y| \geq |T(y) - x_0| - |y - x_0| > \frac{C_0 2\sqrt{2}P}{3} - 4P \qquad \forall\, y \in B.$$

Since $V(0) = \inf_{\mathbb{R}^n} V(x)$ and $\lambda \,\mathrm{Id} \leq D^2 V(x) \leq \Lambda \,\mathrm{Id}$, by restricting V to 1-dimensional lines through the origin we have that

$$V(0) + \frac{\lambda}{2}|x|^2 \leq V(x) \leq V(0) + \frac{\Lambda}{2}|x|^2 \qquad \forall\, x \in \mathbb{R}^n; \tag{4.6}$$

hence, as $B \subseteq B_{6P}$,

$$V(x) \leq V(0) + 18\Lambda P^2 \qquad \forall\, x \in B.$$

We now estimate $\mathrm{cost}(T)$. Since $B_P \cap B = \emptyset$ and $B \subseteq B_{\mathrm{j}}$, we have

$$\begin{aligned} \mathrm{cost}(T) &\geq \int_B |T(x) - x|^2 e^{c_{p,\mathrm{j}} - V^{\mathrm{j}}(x) + c_p}\, dx \\ &\geq \left[\frac{C_0 2\sqrt{2}P}{3} - 4P\right]^2 e^{c_{p,\mathrm{j}} - V(0) - 18\Lambda P^2 + c_p} \mathcal{L}^n(B_P). \end{aligned} \tag{4.7}$$

Furthermore, we claim that the following upper bound on $\mathrm{cost}(T)$ holds:

$$\mathrm{cost}(T) \leq \frac{6}{\lambda}\|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,\mathrm{j}} + c_p + \|p\|_{L^\infty(\mathbb{R}^n)}} \mu(B_P). \tag{4.8}$$

To see this, first, apply the Talagrand-type transport inequality (4.5) with $\mu_1 = \mu$ and $\mu_2 = \mu_{p,\mathrm{j}}$ to find that

$$\mathrm{cost}(T) \leq \frac{2}{\lambda}\int_{\mathbb{R}^n} \big(c_{p,\mathrm{j}} + c_p - p(x)\big)\, e^{c_{p,\mathrm{j}} - V^{\mathrm{j}}(x) + c_p - p(x)}\, dx. \tag{4.9}$$

Second, choose $\mathrm{j}' > 6P$, so that

$$\int_{\mathbb{R}^n \setminus B_{\mathrm{j}'}} e^{-V(0) - \frac{\lambda}{2}|x|^2 + \|p\|_{L^\infty(\mathbb{R}^n)}}\, dx \leq 1 - \exp\left(-\|p\|_{L^\infty(\mathbb{R}^n)} \int_{B_P} e^{-V(0) - \frac{\Lambda}{2}|x|^2}\, dx\right).$$

Notice that $|c_p| \le \|p\|_{L^\infty(\mathbb{R}^n)}$ since

$$e^{-c_p} = \int_{\mathbb{R}^n} e^{-p(x)}\, d\mu(x). \tag{4.10}$$

So, for every $\mathrm{j} \ge \mathrm{j}'$, observe that

$$\begin{aligned} e^{-c_{p,\mathrm{j}}} &= \int_{B_\mathrm{j}} e^{-V(x)+c_p-p(x)}\, dx = 1 - \int_{\mathbb{R}^n \setminus B_\mathrm{j}} e^{-V(x)+c_p}\, dx \\ &\ge 1 - \int_{\mathbb{R}^n \setminus B_{\mathrm{j}'}} e^{-V(0)-\frac{\lambda}{2}|x|^2+\|p\|_{L^\infty(\mathbb{R}^n)}}\, dx \\ &\ge \exp\left(-\|p\|_{L^\infty(\mathbb{R}^n)} \int_{B_P} e^{-V(0)-\frac{\Lambda}{2}|x|^2}\, dx\right), \end{aligned}$$

and then, recalling that $c_{p,\mathrm{j}} > 0$, note

$$\begin{aligned} c_{p,\mathrm{j}} &\le \|p\|_{L^\infty(\mathbb{R}^n)} \int_{B_P} e^{-V(0)-\frac{\Lambda}{2}|x|^2}\, dx \\ &\le \|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,\mathrm{j}}+c_p+\|p\|_{L^\infty(\mathbb{R}^n)}} \mu(B_P). \end{aligned} \tag{4.11}$$

Now, use Jensen's inequality on (4.10) and that p is supported in B_P to deduce that

$$\begin{aligned} &c_p - \int_{B_P} p(x)\, e^{c_{p,\mathrm{j}}-V^\mathrm{j}(x)+c_p-p(x)}\, dx \\ &\le \int_{B_P} p(x)\, e^{c_{p,\mathrm{j}}+c_p}\Big[e^{-c_{p,\mathrm{j}}-c_p} - e^{-p(x)}\Big]\, d\mu(x) \\ &\le 2\|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,\mathrm{j}}+c_p+\|p\|_{L^\infty(\mathbb{R}^n)}} \mu(B_P). \end{aligned} \tag{4.12}$$

Finally, combine (4.9), (4.11), and (4.12) to see that (4.8) holds as claimed.

In particular, since $\mu(B_P) \le e^{-V(0)}\mathcal{L}^n(B_P)$, we have that

$$\mathrm{cost}(T) \le \frac{6}{\lambda}\|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,\mathrm{j}}-V(0)+c_p+\|p\|_{L^\infty(\mathbb{R}^n)}} \mathcal{L}^n(B_P), \tag{4.13}$$

provided that $\mathrm{j} \ge \mathrm{j}'$. Thus, (4.7) and (4.13) imply that

$$C_0 \le C' := 3\sqrt{2} + \frac{9e^{9\Lambda P^2 + \frac{\|p\|_{L^\infty(\mathbb{R}^n)}}{2}}}{2P}\left[\frac{\|p\|_{L^\infty(\mathbb{R}^n)}}{\lambda}\right]^{1/2}.$$

This proves the existence of an upper bound on C_0 depending only on $P,\ \lambda,\ \Lambda$ and $\|p\|_{L^\infty(\mathbb{R}^n)}$.

Taking $P' := (C'+1)P$, we deduce that

$$|T(x_0)| \le |T(x_0) - x_0| + |x_0| \le C_0 P + P \le P',$$

which proves (4.2). □

The following result is a Pogorelov-type a priori estimate on pure second derivatives of the potential associated to our optimal transport. This technique is inspired by Pogorelov's original argument for the classical Monge-Ampère equation [16]. In our case, we face the additional difficulty of constructing an auxiliary function h that compensates for the concavity of our perturbation and the growth of our convex function at infinity. Assuming that our auxiliary function attains a finite maximum, we provide a quantitative estimate on the value of h at its finite maximum. This result contains and overcomes the primary obstacles to demonstrating that our optimal transport is globally Lipschitz.

Before stating the result, we introduce some constants and an auxiliary function ψ, all depending only on the constants $R,\ \lambda,\ \Lambda$, and λ_q that appear in Theorem 1.1. Define the constants $P > 0$ and $Q > 0$ by

$$P := \frac{2\lambda_q + 4\lambda_q R}{\lambda} + 1 + R \qquad \text{and} \qquad Q := \frac{\lambda}{2\lambda_q} + 1 + R; \tag{4.14}$$

let $\overline{\psi} \in C^2([0,\infty))$ be given by

$$\begin{aligned} \overline{\psi}(t) &:= \int_0^t \int_0^s \vartheta(r)\, dr\, ds, \\ \vartheta(r) &:= \begin{cases} \lambda_q & r \in [0, R] \\ -\lambda_q r + \lambda_q + \lambda_q R & r \in [R, Q] \\ \frac{\lambda_q \lambda^2 r}{4\lambda_q^2 + 8\lambda_q^2 R - \lambda^2} - \frac{2\lambda_q^2\lambda + 4\lambda_q^2 \lambda R + \lambda_q \lambda^2 + \lambda_q \lambda^2 R}{4\lambda_q^2 + 8\lambda_q^2 R - \lambda^2} & r \in [Q, P] \\ 0 & r \in [P, \infty); \end{cases} \end{aligned} \tag{4.15}$$

and let $\psi \in C^2(\mathbb{R}^n)$ be defined by

$$\psi(y) := \overline{\psi}(|y|). \tag{4.16}$$

Observe that the function $\overline{\psi}$ is defined in such a way that $\overline{\psi}'' \geq -\lambda/2$ in $[0,\infty)$, $\overline{\psi} = \lambda_q |\cdot|^2/2$ on $[0, R]$, and $\overline{\psi}'$ is supported in B_P (see Figure 4.2).

Proposition 4.3. *Let $V,\ \lambda,\ \Lambda,\ R,\ q,\ \lambda_q$, and c_q be defined as in Theorem* 1.1. *Assume, additionally, that V and q are smooth. Let P, $\overline{\psi}$, and ψ be defined as in* (4.14)*,* (4.15)*, and* (4.16)*. Given $\mathrm{j} > P$, set V^{j} as in* (4.1) *and choose $c_{q,\mathrm{j}} \in (0,\infty)$ such that $e^{c_{q,\mathrm{j}} - V^{\mathrm{j}}(x) + c_q - q(x)}\, dx \in \mathcal{P}(\mathbb{R}^n)$. Also, let $\phi \in C^\infty(\mathbb{R}^n)$ solve*

$$\det D^2\phi = \frac{e^{-V}}{e^{c_{q,\mathrm{j}} - V^{\mathrm{j}}(\nabla\phi) + c_q - q(\nabla\phi)}},$$

and assume that there exist constants $\mathrm{j}', P' > 0$ such that for all $\mathrm{j} \in [\mathrm{j}', \infty]$,

$$\nabla\phi\big(\mathbb{R}^n \setminus B_{P'}\big) \subseteq \mathbb{R}^n \setminus B_P, \tag{4.17}$$

or equivalently, that $[\nabla\phi]^{-1}(B_P) \subseteq B_{P'}$. *If*

$$h(x,\alpha) := \phi_{\alpha\alpha}(x)e^{\psi(\nabla\phi(x))} \tag{4.18}$$

attains a maximum at some point (x_0, α_0) *among all possible* $(x,\alpha) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$, *then there exists a constant* $C = C(R, P', \lambda, \Lambda, \lambda_q) > 0$, *yet independent of* n, *such that*

$$h(x_0, \alpha_0) \leq C.$$

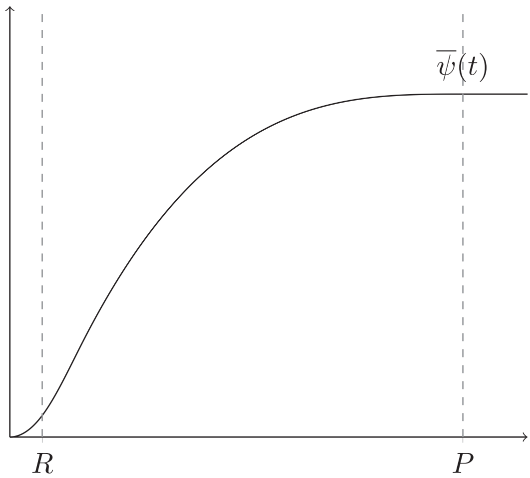

Figure 4.2. The graph of $\overline{\psi}$.

Proof. Since, by assumption, (x_0, α_0) is a maximum point of h, we have $\sup_{|\alpha|=1} \phi_{\alpha\alpha}(x_0) = \phi_{\alpha_0\alpha_0}(x_0)$. This implies that α_0 is an eigenvector of $D^2\phi(x_0)$. Therefore, up to a rotation, we assume that $\alpha_0 = \mathrm{e}_1$ and that $D^2\phi$ is diagonal at x_0. Throughout this proof, the function h is seen as a function of the variable x with α_0 fixed. Then, at x_0 we compute that

$$0 = (\log h)_i = \frac{\phi_{11i}}{\phi_{11}} + \psi_k(\nabla\phi)\phi_{ki}, \tag{4.19}$$

for all $1 \leq i \leq n$, and

$$0 \geq \phi^{ij}(\log h)_{ij} = \phi^{ij}\left[\frac{\phi_{11ij}}{\phi_{11}} - \frac{\phi_{11i}\phi_{11j}}{\phi_{11}^2} + \psi_k(\nabla\phi)\phi_{kij} + \psi_{kl}(\nabla\phi)\phi_{ik}\phi_{jl}\right] \tag{4.20}$$

where we denote the inverse matrix of (ϕ_{ij}) by (ϕ^{ij}).

Let $\tilde{V}^{\mathrm{j}} := V^{\mathrm{j}} - c_{q,\mathrm{j}} + q - c_q$. Using (3.6), we differentiate the equation

$$\log\det D^2\phi = -V + \tilde{V}^{\mathrm{j}}(\nabla\phi) \tag{4.21}$$

in the e_1-direction twice to obtain

$$\phi^{ij}\phi_{1ij} = -V_1 + \tilde{V}_i^{\mathrm{j}}(\nabla\phi)\phi_{1i}$$

and

$$\phi^{ij}\phi_{11ij} - \phi^{il}\phi^{kj}\phi_{1ij}\phi_{1kl} = -V_{11} + \tilde{V}_i^{\mathrm{j}}(\nabla\phi)\phi_{11i} + \tilde{V}_{ij}^{\mathrm{j}}(\nabla\phi)\phi_{1i}\phi_{1j}. \tag{4.22}$$

By (4.20) and (4.22), we deduce that at x_0

$$\begin{aligned} 0 \geq \phi^{il}\phi^{kj}\phi_{1ij}\phi_{1kl} - V_{11} + \tilde{V}_i^{\mathrm{j}}(\nabla\phi)\phi_{11i} + \tilde{V}_{ij}^{\mathrm{j}}(\nabla\phi)\phi_{1i}\phi_{1j} \\ - \frac{\phi^{ij}\phi_{11i}\phi_{11j}}{\phi_{11}} + \phi_{11}\phi^{ij}\psi_k(\nabla\phi)\phi_{kij} + \phi_{11}\phi^{ij}\psi_{kl}(\nabla\phi)\phi_{ik}\phi_{jl}. \end{aligned} \tag{4.23}$$

We estimate each term in (4.23) from below. Recall that (ϕ_{ij}) and (ϕ^{ij}) are diagonal at x_0. Therefore, $\phi^{ii} = 1/\phi_{ii}$, and we see that

$$\phi^{il}\phi^{kj}\phi_{1ij}\phi_{1kl} - \frac{\phi^{ij}\phi_{11i}\phi_{11j}}{\phi_{11}} = \sum_{i=1}^{n}\sum_{k=2}^{n}\phi^{ii}\phi^{kk}\phi_{1ik}^2 \geq 0$$

and

$$\tilde{V}_{ij}^{\mathrm{j}}(\nabla\phi)\phi_{1i}\phi_{1j} = \tilde{V}_{11}^{\mathrm{j}}(\nabla\phi)\phi_{11}^2.$$

Because h has a maximum at e_1 among all directions,

$$\phi_{11}(x_0) \geq \phi_{ii}(x_0), \tag{4.24}$$

and so

$$\phi_{11}\phi^{ij}\psi_{kl}(\nabla\phi)\phi_{ik}\phi_{jl} = \phi_{11}\psi_{ii}(\nabla\phi)\phi_{ii} \geq \psi_{ii}(\nabla\phi)\phi_{ii}^2.$$

Additionally, differentiating (4.21) in the e_k-direction, we have that

$$\phi^{ij}\phi_{kij} = -V_k + \tilde{V}_i^{\mathrm{j}}(\nabla\phi)\phi_{ki}.$$

By (4.19), it then follows that

$$\begin{aligned} \tilde{V}_i^{\mathrm{j}}(\nabla\phi)\phi_{11i} + \phi_{11}\phi^{ij}\psi_k(\nabla\phi)\phi_{kij} &= \tilde{V}_i^{\mathrm{j}}(\nabla\phi)\phi_{11i} + \psi_k(\nabla\phi)(-V_k + \tilde{V}_i^{\mathrm{j}}(\nabla\phi)\phi_{ki})\phi_{11} \\ &= -\psi_k(\nabla\phi)V_k\phi_{11}, \end{aligned}$$

and, consequently, (4.23) becomes

$$0 \geq \tilde{V}_{11}^{\mathrm{j}}(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^{n}\psi_{ii}(\nabla\phi)\phi_{ii}^2 - \psi_k(\nabla\phi)V_k\phi_{11} - \Lambda. \tag{4.25}$$

If $x_0 \in \mathbb{R}^n \setminus B_{P'}$, (4.17) implies that $\nabla\phi(x_0) \in \mathbb{R}^n \setminus B_P$. Then, $\psi_k(\nabla\phi)V_k\phi_{11} = 0$ since the gradient of ψ is zero outside B_P by construction. If, on the other hand, $x_0 \in B_{P'}$, then

$$\psi_k(\nabla\phi)V_k\phi_{11} \leq \Lambda P' \|\nabla\psi\|_{L^\infty(\mathbb{R}^n)}\phi_{11}.$$

(Here, we have used that $V(0) = \inf_{\mathbb{R}^n} V$ and that $D^2V \leq \Lambda\,\mathrm{Id}$ to show V_k is bounded above by $\Lambda P'$.) In both cases, we deduce that

$$\psi_k(\nabla\phi)V_k\phi_{11} \leq \tilde{C}\phi_{11}$$

for a constant $\tilde{C}$ depending only on R, P', λ, Λ, and λ_q. Thus, by (4.25), we have that

$$0 \geq \tilde{V}^{\mathrm{j}}_{11}(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 - \tilde{C}\phi_{11} - \Lambda. \tag{4.26}$$

We claim that

$$\tilde{V}^{\mathrm{j}}_{11}(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 \geq \frac{\lambda}{2}\phi_{11}^2. \tag{4.27}$$

Indeed, let us consider two cases, according to whether or not $\nabla\phi(x_0)$ belongs to B_R. If $\nabla\phi(x_0) \in B_R$, then

$$\tilde{V}^{\mathrm{j}}_{11}(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 \geq \lambda\phi_{11}^2 - \lambda_q\phi_{11}^2 + \lambda_q\phi_{11}^2 = \lambda\phi_{11}^2,$$

and (4.27) follows. In the case that $\nabla\phi(x_0) \notin B_R$, we compute the derivatives of ψ in terms of the derivatives of $\overline{\psi}$. Observe that

$$\psi_i(y) = \frac{\overline{\psi}'(|y|)y_i}{|y|} \qquad \text{and} \qquad \psi_{ii}(y) = \overline{\psi}''(|y|)\frac{y_i^2}{|y|^2} + \frac{\overline{\psi}'(|y|)}{|y|}\left(1 - \frac{y_i^2}{|y|^2}\right).$$

Thus,

$$\psi_{ii}(\nabla\phi) \geq \overline{\psi}''(|\nabla\phi|)\frac{\phi_i^2}{|\nabla\phi|^2} \geq -\frac{\lambda}{2}\frac{\phi_i^2}{|\nabla\phi|^2}$$

since $\overline{\psi}'' \geq -\lambda/2$ in $[0, \infty)$ and $\overline{\psi}' \geq 0$. Then, (4.24) implies that

$$\sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 \geq -\frac{\lambda}{2}\sum_{i=1}^n \frac{\phi_i^2}{|\nabla\phi|^2}\phi_{ii}^2 \geq -\frac{\lambda}{2}\phi_{11}^2\sum_{i=1}^n \frac{\phi_i^2}{|\nabla\phi|^2} = -\frac{\lambda}{2}\phi_{11}^2. \tag{4.28}$$

As $\nabla\phi(x_0) \notin B_R$, we know $\tilde{V}^{\mathrm{j}}_{11}(\nabla\phi(x_0)) = V^{\mathrm{j}}_{11}(\nabla\phi(x_0))$. It follows that

$$\tilde{V}^{\mathrm{j}}_{11}(\nabla\phi)\phi_{11}^2 \geq \lambda\phi_{11}^2. \tag{4.29}$$

By (4.28) and (4.29), we deduce that (4.27) holds in this case as well.

Combining (4.26) and (4.27), we observe that

$$0 \geq \frac{\lambda}{2}\phi_{11}^2 - \tilde{C}\phi_{11} - \Lambda. \tag{4.30}$$

Solving the quadratic equation in (4.30), we find that

$$\phi_{11}(x_0) \leq \frac{\tilde{C} + \sqrt{\tilde{C}^2 + 2\lambda\Lambda}}{\lambda} \leq 2\tilde{C}/\lambda + \sqrt{2\Lambda/\lambda}.$$

As ψ is bounded in $\mathbb{R}^n$ by definition, it follows that

$$h(x) \leq h(x_0) \leq \phi_{11}(x_0)e^{\|\psi\|_{L^\infty(\mathbb{R}^n)}} \leq C$$

for a constant C depending on R, P', λ, Λ, and λ_q, yet independent of n, as desired. □

Notice that if $\lambda_q = 0$, then $\overline{\psi} = 0$. In this case, the constant $\tilde{C}$ found in the proof above is zero, and we recover the global Lipschitz constant obtained by Caffarelli in Theorem 3.2 up to a factor of $\sqrt{2}$ (this is a better bound than the one provided by the proof of Theorem 3.2 before the final bootstrapping argument).

Proof of Theorem 1.1. We first prove the statement assuming that V and q are smooth. For every $\mathrm{j} > R$ set V^{j} as in (4.1), and choose $c_{q,\mathrm{j}} \in (0, \infty)$ such that $e^{c_{q,\mathrm{j}} - V^{\mathrm{j}}(x) + c_q - q(x)}\,dx \in \mathcal{P}(\mathbb{R}^n)$. Let T^{j} be the optimal transport map that takes $e^{-V(x)}\,dx$ to $e^{c_{q,\mathrm{j}} - V^{\mathrm{j}}(x) + c_q - q(x)}\,dx$. Since the density $e^{c_{q,\mathrm{j}} - V^{\mathrm{j}} + c_q - q}$ is supported in a convex set, smooth on its support, and is bounded from above and below by positive constants, by Theorem 2.2, we deduce that $T^{\mathrm{j}} \in C^\infty(\mathbb{R}^n)$. By the stability of optimal transport maps, it suffices to show that for all $\mathrm{j} \geq \mathrm{j}'$ (j' to be chosen possibly depending on n) we have that

$$\left\|\nabla T^{\mathrm{j}}\right\|_{L^\infty(\mathbb{R}^n)} \leq C \tag{4.31}$$

for some constant $C > 0$ depending only on R, λ, Λ, and λ_q.

Let P, ψ, and h be defined as in (4.14), (4.16), and (4.18). Applying Lemma 4.1 to the optimal transport $[T^{\mathrm{j}}]^{-1}$, we see that there exist constants j' and $P' = P'(R, \lambda, \Lambda, \lambda_q) > 0$ (see Remark 4.2) such that $[T^{\mathrm{j}}]^{-1}(B_P) \subseteq B_{P'}$ for all $\mathrm{j} \in [\mathrm{j}', \infty]$; that is, letting $\nabla\phi = T^{\mathrm{j}}$ (for simplicity we omit in ϕ the dependence on j, which can be any number greater than j' in the following),

$$\nabla\phi\big(\mathbb{R}^n \setminus B_{P'}\big) \subseteq \mathbb{R}^n \setminus B_P. \tag{4.32}$$

We split the proof in two cases, according whether or not h achieves a maximum in $\Omega = \mathbb{R}^n \times \mathbb{S}^{n-1}$. If there exists $(x_0, \alpha_0) \in \Omega$ such that

$$h(x_0, \alpha_0) = \sup_{\Omega} h(x, \alpha),$$

then we apply Proposition 4.3 and see that

$$\sup_{\mathbb{S}^{n-1}} \|\phi_{\alpha\alpha}\|_{L^\infty(\mathbb{R}^n)} \le \|h\|_{L^\infty(\Omega)} \le C,$$

which proves (4.31).

Otherwise, we consider the maxima of h in $\Omega_m := \overline{B}_m \times \mathbb{S}^{n-1}$ with $m \in \mathbb{N}$. Let

$$h(x_m, \alpha_m) = \sup_{\Omega_m} h(x, \alpha).$$

Notice that $h(x_m, \alpha_m)$ is nondecreasing (and not definitively constant) and $|x_m| \uparrow \infty$ as $m \to \infty$. Now, consider the functions h^ε approximating h defined by

$$h^\varepsilon(x, \alpha) := [\phi(x + \varepsilon\alpha) + \phi(x - \varepsilon\alpha) - 2\phi(x)]e^{\psi(\nabla\phi(x))} \qquad \forall\,(x, \alpha) \in \Omega.$$

Since ϕ is smooth, we know that $h^\varepsilon \to h$ locally uniformly in Ω as $\varepsilon \to 0$. Furthermore, by Lemma 3.1,

$$\lim_{|x|\to\infty} h^\varepsilon(x, \alpha) = 0 \tag{4.33}$$

uniformly with respect to x and α. Since $h^\varepsilon \ge 0$ (by the convexity of ϕ), the function $h^\varepsilon(x, \alpha)$ has a finite maximum point $(x^\varepsilon, \alpha^\varepsilon)$.

We claim that for sufficiently small ε (possibly depending on n and on the sequence $\{(x_m, \alpha_m)\}_{m\in\mathbb{N}}$)

$$x^\varepsilon \notin B_{P'}. \tag{4.34}$$

Indeed, let m_0 and m_1 be such that $x_{m_0} \notin B_{P'}$ and $h(x_{m_1}, \alpha_{m_1}) > h(x_{m_0}, \alpha_{m_0})$. Since h^ε converges to h locally uniformly, there exists $\varepsilon_0 > 0$ such that

$$\left|h^\varepsilon(x, \alpha) - h(x, \alpha)\right| \le \frac{h(x_{m_1}, \alpha_{m_1}) - h(x_{m_0}, \alpha_{m_0})}{4}$$

for every $x \in \overline{B}_{|x_{m_1}|+1}, \alpha \in \mathbb{S}^{n-1}$, and $\varepsilon \le \varepsilon_0$. So, for every $\varepsilon \le \varepsilon_0$, we have that

$$\begin{aligned} h^\varepsilon(x_{m_1}, \alpha_{m_1}) &\ge h(x_{m_1}, \alpha_{m_1}) - \left|h^\varepsilon(x_{m_1}, \alpha_{m_1}) - h(x_{m_1}, \alpha_{m_1})\right| \\ &\ge \frac{3h(x_{m_1}, \alpha_{m_1}) + h(x_{m_0}, \alpha_{m_0})}{4}. \end{aligned} \tag{4.35}$$

Thus,

$$\begin{aligned} h^\varepsilon(x, \alpha) &\le h(x, \alpha) + \left|h^\varepsilon(x, \alpha) - h(x, \alpha)\right| \\ &\le h(x_{m_0}, \alpha_{m_0}) + \frac{h(x_{m_1}, \alpha_{m_1}) - h(x_{m_0}, \alpha_{m_0})}{4} \\ &= \frac{h(x_{m_1}, \alpha_{m_1}) + 3h(x_{m_0}, \alpha_{m_0})}{4} < \frac{3h(x_{m_1}, \alpha_{m_1}) + h(x_{m_0}, \alpha_{m_0})}{4} \end{aligned} \tag{4.36}$$

for every $x \in \overline{B}_{|x_{m_0}|}$, $\alpha \in \mathbb{S}^{n-1}$, and $\varepsilon \le \varepsilon_0$. Since $B_{P'} \subseteq B_{|x_{m_0}|}$, (4.35) and (4.36) imply that $h^\varepsilon(x, \alpha) \le h^\varepsilon(x_{m_1}, \alpha_{m_1})$ in $B_{P'}$. Therefore, h^ε satisfies (4.34) for every $\varepsilon \le \varepsilon_0$.

Recall that ψ is constant outside B_P. Then, by (4.32) and (4.34), we know that for every $\varepsilon \le \varepsilon_0$, the function $e^{\psi(\nabla\phi(x))}$ is locally constant around x^ε. Therefore, $(x^\varepsilon, \alpha^\varepsilon)$ is also a local maximum point for the incremental quotient $\phi(x + \varepsilon\alpha) + \phi(x - \varepsilon\alpha) - 2\phi(x)$. Moreover, outside B_R the function $V^{\mathrm{j}} - c_{q,\mathrm{j}} + q - c_q$ is convex as it coincides with $V^{\mathrm{j}} - c_{q,\mathrm{j}} - c_q$. So, proceeding as in the proof of Theorem 3.2 (*cf.* Remark 3.3), we conclude that (4.31) is also proved in the case that h is not guaranteed to achieve a maximum in Ω.

In order to remove the smoothness assumptions on V and q, we approximate V and q by convolution (adding a small constant to ensure these approximations define probability measures). Then, from what we have shown above, the approximate transports are all globally and uniformly Lipschitz. Thanks to the stability of optimal transports, passing to the limit, we prove (1.2). □

5. Bounded perturbations in 1-dimension and in the radially symmetric case: proofs of Theorems 1.2 and 1.3

Our goal now is to produce optimal global Lipschitz estimates under strong symmetry but weak regularity assumptions on our log-concave measures. Notice that when our perturbation is zero, we recover that our optimal transport is the identity map (*cf.* Remark 3.4). We begin in 1-dimension and with a technical lemma relating the behavior of our convex base and the cumulative distribution function of the log-concave probability measure it defines.

Lemma 5.1. *Let $V : \mathbb{R} \to \mathbb{R}$ be a convex function such that $e^{-V(x)}\,dx \in \mathcal{P}(\mathbb{R})$ and $x_0 \in \mathbb{R}$ be such that $V(x_0) = \inf_{\mathbb{R}} V$. Define $\Phi, \Psi : \mathbb{R} \to (0, 1)$ by*

$$\Phi(x) := \int_{-\infty}^{x} e^{-V(t)}\,dt \quad \textit{and} \quad \Psi(x) := \int_{x}^{\infty} e^{-V(t)}\,dt = 1 - \Phi(x). \tag{5.1}$$

Then,

$$V(x) - V(y) \le \log\Phi(y) - \log\Phi(x) \qquad \forall\, x \le y \le x_0 \tag{5.2}$$

and

$$V(x) - V(y) \ge \log\Psi(y) - \log\Psi(x) \qquad \forall\, x_0 \le x \le y. \tag{5.3}$$

Proof. Since an analogous argument proves (5.3), we only show (5.2); in other words, we prove that the function $\log\Phi + V$ is nondecreasing in $(-\infty, x_0]$. Let $\hat{x} = \inf\{x : V(x) = V(x_0)\}$. The function $\log\Phi + V$ is clearly nondecreasing in $[\hat{x}, x_0]$, whenever this interval is not a single point. Moreover, it is locally Lipschitz and its derivative is $e^{-V}/\Phi + V'$. Hence, it suffices to show that the derivative is nonnegative in $(-\infty, \hat{x})$. Since V' is nonincreasing in $(-\infty, \hat{x})$ and by the change of variables formula, we have that for a.e. $x \in (-\infty, \hat{x})$

$$V'(x)\Phi(x) \geq \int_{-\infty}^{x} V'(t)e^{-V(t)}\,dt = -e^{-V(x)},$$

which proves our claim. □

Proof of Theorem 1.2. By approximating V with a sequence of convex functions $V_j \to V$ such that $e^{-V_j(x)}\,dx \in \mathcal{P}(\mathbb{R})$ and that are finite on $\mathbb{R}$, we can assume that $V < \infty$ on $\mathbb{R}$. This reduction follows from the stability of optimal transport maps. Recall that, as a consequence of the push-forward condition $T_\#\big(e^{-V(x)}\,dx\big) = e^{-V(x)-q(x)}\,dx$, T satisfies the mass balance equation

$$\int_{-\infty}^{x} e^{-V(t)}\,dt = \int_{-\infty}^{T(x)} e^{-V(t)-q(t)}\,dt, \tag{5.4}$$

which can be also written as

$$\int_{x}^{\infty} e^{-V(t)}\,dt = \int_{T(x)}^{\infty} e^{-V(t)-q(t)}\,dt \tag{5.5}$$

since the measures $e^{-V(x)}\,dx$ and $e^{-V(x)-q(x)}\,dx$ have total mass 1. From (5.4), we deduce that T is differentiable. Indeed, both the functions

$$F(x) := \int_{-\infty}^{x} e^{-V(t)}\,dt \qquad \text{and} \qquad G(x) := \int_{-\infty}^{x} e^{-V(t)-q(t)}\,dt$$

are differentiable and their derivatives do not vanish. So, $T(x) = G^{-1} \circ F(x)$ is differentiable as well. Thus, differentiating with respect to x and then taking the logarithm shows that

$$\log\big(T'(x)\big) = -V(x) + V(T(x)) + q(T(x)) \qquad \forall\, x \in \mathbb{R}.$$

Consequently,

$$\begin{aligned} &V(T(x)) - V(x) - \|q^-\|_{L^\infty(\mathbb{R})} \\ &\leq \log(T'(x)) \leq V(T(x)) - V(x) + \|q^+\|_{L^\infty(\mathbb{R})}. \end{aligned} \tag{5.6}$$

On the other hand, (5.4) implies that

$$e^{-\|q^+\|_{L^\infty(\mathbb{R})}} \int_{-\infty}^{T(x)} e^{-V(t)}\,dt \leq \int_{-\infty}^{x} e^{-V(t)}\,dt \leq e^{\|q^-\|_{L^\infty(\mathbb{R})}} \int_{-\infty}^{T(x)} e^{-V(t)}\,dt$$

since $q \in L^\infty(\mathbb{R})$. Taking the logarithm and defining Φ as in (5.1), we see that

$$-\big\|q^+\big\|_{L^\infty(\mathbb{R})} \leq \log\Phi(x) - \log\Phi(T(x)) \leq \|q^-\|_{L^\infty(\mathbb{R})}. \tag{5.7}$$

Analogously, from (5.5), we deduce that

$$-\left\|q^+\right\|_{L^\infty(\mathbb{R})} \le \log \Psi(x) - \log \Psi(T(x)) \le \|q^-\|_{L^\infty(\mathbb{R})}. \tag{5.8}$$

We claim that

$$-\|q^+\|_{L^\infty(\mathbb{R})} \le V(T(x)) - V(x) \le \|q^-\|_{L^\infty(\mathbb{R})} \qquad \forall x \in \mathbb{R}. \tag{5.9}$$

To prove this claim, let $x_0 \in \mathbb{R}$ be such that $V(x_0) = \inf_{\mathbb{R}} V$ and consider the sets

$$E_1 := \{x : x \le x_0 \text{ and } T(x) \le x_0\} \quad \text{and} \quad E_2 := \{x : x \ge x_0 \text{ and } T(x) \ge x_0\}.$$

Applying (5.2) in E_1 yields that

$$0 \le V(T(x)) - V(x) \le \log \Phi(x) - \log \Phi(T(x))$$

if $T(x) \le x \le x_0$ and

$$\log \Phi(x) - \log \Phi(T(x)) \le V(T(x)) - V(x) \le 0$$

whenever $x \le T(x) \le x_0$. Therefore, (5.9) holds in E_1 by (5.7). Similarly, applying (5.3) gives us that (5.9) holds in E_2 by (5.8). Now, we consider three cases:

1. If $T(x_0) = x_0$, the monotonicity of T implies that $E_1 \cup E_2 = \mathbb{R}$, and (5.9) holds in all of $\mathbb{R}$;
2. If $T(x_0) > x_0$, then $E_1 \cup E_2 \cup E_+ = \mathbb{R}$ where, thanks to the monotonicity of T, we have

   $$E_+ = \{x : x \le x_0 \text{ and } T(x) \ge x_0\} = \left[T^{-1}(x_0), x_0\right].$$

   Since V attains its minimum at x_0, V is decreasing on $(-\infty, x_0]$ and increasing on $[x_0, \infty)$. Consequently,

   $$V(T^{-1}(x_0)) - V(x_0) \le V(x) - V(T(x)) \le V(T(x_0)) - V(x_0) \qquad \forall x \in E_+.$$

   As $T^{-1}(x_0) \in E_1$ and $x_0 \in E_2$, our above analysis shows that (5.9) holds in E_+;
3. If $T(x_0) < x_0$, an analogous argument to one used to prove case 2 demonstrates that $E_1 \cup E_2 \cup E_- = \mathbb{R}$ where $E_- = [x_0, T^{-1}(x_0)]$ and proves (5.9) also in E_-.

Therefore, by (5.6) and (5.9), we deduce (1.3). □

Remark 5.2. From the numerical inequality $|\log(x)| \ge x - 1$, which holds for $x \in [0, e^2]$, we see that if ϕ is the potential associated to T in Theorem 1.2, then provided that $\|q\|_{L^\infty(\mathbb{R})} \le 1$, there exists a constant $C > 0$ such that

$$\|\phi'' - 1\|_{L^\infty(\mathbb{R})} \le C\|q\|_{L^\infty(\mathbb{R})}.$$

We now move to the radially symmetric case in n dimensions.

Proof of Theorem 1.3. Let $\overline{V}$, $\overline{q} : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ be two functions such that $\overline{V} = \overline{q} = \infty$ on $(-\infty, 0)$, and $V(x) = \overline{V}(|x|)$ and $q(x) = \overline{q}(|x|)$ for every $x \in \mathbb{R}^n$. Now, consider the function

$$T(x) := \tilde{T}(|x|)\frac{x}{|x|}$$

where $\tilde{T} : \mathbb{R} \to \mathbb{R}$ is the optimal transport that takes $e^{-\overline{V}(r)}r^{n-1}\,dr$ to $e^{-\overline{V}(r)-\overline{q}(r)}r^{n-1}\,dr$.

Set $\mathbb{R}^+ := [0, \infty)$. We first claim that the optimal transport $\tilde{T}$ is Lipschitz and satisfies

$$\left\|\log \tilde{T}'\right\|_{L^\infty(\mathbb{R}^+)} \le \|q^+\|_{L^\infty(\mathbb{R}^+)} + \|q^-\|_{L^\infty(\mathbb{R}^+)}. \tag{5.10}$$

Indeed, let $\tilde{V} : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ be defined by $\tilde{V}(r) = \overline{V}(r) - (n-1)\log r$ on $\mathbb{R}^+$ and infinity otherwise, and let $\tilde{q} = \overline{q}$ on $\mathbb{R}^+$ and zero elsewhere. Observe that $\tilde{V}$ is convex and $\tilde{q}$ is bounded. Hence, applying Proposition 1.2 with $V = \tilde{V}$ and $q = \tilde{q}$ proves (5.10).

We now conclude the proof. Notice that T is continuous. Furthermore, T is an admissible change of variables from $e^{-V(x)}\,dx$ to $e^{-V(x)-q(x)}\,dx$. To see this, we show that for every bounded, Borel function $\varphi : \mathbb{R}^n \to \mathbb{R}$,

$$\int_{\mathbb{R}^n} \varphi(T(x))e^{-V(x)}\,dx = \int_{\mathbb{R}^n} \varphi(x)e^{-V(x)-q(x)}\,dx. \tag{5.11}$$

The formula (5.11) can be rewritten, using polar coordinates and the definition of T, as

$$\begin{aligned}\int_0^\infty \int_{\mathbb{S}^{n-1}} &\varphi\big(\tilde{T}(r)\alpha\big)\,d\mathcal{H}^{n-1}(\alpha)\,e^{-\overline{V}(r)}r^{n-1}\,dr\\ &= \int_0^\infty \int_{\mathbb{S}^{n-1}} \varphi(r\alpha)\,d\mathcal{H}^{n-1}(\alpha)\,e^{-\overline{V}(r)-\overline{q}(r)}r^{n-1}\,dr,\end{aligned}$$

which is, in turn, satisfied if we use the test function $\overline{\varphi}(r) = \int_{\mathbb{S}^{n-1}} \varphi(r\alpha)\,d\mathcal{H}^{n-1}(\alpha)$ and recall the definition of T.

Now, let $\xi \in \mathbb{S}^{n-1}$ and $x \in \mathbb{R}^n \setminus \{0\}$. Since $\tilde{T}(0) = 0$, we observe that

$$\begin{aligned}\nabla T(x)[\xi] &= \big[\xi|x|^{-1} - x|x|^{-3}\langle x, \xi\rangle\big]\tilde{T}(|x|) + x|x|^{-2}\tilde{T}'(|x|)\langle x, \xi\rangle\\ &= \big[\xi - x|x|^{-2}\langle x, \xi\rangle\big]\tilde{T}'(t) + x|x|^{-2}\tilde{T}'(|x|)\langle x, \xi\rangle,\end{aligned}$$

where $t \in (0, |x|)$. By (5.10), we deduce that

$$e^{-\|q^+\|_{L^\infty(\mathbb{R}^n)}-\|q^-\|_{L^\infty(\mathbb{R}^n)}} \le \langle \xi, \nabla T(x)[\xi]\rangle \le e^{\|q^+\|_{L^\infty(\mathbb{R}^n)}+\|q^-\|_{L^\infty(\mathbb{R}^n)}},$$

which proves (1.4). To conclude, we show that T is the optimal transport taking $e^{-V(x)}\,dx$ to $e^{-V(x)-q(x)}\,dx$. Let $\tilde{\phi}:\mathbb{R}^+\to\mathbb{R}^+$ be the convex potential associated to $\tilde{T}$. By construction, $T(x)=\nabla(\tilde{\phi}(|x|))$ and $\tilde{\phi}(|x|)$ is a convex function. Since optimal transports are characterized by being gradients of convex functions, T is the optimal transport taking $e^{-V(x)}\,dx$ to $e^{-V(x)-q(x)}\,dx$. □

6. Appendix

We now show that the linear bound in Remark 5.2 is specific to the 1-dimensional case.

Proposition 6.1. *Let $n\in\mathbb{N}$ and $V(x)=|x|^2/2+(n/2)\log(2\pi)$, so that e^{-V} is the standard Gaussian density in $\mathbb{R}^n$. Then, for every $C>0$, there exists a bounded, continuous perturbation p such that $\|p\|_{L^\infty(\mathbb{R}^n)}\le 1$ and $e^{-V(x)-p(x)}\,dx\in\mathcal{P}(\mathbb{R}^n)$ and the optimal transport $T=\nabla\phi$ that takes $e^{-V(x)}\,dx$ to $e^{-V(x)-p(x)}\,dx$ satisfies*

$$\left\|D^2\phi-\mathrm{Id}\right\|_{L^\infty(\mathbb{R}^n)}>C\|p\|_{L^\infty(\mathbb{R}^n)}.$$

Proof. Suppose, to the contrary, that for every bounded, continuous function $p:\mathbb{R}^n\to\mathbb{R}$ with $\|p\|_{L^\infty(\mathbb{R}^n)}\le 1$, the optimal transport $T=\nabla\phi$ that takes $e^{-V(x)}\,dx$ to $e^{-V(x)-p(x)}\,dx$ satisfies

$$\left\|D^2\phi-\mathrm{Id}\right\|_{L^\infty(\mathbb{R}^n)}\le C_0\|p\|_{L^\infty(\mathbb{R}^n)} \tag{6.1}$$

for some $C_0>0$. In particular, let $q\in L^\infty(\mathbb{R}^n)\cap C^0(\mathbb{R}^n)$, and for all $\varepsilon\ge 0$, define c_ε by

$$e^{c_\varepsilon}=\int_{\mathbb{R}^n}e^{-V(x)-\varepsilon q(x)}\,dx.$$

By construction, $e^{-V(x)-\varepsilon q(x)-c_\varepsilon}\,dx\in\mathcal{P}(\mathbb{R}^n)$. Thus, let ϕ_ε be the potential associated to the optimal transport that takes $e^{-V(x)}\,dx$ to $e^{-V(x)-\varepsilon q(x)-c_\varepsilon}\,dx$, and remember that ϕ_ε solves the Monge-Ampère equation

$$\det D^2\phi_\varepsilon=e^{-V+V(\nabla\phi_\varepsilon)+\varepsilon q(\nabla\phi_\varepsilon)+c_\varepsilon}. \tag{6.2}$$

Note that $c_\varepsilon\to 0$ as $\varepsilon\to 0$. Also, since

$$\left|c_\varepsilon'\right|=\left|\frac{(e^{c_\varepsilon})'}{e^{c_\varepsilon}}\right|=\left|\int_{\mathbb{R}^n}-q(x)e^{-V(x)-\varepsilon q(x)-c_\varepsilon}\,dx\right|\le\|q\|_{L^\infty(\mathbb{R}^n)},$$

c_ε is Lipschitz as a function of ε and

$$\frac{|c_\varepsilon|}{\varepsilon}\le\|q\|_{L^\infty(\mathbb{R}^n)}. \tag{6.3}$$

In addition, by the dominated convergence theorem,

$$c'_\varepsilon \to \iota_q := \int_{\mathbb{R}^n} -q(x)e^{-V(x)}\,dx \qquad \text{as } \varepsilon \to 0. \tag{6.4}$$

Without loss of generality, we assume that $\phi_\varepsilon(0) = 0$. Now, define

$$\psi_\varepsilon(x) := \frac{\phi_\varepsilon(x) - |x|^2/2}{\varepsilon}.$$

By (6.1) applied to $p = \varepsilon q + c_\varepsilon$ and (6.3), we see that if $\varepsilon \le \frac{1}{2\|q\|_{L^\infty(\mathbb{R}^n)}}$, then

$$\left\| D^2\psi_\varepsilon \right\|_{L^\infty(\mathbb{R}^n)} \le (C_0 + 1)\|q\|_{L^\infty(\mathbb{R}^n)}. \tag{6.5}$$

Recall that, for any $n \times n$ matrix A, there exists a $K > 0$, depending only on $\|A\|$, such that for all ε sufficiently small $|\log\det(\mathrm{Id} + \varepsilon A) - \varepsilon \operatorname{tr} A| \le \varepsilon^2 K$. Therefore, there exist an $\varepsilon_0 > 0$ and a collection of functions g_ε with

$$\sup_{\varepsilon \le \varepsilon_0} \|g_\varepsilon\|_{L^\infty(\mathbb{R}^n)} < \infty \tag{6.6}$$

such that for all $\varepsilon \le \varepsilon_0$,

$$\varepsilon\Delta\psi_\varepsilon(x) + \varepsilon^2 g_\varepsilon(x) = \log\det\left(\mathrm{Id} + \varepsilon D^2\psi_\varepsilon\right) = \log\det D^2\phi_\varepsilon.$$

Thus, by (6.2) and our choice of V,

$$\begin{aligned} \Delta\psi_\varepsilon(x) + \varepsilon g_\varepsilon(x) &= \frac{V(\nabla\phi_\varepsilon(x)) - V(x) + \varepsilon q(\nabla\phi_\varepsilon(x)) + c_\varepsilon}{\varepsilon} \\ &= \int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \nabla\psi_\varepsilon(x)\rangle\,dt + q(\nabla\phi_\varepsilon(x)) + \frac{c_\varepsilon}{\varepsilon} \\ &= \langle x, \nabla\psi_\varepsilon(x)\rangle + \frac{\varepsilon}{2}|\nabla\psi_\varepsilon(x)|^2 + q(\nabla\phi_\varepsilon(x)) + \frac{c_\varepsilon}{\varepsilon}. \end{aligned} \tag{6.7}$$

We claim that, up to a subsequence, there exists a function $\psi_0 \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ such that $\psi_\varepsilon \to \psi_0$ in $C^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $D^2\psi_\varepsilon \rightharpoonup D^2\psi_0$ weakly-$*$ in $L^\infty(\mathbb{R}^n)$ as $\varepsilon \to 0$. To this end, by Arzelà-Ascoli, it suffices to show that ψ_ε are locally bounded in $C^{1,1}$. Since $\psi_\varepsilon(0) = 0$, by (6.5), it is enough to prove that

$$\liminf_{\varepsilon \to 0} |\nabla\psi_\varepsilon(0)| < \infty. \tag{6.8}$$

Assume, to the contrary, that $\lim_{\varepsilon\to 0}|\nabla\psi_\varepsilon(0)| = \infty$. Notice that (6.7) implies that for all $\varepsilon \le \varepsilon_0$ and $x \in \mathbb{R}^n$,

$$\begin{aligned}
&\left|\int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \nabla\psi_\varepsilon(0)\rangle\, dt\right| \\
&\le \left|\int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \nabla\psi_\varepsilon(x)\rangle\, dt\right| + \big(|\nabla\phi_\varepsilon(x)| + |x|\big)|\nabla\psi_\varepsilon(0) - \nabla\psi_\varepsilon(x)| \\
&\le |\Delta\psi_\varepsilon(x)| + \varepsilon|g_\varepsilon(x)| + |q(\nabla\phi_\varepsilon(x))| + \frac{|c_\varepsilon|}{\varepsilon} \\
&\quad + \big(|\nabla\phi_\varepsilon(x)| + |x|\big)|x| \sup_{\varepsilon\le\varepsilon_0}\|D^2\psi_\varepsilon\|_{L^\infty(\mathbb{R}^n)}.
\end{aligned}$$

Let $\alpha_\varepsilon = \nabla\psi_\varepsilon(0)/|\nabla\psi_\varepsilon(0)| \in \mathbb{S}^{n-1}$, and note that up to subsequences $\alpha_\varepsilon \to \alpha_0 \in \mathbb{S}^{n-1}$ as $\varepsilon \to 0$. Furthermore, let $\eta \in C_c^\infty(B_{1/2}(\alpha_0))$ be a nonnegative function that integrates to one. Then, by (6.3), we deduce that

$$\begin{aligned}
&|\nabla\psi_\varepsilon(0)| \int_{\mathbb{R}^n}\left|\int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \alpha_\varepsilon\rangle\, dt\right| \eta(x)\, dx \\
&= \int_{\mathbb{R}^n}\left|\int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \nabla\psi_\varepsilon(0)\rangle\, dt\right| \eta(x)\, dx \\
&\le \sup_{\varepsilon\le\varepsilon_0} \varepsilon\|g_\varepsilon\|_{L^\infty(\mathbb{R}^n)} + 2\|q\|_{L^\infty(\mathbb{R}^n)} \\
&\quad + \sup_{\varepsilon\le\varepsilon_0}\|D^2\psi_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n}\big(|\nabla\phi_\varepsilon(x)||x| + |x|^2 + 1\big)\eta(x)\, dx.
\end{aligned} \tag{6.9}$$

Recall that $D^2\phi_\varepsilon$ converges uniformly to the identity matrix by (6.1) applied to ϕ_ε and εq. By the stability and uniqueness of optimal transports, $\nabla\phi_\varepsilon$ converges locally uniformly to the identity map as $\varepsilon \to 0$. In particular, $|\nabla\phi_\varepsilon(x)| \le 2$ for every $x \in B_{1/2}(\alpha_0)$ and ε sufficiently small, and we obtain that

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\left|\int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \alpha_\varepsilon\rangle\, dt\right| \eta(x)\, dx = \int_{\mathbb{R}^n}\langle x, \alpha_0\rangle\eta(x)\, dx \ge \frac{1}{2}$$

by dominated convergence. Thus, taking the limit in (6.9) and noticing that the right-hand side is bounded as $\varepsilon \to 0$ thanks to (6.5) and (6.6), we see that

$$\infty = \lim_{\varepsilon\to 0}|\nabla\psi_\varepsilon(0)|\left|\int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \alpha_\varepsilon\rangle\, dt\right| < \infty,$$

which, being impossible, proves (6.8) and shows that $\psi_\varepsilon \to \psi_0$ in $C^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $D^2\psi_\varepsilon \rightharpoonup D^2\psi_0$ weakly-$*$ in $L^\infty(\mathbb{R}^n)$ as $\varepsilon \to 0$ for some function $\psi_0 \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$.

Now, reformulating (6.7), we see that for any $\eta \in C_c^\infty(\mathbb{R}^n)$,

$$\begin{aligned}&\int_{\mathbb{R}^n} \Big(\Delta\psi_\varepsilon(x) + \varepsilon g_\varepsilon(x) - q(\nabla\phi_\varepsilon(x)) - \frac{c_\varepsilon}{\varepsilon}\Big)\eta(x)\,dx\\ &= \int_{\mathbb{R}^n} \Big(\langle x, \nabla\psi_\varepsilon(x)\rangle + \frac{\varepsilon}{2}|\nabla\psi_\varepsilon(x)|^2\Big)\eta(x)\,dx.\end{aligned} \tag{6.10}$$

Thus, recalling (6.4) and that q is continuous, we can pass to the limit and obtain that

$$\int_{\mathbb{R}^n} \big(\Delta\psi_0(x) - \langle x, \nabla\psi_0(x)\rangle\big)\eta(x)\,dx = \int_{\mathbb{R}^n} \big(q(x) + \iota_q\big)\eta(x)\,dx$$

for all $\eta \in C_c^\infty(\mathbb{R}^n)$. Since q was arbitrary, we have shown that for every $q \in L^\infty(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$, there exists a function $\psi_0 \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ solution to

$$\Delta\psi_0(x) - \langle x, \nabla\psi_0(x)\rangle = q(x) + \iota_q. \tag{6.11}$$

We now show that this is impossible. Recall that there exists a bounded, continuous g and $\psi \in C^{1,\alpha}_{\mathrm{loc}}(B_2) \cap C^\infty(B_2 \setminus \{0\})$, for any $\alpha \in (0, 1)$, such that $\Delta\psi(x) = g(x)$ in B_2, yet $\psi \notin C^{1,1}(B_2)$. In particular, $\lim_{x\to 0} |D^2\psi(x)| = \infty$. (See [10, Chapter 3].) Define

$$h(x) := \begin{cases} g(x) - \langle x, \nabla\psi(x)\rangle & x \in B_1\\ g(x/|x|) - \langle x/|x|, \nabla\psi(x/|x|)\rangle & x \in \mathbb{R}^n \setminus B_1,\end{cases}$$

and observe that, since $\psi \in C^{1,\alpha}_{\mathrm{loc}}(B_2)$ and g is bounded and continuous, $h \in L^\infty(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$. By construction, there exists a $\psi_0 \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ that solves (6.11) with $q = h$. Then, for $\psi_1 := \psi_0 - \psi$ we have that $\Delta\psi_1(x) - \langle x, \nabla\psi_1(x)\rangle = \iota_h$ in B_1. Thus, $\psi_1 \in C^\infty(B_1)$ by elliptic regularity, a contradiction since $\psi \notin C^{1,1}_{\mathrm{loc}}(B_1)$ and $\psi_0 \in C^{1,1}(B_1)$. □

References

[1] S. ALESKER, S. DAR and V. MILMAN, *A remarkable measure preserving diffeomorphism between two convex bodies in* $\mathbb{R}^n$, Geom. Dedicata **74** (1999), 201–212.

[2] Y. BRENIER, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), 365–417.

[3] L. A. CAFFARELLI, *The regularity of mappings with a convex potential*, J. Amer. Math. Soc. **5** (1992), 99–104.

[4] L. A. CAFFARELLI, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys. **214** (2000), 547–563.

[5] L. A. CAFFARELLI, *Erratum: Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys. **225** (2002), 449–450.

[6] D. CORDERO-ERAUSQUIN, *Some applications of mass transport to Gaussian-type inequalities*, Arch. Ration. Mech. Anal. **161** (2002), 257–269.

[7] D. CORDERO-ERAUSQUIN, M. FRADELIZI and B. MAUREY, *The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems*, J. Funct. Anal. **214** (2004), 410–427.
[8] G. DE PHILIPPIS and A. FIGALLI, *The Monge-Ampère equation and its link to optimal transportation*, Bull. Amer. Math. Soc. (N.S.) **51** (2014), 527–580.
[9] A. FIGALLI, "The Monge-Ampère Equation and its Applications", Zürich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2017.
[10] Q. HAN and F. LIN, "Elliptic Partial Differential Equations", 2nd Edition, Courant Lecture Notes in Mathematics, Vol. 1, New York University, Courant Institute of Mathematical Sciences, New York, American Mathematical Society, Providence, 1997.
[11] G. HARGÉ, *A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces*, Probab. Theory Related Fields **130** (2004), 415–440.
[12] Y.-H. KIM and E. MILMAN, *A generalization of Caffarelli's contraction theorem via (reverse) heat flow*, Math. Ann. **354** (2012), 827–862.
[13] B. KLARTAG, *Marginals of geometric inequalities*, In: "Geometric Aspects of Functional Analysis", V. D. Milman and G. Schechtman (eds.), Lecture Notes in Mathematics, Vol. 1910, Springer Berlin, 2007, 133–166.
[14] A. V. KOLESNIKOV, *Global Hölder estimates for optimal transportation*, (Russian) Mat. Zametki **88** (2010), 708-728; (English) Math. Notes **88** (2010), 678–695.
[15] A. V. KOLESNIKOV, *On Sobolev regularity of mass transport and transportation inequalities*, Theory Probab. Appl. **57** (2013), 243–264.
[16] A.V. POGORELOV, *The regularity of generalized solutions of the equation* $\det(\partial^2 u/\partial x_i \partial x_j) = \varphi(x_1, x_2, \ldots, x_n) > 0$ (Russian), Dokl. Akad. Nauk SSSR **200** (1971), 534–537.
[17] S. I. VALDIMARSSON, *On the Hessian of the optimal transport potential*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **6** (2007), 441–456.
[18] C. VILLANI, "Optimal Transport, Old and New", Grundlehren des Mathematischen Wissenschaften, Vol. 338, Springer-Verlag Berlin Heidelberg, 2009.

Institute for Theoretical Studies
ETH Zürich, Clausiusstrasse 47
CH-8092 Zürich, Switzerland
maria.colombo@eth-its.ethz.ch

The University of Texas at Austin
Mathematics Dept. RLM 8.100
2515 Speedway Stop C1200
Austin, TX 78712, USA
figalli@math.utexas.edu
yjhaveri@math.utexas.edu

Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1521-1549

A moving lemma for cycles with very ample modulus

AMALENDU KRISHNA AND JINHYUN PARK

Abstract. We prove a moving lemma for higher Chow groups with modulus, in the sense of Binda-Kerz-Saito, of projective schemes, when the modulus is given by a very ample divisor. This provides one of the first cases of moving lemmas for cycles with modulus, not covered by the additive higher Chow groups. We apply this to prove a contravariant functoriality of higher Chow groups with modulus. We use our moving techniques to show that the higher Chow groups of a line bundle over a scheme, with the 0-section as the modulus, vanish.

Mathematics Subject Classification (2010): 14C25 (primary); 13F35, 19E15 (secondary).

1. Introduction

The moving lemma is one of the most important technical tools in dealing with algebraic cycles. For usual higher Chow groups, this was established by S. Bloch (see [2,3]). In order to study the relative K-theory of schemes (relative to effective divisors) in terms of algebraic cycles, the theory of additive higher Chow groups (see [5, 9, 10, 14]) and cycles with modulus (see [1, 8]) were recently introduced. But the lack of a moving lemma has been an annoying hindrance in the study of these additive higher Chow groups and the Chow groups with modulus.

A moving lemma for additive higher Chow groups of smooth projective schemes was proven in [10]. A similar moving lemma for the additive higher Chow groups of smooth affine schemes has been very recently established by W. Kai [7], along with some more general results after Nisnevich sheafifications. However, without such modifications, one does not yet know the existence of a moving lemma for the higher Chow groups with modulus which do not arise from additive higher Chow groups.

JP was partially supported by the National Research Foundation of Korea (NRF) grant No. 2015R1A2A2A01004120 funded by the Korean government (MSIP). AK was partially supported by the Swarna Jayanti Fellowship, 2011.

Received September 17, 2015; accepted in revised form September 27, 2016.
Published online December 2017.

1.1. Main results

The goal of this paper is to address the moving lemma problem for the higher Chow groups with modulus of projective schemes when the modulus divisor is very ample. Our main result is the following. The necessary definitions are recalled in Section 2.

Theorem 1.1. *Let X be an equidimensional reduced projective scheme of dimension $d \geq 1$ over a field k. Let $D \subsetneq X$ be a very ample effective Cartier divisor such that $X \setminus D$ is smooth over k. Let $\mathcal{W}$ be a finite collection of locally closed subsets of X. Then, the inclusion $z^q_{\mathcal{W}}(X|D, \bullet) \hookrightarrow z^q(X|D, \bullet)$ is a quasi-isomorphism.*

Our first application of Theorem 1.1 is the following complete solution of the moving lemma for cycles with arbitrary modulus on projective spaces. The analogous question for cycles on affine spaces was solved by W. Kai [7].

Corollary 1.2. *Let k be any field and $r \geq 1$ be any integer. Let $D \subset \mathbb{P}^r_k$ be any effective Cartier divisor. Let $\mathcal{W}$ be a finite collection of locally closed subsets of $\mathbb{P}^r_k$. Then the inclusion $z^q_{\mathcal{W}}(\mathbb{P}^r_k|D, \bullet) \hookrightarrow z^q(\mathbb{P}^r_k|D, \bullet)$ is a quasi-isomorphism.*

In the second application of Theorem 1.1, we prove the following contravariance property of the higher Chow groups with modulus.

Theorem 1.3. *Let $f : Y \to X$ be a morphism of equidimensional reduced quasi-projective schemes over a field k, where X is projective over k. Let $D \subset X$ be a very ample effective Cartier divisor such that $X \setminus D$ is smooth over k. Suppose that $f^*(D)$ is a Cartier divisor on Y (*i.e.*, no minimal or embedded component of Y maps into D). Then there exists a map*

$$f^* : z^q(X|D, \bullet) \to z^q(Y|f^*(D), \bullet)$$

in the derived category of Abelian groups. In particular, for every $p, q \geq 0$, there is a pull-back

$$f^* : \mathrm{CH}^q(X|D, p) \to \mathrm{CH}^q(Y|f^*(D), p).$$

Corollary 1.4. *Let $r \geq 1$ be an integer and let $f : Y \to \mathbb{P}^r_k$ be a morphism of quasi-projective schemes over a field k. Let $D \subset \mathbb{P}^r_k$ be an effective Cartier divisor such that $f^*(D)$ is a Cartier divisor on Y. Then, there exists a pull-back $f^* : \mathrm{CH}^q(\mathbb{P}^r_k|D, p) \to \mathrm{CH}^q(Y|f^*(D), p)$ for every $p, q \geq 0$.*

As a final application of our moving techniques, we prove the following vanishing theorem for the higher Chow groups of a line bundle on a scheme with the modulus given by the 0-section. This provides examples where the higher Chow groups of a variety with a modulus in an effective Cartier divisor are all zero. As one knows, this is not possible for the ordinary higher Chow groups. This also gives an evidence in support of the expectation that the higher Chow groups with modulus are the relative motivic cohomology.

Theorem 1.5. *Let X be a quasi-projective scheme over a field k and let $f : \mathcal{L} \to X$ be a line bundle. Let $\iota : X \hookrightarrow \mathcal{L}$ denote the 0-section embedding. Then, the cycle complex $z_s(\mathcal{L}|X, \bullet)$ is acyclic for all $s \in \mathbb{Z}$.*

1.2. Outline of proofs

We prove Theorem 1.1 by following the classical approach used by Bloch to prove his moving lemma for ordinary higher Chow groups of smooth projective schemes. We first prove the above theorem for projective spaces. The main difficulty here lies in constructing suitable homotopy varieties and to check their modulus condition. We solve this problem by using some blow-up techniques and our homotopy varieties are very different from the one used classically.

To deal with the case of general projective schemes, we use the method of linear projections. However, we need to make more subtle choices of our linear subspaces than in the classical case due to the presence of the modulus.

We show later in this article how this method breaks down if we replace a very ample divisor by just an ample one. We show that the linear projection method cannot be used in general to prove the moving lemma for Chow groups with modulus on either smooth affine or smooth projective schemes, if the modulus divisor is not very ample. This suggests that the general case of the moving lemma for Chow groups with modulus on smooth affine or projective schemes may be a very challenging task.

ACKNOWLEDGEMENTS. The authors are deeply indebted to the referee, who so thoroughly read the paper and suggested many valuable corrections and simplifications.

2. Recalls on cycles with modulus

In this section we recollect some necessary definitions and notation associated with cycles with modulus. Let k be a field and let $\mathbf{Sch}_k$ denote the category of quasi-projective schemes over k. Let $\mathbf{Sm}_k$ denote the full subcategory of $\mathbf{Sch}_k$ consisting of smooth schemes.

2.1. Notation

Set $\mathbb{A}^1_k := \operatorname{Spec} k[t]$, $\mathbb{P}^1_k := \operatorname{Proj} k[Y_0, Y_1]$ and let $y := Y_0/Y_1$ be the coordinate on $\mathbb{P}^1_k$. We set $\square := \mathbb{A}^1_k$ and $\overline{\square} := \mathbb{P}^1_k$. We use the coordinate system $(y_1, \cdots, y_n)$ on $\overline{\square}^n$ with $y_i := y \circ q_i$, where $q_i : \overline{\square}^n \to \overline{\square}$ is the projection onto the i-th $\overline{\square}$. For $i = 1, \dots, n$, let $F^\infty_{n,i}$ be the Cartier divisor on $\overline{\square}^n$ defined by $\{y_i = \infty\}$. Let F^∞_n denote the Cartier divisor $\sum_{i=1}^n F^\infty_{n,i}$ on $\overline{\square}^n$. A *face* of $\overline{\square}^n$ is a closed subscheme defined by a set of equations of the form $\{y_{i_1} = \epsilon_1, \dots, y_{i_s} = \epsilon_s \mid \epsilon_j \in \{0, 1\}\}$. For $\epsilon = 0, 1$, and $i = 1, \cdots, n$, let $\iota_{n,i,\epsilon} : \overline{\square}^{n-1} \hookrightarrow \overline{\square}^n$ be the inclusion

$$\iota_{n,i,\epsilon}(y_1, \dots, y_{n-1}) = (y_1, \dots, y_{i-1}, \epsilon, y_i, \dots, y_{n-1}). \tag{2.1}$$

A face of $\square^n$ is an intersection of $\square^n$ with a face of $\overline{\square}^n$.

2.2. Cycles with modulus

Let $X \in \mathbf{Sch}_k$. Recall ([11, Section 2]) that for effective Cartier divisors D_1 and D_2 on X, we say $D_1 \leq D_2$ if $D_1 + D = D_2$ for some effective Cartier divisor D on X. A *modulus pair* or a *scheme with an effective divisor* is a pair (X, D), where $X \in \mathbf{Sch}_k$ and D an effective Cartier divisor on X. A morphism $f : (Y, E) \to (X, D)$ of modulus pairs is a morphism $f : Y \to X$ in $\mathbf{Sch}_k$ such that $f^*(D)$ is defined as a Cartier divisor on Y and $f^*(D) \leq E$. In particular, $f^{-1}(D) \subset E$. If $f : Y \to X$ is a morphism of k-schemes, and (X, D) is a modulus pair such that $f^{-1}(D) = \emptyset$, then $f : (Y, \emptyset) \to (X, D)$ is a morphism of modulus pairs.

Definition 2.1 ([1,8]). Let (X, D) and $(\overline{Y}, E)$ be two modulus pairs. Let $Y = \overline{Y} \setminus E$. Let $V \subset X \times Y$ be an integral closed subscheme with closure $\overline{V} \subset X \times \overline{Y}$. We say V *has modulus* D *on* $X \times Y$ (relative to E) if $\nu_V^*(D \times \overline{Y}) \leq \nu_V^*(X \times E)$ on $\overline{V}^N$, where $\nu_V : \overline{V}^N \to \overline{V} \hookrightarrow X \times \overline{Y}$ is the normalization followed by the closed immersion.

Definition 2.2 ([1,8]). Let (X, D) be a modulus pair. For $s \in \mathbb{Z}$ and $n \geq 0$, let $\underline{z}_s(X|D, n)$ be the free Abelian group on integral closed subschemes $V \subset X \times \square^n$ of dimension $s + n$ satisfying the following conditions:

(1) (Face condition) for each face $F \subset \square^n$, V intersects $X \times F$ properly;
(2) (Modulus condition) V has modulus D relative to F_n^∞ on $X \times \square^n$.

We usually drop the phrase "relative to F_n^∞" for simplicity. A cycle in $\underline{z}_s(X|D, n)$ is called an *admissible cycle with modulus* D. The following containment lemma is from [11, Proposition 2.4] (see also [1, Lemma 2.1] and [10, Proposition 2.4]).

Proposition 2.3. *Let* (X, D) *and* $(\overline{Y}, E)$ *be modulus pairs and* $Y = \overline{Y} \setminus E$. *If* $V \subset X \times Y$ *is a closed subscheme with modulus* D *relative to* E, *then any closed subscheme* $W \subset V$ *also has modulus* D *relative to* E.

One checks using Proposition 2.3 that $(n \mapsto \underline{z}_s(X|D, n))$ is a cubical Abelian group. In particular, the groups $\underline{z}_s(X|D, n)$ form a complex with the boundary map $\partial = \sum_{i=1}^n (-1)^i (\partial_i^0 - \partial_i^1)$, where $\partial_i^\epsilon = \iota_{n,i,\epsilon}^*$.

Definition 2.4 ([1,8]). The complex $(z_s(X|D, \bullet), \partial)$ is the nondegenerate complex associated to $(n \mapsto \underline{z}_s(X|D, n))$, *i.e.*, $z_s(X|D, n) := \underline{z}_s(X|D, n)/\underline{z}_s(X|D, n)_{\mathrm{degn}}$. The homology $\mathrm{CH}_s(X|D, n) := \mathrm{H}_n(z_s(X|D, \bullet))$ for $n \geq 0$ is called *higher Chow group of* X *with modulus* D. If X is equidimensional of dimension d, for $q \geq 0$, we write $\mathrm{CH}^q(X|D, n) = \mathrm{CH}_{d-q}(X|D, n)$.

The following is a generalization of [11, Proposition 2.12] (see also [1, Lemma 2.7]). The reader can check that the only requirement in the proof of [11, Proposition 2.12] is that the underlying map be flat over the complement of the modulus divisor. This is because of the fact that an admissible cycle lies completely over this complement.

Lemma 2.5. *Let $f : Y \to X$ be a morphism in $\mathbf{Sch}_k$. Let $D \subsetneq X$ be an effective Cartier divisor. Assume that $f^*(D)$ is a Cartier divisor on Y such that the map $f^{-1}(X \setminus D) \to X \setminus D$ is flat of relative dimension d. Then, there is a pull-back map $f^* : z_r(X|D, \bullet) \to z_{d+r}(Y|f^*(D), \bullet)$ such $(f \circ g)^* = g^* \circ f^*$.*

We often use the following result from [11, Lemma 2.2]:

Lemma 2.6. *Let $f : Y \to X$ be a dominant map of normal integral k-schemes. Let D be a Cartier divisor on X such that the generic points of* $\mathrm{Supp}(D)$ *are contained in $f(Y)$. Suppose that $f^*(D) \geq 0$ on Y. Then $D \geq 0$ on X.*

Definition 2.7. Let $\mathcal{W}$ be a finite set of locally closed subsets of X and let $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$ be a set function. Let $\underline{z}^q_{\mathcal{W},e}(X|D, n)$ be the subgroup generated by integral cycles $Z \in \underline{z}^q(X|D, n)$ such that for each $W \in \mathcal{W}$ and each face $F \subset \square^n$, we have $\mathrm{codim}_{W \times F}(Z \cap (W \times F)) \geq q - e(W)$. They form a subcomplex $\underline{z}^q_{\mathcal{W},e}(X|D, \bullet)$ of $\underline{z}^q(X|D, \bullet)$. Modding out by degenerate cycles, we obtain the subcomplex $z^q_{\mathcal{W},e}(X|D, \bullet) \subset z^q(X|D, \bullet)$. We write $z^q_{\mathcal{W}}(X|D, \bullet) := z^q_{\mathcal{W},0}(X|D, \bullet)$. The number $e(W)$ is called the *excess dimension* of the intersection $Z \cap (W \times F)$. Given a function $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$, define $(e - 1) : \mathcal{W} \to \mathbb{Z}_{\geq 0}$ by $(e - 1)(W) = \max\{e(W) - 1, 0\}$. This gives an inclusion $z^q_{\mathcal{W},e-1}(X|D, \bullet) \subset z^q_{\mathcal{W},e}(X|D, \bullet)$.

We also use the following from [12, Proposition 4.3] in our proof of the moving lemma.

Proposition 2.8 (Spreading lemma). *Let $k \subset K$ be a purely transcendental extension. Let (X, D) be a smooth quasi-projective k-scheme with an effective Cartier divisor, and let $\mathcal{W}$ be a finite collection of locally closed subsets of X. Let (X_K, D_K) and $\mathcal{W}_K$ be the base changes via* $\mathrm{Spec}\,(K) \to \mathrm{Spec}\,(k)$. *Let $\mathrm{pr}_{K/k} : X_K \to X_k$ be the base change map. Then for every set function $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$, the pull-back maps*

$$\mathrm{pr}^*_{K/k} : \frac{z^q_{\mathcal{W},e}(X|D, \bullet)}{z^q_{\mathcal{W}}(X|D, \bullet)} \to \frac{z^q_{\mathcal{W}_K,e}(X_K|D_K, \bullet)}{z^q_{\mathcal{W}_K}(X_K|D_K, \bullet)} \tag{2.2}$$

and

$$\mathrm{pr}^*_{K/k} : \frac{z^q_{\mathcal{W},e}(X|D, \bullet)}{z^q_{\mathcal{W},e-1}(X|D, \bullet)} \to \frac{z^q_{\mathcal{W}_K,e}(X_K|D_K, \bullet)}{z^q_{\mathcal{W}_K,e-1}(X_K|D_K, \bullet)} \tag{2.3}$$

are injective on homology.

We remark that Proposition 2.8 is stated in [12, Proposition 4.3] only for (2.2) but the argument given there simultaneously proves (2.3) as well.

3. Moving lemma for projective spaces

In this section we prove our moving lemma for the modulus pair (X, D), where X is a projective space over k and D is a hyperplane in X. We use the following:

Lemma 3.1 ([2, Lemma 1.2]). *Let $X \in \mathbf{Sch}_k$ and let G be a connected algebraic group over k acting on X. Let $A, B \subset X$ be closed subsets. Assume that the fibers of the action map $G \times A \to X$, given by $(g, a) \mapsto g \cdot a$, all have the same dimension and that this map is dominant.*

Assume moreover that there is an overfield $k \hookrightarrow K$ and a K-morphism $\psi : X_K \to G_K$. Let $\emptyset \neq U \subset X$ be open such that for every $x \in U_K$, we have

$$\mathrm{tr.deg}_k\,(\phi \circ \psi(x), \pi(x)) \geq \dim(G),$$

where $\pi : X_K \to X$ and $\phi : G_K \to G$ are the base changes. Define $\theta : X_K \to X_K$ by $\theta(x) = \psi(x) \cdot x$ and assume that θ is an isomorphism. Then, the intersection $\theta(A_K \cap U_K) \cap B_K$ is proper.

Corollary 3.2. *Let $X \in \mathbf{Sch}_k$ and let G be a connected algebraic group over k acting transitively on X. Let $Y \in \mathbf{Sch}_k$ and let $\emptyset \neq A \subset X$ and $B \subset X \times Y$ be closed subsets. Let G act on $X \times Y$ by $g \cdot (x, y) = (g \cdot x, y)$.*

Let $K = k(G)$ and let $\phi : G_K \to G$ be the base change. Suppose $\psi : (X \times Y)_K \to G_K$ is a K-morphism and let $U \hookrightarrow X \times Y$ be an open subset such that:

1. *the image of every point of U_K under the composite map $(X \times Y)_K \xrightarrow{\psi} G_K \xrightarrow{\phi} G$ is the generic point of G;*
2. *the map $\theta : (X \times Y)_K \to (X \times Y)_K$ given by $\theta(z) = \psi(z) \cdot z$, is an isomorphism.*

Then the intersection $\theta((A \times Y)_K \cap U_K) \cap (B_K \cap U_K)$ is proper on U_K.

We let $\mathbb{A}^r_k = \mathrm{Spec}\,(k[x_1, \cdots, x_r])$ and let $\mathbb{P}^r_k = \mathrm{Proj}(k[X_1, \cdots, X_r, X_0])$, where we set $x_i = X_i/X_0$ for $1 \leq i \leq r$. This yields an open immersion $j_0 : \mathbb{A}^r_k \hookrightarrow \mathbb{P}^r_k$. Let $H_\infty = \mathbb{P}^r_k \setminus \mathbb{A}^r_k$ be the hyperplane at infinity. We write the homogeneous coordinates of $\mathbb{P}^r_k$ as $(X_1; \cdots; X_r; X_0)$. We fix this choice of coordinates of $\mathbb{A}^r_k$ and $\mathbb{P}^r_k$. Set $u = \prod_{i=1}^r x_i \in k[x_1, \cdots, x_r]$.

Let $K = k(\mathbb{P}^r_k)$ and consider the point $\eta = (u, \cdots, u) \in \mathbb{P}^r_K$ so that its image under the projection $\mathbb{P}^r_K \to \mathbb{P}^r_k$ is the generic point of $\mathbb{P}^r_k$. Let $U_+ \hookrightarrow \mathbb{P}^r_K \times \overline{\square}_K$ be the open subset $(\mathbb{P}^r_K \times \square_K) \cup (\mathbb{A}^r_K \times \overline{\square}_K)$ and set $\mathcal{Y} = H_\infty \times \{\infty\} = (\mathbb{P}^r_K \times \overline{\square}_K) \setminus U_+$. For K-schemes X and X', we write the product $X \times_K X'$ as $X \times X'$.

Lemma 3.3. *Let $\phi_\eta : \mathbb{A}^r_K \times \square_K \to \mathbb{A}^r_K$ denote the map $\phi_\eta(x, t) = x + \eta \cdot t$. Then, ϕ_η uniquely extends to a morphism $\phi_\eta|_{U_+} : U_+ \to \mathbb{P}^r_K$ such that the following hold:*

1. *U_+ is the largest open subset of $\mathbb{P}^r_K \times \overline{\square}_K$ over which ϕ_η can be extended to a regular morphism;*
2. *The extension of ϕ_η on $\mathbb{P}^r_K \times \square_K$ is a smooth morphism;*
3. *$(\phi_\eta|_{U_+})^{-1}(\mathbb{A}^r_K) = \mathbb{A}^r_K \times \square_K$;*
4. *$(\phi_\eta|_{U_+})^{-1}(H_\infty) = (\mathbb{A}^r_K \times \{\infty\}) + (H_\infty \times \square_K)$.*

Proof. Define the rational map $\phi_\eta : \mathbb{P}^r_K \times \overline{\square}_K \dashrightarrow \mathbb{P}^r_K$ by

$$\begin{aligned} &\phi_\eta((X_1; \cdots; X_r; X_0), (T_0; T_1)) \\ &= (T_1X_1 + uT_0X_0; \cdots; T_1X_r + uT_0X_0; T_1X_0). \end{aligned} \tag{3.1}$$

Note that $\phi_\eta((X_1; \cdots; X_r; 1), (t; 1)) = (X_1 + ut; \cdots; X_r + ut; 1)$ so that ϕ_η restricts to the given map on $\mathbb{A}^r_K \times \square_K$. One checks that (1), (3) and (4) hold from the shape of ϕ_η in (3.1).

To show (2), note that this map is the composite $\mathbb{P}^r_K \times \square_K \to \mathbb{P}^r_K \times \square_K \to \mathbb{P}^r_K$, where the first one is $((X_1; \cdots; X_r; X_0), t) \mapsto ((X_1 + utX_0; \cdots; X_r + utX_0; X_0), t)$ and the second is the projection to $\mathbb{P}^r_K$ (which is smooth). Since the first map is an isomorphism, it follows that ϕ_η is smooth on $\mathbb{P}^r_K \times \square_K$. □

Remark 3.4. The unique extension of ϕ_η to U_+ is not a flat morphism even though it is smooth on $\mathbb{P}^r_K \times \square_K$. If we set $V_i = \{(X_1; \cdots; X_r; X_0) | X_i \neq 0\} \hookrightarrow \mathbb{P}^r_K$ for $i = 1, \cdots, r$, then the map $\phi_\eta^{-1}(V_i) \to V_i$ is not flat because $\mathbb{A}^r_K \times \{0\}$ lies in one fiber but all other fibers have strictly smaller dimensions.

Our idea is to use the rational map $\phi_\eta : \mathbb{P}^r_K \times \overline{\square}_K \dashrightarrow \mathbb{P}^r_K$ to generate a homotopy between an arbitrary admissible cycle in $z^q(\mathbb{P}^r_k | H_\infty, \bullet)$ and a cycle in $z^q_{\mathcal{W},e}(\mathbb{P}^r_k | H_\infty, \bullet)$. In order to do so, we need to extend ϕ_η to an honest morphism of schemes. We achieve this in the following results via a sequence of blow-ups.

Lemma 3.5. *Let $\pi : \Gamma \to \mathbb{P}^r_K \times \overline{\square}_K$ be the blow-up of $\mathbb{P}^r_K \times \overline{\square}_K$ along the closed subscheme $\mathcal{Y} = H_\infty \times \{\infty\}$. Then, there exists a closed point $P_\infty \in \pi^{-1}(\mathcal{Y})$ and a regular map $\overline{\phi}_\eta : \Gamma_+ := \Gamma \setminus \{P_\infty\} \to \mathbb{P}^r_K$ such that $\pi : \Gamma_+ \to \mathbb{P}^r_K \times \overline{\square}_K$ is surjective, and the diagram*

$$\begin{tikzcd} \pi^{-1}(U_+) \arrow[r, hook, "\overline{j}"] \arrow[d, "\simeq"'] & \Gamma_+ \arrow[d, "\pi"'] \arrow[dr, "\overline{\phi}_\eta"] & \\ U_+ \arrow[r, hook, "j"] \arrow[rr, bend right] & \mathbb{P}^r_K \times \overline{\square}_K \arrow[r, dashed, "\phi_\eta"] & \mathbb{P}^r_K \end{tikzcd} \tag{3.2}$$

commutes.

Proof. Let $U_i \subsetneq \mathbb{P}^r_K$ be the open set $\{X_i \neq 0\}$ for $0 \leq i \leq r$. One checks by a direct local calculation the blow-up Γ has the following description. Over U_i, it is defined by

$$\begin{aligned} \pi^{-1}(U_i) = \Big\{ &\big((X_1; \cdots; X_r; X_0), (T_0; T_1), (Y_{1,i}; Y_{0,i})\big) \\ &\in U_i \times \overline{\square}_K \times \mathbb{P}^1_K | X_0T_0Y_{0,i} = X_iT_1Y_{1,i} \Big\} \end{aligned} \tag{3.3}$$

and these blow-ups glue along their intersections to make up Γ via the change of coordinate $Y_{0,i}/Y_{0,j} = (X_i/X_j)(Y_{1,i}/Y_{1,j})$ over $U_i \cap U_j$. The blow-up map $\pi : \pi^{-1}(U_i) \to U_i \times \overline{\square}_K$ is the composite $\pi^{-1}(U_i) \hookrightarrow U_i \times \overline{\square}_K \times \mathbb{P}^1_K \to U_i \times \overline{\square}_K$.

We now define a rational map $\overline{\phi}^i_\eta : \pi^{-1}(U_i) \dashrightarrow \mathbb{P}^r_K$ by

$$\begin{aligned}&\overline{\phi}_\eta\big((X_1; \cdots; X_r; X_0), (T_0; T_1), (Y_{1,i}; Y_{0,i})\big)\\ &= \big(Y_{0,i}X_1 + uX_iY_{1,i}; \cdots; Y_{0,i}X_r + uX_iY_{1,i}; Y_{0,i}X_0\big).\end{aligned} \tag{3.4}$$

The blow-up Γ is glued along $U_i \cap U_j$ via the automorphism $\psi_{i,j} : \pi^{-1}(U_i \cap U_j) \xrightarrow{\simeq} \pi^{-1}(U_i \cap U_j)$:

$$\begin{aligned}&\psi_{i,j}\big((X_1; \cdots; X_r; X_0), (T_0; T_1), (Y_{1,i}; Y_{0,i})\big)\\ &= \Big((X_1; \cdots; X_r; X_0), (T_0; T_1), (X_iX_j^{-1}Y_{1,i}; X_jX_i^{-1}Y_{0,i})\Big).\end{aligned}$$

It is clear from this isomorphism that $\psi_{i,j}(Y_{l,i} \neq 0) = (Y_{l,j} \neq 0)$ for $l = 0, 1$. Over $(Y_{0,i} \neq 0)$, we can let $Y_{0,i} = Y_{0,j} = 1$, $Y_{1,i} = y_i$ and $Y_{1,j} = y_j$. Over this open subset of $\pi^{-1}(U_i \cap U_j)$, we get

$$\begin{aligned}&\overline{\phi}^j_\eta \circ \psi_{i,j}\,((X_1; \cdots; X_r; X_0), (T_0; T_1), y_i)\\ &= \overline{\phi}^j_\eta\Big((X_1; \cdots; X_r; X_0), (T_0; T_1), X_iX_j^{-1}y_i\Big)\\ &= \Big(X_1 + uX_jX_iX_j^{-1}y_i; \cdots; X_r + uX_jX_iX_j^{-1}y_i; X_0\Big)\\ &= (X_1 + uX_iy_i; \cdots; X_r + uX_iy_i; X_0)\\ &= \overline{\phi}^i_\eta\,((X_1; \cdots; X_r; X_0), (T_0; T_1), y_i).\end{aligned} \tag{3.5}$$

Over the intersection of $\pi^{-1}(U_i \cap U_j)$ with the open subset $(Y_{1,i} \neq 0)$, we have

$$\begin{aligned}&\overline{\phi}^j_\eta \circ \psi_{i,j}\,((X_1; \cdots; X_r; X_0), (T_0; T_1), y_i)\\ &= \overline{\phi}^j_\eta\Big((X_1; \cdots; X_r; X_0), (T_0; T_1), X_jX_i^{-1}y_i\Big)\\ &= \Big(X_jX_i^{-1}X_1y_i + uX_j; \cdots X_jX_i^{-1}X_ry_i + uX_j; X_i^{-1}X_jX_0y_i\Big)\\ &= \big(X_1X_jy_i + uX_iX_j; \cdots; X_rX_jy_i + uX_iX_j; X_jX_0y_i\big)\\ &= (X_1y_i + uX_i; \cdots; X_ry_i + uX_i; X_0y_i)\\ &= \overline{\phi}^i_\eta\,((X_1; \cdots; X_r; X_0), (T_0; T_1), y_i).\end{aligned} \tag{3.6}$$

It follows from (3.5) and (3.6) that $\overline{\phi}^j_\eta$'s glue together to yield a rational map $\overline{\phi}_\eta : \Gamma \dashrightarrow \mathbb{P}^r_K$ such that $\overline{\phi}_\eta|_{\pi^{-1}(U_i)} = \overline{\phi}^j_\eta$ for $0 \le i \le r$.

We next show the commutativity of (3.2). The left square of (3.2) commutes by construction. We thus have to show that $\overline{\phi}_\eta \circ \overline{j} = \phi_\eta \circ \pi$, *i.e.*, the trapezoid in (3.2) commutes. It suffices to show this over each open subset $(U_i \times \overline{\square}_K) \cap U_+$.

If $P = \big((X_1; \cdots; X_r; X_0), (T_0; T_1), (Y_{1,i}; Y_{0,i})\big) \in \pi^{-1}(U_+)$, we have $\pi(P) = ((X_1; \cdots; X_r; X_0), (T_0; T_1))$ such that either $T_1 \neq 0$ or $X_0 \neq 0$.

Suppose first that $T_1 \neq 0$. Then, we can take $T_1 = 1$ and $T_0 = t$. In this case, we must have $Y_{0,i} \neq 0$ so that we can assume $Y_{0,i} = 1$. Thus, the equation $X_0 T_0 Y_{0,i} = X_i T_1 Y_{1,i}$ becomes $Y_{1,i} = t X_0 X_i^{-1}$. This yields

$$\overline{\phi}^i_\eta \circ \overline{j}(P) = (X_1 + utX_0; \cdots; X_r + utX_0; X_0)$$

by (3.4) and

$$\phi_\eta \circ \pi(P) = (X_1 + utX_0; \cdots; X_r + utX_0; X_0)$$

by (3.1).

Suppose next that $X_0 \neq 0$. Since the case $T_1 \neq 0$ was already considered, we may suppose $T_0 \neq 0$. Thus, we may take $T_0 = 1$ and $T_1 = t$. In this case, we must have $Y_{1,i} \neq 0$, so that we may take $Y_{1,i} = 1$. Thus, the equation $X_0 T_0 Y_{0,i} = X_i T_1 Y_{1,i}$ becomes $Y_{0,i} = t X_i X_0^{-1}$. This yields

$$\begin{aligned}\overline{\phi}^i_\eta \circ \overline{j}(P) &= (tX_1X_i + uX_0X_i; \cdots; tX_rX_i + uX_0X_i; tX_iX_0)\\ &= (tX_1 + X_0; \cdots; tX_r + X_0; tX_0)\end{aligned}$$

by (3.4). On the other hand, $\phi_\eta \circ \pi(P) = (tX_1 + uX_0; \cdots; tX_r + X_0; tX_0)$ by (3.1). We have thus shown that $\overline{\phi}_\eta \circ \overline{j}(P) = \phi_\eta \circ \pi(P)$ for $P \in \pi^{-1}(U_+)$.

We now show that $\overline{\phi}_\eta$ is regular on $\Gamma \setminus \{P_\infty\}$, where $P_\infty \in \big(\cap_{i=1}^r \pi^{-1}(U_i)\big)$ is the closed point $((1; \cdots; 1; 0), (1; 0), (1; -u))$ in the coordinates of $\pi^{-1}(U_i)$. Let $Q = \big((X_1; \cdots; X_r; X_0), (T_0; T_1), (Y_{1,i}; Y_{0,i})\big) \in \pi^{-1}(U_i)$ be a point so that $X_0 T_0 Y_{0,i} = X_i T_1 Y_{1,i}$. Then $\overline{\phi}_\eta(Q)$ is not defined if and only if all its coordinates are zero, *i.e.*,

$$Y_{0,i} X_j + u X_i Y_{1,i} = 0, \quad \text{for all } 1 \le j \le r, \text{ and } Y_{0,i} X_0 = 0. \tag{3.7}$$

If $Y_{0,i} = 0$ then $u X_i Y_{1,i} = 0$ for $1 \le i \le r$. But $u \in K^\times$ and $Q \in \pi^{-1}(U_i)$ imply that $Y_{1,i} = 0$, which cannot happen since $(Y_{1,i}; Y_{0,i}) \in \mathbb{P}^1_K$. So, $Y_{0,i} \neq 0$ and we must have $X_0 = 0$. Since $X_i \neq 0$, we can assume $X_i = 1$. Since $X_0 = 0$, we also have $T_1 Y_{1,i} = 0$, so that either $Y_{1,i} = 0$ or $T_1 = 0$. If $Y_{1,i} = 0$, then it follows from (3.7) that $Y_{0,i} = -u Y_{1,i} = 0$, which again is absurd because $(Y_{1,i}; Y_{0,i}) \in \mathbb{P}^1_K$. So, $Y_{1,i} \neq 0$, and $T_1 = 0$. We may assume $Y_{1,i} = 1$. Combining this with (3.7), we thus have

$$Y_{0,i} = -u, \;\; Y_{0,i} X_j + u = 0 \quad \text{for all } 1 \le j \neq i \le r \text{ and } X_0 = T_1 = 0. \tag{3.8}$$

We conclude that $\overline{\phi}_\eta(Q)$ is not defined if and only if $Q = ((1; \cdots; 1; 0), (1; 0), (1; -u))$. This proves the regularity of $\overline{\phi}_\eta$ on $\Gamma \setminus \{P_\infty\}$. Since $P_\infty \in \pi^{-1}(\mathcal{Y})$ and since each fiber of π over $\mathcal{Y}$ is 1-dimensional, we conclude that the map $(\Gamma \setminus \{P_\infty\}) \to \mathbb{P}^r_K \times \overline{\square}_K$ is surjective. This finishes the proof of the lemma. $\square$

Remark 3.6. The reader can check that the map $\phi_\eta : \mathbb{P}^r_K \times \overline{\square}_K \dashrightarrow \mathbb{P}^r_K$ is the one defined by the linear system generated by the global sections $S = \{T_1 X_i + uT_0 X_0\}_{1 \le i \le r} \cup \{T_1 X_0\}$ of the line bundle $\mathcal{O}(1,1)$. The sheaf of ideals $\mathcal{I}_\infty$ on $\mathbb{P}^r_K \times \overline{\square}_K$ defining $\mathcal{Y}$ is generated by $\{X_i T_1, X_0 T_0 | 0 \le i \le r\}$. Moreover, $\overline{\phi}_\eta : \Gamma \dashrightarrow \mathbb{P}^r_K$ is the rational map defined by the linear system generated by the global sections $\pi^*(S)$ of the line bundle $\pi^*\mathcal{I}_\infty$.

Let $\pi : \Gamma \to \mathbb{P}^r_K \times \overline{\square}_K$ be the blow-up map as in Lemma 3.5 and let $E = \pi^*(\mathcal{Y})$ denote the exceptional divisor for this blow-up. Note that the map $\pi : E \to \mathcal{Y} \simeq H_\infty$ is the $\mathbb{P}^1_K$-bundle associated to the vector bundle $\mathcal{O}(1) \oplus \mathcal{O}$.

Since $H_\infty \times \overline{\square}_K$ and $\mathbb{P}^r_K \times \{\infty\}$ are smooth schemes, and $\mathcal{Y}$ is a smooth divisor inside these schemes, note that $\mathrm{Bl}_{\mathcal{Y}}(H_\infty \times \overline{\square}_K) \to H_\infty \times \overline{\square}_K$ and $\mathrm{Bl}_{\mathcal{Y}}(\mathbb{P}^r_K \times \{\infty\}) \to \mathbb{P}^r_K \times \{\infty\}$ are isomorphisms.

Lemma 3.7. *Let $\pi : \Gamma \to \mathbb{P}^r_K \times \overline{\square}_K$ be as in Lemma* 3.5. *Then, we have the following.*

(1) $\mathrm{Bl}_{\mathcal{Y}}(H_\infty \times \overline{\square}_K) \cap \{P_\infty\} = \emptyset = \mathrm{Bl}_{\mathcal{Y}}(\mathbb{P}^r_K \times \{\infty\}) \cap \{P_\infty\}$;
(2) $\mathrm{Bl}_{\mathcal{Y}}(H_\infty \times \overline{\square}_K) \cap \mathrm{Bl}_{\mathcal{Y}}(\mathbb{P}^r_K \times \{\infty\}) = \emptyset$ *inside* Γ;
(3) $\pi^*(H_\infty \times \overline{\square}_K) = (H_\infty \times \overline{\square}_K) + E$ *and* $\pi^*(\mathbb{P}^r_K \times \{\infty\}) = (\mathbb{P}^r_K \times \{\infty\}) + E$ *in the group* $\mathrm{Div}(\Gamma)$ *of Weil divisors.*

Proof. It suffices to verify each statement of the lemma over an open subset $\pi^{-1}(U_i)$ with $0 \le i \le r$. On the other hand, (3.3) shows that over U_i, we have

$$\begin{aligned}
&\mathrm{Bl}_{\mathcal{Y}}(H_\infty \times \overline{\square}_K)\\
&= \{((X_1; \cdots; X_r; 0), (T_0; T_1), (Y_{1,i}; Y_{0,i})) \in \mathbb{P}^r_K \times \overline{\square}_K \times \mathbb{P}^1_K | Y_{1,i} = 0\}\\
&= H_\infty \times \overline{\square}_K \times \{0\}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\mathrm{Bl}_{\mathcal{Y}}(\mathbb{P}^r_K \times \{\infty\})\\
&= \{((X_1; \cdots; X_r; X_0), (1; 0), (Y_{1,i}; Y_{0,i})) \in \mathbb{P}^r_K \times \overline{\square}_K \times \mathbb{P}^1_K | Y_{0,i} = 0\}\\
&= \mathbb{P}^r_K \times \{\infty\} \times \{\infty\}.
\end{aligned}$$

Since P_∞ does not map to $\{0, \infty\} \subset \mathbb{P}^1_K$ under the projection $\pi^{-1}(U_i) \to \mathbb{P}^1_K$ for any $0 \le i \le r$, we get (1). The parts (2) and (3) of the lemma are immediate. $\square$

Let $\Gamma_1 \hookrightarrow \Gamma_+ \times \mathbb{P}^r_K$ denote the graph of $\overline{\phi}_\eta$ and let $\overline{\Gamma}_1 \hookrightarrow \Gamma \times \mathbb{P}^r_K$ be its closure. Let $\pi^N : \overline{\Gamma}_1^N \to \overline{\Gamma}_1 \hookrightarrow \Gamma \times \mathbb{P}^r_K$ be the normalization composed with the inclusion, and let $\pi_1 := \mathrm{pr}_1 \circ \pi^N$, $\pi_2 := \mathrm{pr}_2 \circ \pi^N$, where $\mathrm{pr}_1, \mathrm{pr}_2$ are the projections from $\Gamma \times \mathbb{P}^r_K$ to Γ and $\mathbb{P}^r_K$, respectively. Here, π^N is finite and π_1 is projective with $\pi_1^{-1}(\Gamma_+) \xrightarrow{\simeq} \Gamma_+$ such that $\pi_2|_{\Gamma_+} = \overline{\phi}_\eta$.

Since π_1 is a birational projective morphism and Γ is smooth, it follows from [6, Theorem II-7.17, page 166, Exercise II-7.11(c), page 171] that there is a closed subscheme $Z \hookrightarrow \Gamma$ such that $Z_{\text{red}} = \{P_\infty\}$ and $\overline{\Gamma}_1^N = \mathrm{Bl}_Z(\Gamma)$. Let $F \hookrightarrow \overline{\Gamma}_1^N$ denote the exceptional divisor for this blow-up so that $F_{\text{red}} = \pi_1^{-1}(P_\infty)$. Let $E_1 \hookrightarrow \overline{\Gamma}_1^N$ denote the strict transform of E under π_1 so that $\pi_1^*(E) = E_1 + F$.

Letting $\delta := \pi \circ \pi_1 : \overline{\Gamma}_1^N \to \mathbb{P}^r_K \times \overline{\square}_K$ and $E' := \pi_1^*(E) = E_1 + F$, a combination of Lemmas 3.5, 3.7 and the above construction proves the following.

Lemma 3.8. *There exists a commutative diagram*

$$\begin{tikzcd} \delta^{-1}(U_+) \arrow[r, hook, "j_1"] \arrow[d, "\simeq"'] & \overline{\Gamma}_1^N \arrow[d, "\delta"'] \arrow[dr, "\pi_2"] & \\ U_+ \arrow[r, hook, "j"] \arrow[rr, bend right] & \mathbb{P}^r_K \times \overline{\square}_K \arrow[r, dashed, "\phi_\eta"] & \mathbb{P}^r_K \end{tikzcd} \tag{3.9}$$

such that δ is a blow-up, and in the group $\mathrm{Div}(\overline{\Gamma}_1^N)$ *of Weil divisors, we have:*

$$\begin{aligned} \delta^*\left(H_\infty \times \overline{\square}_K\right) &= \left(H_\infty \times \overline{\square}_K\right) + E' \text{ and } \delta^*\left(\mathbb{P}^r_K \times \{\infty\}\right) \\ &= \left(\mathbb{P}^r_K \times \{\infty\}\right) + E'. \end{aligned} \tag{3.10}$$

For any map $f : X \to X'$ of K-schemes, let f_n denote the map

$$f \times \mathrm{Id}_{\overline{\square}^n_K} : X \times \overline{\square}^n_K \to X' \times \overline{\square}^n_K.$$

We now show how the rational map $\phi_\eta : \mathbb{P}^r_K \times \overline{\square}_K \dashrightarrow \mathbb{P}^r_K$ eventually leads to the desired homotopy.

Proposition 3.9. *Let $n \geq 1$ be an integer. Let $V \hookrightarrow \mathbb{P}^r_K \times \square^n_K$ be an integral closed subscheme. Assume that V has modulus H_∞ relative to F_n^∞. Let $\phi_\eta : \mathbb{A}^r_K \times \square_K \to \mathbb{P}^r_K$ be the map as in Lemma* 3.3. *Then, the closure of $\phi_{\eta,n}^{-1}(V)$ in $\mathbb{P}^r_K \times \square^{n+1}_K$ is an integral closed subscheme of $\mathbb{P}^r_K \times \square^{n+1}_K$ which has modulus H_∞ relative to F^∞_{n+1}.*

Proof. We use notations of the paragraph just before Lemma 3.8 and set $E'_n = E' \times \overline{\square}^n_K \hookrightarrow \overline{\Gamma}_1^N \times \overline{\square}^n_K$. Let $\overline{V} \hookrightarrow \mathbb{P}^r_K \times \overline{\square}^n_K$ denote the closure of V and let $\nu_V : \overline{V}^N \to \mathbb{P}^r_K \times \overline{\square}^n_K$ denote the induced map from the normalization of $\overline{V}$. By the modulus condition, we have

$$\nu_V^*\left(\mathbb{P}^r_K \times F_n^\infty\right) \geq \nu_V^*\left(H_\infty \times \overline{\square}^n_K\right) \text{ in } \mathrm{Div}\left(\overline{V}^N\right). \tag{3.11}$$

Condition (3.11) implies that $V \cap (H_\infty \times \square_K^n) = \emptyset$. Set $V' = \phi_{\eta,n}^{-1}(V)$. Since $\phi_{\eta,n}$ is smooth on $\phi_{\eta,n}^{-1}(\mathbb{A}_K^r \times \square_K^n)$ by Lemma 3.3, it follows that V' is an integral closed subscheme of $U_+ \times \overline{\square}_K^n$ with $\dim_K(V') = \dim_K(V) + 1$. Let $\overline{V}' \hookrightarrow \mathbb{P}_K^r \times \overline{\square}_K^{n+1}$ be the Zariski closure of V', and let $\nu_{V'} : \overline{V}'^N \to \overline{V}' \hookrightarrow \mathbb{P}_K^r \times \overline{\square}_K^{n+1}$ be the induced map from the normalization of $\overline{V}'$. Let $W \hookrightarrow \overline{\Gamma}_1^N \times \overline{\square}_K^n$ be the strict transform of $\overline{V}'$. It follows from Lemma 3.5 that $\pi_{2,n}(W \cap \delta_n^{-1}(U_+ \times \square_k^n)) = V$. Since $\pi_{2,n}$ is projective, we must have $\pi_{2,n}(W) = \overline{V}$. This yields a commutative diagram

$$\begin{tikzcd}
W^N \arrow[rr, "f"] \arrow[dr, "\nu_W"] \arrow[dd, "g"'] & & \overline{V}^N \arrow[d, "\nu_V"] \\
& \overline{\Gamma}_1^N \times \overline{\square}_K^n \arrow[r, "\pi_{2,n}"] \arrow[d, "\delta_n"] & \mathbb{P}_K^r \times \overline{\square}_K^n \\
\overline{V}'^N \arrow[r, "\nu_{V'}"'] & \mathbb{P}_K^r \times \overline{\square}_K^{n+1}, &
\end{tikzcd} \tag{3.12}$$

where ν_W is the normalization of W composed with its inclusion into $\overline{\Gamma}_1^N \times \overline{\square}_K^n$, and f and g are the maps induced by the universal property of normalization for dominant maps. Since f is a surjective map of integral schemes, condition (3.11) implies that $(\nu_V \circ f)^*(\mathbb{P}_K^r \times F_n^\infty) \geq (\nu_V \circ f)^*(H_\infty \times \overline{\square}_K^n)$ on W^N. In particular, we get $(\pi_{2,n} \circ \nu_W)^*(\mathbb{P}_K^r \times F_n^\infty) \geq (\pi_{2,n} \circ \nu_W)^*(H_\infty \times \overline{\square}_K^n)$ on W^N. Equivalently,

$$\nu_W^* \left(\overline{\Gamma}_1^N \times F_n^\infty \right) \geq \nu_W^* \left(\pi_2^*(H_\infty) \times \overline{\square}_K^n \right). \tag{3.13}$$

Since $(\phi_\eta|_{U_+})^*(H_\infty) = (\mathbb{A}_K^r \times \{\infty\}) + (H_\infty \times \square_K)$ by Lemma 3.3, we get

$$j_{1,n}^* \circ \pi_{2,n}^* \left(H_\infty \times \overline{\square}_K^n \right) = j_{1,n}^* \left(\mathbb{P}_K^r \times F_{n,n+1}^\infty \right) + j_{1,n}^* \left(H_\infty \times \overline{\square}_K^{n+1} \right),$$

where $j_1 : U_+ \hookrightarrow \overline{\Gamma}_1^N$ is the inclusion. Since $\mathbb{P}_K^r \times F_{n,n+1}^\infty$ and $H_\infty \times \overline{\square}_K^{n+1}$ are irreducible, we get $\pi_2^*(H_\infty) \times \overline{\square}_K^n \geq (\mathbb{P}_K^r \times F_{n,n+1}^\infty) + (H_\infty \times \overline{\square}_K^{n+1})$ on $\overline{\Gamma}_1^N \times \overline{\square}_K^n$. Combining this with (3.13), we get

$$\begin{aligned} \nu_W^* \left(\overline{\Gamma}_1^N \times F_n^\infty \right) &\geq \nu_W^* \left(\mathbb{P}_K^r \times F_{n,n+1}^\infty \right) + \nu_W^* \left(H_\infty \times \overline{\square}_K^{n+1} \right) \\ &\geq \nu_W^* \left(H_\infty \times \overline{\square}_K^{n+1} \right). \end{aligned} \tag{3.14}$$

This in turn implies that

$$
\begin{aligned}
(\delta_n \circ \nu_W)^* \left(\mathbb{P}^r_K \times F^\infty_{n+1}\right) &= (\delta_n \circ \nu_W)^* \left(\mathbb{P}^r_K \times F^\infty_n \times \overline{\square}_K\right) \\
&\quad + (\delta_n \circ \nu_W)^* \left(\mathbb{P}^r_K \times \overline{\square}^n_K \times \{\infty\}\right) \\
&= \nu_W^* \left(\overline{\Gamma}^N_1 \times F^\infty_n\right) + (\delta_n \circ \nu_W)^* \left(\mathbb{P}^r_K \times \overline{\square}^n_K \times \{\infty\}\right) \\
&\geq \nu_W^* \left(H_\infty \times \overline{\square}^{n+1}_K\right) + (\delta_n \circ \nu_W)^* \left(\mathbb{P}^r_K \times \overline{\square}^n_K \times \{\infty\}\right) \\
&=^\dagger \nu_W^* \left(H_\infty \times \overline{\square}^{n+1}_K\right) + \nu_W^*(E'_n) + \nu_W^* \left(\mathbb{P}^r_K \times \overline{\square}^n_K \times \{\infty\}\right) \\
&=^\ddagger (\delta_n \circ \nu_W)^* \left(H_\infty \times \overline{\square}^{n+1}_K\right) + \nu_W^* \left(\mathbb{P}^r_K \times \overline{\square}^n_K \times \{\infty\}\right) \\
W &\geq (\delta_n \circ \nu_W)^* \left(H_\infty \times \overline{\square}^{n+1}_K\right),
\end{aligned}
$$

where $=^\dagger$ and $=^\ddagger$ follow from Lemma 3.8. Using (3.12), this gives $g^*(\nu^*_{V'}(\mathbb{P}^r_K \times F^\infty_{n+1})) \geq g^*(\nu^*_{V'}(H_\infty \times \overline{\square}^{n+1}_K))$. Since g is a surjective map of integral normal schemes, we conclude by Lemma 2.6 that $\nu^*_{V'}(\mathbb{P}^r_K \times F^\infty_{n+1}) \geq \nu^*_{V'}(H_\infty \times \overline{\square}^{n+1}_K)$. $\square$

Theorem 3.10. *Given an integer $r \geq 1$, let $D \hookrightarrow \mathbb{P}^r_k$ be a hyperplane. Let $\mathcal{W} = \{W_1, \cdots, W_s\}$ be a finite collection of locally closed subsets of $\mathbb{P}^r_k$ and let $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$ be a set function. Then, the inclusion $z^q_{\mathcal{W}}(\mathbb{P}^r_k|D, \bullet) \hookrightarrow z^q_{\mathcal{W},e}(\mathbb{P}^r_k|D, \bullet)$ is a quasi-isomorphism. In particular, the inclusion $z^q_{\mathcal{W}}(\mathbb{P}^r_k|D, \bullet) \hookrightarrow z^q(\mathbb{P}^r_k|D, \bullet)$ is a quasi-isomorphism.*

Proof. The second part follows easily from the first part because $z^q(\mathbb{P}^r_k|D, \bullet) = z^q_q(X|D, \bullet)$. We shall prove the first part of the theorem in several steps. We can find a linear automorphism $\tau : \mathbb{P}^r_k \xrightarrow{\simeq} \mathbb{P}^r_k$ such that $\tau(D) = H_\infty$. Replacing $\mathcal{W}$ by $\tau(\mathcal{W})$, we reduce to the case when $D = H_\infty$, condition that we assume from now on. In view of Proposition 2.8, we only need to show that the map

$$
\mathrm{pr}^*_{K/k} : \frac{z^q_{\mathcal{W},e}(\mathbb{P}^r_k|D, \bullet)}{z^q_{\mathcal{W}}(\mathbb{P}^r_k|D, \bullet)} \to \frac{z^q_{\mathcal{W}_K,e}(\mathbb{P}^r_K|D_K, \bullet)}{z^q_{\mathcal{W}_K}(\mathbb{P}^r_K|D_K, \bullet)}
$$

is zero on the homology, where we choose $K = k(\mathbb{P}^r_k)$.

Following the notation we used so far in this section, consider the maps

$$
\mathbb{A}^r_K \times \square^{n+1}_K \xrightarrow{\phi_{\eta,n}} \mathbb{P}^r_K \times \square^n_K \xrightarrow{\mathrm{pr}_{K/k}} \mathbb{P}^r_k \times \square^n_k.
$$

For any irreducible cycle $V \hookrightarrow \mathbb{P}^r_k \times \square^n_k$, let $H^*_n(V) = (\mathrm{pr}_{K/k} \circ \phi_{\eta,n})^{-1}(V)$ and let $\overline{H}^*_n(V)$ be its closure in $\mathbb{P}^r_K \times \square^{n+1}_K$. We can extend this linearly to cycles in $z^q(\mathbb{P}^r_k|D, n)$.

Suppose V is an irreducible cycle in $z^q_{\mathcal{W},e}(\mathbb{P}^r_k|D, n)$. We claim that:

(1) $\overline{H}^*_n(V) \in z^q_{\mathcal{W}_K,e}(\mathbb{P}^r_K|D_K, n+1)$;

(2) $\overline{H}^*_n(V) \in z^q_{\mathcal{W}_K}(\mathbb{P}^r_K|D_K, n+1)$ if $V \in z^q_{\mathcal{W}}(\mathbb{P}^r_k|D, n)$;

(3) $\iota^*_{n+1,n+1,0}(\overline{H}^*_n(V)) = V$ and $\iota^*_{n+1,n+1,1}(\overline{H}^*_n(V)) \in z^q_{\mathcal{W}_K}(\mathbb{P}^r_K|D_K, n)$.

We now prove this claim using the previous results of this section. Since V has modulus D on $\mathbb{P}^r_k \times \square^n_k$, it follows that V is a closed subscheme of $\mathbb{A}^r_k \times \square^n_k$. In particular, $V \in z^q_{\mathcal{W}^0,e}(\mathbb{A}^r_k, n)$, where $\mathcal{W}^0 = \{W_1 \cap \mathbb{A}^r_k, \cdots, W_s \cap \mathbb{A}^r_k\}$. Since $\overline{H}^*_n(V)$ has modulus D on $\mathbb{P}^r_K \times \square^{n+1}_K$ by Proposition 3.9, it follows that $\overline{H}^*_n(V)$ is an integral closed subscheme of $\mathbb{A}^r_K \times \square^{n+1}_K$. In particular, $\overline{H}^*_n(V) = H^*_n(V)$. This shows that we can replace $\mathbb{P}^r_k$, $\overline{H}^*_n(V)$ and $\mathcal{W}$ by $\mathbb{A}^r_k$, $H^*_n(V)$ and $\mathcal{W}^0$ respectively, to prove the claim.

We prove (3) first. By the definition of ϕ_η, we have $\iota^*_{n+1,n+1,0}(H^*_n(V)) = V$. In particular, $H^*_n(V)$ intersects $F_{n+1,n+1,0}$ and all its faces properly. We thus have to show that $\iota^*_{n+1,n+1,1}(H^*_n(V)) \in z^q_{\mathcal{W}^0_K}(\mathbb{A}^r_K|D_K, n)$ to prove (3).

Let $\mathbb{A}^r_k$ act on itself by translation and let it act on $\mathbb{A}^r_k \times \square^n_k$ by acting trivially on $\square^n_k = \square^n_k \times \{1\} \hookrightarrow \square^{n+1}_k$. Consider the map $\psi : \mathbb{A}^r_K \times \square^n_K \to \mathbb{A}^r_K$ defined by $\psi(x, y) = \eta$. The reader can check that the assumptions of Corollary 3.2 are satisfied. Applying this corollary to each $A = \overline{W_i \cap \mathbb{A}^r_k}$ (where the closure is taken in $\mathbb{A}^r_k$) and $B = \mathbb{A}^r_k \times F$ for any face F of $\square^n_k \times \{1\}$, we deduce $\iota^*_{n+1,n+1,1}(H^*_n(V)) \in z^q_{\mathcal{W}^0_K}(\mathbb{A}^r_K|D_K, n)$. We have thus proven (3). Since (2) is a special case of (1) where we take $e = 0$, we are left with proving (1).

To prove (1), it is enough to consider the case when $\mathcal{W} = \{W\}$ is a singleton. Note $V \in z^q_{W,e}(\mathbb{A}^r_k, n)$ and let $F \hookrightarrow \square^{n+1}_K$ be any face. If $F \hookrightarrow \square^n_K \times \{0\}$, then the intersection $H^*_n(V) \cap (W \times F)$ has the desired dimension because $\iota^*_{n+1,n+1,0}(H^*_n(V)) = V$ and $V \in z^q_{W,e}(\mathbb{A}^r_k, n)$. We have already proven in (3) that the intersection $H^*_n(V) \cap (W \times F)$ is proper if $F \hookrightarrow \square^n_K \times \{1\}$. We can thus assume that $F = F'_K \times \square_K$, where F' is a face of $\square^n_k$.

Set $Z = V \cap (\mathbb{A}^r_k \times F')$. Consider the map $\psi : \mathbb{A}^r_K \times \square_K \times F'_K \to \mathbb{A}^r_K$ defined by $\psi(x, t, y) = \eta t$ and let $\theta : \mathbb{A}^r_K \times \square_K \times F'_K \to \mathbb{A}^r_K \times \square_K \times F'_K$ be given by $\theta(x, t, y) = (x + \eta t, t, y)$. Let $\mathbb{A}^r_k$ act by translation on itself and trivially on $\square_k \times F'$. Then $\theta(x, t, y) = \psi(x, t, y) \cdot (x, t, y)$. Applying Lemma 3.1 with $X = \mathbb{A}^r_k \times \square_k \times F'$, $A = \overline{W} \times \square_k \times F'$, $U = \mathbb{A}^r_k \times \mathbb{G}_{m,k} \times F'$, and $B = (V \times \square_k) \cap F_k = Z \times \square_k \hookrightarrow X \times F'$, it follows that the intersection $\theta(A_K) \cap B_K$ is proper away from $\mathbb{A}^r_K \times \{0\} \times F'_K$, *i.e.*, the intersection $(H^*_n(V) \cap F) \cap (W_K \times F)$ is proper away from $\mathbb{A}^r_K \times \{0\} \times F'_K$.

On the other hand, since $V \in z^q_{W,e}(\mathbb{A}^r_k, n)$ and hence V meets $W \times F'$ in excess dimension at most $e(W)$, it follows that $H^*_n(V) \cap F$ must meet $W \times F$ in excess dimension at most $e(W)$ along $\mathbb{A}^r_K \times \{0\} \times F'_K$. Thus $H^*_n(V)$ intersects $W_K \times F_K$

in excess dimension at most $e(W)$ for all faces $F_K \hookrightarrow \square_K^{n+1}$. In other words, $H_n^*(V) \in z^q_{W_K,e}(\mathbb{A}^r_K, n+1)$. This proves (1) and hence the claim.

It follows from the claim that there is a chain homotopy

$$H_\eta^* : \frac{z^q_{\mathcal{W},e}\big(\mathbb{P}^r_k|D,\bullet\big)}{z^q_{\mathcal{W}}\big(\mathbb{P}^r_k|D,\bullet\big)} \to \frac{z^q_{\mathcal{W}_K,e}\big(\mathbb{P}^r_K|D_K,\bullet\big)}{z^q_{\mathcal{W}_K}\big(\mathbb{P}^r_K|D_K,\bullet\big)}[-1],$$

and composed with the restriction map $\{1\} \hookrightarrow \square_k$, there is a chain map

$$H_{\eta,1}^* : \frac{z^q_{\mathcal{W},e}\big(\mathbb{P}^r_k|D,\bullet\big)}{z^q_{\mathcal{W}}\big(\mathbb{P}^r_k|D,\bullet\big)} \to \frac{z^q_{\mathcal{W}_K,e}\big(\mathbb{P}^r_K|D_K,\bullet\big)}{z^q_{\mathcal{W}_K}\big(\mathbb{P}^r_K|D_K,\bullet\big)}$$

such that $H_\eta^* \circ \partial + \partial \circ H_\eta^* = \mathrm{pr}^*_{K/k} - H^*_{\eta,1}$. Since $H^*_{\eta,1} = 0$ by the claim, we see that $\mathrm{pr}^*_{K/k}$ is zero on the homology. The proof of the theorem is complete. □

Corollary 3.11. *Given an integer $r \geq 1$, let $D \hookrightarrow \mathbb{P}^r_k$ be a hyperplane. Let $\mathcal{W} = \{W_1, \cdots, W_s\}$ be a finite collection of locally closed subsets of $\mathbb{P}^r_k$ and let $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$ be a set function. Then the inclusion $z^q_{\mathcal{W},e-1}(\mathbb{P}^r_k|D,\bullet) \hookrightarrow z^q_{\mathcal{W},e}(\mathbb{P}^r_k|D,\bullet)$ is a quasi-isomorphism.*

Proof. For every $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$, there is a short exact sequence of chain complexes

$$0 \to \frac{z^q_{\mathcal{W},e-1}\big(\mathbb{P}^r_k|D,\bullet\big)}{z^q_{\mathcal{W}}\big(\mathbb{P}^r_k|D,\bullet\big)} \to \frac{z^q_{\mathcal{W},e}\big(\mathbb{P}^r_k|D,\bullet\big)}{z^q_{\mathcal{W}}\big(\mathbb{P}^r|D,\bullet\big)} \to \frac{z^q_{\mathcal{W},e}\big(\mathbb{P}^r_k|D,\bullet\big)}{z^q_{\mathcal{W},e-1}\big(\mathbb{P}^r_k|D,\bullet\big)} \to 0. \quad (3.15)$$

The first two quotient complexes are acyclic by Theorem 3.10. Hence the last one must be acyclic as well. □

4. Moving lemma for projective schemes

In this section we prove the moving lemma for the higher Chow groups of projective schemes with very ample modulus. We assume for a while that the base field k is infinite. This is only a temporary assumption and will be removed in the final statement of the moving lemma (see Theorem 4.7).

We fix a closed embedding $\iota_X : X \hookrightarrow \mathbb{P}^N_k$ of an equidimensional reduced projective scheme X over k of dimension $d \geq 1$, with $d < N$. We fix two distinct hyperplanes $H_m, H_\infty \hookrightarrow \mathbb{P}^N_k$ and let $L_{m,\infty} = H_m \cap H_\infty \in \mathrm{Gr}(N-2, \mathbb{P}^N_k)$. We may assume that $X \not\subset H_m \cup H_\infty$. We set

$$X_0 = X \setminus H_\infty \overset{j_0}{\hookrightarrow} X,\ U = X \setminus H_m,\ U_0 = U \cap X_0,\ D = \iota_X^*(H_m) \text{ and } D_0 = j_0^*(D)$$

so that $X = U \cup D$ and $X_0 = U_0 \cup D_0$. We shall assume that U is smooth over k.

Remark 4.1. The hyperplane H_m could have been just called H, but we insist on using the subscript m to keep in mind that H_m later induces the modulus divisor.

Given a locally closed subset $S \subsetneq \mathbb{P}^N_k$, let $\mathrm{Gr}(S, n, \mathbb{P}^N_k)$ denote the set of n-dimensional linear subspaces of $\mathbb{P}^N_k$ which do not intersect S. Denote the set of n-dimensional linear subspaces of $\mathbb{P}^N_k$ containing a locally closed subscheme $S \subsetneq \mathbb{P}^N_k$ by $\mathrm{Gr}_S(n, \mathbb{P}^N_k)$. We let $\dim(\emptyset) = -1$. Given two locally closed subsets $Z_1, Z_2 \hookrightarrow \mathbb{P}^N_k$, let $\mathrm{Sec}(Z_1, Z_2)$ denote the union of all lines $\ell_{xy} \hookrightarrow \mathbb{P}^N_k$, joining $x \in Z_1$ and $y \in Z_2$ with $x \neq y$. One checks that $\dim(\mathrm{Sec}(Z_1, Z_2)) = \dim(Z_1) + \dim(Z_2) - \dim(Z_1 \cap Z_2)$ if Z_1 and Z_2 are linear subspaces of $\mathbb{P}^N_k$. In general, we have $\dim(\mathrm{Sec}(Z_1, Z_2)) \le \dim(Z_1) + \dim(Z_2) + 1$. Given a closed point $x \in X$, let $T_x(X)$ denote the union of lines in $\mathbb{P}^N_k$ which are tangent to X at x. For any locally closed subset $Y \subseteq X$, let $T_Y(X) = \bigcup_{x \in Y} T_x(X)$. It is clear that $\dim(T_Y(X)) \le \dim(Y) + d$ if $Y \subseteq U$. With this notation, we first prove the following:

Lemma 4.2. *Let $W \hookrightarrow \mathbb{P}^N_k$ be a closed subscheme of dimension at most d such that $W \not\subset H_m$. Then,* $\mathrm{Gr}(W, N-d-1, H_m)$ *is a dense open subset of* $\mathrm{Gr}(N-d-1, H_m)$. *If $L_{m,\infty}$ intersects W properly, then* $\mathrm{Gr}(W, N-d-1, L_{m,\infty})$ *is a dense open subset of* $\mathrm{Gr}(N-d-1, L_{m,\infty})$.

Proof. Consider the incidence variety $S = \{(x, L) \in W \times \mathrm{Gr}(N-d-1, H_m) | x \in L\}$. We have the projection maps of projective schemes

$$W \xleftarrow{\pi_1} S \xrightarrow{\pi_2} \mathrm{Gr}(N-d-1, H_m). \tag{4.1}$$

The fiber of π_1 over $W \setminus H_m$ is empty and it is a smooth fibration over $(W \cap H_m)_{\mathrm{red}}$ with each fiber isomorphic to $\mathrm{Gr}(N-d-2, \mathbb{P}^{N-2}_k)$. It follows that

$$\dim(S) = \dim(W \cap H_m) + d(N-d-1) \le d + d(N-d-1) - 1 = d(N-d) - 1.$$

Thus $\pi_2(S)$ is a closed subscheme of $\mathrm{Gr}(N-d-1, H_m)$ of dimension at most $d(N-d)-1$. On the other hand, $\dim(\mathrm{Gr}(N-d-1, H_m)) = d(N-d)$ so that $\mathrm{Gr}(W, N-d-1, H_m)$ is dense and open in $\mathrm{Gr}(N-d-1, H_m) \setminus \pi_2(S)$.

If $L_{m,\infty}$ intersects W properly, then we can argue as above with H_m replaced by $L_{m,\infty}$. We find in this case that

$$\begin{aligned}\dim(\pi_2(S)) \le \dim(S) &= \dim(W \cap L_{m,\infty}) + (d-1)(N-d-1) \\ &\le d + (d-1)(N-d-1) - 2 = (d-1)(N-d) - 1.\end{aligned}$$

Since $\dim(\mathrm{Gr}(N-d-1, L_{m,\infty})) = (d-1)(N-d)$, we get the desired conclusion. □

Given an inclusion of linear subspaces $L \subsetneq L' \subseteq \mathbb{P}^N_k$ such that $\dim(L) \le N-d-1$ and $X \cap L = \emptyset$, the linear projection away from L defines a Cartesian diagram

$$
\begin{tikzcd}
X \setminus L' \arrow[r] \arrow[d] & X \arrow[d, "\phi_L"] & X \cap L' \arrow[l] \arrow[d] \\
\mathbb{P}^d_k \setminus L' \arrow[r] & \mathbb{P}^d_k & \mathbb{P}^d_k \cap L' \arrow[l]
\end{tikzcd}
\tag{4.2}
$$

of finite maps, where $\mathbb{P}^d_k \hookrightarrow \mathbb{P}^N_k$ is a linear subspace complementary to L. Let $R_L(X) \subset X$ denote the ramification locus of ϕ_L.

For an irreducible locally closed subset $A \subsetneq X$, let $L^+(A)$ denote the closure of $\phi_L^{-1}(\phi_L(A)) \setminus A$ in $\phi_L^{-1}(\phi_L(A))$. We linearly extend this definition to all cycles on X. We shall use similar notation for locally closed subsets of $X \times \square^n$ with ϕ_L replaced by $\phi_L \times \mathrm{Id}_{\square^n}$.

For two locally closed subsets $A, C \subset X$, let

$$e(A, C) = \max\{\dim(Z) - \dim(A) - \dim(C) + d\},$$

where the maximum is taken over all irreducible components Z of $A \cap C$, if these numbers are non-negative. We take $e(A, C)$ to be zero it they are not.

Lemma 4.3. *Let $A \subsetneq X \setminus H_m$ be an irreducible locally closed subset and let $C \subsetneq X \setminus H_m$ be any locally closed subset. Let $\Sigma = \{x_1, \cdots, x_s\}$ be a set of distinct closed points of X contained in A. Then, there is a dense open subset $\mathcal{U}_X^{A,C} \hookrightarrow \mathrm{Gr}(N - d - 1, H_m)$ such that the following hold for every $L \in \mathcal{U}_X^{A,C}$:*

1. *$X \cap L = \emptyset$;*
2. *$R_L(X)$ contains no irreducible component of A, C or $A \cap C$;*
3. *$R_L(X) \cap \Sigma = \emptyset$;*
4. *$e(L^+(A) \cap C) \leq \max\{e(A, C) - 1, 0\}$;*
5. *The map $k(\phi_L(x)) \to k(x)$ is an isomorphism for $x \in \Sigma$.*

Proof. Item (1) follows from Lemma 4.2, so we prove the remaining ones. We may assume that C is irreducible. Let $L \in \mathrm{Gr}(X, N - d - 1, H_m)$. Set $T_r^L = R_L(X) \cap A \cap C = R_L(U) \cap A \cap C$ and $T_{ur}^L = (L^+(A) \cap C) \setminus T_r^L$. Note that '$r$' stands for ramified and 'ur' for unramified. Then we must have $L^+(A) \cap C \subseteq T_{ur}^L \cup T_r^L$ and hence $\dim(L^+(A) \cap C) \leq \max\{\dim(T_{ur}^L), \dim(T_r^L)\}$. Since the left square in (4.2) is Cartesian (where $L' = H_m$) and $A, C \subset U = X \setminus H_m$, it follows that the loci T_r^L and T_{ur}^L are contained in U.

Let $S \hookrightarrow ((A \times C) \setminus \Delta_X) \times \mathrm{Gr}(N - d - 1, H_m)$ be the incidence variety $S = \{(a, c, L) | \ell_{ac} \cap L \neq \emptyset\}$. We have the projections $A \times C \xleftarrow{\mathrm{pr}_1} S \xrightarrow{\mathrm{pr}_2} \mathrm{Gr}(N - d - 1, H_m)$. Since $L \cap X = \emptyset$, we see that for any point $(a, c) \in ((A \times C) \setminus \Delta_X)$, $\mathrm{pr}_1^{-1}((a, c)) = \{L \in \mathrm{Gr}(N - d - 1, H_m) | \dim(L \cap \ell_{ac}) = 0\}$. Consider the map $\pi : \mathrm{pr}_1^{-1}((a, c)) \to \ell_{ac}$ given by $\pi(L) = L \cap \ell_{ac}$.

Our hypothesis says that $(A \cup C) \cap H_m = \emptyset$ and this implies that $\ell_{ac} \not\subset H_m$. In particular, $x_{ac} = \ell_{ac} \cap H_m$ is a single closed point of $\mathbb{P}^N_k$. This implies that $\pi^{-1}(\ell_{ac} \setminus \{x_{ac}\}) = \emptyset$ and $\pi^{-1}(\{x_{ac}\}) = \mathrm{pr}_1^{-1}((a,c)) = \{L \in \mathrm{Gr}(N-d-1, H_m) | x_{ac} \in L\} \simeq \mathrm{Gr}(N-d-2, \mathbb{P}^{N-2}_k)$. It follows that $\dim(\mathrm{pr}_1^{-1}((a,c))) = (N-d-1)(N-2-(N-d-2)) = d(N-d-1)$. We conclude from this that

$$\begin{aligned} \dim(S) &\le \dim(A) + \dim(C) + d(N-d-1) \\ &= \dim(A) + \dim(C) + d(N-d) - d \\ &= \dim(A) + \dim(C) + \dim(\mathrm{Gr}(N-d-1, H_m)) - d. \end{aligned} \tag{4.3}$$

Let $p_C : S \to A \times C \to C$ be the composite projection. We now observe that $c \in T^L_{ur}$ if and only if there exists $a \in A$ such that $a \neq c$ and $\ell_{ac} \cap L \neq \emptyset$. Since $c \in C$ as well, this means that $(a,c) \in \mathrm{pr}_2^{-1}(L)$. In other words, $T^L_{ur} \subset p_C(\mathrm{pr}_2^{-1}(L))$. On the other hand, it follows from (4.3) that there is a dense open subset $\mathcal{U}^{A,C}_{ur} \subseteq \mathrm{Gr}(N-d-1, H_m)$ such that $\mathrm{pr}_2^{-1}(L)$ is either empty or has dimension $\dim(A)+\dim(C)-d$ for every $L \in \mathcal{U}^{A,C}_{ur}$. We conclude that:

($\star$) There is a dense open subset $\mathcal{U}^{A,C}_{ur} \subseteq \mathrm{Gr}(N-d-1, H_m)$ such that $\dim(T^L_{ur}) \le \dim(A) + \dim(C) - d$ for each $L \in \mathcal{U}^{A,C}_{ur}$.
Since U is smooth, given any point $x \in A \cap C$, our hypothesis implies that $T_x(X)$ is a locally closed subscheme of $\mathbb{P}^N_k$ of dimension d such that $T_x(X) \not\subset H_m$. We can therefore apply Lemma 4.2 to find a dense open subset of $\mathrm{Gr}(N-d-1, H_m)$ whose elements do not meet $T_x(X)$. But this means that $x \notin R_L(X)$ for every L in this dense open subset. We can repeat this for any chosen point in A and C as well. Since $\Sigma \subset A$, we therefore conclude that:

($\star\star$) There is a dense open subset $\mathcal{U}^{A,C}_{r} \subseteq \mathrm{Gr}(N-d-1, H_m)$ such that $R_L(X)$ does not contain any component of A, C or $A \cap C$ and it does not intersect Σ, whenever $L \in \mathcal{U}^{A,C}_{r}$.
For any $L \in \mathcal{U}^{A,C}_{r}$, we have $\dim(T^L_r) = \dim(R_L(X) \cap A \cap C) \le \max\{\dim(A \cap C) - 1, 0\}$. Combining ($\star$) and ($\star\star$) with Lemma 4.2 and setting $\mathcal{U}^{A,C}_{X} = \mathcal{U}^{A,C}_{ur} \cap \mathcal{U}^{A,C}_{r}$, we conclude that $\mathcal{U}^{A,C}_{X}$ is a dense open subset of $\mathrm{Gr}(N-d-1, H_m)$ such that $e(L^+(A) \cap C) \le \max\{e(A,C) - 1, 0\}$ for $L \in \mathcal{U}^{A,C}_{X}$.

The proof of (5) is clear if k is algebraically closed. In general, let $\overline{k}$ be an algebraic closure of k and let $\pi_Y : Y_{\overline{k}} \to Y$ denote the base change to $\overline{k}$ for any $Y \in \mathbf{Sch}_k$. For any $x \in \Sigma$, let $S_x = \pi_X^{-1}(x)$ and let $S = \bigcup_{x \in \Sigma} S_x$. Then $S \hookrightarrow X_{\overline{k}}$ is a finite set of closed points contained in $A_{\overline{k}}$. Let W' be the union of lines l_{xy} in $\mathbb{P}^N_{\overline{k}}$ such that $x \neq y \in S$. Since $S \subset A_{\overline{k}}$ and $A \cap H_m = \emptyset$, we see that $W' \not\subset H_{m,\overline{k}}$. Since $d \ge 1 = \dim(W')$, we can apply Lemma 4.2 to assume that $W' \cap L = \emptyset$ for all $L \in \mathcal{U}^{A,C}_{X_{\overline{k}}} := \mathcal{U}^{A_{\overline{k}}, C_{\overline{k}}}_{X_{\overline{k}}}$.

Since $\mathrm{Gr}(N-d-1, H_{m,\overline{k}})$ contains an affine space $\mathbb{A}^{d(N-d)}_{\overline{k}}$ as a dense open subset, we can replace $\mathcal{U}^{A,C}_{X_{\overline{k}}}$ by $\mathcal{U}^{A,C}_{X_{\overline{k}}} \cap \mathbb{A}^{d(N-d)}_{\overline{k}}$ and assume that $\mathcal{U}^{A,C}_{X_{\overline{k}}} \subseteq \mathbb{A}^{d(N-d)}_{\overline{k}}$.

Since k is infinite, the set of points in $\mathbb{A}_{\overline{k}}^{d(N-d)}$ with coordinates in k is dense in $\mathbb{A}_{\overline{k}}^{d(N-d)}$. Hence, there is a dense subset of $\mathcal{U}_{X_{\overline{k}}}^{A,C}$ each of whose points L is defined over k, *i.e.*, $L \in \mathrm{Gr}(N-d-1, H_m)$. Let $L \in \mathrm{Gr}(N-d-1, H_m)$ be such that (1) $\sim$ (4) hold and $W' \cap L_{\overline{k}} = \emptyset$. We consider the Cartesian square

$$\begin{tikzcd} X_{\overline{k}} \arrow[r, "\phi_{L_{\overline{k}}}"] \arrow[d, "\pi_X"'] & \mathbb{P}^d_{\overline{k}} \arrow[d, "\pi_{\mathbb{P}^d}"] \\ X \arrow[r, "\phi_L"] & \mathbb{P}^d_k. \end{tikzcd} \tag{4.4}$$

Claim. For a closed point $x \in U$ and $y := \phi_L(x)$, one has $|\pi_{\mathbb{P}^d}^{-1}(y)| \le |\pi_X^{-1}(x)|$, and the equality holds if and only if $[k(x) : k(y)]^{\mathrm{sep}} = 1$. Furthermore, this equality holds if the map $\phi_{L_{\overline{k}}} : \pi_X^{-1}(x) \to \pi_{\mathbb{P}^d}^{-1}(y)$ is injective.

It is an elementary fact that $|\pi_X^{-1}(x)| = [k(x):k]^{\mathrm{sep}}$ and $|\pi_{\mathbb{P}^d}^{-1}(y)| = [k(y):k]^{\mathrm{sep}}$. The inclusions $k \hookrightarrow k(y) \hookrightarrow k(x)$ and therefore the equality $[k(x):k]^{\mathrm{sep}} = [k(y):k]^{\mathrm{sep}} \cdot [k(x):k(y)]^{\mathrm{sep}}$ implies the first assertion. Next, the injectivity of the map $\phi_{L_{\overline{k}}} : \pi_X^{-1}(x) \to \pi_{\mathbb{P}^d}^{-1}(y)$ implies that $|\pi_{\mathbb{P}^d}^{-1}(y)| \ge |\pi_X^{-1}(x)|$. The second part of the Claim follows.

To prove (5) in general, it suffices to show that the finite field extension $k(\phi_L(x)) \hookrightarrow k(x)$ is separable as well as purely inseparable for each $x \in \Sigma$. Now, the separability of this extension is equivalent to the assertion $x \notin R_L(X)$, and this is guaranteed by (3). To prove inseparability, it is enough to show, using the above claim, that $\phi_{L_{\overline{k}}} : \pi_X^{-1}(x) \to \pi_{\mathbb{P}^d}^{-1}(\phi_L(x))$ is injective. But this follows immediately from the fact that $W' \cap L_{\overline{k}} = \emptyset$. The proof of the lemma is complete. $\square$

Lemma 4.4. *Let $\alpha \in z^q(X|H_m, n)$ be an admissible cycle. Let $C \subset X \setminus H_m$ be a locally closed subset as in Lemma 4.3. We can then find a dense open subset $\mathcal{U}_X^{Z,C} \subset \mathrm{Gr}(N-d-1, H_m)$ such that the following hold for every $L \in \mathcal{U}_X^{Z,C}$.*

(1) *$X \cap L = \emptyset$;*
(2) *For every irreducible component Z of α, no irreducible component of the support of the cycle $\phi_L^* \circ \phi_{L*}([Z]) - [Z]$ coincides with Z.*

Proof. It is enough to consider the case when $\alpha = [Z]$ is an irreducible admissible cycle. For any $L \in \mathrm{Gr}(N-d-1, H_m)$ satisfying (1), we need to prove the following to obtain (2):

(i) the ramification locus $R_L^n(X)$ of ϕ_L^n does not contain Z, where $\phi_L^n := \phi_L \times \mathrm{Id}_{\square_k^n}$;
(ii) $\phi_L^n|_Z : Z \to \phi_L^n(Z)$ is birational.

Let $\mathrm{pr}_X : X \times \square_k^n \to X$ and $\mathrm{pr}_{\square_k^n} : X \times \square_k^n \to \square_k^n$ be the projection maps. We fix a closed point $z \in Z$ and set $x = \mathrm{pr}_X(z)$, $y = \mathrm{pr}_{\square_k^n}(z)$, $W = \phi_L^n(Z)$ and $A = \mathrm{pr}_X(Z)$. Then A is a finite disjoint union of locally closed subsets of X. Since Z is an admissible cycle having modulus H_m, we must have $A \cap H_m = \emptyset$. In particular, $x \in U$. It is shown in the proof of Theorem 5.4 that $(\{y\} \times X) \cap Z$ is a finite set of closed points away from $(\{y\} \times H_m)$. In particular, $D := \mathrm{pr}_X((\{y\} \times X) \cap Z)$ is a finite set of closed points of X containing x and contained in A. This implies that $\mathrm{Sec}(x, D)$ is a closed subset of $\mathbb{P}_k^N$ of dimension one which is not contained in H_m. Hence, we conclude from Lemma 4.2 that $\mathrm{Gr}(\mathrm{Sec}(x, D), N-d-1, H_m)$ is dense open in $\mathrm{Gr}(N-d-1, H_m)$.

We have shown in the proof of Lemma 4.3 that there is a dense open subset $\mathcal{U}_{Z,1} \subset \mathrm{Gr}(N-d-1, H_m)$ such that $T_x(X) \cap L = \emptyset$ for each $L \in \mathcal{U}_{Z,1}$. Since the left square in (4.2) is Cartesian and ϕ_L is finite, it follows that its restriction $\phi_L^U : U \to \mathbb{P}_k^d \setminus H_m$ is also finite. Since U is furthermore smooth, it follows that ϕ_L^U is a finite and flat morphism of smooth schemes.

The flatness of ϕ_L^U now implies that there is an open neighborhood $V \subset U$ of x such that $\phi_L : V \to \mathbb{P}_k^d$ is étale. In particular, $\phi_L^n : V \times \square_k^n \to \mathbb{P}_k^d \times \square_k^n$ is étale. This implies that there is an open subset V' of Z containing z such that $\phi_L^n|_{V'} : V' \to W$ is unramified. We set $\mathcal{U}_X^{Z,C} = \mathrm{Gr}(\mathrm{Sec}(x, D), N-d-1, H_m) \cap \mathcal{U}_{Z,1} \cap \mathcal{U}_X^{A,C}$, where $\mathcal{U}_X^{A,C}$ is as in Lemma 4.3.

We fix any $L \in \mathcal{U}_X^{A,C}$. Since $R_L^n(X) = R_L(X) \times \square_k^n$ and no component of A is in $R_L(X)$ by Lemma 4.3, it follows that $Z \not\subset R_L^n(X)$, proving (i). To prove (ii), it suffices to show that $z \notin R_L^n(Z)$, $\{z\} = (\phi_L^n)^{-1}(\phi_L^n(z)) \cap Z$ and $k(\phi_L^n(z)) \xrightarrow{\simeq} k(z)$, because they imply that the map $\mathcal{O}_{W,\phi_L^n(z)} \to \mathcal{O}_{Z,z}$ is an isomorphism, and hence induces isomorphism of the function fields.

We have shown above that $z \notin R_L^n(Z)$. Since the map $k(\phi_L(x)) \to k(x)$ is an isomorphism by Lemma 4.3, it follows that the map $\phi_L^n : \square_{k(x)}^n \to \square_{k(\phi_L(x))}^n$ is also an isomorphism. In particular, the map $k(\phi_L^n(z)) \to k(z)$ is an isomorphism. To show $\{z\} = (\phi_L^n)^{-1}(\phi_L^n(z)) \cap Z$, note that if there is a closed point $z' \in ((\phi_L^n)^{-1}(\phi_L^n(z)) \cap Z) \setminus \{z\}$, then $x' := \mathrm{pr}_X(z') \in D \cap L^+(x)$, where we recall that $L^+(x) = \phi_L^{-1}(\phi_L(x)) \setminus \{x\}$. But this can happen only if $\ell_{xx'} \cap L \neq \emptyset$, which is not the case because $L \in \mathrm{Gr}(\mathrm{Sec}(x, D), N-d-1, H_m)$. This concludes the proof of (ii) and the lemma. □

Remark 4.5. We a make few comments on Lemma 4.3. To some readers, this result may appear similar to [13, Lemma 3.5.4]. But we caution the reader that the context, the underlying hypotheses and the proofs of the two results are different. We explain these differences:

(1) The proof of Lemma 4.3 does not work if we replace X by $X \cap \mathbb{A}_k^N$. The reason is that even if X intersects $L_{m,\infty}$ properly, we may not be able to find points on $A \cap C$ whose tangent spaces will intersect $L_{m,\infty}$ properly, and this breaks the second part of the proof of Lemma 4.3.

Since [13] considers the affine case, Levine cannot therefore use the above argument. Instead, he uses the idea of reimbedding X into a big enough projective space which allows him to take care of the above intersection problem associated to the tangent spaces;

(2) Contrary to [13], we cannot use the reimbedding idea. The reason is that we may not be able to realize our modulus H_m as pull-back of any hypersurface on the bigger projective space under the reimbedding. This in turn may not allow us to realize H_m as pull-back of a hypersurface under a linear projection;

(3) The modulus condition imposes more severe restrictions on the choice of L than in the situation of [13]. Thus we need to make more refined choices and without changing the given embedding of X.

Let $\mathcal{W} = \{W_1, \cdots, W_s\}$ be a finite collection of locally closed subsets of $X \setminus H_m$ and let $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$ be a set function. Let K denote the function field of $\mathrm{Gr}(N-d-1, H_m)$ and let $L_{\mathrm{gen}} \in \mathrm{Gr}(N-d-1, H_m)(K)$ be the generic point of $\mathrm{Gr}(N-d-1, H_m)$. This can be seen as a K-rational point of $\mathrm{Gr}(N-d-1, H_m)$.

Lemma 4.6. *The linear projection away from L_{gen} defines a finite map $\phi_{L_{\mathrm{gen}}} : X_K \to \mathbb{P}^d_K$ satisfying the following conditions:*

(1) *The restriction $\phi^U_{L_{\mathrm{gen}}} : U_K \to \mathbb{P}^d_K \setminus H_{m,K}$ is finite and flat;*

(2) *$D_K = \phi^*_{L_{\mathrm{gen}}}(H_{\mathrm{gen}})$ for the hyperplane $H_{\mathrm{gen}} = (H_m \cap \mathbb{P}^d)_K$ in $\mathbb{P}^d_K$;*

(3) *The pull-back $\phi^*_{L_{\mathrm{gen}}} : z^q(\mathbb{P}^d_K | H_{\mathrm{gen}}, \bullet) \to z^q(X_K | D_K, \bullet)$ is defined;*

(4) *$(\phi^*_{L_{\mathrm{gen}}} \circ \phi_{L_{\mathrm{gen}}*} \circ \mathrm{pr}^*_{K/k} - \mathrm{pr}^*_{K/k})$ maps $z^q_{\mathcal{W},e}(X|D, \bullet)$ to $z^q_{\mathcal{W}_K, e-1}(X_K|D_K, \bullet)$.*

Proof. Having established Lemmas 4.3 and 4.4, the proof of this lemma is identical to that of [13, Lemma 3.5.6]. The modulus condition plays no role in this deduction. Using Lemmas 4.3 and 4.4 and the argument of [13] *verbatim*, one shows that given a cycle $\alpha \in z^q_{\mathcal{W},e}(X|D, p)$, there exists a dense open subset $\mathcal{U}^\alpha_X \subseteq \mathrm{Gr}(N-d-1, H_m)$ such that for each $L \in \mathcal{U}^\alpha_X$, the linear projection away from L defines a finite map $\phi_L : X \to \mathbb{P}^d_k$ satisfying the required conditions. This map is flat on U as shown in the proof of Lemma 4.4. Taking $L = L_{\mathrm{gen}}$ and using Lemma 2.5, we get (1), (3) and (4). The map $\phi_{L_{\mathrm{gen}}*}$ is defined by [11, Proposition 2.10].

Item (2) follows at once from our choice of L_{gen} and the elementary property of linear projection that a hyperplane section $X \cap H$ in $\mathbb{P}^N_k$ is a pull-back of a hyperplane of $\mathbb{P}^d_k$ via ϕ_L if and only if $L \subset H$. □

We are now ready to prove our main theorem on the moving lemma for the higher Chow groups of projective schemes with very ample modulus.

Theorem 4.7. *Let k be any field and let X be an equidimensional reduced projective scheme of dimension $d \geq 1$ over k. Let $D \subset X$ be a very ample effective Cartier divisor such that $X \setminus D$ is smooth over k. Let $\mathcal{W} = \{W_1, \cdots, W_s\}$ be a finite collection of locally closed subsets of X and let $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$ be a set function.*

Then, the inclusion $z^q_{\mathcal{W},e-1}(X|D,\bullet) \hookrightarrow z^q_{\mathcal{W},e}(X|D,\bullet)$ *is a quasi-isomorphism. In particular, the inclusion* $z^q_{\mathcal{W}}(X|D,\bullet) \hookrightarrow z^q(X|D,\bullet)$ *is a quasi-isomorphism.*

Proof. The second part easily follows from the first part by induction because $z^q_{\mathcal{W}}(X|D,\bullet) = z^q_{\mathcal{W},0}(X|D,\bullet)$ and $z^q(X|D,\bullet) = z^q_{\mathcal{W},q}(X|D,\bullet)$. We thus need to show that the quotient complex $\frac{z^q_{\mathcal{W},e}(X|D,\bullet)}{z^q_{\mathcal{W},e-1}(X|D,\bullet)}$ is acyclic.

First suppose that the theorem is true for all infinite fields and let k be a finite field. Take a homology class α in this quotient. We choose two distinct primes ℓ_1 and ℓ_2, other than $\mathrm{char}(k)$, and take pro-ℓ_i-extensions $\iota_i : \mathrm{Spec}\,(k_i) \to \mathrm{Spec}\,(k)$ for $i = 1, 2$. Then the case of infinite fields tells us that $\iota_i^*(\alpha) = 0$ for $i = 1, 2$. Hence, a descent argument implies that there are finite extensions $\tau_i : \mathrm{Spec}\,(k_i') \to \mathrm{Spec}\,(k)$ of relatively prime degrees such that $\tau_i^*(\alpha) = 0$ for $i = 1, 2$. Using the projection formula for finite and flat morphisms (see [11, Theorem 3.12]), this implies that $d_1\alpha = 0 = d_2\alpha$, where $(d_1, d_2) = 1$. We conclude that $\alpha = 0$.

We can now assume that k is infinite. We set $\mathcal{W}^0 = \{W_1 \setminus D, \cdots, W_s \setminus D\}$. Since a cycle in $z^q(X|D,p)$ does not intersect $D \times \square^p$, we see that $z^q_{\mathcal{W}}(X|D,\bullet) = z^q_{\mathcal{W}^0}(X|D,\bullet)$, and we may assume that $W \cap D = \emptyset$ for each $W \in \mathcal{W}$.

Since D is very ample, we can choose a closed embedding $\iota_X : X \hookrightarrow \mathbb{P}^N_k$ and a hyperplane $H_m \subset \mathbb{P}^n_k$ such that $D = \iota^*(H_m)$. If $X = \mathbb{P}^N_k$, we are done by Theorem 3.10. So we can assume that $1 \le d \le N-1$.

It follows from Lemma 4.6 that the map

$$\left(\phi^*_{L_{\mathrm{gen}}} \circ \phi_{L_{\mathrm{gen}}*} \circ \mathrm{pr}^*_{K/k} - \mathrm{pr}^*_{K/k}\right) : \frac{z^q_{\mathcal{W},e}(X|D,\bullet)}{z^q_{\mathcal{W},e-1}(X|D,\bullet)} \to \frac{z^q_{\mathcal{W}_K,e}(X_K|D_K,\bullet)}{z^q_{\mathcal{W}_K,e-1}(X_K|D_K,\bullet)} \tag{4.5}$$

is zero. On the other hand, each $\phi^*_{L_{\mathrm{gen}}} \circ \phi_{L_{\mathrm{gen}}*}$ factors as

$$\frac{z^q_{\mathcal{W}_K,e}(X_K|D_K,\bullet)}{z^q_{\mathcal{W}_K,e-1}(X_K|D_K,\bullet)} \xrightarrow{\phi_{L_{\mathrm{gen}}*}} \frac{z^q_{\phi_{L_{\mathrm{gen}}}(\mathcal{W}_K),e'}\left(\mathbb{P}^d_K|H_{\mathrm{gen}},\bullet\right)}{z^q_{\phi_{L_{\mathrm{gen}}}(\mathcal{W}_K),e'-1}\left(\mathbb{P}^d_K|H_{\mathrm{gen}},\bullet\right)} \xrightarrow{\phi^*_{L_{\mathrm{gen}}}} \frac{z^q_{\mathcal{W}_K,e}(X_K|D_K,\bullet)}{z^q_{\mathcal{W}_K,e-1}(X_K|D_K,\bullet)}$$

for some e' (see [10, Section 6C]). It follows from Corollary 3.11 that the middle complex is acyclic. This in turn implies that $\phi^*_{L_{\mathrm{gen}}} \circ \phi_{L_{\mathrm{gen}}*} = 0$ is zero on the level of homology. Combining this with (4.5), we conclude that $\mathrm{pr}^*_{K/k}$ is zero on the level of homology. By Proposition 2.8, the complex $\frac{z^q_{\mathcal{W},e}(X|D,\bullet)}{z^q_{\mathcal{W},e-1}(X|D,\bullet)}$ is acyclic. This concludes the proof of the theorem. $\square$

5. Applications and remarks

In this section we apply our moving lemma to prove certain contravariant functoriality for higher Chow groups with modulus. We prove a vanishing theorem on higher Chow groups with ample modulus. We end the section by explaining why the very ampleness condition is crucial for proving the moving lemma.

5.1. Contravariance

Let X be a quasi-projective scheme over a field k and let $D \subset X$ be a very ample effective Cartier divisor. Recall from [11, Theorem 3.12] if that X is smooth, there is a cap product $\cap_X : \mathrm{CH}^q(X, p) \otimes_{\mathbb{Z}} \mathrm{CH}^{q'}(X|D, p') \to \mathrm{CH}^{q+q'}(X|D, p+p')$. We prove the following contravariant functoriality for cycles with modulus.

Theorem 5.1. *Let $f : Y \to X$ be a morphism of quasi-projective schemes over a field k, where X is projective over k. Let $D \subset X$ be a very ample effective Cartier divisor such that $X \setminus D$ is smooth over k. Suppose that $f^*(D)$ is a Cartier divisor on Y* (i.e., *no minimal or embedded component of Y maps into D*). *Then there exists a map*

$$f^* : z^q(X|D, \bullet) \to z^q\big(Y|f^*(D), \bullet\big)$$

in the derived category of Abelian groups. In particular, there is a pull-back f^* : $\mathrm{CH}^q(X|D, p) \to \mathrm{CH}^q(Y|f^*(D), p)$ for every $p, q \geq 0$. If X and Y are smooth and projective, then for every $a \in \mathrm{CH}^*(Y, \bullet)$ and $b \in \mathrm{CH}^*(X|D, \bullet)$, there is a projection formula $f_*(a \cap_Y f^*(b)) = f_*(a) \cap_X b$.

Proof. The proof is a standard application of the moving lemma for Chow groups. Set $E = f^*(D)$. For $0 \leq i \leq \dim(Y)$, let X_i be the set of points $x \in X$ such that $\dim(f^{-1}(x)) \geq i$, where we assume $\dim(\emptyset) = -1$. Let $\mathcal{W}$ be the collection of the irreducible components of all X_i. The reader can check that $\mathcal{W}$ is a finite collection and the pull-back $f^* : z^q_{\mathcal{W}}(X|D, \bullet) \to z^q(Y|E, \bullet)$ is defined (see [10, Theorem 7.1]). We thus have maps $z^q(X|D, \bullet) \overset{q.iso}{\leftarrow} z^q_{\mathcal{W}}(X|D, \bullet) \overset{f^*}{\to} z^q(Y|E, \bullet)$ and Theorem 4.7 says that the arrow on the left is a quasi-isomorphism. This proves the first part of the theorem.

To prove the projection formula, we can assume using Theorem 4.7 that $b \in \mathrm{CH}^*(X|D, \bullet)$ is represented by a cycle $Z \in z^q_{\mathcal{W}}(X|D, \bullet)$, where $\mathcal{W}$ is as constructed above. By [11, Lemma 3.10], there is a finite collection of locally closed subsets $\mathcal{C}$ of Y such that $Z' \boxtimes f^*(Z) \in z^q_{\Delta_Y}(Y|E, \bullet)$ for all $Z' \in z^q_{\mathcal{C}}(Y, \bullet)$. By the moving lemma for Bloch's higher Chow groups, we can assume that $a \in \mathrm{CH}^*(Y, \bullet)$ is represented by a cycle $Z' \in z^q_{\mathcal{C}}(Y, \bullet)$. In this case, it is straightforward to check that $f_*(Z') \boxtimes Z \in z^q_{\Delta_X}(X|D, \bullet)$ and $f_* \circ \Delta^*_Y(Z' \boxtimes f^*(Z)) = \Delta^*_X(f_*(Z') \boxtimes Z)$. This finishes the proof. $\square$

Remark 5.2. We remark that a pull-back map on higher Chow groups with modulus was constructed in [11, Theorem 4.3]. But Theorem 5.1 cannot be deduced from [11, Theorem 4.3]. The reason is that we make no assumption on the map f while [11] assumes D and E to be the pull-backs of a divisor on a base scheme S over which both X and Y should be smooth.

We also remark that Theorem 5.1 proves a stronger statement than giving a pull-back map on the higher Chow groups with modulus. This stronger version of [11, Theorem 4.3] is not yet known.

Corollary 5.3. *Let $r \geq 1$ be an integer and let $f : Y \to \mathbb{P}^r_k$ be a morphism of quasi-projective schemes over a field k. Let $D \subset \mathbb{P}^r_k$ be an effective Cartier divisor such that $f^*(D)$ is a Cartier divisor on Y. Then, there exists a pull-back $f^* : \mathrm{CH}^q(\mathbb{P}^r_k|D, p) \to \mathrm{CH}^q(Y|f^*(D), p)$ for every $p, q \geq 0$. If Y is also smooth and projective, then for every $a \in \mathrm{CH}^*(Y, \bullet)$ and $b \in \mathrm{CH}^*(\mathbb{P}^r_k|D, \bullet)$, there is a projection formula $f_*(a \cap_Y f^*(b)) = f_*(a) \cap_X b$.*

Proof. If $D = 0$, then it is just an application of the moving lemma for usual higher Chow groups. If $D \neq 0$ then it is very ample, so that Theorem 5.1 applies. $\square$

5.2. A vanishing theorem

The following result shows that the higher Chow groups of projective schemes (not necessarily smooth) with ample modulus are nontrivial only in high codimension. More precisely:

Theorem 5.4. *Let X be a projective scheme of dimension $d \geq 1$ over a field k. Let $D \subset X$ be an ample effective Cartier divisor. Then $z_s(X|D, p) = 0$ for $s > 0$. In particular, $\mathrm{CH}_s(X|D, p) = 0$ for $s > 0$.*

Proof. We can find a closed embedding $\iota_X : X \hookrightarrow \mathbb{P}^N_k$ and a hyperplane $H \hookrightarrow \mathbb{P}^N_k$ such that $nD = \iota_X^*(H)$ for some $n \gg 0$. Suppose $z_s(X|D, p) \neq 0$ for some $s \in \mathbb{Z}$. Let $\alpha \in z_s(X|D, p)$ be a nonzero admissible cycle and let Z be an irreducible component of α. Let $\mathrm{pr}_{\mathbb{P}^N_k} : \mathbb{P}^N_k \times \square^p_k \to \mathbb{P}^N_k$ and $\mathrm{pr}_{\square^p_k} : \mathbb{P}^N_k \times \square^p_k \to \square^p_k$ denote the projection maps. Let $y \in \square^p_k$ be any scheme point. For any map $W \to \square^p_k$, let W_y denote the fiber $\mathrm{Spec}\,(k(y)) \times_{\square^p_k} W$ over y. The modulus condition for Z implies that Z_y is a closed subscheme of $\mathbb{P}^N_y$ disjoint from H_y. In particular, Z_y is a projective $k(y)$-scheme which is a closed subscheme of $(\mathbb{P}^N_y \setminus H_y) \simeq \mathbb{A}^N_{k(y)}$. Hence, it must be finite. We have thus shown that the projection map $Z \to \square^p_k$ is projective and quasi-finite, and hence finite. In other words, we must have $\dim(Z) = s + p \leq p$, *i.e.*, $s \leq 0$. Thus $z_s(X|D, p) = 0$ if $s > 0$, as desired. $\square$

5.3. Sharpness of the very ampleness condition

We now show by an example that we cannot weaken the very ampleness condition to mere ampleness for the modulus divisor $D \subset X$. This also shows that the moving lemma for cycles with modulus on smooth affine schemes cannot be proven using the method of linear projections, in general. This partly explains the need for the Nisnevich sheafification of the cycle complex for the moving lemma of W. Kai [7].

Let X be an elliptic curve over an algebraically closed field k and let $D \subset X$ be a closed point. It is clear that $\mathcal{O}_X(D)$ is ample. We claim that there exists no pair (f, D') consisting of a map $f : X \to \mathbb{P}^1_k$ and an effective Cartier divisor $D' \in \mathrm{Div}(\mathbb{P}^1_k)$ such that $D = f^*(D')$.

Suppose there does exist such a pair (f, D'). Observe that we must have $d := \deg(D') > 0$ and D' is very ample. Let $\iota : \mathbb{P}^1_k \hookrightarrow \mathbb{P}^d_k$ denote the closed embedding such that $\mathcal{O}_{\mathbb{P}^1_k}(D') \simeq \iota^*(\mathcal{O}_{\mathbb{P}^d_k}(1))$. This gives a regular map $\iota \circ f : X \to \mathbb{P}^d_k$ such that $(\iota \circ f)^*(\mathcal{O}_{\mathbb{P}^d_k}(1)) = \mathcal{O}_X(D)$. This implies that $\mathcal{O}_X(D)$ is globally generated. However, by Riemann-Roch, one checks immediately that $h^0(D) = 1$ in our case, *i.e.*, $\dim(|D|) = 0$ and the unique element of $|D|$ vanishes at D, a contradiction.

The only technique yet available in the literature to prove the moving lemma for Bloch's higher Chow groups of smooth affine schemes is the method of *linear projections*. Bloch proved the moving lemma for higher Chow groups of all smooth quasi-projective schemes (see [3] and [4, Proposition 2.5.2]). But his proof depends on the moving lemma for smooth affine schemes proven in [2] using linear projections.

Let us now consider the case of moving lemma for higher Chow groups with modulus on smooth affine schemes. Let U be a smooth affine scheme over an algebraically closed field k of characteristic zero. Let $D \subset U$ be a principal effective divisor (u) such that the induced map $u : U \setminus D \to \mathbb{A}^1_k$ is smooth. We use the above example to show that even in this special case, the method of linear projections cannot be used to prove the moving lemma for the higher Chow groups on U with modulus D. This makes proving the moving lemma for cycles with modulus on smooth affine or projective schemes very subtle and challenging.

Let X be an elliptic curve over k as above and let $D \hookrightarrow X$ be a closed point. There exists an affine neighborhood $V \hookrightarrow X$ of D such that $D = (u)$ is principal on V. Let $u : V \to \mathbb{A}^1_k$ be the induced dominant map. We can find an affine neighborhood $U \hookrightarrow V$ of D such that $u : U \setminus D \to \mathbb{A}^1_k$ is étale.

Proposition 5.5. *There exists no pair (f, D') consisting of a finite map $f : U \to \mathbb{A}^1_k$ and effective Cartier divisor $D' \hookrightarrow \mathbb{A}^1_k$ such that $D = f^*(D')$.*

Proof. If such pair (f, D') exists, then we get a commutative diagram

$$\begin{tikzcd} U \arrow[r, "j'"] \arrow[d, "f"'] & X \arrow[d, "f'"] \\ \mathbb{A}^1_k \arrow[r, "j"] & \mathbb{P}^1_k, \end{tikzcd} \tag{5.1}$$

where the horizontal maps are open inclusions and the vertical maps are finite. This finiteness implies that the above square is Cartesian. This in turn implies that we have a finite map $f' : X' \to \mathbb{P}^1_k$ and effective Cartier divisor $D' \hookrightarrow \mathbb{P}^1_k$ such that $D = f'^*(D')$ on X. But we have previously shown that this is not possible. ☐

6. Higher Chow groups with modulus of a line bundle

Let X be a quasi-projective scheme of dimension $d \geq 0$ over a field k. Let $f : \mathcal{L} \to X$ be a line bundle and let $\iota : X \hookrightarrow \mathcal{L}$ be the 0-section embedding. In this case, one knows that there is an isomorphism $\iota^* : \mathrm{CH}_*(\mathcal{L}, \bullet) \xrightarrow{\simeq} \mathrm{CH}_*(X, \bullet)$ (up to a shift in dimension) of ordinary higher groups. Since the Chow groups with modulus are supposed to be the 'relative motivic cohomology' of the pair $(\mathcal{L}, \iota(X))$, one expects $\mathrm{CH}_*(\mathcal{L}|X, \bullet)$ to be trivial.

As an application of the moving techniques of Section 3, we show in this section that every cycle in $z_s(\mathcal{L}|X, \bullet)$ can be moved to a trivial cycle so that this complex is acyclic. This gives an evidence supporting the expectation that the Chow groups with modulus are the relative motivic cohomology. It also provides examples where the higher Chow groups of a variety with a modulus in an effective Cartier divisor are all zero. Note that this can never happen for the ordinary higher groups. The proof closely follows the arguments of Lemmas 3.5 and 3.8, and Proposition 3.9.

Let $H : \mathcal{L} \times \mathbb{A}^1_k \to \mathcal{L}$ be the standard fiberwise contraction given explicitly as follows: for an affine open subset $U = \mathrm{Spec}\,(R) \subset X$ such that $f|_U$ is trivial, *i.e.*, of the form $f|_U : U \times \mathbb{A}^1_k \to U$, write $\mathcal{L}|_U = \mathrm{Spec}\,(R[t])$. Then, $H|_U : U \times \mathbb{A}^1_k \times \mathbb{A}^1_k \to U \times \mathbb{A}^1_k$ is induced by the polynomial map $R[x] \to R[t, x]$, given by $x \mapsto tx$.

For $n \geq 0$, let $H_n : \mathcal{L} \times \mathbb{A}^1_k \times \overline{\square}^n_k \to \mathcal{L} \times \overline{\square}^n_k$ be the map $H \times \mathrm{Id}_{\overline{\square}^n_k}$. For any irreducible closed admissible cycle $V \in z_s(\mathcal{L}|X, n)$, let $H^*(V)$ denote the cycle associated to the flat pull-back $H_n^{-1}(V)$. Set $V' = (H^*(V))_{\mathrm{red}}$. We extend H^* linearly to all cycles. Let $\overline{V} \hookrightarrow \mathcal{L} \times \overline{\square}^n_k$ denote the closure of V and let $\nu_V : \overline{V}^N \to \mathcal{L} \times \overline{\square}^n_k$ be the composition of the normalization and the inclusion. Let $\overline{V}'$ denote the closure of V' in $\mathcal{L} \times \overline{\square}^{n+1}_k$ and let $\nu_{V'} : \overline{V}'^N \to \mathcal{L} \times \overline{\square}^{n+1}_k$ denote the map induced by the normalization of $\overline{V}'$.

Lemma 6.1. *$V' \hookrightarrow \mathcal{L} \times \square^{n+1}_k$ has modulus X.*

Proof. Since the modulus condition is local on $\mathcal{L}$, it is enough to show that $V' \cap (f^{-1}(U) \times \square^{n+1}_k)$ has modulus U for every affine open subset $U \subset X$ over which f is trivial. So we may assume $X = \mathrm{Spec}\,(R)$ is affine and $\mathcal{L} = \mathrm{Spec}\,(R[X])$ is trivial. In this case, $H : U \times \mathbb{A}^1_k \times \mathbb{A}^1_k \to U \times \mathbb{A}^1_k$ is given by $H(u, x, y) = (u, xy)$. Since U plays no role in this map, we can drop it and assume $U = \mathrm{Spec}\,(k)$ so that $H : \mathbb{A}^1_k \times \mathbb{A}^1_k \to \mathbb{A}^1_k$ is the multiplication map. This map uniquely extends to a rational map $H : \mathbb{P}^1_k \times \mathbb{P}^1_k \dashrightarrow \mathbb{P}^1_k$, given by $H((X_0; X_1), (T_0; T_1)) = (X_0T_0; X_1T_1)$, which is regular on $W = (\mathbb{P}^1_k \times \mathbb{P}^1_k) \setminus \{(0, \infty), (\infty, 0)\}$.

We next observe that since the modulus divisor is $U = \{0\} \hookrightarrow \mathbb{A}^1_k$, to check the modulus condition for $H^{-1}(V)$ is equivalent to check the modulus $(\{0\} \times \mathbb{A}^1_k)$ for $(H|_{W \times \square^n_k})^{-1}(V_1)$, where V_1 is the closure of V in $\mathbb{P}^1_k \times \square^n_k$. We can thus replace $\mathbb{A}^1_k$ by $\mathbb{P}^1_k$ as the target space of H and $\overline{V}'$ by its closure in $\mathbb{P}^1_k \times \overline{\square}^{n+1}_k$ in order to check the modulus condition for V'.

Let $\pi : \Gamma \to \mathbb{P}^1_k \times \mathbb{P}^1_k$ be the blow-up along $\Sigma = \{(0, \infty), (\infty, 0)\}$. It is easily checked (see the proof of Lemma 3.5) that $\Gamma \hookrightarrow \mathbb{P}^1_k \times \mathbb{P}^1_k \times \mathbb{P}^1_k$ is the closed subscheme given by $\Gamma = \{((X_0; X_1), (T_0; T_1), (Y_1; Y_0)) \,|\, X_0T_0Y_0 = X_1T_1Y_1\}$. Define a map $\overline{H} : \Gamma \to \mathbb{P}^1_k$ by $\overline{H}\,((X_0; X_1), (T_0; T_1), (Y_1; Y_0)) = (Y_1; Y_0)$.

We claim that $\overline{H}|_W = H$. To check this, let $U_1 = \{((X_0; X_1), (T_0; T_1)) \,|\, X_1 \neq 0 \neq T_0\}$ and $U_2 = \{((X_0; X_1), (T_0; T_1)) \,|\, X_0 \neq 0 \neq T_1\}$ be two open subsets of $\mathbb{P}^1_k \times \mathbb{P}^1_k$. In the affine coordinates $(x_0, t_1) \in U_1 \simeq \mathbb{A}^2_k$, the restriction of H on $U_1 \cap W$ is given by $H(x_0, t_1) = (x_0; t_1)$ and the restriction of $\overline{H}$ on $\pi^{-1}(U_1) \cap W \cap (x_0 \neq 0)$ is given by $\overline{H}\big((x_0, t_1, (1; x_0^{-1}t_1)\big) = (1; x_0^{-1}t_1) = (x_0; t_1) = H(x_0, t_1)$. The restriction of $\overline{H}$ on $\pi^{-1}(U_1) \cap W \cap (t_1 \neq 0)$ is given by $\overline{H}\big((x_0, t_1, (x_0t_1^{-1}; 1)\big) = (x_0t_1^{-1}; 1) = (x_0; t_1) = H(x_0, t_1)$.

The restriction of H on $U_2 \cap W$ is given by $H(x_1, t_0) = (t_0; x_1)$ and the restriction of $\overline{H}$ on $\pi^{-1}(U_2) \cap W \cap (x_1 \neq 0)$ is given by $\overline{H}\big((x_1, t_0, (x_1^{-1}t_0; 1)\big) = (x_1^{-1}t_0; 1) = (t_0; x_1) = H(x_1, t_0)$. The restriction of $\overline{H}$ on $\pi^{-1}(U_1) \cap W \cap (t_0 \neq 0)$ is given by $\overline{H}\big((x_1, t_0, (1; x_1t_0^{-1})\big) = (1; x_1t_0^{-1}) = (t_0; x_1) = H(x_1, t_0)$. Since π is an isomorphism away from $U_1 \cup U_2$, we have shown that $\overline{H}|_W = H$.

It follows from the claim that there is a commutative diagram

$$\begin{tikzcd}
\pi^{-1}(W) \arrow[r, hook, "j_1"] \arrow[d, "\simeq"'] & \Gamma \arrow[d, "\pi"'] \arrow[dr, "\overline{H}"] & \\
W \arrow[r, hook, "j"] \arrow[rr, bend right, ""] & \mathbb{P}^1_k \times \mathbb{P}^1_k \arrow[r, dotted, "H"] & \mathbb{P}^1_k.
\end{tikzcd} \tag{6.1}$$

Let $E = \pi^*((0, \infty))$ denote one of the two components of the exceptional divisor for π and let $D = U = \{0\} \hookrightarrow \mathbb{P}^1_k$. We have $\pi^*(D \times \mathbb{P}^1_k) = (D \times \mathbb{P}^1_k) + E$. Similarly, we have $\pi^*(\mathbb{P}^1_k \times \{\infty\}) = (\mathbb{P}^1_k \times \{\infty\}) + E$ in $\mathrm{Div}(\Gamma)$. Set $E_n = E \times \overline{\square}^n_k$.

Let $Z \hookrightarrow \Gamma \times \overline{\square}^n_k$ denote the strict transform of $\overline{V}'$. Since $\overline{H}_n(Z \cap (\pi^{-1}(W) \times \square^n_k)) = V$ and since $\overline{H}_n$ is projective, we must have $\overline{H}_n(Z) = \overline{V}$. We remark at this stage that ensuring the projectivity of $\overline{H}_n$ was the reason for us to replace $\mathbb{A}^1_k \times \mathbb{A}^1_k$ by $\mathbb{P}^1_k \times \mathbb{P}^1_k$ and $\mathbb{A}^1_k$ by $\mathbb{P}^1_k$ as the source and the target of H.

We now have a commutative diagram

$$\begin{tikzcd}
Z^N \arrow[rr, "f"] \arrow[dr, "\nu_Z"] \arrow[dd, "g"'] & & \overline{V}^N \arrow[d, "\nu_V"] \\
& \Gamma \times \overline{\square}^n_k \arrow[r, "\overline{H}_n"] \arrow[d, "\pi_n"] & \mathbb{P}^1_k \times \overline{\square}^n_k \\
\overline{V}'^N \arrow[r, "\nu_{V'}"'] & \mathbb{P}^1_k \times \overline{\square}^{n+1}_k, &
\end{tikzcd} \tag{6.2}$$

where f and g are the unique maps induced by the universal property of normalization for dominant maps. Since f is a surjective map of integral schemes, the modulus condition for V implies that $(\nu_V \circ f)^*(\mathbb{P}^1_k \times F_n^\infty) \geq (\nu_V \circ f)^*(D \times \overline{\square}^n_k)$ on Z^N. In particular, we get $(\overline{H}_n \circ \nu_Z)^*(\mathbb{P}^1_k \times F_n^\infty) \geq (\overline{H}_n \circ \nu_Z)^*(D \times \overline{\square}^n_k)$ on Z^N. Equivalently, we have

$$\nu_Z^*\big(\Gamma \times F_n^\infty\big) \geq \nu_Z^* \left(\overline{H}^*(D) \times \overline{\square}^n_K\right). \tag{6.3}$$

Since $H^*(D) = (\mathbb{A}^1_k \times \{0\}) + (\{0\} \times \square_k)$, we get $j^*_{1,n} \circ \overline{H}^*_n(D \times \overline{\square}^n_k) = j^*_{1,n}(\mathbb{A}^1_k \times F^0_{n,n+1}) + j^*_{1,n}(D \times \overline{\square}^{n+1}_k)$, where $j_1 : W \hookrightarrow \Gamma$ is the inclusion. Since $\mathbb{A}^1_k \times F^0_{n,n+1}$ and $D \times \overline{\square}^{n+1}_k$ are irreducible, we get $\overline{H}^*(D) \times \overline{\square}^n_k \geq (\mathbb{P}^1_k \times F^0_{n,n+1}) + (D \times \overline{\square}^{n+1}_k)$ on $\Gamma \times \overline{\square}^n_k$. Combining this with (6.3), we get

$$\nu_Z^*\big(\Gamma \times F_n^\infty\big) \geq \nu_Z^* \left(D \times \overline{\square}^{n+1}_k\right). \tag{6.4}$$

This in turn implies that

$$\begin{aligned}
(\pi_n \circ \nu_Z)^* \left(\mathbb{P}^1_k \times F^\infty_{n+1}\right) &= (\pi_n \circ \nu_Z)^* \left(\mathbb{P}^1_k \times F^\infty_n \times \overline{\square}_k\right) \\
&\quad + (\pi_n \circ \nu_Z)^* \left(\mathbb{P}^1_k \times \overline{\square}^n_k \times \{\infty\}\right) \\
&= \nu_Z^*\big(\Gamma \times F^\infty_n\big) + (\pi_n \circ \nu_Z)^* \left(\mathbb{P}^1_k \times \overline{\square}^n_k \times \{\infty\}\right) \\
&\geq \nu_Z^* \left(D \times \overline{\square}^{n+1}_k\right) + (\pi_n \circ \nu_Z)^* \left(\mathbb{P}^1_k \times \overline{\square}^n_k \times \{\infty\}\right) \\
&= \nu_Z^* \left(D \times \overline{\square}^{n+1}_k\right) + \nu_Z^*(E_n) + \nu_Z^* \left(\mathbb{P}^1_k \times \overline{\square}^n_k \times \{\infty\}\right) \\
&= (\pi_n \circ \nu_Z)^* \left(D \times \overline{\square}^{n+1}_k\right) + \nu_Z^* \left(\mathbb{P}^1_k \times \overline{\square}^n_k \times \{\infty\}\right) \\
&\geq (\pi_n \circ \nu_Z)^* \left(D \times \overline{\square}^{n+1}_k\right).
\end{aligned}$$

Using (6.2), this gives $g^*(\nu^*_{V'}(\mathbb{P}^1_k \times F^\infty_{n+1})) \geq g^*(\nu^*_{V'}(D \times \overline{\square}^{n+1}_k))$. We now apply Lemma 2.6 to conclude that $\nu^*_{V'}(\mathbb{P}^1_k \times F^\infty_{n+1}) \geq \nu^*_{V'}(D \times \overline{\square}^{n+1}_k)$ and this is the modulus condition for V'. $\square$

Lemma 6.2. *$V' \hookrightarrow \mathcal{L} \times \square^{n+1}_k$ intersects all the faces properly.*

Proof. Since H is flat, V' intersects properly all the faces of $\square^{n+1}_k$ of the form $F \times \square_k$. Since $\iota^*_{n+1,n+1,1}(V') = V$, which intersects the faces of $\square^n_k$ properly, we see that V' intersects $F^1_{n+1,n+1}$ properly. Since $V \cap (X \times \square^n_k) = \emptyset$, we must have $\iota^*_{n+1,n+1,0}(V') = 0$. We have thus shown that V' satisfies the face condition. $\square$

Theorem 6.3. *Let X be a quasi-projective scheme over a field k and let $f : \mathcal{L} \to X$ be a line bundle. Let $\iota : X \hookrightarrow \mathcal{L}$ denote the 0-section embedding. Then the cycle complex $z_s(\mathcal{L}|X, \bullet)$ is acyclic for all $s \in \mathbb{Z}$.*

Proof. It follows from Lemmas 6.1 and 6.2 that $H : \mathcal{L} \times \mathbb{A}^1_k \to \mathcal{L}$ defines a chain homotopy $H^* : z_s(\mathcal{L}|X, \bullet) \to z_s(\mathcal{L}|X, \bullet)[-1]$ between $H^*_0 = (H|_{\mathcal{L}\times 0})^*$ and $H^*_1 = (H|_{\mathcal{L}\times 1})^*$. It is clear that $H^*_1 = \mathrm{Id}_{z_s(\mathcal{L}|X,\bullet)}$ and the modulus condition implies that $H^*_0 = 0$. It follows that $z_s(\mathcal{L}|X, \bullet)$ is acyclic. □

References

[1] F. BINDA and S. SAITO, *Relative cycles with moduli and regulator maps*, J. Inst. Math. Jussieu (2017), to appear.
[2] S. BLOCH, *Algebraic cycles and higher K-theory*, Adv. Math. **61** (1986), 267–304.
[3] S. BLOCH, *The moving lemma for higher Chow groups*, J. Algebraic Geom. **3** (1994), 537–568.
[4] S. BLOCH, “Basic Definition and Properties of Higher Chow Groups”, web page of Spencer Bloch, University of Chicago, 2005.
[5] S. BLOCH and H. ESNAULT, *The additive dilogarithm*, Doc. Math. **Extra Vol.** (2003), 131–155.
[6] R. HARTSHORNE, “Algebraic Geometry”, Graduate Texts in Math., Vol. 52, Springer-Verlag, New York., 1977.
[7] W. KAI, *A moving lemma for algebraic cycles with modulus and contravariance*, arXiv:1507.07619v1, (2015).
[8] M. KERZ and S. SAITO, *Chow group of 0-cycles with modulus and higher dimensional class field theory*, Duke Math. J. **165** (2016), 2811–2897.
[9] A. KRISHNA and M. LEVINE, *Additive higher Chow groups of schemes*, J. Reine Angew. Math. **619** (2008), 75–140.
[10] A. KRISHNA and J. PARK, *Moving lemma for additive higher Chow groups*, Algebra Number Theory **6** (2012), 293–326.
[11] A. KRISHNA and J. PARK, *A module structure and a vanishing theorem for cycles with modulus*, Math. Res. Lett., to appear, arXiv:1412.7396v2, (2015).
[12] A. KRISHNA and J. PARK, *On the additive higher Chow groups of affine schemes*, Doc. Math. **21** (2016), 49–89.
[13] M. LEVINE, “Mixed Motives”, Mathematical Surveys and Monographs, Vol. 57, American Mathematical Society, Providence, RI, 1998.
[14] J. PARK, *Regulators on additive higher Chow groups*, Amer. J. Math. **131** (2009), 257–276.

School of Mathematics
Tata Institute of Fundamental Research
1 Homi Bhabha Road
Colaba, Mumbai, India
amal@math.tifr.res.in

Department of Mathematical Sciences
KAIST, 291 Daehak-ro Yuseong-gu
Daejeon, 305-701, Republic of Korea (South)
jinhyun@mathsci.kaist.ac.kr

INDICE DEL VOLUME XVII

pag.

PUBLICATIONS OF CLASSE DI SCIENZE DELLA SCUOLA NORMALE SUPERIORE

QUADERNI

E. De Giorgi, F. Colombini, L. C. Piccinini, *Frontiere orientate di misura minima e questioni collegate,* 1972.
C. Miranda, *Su alcuni problemi di geometria differenziale in grande per gli ovaloidi,* 1973.
G. Prodi, A. Ambrosetti, *Analisi non lineare,* 1973.
C. Miranda, *Problemi di esistenza in analisi funzionale,* 1975 (out of print).
I. T. Todorov, M. Mintchev, V. B. Petkova, *Conformal Invariance in Quantum Field Theory,* 1978.
A. Andreotti, M. Nacinovich, *Analytic Convexity and the Principle of Phragmén-Lindelöf,* 1980.
S. Campanato, *Sistemi ellittici in forma divergenza. Regolarità all'interno,* 1980.
Topics in Functional Analysis 1980-81, Contributors: F. Strocchi, E. Zarantonello, E. De Giorgi, G. Dal Maso, L. Modica, 1981.
G. Letta, *Martingales et intégration stochastique,* 1984.
Old and New Problems in Fundamental Physics, Meeting in honour of Gian Carlo Wick, 1986.
Interaction of Radiation with Matter, A Volume in honour of Adriano Gozzini, 1987.
M. Métivier, *Stochastic Partial Differential Equations in Infinite Dimensional Spaces,* 1988.
Symmetry in Nature, A Volume in honour of Luigi A. Radicati di Brozolo, 2 voll., 1989 (reprint 2005).
Nonlinear Analysis, A Volume in honour of Giovanni Prodi, 1991 (out of print).
C. Laurent-Thiébaut, J. Leiterer, *Andreotti-Grauert Theory on Real Hypersurfaces,* 1995.
J. Zabczyk, *Chance and Decision. Stochastic Control in Discrete Time,* 1996 (reprint 2000).
I. Ekeland, *Exterior Differential Calculus and Applications to Economic Theory,* 1998 (2000).
Electrons and Photons in Solids, A Volume in honour of Franco Bassani, 2001.
J. Zabczyk, *Topics in Stochastic Processes,* 2004.
N. Touzi, *Stochastic Control Problems, Viscosity Solutions and Application to Finance,* 2004.

Nuova serie

1. *Highlights in the quantum theory of condensed matter.* A symposium to honour Mario Tosi on his 72nd birthday, 2005.
2. *No-nonsense Physicist. An overview of Gabriele Giuliani's work and life*. M. Polini, G. Vignale, V. Pellegrini and J. K. Jain (eds.), 2016.

CATTEDRA GALILEIANA

P. L. Lions, *On Euler Equations and Statistical Physics,* 1998.
T. Björk, *A Geometric View of the Term Structure of Interest Rates,* 2000 (2001).
F. Delbaen, *Coherent Risk Measures,* 2002.
W. Schachermayer, *Portfolio Optimization in Incomplete Financial Markets,* 2004.
D. Duffie, *Credit Risk Modeling with Affine Processes,* 2004.
H. M. Soner, *Stochastic Optimal Control in Finance,* 2004 (2005).

LEZIONI LAGRANGE

C. Voisin, *Variations of Hodge Structure of Calabi-Yau Threefolds,* 1998.

LEZIONI FERMIANE

R. Thom, *Modèles mathématiques de la morphogénèse,* 1971.
S. Agmon, *Spectral Properties of Schrödinger Operators and Scattering Theory,* 1975.

M. F. ATIYAH, *Geometry of Yang-Mills Fields,* 1979.
KAC M., *Integration in Function Spaces and some of its Applications,* 1983 (out of print).
J. MOSER, *Integrable Hamiltonian Systems and Spectral Theory,* 1983.
T. KATO, *Abstract Differential Equations and Nonlinear Mixed Problems,* 1988.
W. H. FLEMING, *Controlled Markov Processes and Viscosity Solution of Nonlinear Evolution Equations,* 1988.
V. I. ARNOLD, *The Theory of Singularities and its Applications,* 1991.
J. P. OSTRIKER, *Development of Larger-Scale Structure in the Universe,* 1993.
S. P. NOVIKOV, *Solitons and Geometry,* 1994.
L. A. CAFFARELLI, *The Obstacle Problem,* 1999.
J. CHEEGER, *Degeneration of Riemannian Metrics under Ricci Curvature Bounds,* 2001.

NUOVA SERIE

1. A. HOLEVO, *Probabilistic and Statistical Aspects of Quantum Theory,* 2011.
2. *On Some Applications of Diophantine Approximations* (a translation of Carl Ludwig Siegel's *Über einige Anwendungen diophantischer Approximationen* by Clemens Fuchs), U. ZANNIER (ed.), 2014.

CRM SERIES

Dynamical Systems. Proceedings, 2002 (2003).
Part I: *Hamiltonian Systems and Celestial Mechanics*.
Part II: *Topological, Geometrical and Ergodic Properties of Dynamics*.
Matematica, Cultura e Società 2003, 2004.
Ricordando Franco Conti, 2004.
N. KRYLOV, *Probabilistic Methods of Investigating Interior Smoothness of Harmonic Functions Associated with Degenerate Elliptic Operators,* 2004.
Phase Space Analysis of Partial Differential Equations. Proceedings, 2004 (2005).

NUOVA SERIE

1. *Matematica, cultura e società 2004,* 2005.
2. *Matematica, cultura e società 2005,* 2006.
3. M. GIAQUINTA, D. MUCCI, *Maps into manifolds and currents: area and $W^{1,2}$-, $W^{1/2}$-, BV-energies,* 2006.
4. *Diophantine Geometry.* Proceedings, 2007.
5. G. MÉTIVIER, *Para-Differential Calculus and Nonlinear Systems,* 2008.
6. F. GUERRA, N. ROBOTTI, *Ettore Majorana. Aspects of his Scientific and Academic Activity,* 2008.
7. *Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT),* Y. CENSOR, M. JIANG, A. K. LOUIS (eds.), 2008.
8. *Quantum Information and Many Body Quantum System*. Proceedings, MARIE ERICSSON, SIMONE MONTANGERO (eds.), 2008.
9. *Singularities in Nonlinear Evolution Phenomena and Applications*. Proceedings, M. NOVAGA, G. ORLANDI (eds.), 2008 (2009).

   Matematica, cultura e società 2006, 2009.
10. *Probability, Uncertainty and Rationality,* Proceedings, H. HOSNI, F. MONTAGNA (eds.), 2010.
11. *Optimal Transportation, Geometry and Functional Inequalities,* L. AMBROSIO (ed.), 2010.

[#] These volumes are distributed by Cambridge University Press.

12*. *Asymptotics in Dynamics, Geometry and PDEs; Generalized Borel Summation,* Vol. I, Proceedings, O. COSTIN, F. FAUVET, F. MENOUS, D. SAUZIN (eds.), 2011.

12**. *Asymptotics in Dynamics, Geometry and PDEs; Generalized Borel Summation,* Vol. II, Proceedings, O. COSTIN, F. FAUVET, F. MENOUS, D. SAUZIN (eds.), 2011.

13. *Topics in Modern Regularity Theory*, Proceedings, G. MINGIONE (ed.), 2012.
Matematica, cultura e società 2007-2008, 2012.

14. *Configuration Spaces; Geometry, Combinatorics and Topology,* Proceedings, A. BJORNER, F. COHEN, C. DE CONCINI, C. PROCESI, M. SALVETTI (eds.), 2012.

15. *Geometric Partial Differential Equations*, A. CHAMBOLLE, M. NOVAGA, E. VALDINOCI (eds.), 2013.

16. *The Seventh European Conference on Combinatorics, Graph Theory and Applications*, EuroComb 2013. J. NEŠETŘIL, M. PELLEGRINI (eds.), 2013.

17. *Geometric Measure Theory and Real Analysis*, L. AMBROSIO (ed.), 2014.

18. *Geometry, Structure and Randomness in Combinatorics*, J. MATOUŠEK, J. NEŠETŘIL, M. PELLEGRINI (eds.), 2015.

19. *Free Discontinuity Problems*, N. FUSCO, A. PRATELLI (eds.), 2016.

COLLOQUIA

1. *Colloquium De Giorgi 2006,* 2007.
2. *Colloquium De Giorgi 2007 and 2008,* 2009.
3. *Colloquium De Giorgi 2009,* 2012.
4. *Colloquium De Giorgi 2010-2012,* 2013.
5. *Colloquium De Giorgi 2013 and 2014,* 2015.

LECTURE NOTES

G. DA PRATO, *Introduction to Differential Stochastic Equations*, 1995 (second edition 1998).

L. AMBROSIO, *Corso introduttivo alla Teoria Geometrica della Misura ed alle Superfici Minime*, 2000 (reprint 2010).

E. VESENTINI, *Introduction to Continuous Semigroups*, 1996 (second edition 2002).

C. PETRONIO, *A Theorem of Eliashberg and Thurston on Foliations and Contact Structures*, 1997.

Quantum cohomology at the Mittag-Leffler Institute, a cura di Paolo Aluffi, 1998.

G. BINI, C. DE CONCINI, M. POLITO, C. PROCESI, *On the Work of Givental Relative to Mirror Symmetry*, 1998.

H. PHAM, *Imperfections de Marchés et Méthodes d'Evaluation et Couverture d'Options*, 1998.

H. CLEMENS, *Introduction to Hodge Theory*, 1998.

Seminari di Geometria Algebrica 1998-1999, 1999.

A. LUNARDI, *Interpolation Theory*, 1999.

R. SCOGNAMILLO, *Rappresentazioni dei gruppi finiti e loro caratteri*, 1999.

S. RODRIGUEZ, *Symmetry in Physics*, 1999.

F. STROCCHI, *Symmetry Breaking in Classical Systems*, 1999 (2000).

L. AMBROSIO, P. TILLI, *Selected Topics on* "Analysis in Metric Spaces", 2000.

A. C. G. MENNUCCI, S. K. MITTER, *Probabilità ed Informazione*, 2000.

S. V. BULANOV, *Lectures on Nonlinear Physics*, 2000 (2001).

Lectures on Analysis in Metric Spaces, a cura di Luigi Ambrosio e Francesco Serra Cassano, 2000 (2001).

L. CIOTTI, *Lectures Notes on Stellar Dynamics*, 2000 (2001).

S. RODRIGUEZ, *The Scattering of Light by Matter*, 2001.

G. DA PRATO, *An Introduction to Infinite Dimensional Analysis*, 2001.

S. SUCCI, *An Introduction to Computational Physics*: – Part I: *Grid Methods*, 2002.

D. BUCUR, G. BUTTAZZO, *Variational Methods in Some Shape Optimization Problems*, 2002.

A. MINGUZZI, M. TOSI, *Introduction to the Theory of Many-Body Systems*, 2002.

S. SUCCI, *An Introduction to Computational Physics*: – Part II: *Particle Methods*, 2003.

A. MINGUZZI, S. SUCCI, F. TOSCHI, M. TOSI, P. VIGNOLO, *Numerical Methods for Atomic Quantum Gases*, 2004.

NUOVA SERIE

1. M. TOSI, P. VIGNOLO, *Statistical Mechanics and the Physics of Fluids*, 2005 (second edition).
2. M. GIAQUINTA, L. MARTINAZZI, *An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs*, 2005.
3. G. DELLA SALA, A. SARACCO, A. SIMIONIUC, G. TOMASSINI, *Lectures on Complex Analysis and Analytic Geometry*, 2006.
4. M. POLINI, M. TOSI, *Many-Body Physics in Condensed Matter Systems*, 2006.

   P. AZZURRI, *Problemi di Meccanica*, 2007.
5. R. BARBIERI, *Lectures on the ElectroWeak Interactions*, 2007.
6. G. DA PRATO, *Introduction to Stochastic Analysis and Malliavin Calculus*, 2007.

   P. AZZURRI, *Problemi di meccanica*, 2008 (second edition).

   A. C. G. MENNUCCI, S. K. MITTER, *Probabilità e informazione*, 2008 (second edition).
7. G. DA PRATO, *Introduction to Stochastic Analysis and Malliavin Calculus*, 2008 (second edition).
8. U. ZANNIER, *Lecture Notes on Diophantine Analysis*, 2014 (revised edition).
9. A. LUNARDI, *Interpolation Theory*, 2009 (second edition).

   G. FOCHI, *Chimica da capire,* Compendio di chimica generale con brevi cenni di chimica inorganica descrittiva, 2009 (second edition).
10. L. AMBROSIO, G. DA PRATO, A. MENNUCCI, *Introduction to Measure Theory and Integration*, 2011.
11. M. GIAQUINTA, L. MARTINAZZI, *An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs*, 2012 (second edition).

    G. PRADISI, *Lezioni di metodi matematici della fisica*, 2012.
12. G. BELLETTINI, *Lecture Notes on Mean Curvature Flow, Barriers and Singular Perturbations*, 2013.
13. G. DA PRATO, *Introduction to Stochastic Analysis and Malliavin Calculus*, 2014 (third edition).
14. R. SCOGNAMILLO, U. ZANNIER, *Introductory Notes on Valuation Rings and Function Fields in One Variable*, 2014.
15. S. DIPIERRO, M. MEDINA, E. VALDINOCI, *Fractional Elliptic Problems with Critical Growth in the Whole of* $\mathbb{R}^n$, 2017.

OTHER PUBLICATIONS

A. ANDREOTTI, *SELECTA vol. I, Geometria algebrica,* 1982.

A. ANDREOTTI, *SELECTA vol. II, Analisi complessa,* Tomo I, 1982 (reprint 2005) e Tomo II, 1994.

A. ANDREOTTI, *SELECTA vol. III, Complessi di operatori differenziali,* 1999.

S. HILDEBRANDT, A. TROMBA, *Principi di minimo. Forme ottimali in natura,* 2007.

Instructions to order these publications are available on the website:

`http://edizioni.sns.it`

INFORMATION FOR AUTHORS

The Annali della Scuola Normale Superiore, Classe di Scienze, publishes papers that contribute to the development of Mathematics both from the theoretical and the applied point of view. Research papers or papers of expository type are considered for publication.

How to submit a paper

Authors should submit their manuscript as a PDF file to the address:

publ_sci@sns.it

An abstract should also be enclosed. The abstract must be in English, in the form of a text file containing as little mathematical formulae as possible. If formulae cannot be avoided they should be written using the customary $\TeX$ or $\LaTeX$-like method.
The appropriate Mathematics Subject Classification (MSC) number(s) should also be provided.

Papers will not be considered for publication unless the abstract and MSC numbers have been sent by e-mail.

The sender's e-mail address must be the one where authors wish to receive all correspondence concerning their paper. Authors should also indicate the postal address where they wish to receive ordinary mail. The editorial office (Redazione) should be informed immediately of any change of address.
Prospective contributors are expected to keep a copy of their own paper identical to the copy they submitted.

On submitting a paper authors bind themselves neither to submit the same or essentially the same paper to other journals nor to do so before the Editorial Board has come to a decision concerning publication.

How to prepare manuscripts

Papers submitted for publication should be preferably in English.

Authors are encouraged to prepare their manuscripts using $\TeX$ or one of the systems derived from it. Whatever system they use, authors will be requested, in case of acceptance, to send the source file(s), in order to speed up publication. Figures and complex diagrams should be sent as PostScript or Encapsulated PostScript files, or printed on a separate sheet in a form suitable for reproduction.

The first page of each paper should contain: title, full name of each author (first names should not be abbreviated), abstract and MSC number(s); key-words need not be indicated. Acknowledgements should appear at the end of the introduction. The last page should contain the affiliation, postal address and e-mail address of each author.

If the paper title exceeds 55 characters, including spaces, a short title to be used as a running-head should be provided.

Literature references should be listed in alphabetical order at the end of the paper. Authors should imitate the following examples, which correspond respectively to papers published in journals, papers published in proceedings or other edited volumes, and books.

[1] E. De Giorgi, *Nuovi teoremi relativi alle misure* $(r-1)$*-dimensionali in uno spazio ad* r *dimensioni*, Ricerche Mat. **4** (1955), 95–113.

[2] E. Bombieri and A. J. van der Poorten, *Continued fractions of algebraic numbers*, In: "Computational Algebra and Number Theory", W. Bosma and A. van der Poorten (eds.), Kluwer Academic Publishers, Dordrecht, 1995, 137–158.

[3] S. K. Donaldson and P. B. Kronheimer, "The Geometry of Four-Manifolds", Oxford Science Publications, Clarendon Press, Oxford, 1990.

Authors are requested to make sure that they are including all the appropriate information in each reference item, and to pay special attention to the capitalization and italicization rules. The Abbreviations of Names of Serials, available from the American Mathematical Society, should be used as a reference for all abbreviations. If several papers by the same set of authors are included in the references, name(s) should be repeated.

Open access policy

A paper accepted for publication in the Annali della Scuola Normale Superiore, Classe di Scienze, can be posted on personal web pages, institutional servers, or public repositories, provided:

1) All the publication details (journal name, year, volume, page numbers, DOI code) are indicated as soon as they become available;
2) The article's version is identical to the final one submitted by the author(s) and accepted by the journal, before undergoing copy-editing and typsetting; posting the PDF file that the journal sends to the author(s) is strictly forbidden.

If a paper was already posted before acceptance, it can be kept available provided the above two conditions are met.

Other information

As a rule, authors will be asked to correct only one set of proofs. Contributors are invited to limit corrections as far as possible to typographical errors only.

Fifteen off-prints and a high-resolution PDF file will be supplied, free of charge, for each article. Additional copies may be requested by authors when returning the corrected proofs and will be paid separately.

ISSN 0391-173x (print version)
ISSN 2036-2145 (online version)

UMBERTO ZANNIER Direttore responsabile

Autorizzazione N. 8 del Tribunale di Pisa in data 12 giugno 1964
Fotocomposizione: CompoMat srl, Loc. Braccone, 02040 Configni (RI), Italia
Finito di stampare presso le Industrie Grafiche della Pacini Editore S.r.l.
Via A. Gherardesca, 56121 Ospedaletto, Pisa, nel dicembre 2017