Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Global minimization of a mutivariate polynomial using matrix methods

B. Hanzon, D. Jibetean

Probability, Networks and Algorithms (PNA)

PNA-R0109 July 31, 2001

Report PNA-R0109
ISSN 1386-3711

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Global Minimization of a Multivariate Polynomial using Matrix Methods

Bernard Hanzon
Vrije Universiteit
De Boelelaan 1105, 1081 HV Amsterdam
The Netherlands
bhanzon@econ.vu.nl

Dorina Jibetean
Centrum voor Wiskunde en Informatica
P.O. Box 94079, 1090 GB Amsterdam
The Netherlands
d.jibetean@cwi.nl

ABSTRACT
The problem of minimizing a polynomial function in several variables over $\mathbf{R}^n$ is considered and an algorithm is given. When the polynomial has a minimum the algorithm returns the global minimum and finds at least one point in every connected component of the set of minimizers. A characterization of such points is given. When the polynomial does not have a minimum the algorithm can compute its infimum. No assumption is made on the polynomial. The algorithm can be applied for solving a system of polynomial equations.

2000 Mathematics Subject Classification: 90C30,13P10.
Keywords and Phrases: Polynomial optimization, global minimum, infimum, connected components, Gröbner basis, eigenvalue problems, polynomial matrices, algebraic functions.
Note: Research is supported in part by the NWO Project *Constructive Algebra for System Identification and Model Reduction.*

1 Introduction

Consider the problem of minimizing a given polynomial $p \in \mathbf{R}^n[x_1, x_2, ..., x_n]$ over $\mathbf{R}^n$, p having the degree larger than 1. The problem of finding the global minimum of a function is not solved for the general case. Known algorithms find local minima but there are no guarantees that the global minimum is found, except for very special cases. However for polynomial functions the situation is different. It is known that the polynomials have a finite number of critical values (although the number of critical points can be infinite!). Hence, if the polynomial has a minimum, it is possible, at least in principle, to find all critical values and by comparison to decide which one is the global minimum.

In this paper we give an algorithmic solution to the problem. The algorithm is guaranteed to find the minimal value of p, when this exists. When the polynomial has a finite number of points where the minimum is attained, the algorithm finds all of them. In case the number of points of (global) minimum is infinite, there is still a finite number of connected components composing the set $p^{-1}(\{\min_{x\in R^n} p(x)\})$. The algorithm will return at least one point in every connected component. In case the polynomial has a finite infimum, the algorithm will return this value.

No assumptions are made on the polynomial p. Note that we do not include in this setting any domain constraints.

Although optimization problems have been treated extensively, the particular case of polynomial optimization did not received much attention. However some approaches to the problem of (constrained) polynomial optimization can be found in [10] and [14]. The first paper looks at the first order conditions. They form a system of polynomial equations that can be solved for example by using Gröbner basis techniques although in general the computation of a Gröbner basis can be time-costly. Moreover, in the case of infinite number of critical points, even when the Gröbner basis can be computed, its elements may describe very complicated sets of points. It is not clear how one would proceed from there.
The second paper mentioned makes some assumptions on the given polynomial restricting in this way its applicability.

A different algorithm, considered for a different problem but which can be applied for polynomial optimization can be found in [1]. Here the problem of solving a system of polynomial equations over $\mathbf{R}^n$ is considered and an algorithm which returns a point in every connected component of the solution set is given. For solving a polynomial optimization problem, the algorithm could be applied to the first order conditions' system.
The connection between the two problems is stronger than that. Given a set of polynomial equations $f_i(x_1, \ldots, x_n) = 0, \ i = 1, \ldots s$, with our algorithm we can find a point in every connected component of the solution set by applying our algorithm to the polynomial $f = \sum_{i=1}^{s} f_i^2$.

All the algorithms mentioned above work when the given polynomial has a minimum, without considering an approach for finding the infimum.

The remainder of the paper is organized as follows. Section 2 introduces few notions used for systems of equations and describes the Stetter-Möller method for finding all critical points of a polynomial when the number of critical points is finite. In Section 3 we propose a certain perturbation on the original problem which would include the case of infinite number of critical points and we give some theoretical results. Section 4 deals with the actual computations, describing in more detail the output of the algorithm. In the end, Sections 5 and 6, we discuss the algorithm in a particular case and draw the conclusions. Part of intermediate computations of the algorithm can be found in the appendix.

2 Solving polynomial equations

When minimizing a function in n variables one often looks at the first order conditions. They form a system of (nonlinear) equations in n variables. The case of systems of polynomial equations received much attention and methods like Gröbner bases calculation and Stetter-Möller method were proposed for solving them. We assume that the reader is familiar with the theory of Gröbner bases but we discuss the second mentioned method.

2.1 Preliminary notions

To begin, we recall some definitions and results regarding the solutions of a system of polynomial equations. Let K be a field. Given a set of polynomials $f_1, \ldots, f_s \in K[x_1, \ldots, x_n]$ we define

$$< f_1, \ldots, f_s >= \{p_1 f_1 + \ldots + p_s f_s \ : \ p_i \ \in \ K[x_1, \ldots, x_n], i = 1, \ldots, s\}$$

to be the ideal generated by $f_1, \ldots, f_s$. The set of all simultaneous solutions in K^n of a system of equations

$$\begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ f_2(x_1, \ldots, x_n) = 0 \\ \vdots \\ f_s(x_1, \ldots, x_n) = 0 \end{cases}$$

is called the affine or algebraic variety defined by $f_1, \ldots, f_s$ and denoted by $V(f_1, \ldots, f_s)$.
It can be shown easily that the set $< f_1, \ldots, f_s >$ is indeed an ideal. Given a polynomial ideal I one can define the associated affine variety

$$V(I) = \{(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n] | \forall f \in I,\ f(x_1, \ldots, x_n) = 0\}.$$

Note that if $I =< f_1, \ldots, f_s >$, then $V(I) = V(f_1, \ldots, f_s)$.

The generating sets of polynomials for a polynomial ideal I play an important role. Obviously, any finite set of polynomials define a polynomial ideal. The converse is also true: given a polynomial ideal, there always exists a finite set of polynomials which generates it. Note however that the generating set is not unique. The uniqueness of a basis can be obtained by imposing some supplementary conditions on its elements. Given a polynomial ideal by a set of generating polynomials, one can find another generating set which is simpler in some sense. Simpler means for example that we can realize a reduction on the number of variables that appear in certain equations or a reduction on the polynomials' degree while the set of solutions remains invariant. This reduction procedures correspond to computations of what is called a *Gröbner basis*. There are algorithms for computing such a Gröbner basis but in general they have high computational complexity(see [6]).

More details about Gröbner bases and their properties can be found for example in [4].

2.2 Stetter-Möller method for solving systems of polynomial equations

This Section is based on [5], [7], [13].

Given a polynomial ideal I we can define the quotient space $K[x_1, \ldots, x_n]/I$. This set together with an internal addition operation and a scalar multiplication operation has a vector space structure. The elements of this space are classes of polynomials of the form $[f] = \hat{f} + I$. If G is a reduced Gröbner basis for I, then for every polynomial f we have $f = f_1 g_1 + \ldots + f_n g_n + \hat{f}$ where the remainder $\hat{f}$ is unique. Obviously, the remainder is zero if and only if $f \in I$ and polynomials in the same class have the same remainder. The following theorem, characterizing the finite dimensional quotient spaces, is of importance for us.

Theorem 2.1 *Let $K \subseteq \mathbf{C}$ and $I \subseteq K[x_1, \ldots, x_n]$ be an ideal. The following conditions are equivalent:*
a. The vector space $K[x_1, \ldots, x_n]/I$ is finite dimensional over K.
b. The associated variety $V(I)$ is a finite set.
c. If G is a Gröbner basis for I, then for each i, $1 \leq i \leq n$, there is an $m_i \geq 0$ such that $x_i^{m_i}$ is the leading term of g for some $g \in G$.

Such an ideal is called zero-dimensional.

From now on we take the field K equal to the field of complex numbers $\mathbf{C}$. Next we recall the Stetter-Möller method for solving a system of polynomial equations or, in other words, for finding

the points of the variety associated to the generated ideal. When the system of equations has finitely many solutions, that is when $\mathbf{C}[x_1, \ldots, x_n]/I$ is a finite dimensional space over $\mathbf{C}$, the method evaluates an arbitrary polynomial at the points of $V(I)$. In particular, considering $f = x_i$, the method gives the coordinates of the points in $V(I)$. Define $A_f : \mathbf{C}[x_1, \ldots, x_n]/I \to \mathbf{C}[x_1, \ldots, x_n]/I$ by $A_f([g]) = [f][g] = [fg]$.
Note that the multiplication is well defined on $\mathbf{C}[x_1, \ldots, x_n]/I$ due to the fact that I is an ideal. As A_f is a linear mapping from a finite dimensional space to itself, there exists a matrix representation of it with respect to a basis of $\mathbf{C}[x_1, \ldots, x_n]/I$. As such a basis we choose the normal set associated to the reduced Gröbner basis, i.e. the set of monomials which are not divisible by any leading term of the Gröbner basis, $B = \{x^{\alpha(1)}, \ldots, x^{\alpha(m)}\}$. In the following we denote in the same way the linear mapping A_f as well as the matrix associated to it. The following properties hold for the matrices A_f.

Proposition 2.2 *Let $f, g \in \mathbf{C}[x_1, \ldots, x_n]$. Then:*
a. $A_f = 0$ if and only if $f \in I$.
b. $A_{f+g} = A_f + A_g$.
c. $A_{fg} = A_f A_g$.
d. Given a polynomial $h \in \mathbf{C}[t]$ we have $A_{h(f)} = h(A_f)$

Consider the special cases $f = x_i, i = 1, \ldots, n$. Using the properties above it is not difficult to see that $(A_{x_1}, \ldots, A_{x_n})$ is in fact a *matrix element* of $V(I)$, that is $\forall f \in I, f(A_{x_1}, \ldots, A_{x_n}) = 0$. Here 0 denotes the zero matrix and $f(A_{x_1}, \ldots, A_{x_n})$ is well-defined due to the commutativity of the matrices.
Since matrices $A_{x_1}, \ldots, A_{x_n}$ are pairwise commutative, they have common eigenvectors and the n-tuple $(\xi_1 \ldots, \xi_n)$ of eigenvalues of $A_{x_1}, \ldots, A_{x_n}$ respectively, corresponding to the same common eigenvector will be an element of $V(I)$. Moreover, *all* the points in $V(I)$ are found as n-tuples of eigenvalues corresponding to the same common eigenvector ([7]).
For a general polynomial f we have:

Theorem 2.3 *Let $I \subseteq \mathbf{C}[x_1, \ldots, x_n]$ be zero-dimensional, let $f \in \mathbf{C}[x_1, \ldots, x_n]$ and A_f the associated matrix. Then z is an eigenvalue of A_f if and only if z is a value of the function f on $V(I)$.*

In their papers ([12], [13]), Stetter and Möller use instead of A_f the so-called *multiplication table* which is in fact the transpose of our matrix. By looking at the eigenvectors (which in our case become the left eigenvectors) Stetter makes the interesting remark that if the eigenspace associated to a certain eigenvalue of A_f is 1-dimensional, then the eigenvector is $(\xi^{\alpha(1)}, \ldots, \xi^{\alpha(m)})$, where ξ is a solution of the system. In that case we call an eigenvector a Stetter vector. Hence, the solutions of the system can be retrieved from the (left) eigenvectors of A_f.
In [5] a method is proposed for choosing the polynomial f such that the left-eigenspaces of A_f are 1-dimensional, so that one can "read" immediately not only the values of f on $V(I)$ but also the points where the value is obtained.

3 Construction of an auxiliary polynomial

Recall that want to solve the first order conditions of the polynomial p. They form a system of n polynomial equations in the variables $x_1, x_2, ..., x_n$. One approach would be to compute a Gröbner basis and apply for it the Stetter-Möller method, but this is in general time-costly.

In the following, we propose to avoid the computations of a Gröbner basis by looking at a different problem. Consider a family of polynomials depending on the real positive parameter λ given by

$$q_\lambda(x_1, x_2, ..., x_n) = p(x_1, x_2, ..., x_n) + \lambda(x_1^{2m} + x_2^{2m} + \ldots + x_n^{2m}) = p(x_1, x_2, ..., x_n) + \lambda\|x\|^{2m},$$

where $\|x\|$ denotes the Minkowski $2m$ norm of $x = (x_1, x_2, ..., x_n)$ and $m > \texttt{tdeg(p)}/2$ is a fixed positive integer.
When λ goes to zero, from the family of polynomials q_λ we obtain again the polynomial p. We will study the relation between the minima of the polynomials q_λ and the infimum of p. Our claim is that $\inf_{x\in R^n} p(x_1, x_2, ..., x_n) = \lim_{\lambda\to 0} \min_{x\in R^n} q_\lambda(x_1, x_2, ..., x_n)$. Therefore, the new problem is to minimize a family of polynomials $\min_{x\in R^n} q_\lambda(x_1, x_2, ..., x_n)$
The new problem may be easier to solve in the sense that when m is chosen as above, the set of the first order conditions of the new polynomial is a reduced Gröbner basis. So we can avoid the computations of the Gröbner basis by constructing one.

Proposition 3.1 *The first order condition of the polynomial q_λ form a reduced Gröbner basis for the ideal generated by themselves.*

Proof
The partial derivatives of q_λ are $\partial q_\lambda(x)/\partial x_i = 2m\lambda x_i^{2m-1} + \partial p(x)/\partial x_i,\ \forall i = 1, \ldots, n$. With our choice of m, we have $2m > \texttt{tdeg}(p)$ hence $2m-1 > \texttt{tdeg}(\partial p(x)/\partial x_i),\ \forall i = 1, \ldots, n$. In other words, the leading term of $\partial q_\lambda(x)/\partial x_i$ is $2m\lambda x_i^{2m-1}$ and it depends on x_i alone.
According to [4], ch. 2, § 9, Theorem 3 and Proposition 4, the set $\{\partial q_\lambda(x)/\partial x_i \mid i = 1, \ldots, n\}$ is a Gröbner basis. It is obvious that G is in fact a reduced Gröbner basis. □

In the following we discuss the relation between the infimum of the polynomial p and the minima of the polynomials q_λ.

Remark 3.2 The following example is considered in [14]: $p(x_1, x_2) = x_1^2 x_2^4 + x_1 x_2^2 + x_1^2$ for which we have $\inf_{(x_1,x_2)\in\mathbf{R}^2} p(x_1, x_2) = -1/4$. However the infimum is "reached" at infinity. In the following, we'll distinguish between the cases when p has a minimum, has a finite infimum or an infinite infimum.

Lemma 3.3 *For every positive λ, the polynomial q_λ has a minimum.*

Proof
We want to show that for every $\lambda > 0$ there exists an r_λ such that the minimum of q_λ is reached inside the Minkowski ball $B(0, r_\lambda)$.
Let $x \in \mathbf{R}^n$ with the norm $\| x \| = r$. Then for every component of x we have $-r \le x_i \le r,\ i = 1, \ldots, n$ and

$$q_\lambda(x) = \| x \|^{2m} (\lambda + p(x)/ \| x \|^{2m})$$

But $-p_{abs}(r) \le p(x) \le p_{abs}(r)$ for all x with $\| x \| = r$ implies

$$r^{2m}(\lambda - p_{abs}(r)/r^{2m}) \quad \le \quad q_\lambda(x)$$

Here p_{abs} is the polynomial obtained from p by replacing all its coefficients by their absolute value and taking all its variables equal. By construction we have that $2m$ is strictly larger than the total degree of the polynomial p (and also p_{abs}), therefore $p_{abs}(r)/r^{2m}$ is a rational function in the variable r having the degree of the numerator strictly smaller than the degree of the denominator. Hence $\lim_{r\to\infty} p_{abs}(r)/r^{2m} = 0$ and so there exists an $r_\lambda^1 > 0$ such that for every $r \ge r_\lambda^1$ we have $\lambda > p_{abs}(r)/r^{2m}$. That means that for every x with $\| x \| = r \ge r_\lambda^1$ we have

$$0 \quad < \quad r^{2m}(\lambda - p_{abs}(r)/r^{2m}) \le q_\lambda(x). \tag{1}$$

Note that at the point $x = 0$ we have $q_\lambda(0) = p(0)$.
From (1) we see that $q_\lambda(x)$ goes to infinite for $r \to \infty,\ r = \|x\|$. Hence $\exists r_\lambda \geq r_\lambda^1$ such that $\forall r \geq r_\lambda$ and $x,\ \|x\| = r$, we have $q_\lambda(x) > q_\lambda(0)$, where $q_\lambda(0) = p(0)$ is a fixed number.
Hence $\forall x,\ \|x\| \geq r_\lambda$ we have $q_\lambda(x) > q_\lambda(0)$ which implies that the minimum of q_λ must be attained inside the Minkowski ball $B(0, r_\lambda)$.
That completes our proof. □

Denote by X_λ the set of real points where the minimum of q_λ is attained

$$X_\lambda = \{x_\lambda \in \mathbf{R}^n \mid q_\lambda(x_\lambda) = \min_{x \in \mathbf{R}^n} q_\lambda(x)\}.$$

Elements of X_λ will be denoted by x_λ. From Theorem 2.1 we know that X_λ is a finite set for every λ positive. From the previous proposition we also know that X_λ in nonempty for every λ positive. In the following we'll use the notion of limit set as defined below. The set L given by

$$L = \{x \in \mathbf{R}^n \mid \forall \varepsilon > 0\ \exists\ \lambda_\varepsilon\ s.t.\ \forall \lambda,\ 0 < \lambda < \lambda_\varepsilon,\ X_\lambda \cap B(x, \varepsilon) \neq \emptyset\}$$

is called the limit set of X_λ. For a multi-valued function with branches, by definition, the limit set will be simply the set of limits on every branch, assuming they exist.

Theorem 3.4 *The following statements are true:*

(i) $\lim_{\lambda \to 0} \min_{x \in \mathbf{R}^n} q_\lambda(x) = \inf_{x \in \mathbf{R}^n} p(x)$

(ii) $\lim_{\lambda \to 0} p(x_\lambda) = \inf_{x \in \mathbf{R}^n} p(x),\ \forall x_\lambda \in X_\lambda$

(iii) If the polynomial p has a minimum then $L \subseteq \{x \in \mathbf{R}^n \mid p(x) = \min_{x \in \mathbf{R}^n} p(x)\}$.

Proof
(i) We consider two cases. First, we treat the case when p has a minimum attained at some point $\underline{x}$. Then

$$p(\underline{x}) = \inf_{x \in \mathbf{R}^n} p(x) \leq \inf_{x \in \mathbf{R}^n} (p(x) + \lambda \|x\|^{2m}) \leq p(\underline{x}) + \lambda \|\underline{x}\|^{2m}.$$

The above relation holds for every $\lambda > 0$, hence the relation is also valid at the limit $\lambda \downarrow 0$:

$$p(\underline{x}) \leq \lim_{\lambda \downarrow 0} \inf_{x \in \mathbf{R}^n} q_\lambda(x) \leq p(\underline{x})$$

which proves our statement.
Suppose now that $\inf_{x \in \mathbf{R}^n} p(x) = INF$, where INF may be finite or infinite. Let M be a real number $M > INF$, arbitrarily close to INF. As p does not reach INF, there exists an $\underline{x}$ such that $p(\underline{x}) < M$; then there is a $\underline{\varepsilon} > 0$ such that $p(\underline{x}) + \underline{\varepsilon} < M$. Define $\lambda_{\underline{\varepsilon}} = \underline{\varepsilon}/\|\underline{x}\|^{2m}$, where $\|x\|$ is the Minkowski norm. Then we have that for every $\lambda < \lambda_{\underline{\varepsilon}}$

$$\min_{x \in \mathbf{R}^n} [p(x) + \lambda \|x\|^{2m}] \leq p(\underline{x}) + \lambda \|\underline{x}\|^{2m} < M.$$

Since for every positive λ_1, λ_2 with $\lambda_1 < \lambda_2$ we have $q_{\lambda_1}(x) \leq q_{\lambda_2}(x),\ \forall x \in \mathbf{R}^n$, the limit exists and

$$\inf_{x \in R^n} p(x) \leq \lim_{\lambda \downarrow 0} [\min_{x \in R^n} [p(x) + \lambda \|x\|^{2m}]] \leq M$$

As M is arbitrarily close to INF,

$$\lim_{\lambda \downarrow 0} [\min_{x \in R^n} [p(x) + \lambda \|x\|^{2m}]] = INF$$

(ii) It follows immediately from (i) since $\inf_{x\in\mathbf{R}^n} p(x) \le p(x_\lambda) \le q_\lambda(x_\lambda),\ \forall x_\lambda \in X_\lambda$.
(iii) Define $S = \{x \in \mathbf{R}^n \mid p(x) = \min_{x\in\mathbf{R}^n} p(x)\}$. We want to show $L \subseteq S$. By contradiction, suppose $\exists x \in L,\ x \notin S$. Clearly $x \notin S$ is equivalent to $p(x) \neq \min_{x\in\mathbf{R}^n} p(x)$. From the definition of the limit set L, we can construct a function which associates to every $\lambda > 0$ an $x_\lambda \in X_\lambda$ such that

$$\forall\varepsilon > 0\ \exists\lambda_\varepsilon > 0,\ s.t.\ \forall\lambda,\ 0 < \lambda < \lambda_\varepsilon \quad x_\lambda \in B(x, \varepsilon)$$

But this says exactly that $\lim_{\lambda\downarrow 0} x_\lambda = x$. As p is a continuous function we have that $\lim_{\lambda\downarrow 0} p(x_\lambda) = p(x)$.
From part (ii) we have $\lim_{\lambda\downarrow 0} p(x_\lambda) = \min_{x\in\mathbf{R}^n} p(x)$.
This is however in contradiction with our assumption that $p(x) \neq \min_{x\in\mathbf{R}^n} p(x)$. That concludes our proof. □

According to the theorem, one can obtain the infimum of p from the minima of the family of polynomials q_λ and, in case the minimum exists, can also obtain some set, the limit set denoted here by L, of points at which the minimum is attained. To complete the discussion, we need to prove that L is always a nonempty set and moreover is finite.

Proposition 3.5 *The set L is finite.*

Proof
It is known that $\forall\lambda > 0,\ q_\lambda$ has at most N critical points (see Theorem 2.1). Therefore the cardinality of X_λ is also bounded by N. We will show that L has at most N points. Suppose that L has more that N distinct points and consider $N + 1$ of them $l_1, \ldots, l_{N+1}$. Let $\delta > 0$ denote the smallest distance between any two of these points. For every $i = 1, \ldots, N + 1$ construct the pairwise disjoint balls $B(l_i, \delta/2)$. By definition of L we have that there exists a $\lambda_{\delta/2} > 0$ such that every $B(l_i, \delta/2)$ has a nonempty intersection with X_λ, for each $\lambda \in (0, \lambda_{\delta/2})$. But for every $\lambda > 0$ X_λ has at most N elements, hence for each $\lambda \in (0, \lambda_{\delta/2})$, N elements must belong to $N + 1$ disjoint balls which is impossible.
Therefore L has at most N points. □

For our purposes, the non-emptiness is the most interesting part. In this way we have a guarantee that at least one point of global minimum is *always* obtained with our procedure.

Proposition 3.6 *If the polynomial p has a minimum, then L is nonempty.*

The proof of this proposition is given in the next section.

So far we have shown that with this method we can find the minimum of every polynomial and some of the points in which the minimum is attained. In general one cannot find all such critical points, especially when their number is infinite. One may wonder then which points we *do* find and the answer is partially given in the next proposition.

Proposition 3.7 *The set L is a subset of* $\{x_0 \mid \|x_0\| = \min_{\{x \ \mid\ p(x)=\min p(\tilde{x})\}} \|x\|\}$.

Proof
i) Let x_* be a point where the minimum of p is attained of minimal Minkowski norm. We prove that $\|x_\lambda\| \le \|x_*\|,\ \forall\lambda > 0, \forall\ x_\lambda \in X_\lambda$.

As p has a minimum, the norm $\|x_*\|$ is finite. From

$$q_\lambda(x_\lambda) = p(x_\lambda) + \lambda\|x_\lambda\|^{2m},\ q_\lambda(x_*) = p(x_*) + \lambda\ \|x_*\|^{2m}$$

and $q_\lambda(x_\lambda) \le q_\lambda(x_*)$ we have

$$\lambda\ [\|x_\lambda\|^{2m} - \|x_*\|^{2m}]\ \le\ p(x_*) - p(x_\lambda)\ \le\ 0$$

and therefore $\|x_\lambda\| \ \leq \ \|x_*\|, \ \forall\lambda > 0$.

ii) Suppose L is a non-empty set, otherwise the result is trivial. As the norm is a continuous function, using the result of part i) we have

$$\forall x \in L \ \ , \ \ \|x\| \ = \ \|\lim_{\lambda \downarrow 0} x_\lambda\| \leq \|x_*\|$$

But $\forall x \in L$ we have from Theorem 3.4, part (iii) that $\|x\| \geq \|x_*\|$. Hence $\|x\| \ = \ \|x_*\|$ which implies $x \in \{x_0 \mid \|x_0\| = \min_{\{x \ | \ p(x)=\min p(\tilde{x})\}} \|x\|\}$ for every $x \in L$, or else $L \subseteq \{x_0 \mid \|x_0\| = \min_{\{x \ | \ p(x)=\min p(\tilde{x})\}} \|x\|\}$. □

Denote by X the multi-valued function defined on $(0, \lambda)$ which associates to each λ the set X_λ. To give more insight into the properties of the branches of X, we prove their monotonicity. However, it has no relevance for our purposes so it may be skipped.

Proposition 3.8 *The multi-valued function X satisfies:*
$\forall\lambda_1, \lambda_2$ with $0 < \lambda_1 < \lambda_2$ and $\forall x_{\lambda_1} \in X_{\lambda_1}, x'_{\lambda_2} \in X_{\lambda_2}$ we have

$$\|x_{\lambda 1}\| \geq \|x'_{\lambda_2}\|$$

In particular, for one branch ($x = x'$) the lemma tells that the branch is monotonously decreasing with respect to λ in Minkowski norm.

Proof
Given $\lambda_1 < \lambda_2$ we have

$$\begin{cases} q_{\lambda_1}(x_{\lambda_1}) \leq q_{\lambda_1}(x'_{\lambda_2}) \\ q_{\lambda_2}(x'_{\lambda_2}) \leq q_{\lambda_2}(x_{\lambda_1}) \end{cases}$$

or equivalently

$$\begin{cases} p(x_{\lambda_1}) + \lambda_1\|x_{\lambda_1}\|^{2m} - p(x'_{\lambda_2}) - \lambda_1\|x'_{\lambda_2}\|^{2m} \leq 0 \\ p(x'_{\lambda_2}) + \lambda_2\|x'_{\lambda_2}\|^{2m} - p(x_{\lambda_1}) - \lambda_2\|x_{\lambda_1}\|^{2m} \leq 0 \end{cases}$$

By adding the two inequalities we obtain

$$(\lambda_1 - \lambda_2)(\|x_{\lambda_1}\|^{2m} - \|x'_{\lambda_2}\|^{2m}) \leq 0$$

which implies $\|x_{\lambda 1}\| \geq \|x'_{\lambda_2}\|$ □

To summarize, we have constructed a family of polynomials q_λ, such that the infimum of our initial polynomial p can be obtained from the minima of the polynomials in the family, by letting the parameter λ to tend to 0. If the original polynomial has a minimum, the method will find at least one point in which the minimum is attained. We also have the Stetter-Möller method for solving our system of first order equations which is by construction a reduced Gröbner basis. Hence, we need to compute the limits of the eigenvalues of a matrix A_λ associated to the polynomial q_λ for λ going to 0.

In the following section, we will propose a method for computing these limits.

4 Computing the minimum

From the previous section we know that we can find the minimum of the original polynomial p by computing the limits when λ goes to 0 of the eigenvalues of the matrix A_λ.

Before discussing the computation of the matrix A_λ, associated to q_λ, we would like to stress that all the results obtained for A_λ remain valid for any matrix A_f associated to a polynomial $f \in \mathbf{C}[x_1, \ldots, x_n]$.

Proposition 4.1 *The matrix A_λ is polynomial in $1/\lambda$.*

Proof
The proof goes by induction on the number of reduction steps. Recall that our Gröbner basis has a particular form in which the leading monomials are pure powers of the variables and λ appears only in the leading monomial. Hence we start with constant entries but, due to the particular form of the Gröbner basis, whenever we make a reduction step ([5]), we introduce a $1/\lambda$ or a power of it in every entry. Therefore, the entries of the final matrix will be polynomials in $1/\lambda$. □

Remark that the size of A_λ is given by the dimension of the basis B consisting of all monomials in the variables $x_1, \ldots, x_n$ in which the exponent of every variable is an integer number, larger or equal to 0 and strictly smaller that $2m$. Hence the dimension of B, and therefore the size of A_λ, is $N = (2m)^n$.

We want to find the limits, when λ decreases to 0, of the eigenvalues of A_λ, i.e. the solutions of the equation $\det(A_\lambda - zI) = 0$. Recall the interpretation of the eigenvalues in the Stetter-Möller method. They represent the values of the polynomial q_λ in the critical points. Hence the minimum of q_λ will be in the set of eigenvalues. The eigenvectors of A_λ will give the corresponding points and their limits for $\lambda \downarrow 0$ will allow us to read off a critical point where the minimum is attained. However if the critical point set is not finite we are not able in general to find the whole set, but we find a finite subset of it.

The equation

$$\det(A_\lambda - zI) = 0 \quad , \quad \lambda > 0 \quad , \quad z \in \mathbf{C}$$

is satisfied if and only if

$$\lambda^k \det(A_\lambda - zI) = 0 \quad , \quad \lambda > 0 \quad , \quad z \in \mathbf{C} \tag{2}$$

where k is the highest power of $1/\lambda$ appearing in the determinant. The second equation, polynomial in both z and λ, was studied extensively in the literature. Its solutions $z(\lambda)$ which satisfy the equation for every positive λ are known as *algebraic functions* (see [2]). An algebraic function is a multi-valued function having a finite number of branches $\zeta_i(\lambda)$, $i = 1, \ldots N$. The values of each branch around an arbitrary $\lambda_0 \geq 0$ are given by a Puiseux expansion in rational powers of $\lambda - \lambda_0$. To be more precise, the following proposition holds.

Proposition 4.2 *In a neighborhood V of every finite point $\lambda = \lambda_0$ all values of an algebraic function $z(\lambda)$ are determined by branches of the form*

$$\lambda = \lambda_0 + t^r \quad , \qquad z = z_{-\kappa} t^{-\kappa} + z_{-\kappa+1} t^{-\kappa+1} + \ldots + z_0 + z_1 t + \ldots \tag{3}$$

in which r is a positive integer, the coefficients $z_{-\kappa}, z_{-\kappa+1}, \ldots$ indicated are complex, possibly zero, and κ is a non-negative integer. For a value $\lambda \neq \lambda_0$ in V, (3) determines r distinct values of $z(\lambda)$ when the r values of the root $t = (\lambda - \lambda_0)^{1/r}$ are substituted in the series for z.

Proof of Proposition 3.6
In the definition of L, X_λ denotes the set of real points where the minimum of q_λ is attained. To show that L is nonempty it is enough to prove that X_λ is continuous on branches on an interval $(0, \lambda_0)$ for λ_0 sufficiently small. For that, we refer to Stetter-Möller theory. It is known that the coordinates of the point in $\mathbf{R}^N$ where the minimum of q_λ is attained, i.e. the coordinates of X_λ, can be obtained as the eigenvalues of the matrices A_{x_i} for $i = 1, \ldots, n$, where A_{x_i} denotes the linear mapping associated to the polynomial x_i (see section 2).
From Proposition 4.1 we have that the matrices A_{x_i} are polynomial matrices in $1/\lambda$. So, the eigenvalues of A_{x_i} are the solutions of the equation in x $\det(A_{x_i} - xI) = 0$ or equivalently, $\lambda^k \det(A_{x_i} - xI) = 0$

where k is the highest power of $1/\lambda$ appearing in the determinant. As the equation is polynomial in x and λ, the solutions $X_i(\lambda)$ are algebraic functions. For every fixed positive λ the algebraic functions admit in a neighborhood of λ an expansion in which radicals or (a finite number of) terms with negative exponent may be involved (see Theorem 4.2) This implies in particular that the branches of X_i as functions of λ are continuous in a right neighborhood of $\lambda = 0$. Since $X_i(\lambda)$ are coordinates of X_λ, then also X_λ is continuous in a right neighborhood.
Next we argue that when p has a minimum, there will be a branch of X_λ which does not contain negative powers of λ in its expansion. As p has a minimum, there exists a finite point in which the minimum is attained. We know that the branches of X_λ are bounded in the Minkowski norm by such a finite point (see Theorem 3.7, first part of the proof). Hence X_λ will have finite limits on the branches when $\lambda \downarrow 0$ and all these limits belong to the limit set L which is therefore nonempty. □

Recall that we want to compute the limits of the branches when $\lambda \downarrow 0$ so in our case $\lambda_0 = 0$ and V is the neighborhood of 0. The expansion of a branch of an algebraic function may have a finite number of terms containing negative powers of λ. We say that a branch has an *infinite limit* when $\lambda \downarrow 0$ if its expansion contains negative powers of λ. Otherwise we say that it has *finite limit*. The branches that have finite limits will tend, when $\lambda \downarrow 0$, to z_0, the term of the expansion which does not depend on λ.

Let

$$\det(A_\lambda - zI) = f(\lambda, z) = 1/\lambda^k f_0(z) + 1/\lambda^{k-1} f_1(z) + \ldots + f_k(z).$$

where $f_0, f_1, \ldots, f_k$ are polynomials in z. Then equation (2) becomes

$$f_0(z) + \lambda f_1(z) + \ldots + \lambda^k f_k(z) = 0$$

We can easily see from Proposition 4.2 that the finite limits are solutions of the equation $f_0(z) = 0$. In fact one can show a bit more.

Proposition 4.3 *The critical values of the polynomial q_λ define a finite number of branches having, when $\lambda \downarrow 0$, finite or infinite limits. The set of finite limits of q_λ coincides with the set of solutions of $f_0(z) = 0$.*

Proof
The first part of the theorem was already discussed.
For the last part, consider $\zeta(\lambda)$ a branch having a finite limit. By replacing $\zeta(\lambda)$ by its expansion, one can easily see that the lambda-free term in the expansion, is a solution of $f_0(z) = 0$. Hence the number of branches having a finite limit is at most equal to the degree of f_0, denoted by d. We'll show that in fact the equality holds, hence the two sets must be equal. For this purpose we consider next the branches having infinite limits, i.e. their expansion contains negative powers of λ. Let $\zeta(\lambda)$ be a solution of (2) whose expansion contains negative powers of λ. Then $\omega(\lambda) = 1/\zeta(\lambda)$ is a solution of the equation $f(\lambda, 1/w) = 0$ or equivalently

$$w^N f(\lambda, 1/w) = 0. \tag{4}$$

Note that the second equation was obtained by bringing the terms in $f(\lambda, 1/w)$ to the common denominator w^N and taking afterwards the nominator equal to 0. Remark that $\lim_{\lambda\downarrow 0} \omega(\lambda) = 0$ as can be seen for example from the expansion of $\zeta(\lambda)$. Hence $\omega(\lambda)$ is solution of the polynomial equation (4) and having limit 0 is a finite solution of the equation. Rewriting the equation (4) we have

$$w^N[f_0(1/w) + \lambda f_1(1/w) + \ldots + \lambda^k f_k(1/w)] = 0$$

and we need to compute the number of branches $w(\lambda)$ that tend to 0 when $\lambda \downarrow 0$. But as we have argued before, every 0 limit of a branch of $w(\lambda)$ is a root of the λ-free term, $w^N f_0(1/w)$. But $w^N f_0(1/w)$ has exactly $N-d$ zero roots, where d was the degree of f_0. Hence the number of branches of $w(\lambda)$ having the limit 0, which equals the number of branches of $z(\lambda)$ having infinite limits, is at most $N-d$. To conclude, we have exactly N branches having either finite or infinite limit and we have shown that among them at most d have finite limits and at most $N-d$ have infinite limits. Hence the inequalities must be in fact equalities.□

Lemma 4.4 *Suppose that p has a minimum and $\underline{x}$ is an isolated point of minimum of the polynomial p. There exists a branch x_λ of (local) minima of q_λ convergent to $\underline{x}$ for $\lambda \downarrow 0$.*

Proof
As p is a polynomial and $\underline{x}$ is an isolated point of (global) minimum, there exists a convex neighborhood V of $\underline{x}$, where p is strictly convex. The function $x_1^{2m} + x_2^{2m} + \ldots + x_n^{2m}$ is strictly convex on R^n. It follows immediately that for every $\lambda > 0$, q_λ is strictly convex in V.
Let $\varepsilon > 0$ such that $B(\underline{x}, \varepsilon) \subset V$. Next, we show that the unique point of minimum of q_λ on $\overline{B(\underline{x}, \varepsilon)}$ is, for every λ sufficiently small, a point of $B(\underline{x}, \varepsilon)$.
Remark that $\min_{x \in \partial B(\underline{x},\varepsilon)} p(x) > \min_{x \in B(\underline{x},\varepsilon)} p(x) = p(\underline{x})$. From the above inequality and Theorem 3.4, part (i), we have

$$\exists \lambda_\varepsilon > 0 \text{ such that } \forall \lambda, 0 < \lambda < \lambda_\varepsilon \ , \quad q_\lambda(\underline{x}) < \min_{x \in \partial B(\underline{x},\varepsilon)} p(x) \leq \min_{x \in \partial B(\underline{x},\varepsilon)} q_\lambda(x).$$

That implies that the minimum of q_λ does not lie on the border. Hence, for every $\lambda < \lambda_\varepsilon$ we have a unique $x_\lambda \in B(\underline{x}, \varepsilon)$ such that $q_\lambda(x_\lambda) = \min_{x \in B(\underline{x},\varepsilon)} q_\lambda$. In other words, x_λ is a local minimum of q_λ and $(x_\lambda)_{\lambda > 0}$ a branch of local minima of q_λ, convergent to $\underline{x}$. □

Theorem 4.5 *If p has a minimum, then the algorithm finds all isolated points of (global) minimum of the polynomial p. In particular, if p has a finite number of points of minimum, the algorithm finds them all.*

Proof
The algorithm computes the solutions of equation (2), where A_λ may be the matrix associated to q_λ. On the other hand, the solutions are finite limits of local extrema of q_λ. Hence we also find the local extrema of q_λ, convergent to the global minimum of p. Remark that, if p has a finite number of points of minimum, they are all isolated. □

Theorem 4.6 *If p has a minimum then the set $p^{-1}(\{\min_{x \in \mathbf{R}^n} p(x)\})$ consists of one or more connected components. The algorithm finds at least one point in each connected component. In fact these are points having minimal Minkowski norm inside the component .*

Proof
Note that the number of connected components of $p^{-1}(\{p_{min}\})$ is finite (see [3], Th 2.4.5), where $p_{min} = \min_{x \in \mathbf{R}^n} p(x)$.
Pick a point, say $x(j)$, in each component C_j, where

$$C = \bigcup_{j \in J} C_j = \{x \in \mathbf{R}^n \mid p(x) = p_{min}\}.$$

Let $M_j = \|x(j)\|$ and $M > \max_{j \in J} M_j$.
We want to show that for every $j \in J$, there will be a local minimum of q_λ whose points of minimum are in the Minkowski ball $B(0, M)$ and converge to an element of C_j. If this holds, then from the

local minima of q_λ we obtain at least one point in each component C_j.
Note that in each component C_j there is a point, namely $x(j)$, such that

$$q_\lambda(x(j)) < p_{min} + \lambda M^{2m} \leq q_\lambda(x),\ \forall x \notin B(0,M).$$

Hence

$$q_\lambda(x(j)) < q_\lambda(x),\ \forall x \notin B(0,M)$$

and the local minima of q_λ corresponding to every component C_j, provided they exist, are in the Minkowski ball $B(0,M)$.
Consider $q_\lambda\Big|_{\overline{B(0,M)}}$. The number of connected components of $C \bigcap \overline{B(0,M)}$ is still finite since the set $\{x \in \mathbf{R}^n \mid \|x\|^{2m} \leq M^{2m}, p(x) = p_{min}\}$ is a semi-algebraic set ([3], Th 2.4.5). Denote them by D_l.
Since $\overline{B(0,M)}$ is a compact set and the sets D_l are disjoint, it follows that $\exists \varepsilon_0 > 0$ such that $\forall l_1 \neq l_2$ $d(D_{l_1}, D_{l_2}) > \varepsilon_0$, where d denotes the Minkowski distance between sets.
Define the neighborhood of a component D_l as

$$N_{\varepsilon_0/3}(D_l) = \{x \in B(0,M) |\ d(x, D_l) < \varepsilon_0/3\}.$$

We want to show that the minimum of $q_\lambda\Big|_{\overline{N_{\varepsilon_0/3}(D_l)}}$ is not attained on the border of $\overline{N_{\varepsilon_0/3}(D_l)}$. Note that any point on the border satisfies one of the relations $\|x\| = M$ or $d(x, D_l) = \varepsilon_0/3$. We already know that the points on the border of $B(0,M)$ are not local minima of q_λ.
Let $\bar{p} = \min_{\bigcup_l (\partial N_{\varepsilon_0/3}(D_l) \bigcap B(0,M))} p(x)$. Then $\bar{p} > p_{min}$. We have $q_\lambda\Big|_{\partial N_{\varepsilon_0/3}(D_l) \cap B(0,M)} \geq \bar{p}$.
On the other hand, for any $x \in D_l$ we have $q_\lambda(x) = p_{min} + \lambda\|x\|^{2m} \leq p_{min} + \lambda M^{2m} < \bar{p}$ for λ sufficiently small, namely $\lambda < (\bar{p} - p_{min})/M^{2m}$. Therefore, if $\lambda < (\bar{p} - p_{min})/M^{2m}$ then $\min_{x \in \overline{N_{\varepsilon_0/3}(D_l)}} q_\lambda$ is attained in the open set, not on the boundary.
We have proved that for λ smaller than a certain value, for every component D_l there exists an open neighborhood of it containing the points of local minimum of $q_\lambda\Big|_{\overline{B(0,M)}}$.
Let x_λ^l be a global minimizer of $q_\lambda\Big|_{N_{\varepsilon_0/3}(D_l)}$. Then x_λ^l is a local minimizer of q_λ (on R^n). Since x_λ^l is local minimizer, it is convergent as in Proposition 3.6 to a point, say $x_* \in \overline{N_{\varepsilon_0/3}(D_l)}$.
We want to show that $x_* \in D_l$.
We have $p(x) \leq q_\lambda(x)$ and $\lim_{\lambda \downarrow 0} q_\lambda(x) = p(x),\ \forall x \in R^n$. Hence $p(x_\lambda^l) \leq q_\lambda(x_\lambda^l) \leq q_\lambda(x_*)$. When $\lambda \downarrow 0$ we obtain $\lim_{\lambda \downarrow 0} q_\lambda(x_\lambda^l) = p(x_*)$.
Take $x_0 \in D_l$. We have $q_\lambda(x_\lambda^l) \leq q_\lambda(x_0)$ and at the limit it becomes $p(x_*) \leq p_{min}$ or in fact $p(x_*) = p_{min}$. This implies that $x_* \in D_l$. $\square$

4.1 Case: the polynomial p has a minimum

From Theorem 3.4 we know that $\min_{x \in R^n} q_\lambda(x) = q_\lambda(x_\lambda)$ converges to $\min_{x \in R^n} p(x)$. But $q_\lambda(x_\lambda)$ satisfies the equation (2), so it is a branch of the algebraic function associated to the equation. Moreover, we know it has a finite limit. Hence $\lim_{\lambda \downarrow 0} q_\lambda(x_\lambda)$ will be a root of f_0. The smallest real root is our candidate for the minimum of p. Note that we have been working over the field of complex numbers and it is possible that the smallest real root is a value of p attained in a complex point. Hence, before deciding that the smallest real root is the minimum of p, we need to do a check on the point where the minimum is attained. We'll discuss this issue later, but until then, in order to make the discussion easier, we'll assume that the smallest real eigenvalue is indeed the minimum.

The way to compute $\min_{x\in R^n} p(x)$ becomes more clear now. Having constructed the matrix A_λ, one can calculate $\det(A_\lambda - zI)$ polynomial in $1/\lambda$ and z, then isolate the coefficient of the largest power of $1/\lambda$. This is a polynomial in z whose smallest real root gives us the minimum of p.

We have now a straightforward way to compute the minimum of our polynomial p. However, the drawback of using the determinant is, besides the high computational complexity, that it will not tell us anything about the corresponding eigenvectors. As we already remarked, knowing the eigenvectors may be helpful in finding not only the minimum but also (at least) a point in which the minimum is attained. Hence we need a more "sophisticated" method for the actual calculations.
We describe here a method for computing the finite limits of the eigenvalues, without actually computing the determinant. It will be clear that with this new method, we can not only find the corresponding eigenvectors but also we do less calculations, as we only need one term in the determinant.

The method is the well-known algorithm of Forney ([8]) for minimizing the sum of the row degrees of a polynomial matrix over an equivalence class of polynomial matrices. With this method we obtain the coefficient of the highest power of $1/\lambda$ in the expression of the determinant $\det(A_\lambda - zI)$ as a polynomial matrix in z. Then, according to Proposition 4.3, the minimal value of p is among the values of z for which this new matrix becomes singular.
After applying linearization techniques (see [9], § 7.2) we reduce it to the problem of finding the eigenvalues of a pencil. Since the original matrix is nonsingular and the linearization procedure leaves the determinant unchanged, the generalized eigenvalue problem obtained is always nonsingular.

Remark that the problem of finding the minimum of a polynomial and some point where this is attained is reduced to solving a generalized eigenvalue problem. For this new problem, a large variety of algorithms exist and these can handle quite large matrices.

Let us describe shortly how to find the coefficient of the highest power of $1/\lambda$ in the expression of the determinant $\det(A_\lambda - zI)$.

Let us denote for simplicity $\mu = 1/\lambda$. We first introduce few notions. Let B_μ be a polynomial matrix in μ.
The degree of the i-th row, denoted d_i, is the highest degree in μ of all its entries. The total row degree of the matrix is the sum of its row degrees.
The associated high order coefficient matrix, denoted HOCM, is constructed by retaining from each entry of the i-th row, the coefficient of μ^{d_i}.
The algorithm for finding the leading term of $\det(B_\mu)$, i.e. the term containing the highest power of μ in the expression of the determinant $\det(B_\mu)$, is based on the following:

Proposition 4.7 *Let B_μ be a polynomial matrix in μ. If its* HOCM *is nonsingular, then the leading term of* $\det(B_\mu)$ *is* $\det(\text{HOCM})\mu^d$*, where d is the total row degree of B_μ.*

Proof
From Cramer's formula for computing determinants we know that the degree of $\det(B_\mu)$ cannot be larger than d. If HOCM is nonsingular, then the term $\det(\text{HOCM})\mu^d$ appears in the determinant. There is no other term having the same power of μ since if one row contains an element of degree strictly smaller than the row degree, all the terms of $\det(B_\mu)$ which contain this element will have the exponent strictly smaller than d. □

We can now give the algorithm for finding the leading coefficient of $\det(A_\mu - zI)$, seen as polynomial in μ. Note that by construction the total row degree of $(A_\mu - zI)^T$ is in general much smaller than the total row degree of $A_\mu - zI$. Therefore, for computational reasons, we work with $(A_\mu - zI)^T$.

Algorithm 4.8 *The following procedure returns a matrix, polynomial in μ and rational in z, of minimal total row degree, equivalent to the input matrix $(A_\mu - zI)^T$.*

Step 1. Input: $B_\mu \leftarrow (A_\mu - zI)^T, \Delta \leftarrow 1$

Step 2. If $\mathrm{HOCM}(B_\mu)$ *is nonsingular, then go to Step 7.*

Step 3. Else compute a nonzero vector $v = (v_1, \ldots, v_N)$ *in the left kernel of* $\mathrm{HOCM}(B_\mu)$. *The vector can be chosen polynomial in z.*

Step 4. Construct the vector $\tilde{v} = (v_1\mu^{d_*-d_1}, \ldots, v_N\mu^{d_*-d_N})$, *where* $d_* = \max_{i=1,\ldots,N} d_i$.

Step 5. Construct a matrix $L_{\mu,z}$ from the identity matrix by replacing its i-th row by $\tilde{v}$, where i is chosen such that $d_i = d_*$.

Step 6. $B_\mu \leftarrow L_{\mu,z} B_\mu$, $\Delta \leftarrow \Delta \det(L_{\mu,z})$. *Go to Step 2.*

Step 7. Output: B_μ, *with* $\mathrm{HOCM}((A_\mu - zI)^T) = \frac{1}{\Delta}\,\mathrm{HOCM}(B_\mu)$

To show that the algorithm finishes after a finite number of steps, remark that at Step 5 the total row degree of the matrix is decreased by 1 at least. As $A_\mu - zI$ is nonsingular, i.e. its determinant is non-identically zero, the power of μ in the leading term is nonnegative and the algorithm stops when it reaches this value.
Remark that $\mathrm{HOCM}(B_\mu)$ is polynomial matrix in z hence a vector as in Step 3 always exists. Remark also that the determinant of $L_{\mu,z}$ from Step 5 does not depend on μ. It may depend on z, therefore we need the corrections Δ. Matrices like $L_{\mu,z}$ depending on a parameter μ, whose determinant does not depend on μ are called z-modular or unimodular over $\mathbf{R}[z]$.
Since at Step 6 we multiply with z-modular matrices, our HOCM may become polynomial, not linear, in z. The nonsingular polynomial matrix in z can be brought by a linearization procedure (see [9], § 7.2) into an equivalent matrix, linear in z of a larger dimension. Note however that in the reduction process while multiplying on the left with z-modular matrices we introduce some new solutions. Hence we must keep track of the solutions we introduce and subtract them in the end.
To be more precise, after running the algorithm we have

$$R_{\mu,z} = L_{\mu,z}(A_\mu - zI)^T \ ,$$

where $R_{\mu,z}$ has a nonsingular HOCM and $L_{\mu,z}$ is z-modular. For their determinants, the following holds:

$$\det(R_{\mu,z}) = \det(L_{\mu,z})\det(A_\mu - zI)$$

and using Proposition 4.7 and the fact that $\det(L_{\mu,z})$, which equals our final value of Δ in the algorithm, does not depend on μ it follows that the leading term in μ of $\det(A_\mu - zI)$ satisfies

$$\mathrm{lt}(\det(A_\mu - zI)) = (\det(L_{\mu,z}))^{-1}\ \det(\mathrm{HOCM}(R_{\mu,z})).$$

The roots of $\det(L_{\mu,z})$ are artificially introduced so we must eliminate them.

The algorithm can be applied in general for finding a minimal total row degree, (left-)equivalent representation of a matrix. In the following we give a small example to illustrate how the algorithm works.

Example 4.9 Consider a matrix M_μ, polynomial in μ, of non-minimal total row degree. M_μ plays the role of A_μ, the difference being that M_μ is not associated to a polynomial. Let

$$M_\mu = \begin{pmatrix} \mu^2 & 0 & \mu \\ 1 & 0 & -2 \\ \mu^3 & \mu & \mu^2 \end{pmatrix}.$$

The matrix $B_\mu = M_\mu - zI$ becomes

$$B_\mu = \begin{pmatrix} \mu^2 - z & 1 & \mu^3 \\ 0 & -z & \mu \\ \mu & -2 & \mu^2 - z \end{pmatrix}$$

with the row degree vector $(3, 1, 2)$, hence the total row degree 6. However its HOCM is singular,

$$\mathrm{HOCM}(B_\mu) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

hence its total row degree is not minimal. Pick up a vector in the left kernel of $\mathrm{HOCM}(B_\mu)$, say $v = (-1, 1, 0)$ and construct $\tilde{v} = (-1, \mu^2, 0)$. The matrix $L_{\mu,z}$ becomes

$$L_{\mu,z} \leftarrow \begin{pmatrix} -1 & \mu^2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and by multiplication on the right with B_μ,

$$B_\mu \leftarrow \begin{pmatrix} -\mu^2 + z & -1 - z\mu^2 & 0 \\ 0 & -z & \mu \\ \mu & -2 & \mu^2 - z \end{pmatrix} \quad \text{and} \quad \Delta \leftarrow -1.$$

Since the new matrix has a singular HOCM, we return to Step 3 and continue the reduction procedure. Hence

$$B_\mu \leftarrow \begin{pmatrix} -\mu^2 + z & -1 - z\mu^2 & 0 \\ 0 & -z & \mu \\ \mu & \mu z - 2 & -z \end{pmatrix} \quad \text{and} \quad \Delta \leftarrow -1.$$

Remark that another reduction step is necessary and finally we obtain

$$B_\mu = \begin{pmatrix} z & -2\mu - 1 & -\mu z \\ 0 & -z & \mu \\ \mu & \mu z - 2 & -z \end{pmatrix}$$

whose HOCM is nonsingular. Remark that the determinant of this matrix is $-2\mu^2 z^2 + z^3 + 2z\mu - 2\mu^3 - \mu^2$ and it is equal to $\Delta \det(M_\mu - zI)$. In this example, the total row degree was reduced from 6 to the minimal row degree which is 3.

In general, when Δ depends on z we introduce false solutions for $\det(A_\mu - zI)$ during the reduction procedure.
An improvement on the algorithm would be to avoid introducing such solutions or if we do, to eliminate them in a smarter way. The problem reduces basically to the following one: Having a polynomial matrix $B(z)$ and a polynomial in $b(z)$ which divides its determinant, find a polynomial matrix whose determinant is $\det B(z)/b(z)$. Obviously, such a matrix exists as well as an algorithm to compute it. The question is whether we can compute such a matrix in an efficient way.

Note that the eigenvectors of the polynomial matrix HOCM preserve the property of the Stetter vectors. Namely when the eigenspace is 1-dimensional, the eigenvector is the basis vector evaluated at the critical point.

This follows from the fact that we multiply the matrix $(A_\lambda - zI)^T$ only on the left-hand side, hence its Stetter eigenvectors are preserved. In the end we obtain,

$$L_{1/\lambda,z} A_{1/\lambda,z} v_\lambda = 0, \ \forall \lambda > 0.$$

By premultiplying with $\mathtt{diag}(\lambda^{d_1}, \ldots, \lambda^{d_N})$, where d_j is the (minimal) row degree of row j we obtain a N-dimensional equation in λ, valid for every $\lambda > 0$ and well-defined in $\lambda = 0$. Then the equation must hold also for $\lambda = 0$, but that is exactly HOCM $\lim_{\lambda \downarrow 0} v_\lambda = 0$. That insures us that the eigenvectors of HOCM will indeed correspond to critical points of p.

4.2 Case: the polynomial p has an infimum

At this point we do not have a direct way of deciding whether the polynomial p has a minimum or not or, in the latter case, whether its infimum is bounded or not. However, the following procedure can in principle be used to decide this. Compute the candidate for the minimum by running the algorithm described above. Let us denote the obtained value by c. Then form the polynomial $(p - c + \alpha)^2$, with $\alpha > 0$ and run the algorithm again. If c was indeed the minimum of p, then the new polynomial we must have minimum α^2. If there are values of p strictly smaller than c, then due to the continuity of p there must exist a point x such that $p(x) = c - \alpha$. Hence the new polynomial will have the minimum equal to 0. Further research can be done in the direction of finding a direct way to decide upon this matter. Note that Theorem 3.4 suggests that this might be possible if we use a different algorithm for the actual computations.

5 Example

We consider here a rather small example. Let

$$p(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2.$$

It is easy to see that the minimum of p is zero and it is attained for all the points of the circle of radius 1, centered in the origin. There are few reasons for our choice. The first one is that the method we have proposed requires a number of calculations that increases rapidly with the degree of the polynomial and the number of variables. The second, and more important reason, is that in this case we already know the minimum and the set of points where it is attained, therefore it is possible to analyze the algorithm in this specific example.

First we construct the family of polynomials

$$q_\lambda = (x_1^2 + x_2^2 - 1)^2 + \lambda(x_1^6 + x_2^6).$$

The power in the extra-term was chosen to be an even number, strictly larger than 4, the total degree of p.
Next we construct using the Stetter-Möller method, the matrix A_λ whose eigenvalues are the critical values of q_λ. The matrix can be seen in the appendix.
One can easily see that the total row degree of the matrix equals 44. However it is not minimal, i.e. the highest power of $1/\lambda$ appearing in the determinant of $A_\lambda - zI$ is actually smaller than 44. This, of course, can be seen by running the total row degree reduction algorithm of Forney on $A_\lambda - zI$ which will return the matrix $\bar{A}_\lambda(z)$ having the total row degree minimal, equal to 16 (see appendix). At this point we have also obtained the coefficient of the highest power of $1/\lambda$ in the expression $\det(A_\lambda - zI)$. This is the determinant of the HOCM of $\bar{A}_\lambda(z)$, divided by the polynomial that we have introduced in the reduction process. In this particular case, the polynomial is $(z-1)^4$. Their ratio is

$$-z^8 \,(z-1)^9$$

and the roots of this polynomial are the finite limits we were looking for. The solution set is $S = \{0, 1\}$, 0 with multiplicity 8, and 1 with multiplicity 9. As in the solution set, 0 is the smallest, it will be our candidate for the minimum of p. Note however that the multiplicity of 0 is strictly

larger than 1, hence we can not decide yet whether this is indeed the minimum. It could still be that the value 0 of p is obtained for some complex values of x. If the multiplicity of the smallest real value (actually eigenvalue of HOCM of $\bar{A}_\lambda(z)$) is 1, than we know that there is a real eigenvector corresponding to it, eigenvector that gives us the value of the critical point. But in this case, before concluding that $\min_{x\in R^n} p(x) = 0$, we need to compute the points where zero is attained and check that they are real numbers.
For that we consider the matrices associated to x_1 and x_2 and repeat the algorithm. Thus, we obtain for both x_1 and x_2 the values $\{0, -1, 1, 1/2\sqrt{2}, -1/2\sqrt{2}\}$ with respective multiplicities 5,1,1,2,2. There exist a combination of the values of x_1 and the values of x_2 such that, by computing the value of p at these points we obtain the set S. By inspection we conclude that the value 0 of p is obtained for

$$\{(\pm 1/2\sqrt{2}, \pm 1/2\sqrt{2}), (\pm 1, 0), (0, \pm 1)\}.$$

They are real hence the minimum of p is indeed 0. Remark that the values $(\pm 1/2\sqrt{2}, \pm 1/2\sqrt{2})$ are points where the minimum of p is attained, of minimal Minkowski norm. This was predicted in Proposition 3.7. However we obtain some extra points which in this case are points of maximal Minkowski norm. It is an open question whether we find points of maximal Minkowski norm in every connected component whenever the component is bounded.

6 Conclusions

The proposed method is guaranteed to find the global minimum of a general polynomial. Moreover, if the minimum does not exist, we can decide if the infimum is finite or not, and give its value in the first case. To the best of our knowledge this problem did not receive until now a solution in the general case.
The approach translates the original problem into a generalized eigenvalue problem. This may open up the possibility for numerical calculations.

Another very important feature of the algorithm is that it returns a point in every connected component of the set of (global) minimizers. Using the algorithm we can in fact answer a different problem as well. Given a set of polynomial equations $f_i(x_1, \ldots, x_n) = 0,\ i = 1, \ldots s$, we can find a point in every connected component of the solution set as described in the introduction. Such problems received a lot of attention (see [1] and the references contained therein).

References

[1] S. Basu, R. Pollack, M-F. Roy, *A new algorithm to find a point in every cell defined by a family of polynomials,* in: B.F. Caviness,J.R. Johnson(eds), *Quantifier Elimination and Cylindrical Algebraic Decomposition,* 1998.

[2] G.A. Bliss, *Algebraic functions*, American mathematical society, Colloquium Publications, vol. XVI, 1993

[3] J. Bochnak, M. Coste, M-F. Roy, *Géométry algébrique réelle*, Springer-Verlag, 1987

[4] D. Cox, J. Little, D. O'Shea, *Ideals, Varieties, and Algorithms,* 2nd ed., Springer, New York, 1997.

[5] D. Cox, J. Little, D. O'Shea, *Using algebraic geometry*, Springer-Verlag, New-York, 1998

[6] K.O. Geddes, S.R. Czapor, G. Labahn *Algorithms for computer algebra*, Kluwer Academic Publishers, 1992

[7] B.Hanzon, J.M.Maciejowski, C.T.Chou, *Model reduction in H_2 using matrix solutions of polynomial equations,* Cambridge University Engineering Department, Technical Report CUED/F-INFENG/TR314, 1998.

[8] G. D. Forney, *Minimal bases of rational vector spaces, with applications to multivariable linear systems*, SIAM J. Control, vol. 13, no.3, pp. 493-520, 1975

[9] I. Gohberg, P. Lancaster, L. Rodman *Matrix polynomials*, Academic Press, 1982

[10] K. Hägglöf, P.O. Lindberg, L. Stevenson, *Computing global minima to polynomial optimization problems using Gröbner bases*, Journal of global optimization, vol. 7, no.2 pp. 115-125, 1995

[11] H.Möller, *Systems of algebraic equations solved by means of endomorphisms,* in: Cohen, Mora, Moreno(eds), LNComp. Sc. 673, Springer Verlag, New York, pp. 43-56, 1993.

[12] H.M. Möller, H.J. Stetter, *Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems*, Numer. Math., vol.70, pp.311-329, 1995

[13] H.J.Stetter, *Matrix eigenproblems are at the heart of polynomials systems solving,* SIGSAM Bulletin, vol.30, pp.22-25, 1996.

[14] A.Y. Uteshev, T.M. Cherkasov, *Polynomial optimization problem*, Doklady Mathematics, vol. 58, no. 1, pp. 46-48, 1998

7 Appendix

The matrices A_λ, associated to the polynomial q_λ and $\bar{A}_\lambda(z)$, obtained after the running the Forney algorithm on $A_\lambda - zI$ are

$$
\left[\begin{array}{ccccccccccccccccccccccccc}
1 & 0 & -\frac{4}{3} & 0 & \frac{1}{3} & 0 & -\frac{4}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{9}\frac{2+9\lambda}{\lambda} & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{9}\frac{2+9\lambda}{\lambda} & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & 0 & 0 \\
0 & -\frac{4}{27}\frac{1+6\lambda}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+30\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+1}{\lambda^2} & 0 & -\frac{4}{9}\frac{1}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{2}{3}\frac{2\lambda+1}{\lambda} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\
0 & 0 & -\frac{4}{27}\frac{1+6\lambda}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+30\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+1}{\lambda^2} & 0 & -\frac{4}{9}\frac{1}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{2}{3}\frac{2\lambda+1}{\lambda} & 0 & \frac{1}{3} \\
0 & 0 & 0 & 0 & 0 & \frac{1}{9}\frac{2+9\lambda}{\lambda} & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{9}\frac{2+9\lambda}{\lambda} & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{2}{9}\frac{1+6\lambda}{\lambda} & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{4}{27}\frac{1+6\lambda}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+30\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+1}{\lambda^2} & 0 & -\frac{2}{3}\frac{2\lambda+1}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{4}{9}\frac{1}{\lambda} & 0 & \frac{1}{3} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{4}{27}\frac{1+6\lambda}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+30\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+1}{\lambda^2} & 0 & -\frac{2}{3}\frac{2\lambda+1}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{4}{9}\frac{1}{\lambda} & 0 & \frac{1}{3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{9}\frac{9\lambda+4}{\lambda} & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{9}\frac{9\lambda+4}{\lambda} & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{9}\frac{9\lambda+4}{\lambda} & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & \frac{2}{3} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{9}\frac{9\lambda+4}{\lambda} & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & 0 & 0 & 0 & 0 & -\frac{4}{9}\frac{1+3\lambda}{\lambda} & 0 & \frac{2}{3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{81}\frac{15\lambda+8}{\lambda^3} & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & 0 & 0 & 0 & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & \frac{1}{27}\frac{27\lambda^2+60\lambda+32}{\lambda^2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{81}\frac{15\lambda+8}{\lambda^3} & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & 0 & 0 & 0 & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & \frac{1}{27}\frac{27\lambda^2+60\lambda+32}{\lambda^2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{8}{27}\frac{3\lambda+2}{\lambda^2} & 0 & \frac{4}{27}\frac{9\lambda+4}{\lambda^2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{27}\frac{27\lambda^2+16+36\lambda}{\lambda^2} & 0 & -\frac{4}{9}\frac{3\lambda+2}{\lambda} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{81}\frac{15\lambda+8}{\lambda^3} & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & 0 & 0 & 0 & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & \frac{1}{27}\frac{27\lambda^2+60\lambda+32}{\lambda^2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{81}\frac{15\lambda+8}{\lambda^3} & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & 0 & 0 & 0 & 0 & -\frac{8}{81}\frac{9\lambda^2+8+21\lambda}{\lambda^3} & 0 & \frac{1}{27}\frac{27\lambda^2+60\lambda+32}{\lambda^2}
\end{array}\right]
$$

$$
\left[\begin{array}{l}
1-\varepsilon, 0, \frac{-4}{3}, 0, \frac{1}{3}, 0, \frac{-4}{3}, 0, \frac{1}{3}, 0, 0, 0, 0, 0, \frac{2}{3}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 \\
0, 1+\frac{2}{9}\frac{1}{\lambda}-\varepsilon, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, 0, 0, 0, 0, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, \frac{1}{3}, 0, 0, 0, 0, 0, \frac{2}{3}, 0, 0, 0, 0, 0, 0 \\
0, 0, 1+\frac{2}{9}\frac{1}{\lambda}-\varepsilon, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, \frac{1}{3}, 0, 0, 0, 0, 0, \frac{2}{3}, 0, 0 \\
0, -\frac{1}{3}\frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, \frac{3}{2}+\frac{1}{3}\frac{1}{\lambda}-\frac{3}{2}\varepsilon, 0, 0, 0, 0, 0, 0, \frac{2}{3}\frac{1}{\lambda}, 0, -\frac{1}{3}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -2-\frac{1}{3}\frac{1}{\lambda}, 0, \frac{1}{2}, 0, 0, 0, 0 \\
0, 0, -\frac{1}{3}\frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, \frac{3}{2}+\frac{1}{3}\frac{1}{\lambda}-\frac{3}{2}\varepsilon, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{2}{3}\frac{1}{\lambda}, 0, -\frac{1}{3}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -2-\frac{1}{3}\frac{1}{\lambda}, 0, \frac{1}{2} \\
0, 0, 0, 0, 0, 1+\frac{2}{9}\frac{1}{\lambda}-\varepsilon, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, \frac{2}{3}, 0, 0, 0, 0, 0, \frac{1}{3}, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 1+\frac{2}{9}\frac{1}{\lambda}-\varepsilon, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -\frac{4}{3}-\frac{2}{9}\frac{1}{\lambda}, 0, \frac{2}{3}, 0, 0, 0, 0, 0, \frac{1}{3}, 0, 0 \\
0, 0, 0, 0, 0, -\frac{1}{3}\frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, \frac{3}{2}+\frac{1}{3}\frac{1}{\lambda}-\frac{3}{2}\varepsilon, 0, 0, 0, 0, 0, \frac{2}{3}\frac{1}{\lambda}, 0, -2-\frac{1}{3}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -\frac{1}{3}\frac{1}{\lambda}, 0, \frac{1}{2}, 0 \\
0, 0, 0, 0, 0, 0, -\frac{1}{3}\frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, \frac{3}{2}+\frac{1}{3}\frac{1}{\lambda}-\frac{3}{2}\varepsilon, 0, 0, 0, 0, 0, \frac{2}{3}\frac{1}{\lambda}, 0, -2-\frac{1}{3}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -\frac{1}{3}\frac{1}{\lambda}, 0, \frac{1}{2} \\
0, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{9}{4}\varepsilon^2+\frac{9}{2}\varepsilon-\frac{9}{4}, 0, 3-3\varepsilon, 0, 0, 0, 0, 0, 3-3\varepsilon, 0, -3+3\varepsilon, 0, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{9}{4}\varepsilon^2-\frac{9}{2}\varepsilon+\frac{9}{4}, 0, -3+3\varepsilon, 0, 0, 0, 0, 0, -3+3\varepsilon, 0, 3-3\varepsilon, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \varepsilon-1, 0, 0, 0, 0, 0, 1-\varepsilon, 0, 0, 0, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \varepsilon-1, 0, 0, 0, 0, 0, 1-\varepsilon, 0, 0, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{9}{4}\varepsilon^2-\frac{9}{2}\varepsilon+\frac{9}{4}, 0, -3+3\varepsilon, 0, 0, 0, 0, 0, -3+3\varepsilon, 0, 3-3\varepsilon, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{9}{4}\varepsilon^2+\frac{9}{2}\varepsilon-\frac{9}{4}, 0, 3-3\varepsilon, 0, 0, 0, 0, 0, 3-3\varepsilon, 0, -3+3\varepsilon \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \varepsilon-1, 0, 0, 0, 0, 0, 1-\varepsilon, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1-\varepsilon, 0, 0, 0, 0, 0, \varepsilon-1, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{4}{9}\frac{1}{\lambda}-\frac{4}{3}\frac{\varepsilon}{\lambda}, 0, -\frac{4}{9}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -\frac{4}{9}\frac{1}{\lambda}+1-\varepsilon, 0, \frac{-4}{3}, 0, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{8}{9}\frac{1}{\lambda}, 0, \frac{20}{9}\frac{1}{\lambda}-\frac{8}{3}\frac{\varepsilon}{\lambda}, 0, 0, 0, 0, 0, -\frac{4}{9}\frac{1}{\lambda}+1-\varepsilon, 0, \frac{2}{3}-2\varepsilon, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, 0, 0, 0, 0, 0, 0, -2\frac{1}{\lambda}+2\frac{\varepsilon}{\lambda}, 0, -\frac{3}{2}+\frac{3}{2}\varepsilon, 0, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, -2\frac{1}{\lambda}+2\frac{\varepsilon}{\lambda}, 0, 0, 0, 0, 0, 0, 0, -\frac{3}{2}+\frac{3}{2}\varepsilon, 0, 0, 0, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{8}{9}\frac{1}{\lambda}, 0, \frac{20}{9}\frac{1}{\lambda}-\frac{8}{3}\frac{\varepsilon}{\lambda}, 0, 0, 0, 0, 0, -\frac{4}{9}\frac{1}{\lambda}+1-\varepsilon, 0, \frac{2}{3}-2\varepsilon, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{4}{9}\frac{1}{\lambda}-\frac{4}{3}\frac{\varepsilon}{\lambda}, 0, -\frac{4}{9}\frac{1}{\lambda}, 0, 0, 0, 0, 0, -\frac{4}{9}\frac{1}{\lambda}+1-\varepsilon, 0, \frac{-4}{3} \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, -2\frac{1}{\lambda}+2\frac{\varepsilon}{\lambda}, 0, 0, 0, 0, 0, 0, 0, -\frac{3}{2}+\frac{3}{2}\varepsilon, 0 \\
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{\lambda}-\frac{\varepsilon}{\lambda}, 0, 0, 0, 0, 0, 0, 0, -2\frac{1}{\lambda}+2\frac{\varepsilon}{\lambda}, 0, -\frac{3}{2}+\frac{3}{2}\varepsilon
\end{array}\right]
$$