



# OPTIMAL TRANSPORT BETWEEN ALGEBRAIC HYPERSURFACES

PAOLO ANTONINI, FABIO CAVALLETTI, AND ANTONIO LERARIO

**Abstract.** What is the optimal way to deform a projective hypersurface into another one? In this paper we will answer this question adopting the point of view of measure theory, introducing the optimal transport problem between complex algebraic projective hypersurfaces.

First, a natural topological embedding of the space of hypersurfaces of a given degree into the space of measures on the projective space is constructed. Then, the optimal transport problem between hypersurfaces is defined through a constrained dynamical formulation, minimizing the energy of absolutely continuous curves which lie on the image of this embedding. In this way an inner Wasserstein distance on the projective space of homogeneous polynomials is introduced. This distance is finer than the Fubini–Study one.

The inner Wasserstein distance is complete and geodesic: geodesics corresponds to optimal deformations of one algebraic hypersurface into another one. Outside the discriminant this distance is induced by a smooth Riemannian metric, which is the real part of an explicit Hermitian structure. Moreover, this Hermitian structure is Kähler and the corresponding metric is of Weil–Peterson type.

To prove these results we develop new techniques, which combine complex and symplectic geometry with optimal transport, and which we expect to be relevant on their own.

We discuss applications on the regularity of the zeroes of a family of multivariate polynomials and on the condition number of polynomial systems solving.

## Contents

1	Introduction	44
2	A quick overview on optimal transport	56
3	Measure theory and algebraic geometry	62
4	Symplectic geometry and metric speed of Hamiltonian isotopies	72
5	The optimal transport problem between algebraic hypersurfaces	76
6	Absolute continuity of curves of roots	92
7	A Wasserstein approach to condition numbers	101
	Appendix: Projective geometry	105
	References	109

## 1 Introduction

What is the optimal way to deform an algebraic hypersurface in the  $n$ -dimensional projective space into another one? In this paper we will answer this question adopting the point of view of measure theory, associating to an algebraic hypersurface a probability measure on the projective space and formulating the problem with the language of *optimal transport*.

Optimal Transport is nowadays a pillar of modern mathematics. In a nutshell, it is a general tool to measure how expensive the transportation between two given probability measures  $\mu_0, \mu_1$  over a fixed metric space  $(M, \mathbf{d})$  is. This is done by a minimization procedure: for  $q \geq 1$ , the  $q$ -optimal transport problem between  $\mu_0$  and  $\mu_1$  is defined as

$$W_q(\mu_0, \mu_1) := \inf_{\xi \in \Pi(\mu_0, \mu_1)} \left( \int_{M \times M} \mathbf{d}(x, y)^q \xi(\mathrm{d}x \mathrm{d}y) \right)^{\frac{1}{q}}, \quad (1.1)$$

where  $\Pi(\mu_0, \mu_1)$  consists of the set of probability measures  $\xi$  over  $M \times M$  with marginals  $\mu_0$  and  $\mu_1$  (see Sect. 2.1). The elements of  $\Pi(\mu_0, \mu_1)$  are called *transport plans*, capturing the idea of transporting  $\mu_0$  to  $\mu_1$ . The quantity in (1.1) defines indeed a distance, the  $q$ -Wasserstein distance, on the set of Borel probability measures on  $M$ . The resulting metric space is denoted by  $\mathcal{P}_q(M)$  and called the  $q$ -Wasserstein space. When  $(M, \mathbf{d})$  is compact and geodesic,  $\mathcal{P}_q(M)$  is also compact and geodesic: in this case the optimal transport problem (1.1) has a natural interpretation as a geodesic problem.

Thanks to the generality of this formulation, during the last 30 years, Optimal Transport and the Wasserstein distance have played a central role in a large variety of different areas of mathematics. To list few notable examples we mention: PDEs [JKO98], geometric inequalities [FMP10], Random matrices [FG16] and Quantum mechanics [CM14]. For a comprehensive introduction to the subject we refer to the monographs [Vil03, Vil09]. In differential geometry, recently Optimal Transport has been at the core of the synthetic treatment of lower Ricci curvature bounds, culminated with the successful theory of J. Lott, K.-T. Sturm and C. Villani [LV09, Stu062, Stu061] completing Gromov's program of describing lower bounds on the Ricci curvature in an intrinsic manner for metric spaces.

In this paper we will take  $(M, \mathbf{d}) = (\mathbb{C}\mathbb{P}^n, \mathbf{d}_{\text{FS}})$ , the  $n$ -dimensional complex projective space with the *Fubini-Study* distance (see the Appendix). We will treat a hypersurface of degree  $d$  as a measure on  $\mathbb{C}\mathbb{P}^n$ : for a smooth hypersurface this measure is given by integrating continuous functions on it, but some care has to be done to extend this definition to singular hypersurfaces, see Sect. 1.1.1 below. In the main part of the paper we will study the intrinsic Wasserstein geometry of the space of hypersurfaces, building the tools for answering the question posed at the beginning of the paper, establishing connections between optimal transport, symplectic geometry and singularity theory. In the final part we illustrate how to apply our techniques and formalism: first we generalize to higher dimensions the result of A. Parusinski and A.

Rainer on the sharp regularity of roots of families of polynomials. We use this result to prove the compactness of the space of hypersurfaces with the inner  $q$ -Wasserstein distance for  $q \in [1, q_0]$  with  $q_0 > 1$ . Then we propose a solution to a problem posed by C. Beltrán, J.-P. Dedieu, G. Malajovich and M. Shub on the condition length of a certain class of curves.

Finally, we expect our techniques to be applicable beyond the context of complex algebraic geometry, for example in the more flexible framework of symplectic geometry.

**1.1 Optimal transport between algebraic hypersurfaces.** As a warm up we first illustrate some of the ideas in the case  $n = 1$ .

Let  $\mathbb{C}^d$  denote the space of coefficients of monic polynomials  $p(z)$  of degree  $d$  in one complex variable,  $p(z) = z^d + c_1 z^{d-1} + \dots + c_{d-1} z + c_d$ ,  $c_k \in \mathbb{C}$ . To each polynomial  $p \in \mathbb{C}^d$  we can associate in a natural way a probability measure  $\mu(p) \in \mathcal{P}(\mathbb{C})$  defined by

$$\mu(p) := \frac{1}{d} \sum_{p(z)=0} m_z(p) \cdot \delta_z, \quad (1.2)$$

where  $m_z(p) \in \mathbb{N}$  denotes the multiplicity of  $z$  as a zero of  $p$ . This defines a map

$$\mu : \mathbb{C}^d \rightarrow \mathcal{P}_2(\mathbb{C}). \quad (1.3)$$

We can think of  $\mu$  as a “solution map”: given a polynomial  $p$ , this map gives all the solutions to  $p(z) = 0$ , condensed into a single object, the measure  $\mu(p)$ . The image of  $\mu$  can be identified with  $\text{SP}^d(\mathbb{C})$ , the  $d$ -th symmetric product of  $\mathbb{C}$ , i.e. the quotient of  $\mathbb{C}^d$  under the action of the symmetric group  $S_d$  by permutations of the factors. The Fundamental Theorem of Algebra implies that  $\mu$  is in fact a homeomorphism onto its image. It turns out that the restriction of the  $W_2$ -metric to  $\text{SP}^d(\mathbb{C}) \subset \mathcal{P}(\mathbb{C})$  is given by

$$W_2([x], [y])^2 = \min_{\sigma \in S_d} \frac{1}{d} \sum_{j=1}^d |x_j - y_{\sigma(j)}|^2, \quad (1.4)$$

where  $x = (x_1, \dots, x_d)$  and  $y = (y_1, \dots, y_d)$ , with a similar expression for  $q \neq 2$ .

For every  $q \geq 1$ ,  $\text{SP}^d(\mathbb{C})$  is geodesically convex in  $\mathcal{P}_q(\mathbb{C})$ . Geodesics of  $(\text{SP}^d(\mathbb{C}), W_2)$  have a natural description: once we have the optimal matching between  $[x]$  and  $[y]$  (obtained minimizing the sum of the square of their mutual distances in  $\mathbb{C}$ ) then we let evolve the point  $x_i$  along the segment joining  $x_i$  to the optimal  $y_{j_i}$ . As we will see, these geodesics have some special behavior with respect to the *discriminant*, the set of polynomials with multiple roots. For instance if both  $p_0$  and  $p_1$  have simple roots, along the 2-Wasserstein geodesic joining  $\mu(p_0)$  with  $\mu(p_1)$  the roots remain distinct. Optimal Transport allows to make this statement quantitative (Theorem 7.1). We will come back to this result later, in the context of condition numbers of polynomial system solving.

1.1.1 *The measure associated to an algebraic hypersurface in the projective space.*

In higher dimension the situation is much more intricate. Given natural numbers  $n, d \in \mathbb{N}$ , we denote by  $H_{n,d}$  the space of complex, *homogeneous* polynomials  $p(z)$  in  $n + 1$  variables and of degree  $d$ ,

$$p(z_0, \dots, z_n) = \sum_{|\alpha|=d} c_\alpha z_0^{\alpha_0} \cdots z_n^{\alpha_n}, \quad c_\alpha \in \mathbb{C}.$$

Every nonzero polynomial  $p \in H_{n,d}$ , being homogeneous, defines a (possibly very singular) complex hypersurface  $Z(p)$  in the complex projective space  $\mathbb{CP}^n$ ,

$$Z(p) := \{[z] \in \mathbb{CP}^n \mid p(z) = 0\} \subset \mathbb{CP}^n.$$

Since nonzero multiples of the same polynomial define the same hypersurface, it is more convenient to work with the projectivization  $P_{n,d}$  of the space  $H_{n,d}$ . This is itself a complex projective space

$$P_{n,d} := \mathbb{P}(H_{n,d}) \simeq \mathbb{CP}^N, \quad N = \binom{n+d}{d} - 1.$$

Elements of  $P_{n,d}$  are therefore polynomials up to nonzero multiples but, in order to simplify the exposition, we will often call them “polynomials” and use the notation “ $p \in P_{n,d}$ ” (it would be more precise to write “ $[p] \in P_{n,d}$ ”).

With this notation, we denote by  $\Delta_{n,d} \subset P_{n,d}$  the discriminant, i.e. the set of polynomials  $p \in P_{n,d}$  such that the equation  $p = 0$  is *not* regular on  $\mathbb{CP}^n$ :

$$\Delta_{n,d} := \left\{ p \in P_{n,d} \mid \exists z \in \mathbb{C}^{n+1} \setminus \{0\}, p(z) = \frac{\partial p}{\partial z_0}(z) = \cdots = \frac{\partial p}{\partial z_n}(z) = 0 \right\}.$$

If  $p \in P_{n,d} \setminus \Delta_{n,d}$ , then  $Z(p)$  is a smooth complex submanifold of  $\mathbb{CP}^n$ , of real dimension  $2n - 2$ .

When  $n = 1$ , this construction coincides with the one from Sect. 1.1, except that we view our zeroes in the compact space  $\mathbb{CP}^1$ , rather than in  $\mathbb{C}$ . When  $n > 1$ , algebraic hypersurfaces in  $\mathbb{C}^n$  are never compact and it is more natural to work with  $\mathbb{CP}^n$  as ambient space.

As we have seen above, in the one-dimensional case we can treat the zero set of a nonzero polynomial  $p$  as a sum of Dirac’s delta measures, as in (1.2). The multi-dimensional analogue of this construction is the following.

Let us first look at the case  $p \in P_{n,d} \setminus \Delta_{n,d}$ . Restricting the Fubini–Study Riemannian metric of  $\mathbb{CP}^n$  to  $Z(p)$  turns it into an oriented Riemannian manifold, with volume form  $\text{vol}_{Z(p)}$ . We define the associated measure  $\mu(p) \in \mathcal{P}(\mathbb{CP}^n)$  to be the volume measure of the smooth submanifold  $Z(p)$ , normalized to total mass 1:

$$\mu(p) := \frac{1}{\text{vol}(Z(p))} \text{vol}_{Z(p)}. \quad (1.5)$$

Here we are tacitly agreeing that a Borel probability measure on  $Z(p)$  is also a Borel probability measure on  $\mathbb{CP}^n$ .

Notice that all smooth hypersurfaces of degree  $d$  have the same volume  $\text{vol}(Z) = d \cdot \text{vol}(\mathbb{CP}^{n-1})$  (when  $n = 1$ , this reduces to  $d \cdot \text{vol}(\mathbb{CP}^0) = d$ ), so that the renormalization factor does not depend on  $p$ ; this is a standard fact in complex algebraic geometry, see Lemma 3.7.

Hence the correspondence  $p \mapsto \mu(p)$  defines a “family” of probability measures on  $\mathbb{CP}^n$ , depending on the “parameter”  $p \in P_{n,d} \setminus \Delta_{n,d}$ . It is not difficult to show that  $\mu(p)$  depends continuously on  $p \in P_{n,d} \setminus \Delta_{n,d}$ . Less trivial is the fact that this map can be continuously extended to the whole  $P_{n,d}$ , i.e. also on the discriminant. For instance, the Hausdorff measure of  $Z(p)$  fails to be continuous even for  $n = 1$ . Moreover it is in general not true that similarly defined maps can be continuously extended also on discriminants, even in the simplest cases (see Remark 3.12).

For the continuous extension we will use the notion of multiplicity, together with an integral geometric argument. We denote by  $\mathbb{G}(1, n)$  the Grassmannian of lines in  $\mathbb{CP}^n$ , endowed with the uniform probability measure. For every  $\ell \in \mathbb{G}(1, n)$ , the restriction  $p|_\ell$  is a polynomial in one variable, and one can talk about the multiplicity  $m_z(p|_\ell)$  of its zeroes  $z \in \ell \cap Z(p)$ .

**Theorem 1.1** (The measure associated to an algebraic hypersurface). *For every  $q \geq 1$ , the map  $\mu : P_{n,d} \rightarrow \mathcal{P}_q(\mathbb{CP}^n)$  defined for  $f \in \mathcal{C}^0(\mathbb{CP}^n)$  by*

$$\int_{\mathbb{CP}^n} f d\mu(p) := \frac{1}{d} \int_{\mathbb{G}(1, n)} \left( \sum_{z \in Z(p) \cap \ell} m_z(p|_\ell) f(z) \right) \text{vol}_{\mathbb{G}(1, n)}(d\ell)$$

*is continuous and injective; restricted to  $P_{n,d} \setminus \Delta_{n,d}$ , it coincides with (1.5).*

1.1.2 *The Wasserstein distance between algebraic hypersurfaces.* At this point, using the map  $\mu$  from Theorem 1.1, we have associated to each element of  $P_{n,d}$  a unique element of  $\mathcal{P}(\mathbb{CP}^n)$ , getting in this way the  $n$ -dimensional analogue of the “solution map” (1.3),

$$\{\text{coefficients}\} = P_{n,d} \xrightarrow{\mu} \mu(P_{n,d}) = \{\text{zeroes}\} \subset \mathcal{P}(\mathbb{CP}^n).$$

Our goal now is to minimize the cost (1.1), when moving one hypersurface to another one *within* the (finite-dimensional) space of hypersurfaces of a given degree. However, unless  $n = 1$ , the image of  $\mu$  is *not* geodesically convex for the  $W_q$ -metric and, given  $p_0, p_1 \in P_{n,d}$ , the geodesic joining  $\mu(p_0)$  to  $\mu(p_1)$  in  $\mathcal{P}_q(\mathbb{CP}^n)$  will in general not stay on  $\mu(P_{n,d})$ . To overcome this issue and to formulate the optimal transport problem between complex hypersurfaces, we adopt a “constrained” version of the dynamic formulation of J.-D. Benamou and Y. Brenier [BB00].

To this end, recall first that a curve  $\mu_t : [0, 1] \rightarrow \mathcal{P}_q(\mathbb{CP}^n)$  is said to be in  $AC^q(I, \mathcal{P}_q(\mathbb{CP}^n))$  if for almost every  $t \in I$

$$\exists \lim_{h \rightarrow 0} \frac{W_q(\mu_{t+h}, \mu_t)}{|h|} =: |\dot{\mu}_t| \quad \text{and} \quad \int_I |\dot{\mu}_t|^q dt < \infty.$$

The function  $|\dot{\mu}_t|$  is called the *metric speed* of  $\mu_t$  (notice that this notion makes sense for curves with values in general metric spaces, see Sect. 2.2).

For the rest of this section we focus on the case  $q = 2$ ; similar definitions can be given for  $q \geq 1$ ,  $q \neq 2$ . With the above notation, we define the set of *admissible curves*

$$\Omega_{n,d} := \left\{ \gamma : I \rightarrow P_{n,d} \mid \mu \circ \gamma \in \text{AC}^2(I, \mathcal{P}_2(\mathbb{C}P^n)) \right\}.$$

If  $\gamma \in \Omega_{n,d}$  is an admissible curve, we write  $\mu_t := \mu(\gamma(t))$  and define the *Energy* of  $\gamma$  by

$$\mathcal{E}(\gamma) := \int_I |\dot{\mu}_t|^2 dt. \quad (1.6)$$

Given  $p_0, p_1 \in P_{n,d}$  we denote by  $\Omega_{n,d}(p_0, p_1) \subset \Omega_{n,d}$  the set of admissible curves  $\gamma : I \rightarrow P_{n,d}$  such that  $\gamma(0) = p_0$ ,  $\gamma(1) = p_1$  and we define

$$W_2^{\text{in}}(p_0, p_1) := \inf_{\gamma \in \Omega_{n,d}(p_0, p_1)} \mathcal{E}(\gamma)^{\frac{1}{2}}. \quad (1.7)$$

The superscript in (1.7) refers to the fact that  $W_2^{\text{in}}$  is the *inner* metric inherited on  $P_{n,d}$  from its embedding in  $\mathcal{P}_2(\mathbb{C}P^n)$  via Theorem 1.1, not to be confused with the restriction of the ambient  $W_2$ -metric, called the *outer* metric.

The definition of inner metrics via the minimization of the Energy of admissible curves, mimicking the familiar construction from Riemannian geometry, is quite natural in metric geometry, where it is used to build a geodesic distance on non-geodesically convex subsets [BBI01]. To obtain therefore an insight on  $W_2^{\text{in}}$  one needs to study the metric speed of admissible curves, a problem which reduces to studying the infinitesimal behaviour of the  $W_2$  distance between two algebraic hypersurfaces. For this we will need a novel combination of techniques from complex, symplectic and metric geometry, which we explain in the next section.

We note that, though being quite natural to address, the general optimal transport problem between measures supported on submanifolds in an arbitrary Riemannian manifold seems to have a short bibliography. Gangbo–McCann [GM00] considered optimal transport for measures supported on hypersurfaces in the Euclidean space; McCann–Sosio [MS11] and Kitagawa–Warren [KW12] gave more refined results about optimal transport between measures supported on a codimension one sphere in Euclidean spaces. Castillon [Cas10] considered optimal transport between a measure supported on a submanifold of Euclidean spaces and a measure supported on a linear subspace. Lott [Lot17] characterized the tangent cone (in the  $W_2$ -metric) to a probability measure supported on a smooth submanifold of a Riemannian manifold. Finally Ketterer–Mondino [KM18] linked the convexity properties of entropy functionals along  $W_2$ -geodesics concentrated on rectifiable submanifold to bounds on the partial traces of the Riemann tensor. None of these contributions constrain the paths along which the measures are allowed to move.

1.1.3 *The metric space  $(P_{n,d}, W_2^{\text{in}})$ .* To see that the infimum in (1.7) is finite for every pair of points  $p_0, p_1 \in P_{n,d}$ , and in particular that it defines a distance, requires a certain amount of work. We overview the main ideas in this section.

We will use the following notation: whenever a curve  $\gamma : I \rightarrow P_{n,d}$  will be given, we write  $p_t := \gamma(t)$ . With a slight abuse of notation, we sometimes write  $p_t : I \rightarrow P_{n,d}$  instead of  $\gamma$ .

First assume  $p_0, p_1 \notin \Delta_{n,d}$ . Since the discriminant has real codimension 2, we can find a smooth curve  $\gamma : I \rightarrow P_{n,d} \setminus \Delta_{n,d}$  joining  $p_0$  and  $p_1$ . For this curve, all the hypersurfaces  $Z(p_t)$  are embedded in  $\mathbb{C}P^n$  in the same way (they are “isotopic” in the language of differential topology). This is a consequence of the *Thom’s Isotopy Lemma*: under the assumption  $p_t \in P_{n,d} \setminus \Delta_{n,d}$ , there exists a smooth family of diffeomorphisms  $\varphi_t : \mathbb{C}P^n \rightarrow \mathbb{C}P^n$  such that  $\varphi_t(Z(p_0)) = Z(p_t)$  for every  $t \in I$ . There is also a Hamiltonian version of Thom’s Isotopy Lemma, which in our case takes the following form (variations of this result can be found in [Sil06, Aur97, ST05], for a self-contained detailed proof see [ACL23]).

**Theorem 1.2** (Hamiltonian Thom’s Isotopy Lemma). *Let  $\gamma : I \rightarrow P_{n,d} \setminus \Delta_{n,d}$  be a smooth curve. Then there exists a Hamiltonian flow  $\varphi_t : \mathbb{C}P^n \rightarrow \mathbb{C}P^n$  such that  $\varphi_t(Z(p_0)) = Z(p_t)$  for every  $t \in I$ .*

Using Theorem 1.2, we can write the measures  $\mu_t := \mu(p_t)$  as  $\mu_t = (\varphi_t)_\# \mu_0$ . In general, if  $v_t \in L^2(\mu_t; TM)$  is a time-dependent vector field solving (in the sense of distributions) the continuity equation

$$\partial_t \mu_t + \text{div}(v_t \mu_t) = 0, \quad (1.8)$$

then, for almost every  $t \in I$ , the metric speed of  $\mu_t$  equals:

$$|\dot{\mu}_t| = \|\mathbf{P}(v_t)\|_{L^2_{\mu_t}},$$

where  $\mathbf{P} : L^2(\mu; TM) \rightarrow T_\mu(\mathcal{P}_2(M))$  denotes the projection on the “space of gradients” (see Sect. 2.3). In our case the crucial point is that, since  $v_t$  is Hamiltonian, for every  $z \in Z(p_t)$  we have  $\mathbf{P}(v_t)(z) = N(v_t(z))$ , where  $N$  denotes the orthogonal projection of the field on the normal bundle of  $Z(p_t)$ , see Proposition 4.5.

Summing up: if  $\gamma : I \rightarrow P_{n,d} \setminus \Delta_{n,d}$  is a smooth curve and  $v_t$  is the time-dependent vector field generating the Hamiltonian isotopy from Theorem 1.2, the Energy of the curve can be written as

$$\mathcal{E}(\gamma) = \int_I |\dot{\mu}_t|^2 = \int_I \int_{\mathbb{C}P^n} \|N(v_t)\|^2 d\mu_t dt.$$

There are plenty of Hamiltonian fields  $v_t$  solving the continuity equation (1.8), but the condition  $p_t \in P_{n,d} \setminus \Delta_{n,d}$  determines the normal component of such fields and allows to prove the following theorem.

**Theorem 1.3** (The energy of an admissible curve avoiding the discriminant). *Let  $\gamma : I \rightarrow P_{n,d} \setminus \Delta_{n,d}$  be a  $\mathcal{C}^1$  curve. Then  $\gamma$  is admissible and*

$$\mathcal{E}(\gamma) = \frac{1}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} \int_I \int_{Z(p_t)} \frac{|\dot{p}_t(z)|^2}{\|\nabla^{\mathbb{C}} p_t(z)\|^2} \text{vol}_{Z(p_t)}(dz) dt. \quad (1.9)$$

Here the innermost integrand in (1.9) has to be thought as a function on the zero set of  $p_t$  in the unit sphere  $S^{2n+1}$ ; this function is invariant under  $U(1)$  and therefore it is well defined on the zero set of  $p_t$  in  $\mathbb{C}P^n$ . (The same remark applies to similar integrals below, see Remark 5.4.)

We observe two interesting consequences of the previous result, which are peculiar of the case  $q = 2$ . First, the innermost integral in (1.9) defines a quadratic form on the tangent spaces to  $P_{n,d} \setminus \Delta_{n,d}$ . This is induced by the real part of a Hermitian form  $h$ , which we call the *Wasserstein–Hermitian structure*:

$$h_p(q_1, q_2) := \int_{Z(p)} \frac{q_1(z) \overline{q_2(z)}}{\|\nabla^{\mathbb{C}} p(z)\|^2} \text{vol}_{Z(p)}(dz), \quad q_1, q_2 \in T_p(P_{n,d} \setminus \Delta_{n,d}). \quad (1.10)$$

In particular, the Energy of a  $\mathcal{C}^1$  curve  $\gamma : I \rightarrow P_{n,d} \setminus \Delta_{n,d}$  can be written in the usual Riemannian way as:

$$\mathcal{E}(\gamma) = \frac{1}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} \int_I h_{p_t}(\dot{p}_t, \dot{p}_t) dt.$$

Notice that  $p \in P_{n,d}$  denotes a class of polynomials in  $H_{n,d}$  up to nonzero multiples, but the above definition does not depend on the choice of representatives, see Lemma 5.5 and Sect. 5.1 for more details. Second, this Hermitian structure is not defined on  $\Delta_{n,d}$  (this is because the integrand in (1.10) is not defined when  $p \in \Delta_{n,d}$ ), but its explicit form allows us to perform a semialgebraic reparametrization argument for curves with endpoints on  $\Delta_{n,d}$  (Theorem 5.13), and to prove that any two points  $q_0, q_1$ , regardless being on the discriminant or not, can be joined by an admissible curve with *finite* Energy (Corollary 5.14). In particular, the minimization problem (1.6) always admits a solution:

$$W_2^{\text{in}}(p_0, p_1) = \min_{\gamma \in \Omega_{p_0, p_1}} \mathcal{E}(\gamma)^{\frac{1}{2}}.$$

**Theorem 1.4** (The Wasserstein space of hypersurfaces). *The function  $W_2^{\text{in}}$  defines a distance on  $P_{n,d}$  and the metric space  $(P_{n,d}, W_2^{\text{in}})$  is complete and geodesic, i.e. for any couple  $p_0, p_1$ , there exists an admissible curve  $\gamma$  from  $p_0$  to  $p_1$  such that  $W_2^{\text{in}}(\gamma(s), \gamma(t)) = |t - s| W_2^{\text{in}}(p_0, p_1)$ .*

Notice that  $P_{n,d}$  is compact for the ambient  $W_2$ -metric, by Theorem 1.1, but the compactness of  $(P_{n,d}, W_2^{\text{in}})$  is still an open problem see [PR23, Open problem 6.11]. On the other hand, denoting by  $d_{\text{FS}}$  the metric induced by the Fubini–Study structure on  $P_{n,d} \simeq \mathbb{C}P^N$ , the compactness of  $(P_{n,d}, W_2^{\text{in}})$  is equivalent to the continuity of

$$\text{id} : (P_{n,d}, d_{\text{FS}}) \longrightarrow (P_{n,d}, W_2^{\text{in}}).$$

It is typically hard to predict compactness for the inner metric of a compact set in the ambient topology (and in fact, in general, not even true for very simple examples, see [BBI01, Fig. 2.2]). However, in a fashion similar to the one-dimensional case, we prove that the above map is continuous (in fact Lipschitz) “away” from the discriminant, i.e. where we have the Riemannian structure (1.10). More precisely, setting

$$P_{n,d}(\epsilon) := \{p \in P_{n,d} \mid d_{\text{FS}}(p, \Delta_{n,d}) \geq \epsilon\},$$

we prove the following result.

**Theorem 1.5** (Wasserstein and Fubini–Study). *For every  $\epsilon > 0$  the identity map*

$$\text{id} : (P_{n,d}(\epsilon), d_{\text{FS}}) \longrightarrow (P_{n,d}(\epsilon), W_2^{\text{in}})$$

*is Lipschitz. In particular  $P_{n,d}(\epsilon)$  is compact in the  $W_2^{\text{in}}$ -topology.*

It turns out that, for  $\delta(n) > 0$  small enough and for  $1 \leq q < 1 + \delta(n)$ , the space  $(P_{n,d}, W_q^{\text{in}})$ , is compact (see Theorem 1.9 below). The proof of this result requires some additional work and for it we will need higher dimensional versions of results on the regularity of roots of polynomials. We will discuss this in Sect. 1.3.

## 1.2 The Kähler structure and a Weil–Peterson type metric.

Quite remarkably on  $P_{n,d} \setminus \Delta_{n,d}$  the Wasserstein–Hermitian structure from (1.10) is Kähler and the corresponding metric is a Weil–Peterson type metric. More precisely, consider the solution variety  $V := \{(p, z) \mid p(z) = 0\} \subset P_{n,d} \times \mathbb{C}P^n$  together with the restrictions of the projections on the two factors:

$$\begin{array}{ccc} & V & \\ \pi_1 \swarrow & & \searrow \pi_2 \\ P_{n,d} & & \mathbb{C}P^n \end{array}$$

From this diagram we have two induced maps at the level of  $k$ -forms: the usual pull-back  $\pi_2^* : \Omega^k(\mathbb{C}P^n) \rightarrow \Omega^k(V)$  and the integration along the fibers

$$(\pi_1)_* : \Omega^k(V \setminus \text{crit}(\pi_1)) \rightarrow \Omega^{k-2n-2}(P_{n,d} \setminus \Delta_{n,d}).$$

On  $\mathbb{C}P^n$  we have the Fubini–Study form  $\omega_{\text{FS}} \in \Omega^{1,1}(\mathbb{C}P^n) \subset \Omega^2(\mathbb{C}P^n)$ , which we can pull-back to  $V$  by  $\pi_2$ , and the integral along the fibers of  $\pi_2^* \omega_{\text{FS}}^n / n!$  (which is the volume form of  $\mathbb{C}P^n$ ) turns out to be the imaginary part of  $\mathfrak{h}$ .

**Theorem 1.6** (The Wasserstein–Kähler structure). *Let  $\mathfrak{h}$  be the Wasserstein–Hermitian structure (1.10) and let  $\sigma := \text{Im}(\mathfrak{h}) \in \Omega^2(P_{n,d} \setminus \Delta_{n,d})$ . Then*

$$\sigma = (\pi_1)_* \pi_2^* \left( \frac{\omega_{\text{FS}}^n}{n!} \right) \in \Omega^{1,1}(P_{n,d} \setminus \Delta_{n,d}). \quad (1.11)$$

*In particular  $(P_{n,d} \setminus \Delta_{n,d}, \mathfrak{h})$  is a Kähler manifold.*

The Kähler metrics on the parameter space of a family of algebraic varieties which can be obtained as pushforwards of objects on the total space of the family (as in (1.11)) are recurring objects in complex algebraic geometry, and they are usually called *Weil–Petersson type metrics*. For instance, the Weil–Petersson metric on the moduli space of curves can be obtained in this way, and a version for a family of Calabi–Yau varieties appears for instance in [ST12, Tos10].

**1.3 The regularity of a curve of measures and a compactness result.** Let us go back to the metric space  $(P_{n,d}, W_q^{\text{in}})$ . In order to prove its compactness for  $1 \leq q \leq 1 + \delta(n)$  we will use a higher dimensional generalization of the results from [PR18]. We first discuss this generalization, which is interesting in its own, as it addresses and solves a natural question on the regularity of zeroes of families of *multivariate* polynomials. We then explain how it can be used to prove the compactness of the  $q$ -Wasserstein inner metric.

Given a curve of polynomials  $p_t : I \rightarrow P_{n,d}$ , it is natural to ask what regularity we can expect for the curve of measures  $\mu(p_t)$ . We go back to the case  $n = 1$ . By the homeomorphism

$$\{\text{coefficients}\} = \mathbb{C}^d \xrightarrow{\mu} \text{SP}^d(\mathbb{C}) = \{\text{zeroes}\},$$

it follows that the roots of a polynomial depend continuously on its coefficients but, in general, no more regularity can be assumed. We can explain this phenomenon by looking at the previous homeomorphism from a metric point of view, endowing  $\text{SP}^d(\mathbb{C})$  with the Wasserstein distance, as in (1.4). The loss of regularity of the roots of a family of polynomials is due to the fact that the above map  $\mu$ , seen as a map between metric spaces, is *not* Lipschitz. This failure of regularity happens precisely on the discriminant, where we have polynomials with multiple roots. For instance, the zeroes  $e^{\frac{i2\pi k}{d}} t^{\frac{1}{d}}$  of the polynomial  $z^d - t$  involve a  $d$ -th square root and cannot be Lipschitz functions of  $t$  near  $t = 0$ . (Notice that, in this case,  $\mu^{-1} : \text{SP}^d(\mathbb{C}) \rightarrow \mathbb{C}^d$  is instead Lipschitz, as the coefficients of a polynomial depends smoothly on its zeroes.)

Still, A. Parusinski and A. Rainer in a series of works culminated with [PR18] have recently obtained the sharp result that the zeroes of a curve of polynomials (for  $n = 1$ ) enjoy a Sobolev regularity: in particular any continuous choice of a root of  $p_t$  has first derivative in  $L^q$  for  $1 \leq q < \frac{d}{d-1}$ , provided  $p_t : I \rightarrow \mathbb{C}^d$  is a  $\mathcal{C}^{d-1,1}$  curve.

In our formalism the results from [PR18] are equivalently stated by saying that if  $p_t : I \rightarrow \mathbb{C}^d$  is a  $\mathcal{C}^{d-1,1}$  curve, then  $\mu_t := \mu(p_t)$  is in  $AC^q(I, \mathcal{P}_q(\mathbb{C}))$  for every  $1 \leq q < \frac{d}{d-1}$ . It is then natural to adopt this formalism to address the same question in the more general setting of  $n \geq 2$ ; to the best of our knowledge no other approach is at disposal to address this question.

Quite remarkably, in higher dimensions, the existence of rational curves on the zero set of some of the polynomials in the curve  $p_t$  (a very classical problem in algebraic geometry!) seems to be an obstruction for the regularity of  $\mu(p_t)$ . In this direction we prove two results, the first being a direct generalization of [PR18, Theorem 1]. It requires the same regularity as [PR18, Theorem 1] on the curve  $p_t$ , but

at the same time it needs the curve to miss (what we call) the *Fano discriminant* (see Definition 6.2). It can be thought of as the set  $F_{n,d} \subset P_{n,d}$  consisting of polynomials  $p$  for which the set of lines on  $Z(p)$  is not defined by regular equations on the Grassmannian  $\mathbb{G}(1, n)$ .

The Fano discriminant is contained in a proper algebraic set of  $P_{n,d}$ , see [EH16, Theorem 6.34]. When  $n = 1$ , it is empty: in this case the only line is  $\mathbb{C}P^1$  itself, where no nonzero polynomial can vanish identically. In general, if  $d > 2n - 3$  and  $p \notin F_{n,d}$ , then no line can be found on  $Z(p)$  and one can use integral geometry to perform a coupling between measures “uniformly” on the set of lines. Notice that, when  $n = 1$ , the conditions  $d > 2n - 3$  and  $p_t \notin F_{n,d}$  are empty, and the statement below reduces to [PR18, Theorem 1].

**Theorem 1.7** (Regularity of curves of measures I). *Let  $d > 2n - 3$ . Let  $p_t \in \mathcal{C}^{d-1,1}(I, P_{n,d})$  be a curve such that  $p_t \notin F_{n,d}$  for every  $t \in I$ . Then, for every  $1 \leq q < d/(d-1)$  the curve  $\mu_t := \mu(p_t) \in AC^q(I, \mathcal{P}_q(\mathbb{C}P^n))$ .*

The second regularity result that we prove allows to drop the restriction  $p_t \notin F_{n,d}$ , but requires to assume more regularity on the curve  $p_t$ , at the price of losing some regularity on the curve of measures  $\mu_t$ .

**Theorem 1.8** (Regularity of curves of measures II). *For every  $n \in \mathbb{N}$  there exists  $e(n) > 0$  such that for every curve  $p_t \in \mathcal{C}^{k,1}(I, P_{n,d})$ , with  $k \geq e(n)d - 1$ , and for every  $1 \leq q < e(n)d/(e(n)d - 1)$  the curve  $\mu_t := \mu(p_t) \in AC^q(I, \mathcal{P}_q(\mathbb{C}P^n))$ .*

The main idea for the proof of Theorem 6.5 is to replace the use of lines, in the integral geometry approach, with the use of rational curves, see Proposition 6.7. The crucial point for our argument is to find at least one rational curve, possibly of very high degree, with the property that neither this curve nor any of its translates under the natural unitary group action lie on any of the hypersurfaces  $Z(p_t)$ ,  $t \in I$ . The existence of such a rational curve follows from a recent result of B. Lehmann, E. Riedl and S. Tanimoto, privately communicated to the authors (Proposition 6.8). The constant  $e(n)$  in Theorem 1.8 is the degree of this rational curve, and is related to the so-called Fujita invariants.

Using Theorem 1.8, we can finally prove the following compactness result.

**Theorem 1.9** (Compactness). *For every  $1 \leq q < 1 + \frac{1}{e(n)d-1}$ , the metric space  $(P_{n,d}, W_q^{\text{in}})$  is compact.*

**1.4 A Wasserstein approach to condition numbers.** In an influential sequence of papers, the “Complexity of Bézout’s Theorem” series [SS931, SS932, SS93, SS96, SS94], M. Shub and S. Smale opened to a geometric investigation of the complexity of polynomial systems solving. This line of research has received a lot of attention, and culminated with the positive solution of Smale’s 17th Problem [Sma00]: “Can a zero of  $n$  complex polynomial equations in  $n$  unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?”. The solution to this

problem follows from a combination of works by C. Beltrán and L. M. Pardo [BP09], M. Shub [Shu09], P. Bürgisser and F. Cucker [BC11] and finally by P. Lairez [Lai17].

Here we propose a conceptual application of our approach: to go back to the original question of Smale, reinterpreting it in the language of Optimal Transport. We do not venture into this full program, which would be beyond the purposes of this paper. Nonetheless, we will show how our point of view already gives a solution to [BCC13, Problem 14] in the case  $n = 1$  (see Theorem 1.10 below).

Let us explain the context behind [BCC13, Problem 14]. One of the main ideas underlying the algorithm solving Smale’s problem is *continuation along linear paths*: given a polynomial system  $p_1 = 0$ , which is assumed to be regular, one builds another regular polynomial system  $p_0$  of which one solution  $z_0$  is known, and then “continues” this solution along the linear homotopy  $p_t := (1 - t)p_0 + tp_1$ . The fact that one can choose  $p_0$  so that the linear path  $p_t$  does not hit the discriminant follows from dimensional considerations (the discriminant has real codimension two); the complexity of the algorithm is related to the possibility of estimating the distance to the discriminant throughout the path  $p_t$ . In order to measure this distance, one introduces the notion of *normalized condition number* of a system at one of its solutions. In the case of a single polynomial  $p \in P_{1,d}$  at one of its zeroes  $z \in \mathbb{CP}^1$ , this is the quantity:

$$\nu_{\text{norm}}(p, z) := d_{\text{BW}}(p, \Delta_z)^{-1},$$

where  $\Delta_z \subset \Delta_{1,d}$  denotes the set of polynomials with a multiple root at  $z$  (see Sect. 7 for more details), and  $d_{\text{BW}}$  denotes the Bombieri–Weyl distance on the projective space of polynomials (see the Appendix).

In order to address Smale’s problem, prior to its solution, M. Shub [Shu09] suggested the possibility of changing the class of paths, switching from linear paths to special homotopies  $(p_t, z_t)$ , with  $p_t(z_t) \equiv 0$ , called *condition geodesics*. In order to describe them, in the case  $n = 1$ , let us introduce the *solution variety*  $V := \{(p, z) \in P_{1,d} \times \mathbb{CP}^1 \mid p(z) = 0\}$  together with the projection  $\pi_1 : V \rightarrow P_{1,d}$  on the first factor. Denoting by  $\Sigma'$  the set of critical points of  $\pi_1$ , the normalized condition number can be used to turn  $V \setminus \Sigma'$  into a *Lipschitz* Riemannian manifold as follows. First, we denote by  $g_V$  the Riemannian structure on  $V \subset P_{1,d} \times \mathbb{CP}^1$  induced by the product Riemannian structure. Then, for every absolutely continuous curve  $\gamma_t = (p_t, z_t) : I \rightarrow V \setminus \Sigma'$  we define its *condition length* by

$$L_{\text{cond}}(\gamma) := \int_I \nu_{\text{norm}}(p_t, z_t) \cdot \|\dot{\gamma}_t\|_V dt.$$

Taking the infimum of the length over all absolutely continuous curves joining two points endows  $V \setminus \Sigma'$  with a metric structure, called the *condition metric*. A condition geodesic is a curve  $\gamma_t$  in  $V \setminus \Sigma'$  minimizing the condition length between any two of its points.

A similar construction can be introduced in the case of polynomial systems, and in [Shu09] it is shown that  $O(d^{3/2} L_{\text{cond}}(\gamma))$  Newton steps are sufficient to con-

tinue the zero  $z_0$  from  $p_0$  to  $p_1$  along  $\gamma_t$ . Following this result, C. Beltrán, J.-P. Dedieu, G. Malajovich and M. Shub, in [B+12, B+09], asked whether the function  $t \mapsto \log(\nu_{\text{norm}}(\gamma_t))$  is convex along a condition geodesic. An affirmative answer to this question would imply that, for any condition geodesic  $\gamma_t$ , one has  $L_{\text{cond}}(\gamma_t) \leq L_V(\gamma_t) \max\{\nu_{\text{norm}}(p_0, z_0), \nu_{\text{norm}}(p_1, z_1)\}$ , where  $L_V(\gamma_t)$  is the length of  $\gamma_t$  in the Riemannian metric of  $V$  defined above. In particular, it would be possible to estimate the distance from the discriminant along the whole path  $p_t$  by only knowing this distance at the starting and the ending system. For more details on this problem, we refer the reader to [BCC13, Problem 14], from where this formulation is taken.

Here we propose a solution to this problem in the case  $n = 1$ , changing the class of paths, and replacing *condition geodesics* with lifts of *Wasserstein geodesics*. We expect a similar result to be true in the case  $n > 1$ , but we leave this question to future investigation (this will require to work with the  $W_2^{\text{in}}$  metric on the space of polynomial systems).

**Theorem 1.10** (The condition length of Wasserstein geodesics). *For every  $d \in \mathbb{N}$  there exist  $\beta_1, \beta_2, \beta_3, \beta_4 > 0$  such that, if  $p_t : I \rightarrow P_{1,d}$  is a  $W_2$ -geodesic between  $p_0, p_1 \in P_{1,d} \setminus \Delta_{1,d}$  then the following is true:*

(1) for every  $t \in I$

$$\text{dist}(p_t, \Delta_{1,d})^2 \geq \beta_1 \min\{\text{dist}(p_0, \Delta_{1,d}), \text{dist}(p_1, \Delta_{1,d})\}^{\beta_2}.$$

*In particular,  $P_{1,d} \setminus \Delta_{1,d}$  is  $W_2$ -geodesically convex;*

(2) if  $\gamma : I \rightarrow V \setminus \Sigma'$  is a lift of  $p_t$ , i.e.  $\gamma(t) = (p_t, z_t)$ , then:

$$L_{\text{cond}}(\gamma) \leq \beta_3 \cdot L_V(\gamma) \cdot \max\{\nu_{\text{norm}}(p_0, z_0), \nu_{\text{norm}}(p_1, z_1)\}^{\beta_4}.$$

The proof of this result is based on the structure of Wasserstein geodesics, see for instance the section on Monge–Mather shortening principle in [Vil09], which guarantees a quantitative separation of roots along such curves (Proposition 7.3). (The constants  $\beta_i$  from the statement come from the use of Lojasiewicz’s inequality, and we expect that actually  $\beta_2 = \beta_4 = 1$ .)

We remark that recently P. Lairez [Lai20] has proposed to consider yet another class of paths, called *rigid paths*, that have a quasi-optimal polynomial behavior in terms of the complexity of Smale’s 17th Problem. We wonder what could be the behavior of an algorithm based on our Wasserstein approach. In this direction, we observe that, compared to the condition metric, which is only Lipschitz, the  $W_2^{\text{in}}$  metric, away from the discriminant, comes from a smooth Riemannian metric (Proposition 5.7); in particular Wasserstein geodesics can be found by means of a smooth Riemannian optimization problem.

**1.5 Structure of the paper.** Section 2 recalls basic facts about the optimal transport problem and Wasserstein distances on Riemannian manifolds reserving special care to the dynamical formulation à la Benamou–Brenier. The section also covers some needed tools from metric geometry.

Section 3 is the heart of the construction of the correspondence  $\mu$  (in all its incarnations) associating a measure to a polynomial. This is the map that appears in Theorem 1.1 before, proved in this section (this is precisely Theorem 3.8). We begin with a careful study of the zero dimensional case (illustrated in the introduction) then, using integral geometry tools, we move to the higher dimensional case.

In Sect. 4, we use instruments from symplectic geometry to compute the metric speed of a curve of polynomials that stays away from the discriminant. In particular this is a consequence of Theorem 1.2.

In Sect. 5, the optimal transport problem for hypersurfaces in the form of zero sets of polynomials is introduced and studied. We define the notion of admissible curves, energy and inner metric. Here we prove Theorem 1.3 (corresponding to Theorem 5.3), we define the Wasserstein–Hermitian structure and prove Theorem 1.4 (Theorem 5.15 in the body of the paper) and Theorem 1.5 (corresponding to Theorem 5.17) of comparison between the Wasserstein and Fubini–Study metric. In Sect. 5.2 we prove that the Wasserstein–Hermitian metric is Kähler (Theorem 5.10 in the body of the paper).

In Sect. 6 we study the problem of the regularity of curves of solutions of polynomials in dimension greater than one. We prove here Theorem 1.7 (Theorem 6.3 in the body) and Theorem 1.8 (Theorem 6.5 in the body). Theorem 1.9 corresponds to Theorem 6.11 in the body of the paper. In Sect. 7 a Wasserstein approach to condition number is discussed; here we prove Theorem 1.10 (Theorem 7.1 in the body).

## 2 A quick overview on optimal transport

In order to introduce an optimal transport distance over the space of homogeneous polynomials, a careful analysis on the infinitesimal behaviour of the Wasserstein distance near a general probability measure is required. In this section we will therefore provide a quick introduction to Optimal Transport and review the needed material for the rest of the paper. We refer to the monographs [Vil03, Vil09] and references therein for a complete account on the theory.

**2.1 Wasserstein distance.** Let  $(M, g)$  be a smooth compact Riemannian manifold. We will denote by  $\mathcal{P}(M)$  the space of Borel probability measures over  $M$ . On  $\mathcal{P}(M)$  we will consider the following distances induced by Optimal Transport.

**DEFINITION 2.1** ( $q$ -Wasserstein distance). Fix any  $q \geq 1$ . For any  $\mu_0, \mu_1 \in \mathcal{P}(M)$ , the  $q$ -Wasserstein distance is defined by

$$W_q(\mu_0, \mu_1) := \left( \inf_{\xi \in \Pi(\mu_0, \mu_1)} \int_{M \times M} d_g(x, y)^q \xi(dxdy) \right)^{\frac{1}{q}},$$

where  $d_g$  is the geodesic distance induced by the Riemannian metric  $g$  (making  $(M, d_g)$  a compact metric space) and

$$\Pi(\mu_0, \mu_1) := \left\{ \xi \in \mathcal{P}(M \times M) : (\pi_1)_\# \xi = \mu_0, (\pi_2)_\# \xi = \mu_1 \right\}$$

is the (convex) set of *transport plans* between  $\mu_0$  and  $\mu_1$ ; finally  $\pi_i : M \times M \rightarrow M$  denotes the projection on the  $i$ -th component, for  $i = 1, 2$ .

Roughly, transport plans represent the way in which  $\mu_0$  is moved to  $\mu_1$  while  $d_g^q$  is the chosen cost function for transporting an infinitesimal amount of mass. The subset of those transport plans realising the infimum in the definition of the  $W_q$  distance are called optimal transport plans and are denoted by  $\Pi_{opt}^q(\mu_0, \mu_1)$ .

As we will always stick to the case of compact manifolds,  $W_q$  will be well defined between any couple of elements of  $\mathcal{P}(M)$  with a non-empty  $\Pi_{opt}^q(\mu_0, \mu_1)$ . In particular  $(\mathcal{P}(M), W_q)$  is a complete and separable metric space; from the compactness of  $M$  it follows moreover that it is compact as well. For ease of notation we will sometimes write  $\mathcal{P}_q(M)$  to intend the metric space  $(\mathcal{P}(M), W_q)$ .

**2.2 Length structures and inner distances.** Here we follow [BBI01] to which we refer for the missing details. If  $(X, d)$  is a metric space one can define a length structure in the following way. Firstly consider a class of admissible curves.

**DEFINITION 2.2** (Absolutely continuous curves with values in a metric space). Let  $(X, d)$  be a metric space and let  $I$  be a real interval. We say that a curve  $\gamma : I \rightarrow X$  belongs to  $AC^q(I; X)$ ,  $q \geq 1$ , if there exists  $m \in L^q(I)$  such that

$$d(\gamma_s, \gamma_t) \leq \int_s^t m(\tau) d\tau, \quad \forall s \leq t. \quad (2.1)$$

When  $q = 1$ , we simply denote  $AC = AC^1$ .

**Theorem 2.3** (Theorem 1.1.2 of [AGS05]). *If  $\gamma \in AC^q(I; X)$ ,  $q \geq 1$ , then for  $\mathcal{L}^1$ -a.e.  $t \in I$  there exists the limit*

$$\lim_{h \rightarrow 0} \frac{d(\gamma_{t+h}, \gamma_t)}{|h|}.$$

*We denote the value of this limit by  $|\dot{\gamma}|(t)$  and we call it metric derivative, or metric speed, of  $\gamma$  at  $t$ . The function  $t \mapsto |\dot{\gamma}|(t)$  belongs to  $L^q(I)$  and*

$$d(u(s), u(t)) \leq \int_s^t |\dot{\gamma}|(\tau) d\tau, \quad \forall s \leq t.$$

*Moreover  $|\dot{\gamma}|(t) \leq m(t)$  for  $\mathcal{L}^1$ -a.e.  $t \in I$ , for each  $m$  such that (2.1) holds.*

The space of absolutely continuous curves  $AC^q(I; X)$  can be seen as a subset of the metric space  $\mathcal{C}^0(I; X)$  (endowed with the natural sup distance); recall that  $\mathcal{C}^0(I; X)$  is complete and separable provided the same is true for  $X$ .

One then defines the length of absolutely continuous curves, that we can assume without loss of generality to be defined on  $I = [0, 1]$ , by

$$L(\gamma) = \int_0^1 |\dot{\gamma}|(t) dt.$$

The length  $L$  induces an “intrinsic” distance, called *inner distance*, by minimizing the length functional among all the sets of admissible curves:

$$\widehat{d}(x, y) = \inf \left\{ L(\gamma) : \gamma \in AC([0, 1]; X), \gamma_0 = x, \gamma_1 = y \right\}.$$

If  $\widehat{d} = d$ , then  $(X, d)$  is called a length space. It is called a geodesic space if the infimum is attained by at least one admissible curve, hereafter called a geodesic.

Our analysis will naturally bring us to study non-geodesically convex subsets of a geodesic space  $X$ . For our scope it will be sufficient to assume  $(X, d)$  to be complete and separable and  $Y \subset X$  to be a closed set so that  $(Y, d)$  is still complete and separable. Clearly if  $Y$  fails to be geodesically convex then  $(Y, d)$  will not be anymore a geodesic space. Nonetheless one can consider the length structure induced by  $L$  over  $Y$  and define the inner distance of  $Y$  as follows:

$$d_Y^{\text{in}}(x, y) := \inf \left\{ L(\gamma) : \gamma \in AC([0, 1]; Y), \gamma_0 = x, \gamma_1 = y \right\}. \quad (2.2)$$

If for any couple  $x, y \in Y$  there exists an admissible curve  $\gamma \in AC([0, 1]; Y)$  of finite length  $L(\gamma) < \infty$ , then  $d_Y^{\text{in}}$  is a distance making  $(Y, d_Y^{\text{in}})$  a length space.

Through a standard reparametrization argument, see for instance [AGS05, Lemma 1.1.4], one notices that for any  $q > 1$

$$d_Y^{\text{in}}(x, y)^q = \inf \left\{ \int_0^1 |\dot{\gamma}|^q(t) dt : \gamma \in AC^q([0, 1]; Y), \gamma_0 = x, \gamma_1 = y \right\}.$$

Finally we mention that, defining the energy functional  $\mathcal{E} : \mathcal{C}^0(I; Y) \rightarrow [0; \infty]$  by

$$\mathcal{E}(\gamma) := \begin{cases} \int_0^1 |\dot{\gamma}|^2(t) & u \in AC^2(I; Y); \\ +\infty, & \text{otherwise,} \end{cases}$$

then  $\mathcal{E}$  is lower semicontinuous (see for instance [GP20, Proposition 1.2.7]).

### 2.3 Absolutely continuous curves, the continuity equation and the space of gradients.

We continue the short review of Optimal Transport needed for this paper. Since  $(M, d_g)$  is a geodesic space, the  $q$ -Wasserstein space  $\mathcal{P}_q(M)$  is a geodesic space as well (see [Stu062, Proposition 2.10]): for any  $\mu_0, \mu_1 \in \mathcal{P}(M)$  there exists a curve  $(\mu_t)_{t \in [0, 1]} \subset \mathcal{P}(M)$  from  $\mu_0$  to  $\mu_1$  such that

$$W_q(\mu_t, \mu_s) = |t - s| W_q(\mu_1, \mu_0).$$

Wasserstein geodesics are strongly linked to the geometry of the underlying space  $M$ . Denote by  $\text{Geo}(M)$  the set of minimal geodesics of  $M$  seen as a closed subspace of

$\mathcal{C}^0([0, 1]; M)$ ; then in the strictly convex case ( $q > 1$ )  $W_q$ -geodesics have the following lifting property (see [Stu062, Proposition 2.10]): for each  $W_q$ -geodesic  $(\mu_t)_{t \in [0, 1]}$  there exists a probability measure  $\eta$  on  $\text{Geo}(M)$ , such that

$$(e_t)_\# \eta = \mu_t, \quad \forall s, t \in [0, 1] \quad (e_s, e_t)_\# \eta \in \Pi_{opt}^q(\mu_s, \mu_t),$$

where  $e_t : \mathcal{C}^0([0, 1]; M) \rightarrow M$  is the evaluation map ( $e_t(f) := f(t)$ ). The set of  $\eta \in \mathcal{P}(\text{Geo}(M))$  verifying the previous conditions is called the set of dynamical  $q$ -optimal plans and is denoted by  $\text{OptGeo}_q(\mu_0, \mu_1)$ .

As discussed, our interest will be in studying the inner distance induced by  $W_q$  on a distinguished closed subset of  $\mathcal{P}(M)$  and to define its inner Wasserstein distance as in (2.2). It will be therefore necessary to discuss the metric speed of curves in  $AC^q([0, 1]; \mathcal{P}_q(M))$ . As for  $W_q$ -geodesics, also  $q$ -absolutely continuous curves  $\mathcal{P}_q(M)$  have a geometric interpretation in terms of family of curves of  $M$  firstly pioneered by J.-D. Benamou and Y. Brenier [BB00].

If  $\mu_t : I \rightarrow \mathcal{P}_q(M)$  is a  $q$ -absolutely continuous curve with  $I \subset \mathbb{R}$  an interval, there exists a Borel vector field  $v : (x, t) \mapsto v_t(x) \in T_x M$  such that  $\mu_t$  solves the continuity equation with vector field  $v_t$ :

$$\partial_t \mu_t + \text{div}(v_t \mu_t) = 0, \tag{2.3}$$

in the sense of distributions (see (2.5) below). One can check that (2.3) implies that  $|\dot{\mu}_t| \leq \|v_t\|_{L^q_{\mu_t}}$  for  $\mathcal{L}^1$ -a.e.  $t \in I$ .

The vector field  $v_t$  is highly non-unique and in order to restore the identity between the metric speed of  $\mu_t$  and the norm of  $v_t$  a subclass of special vector fields for (2.3) has to be considered. This class is introduced by analogy with the infinitesimal behaviour of  $q$ -Wasserstein geodesics. As we will be mostly dealing with the case  $q = 2$ , where an infinitesimal Riemannian structure appears naturally, we will continue this review only for the case  $q = 2$ . We refer to [AGS05] and references therein for the other cases.

For every  $\mu \in \mathcal{P}_2(M)$ , the so-called ‘‘space of gradients’’ is defined by

$$T_\mu(\mathcal{P}_2(M)) := \overline{\{\nabla f : f \in \mathcal{C}^\infty(M)\}}^{L^2_\mu}.$$

This definition is motivated by the classical results of Y. Brenier and R. McCann on the shape of optimal maps and on 2-cyclically monotone sets. For more details on all the possible notions of tangent space to  $\mathcal{P}_2(M)$  we refer to [Gig11, Oht09].

We now state the main result concerning the Wasserstein metric speed in  $\mathcal{P}_q(M)$ . For the analogous and earlier result when  $M$  is an Hilbert space we refer to [AGS05, Theorem 8.3.1] and [AGS05, Proposition 8.4.5], where also the case  $q \neq 2$  is considered.

**Theorem 2.4** (Proposition 2.5 of [Erb10]). *Let  $(\mu_t)_t \in AC^2(I; \mathcal{P}_2(M))$  be a 2-absolutely continuous curve with  $I \subset \mathbb{R}$  an interval. Then there exists a Borel vector*

field  $v : (x, t) \mapsto v_t(x) \in T_x M$  such that

$$v_t \in L^2(\mu_t, TM), \quad \|v_t\|_{L^2_{\mu_t}} = |\dot{\mu}_t|, \quad \mathcal{L}^1 - a.e. t \in I,$$

and the continuity equation

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0, \quad (2.4)$$

holds in the sense of distributions on  $I \times M$ . Moreover for  $\mathcal{L}^1$ -a.e.  $t \in I$ ,  $v_t \in T_{\mu_t}(\mathcal{P}_2(M))$ .

The vector field  $v_t$  is uniquely determined in  $L^2(\mu_t, TM)$  by (2.4) and by being an element of  $T_{\mu_t}(\mathcal{P}_2(M))$ .

Conversely, if a narrowly continuous curve  $\mu_t : I \rightarrow \mathcal{P}_2(M)$  satisfies the continuity equation for some Borel velocity field  $v_t$  with  $\|v_t\|_{L^2_{\mu_t}} \in L^2(I)$ , then  $(\mu_t)_t \in AC^2(I; \mathcal{P}_2(M))$  and  $|\dot{\mu}_t| \leq \|v_t\|_{L^2_{\mu_t}}$  for  $\mathcal{L}^1$ -a.e.  $t \in I$ .

Notice that [Erb10, Proposition 2.5] and the discussion of Sect. 2.2 permit to reobtain the so-called Benamou-Brenier formulation of the Optimal Transport problem: for any  $\mu_0, \mu_1 \in \mathcal{P}_2(M)$  it holds

$$W_2(\mu_0, \mu_1)^2 = \min \left\{ \int_0^1 \|v_t\|_{L^2_{\mu_t}}^2 dt \right\},$$

where the minimum is taken over all absolutely continuous curves  $(\mu_t)_{t \in [0,1]}$  connecting  $\mu_0$  to  $\mu_1$  with tangent vector field  $(v_t)_t$  in the sense of (2.4).

By definition,  $T_{\mu}(\mathcal{P}_2(M))$  is a closed subspace of  $L^2(\mu; TM)$  and therefore has a uniquely defined projection

$$\mathbf{P} : L^2(\mu; TM) \rightarrow T_{\mu}(\mathcal{P}_2(M)).$$

The orthogonal space to  $T_{\mu}(\mathcal{P}_2(M))$  consists of divergence-free vector fields, giving the following characterization for  $\mathbf{P}$ .

LEMMA 2.5 (Lemma 2.4 of [Erb10]). *A vector  $v \in L^2(\mu; TM)$  belongs to  $T_{\mu}(\mathcal{P}_2(M))$  if and only if*

$$\|v + w\|_{L^2_{\mu}} \geq \|v\|_{L^2_{\mu}}, \quad \forall w \in L^2(\mu; TM) \text{ such that } \operatorname{div}(w\mu) = 0.$$

*In particular, for every  $v \in L^2(\mu; TM)$  there exists a unique  $\mathbf{P}(v) \in T_{\mu}(\mathcal{P}_2(M))$  in the equivalence class of  $v$  modulo divergence-free vector fields: it is the element of minimal  $L^2$ -norm in this class.*

In order to highlight the relevance of the projection  $\mathbf{P}$  we reformulate Theorem 2.4.

PROPOSITION 2.6 (Proposition 2.5 of [Erb10]). *Let  $\mu_t : I \rightarrow \mathcal{P}_2(M)$  be an absolutely continuous curve and let  $v_t \in L^2(\mu_t; TM)$  be such that*

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0,$$

holds in the sense of distributions. Then  $v_t$  satisfies

$$\|v_t\|_{L^2_{\mu_t}} = |\dot{\mu}_t|, \quad \mathcal{L}^1 - a.e.,$$

if and only if  $v_t = \mathbf{P}(v_t) \in T_{\mu_t}(\mathcal{P}_2(M))$  for  $\mathcal{L}^1$ -a.e.  $t \in I$ .

Relevant for this paper will be curves obtained via the pushforward of families of diffeomorphisms: let  $\varphi_t : M \rightarrow M$  be a family of diffeomorphisms which is the flow of a time dependent vector field  $v_t \in \Gamma(TM)$ . If we define  $\mu_t := (\varphi_t)_\# \mu_0$ , then

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0.$$

Indeed by definition  $\frac{d}{dt} \varphi_t(x) = v_t(\varphi_t(x))$  and for any smooth test function  $\psi$ :

$$\frac{d}{dt} \int_M \psi d\mu_t = \frac{d}{dt} \int_M \psi(\varphi_t) d\mu_0 = \int g(\nabla \psi(\varphi_t(x)), v_t(\varphi_t(x))) d\mu_0 = \int g(\nabla \psi, v_t) d\mu_t. \quad (2.5)$$

Moreover since  $\operatorname{div}((v_t - \mathbf{P}(v_t))\mu_t) = 0$ ,

$$\partial_t \mu_t + \operatorname{div}(\mathbf{P}(v_t)\mu_t) = 0.$$

Since  $\mathbf{P}(v_t) \in T_{\mu_t}(\mathcal{P}_2(M))$ , by Proposition 2.6 we have therefore obtained the following statement.

**COROLLARY 2.7.** *Let  $\varphi_t : M \rightarrow M$  be a family of diffeomorphisms for  $t \in I$ , with  $\varphi_t$  be the flow map associated to a vector field  $v_t \in \Gamma(TM)$  and  $I$  any real interval. Define for  $t \in I$  the curve of probability measures  $\mu_t := (\varphi_t)_\# \mu_0$ . Then, for  $\mathcal{L}^1$ -a.e.  $t \in I$ :*

$$|\dot{\mu}_t| = \|\mathbf{P}(v_t)\|_{L^2_{\mu_t}}.$$

The description of the space of gradients is therefore a natural problem to address. The following proposition describes the space of gradients in the case of a measure  $\mu$  supported on a smooth submanifold  $S \subseteq M$ . For other descriptions of the tangent cone to the 2-Wasserstein space and similar results, we refer to [Lot17].

**PROPOSITION 2.8.** *Let  $S \subset M$  be a submanifold and  $\mu \in \mathcal{P}(M)$  such that  $\operatorname{supp}(\mu) \subset S$ . Then*

$$\overline{\{\nabla f : f \in \mathcal{C}^\infty(M)\}}^{L^2_\mu} = \overline{\{\nabla g : g \in \mathcal{C}^\infty(S)\}}^{L^2_\mu} \oplus L^2(NS, \mu),$$

where  $NS \subset TM$  is the normal bundle to  $S$  in  $M$ .

*Proof.* Since  $\mu$  is supported on  $S$  this arises by separating tangential and normal components at  $S$  of the gradients of functions on  $M$ . The unique non trivial inclusion to justify is:  $L^2(NS, \mu) \subseteq \overline{\{\nabla f : f \in \mathcal{C}^\infty(M)\}}^{L^2_\mu}$ . For this we just have to prove that any smooth section  $\nu \in \Gamma(S, NS)$  is the normal component at  $S$  of the gradient

of a smooth function defined around  $S$ . This can be easily done using the tubular neighborhood theorem. Identify  $NS \cong TS^\perp$ ; the map

$$\varphi : NS \longrightarrow M, \quad v \longmapsto \exp v,$$

maps diffeomorphically a neighborhood of  $S$  inside  $NS$  to a neighborhood  $U$  of  $S$  in  $M$ . Moreover  $d\varphi|_S = \text{Id}$ . Let  $\pi : NS \rightarrow S$  be the projection and define the function  $g(x) := \langle \nu(\pi(\varphi^{-1}(x))), \varphi^{-1}(x) \rangle$ , for  $x \in U$ . Denote with  $\nabla^\perp g$  the normal component of the gradient over  $S$ . This is computed by differentiating  $g$  along curves in the form  $t \mapsto \exp(tw) = \varphi(tw)$  for  $w \in NS$ . It follows immediately:  $\nabla^\perp g = \nu$ .  $\square$

### 3 Measure theory and algebraic geometry

In this section we will define the correspondence between homogeneous polynomials and probability measures on their zero sets. We will begin by studying families of measures depending on a finite dimensional parameter space.

The setting we will consider is, roughly speaking, the following. We are given a compact Riemannian manifold  $(M, g)$ , a manifold  $P$ , and a smooth map  $\pi : M \rightarrow P$ . If  $p \in P$  is a regular value of  $\pi$ , then  $\pi^{-1}(p)$  is a smooth, compact submanifold of  $M$ , which inherits a Riemannian structure with its (normalized) volume measure  $\mu(p) \in \mathcal{P}(M)$  that equals the Hausdorff measure of its dimension. Some care has to be put if one wants to extend the map  $p \mapsto \mu(p)$  to the whole  $P$  (i.e. also to the set of critical values of  $\pi$ ) in a nice way. In fact, even if  $\mu : \{\text{regular values of } \pi\} \rightarrow \mathcal{P}(M)$  is continuous, it is not always possible to extend it to a continuous map without imposing additional conditions on  $\pi$  (see the example in Remark 3.12).

In the case  $M$  and  $P$  are complex manifolds of the same dimension and  $\pi$  is holomorphic, the so-called “finiteness” of  $\pi$  guarantees a continuous extension of  $\mu$  (we discuss this in Sect. 3.1). In fact, by Remark 3.12, the finiteness condition is also necessary for the continuous extension. The key ingredient here is the notion of *multiplicity* from complex algebraic geometry.

In the case  $\dim(M) > \dim(P)$ , in general, the situation is more involved. However, when  $M = \mathbb{C}P^n$ , using integral geometry arguments, one can provide the continuous extension for interesting classes of maps, including the one parametrizing complex hypersurfaces of a given degree (see Sect. 3.3).

In this framework, the problem of the continuous extension can also be treated directly using a variation of the arguments from [Sto67], but we prefer to keep the exposition self-contained as the description of the extension using integral geometry will become useful in Sect. 6.

**3.1 Families of zero-dimensional measures.** We say that a holomorphic map  $\pi : V \rightarrow P$  between smooth, complex, compact and connected manifolds of the same dimension is *finite* if for every  $p \in P$  we have  $\#\pi^{-1}(p) < \infty$ . We denote by  $\Sigma(\pi) \subset V$  the set of critical points of  $\pi$ , and by  $\Delta(\pi) := \pi(\Sigma(\pi))$  the set of its critical values (the

“discriminant”). The degree  $\deg(\pi)$  is defined as  $\deg(\pi) := \#\pi^{-1}(p)$  for  $p \in P \setminus \Delta(\pi)$  (the set of critical points is an analytic set of complex codimension one, therefore  $P \setminus \Delta(\pi)$  is connected). For every  $v \in V$  we denote by  $\text{mult}_v(\pi)$  the multiplicity of  $\pi$  at  $v$ . The multiplicity satisfies

$$\deg(\pi) = \sum_{v \in \pi^{-1}(p)} \text{mult}_v(\pi), \quad \forall p \in P.$$

We refer to [GH94, Chap. 5.2] for elementary properties of finite maps.

LEMMA 3.1. *Let  $\pi : V \rightarrow P$  be a finite map between smooth, compact, complex manifolds of the same dimension. Then the map  $\alpha : P \rightarrow \mathcal{P}(V)$ , defined by*

$$\alpha(p) := \frac{1}{\deg(\pi)} \sum_{v \in \pi^{-1}(p)} \text{mult}_v(\pi) \delta_v,$$

is continuous.

*Proof.* The topology on  $\mathcal{P}(V)$  is metrizable and it is enough to prove sequential continuity of  $\alpha$ : given  $\{p_n\}_{n \in \mathbb{N}} \subset P$  converging to  $p \in P$ , the sequence  $\{\alpha(p_n)\}_{n \in \mathbb{N}}$  converges to  $\alpha(p)$ . In other words, we need to prove that  $\forall f \in \mathcal{C}^0(V)$

$$\lim_{n \rightarrow \infty} \frac{1}{\deg(\pi)} \sum_{v \in \pi^{-1}(p_n)} \text{mult}_v(\pi) f(v) = \frac{1}{\deg(\pi)} \sum_{v \in \pi^{-1}(p)} \text{mult}_v(\pi) f(v) \quad (3.1)$$

Let us consider first the case  $\{p_n\}_{n \in \mathbb{N}} \subset P \setminus \Delta(\pi)$ . Let  $p = \lim_n p_n$  and set  $\pi^{-1}(p) = \{v_1, \dots, v_r\}$ , with  $m_j := \text{mult}_{v_j}(\pi)$ . For every  $\epsilon > 0$  there exists an open neighborhood  $U$  of  $p$  such that  $\pi^{-1}(U) = V_1 \sqcup \dots \sqcup V_r$  and for every  $j = 1, \dots, r$  and for every  $v \in V_j$

$$|f(v) - f(v_j)| \leq \epsilon.$$

For every  $n \in \mathbb{N}$  we denote by  $v_{n,j}^1, \dots, v_{n,j}^{m_j}$  the (distinct!) preimages of  $p_n$  in  $V_j$ . Notice in particular that, since  $p_n \notin \Delta(\pi)$ , we have  $\text{mult}_{v_{n,j}^k}(\pi) \equiv 1$ . Then

$$\begin{aligned} \left| \sum_{v \in \pi^{-1}(p_n)} \text{mult}_v(\pi) f(v) - \sum_{v \in \pi^{-1}(p)} \text{mult}_v(\pi) f(v) \right| &= \left| \sum_{j=1}^r \left[ \left( \sum_{k=1}^{m_j} f(v_{n,j}^k) \right) - m_j f(v_j) \right] \right| \\ &= \left| \sum_{j=1}^r \sum_{k=1}^{m_j} \left( f(v_{n,j}^k) - f(v_j) \right) \right| \\ &\leq \deg(\pi) \epsilon. \end{aligned}$$

This proves (3.1) in the case  $\{p_n\}_{n \in \mathbb{N}} \subset P \setminus \Delta(\pi)$ . For the general case we argue as follows. For every  $n \in \mathbb{N}$  let  $\{p_{n,m}\}_{m \in \mathbb{N}} \subset P \setminus \Delta(\pi)$  be a sequence such that  $\lim_m p_{n,m} = p_n$ . For every  $\epsilon > 0$  let  $p_{n,m_\epsilon} =: \tilde{p}_n$  such that  $\lim_n \tilde{p}_n = p$  and

$$\left| \sum_{v \in \pi^{-1}(\tilde{p}_n)} f(v) - \sum_{v \in \pi^{-1}(p_n)} \text{mult}_v(\pi) f(v) \right| \leq \epsilon.$$

Then:

$$\begin{aligned} \left| \sum_{v \in \pi^{-1}(p_n)} \text{mult}_v(\pi)f(v) - \sum_{v \in \pi^{-1}(p)} \text{mult}_v(\pi)f(v) \right| &\leq \left| \sum_{v \in \pi^{-1}(\tilde{p}_n)} \text{mult}_v(\pi)f(v) + \right. \\ &\quad \left. - \sum_{v \in \pi^{-1}(p_n)} \text{mult}_v(\pi)f(v) \right| + \\ &\quad + \left| \sum_{v \in \pi^{-1}(p_n)} \text{mult}_v(\pi)f(v) + \right. \\ &\quad \left. - \sum_{v \in \pi^{-1}(p)} \text{mult}_v(\pi)f(v) \right|. \end{aligned}$$

By the previous part of the proof, both summands go to zero as  $n \rightarrow \infty$ , and this concludes the proof.  $\square$

**3.2 Polynomials in one variable.** We will now focus on the case of the family of measures given by the normalized sum of the delta measures centered at the zeroes of a polynomial. Our “parameter” space will be the (projectivization of the) complex vector space  $\mathbb{C}[z_0, z_1]_{(d)}$  of homogeneous polynomials of degree  $d$  in two variables.

Every nonzero homogeneous polynomial  $p \in \mathbb{C}[z_0, z_1]_{(d)}$  can be written as

$$p(z_0, z_1) = \prod_{j=1}^r (\lambda_{j,1}z_0 - \lambda_{j,0}z_1)^{m_j}, \quad (3.2)$$

where  $[\lambda_{j,0}, \lambda_{j,1}]$ ,  $j = 1, \dots, r$  are distinct points in  $\mathbb{CP}^1$  and  $m_1 + \dots + m_r = d$ .

With this notation, if  $w = [w_0, w_1] \in \mathbb{CP}^1$ , we denote by  $m_w(p)$  the *multiplicity* of  $p$  at  $w$ . In other words,  $m_w(p) = m_j$ , if  $w = [\lambda_{j,0}, \lambda_{j,1}]$  for some  $j = 1, \dots, r$ , and  $m_w(p) = 0$  otherwise. In particular,  $m_w(p) \neq 0$  only if  $w$  is a zero of  $p$ , in which case it is the multiplicity of it as a projective zero of  $p$ . For a nonzero polynomial  $p \in \mathbb{C}[z_0, z_1]_{(d)}$  we denote by  $Z(p) \subset \mathbb{CP}^1$  the set of its zeroes, i.e., writing  $p$  as in (3.2), we set

$$Z(p) := \{[\lambda_{j,0}, \lambda_{j,1}] \in \mathbb{CP}^1, j = 1, \dots, r\}.$$

**REMARK 3.2** (Homogeneous versus non-homogeneous polynomials). The reader might prefer to work with polynomials in one variable of degree *at most*  $d$ , whose zero set lives in the complex line  $\mathbb{C}$ . Every monic polynomial of the form

$$q(z) = z^d + a_{d-1}z^{d-1} + \dots + a_1z + a_0$$

can be written (by the Fundamental Theorem of Algebra) as

$$q(z) = \prod_{j=1}^d (z - \lambda_j),$$

for some numbers  $\lambda_1, \dots, \lambda_d \in \mathbb{C}$  (the zeroes of  $q$ ). Homogenizing  $q$  we get the polynomial

$$p(z_0, z_1) = z_1^d + a_{d-1}z_0z_1^{d-1} + \dots + a_1z_0^{d-1}z_1 + a_0z_0^d,$$

whose zeroes in  $\mathbb{CP}^1$  are  $[1, \lambda_1], \dots, [1, \lambda_d]$ . Notice that  $p$  can be written as  $p(z_0, z_1) = \prod_{j=1}^d (z_1 - \lambda_j z_0)$ . The advantage of working with homogenous polynomials is that their zeroes live in a compact space (the projective line). In the case of homogeneous polynomials in two variables it is still safe to think at their dehomogenization and look at their zeroes in  $\mathbb{C}$  (notice, though, that if we start with a polynomial with  $a_d = 0$  some zeroes are lost in this procedure). However, for our purposes, when moving to more variables the reader should really think in terms of projective zeroes and pay attention to the fact that the set of zeroes of a complex polynomial in  $\mathbb{C}^n$ ,  $n \geq 2$ , is not compact (see also Sect. A.3).

From Lemma 3.1 we deduce the following corollary (this also proved in [Fed69, Sect. 4.3.12]).

**COROLLARY 3.3.** *Let  $\mathbb{CP}^d$  be the projectivization of  $\mathbb{C}[z_0, z_1]_{(d)} \simeq \mathbb{C}^{d+1}$ . The map  $\mu : \mathbb{CP}^d \rightarrow \mathcal{P}(\mathbb{CP}^1)$  defined by*

$$\mu(p) := \frac{1}{d} \sum_{z \in Z(p)} m_z(p) \delta_z$$

*is continuous.*

*Proof.* Let  $P = \mathbb{CP}^d$  and consider the smooth algebraic set  $V = \{(p, z) \in P \times \mathbb{CP}^1 \mid p(z) = 0\}$  and the diagram:

$$\begin{array}{ccc} & V & \\ \pi_1 \swarrow & & \searrow \pi_2 \\ P & & \mathbb{CP}^1 \end{array}$$

where  $\pi_1$  and  $\pi_2$  are the restrictions of the projections on the two factors. The map  $\pi_1 : V \rightarrow P$  is a finite map between smooth, compact, complex manifolds and  $\deg(\pi_1) = d$ . Moreover, using the notation of Lemma 3.1, we see that  $\mu(p) = (\pi_2)_\# \alpha(p)$ . Since  $\pi_2$  is continuous, the continuity of  $\mu$  follows from Lemma 3.1.  $\square$

**3.3 Families of hypersurfaces.** We now generalize the previous construction to homogeneous polynomials of more than two variables, whose zero sets are hypersurfaces in the projective space.

For a smooth Riemannian manifold  $(M, g)$ , we denote by  $\text{vol}_M$  the corresponding Riemannian volume measure (hence omitting the dependence on the Riemannian metric  $g$  in the notation). In the case  $M = \mathbb{CP}^n$ , we denote by  $g_{\text{FS}}$  the Fubini–Study metric (see the Appendix).

DEFINITION 3.4 (The space of polynomials and the discriminant). For  $n, d \in \mathbb{N}$ , we denote by

$$H_{n,d} := \mathbb{C}[z_0, \dots, z_n]_{(d)}$$

the space of homogeneous polynomials of degree  $d$ , a vector space of complex dimension  $N + 1 = \binom{n+d}{d}$ . We also consider the *discriminant*  $D_{n,d} \subset H_{n,d}$  defined by

$$D_{n,d} := \left\{ p \in H_{n,d} \mid \exists z \in \mathbb{C}^{n+1} \setminus \{0\}, p(z) = \frac{\partial p}{\partial z_0}(z) = \dots = \frac{\partial p}{\partial z_n}(z) = 0 \right\},$$

i.e. the set of polynomials  $p$  such that the equation  $p = 0$  is not regular in  $\mathbb{C}P^n$ . Notice that the discriminant is an algebraic set in  $H_{n,d}$  defined by a polynomial of degree  $(n+1)(d-1)^{n-1}$ . We also denote by  $P_{n,d}$  and  $\Delta_{n,d}$  the projectivization of  $H_{n,d}$  and  $D_{n,d}$ :

$$P_{n,d} := P(H_{n,d}) \simeq \mathbb{C}P^N \quad \text{and} \quad \Delta_{n,d} := P(D_{n,d}).$$

Notice that, given a point  $[p] \in P_{n,d}$ , its projective zero set

$$Z(p) := \{[x] \in \mathbb{C}P^n \mid p(x) = 0\}$$

is well-defined (for this reason we sometimes simply write  $p$  for  $[p]$ ).

Let us make a couple of comments on the previous definitions. If  $p \in P_{n,d}$ , regardless of it being on the discriminant or not, the set  $Z(p) \subset \mathbb{C}P^n$  is a (possibly singular) compact, complex hypersurface, which has Hausdorff dimension  $2n - 2$ . If  $p \in P_{n,d} \setminus \Delta_{n,d}$ , then the set  $Z(p)$  is a smooth complex manifold. If  $p \in \Delta_{n,d}$  then, as a subset of  $\mathbb{C}P^n$ ,  $Z(p)$  can still be smooth. For example, if  $p \in P_{n,d} \setminus \Delta_{n,d}$ , then  $p^2 \in \Delta_{n,2d}$ , but  $Z(p^2) = Z(p)$ , as a set, is a smooth complex submanifold. In order to emphasize this phenomenon, we will say that  $Z(p)$  is *regular* if  $p \in P_{n,d} \setminus \Delta_{n,d}$  and *singular* otherwise (thus being “regular” has to do with the equation defining a set, and not with the set itself). If  $p \in \Delta_{n,d}$  we will denote by  $Z(p)^{\text{sing}}$  the set

$$Z(p)^{\text{sing}} = \left\{ [z] \in Z(p) \mid \frac{\partial p}{\partial z_0}(z) = \dots = \frac{\partial p}{\partial z_n}(z) = 0 \right\}.$$

It can be that  $Z(p)^{\text{sing}} = Z(p)$  (for instance, this happens again for the polynomials of the form  $p^2$ ). On the other hand, if  $p \in P_{n,d}$  is irreducible, then  $Z(p)^{\text{sing}}$  is a proper algebraic subset of  $Z(p)$ , therefore of complex codimension one in it, and

$$Z(p)^{\text{reg}} := Z(p) \setminus Z(p)^{\text{sing}}$$

is a smooth submanifold of  $\mathbb{C}P^n$ .

In the rest we will often use *semialgebraic* objects, whose definition we recall now. For more details the reader is referred to [BCR98].

**DEFINITION 3.5** (Semialgebraic sets and functions). A semialgebraic set  $S \subseteq \mathbb{R}^n$  is a set obtained by taking finite unions, finite intersections of sets of the form  $\{p \leq 0\}$  and their complements, where  $p \in \mathbb{R}[x_1, \dots, x_n]$  is a polynomial. A function  $f : A \rightarrow B$ , where  $A \subset \mathbb{R}^n$ ,  $B \subset \mathbb{R}^m$  are semialgebraic, is said to be semialgebraic if its graph  $\Gamma(f) \subset A \times B \subset \mathbb{R}^{n+m}$  is a semialgebraic set.

By [BCR98, Proposition 2.4.6], there exists a smooth, real algebraic embedding  $\varphi : \mathbb{C}\mathbb{P}^n \rightarrow \mathbb{R}^\ell$ , for some  $\ell \in \mathbb{N}$  sufficiently large. Therefore we can view  $\mathbb{C}\mathbb{P}^n$  as a semialgebraic set itself. Since complex algebraic sets are in particular real algebraic sets (every complex equation corresponds to a pair of algebraic equations), every complex algebraic set  $Z \subseteq \mathbb{C}\mathbb{P}^n$  is therefore semialgebraic. Therefore we can stratify  $Z$  as a disjoint union

$$Z = \bigsqcup_{j=1}^s Z_j, \quad (3.3)$$

where each  $Z_j \subset \mathbb{C}\mathbb{P}^n$  is a smooth semialgebraic set (see [BCR98, Theorem 9.1.8]).

**DEFINITION 3.6** (Volume of a complex algebraic set). Let  $Z \subset \mathbb{C}\mathbb{P}^n$  be a complex algebraic set, stratified as (3.3). Each stratum inherits a Riemannian metric  $g_{\text{FS}}|_{Z_j}$  and hence a volume measure  $\text{vol}_{Z_j}$ . We define the volume measure  $\text{vol}_Z \in \mathcal{P}(\mathbb{C}\mathbb{P}^n)$  as follows. For a continuous function  $f : Z \rightarrow \mathbb{R}$  we set

$$\int_Z f(z) \text{vol}_Z(dz) := \sum_{\dim(Z_j)=2n-2} \int_{Z_j} f(z) \text{vol}_{Z_j}(dz).$$

Since any two semialgebraic stratifications can be commonly refined (see [BCR98, Theorem 9.1.8]), this definition does not depend on the choice of the stratification from (3.3). We denote by  $\text{vol}(Z)$  the total mass of  $\text{vol}_Z$ .

In the case  $p \in P_{n,d} \setminus \Delta_{n,d}$  (i.e. if the hypersurface  $Z(p) \subset \mathbb{C}\mathbb{P}^n$  is *regular*), then for every Borel set  $A \subseteq Z(p)$ ,

$$\text{vol}_{Z(p)}(A) = \frac{1}{(n-1)!} \int_A \omega_{\text{FS}}^{n-1},$$

where  $\omega_{\text{FS}}$  denotes the Kähler form of  $\mathbb{C}\mathbb{P}^n$  (see the [Appendix](#)).

We observe that Definition 3.6 does not produce a continuous map in the Wasserstein topology, i.e. the correspondence  $p \mapsto \frac{1}{\text{vol}(Z(p))} \text{vol}_{Z(p)}$  is not continuous on  $\Delta_{n,d}$ . As we have seen in Corollary 3.3, a continuous extension on  $\Delta_{1,d}$  should take into account the multiplicities of zeroes. In the general case we proceed as follows. First, in Lemma 3.7, we give an integral geometric formulation of the volume, for smooth submanifolds of  $\mathbb{C}\mathbb{P}^n$ . Then, in Theorem 3.8, we use this formulation to extend the map  $\mu$  to  $\Delta_{n,d}$ .

We denote by  $\mathbb{G}(1, n)$  the Grassmannian of projective lines in  $\mathbb{C}\mathbb{P}^n$  (notice that  $\mathbb{G}(1, n)$  is isomorphic to  $G(2, n+1)$ , the Grassmannian of complex 2-planes in

$\mathbb{C}^{n+1}$ ). There is a natural action of  $U(n+1)$  on  $\mathbb{G}(1, n)$ , and we endow  $\mathbb{G}(1, n)$  with a Riemannian metric which is invariant under this action, normalized such that  $\text{vol}_{\mathbb{G}(1, n)}(\mathbb{G}(1, n)) = 1$ .

LEMMA 3.7. *Let  $f \in \mathcal{C}^0(\mathbb{C}P^n)$  and  $Z \subset \mathbb{C}P^n$  be a complex submanifold of complex dimension  $n-1$ . Then for every  $U \subseteq Z$  open we have*

$$\int_U f(z) \text{vol}_Z(dz) = \text{vol}(\mathbb{C}P^{n-1}) \int_{\mathbb{G}(1, n)} \left( \sum_{z \in U \cap \ell} f(z) \right) \text{vol}_{\mathbb{G}(1, n)}(d\ell), \quad (3.4)$$

the integrand being a.e. finite. In particular, if  $p \in P_{n, d} \setminus \Delta_{n, d}$ , then  $\text{vol}_{Z(p)}(Z(p)) = d \cdot \text{vol}(\mathbb{C}P^{n-1})$ .

*Proof.* Let us recall first the kinematic formula in complex projective spaces from [How93, Example 3.12 (b)]. Given two open subsets  $A \subseteq Z$ ,  $B \subseteq L$  of complex submanifolds  $Z, L$  of complementary dimensions in  $\mathbb{C}P^n$ ,  $\dim_{\mathbb{C}}(Z) = a$ ,  $\dim_{\mathbb{C}}(L) = b$ , we have:

$$\int_{U(n+1)} \#(A \cap gB) \text{vol}_{U(n+1)}(dg) = \frac{\text{vol}_Z(A)}{\text{vol}_{\mathbb{C}P^a}(\mathbb{C}P^a)} \frac{\text{vol}_L(B)}{\text{vol}_{\mathbb{C}P^b}(\mathbb{C}P^b)}, \quad (3.5)$$

where “ $dg$ ” denotes the integration with respect to the normalized Haar measure  $\int_{U(n+1)} dg = 1$  (which is obtained as the volume form of a left-invariant Riemannian metric on  $U(n+1)$ ). Applying this to the special case  $B = \mathbb{C}P^1$ , we get

$$\int_{U(n+1)} \#(A \cap gB) \text{vol}_{U(n+1)}(dg) = \frac{\text{vol}_Z(A)}{\text{vol}_{\mathbb{C}P^{n-1}}(\mathbb{C}P^{n-1})}.$$

In the case  $A = Z = Z(p)$  is regular, the integrand is equal to  $d$  for a.e.  $g$  (see [GH94]) and

$$d = \frac{\text{vol}_Z(Z)}{\text{vol}_{\mathbb{C}P^{n-1}}(\mathbb{C}P^{n-1})},$$

proving the second part of the statement.

Observe now that the map  $U(n+1) \rightarrow \mathbb{G}(1, n)$  given by  $g \mapsto g(\ell_0)$ , where  $\ell_0 \in \mathbb{G}(1, n)$  is a fixed line, is a Riemannian submersion. Therefore, in the case  $A \subseteq Z$  is open and  $B = \mathbb{C}P^1$ , using the smooth coarea formula, (3.5) can be written as

$$\int_{\mathbb{G}(1, n)} \#(A \cap \ell) \text{vol}_{\mathbb{G}(1, n)}(d\ell) = \frac{\text{vol}_Z(A)}{\text{vol}_{\mathbb{C}P^{n-1}}(\mathbb{C}P^{n-1})}. \quad (3.6)$$

Consider now the linear functional  $\lambda : \mathcal{C}^0(Z) \rightarrow \mathbb{R}$  defined by:

$$f \mapsto \text{vol}_{\mathbb{C}P^{n-1}}(\mathbb{C}P^{n-1}) \int_{\mathbb{G}(1, n)} \left( \sum_{z \in Z \cap \ell} f(z) \right) \text{vol}_{\mathbb{G}(1, n)}(d\ell).$$

This is a continuous and positive linear functional on  $\mathcal{C}^0(Z)$ , therefore by Riesz's Representation Theorem [Rud87, Theorem 2.14], there exists a unique nonnegative

Radon measure  $\nu$  such that for every  $f : Z \rightarrow \mathbb{R}$  Borel measurable:

$$\lambda(f) = \int_Z f(z) \nu(dz).$$

Let now  $U$  be an open set of  $Z$  and  $f = \chi_U$  be its characteristic function. Approximating  $f$  with continuous functions, we see that

$$\begin{aligned} \nu(U) &= \lambda(f) = \text{vol}_{\mathbb{C}P^{n-1}}(\mathbb{C}P^{n-1}) \int_{\mathbb{G}(1,n)} \left( \sum_{z \in Z \cap \ell} f(z) \right) \text{vol}_{\mathbb{G}(1,n)}(d\ell) \\ &= \text{vol}_{\mathbb{C}P^{n-1}}(\mathbb{C}P^{n-1}) \int_{\mathbb{G}(1,n)} \#(U \cap \ell) \text{vol}_{\mathbb{G}(1,n)}(d\ell) \\ &= \text{vol}_Z(U). \quad (\text{by (3.6)}) \end{aligned}$$

This shows that, as measures,  $\nu = \text{vol}_Z$  and (3.4) follows.  $\square$

We have now a well defined map

$$\mu : P_{n,d} \setminus \Delta_{n,d} \longrightarrow \mathcal{P}(\mathbb{C}P^n)$$

given by

$$\mu(p)(f) := \frac{1}{\text{vol}(Z(p))} \int_{Z(p)} f(z) \text{vol}_Z(dz), \quad f \in \mathcal{C}^0(\mathbb{C}P^n). \quad (3.7)$$

We prove now that this map can be continuously extended to the whole  $P_{n,d}$ , generalizing Corollary 3.3. Notice that the normalization factor  $\frac{1}{\text{vol}(Z(p))}$  in (3.7) does not actually depend on the hypersurface, because of Lemma 3.7.

**Theorem 3.8.** *The map  $\mu : P_{n,d} \rightarrow \mathcal{P}(\mathbb{C}P^n)$  defined for  $f \in \mathcal{C}^0(\mathbb{C}P^n)$  by*

$$\int_{\mathbb{C}P^n} f d\mu(p) := \frac{1}{d} \int_{\mathbb{G}(1,n)} \left( \sum_{z \in Z(p) \cap \ell} m_z(p|_\ell) f(z) \right) \text{vol}_{\mathbb{G}(1,n)}(d\ell) \quad (3.8)$$

*is continuous and, when restricted to  $P_{n,d} \setminus \Delta_{n,d}$  coincides with (3.7).*

*Proof.* The fact that the restriction of  $\mu$  to  $P_{n,d} \setminus \Delta_{n,d}$  coincides with (3.7) is an immediate consequence of Lemma 3.7. For the continuity of  $\mu$ , as above, it is enough to prove sequential continuity. Let therefore  $\{p_k\}_{k \in \mathbb{N}} \subset P_{n,d}$  be a sequence converging to  $p$ . For every  $k \in \mathbb{N}$  denote by  $F_k \subset \mathbb{G}(1,n)$  the set of lines contained in  $Z(p_k)$  and by  $F$  the set of lines contained in  $Z(p)$ . Notice that these sets have measure zero in  $\mathbb{G}(1,n)$ , therefore the complement of their union

$$\tilde{\mathbb{G}} := \mathbb{G}(1,n) \setminus \left( F \cup \bigcup_{k \in \mathbb{N}} F_k \right)$$

has full measure in  $\mathbb{G}(1, n)$ . Given  $f \in \mathcal{C}^0(\mathbb{C}\mathbb{P}^n)$  and for every  $k \in \mathbb{N}$  denote by  $g_k, g : \mathbb{G}(1, n) \rightarrow \mathbb{R}$  the functions defined for  $\ell \in \tilde{\mathbb{G}}$  by:

$$g_k(\ell) := \frac{1}{d} \sum_{z \in Z(p_k) \cap \ell} m_z(p|_\ell) f(z) \quad \text{and} \quad g(\ell) := \frac{1}{d} \sum_{z \in Z(p) \cap \ell} m_z(p|_\ell) f(z),$$

and set these function to zero on  $\tilde{\mathbb{G}}^c$ . Then for every  $k \in \mathbb{N}$  we have

$$\begin{aligned} \int_{\mathbb{C}\mathbb{P}^n} f d\mu(p_k) &= \int_{\mathbb{G}(1, n)} g_k(\ell) \text{vol}_{\mathbb{G}(1, n)}(d\ell) \quad \text{and} \\ \int_{\mathbb{C}\mathbb{P}^n} f d\mu(p) &= \int_{\mathbb{G}(1, n)} g(\ell) \text{vol}_{\mathbb{G}(1, n)}(d\ell). \end{aligned}$$

The sequence of functions  $\{g_k\}_{k \in \mathbb{N}}$  converges pointwise to  $g$  by Corollary 3.3. Moreover the whole sequence is dominated by  $|g_k| \leq \max_{z \in \mathbb{C}\mathbb{P}^n} |f(z)|$ , therefore the convergence  $\mu(p_k) \rightarrow \mu(p)$  follows from the Dominated Convergence Theorem.  $\square$

We prove now that  $\mu : P_{n, d} \rightarrow \mathcal{P}(\mathbb{C}\mathbb{P}^n)$  is injective. We need a couple of lemmas first.

LEMMA 3.9. *If  $p \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$  is irreducible, then  $\mu(p) = \frac{1}{d \cdot \text{vol}(\mathbb{C}\mathbb{P}^{n-1})} \text{vol}_{Z(p)^{\text{reg}}}$ .*

*Proof.* If  $p$  is irreducible, then  $Z(p)^{\text{sing}}$  has complex codimension at least 2 in  $\mathbb{C}\mathbb{P}^n$  (see [GH94, Chap. 0.2]). In particular, almost every projective line  $\ell \in \mathbb{G}(1, n)$  misses it; since the condition “ $\ell$  misses  $Z(p)^{\text{sing}}$ ” is semialgebraic, it follows that the set  $D_1 \subset \mathbb{G}(1, n)$  of lines missing  $Z(p)^{\text{sing}}$  is a full-measure semialgebraic set. Moreover, by Bertini’s Theorem (see [GH94, Chap. 1.1]), for the generic line  $\ell$  the equation  $p|_\ell = 0$  is regular. Again, since the condition “ $p|_\ell = 0$  is regular” is semialgebraic, the set  $D_2$  of lines  $\ell$  such that  $p|_\ell = 0$  is regular is also full-measure and semialgebraic. (Notice that  $p|_\ell = 0$  is regular if and only if for every  $z \in Z(p) \cap \ell$  we have  $m_z(p|_\ell) = 1$ ).

In particular, given  $f \in \mathcal{C}^0(\mathbb{C}\mathbb{P}^n)$ , we have

$$\begin{aligned} \int_{\mathbb{C}\mathbb{P}^n} f d\mu(p) &= \frac{1}{d} \int_{\mathbb{G}(1, n)} \left( \sum_{z \in Z(p) \cap \ell} m_z(p|_\ell) f(z) \right) \text{vol}_{\mathbb{G}(1, n)}(d\ell) \\ &= \frac{1}{d} \int_{D_1 \cap D_2} \left( \sum_{z \in Z(p) \cap \ell} m_z(p|_\ell) f(z) \right) \text{vol}_{\mathbb{G}(1, n)}(d\ell), \end{aligned}$$

since  $D_1 \cap D_2$  is full measure. This is also equal to

$$\begin{aligned} &= \frac{1}{d} \int_{D_1 \cap D_2} \left( \sum_{z \in Z(p) \cap \ell} f(z) \right) \text{vol}_{\mathbb{G}(1, n)}(d\ell) && (\ell \in D_1) \\ &= \frac{1}{d} \int_{D_1 \cap D_2} \left( \sum_{z \in Z(p)^{\text{reg}} \cap \ell} f(z) \right) \text{vol}_{\mathbb{G}(1, n)}(d\ell) && (\ell \in D_2) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{d} \int_{\mathbb{G}(1,n)} \left( \sum_{z \in Z(p)^{\text{reg}} \cap \ell} f(z) \right) \text{vol}_{\mathbb{G}(1,n)}(d\ell) && (D_1 \cap D_2 \text{ is full measure}) \\
&= \frac{1}{d} \int_{Z(p)^{\text{reg}}} f(z) \text{vol}_{Z(p)^{\text{reg}}}(dz) && (\text{by Lemma 3.7}).
\end{aligned}$$

This concludes the proof.  $\square$

We need also the following lemma.

**LEMMA 3.10.** *Let  $p = p_1^{m_1} \cdots p_k^{m_k} \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$ , with  $p_j \in \mathbb{C}[z_0, \dots, z_n]_{(d_j)}$  irreducible and  $m_j \in \mathbb{N}$ . Then*

$$\mu(p) = \frac{1}{d \cdot \text{vol}(\mathbb{CP}^{n-1})} \sum_{j=1}^k m_k \cdot \text{vol}_{Z(p_j)^{\text{reg}}}.$$

*Proof.* Using (3.8), for a given  $f \in \mathcal{C}^0(\mathbb{CP}^n)$ , we have:

$$\begin{aligned}
\int_{\mathbb{CP}^n} f d\mu(p) &= \frac{1}{d} \int_{\mathbb{G}(1,n)} \left( \sum_{z \in Z(p) \cap \ell} m_z(p|\ell) f(z) \right) \text{vol}_{\mathbb{G}(1,n)}(d\ell) \\
&= \frac{1}{d} \int_{\mathbb{G}(1,n)} \left( \sum_{j=1}^k \sum_{z \in Z(p_j) \cap \ell} m_k m_z(p_j|\ell) f(z) \right) \text{vol}_{\mathbb{G}(1,n)}(d\ell) \\
&= \sum_{j=1}^k m_k \frac{1}{d} \int_{\mathbb{G}(1,n)} \left( \sum_{z \in Z(p_j) \cap \ell} m_z(p_j|\ell) f(z) \right) \text{vol}_{\mathbb{G}(1,n)}(d\ell) \\
&= \sum_{j=1}^k m_k \int_{Z(p_j)^{\text{reg}}} f(z) \text{vol}_{Z(p_j)^{\text{reg}}}(dz) && (\text{by Lemma 3.9}).
\end{aligned}$$

Since almost every lines misses the common intersection  $Z(p_j) \cap Z(p_j)$  (which have complex codimension at least 2), this proves the claim.  $\square$

**PROPOSITION 3.11.** *The map  $\mu: P_{n,d} \rightarrow \mathcal{P}(\mathbb{CP}^n)$  given by (3.8) is injective. In particular  $\mu$  is a homeomorphism onto its image.*

*Proof.* Let  $[p] \neq [q]$ , so that  $p \neq \lambda q$  for  $\lambda \in \mathbb{C} \setminus \{0\}$ . Let us write  $p = p_1^{m_1} \cdots p_k^{m_k}$  and  $q = q_1^{r_1} \cdots q_h^{r_h}$  with the  $p_j$  and the  $q_i$  irreducible and coprime.

Then the sets  $\{Z(p_i)\}$  and  $\{Z(q_j)\}$  are the irreducible components of  $Z(p)$  and  $Z(q)$  respectively. The irreducible components are uniquely determined, therefore, if the sets  $\{[p_1], \dots, [p_k]\}$  and  $\{[q_1], \dots, [q_h]\}$  are different, then the supports of the measures  $\mu(p)$  and  $\mu(q)$  are different (as the two polynomials have different zero sets) and  $\mu([p]) \neq \mu([q])$ .

If  $\{[p_1], \dots, [p_k]\} = \{[q_1], \dots, [q_h]\}$ , then  $h = k$  and, up to relabeling,  $[p_j] = [q_j]$ . Since  $[p] \neq [q]$  there must be  $j \in \{1, \dots, k\}$  such that  $m_j \neq r_j$  and the fact that  $\mu([p]) \neq \mu([q])$  follows from Lemma 3.10.

Since  $\mu$  is continuous and injective and  $P_{n,d}$  is compact, it is a homeomorphism onto its image.  $\square$

REMARK 3.12. The construction for algebraic hypersurfaces is a special case of a more general framework. Consider the set  $V_{n,d} := \{([p], [z]) \in P_{n,d} \times \mathbb{CP}^n \mid p(z) = 0\}$ . The projection  $\pi_2 : V_{n,d} \rightarrow \mathbb{CP}^n$ , whose fibers consist of the set of polynomials vanishing at a given point (a projective hyperplane in the space of polynomials), turns  $V$  into a fiber bundle over  $\mathbb{CP}^n$ ; in particular  $V_{n,d} \subset P_{n,d} \times \mathbb{CP}^n$  is a smooth complex submanifold. The projection on the first factor  $\pi_1 : V_{n,d} \rightarrow P_{n,d}$  has fibers which are precisely the algebraic hypersurfaces of degree  $d$ . Endow  $\mathbb{CP}^n$  and  $P_{n,d}$  (which are both complex projective spaces) with the Fubini–Study Riemannian structure (see the [Appendix](#)), and  $V_{n,d}$  with the induced Riemannian structure  $g$ . Then we can view the map  $\mu : P_{n,d} \setminus \Delta_{n,d} \rightarrow \mathcal{P}(\mathbb{CP}^n)$  from (3.7) as:

$$\mu(p) = (\pi_2)_\# \lambda_{\pi_1^{-1}(p)},$$

were, in general, if  $X \hookrightarrow V$  is a  $m$ -dimensional, compact and smooth submanifold of a Riemannian manifold  $(V, g)$ , we denote by  $\lambda_X \in \mathcal{P}(V)$  the measure defined on  $f \in \mathcal{C}^0(V)$  by

$$\lambda_X(f) = \frac{1}{\text{vol}_m(X)} \int_X f \, \text{dvol}_X.$$

More generally, one can consider a smooth map  $\pi : V \rightarrow P$ , where  $(V, g)$  is a Riemannian manifold, and define  $\alpha : P \setminus \Delta(\pi) \rightarrow \mathcal{P}(V)$  by  $\alpha(p) := \lambda_{\pi^{-1}(p)}$ . Even under the assumptions that  $V$  and  $P$  are complex of the same dimension and  $\pi : V \rightarrow P$  is holomorphic, it is not possible in general to continuously extend  $\alpha$  to the whole  $P$ . For instance, the conclusions of Lemma 3.1 are false if we drop the finiteness assumption on  $\pi : V \rightarrow P$ . To see this, take  $P = \mathbb{CP}^2$  and  $V = \text{Bl}_{p_0}(\mathbb{CP}^2)$  (the blow-up at one point). Both  $P$  and  $V$  are compact complex manifolds of the same complex dimension. We have  $\Delta(\pi) = \{p_0\}$  and the map  $\pi : V \rightarrow P$  is finite (one-to-one) only on  $V \setminus \pi^{-1}(p_0)$ , since  $\pi^{-1}(p_0) \simeq \mathbb{CP}^1$ . In this case, for  $p \neq p_0$ , we have  $\alpha(p) = \delta_{\pi^{-1}(p)}$ , but for  $p_n \rightarrow p_0$  the limit  $\alpha(p_n)$  depends on the sequence  $\{p_n\}_{n \in \mathbb{N}}$ . This is because in general there is no notion of “multiplicity” for the fibers. In the case of Theorem 3.8 the manifolds  $V_{n,d}$  and  $P_{n,d}$  do not have the same dimension (unless  $n = 1$ ) but we can still continuously extend  $\mu = (\pi_2)_\# \circ \alpha$  using integral geometry.

## 4 Symplectic geometry and metric speed of Hamiltonian isotopies

To define an optimal transport distance between algebraic hypersurfaces we will need to study curves inside the image of  $\mu$ . This will be done relying on techniques from symplectic geometry. For the reader’s convenience, we first discuss some basic facts on symplectic manifolds and isotopies (Sect. 4.1); the expert reader can safely jump to Sect. 4.2.

Recall that a *symplectic manifold*  $(M, \omega)$  is a smooth manifold  $M$  equipped with a closed, nondegenerate 2-form. The nondegeneracy assumption on  $\omega$  implies that for every smooth function  $H : M \rightarrow \mathbb{R}$  there is a unique vector field  $V_H \in \Gamma(TM)$  such that

$$dH = \omega(V_H, \cdot) = \iota_{V_H} \omega.$$

A non-autonomous vector field  $v_t$  on a symplectic manifold  $(M, \omega)$  is said to be *Hamiltonian* if there exists a smooth function  $H : M \times I \rightarrow \mathbb{R}$  (the Hamiltonian) such that, denoting by  $h_t := H(\cdot, t) : M \rightarrow \mathbb{R}$ , we have  $v_t = V_{h_t}$ , i.e.  $dh_t = \iota_{V_{h_t}} \omega$ . Notice that the flow  $\varphi_t$  of a Hamiltonian vector field  $V_{h_t}$  preserves the symplectic form: in fact, by Cartan's magic formula,

$$\frac{d}{dt} \varphi_t^* \omega = \varphi_t^* \mathcal{L}_{X_{h_t}} \omega = \varphi_t^* (\iota_{V_{h_t}} d\omega + d(\iota_{V_{h_t}} \omega)) = 0,$$

since  $\omega$  is closed and  $\iota_{V_{h_t}} \omega = dh_t$ .

Recall also that a *Kähler manifold*  $(M, \omega, J)$  is a symplectic manifold  $(M, \omega)$  equipped with an integrable almost-complex structure (i.e. endomorphisms  $J : TM \rightarrow TM$  with  $J^2 = -\text{id}_{TM}$  and admitting a holomorphic atlas) which is compatible with the symplectic form. This means that the bilinear form  $g : TM \times TM \rightarrow \mathbb{R}$  defined by

$$g(u, v) := \omega(u, Jv)$$

is symmetric and positive definite. In particular  $g$  defines a Riemannian metric on  $M$ , which we call the *Kähler Riemannian metric*.

#### 4.1 Hamiltonian isotopies.

**DEFINITION 4.1** (Smooth family of submanifolds). We say that a family  $\{S_t\}_{t \in I}$  of submanifolds of a manifold  $M$  is a *smooth family* if there exist a smooth map  $F : M \times I \rightarrow L$ , where  $I = [0, 1]$  and  $L$  is a smooth manifold, and a submanifold  $A \hookrightarrow L$  such that for every  $t \in I$  the map  $f_t := F(\cdot, t) : M \rightarrow L$  is transverse<sup>1</sup> to  $A$  and  $S_t = f_t^{-1}(A)$ .

Recall that  $f_t : M \rightarrow L$  transverse to  $A$  means that  $df_t(T_x M) + T_{f(x)} A = T_{f_t(x)} L$  at every point  $x \in f_t^{-1}(A)$ . In this case  $f_t^{-1}(A)$  is a submanifold of the same codimension as  $A$ .

**REMARK 4.2.** A classical result in differential topology (Thom's isotopy lemma) states that if  $M$  is compact and  $A$  is closed, given a smooth family  $\{S_t \hookrightarrow M\}_{t \in I}$  there exists a (not unique) isotopy (i.e. a smooth family of diffeomorphisms with  $\varphi_0 = \text{id}_M$ )  $\varphi_t : M \rightarrow M$  such that  $S_t = \varphi_t(S_0)$ . We recall now how the construction of the isotopy  $\varphi_t$  is made, using the notion of flow of a *non-autonomous* vector field

---

<sup>1</sup> If  $f : M \rightarrow L$  is transverse to  $A$ , we will write  $f \pitchfork A$ .

on  $M$ . First observe that we can associate to a non-autonomous vector field  $v_t$  on  $M$ , an autonomous vector field  $\widehat{v}(x, t) := v_t(x) + \partial_t$  on  $M \times I$ . The isotopy  $\varphi_t$  is obtained by restricting the time- $t$  flow  $\widehat{\varphi}_t$  of  $\widehat{v}$  to the slice  $M \times \{0\}$  (which is sent to the slice  $M \times \{t\} \simeq M$ ). In order to understand what properties of  $v_t$  would guarantee that  $\varphi_t(S_0) = S_t$ , denote by  $\widehat{S} := F^{-1}(A) \subseteq M \times I$ . The transversality condition  $f_t \pitchfork A$  for every  $t \in I$  implies that  $\widehat{S}$  is a smooth submanifold of  $M \times I$ . If the vector field  $v_t$  is such that the corresponding vector field  $\widehat{v}$  on  $M \times I$  is tangent to  $\widehat{S}$ , then  $\widehat{S}$  would be preserved by the flow  $\widehat{\varphi}_t$  and  $\varphi_t(S_0) = \widehat{\varphi}_t(\widehat{S} \cap (M \times \{0\})) \subseteq \widehat{S} \cap (M \times \{t\}) \simeq S_t$ . (The reverse inclusion is obtained by considering the flow of  $-\widehat{v}$ .) In other words: in order to find an isotopy  $\varphi_t : M \rightarrow M$  such that  $\varphi_t(S_0) = S_t$ , it is enough to find a non-autonomous vector field  $v_t$  on  $M$  such that  $\widehat{v}$  is tangent to  $\widehat{S}$ . A crucial observation is that the condition “ $\widehat{v}$  is tangent to  $\widehat{S}$ ” behaves well under partition of unities: if we can find, for every  $z = (x, t)$  a neighborhood  $U_z$  of  $z$  in  $M \times I$  and a vector field  $\widehat{v}_z$  on  $U_z$  of the form  $\widehat{v}_z(x, t) = v_{z,t}(x) + \partial_t$  such that “ $\widehat{v}_z$  is tangent to  $\widehat{S} \cap U_z$ ”, then, if  $\{\rho_z\}_{z \in M \times I}$  is a partition of unity subordinated to  $\{U_z\}_{z \in M \times I}$ , the field  $\widehat{v} := \sum_z \rho_z v_z$  is of the form  $\widehat{v}(x, t) = \sum_z \rho_z(x, t) v_{z,t}(x) + \partial_t$  and “ $\widehat{v}$  is tangent to  $\widehat{S}$ ”.

There is also an Hamiltonian version of Thom’s isotopy lemma (see [Sil06, Aur97, ST05]) which is well known by the symplectic community. A complete proof can be found in [ACL23].

**Theorem 4.3.** *Let  $(M, \omega)$  be a smooth, compact symplectic manifold and  $\{S_t\}_{t \in I}$  be a smooth family of symplectic submanifolds. Then there exists a Hamiltonian flow  $\varphi_t : M \rightarrow M$  such that  $\varphi_t(S_0) = S_t$  for every  $t \in I$ .*

**COROLLARY 4.4.** *Let  $(M, \omega)$  be a smooth, compact symplectic manifold and  $\{S_t\}_{t \in I}$  be a smooth family of symplectic submanifolds of dimension  $2k$ . For every  $t \in I$  denote by  $\mu_t \in \mathcal{P}(M)$  the measure defined by setting for every  $f \in \mathcal{C}^0(M, \mathbb{R})$ :*

$$\int_M f d\mu_t := \frac{\int_{S_t} f \omega^k}{\int_{S_t} \omega^k}. \quad (4.1)$$

*Then there exists a Hamiltonian flow  $\varphi_t : M \rightarrow M$  such that  $\mu_t = (\varphi_t)_\# \mu_0$  for every  $t \in I$ .*

*Proof.* This follows immediately from Theorem 4.3, using the fact that  $\varphi_t^* \omega = \omega$ .  $\square$

**4.2 The metric speed of the isotopy.** We now proceed calculating the metric speed of the curve  $(\mu_t)_{t \in [0,1]}$  from Corollary 4.4 using the results of Sect. 2, under the additional assumption that  $M$  is Kähler. The connection with Sect. 2 is given by Corollary 2.7.

**PROPOSITION 4.5.** *Let  $(M, \omega, J)$  be a smooth, compact, Kähler manifold and  $\{S_t\}_{t \in I}$  be a smooth family of codimension-two submanifolds such that  $(S_t, \omega|_{S_t}, J|_{S_t})$  is Kähler for every  $t \in I$ . Denote by  $\mu_t \in \mathcal{P}(M)$  the measure defined by (4.1) and*

by  $v_t$  the time-dependent vector field whose flow generates the corresponding Hamiltonian isotopy given by Corollary 4.4. For every  $t \in I$  let also denote by  $T(v_t)$  and  $N(v_t)$  the orthogonal projections (with respect to the Kähler Riemannian structure) of  $v_t$  onto  $TS_t$  and  $NS_t$  (the normal bundle of  $S_t$  in  $M$ ), respectively. Then,  $t \mapsto \mu_t$  is absolutely continuous and

$$|\dot{\mu}_t|^2 = \int_{S_t} \|N(v_t)\|^2 d\mu_t.$$

*Proof.* Using Corollary 2.7 we get that  $|\dot{\mu}_t|^2 = \|\mathbf{P}(v_t)\|_{L^2(\mu_t)}^2$  for almost every  $t \in I$ , where  $\mathbf{P}(v_t)$  denotes the projection of  $v_t$  onto

$$\overline{\{\nabla f : f \in \mathcal{C}^\infty(M)\}^{L^2(\mu_t)}} = \overline{\{\nabla g : g \in \mathcal{C}^\infty(S)\}^{L^2(\mu_t)}} \oplus L^2(NS, \mu_t),$$

where this last decomposition follows from Proposition 2.8.

We will prove that  $T(v_t)$  is  $\mu_t$ -divergence free, which implies that  $\mathbf{P}(v_t) = N(v_t)$ , from which the conclusion follows. To this end, denoting by  $h_t : M \rightarrow M$  the time-dependent Hamiltonian corresponding to the Hamiltonian flow  $\varphi_t$ , it will be sufficient to show that  $T(v_t)$  (a vector field on  $S_t$ ) is the Hamiltonian vector field of  $h_t|_{S_t}$ .

Let therefore  $w \in TS_t$ . We have

$$\begin{aligned} dh_t|_{S_t}(w) &= dh_t(w) = \omega(T(v_t) + N(v_t), w) \\ &= \omega(T(v_t) + N(v_t), -JJw) \quad (\text{using the fact that } J \text{ is a complex structure}) \\ &= -g(T(v_t) + N(v_t), Jw) \quad (\text{by definition of } g) \\ &= -g(T(v_t), Jw) \quad (\text{since } J|_{S_t} \text{ is a complex structure on } S_t) \\ &= \omega(T(v_t), w) \\ &= \omega|_{S_t}(T(v_t), w). \end{aligned}$$

This shows that  $T(v_t)$  is the Hamiltonian vector field of  $h_t|_{S_t}$  and concludes the proof.  $\square$

**REMARK 4.6.** In the previous proposition the compactness of  $M$  is used only to know that an Hamiltonian isotopy determining the family of submanifolds can be found. If we know, a priori, such an Hamiltonian isotopy (and the submanifolds of the family stay compact) the result remains of course true.

**LEMMA 4.7.** *Let  $\pi : L \rightarrow M$  be a rank- $k$  vector bundle on a Riemannian manifold  $(M, g)$ . Let also  $F : M \times I \rightarrow L$  be a smooth map such that for every  $t \in I$  the map  $f_t := F(\cdot, t) : M \rightarrow L$  is a section of  $L$  which is transversal to the zero section  $A \hookrightarrow L$ . (Using the language of Definition 4.1 this implies that  $\{S_t := f_t^{-1}(A)\}_{t \in I}$  is a smooth family of submanifolds of  $M$ .) Let  $v_t$  be a time-dependent vector field on  $M$  such that  $v_t + \partial_t$  is tangent to  $\widehat{S} := F^{-1}(A)$ . For every  $(x, t) \in \widehat{S}$ , denoting by  $N(v_t(x))$  the orthogonal projection of  $v_t(x)$  on the normal space to  $S_t$ , we have:*

$$N(v_t(x)) = -\left(\pi_2|_{f_t(x)} \circ d_x f_t|_{N_x S_t}\right)^{-1} \frac{\partial F}{\partial t}(x, t),$$

where, for every  $a \in A$ , the linear map  $\pi_2|_a : T_a L \simeq T_a A \oplus L_a \rightarrow L_a$  is the natural projection and  $L_a := \pi^{-1}(a)$ .

*Proof.* Assume  $v_t + \partial_t$  is tangent to  $\widehat{S}$  at  $(x, t)$ . Then  $D_{(x,t)}F(v_t(x) + \partial_t) \in T_{f_t(x)}A \subset T_{f_t(x)}L$  and, in particular,

$$\pi_2|_{f_t(x)}(D_{(x,t)}F(v_t(x) + \partial_t)) = 0.$$

Since for every  $t \in I$  the map  $f_t : M \rightarrow L$  is a section, we have  $\frac{\partial F}{\partial t}(x, t) \in L_{f_t(x)}$ , which implies  $\pi_2|_{f_t(x)}\left(\frac{\partial F}{\partial t}(x, t)\right) = \frac{\partial F}{\partial t}(x, t)$ . Substituting this in the above equation gives:

$$\begin{aligned} -\frac{\partial F}{\partial t}(x, t) &= \pi_2|_{f_t(x)}(D_{(x,t)}F(v_t(x))) \\ &= \pi_2|_{f_t(x)}(d_x f_t(v_t(x))) \\ &= \pi_2|_{f_t(x)}(d_x f_t(N(v_t(x)) + d_x f_t T(v_t(x)))) \\ &= \pi_2|_{f_t(x)}(d_x f_t(N(v_t(x)))) \quad (\text{since } d_x f_t(T_x S_t) \subset T_{f_t(x)}A). \end{aligned} \tag{4.2}$$

The transversality assumption  $f_t \pitchfork A$  implies that  $\pi_2|_{f_t(x)} \circ d_x f_t|_{N_x S_t}$  is invertible and (4.2) gives the statement.  $\square$

It will be useful for us to have a formulation of the previous result using vector bundles trivializations.

LEMMA 4.8. *Under the same assumptions and notations of Lemma 4.7, let  $z \in M$ ,  $U$  be a neighborhood of  $z$  and  $\psi : L|_U \rightarrow U \times \mathbb{R}^k$  be a vector bundle trivialization. Denote by  $\widetilde{F} : U \times I \rightarrow \mathbb{R}^k$  the map defined by*

$$\psi(F(x, t)) = (x, \widetilde{F}(x, t)) \quad \forall (x, t) \in U \times I. \tag{4.3}$$

For every  $t \in I$  let also  $\widetilde{f}_t := \widetilde{F}(\cdot, t) : U \rightarrow \mathbb{R}^k$ . Then, for every  $(x, t) \in \widehat{S} \cap (U \times I)$  we have:

$$N(v_t(x)) = \left(d_x \widetilde{f}_t|_{N_x S_t}\right)^{-1} \frac{\partial \widetilde{F}}{\partial t}(x, t).$$

*Proof.* The proof follows easily by rewriting the previous proof in coordinates.  $\square$

## 5 The optimal transport problem between algebraic hypersurfaces

We now introduce the optimal transport problem between two algebraic hypersurfaces adopting the dynamical formulation. This will produce a new metric on the space of homogeneous polynomials.

**DEFINITION 5.1** (Admissible curves and their Energy). Let  $I = [0, 1]$  and  $\mu : P_{n,d} \rightarrow \mathcal{P}(\mathbb{C}P^n)$  be the map from Theorem 3.8. We define the set of *admissible curves* to be

$$\Omega := \{\gamma : I \rightarrow P_{n,d} \mid \mu \circ \gamma \in \text{AC}^2(I, \mathcal{P}_2(\mathbb{C}P^n))\}.$$

If  $\gamma \in \Omega$  is an admissible curve, we write  $t \mapsto \mu_t := \mu(\gamma(t))$  and define the *Energy* of  $\gamma$  by

$$\mathcal{E}(\gamma) := \int_I |\dot{\mu}_t|^2 dt.$$

It is worth stressing again that the metric speed  $|\dot{\mu}_t|$ , defined in Sect. 2, is calculated with respect to the  $W_2$ -distance as denoted by  $\mathcal{P}_2(\mathbb{C}P^n)$ .

**DEFINITION 5.2** (The inner Wasserstein distance on  $P_{n,d}$ ). We define the inner Wasserstein distance  $W_2^{\text{in}}$  on  $P_{n,d}$  by

$$W_2^{\text{in}}(p_0, p_1) := \inf\{\mathcal{E}(\gamma)^{1/2} \mid \gamma \in \Omega, \gamma(0) = p_0, \gamma(1) = p_1\}.$$

We will prove below (Corollary 5.14) that the function  $W_2^{\text{in}}$  is finite, i.e. that any two points in  $P_{n,d}$  can be joined by an admissible curve, and that it is a distance. In order to do this we will first need to investigate the behaviour of the Energy on curves in  $P_{n,d} \setminus \Delta_{n,d}$ , as done in the next section.

Similarly to Definition 5.2, for any other  $q \geq 1$ , over  $P_{n,d}$  one can consider also the inner  $W_q$  distance defined by

$$W_q^{\text{in}}(p_0, p_1)^q := \inf \int_I |\dot{\mu}_t|_q^q dt,$$

where the infimum runs over the set of curves  $\gamma \in \text{AC}^q(I; \mathcal{P}_q(\mathbb{C}P^n))$  such that  $\gamma(0) = p_0$  and  $\gamma(1) = p_1$ . Recall that if no exponent is specified, the metric speed has to be considered in  $\mathcal{P}_2(\mathbb{C}P^n)$ .

### 5.1 The Wasserstein–Hermitian structure on the set of regular hypersurfaces.

**Theorem 5.3.** Let  $c : I \rightarrow \mathbb{C}[z_0, \dots, z_n]_{(d)} \setminus D_{n,d}$  be a  $\mathcal{C}^1$  curve. For every  $t \in I$  denote by  $\mu_t = \mu([c_t]) \in \mathcal{P}(\mathbb{C}P^n)$  the probability measure defined by (3.7) and by

$$\widehat{Z}(c_t) := \{z \in S^{2n+1} \mid c_t(z) = 0\}$$

the zero set of  $c_t$  on the sphere (this is a smooth manifold by assumption). Then

$$|\dot{\mu}_t|^2 = \frac{1}{d \cdot \text{vol}(S^{2n-1})} \int_{\widehat{Z}(c_t)} \frac{|\dot{c}_t(b)|^2}{\|\nabla^{\mathbb{C}} c_t(b)\|^2} \text{vol}_{\widehat{Z}(c_t)}(db), \quad (5.1)$$

where  $\nabla^{\mathbb{C}} p(b) := (\frac{\partial p}{\partial z_0}(b), \dots, \frac{\partial p}{\partial z_n}(b))$  denotes the “complex gradient” of  $p$ .

REMARK 5.4. Given  $p \in \mathbb{C}[z_0, \dots, z_n]_{(d)} \setminus D_{n,d}$  and  $q_1, q_2 \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$ , the function  $I : S^{2n+1} \supset \widehat{Z}(p) \rightarrow \mathbb{C}$  defined by

$$I(z) := \frac{q_1(z)\overline{q_2(z)}}{\|\nabla^{\mathbb{C}} p(z)\|^2} \quad (5.2)$$

is  $U(1)$ -invariant, i.e.  $I(z) = I(e^{i\theta}z)$ , and therefore it descends to a well defined function, that we still denote by the same symbol  $I : \mathbb{CP}^n \supset Z(p) \rightarrow \mathbb{C}$ . By the smooth coarea formula, since  $\widehat{Z}(p) \rightarrow Z(p)$  is a Riemannian submersion,

$$\int_{\widehat{Z}(p)} I(z) \operatorname{vol}_{Z(p)}(dz) = 2\pi \int_{Z(p)} I(z) \operatorname{vol}_{Z(p)}(d[z]). \quad (5.3)$$

In the sequel we will often simply write these integrals in the form of the right hand side of (5.3), i.e. with the convention that the representatives for the evaluation of the integrand should be taken in the unit sphere. Alternatively, one can notice that the function (5.2) is itself the restriction to  $\widehat{Z}(p)$  of a function  $\{z \in \mathbb{C}^{n+1} \setminus \{0\} \mid p(z) = 0\} \rightarrow \mathbb{C}$  that is homogenous of degree  $-2$ . Hence, dividing by  $\|z\|^2$ , we get a function homogeneous of degree 0, that descends to the same function on  $Z(p) \subset \mathbb{CP}^n$ , and the integral in (5.3) can also be written as

$$2\pi \int_{Z(p)} I(z) \operatorname{vol}_{Z(p)}(d[z]) = 2\pi \int_{Z(p)} \frac{q_1(z)\overline{q_2(z)}}{\|\nabla^{\mathbb{C}} p(z)\|^2 \|z\|^2} \operatorname{vol}_{Z(p)}(d[z]).$$

*Proof of Theorem 5.3.* For this proof we make extensive use of the notation from the [Appendix](#).

**Step 1.** Let us observe first that  $\{Z_t := Z(c_t)\}_{t \in I}$  is a smooth family of submanifolds in  $\mathbb{CP}^n$ , in the sense of Definition 4.1. To this end, denote by  $\sigma_t := \sigma_{c_t} \in \Gamma(L_{n,d}, \mathbb{CP}^n)$  the section defined by the polynomial  $c_t$  (see Sect. A.3). We define the map  $F : \mathbb{CP}^n \times I \rightarrow L_{n,d}$  by

$$F([z], t) := \sigma_t([z]).$$

The condition  $c_t \notin D_{n,d}$  implies that for every  $t \in I$  we have  $\sigma_t \pitchfork A$  and  $Z_t = \sigma_t^{-1}(A)$  is a smooth submanifold. Moreover, for every  $t \in I$  the manifold  $Z_t$  is a smooth complex hypersurface, hence Kähler. We are therefore in the setting of Proposition 4.5, which gives:

$$|\dot{\mu}_t|^2 = \int_{Z_t} \|N(v_t)\|^2 d\mu_t = \frac{1}{\operatorname{vol}(Z_t)} \int_{Z_t} \|N(v_t)\|^2 \omega_{\mathbb{F}\mathbb{S}}^{n-1}, \quad (5.4)$$

where  $v_t$  is the time-dependent vector field whose flow generates the Hamiltonian isotopy given by Corollary 4.4 and  $N(v_t)$  is its orthogonal projection on  $NZ_t$ .

**Step 2.** Without loss of generality, we prove (5.1) for  $t = 0$ . We denote by  $p_0 := c_0$  and  $p_1 := \dot{c}_0$  so that  $c_t = p_0 + tp_1 + o(t)$ . Since the map associating to a polynomial  $p$  the corresponding section  $\sigma_p$  of  $L_{n,d}$  is linear (see Sect. A.3) we have  $\sigma_t = \sigma_0 + t\sigma_1 + o(t)$ , where  $\sigma_0 := \sigma_{p_0}$  and  $\sigma_1 := \sigma_{p_1}$ .

Let  $[b] \in Z_0$  with  $b \in S^{2n+1}$ . In order to compute  $N(v_0([b]))$  we use Lemma 4.8. Denote by  $\psi_b : L|_{U_b} \rightarrow U_b \times \mathbb{C}$  the trivialization given by (A.5). Then, by (A.7),

$$\tilde{F}([z], t) = \left( [z], \frac{\sigma_0([z]) + t\sigma_1([z]) + o(t)}{\sigma_b([z])} \right) = \left( [z], \frac{p_0(z) + tp_1(z) + o(t)}{\langle z, b \rangle^d} \right),$$

where  $\sigma_b$  is the section defined in (A.6). By Lemma 4.8, we have

$$N(v_0([b])) = - \left( D_{[b]} \tilde{f}_0|_{N_{[b]}Z_0} \right)^{-1} \frac{\partial \tilde{F}}{\partial t}([b], 0),$$

where, according to the definition (4.3),

$$\tilde{f}_t([z]) = \frac{p_0(z) + tp_1(z) + o(t)}{\langle z, b \rangle^d}.$$

In order to perform the computation of the norm of  $N(v_0([b]))$ , we use the parametrization  $\phi_b : \mathbb{C}^n \rightarrow U_b$  introduced in (A.1). In fact, by (A.2), we have

$$\|N(v_0([b]))\|_{(T_{[b]}\mathbb{C}P^n, h_{\text{FS}, [b]})} = \|D_0\phi_b^{-1}N(v_0([b]))\|_{(\mathbb{C}^n, h_{\text{std}})}. \quad (5.5)$$

Set now  $\tilde{g}_t(v) := \tilde{f}_t(\phi_b(v))$  so that

$$\tilde{G}(v, t) := \tilde{F}(\phi_b(v), t) = \tilde{g}_t(v) = \frac{p_0 \left( R \begin{pmatrix} 1 \\ v \end{pmatrix} \right) + tp_1 \left( R \begin{pmatrix} 1 \\ v \end{pmatrix} \right) + o(t)}{\left\langle R \begin{pmatrix} 1 \\ v \end{pmatrix}, b \right\rangle^d}, \quad (5.6)$$

where  $R$  is a unitary matrix such that  $Re_0 = b$ . Denote also by

$$V := D_0\phi_b^{-1}(N_{[b]}S_0) = (T_0\tilde{g}^{-1}(0))^{\perp_{\text{std}}}.$$

With these notations, by (5.5), we have

$$\|N(v_0([b]))\|_{(T_{[b]}\mathbb{C}P^n, h_{\text{FS}, [b]})} = \left\| (D_0\tilde{g}_0|_V)^{-1} \frac{\partial \tilde{G}}{\partial t}(0, 0) \right\|_{(\mathbb{C}^n, h_{\text{std}})}.$$

To start with

$$\frac{\partial \tilde{G}}{\partial t}(0, 0) = \frac{p_1 \left( R \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)}{\left\langle R \begin{pmatrix} 1 \\ 0 \end{pmatrix}, b \right\rangle^d} = \frac{p_1(b)}{\|b\|^2} = p_1(b),$$

since  $Re_0 = b$  and  $b \in S^{2n+1}$ .

Notice now that, the complex-linear operator  $D_0\tilde{g}_0 : \mathbb{C}^n \rightarrow \mathbb{C}$  acts as:

$$w = (w_1, \dots, w_n) \mapsto \sum_{j=1}^n \frac{\partial \tilde{g}_0}{\partial v_j}(0) w_j.$$

Therefore, defining

$$\nabla^{\mathbb{C}} \tilde{g}_0(0) := \left( \frac{\partial \tilde{g}_0}{\partial v_1}(0), \dots, \frac{\partial \tilde{g}_0}{\partial v_n}(0) \right),$$

using the standard Hermitian structure we can write

$$D_0\tilde{g}_0 : w \mapsto \langle w, \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)} \rangle_{\mathbb{C}^n}.$$

Consequently, since  $T_0\tilde{g}_0^{-1}(0) = \ker(D_0\tilde{g}_0) = \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)}^\perp$ , we see that  $V = (T_0\tilde{g}_0^{-1}(0))^\perp$  can be identified with

$$V = \text{span}_{\mathbb{C}} \{ \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)} \}.$$

In particular,  $D_0\tilde{g}_0|_V$  can be described as the linear map

$$D_0\tilde{g}_0|_V : \lambda \cdot \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)} \mapsto \langle \lambda \cdot \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)}, \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)} \rangle_{\mathbb{C}^n} = \lambda \|\nabla^{\mathbb{C}} \tilde{g}_0(0)\|^2.$$

This implies that

$$(D_0\tilde{g}_0|_V)^{-1} \frac{\partial \tilde{G}}{\partial t}(0,0) = \frac{p_1(b)}{\|\nabla^{\mathbb{C}} \tilde{g}_0(0)\|^2} \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)},$$

and, consequently:

$$\|N(v_0([b]))\|^2 = \frac{|p_1(b)|^2}{\|\nabla^{\mathbb{C}} \tilde{g}_0(0)\|^2}. \quad (5.7)$$

Given  $w \in \mathbb{C}^n$ , using (5.6), for every  $w \in \mathbb{C}^n$  we have

$$\begin{aligned} D_0\tilde{g}_0 w &= \langle w, \overline{\nabla^{\mathbb{C}} \tilde{g}_0(0)} \rangle_{\mathbb{C}^n} \\ &= D_0 p_0 R \begin{pmatrix} 0 \\ w \end{pmatrix} \\ &= \langle R \begin{pmatrix} 0 \\ w \end{pmatrix}, \overline{\nabla^{\mathbb{C}} p_0(b)} \rangle_{\mathbb{C}^{n+1}} = (*). \end{aligned} \quad (5.8)$$

Observe now that, since  $Re_0 = b$  and  $R$  is a unitary matrix,  $R \begin{pmatrix} 0 \\ w \end{pmatrix} \in b^\perp$  for every  $w \in \mathbb{C}^n$  and, consequently,

$$(*) = \langle R \begin{pmatrix} 0 \\ w \end{pmatrix}, \text{proj}_{b^\perp} \overline{\nabla^{\mathbb{C}} p_0(b)} \rangle_{\mathbb{C}^{n+1}} = \langle \begin{pmatrix} 0 \\ w \end{pmatrix}, R^* \text{proj}_{b^\perp} \overline{\nabla^{\mathbb{C}} p_0(b)} \rangle_{\mathbb{C}^{n+1}}.$$

Comparing with (5.8), and using the fact that  $R^*$  is a unitary matrix, we have:

$$\|\nabla^{\mathbb{C}} \tilde{g}_0(0)\|^2 = \|\text{proj}_{b^\perp} \overline{\nabla^{\mathbb{C}} p_0(b)}\|^2.$$

Observe now that, for a homogeneous polynomial  $p$ , Euler's identity  $\sum_j z_j \partial_j p(z) = \deg(p)p(z)$  can be written as

$$\deg(p)p(z) = \langle z, \overline{\nabla^{\mathbb{C}} p(z)} \rangle.$$

In particular, if  $z = b$  is a zero of  $p$ , we have  $\langle b, \overline{\nabla^{\mathbb{C}} p(b)} \rangle = 0$ . Consequently

$$\begin{aligned} \overline{\nabla^{\mathbb{C}} p(b)} &= \langle \overline{\nabla^{\mathbb{C}} p(b)}, b \rangle \cdot b + \text{proj}_{b^\perp} \overline{\nabla^{\mathbb{C}} p(b)} \\ &= \overline{\langle b, \overline{\nabla^{\mathbb{C}} p(b)} \rangle} \cdot b + \text{proj}_{b^\perp} \overline{\nabla^{\mathbb{C}} p(b)} \\ &= \text{proj}_{b^\perp} \overline{\nabla^{\mathbb{C}} p(b)}. \end{aligned}$$

Since  $\|\overline{\nabla^{\mathbb{C}} p(b)}\| = \|\nabla^{\mathbb{C}} p(b)\|$ , together with (5.7), the expression (5.4) gives:

$$|\dot{\mu}_t|^2 = \frac{1}{\text{vol}(Z_t)} \int_{Z_t} \frac{|\dot{c}_t(b)|^2}{\|\nabla^{\mathbb{C}} c_t(b)\|^2} \text{vol}_{Z_t}(d[b]),$$

where in the integrand, which is a function of  $[b]$ , we have chosen representatives  $b \in S^{2n+1}$  (see Remark 5.4). Using the fact that the quotient map  $S^{2n+1} \rightarrow \mathbb{C}P^n$  is a Riemannian submersion with fibers which are isometric to circles and the fact that  $\text{vol}(Z_t) = d \cdot \text{vol}(\mathbb{C}P^{n-1})$  (Lemma 3.7), one can immediately rewrite the last integral as in (5.1). This concludes the proof.  $\square$

Let now  $p \in \mathbb{C}[z_0, \dots, z_n]_{(d)} \setminus D_{n,d}$ . Motivated by (5.1), we define a Hermitian form  $\widehat{h}_p$  on  $T_p \mathbb{C}[z_0, \dots, z_n]_{(d)} \simeq \mathbb{C}[z_0, \dots, z_n]_{(d)}$  by setting:

$$\widehat{h}_p(q_1, q_2) := \int_{\widehat{Z}(p)} \frac{q_1(b) \overline{q_2(b)}}{\|\nabla^{\mathbb{C}} p(b)\|^2} \text{vol}_{\widehat{Z}(p)}(db). \quad (5.9)$$

This Hermitian form *is not* positive definite, since  $\widehat{h}_p(p, p) = 0$ . In fact we have the following.

LEMMA 5.5. *For every  $p \in \mathbb{C}[z_0, \dots, z_n]_{(d)} \setminus D_{n,d}$  we have  $\ker(\widehat{h}_p) = \text{span}_{\mathbb{C}}\{p\}$ .*

*Proof.* Let  $q \in \ker(\widehat{h}_p)$ . Then for all  $u \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$  we have  $\widehat{h}_p(q, u) = 0$ . In particular this holds true for  $u = q$ . Observe that, since  $p \notin D_{n,d}$ , the function  $b \mapsto \|\nabla^{\mathbb{C}} p(b)\|^2$  is never zero on  $\widehat{Z}(p)$ , otherwise  $p$  would have a singular projective zero. In order for the integral in (5.9) to vanish, we must therefore have  $|q|^2|_{\widehat{Z}(p)} \equiv 0$ . This means that  $Z(p) \subseteq Z(q)$ . Writing  $q = q_1 \cdots q_\ell$  as a product of irreducible factors, we must have  $Z(p) = Z(q_j)$  for some  $q_j$  and, since  $p$  is irreducible (because  $Z(p)$  is smooth),  $q_j = c_j p$  for some  $c_j \in \mathbb{C}$  nonzero. But then, since  $\deg(q) = \deg(p)$  all the other irreducible factors of  $q$  must have zero degree, which implies that  $q$  is a multiple of  $p$ .  $\square$

We define now a Hermitian structure on  $P_{n,d} \setminus \Delta_{n,d}$ , as follows. Denote by

$$\pi : \mathbb{C}[z_0, \dots, z_n]_{(d)} \setminus \{0\} \rightarrow P_{n,d}$$

the quotient map. Given  $[p] \in P_{n,d} \setminus \Delta_{n,d}$  and  $v_1, v_2 \in T_{[p]}(P_{n,d})$ , let  $p \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$  be a representative and  $q_1, q_2 \in T_p(\mathbb{C}[z_0, \dots, z_n]_{(d)})$  such that  $D_p \pi q_i = v_i$ . We define

$$(h)_{[p]}(v_1, v_2) := \widehat{h}_p(q_1, q_2). \quad (5.10)$$

This definition does not depend on the choice of the representatives  $p$  and  $q_1, q_2$ . In fact, let  $\lambda p \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$  be another representative for  $[p]$ , where  $\lambda \in \mathbb{C} \setminus \{0\}$ . Then  $\lambda q_1$  and  $\lambda q_2$  are two vectors in  $T_{\lambda p}(\mathbb{C}[z_0, \dots, z_n]_{(d)})$  such that  $D_{\lambda p} \pi \lambda q_i = v_i$  and

$$\begin{aligned} \widehat{h}_{\lambda p}(\lambda q_1, \lambda q_2) &= \int_{\widehat{Z}(\lambda p)} \frac{\lambda q_1 \overline{\lambda q_2}}{\|\nabla^{\mathbb{C}}(\lambda p)\|^2} \text{vol}_{\widehat{Z}(\lambda p)} \\ &= \int_{\widehat{Z}(p)} \frac{|\lambda|^2 q_1 \overline{q_2}}{\|\lambda \nabla^{\mathbb{C}} p\|^2} \text{vol}_{\widehat{Z}(p)} \\ &= \widehat{h}_p(q_1, q_2). \end{aligned}$$

Lemma 5.5 implies now that the Hermitian form  $h$  is positive definite.

DEFINITION 5.6 (The Wasserstein–Hermitian structure on the set of nonsingular hypersurfaces). We call the form  $h$  defined in (5.10) the *Wasserstein–Hermitian structure* on  $P_{n,d} \setminus \Delta_{n,d}$ .

We notice that the integrand in (5.9) is a well defined function on  $Z(p)$ , therefore, using the fact that the Riemannian submersion  $S^{2n+1} \rightarrow \mathbb{C}P^n$  restricts to a Riemannian submersion  $\widehat{Z}(p) \rightarrow Z(p)$  (always with fibers that are circles), we can rewrite it as:

$$\begin{aligned} \widehat{h}_p(q_1, q_2) &= \int_{\widehat{Z}(p)} \frac{q_1(b) \overline{q_2(b)}}{\|\nabla^{\mathbb{C}} p(b)\|^2} \text{vol}_{\widehat{Z}(p)}(db) \\ &= 2\pi \int_{Z(p)} \frac{q_1(b) \overline{q_2(b)}}{\|\nabla^{\mathbb{C}} p(b)\|^2} \text{vol}_{Z(p)}(db) \\ &= \frac{d \cdot \text{vol}(S^{2n-1})}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} \int_{Z(p)} \frac{q_1(b) \overline{q_2(b)}}{\|\nabla^{\mathbb{C}} p(b)\|^2} \text{vol}_{Z(p)}(db). \end{aligned} \quad (5.11)$$

The last expression reconciles with the notation adopted in Sect. 1, but the expression in (5.9) is more practical to work with.

PROPOSITION 5.7. *The Hermitian form  $h$  on  $P_{n,d} \setminus \Delta_{n,d}$  is smooth.*

*Proof.* Put  $\Omega = \mathbb{C}[z_0, \dots, z_n]_{(d)} \setminus D_{n,d}$ . Then it is clear that it is sufficient to show that the form

$$\widehat{h} : \Omega \times \mathbb{C}[z_0, \dots, z_n]_{(d)}^2 \longrightarrow \mathbb{C}, \quad (p, q_1, q_2) \mapsto \widehat{h}_p(q_1, q_2),$$

is smooth with respect to all its variables. Consider the solution manifold on the sphere:

$$\widehat{V} = \left\{ (p, x) \in \Omega \times S^{2n+1} : x \in \widehat{Z}(p) \right\}.$$

It is a smooth manifold equipped with the projection  $\pi_1 : \widehat{V} \rightarrow \Omega$  on the first factor. Now  $\pi_1$  is proper (smooth with compact fibers) and by the Ehresmann Fibration Theorem is a locally trivial fibration. In other words, around any point  $p_0 \in \Omega$  we can find a small open set  $U$  and a diffeomorphism  $\Phi : \pi_1^{-1}(U) \rightarrow U \times \widehat{Z}(p_0)$  mapping the fibers of  $\pi_1$  to the fibers of the projection on  $U$ . In particular the integration domains vary smoothly together with the volume forms  $p \mapsto \text{vol}_{\widehat{Z}(p)}$ . Indeed, in  $U$ , the maps  $\Psi_p := \Phi^{-1}(p, \cdot) : \widehat{Z}(p_0) \rightarrow \widehat{Z}(p)$  give a smooth family of diffeomorphisms. Defining  $\sigma_p := \Psi_p^* \text{vol}_{\widehat{Z}(p)}$  we get a smooth family of volume forms on  $\widehat{Z}(p_0)$  such that

$$\widehat{h}_p(q_1, q_2) = \int_{\widehat{Z}(p_0)} \frac{q_1 \bar{q}_2}{\|\nabla^{\mathbb{C}p}\|^2 \circ \Psi_p} \sigma_p.$$

This implies the required smoothness.  $\square$

**COROLLARY 5.8.** *Let  $\gamma : I \rightarrow P_{n,d} \setminus \Delta_{n,d}$  be a  $\mathcal{C}^1$  curve. Then  $\gamma$  is admissible in the sense of Definition 5.1 and, denoting by  $t \mapsto \mu_t := \mu(\gamma(t))$ , we have*

$$\mathcal{E}(\gamma) = \frac{1}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} \int_0^1 (\mathbf{h})_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) dt.$$

*Proof.* Let  $c : I \rightarrow \mathbb{C}[z_0, \dots, z_n]_{(d)}$  be a smooth curve lifting  $\gamma$  (i.e. such that  $[c(t)] = \gamma(t)$ ). Then, by Theorem 5.3, we have

$$\begin{aligned} \mathcal{E}(\gamma) &= \int_0^1 |\dot{\mu}_t|^2 dt \\ &= \frac{1}{d \cdot \text{vol}(S^{2n-1})} \int_0^1 \widehat{h}_{c(t)}(\dot{c}(t), \dot{c}(t)) dt \\ &= \frac{1}{d \cdot \text{vol}(S^{2n-1})} \int_0^1 \frac{d \cdot \text{vol}(S^{2n-1})}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} (\mathbf{h})_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) dt \quad (\text{by (5.11)}) \\ &= \frac{1}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} \int_0^1 (\mathbf{h})_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)), \end{aligned}$$

where the last identity follows from the definition of the Wasserstein–Hermitian structure and the fact that  $D_{c(t)} \pi \dot{c}(t) = \dot{\gamma}(t)$  for every  $t \in I$ .  $\square$

**REMARK 5.9.** In the case  $d = 1$  the discriminant  $\Delta_{n,1}$  is empty and the Wasserstein metric coincides with the Fubini–Study metric (up to multiples). To see this, we identify  $\mathbb{C}P^n \simeq (\mathbb{C}P^n)^*$  and observe that the Fubini–Study metric is  $U(n+1)$ -invariant; moreover, since the action of  $U(n) = \text{Stab}_{[e_0]}$  on  $T_{[e_0]}\mathbb{C}P^n \simeq U(n)$  is the standard action, this action is irreducible. In particular, by Schur’s Lemma, the Fubini–Study metric is the unique  $U(n+1)$ -invariant Riemannian metric on  $\mathbb{C}P^n$  (up to multiples). Since our Wasserstein–Hermitian structure is  $U(n+1)$ -invariant, it must therefore be a multiple of the Fubini–Study one.

**5.2 The Wasserstein–Kähler structure.** In this section we prove that the imaginary part of the Wasserstein–Hermitian structure  $\mathfrak{h}$  from Definition 5.6 is a closed 2-form, and therefore  $(P_{n,d} \setminus \Delta_{n,d}, \mathfrak{h})$  is a Kähler manifold.

Before proceeding, let us recall the notion of *fiber integral* of a differential form. Let  $\pi : W \rightarrow U$  be a proper submersion with  $f$ -dimensional fibers and  $\alpha \in \Omega^a(W)$  be a differential form. The fiber integral of  $\alpha$ , denoted by  $\pi_*\alpha \in \Omega^{a-f}(U)$ , is the differential form defined as follows. Given  $u \in U$  and  $v_1, \dots, v_{a-f} \in T_u U$ , we first choose vector fields  $\tilde{v}_1, \dots, \tilde{v}_{a-f}$  on  $W$  along  $\pi^{-1}(u)$  such that  $D\tilde{v}_j = v_j$  for every  $j = 1, \dots, a-f$ . (These vector fields exist, since  $\pi$  is a submersion; the construction will not depend on this choice.) Then we consider the interior product  $\iota_{\tilde{v}_1} \cdots \iota_{\tilde{v}_{a-f}} \alpha$ , which is a top form on  $\pi^{-1}(u)$ , and we set

$$(\pi_*\alpha)_u(v_1, \dots, v_{a-f}) := \int_{\pi^{-1}(u)} \iota_{\tilde{v}_1} \cdots \iota_{\tilde{v}_{a-f}} \alpha.$$

This defines a smooth differential form on  $U$ .

Let now  $V = \{(p, z) \mid p(z) = 0\} \subset P_{n,d} \times \mathbb{C}P^n$  be the solution variety, together with the restrictions

$$\pi_1 : V \rightarrow P_{n,d} \quad \text{and} \quad \pi_2 : V \rightarrow \mathbb{C}P^n$$

of the projections on the two factors. Set  $W := V \setminus \pi_1^{-1}(\Delta_{n,d})$  and  $U = P_{n,d} \setminus \Delta_{n,d}$ . We will apply the previous construction to  $\pi_2^* \omega_{\mathbb{C}P^n}^n / n! \in \Omega^{n,n}(W) \subset \Omega^{2n}(W)$  (note that this is the pull-back of the volume form of  $\mathbb{C}P^n$ ).

**Theorem 5.10** (The Wasserstein–Kähler structure). *Let  $\mathfrak{h}$  be the Wasserstein–Hermitian structure and let  $\sigma := \text{Im}(\mathfrak{h}) \in \Omega^2(P_{n,d} \setminus \Delta_{n,d})$ . Then*

$$\sigma = (\pi_1)_* \pi_2^* \left( \frac{\omega_{\mathbb{C}P^n}^n}{n!} \right) \in \Omega^{1,1}(P_{n,d} \setminus \Delta_{n,d}). \quad (5.12)$$

*In particular  $(P_{n,d} \setminus \Delta_{n,d}, \mathfrak{h})$  is a Kähler manifold.*

*Proof.* Let  $p \in P_{n,d} \setminus \Delta_{n,d}$  and  $q_1, q_2 \in T_p P_{n,d}$ . In order to compute  $\sigma_p(q_1, q_2)$ , we need to lift  $q_1, q_2$  to vector fields on  $V$  along  $\pi_1^{-1}(p) \simeq Z(p)$ . Working in homogenous coordinates, since  $TV = \{(\dot{p}, \dot{z}) \mid \dot{p}(z) + D_z p \dot{z} = 0\}$ , it is immediate to see that the vector fields defined by

$$\tilde{q}_i(p, z) := (q_i, -q_i(z)\xi(z)), \quad \text{where} \quad \xi(z) := \frac{\overline{\nabla^{\mathbb{C}} p(z)}}{\|\nabla^{\mathbb{C}} p(z)\|},$$

are two such lifts.

Now, recall the following basic fact from differential geometry. Given a 2-form  $\beta$  and two vector fields  $X, Y$ , we have

$$\iota_X \iota_Y \beta^n = n\beta(X, Y)\beta^{n-1} + n(n-1)\iota_X \beta \wedge \iota_Y \beta \wedge \beta^{n-2}. \quad (5.13)$$

In the case  $\beta = \pi_2^* \omega_{\text{FS}}$ , we have

$$\iota_{\tilde{q}_i} \beta = \omega_{\text{FS}}(-q_i(z)\xi(z), \cdot).$$

Since  $\xi(z) \perp T_z Z(p)$ , it follows that  $\omega_{\text{FS}}(-q_i(z)\xi(z), \tau) = \text{Im}(h_{\text{FS}}(-q_i(z)\xi(z), \tau)) = 0$  for every  $\tau \in T_z Z(p)$ , i.e.

$$\iota_{\tilde{q}_i} \beta|_{\pi^{-1}(p)} = 0. \quad (5.14)$$

In particular, using (5.13) and (5.14),

$$\begin{aligned} \sigma_p(q_1, q_2) &= \int_{\pi_2^{-1}(p)} \frac{1}{n!} \iota_{\tilde{q}_1} \iota_{\tilde{q}_2} \beta^n \\ &= \frac{n}{n!} \int_{Z(p)} \omega_{\text{FS}}(-q_1(z)\xi(z), -q_2(z)\xi(z)) \omega_{\text{FS}}^{n-1} \\ &= \int_{Z(p)} \frac{\text{Im}(q_1(z)\overline{q_2(z)})}{\|\nabla^{\mathbb{C}} p(z)\|^2} \frac{\omega_{\text{FS}}^{n-1}}{(n-1)!} \\ &= \int_{Z(p)} \frac{\text{Im}(q_1(z)\overline{q_2(z)})}{\|\nabla^{\mathbb{C}} p(z)\|^2} \text{vol}_{Z(p)}(dz) \\ &= \text{Im}(\mathbf{h}_p(q_1, q_2)), \end{aligned}$$

which proves (5.12). The fact that  $\sigma$  is closed follows now from the fact that integration along the fibers commutes with exterior derivative.  $\square$

**5.3 Properties of  $W_2^{\text{in}}$ .** We are now in the position of proving that for any two points  $p_0, p_1 \in P_{n,d}$ , the function  $W_2^{\text{in}}(p_0, p_1)$  is finite and it defines a distance. We prove first some auxiliary results, using tools from semialgebraic geometry (see Definition 3.5).

LEMMA 5.11. *Let  $f : [0, r] \rightarrow \mathbb{R}^\ell$  be a continuous semialgebraic curve. Then there exists  $m \in \mathbb{N}$  and  $0 < r_0 \leq r$  such that the curve  $\tilde{f}(s) := f(s^m)$  is  $\mathcal{C}^1$  on  $[0, r_0]$ .*

*Proof.* Denote by  $(f_1, \dots, f_\ell)$  the components of  $f$ . Since each of them is continuous and semialgebraic, for every  $i$  there exists  $q_i \in \mathbb{N}$ ,  $r_i \leq r$  and a  $\mathcal{C}^1$  function  $g_i : [0, r_i] \rightarrow \mathbb{R}$  such that:

$$f_i(t) = f_i(0) + t^{\frac{1}{q_i}} g_i(t) \quad \text{for all } 0 \leq t \leq r_i.$$

Letting  $m = \max_i q_i$  and  $r_0 = \min_i r_i$  we see that

$$f_i(s^m) = f_i(0) + t^{\frac{m}{q_i}} g_i(s^m) \quad \text{for all } 0 \leq t \leq r_0.$$

Since  $\frac{m}{q_i} \geq 1$  and  $g_i$  is  $\mathcal{C}^1$ , this implies that  $\tilde{f}$  is  $\mathcal{C}^1$  on  $[0, r_0]$ .  $\square$

LEMMA 5.12. *Let  $\gamma : [0, 1] \rightarrow P_{n,d}$  be a continuous semialgebraic curve such that  $\gamma(0) \in \Delta_{n,d}$  and  $\gamma(t) \in P_{n,d} \setminus \Delta_{n,d}$  for all  $t \in (0, 1]$ . Then there exist  $0 < r \leq 1$  and a continuous semialgebraic curve  $b : [0, r] \rightarrow \mathbb{C}P^n$  such that  $b|_{(0,r]}$  is  $\mathcal{C}^1$  and*

$$b(t) \in \operatorname{argmax}_{b \in Z(\gamma_t)} \frac{|\dot{\gamma}_t(b)|^2}{\|\nabla^{\mathbb{C}} \gamma_t(b)\|^2}. \quad (5.15)$$

Moreover, for every  $0 < t \leq r$  we have:

$$\frac{|\dot{\gamma}_t(b(t))|^2}{\|\nabla^{\mathbb{C}} \gamma_t(b(t))\|^2} \leq |\dot{b}(t)|^2. \quad (5.16)$$

*Proof.* Notice first that the function

$$t \mapsto \max_{\theta \in Z(\gamma_t)} \frac{|\dot{\gamma}_t(\theta)|^2}{\|\nabla^{\mathbb{C}} \gamma_t(\theta)\|^2}$$

is well defined on  $(0, 1)$ , since  $\gamma(0, 1] \notin \Delta_{n,d}$ , and semialgebraic (its graph is semialgebraic). In particular the following set  $S \subset (0, 1) \times \mathbb{C}P^n$  is semialgebraic:

$$S := \left\{ (t, b) \mid b \in Z(\gamma_t) \quad \text{and} \quad \frac{|\dot{\gamma}_t(b)|^2}{\|\nabla^{\mathbb{C}} \gamma_t(b)\|^2} = \max_{\theta \in Z(\gamma_t)} \frac{|\dot{\gamma}_t(\theta)|^2}{\|\nabla^{\mathbb{C}} \gamma_t(\theta)\|^2} \right\}.$$

Let now  $\pi : S \rightarrow (0, 1)$  be the restriction of the projection on the first factor. By the Semialgebraic Triviality Theorem [BCR98, Theorem 9.3.2] there exist  $0 < \delta < 1$ , a semialgebraic set  $F \subset \mathbb{C}P^n$  and a semialgebraic homeomorphism  $\psi : \pi^{-1}((0, \delta)) \rightarrow (0, \delta) \times F$  such that, denoting by  $p_1 : (0, \delta) \times F \rightarrow (0, \delta)$  the projection on the first factor, we have  $\pi = p_1 \circ \psi$ .

Pick  $f_0 \in F$  and define the curve  $b : [0, \delta) \rightarrow \mathbb{C}P^n$  by:

$$b(t) := \begin{cases} \psi^{-1}(t, f_0) & t \in (0, \delta) \\ \lim_{t \rightarrow 0} \psi^{-1}(t, f_0) & t = 0. \end{cases}$$

The limit  $\lim_{t \rightarrow 0} \psi^{-1}(t, f_0)$  exists since  $\psi$  is semialgebraic with codomain  $\mathbb{C}P^n$  (which is compact). Such curve  $b : [0, \delta] \rightarrow \mathbb{C}P^n$  is continuous and is semialgebraic, therefore there exists  $0 < r \leq \delta$  such that  $b|_{(0,r]}$  is  $\mathcal{C}^1$ . Since  $(t, b(t)) \in S$ , then  $b(t)$  verifies (5.15) for all  $0 < t \leq r$ , this proves the first part of the statement.

For the second part of the statement, we use the fact that  $b(t) \in Z(\gamma_t)$  for all  $0 < t \leq r$ , hence  $\gamma_t(b(t)) \equiv 0$ . Differentiating this last equation (viewed in  $S^{2n+1} \subset \mathbb{C}^{n+1}$ ) with respect to  $t$  gives  $\dot{\gamma}_t(b(t)) + \langle \dot{b}(t), (\nabla^{\mathbb{C}} \gamma_t)(b(t)) \rangle \equiv 0$  which implies

$$|\dot{\gamma}_t(b(t))| \leq \|\dot{b}(t)\| \cdot \|\nabla^{\mathbb{C}} \gamma_t(b(t))\|.$$

The last inequality implies (5.16), and this concludes the proof.  $\square$

**Theorem 5.13.** *Let  $\gamma : [0, 1] \rightarrow P_{n,d}$  be a continuous semialgebraic curve such that  $\gamma(0) \in \Delta_{n,d}$  and  $\gamma(t) \notin \Delta_{n,d}$  for all  $t \in (0, 1]$ . Then there exists  $m \in \mathbb{N}$  such that the curve  $\alpha : [0, 1] \rightarrow P_{n,d}$ , defined by  $\alpha(s) := \gamma(s^m)$ , has finite energy, i.e.  $\mathcal{E}(\alpha) < \infty$ .*

*Proof.* Let  $b : [0, r] \rightarrow \mathbb{C}P^n$  be the curve given by Lemma 5.12. By [BCR98, Proposition 2.4.6], there exists a smooth, real algebraic embedding  $\varphi : \mathbb{C}P^n \rightarrow \mathbb{R}^\ell$ , for some  $\ell \in \mathbb{N}$  sufficiently large. Denote by  $f(t) := \varphi(b(t))$ . By Lemma 5.11 there exists  $0 < r_0 \leq r \leq 1$  and  $m \in \mathbb{N}$  such that the curve  $\tilde{f}(s) := f(s^m)$  is  $\mathcal{C}^1$  on  $[0, r_0]$ . In particular, since  $\varphi$  is a semialgebraic diffeomorphism, the curve

$$\beta(s) := b(s^m) = \varphi^{-1}(\tilde{f}(s))$$

is semialgebraic and of class  $\mathcal{C}^1$  on  $[0, r_0]$ .

Consider now the curve  $\alpha(s) := \gamma(s^m)$ . By Corollary 5.8 we have

$$\begin{aligned} \mathcal{E}(\alpha) &= \int_0^1 \frac{1}{\text{vol}(Z(\alpha(s)))} \int_{Z(\alpha(s))} \frac{|\dot{\alpha}_s(b)|^2}{\|\nabla^{\mathbb{C}} \alpha_s(b)\|^2} \text{vol}_{Z(\alpha(s))}(d[b]) ds \\ &= \left( \int_0^{r_0} + \int_{r_0}^1 \right) \left( \frac{1}{\text{vol}(Z(\alpha(s)))} \int_{Z(\alpha(s))} \frac{|\dot{\alpha}_s(b)|^2}{\|\nabla^{\mathbb{C}} \alpha_s(b)\|^2} \text{vol}_{Z(\alpha(s))}(d[b]) \right) ds. \end{aligned}$$

Since for  $s \geq r_0$  the set  $Z(\alpha(s)) = Z(\gamma(s^m))$  is defined by a regular equation, the integrand is bounded on  $[r_0, 1]$ . Therefore it remains to estimate the integral in  $[0, r_0]$ . Using (5.16), we have:

$$\begin{aligned} & \int_0^{r_0} \frac{1}{\text{vol}(Z(\alpha(s)))} \int_{Z(\alpha(s))} \frac{|ms^{m-1}\dot{\gamma}_{s^m}(b)|^2}{\|\nabla^{\mathbb{C}} \gamma_{s^m}(b)\|^2} \text{vol}_{Z(\alpha(s))}(d[b]) ds \\ & \leq \int_0^{r_0} \frac{1}{\text{vol}(Z(\alpha(s)))} \int_{Z(\alpha(s))} |ms^{m-1}\dot{b}(s^m)|^2 \text{vol}_{Z(\alpha(s))}(d[b]) ds \\ & = \int_0^{r_0} \frac{1}{\text{vol}(Z(\alpha(s)))} \int_{Z(\alpha(s))} |\dot{\beta}(s)|^2 \text{vol}_{Z(\alpha(s))}(d[b]) ds \\ & = \int_0^{r_0} |\dot{\beta}(s)|^2 ds < \infty \quad (\text{since } \beta \text{ is } \mathcal{C}^1 \text{ on } [0, r_0]). \end{aligned}$$

This concludes the proof.  $\square$

**COROLLARY 5.14.** *For any two  $p_0, p_1 \in P_{n,d}$  there exists a curve  $\gamma : I \rightarrow P_{n,d}$  such that  $\gamma(0) = p_0$ ,  $\gamma(1) = p_1$  and  $\mathcal{E}(\gamma) < \infty$ .*

*Proof.* If  $p_0, p_1 \notin \Delta_{n,d}$ , then any smooth curve  $\gamma : I \rightarrow P_{n,d} \setminus \Delta_{n,d}$  connecting  $p_0$  with  $p_1$  (which exists because the real codimension of  $\Delta_{n,d}$  is two) has finite energy by Corollary 5.8.

Given  $p_0 \in \Delta_{n,d}$  and  $p_1 \in P_{n,d} \setminus \Delta_{n,d}$  there exists a curve  $\gamma$  joining them with finite energy, by Theorem 5.13. The result for general  $p_0, p_1$  (possibly both on  $\Delta_{n,d}$ ) follows by concatenating two such curves.  $\square$

We now prove that  $W_2^{\text{in}}$  is a metric on  $P_{n,d}$ .

**Theorem 5.15.** *The function  $W_2^{\text{in}}$  defines a distance on  $P_{n,d}$  and the metric space  $(P_{n,d}, W_2^{\text{in}})$  is complete and geodesic, i.e. for any couple  $p_1, p_2$ , there exists an admissible curve  $\gamma$  from  $p_1$  to  $p_2$  such that  $W_2^{\text{in}}(\gamma(s), \gamma(t)) = |t - s|W_2^{\text{in}}(p_1, p_2)$ .*

*Proof.* By construction  $W_2^{\text{in}}(p_0, p_1) = W_2^{\text{in}}(p_1, p_0)$ . Moreover, given  $p_0, p_1, p_2 \in P_{n,d}$ , by Corollary 5.14 for every  $\epsilon > 0$  there exist admissible curves  $\gamma_1, \gamma_2 : I \rightarrow P_{n,d}$  such that:  $\gamma_1(0) = p_0$ ,  $\gamma_1(1) = p_1$ ,  $\gamma_2(0) = p_1$ ,  $\gamma_2(1) = p_2$  and

$$\mathcal{E}(\gamma_1) \leq W_2^{\text{in}}(p_0, p_1)^2 + \epsilon \quad \text{and} \quad \mathcal{E}(\gamma_2) \leq W_2^{\text{in}}(p_1, p_2)^2 + \epsilon.$$

We use now [AGS05, Lemma 1.1.4] to get reparametrizations  $\eta_1 : [0, L_1] \rightarrow P_{n,d}$  and  $\eta_2 : [0, L_2] \rightarrow P_{n,d}$  of these curves, with  $L_1^2 \leq W_2^{\text{in}}(p_0, p_1)^2 + \epsilon$  and  $L_2^2 \leq W_2^{\text{in}}(p_1, p_2)^2 + \epsilon$ , such that, denoting by  $\zeta_i := \mu \circ \eta_i$ , we have

$$|\dot{\zeta}_i| \equiv 1, \quad i = 1, 2.$$

Define now  $\eta : [0, L_1 + L_2] \rightarrow P_{n,d}$  by concatenating  $\eta_1$  and  $\eta_2$  and  $\alpha : [0, 1] \rightarrow P_{n,d}$  by:

$$\alpha(t) = \eta((L_1 + L_2)t).$$

Then, denoting by  $\beta := \mu \circ \alpha$ , we have

$$W_2^{\text{in}}(p_0, p_2)^2 \leq \mathcal{E}(\alpha) = \int_0^1 |\dot{\beta}|^2 = (L_1 + L_2)^2,$$

which, letting  $\epsilon$  go to zero, implies the triangle inequality. Finally, since if  $\gamma$  is admissible,  $\mu \circ \gamma \in AC^2(I, \mathcal{P}_2(\mathbb{C}P^n))$ , it follows that  $W_2^{\text{in}}(p_0, p_1) = 0$  implies  $p_0 = p_1$ . Together with Corollary 5.14, this proves that  $W_2^{\text{in}}$  defines a distance on  $P_{n,d}$ .

In order to prove that  $(P_{n,d}, W_2^{\text{in}})$  is geodesic, we observe that the map  $\mu : P_{n,d} \rightarrow \mathcal{P}_2(\mathbb{C}P^n)$  is injective and continuous by Theorem 3.8, where  $P_{n,d}$  is taken with the manifold topology. In particular  $E := \mu(P_{n,d})$  is a compact subset of  $\mathcal{P}_2(\mathbb{C}P^n)$ , and  $E$  is homeomorphic to  $P_{n,d}$  itself, through  $\mu$ . Admissible curves on  $P_{n,d}$  are precisely curves  $\gamma : I \rightarrow P_{n,d}$  such that there exists an admissible curve  $\beta : I \rightarrow \mathcal{P}_2(\mathbb{C}P^n)$ , entirely contained in  $E$  and such that  $\gamma = \mu^{-1} \circ \beta$ . The fact that  $(P_{n,d}, W_2^{\text{in}})$  is geodesic follows now from Corollary 5.14, [AT04, 4.3.2] and [GP20, Proposition 1.2.10].

The completeness is a direct consequence of  $W_2^{\text{in}}$  being lower semicontinuous with respect to the Fubini–Study metric. The proof of the lower semicontinuity of  $W_2^{\text{in}}$  is in the subsequent Lemma 5.16. Consider therefore a Cauchy sequence  $\{p_k\}_{k \in \mathbb{N}} \subset (P_{n,d}, W_2^{\text{in}})$ . By compactness of  $P_{n,d}$  in the Fubini–Study metric, it follows that  $p_k$  has a converging subsequence (still denoted by  $p_k$ ) in the Fubini–Study metric to a certain  $p_\infty \in P_{n,d}$ . Hence for each  $j \in \mathbb{N}$  the lower semicontinuity of  $W_2^{\text{in}}$  implies that

$$\liminf_{k \rightarrow \infty} W_2^{\text{in}}(p_k, p_j) \geq W_2^{\text{in}}(p_\infty, p_j).$$

Letting also  $j \rightarrow \infty$  it follows that  $W_2^{\text{in}}(p_\infty, p_j) \rightarrow 0$ . Since for Cauchy sequences the convergence of a subsequence is enough to deduce convergence of the whole sequence, the completeness is proved.  $\square$

**LEMMA 5.16.** *For every fixed  $p \in P_{n,d}$ , the function  $W_2^{\text{in}}(p, \cdot) : P_{n,d} \rightarrow [0, \infty)$  is lower semicontinuous with respect to the Fubini–Study metric.*

*Proof.* Fix  $p \in P_{n,d}$  and consider a sequence  $\{p_k\}_{k \in \mathbb{N}} \subset P_{n,d}$  with  $p_k \rightarrow p_\infty \in P_{n,d}$  in the Fubini–Study metric. By Theorem 5.15 there exists a sequence of admissible curves  $\gamma_k$  such that  $\mu(\gamma_k) \in AC^2([0, 1]; \mathcal{P}_2(\mathbb{C}P^n))$  and  $\mathcal{E}(\gamma_k) = W_2^{\text{in}}(p_k, p)$ . From Corollary 5.14 we know that  $W_2^{\text{in}}(p_\infty, p) < \infty$ ; hence if  $\liminf_{k \rightarrow \infty} W_2^{\text{in}}(p_k, p) = \infty$  the claim is trivially verified. We assume therefore that  $\liminf_{k \rightarrow \infty} W_2^{\text{in}}(p_k, p) < \infty$  and, up to relabeling, that the liminf is indeed a limit. In particular, we deduce that  $\mathcal{E}(\gamma_k) \leq C$ , for some positive constant  $C > 0$ . Therefore applying Hölder inequality for any  $s < t \in (0, 1)$  and  $k \in \mathbb{N}$  we obtain:

$$W_2(\mu(\gamma_k)_t, \mu(\gamma_k)_s) \leq \int_s^t |\mu(\dot{\gamma}_k)|(\tau) d\tau \leq C\sqrt{t-s}.$$

The previous inequality implies that the family of maps  $[0, 1] \ni t \mapsto \mu(\gamma_k)_t \in \mathcal{P}_2(\mathbb{C}P^n)$  is equicontinuous. Since  $\mathcal{P}_2(\mathbb{C}P^n)$  is compact we can invoke Ascoli–Arzelá Theorem to deduce the existence of a subsequence  $\mu(\gamma_{k_j})$  and of a limit curve  $\eta \in \mathcal{C}^0([0, 1]; \mathcal{P}_2(\mathbb{C}P^n))$  such that  $\mu(\gamma_{k_j}) \rightarrow \eta$  in the  $\mathcal{C}^0$  distance. We can invoke then the lower semicontinuity of the energy (see [GP20, Proposition 1.2.7]) to deduce

$$\mathcal{E}(\eta) \leq \liminf_{k \rightarrow \infty} \mathcal{E}(\mu(\gamma_{k_j})).$$

Finally, since  $\mu(P_{n,d})$  is a closed subset of  $\mathcal{P}_2(\mathbb{C}P^n)$ , we deduce the existence of a curve  $\gamma_\infty$  such that  $\eta = \mu(\gamma_\infty)$ . Then the above inequality can be rewritten as

$$\mathcal{E}(\gamma_\infty) \leq \liminf_{k \rightarrow \infty} \mathcal{E}(\gamma_{k_j}) = \lim_{j \rightarrow \infty} W_2^{\text{in}}(p_{k_j}, p) = \liminf_{k \rightarrow \infty} W_2^{\text{in}}(p_k, p),$$

implying in particular that  $\gamma_\infty$  is an admissible curve. Moreover is trivially verified (recall that  $\mu$  is an homeomorphism) that  $\gamma_\infty$  is joining the fixed  $p \in P_{n,d}$  with  $p_\infty \in P_{n,d}$ . By taking the infimum on the first term, the claim follows.  $\square$

**5.4 Comparing the inner Wasserstein distance with the Fubini–Study one.** As one can easily verify, the continuity of

$$\text{id} : (P_{n,d}, \mathbf{d}_{\text{FS}}) \rightarrow (P_{n,d}, W_2^{\text{in}})$$

is equivalent to the compactness of  $(P_{n,d}, W_2^{\text{in}})$ ; notice that  $P_{n,d}$  is compact for the ambient  $W_2$ –topology, by Proposition 3.11. Compactness for the inner distance topology of a compact space in the ambient topology is in general not true, see [BBI01, Fig. 2.2]. However, building on the results from the previous section, here we will prove that the above map is continuous (in fact Lipschitz) “away” from the discriminant, where the structure is Riemannian. More precisely, we will prove the following result. For every  $\epsilon > 0$  denote by

$$P_{n,d}(\epsilon) := \{p \in P_{n,d} \mid \mathbf{d}_{\text{FS}}(p, \Delta_{n,d}) \geq \epsilon\}$$

**Theorem 5.17.** *For every  $\epsilon > 0$  the identity map*

$$\text{id} : (P_{n,d}(\epsilon), \mathbf{d}_{\text{FS}}) \rightarrow (P_{n,d}(\epsilon), W_2^{\text{in}})$$

*is Lipschitz. In particular  $P_{n,d}(\epsilon)$  is compact in the  $W_2^{\text{in}}$ –topology.*

REMARK 5.18. The above statement is phrased using the Fubini–Study distance on  $P_{n,d} \simeq \mathbb{C}P^N$ . Recall that the Bombieri–Weyl structure (see Sect. A.4) is a special case of the Fubini–Study one, i.e. the case when we identify  $\mathbb{C}^{\binom{n+d}{d}}$  with  $\mathbb{C}[z_0, \dots, z_n]_{(d)}$  through the map that sends the standard basis to the basis (A.8); therefore the reader can also think of this as a statement involving the Bombieri–Weyl distance.

In order to prove Theorem 5.17 we need some preliminary lemmas.

LEMMA 5.19. *There exist two constants  $c_1, c_2 > 0$  such that every  $p \notin \Delta_{n,d}$  has a neighborhood  $U_p$  with  $\text{clos}(U_p) \subset P_{n,d} \setminus \Delta_{n,d}$ , which is geodesically convex in the Fubini–Study metric, and such that for every  $q_1, q_2 \in U_p$*

$$W_2^{\text{in}}(q_1, q_2) \leq \left( c_1 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_2} \right) \mathbf{d}_{\text{FS}}(q_1, q_2).$$

*Proof.* Let  $U_p$  be a geodesically convex neighborhood of  $p$  in the Fubini–Study metric such that  $\text{clos}(U_p) \cap \Delta_{n,d} = \emptyset$  and such that for every  $q \in U_p$  we have

$$\mathbf{d}_{\text{FS}}(q, \Delta_{n,d}) \geq \frac{1}{2} \mathbf{d}_{\text{FS}}(p, \Delta_{n,d}). \quad (5.17)$$

Consider also the semialgebraic functions  $\alpha, \beta : P_{n,d} \rightarrow \mathbb{R}$  defined by

$$\alpha(p) := \mathbf{d}_{\text{FS}}(p, \Delta_{n,d}) \quad \text{and} \quad \beta(p) := \min_{b \in Z(p)} \frac{\|\nabla^{\mathbb{C}} p(b)\|^2}{\|p\|_{\text{FS}}^2}.$$

(Notice that  $\beta$  is well defined, since scaling  $p$  by a nonzero multiple does not change its value.) Notice that both  $\alpha$  and  $\beta$  vanish precisely on  $\Delta_{n,d}$ , therefore, since  $P_{n,d}$  is compact, Lojasiewicz’s inequality tells that there exist  $c_3, c_4 > 0$  such that  $\alpha^{c_3} \leq c_4 \beta$ , i.e.

$$\mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{c_3} \leq c_4 \min_{b \in Z(p)} \frac{\|\nabla^{\mathbb{C}} p(b)\|^2}{\|p\|_{\text{FS}}^2}. \quad (5.18)$$

Let now  $q_1, q_2 \in U_p$  and consider the Fubini–Study geodesic  $\gamma : I \rightarrow P_{n,d}$  joining them, parametrized with  $\|\dot{\gamma}(t)\|_{\text{FS}} \equiv \mathbf{d}_{\text{FS}}(q_1, q_2)$ . Since  $U_p$  is geodesically convex, we have  $\gamma_t \in U_p$  for every  $t \in I$ . Let us now compute the length of the curve  $\gamma$  in  $W_2^{\text{in}}$ , using Corollary 5.8:

$$\begin{aligned} L_{W_2^{\text{in}}}(\gamma) &= \int_0^1 (\mathbf{h})_{\gamma(t)}(\dot{\gamma}_t, \dot{\gamma}_t)^{\frac{1}{2}} dt \\ &= \frac{1}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} \int_0^1 \left( \int_{Z(\gamma_t)} \frac{|\dot{\gamma}_t(b)|^2}{\|\nabla^{\mathbb{C}} \gamma_t(b)\|^2} \text{vol}_{Z(\gamma_t)}(db) \right)^{\frac{1}{2}} dt \\ &\leq \frac{1}{d \cdot \text{vol}(\mathbb{C}P^{n-1})} \int_0^1 \left( \int_{Z(\gamma_t)} \frac{c_4 |\dot{\gamma}_t(b)|^2}{\|\gamma_t\|_{\text{FS}}^2 \cdot \mathbf{d}_{\text{FS}}(\gamma_t, \Delta_{n,d})^{c_3}} \text{vol}_{Z(\gamma_t)}(db) \right)^{\frac{1}{2}} dt \\ &\leq \left( c_5 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_3/2} \right) \int_0^1 \left( \int_{Z(\gamma_t)} \frac{|\dot{\gamma}_t(b)|^2}{\|\gamma_t\|_{\text{FS}}^2} \text{vol}_{Z(\gamma_t)}(db) \right)^{\frac{1}{2}} dt = (*), \end{aligned}$$

where we have used (5.17) and (5.18).

Observe now that, since  $\|b\|^2 = 1$ , we have  $|\dot{\gamma}_t(b)|^2 \leq c_6 \|\dot{\gamma}_t\|_{\text{FS}}^2$  and therefore:

$$\begin{aligned}
(*) &\leq \left( c_5 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_3/2} \right) \int_0^1 \left( \int_{Z(\gamma_t)} \frac{c_6 \|\dot{\gamma}_t\|_{\text{FS}}^2}{\|\gamma_t\|_{\text{FS}}^2} \text{vol}_{Z(p_t)}(db) \right)^{\frac{1}{2}} dt \\
&= \left( c_7 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_3/2} \right) \int_0^1 \frac{\|\dot{\gamma}_t\|_{\text{FS}}}{\|\gamma_t\|_{\text{FS}}} \left( \int_{Z(\gamma_t)} \text{vol}_{Z(\gamma_t)}(db) \right)^{\frac{1}{2}} dt \\
&\leq (c_1 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_2}) \int_0^1 \frac{\|\dot{\gamma}_t\|_{\text{FS}}}{\|\gamma_t\|_{\text{FS}}} dt \\
&= (c_1 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_2}) L_{\text{FS}}(\gamma) = (c_1 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_2}) \cdot \mathbf{d}_{\text{FS}}(q_1, q_2).
\end{aligned}$$

Since  $W_2^{\text{in}}(q_1, q_2) \leq L_{W_2}(\gamma)$ , this concludes the proof.  $\square$

LEMMA 5.20. *For every  $\epsilon > 0$  there exists  $c(\epsilon) > 0$  such that*

$$W_2^{\text{in}}|_{P_{n,d}(\epsilon) \times P_{n,d}(\epsilon)} \leq c(\epsilon).$$

*Proof.* For every point  $q \in P_{n,d}(\epsilon)$  let  $U_q$  be the Fubini–Study convex neighborhood centered at  $q$  given by Lemma 5.19. Possibly shrinking this neighborhood, we can assume that it is contained in  $P_{n,d}(\epsilon/2)$ . Since  $P_{n,d}(\epsilon)$  is compact in  $\mathbf{d}_{\text{FS}}$ , it can be covered by finitely many such neighborhoods  $U_{q_1}, \dots, U_{q_r}$ .

Given  $p_0, p_1 \in P_{n,d}(\epsilon)$ , let  $\gamma : I \rightarrow P_{n,d}$  be a piecewise smooth curve joining  $p_0$  and  $p_1$ , entirely contained in  $\bigcup U_{q_j}$  and such that for every  $j = 1, \dots, r$  the set  $\gamma^{-1}(U_{q_j})$  is an interval. Then the length of  $\gamma$  is bounded by

$$L_{\text{FS}}(\gamma) \leq r \max_j \text{diam}_{\text{FS}}(U_{q_j}) \leq c_1(\epsilon).$$

Arguing now as in the proof of Lemma 5.19, we see that

$$W_2^{\text{in}}(p_0, p_1) \leq L_{W_2^{\text{in}}}(\gamma) \leq \max_j \{c_1 \mathbf{d}_{\text{FS}}(q_j, \Delta_{n,d})^{-c_2}\} L_{\text{FS}}(\gamma) \leq c(\epsilon). \quad \square$$

*Proof of Theorem 5.17.* Assume by contradiction that  $\text{id}$  is not Lipschitz. Then, for every  $n \in \mathbb{N}$  there exist  $a_m, b_m \in P_{n,d}(\epsilon)$  such that

$$\frac{W_2^{\text{in}}(a_m, b_m)}{\mathbf{d}_{\text{FS}}(a_m, b_m)} \geq m. \quad (5.19)$$

Since  $W_2^{\text{in}}$  is bounded on  $P_{n,d}(\epsilon) \times P_{n,d}(\epsilon)$  (by Lemma 5.20), it follows that we must have  $\mathbf{d}_{\text{FS}}(a_m, b_m) \rightarrow 0$ . In particular, since  $P_{n,d}(\epsilon)$  is compact for the Fubini–Study topology, up to subsequences, we must have  $a_{m_k}, b_{m_k} \rightarrow a$  as  $k \rightarrow \infty$  in the Fubini–Study topology, with  $a \in P_{n,d}(\epsilon)$ .

Let now  $U_a$  be the neighborhood of  $a$  given by Lemma 5.19. The sequences  $\{a_{m_k}\}$  and  $\{b_{m_k}\}$  lie definitely in  $U_p$ , therefore by Lemma 5.19:

$$W_2^{\text{in}}(a_{m_k}, b_{m_k}) \leq \left( c_1 \mathbf{d}_{\text{FS}}(p, \Delta_{n,d})^{-c_2} \right) \mathbf{d}_{\text{FS}}(a_{m_k}, b_{m_k}). \quad (5.20)$$

However, as  $k \rightarrow \infty$ , (5.20) contradicts (5.19).

The second part of the statement follows from the fact that, being Lipschitz,  $\text{id} : (P_{n,d}(\epsilon), \mathbf{d}_{\text{FS}}) \rightarrow (P_{n,d}(\epsilon), W_2^{\text{in}})$  is in particular continuous (and  $(P_{n,d}(\epsilon), \mathbf{d}_{\text{FS}})$  is compact).  $\square$

## 6 Absolute continuity of curves of roots

In this section we deduce a first application of the new definition of Wasserstein distance introduced in Sect. 5, generalizing a result by A. Parusinski and A. Rainer [PR18] on the regularity of roots of univariate polynomials, to the case of hypersurfaces. As we will see, for the generalization the existence of rational curves on the zero set of our curve of polynomials will play a role.

Let us start by recalling [PR18, Theorem 1].

**Theorem 6.1** (Parusinski–Rainer). *Let  $(\alpha, \beta) \subset \mathbb{R}$  be a bounded open interval and  $P_a$  a monic polynomial*

$$P_a(t)(z) = P_{a(t)}(z) = z^d + \sum_{j=1}^d a_j(t)z^{d-j},$$

with coefficients  $a_j \in \mathcal{C}^{d-1,1}([\alpha, \beta])$ , with  $j = 1, \dots, d$ . Let  $\lambda \in \mathcal{C}((\alpha, \beta))$  be a continuous root of  $P_a$  on  $(\alpha, \beta)$ . Then  $\lambda$  is absolutely continuous on  $(\alpha, \beta)$  and belongs to the Sobolev space  $W^{1,p}((\alpha, \beta))$  for every  $1 \leq p < d/(d-1)$ . The derivative  $\lambda'$  satisfies

$$\|\lambda'\|_{L^p((\alpha, \beta))} \leq C(d, p) \max\{1, (\beta - \alpha)^{1/p}\} \max_{1 \leq j \leq d} \|a_j\|_{\mathcal{C}^{d-1,1}([\alpha, \beta])}^{1/j}.$$

[PR18, Theorem 1] is optimal: in general the derivatives of the roots of a smooth curve of monic polynomials of degree  $d$  are not locally  $d/(d-1)$ -integrable (take for instance  $z^d + t$ ), and the roots may have locally unbounded variation if the coefficients are only of class  $\mathcal{C}^{d-1,\alpha}$  for  $\alpha < 1$ .

We now reformulate [PR18, Theorem 1] in our formalism. The natural generalization to higher degree will be then discussed.

Given a monic polynomial  $P_{a(t)}$  as in [PR18, Theorem 1] (for simplicity defined over  $[0, 1]$ ), we consider the associated homogeneous polynomial

$$p_t(z_0, z_1) = z_1^d + \sum_{j=1}^d a_j(t)z_1^{d-j}z_0^j \in \mathbb{C}[z_0, z_1]_{(d)}.$$

The corresponding measure  $\mu_t := \mu(p_t)$  will then be a probability measure over  $\mathbb{CP}^1$ :

$$\mu_t = \frac{1}{d} \sum_{z \in Z(P_{a(t)})} m_z(P_{a(t)}) \delta_{[1, z]}.$$

We now compute the metric speed of the curve  $[0, 1] \ni t \mapsto \mu_t \in \mathcal{P}_p(\mathbb{CP}^1)$ , with  $p \geq 1$ . Suppose additionally the existence of continuous, pairwise distinct roots  $\lambda_1, \dots, \lambda_r$

of  $P_{a(t)}$  with  $r \leq d$ . For  $h$  sufficiently small we then have

$$W_p(\mu_t, \mu_{t+h})^p = \frac{1}{d} \sum_{j=1}^r m_{[1, \lambda_j(t)]}(p_t) \mathbf{d}_{FS}([1, \lambda_j(t)], [1, \lambda_j(t+h)])^p.$$

Hence taking the limit as  $h \rightarrow 0$  we obtain

$$\frac{C_1}{d} \sum_{j=1}^r m_{[1, \lambda_j(t)]}(p_t) |\lambda_j(t)'|^p \leq |\dot{\mu}_t|_p^p \leq \frac{C_2}{d} \sum_{j=1}^r m_{[1, \lambda_j(t)]}(p_t) |\lambda_j(t)'|^p,$$

where  $C_1, C_2$  are two positive constants given by the equivalence of the metric speeds calculated in  $\mathbb{CP}^1$  or in  $\mathbb{C}$ . We see therefore that the absolute continuity of the roots and the integrability of their derivative  $|\lambda_j(t)'|^p$  is actually equivalent to the  $p$ -absolute continuity of the curve  $t \mapsto \mu_t$  inside  $\mathcal{P}_p(\mathbb{CP}^1)$ . In particular to  $(\mu_t)_{[0,1]} \in AC^p([0, 1]; \mathcal{P}_p(\mathbb{CP}^1))$ .

In order to formulate and prove the first generalization of [PR18, Theorem 1], we need one extra definition. Let  $\tau^* \rightarrow \mathbb{G}(1, n)$  be the dual of the tautological bundle and

$$E := \text{sym}^{(d)}(\tau^*) \rightarrow \mathbb{G}(1, n)$$

be its  $d$ -th symmetric power. Recall that  $E$  is a complex vector bundle on  $\mathbb{G}(1, n)$ ; given  $\ell \in \mathbb{G}(1, n)$ , the fiber  $E|_\ell$  is naturally isomorphic to the set of homogeneous polynomials of degree  $d$  on  $\ell$ ; in particular the complex rank of  $E$  is  $d + 1$ . Every homogeneous polynomial  $p \in H_{n,d}$  gives rise to a section  $\sigma_p$  of  $E$ , defined by

$$\sigma_p(\ell) = p|_\ell.$$

A line  $\ell$  is contained in the zero set of  $Z(p)$  if and only if  $\sigma_p(\ell) = 0$ .

**DEFINITION 6.2** (Fano discriminant). We denote by  $F_{n,d} \subset P_{n,d}$  and call it the *Fano discriminant* the set of polynomials  $[p] \in P_{n,d}$  such that the section  $\sigma_p$  is not transversal to the zero section of  $E$ .

When  $n = 1$ , the Fano discriminant is empty: in this case the only line is  $\mathbb{CP}^1$  itself, and no nonzero polynomial  $p \in \mathbb{C}[z_0, z_1]_{(d)}$  can vanish identically on this line. If  $p \in P_{n,d} \setminus F_{n,d}$ , then the set of lines contained in  $Z(p)$  is a smooth, complex subset of the Grassmannian, of complex codimension  $d + 1$  (the rank of  $E$ ). In particular, if  $d = 2n - 3$  and  $p \notin F_{n,d}$ , then there are finitely many lines on  $Z(p)$ . If  $d > 2n - 3$  and  $p \notin F_{n,d}$ , then there is no line on  $Z(p)$  (the “generic” section misses the zero section). This explains the assumptions in next theorem, which gives a regularity for a curve of measures of the form  $\mu_t = \mu(p_t)$ . Of course, the relevant fact is that  $F_{n,d}$  is contained in a proper algebraic set [EH16, Theorem 6.34].

**Theorem 6.3.** *Let  $d > 2n - 3$ . For every curve  $p_t \in \mathcal{C}^{d-1,1}(I, \mathbb{C}[z_0, \dots, z_n]_{(d)})$ , such that  $p_t \notin F_{n,d}$  for every  $t \in I$ , and for every  $1 \leq q < d/(d - 1)$  the curve  $\mu_t := \mu(p_t) \in AC^q(I, \mathcal{P}_q(\mathbb{CP}^n))$ .*

*Proof.* Fix any  $q$  with  $1 \leq q < d/(d-1)$ . We will provide an upper bound on the increment of  $\mu_t$  via an explicit coupling between  $\mu_{t+h}$  and  $\mu_t$  constructed by means of integral geometry. For any  $t \in [0, 1]$ , the claim of Theorem 3.8 can be rewritten as

$$\mu_t = \int_{\mathbb{G}(1,n)} \mu(\gamma(t)|_\ell) \text{vol}_{\mathbb{G}(1,n)}(d\ell),$$

where, in particular, for any continuous function  $f$  of  $\mathbb{C}\mathbb{P}^n$ , the map

$$\mathbb{G}(1,n) \ni \ell \longmapsto \int f \mu(\gamma(t)|_\ell) = \frac{1}{d} \sum_{z \in Z(\gamma(t)) \cap \ell} m_z(\gamma(t)|_\ell) f(z),$$

is measurable and bounded. For ease of notation we will write  $\mu_{t,\ell}$  in place of  $\mu(\gamma(t)|_\ell)$ .

We consider then the optimal transport problem between  $\mu_{t,\ell}$  and  $\mu_{t+h,\ell}$  with cost function given by  $d_{std}^q$  for which, by compactness of  $\Pi(\mu_{t,\ell}, \mu_{t+h,\ell})$  and linearity of the functional, there exists a minimizer that we denote by  $\xi_{t,h,\ell} \in \Pi(\mu_{t,\ell}, \mu_{t+h,\ell})$ . By a classical selection argument, the map  $\mathbb{G}(1,n) \ni \ell \mapsto \xi_{t,h,\ell} \in \mathcal{P}(\mathbb{C}\mathbb{P}^n \times \mathbb{C}\mathbb{P}^n)$  can be assumed to be measurable; we can then define

$$\xi_{t,h} := \int_{\mathbb{G}(1,n)} \xi_{t,h,\ell} \text{vol}_{\mathbb{G}(1,n)}(d\ell) \in \Pi(\mu_t, \mu_{t+h}),$$

and compute:

$$\begin{aligned} W_q(\mu_{t+h}, \mu_t)^q &\leq \int_{\mathbb{C}\mathbb{P}^n \times \mathbb{C}\mathbb{P}^n} d_{std}^q(z, w) \xi_{t,h}(dzdw) \\ &= \int_{\mathbb{G}(1,n)} \int_{\mathbb{C}\mathbb{P}^n \times \mathbb{C}\mathbb{P}^n} d_{std}^q(z, w) \xi_{t,h,\ell} \text{vol}_{\mathbb{G}(1,n)}(d\ell) \\ &= \int_{\mathbb{G}(1,n)} W_q(\mu_{t+h,\ell}, \mu_{t,\ell})^q \text{vol}_{\mathbb{G}(1,n)}(d\ell). \end{aligned}$$

We will now estimate  $W_q(\mu_{t+h,\ell}, \mu_{t,\ell})^q$ .

To this end, fix a line  $\ell_0$  and denote by  $L_0 \simeq \mathbb{C}^2 \subset \mathbb{C}^{n+1}$  a complex vector space such that  $P(L_0) = \ell_0$ . On a small enough neighborhood  $U_1(\ell_0)$  of  $L_0$  in  $G(2, n+1) = \mathbb{G}(1, n)$ , by Lemma 6.4 (proved below), we can find smooth maps  $e_0, e_1 : U_1(\ell_0) \rightarrow \mathbb{C}^{n+1}$  such that: (i) for every  $L \in U_1(\ell_0)$ , where  $P(L) = \ell$ , the set  $\{e_0(\ell), e_1(\ell)\}$  is a Hermitian orthonormal basis of  $L$  and (ii) for every  $t \in I$  the point  $e_1(\ell_0)$  is not a zero of  $p_t$ .

For every  $\ell \in U_1(\ell_0)$  define the polynomial:

$$a_{t,\ell}(z_0, z_1) := p_t(z_0 e_0(\ell) + z_1 e_1(\ell)) = \sum_{k=0}^d a_k(t, \ell) z_0^k z_1^{d-k} \in \mathbb{C}[z_0, z_1]_{(d)}.$$

Notice that, the functions  $t \mapsto a_k(t, \ell)$  are  $\mathcal{C}^{d-1,1}$ . Moreover, since  $a_0(t, \ell_0) = p_t(e_1(\ell_0)) \neq 0$  for all  $t \in I$ , then  $a_0(t, \ell_0) \neq 0$  for all  $t \in I$ . In particular (by continuity of  $a_0$ ) there is a neighborhood  $U_2(\ell_0)$  with  $\text{clos}(U_2(\ell_0)) \subseteq U_1(\ell_0)$  such that

$a_0(t, \ell) \neq 0$  for every  $\ell \in \text{clos}(U_2(\ell_0))$  and for every  $t \in I$ . For  $(t, \ell) \in I \times U_2(\ell)$  define:

$$b_k(t, \ell) := \frac{a_k(t, \ell)}{a_0(t, \ell)}.$$

For every  $\ell \in U_2(\ell_0)$  and  $t \in I$  we consider now the polynomial  $b_{t, \ell} \in \mathbb{C}[z_0, z_1]_{(d)}$  defined by

$$b_{t, \ell}(z_0, z_1) := \sum_{k=0}^d b_k(t, \ell) z_0^k z_1^{d-k} = z_1^d + \sum_{k=1}^d b_k(t, \ell) z_0^k z_1^{d-k}.$$

Notice that  $b_{t, \ell}$  has the same zeroes of  $a_{t, \ell}$  and that the functions  $t \mapsto b_k(t, \ell)$  are again  $\mathcal{C}^{d-1, 1}$  (moreover,  $b_0 \equiv 1$ ).

For every  $\ell \in U_1(\ell_0)$  consider now the map  $\varphi_\ell : \mathbb{CP}^1 \rightarrow \ell$  given by

$$\varphi_\ell([z_0, z_1]) := [z_0 e_0(\ell) + z_1 e_1(\ell)].$$

Such map is an isometry, since  $\{e_0(\ell), e_1(\ell)\}$  is a Hermitian orthonormal basis. In particular, if  $\lambda_1(t, \ell), \dots, \lambda_d(t, \ell)$  are continuous selections of the zeroes of  $b_{t, \ell}$ , we have:

$$W_q(\mu_{t+h, \ell}, \mu_{t, \ell})^q \leq \frac{1}{d} \sum_{j=1}^d \mathbf{d}_{std}^q(\lambda_j(t+h, \ell), \lambda_j(t, \ell)).$$

Denote now by  $\sigma : \mathbb{CP}^1 \setminus \{[0, 1]\} \rightarrow \mathbb{C}$  the stereographic projection,  $\sigma([z_0, z_1]) = \frac{z_1}{z_0}$ , and by  $\tilde{\lambda}_k(t, \ell) := \sigma(\lambda_k(t, \ell))$ . Then  $\tilde{\lambda}_1(t, \ell), \dots, \tilde{\lambda}_d(t, \ell)$  are continuous selections of the zeroes of the polynomial  $\tilde{b}_{t, \ell}$  given by

$$\tilde{b}_{t, \ell}(z) := b_{t, \ell}(1, z) = z^d + \sum_{k=1}^d b_k(t, \ell) z_1^{d-k}.$$

Since there exists  $C_1 > 0$  such that  $\text{dist}_{\mathbb{CP}^1}(x, y) \leq C_1 \cdot \text{dist}_{\mathbb{C}}(\sigma(x), \sigma(y))$  for every  $x, y \in \mathbb{CP}^1 \setminus \{[0, 1]\}$ , then for some  $C_2 > 0$

$$\mathbf{d}_{std}(\lambda_j(t+h, \ell), \lambda_j(t, \ell)) \leq C_2 \cdot |\tilde{\lambda}_j(t+h, \ell) - \tilde{\lambda}_j(t, \ell)|_{(\mathbb{C}, g_{std})}.$$

We use now [PR18, Theorem 1], which guarantees that each  $\tilde{\lambda}_j(\cdot, \ell) \in AC^q(I, \mathbb{C})$  and that

$$\|\partial_t \tilde{\lambda}_j(\cdot, \ell)\|_{L^q} = \left( \int_I |\partial_t \tilde{\lambda}_j(t, \ell)|_{(\mathbb{C}, g_{std})}^q dt \right)^{1/q} \leq C(d, q) \max_{1 \leq j \leq d} \|b_j(\cdot, \ell)\|_{C^{d-1, 1}}^{1/j}.$$

Since  $a_0(t, \ell) \neq 0$  for all  $t \in I$  and  $\ell \in U_2(\ell_0)$ , we deduce that

$$\max_{1 \leq j \leq d} \|b_j(\cdot, \ell)\|_{C^{d-1, 1}}^{1/j} \leq C(\gamma, \ell_0, q) < \infty,$$

for all  $t \in I$  and  $\ell \in U_2(\ell_0)$ . Hence for each  $\ell \in U_2(\ell_0)$ :

$$\begin{aligned}
\int_{[0,1-h]} \left( \frac{W_q(\mu_{t+h,\ell}, \mu_{t,\ell})}{h} \right)^q dt &\leq \frac{1}{d} \sum_{j=1}^d \int_{[0,1-h]} \left( \frac{\mathbf{d}_{std}(\lambda_j(t+h,\ell), \lambda_j(t,\ell))}{h} \right)^q dt \\
&\leq \frac{1}{d} \sum_{j=1}^d \int_{[0,1-h]} \left( \frac{|\tilde{\lambda}_j(t+h,\ell) - \tilde{\lambda}_j(t,\ell)|}{h} \right)^q dt \\
&\leq \frac{1}{d} \sum_{j=1}^d \int_{[0,1-h]} \left( \frac{1}{h} \int_{[t,t+h]} |\partial_t \tilde{\lambda}_j(s,\ell)|^q ds \right) dt \\
&= \frac{1}{d} \sum_{j=1}^d \int_{[h,1]} |\partial_t \tilde{\lambda}_j(s,\ell)|^q ds \\
&\leq C(\gamma, \ell_0, q)
\end{aligned}$$

where the estimate is uniform for  $\ell \in U_2(\ell_0)$ .

By compactness, we can cover  $\mathbb{G}(1, n)$  with finitely many open sets of the form  $U(q, \gamma, \ell_1), \dots, U(q, \gamma, \ell_r)$  and, setting  $c(q, \gamma) := \max_k C(\gamma, \ell_k, q)$  we have:

$$\begin{aligned}
\int_{[0,1-h]} \left( \frac{W_q(\mu_{t+h}, \mu_t)}{h} \right)^q dt &= \int_{[0,1-h]} \int_{\mathbb{G}(1,n)} \left( \frac{W_q(\mu_{t+h,\ell}, \mu_{t,\ell})}{h} \right)^q \text{vol}_{\mathbb{G}(1,n)}(d\ell) dt \\
&\leq \int_{\mathbb{G}(1,n)} C(\gamma, \ell, q) \text{vol}_{\mathbb{G}(1,n)}(d\ell) < C,
\end{aligned}$$

for any  $h < 1$  and  $C$  independent of  $h$  and  $t$ . This estimate is known to be sufficient to prove the claim  $\mu \circ \gamma \in AC^q(I, \mathcal{P}_q(\mathbb{C}P^n))$ , see for instance [Lis07, Lemma 1].  $\square$

It remains to prove the Lemma used in the proof.

**LEMMA 6.4.** *Let  $p_t : I \rightarrow \mathbb{C}[z_0, \dots, z_n]_{(d)}$  be a curve satisfying the assumptions of Theorem 6.3. For every  $\ell_0 \in \mathbb{G}(1, n)$  denote by  $L_0 \simeq \mathbb{C}^2 \subset \mathbb{C}^{n+1}$  the complex vector space such that  $P(L_0) = \ell_0$ . There exists a neighborhood  $U_1(\ell_0)$  of  $L_0$  in  $G(2, n+1) = \mathbb{G}(1, n)$  and smooth maps  $e_0, e_1 : U(\ell_0) \rightarrow \mathbb{C}^{n+1}$  such that:*

- (i) *for every  $L \in U_1(\ell_0)$ , where  $P(L) = \ell$ , the set  $\{e_0(\ell), e_1(\ell)\}$  is a Hermitian orthonormal basis of  $L$ ;*
- (ii) *for every  $t \in I$  the point  $e_1(\ell_0)$  is not a zero of the polynomial  $p_t$ .*

*Proof.* Consider the set

$$U := \{(z, t) \in \ell_0 \times I \mid p_t(z) = 0\} \subset L_0 \times I,$$

together with the restrictions  $\pi_1 : U \rightarrow \ell_0$  and  $\pi_2 : U \rightarrow I$  of the projections on the two factors. Since  $p_t \notin \Delta_{n,d}^{\text{Fano}}$ , the hypothesis  $d > 2n - 3$  imply that there is no line on  $Z(p_t)$  and in particular for every  $t \in I$  the polynomial  $p_t|_{L_0}$  is not identically zero. For every  $t \in I$ , the set  $\pi_2^{-1}(t)$  consists of finitely many points, therefore it has Hausdorff

dimension zero; consequently  $U$  has Hausdorff dimension at most 1 in  $\ell_0 \times I \simeq \mathbb{CP}^1 \times I$ . Since  $\pi_1$  is a smooth map, then also  $\pi_1(U)$  has Hausdorff dimension at most 1. In particular there exists a point  $x \in \ell_0 \setminus \pi_1(U)$ . This point  $x$  is the projectivization of some complex one-dimensional vector space  $V \subset L_0$  and we can therefore pick a norm one vector  $e_1(\ell_0) \in V$ . We set  $e_0(\ell_0) \in L_0$  to be any norm-one vector in the Hermitian orthogonal complement of  $e_1(\ell_0)$ . There exists now a neighborhood  $U(\ell_0)$  such that for every line  $\ell = P(L) \in U(\ell_0)$ , the orthogonal projections  $\{\text{proj}_L e_0(\ell_0), \text{proj}_L e_1(\ell_0)\}$  form a basis for  $L$ . Applying the Gram–Schmidt procedure to this basis yields the desired functions  $e_0(\cdot), e_1(\cdot) : U(\ell_0) \rightarrow \mathbb{C}^{n+1}$ .  $\square$

In the way of proving Theorem 6.3 we have also proved the following upper bound on the metric speed via integral geometry: using the notation of the above proof, for every curve  $\gamma \in C^{d-1,1}(I, \mathbb{C}[z_0, \dots, z_n]_{(d)})$ , and for every  $1 \leq q < d/(d-1)$ , we have

$$|\dot{\mu}_t|_q^q \leq \int_{\mathbb{G}(1,n)} |\mu_{t,\ell}|_q^q \text{vol}_{\mathbb{G}(1,n)}(d\ell),$$

for almost all  $t \in [0, 1]$ ; here  $|\dot{\mu}_t|_q$  denotes the metric speed of  $\mu_t$  in the  $q$ -Wasserstein space.

We now address the compatibility of Theorem 6.3 with the sharp result obtained in [PR18] by showing that, in the case  $n = 1$ , Theorem 6.3 recovers [PR18].

**6.1 Rational maps.** In this section we prove a further generalization of [PR18, Theorem 1]. The main assumption of Theorem 6.3 was the condition that the curve  $p_t$  does not hit the Fano discriminant  $F_{n,d}$ ; under this assumption, if  $p_t \in \mathcal{C}^{d-1,1}$ , then  $\mu_t \in AC^q(I, \mathcal{P}_q(\mathbb{CP}^n))$  for all  $1 \leq q < d/(d-1)$ . We can remove the condition  $p_t \notin F_{n,d}$ , asking for more regularity on the curve  $p_t$  at the price of losing some regularity on the curve of measures  $\mu_t$ .

**Theorem 6.5.** *For every  $n \in \mathbb{N}$  there exists  $e(n) > 0$  such that for every curve  $p_t \in \mathcal{C}^{k,1}(I, \mathbb{C}[z_0, \dots, z_n]_{(d)})$ , with  $k \geq e(n)d - 1$ , and for every  $1 \leq q < e(n)d/(e(n)d - 1)$  the curve  $\mu_t := \mu(p_t) \in AC^q(I, \mathcal{P}_q(\mathbb{CP}^n))$ .*

The main idea for the proof of Theorem 6.5 is to replace the use of lines, in the integral geometry approach, with the use of rational maps.

**DEFINITION 6.6.** For every  $n, e \in \mathbb{N}$  we denote by  $R_e(\mathbb{CP}^n) \subset \mathbb{CP}^{(n+1)(e+1)-1}$  the set of  $(n+1)$ -tuples  $[q_0, \dots, q_n]$  where  $q_0, \dots, q_n \in H_{1,e}$  are polynomials with no common zeroes. (The set  $R_e(\mathbb{CP}^n)$  is a Zariski open set.) We call elements of  $R_e(\mathbb{CP}^n)$  *rational maps*, since each of them defines a map  $\nu : \mathbb{CP}^1 \rightarrow \mathbb{CP}^n$  by

$$\nu(w) := [q_0(w), \dots, q_n(w)].$$

If  $Z \subset \mathbb{CP}^n$  is a hypersurface, we denote by  $R_e(Z) \subset R_e(\mathbb{CP}^n)$  the set of rational maps such that  $\nu(\mathbb{CP}^1) \subset Z$ . Notice that, given a polynomial  $p \in H_{n,d}$  and a rational map  $\nu : \mathbb{CP}^1 \rightarrow \mathbb{CP}^n$ , the composition  $p \circ \nu$  defines an element of  $H_{1,de}$  (up to multiples) and  $\nu(\mathbb{CP}^1) \subset Z(p)$  if and only if  $p \circ \nu \equiv 0$ .

We can generalize the construction from Sect. 3.3 and recover the map  $\mu : P_{n,d} \rightarrow \mathcal{P}(\mathbb{C}P^n)$  integrating over rational maps of high degree (instead of lines, which are images of rational maps of degree one).

**PROPOSITION 6.7.** *Let  $\nu \in R_e(\mathbb{C}P^n)$  such that  $\nu : \mathbb{C}P^1 \rightarrow \mathbb{C}P^n$  is an embedding. Then, for every  $p \in P_{n,d}$  and  $f \in \mathcal{C}^0(\mathbb{C}P^n)$  we have*

$$\int_{\mathbb{C}P^n} f d\mu(p) = \frac{1}{d \cdot e} \int_{U(n+1)} \sum_{p(g(\nu(w)))=0} m_w(p \circ g \circ \nu) f(g(\nu(w))) dg. \quad (6.1)$$

*Proof.* Observe first that, if  $Z \subset \mathbb{C}P^n$  is a smooth complex hypersurface and  $U \subseteq Z$  is an open set, the integral geometry formula (3.5) applied with the choice  $A = U$  and  $B = \nu(\mathbb{C}P^1)$  gives

$$\int_{U(n+1)} \#U \cap g\nu(\mathbb{C}P^1) dg = \frac{\text{vol}(U)}{\text{vol}(\mathbb{C}P^{n-1})} \frac{\text{vol}(\nu(\mathbb{C}P^1))}{\text{vol}(\mathbb{C}P^1)}.$$

Moreover, letting  $U = Z = \mathbb{C}P^{n-1}$ , we see that  $\frac{\text{vol}(\nu(\mathbb{C}P^1))}{\text{vol}(\mathbb{C}P^1)} = e$ .

Still letting  $Z \subset \mathbb{C}P^n$  be smooth, arguing as in Lemma 3.7 and denoting by  $\lambda : \mathcal{C}^0(Z) \rightarrow \mathbb{R}$  the linear functional

$$\lambda(f) := \int_{U(n+1)} \frac{1}{d \cdot e} \sum_{z \in Z \cap \nu(\mathbb{C}P^1)} f(z) dg,$$

we see that  $\lambda \geq 0$  and we find a unique measure  $\sigma$  on  $Z$  such that  $\lambda(f) = \int_Z f(z) \sigma(dz)$ . Approximating the characteristic function of an open set  $U$  by continuous functions, as in Lemma 3.7, we see that  $\sigma = \text{vol}_Z$ .

Observe now that, if  $p \in P_{n,d} \setminus \Delta_{n,d}$ , then for almost every  $g \in U(n+1)$  the set  $\nu(\mathbb{C}P^1)$  and  $Z(p)$  intersect transversally and for every point of intersection  $z = \nu(w) \in \nu(\mathbb{C}P^1) \cap Z$  we have  $m_w(p \circ g \circ \nu) = 1$ . This proves the identity (6.1) for  $p \in P_{n,d} \setminus \Delta_{n,d}$ . Finally, as in the proof of Theorem 3.8, we see that (6.1) is continuous on  $P_{n,d}$  and, since  $P_{n,d} \setminus \Delta_{n,d}$  is dense in it, then (6.1) agrees with (3.8).  $\square$

We will need the following recent result of B. Lehmann, E. Riedl and S. Tanimoto, privately communicated to the authors. Roughly speaking, the result says that, given positive integers  $n$  and  $T$ , there exists  $\xi(n, T) > 0$  such that for any family  $X \subset R_e(\mathbb{C}P^n)$  with  $e \geq \xi$  and with  $\dim_{\mathbb{C}}(X) \geq (e+1)(n+1) - T$ , we have

$$\bigcup_{\nu \in X} \nu(\mathbb{C}P^1) = \mathbb{C}P^n.$$

In particular such family of curves cannot lie on a single hypersurface.

**PROPOSITION 6.8** (Lehmann–Riedl–Tanimoto). *For every  $n, T \in \mathbb{N}$  there exists  $\xi(n, T) > 0$  such that for every  $e \geq \xi(n, T)$  and for every  $p \in P_{n,d}$  we have:*

$$\dim_{\mathbb{C}}(R_e(Z(p))) \leq (n+1)(e+1) - T.$$

In the special case  $T(n) := \lfloor \frac{(n+1)^2}{2} \rfloor + 1$ , using Proposition 6.8, we get the number

$$e(n) := \xi(n, T(n)). \quad (6.2)$$

The next proposition ensures that, given a curve  $p_t$  of polynomials, we can find a “good” rational map to perform integral geometry on the family  $\{Z(p_t)\}_{t \in I}$ .

**PROPOSITION 6.9.** *Let  $p_t : I \rightarrow P_{n,d}$  be a  $\mathcal{C}^k$  curve, with  $k \geq 1$  and  $e(n) > 0$  be given by (6.2). For every  $e \geq e(n)$  there exists  $\nu \in R_e(\mathbb{C}P^n)$  such that*

- (1) *for every  $t \in I$  and for every  $g \in U(n+1)$  the image of  $g \circ \nu$  is not contained in  $Z(p_t)$ ;*
- (2)  *$\nu : \mathbb{C}P^1 \rightarrow \mathbb{C}P^n$  is an embedding.*

*Proof.* Denote by  $Y \subset I \times R_e(\mathbb{C}P^n)$  the set

$$\begin{aligned} Y &= \left\{ (t, \nu) \mid \exists g \in U(n+1), g\nu(\mathbb{C}P^1) \subset Z(p_t) \right\} \\ &= \left\{ (t, \nu) \mid \nu \in \bigcup_{g \in U(n+1)} gR_e(Z(p_t)) \right\}. \end{aligned}$$

Let also  $\pi_1 : Y \rightarrow I$  and  $\pi_2 : Y \rightarrow R_e(\mathbb{C}P^n)$  denote the projections on the two factors. Notice that every element  $\nu \notin \pi_2(Y)$  satisfies condition (1) of the statement. We will first prove that there is plenty of such elements.

For every  $t \in I$ , the set  $\pi_1^{-1}(t)$  is the continuous semialgebraic image of the semialgebraic set  $U(n+1) \times R_e(Z(p_t))$  under the map  $(g, \nu) \mapsto g\nu$ . In particular

$$\dim(\pi_1^{-1}(t)) \leq \dim(U(n+1)) + 2 \dim_{\mathbb{C}}(R_e(Z(p_t))) \leq \dim_{\mathbb{R}}(R_e(\mathbb{C}P^n)) - 2.$$

This implies that the Hausdorff dimension of  $Y$  is at most  $\dim_{\mathbb{R}}(R_e(\mathbb{C}P^n)) - 1$  and, since  $\pi_2$  is a smooth map, the Hausdorff dimension of  $\pi_2(Y)$  is also at most  $\dim_{\mathbb{R}}(R_e(\mathbb{C}P^n)) - 1$ . This means that  $\pi_2(Y)$  is a semialgebraic set of codimension at least one, in particular its complement  $A_1$  contains an open, dense, semialgebraic set all of whose elements satisfy condition (1) of the statement.

The set  $A_2$  of  $\nu \in R_e(\mathbb{C}P^n)$  satisfying condition (2) of the statement is open and dense. Therefore it is enough to pick  $\nu \in A_1 \cap A_2 \neq \emptyset$  to conclude the proof.  $\square$

We are now ready to prove Theorem 6.5.

*Proof of Theorem 6.5.* We are now in position to revisit the proof of Theorem 6.3.

Consider  $p_t \in \mathcal{C}^{k,1}(I, \mathbb{C}[z_0, \dots, z_n]_{(d)})$  for some natural number  $k$ . By Proposition 6.8 and Proposition 6.9, letting  $e(n)$  be defined by (6.2), for every  $k \geq e(n)$  we have the existence of a  $\nu \in R_e(\mathbb{C}P^n)$  such that  $\nu : \mathbb{C}P^1 \rightarrow \mathbb{C}P^n$  is an embedding and for every  $t \in I$  and for every  $g \in U(n+1)$  the image of  $g \circ \nu$  is not contained in  $Z(p_t)$ . In particular,  $p_{t,g} := p_t \circ g \circ \nu \in H_{1,de}$ , up to multiplies. We can therefore define

$$\mu_{t,g} := (g \circ \nu)_\#(\mu(p_{t,g})) \in \mathcal{P}(\mathbb{C}P^n).$$

Notice indeed that  $\mu(p_{t,g}) \in \mathcal{P}(\mathbb{CP}^1)$  and  $(g \circ \nu) : \mathbb{CP}^1 \rightarrow \mathbb{CP}^n$ . Then Proposition 6.7 can be restated by saying that for any measurable function  $f$

$$\begin{aligned} \int_{\mathbb{CP}^n} f(z) \mu_t(dz) &= \frac{1}{de} \int_{U(n+1)} \int_{\mathbb{CP}^1} f(g(\nu(z))) \left( \sum_{w \in Z(p_{t,g})} m_w(p_{t,g}) \delta_w \right) (dz) dg \\ &= \int_{U(n+1)} \int_{\mathbb{CP}^1} f(g(\nu(z))) \mu(p_{t,g})(dz) dg \\ &= \int_{U(n+1)} \int_{\mathbb{CP}^n} f(z) (g \circ \nu)_\# \mu(p_{t,g})(dz) dg, \end{aligned}$$

giving the concise identity:

$$\mu_t = \int_{U(n+1)} \mu_{t,g} dg.$$

Reasoning as in the proof of Theorem 6.3 where we constructed a coupling between  $\mu_t$  and  $\mu_{t+h}$  by gluing the couplings line by line, one obtains that for any  $q > 1$

$$|\dot{\mu}_t|_q^q \leq \int_{U(n+1)} |\dot{\mu}_{t,g}|_q^q dg.$$

Next we write

$$p_t(g(\nu))(z_0, z_1) = \sum_{k=0}^{de} a_k(t, g) z_0^k z_1^{de-k} \in \mathbb{C}[z_0, z_1]_{(de)};$$

the functions  $t \mapsto a_k(t, g)$  are  $\mathcal{C}^{k-1,1}(I; \mathbb{C})$  and since the image of  $g \circ \nu$  is not contained in  $Z(p_t)$ ,  $p_t(g(\nu))$  is not identically zero. Hence up to pre-composing  $\nu$  with  $h \in U(2)$ , we can assume that

$$a_0(t, g) = p_t(g(\nu))(0, 1) \neq 0.$$

We can now proceed as in the proof of Theorem 6.3 by dehomogenizing the polynomial  $p_{t,g}$  and invoking [PR18, Theorem 1] to deduce that, for any  $1 \leq q < e(n)d/(e(n)d - 1)$ , the curve  $\mu(p_{t,g}) \in AC^q(I; \mathcal{P}_q(\mathbb{CP}^1))$  with the  $L^q$ -norm in time of the  $q$ -metric speed bounded locally uniformly in  $g \in U(n+1)$ .

In order to deduce the same regularity for  $\mu_{t,g}$  we reason as follows. By the isometric injection of each (metric) space  $X$  into the  $q$ -Wasserstein space over itself, the push-forward of the smooth map  $g \circ \nu : \mathbb{CP}^1 \rightarrow \mathbb{CP}^n$  induces a Lipschitz map between  $\mu(P_{1,ed})$  and  $\mathcal{P}(\mathbb{CP}^n)$ ; in particular

$$|\dot{\mu}_{t,g}|_q \leq C(\nu) |\dot{\mu}(p_{t,g})|_q,$$

where  $C(\nu)$  only depends on the Lipschitz constant of  $\nu$ ; hence the proof is completed.  $\square$

This regularity is sufficient to deduce the compactness of the geodesic space  $(P_{n,d}, W_q^{\text{in}})$  for  $q$  in the range of Theorem 6.5.

REMARK 6.10. In the range of Theorem 6.5 we will have  $q \leq 2$ . Hence by Jensen inequality for any couple of probability measures  $W_q(\mu_0, \mu_1) \leq W_2(\mu_0, \mu_1)$  and therefore also the metric speeds will have the same ordering: if  $\mu_t \in AC^2(I; \mathcal{P}_2(\mathbb{C}\mathbb{P}^n))$  then  $\mu_t \in AC^q(I; \mathcal{P}_q(\mathbb{C}\mathbb{P}^n))$  and  $|\dot{\mu}_t|_q \leq |\dot{\mu}_t|_2$ . Hence Corollary 5.14 implies that  $W_q^{\text{in}}$  is finite over  $P_{n,d}$ . Moreover all the conclusions of Theorem 5.15 will remain valid for  $(P_{n,d}, W_q^{\text{in}})$ .

**Theorem 6.11.** *Let  $e(n) \in \mathbb{N}$  be given by (6.2). Then, for every  $1 \leq q < e(n)d/(e(n)d-1)$ , the complete and separable geodesic space  $(P_{n,d}, W_q^{\text{in}})$  is compact.*

*Proof.* Consider a sequence  $p_k \in P_{n,d}$ . Then there exists a subsequence, that we omit in the notation, and  $p \in P_{n,d}$  such that  $d_{\text{FS}}(p_k, p) \rightarrow 0$  as  $k \rightarrow \infty$ . Then, by [KM97, Lemma 12.2], there exists a smooth curve  $\gamma: I \rightarrow P_{n,d}$  and  $t_k \in I$  such that for each  $k \in \mathbb{N}$   $\gamma(t_k) = p_k$  and  $t_k \rightarrow 0$  as  $k \rightarrow \infty$ . Then by Theorem 6.5 the curve  $\mu_t := \mu(\gamma)$  is in  $AC^q(I; \mathcal{P}_q(\mathbb{C}\mathbb{P}^n))$  for  $1 \leq q < e(n)d/(e(n)d-1)$ . Hence

$$W_q^{\text{in}}(p_0, p_k) = W_q^{\text{in}}(\mu_0, \mu_{t_k}) \leq \left( \int_0^{t_k} |\dot{\mu}_s|_q^q ds \right)^{1/q},$$

proving that  $W_q^{\text{in}}(p_0, p_k) \rightarrow 0$  as  $k \rightarrow \infty$  and therefore the claim.  $\square$

## 7 A Wasserstein approach to condition numbers

In this section we adopt the approach of Wasserstein metric geometry for the study of condition numbers of polynomial systems. Let us recall the standard framework from [BCC13].

Given a list of degrees  $\underline{d} = (d_1, \dots, d_n) \in \mathbb{N}^n$  we define  $P_{n,\underline{d}} := P_{n,d_1} \times \dots \times P_{n,d_n}$  and consider the *solution variety*

$$V := \{((p_1, \dots, p_n), [z]) \in P_{n,\underline{d}} \times \mathbb{C}\mathbb{P}^n \mid p_1(z) = \dots = p_n(z) = 0\},$$

together with the two projections  $\pi_1: V \rightarrow P_{n,\underline{d}}$  and  $\pi_2: V \rightarrow \mathbb{C}\mathbb{P}^n$ . We denote by  $\Sigma' \subset V$  the set of critical points of  $\pi_1$ .

The *discriminant*  $\Delta_{n,\underline{d}} \subset P_{n,\underline{d}}$  is defined by

$$\Delta_{n,\underline{d}} := \{(p_1, \dots, p_n) \in P_{n,\underline{d}} \mid p_1(z) = \dots = p_n(z) = 0 \text{ is not regular on } \mathbb{C}^{n+1} \setminus \{0\}\}.$$

Given  $[w] \in \mathbb{C}\mathbb{P}^n$  we also consider  $\Delta_{n,\underline{d},[w]} \subset \Delta_{n,\underline{d}}$  defined by

$$\Delta_{n,\underline{d},[w]} := \{(p_1, \dots, p_n) \in P_{n,\underline{d}} \mid p_1(z) = \dots = p_n(z) = 0 \text{ is not regular at } w\}.$$

The *normalized condition number*  $\nu_{\text{norm}}: V \setminus \Sigma' \rightarrow [1, \infty)$  is defined by

$$\nu_{\text{norm}}(p, [w]) := \frac{1}{d_{\text{BW}}(p, \Delta_{n,\underline{d},[w]})},$$

where  $\mathbf{d}_{\text{BW}}$  denotes the metric induced by the Bombieri–Weyl structure, see Sect. A.4. The normalized condition number can be used to turn  $V \setminus \Sigma'$  into a *Lipschitz* Riemannian manifold as follows. First, we denote by  $g_V$  the Riemannian structure on  $V \subset \mathbb{P}(H_{n,d}) \times \mathbb{C}\mathbb{P}^n$  induced by the product Riemannian structure  $g_{\text{BW}} \times g_{\text{FS}}$ . Then, for every absolutely continuous curve  $\gamma_t = (p_t, z_t) : I \rightarrow V \setminus \Sigma'$  we define its *condition length* by

$$L_{\text{cond}}(\gamma_t) := \int_I \nu_{\text{norm}}(p_t, z_t) \cdot \|\dot{\gamma}_t\|_V dt.$$

Taking the infimum of the length over all absolutely continuous curves joining two points endows  $V \setminus \Sigma'$  with a metric structure, called the *condition metric*. We call a curve  $\gamma_t$  in  $V \setminus \Sigma'$  a condition geodesic if it minimizes the condition length between any two of its points. Denoting by  $D := \max\{d_1, \dots, d_n\}$ , in [Shu09] it is shown that  $O(d^{3/2} L_{\text{cond}}(\gamma))$  Newton steps are sufficient to continue the zero  $z_0$  from  $p_0$  to  $p_1$  along  $\gamma_t$ .

In [B+12, B+09] it is asked whether the function  $t \mapsto \log(\nu_{\text{norm}}(\gamma_t))$  is convex along a condition geodesic. An affirmative answer to this question would imply that, for any condition geodesic  $\gamma_t$ , one has  $L_{\text{cond}}(\gamma_t) \leq L_V(\gamma_t) \max\{\nu_{\text{norm}}(p_0, z_0), \nu_{\text{norm}}(p_1, z_1)\}$ , where  $L_V(\gamma_t)$  is the length of  $\gamma_t$  in the Riemannian metric of  $V$  defined above. This problem is also formulated as an open problem in [BCC13, Problem 14].

Here we propose a solution to this problem in the case  $n = 1$ , replacing  $\gamma_t$  with a lift of a Wasserstein geodesic. We expect a similar result to be true in the case  $n > 1$ , but this will be the subject of future work.

**Theorem 7.1.** *For every  $d \in \mathbb{N}$  there exist  $\beta_1, \beta_2, \beta_3, \beta_4 > 0$  such that, if  $p_t : I \rightarrow P_{1,d}$  is a  $W_2$ -geodesic between  $p_0, p_1 \in P_{1,d} \setminus \Delta_{1,d}$  then the following is true:*

(1) *for every  $t \in I$*

$$\text{dist}(p_t, \Delta_{1,d})^2 \geq \beta_1 \min\{\text{dist}(p_0, \Delta_{1,d}), \text{dist}(p_1, \Delta_{1,d})\}^{\beta_2}.$$

*In particular,  $P_{1,d} \setminus \Delta_{1,d}$  is  $W_2$ -geodesically convex;*

(2) *if  $\gamma : I \rightarrow V \setminus \Sigma'$  is a lift of  $p_t$ , i.e.  $\gamma(t) = (p_t, z_t)$ , then:*

$$L_{\text{cond}}(\gamma) \leq \beta_3 \cdot L_V(\gamma) \cdot \max\{\nu_{\text{norm}}(p_0, z_0), \nu_{\text{norm}}(p_1, z_1)\}^{\beta_4}.$$

Before proving the theorem, we need to introduce a couple more ingredients. For the ease of notation, we simply denote by  $p \in P_{1,d}$  the class of a polynomial and by  $w \in \mathbb{C}\mathbb{P}^1$  the class of a point (i.e. we omit the square brackets in the notation). Consider the following two functions  $\alpha_1, \alpha_2 : P_{1,d} \times \mathbb{C}\mathbb{P}^1 \rightarrow [0, \infty)$

$$\alpha_1(p, w) := \nu_{\text{norm}}(p, w)^{-1} \quad \text{and} \quad \alpha_2(p, w) := \min_{p(z_i)=p(z_j)=0, z_i \neq z_j} \mathbf{d}_{\text{FS}}(w, z_i)^2 + \mathbf{d}_{\text{FS}}(w, z_j)^2.$$

Using the semialgebraic homeomorphism  $P_{1,d} \simeq \text{SP}^d(\mathbb{C}\mathbb{P}^1)$ , it is easy to see that  $\alpha_2$  (which a priori is defined only  $P_{1,d} \setminus \Delta_{1,d}$ ) is actually continuous on  $P_{1,d}$ .

It will be also convenient to consider the following functions  $\alpha_3, \alpha_4 : P_{1,d} \rightarrow [0, \infty)$

$$\alpha_3(p) := \text{dist}(p_t, \Delta_{1,d})^2 \quad \text{and} \quad \alpha_4(p) := \min_{p(z_i)=p(z_j)=0, z_i \neq z_j} \mathbf{d}_{\text{FS}}(z_i, z_j)^2.$$

Functions  $\alpha_3$  and  $\alpha_4$  are the analogous of  $\alpha_1$  and  $\alpha_2$  where we do not fix a specified point  $w \in \mathbb{CP}^1$ . As above,  $\alpha_4$  (which a priori is defined only  $P_{1,d} \setminus \Delta_{1,d}$ ) is actually continuous on  $P_{1,d}$ .

The key tools for the proof of Theorem 7.1 are Lemma 7.2 and Proposition 7.3 below. The first one relates the function  $\alpha_1$  to  $\alpha_2$  and, in the same spirit, the function  $\alpha_3$  to  $\alpha_4$ . The second one studies concavity properties of  $\alpha_2$  and  $\alpha_4$  – this is where optimal transport plays the key role giving the main estimate.

LEMMA 7.2. *There exists  $c_1, c_2, c_3, c_4 > 0$  such that*

$$\alpha_2^{c_2} \leq c_1 \alpha_1 \quad \text{and} \quad \alpha_1^{c_3} \leq c_4 \alpha_2.$$

Analogously, there exists  $c_5, c_6, c_7, c_8 > 0$  such that

$$\alpha_4^{c_6} \leq c_5 \alpha_3 \quad \text{and} \quad \alpha_3^{c_7} \leq c_8 \alpha_4.$$

*Proof.* All the functions  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  are continuous and semialgebraic. Moreover,  $\alpha_1$  and  $\alpha_2$  have the same zero sets, and are defined on the compact set  $P_{1,d} \times \mathbb{CP}^1$ . Therefore the first result follows from Lojasiewicz’s inequality. Analogously,  $\alpha_3$  and  $\alpha_4$  have the same zero sets, and are defined on the compact set  $P_{1,d}$  hence Lojasiewicz’s inequality proves the second claim.  $\square$

PROPOSITION 7.3. *The functions  $\alpha_2$  and  $\alpha_4$  are “quasi-concave along  $W_2$ -geodesics”, i.e. given a curve  $\gamma_t = (p_t, z_t) : I \rightarrow V$  with  $p_t : I \rightarrow P_{1,d}$  a  $W_2$ -geodesic, then for every  $t \in [0, 1]$*

$$\alpha_2(p_t, z_t) \geq (1-t)^2 \alpha_2(p_0, z_0) + t^2 \alpha_2(p_1, z_1),$$

and

$$\alpha_4(p_t) \geq (1-t)^2 \alpha_4(p_0) + t^2 \alpha_4(p_1).$$

*Proof.* For every  $t \in I$ , let  $z_i(t)$  be a zero of  $p_t$ . Since  $z_t$  is also a zero of  $p_t$  and  $p_t$  is a  $W_2$ -geodesic, the optimality of the coupling gives the following estimate (see [Vil09, Eq. (8.45)] and [FJ08, Remark 2.1])

$$\mathbf{d}_{\text{FS}}(z_t, z_i(t))^2 \geq (1-t)^2 \mathbf{d}_{\text{FS}}(z_0, z_i(0))^2 + t^2 \mathbf{d}_{\text{FS}}(z_1, z_i(1))^2.$$

For every  $t \in I$ , let now  $z_i(t), z_j(t)$  be two zeros of  $p_t$  such that

$$\alpha_2(p_t, z_t) = \mathbf{d}_{\text{FS}}(z_t, z_i(t))^2 + \mathbf{d}_{\text{FS}}(z_t, z_j(t))^2.$$

Then

$$\begin{aligned}
\alpha_2(p_t, z_t) &= \mathbf{d}_{\text{FS}}(z_t, z_i(t))^2 + \mathbf{d}_{\text{FS}}(z_t, z_j(t))^2 \\
&\geq (1-t)^2 \left( \mathbf{d}_{\text{FS}}(z_0, z_i(0))^2 + \mathbf{d}_{\text{FS}}(z_0, z_j(0))^2 \right) \\
&\quad + t^2 \left( \mathbf{d}_{\text{FS}}(z_1, z_i(1))^2 + \mathbf{d}_{\text{FS}}(z_0, z_j(0))^2 \right) \\
&\geq (1-t)^2 \min_{i \neq j} \left( \mathbf{d}_{\text{FS}}(z_0, z_i(0))^2 + \mathbf{d}_{\text{FS}}(z_0, z_j(0))^2 \right) \\
&\quad + t^2 \min_{i \neq j} \left( \mathbf{d}_{\text{FS}}(z_1, z_i(1))^2 + \mathbf{d}_{\text{FS}}(z_0, z_j(0))^2 \right) \\
&= (1-t)^2 \alpha_1(p_0, z_0) + t^2 \alpha_2(p_1, z_1).
\end{aligned}$$

The same argument works also for  $\alpha_4$ . □

*Proof of Theorem 7.1.* For the first part of the statement follows by direct application of Lemma 7.2 and Proposition 7.3:

$$\begin{aligned}
(\alpha_3(p_t) c_5)^{1/c_6} &\geq \alpha_4(p_t) \\
&\geq (1-t)^2 \alpha_4(p_0) + t^2 \alpha_4(p_1) \\
&\geq \frac{1}{c_8} \left( (1-t)^2 \alpha_3(p_0)^{c_7} + t^2 \alpha_3(p_1)^{c_7} \right).
\end{aligned}$$

For the second part of the statement, we observe that, by Lemma 7.2,

$$\nu_{\text{norm}}(p_t, z_t) = \alpha_1(p_t, z_t)^{-1} \leq c_1 \alpha_2(p_t, z_t)^{-c_2}.$$

Using Proposition 7.3, and the convexity of  $x \mapsto 1/x$ , we get

$$\alpha_2(p_t, z_t)^{-1} \leq \left( \alpha_2(p_0, z_0)^{-1} + \alpha_2(p_1, z_1)^{-1} \right),$$

which in particular gives

$$\nu_{\text{norm}}(p_t, z_t) \leq c_1 \left( \alpha_2(p_0, z_0)^{-1} + \alpha_2(p_1, z_1)^{-1} \right)^{c_2}.$$

Using again Lemma 7.2, we get

$$\begin{aligned}
c_1 \left( \alpha_2(p_0, z_0)^{-1} + \alpha_2(p_1, z_1)^{-1} \right)^{c_2} &\leq c_1 \left( c_4 \alpha_1(p_0, z_0)^{-c_3} + c_4 \alpha_1(p_1, z_1)^{-c_3} \right)^{c_2} \\
&\leq (c_1 c_4^{c_2}) \max \{ \alpha_1(p_0, z_0)^{-1}, \alpha_1(p_1, z_1)^{-1} \}^{c_3 c_2} \\
&= \beta_1 \max \{ \nu_{\text{norm}}(p_0, z_0), \nu_{\text{norm}}(p_1, z_1) \}^{\beta_2}.
\end{aligned}$$

This proves that along a  $W_2$ -geodesic:

$$\nu_{\text{norm}}(p_t, z_t) \leq \beta_1 \max \{ \nu_{\text{norm}}(p_0, z_0), \nu_{\text{norm}}(p_1, z_1) \}^{\beta_2}$$

and therefore,

$$\begin{aligned} L_{\text{cond}}(\gamma_t) &= \int_I \nu_{\text{norm}}(p_t, z_t) \cdot \|\dot{\gamma}_t\|_V dt \\ &\leq \beta_1 \max\{\nu_{\text{norm}}(p_0, z_0), \nu_{\text{norm}}(p_1, z_1)\}^{\beta_2} \int_I \|\dot{\gamma}_t\|_V dt. \quad \square \end{aligned}$$

REMARK 7.4. Two comments are in order. The first is that from the proof of the first part claim of Theorem 7.1 it follows a slightly stronger statement than the convexity of  $P_{1,d} \setminus \Delta_{1,d}$ . Indeed the quasi-concavity of  $\alpha_4$  yields that if  $p_t$  is a  $W_2$ -geodesic with  $p_0$  or  $p_1$  not in  $\Delta_{1,d}$  then  $p_t \notin \Delta_{1,d}$  for all  $t \in (0, 1)$ .

The second one is that quasi-concavity implies the following inequality that we might call quasi-log-concavity: from

$$\alpha(p_t) \geq (1-t)^2 \alpha(p_0) + t^2 \alpha(p_1),$$

one deduces from the concavity of the log

$$\log(\alpha(p_t)) \geq (1-t) \log(\alpha(p_0)) + t \log(\alpha(p_1)) + (1-t) \log(1-t) + t \log(t).$$

Applying this inequality to  $\alpha_2$  and invoking Lemma 7.2 one obtains that there exists universal constant  $\tilde{c}_1, \tilde{c}_2, \tilde{c}_3$ , such that

$$\begin{aligned} \tilde{c}_1 \log(\nu_{\text{norm}}(\gamma_t)) &\leq \tilde{c}_2 (1-t) \log(\nu_{\text{norm}}(\gamma_0)) + \tilde{c}_3 t \log(\nu_{\text{norm}}(\gamma_1)) \\ &\quad + (1-t) \log\left(\frac{1}{1-t}\right) + t \log\left(\frac{1}{t}\right), \end{aligned}$$

for any  $\gamma_t = (p_t, z_t) : I \rightarrow V \setminus \Sigma'$  lift of a  $W_2$ -geodesic  $p_t : I \rightarrow P_{1,d}$ .

## Appendix: Projective geometry

### A.1: The Fubini–Study structure

In this section we recall the Fubini–Study construction, which is a procedure to endow the projective space  $\mathbb{C}P^n = \mathbb{P}(\mathbb{C}^{n+1})$  with a Riemannian Hermitian structure, once a Hermitian product is chosen on  $\mathbb{C}^{n+1}$ . We refer the reader to [GH94] for more details.

The construction works for any Hermitian product  $h$  on  $\mathbb{C}^{n+1}$ , but here we consider simply the standard one  $h_{\text{std}}(v, w) = \langle v, w \rangle := \sum v_j \bar{w}_j$ . Denote by  $S^{2n+1} \hookrightarrow \mathbb{C}^{n+1}$  the unit sphere with respect to the induced scalar product and by  $\pi : S^{2n+1} \rightarrow \mathbb{C}P^n$  the quotient map. Given  $b \in S^{2n+1}$  we can write:

$$T_b S^{2n+1} = b^\perp \oplus \text{span}_{\mathbb{R}}\{ib\},$$

where this decomposition is orthogonal with respect to the scalar product induced by the Hermitian structure.

Recall also that given  $[b] \in \mathbb{C}P^n$  we have

$$T_{[b]}\mathbb{C}P^n \simeq \text{Hom}(\text{span}_{\mathbb{C}}\{b\}, \mathbb{C}^{n+1}/\text{span}_{\mathbb{C}}\{b\}) \simeq b^\perp.$$

In particular, picking the representative  $b \in S^{2n+1}$ , for two tangent vectors  $v_1, v_2 \in T_{[b]}\mathbb{C}P^n \simeq b^\perp$  we can define

$$h_{\text{FS},[b]}(v_1, v_2) := \langle v_1, v_2 \rangle.$$

This defines a Hermitian structure on each tangent space to  $\mathbb{C}P^n$ , which we call the *Fubini–Study* structure (hence the subscript “FS” in the notation  $h_{\text{FS}}$ ). The real part of  $h_{\text{FS}}$  is a Riemannian metric  $g_{\text{FS}} := \text{Re}(h_{\text{FS}})$  that we call the *Fubini–Study metric* and its imaginary part  $\omega_{\text{FS}} := \text{Im}(h_{\text{FS}})$  is a Kähler form that we call the *Fubini–Study Kähler form*. With this definition the map  $\pi : S^{2n+1} \rightarrow \mathbb{C}P^n$  becomes a Riemannian submersion. Moreover, for every  $R \in \text{U}(n+1)$  the map  $[v] \mapsto [Rv]$  is an isometry of  $(\mathbb{C}P^n, h_{\text{FS}})$ .

Let now  $[b] \in \mathbb{C}P^n$  with  $b \in S^{2n+1}$ . We define a parametrization  $\phi_b : \mathbb{C}^n \rightarrow U_b$  (i.e. the inverse of a chart) of a neighborhood  $U_b \subset \mathbb{C}P^n$  of  $[b]$ . First, let  $U_b := \{[z] \in \mathbb{C}P^n \mid \langle z, b \rangle \neq 0\}$  (i.e. the complement of the hyperplane  $b^\perp$ ). Let also  $R \in \text{U}(n+1)$  be such that  $Re_0 = b$  and define

$$\phi_b(v) := \left[ R \begin{pmatrix} 1 \\ v \end{pmatrix} \right]. \quad (\text{A.1})$$

Notice that  $\phi_b(0) = [b]$ . A useful property of this map is that  $D_0\phi_b$  gives an isometry

$$D_0\phi_b : (\mathbb{C}^n, h_{\text{std}}) \xrightarrow{\simeq} (T_{[b]}\mathbb{C}P^n, h_{\text{FS},[b]}). \quad (\text{A.2})$$

## A.2: Vector bundles

In this section we explain why homogeneous polynomials are sections of an appropriate vector bundle (and not functions) on the projective space. Let us recall some basics notions first.

Recall that if  $\pi : L \rightarrow M$  is a rank- $k$  vector bundle on  $M$ , then there exists an open cover  $\{U_i\}_{i \in I}$  for  $M$  and, for every  $i \in I$ , a smooth map  $\psi_i : L|_{U_i} = \pi^{-1}(U_i) \rightarrow U_i \times \mathbb{R}^k$  such that

- (1)  $\pi = p_1 \circ \psi_i$ , where  $p_1 : U_i \times \mathbb{R}^k$  denotes the projection on the first factor;
- (2) for every  $i, j \in I$  such that  $U_i \cap U_j \neq \emptyset$ , and for every  $(x, v) \in (U_i \cap U_j) \times \mathbb{R}^k$  we have

$$\psi_i \circ \psi_j^{-1}(x, v) = (x, g_{ij}(x)v),$$

where  $g_{ij}(x) \in \text{GL}(\mathbb{R}^k)$ .

The family  $\{(U_i, \psi_i)\}_{i \in I}$  is called a *vector bundle atlas* and the family  $\{g_{ij}\}_{i, j \in I}$  is called the *cocycle* of the vector bundle atlas. The cocycle satisfies  $g_{ij}g_{jk} = g_{ik}$  (on the

common domains of definition) and  $g_{ii} \equiv \mathbf{1}$ . It turns out that in order to construct a vector bundle it is enough to assign its cocycle. A *complex* rank- $k$  vector bundle is obtained by requiring, in the above construction, that  $g_{ij}(x) \in \mathrm{GL}(\mathbb{C}^k)$ .

We denote by  $\Gamma(L, U)$  the space of sections of  $L$  over  $U \subseteq M$ . Let now  $\sigma \in \Gamma(L, M)$  be a global section of  $L$  (a real or complex vector bundle) and  $\{(U_i, \psi_i)\}_{i \in I}$  be a vector bundle atlas (real or complex). Then, for every  $i \in I$ , the map  $\sigma_i := \psi_i \circ \sigma|_{U_i} \in \Gamma(L, U_i)$ , being a section, has the form

$$\sigma_i(x) = (x, \tilde{\sigma}_i(x)),$$

for some map  $\tilde{\sigma}_i : U_i \rightarrow \mathbb{R}^k$ . By definition of the cocycle, the family of maps  $\{\tilde{\sigma}_i\}_{i \in I}$  satisfies

$$\tilde{\sigma}_i(x) = g_{ij}(x) \tilde{\sigma}_j(x) \quad \forall x \in U_i \cap U_j. \quad (\text{A.3})$$

Viceversa, given a family of smooth functions  $\{\tilde{\sigma}_i\}_{i \in I}$  satisfying (A.3), there exists a section  $\sigma : M \rightarrow L$  such that  $\psi_i(\sigma(x)) = (x, \tilde{\sigma}_i(x))$  for all  $i \in I$  and  $x \in U_i$ .

### A.3: Homogeneous polynomials as sections of a vector bundle

Let us move now to homogeneous polynomials. Consider the open cover  $\{U_i\}_{i=0, \dots, n}$  of  $\mathbb{C}\mathbb{P}^n$  given by  $U_i := \{[z_0, \dots, z_n] \mid z_i \neq 0\}$ . Given a homogeneous polynomial  $p \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$ , we can define the zero set

$$Z(p) := \{[z_0, \dots, z_n] \mid p(z) = 0\} \subset \mathbb{C}\mathbb{P}^n.$$

This makes sense because  $p$  vanishes at  $z \in \mathbb{C}^{n+1} \setminus \{0\}$  if and only if it vanishes at  $\lambda z$ , for every  $\lambda \neq 0$ , while the value of  $p$  at a point  $[z] \in \mathbb{C}\mathbb{P}^n$  is not well defined. On the other hand, for every  $i = 0, \dots, n$  we can consider the function  $\tilde{\sigma}_{p,i} : U_i \rightarrow \mathbb{C}$  given by

$$\tilde{\sigma}_{p,i}([z]) := p(z_0/z_i, \dots, z_n/z_i). \quad (\text{A.4})$$

Notice that  $Z(p) \cap U_i = \tilde{\sigma}_{p,i}^{-1}(0)$ . Moreover, for every  $i \neq j$  and  $[z] \in U_i \cap U_j$  we have

$$\tilde{\sigma}_{p,i}([z]) = (z_j/z_i)^d \tilde{\sigma}_{p,j}([z]).$$

Now, observe that the family of functions  $\{g_{ij}([z]) := (z_j/z_i)^d\}_{ij}$ , which are defined on  $U_i \cap U_j$  and take values in  $\mathrm{GL}(\mathbb{C}) = \mathbb{C}^*$ , satisfy the cocycle condition. Therefore, by what we explained above, they define a complex line bundle that algebraic geometers denote by  $\mathcal{O}_{\mathbb{C}\mathbb{P}^n}(d)$ ; we will use the notation  $L_{n,d}$  for this bundle. Because of (A.4), the family of functions  $\{\tilde{\sigma}_{p,i}\}_{i=0, \dots, n}$  glue together to a section of  $L_{n,d}$  which we denote by  $\sigma_p$ .

In this way we have constructed a linear map  $\sigma_{(\cdot)} : \mathbb{C}[z_0, \dots, z_n]_{(d)} \rightarrow \Gamma(L_{n,d}, \mathbb{C}\mathbb{P}^n)$  that associates to a polynomial  $p$  the section  $\sigma_p$  which in coordinates is written as in (A.4). One can show that this map is injective and that its image coincides with the

set of global holomorphic sections of  $L_{n,d}$ . Moreover, denoting by  $A \hookrightarrow L$  the zero section, we have  $\sigma_p^{-1}(A) = Z(p)$ .

In particular, for every  $i = 0, \dots, n$  the polynomial  $z_i^d \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$  corresponds to a section  $\sigma_{z_i^d}$  that, by construction (A.4) satisfies:

$$\tilde{\sigma}_{z_i^d, i}([z_0, \dots, z_n]) = 1. \quad (\text{A.5})$$

Notice that, using the standard Hermitian structure, we can write  $z_i^d = \langle z, e_i \rangle^d$ , where  $e_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{C}^{n+1}$  (the “1” is in position  $i$ ). In other words,  $\sigma_{z_i^d}$  is a section of  $L_{n,d}$  that does not vanish on a neighborhood  $U_i = \{z_i \neq 0\}$  of  $[e_i]$ . We can build similar sections for every other point  $[b] \in \mathbb{C}\mathbb{P}^n$  as follows. Let  $R \in \text{GL}(n+1, \mathbb{C})$  be such that  $Re_0 = b$  and consider the polynomial  $p_b(z) := \langle z, b \rangle^d$ . Since  $p_b(b) = \|b\|^2 \neq 0$ , the section

$$\sigma_b := \sigma_{p_b} \quad (\text{A.6})$$

does not vanish at  $[b]$  and, consequently, it does not vanish on the neighborhood  $U_b := \{\langle z, b \rangle \neq 0\}$  of  $[b]$ . We can use this section to produce a trivialization of  $L_{n,d}$  on  $U_b$ , defining  $\psi_b : L_{n,d}|_{U_b} \rightarrow U_b \times \mathbb{C}$  by

$$\psi_b(w) := \left( \pi(w), \frac{w}{\sigma_b(\pi(w))} \right).$$

Observe that, given any other section  $\sigma = \sigma_p \in \Gamma(L_{n,d}, U_b)$ , with  $p \in \mathbb{C}[z_0, \dots, z_n]_{(d)}$ , we have:

$$\psi_b(\sigma([z])) = \left( [z], \frac{\sigma_p([z])}{\sigma_b([z])} \right) = \left( [z], \frac{p(z)}{\langle z, b \rangle^d} \right). \quad (\text{A.7})$$

#### A.4: The Bombieri–Weyl Hermitian structure

Consider now the *Bombieri–Weyl* Hermitian product on  $\mathbb{C}[z_0, \dots, z_n]_{(d)}$ , defined by:

$$\langle p_1, p_2 \rangle_{\text{BW}} := \frac{1}{\pi^{n+1}} \int_{\mathbb{C}^{n+1}} p_1(z) \overline{p_2(z)} e^{-\|z\|^2} dz,$$

where  $dz := (i/2)^{n+1} dz_0 d\bar{z}_0 \cdots dz_n d\bar{z}_n$  is the Lebesgue measure. It is not difficult to show that the Bombieri–Weyl Hermitian structure is invariant under the action of  $U(n+1)$  by change of variables. Moreover, since this action is irreducible (over  $\mathbb{C}$ ), Schur’s Lemma implies that this is the unique Hermitian structure with this property (up to multiples). A Hermitian orthonormal basis for the Bombieri–Weyl product is given by

$$\left\{ \left( \frac{d!}{\alpha_0! \cdots \alpha_n!} \right)^{\frac{1}{2}} z_0^{\alpha_0} \cdots z_n^{\alpha_n} \right\}_{\{|\alpha|=d\}}. \quad (\text{A.8})$$

If we apply now the Fubini–Study construction above to the space  $P(\mathbb{C}[z_0, \dots, z_n]_{(d)})$  we obtain a Hermitian structure on it, which we still call the Bombieri–Weyl structure. The unitary group  $U(n+1)$  acts on the projective space  $P(\mathbb{C}[z_0, \dots, z_n]_{(d)})$  by change of variables and the action is by isometries.

## Funding

Open access funding provided by Università del Salento within the CRUI-CARE Agreement.

**Open Access.** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- [AGS05] AMBROSIO, L., GIGLI, N., SAVARÉ, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel (2005).
- [AT04] AMBROSIO, L., TILLI, P.: Topics on Analysis in Metric Spaces. Oxf. Lect. Ser. Math. Appl., vol. 25. Oxford University Press, Oxford (2004).
- [ACL23] ANTONINI, P., CAVALLETTI, F., LERARIO, A.: A proof of the Hamiltonian Thom isotopy lemma. *Note Mat.* **43**(2), 13–26 (2023).
- [Aur97] AUROUX, D.: Asymptotically holomorphic families of symplectic submanifolds. *Geom. Funct. Anal.* **7**(6), 971–995 (1997).
- [B+09] BELTRÁN, C., DEDIEU, J.-P., MALAJOVICH, G., SHUB, M.: Convexity properties of the condition number. *SIAM J. Matrix Anal. Appl.* **31**(3), 1491–1506 (2009).
- [B+12] BELTRÁN, C., DEDIEU, J.-P., MALAJOVICH, G., SHUB, M.: Convexity properties of the condition number II. *SIAM J. Matrix Anal. Appl.* **33**(3), 905–939 (2012).
- [BP09] BELTRÁN, C., PARDO, L.M.: Smale’s 17th problem: average polynomial time to compute affine and projective solutions. *J. Am. Math. Soc.* **22**(2), 363–385 (2009).
- [BB00] BENAMOU, J.-D., BRENIER, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84**(3), 375–393 (2000).
- [BCR98] BOCHNAK, J., COSTE, M., ROY, M.-F.: Real Algebraic Geometry. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 36. Springer, Berlin (1998). Translated from the 1987 French original, Revised by the authors.
- [BBI01] BURAGO, D., BURAGO, Y., IVANOV, S.: A Course in Metric Geometry. Graduate Studies in Mathematics, vol. 33. Am. Math. Soc., Providence (2001).
- [BC11] BÜRGISSER, P., CUCKER, F.: On a problem posed by Steve Smale. *Ann. Math. (2)* **174**(3), 1785–1836 (2011).

- [BCC13] BÜRGISSER, P., CUCKER, F.: Condition: The Geometry of Numerical Algorithms. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 349. Springer, Heidelberg (2013).
- [Sil06] CANNAS DA SILVA, A.: Symplectic geometry. In: Handbook of Differential Geometry. Vol. II, pp. 79–188. Elsevier, Amsterdam (2006).
- [CM14] CARLEN, E.A., MAAS, J.: An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker-Planck equation is gradient flow for the entropy. *Commun. Math. Phys.* **331**(3), 887–926 (2014).
- [Cas10] CASTILLON, P.: Submanifolds, isoperimetric inequalities and optimal transportation. *J. Funct. Anal.* **259**(1), 79–103 (2010).
- [EH16] EISENBUD, D., HARRIS, J.: 3264 and All That—A Second Course in Algebraic Geometry. Cambridge University Press, Cambridge (2016).
- [Erb10] ERBAR, M.: The heat equation on manifolds as a gradient flow in the Wasserstein space. *Ann. Inst. Henri Poincaré Probab. Stat.* **46**, 1 (2010).
- [Fed69] FEDERER, H.: Geometric Measure Theory. Springer, New York (1969).
- [FG16] FIGALLI, A., GUIONNET, A.: Universality in several-matrix models via approximate transport maps. *Acta Math.* **217**(1), 81–176 (2016).
- [FJ08] FIGALLI, A., JUILLET, N.: Absolute continuity of Wasserstein geodesics in the Heisenberg group. *J. Funct. Anal.* **255**(1), 133–141 (2008).
- [FMP10] FIGALLI, A., MAGGI, F., PRATELLI, A.: A mass transportation approach to quantitative isoperimetric inequalities. *Invent. Math.* **182**(1), 167–211 (2010).
- [GM00] GANGBO, W., MCCANN, R.J.: Shape recognition via Wasserstein distance. *Q. Appl. Math.* **58**(4), 705–737 (2000).
- [Gig11] GIGLI, N.: On the inverse implication of Brenier-McCann theorems and the structure of  $(P_2(M), W_2)$ . *Methods Appl. Anal.* **18**(2), 127–158 (2011).
- [GP20] GIGLI, N., PASQUALETTO, E.: Lectures on Nonsmooth Differential Geometry. SISSA Springer Series, vol. 2. Springer, Cham (2020).
- [GH94] GRIFFITHS, P., HARRIS, J.: Principles of Algebraic Geometry. Wiley Classics Library. Wiley, New York (1994). Reprint of the 1978 original.
- [How93] HOWARD, R.: The kinematic formula in Riemannian homogeneous spaces. *Mem. Am. Math. Soc.* **106**(509), vi+69 (1993).
- [JKO98] JORDAN, R., KINDERLEHRER, D., OTTO, F.: The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29**(1), 1–17 (1998).
- [KM18] KETTERER, C., MONDINO, A.: Sectional and intermediate Ricci curvature lower bounds via optimal transport. *Adv. Math.* **329**, 781–818 (2018).
- [KW12] KITAGAWA, J., WARREN, M.: Regularity for the optimal transportation problem with Euclidean distance squared cost on the embedded sphere. *SIAM J. Math. Anal.* **44**(4), 2871–2887 (2012).
- [KM97] KRIEGL, A., MICHOR, P.W.: The Convenient Setting of Global Analysis. Mathematical Surveys and Monographs, vol. 53. Am. Math. Soc., Providence (1997).
- [Lai17] LAIREZ, P.: A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time. *Found. Comput. Math.* **17**(5), 1265–1292 (2017).
- [Lai20] LAIREZ, P.: Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems. *J. Am. Math. Soc.* **33**(2), 487–526 (2020).
- [Lis07] LISINI, S.: Characterization of absolutely continuous curves in Wasserstein spaces. *Calc. Var. Partial Differ. Equ.* **28**(1), 85–120 (2007).
- [Lot17] LOTT, J.: On tangent cones in Wasserstein space. *Proc. Am. Math. Soc.* **145**(7), 3127–3136 (2017).

- [LV09] LOTT, J., VILLANI, C.: Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math. (2)* **169**(3), 903–991 (2009).
- [MS11] MCCANN, R.J., SOSIO, M.: Hölder continuity for optimal multivalued mappings. *SIAM J. Math. Anal.* **43**(4), 1855–1871 (2011).
- [Oht09] OHTA, S.-I.: Gradient flows on Wasserstein spaces over compact Alexandrov spaces. *Am. J. Math.* **131**(2), 475–516 (2009).
- [PR18] PARUSIŃSKI, A., RAINER, A.: Optimal Sobolev regularity of roots of polynomials. *Ann. Sci. Éc. Norm. Supér. (4)* **51**(5), 1343–1387 (2018).
- [PR23] PARUSIŃSKI, A., RAINER, A.: Perturbation theory of polynomials and linear operators (2023).
- [Rud87] RUDIN, W.: Real and Complex Analysis, 3rd edn. McGraw-Hill, New York (1987).
- [Shu09] SHUB, M.: Complexity of Bezout’s theorem. VI. Geodesics in the condition (number) metric. *Found. Comput. Math.* **9**(2), 171–178 (2009).
- [SS931] SHUB, M., SMALE, S.: Complexity of Bézout’s theorem. I. Geometric aspects. *J. Am. Math. Soc.* **6**(2), 459–501 (1993).
- [SS932] SHUB, M., SMALE, S.: Complexity of Bezout’s theorem. II. Volumes and probabilities. In: Computational Algebraic Geometry (Nice, 1992). *Progr. Math.*, vol. 109, pp. 267–285. Birkhäuser, Boston (1993).
- [SS93] SHUB, M., SMALE, S.: Complexity of Bezout’s theorem. III. Condition number and packing. *J. Complex.* **9**, 4–14 (1993). Festschrift for Joseph F. Traub, Part I.
- [SS94] SHUB, M., SMALE, S.: Complexity of Bezout’s theorem. V. Polynomial time. *Theor. Comput. Sci.* **133**, 141–164 (1994). Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).
- [SS96] SHUB, M., SMALE, S.: Complexity of Bezout’s theorem. IV. Probability of success; extensions. *SIAM J. Numer. Anal.* **33**(1), 128–148 (1996).
- [ST05] SIEBERT, B., TIAN, G.: On the holomorphicity of genus two Lefschetz fibrations. *Ann. Math. (2)* **161**(2), 959–1020 (2005).
- [Sma00] SMALE, S.: Mathematical problems for the next century. In: Mathematics: Frontiers and Perspectives, pp. 271–294. Am. Math. Soc., Providence (2000).
- [ST12] SONG, J., TIAN, G.: Canonical measures and Kähler-Ricci flow. *J. Am. Math. Soc.* **25**(2), 303–353 (2012).
- [Sto67] STOLL, W.: The continuity of the fiber integral. *Math. Z.* **95**, 87–138 (1967).
- [Stu061] STURM, K.-T.: On the geometry of metric measure spaces. I. *Acta Math.* **196**(1), 65–131 (2006).
- [Stu062] STURM, K.-T.: On the geometry of metric measure spaces. II. *Acta Math.* **196**(1), 133–177 (2006).
- [Tos10] TOSATTI, V.: Adiabatic limits of Ricci-flat Kähler metrics. *J. Differ. Geom.* **84**(2), 427–453 (2010).
- [Vil03] VILLANI, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics, vol. 58. Am. Math. Soc., Providence (2003).
- [Vil09] VILLANI, C.: Optimal Transport: Old and New. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338. Springer, Berlin (2009).

Paolo Antonini

Dipartimento di Matematica e Fisica “E. De Giorgi”, Università del Salento, Lecce, Italy.

[paolo.antonini@unisalento.it](mailto:paolo.antonini@unisalento.it)

Fabio Cavalletti

Department of Mathematics, University of Milan, Milan, Italy.

[fabio.cavalletti@unimi.it](mailto:fabio.cavalletti@unimi.it)

Antonio Lerario

Mathematics Area, SISSA, Trieste, Italy.

[lerario@sissa.it](mailto:lerario@sissa.it)

Received: 16 October 2023

Accepted: 25 July 2024