Groups Geom. Dyn. 5 (2011), 553–565
DOI 10.4171/GGD/139

Groups, Geometry, and Dynamics

Reduction theory of point clusters in projective space

Michael Stoll

Dedicated to the memory of Fritz Grunewald

Abstract. We generalise earlier results of John Cremona and the author on the reduction theory of binary forms, whose zeros give point clusters in $\mathbb{P}^1$, to point clusters in projective spaces $\mathbb{P}^n$ of arbitrary dimension. In particular, we show how to find a reduced representative in the $\mathrm{SL}(n+1, \mathbb{Z})$-orbit of a given cluster. As an application, we show how one can find a unimodular transformation that produces a small equation for a given smooth plane curve.

Mathematics Subject Classification (2010). 11E10, 11H46, 11E39.

Keywords. Reduction theory, point clusters.

1. Introduction

In this paper, we generalise the results of [8] on the reduction theory of binary forms, which describe positive zero-cycles in $\mathbb{P}^1$, to positive zero-cycles (or point clusters) in projective spaces of arbitrary dimension. This should have applications to more general projective varieties in $\mathbb{P}^n$, by associating a suitable positive zero-cycle to them in an $\mathrm{PGL}(n+1)$-invariant way. We discuss this in the case of (smooth) plane curves.

The basic problem motivating this work is as follows. Consider projective varieties over $\mathbb{Q}$ in some $\mathbb{P}^n$, with fixed discrete invariants. On this set, there is an action of $\mathrm{SL}(n+1, \mathbb{Z})$ by linear substitution of the coordinates. We would like to be able to select a specific representative of each orbit, which we will call *reduced*, in a way that is as canonical as possible. Hopefully, this representative will then also allow a description as the zero set of polynomials with fairly small integer coefficients.

Recall the main ingredients of the approach taken in [8]. The key role is played by a map z from binary forms of degree d into the symmetric space of $\mathrm{SL}(2, \mathbb{R})$ (which is the hyperbolic plane $\mathcal{H}$ in this case) that is equivariant with respect to the action of $\mathrm{SL}(2, \mathbb{Z})$. We then define a form F to be *reduced* if $z(F)$ is in the standard fundamental domain for $\mathrm{SL}(2, \mathbb{Z})$ in $\mathcal{H}$. In order to make the map z as canonical as possible, we use a larger group than $\mathrm{SL}(2, \mathbb{Z})$, namely $\mathrm{SL}(2, \mathbb{C})$; we then look for a map z from binary forms with complex coefficients into the symmetric space $\mathcal{H}_{\mathbb{C}}$ for $\mathrm{SL}(2,\mathbb{C})$ that is $\mathrm{SL}(2,\mathbb{C})$-equivariant and commutes with complex conjugation. This map restricted to real forms will have image contained in $\mathcal{H}$ and satisfy our initial requirement.

Now there are in general many possible such maps z (for exceptions, see Remark 12 below). We therefore need to pick one of them. In [8] this is achieved by a geometric property: we define a function on $\mathcal{H}_{\mathbb{C}}$, depending on F, that measures how far a point is from the roots of F (up to an arbitrary additive constant); the covariant $z(F)$ is then the unique point in $\mathcal{H}_{\mathbb{C}}$ minimising this distance. This is essentially the same approach (but in a different interpretation) as that used by Julia in his thesis [5], who works out what $z(F)$ is for F of degree 3 or 4, but defines it more generally. He did not prove that his covariant is always well-defined, though. Julia was building on previous work by Hermite [3], [4]. For a more detailed discussion, see [8].

In our more general situation, we work with the space $\mathcal{H}_{n,\mathbb{R}}$ of positive definite quadratic forms in $n+1$ variables, modulo scaling, and the space $\mathcal{H}_{n,\mathbb{C}}$ of positive definite Hermitian forms in $n+1$ variables, modulo scaling (by positive real factors). There is a natural action of complex conjugation on $\mathcal{H}_{n,\mathbb{C}}$; the subset fixed by it can be identified with $\mathcal{H}_{n,\mathbb{R}}$.

We use the formula for the distance function mentioned above to obtain a similar function on $\mathcal{H}_{n,\mathbb{C}}$, depending on a collection of points in $\mathbb{P}^n(\mathbb{C})$. Under a suitable condition on the point cluster or zero-cycle Z, this distance function has a unique critical point, which provides a global minimum. We assign this point to Z as its covariant $z(Z)$, thus solving our problem.

2. Basics

In all of the paper, we fix $n \geq 0$.

We consider the group $G = \mathrm{SL}(n+1,\mathbb{C})$ and its natural action on forms (homogeneous polynomials) in $n+1$ variables $X_0,\ldots,X_n$ by linear substitutions; this action will be on the right:

$$F(X_0, X_1, \ldots, X_n) \cdot (a_{ij})_{0\le i,j\le n} = F\bigl(\textstyle\sum_{j=0}^n a_{0j}X_j, \ldots, \sum_{j=0}^n a_{nj}X_j\bigr).$$

The same action is used for Hermitian forms in $X_0,\ldots,X_n$. A Hermitian form can be considered as a bihomogeneous polynomial of bidegree $(1,1)$ in two sets of variables $X_0,\ldots,X_n$ and $\bar{X}_0,\ldots,\bar{X}_n$, where the action on the second set is through the complex conjugate of the matrix. The form Q is Hermitian if $Q(\bar{X};X) = \bar{Q}(X;\bar{X})$, where $\bar{Q}$ denotes the form obtained from Q by replacing the coefficients with their complex conjugates. Hermitian forms can also be identified with Hermitian matrices, i.e., matrices A such that $A^\top = \bar{A}$, where A corresponds to Q if $Q(\boldsymbol{x}) = \bar{\boldsymbol{x}} A \boldsymbol{x}^\top$; then the action of G is given by $A \cdot \gamma = \bar{\gamma}^\top A \gamma$.

The group G also acts on coordinates $(\xi_0, \dots, \xi_n)$ on the right via the contragredient representation,

$$(\xi_0, \dots, \xi_n) \cdot \gamma = (\xi_0, \dots, \xi_n)\gamma^{-\top}.$$

These actions are compatible in the sense that

$$(Q \cdot \gamma)(\boldsymbol{x} \cdot \gamma) = Q(\boldsymbol{x})$$

for Hermitian forms Q and coordinate vectors $\boldsymbol{x}$.

3. Point clusters

The actions described above induce actions of $\mathrm{PSL}(n+1, \mathbb{C}) = \mathrm{PGL}(n+1, \mathbb{C})$ on projective schemes over $\mathbb{C}$ and points in projective space $\mathbb{P}^n(\mathbb{C})$. The first specialises and the second generalises to an action on positive zero-cycles.

Definition 1. A *positive zero-cycle* or *point cluster* is a formal sum $Z = \sum_{j=1}^m P_j$ of points $P_j \in \mathbb{P}^n$. The number m of points is the *degree* of Z, written $\deg Z$. If $L \subset \mathbb{P}^n$ is a linear subspace, we let $Z|_L$ be the sum of those points in Z that lie in L.

Definition 2. Let Z be a point cluster in $\mathbb{P}^n$.

(1) Z is *split* if there are two disjoint and nonempty linear subspaces L_1, L_2 of $\mathbb{P}^n$ such that $Z = Z|_{L_1} + Z|_{L_2}$. Otherwise, Z is *non-split*.

(2) Z is *semi-stable* if for every linear subspace $L \subset \mathbb{P}^n$, we have

$$(n+1) \deg Z|_L \leq (\dim L + 1) \deg Z.$$

(3) Z is *stable* if for every linear subspace $\emptyset \neq L \subsetneq \mathbb{P}^n$, we have

$$(n+1) \deg Z|_L < (\dim L + 1) \deg Z.$$

Remark 3. Note that a split point cluster cannot be stable.

If we identify the cluster $Z = \sum_{j=1}^m P_j$, where $P_j = (a_{j0} : a_{j1} : \cdots : a_{jn})$, with the form $F(Z) = \prod_{j=1}^m (a_{j0}x_0 + a_{j1}x_1 + \cdots + a_{jn}x_n)$ (up to scaling), then Z is (semi-)stable if and only if $F(Z)$ is (semi-)stable in the sense of Geometric Invariant Theory; see [7].

If $n = 1$, then the notions of stable and semi-stable defined here coincide with those defined in [8] (in Def. 4.1 and before Prop. 5.2) for binary forms.

Definition 4. Let $\mathcal{Z}_m$ denote the set of point clusters of degree m in $\mathbb{P}^n(\mathbb{C})$, $\mathcal{Z}_m^{\mathrm{sst}}$ the subset of semi-stable and $\mathcal{Z}_m^{\mathrm{st}}$ the subset of stable point clusters. We denote by $\mathcal{Z}_m(\mathbb{R})$ etc. the subset of point clusters fixed by complex conjugation, which acts via $\sum_j P_j \mapsto \sum_j \bar{P}_j$.

For notational convenience, for a point cluster Z and $-1 \le k \le n$ we define

$$\varphi_Z(k) = \max\{\deg Z|_L : L \subset \mathbb{P}^n \text{ a } k\text{-dimensional linear subspace}\}.$$

Then Z is semi-stable if and only if $\varphi_Z(k) \le \frac{k+1}{n+1} \deg Z$ for all $0 \le k \le n$, and Z is stable if and only if the inequality is strict for $0 \le k < n$.

We let $\langle P, P' \rangle = \bar{P}(P')^\top$ denote the standard Hermitian inner product on row vectors and $\|P\|^2 = \langle P, P \rangle$ the corresponding norm. The next lemma is the basis for most of what follows.

Lemma 5. *Let $Z \in \mathcal{Z}_m$. Fix row vectors P_j, $j \in \{1, \dots, m\}$, representing the points in Z, such that $\|P_j\|^2 = 1$. Then there is a constant $c > 0$ such that for every positive definite Hermitian matrix Q with eigenvalues $0 < \lambda_0 \le \lambda_1 \le \dots \le \lambda_n$, we have*

$$\prod_{j=1}^{m} \left(\bar{P}_j Q P_j^\top\right) \ge c \prod_{k=0}^{n} \lambda_k^{\varphi_Z(k) - \varphi_Z(k-1)}.$$

Proof. Let $B = b_0, \dots, b_n$ be a unitary basis of $\mathbb{C}^{n+1}$. Let $E_k = \langle b_0, \dots, b_k \rangle$ the subspace generated by the first $k+1$ basis vectors. By definition of φ_Z, the set $\Sigma(B) \subset S_m$ of permutations σ with the following property is nonempty:

$$P_{\sigma(j)} \notin E_k \quad \text{if } j > \varphi_Z(k).$$

Define $k_\sigma(j) = \min\{k : \sigma(j) \le \varphi_Z(k)\}$; then $P_{\sigma(j)} \notin E_{k_\sigma(j)-1}$ if $\sigma \in \Sigma(B)$. Write $P_j = \sum_{i=0}^{n} \xi_{ji} b_i$ and define

$$f_\sigma(B) = \prod_{j=1}^{m} \Bigg(\sum_{i=k_\sigma(j)}^{n} |\xi_{\sigma(j),i}|^2 \Bigg) = \prod_{j=1}^{m} \Bigg(\sum_{i=k_\sigma(j)}^{n} |\langle P_{\sigma(j)}, b_i \rangle|^2 \Bigg)$$

and

$$f(B) = \max\{f_\sigma(B) : \sigma \in S_m\}.$$

It is clear that f_σ is continuous on the set of unitary bases and that $f_\sigma(B) > 0$ if $\sigma \in \Sigma(B)$. This implies that f is continuous and positive. Since the set of all unitary bases (i.e., $\mathrm{U}(n+1)$) is compact, there is some $c > 0$ such that $f(B) \ge c$ for all B.

Now let Q be a positive definite Hermitian matrix as in the statement of the Lemma. Let $B = b_0, \dots, b_n$ be a unitary basis of eigenvectors of Q such that $b_j Q = \lambda_j b_j$. We then have for $\sigma \in S_m$ and using notation introduced above

$$\begin{aligned}
\prod_{j=1}^{m} \left(\bar{P}_j Q P_j^\top\right) &= \prod_{j=1}^{m} \left(\bar{P}_{\sigma(j)} Q P_{\sigma(j)}^\top\right) = \prod_{j=1}^{m} \Bigg(\sum_{i=0}^{n} \lambda_i |\xi_{\sigma(j),i}|^2 \Bigg) \\
&\ge \prod_{j=1}^{m} \Bigg(\lambda_{k_\sigma(j)} \sum_{i=k_\sigma(j)}^{n} |\xi_{\sigma(j),i}|^2 \Bigg) \\
&= f_\sigma(B) \prod_{j=1}^{m} \lambda_{k_\sigma(j)} = f_\sigma(B) \prod_{k=0}^{n} \lambda_k^{\varphi_Z(k) - \varphi_Z(k-1)}.
\end{aligned}$$

Taking the maximum over all $\sigma \in S_m$ now shows that

$$\prod_{j=1}^{m} \big(\bar{P}_j Q P_j^\top\big) \geq f(B) \prod_{k=0}^{n} \lambda_k^{\varphi_Z(k)-\varphi_Z(k-1)} \geq c \prod_{k=0}^{n} \lambda_k^{\varphi_Z(k)-\varphi_Z(k-1)}. \qquad \square$$

4. The covariant

Definition 6. Let $\widetilde{\mathcal{Z}}_m$ denote the set of point clusters of degree m with a choice of coordinates for the points, up to scaling the coordinates of the points with factors whose product is 1. We will call $\widetilde{Z} \in \widetilde{\mathcal{Z}}_m$ a *point cluster with scaling*. We define $\widetilde{\mathcal{Z}}_m^{\text{st}}$ and $\widetilde{\mathcal{Z}}_m^{\text{sst}}$ analogously.

For $\lambda \in \mathbb{C}^\times$ and $\widetilde{Z} \in \widetilde{\mathcal{Z}}_m$, we write $\lambda\widetilde{Z}$ for the cluster with scaling that we obtain by scaling one of the points in $\widetilde{Z}$ by λ. This defines an action of $\mathbb{C}^\times$ on $\widetilde{\mathcal{Z}}_m$ such that the quotient $\mathbb{C}^\times \backslash \widetilde{\mathcal{Z}}_m$ is $\mathcal{Z}_m$. If $\widetilde{Z} \in \widetilde{\mathcal{Z}}_m$, then we write Z for the image of $\widetilde{Z}$ in $\mathcal{Z}_m$.

Definition 7. For a point cluster with scaling $\widetilde{Z} \in \widetilde{\mathcal{Z}}_m$, pick a representative $\sum_{j=1}^m P_j$ with row vectors P_j. Then, for $Q \in \mathcal{H}_{n,\mathbb{C}}$, represented by a Hermitian matrix, we define

$$D_{\widetilde{Z}}(Q) = D(\widetilde{Z}, Q) = \sum_{j=1}^{m} \log(\bar{P}_j Q P_j^\top) - \frac{m}{n+1} \log\det Q.$$

$D(\widetilde{Z}, Q)$ is clearly invariant under scaling of Q, and it does not depend on the choice of representative for $\widetilde{Z}$. Note also that for $\gamma \in G$,

$$D(\widetilde{Z} \cdot \gamma, Q \cdot \gamma) = D(\widetilde{Z}, Q).$$

Furthermore, we have $D(\bar{\widetilde{Z}}, \bar{Q}) = D(\widetilde{Z}, Q)$ and $D(\lambda\widetilde{Z}, Q) = \log|\lambda|^2 + D(\widetilde{Z}, Q)$.

This function generalises the distance function used in Prop. 5.3 of [8]. We will now proceed to show that for stable clusters, there is a unique form $Q \in \mathcal{H}_{n,\mathbb{C}}$ that minimises this distance.

To that end, we now identify $\mathcal{H}_{n,\mathbb{C}}$ with the set of positive definite Hermitian matrices of determinant 1. This is a real $n(n+2)$-dimensional submanifold of the space of all complex $(n+1) \times (n+1)$-matrices. $\mathrm{SL}(n+1, \mathbb{C})$ acts transitively on this space, and the tangent space T at the identity matrix I consists of the Hermitian matrices of trace zero. We say that a twice continuously differentiable function on $\mathcal{H}_{n,\mathbb{C}}$ is *convex* if its second derivative is positive semidefinite, and *strictly convex* if its second derivative is positive definite. Then the usual conclusions on convex functions apply.

Lemma 8. *Let $\widetilde{Z} \in \widetilde{\mathcal{Z}}_m$ be a point cluster with scaling.*

(1) *The function $D_{\widetilde{Z}}$ is convex.*

(2) *If Z is non-split, then $D_{\tilde{Z}}$ is strictly convex.*
(3) *If Z is semi-stable, then $D_{\tilde{Z}}$ is bounded from below.*
(4) *If Z is stable, then the sets $\{Q \in \mathcal{H}_{n,\mathbb{C}} : D_{\tilde{Z}}(Q) \le B\}$ are compact for all $B \in \mathbb{R}$.*

Proof. Since scaling $\tilde{Z}$ only changes $D_{\tilde{Z}}$ by an additive constant, we can assume that $\tilde{Z} = P_1 + \cdots + P_m$ with row vectors P_j satisfying $\|P_j\|^2 = 1$.

(1) Since $D_{\tilde{Z}}(Q \cdot \gamma) = D_{\tilde{Z}\cdot\gamma^{-1}}(Q)$, we can assume that $Q = I$. We compute the second derivative at $\lambda = 0$ of $\lambda \mapsto f(\lambda) = D_{\tilde{Z}}(\exp(\lambda A))$, where $A \neq 0$ is a Hermitian trace-zero matrix (i.e., $A \in T$). We have

$$\begin{aligned} D_{\tilde{Z}}(\exp(\lambda A)) &= \sum_j \log(1 + \bar{P}_j A P_j^\top \cdot \lambda + \bar{P}_j A^2 P_j^\top \cdot \lambda^2/2 + \dots) \\ &= \sum_j (\bar{P}_j A P_j^\top \cdot \lambda + (\bar{P}_j A^2 P_j^\top - (\bar{P}_j A P_j^\top)^2) \cdot \lambda^2/2 + \dots) \end{aligned}$$

The second derivative therefore is

$$\sum_j (\bar{P}_j A^2 P_j^\top - (\bar{P}_j A P_j^\top)^2) = \sum_j (\|P_j \bar{A}\|^2 \|P_j\|^2 - |\langle P_j \bar{A}, P_j\rangle|^2) \ge 0$$

by the Cauchy–Schwarz inequality. This shows that the second derivative is positive semidefinite, whence the first claim.

(2) As in (1), it suffices to consider the case $Q = I$, since the condition for Z to be non-split is invariant under the action of $\mathrm{SL}(n+1, \mathbb{C})$. The second derivative in (1) vanishes exactly when P_j is an eigenvector of A, for all j. Since Z is non-split, this is only possible if A is a scalar matrix: the P_j must all be in the same eigenspace, and their span is the whole space. But $A \neq 0$ has trace zero, so A cannot be a scalar matrix. So the second derivative at I must be positive definite.

(3) By Lemma 5, we find some $c > 0$ such that for $Q \in \mathcal{H}_{n,\mathbb{C}}$ with eigenvalues $\lambda_0 \le \cdots \le \lambda_n$, we have

$$\prod_{j=1}^{m} \left(\bar{P}_j Q P_j^\top\right) \ge c \prod_{k=0}^{n} \lambda_k^{\varphi_Z(k) - \varphi_Z(k-1)}.$$

With $\varphi_Z(k) \le (k+1)\frac{m}{n+1}$, we obtain

$$\begin{aligned} D_{\tilde{Z}}(Q) &\ge \log c + \sum_{k=0}^{n} (\varphi_Z(k) - \varphi_Z(k-1)) \log \lambda_k \\ &= \log c + m \log \lambda_n - \sum_{k=1}^{n} \varphi_Z(k-1)(\log \lambda_k - \log \lambda_{k-1}) \end{aligned}$$

$$\begin{aligned}
&\geq \log c + m \log \lambda_n - \frac{m}{n+1} \sum_{k=1}^{n} k(\log \lambda_k - \log \lambda_{k-1}) \\
&= \log c + \frac{m}{n+1} \sum_{k=0}^{n} \log \lambda_k \\
&= \log c
\end{aligned}$$

(recall that $\sum_k \log \lambda_k = \log \det Q = 0$).

(4) We now use that $\varphi_Z(k) \leq (k+1)\frac{m}{n+1} - \frac{1}{n+1}$ for $0 \leq k \leq n-1$. The computation in the proof of (3) above then yields

$$\begin{aligned}
D_{\tilde{Z}}(Q) &\geq \log c + m \log \lambda_n - \sum_{k=1}^{n} \varphi_Z(k-1)(\log \lambda_k - \log \lambda_{k-1}) \\
&\geq \log c + m \log \lambda_n \\
&\qquad - \frac{m}{n+1} \sum_{k=1}^{n} k(\log \lambda_k - \log \lambda_{k-1}) + \frac{1}{n+1} \sum_{k=1}^{n} (\log \lambda_k - \log \lambda_{k-1}) \\
&= \log c + \frac{1}{n+1}(\log \lambda_n - \log \lambda_0).
\end{aligned}$$

So $D_{\tilde{Z}}(Q) \leq B$ implies that λ_n/λ_0 is bounded, but this implies that the subset of $Q \in \mathcal{H}_{n,\mathbb{C}}$ satisfying $D_{\tilde{Z}}(Q) \leq B$ is also bounded. Since it is obviously closed, it must be compact. □

Remark 9. Note that if Z is not stable, then there are sets $\{Q : D_{\tilde{Z}}(Q) \leq B\}$ that are not compact. Indeed, there is a linear subspace $L_0 \subset \mathbb{C}^{n+1}$ of some dimension $0 < k+1 < n+1$ containing at least $(k+1)m/(n+1)$ points of Z. Let L_1 be its orthogonal complement. Let Q_λ be the Hermitian matrix with eigenvalue $\lambda^{-(n-k)}$ on L_0 and eigenvalue λ^{k+1} on L_1. Then we have for $\lambda \geq 1$ that

$$D_{\tilde{Z}}(Q_\lambda) \leq \text{const.} + (k+1)\frac{m}{n+1} \log \lambda^{-(n-k)} + (n-k)\frac{m}{n+1} \log \lambda^{k+1} = \text{const.};$$

but the set $\{Q_\lambda : \lambda \geq 1\}$ is not relatively compact.

We also see that $D_{\tilde{Z}}$ is not bounded from below when Z is not semi-stable, since using the corresponding strict inequality, we find with a similar argument that

$$D_{\tilde{Z}}(Q_\lambda) \leq \text{const.} - \varepsilon \log \lambda$$

for some $\varepsilon > 0$.

Corollary 10. *If $\tilde{Z} \in \tilde{\mathcal{Z}}^{\text{st}}_m$, then the function $D_{\tilde{Z}}$ has a unique critical point $z(Z)$ on $\mathcal{H}_{n,\mathbb{C}}$, and at this point $D_{\tilde{Z}}$ achieves its global minimum $\log \theta(\tilde{Z})$ (for some $\theta(\tilde{Z}) \in \mathbb{R}_{>0}$).*

Proof. By Lemma 8, we know that $D_{\tilde{Z}}$ is strictly convex and also that for all B the set $\{Q \in \mathcal{H}_{n,\mathbb{C}} : D_{\tilde{Z}}(Q) \leq B\}$ is compact. The first property implies that every critical point must be a local minimum. By the second property, there exists a global minimum. If there were two distinct local minima, then on a path joining the two, there would have to be a local maximum, but then the second derivative would not be positive definite in this point, a contradiction. Hence there is a unique local minimum, which must then also be the global minimum and the unique critical point.

Since $D_{\lambda\tilde{Z}} = \log|\lambda|^2 + D_{\tilde{Z}}$, the minimising point in $\mathcal{H}_{n,\mathbb{C}}$ does not depend on the scaling, so it only depends on Z, and the notation $z(Z)$ is justified. □

Note that we have $\theta(\lambda\tilde{Z}) = |\lambda|^2\theta(\tilde{Z})$.

Corollary 10 defines $z\colon \mathcal{Z}_m^{\mathrm{st}} \to \mathcal{H}_{n,\mathbb{C}}$ and $\theta\colon \tilde{\mathcal{Z}}_m^{\mathrm{st}} \to \mathbb{R}_{>0}$. The latter extends to

$$\theta\colon \tilde{\mathcal{Z}}_m \longrightarrow \mathbb{R}_{\geq 0}$$

with the definition $\theta(\tilde{Z}) = \inf_{Q\in\mathcal{H}_{n,\mathbb{C}}} \exp(D(\tilde{Z}, Q))$. By Lemma 8 (3) we have $\theta(\tilde{Z}) > 0$ if $\tilde{Z} \in \tilde{\mathcal{Z}}_m^{\mathrm{sst}}$, and by the preceding remark, $\theta(\tilde{Z}) = 0$ if $\tilde{Z}$ is not semi-stable.

Corollary 11. *The function $z\colon \mathcal{Z}_m^{\mathrm{st}} \to \mathcal{H}_{n,\mathbb{C}}$ is $\mathrm{SL}(n+1,\mathbb{C})$-equivariant. It also satisfies $z(\bar{Z}) = \overline{z(Z)}$. In particular, z restricts to $z : \mathcal{Z}_m^{\mathrm{st}}(\mathbb{R}) \to \mathcal{H}_{n,\mathbb{R}}$.*

The function $\theta\colon \tilde{\mathcal{Z}}_m \to \mathbb{R}_{\geq 0}$ is invariant under $\mathrm{SL}(n+1,\mathbb{C})$ and under complex conjugation.

Proof. The first statement follows from the invariance of D (under the action of both $\mathrm{SL}(n+1,\mathbb{C})$ and complex conjugation) and the uniqueness of $z(Z)$. The second statement follows from the invariance of D. □

Remark 12. In some cases the point $z(Z)$ is uniquely determined by symmetry considerations. Namely if the point cluster $Z \in \mathcal{Z}_m^{\mathrm{st}}$ is stabilised by a subgroup of $\mathrm{SL}(n+1,\mathbb{C})$ that fixes a unique point in $\mathcal{H}_{n,\mathbb{C}}$, then $z(Z)$ must be this point. See Lemma 3.1 in [8] for a precise statement. This observation facilitates the numerical computation of $z(Z)$, since it eliminates the need for finding numerically the minimum of the distance function on $\mathcal{H}_{n,\mathbb{C}}$.

Example 13. Consider a sum Z of $n+2$ points in general position in $\mathbb{P}^n(\mathbb{C})$. Then Z is stable. Since $\mathrm{PGL}(n+1,\mathbb{C})$ acts transitively on $(n+2)$-tuples of points in general position, we can assume that the points in Z are the coordinate points together with the point $(1 : \cdots : 1)$. Let this specific cluster be Z_0. The stabiliser of Z_0 in $\mathrm{PGL}(n+1)$ is isomorphic to the symmetric group S_{n+2}; its preimage Γ in $\mathrm{SL}(n+1,\mathbb{C})$ acts irreducibly on $\mathbb{C}^{n+1}$. By Schur's lemma, there is a unique (up to scaling) Γ-invariant positive definite Hermitian form. It can be checked that

$$Q_0(x_0,\ldots,x_n) = \sum_{i=0}^{n}|x_i|^2 + \sum_{0\leq i<j\leq n}|x_i - x_j|^2 = (n+2)\sum_{i=0}^{n}|x_i|^2 - \Bigl|\sum_{i=0}^{n}x_i\Bigr|^2$$

is invariant under Γ, hence $z(Z_0) = Q_0$. In general, we just have to find a matrix γ such that $Z_0 \cdot \gamma^{-\top} = Z$; then

$$z(Z) = z(Z_0 \cdot \gamma^{-\top}) = Q_0 \cdot \gamma^{-\top}.$$

Note that $Z_0 \cdot \gamma^{-\top} = \sum_j P_{0,j}\gamma$ if $Z_0 = \sum_j P_{0,j}$ and we think of the $P_{0,j}$ as row vectors. So if $Z = \sum_j P_j$, then the rows of γ are coordinate vectors for the first $n+1$ points in Z, scaled in such a way that their sum is a coordinate vector for the last point.

5. Reduction of point clusters

We can now define when a point cluster is reduced.

Definition 14. Let $Z \in \mathcal{Z}_m^{\mathrm{st}}(\mathbb{R})$. We say that Z is *LLL-reduced*, resp., *Minkowski-reduced* if the positive definite real quadratic form corresponding to $z(Z)$ is LLL-reduced, resp., Minkowski-reduced.

By definition, there is an essentially unique Minkowski-reduced representative in the $\mathrm{SL}(n+1,\mathbb{Z})$-orbit of a given point cluster $Z \in \mathcal{Z}_m^{\mathrm{st}}(\mathbb{R})$. On the other hand, for computational purposes, it is usually more convenient to work with LLL-reduced representatives. In order to find an LLL-reduced representative of Z's orbit, we compute the covariant $Q = z(Z)$. Then we use the LLL algorithm [6] to find $\gamma \in \mathrm{SL}(n+1,\mathbb{Z})$ such that $Q \cdot \gamma$ is LLL-reduced. Then $Z \cdot \gamma$ is an LLL-reduced representative of the orbit of Z.

Example 15. We can use our results to reduce pencils of quadrics in three variables whose generic member is smooth. These correspond to four points in general position in $\mathbb{P}^2$. We illustrate the method with a concrete example. Let

$$\begin{aligned} Q_1(x,y,z) &= 857211194051x^2 - 10879213981695xy - 1296007209476xz \\ &\quad + 34518126244996y^2 + 8224075847095yz + 489854396055z^2, \end{aligned}$$

$$\begin{aligned} Q_2(x,y,z) &= 2274418654562x^2 - 28865567091425xy - 3438665984061xz \\ &\quad + 91586146842213y^2 + 21820750429746yz + 1299719350945z^2 \end{aligned}$$

be a pair of quadrics. We first determine a good basis of the pencil spanned by Q_1 and Q_2 by reducing the binary cubic

$$\det(xM_1 + yM_2) = 27348x^3 + 215720x^2y + 567184xy^2 + 497080y^3$$

with the approach described in [8]. Here M_1 and M_2 are the matrices of second partial derivatives of Q_1 and Q_2, respectively. This suggests the new basis

$$Q_1' = -21Q_1 + 8Q_2, \quad Q_2' = -8Q_1 + 3Q_2$$

with already somewhat smaller coefficients; the new binary cubic is

$$-4x^3 + 88x^2y + 112xy^2 - 24y^3.$$

Now we find the four points of intersection numerically. We obtain

$$P_1 = (0.3038054131 + 0.0003625989i : -0.0712511408 + 0.0000571409i : 1),$$
$$P_2 = (0.3038054131 - 0.0003625989i : -0.0712511408 - 0.0000571409i : 1),$$
$$P_3 = (0.3038639670 + 0.0003672580i : -0.0712419135 + 0.0000578751i : 1),$$
$$P_4 = (0.3038639670 - 0.0003672580i : -0.0712419135 - 0.0000578751i : 1),$$

and from this a matrix $\gamma \in \mathrm{SL}(3, \mathbb{C})$ that brings these points in standard position:

$$\gamma^{-1} = \begin{pmatrix} -13584.01 - 1762.69i & 3186.66 + 407.04i & -44719.72 - 5748.66i \\ 8318.54 + 10882.75i & -1945.84 - 2556.21i & 27338.35 + 35854.08i \\ 14176.55 + 2104.80i & -3324.73 - 486.76i & 46662.58 + 6870.37i \end{pmatrix}.$$

From this, we obtain a matrix representing $z(P_1 + P_2 + P_3 + P_4)$ as

$$\bar{\gamma} \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} \gamma^{\top} = \begin{pmatrix} 241474533625.0 & -1532325529959.9 & -182541212588.9 \\ -1532325529959.9 & 9723681808257.5 & 1158352212636.4 \\ -182541212588.9 & 1158352212636.4 & 137990925143.2 \end{pmatrix}$$

(For the actual computation, more precision is needed than indicated by the numbers above.) An LLL computation applied to this Gram matrix suggests the transformation given by

$$g = \begin{pmatrix} 3780 & 19276 & -12561 \\ -889 & -4515 & 2953 \\ 12463 & 63400 & -41405 \end{pmatrix}$$

and indeed, if we apply the corresponding substitution to Q'_1 and Q'_2, we obtain the nice and small quadrics

$$2x^2 - xy + xz + 2z^2 \quad \text{and} \quad -2xz + 3y^2 - yz + 2z^2.$$

6. Reduction of ternary forms

In this section, we apply the reduction theory of point clusters to ternary forms. The idea is to associate to a ternary form, or rather, to the plane curve it defines, a stable point cluster in a covariant way. This should be a purely geometric construction working over any base field of characteristic zero.

We will only consider irreducible ternary forms F of degree d. Assume that the curve defined by F has r nodes and no other singularities; then its genus is

$$g = \tfrac{1}{2}(d-1)(d-2) - r,$$

and by [2], Exercise IV.4.6, p. 337, the number of inflection points is

$$6(g-1) + 3d = 3d(d-2) - 6r.$$

We let $Z(F)$ be the sum of the inflection points, counted with multiplicity. When is $Z(F)$ stable? The first condition is that the multiplicity of any point must be less than $d(d-2) - 2r$. Now the multiplicity is 2 less than the order of tangency of the inflectional tangent, so it is at most $d-2$. Hence the condition is satisfied if $d-2 < d(d-2)-2r$, i.e., if $0 < (d-1)(d-2)/2 - r = g$. The second condition is that the multiplicities of points on a line add up to less than $2d(d-2)-4r$. Since there are at most d points on the curve on a line, this sum is at most $d(d-2)$. Hence the condition is satisfied if $r < d(d-2)/4$.

In any case, if F defines a nonsingular plane curve of positive genus, then $Z(F)$ is stable, and we can set $z(F) = z(Z(F))$. We then define F to be *reduced* if $z(F)$ is reduced (i.e., if $Z(F)$ is reduced).

Example 16. If F is a nonsingular cubic, then it defines a smooth curve C of genus 1, with Jacobian elliptic curve E. The 3-torsion subgroup $E[3]$ acts on C by linear automorphisms of the ambient $\mathbb{P}^2$. The preimage of $E[3]$ in $\mathrm{SL}(3,\mathbb{C})$ is a nonabelian group Γ of order 27 that acts irreducibly on $\mathbb{C}^3$. Therefore there is a unique $Q \in \mathcal{H}_{2,\mathbb{C}}$ that is invariant under the action of $E[3]$. This Q is then $z(F)$. If we know explicit matrices $M_T \in \mathrm{SL}(3,\mathbb{C})$ for $T \in E[3]$ that give the action of $E[3]$ on $\mathbb{P}^2$, then we can compute a representative of Q as a Hermitian matrix as

$$Q = \sum_{T \in E[3]} \overline{M_T}^{\top} M_T \,,$$

compare [1], §6.

We get the same result if we consider the cluster of inflection points on C, since this cluster (which is a principal homogeneous space for the action of $E[3]$) is invariant under the same group Γ. Numerically, however, the method using the action of $E[3]$ seems to be more stable. See [1], §6, for some more discussion and details.

In general, we have to find the inflection points numerically and then find the minimum of $D_{\tilde{Z}}$, also numerically. This can be done by a steepest descent method. We will illustrate this by reducing a ternary quartic.

Example 17. Let

$$\begin{aligned}F(x,y,z) = 390908548757x^4 &- 1083699236751x^3y + 835578482044x^3z \\ &+ 1126610184312x^2y^2 - 1737329379412x^2yz \\ &+ 669777678687x^2z^2 - 520542386163xy^3 \\ &+ 1204081445939xy^2z - 928398396271xyz^2 \\ &+ 238611653627xz^3 + 90192376558y^4 - 278168756247y^3z \\ &+ 321720059816y^2z^2 - 165373310794yz^3 + 31877479532z^4.\end{aligned}$$

We compute the inflection points as the intersection points of $F = 0$ and $H = 0$, where H is the Hessian of F. This gives 24 coordinate vectors and defines the point cluster $\tilde{Z}$. We then use a steepest descent method to find (an approximation to) $z(Z)$, represented by the matrix

$$\begin{pmatrix} 367751.9942 & -254909.8720 & 196557.1210 \\ -254909.8720 & 176692.9800 & -136245.3974 \\ 196557.1210 & -136245.3974 & 105056.8935 \end{pmatrix}.$$

LLL applied to this Gram matrix suggests the transformation

$$\begin{pmatrix} -7 & 23 & -89 \\ -34 & 118 & -443 \\ -31 & 110 & -408 \end{pmatrix},$$

which turns F into

$$3x^4-3x^3y+3x^3z+x^2y^2-2x^2z^2+xy^2z-xyz^2-2xz^3+3y^4-3y^3z+y^2z^2-3z^4.$$

References

[1] J. E. Cremona, T. A. Fisher, and M. Stoll, Minimisation and reduction of 2-, 3- and 4-coverings of elliptic curves. *Algebra Number Theory* **4** (2010), 763–820. Zbl 05809198 MR 2728489

[2] R. Hartshorne, *Algebraic geometry*. Grad. Texts in Math. 52, Springer-Verlag, New York 1977. Zbl 0367.14001 MR 0463157

[3] C. Hermite, Note sur la réduction des fonctions homogènes à coefficients entiers et à deux indéterminées. *J. Reine Angew. Math.* **36** (1848), 357–364; also in *Œuvres de Charles Hermite*, publiés par Émile Picard, Tome I, Gauthier-Villars, Paris 1905, 84–93.

[4] C. Hermite, Sur l'introduction des variables continues dans la théorie des nombres. *J. Reine Angew. Math.* **41** (1850), 191–216; also in: *Œuvres de Charles Hermite*, publiés par Émile Picard, Tome I, Gauthier-Villars, Paris 1905, 164–192, Sections V and VI.

[5] G. Julia, Étude sur les formes binaires non quadratiques à indéterminées réelles ou complexes, ou à indéterminées conjugées. *Mém. Acad. Sci. Inst. France* (2) **55** (1917), 1–296; also in *Œuvres de Gaston Julia*, Vol. V, Gauthier-Villars, Paris 1970, 51–342.

[6] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, Factoring polynomials with rational coefficients. *Math. Ann.* **261** (1982), 515–534. Zbl 0488.12001 MR 682664

[7] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*. Ergeb. Math. Grenzgeb. (2) 34, 3rd ed., Springer-Verlag, Berlin 1994. Zbl 0797.14004 MR 1304906

[8] M. Stoll and J. E. Cremona, On the reduction theory of binary forms. *J. Reine Angew. Math.* **565** (2003), 79–99. Zbl 1153.11317 MR 2024647

Received August 31, 2009; revised March 8, 2010

M. Stoll, Mathematisches Institut, Universität Bayreuth, Universitätsstr. 30,
95447 Bayreuth, Germany

E-mail: Michael.Stoll@uni-bayreuth.de