Arnold Mathematical Journal (2022) 8:513–516
https://doi.org/10.1007/s40598-022-00207-0

RESEARCH CONTRIBUTION

Maximum Likelihood Degree of Surjective Rational Maps

Ilya Karzhemanov[1]

Received: 6 April 2021 / Revised: 4 March 2022 / Accepted: 2 May 2022 /
Published online: 25 May 2022
© Institute for Mathematical Sciences (IMS), Stony Brook University, NY 2022

Abstract
With any *surjective rational map* $f : \mathbb{P}^n \dashrightarrow \mathbb{P}^n$ of the projective space, we associate a numerical invariant (*ML degree*) and compute it in terms of a naturally defined vector bundle $E_f \longrightarrow \mathbb{P}^n$.

Keywords Surjective rational map · Vector bundle · Chern number

Mathematics Subject Classification 14E05 · 14N10 · 14F10

With a collection of effective divisors $D_0, \ldots, D_m$ in the projective space $\mathbb{P} := \mathbb{CP}^n$ is associated the *maximum likelihood degree* $(-1)^n e_{\mathrm{top}}(\mathbb{P}\backslash D)$, $D := \bigcup_{i=0}^m D_i$. Alternatively, letting $\Omega^1_{\mathbb{P}}(\log D)$ be the *Saito's sheaf*, i.e., the *double dual* of the sheaf of logarithmic differential 1-forms, one computes the ML degree as the top Chern class $c_n(\Omega^1_{\mathbb{P}}(\log D))$ (we refer to [2, 4] for basic properties of ML degree, its connections with algebraic statistics, topology of arrangements, combinatorics, etc.). Note, however, that it is difficult to compute $c_n(\Omega^1_{\mathbb{P}}(\log D))$ in general (when D is not SNC).

In the present note, we study the ML degree under the condition that defining polynomials f_i of D_i, $0 \le i \le m = n$, span the linear system of a *surjective rational map* $f : \mathbb{P} \dashrightarrow \mathbb{P}$ (see [1] and [6] for some aspects of such maps). Our main result (proved along the lines that follow) is the next.

Theorem 1 *In the previous setting, the ML degree $c_n(\Omega^1_{\mathbb{P}}(\log D))$ is equal to the coefficient of z^n in $\frac{(1-z\mathcal{O}_{\mathbb{P}}(1))^{n+1}}{\prod_{i=0}^n(1-z\mathcal{O}_{\mathbb{P}}(D_i'))}$, where $\bigcup_{i=0}^n D_i' =: D_{red}$ is the reduced scheme associated with D (so that $D = D_{red}$ as sets).*

For a vector bundle E over $\mathbb{P}$, given by an affine open cover $\mathbb{P} = \cup_\alpha U_\alpha$ and transition functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \longrightarrow \mathrm{GL}(r, \mathbb{C})$, the *pullback* $f^*(E)$ on $\mathbb{P}\backslash\{\Sigma :=$ base locus of $f\}$ is defined as usual (due to the surjectivity of f), via

✉ Ilya Karzhemanov
karzhemanov.iv@mipt.ru

1 Laboratory of AGHA, Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region 141701, Russia

$f^{-1}(U_\alpha)$ and $f^*(g_{\alpha\beta})$. Note that all $f^{-1}(U_\alpha)$ are affine open in $\mathbb{P}$. Let $\cup_k U_{\alpha,k}$ be an affine open cover of $f^{-1}(U_\alpha)$, such that $\mathbb{P} = \cup_{\alpha,k} U_{\alpha,k}$. Then, since codim $\Sigma > 1$ and $f^*(g_{\alpha\beta})$ are *algebraic*, every $f^*(g_{\alpha\beta})$ extends through $U_{\alpha,k} \cap U_{\beta,m} \cap \Sigma$ to each $U_{\alpha,k} \cap U_{\beta,m}$. Furthermore, the 1-cocyle property of $f^*(g_{\alpha\beta})$ (considered on $(\cup_k U_{\alpha,k}) \cap (\cup_m U_{\beta,m}) \supseteq f^{-1}(U_\alpha) \cap f^{-1}(U_\beta)$) is preserved and one gets a vector bundle, over $\mathbb{P}$, which we denote again by $f^*(E)$.

Furthermore, let $x_0, \dots, x_n$ be projective coordinates on $\mathbb{P}$, such that $f^*(x_i) = f_i$. Denote by H the union of coordinate hyperplanes $H_i := (x_i = 0) \subset \mathbb{P}$. There is an exact sequence

$$0 \longrightarrow \Omega^1_{\mathbb{P}} \xrightarrow{\psi_H} \Omega^1_{\mathbb{P}}(\log H) \xrightarrow{\varphi_H} \bigoplus_{i=0}^{n} \mathcal{O}_{H_i} \longrightarrow 0 \tag{2}$$

(see, e.g., [2, Lemma 2]). We have $f^*(\mathcal{O}_{\mathbb{P}}) = \mathcal{O}_{\mathbb{P}}$ and $f^*(\mathcal{O}_{\mathbb{P}}(H)) = \mathcal{O}_{\mathbb{P}}(D)$ by construction. Then, (2) pulls back to an exact sequence

$$f^*(\Omega^1_{\mathbb{P}}) \xrightarrow{\psi_D} f^*(\Omega^1_{\mathbb{P}}(\log H)) \xrightarrow{\varphi_D} f^*\left(\bigoplus_{i=0}^{n} \mathcal{O}_{H_i}\right) = \bigoplus_{i=0}^{n} \mathcal{O}_{D_i}. \tag{3}$$

Note, however, that the morphism $\psi_D := f^*(\psi_H)$ (resp. $\varphi_D := f^*(\varphi_H)$) need not be injective (resp. surjective)—see below.

Lemma 4 $f^*(\Omega^1_{\mathbb{P}}(\log H)) = \Omega^1_{\mathbb{P}}(\log D)$.

Proof The bundle $\Omega^1_{\mathbb{P}}(\log H)$ (resp. $\Omega^1_{\mathbb{P}}(\log D)$) is trivial over an affine open set not containing H (resp. D). Hence, as $f^*(\mathcal{O}_{\mathbb{P}}) = \mathcal{O}_{\mathbb{P}}$, it suffices to restrict to an affine open $U \subset \mathbb{P}^n$ (resp. $f^{-1}(U)$), such that $U \cap H \neq \emptyset$ (we may also assume that $x_0 \neq 0$ on U). Then, $\Omega^1_{\mathbb{P}}(\log H)\big|_U$ is generated by the local sections $\sum_{i=1}^n c_i \log x_i$, $c_i \in \mathbb{C}$, whereas $f^*(\Omega^1_{\mathbb{P}}(\log H))\big|_{f^{-1}(U)}$ is generated by $\sum_{i=1}^n c_i \log f^*(x_i)$ (as usual we take double duals when needed). This yields $f^*(\Omega^1_{\mathbb{P}}(\log H))\big|_{f^{-1}(U)} = \Omega^1_{\mathbb{P}}(\log D)\big|_{f^{-1}(U)}$ and the result follows. □

Before finding $f^*(\Omega^1_{\mathbb{P}})$, we need an auxiliary construction. Namely, put $d_f :=$ deg f_i and consider the subspace $V \subset H^0(\mathbb{P}, \mathcal{O}_{\mathbb{P}}(d_f))$ spanned by $f_0, \dots, f_n$. Recall that by Kodaira's construction of rational maps via linear systems, every point $p \in f(\mathbb{P}\backslash\Sigma)$ is represented by hyperplane $H_p \subset V$, which consists of all polynomials from V vanishing at $f^{-1}(p)$. Then, since f is surjective, this defines a *vector bundle* $E_f \longrightarrow \mathbb{P} = f(\mathbb{P}\backslash\Sigma)$, with fibers $E_{f,p} = H_p$ for all p, and an exact sequence

$$0 \longrightarrow \mathcal{L} \longrightarrow \mathbb{C}^{n+1} \longrightarrow E_f \longrightarrow 0 \tag{5}$$

for some line bundle $\mathcal{L}$. It is easy to prove (by induction on n) that $\mathcal{L} = \mathcal{O}_{\mathbb{P}}(-n-1)$. This implies that both E_f and $f^*(E_f)$ are generated by global sections.

We now prove the following ("Hurwitz-type"):

Lemma 6 $f^*(\Omega^1_{\mathbb{P}}) \subseteq \Omega^1_{\mathbb{P}} \otimes_{\mathcal{O}_{\mathbb{P}}} \mathcal{O}_{\mathbb{P}}(-d_f + 1)$.

Proof Each global section of $f^*(E_f)$ is given by some choice of a basis ($= \{x_0, \ldots, x_n\}$) in $\mathbb{C}^{n+1}$ and a way every $p \in \mathbb{P}$ (identified with $\sum p_i x_i$ for $p_i \in \mathbb{C}$) is represented by a point in $V \simeq \mathbb{C}^{n+1} = H^0(\mathbb{P}, E_f)$. This yields a *surjection*

$$\mathcal{H}om_{\mathcal{O}_{\mathbb{P}}}(\mathcal{O}_{\mathbb{P}}(1), E_f \otimes_{\mathcal{O}_{\mathbb{P}}} \mathcal{O}_{\mathbb{P}}(d_f)) = E_f \otimes_{\mathcal{O}_{\mathbb{P}}} \mathcal{O}_{\mathbb{P}}(d_f - 1) \twoheadrightarrow f^*(E_f)$$

of vector bundles generated by global sections.

Now, observe that $E_f \simeq T_{\mathbb{P}}$ (= the dual of $\Omega^1_{\mathbb{P}}$) by (5) and [5, Theorem 3.1]. Hence, $f^*(\Omega^1_{\mathbb{P}})$ embeds into $\Omega^1_{\mathbb{P}} \otimes_{\mathcal{O}_{\mathbb{P}}} \mathcal{O}_{\mathbb{P}}(-d_f + 1)$ by duality. ◻

Note that $\Omega^1_{\mathbb{P}}(\log D) = \Omega^1_{\mathbb{P}}(\log D_{\text{red}})$ (cf. the proof of Lemma 4). Hence, $\varphi_D(\Omega^1_{\mathbb{P}}(\log D)) = \bigoplus_{i=0}^n \mathcal{O}_{D'_i}$. Further, it follows from (3) and Lemma 6 that the kernel of φ_D is a subsheaf of $\psi_D(\Omega^1_{\mathbb{P}} \otimes \mathcal{O}_{\mathbb{P}}(-d_f + 1))$, whose general local section is easily seen (by restricting on $\mathbb{P}\backslash\Sigma$) to coincide with a holomorphic 1-form, which vanishes *at most* on D_{red}. One actually finds that this is a *subbundle* of $\Omega^1_{\mathbb{P}}$ generated by all such 1-forms. Thus, we get $\operatorname{Ker} \varphi_D = \Omega^1_{\mathbb{P}}$ and an exact sequence

$$0 \longrightarrow \Omega^1_{\mathbb{P}} \longrightarrow \Omega^1_{\mathbb{P}}(\log D) \longrightarrow \bigoplus_{i=0}^{n} \mathcal{O}_{D'_i} \longrightarrow 0.$$

Taking the total Chern class of the latter concludes the proof of Theorem 1.

Remark 7 We summarize that any f defines, *canonically*, a fiberwise non-degenerate element $e \in \operatorname{Hom}(\mathbb{P}; T_{\mathbb{P}}, T_{\mathbb{P}} \otimes \mathcal{O}_{\mathbb{P}}(d_f - 1))$. This can also be seen as follows. Namely, the embedding $\mathcal{L} \subset \mathbb{C}^{n+1}$ in (5) is given by some global sections $s_0, \ldots, s_n \in H^0(\mathbb{P}, \mathcal{O}_{\mathbb{P}}(n+1))$, so that $x_i \mapsto s_i$, $0 \le i \le n$, defines a *regular surjective* self-map of $\mathbb{P}$. This yields a family (a "field") $\{H_p\}$ of hyperlines on $\mathbb{P} \ni p$. After choosing e, one gets another family $\{H'_p\}$, where $H'_p \simeq H_p$ are spaces of forms of degree d_f and $\bigcup H'_p = V$. Identify H'_p with the set of corresponding hypersurfaces that vanish at p. The map f is now obtained by sending each $p \in H'_p$ to H_p (it is defined exactly on $\mathbb{P}\backslash \bigcap H'_p$). One thus obtains a description of the moduli spaces of surjective maps f. It would be interesting to relate this picture with [3], where the moduli of degree k rational self-maps of $\mathbb{P}^1$ were interpreted as the moduli of (pairs of) *monopoles*, having magnetic charge k.

Example 8 The need for D_{red} in Theorem 1 is justified by the *Frobenius map* f, given by $f_i := x_i^{d_f}$, $0 \le i \le n$; ML degree of f equals $(-1)^n e_{\text{top}}((\mathbb{C}^*)^n) = \mathbf{0}$ in this case. Furthermore, one computes the ML degree of f in [6, Example1.6] to be $\mathbf{9}$, which can also be seen directly from [2, Corollary 6] (here, the divisors D_i satisfy the *GNC condition*). Indeed, in this case, D_i are reduced and $\deg f_i = 2$ for all i, so that the expression with Chern classes from Theorem 1 becomes

$$\begin{aligned}\frac{(1 - z\mathcal{O}_{\mathbb{P}}(1))^3}{(1 - z\mathcal{O}_{\mathbb{P}}(2))^3} &= (1 - z\mathcal{O}_{\mathbb{P}}(1))^3(1 + z\mathcal{O}_{\mathbb{P}}(2) + 4z^2)^3 \\ &= (1 + z\mathcal{O}_{\mathbb{P}}(1) + 2z^2)^3 = 1 + 3(z\mathcal{O}_{\mathbb{P}}(1) + 2z^2) \\ &\quad +3(z\mathcal{O}_{\mathbb{P}}(1) + 2z^2)^2 = 1 + z\mathcal{O}_{\mathbb{P}}(3) + \mathbf{9}z^2.\end{aligned}$$

Acknowledgements I am grateful to anonymous referee for helpful comments and corrections.

References

1. Arutyunov, A.V.: Smooth abnormal problems in extremum theory and analysis. Russian Math. Surveys. **67**(3), 403–457 (2012)
2. Catanese, F., Hosten, S., Khetan, A., Sturmfels, B.: The maximum likelihood degree. Amer. J. Math. **128**(3), 671–697 (2006)
3. Donaldson, S.: Nahm's equations and the classification of monopoles. Commun. Math. Phys. **96**(3), 387–407 (1984)
4. Huh, J.: The maximum likelihood degree of a very affine variety. Compos. Math. **149**(8), 1245–1266 (2013)
5. Hwang, J.-M.: Deformations of the tangent bundle and their restriction to standard rational curves. Acta Math. Vietnam. **35**(1), 71–77 (2010)
6. Karzhemanov, I., Zhdanovskiy, I.: Some properties of surjective rational maps. Eur. J. Math. **4**(1), 326–329 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.