Journal of Global Optimization **30:** 405–433, 2004.

Extrema of a Real Polynomial

LIQUN QI*
Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: maqilq@polyu.edu.hk)

(Received: 2 December 2002; revised: 26 May 2003; accepted: 4 January 2004)

Abstract. In this paper, we investigate critical point and extrema structure of a multivariate real polynomial. We classify critical surfaces of a real polynomial f into three classes: repeated, intersected and primal critical surfaces. These different critical surfaces are defined by some essential factors of f, where an essential factor of f means a polynomial factor of $f - c_0$, for some constant c_0. We show that the degree sum of repeated critical surfaces is at most $d-1$, where d is the degree of f. When a real polynomial f has only two variables, we give the minimum upper bound for the number of other isolated critical points even when there are nondegenerate critical curves, and the minimum upper bound of isolated local extrema even when there are saddle curves. We show that a normal polynomial has no odd degree essential factors, and all of its even degree essential factors are normal polynomials, up to a sign change. We show that if a normal quartic polynomial f has a normal quadratic essential factor, a global minimum of f can be either easily found, or located within the interior(s) of one or two ellipsoids. We also show that a normal quartic polynomial can have at most one local maximum.

Key words. critical surface, extrema, extrema surface, normal polynomial, polynomial, quadratic polynomial, quartic polynomial, tensor.

1. Introduction

We assume that $f: \Re^n \to \Re$ is a real polynomial of degree d and n variables. We are interested in minimizing f. Throughout this paper, we assume $d \geqslant 1$ as we are not interested in constant polynomials.

The multivariate polynomial optimization problem has attracted some attention recently [6, 8–10, 12]. It has applications in signal processing [12, 16], merit functions of polynomial equations [6], $0-1$ integer linear and quadratic programs [8], nonconvex quadratic programs [8] and bilinear matrix inequalities [8]. It is related with Hilbert's 17th problem on the representation of nonnegative polynomials [9, 13].

What distinguishes a polynomial from an arbitrary smooth function? If a polynomial only has isolated critical points, then it can only have a finite number of isolated extrema. This number is bounded above by a function which depends

* The work of this author was supported by the Research Grant Council of Hong Kong.

upon the degree and the number of variables of the polynomial. A polynomial of two or more variables may have manifold extrema. The shape of the manifold extrema of a polynomial is also not arbitrary.

Let C_f be the number of isolated critical points of f. When f has only isolated critical points, by the Bézout Theorem [1, 7], we have [12]

$$C_f \leqslant (d-1)^n. \tag{1}$$

Let E_f be the number of isolated extrema of f. Since critical points include extrema and saddle points, we may expect that there is an upper bound for E_f, which is smaller than $(d-1)^n$. In 1993, Durfee et al. [4] studied this problem for $n=2$. They proved that for $n=2$, when f has only isolated critical points,

$$E_f \leqslant \frac{1}{2}d^2 - d + 1. \tag{2}$$

They pointed out that this problem is closely related to Hilbert's 16th problem on the arrangements of ovals of real algebraic curves. Shustin [14, 15] further studied this problem for $n=2$.

From point of view of basic mathematics, this topic is closely related with algebraic geometry and topology, [1–3, 5, 7, 14, 15]. Actually, algebraic geometry is the study of geometric objects defined by polynomial equations, using algebraic means [2], while the number counting for isolated critical points and extrema involves topological degrees and other topological tools [3, 4, 14, 15].

However, from point of view of multivariate polynomial optimization, the above limited knowledge on the numbers of isolated critical points and extrema is vague and seems not very helpful. When $n \geqslant 2$, f may have manifold critical points and extrema. Then we have three immediate questions:

(A) How can we judge if f has only isolated critical points or not?
(B) If f has manifold critical points or extrema, what are their characteristics? Are there upper bounds for the numbers of such critical point manifolds or extrema manifolds?
(C) Even if f has manifold critical points or extrema, may be only isolated critical points or extrema are useful in applications. Are there upper bounds for the numbers of isolated critical points or extrema even if manifold critical points exist?

In this paper, we investigate critical point and extrema structure of a multivariate real polynomial, and in particular, a normal polynomial, which has engineering applications [12]. Normal quartic optimization is the simplest nontrivial case of nonconvex global optimization.

In Section 2, we summarize the current knowledge and questions on isolated critical points and extrema of a general real polynomial f.

In Section 3, we characterize critical surfaces of f. We classify critical surfaces of f into three classes: repeated, intersected and primal critical surfaces. These critical surfaces are defined by some essential factors of f. If we may write

$$f(x) = g(x)h(x) + c_0$$

for some polynomials g and h, and a constant c_0, then we call g and h essential factors of f. If h is irreducible and the zero set of h is nonempty, we call h a substantial polynomial. Repeated and primal critical surfaces are defined by some essential substantial factors of f, while intersected critical surfaces are defined by some essential substantial factor pairs of f. We show that the degree sum of repeated critical surfaces is at most $d-1$.

We then characterize extrema and saddle surfaces of f in Section 4.

In Section 5, we give the minimum upper bound on the number of isolated critical points of a two-variable real polynomial f, even when f has nondegenerate critical curves:

$$\widehat{C}_f \leqslant (d-d_f-1)^2,$$

where $\widehat{C}_f$ be the number of isolated critical points of f, which are not in some repeated critical curves, and d_f is the repeated critical degree sum of f, which we will define later. Clearly, this bound $\widehat{C}_f$ extends (1) in the case of $n=2$. We also give the minimum upper bound on the number of isolated extrema of f, even when f has no nondegenerate extrema curves but f may have some saddle curves and degenerate extrema curves:

$$\widehat{E}_f \leqslant \frac{1}{2}(d-d_f)^2-(d-d_f)+1,$$

where $\widehat{E}_f$ be the number of isolated extrema of f, which are not in some repeated critical curves. Clearly, this bound $\widehat{E}_f$ extends (2), which was given when f has only isolated critical points.

An even degree polynomial is called a normal polynomial if the coefficient tensor of its leading degree term is positive definite. It has engineering applications [12]. In Section 6, we show that a normal polynomial has no odd degree essential factors, and all of its even degree essential factors are normal polynomials, up to a sign change.

In Section 7, we investigate the extrema structure of a normal quartic polynomial f. We show that if it has a repeated quadratic essential factor, or an unrepeated quadratic substantial factor, a global minimum of f can be either easily found, or located within the interior(s) of one or two ellipsoids. Finally, we show that a normal quartic polynomial can have at most one local maximum.

Some final remarks are given in Section 8.

2. Isolated Critical Points and Extrema

We denote F: $\Re^n \to \Re^n$ as the gradient function of f, i.e., $F=\nabla f$. The results of this section are either summarized or derived from some existing results, or observed by counterexamples.

Define $C(n,d)$, $E(n,d)$ and $N(n,d)$ as the minimum upper bounds of numbers of isolated critical points, local extrema and local minima of a real multivariate polynomial f of degree d and n variables. Then we have

$$N(n,d) \leqslant E(n,d) \leqslant C(n,d).$$

Clearly these three functions only take positive integer values. It is clear that

$$E(1,d) = C(1,d) = d-1$$

and

$$N(1,d) = r,$$

where $d = 2r$ or $2r+1$. We hence consider the case that $n \geqslant 2$.

We distinguish two cases: (i) the general case in which f may have manifold critical points; (ii) the special case in which f has only isolated critical points.

We first consider the function $C(n,d)$.

PROPOSITION 1. *In general, we have*

$$C(n,d) \geqslant (d-1)^n. \tag{3}$$

If f has only isolated critical points, we have

$$C(n,d) = (d-1)^n. \tag{4}$$

Proof. Let

$$\phi(t) = \prod_{i=1}^{d-1}(t-i),$$

$$\psi(t) = \int_0^t \phi(t)\mathrm{d}t$$

and

$$f(x) = \sum_{i=1}^{n} \psi(x_i).$$

Then (3) holds. If f has only isolated critical points, by the Bézout Theorem [1, 7],

$$C(n,d) \leqslant (d-1)^n. \tag{5}$$

Hence, by (3), we have (4). □

QUESTION 1. Does (5) also hold even when f has manifold critical points?

In Section 5, we will show that (5) is true for $n=2$ even when f has manifold critical points. Thus, we conjecture that (5) may also be true for $n \geqslant 3$ even when f has manifold critical points.

We now consider the function $E(n,d)$.

PROPOSITION 2. *In general, we have*

$$E(n,2r) \geqslant r^n + (r-1)^n \tag{6}$$

and

$$E(n,2r+1) \geqslant 2r^n \tag{7}$$

for $r \geqslant 1$. *The equality holds in* (6) *and* (7) *when* $r=1$.

Proof. The inequalities can be seen by the example in the proof of Proposition 1. When $r=1$, it is obvious that the equality holds in (6). By [12, 14], a cubic polynomial has at most one isolated local minimum, thus also at most one isolated local maximum. Hence, when $r=1$, the equality holds in (7). □

PROPOSITION 3. *If* f *has only isolated critical points, we have*

$$E(2,2r) = 2r^2 - 2r + 1 \tag{8}$$

and

$$E(2,2r+1) = 2r^2 \tag{9}$$

for $r \geqslant 1$.

Proof. By (2), (6) and (7), we have (8) and (9). □

In Section 5, we will show that (2) holds even when $n=2$ and f has saddle curves.

QUESTION 2. By Propositions 2 and 3, we see that the equality holds in (6) and (7) when $r=1$ or $n=2$ and f has only isolated critical points. Does the equality hold in (6) and (7) when $n \geqslant 3$, $r \geqslant 2$, and f has only isolated critical points?

Finally, we discuss the function $N(n,d)$.

PROPOSITION 4. *In general, we have*

$$N(n,2) = N(n,3) = 1, \tag{10}$$

$$N(2,2r) \geqslant 2r^2 - 2r + 1 \tag{11}$$

for $r \geqslant 1$ *and*

$$N(2,2r+1) \geqslant 2r^2 - r - 1 \tag{12}$$

for $r \geqslant 2$. *When* f *has only isolated critical points, equality holds in* (11).

Proof. The first equality of (10) is obvious. By [12, 14], we have the second equality of (10). By Theorem 3.1.6 of [14], we have (11). By (9), equality holds in (11) when f has only isolated critical points. By Theorem 3.1.6 of [14], when $d = 2r + 1$ is odd, examples can be constructed such that (3.1.2), (3.1.3) and (3.1.5) of [14] hold except the case that (3.1.7) of [14] holds. Let N_f and M_f be the numbers of local minima and maxima of f, α be the number of real intersection points of the projective closure of $\{F(x) = 0\}$ by the line at infinity. If we have examples such that (3.1.2), (3.1.3) and (3.1.5) of [14] hold, we have

$$N_f + M_f = \frac{(d-1)^2 + 1 - \alpha}{2}.$$

However, the rule that (3.1.7) of [14] must not hold implies that

$$M_f \geqslant \frac{d - \alpha}{2}.$$

These two expressions imply that for Shustin's example,

$$N_f = 2r^2 - r - 1.$$

Hence, we have (12). □

Comparing (8) and (9) with (11) and (12), we see that the minimum bound for isolated local minima is very close to the minimum bound for isolated local extrema. We cannot expect a great reduction in the number of local extrema if we only consider local minima.

3. Critical Surfaces

We now cite the following theorem from [1, 7].

THEOREM 1 (Unique Factorization Theorem). *Any non-constant polynomial f over a field can be written uniquely (up to order and non-zero scalars) in the form*

$$f = c f_1^{r_1} \cdots f_s^{r_s}$$

where c is a scalar, $f_1, \ldots, f_s$ are irreducible, *and for $i \neq j$ no f_i is a factor of f_j.*

We discuss on the real field.

Thus, for any real number c_0, f can be written uniquely (up to order and non-zero scalars) in the form

$$f = c f_1^{r_1} \cdots f_s^{r_s} + c_0, \tag{13}$$

where c is a scalar, $f_1, \ldots, f_s$ are irreducible, and for $i \neq j$ no f_i is a factor of f_j, with $c, s, f_1, \ldots, f_s, r_1, \ldots, r_s$ depends upon c_0.

The polynomials $f_1, \ldots, f_s$ are irreducible factors of $f - c_0$, and the numbers $r_1, \ldots, r_s$ are their *multiplicities*. A factor of multiplicity $\geqslant 2$ is said to be *repeated*.

Denote

$$Z := \{x \in \Re^n : F(x) = 0\}.$$

Then Z is the critical point set of f. We use ∂_i to denote $\partial/\partial x_i$.

Let C be a connected component of Z. Then f takes the same value on C. This can be seen by the fact that for any two points z and y in C, $f(z) - f(y)$ should be equal to the line integral

$$\int_\Gamma \partial_1 f(x) \mathrm{d}x_1 + \cdots + \partial_n f(x) \mathrm{d}x_n,$$

where Γ is a smooth curve from y to z on C. Then this line integral is zero as $F(x) \equiv 0$ on C.

We call c_0 a *critical level* of f if there is a critical point x^* of f such that $f(x^*) = c_0$.

QUESTION 3. Let $L(n,d)$ denote the minimum upper bound of the number of such critical levels for a real polynomial of degree d and n variables. Is $L(n,d)$ finite? If so, how can we estimate $L(n,d)$?

Let c_0 be a critical level of f. Then there is at least one factor f_i of $f - c_0$ such that there is $x^* \in \Re^n$, which is a critical point of f and $f_i(x^*) = 0$.

We now wish to know, if h is a factor of $f - c_0$, and $h(x^*) = 0$, is x^* a critical point of f?

To answer this question, we need to define substantial and primal polynomials.

Suppose that h: $\Re^n \to \Re$ is a non-constant polynomial. Denote

$$D_0(h) = \{x \in \Re^n : h(x) = 0\}$$

and let $Z(h)$ be the critical point set of h. Denote

$$ZD_0(h) = D_0(h) \cap Z(h).$$

If h is irreducible and $D_0(h)$ is nonempty, h is called a *substantial polynomial*. If furthermore $ZD_0(h)$ is nonempty, h is called a *primal polynomial*.

For example,

$$h(x) = \sum_{i=1}^n x_i^2$$

is a primal polynomial, while

$$h(x)=\sum_{i=1}^{n}x_i^2+1$$

is not a substantial polynomial.

THEOREM 2 (Critical Points). *Suppose that c_0 is a real constant.*

(i) *If h is a repeated substantial factor of $f-c_0$, then all points in $D_0(h)$ are critical points of f.*
(ii) *If h_1 and h_2 are two substantial factors of $f-c_0$, they do not contain each other as a factor, and $D_0(h_1)\cap D_0(h_2)$ is nonempty, then all points in $D_0(h_1)\cap D_0(h_2)$ are critical points of f.*
(iii) *If h is a primal factor of $f-c_0$, then all points in $ZD_0(h)$ are critical points of f.*
(iv) *Any critical point of f should be in one of the above three types for some c_0.*

Proof. (i) Denote

$$f(x)=g(x)[h(x)]^{p+1}+c_0,$$

where h is substantial and g does not contain h as a factor. Let $x^*\in D_0(h)$. If $p\geqslant 1$, then we have

$$F_i(x^*)=\partial_i g(x^*)[h(x^*)]^{p+1}+(p+1)g(x^*)[h(x^*)]^p\partial_i h(x^*)=0$$

for all i. This proves (i).

(ii) We now denote

$$f(x)=g(x)h_1(x)h_2(x)+c_0.$$

If $h_1(x^*)=h_2(x^*)=0$, we have

$$\begin{aligned}F_i(x^*) &= h_1(x^*)h_2(x^*)\partial_i g(x^*)+g(x^*)h_2(x^*)\partial_i h_1(x^*)\\ &\quad +g(x^*)h_1(x^*)\partial_i h_2(x^*)=0\end{aligned}$$

for all i. This proves (ii).

(iii) Denote

$$f(x)=g(x)h(x)+c_0.$$

If $h(x^*)=0$ and $\nabla h(x^*)=0$, we have

$$F_i(x^*)=h(x^*)\partial_i g(x^*)+g(x^*)\partial_i h(x^*)=0$$

for all i. This proves (iii).

(iv) Finally, assume that x^* is a critical point of f. Let $c_0 = f(x^*)$. Then we may write

$$f(x) = g(x)h(x) + c_0,$$

where h is an irreducible factor of $f - c_0$ and $h(x^*) = 0$. Since x^* is a critical point of f, we have

$$0 = F_i(x^*) = h(x^*)\partial_i g(x^*) + g(x^*)\partial_i h(x^*) = g(x^*)\partial_i h(x^*)$$

for all i. If x^* is not in the types of (i) and (ii), then we have

$$g(x^*) \neq 0.$$

This implies that

$$\partial_i h(x^*) = 0$$

for all i. This shows that $x^* \in ZD_0(h)$, i.e., x^* is a critical point in the type of (iii). This proves (iv). □

LEMMA 1. *Suppose h_1 and h_2 are irreducible factors of $f - c_1$ and $f - c_2$ respectively, $c_1 \neq c_2$. Then h_1 and h_2 do not contain each other as a factor.*

Proof. We have

$$f(x) = g_1(x)h_1(x) + c_1 = g_2(x)h_2(x) + c_2$$

for some polynomials g_1 and g_2. If h_1 is a factor of h_2, then it is a factor of $c_2 - c_1$, which is impossible. □

We say that h is an *essential factor* of f if it is a factor of $f - c_0$ for some constant c_0. By Lemma 1, for an essential factor h of f, the associated constant c_0 is unique.

If h is a repeated essential substantial factor of f, then by Theorem 2 (i), all points of $D_0(h)$ are critical points of f. In this case, we call $D_0(h)$ a *repeated critical surface* of f.

A question is: how many distinct repeated critical surfaces may f have at most? An alternative way to state this question is: how many distinct repeated essential substantial factors of f (up to order and non-zero scalars) may f have at most?

Actually, we may bound the number of repeated essential (may be not substantial) factors of f.

If h is a repeated essential factor of f, then we may write f as:

$$f(x) = g(x)[h(x)]^{p+1} + c_0,$$

where $p \geqslant 1$, h is an irreducible polynomial and g does not contain h as a factor. Then we say that h is a repeated essential factor of f, with multiplicity $p+1$.

Let $\{h_i\colon i \in I\}$ be the set of all distinct repeated essential factors of f (up to order and non-zero scalars) such that h_i is not a factor of h_j for $i \neq j$. Let the degree of h_i be d_i and the multiplicity of h_i as a repeated essential factor of f be $p_i + 1$. Denote

$$d_f = \sum_{i \in I} d_i p_i.$$

We call d_f the *repeated critical degree sum* of f.

THEOREM 3 (Repeated Essential Factor and Repeated Critical Degree Sum). *Suppose that h is an irreducible polynomial.*

(i) *If h is a repeated essential factor of f, with multiplicity* $p+1$, $p \geqslant 1$, *then h is a factor of* F_i, *exactly with multiplicity p, for* $i = 1, \ldots, n$.
(ii) $d_f \leqslant d - 1$.

Proof. Let

$$f(x) = g(x)[h(x)]^{p+1} + c_0,$$

where $p \geqslant 1$ and g does not contain h as a factor.

(i) We now have

$$F_i(x) = [\partial_i g(x) h(x) + (p+1) g(x) \partial_i h(x)] \cdot [h(x)]^p$$

for $i = 1, \ldots, n$. Since the degree of $\partial_i h$ is less than the degree of h, $\partial_i h$ does not contain h as a factor. Since g does not contain h as a factor, the multiplicity of h in F_i is exactly p. This proves (i).

(ii) By (i), since h_i and h_j do not contain each other as a factor for $i \in I, j \in I$ and $i \neq j$, we may write

$$F_i(x) = g_i(x) \prod_{j \in I} [h_j(x)]^{p_j}$$

for some polynomial g_i for $i = 1, \ldots, n$. Since the degree of F_i is not greater than $d - 1$, we have

$$d_f \leqslant d - 1.$$

This proves (ii). □

Let

$$f(x) = [h(x)]^d,$$

where h is a linear function. Then we have $d_f = d-1$. Hence, if we let $\mathrm{RCDS}(n,d)$ denote the minimum upper bound of the repeated critical degree sum of a polynomial of degree d with n variables, then we have

$$\mathrm{RCDS}(n,d) = d-1,$$

which is independent from n.

Write

$$F_i(x) = g_i(x)\prod_{j\in I}[h_j(x)]^{p_j},$$

where $\{h_j\colon j \in I\}$ is the set of all distinct repeated essential factors of f (up to order and non-zero scalars), with multiplicity $p_i + 1$. Then the degree of g_i is not greater than $d-1-d_f$. Clearly, all other critical points which are not in some repeated critical surfaces are solutions of

$$g(x) = 0,$$

where $g\colon \Re^n \to \Re^n$ and the components of g are g_i. By the Bézout theorem, if g has only isolated solutions, the number of these isolated solutions is not greater than $(d-1-d_f)^n$. Hence, we have the following corollary.

COROLLARY 1. *If all the other critical points of f, which are not in some repeated critical surfaces of f, are isolated, then the number of these isolated solutions is not greater than $(d-1-d_f)^n$.*

Let $\mathrm{REF}(n,d)$ denote the minimum upper bound of the number of distinct (up to order and non-zero scalars) repeated essential factors of a polynomial of degree d with n variables. Then we have

$$\mathrm{REF}(n,d) \leqslant \mathrm{RCDS}(n,d) = d-1.$$

Consider

$$f(x) = \prod_{i=1}^{r}[h_i(x)]^2,$$

where $h_1,\ldots,h_r$ are distinct (up to order and non-zero scalars) linear functions. We see that

$$\mathrm{REF}(n,2r+1) \geqslant \mathrm{REF}(n,2r) \geqslant r.$$

QUESTION 4.

$$\mathrm{REF}(n,d) = ?$$

We believe that $\mathrm{REF}(n,d)$ is also independent from n.

QUESTION 5. Can different repeated essential factors of f have different levels? For example, is it possible that

$$f(x)=g_1(x)[h_1(x)]^{p_1}+c_1=g_2(x)[h_2(x)]^{p_2}+c_2,$$

where h_i is irreducible and g_i does not contain h_i as a factor for $i=1,2$, $c_1\neq c_2$?

We now study other types of critical points. Suppose that h_1 and h_2 are two unrepeated substantial factors of $f-c_0$ for some constant c_0, h_1 and h_2 do not contain each other as a factor, and $D_0(h_1)\cap D_0(h_2)$ is nonempty. By Theorem 2, all points of $D_0(h_1)\cap D_0(h_2)$ are critical points of f. In this case, we call (h_1,h_2) a *substantial pair of essential factors* of f, and $D_0(h_1)\cap D_0(h_2)$ an *intersected critical surface* of f.

Let $\{(h_i,h_j):(i,j)\in I_1\}$ be the set of all distinct substantial pairs of essential factors (up to order and non-zero scalars for h_i and h_j) of f. Let the degrees of h_i and h_j be d_i and d_j respectively. We call

$$id_f=\sum_{(i,j)\in I_1} d_i d_j$$

the *intersected critical degree product sum* of f. Let the minimum upper bound of the intersected critical degree product sum of a real polynomial of degree d with n variables be $\mathrm{ICDPS}(n,d)$. Let the minimum upper bound of the number of distinct substantial pairs of essential factors (up to order and non-zero scalars for h_i and h_j) of a real polynomial of degree d with n variables be $\mathrm{SPEF}(n,d)$.

Let

$$f=h_1,\ldots,h_d,$$

where h_i are linear functions, no two of them are proportional, and any two of them have common zero set. Then we see that

$$\mathrm{id}_f=\frac{1}{2}d(d-1).$$

Hence,

$$\mathrm{ICDPS}(n,d)\geqslant\frac{1}{2}d(d-1).$$

We also see that

$$\mathrm{SPEF}(n,d)\geqslant\frac{1}{2}d(d-1).$$

QUESTION 6.

$$\mathrm{ICDPS}(n,d)=?$$

$$\mathrm{SPEF}(n,d)=?$$

QUESTION 7. Can two substantial pairs of essential factors of f be in two different levels, i.e., are there $g_1, g_2, h_1, h_2, h_3, h_4, c_1, c_2$ such that

$$f(x) = g_1(x)h_1(x)h_2(x) + c_1 = g_2(x)h_3(x)h_4(x) + c_2,$$

where h_i are substantial, $D_0(h_1) \cap D_0(h_2)$ and $D_0(h_3) \cap D_0(h_4)$ are not empty, g_1 does not contain h_1 or h_2 as a factor, g_2 does not contain h_3 or h_4 as a factor, $c_1 \neq c_2$?

We call two linear functions are parallel if the hyperplanes which they define are parallel. We have the following proposition:

PROPOSITION 5 (Linear Essential Factors). *Suppose that f has some linear essential factors. Then there are only two cases*:

(i) *All linear essential factors are in the same critical level.*
(ii) *All linear essential factors are parallel.*

Proof. Suppose that $h_1, \ldots, h_{s_l}$ are distinct (up to order and non-zero scalars) linear essential factors of f. If h_1 and h_2 are not parallel, then $D_0(h_1) \cap D_0(h_2)$ is nonempty. Hence, they must be in the same critical level. Then other h_i must not be parallel to at least one of h_1 and h_2. Thus, all of them are in the same critical level. This proves the proposition. □

A trivial example for (ii) is that f itself is a linear function.

QUESTION 8. Is there a nontrivial example for (ii) in this proposition?

Finally, if h is an unrepeated primal essential factor of f, by Theorem 2, all points of $ZD_0(h)$ are critical points of f. In this case, we call $ZD_0(h)$ a *primal critical surface* of f, and call a point in $ZD_0(h)$ a *primal critical point* of f. In many cases, f has only isolated primal critical points.

For example, consider $n = 2$ and

$$f(x) = x_1^3 - 3x_1 - x_2^2.$$

Then f has two critical levels -2 and 2. For $c_0 = -2$, we have a primal essential factor $h(x) = x_1^3 - 3x_1 - x_2^2 + 2$. The only point in $ZD_0(h)$ is $(1, 0)$. Hence, $(1, 0)$ is a primal critical point of f. For $c_0 = 2$, we have a primal essential factor $h(x) = x_1^3 - 3x_1 - x_2^2 - 2$. In this case, the only point in $ZD_0(h)$ is $(-1, 0)$. Hence, $(-1, 0)$ is another primal critical point of f, and f has only two primal critical points.

Denote the minimum bound of the number of distinct primal essential factors (up to order and non-zero scalars) of a real polynomial of degree d with n variables as $\mathrm{PEF}(n, d)$.

QUESTION 9.

$$\mathrm{PEF}(n,d)=?$$

4. Extrema and Saddle Surfaces

In the last section we characterized critical surfaces. If x^* is a point in a repeated or intersected or primal critical surface of f, is it a local minimum or a local maximum or a saddle point of f?

We need to investigate the zero set of an essential factor of f more carefully.

Suppose that h: $\Re^n \to \Re$ is a non-constant polynomial. Beside $D_0(h)$, we now denote

$$D_1(h)=\{x\in\Re^n : h(x)>0\},$$

$$D_2(h)=\{x\in\Re^n : h(x)<0\},$$

$$D_3(h)=\{x\in\Re^n : h(x)\neq 0\},$$

$$D_{00}(h)=D_0(h)\cap \mathrm{cl}D_1(h)\cap \mathrm{cl}D_2(h),$$

$$D_{01}(h)=(D_0(h)\cap \mathrm{cl}D_1(h))\setminus D_{00}(h),$$

$$D_{02}(h)=(D_0(h)\cap \mathrm{cl}D_2(h))\setminus D_{00}(h)$$

and

$$D_{03}(h)=D_{01}(h)\cup D_{02}(h).$$

We call $D_{00}(h)$ the *nondegenerate part* of $D_0(h)$, while $D_{03}(h)$ the *degenerate part* of $D_0(h)$. If $D_{00}(h)$ is nonempty, we say that h has a nondegenerate part. If $D_{03}(h)$ is nonempty, we say that h has a degenerate part. Clearly, all points in $D_{03}(h)$ are extrema of h. Hence, if h has a degenerate part, h is a primal polynomial.

We now characterize the case that x^* is a zero of only one (may be repeated) essential factor of f.

PROPOSITION 6. *Suppose that h is a substantial factor of $f-c_0$ for some constant c_0. Denote*

$$f(x)=g(x)[h(x)]^p+c_0,$$

where g does not contain h as a factor.

(i) *If $p\geqslant 2$ is even, then all points in $D_1(g)\cap D_0(h)$ are local minima of f, and all points in $D_2(g)\cap D_0(h)$ are local maxima of f.*
(ii) *If p is odd and h has a degenerate part, then all points in $D_1(g)\cap D_{01}(h)$ and $D_2(g)\cap D_{02}(h)$ are local minima of f, and all points in $D_1(g)\cap D_{02}(h)$ and $D_2(g)\cap D_{01}(h)$ are local maxima of f.*

(iii) *If* $p \geqslant 3$ *is odd and* h *has a nondegenerate part, then all points in* $D_3(g) \cap D_{00}(h)$ *are saddle points of* f.
(iv) *If* $p = 1$ *and* h *is a primal essential factor of* f *and* h *has a nondegenerate part, then all points in* $D_3(g) \cap ZD_0(h) \cap D_{00}(h)$ *are saddle points of* f.

Proof. These may be directly seen by the definitions of $D_0, D_1, D_2, D_3, D_{00}, D_{01}$ and D_{02}. □

For the example at the end of last section with $n = 2$ and

$$f(x) = x_1^3 - 3x_1 - x_2^2,$$

we see that $h(x) = f(x) + 2$ is a primal essential factor of f. Here, $g(x) = 1$, h has only a nondegenerate part. The only point in $D_3(g) \cap ZD_0(h)$ is $(1,0)$. Hence, $(1,0)$ is a saddle point of f. On the other hand, for $c_0 = 2$, the only point in $D_1(g) \cap D_{02}(h)$ is $(-1,0)$. Hence, $(-1,0)$ is a local maximum of f.

Based upon this proposition, we may characterize extrema surfaces of f.

PROPOSITION 7 (Extrema Surfaces). *Let* c_0 *be a real constant. Suppose that*

$$f(x) = g(x) \prod_{i=1}^{s_0} [h_{0i}(x)]^{2p_{0i}} \cdot \prod_{i=1}^{s_1} [h_{1i}(x)]^{2p_{1i}-1} \cdot \prod_{i=1}^{s_2} [h_{2i}(x)]^{2p_{2i}-1} + c_0,$$

where $p_{ji} \geqslant 1$ *and* h_{ji} *are substantial factors of* $f - c_0$, h_{1i} *and* h_{2i} *have degenerate parts, for* $i = 1, \ldots, s_j$, $j = 0, 1$ *and* 2, $s_0 + s_1 + s_2 \geqslant 1$, h_{ji} *does not contain each other, and* g *does not contain any* h_{ji} *as a factor.*

Then all points in

$$D_3(g) \bigcap \left(\bigcap_{i=1}^{s_0} D_0(h_{0i}) \right) \bigcap \left(\bigcap_{i=1}^{s_1} D_{01}(h_{1i}) \right) \bigcap \left(\bigcap_{i=1}^{s_2} D_{02}(h_{2i}) \right)$$

are local extrema of f. *If* s_2 *is even, then all points in*

$$E_1 = D_1(g) \bigcap \left(\bigcap_{i=1}^{s_0} D_0(h_{0i}) \right) \bigcap \left(\bigcap_{i=1}^{s_1} D_{01}(h_{1i}) \right) \bigcap \left(\bigcap_{i=1}^{s_2} D_{02}(h_{2i}) \right)$$

are local minima of f, *and all points in*

$$E_2 = D_2(g) \bigcap \left(\bigcap_{i=1}^{s_0} D_0(h_{0i}) \right) \bigcap \left(\bigcap_{i=1}^{s_1} D_{01}(h_{1i}) \right) \bigcap \left(\bigcap_{i=1}^{s_2} D_{02}(h_{2i}) \right)$$

are local maxima of f. *If* s_2 *is odd, then all points in* E_1 *are local maxima of* f, *and all points in* E_2 *are local minima of* f. *In these cases, we call* E_1 *and* E_2 extrema surfaces *of* f.

Proof. Again, these may be directly seen by the definitions of $D_0, D_1, D_2, D_3, D_{00}, D_{01}$ and D_{02}. □

For saddle surfaces, we need to think about the following phenomenon.

QUESTION 10. Suppose that

$$h = h_1, \ldots, h_s,$$

where $h_1, \ldots, h_s$ are real polynomials in $\Re^n$. Is it possible that there is a point x^* in $\Re^n$ such that

$$x^* \in D_{03}(h) \bigcap \left(\bigcap_{i=1}^{s} D_{00}(h_i) \right)?$$

If the above phenomenon exists, we say that $\{h_1, \ldots, h_s\}$ forms a *clique* for x^*.

With this definition and definitions for $D_0, D_1, D_2, D_3, D_{00}$ and D_{03}, we have the following proposition which characterizes saddle surfaces of f.

PROPOSITION 8 (Saddle Surfaces). *Let c_0 be a real constant. Suppose that*

$$f(x) = g(x) \prod_{i=1}^{s_0} [h_{0i}(x)]^{2p_{0i}} \cdot \prod_{i=1}^{s_1} [h_{1i}(x)]^{2p_{1i}-1} \cdot \prod_{i=1}^{s_2} [h_{2i}(x)]^{2p_{2i}-1} + c_0,$$

where $p_{ji} \geqslant 1$ and h_{ji} are substantial factors of $f - c_0$, h_{1i} have degenerate parts, and h_{2i} have nondegenerate parts, for $i = 1, \ldots, s_j$, $j = 0, 1$ and 2, $s_2 \geqslant 1$, h_{ji} does not contain each other, and g does not contain any h_{ji} as a factor.

(i) *If $s_2 = 1$, then all points in*

$$S_1 = D_3(g) \bigcap \left(\bigcap_{i=1}^{s_0} D_0(h_{0i}) \right) \bigcap \left(\bigcap_{i=1}^{s_1} D_{03}(h_{1i}) \right) \bigcap D_{00}(h_{21})$$

are saddle points of f.

(ii) *If $s_2 \geqslant 2$, then all points in*

$$S_2 = D_3(g) \bigcap \left(\bigcap_{i=1}^{s_0} D_0(h_{0i}) \right) \bigcap \left(\bigcap_{i=1}^{s_1} D_{03}(h_{1i}) \right) \bigcap \left(\bigcap_{i=1}^{s_2} D_{00}(h_{2i}) \right)$$

are saddle points of f, unless $h_{21}, \ldots, h_{2s_2}$ form a clique for that x. In these cases, we call S_1 and S_2 saddle surfaces *of f.*

Hence, if such cliques do not exist for polynomials, all points in S_2 are saddle points of f. In particular, all points in intersected critical surfaces are saddle points of f. Therefore, if such cliques do not exist, things will be very simple.

PROPOSITION 9. *A local extremum of f is either a strict isolated extremum of f or it is on an extrema surface of f.*

Proof. If x is a non-isolated local extremum of f, then there are points $x^{\{k\}} \to x$ as $k \to \infty$ and

$$f(x^{\{k\}}) = f(x).$$

This implies that x is not a strict local extremum of f and x is on an extrema surface of f. □

5. Isolated Critical Points and Extrema when $n=2$

In this section, we assume that $f\colon \Re^2 \to \Re$ is a real polynomial of degree d and two variables. As stated before, we assume that $d \geqslant 1$.

Since $n=2$, all the terms used in the last two sections with 'surfaces' should be changed to curves. For a substantial polynomial $h\colon \Re^2 \to \Re$, $D_{00}(h)$ defines a curve in the plane, while $D_{03}(h)$ may only contain some isolated points.

For two substantial polynomials $h_1, h_2\colon \Re^2 \to \Re$, if they do not contain each other as a factor, $D_0(h_1) \cap D_0(h_2)$ only contains isolated points, as the intersection of two distinct algebraic curves can only have some isolated points. Since

$$ZD_0(h) = D_0(h) \cap D_0(\partial_1(h)) \cap D_0(\partial_2(h)),$$

$ZD_0(h)$ is also a set of isolated points. Hence, by Theorem 2, for manifold critical points of f, we only need to consider repeated critical curves.

Assume h is a repeated substantial factor of $f - c_0$ for some constant c_0, and the multiplicity of h in $f - c_0$ is even, i.e., we may write

$$f(x) = g(x)[h(x)]^{2p} + c_0,$$

where $p \geqslant 1$ and g does not contain h as a factor. Then by Proposition 6, all points in $D_3(g) \cap D_0(h)$ are local extrema of f. Assume that $D_{00}(h)$ is nonempty, in this case, we call $D_3(g) \cap D_0(h)$ a *nondegenerate extrema curve* of f.

By Theorem 3, we may write

$$F_1 = g_1 h_1^{p_1} \cdots h_s^{p_s} \tag{14}$$

and

$$F_2 = g_2 h_1^{p_1} \cdots h_s^{p_s}, \tag{15}$$

where $h_1, \ldots, h_s$ are all distinct repeated essential factors of f, g_1 and g_2 do not contain any of $h_1, \ldots, h_s$ as a factor. Let the degrees of $h_1, \ldots, h_s$ be $d_1, \ldots, d_s$. Let

$$d_f = \sum_{i=1}^{s} p_i d_i.$$

Then d_f is the repeated critical degree sum of f. We also let

$$H = h_1^{p_1} \cdots h_s^{p_s}.$$

By Corollary 1, we have the following result.

PROPOSITION 10. *Let $\widehat{C}_f$ be the number of isolated critical points of f, which are not in some repeated critical curves. Then*

$$\widehat{C}_f \leqslant (d - d_f - 1)^2.$$

Proposition 10 extends (1) in the case that $n = 2$.

QUESTION 11. How can we bound the number of isolated critical points which are in some repeated essential critical curves of f?

In the rest of this section, we will give the minimum upper bound of the number of isolated extrema of f, when f has no nondegenerate extrema curves. In this case, f may have saddle curves and degenerate extrema curves. Then, if p_i is odd, $D_{00}(h_i)$ must be empty. In this case, if $D_{02}(h_i)$ is nonempty, we may reverse the sign of h_i and absorb this sign change to g_1 and g_2. Then, H is always nonnegative and is positive at an isolated extremum of f, which is not in a repeated critical curve of f.

As Proposition 10 reveals, such an isolated critical point of f is an isolated solution of the polynomial system

$$G(x) \equiv \begin{pmatrix} g_1(x) \\ g_2(x) \end{pmatrix} = 0,$$

where g_1 and g_2 are defined by (14) and (15). Thus, we may study the index of G to estimate the number of isolated extrema of f.

PROPOSITION 11. *Suppose that f has no nondegenerate extrema curves. Let γ be a closed counterclock-oriented curve which does not pass a zero of F and does not encircle repeated critical curves of f. Then the index of F around γ is the same as the index of G around γ.*

Proof. Actually, the index of F around γ is the topological degree of the map u: $\gamma \to S^1$ given by

$$u(P) = \frac{F(P)}{\|F(P)\|} = \frac{G(P)}{\|G(P)\|},$$

while the topological degree of the map given by the right hand side of the above equality is the index of G around γ [4, 11]. □

Thus, we may define the *index* of G as the index of G around a circle (oriented counterclockwise) containing all isolated critical points of f, which are not in some repeated critical curves of f. In this way, we may avoid the critical curves of f which may extend to infinity. Under our assumption, the only nondegenerate critical curves of f are saddle curves.

Let i_G be the index of G, $\widehat{E}_f$ and $\widehat{S}_f$ be the numbers of isolated extrema and isolated saddle points of f, which are not in some repeated critical curves of f, respectively.

Suppose that f has no nondegenerate extrema curves. By [4, 11] and Proposition 11, at a nondegenerate extremum of f, the index of G is $+1$, at a nondegenerate saddle point of f, the index of G is -1. Actually, by perturbation analysis, we may see that at an isolated extremum of f, the index of G is $+p$, at an isolated saddle point of f, the index of G is $-p$, where p is the multiplicity of the extremum or the saddle point in G respectively.

Under our assumption, as in [4], we have

$$i_G = \widehat{E}_f - \widehat{S}_f,$$

counted with multiplicities.

PROPOSITION 12. *The index i_G about a circle C in the plane satisfies*

$$|i_G| \leqslant d - d_f - 1.$$

Proof. By Proposition 11, this proof is the same as the proof of Proposition 2.5 of [4], with F replaced by $G, d-1$ replaced by $d-d_f-1$, and i replaced by i_G. □

By Proposition 11, with F replaced by G, $d-1$ replaced by $d-d_f-1$, and i replaced by i_G, we may follow the discussion of Section 6 of [4] word by word. As Corollary 6.9 of [4], we have the following result.

THEOREM 4. *Suppose that f has no nondegenerate extrema curves. Then,*

$$\widehat{E}_f \leqslant \frac{1}{2}(d-d_f)^2 - (d-d_f) + 1.$$

Clearly, this proposition extends (2).

QUESTION 12. When f has nondegenerate extrema curves, can we give a better upper bound for $\widehat{E}_f$? Now its known upper bound in this case is actually $\widehat{C}_f$, which is bounded by $(d-d_f-1)^2$.

QUESTION 13. How can we extend the results in this section to the case that $n \geqslant 3$?

6. Positive Definite Tensor and Normal Polynomial

Let $r \geqslant 1$ and $n \geqslant 2$ in this section.

We use $A^{(k)}$ to denote a kth order tensor and use $A^{(k)}_{i_1 \ldots i_k}$ to denote its elements. We assume $i_l \in \{1, \ldots, n\}$ for $l = 1, \ldots, k$. We assume that $A^{(k)}$ *totally symmetric*, i.e.,

$$A^{(k)}_{i_1 \ldots i_k} = A^{(k)}_{j_1 \ldots j_k}$$

if $\{i_1, \ldots, i_k\}$ is any reordering of $\{j_1, \ldots, j_k\}$. Let $x \in \Re^n$. Define

$$A^{(k)} x^k := \sum_{i_1, \ldots, i_k = 1}^{n} A^{(k)}_{i_1 \ldots i_k} x_{i_1} \ldots x_{i_k}.$$

Let $\| \cdot \|$ be the 2-norm in $\Re^n$. Denote

$$S := \{x \in \Re^n : \|x\| = 1\}.$$

We say that an even-order tensor $A^{(2r)}$ is *positive definite* if

$$A^{(2r)} x^{2r} > 0$$

for all $x \in S$. We say that an even-order tensor $A^{(2r)}$ is *negative definite* if

$$A^{(2r)} x^{2r} < 0$$

for all $x \in S$. These two definitions extend the definitions for positive and negative definite matrices when $r = 1$.

A $2r$ degree normal polynomial $f \colon \Re^n \to \Re$ can be written as

$$f(x) = \sum_{i=0}^{2r} A^{(2r-i)} x^{2r-i}, \tag{16}$$

where $A^{(i)}$ is an ith order totally symmetric tensor for $i = 0, 1, \ldots, 2r$, and $A^{(2r)}$ is positive definite. Normal polynomials have many nice properties [12]. When $\|x\|$ tends to infinity, the value of a normal polynomial f will also tend to infinity. A normal polynomial always has a global minimum. A bound for the norms of global minima of a normal polynomial was given in [12].

Now, we will study what kinds of essential factors a normal polynomial f may have. We first study the properties of positive definite and negative definite tensors.

THEOREM 5. *Suppose that $n \geqslant 2$, and $A^{(k)}$ is a kth order totally symmetric tensor. Then there is a point $x \in S$ such that*

$$A^{(k)} x^k = 0,$$

if either k is odd, or if k is even but $A^{(k)}$ is neither a positive nor a negative definite tensor.

Proof. If k is odd, let $y \in S$. If $A^{(k)}y^k \neq 0$, then $A^{(k)}z^k$ has different sign from $A^{(k)}y^k$, where $z = -y$. Let C be the shortest circular curve on S which connects y and z. This is possible when $n \geqslant 2$. Since $A^{(k)}x^k$ is continuous as a function of x on C, there is a point $x \in C \subset S$ such that $A^{(k)}x^k = 0$. This proves the theorem when k is odd.

Suppose that k is even but $A^{(k)}$ is neither positive nor negative definite. If there is no $x \in S$ such that $A^{(k)}x^k = 0$, then there are $y, z \in S$ such that $A^{(k)}y^k$ and $A^{(k)}z^k$ have different signs. But as argued before, this implies that there is $x \in S$ such that $A^{(k)}x^k = 0$, which forms a contradiction. This completes the proof of the theorem. □

This justifies that there are only even order positive and negative definite tensors. We now turn to normal polynomials.

THEOREM 6 (Only Normal Essential Factors). *When $n \geqslant 2$, a $2r$ degree normal polynomial f has no odd degree essential factors, and all of its even degree essential factors are normal polynomials, up to a sign change.*

Proof. Let

$$f(x) = g(x)h(x) + c_0,$$

where c_0 is a constant, h is a k degree polynomial and g is a $2r - k$ degree polynomial, $1 \leqslant k \leqslant 2r$ Denote

$$h(x) = \sum_{i=0}^{k} Q^{(k-i)}x^{k-i}$$

and

$$g(x) = \sum_{i=0}^{2r-k} B^{(2r-k-i)}x^{2r-k-i},$$

where $Q^{(i)}$ is an ith order tensor for $i = 0, \ldots, k$, $B^{(i)}$ is an ith order tensor for $i = 0, \ldots, 2r - k$. Comparing with (16), we have

$$A^{(2r)}x^{2r} = (B^{(2r-k)}x^{2r-k})(Q^{(k)}x^k).$$

If either k is odd, or k is even but $Q^{(k)}$ is neither a positive nor a negative definite tensor, then by Theorem 5, there is $x \in S$ such that

$$Q^{(k)}x^k = 0.$$

Then we have

$$A^{(2r)}x^{2r} = (B^{(2r-k)}x^{2r-k})(Q^{(k)}x^k) = 0,$$

contradicting the assumption that $A^{(2r)}$ is positive definite.

If $Q^{(k)}$ is negative definite, then $B^{(2r-k)}$ is also negative definite. Replace g and h by $-g$ and $-h$. Then both $Q^{(k)}$ and $B^{(2r-k)}$ are positive definite. This proves the theorem. □

This theorem reveals that a normal polynomial has very special critical surface structure.

7. Normal Quartic Polynomial

We now consider a normal quartic polynomial $f: \Re^n \to \Re$.

According to the last section, a normal quartic polynomial f may only have quadratic and quartic essential factors. It turns out that if f has a repeated quadratic essential factor, or an unrepeated quadratic substantial factor h, its extrema structure will be simple. In this case, a global minimum of f can be either easily found, or located within the interior(s) of one or two ellipsoids.

By linear algebra, we have the following lemma.

LEMMA 2. *If h is a normal quadratic polynomial, then we may write*

$$h(x) = x^T Qx + a^T x + a_0, \tag{17}$$

where Q is a positive definite $n \times n$ symmetric matrix, $a \in \Re^n$ and $a_0 \in \Re$. Let

$$x^* = -\frac{1}{2}Q^{-1}a. \tag{18}$$

Then

$$h(x^*) = a_0 - \frac{1}{4}a^T Q^{-1}a.$$

If $h(x^*) < 0$, then $D_0(h)$ is an ellipsoid in $\Re^n$. If $h(x^*) = 0$, then $D_0(h) = \{x^*\}$ and $h(x) \geqslant 0$ for all x. If $h(x^*) > 0$, then $h(x) > 0$ for all x.

In the following, if $h(x^*) < 0$, then we use the word 'ellipsoid' to mean the surface $D_0(h)$, while the phrase 'ellipsoid ball' to mean the set

$$\{x \in \Re^n : h(x) \leqslant 0\}.$$

We now discuss different cases in which f has a repeated quadratic essential factor or an unrepeated quadratic substantial factor. The conclusions are made by judging the sign of $f - c_0$ in different regions.

PROPOSITION 13 (Repeated Quadratic Essential Factor). *If a normal quartic polynomial f has a repeated quadratic essential factor h, then we may write*

$$f(x) = [h(x)]^2 + c_0,$$

where h is a normal quadratic polynomial and c_0 is a constant. Let h be expressed by (17) *and x^* be defined by* (18).

If $h(x^*) < 0$, then all points in $D_0(h)$ are global extrema of f, with the global minimum value of f as c_0, and f has a local maximum at x^*.

If $h(x^*) \geqslant 0$, then f has a unique global minimum at x^*. f has no other critical points.

Proof. We have $F(x) = 2h(x)\nabla h(x)$.

If $h(x^*) < 0$, then $D_0(h)$ defines an ellipsoid, $Z = D_0(h) \cup \{x^*\}$, $f(x) \geqslant c_0$ for all x and $f(x) = c_0$ for $x \in D_0(h)$. The conclusion follows.

If $h(x^*) \geqslant 0$, then $Z = \{x^*\}$. Since f has a global minimum as f is a normal polynomial. The conclusion follows. □

The proofs of the following propositions are also simple. We omit their proofs.

PROPOSITION 14 (Two Nondegenerate Ellipsoids). *Suppose that a normal quartic polynomial f has a pair of quadratic substantial factors h and g in the same level, i.e.,*

$$f(x) = g(x)h(x) + c_0$$

for some constant c_0, where both g and h are normal substantial quadratic polynomials, and $D_0(g)$ and $D_0(h)$ define two nondegenerate ellipsoids.

If the two ellipsoid balls are intersected with a nonempty interior, then all points in $D_0(g) \cap D_0(h)$ are saddle points of f, there are a local maximum in the interior of the intersection of these two ellipsoid balls, and two minima in the remaining parts of the interiors of these two ellipsoid balls. Comparing the values of f at these two minima, we may find a global minimum of f.

If these two ellipsoid balls are only touched but do not contain each other, then the touch point is a saddle point of f, and there are two local minima in the interiors of these two ellipsoid balls. Comparing the values of f at these two local minima, we may find a global minimum of f.

If these two ellipsoid balls are not intersected and do not contain each other, then there are two local minima in the interiors of these two ellipsoid balls. Comparing the values of f at these two local minima, we may find a global minimum of f.

If these two ellipsoid balls are touched and one ellipsoid ball contains the other in its interior, then the touch point is a saddle point of f, there is a local maximum in the interior of the smaller ellipsoid ball and a global minimum in the interior of the bigger ellipsoid ball subtracted the smaller ellipsoid ball.

If these two ellipsoid balls are not intersected and one ellipsoid ball contains the other in its interior, then there is a local maximum in the interior of the smaller ellipsoid ball and a global minimum in the interior of the bigger ellipsoid ball subtracted by the smaller ellipsoid ball.

For example, let

$$f(x)=g(x)h(x)+c_0$$

with

$$g(x)=\sum_{i=1}^{n}x_i^2-1$$

and

$$h(x)=(x_1-1)^2+\sum_{i=2}^{n}x_i^2-1.$$

Then $D_0(g)$ and $D_0(h)$ define two spheres. We have

$$F_1(x)=2(2x_1-1)\left[x_1(x_1-1)+\sum_{i=2}^{n}x_i^2-1\right]$$

and

$$F_i(x)=2x_i\left[x_1^2+(x_1-1)^2+2\sum_{i=2}^{n}x_i^2-2\right]$$

for $i=2,\dots,n$. Then we see that $D_0(g)\cap D_0(h)$ defines an $n-1$ dimensional sphere surface

$$\left\{x\in\Re^n : x_1=\frac{1}{2},\ \sum_{i=2}^{n}x_i^2=\frac{3}{4}\right\},$$

which is the saddle surface of f. The value of f at this saddle surface is c_0. f has a local maximum $(\frac{1}{2},0,\dots,0)$, which is at the center of the intersection of the two balls. The value of f at this point is $c_0+\frac{9}{16}$. f has two global minima $((1+\sqrt{5})/2,0,\dots,0)$ and $((1-\sqrt{5})/2,0,\dots,0)$. They are in the interiors of the non-intersected parts of the two balls. The value of f at this point is c_0-1. f has no other critical points.

QUESTION 14. In this example, each F_i is the product of a linear function and a quadratic polynomial. This makes the task to find all critical points of f much easier. In which case can we have such luck?

Note that in the last case of the proposition it is possible that there is a global minima surface. For example, if the two ellipsoids are proportional and have the same axes, etc., then this case can be reformulated as the case described by Proposition 13 and the extrema surface is another ellipsoid surface in the interior of the bigger ellipsoid ball subtracted by the smaller ellipsoid ball.

PROPOSITION 15 (One Nondegenerate Ellipsoid and One Point). *If a normal quartic polynomial f has a pair of quadratic substantial factors h and g in the same level, i.e.,*

$$f(x) = g(x)h(x) + c_0$$

for some constant c_0, where both g and h are substantial quadratic polynomials, and one of $D_0(g)$ and $D_0(h)$ defines a point, say $D_0(h) = \{x^*\}$, and the other of them, say $D_0(g)$, defines a nondegenerate ellipsoid. Then f has a global minimum in the interior of the ellipsoid ball defined by g.

If $g(x^*) > 0$, then x^* is a local minimum of f.

If $g(x^*) = 0$, then x^* is a saddle point of f.

If $g(x^*) < 0$, then x^* is a local maximum of f.

For example, for f defined by

$$f(x) = \left(\sum_{i=1}^{n} x_i^2 - 1\right)\left[(x_1 - 1)^2 + \sum_{i=2}^{n} x_i^2\right],$$

it has only two critical points, a global minimum $(1/2, 0, \ldots, 0)$ with function value $-\frac{27}{16}$, and a saddle point $(1, 0, \ldots, 0)$.

PROPOSITION 16 (Two Single Points). *If a normal quartic polynomial f has a pair of quadratic substantial factors h and g in the same level, i.e.,*

$$f(x) = g(x)h(x) + c_0$$

for some constant c_0, where both g and h are substantial quadratic polynomials, and each of $D_0(g)$ and $D_0(h)$ defines a point, say $D_0(h) = \{x^*\}$ and $D_0(g) = \{y^*\}$. Then x^* and y^* are two global minima of f and the global minimum value of f is c_0.

PROPOSITION 17 (One Quadratic Substantial Factor). *If a normal quartic polynomial f has a pair of quadratic essential factors h and g in the same level, i.e.,*

$$f(x) = g(x)h(x) + c_0$$

for some constant c_0, where one of g and h, say h is a normal substantial quadratic polynomial, and the other, say g, is always positive, i.e., $g(x)>0$, for all x.

If $D_0(h)$ defines a nondegenerate ellipsoid, then f has a global minimum in the interior point of the ellipsoid ball defined by h. If $D_0(h)=\{x^*\}$, then f has the unique global minimum x^* with function value c_0.

The only case which has not discussed so far is the case that f has only primal quartic essential factors. We now see an example of a primal quartic essential factor. Let

$$f(x)=\sum_{i=1}^{p} g_i(x)^2+\sum_{i=1}^{q} h_i(x)^2+c_0,$$

where g_i are normal quadratic polynomials, h_i are linear functions. Let $p\geqslant 1$. Then f is a normal quartic polynomial. Assume that

$$Z=\left(\bigcap_{i=1}^{p} D_0(g_i)\right)\bigcap\left(\bigcap_{i=1}^{q} D_0(h_i)\right)$$

is nonempty. Then we see that Z is the global extrema surface of f. More specifically, let

$$f(x)=\left(\sum_{i=1}^{n} x_i^2-1\right)^2+\sum_{i=1}^{k} x_i^2,$$

where $1\leqslant k\leqslant n-1$. Then we have

$$F_i(x)=4x_i\left(\sum_{i=1}^{n} x_i^2-\frac{1}{2}\right),$$

for $i=1,\dots,k$, and

$$F_i(x)=4x_i\left(\sum_{i=1}^{n} x_i^2-1\right),$$

for $i=k+1,\dots,n$. It is not difficult to see that f has a global minima sphere surface

$$S_1=\left\{x\in\Re^n:\ \sum_{i=k+1}^{n} x_i^2=1,\ x_i=0,\ \text{for } i=1,\dots,k\right\},$$

with function value 0, a saddle sphere surface

$$S_2=\left\{x\in\Re^n:\sum_{i=1}^{k} x_i^2=\frac{1}{2},\ x_i=0,\ \text{for } i=k+1,\dots,n\right\},$$

with function value 3/4, and a local maximum at the origin, with function value 1.

QUESTION 15. Is any critical surface of a normal quartic polynomial the intersection of one or several ellipsoids and some hyperplanes?

We may also see the following example:

$$f(x)=\sum_{i=1}^{n}(x_i^4-2x_i^2).$$

We see that f has 2^n global minima

$$\{x\in\Re^n : x_i=1 \text{ or } -1, \text{ for } i=1,\ldots,n\},$$

with function value $-n$, 3^n-2^n-1 saddle points with some x_i being replaced by zero in the above set, and a local maximum at the origin with function value 0.

QUESTION 16. What is the minimum upper bound of the number of isolated local minima of a normal quartic polynomial? Here we see this bound is at least 2^n.

Finally, we prove the following theorem:

THEOREM 7 (Uniqueness of the Local Maximum). *A normal quartic polynomial f has at most one local maximum.*

Proof. Suppose f has two distinct local maxima. It is easy to see that the restriction of a normal quartic polynomial f in a line is a one dimensional normal quartic polynomial, and in the line connecting these two local maxima, these two local maxima of f are two local maxima of the restriction of f in this line. But a one dimensional normal quartic polynomial has at most one local maxima. This forms a contradiction and the theorem is proved. □

QUESTION 17. In some examples before, we see that f may have no local maximum at all. In which case f has no local maximum? One conjecture is that if

$$f(x)=g(x)h(x)+c_0,$$

g and h are two normal quadratic polynomials, the two ellipsoid balls (may be degenerate) defined by g and h only touch at one point and none of them contains the other in its interior, then f has no local maximum. Is this conjecture true? Is this the only case in which f has no local maximum?

Suppose that f has a unique local maximum x^*. Then the restriction of f on any line passing x^* has two local minima which are on the two sides of x^*. Connecting all these local minima of the line restrictions of f, we have a solid with x^* in its center. We see all the other critical points of f are on the boundary surface of this solid. Thus, in global optimization methods for finding a global minimum of f, we may working on the surface of that solid. We call such a solid the *critical solid* of f.

QUESTION 18. Is such a critical solid convex? If we consider the convex hull of all critical points of a normal quartic polynomial f, are all local maxima of f always located on the boundary of this convex hull, while the local maximum, if exists, is always located in the interior of this convex hull? The examples what we have seen imply a positive answer to this question.

8. Concluding Remarks

In this paper, we explored the critical point and extrema structure of a real polynomial, in particular, a normal quartic polynomial. It is shown that repeated critical surfaces are relatively easy to handle. It is also shown that normal quartic optimization should have more efficient methods than general global optimization because of its special structure. We raised 18 questions along with our theorems and propositions. We hope that more properties of critical and extrema surfaces can be discovered and more efficient algorithms for solving normal quartic optimization can be established later.

Acknowledgments

The author would like to thank Zhong Wan, Yufei Yang, Chen Ling and two referees for their comments.

References

1. Cox, D., Little, J. and O'Shea, D. (1992), *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer-Verlag, New York.
2. Cox, D., Little, J. and O'Shea, D. (1998), *Using Algebraic Geometry*, Springer-Verlag, New York.
3. Durfee, A. (1998), The Index of grad $f(x, y)$, *Topology*, 37, 1339–1361.
4. Durfee, A., Kronenfeld, N., Munson, H., Roy, J. and Westby, I. (1993), Counting critical points of real polynomials in two variables, *Amer. Math. Monthly*, 100, 255–271.
5. Fischer, G. (2001), *Plane Algebraic Curves*, American Mathematical Society, USA.
6. Floudas, C.A. (2000), *Deterministic Global Optimization: Theory, Methods and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
7. Gibson, C.G. (1998), *Elementary Geometry of Algebraic Curves*, Cambridge University Press, Cambridge, UK.
8. Kojima, M., Kim, S. and Waki, H. (2002), A general framework for convex relaxation of polynomial optimization problems over cones, In: Tamura, A. and Ito, H. (eds.), *Proceedings of The Fourteenth RAMP Symposium*, RAMP, Kyoto, pp. 115–132.
9. Lasserre, J.B. (2001), Global optimization with polynomials and the problem of moments, *SIAM J. Opt.*, 11, 796–817.
10. Nesterov, Y. (2000), Squared functional systems and optimization problems, In: Frenk, H., Roos, K. Terlaky, T. and Zhang, S. (eds.), *High Performance Optimization*, Dordrecht, The Netherlands, pp. 405–440.
11. Ortega, J.M. and Rheinboldt, W.C. (1970), *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, Inc., New York.

12. Qi, L. and Teo, K.L. (2003), Multivariate polynomial minimization and its application in signal processing, J. Global Optim. 26, 419–433.
13. Shor, N.Z. (1998), *Nondifferentiable Optimization and Polynomial Problems*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
14. Shustin, E. (1996), Critical points of real polynomials, subdivisions of Newton polyhedra and topology of real algebraic hypersurfaces, *Amer. Math. Soc. Transl.*, 173, 203–223.
15. Shustin, E. (1998), Gluing of singular and critical points, *Topology*, 37, 195–217.
16. Thng, I., Cantoni, A. and Leung, Y.H. (1996), Analytical solutions to the optimization of a quadratic cost function subject to linear and quadratic equality constraints, *Applied Math. Optim.*, 34, 161–382.