MARTÍN SOMBRA Talk at FoCM 2005, Santander

On the size of the solutions of sparse polynomial systems

Joint work with Patrice Philippon (Paris)

1. Root counting over the torus $(\mathbb{C}^\times)^n$

Let

$$f = (t-1)+(t-1)^2\, x - t\, x^2, \quad g = (t-1)+2\,(t-1)^2\, x - 2\, t\, x^2 \quad \in \mathbb{C}[t,x]$$

Solve $f = g = 0$:

$$g - f = (t-1) - (t-1)^2\, x \Longrightarrow t - 1 = 0 \text{ or } x = \frac{1}{t-1}$$

If $t-1=0$ then $f = x^2 \neq 0$. Otherwise

$$f\left(t, \frac{1}{t-1}\right) = (t-1)+(t-1)^2\,\frac{1}{t-1} - t\,\left(\frac{1}{t-1}\right)^2 = (t-1)^{-2}\,(2\,(t-1)^3 + t)$$

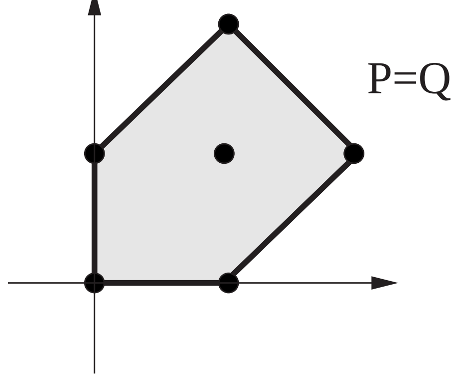
hence $\mathrm{Card}(f = g = 0) = 3$

By the *Bernstein-Kushnirenko thm*

$$\mathrm{Card}(f=g=0) \leq \mathrm{MV}(P,Q) = \mathrm{Vol}(P+Q) - \mathrm{Vol}(P) - \mathrm{Vol}(Q)$$

where $P := \mathrm{Conv}((i,j) : \alpha_{i,j} \neq 0), Q := \mathrm{Conv}((i,j) : \beta_{i,j} \neq 0) \subset \mathbb{R}^2$ are the *Newton polytopes* of f and g resp.; generically we have $=$ instead of $\leq$.

In this case $\mathrm{MV}(P,Q) = 5$



Take any $f, g \in \mathbb{C}[t^{\pm 1}, x^{\pm 1}]$ and write them as

$$f = \sum_{j \in \mathbb{Z}} \alpha_j(t)\, x^j \quad , \quad g = \sum_{j \in \mathbb{Z}} \beta_j(t)\, x^j$$

with $\alpha_i(t), \beta_j(t) \in \mathbb{C}[t^{\pm 1}]$, assume f, g *primitive*

Set

$$Q_0 := \mathrm{Conv}(j : \alpha_j \neq 0) \quad , \quad Q_1 := \mathrm{Conv}(j : \beta_j \neq 0) \quad \subset \mathbb{R}$$

and for $v \in \mathbb{P}^1$ consider the *v-adic Newton polytopes*

$$\begin{aligned} Q_{0,v} &:= \mathrm{Conv}((j, -\mathrm{ord}_v(\alpha_j)) : \alpha_j \neq 0) \quad \subset \mathbb{R}^2 \\ Q_{,v} &:= \mathrm{Conv}((j, -\mathrm{ord}_v(\beta_j)) : \beta_j \neq 0) \quad \subset \mathbb{R}^2 \end{aligned}$$

with $\mathrm{ord}_v(\alpha_j)$ the *order of vanishing* of α_j at v (note that $-\mathrm{ord}_{(0:1)}(\alpha_j) = \deg(\alpha_j)$) and $\vartheta_i : Q_i \to \mathbb{R}$ parametrization of the upper envelope of $Q_{i,v}$ $(i = 0, 1)$.

Let $Z(f, g) := \{\xi \in (\mathbb{C}^\times)^2 : f(\xi) = g(\xi) = 0\} \subset (\mathbb{C}^\times)^2$ denote the *solution set* and Z_0 the geometrically isolated points of Z

Thm 1. Suppose $\vartheta_0 = \vartheta_1$ (and in particular $Q_0 = Q_1$), then

$$\mathrm{Card}(Z(f, g)_0) \leq 2! \sum_{v \in \mathbb{P}^1} \int_{Q_0} \vartheta_0(x)\, dx$$

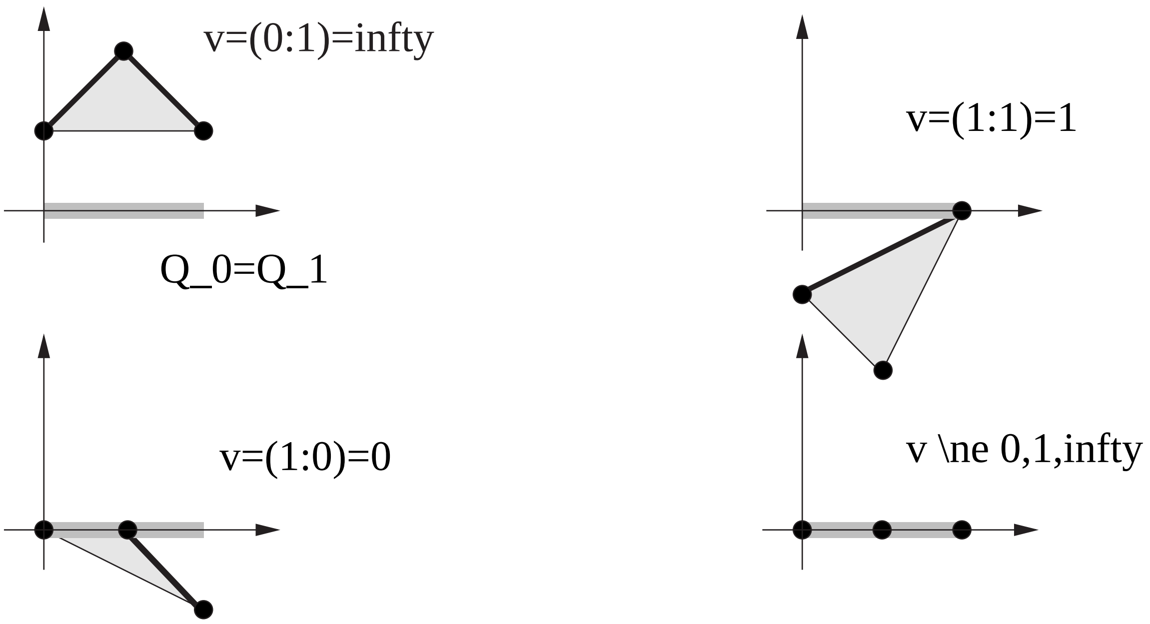
Multiplying $\alpha_j(t)$ and $\beta(t)$ by $\lambda_j, \lambda'_j \in \mathbb{C}^\times$ resp., we have $=$ for generic λ_j, λ'_j

Back to the example

$$f = (t-1)+(t-1)^2\,x-t\,x^2, \quad g = (t-1)+2\,(t-1)^2\,x-2\,t\,x^2 \quad \in \mathbb{C}[t,x]$$

Thm 1 gives the estimate

$$3 = \mathrm{Card}(Z(f,g)_0) \leq 2 \sum_{v\in\mathbb{P}^1} \int_{Q_0} \vartheta_0(x)\,dx = 2\,(3-1-\frac{1}{2}) = 3$$



Note that the BK thm corresponds to taking the 0-th and ∞-th contributions (and that all of the others terms are ≤ 0)

2.Mixed integrals

Let $\rho : R \to \mathbb{R}, \quad \sigma : S \to \mathbb{R}$ concave functions over convex sets $R, S \subset \mathbb{R}$
Set

$$\rho \boxplus \sigma : R+S \to \mathbb{R} \ , \qquad x \mapsto \max\{\rho(y)+\sigma(z) \ : \ y \in R, \ z \in S, \ y+z = x\}$$

concave function over the Minkowski sum $R + S$

The *mixed integral* is

$$\mathrm{MI}(\rho, \sigma) := \int_{R+S} \rho \boxplus \sigma \, dx \ - \int_R \rho \, dx \ - \int_S \sigma \, dx$$

Very close to the notion of mixed volume of convex sets. Basic properties:

- symmetric and linear in ρ and σ w.r. $\boxplus$
- if $\rho = \sigma$ (and *a fortiori* $R = S$) then $\mathrm{MI}(\rho, \sigma) = 2! \displaystyle\int_R \rho \, dx$
- monotonicity: if $\rho_1 \leq \rho_2$ and $\sigma_1 \leq \sigma_2$ then $\mathrm{MI}(\rho_1, \sigma_1) \leq \mathrm{MI}(\rho_2, \sigma_2)$
- decomposes following the geometry of R and S:

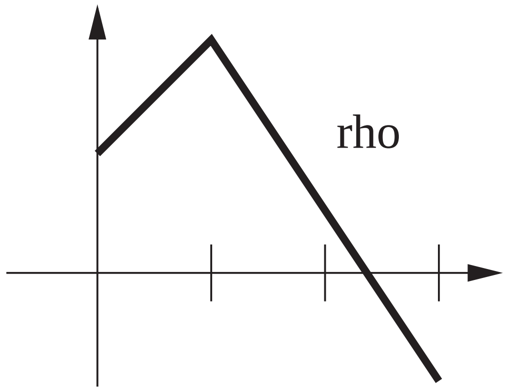
  $$\mathrm{MI}(\rho, \sigma) = \sum_{u \in \mathbb{R}^n} w_{Q_\rho}(u, 1) \, \mathrm{MV}(\pi(Q_\sigma^u)) \ + \sum_{u = \pm 1} w_R(u) \, \mathrm{MI}(\sigma|_{Q^u})$$

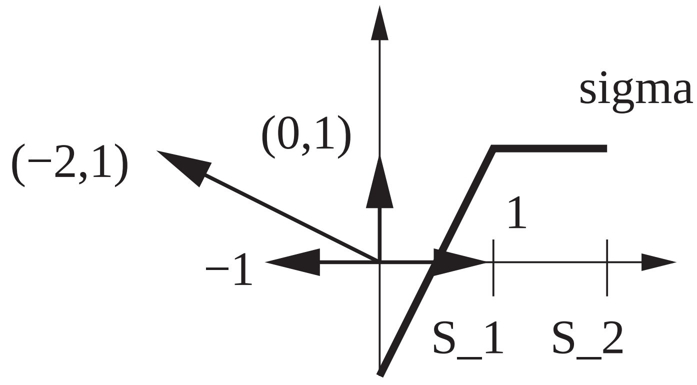
  where $Q_\rho, Q_\sigma \subset \mathbb{R}^2$ are the *convex hull of the graphs* of ρ and σ resp.

  $w_{Q_\rho}(u, 1) := \max_{\mathbf{x} \in Q_\rho} \langle (u, 1), \mathbf{x} \rangle$ *support function*

  $\pi(Q_\sigma^{(u,1)}) \subset \mathbb{R}$ projection to $\mathbb{R}$ of the face of Q_σ in the $(u, 1)$-direction

Example





$$\begin{aligned}\mathrm{MI}(\rho,\sigma) &= (w_{Q_\rho}(-2,1)\,\mathrm{MV}(S_1) + w_{Q_\rho}(1,1)\,\mathrm{MV}(S_2)) \\ &\quad + (w_{[0,3]}(-1)\,\mathrm{MI}(\sigma|_{\{0\}}) + w_{[0,3]}(1)\,\mathrm{MI}(\sigma|_{\{3\}})) \\ &= (1+2)+(0+3) = 6\end{aligned}$$

Thm 1 ("mixed" version). Write $f = \sum_{j\in\mathbb{Z}} \alpha_j(t)\,x^j$, $g = \sum_{j\in\mathbb{Z}} \beta_j(t)\,x^j$ with $\alpha_i(t), \beta_j(t) \in \mathbb{C}[t^{\pm 1}]$, and assume f, g *primitive* Set $Q_i, Q_{i,v}, \vartheta_{i,v}$ as before, then

$$\mathrm{Card}(Z(f,g)_0) \leq \sum_{v\in\mathbb{P}^1} \mathrm{MI}(\vartheta_{1,v}, \vartheta_{2,v})$$

Multiplying $\alpha_j(t)$ and $\beta(t)$ by $\lambda_j, \lambda'_j \in \mathbb{C}^\times$ resp., we have $=$ for generic λ_j, λ'_j

For all $v \neq 0, \infty$ we have $\vartheta_{i,v} \leq 0$ and so

$$\mathrm{MI}(\vartheta_{0,v}, \vartheta_{1,v}) \leq 0$$

For comparison: the BK thm gives the bound

$$\mathrm{Card}(Z(f,g)_0) \leq \mathrm{MV}(Q_{0,\infty}, Q_{1,\infty}) = \mathrm{MI}(\vartheta_{0,\infty}, \vartheta_{1,\infty}) + \mathrm{MI}(\vartheta_{0,0}, \vartheta_{1,0})$$

3. General problem

Let A integral domain equipped with a *height* (complexity measure) $h : A \setminus \{0\} \to \mathbb{R}$ and field of fractions K. Typical examples

$$A = \mathbb{Z} \quad (h(m) = \log|m|) \quad , \quad A = \mathbb{C}[t] \quad (h(f) = \deg(f))$$

Let $f_1, \ldots, f_n \in A[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ which is the height of

$$Z(f_1, \ldots, f_n)_0 \subset (\overline{K}^{\times})^n \ \ ?$$

We consider $A = \mathbb{C}[t]$: $f_1, \ldots, f_n$ system depending on one parameter t

4. Height of varieties over $\mathbb{C}(t)$

Let $Y \subset \mathbb{P}^N(\overline{\mathbb{C}(t)})$ be a $\mathbb{C}(t)$-variety and $\mathcal{Y} \subset \mathbb{P}^1(\mathbb{C}) \times \mathbb{P}^N(\mathbb{C})$

$$\pi \downarrow$$

$$\mathbb{P}^1(\mathbb{C})$$

model for Y, that is $\mathcal{Y}$ is the Zariski closure of

$$\cup_{\xi \in Y} \{((1:t), \xi(t)) : t \in \mathbb{C}\}$$

Then

$$\deg(Y) = \deg(\pi^{-1}(\eta)) \qquad \text{for } \eta \in \mathbb{P}^1 \text{ generic}$$

$$h(Y) := \deg_{\mathbb{P}^N}(\mathcal{Y})$$

- $h(Y) = 0$ iff Y is constant w.r. $\mathbb{P}^1$

For $\xi(t) = (\xi_0(t) : \cdots : \xi_N(t)) \in \mathbb{P}^N(\overline{\mathbb{C}(t)})$ a $\mathbb{C}(t)$-point written in primitive homogeneous coordinates ($\xi(t) \in \mathbb{C}[t]$ and coprime)

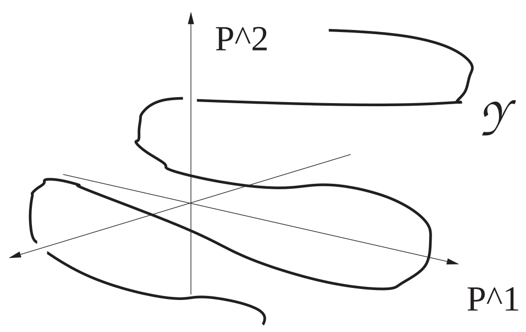
$$h(\xi) = \max_{0 \le j \le N} \deg(\xi_j(t))$$

More generally, for $Y \subset \mathbb{P}^N(\overline{\mathbb{C}(t)})$ a 0-dim $\mathbb{C}(t)$-variety, consider

$$\mathrm{Ch}_Y(\mathbf{U}) = \delta(t) \cdot \prod_{\xi \in Y} (U_0\, \xi_0 + \cdots + U_N\, \xi_N) \quad \in \mathbb{C}[t][\mathbf{U}]$$

its primitive *Chow form* (or $\mathbf{U}$-*resultant*), then

$$\deg(Y) = \deg_{\mathbf{U}}(\mathrm{Ch}_Y) \quad , \quad h(Y) = \deg_t(\mathrm{Ch}_Y)$$



$\deg(Y) =$ number of points for a generic specialization of t

$h(Y) =$ complexity of the description of the curve $\mathcal{Y}$

Consider a *monomial* map of the torus into projective space

$$\begin{aligned}\varphi_{\mathcal{A}_0,\alpha_0}: \quad & (\overline{\mathbb{C}(t)}^{\times})^n \to \mathbb{P}^N(\overline{\mathbb{C}(t)}) \\ & \mathbf{x} = (x_1, \dots, x_n) \mapsto (\alpha_{0,0}(t)\, \mathbf{x}^{a_{0,0}} : \cdots : \alpha_{0,N_0}(t)\, \mathbf{x}^{a_{0,N_0}})\end{aligned}$$

and set

$\mathcal{A}_0 := (a_{0,0}, \dots, a_{0,N_0}) \in (\mathbb{Z}^n)^{N_0+1} \quad , \quad \alpha_0 := (\alpha_{0,0}, \dots, \alpha_{0,N_0}) \in (\mathbb{C}(t)^{\times})^{N_0+1}$

then pose

$$h_{\mathcal{A}_0,\alpha_0} := h \circ \varphi_{\mathcal{A}_0,\alpha_0}$$

height function for $\mathbb{C}(t)$-varieties of $(\overline{\mathbb{C}(t)}^{\times})^n$

The standard inclusion $\mathbf{x} \mapsto (1 : \mathbf{x})$ corresponds to $\mathcal{A}_0 = (\mathbf{0}, (1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)) \in (\mathbb{Z}^n)^{n+1}$ et $\alpha_0 = \mathbf{1} \in (\mathbb{C}[t]^{\times})^{n+1}$

For $i = 1, \dots, n$ let also

$\mathcal{A}_i = (a_{i,0}, \dots, a_{i,N_i}) \in (\mathbb{Z}^n)^{N_i+1} \quad , \quad \alpha_i := (\alpha_{i,0}, \dots, \alpha_{i,N_i}) \in (\mathbb{C}(t)^{\times})^{N_i+1}$

and for $i = 0, \dots, n$ set $f_i = \sum_{j=0}^{N_i} \alpha_{i,j}(t)\, \mathbf{x}^{a_{ij}}$ and $Q_i := \mathrm{Conv}(\mathcal{A}_i) \subset \mathbb{R}^N$

For each $v \in \mathbb{P}^1$ and $0 \leq i \leq n$ consider the v-adic polytope

$$Q_{i,v} := \mathrm{Conv}\Big((a_{i,j}, -\mathrm{ord}_v(\alpha_{i,j}) : 0 \leq j \leq N_i\Big) \subset \mathbb{R}^{n+1}$$

and take $\vartheta_{i,v} : Q_i \to \mathbb{R}$ the parametrization of its upper envelope

Thm 1 (general form).

$$h_{\mathcal{A}_0,\alpha_0}(Z(f_1, \dots, f_n)_0) \leq \sum_{v \in \mathbb{P}^1} \mathrm{MI}(\vartheta_{0,v}, \dots, \vartheta_{n,v})$$

Multiplying each $\alpha_{i,j}(t)$ by $\lambda_{i,j} \in \mathbb{C}^{\times}$, we have $=$ for generic $\lambda_{i,j}$

If $f_0, \dots, f_n$ are primitive then Thm 1 is equivalent to

$$\mathrm{Card}(Z(f_0, \dots, f_n)_0) \leq \sum_{v \in \mathbb{P}^1} \mathrm{MI}(\vartheta_{0,v}, \dots, \vartheta_{n,v})$$

5.Some words about the proof

of the case

$$f=\lambda_0\,(t{-}1){+}\lambda_1\,(t{-}1)^2\,x{+}\lambda_2\,t\,x^2 \quad , \quad g=\lambda'_0\,(t{-}1){+}\lambda'_1\,(t{-}1)^2\,x{+}\lambda_2\,t\,x^2$$

Consider the map

$$(\mathbb{C}^\times)^2\to\mathbb{P}^1\times\mathbb{P}^2 \quad , \quad (t,x)\mapsto((1:t),(t-1:(t-1)^2\,x:t\,x^2))$$

and denote $\mathcal{X}\subset\mathbb{P}^1\times\mathbb{P}^2$ the Zariski closure of its image. Then

$$\mathrm{Card}(Z(f,g)_0)\leq\deg_{\mathbb{P}^2}(\mathcal{X})$$

with $=$ for generic λ_j, λ'_j. The generic fiber of $\mathcal{X}$ over $\mathbb{P}^1$ is a variety

$$\alpha(t)\cdot X\subset\mathbb{P}^2(\overline{\mathbb{C}(t)})$$

translate by a point $\alpha:=(t{-}1,(t{-}1)^2:t)$ of a "constant" toric variety $X\subset\mathbb{P}^N(\mathbb{C})$, Zariski closure of the image of the map $x\mapsto(1:x:x^2)$.
We show that

$$h(\alpha\cdot X)=\sum_{v\in\mathbb{P}^1}e_X(-\mathrm{ord}_v(\alpha))$$

where $e_X(\tau)\in\mathbb{Z}$ denotes the *Chow weight* of X with respect to $\tau\in\mathbb{Z}^3$ [Mumford 1977]
Chow weights of toric varieties can be explicited, see e.g. [Donaldson 2002], [Philippon-S. 2004]: let $Q_\tau:=\mathrm{Conv}((0,\tau_0),(1,\tau_1),(2,\tau_3))\subset\mathbb{R}^2$ and $\vartheta_\tau:Q=[0,2]\to\mathbb{R}$ parametrization of the upper envelope of Q_τ w.r. to Q, then

$$e_X(\tau)=2!\int_{[0,2]}\vartheta_\tau(x)\,dx$$

□

Interesting problems:

- find a Bernstein's type proof, allowing a homothopy continuation algorithm exploiting the geometry of $\vartheta_{i,v}$
- generalisation to any number of parameter variables

6. $\mathbb{Z}$ instead of $\mathbb{C}[t]$

Let $f_1, \ldots, f_n \in \mathbb{Z}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$, which is the complexity of the 0-dimensional variety

$$Z(f_1, \ldots, f_n)_0 \subset (\overline{\mathbb{Q}}^{\times})^n \text{ ?}$$

For $X \subset \mathbb{P}^N(\overline{\mathbb{Q}})$ a 0-dim $\mathbb{Q}$-variety we consider $h(X) \in \mathbb{R}$ its *Weil height*

Basic properties:

- $\xi = (\xi_0 : \cdots : \xi_N) \in \mathbb{P}^N$ a $\mathbb{Q}$-point in primitive homogeneous coordinates ($\xi_j \in \mathbb{Z}$, $\gcd(\xi_0, \ldots, \xi_N) = 1$) then

$$h(\xi) = \max_{0 \le j \le N} \log |\xi_j|$$

- $\mathrm{Ch}_X(\mathbf{U}) = \delta \cdot \prod_{\xi \in X} (U_0 \, \xi_0 + \cdots + U_N \, \xi_N) \in \mathbb{Z}[\mathbf{U}]$ primitive Chow form of X, then

$$\left| h(X) - \max\left(\log |\text{Coeffs of } \mathrm{Ch}_X|\right) \right| \le \log(N+1) \, \mathrm{Card}(X)$$

- $h(X) \ge 0$, and $h(X) = 0$ iff X is torsion

For $i = 0, \ldots, n$ let

$$\mathcal{A}_i = (a_{i,0}, \ldots, a_{i,N_i}) \in (\mathbb{Z}^n)^{N_i+1} \quad , \quad \alpha_i := (\alpha_{i,0}, \ldots, \alpha_{i,N_i}) \in (\mathbb{Q}^\times)^{N_i+1}$$

The vectors $\mathcal{A}_0$ and α_0 define a monomial map

$$\varphi_{\mathcal{A}_0,\alpha_0} : (\overline{\mathbb{Q}}^\times)^n \to \mathbb{P}^N(\overline{\mathbb{Q}}) \quad , \quad \mathbf{x} \mapsto (\alpha_{0,0}\, \mathbf{x}^{a_{0,0}} : \cdots : \alpha_{0,N_0}\, \mathbf{x}^{a_{0,N_0}})$$

and a height function $h_{\mathcal{A}_0,\alpha_0} := h \circ \varphi_{\mathcal{A}_0,\alpha_0}$ for 0-dim $\mathbb{Q}$-varieties of the torus

Also set $f_i = \sum_{j=0}^{N_i} \alpha_{i,j}\, \mathbf{x}^{a_{ij}}$ and $Q_i := \mathrm{Conv}(\mathcal{A}_i) \subset \mathbb{R}^N$

Thm 2. Let $\mathcal{A}_0 \in (\mathbb{Z}^n)^{N_0+1}$ and $\alpha_0 \in \mathbb{Z}^{N_0+1}$, then

$$h_{\mathcal{A}_0,\alpha_0}(Z(f_1, \ldots, f_n)_0) \leq \sum_{v \in M_\mathbb{Q}} \mathrm{MI}(\vartheta_{0,v}, \ldots, \vartheta_{n,v})$$

where $\vartheta_{i,v} : Q_i \to \mathbb{R}$ is some concave function defined

$M_\mathbb{Q} = \{|\cdot|_\infty\} \cup \{|\cdot|_p \ : \ p \text{ prime}\}$ is the *canonical set* of absolute values of $\mathbb{Q}$, where $|\cdot|_\infty$ is the ordinary absolute value and $|\cdot|_p$ is the p-adic absolute value defined by

$$|\alpha|_p := p^{-\mathrm{ord}_p(\alpha)} \quad , \quad \text{for } \alpha \in \mathbb{Q}^\times$$

Should be think of as $\mathrm{Spec}(\mathbb{Z})$ compactified with a point ∞, analogous to $\mathbb{P}^1(\mathbb{C}) = \mathrm{Spec}(\mathbb{C}[t]) \cup (0:1)$

The integrals cannot be easily calculated for $n \geq 2$ though, anyway the estimate is not exact in the general case... Estimating the $\vartheta_{i,v}$'s we re-obtain [S. 2002]

$$h_{\mathcal{A}_0,\alpha_0}(Z(f_1, \ldots, f_n)_0) \leq \sum_{i=0}^{n} \Big(\mathrm{MV}(Q_0, \ldots, Q_{i-1}, Q_{i+1}, \ldots, Q_n) \sum_{j=0}^{N_i} |\alpha_{i,j}|\Big)$$

7. Construction of $\vartheta_{i,v}$

For $v = p$ a prime the construction is the same as before: for eqch $0 \leq i \leq n$ consider the v-adic polytope

$$Q_{i,v} := \mathrm{Conv}\Big((a_{i,j}, -\log|\alpha_{i,j}|_p : j = 0, \ldots, N_i\Big) \subset \mathbb{R}^{n+1}$$

then $\vartheta_{i,v} : Q_i \to \mathbb{R}$ parametrization of its upper envelope

For $v = \infty$, the functions are no longer piecewise affine, but $\mathcal{C}^\infty$. Forget the index i: let

$$T_N := \{\mathbf{t} = (t_0, \ldots, t_N) : t_j \geq 0, \sum_{j=0}^{N} t_j = 1\} \subset \mathbb{R}^{N+1}$$

with the "entropy" map

$$\varepsilon : T_N \to \mathbb{R} \quad , \quad \mathbf{t} \mapsto -\sum_{j=0}^{N} t_i \, \log(t_i)$$

Let

$$X_+ := \Big\{\frac{1}{\sum_{j=0}^{N} |\alpha_j| \, \mathbf{x}^{a_j}} \left(|\alpha_0| \, \mathbf{x}^{a_0}, \ldots, |\alpha_N| \, \mathbf{x}^{a_N}\right) : \mathbf{x} \in (\mathbb{R}_{>0})^n\Big\} \quad \subset T_N$$

a *"positive" toric variety* and

$$\mu : T_N \to Q \quad , \quad \mathbf{t} \mapsto a_0 \, t_0 + \cdots + a_N \, t_N$$

moment map, analytic isomorphism between X_+ and Q°; then

$$\vartheta_\infty := \varepsilon \circ \mu^{-1}$$