DISTANCE TO THE DISCRIMINANT

C. RAFFALLI

Abstract. The main contribution of this article is to establish that for an homogeneous polynomial P of degree d with n variables, every component of the complement of the 0 level of P in $\mathcal{P}^n(\mathbb{R})$ contains a sphere whose radius is proportional to the square root of the distance between P and the discriminant (the set of polynomial with a non smooth zero level).

The distance we use between polynomials is induced by the Bombieri norm [1] for which we establish a nice formula for the distance to the discriminant which is the main tool of our proof.

1. Notation

Let $\mathcal{S}^{n-1}$ be the unit sphere of $\mathbb{R}^n$. In $\mathbb{R}^n$, we write $\|x\|$ the usual Euclidean norm.

We consider $\mathbb{E} = \mathbb{R}[X_1, \dots, X_n]_d$ the vector space of homogeneous polynomials in $n > 1$ variables of degree $d > 1$. Let N be the dimension of this vector space, we have $N = \binom{d+n-1}{n-1} \geq n$

Let $\langle \_, \_ \rangle$ be a scalar product on $\mathbb{E}$ and $\|\_\|$ the associated norm. We use the same notation for the scalar product and the norm of $\mathbb{E}$ and the euclidian one's on $\mathbb{R}^n$, the context should make it clear what norm we are using.

Let $\mathcal{B} = (E_1, \dots, E_N)$ be an orthonormal basis of $\mathbb{E}$.

For $x \in \mathbb{R}^n$, $C(x)$ denotes the line vector $(E_1(x), \dots, E_N(x))$ and $B_i(x)$ for $i \in \{1, \dots, n\}$ denotes the line vector $(\frac{\partial E_1(x)}{\partial x_i}, \dots, \frac{\partial E_N(x)}{\partial x_i})$. Let $B(x)$ be the $n \times N$ matrix whose lines are $B_i(x)$ for $i \in \{1, \dots, n\}$.

For $P \in \mathbb{E}$, let $P_{\mathcal{B}}$ be the column vector coordinate of P in the basis $\mathcal{B}$. We may write:

$$P(x) = C(x)P_{\mathcal{B}}, \frac{\partial P(x)}{\partial x_i} = B_i(x)P_{\mathcal{B}} \text{ and } \nabla P(x) = B(x)P_{\mathcal{B}}$$

We will also use the following notation for the normal and tangent component of $\nabla P(x)$:

$$\begin{array}{rcl} \nabla^N P(x) & = & \langle x | \nabla P(x) \rangle x \\ \nabla^T P(x) & = & \nabla P(x) - \nabla^N P(x) \end{array}$$

We assume the matrix $B(x)$ of maximal rank (i.e. of rank n) for all $x \neq 0$. It may not be true for all basis, but we prove below that it is true when the elements of the basis are monomials (with arbitrary coefficients). This includes the canonical norm and the Bombieri norm which is sufficient for our purpose because the latest

---

Date: April 30, 2007.

is invariant by the action of the orthogonal group which makes it the most natural norm to study the distance to the discriminant.

Let us prove that $B(x)$ is of maximal rank when the elements of $\mathcal{B}$ are monomials with arbitrary coefficients. By symmetry, we may assume that $x_1 \neq 0$. Thus, $B(x)$ contains the following columns coming from the partial derivatives of $a_1 x_1^n$ and $a_i x_1^{n-1} x_i$:

$$\begin{pmatrix} a_1 n x_1^{n-1} & a_2(n-1)x_1^{n-2}x_2 & a_3(n-1)x_1^{n-2}x_3 & \dots & a_n(n-1)x_1^{n-2}x_n \\ 0 & a_2 x_1^{n-1} & 0 & \dots & 0 \\ 0 & 0 & a_3 x_1^{n-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_n x_1^{n-1} \end{pmatrix}$$

And this proves that the n lines of $B(x)$ are linearly independent.

2. The real discriminant

The real discriminant for the degree d and n variables is the set of polynomials $P \in \mathbb{E}$ such that there exists $x \in \mathcal{S}^{n-1}$ where $P(x) = 0$ and $\nabla P(x) = 0$. This can be written

$$\Delta = \bigcup_{x \in \mathcal{S}^{n-1}} \Delta_x \text{ where } \Delta_x = \{P \in \mathbb{E}; B(x)P_{\mathcal{B}} = 0 \text{ and } C(x)P_{\mathcal{B}} = 0\}$$

In fact the equation $C(x)P_{\mathcal{B}} = 0$ is redundant because $C(x) = \frac{1}{d}(x_1, \dots, x_n)B(x)$. This follows from:

$$\sum_{i=1}^{n} x_i \frac{\partial x^\alpha}{\partial x_i} = |\alpha| x^\alpha = d x^\alpha$$

where x^α means $x_1^{\alpha_1} \dots x_n^{\alpha_n}$ and $|\alpha| = \alpha_1 + \dots + \alpha_n = \mathrm{d}$.

Therefore, the discriminant Δ is a union of sub-vector spaces of $\mathbb{E}$ of codimension n (given that $B(x)$ is of maximal rank).

3. Distance to the real discriminant

Let P be a given polynomial in $\mathbb{E}$. We give a way to compute the distance between P and Δ.

We first choose $x_0 \in \mathcal{S}^{n-1}$ and we compute the distance from P to Δ_{x_0}. Therefore, we look for $Q \in \mathbb{E}$, such that:

- $P + Q \in \Delta_{x_0}$.
- $\|Q\|$ minimal.

The first condition may be written

$$B(x_0)(P_{\mathcal{B}} + Q_{\mathcal{B}}) = 0$$

The second condition is equivalent to Q orthogonal to Δ_{x_0}, which means that $Q_{\mathcal{B}}$ is a linear combination of the vectors ${}^t B_i(x_0)$, the columns of ${}^t B(x_0)$.

This means that there exists a column vector H of size n such that

$$Q_{\mathcal{B}} = {}^t B(x_0) H.$$

This gives:

$$B(x_0) P_{\mathcal{B}} + B(x_0)^t B(x_0) H = 0$$

Let us define

$$A(x) = B(x)^t B(x) \text{ and } M(x) = A(x)^{-1}$$

$B(x)$ is a $n \times N$ matrix of maximal rank with $n \leq N$. This implies that $A(x)$ is an $n \times n$ symmetrical and definite matrix for all $x \neq 0$. Hence, $M(x)$ is well defined and symmetrical.

We have

$$B(x_0) P_{\mathcal{B}} + A(x_0) H = 0 \text{ which implies } H = -M(x_0) B(x_0) P_{\mathcal{B}}$$

and

$$Q_{\mathcal{B}} = -{}^t B(x_0) M(x_0) B(x_0) P_{\mathcal{B}}$$

We can now write the distance to Δ_{x_0} by

$$\begin{aligned}
\operatorname{dist}^2(P, \Delta_{x_0}) &= \|Q\|^2 \\
&= {}^t Q_{\mathcal{B}} Q_{\mathcal{B}} \\
&= {}^t P_{\mathcal{B}} {}^t B(x_0) M(x_0) B(x_0) {}^t B(x_0) M(x_0) B(x_0) P_{\mathcal{B}} \\
&= {}^t P_{\mathcal{B}} {}^t B(x_0) M(x_0) A(x_0) M(x_0) B(x_0) P_{\mathcal{B}} \\
&= {}^t P_{\mathcal{B}} {}^t B(x_0) M(x_0) B(x_0) P_{\mathcal{B}} \\
&= {}^t \nabla P(x_0) M(x_0) \nabla P(x_0)
\end{aligned}$$

The above formula is established for any $x_0 \neq 0$ and we can therefore state our first lemma:

Lemma 1. *Let* $(E_1, \ldots, E_N)$ *be an orthornomal basis of* $\mathbb{E} = \mathbb{R}[X_1, \ldots, X_n]_d$ *for a given scalar product. Let* $B(x)$ *be the* $n \times N$ *matrix defined by:*

$$B(x) = \left(\frac{\partial E_j(x)}{\partial x_i} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq N}}$$

Assume $B(x)$ *of rank* n*, which is at least true when the* E_i *are all monomials. Then, for any homogeneous polynomial* $P \in \mathbb{E}$*, the distance to the discriminant* Δ *associated to the given scalar product is given by*

$$dist(P, \Delta) = \min_{x \in \mathcal{S}^{n-1}} \sqrt{{}^t \nabla P(x) M(x) \nabla P(x)} \text{ with } M(x) = (B(x)^t B(x))^{-1}$$

4. Distance with Bombieri norm

The above lemma can be simplified in the particular case of Bombieri norm [1]. To do so, we recall the definition of Bombieri norm and scalar product. Then, we compute the matrix $M(x)$ for this norm.

Notation: let $\alpha = (\alpha_i, \ldots, \alpha_n)$ be a vector in $\mathbb{N}^n$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we write:

- $|\alpha| = \Sigma_{i=1}^n \alpha_i = d$,

- $\alpha! = \Pi_{i=1}^{n} \alpha_i!$,
- $x^\alpha = \Pi_{i=1}^{n} x_i^{\alpha_i}$ for $x \in \mathbb{R}^n$,
- $\chi_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ where the index of 1 is i.

Definition 2 (Bombieri norm and scalar product)**.** *The Bombieri scalar product* [1] *for homogeneous polynomial of degree d is defined by*

$$\|x^\alpha\|^2 = \frac{\alpha!}{|\alpha|!} \text{ and } \langle x^\alpha | x^\beta \rangle = 0 \text{ if } \alpha \neq \beta$$

The Bombieri scalar product and the associated norm have the remarkable property to be invariant by the action of the orthogonal group or $\mathbb{R}^n$. It also verify the Bombieri inequalities for product of polynomials. However, we do not use these properties here. We will give later a shorter proof using the invariance. Nevertheless, we kept the longer proof because it was our first proof and we like this calculation.

Remark: we index the coefficients of the $n \times N$ matrix $B(x)$ by $i \in \{1, \ldots, n\}$ and $\alpha \in \mathbb{Z}^n$ with $|\alpha| = d$. Then, we choose the monomials multiplied by the proper coefficients as our orthonormal base and we have for all $x \in \mathbb{R}^n$:

$$B(x) = (b_{i,\alpha}(x))_{1 \leq i \leq n, |\alpha| = d} \text{ with } b_{i,\alpha}(x) = \sqrt{\frac{|\alpha|!}{\alpha!}} \frac{\partial x^\alpha}{x_i} = \sqrt{\frac{|\alpha|!}{\alpha!}} \alpha_i x^{\alpha - \chi_i}$$

Now, we compute the coefficient of the $n \times n$ matrix $A(x) = B(x)^t B(x)$, when $\|x\| = 1$. We start with the diagonal coefficient $a_{i,i}(x)$:

$$\begin{aligned}
a_{i,i}(x) &= \sum_{|\alpha|=d} \frac{|\alpha|!}{\alpha!} \alpha_i^2 (x^{\alpha - \chi_i})^2 \\
&= \sum_{|\beta|=d-1} \frac{(|\beta|+1)!}{(\beta+\chi_i)!} (\beta_i + 1)^2 x^{2\beta} \text{ change of var. } \beta = \alpha - \chi_i \\
&= d \left(\sum_{|\beta|=d-1} \frac{|\beta|!}{\beta!} (\beta_i + 1) x^{2\beta} \right) \\
&= d \left(\sum_{|\beta|=d-1} \frac{|\beta|!}{\beta!} \beta_i x^{2\beta} + \sum_{|\beta|=d-1} \frac{|\beta|!}{\beta!} x^{2\beta} \right) \\
&= d \left(\sum_{|\beta|=d-1} \frac{|\beta|!}{\beta!} \frac{x_i}{2} \frac{\partial x^{2\beta}}{\partial x_i} + \|x\|^{2(d-1)} \right) \\
&= d \left(\frac{x_i}{2} \frac{\partial}{\partial x_i} \sum_{|\beta|=d-1} \frac{|\beta|!}{\beta!} x^{2\beta} + \|x\|^{2(d-1)} \right) \\
&= d \left(\frac{x_i}{2} \frac{\partial \|x\|^{2(d-1)}}{\partial x_i} + \|x\|^{2(d-1)} \right) \\
&= d \left((d-1) x_i^2 \|x\|^{2(d-2)} + \|x\|^{2(d-1)} \right) \\
&= d(d-1) x_i^2 + d \text{ because } \|x\| = 1
\end{aligned}$$

We now compute the non diagonal coefficient $a_{i,j}(x)$ when $i \neq j$:

$$\begin{aligned}
a_{i,j}(x) &= \sum_{|\alpha|=d} \frac{|\alpha|!}{\alpha!}\alpha_i\alpha_j(x^{\alpha-\chi_i})(x^{\alpha-\chi_j}) \\
&= \sum_{|\beta|=d-2} \frac{(|\beta|+2)!}{\beta!}(x^{\beta+\chi_j})(x^{\beta+\chi_i}) \text{ change of var. } \beta=\alpha-\chi_i-\chi_j \\
&= d(d-1)x_ix_j \sum_{|\beta|=d-2} \frac{|\beta|!}{\beta!}x^{2\beta} \\
&= d(d-1)x_ix_j\|x\|^{2(d-2)} \\
&= d(d-1)x_ix_j \text{ because } \|x\|=1
\end{aligned}$$

This means that we have

$$A(x) = d(d-1)x\,{}^t x + dI$$

Therefore, when $\|x\| = 1$, $A(x)$ has one eigenvalue $d(d-1)+d = d^2$ for the eigenvector x normal to the sphere and all the vectors tangent to the sphere at x are eigenvectors with eigenvalue d. Thus, the matrix $M(x) = A(x)^{-1}$ has one eigenvalue d^{-2} for the eigenvector x normal to the sphere and d^{-1} for the tangent vectors.

Then, the length of the normal component $\nabla^N P(x)$ of $\nabla P(x)$ at x is $\|\nabla^N P(x)\| = {}^t x\nabla P(x) = d\,P(x)$. Moreover, by Pythagorean theorem, the square norm of the tangent component $\nabla^T P(x)$ of $\nabla P(x)$ at x is $\|\nabla^T P(x)\|^2 = \|\nabla P(x)\|^2 - d^2P(x)^2$.

This means that we have

$$\begin{aligned}
\mathrm{dist}^2(P,\Delta_x) &= {}^t\nabla P(x)M(x)\nabla P(x) \\
&= \frac{\|\nabla^N P(x)\|^2}{d^2} + \frac{\|\nabla^T P(x)\|^2}{d} \\
&= \frac{d^2P(x)^2}{d^2} + \frac{\|\nabla P(x)\|^2 - d^2P(x)^2}{d} \\
&= \frac{\|\nabla P(x)\|^2}{d} - (d-1)P(x)^2
\end{aligned} \tag{1}$$

This leads to the following formula:

Theorem 3. *Let $P \in \mathbb{E}$ be an homogeneous polynomial of degree d with n variables. The distance to the real discriminant Δ for the Bombieri norm is given by:*

$$\mathit{dist}(P,\Delta) = \min_{x\in\mathcal{S}^{n-1}} \sqrt{\frac{\|\nabla P(x)\|^2}{d} - (d-1)P(x)^2}$$

Let us show a shorter proof using the invariance by the orthogonal group:

Proof. (short proof of theorem 3) Consider $c \in \mathcal{S}^{n-1}$. We want to compute $\mathrm{dist}(P,\Delta_c)$ One can always find h an element of the orthogonal group such that

$$h(0,\dots,0,1) = c \text{ and } P\circ h(x) = P(c)x_n^d + \nabla^T P(c)x_1x_n^{d-1} + Q(x)$$

where the monomials x_n^d and $x_ix_n^{d-1}$ for $i \in \{1,\dots,n\}$ do not appear in Q.

Then, using the fact that the Bombieri norm is invariant by isometry, the fact that distinct monomials are othogonal and the fact that $Q(x) \in \Delta_{(0,\ldots,0,1)}$ which implies that $Q \circ h^{-1} \in \Delta_c$, we have:

$$\begin{aligned}
\mathrm{dist}^2(P, \Delta_c) &= \mathrm{dist}^2(P \circ h, \Delta_0) \\
&= ||P(c)x_n^d + \nabla^T P(c) x_1 x_n^{d-1}||^2 \\
&= P(c)^2 + \frac{1}{d}||\nabla^T P(c)||^2 \\
&= \frac{1}{d^2}||\nabla^N P(c)||^2 + \frac{1}{d}||\nabla^T P(c)||^2
\end{aligned}$$

Which is the equation (1) at the end of the previous proof of theorem 3 and allows us to end the proof in the same way. □

5. Further simplification

Let us define $\delta_P(x) = \frac{\|\nabla P(x)\|^2}{d} - (d-1)P(x)^2$. In the theorem 3, it is enough to consider the critical points of δ_d, that is points where $\nabla^T \delta_P(x) = 0$. This means we have:

$$\mathrm{dist}(P, \Delta) = \min_{x \in \mathcal{S}^{n-1}, \nabla^T \delta_P(x) = 0} \sqrt{\delta_P(x)}$$

We compute:

$$\begin{aligned}
\frac{d}{2}\nabla \delta_P(x) &= \mathcal{H}P(x)\nabla P(x) - d(d-1)P(x)\nabla P(x) \\
&= \mathcal{H}P(x)\nabla P(x) - \langle \nabla P(x)|x\rangle \mathcal{H}P(x)x \\
&= \mathcal{H}P(x)(\nabla P(x) - \langle \nabla P(x)|x\rangle x) \\
&= \mathcal{H}P(x)\nabla^T P(x)
\end{aligned} \tag{2}$$

Definition 4 (quasi singular points, contact points, contact radius)**.** *We will call quasi singular points for P the critical points of δ_d with norm 1 where the distance to the discriminant is reached. This means that $c \in \mathcal{S}^{n-1}$ is a contact point iff* $\mathrm{dist}(P, \Delta) = \delta_P(c)$ *which implies* $\nabla^T \delta_P(c) = 0$.

We will say that Q is a contact point or contact polynomial for P at c if c is a quasi-singular point for P, $Q \in \Delta_c$ (this means that $\{x; Q(x) = 0\}$ has a singularity at c) and $\mathrm{dist}(P, \Delta) = ||Q - P||$.

When Q is contact polynomial for P at c, we will say that $R = Q - P$ is a contact radius for P at c.

Then, we distinguish two kinds of quasi-singular points for P (their names will be explaned later):

- **quasi-double points:** The critical points of P which satisfies $\nabla^T P(c) = 0$ and therefore are critical points of δ_P.
- **quasi-cuspidal points:** The critical points of δ_P which are not critical points of P.

First, using the quasi-double points, we can find a very simple inequality for the distance to the discriminant:

Theorem 5. *Let $P \in \mathbb{E}$ be an homogeneous polynomial of degree d with n variables. The distance to the real discriminant Δ for the Bombieri norm satisfies:*

$$dist(P, \Delta) \leq \min_{x \in \mathcal{S}^{n-1}, \nabla^T P(x)=0} |P(x)|$$

The condition $\nabla^T P(x) = 0$ means that x is a critical point of P and our theorem means that the distance to the discriminant is less or equal to the minimal value of $|P(x)|$ at the critical points of P.

Proof. We use the theorem 3 and the fact that $\|\nabla P(x)\| = d|P(x)|$ when x is a critical point of P (which means that $\nabla^T P(x) = 0$), we have:

$$\begin{aligned} \mathrm{dist}(P, \Delta)^2 &= \min_{x \in \mathcal{S}^{n-1}} \frac{\|\nabla P(x)\|^2}{d} - (d-1)P(x)^2 \\ &\leq \min_{x \in \mathcal{S}^{n-1}, \nabla^T P(x)=0} \frac{\|\nabla P(x)\|^2}{d} - (d-1)P(x)^2 \\ &= \min_{x \in \mathcal{S}^{n-1}, \nabla^T P(x)=0} \frac{d^2 P(x)^2}{d} - (d-1)P(x)^2 \\ &= \min_{x \in \mathcal{S}^{n-1}, \nabla^T P(x)=0} P(x)^2 \end{aligned}$$

□

Theorem 6. *Let $P \in \mathbb{E}$ be an homogeneous polynomial of degree d with n variables. Let c be a quasi-double point for P. Then, the contact radius at c is the polynomial*

$$R(x) = -P(c)\langle x \mid c\rangle^d.$$

Proof. Let c be a quasi-double point for P. The Bombieri norm being invariant by the orthogonal group, using a rotation we can assume that $c = (0, \ldots, 0, 1)$. Then, we know that $\nabla^T P(c) = 0$ which means that

$$P(x) = P(c)x_n^d + Q(x) \text{ with no monomial } x_n^d \text{ nor } x_i x_n^{d-1} \text{ in } Q(x)$$

Then, the contact radius is the smallest polynomial for the Bombieri norm such that $\{x; P(x) + R(x) = 0\}$ has a singularity at c. Distinct monomials being orthogonal the Bombieri scalar product, this means that

$$R(x) = -P(c)x_n^d = -P(c)\langle x \mid c\rangle^d$$

Now, if we remove the assumption that $c = (0, \ldots, 0, 1)$, we consider h an element of the orthogonal group such that $h(c) = c' = (0, \ldots, 0, 1)$ for c a quasi-double point of P. This implies that c' is a quasi-double point for $P \circ h^{-1}$ at c' and that the corresponding contact radius $R(x)$ is

$$R(x) = -(P \circ h^{-1})(c')\langle x \mid c'\rangle^d = -P(c)\langle h^{-1}(x) \mid c\rangle^d$$

And, the invariance of the Bombieri norm by h implies that the contact radius for P at c is $R \circ h$ which gives the wanted result. □

Corollary 7. *Let $P \in \mathbb{E}$ be an homogeneous polynomial of degree d with n variables. Let c be a quasi-double point for P and Q a corresponding contact polynomial at c. Then, the singularity for $\{x; Q(x) = 0\}$ at c is generically a double point. This justifies the name of "quasi-double" points because these points are the easiest points to make $x; P(x) = 0$ singular.*

Proof. In the above proof, the monomial x^α of Q with $\alpha_n < d-1$ have no reason to be absent. More precisely, for any Q, if

$$P(x) = P(c)x_n^d + Q(x) \text{ with } c = (0, \ldots, 0, 1) \text{ and no monomial } x_n^d \text{ nor } x_i x_n^{d-1} \text{ in } Q(x)$$

and $P(c) = d(P, \Delta)$, then, it is clear that c is a quasi-double contact point for P and generically a double point for $Q(x)$. □

Theorem 8. *Let $P \in \mathbb{E}$ be an homogeneous polynomial of degree d with n variables. Let c be a quasi-cuspidal point for P. Then, the contact radius at c is the polynomial*

$$R(x) = -P(c)\langle x \mid c\rangle^d - \langle x \mid \nabla^T P(c)\rangle\langle x \mid c\rangle^{d-1}.$$

Moreover,

$$\mathcal{H}P(c).\nabla^T P(c) = (d-1)||\nabla^T P(c)||^2.c \text{ recall that } ||c|| = 1.$$

Proof. Let c be a quasi-cuspidal point for P. The Bombieri norm being invariant by the orthogonal group, using a rotation we can assume that $c = (0, \ldots, 0, 1)$ and $\nabla^T P(c) = (||\nabla^T P(c)||, 0, \ldots, 0)$.

We have $\mathcal{H}P.\nabla^T P(c)$ normal to the sphere because c is a critical point of δ_d (see equation (2) page 6). This means that $\mathcal{H}P(c).\nabla^T P(c) = \lambda c$. To compute λ, we use the fact that $||c|| = 1$ and $\langle \mathcal{H}P(c)|c\rangle = (d-1)\nabla P(c)$. This means that

$$\begin{array}{rcl} \lambda & = & \langle \mathcal{H}P(c).\nabla^T P(c)|c\rangle \\ & = & \langle \nabla^T P(c)|\mathcal{H}P(c).c\rangle \\ & = & \langle \nabla^T P(c)|(d-1)\nabla P(c)\rangle \\ & = & (d-1)||\nabla^T P(c)||^2. \end{array}$$

We also have:

$$P(x) = P(c)x_n^d + ||\nabla^T P(c)||x_1 x_n^{d-1} + Q(x) \text{ with no } x_n^d,\ x_i x_n^{d-1} \text{ nor } x_1^2 x_n^{d-2} \text{ in } Q(x)$$

The absence of $x_1^2 x_n^{d-2}$ in $Q(x)$ is a consequence of $\mathcal{H}P(c).\nabla^T P(c) = \lambda c$, because the coefficient of $x_1^2 x_n^{d-2}$ is proportional to $\frac{1}{2}\langle \nabla^T P(c)|\mathcal{H}P(c).\nabla^T P(c)\rangle = \frac{1}{2}\langle \nabla^T P(c)|\lambda c\rangle = 0$.

Then, the contact radius is the smallest polynomial for the Bombieri norm such that $P(x) + R(x)$ has a singularity at c. Distinct monomials being orthogonal the Bombieri scalar product, this means that

$$R(x) = -P(c)x_n^d + ||\nabla^T P(c)||x_1 x_n^{d-1} = -P(c)\langle x \mid c\rangle^d - \langle x \mid \nabla^T P(c)\rangle\langle x \mid c\rangle^{d-1}$$

Then, as in the proof of the previous theorem, the invariance of the Bombieri norm by the orthogonal group implies that the above equality written with scalar product is true for any P and any quasi-cuspidal point c. □

Corollary 9. *Let $P \in \mathbb{E}$ be an homogeneous polynomial of degree d with n variables. Let c be a quasi-cuspidal point for P and Q a corresponding contact polynomial at c. Then, the singularity for Q at x is generically a cusp.*

Proof. Similar to the proof of corollary 7. □

Lemma 10. *When Q is a contact point for P at c, c is the only singularity of $\{x; Q(x) = 0\}$.*

Proof. If this was not the case, Q would be a contact points for two distinct quasi-singular points c and c'. This would mean that the corresponding contact radiuses would be equal which is impossible given the formula established for contact radiuses. $\square$

Theorem 11. *Let $P \in \mathbb{E}$ be an homogeneous polynomial of degree d with n variables. Assume that the zero level of P is extremal (that is no connected component can be added without changing the degree). Then, P has no quasi-cuspidal point and therefore the inequality of theorem 5 is an equality.*

Proof. Let c be a quasi-cuspidal point for P. As in the previous theorem, we consider only the case where $c = (0, \dots, 0, 1)$ and $\nabla^T P(c) = (\varepsilon||\nabla^T P(c)||, 0, \dots, 0)$ with $\varepsilon \in \{1, -1\}$ and:

$P(x) = P(c)x_n^d + \varepsilon||\nabla^T P(c)||x_1x_n^{d-1} + Q(x)$ with no x_n^d, $x_ix_n^{d-1}$ nor $x_1^2x_n^{d-2}$ in $Q(x)$

We also consider that $Q(x) = ax_1^3x_n^{d-3} + \Sigma_{i=2}^{n-1} b_ix_i^2x_n^{d-2} + H_a(x)$ where H_a is a polynomial with only monomial distinct from x_1^3 and of degree $\leq d-3$ with respect to x_n. To do this, we use an element of the orthogonal group to diagonalize the quadratic part of Q when $x_n = 1$.

Finally, we consider that Q is "generic", that is all the coefficient of Q that can be present are non nul. This means that we assume that a and the b_i are non nul.

Then, we define:

$$Q_h(x) = -Q(h, 0 \dots, 0, 1)x_n^d - \frac{\partial Q}{\partial x_1}(h, 0 \dots, 0, 1)x_1x_n^{d-1} + Q(x).$$

It is clear that $Q_h(x)$ has a singularity in $(h, 0, \dots, 0, 1)$.

We remark that:

$$\begin{array}{rcl} Q(h, 0 \dots, 0, 1) & = & ah^3 + O(h^4) \\ \frac{\partial Q}{\partial x_1}(h, 0 \dots, 0, 1) & = & 3ah^2 + O(h^3) \end{array}$$

Now, let us compute $d(P, Q_h)$ and $d(P, Q)$:

$$\begin{array}{rcl} \operatorname{dist}^2(P, Q) & = & P(c)^2 + d^{-1}||\nabla^T P(c)||^2 \\ \operatorname{dist}^2(P, Q_h) & = & (P(c) + (ah^3 + O(h^4)))^2 \\ & + & (d)^{-1}(\varepsilon||\nabla^T P(c)|| + (3ah^2 + O(h^3)))^2 \end{array}$$

Therefore, if $a\varepsilon < 0$, by choosing $|h|$ small enough with a sign such that $P(c)ah < 0$ we have $\operatorname{dist}(P, Q_h) < \operatorname{dist}(P, Q)$. But, by hypothesis, c was a quasi-cuspidal point for P, and Q was the associated contact polynomial. This means that $\operatorname{dist}(P, Q) = \operatorname{dist}(P, \Delta)$ and therefore $\operatorname{dist}(P, Q_h) \geq \operatorname{dist}(P, Q)$ since $Q_h \in \Delta$.

Therefore, we know that $a\varepsilon > 0$ (recall that we assumed $a \neq 0$).

Now, consider:

$$P_h(x) = hP(c)x_n^d + h\varepsilon||\nabla^T P(c)||x_1x_n^{d-1} + Q(x)$$

We have $P_1(x) = P(x)$ and $P_0(x) = Q(x)$

Then, for $x \in \{(x_1, 0, \dots, 0, 1); x_1 \in \mathbb{R}\}$, we have

$$P_h(x) = hP(c) + h\varepsilon||\nabla^T P(c)||x_1 + ax_1^3 + O(x_1^4).$$

The discriminant of the part of P_h of degree less or equal to 3 is

$$27\frac{h^2P^2(x)}{a^2} + 4\frac{h\varepsilon||\nabla^T P(c)||^3}{a^3}$$

This means that this discriminant is always positive for $h > 0$, small enough and $P_h(x)$ as only one real root for $x \in \{(x_1, 0, \dots, 0, 1); x_1 \in \mathbb{R}\}$.

This means that the surface $C_h = \{x; P_h(x) = 0\}$ for $h \in [0, 1]$ has the same number of components than both $\{x; Q(x) = 0\}$ and $\{x; P(x) = 0\}$. Indeed, $\{C_h; h \in [0, 1]\}$ is the segment $[P, Q]$ in $\mathbb{E}$. All its element have the same topology except Q because $\mathrm{dist}(P, \Delta) = \mathrm{dist}(P, Q)$.

It remains to show that for $h > 0$ small enough, C_h and Q have the same number of components. By considring only the dominant term in P_h, for $h > 0$ small enough, there is a neighbourhood V of 0 such that $C_{h\restriction V} = \{x \in V; P_h(x) = 0\}$ has the same topology than $A_{h\restriction V} = \{x \in V; T_h(x) = 0\}$ with

$$T_h(x) = hP(c) + h\varepsilon||\nabla^T P(c)||x_1 + \Sigma_{i=1}^{n-1} b_i x_i^2 x_n^{d-2} + ax_1^3$$

By using the proper homotety on the axis $x_2, \dots, x_{n-1}$, this surface has a symmetry around the x_1 axe and if it had two connected components, they would meet this axe in three points. This means that on the segment $[P, Q]$, we desingularize the cusp of C_0 at c without adding a new component, which is what we had to show since c is the only singularity of C_0 by the previous lemma.

However, it is always possible to desingularize a cusp by adding one component. this contradicts the extremality of P and therefore P cannot have quasi-cuspidal point.

This proves the wanted equality for P extremal and generic. But, the distance to the discriminant being a continuous function, the equality is true for all extremal polynomials. Indeed, non generic here means that some coefficient are zero after composition with a specific element h of the orthogonal group such that $h(0)$ is a quasi singular points. These are closed conditions that are easy to remove one by one using smaller and smaller perturbations. $\square$

6. Some inequalities for the Bombieri norm

For our next results, we will use the following inequality:

Lemma 12. *For all $P \in \mathbb{E}$ and all $x \in \mathbb{R}^n$, we have:*

$$\begin{array}{rcl} |P(x)| & \leq & \|P\|\|x\|^d \\ \|\nabla P(x)\| & \leq & d\,\|P\|\,\|x\|^{d-1} \\ \|\mathcal{H}P(x)\|_2 \leq \|\mathcal{H}P(x)\|_F & \leq & d(d-1)\,\|P\|\,\|x\|^{d-2} \end{array}$$

Using the following norms:

- *The Euclidian norm on $\mathbb{R}^n$ (for x and $\nabla P(x)$),*
- *The Bombieri norm for polynomials (for P)*
- *The Frobenius norm written $\|\_\|_F$ which is the square root of the sum of the squares of the matrix coefficients (for the Hessian $\mathcal{H}P(x)$).*

- *The spectral norm written* $\|\_\|_2$ *which is the largest absolute value of the eigenvalues of the matrix (also for the Hessian* $\mathcal{H}P(x)$*).*

All this inequalities are equalities for the monomial x_i^d for $1 \leq i \leq n$. In this case, the Hessian matrix will have only one non null eigenvalue which implies that $\|\mathcal{H}P(x)\|_2 = \|\mathcal{H}P(x)\|_F$.

We now prove the inequalities. We consider that $P_{\mathcal{B}} = (a_\alpha)_{|\alpha|=d}$ and therefore, $P(x) = \sum_{|\alpha|=d} a_\alpha \sqrt{\frac{d!}{\alpha!}} x^\alpha$:

(1) For the first inequality, the proof is easy:

$$\begin{aligned} P(x)^2 &= \left(\sum_{|\alpha|=d} a_\alpha \sqrt{\frac{d!}{\alpha!}} x^\alpha\right)^2 \\ &\leq \sum_{|\alpha|=d} a_\alpha^2 \sum_{|\alpha|=d} \frac{d!}{\alpha!} x^{2\alpha} \text{ by Cauchy-Schwartz inequality} \\ &= \|P\|^2 \|x\|^{2d} \end{aligned}$$

(2) For the second inequality, we first consider the partial derivative $\frac{\partial P(x)}{\partial x_i}$:

$$\begin{aligned} \left(\frac{\partial P(x)}{\partial x_i}\right)^2 &= \left(\sum_{|\alpha|=d} a_\alpha \sqrt{\frac{d!}{\alpha!}} \alpha_i x^{\alpha-\chi_i}\right)^2 \\ &\leq \sum_{|\alpha|=d} \alpha_i a_\alpha^2 \sum_{|\alpha|=d} \frac{d!}{\alpha!} \alpha_i x^{2(\alpha-\chi_i)} \text{ by Cauchy-Schwartz} \\ &= d \sum_{|\alpha|=d} \alpha_i a_\alpha^2 \sum_{|\alpha|=d, \alpha_i \neq 0} \frac{(d-1)!}{(\alpha-\chi_i)!} x^{2(\alpha-\chi_i)} \\ &= d \sum_{|\alpha|=d} \alpha_i a_\alpha^2 \sum_{|\beta|=d-1} \frac{(d-1)!}{\beta!} x^{2\beta} \\ &= d\|x\|^{2(d-1)} \sum_{|\alpha|=d} \alpha_i a_\alpha^2 \end{aligned}$$

This means that:

$$\begin{aligned} \|\nabla P(x)\|^2 &= \sum_{1\leq i\leq n} \left(\frac{\partial P(x)}{\partial x_i}\right)^2 \\ &\leq \sum_{1\leq i\leq n} \left(d\|x\|^{2(d-1)} \sum_{|\alpha|=d} \alpha_i a_\alpha^2\right) \\ &= d\|x\|^{2(d-1)} \sum_{1\leq i\leq n} \sum_{|\alpha|=d} \alpha_i a_\alpha^2 \\ &= d\|x\|^{2(d-1)} \sum_{|\alpha|=d} \left(\sum_{1\leq i\leq n} \alpha_i\right) a_\alpha^2 \\ &= d^2\|x\|^{2(d-1)} \sum_{|\alpha|=d} a_\alpha^2 \\ &= d^2\|P\|^2\|x\|^{2(d-1)} \end{aligned}$$

(3) For the last inequality, we consider the partial derivative $\dfrac{\partial^2 P(x)}{\partial x_i x_j}$ when $i \neq j$:

$$\begin{aligned}
\left(\frac{\partial^2 P(x)}{\partial x_i x_j}\right)^2 &= \left(\sum_{|\alpha|=d} a_\alpha \sqrt{\frac{d!}{\alpha!}} \alpha_i \alpha_j x^{\alpha-\chi_i-\chi_j}\right)^2 \\
&\leq \sum_{|\alpha|=d} \alpha_i \alpha_j a_\alpha^2 \sum_{|\alpha|=d} \frac{d!}{\alpha!} \alpha_i \alpha_j x^{2(\alpha-\chi_i-\chi_j)} \text{ by Cauchy-Schwartz} \\
&= d(d-1) \sum_{|\alpha|=d} \alpha_i \alpha_j a_\alpha^2 \sum_{|\alpha|=d, \alpha_i \neq 0, \alpha_j \neq 0} \frac{(d-2)!}{\alpha - \chi_i - \chi_j!} x^{2(\alpha-\chi_i-\chi_j)} \\
&= d(d-1) \sum_{|\alpha|=d} \alpha_i \alpha_j a_\alpha^2 \sum_{|\beta|=d-2} \frac{(d-2)!}{\beta!} x^{2\beta} \\
&= d(d-1) \|x\|^{2(d-2)} \sum_{|\alpha|=d} \alpha_i \alpha_j a_\alpha^2
\end{aligned}$$

Now, we consider the partial derivative $\frac{\partial^2 P(x)}{\partial x_i^2}$:

$$\begin{aligned}
\left(\frac{\partial^2 P(x)}{\partial x_i^2}\right)^2 &= \left(\sum_{|\alpha|=d} a_\alpha \sqrt{\frac{d!}{\alpha!}} \alpha_i (\alpha_i - 1) x^{\alpha-2\chi_i}\right)^2 \\
&\leq \sum_{|\alpha|=d} \alpha_i (\alpha_i - 1) a_\alpha^2 \sum_{|\alpha|=d} \frac{d!}{\alpha!} \alpha_i (\alpha_i - 1) x^{2(\alpha-2\chi_i)} \text{ by Cauchy-Schwartz} \\
&= d(d-1) \sum_{|\alpha|=d} \alpha_i (\alpha_i - 1) a_\alpha^2 \sum_{|\alpha|=d, \alpha_i \geq 2} \frac{(d-2)!}{(\alpha - 2\chi_i)!} x^{2(\alpha-2\chi_i)} \\
&= d(d-1) \sum_{|\alpha|=d} \alpha_i (\alpha_i - 1) a_\alpha^2 \sum_{|\beta|=d-2} \frac{(d-2)!}{\beta!} x^{2\beta} \\
&= d(d-1) \|x\|^{2(d-2)} \sum_{|\alpha|=d} \alpha_i (\alpha_i - 1) a_\alpha^2
\end{aligned}$$

Let us define $\iota_{i,j} = 0$ when $i \neq j$ and $\iota_{i,i} = 1$. Then, we have:

$$\begin{aligned}
\|\mathcal{H}P(x)\|_F^2 &= \sum_{1 \leq i,j \leq n} \left(\frac{\partial^2 P(x)}{x_i x_j}\right)^2 \\
&\leq \sum_{1 \leq i,j \leq n} \left(d(d-1) \|x\|^{2(d-2)} \sum_{|\alpha|=d} \alpha_i (\alpha_j - \iota_{i,j}) a_\alpha^2\right) \\
&= d(d-1) \|x\|^{2(d-2)} \sum_{1 \leq i,j \leq n} \sum_{|\alpha|=d} \alpha_i (\alpha_j - \iota_{i,j}) a_\alpha^2 \\
&= d(d-1) \|x\|^{2(d-2)} \sum_{1 \leq i \leq n} \sum_{|\alpha|=d} \alpha_i (d-1) a_\alpha^2 \\
&= d(d-1) \|x\|^{2(d-2)} \sum_{|\alpha|=d} d(d-1) a_\alpha^2 \\
&= d^2 (d-1)^2 \|P\|^2 \|x\|^{2(d-2)}
\end{aligned}$$

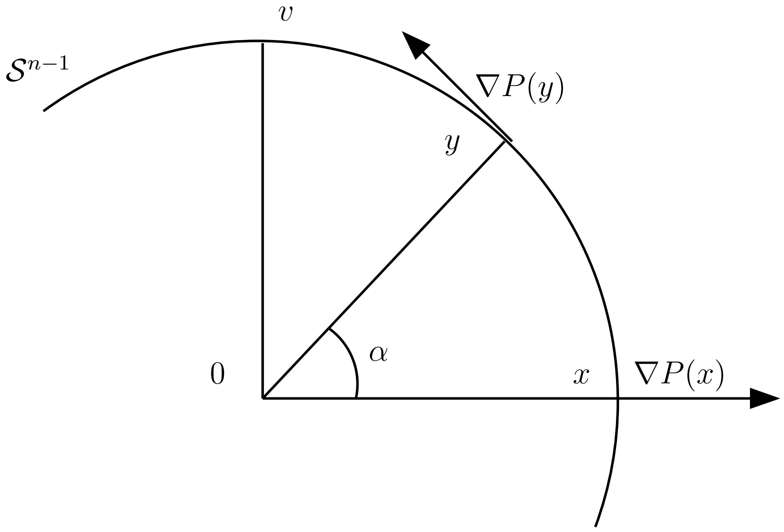

Figure 1.

7. Large components far from the discriminant

Let us consider $x \in \mathcal{S}^{n-1}$ a critical point of P and $y \in \mathcal{S}^{n-1}$ such that $P(y) = 0$. Consider α the measure of the angle $x0y$, and consider $v \in \mathcal{S}^{n-1}$ such that v orthogonal to x and $y = \cos(\alpha)x + \sin(\alpha)v$. See figure 1.

Then, we define:

$$f(\theta) = P(\cos(\theta)x + \sin(\theta)v)$$

We have:

$$\begin{array}{rcl}
f(0) & = & P(x) \\
f(\alpha) & = & P(y) = 0 \\
f'(\theta) & = & {}^t(-\sin(\theta)x + \cos(\theta)v)\nabla P(\cos(\theta)x + \sin(\theta)v) \\
f'(0) & = & {}^t v \nabla P(x) = 0 \text{ because } v \text{ orthogonal to } x \text{ and } \nabla P(x) \\
f''(\theta) & = & {}^t(-\sin(\theta)x + \cos(\theta)v)\mathcal{H}P(\cos(\theta)x + \sin(\theta)v)(-\sin(\theta)x + \cos(\theta)v) \\
 & & + {}^t(-\cos(\theta)x - \sin(\theta)v)\nabla P(\cos(\theta)x + \sin(\theta)v)
\end{array}$$

Using the inequality of lemma 12 and the fact that $\cos(\theta)x + sin(\theta)v \in \mathcal{S}^{n-1}$, we have:

$$\begin{array}{rcl}
|f''(\theta)| & \leq & \|\mathcal{H}P(\cos(\theta)x + \sin(\theta)v)\|_2 \, \|-\sin(\theta)x + \cos(\theta)v\|^2 \\
 & & + \|\nabla P(\cos(\theta)x + \sin(\theta)v)\| \, \|-\cos(\theta)x - \sin(\theta)v\| \\
 & = & \|\mathcal{H}P(\cos(\theta)x + \sin(\theta)v)\|_2 + \|\nabla P(\cos(\theta)x + \sin(\theta)v)\| \\
 & \leq & d(d-1)\,\|P\|\,\|\cos(\theta)x + \sin(\theta)v)\|^{d-2} + d\,\|P\|\,\|\cos(\theta)x + \sin(\theta)v)\|^{d-1} \\
 & = & d^2\|P\|
\end{array}$$

Then, using Taylor-Lagrange equality, we find $\theta \in [0, \alpha]$ such that

$$0 = f(\alpha) = f(0) + \alpha f'(0) + \frac{\alpha^2}{2} f''(\theta) = P(x) + \frac{\alpha^2}{2} f''(\theta)$$

This implies:

$$|P(x)| \leq \frac{d^2\alpha^2}{2}\|P\|$$

and therefore with theorem 5, we have

$$\alpha \geq \frac{1}{d}\sqrt{\frac{2\,\mathrm{dist}(P,\Delta)}{\|P\|}}$$

Therefore, we can state the following result:

Theorem 13. *Let $P \in \mathbb{S}$ (this means that P is an homogeneous polynomial of degree d with n variables and Bombieri norm 1). Let $dist(P,\Delta) \neq 0$ be the distance between P and the real discriminant of $\mathbb{E}$, then for any $x \in \mathcal{S}^{n-1}$ a critical point of P, the open spherical cap of $\mathcal{S}^{n-1}$ with center x and radius angle $\alpha = \frac{1}{d}\sqrt{2dist(P,\Delta)}$ does not meet the zero level of P.*

Corollary 14. *Let $P \in \mathbb{S}$ and $dist(P,\Delta) \neq 0$ be the distance between P and the discriminant for $\mathbb{E}$. Every connected component of the complement of the zero level of P in $\mathcal{S}^{n-1}$ contains an open spherical cap of $\mathcal{S}^{n-1}$ with center x and radius angle $\alpha = \frac{1}{d}\sqrt{2dist(P,\Delta)}$*

This is true, because every connected component of the complement of the zero level of P contains at least one extrema of P which is a critical point of P.

These two last results can also be used in the projective space $\mathcal{P}^{n-1}(\mathbb{R})$ with the metric induced by the metric on the sphere $\mathcal{S}^{n-1}$ because the radius angle of the spherical cap in $\mathcal{S}^{n-1}$ is the radius of a disk is $\mathcal{P}^{n-1}(\mathbb{R})$.

Theorem 15. *Let $P \in \mathbb{S}$ and $dist(P,\Delta) \neq 0$ be the distance between P and the discriminant for $\mathbb{E}$. The distance (measured as an arc length) on between two distinct connected components of the zero level of P in $\mathcal{S}^{n-1}$ is greater of equal to $\alpha = \frac{2}{d}\sqrt{2dist(P,\Delta)}$*

Proof. Consider an arc $[A,B]$ on $\mathcal{S}^{n-1}$ joining two distinct connected components of the zero level of P. By the Ehresmann theorem [2], There is a point C on $[A,B]$ where P reaches a value greater, absolute value, than a critical value of P. We can take C the points where $|P(x)|$ is maximum on $[A,B]$.

Then, by exactly the same computation than for the previous theorem, using the fact that $\nabla P(C)$ is zero in the direction of the segment, we find that the arc length of $[A,C]$ and $[C,B]$ are greater or equal to $\frac{1}{d}\sqrt{2\mathrm{dist}(P,\Delta)}$ which end the proof. □

References

[1] B. Beauzamy, E. Bombieri, P. Enflo, and H. L. Montgomery. Products of polynomials in many variables. *J. Number Th.*, pages 219–245, 1990.

[2] C. Ehresmann. Les connexions infinitsimales dans un espace fibr diffrentiable. In *Colloque de Topologie*, pages 29–55, Bruxelles, 1950. Masson.