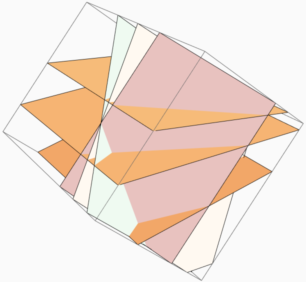
Logarithmic Discriminants of Hyperplane Arrangements

Positive Geometry Seminar

MAX PLANCK INSTITUTE
FOR MATHEMATICS
IN THE SCIENCES

Leonie Kayser, Andreas Kretschmer & Simon Telen
leokayser.github.io

October 22, 2024

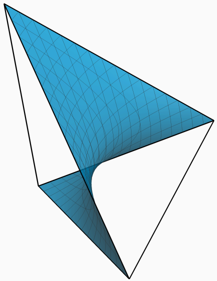


Maximum Likelihood Estimation in Algebraic Statistics

- Let $X \hookrightarrow (\mathbb{C}^\times)^{n+1}$ be a d-dimensional smooth variety
- Discrete statistical model $X \cap \Delta_n = \{\, p \in X \cap \mathbb{R}^{n+1} \mid p_i > 0,\ p_0 + \cdots + p_n = 1 \,\}$
- Given data points $u \in \mathbb{N}^{n+1}$, which parameter maximizes the log-likelihood

$$\mathcal{L}_u(x) = \log x_0^{u_0} \cdots x_n^{u_n}, \qquad x \in X \cap \Delta_n?$$

- Critical equations: $x \in \mathrm{Crit}_X(u) = \{\, x \in X \mid \nabla \mathcal{L}_u(x) = 0 \,\}$
- $\mathrm{Crit}_X(u)$ is a finite set of $\mathrm{MLdeg}(X)$ non-degenerate critical points for general data $u \in \mathbb{N}^{n+1}$ (or $\mathbb{C}^{n+1}$)
- [Huh13] $\mathrm{MLdeg}(X) = (-1)^d \cdot \chi(X)$
- Extensively studied for toric models (exponential families), linear models, determinantal varieties, . . .



Linear models and scattering amplitudes

- Let be an essential arrangement of $n+1$ hyperplanes in $\mathbb{C}^d$

$$\mathcal{A} = \mathrm{V}(\ell_0 \cdots \ell_n) \subseteq \mathbb{C}^d, \qquad (\ell_0(x), \dots, \ell_n(x))^{\mathsf{T}} = Ax + b, \quad L^{\mathsf{T}} = [b \mid A]$$

- Parametrizes linear model $X := \mathbb{C}^d \setminus \mathcal{A} \xhookrightarrow{\ell} (\mathbb{C}^\times)^{n+1}$, can assume $\sum_j \ell_j = 1$
- Log-likelihood function or master function given by

$$\mathcal{L}_u(x) = u_0 \log \ell_0(x) + \cdots + u_n \log \ell_n(x)$$

- If the ℓ_j are real, then $\mathrm{MLdeg}(X)$ is the number of bounded chambers of $\mathcal{A} \cap \mathbb{R}^d$
- Critical equations appear as scattering equations in bi-adjoint scalar ϕ^3-theories (Cachazo, He & Yuan [CHY14])

What is "general" data?

- Moving from general to special $u \in \mathbb{P}^n = \mathbb{P}(\mathbb{C}^{n+1})$, what can happen to $\mathrm{Crit}_X(u)$?
  1. Two critical points collide to form a non-reduced/degenerate point
  2. A positive-dimensional component appears
  3. A critical point disappears to infinity
- The closure of 1.-3. was called the *data discriminant* in [RT15]
- 3. was studied from a tropical and a Bernstein–Sato perspective by (Sattelberger & van der Veer [SvdV23])

Definition

The *logarithmic discriminant* of a (smooth) variety $X \hookrightarrow (\mathbb{C}^\times)^{n+1}$ is

$$\nabla_{\log}(X) := \overline{\{\, u \in \mathbb{P}^n \mid \mathrm{Crit}_X(u) \text{ is infinite or non-reduced} \,\}}.$$

$\rightsquigarrow$ Goal: Understand logarithmic discriminants of hyperplane arrangements!

Three points enter a bar

- Three points on a line $\mathcal{A} = \mathrm{V}(x\,(x+1)\,(x+b)) = \{0,-1,-b\} \subseteq \mathbb{C}^1$ $(b \notin \{0,1\})$
- Model is a line $X \subseteq (\mathbb{C}^\times)^3$ parametrized by $(x, x+1, x+b)$,

$$\mathcal{L}_u(x) = u_0 \log x + u_1 \log(x+1) + u_2 \log(x+b)$$

- A single critical equation in $x \in \mathbb{C}^1 \setminus \mathcal{A}$

$$\frac{u_0}{x} + \frac{u_1}{x+1} + \frac{u_2}{x+b} = 0 \quad \Longleftrightarrow \quad u_0(x+1)(x+b) + u_1 x(x+b) + u_2 x(x+1) = 0$$

- When does this quadric have a double root in x? Highschool discriminant!

$$\Delta_{\log}(X) = (b-1)^2\, u_0^2 + 2b(b-1)\, u_0 u_1 + b^2\, u_1^2 - 2(b-1)\, u_0 u_2 + 2b\, u_1 u_2 + u_2^2$$

- $\Delta_{\log}(X)$ itself is a smooth quadric in u with discriminant $-4b^2(b-1)^2$

Ramification and its consequences

- Let $f\colon V \to W$ be a dominant map of smooth irreducible varieties of dimension n
- The ramification locus $\mathrm{Ram}(f) \subseteq V$ is the hypersurface

$$\mathrm{Ram}(f) = \left\{ x \in V \;\middle|\; x \in f^{-1}(f(x)) \text{ is not isolated or reduced} \right\} = \mathrm{V}(\det J_f(x))$$

- The branch locus is the scheme-theoretic image $\mathrm{Branch}(f) = \overline{f(\mathrm{Ram}(f))} \subseteq W$
- Apply this to the *likelihood correspondence*

$$f\colon \mathcal{L}_X^\circ = \{ (u,x) \in \mathbb{P}^n \times X \mid \nabla \mathcal{L}_u(x) = 0 \} \to \mathbb{P}^n$$

Definition (Scheme-theoretic definition of $\nabla_{\log}(X)$)

The logarithmic discriminant is the branch locus of the projection f. The ramification locus is defined in $\mathbb{P}^n \times X$ by

$$\nabla \mathcal{L}_u(x) = 0, \qquad \det \mathrm{Hess}_x(\mathcal{L}_u(x)) = 0.$$

Irreducibility of $\mathrm{Ram}(f)$

- $X = \mathbb{C}^d \setminus \mathrm{V}(\ell_0 \cdots \ell_n)$, $(\ell_0(x), \ldots, \ell_n(x))^\mathsf{T} = Ax + b$
- Here the equations of the ramification locus have a very concrete form

$$\nabla \mathcal{L}_u(x) = A^\mathsf{T} \cdot \mathrm{diag}(1/\ell_0, \ldots, 1/\ell_n) \cdot u = 0$$

$$h = \det\left(A^\mathsf{T} \cdot \mathrm{diag}\left(\frac{u_0}{\ell_0^2}, \ldots, \frac{u_n}{\ell_n^2}\right) \cdot A\right) = \sum_{\substack{I \subseteq \{0,\ldots,n\} \\ |I|=d}} |A_I|^2 \frac{u^I}{(\ell^I)^2}$$

- Critical equations are linear in the u_j $\rightsquigarrow$ substitute them in h to obtain

$$\tilde{h} \in \mathbb{C}[u_d, \ldots, u_n; x], \qquad \mathrm{Ram}(f) \cong \mathrm{V}(\tilde{h}) \subseteq \mathbb{P}^{n-d} \times X$$

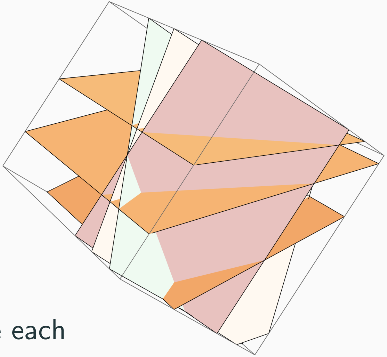
Theorem

If the arrangement contains a subset of $d+2$ hyperplanes which is bi-uniform (to be defined in a moment), then $\mathrm{Ram}(f)$ and hence $\nabla_{\log}(X)$ are irreducible varieties.

A split discriminant!

- Consider the arrangement $\mathcal{A}$ of six planes

$$(1, x_1, x_2, x_3) \cdot \begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 2 & 1 & 0 & 1 \\ 1 & \frac{3}{2} & \frac{3}{2} & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix}$$



- The first and the last three planes intersect in a line each
- The logarithmic discriminant decomposes as

$$\begin{aligned} \nabla_{\log} = {} & \mathrm{V}(144u_0^2 + 120u_0u_1 + 168u_0u_2 + 25u_1^2 - 70u_1u_2 + 49u_2^2) \\ & \cup \mathrm{V}(u_3^2 - 2u_3u_4 + 4u_3u_5 + u_4^2 + 4u_4u_5 + 4u_5^2) \\ & \cup \mathrm{V}(u_0 + u_1 + u_2, u_3 + u_4 + u_5). \end{aligned}$$

A complete answer in $\mathbb{C}^1$

Theorem

Let $\mathcal{A} \subseteq \mathbb{C}^1$ be an arrangement of $n+1 \geq 3$ distinct points.

1. *The ramification locus is a smooth irreducible hypersurface in $\mathbb{P}^n \times (\mathbb{C}^1 \setminus \mathcal{A})$.*
2. *The class of $\overline{\mathrm{Ram}(f)}$ in the Chow ring $\mathrm{A}^\bullet(\mathbb{P}^n \times \mathbb{P}^1) = \mathbb{Z}[\alpha, \beta]/\langle \alpha^{n+1}, \beta^2 \rangle$ is $\alpha^2 + 2(n-1)\alpha\beta$.*
3. *The projection $f\colon \mathrm{Ram}(f) \to \nabla_{\log}$ is birational.*
4. *$\nabla_{\log}(X) \subseteq \mathbb{P}^n$ is an irreducible reduced hypersurface of degree $2(n-1)$.*

- Explicit formula

$$\Delta_{\log} = \mathrm{Disc}_x \left(\sum_{i=0}^{n} u_i \prod_{k \neq i} (x + b_k) \right)$$

- For $n+1 = 4$ points $\nabla_{\log} \subseteq \mathbb{P}^3$ is always a singular quartic surface

A positivity result

- Let $H_i := \overline{\mathrm{V}(\ell_i)} \subseteq \mathbb{P}^n$ be the closures of the affine hyperplanes
- *Flats* of the matroid $\mathrm{M}(\mathcal{A})$ are the linear spaces obtained as intersections of subsets of the H_i
- *$\mathcal{A}$ has no flats at infinity* if no non-empty flats are contained in $\mathbb{P}^d \setminus \mathbb{C}^d$.

Theorem

If $\mathcal{A}$ has no flats at infinity and if $u \in (\mathbb{C}^\times)^{n+1}$ is such that $\mathrm{Crit}_X(u)$ consists of $\mathrm{MLdeg}(X)$ reduced points, then $u \notin \nabla_{\log}(X)$.

Corollary (An application of Varchenko's theorem)

Let $\mathcal{A} \subseteq \mathbb{C}^d$ be a real arrangement. If $\mathcal{A}$ has no flats at infinity, then $\nabla_{\log} \cap \mathbb{R}_+^{n+1} = \emptyset$.

A link to reciprocal linear spaces

- The critical equations can be rearranged as

$$x \in \mathrm{Crit}_X(u) \quad \text{if and only if} \quad (1/\ell_0(x), \ldots, 1/\ell_n(x))^{\mathsf{T}} \in \mathrm{Ker}(A^{\mathsf{T}} \operatorname{diag}(u_0, \ldots, u_n)).$$

- Let $\mathcal{R}_L \subseteq \mathbb{P}^n$ be the image closure of the locally closed embedding

$$\gamma \colon \mathbb{C}^d \setminus \mathcal{A} \to \mathbb{P}^n, \qquad (x_1, \ldots, x_d) \mapsto (\ell_0(x)^{-1} : \cdots : \ell_n(x)^{-1})$$

- Considering the kernel as a point in the Grassmannian $\mathbb{G}(n-d, \mathbb{P}^n)$

$$\varphi \colon \mathbb{P}^n \setminus \mathrm{V}(u_0 \cdots u_n) \to \mathbb{G}(n-d, \mathbb{P}^n), \qquad u \mapsto [\mathrm{Ker}(A^{\mathsf{T}} \operatorname{diag}(u_0, \ldots, u_n))]$$

- Then the critical equations become (a subset of) a linear section of $\mathcal{R}_L$

$$x \in \mathrm{Crit}_X(u) \qquad \text{if and only if} \qquad \gamma(x) \in \varphi(u) \cap \mathcal{R}_L \subseteq \mathbb{P}^n$$

Hurwitz discriminant

- Consider the incidence

$$\mathcal{I}^\circ = \{ \, (\Lambda, y) \mid y \in \Lambda \cap \mathrm{Im}(\gamma) \, \} \subseteq \mathbb{G}(n-d, \mathbb{P}^n) \times \mathcal{R}_L$$

- The branch locus $\mathcal{I}^\circ \to \mathbb{G}(n-d, \mathbb{P}^n)$ is the first associacted hypersurface $\mathcal{Z}_1(\mathcal{R}_L)$
- Its defining equation in the Plücker ring is the *Hurwitz form* $\mathrm{Hu}_{\mathcal{R}_L}$
- If $\mathcal{A}$ is an uniform arrangement, then $\deg \mathrm{Hu}_{\mathcal{R}_L} = 2(n-d)\binom{n}{d-1}$

Definition

The pullback along $\varphi\colon \mathbb{P}^n \dashrightarrow \mathbb{G}(n-d, \mathbb{P}^n)$ is the *Hurwitz discriminant*

$$\nabla_{\mathrm{Hu}}(\mathcal{A}) = \varphi^{-1}(\mathcal{Z}_1(\mathcal{R}_L)).$$

- Always have $\nabla_{\log}(\mathcal{A}) \subseteq \nabla_{\mathrm{Hu}}(\mathcal{A})$, equality does not necessarily hold

General arrangements

- $\mathcal{A}$ defined by $(\ell_0(x), \ldots, \ell_n(x))^\mathsf{T} = Ax + b$
- $k \times n$ matrix is *uniform* if all sets of k columns are linearly independent
- *Little matroid* $\mathrm{M}(A^\mathsf{T})$, *big matroid* $\mathrm{M}(\,[b \mid A]^\mathsf{T}\,)$, bi-uniform = both are uniform

Theorem

Let $\mathcal{A}$ be a bi-uniform arrangement of $n+1 \geq d+2$ hyperplanes in $\mathbb{C}^d$.

1. $\nabla_{\mathrm{Hu}}(\mathcal{A})$ *is a hypersurface of degree* $2d\binom{n-1}{d}$ *with full Newton polytope*
2. $\nabla_{\log}(\mathcal{A})$ *is an irreducible and reduced hypersurface.*
3. $\nabla_{\log} \subseteq \nabla_{\mathrm{Hu}}$ *coincide as sets, so* $\nabla_{\mathrm{Hu}} = \mathrm{V}((\Delta_{\log})^e)$ *for some* $e \geq 1$.
4. *If the arrangement is defined by real affine linear forms, then* $\nabla_{\log} \cap \mathbb{R}_+^{n+1} = \emptyset$.

- We know that equality hold for $d = 1$ and expect this to always hold true

The discriminant of $\mathcal{M}_{0,m}$

- $\mathcal{M}_{0,m}$ parametrizes tuples of m points on the projective line $\mathbb{P}^1$
- By fixing $(0, 1, x_1, \ldots, x_{m-3}, \infty)$, it can be realized as the complement in $\mathbb{C}^{m-3}$ of the $n = \binom{m-1}{2} - 1$ minors of

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 0 \\ 0 & 1 & x_1 & x_2 & \cdots & x_{m-3} & 1 \end{pmatrix}$$

- Variable corresponding to minor (i, j) are *Mandelstam invariants* s_{ij}
- Discriminant for $m = 5$ has degree $4 < 2 \cdot 2 \cdot \binom{5-2}{2} = 12$

$$\Delta_{\log}(\mathcal{M}_{0,5}) = (s_{13}s_{24} + s_{13}s_{34} + s_{14}s_{34} + s_{14}s_{23} + s_{23}s_{34} + s_{24}s_{34} + s_{34}^2)^2 - 4s_{13}s_{14}s_{23}s_{24}$$

- The Hurwitz discriminant has the extra factors

$$\Delta_{\mathrm{Hu}}(\mathcal{M}_{0,5}) = (s_{13} + s_{23} + s_{34})^2 \cdot (s_{14} + s_{24} + s_{34})^2 \cdot \Delta_{\log}(\mathcal{M}_{0,5})$$

- Conjecturally rich nested structure, degrees of $\nabla_{\log}(\mathcal{M}_{0,m})$ are $4, 30, 208, 1540, \ldots$

Many open questions

- Missing piece of the puzzle: Is ∇_{Hu} reduced for a bi-uniform arrangement?
- (When) is the projection $\mathrm{Ram}(f) \to \nabla_{\log}$ generically finite? When is it birational?
- Is there any arrangement such that $\nabla_{\log}(\mathcal{A})$ is *not* reduced?
- Is there an arrangement of lines whose logarithmic discriminant is reducible?
- Is the degree of $\nabla_{\log}(\mathcal{A})$ an invariant of the little and big matroid?
- What is the meaning of the components $\nabla_{\mathrm{Hu}} \setminus \nabla_{\log}$?
- Can the assumption "no flats at infinity" be dropped from the positivity result?
- Is there a closed expression for the degree of $\nabla_{\log}(\mathcal{M}_{0,m})$?

Beyond hyperplane arrangements

- Let $f_0, \dots, f_n \in \mathbb{C}[x]$ be polynomials parametrizing a model

$$X \cong \mathbb{C}^d \setminus \mathrm{V}(f_0 \cdots f_n) \hookrightarrow (\mathbb{C}^\times)^{n+1}, \qquad x \mapsto (f_0(x), \dots, f_n(x))$$

- Case $(f, x_1, \dots, x_d)$ closely related to toric models
- $d = 1$: $\nabla_{\log}(X)$ is an irreducible hypersurface of degree $2(\# \mathrm{V}(f_0 \cdots f_n) - 2)$
- Consider a family X_z of conics in $\mathbb{C}^2$ given by polynomials in $\mathbb{C}[z][x_1, x_2]$

  `https://www.geogebra.org/calculator/rjxgakbv`

$$f_0 = (x_1 + x_2 + 1)(-x_1 + x_2 - 2) + z, \quad f_1 = x_1, \quad f_2 = x_2$$

- X_0 is a bi-uniform arrangement of 4 lines, hence $\deg \nabla_{\log}(X_0) = 2 \cdot 2 \cdot \binom{4-2}{2} = 4$

$$\begin{aligned}\Delta_{\log}(X_0) = {} & 36\, u_0^4 + 44\, u_0^3 u_1 + 21\, u_0^2 u_1^2 + 6\, u_0 u_1^3 + u_1^4 + 684\, u_0^3 u_2 + 198\, u_0^2 u_1 u_2 \\ & + 90\, u_0 u_1^2 u_2 + 981\, u_0^2 u_2^2 + 90\, u_0 u_1 u_2^2 - 18\, u_1^2 u_2^2 + 486\, u_0 u_2^3 + 81\, u_2^4\end{aligned}$$

Men who stare at ~~goats~~ sextics

$$
\begin{aligned}
\Delta_{\log}(X_z) = {} & 324\, u_0^6 + (576\, z + 396) u_0^5 u_1 + (256\, z^2 + 928\, z + 189) u_0^4 u_1^2 + (512\, z^2 + 560\, z + 54) u_0^3 u_1^3 \\
& + (384\, z^2 + 168\, z + 9) u_0^2 u_1^4 + (128\, z^2 + 32\, z) u_0 u_1^5 + (16\, z^2 + 4\, z) u_1^6 + (-5184\, z + 6156) u_0^5 u_2 \\
& + (-9\,216\, z^2 + 6\,912\, z + 1\,782) u_0^4 u_1 u_2 + (-4096\, z^3 - 5632\, z^2 + 5760\, z + 810) u_0^3 u_1^2 u_2 \\
& + (-6144\, z^3 + 768\, z^2 + 1224\, z) u_0^2 u_1^3 u_2 + (-3072\, z^3 + 384\, z^2 + 360\, z) u_0 u_1^4 u_2 + (-512\, z^3 - 128\, z^2) u_1^5 u_2 \\
& + (20736\, z^2 - 54432\, z + 8829) u_0^4 u_2^2 + (36864\, z^3 - 87552\, z^2 - 5760\, z + 810) u_0^3 u_1 u_2^2 \\
& + (16384\, z^4 - 16384\, z^3 - 55808\, z^2 - 2304\, z - 162) u_0^2 u_1^2 u_2^2 + (16384\, z^4 - 34816\, z^3 - 12032\, z^2 + 72\, z) u_0 u_1^3 u_2^2 \\
& + (4096\, z^4 - 8704\, z^3 - 3008\, z^2 - 108\, z) u_1^4 u_2^2 + (41472\, z^2 - 97200\, z + 4374) {u_0}^3 u_2^3 \\
& + (55296\, z^3 - 122112\, z^2 - 11016\, z) u_0^2 u_1 u_2^3 + (16384\, z^4 - 14336\, z^3 - 52992\, z^2 - 648\, z) u_0 u_1^2 u_2^3 \\
& + (8192\, z^4 - 16384\, z^3 - 5472\, z^2) u_1^3 u_2^3 + (31104\, z^2 - 70632\, z + 729) u_0^2 u_2^4 \\
& + (27648\, z^3 - 61056\, z^2 - 3240\, z) u_0 u_1 u_2^4 + (4096\, z^4 - 3584\, z^3 - 13248\, z^2 + 972\, z) u_1^2 u_2^4 \\
& + (10368\, z^2 - 23328\, z) u_0 u_2^5 + (4608\, z^3 - 10368\, z^2) u_1 u_2^5 + (1296\, z^2 - 2916\, z) u_2^6
\end{aligned}
$$

$$\boxed{\Delta_{\log}(X_z)|_{z=0} = \Delta_{\log}(X_0) \cdot u_0^2}$$

Thank you! Questions?

arXiv:2410.11675

References i

Freddy Cachazo, Song He, and Ellis Ye Yuan.
Scattering equations and Kawai-Lewellen-Tye orthogonality.
Physical Review D, 90(6):065001, 2014.

June Huh.
The maximum likelihood degree of a very affine variety.
Compositio Mathematica, 149(8):1245–1266, 2013.

Leonie Kayser, Andreas Kretschmer, and Simon Telen.
Logarithmic discriminants of hyperplane arrangements, 2024.

Jose Israel Rodriguez and Xiaoxian Tang.
Data-discriminants of likelihood equations.
In *Proceedings of the 2015 ACM on international symposium on symbolic and algebraic computation*, pages 307–314, 2015.

Anna-Laura Sattelberger and Robin van der Veer.
Maximum likelihood estimation from a tropical and a Bernstein–Sato perspective.
International Mathematics Research Notices, 2023(6):5263–5292, 2023.