ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Approximation Theory 166 (2013) 19–41

Journal of Approximation Theory

www.elsevier.com/locate/jat

Full length article

An algorithm to find a maximum of a multilinear map over a product of spheres

César Massri

Department of Mathematics, FCEN, University of Buenos Aires, Argentina

Received 12 March 2012; received in revised form 23 March 2012; accepted 28 September 2012
Available online 26 October 2012

Communicated by Paul Nevai

Abstract

We provide an algorithm to compute the 2-norm maximum of a multilinear map over a product of spheres. As a corollary we give a method to compute the first singular value of a linear map and an application to the theory of entangled states in quantum physics. Also, we give an application to find a closest rank-one tensor of a given one.

Keywords: Maximum; Product of spheres; Algorithm; Multilinear map; First singular value

0. Introduction

Many problems in mathematics need to maximize a bilinear form over a product of spheres. For example, the 2-norm of a matrix is given by the maximum of the bilinear form $(x, y) \to x^t Ay$, where $\|x\| = \|y\| = 1$. Another interesting problem is to find a closest rank-one tensor of a given tensor, $\sum a_{ijk} x_i \otimes y_j \otimes z_k$. To answer this problem one has to find the maximum of a trilinear form over a product of three spheres (see the examples).

This article provides an algorithm to find the maximum of a multilinear map over a product of spheres,

$$\ell : \mathbb{R}^{n_1+1} \times \cdots \times \mathbb{R}^{n_r+1} \to \mathbb{R}^{n_{r+1}+1}, \qquad \max_{\|x_1\|=\cdots=\|x_r\|=1} \|\ell(x_1, \ldots, x_r)\|.$$

E-mail address: cmassri@dm.uba.ar.

0021-9045/$ - see front matter
doi:10.1016/j.jat.2012.09.007

We have reduced the problem of finding the maximum of ℓ to a problem of finding fixed points of a map $\nabla\ell : \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_{r+1}} \to \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_{r+1}}$. The advantage of this reduction is the possibility to count the number of extreme points of ℓ, and also, to find the fixed points of $\nabla\ell$ solving a system of polynomial equations. There are standard algebro-geometric tools to solve systems of polynomial equations.

In Section 1 we review some concepts and definitions in algebraic geometry such as projective spaces, maps, products of projective spaces and maps between them.

In Section 2, using Lagrange's method of multipliers, see [1, Section 13.7], we reduce the problem of finding the maximum of a multilinear map ℓ to the problem of finding fixed points of a map $\nabla\ell$. We compare our approach with the ones in the literature.

In Section 3 we make a digression to discuss the number of extreme points of a multilinear map over a product of spheres. We use intersection theory to count the number of fixed points of the map $\nabla\ell : \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_{r+1}} \to \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_{r+1}}$. Recall that the number of fixed points of a generic map $F : \mathbb{P}^N \to \mathbb{P}^N$, of degree d, is $1 + d + \cdots + d^N$. In this section we give a formula to compute the number of extreme points of a multilinear map over a product of spheres. If the map is generic, this number is achieved over $\mathbb{C}$, and if it is not generic, this number is a bound when the extreme points are finite. In the literature, the extreme points of ℓ are called singular vectors (see [17]) and in this section we count them.

In Section 4 we use our approach to find the maximum of a bilinear form over a product of spheres. In the bilinear case, the map $\nabla\ell$, induces a linear map $L : \mathbb{P}^N \to \mathbb{P}^N$, where N is a natural number, and we prove that for a generic $q \in \mathbb{P}^N$, the sequence $\{q, L(q), L^2(q), \ldots\}$ converges to the absolute maximum. In other words, the absolute maximum is an attractive fixed point of L. Also, with the same tools, we give an algorithm to find the spectral radius of a square matrix.

In Section 5 we use the theory developed to present the algorithm. We take advantage of a result in Section 2; the classes of extreme points of a multilinear form ℓ, are in bijection with the fixed points of $\nabla\ell$. We reduce the problem of finding fixed points of $\nabla\ell$ to solve a system of polynomial equations with finitely many solutions. In the literature about computational aspects of algebraic geometry, there exists a lot of algorithms to solve a system of polynomial equations with finitely many solutions; see [10]. This gives us the ability to find the absolute maximum of ℓ. It is important to mention that the system of polynomial equations obtained with our approach is slightly different from the system of polynomial equations obtained naively from the method of Lagrange's multipliers. Our approach in projective geometry allows us to find the correct solution removing some constraints. In the first part of the section, we present a direct method to find the maximum value of a generic multilinear form over a product of spheres. Basically, it reduces to finding the spectral radius of a matrix. In the second part of the section, we give an algorithm to find the point $(x_1, \ldots, x_r) \in \mathbb{R}^{n_1+1} \times \cdots \times \mathbb{R}^{n_r+1}$, where $\|x_i\| = 1, 1 \le i \le r$, such that $\|\ell(x_1, \ldots, x_r)\|$ is maximum.

In Section 6 we use the theory developed to compute a lot of examples and applications. One of them is the ability to find a closest rank-one tensor of a given tensor. We prove that this problem is well posed and we apply our algorithm to solve it. Another application is related to quantum physics. It is a criterion of separability. Given a quantum state, we can say if it is separable (see Remark 22 for definitions and related concepts).

1. Review on projective geometry

In this section we give some definitions that we are going to use such as projective spaces, maps, projective tangent spaces, product of projective spaces and maps between them. We are assuming that the base field is $\mathbb{R}$, but all the definitions are true in the complex case. All the notions in this section may be found in [14].

Definition 1. Let n be a natural number and let $\mathbb{R}^{n+1}$ be a real vector space of dimension $n+1$. The *projective space*, $\mathbb{P}^n$, is the space of lines passing through the origin in $\mathbb{R}^{n+1}$. We say that the *dimension* of $\mathbb{P}^n$ is n. Every non-zero vector v in $\mathbb{R}^{n+1}$ determines the line $[v]$ that joins v with the origin $0 \in \mathbb{R}^{n+1}$. The vectors v and λv, $\lambda \in \mathbb{R}$, $\lambda \neq 0$, determine the same point $[v] \in \mathbb{P}^n$.

Let us fix a basis $\{v_0, \ldots, v_n\}$ of $\mathbb{R}^{n+1}$. If the coordinates in this basis of v are $(a_0, \ldots, a_n)$, then the coordinates of the point $[v]$ are

$$[v] = (a_0 : \ldots : a_n) = (\lambda a_0 : \ldots : \lambda a_n), \quad \lambda \in \mathbb{R},\ \lambda \neq 0.$$

In general, we denote $[v] \in \mathbb{P}^n$ to remark that the point $[v]$ is represented by the vector $v \in \mathbb{R}^{n+1}$. Also, we denote an arbitrary point in the projective space, as $p \in \mathbb{P}^n$. The projective space $\mathbb{P}^n$ is a compact space.

Let n and m be two natural numbers. We say that a polynomial P in $n+1$ variables is *homogeneous* of *degree* d, where d is a natural number, if

$$P(\lambda x_0, \ldots, \lambda x_n) = \lambda^d P(x_0, \ldots, x_n), \quad \lambda \in \mathbb{R},\ \lambda \neq 0.$$

For example, a linear form is homogeneous of degree 1.

A map F from $\mathbb{P}^n$ to $\mathbb{P}^m$, denoted by $F : \mathbb{P}^n \to \mathbb{P}^m$, is given by $m+1$ homogeneous polynomials, $F_0, \ldots, F_m$ of degree d,

$$F = (F_0 : \ldots : F_m) : \mathbb{P}^n \to \mathbb{P}^m, \qquad F(x) = (F_0(x) : \ldots : F_m(x)), \quad x \in \mathbb{P}^n.$$

The homogeneity of the polynomials $F_0, \ldots, F_m$, implies that the value of F at $[v]$ and at $[\lambda v]$ is the same in $\mathbb{P}^m$. We say that F has *degree* d. When $d = 1$ we say that F is *linear*.

Let $n_1, \ldots, n_r$ be a list of natural numbers. A *multihomogeneous polynomial* is a polynomial P in variables $x_{i,0}, \ldots, x_{i,n_i}$, $1 \leq i \leq r$, such that

$$P(\lambda_1 x_1, \ldots, \lambda_r x_r) = \lambda_1^{d_1} \cdots \lambda_r^{d_r} P(x_1, \ldots, x_r), \quad x_i = (x_{i,0}, \ldots, x_{i,n_i}),\ \lambda_i \in \mathbb{R}.$$

The vector $(d_1, \ldots, d_r)$ is called the *multidegree* of P. For example, a multilinear form is a multihomogeneous polynomial of multidegree $(1, \ldots, 1)$.

A map $F : \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_r} \to \mathbb{P}^m$, where $m \in \mathbb{N}$, is given by $m+1$ multihomogeneous polynomials, $F_0, \ldots, F_m$ of multidegree $(d_1, \ldots, d_r)$,

$$F(x_1, \ldots, x_r) = (F_0(x_1, \ldots, x_r) : \ldots : F_m(x_1, \ldots, x_r)), \quad x_i \in \mathbb{P}^{n_i},\ 1 \leq i \leq r.$$

The multi-homogeneity of the polynomials $F_0, \ldots, F_m$, implies that the value of F at $([v_1], \ldots, [v_r])$ and at $([\lambda_1 v_1], \ldots, [\lambda_r v_r])$ is the same in $\mathbb{P}^m$. We say that F has *multidegree* $(d_1, \ldots, d_r)$.

Finally, a map $F : \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_r} \to \mathbb{P}^{m_1} \times \cdots \times \mathbb{P}^{m_s}$ is given by s maps $F = (F_1, \ldots, F_s)$,

$$F_i : \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_r} \to \mathbb{P}^{m_i}, \quad 1 \leq i \leq s.$$

Note that the multidegree of F_i may differ from the multidegree of F_j, $i \neq j$. When all the forms $\{F_1, \ldots, F_s\}$ are multilinear, we say that F is a *multilinear* map.

Definition 2. Let n and m be two natural numbers and fix bases for $\mathbb{R}^{n+1}$ and $\mathbb{R}^{m+1}$. Every vector $v \in \mathbb{R}^{n+1}$ has associated a vector space of dimension $n+1$, the *tangent space*, denoted by $T_v\mathbb{R}^{n+1}$.

A polynomial map $F = (F_0, \dots, F_m) : \mathbb{R}^{n+1} \to \mathbb{R}^{m+1}$ such that $F(v) = w$ determines a linear map, $\widehat{dF_v}$, called the differential of F at v,

$$\widehat{dF_v} : T_v\mathbb{R}^{n+1} \to T_w\mathbb{R}^{m+1},$$
$$\widehat{dF_v}(a_0, \dots, a_n) = \left(\sum_{i=0}^{n} \frac{\partial F_0}{\partial x_i}(v)a_i, \dots, \sum_{i=0}^{n} \frac{\partial F_m}{\partial x_i}(v)a_i \right).$$

In the projective space the situation is similar [14, p. 181]. Every point, $x \in \mathbb{P}^n$, has associated a projective space of dimension n, the *projective tangent space*, denoted by $\mathbb{T}_x\mathbb{P}^n$. A map $F = (F_0 : \dots : F_m) : \mathbb{P}^n \to \mathbb{P}^m$ of degree d such that $F(x) = y$ induces a linear map between projective tangent spaces,

$$dF_x : \mathbb{T}_x\mathbb{P}^n \to \mathbb{T}_y\mathbb{P}^m,$$
$$dF_x(a_0 : \dots : a_n) = \left(\sum_{i=0}^{n} \frac{\partial F_0}{\partial x_i}(x)a_i : \dots : \sum_{i=0}^{n} \frac{\partial F_m}{\partial x_i}(x)a_i \right).$$

Given that the partial derivative of a homogeneous polynomial is also homogeneous, the map dF_x is well defined.

Remark 3. Recall the *Euler relation* for a homogeneous polynomial P of degree d, [14, p. 182],

$$\sum_{i=0}^{N} \frac{\partial P}{\partial x_i}(v)v_i = d \cdot P(v), \quad v = (v_0, \dots, v_N) \in \mathbb{R}^{N+1}.$$

The relation follows at once by differentiating both sides of the equation $P(\lambda v) = \lambda^d P(v)$.

If $F = (F_0 : \dots : F_N) : \mathbb{P}^N \to \mathbb{P}^N$ is a map of degree d and $x \in \mathbb{P}^N$ is a point such that $F(x) = x$, then using the Euler relation, we get $dF_x(x) = x$,

$$dF_x : \mathbb{T}_x\mathbb{P}^N \to \mathbb{T}_x\mathbb{P}^N,$$
$$dF_x(a_0 : \dots : a_N) = \left(\sum_{i=0}^{N} \frac{\partial F_0}{\partial x_i}(x)a_i : \dots : \sum_{i=0}^{N} \frac{\partial F_N}{\partial x_i}(x)a_i \right).$$

In particular, if the vector $v \in \mathbb{R}^{N+1}$ represents $x \in \mathbb{P}^N$, $x = [v]$, and the matrix $\widehat{dF_x}$ represent the linear map dF_x,

$$\left(\widehat{dF_x}\right)_{i+1,j+1} = \frac{\partial F_i}{\partial x_j}(v), \quad 0 \le i, j \le N,$$

then v is an eigenvector of $\widehat{dF_x}$. Let us compute the eigenvalue of the eigenvector v. Given that $F(x) = x$, there exists a non-zero real number λ such that $(F_0(v), \dots, F_N(v)) = \lambda v$. Then

$$\lambda v_j = F_j(v) = \frac{1}{d} \sum_{i=0}^{N} \frac{\partial F_j}{\partial x_i}(v)v_i, \quad 0 \le j \le N.$$

Then the eigenvalue of v is $d \cdot \lambda$, where d is the degree of the map F.

2. Theory for a multilinear map

In this section we translate the problem of finding a maximum of a multilinear map to a problem of finding fixed points. Let us present the notation and some preliminaries.

Let $\mathbb{S}^n$ be the sphere in $\mathbb{R}^{n+1}$,

$$\mathbb{S}^n = \left\{ u \in \mathbb{R}^{n+1} \colon \|u\| = \sqrt{|u_0|^2 + \cdots + |u_n|^2} = 1 \right\},$$

and let $\langle -, - \rangle : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$ be the inner product, $\langle x, y \rangle = x_0 y_0 + \cdots + x_n y_n$. The norm assigned to this inner product is the usual 2-norm, $\langle u, u \rangle = \|u\|^2$.

When the codomain of a map is $\mathbb{R}$, we say that the map is a *form*.

Lemma 4. *Given a multilinear map $\ell : \mathbb{R}^{n_1+1} \times \cdots \times \mathbb{R}^{n_r+1} \longrightarrow \mathbb{R}^{s+1}$ there exists a multilinear form $\widehat{\ell}$,*

$$\widehat{\ell} : \mathbb{R}^{n_1+1} \times \cdots \times \mathbb{R}^{n_r+1} \times \mathbb{R}^{s+1} \longrightarrow \mathbb{R}, \qquad \widehat{\ell}(x_1, \ldots, x_r, y) = \langle \ell(x_1, \ldots, x_r), y \rangle,$$

such that

$$\max_{\|x_1\|=\cdots=\|x_r\|=1} \|\ell(x_1, \ldots, x_r)\| = \max_{\|x_1\|=\cdots=\|x_r\|=\|y\|=1} |\widehat{\ell}(x_1, \ldots, x_r, y)|.$$

Proof. The proof is based on the compactness of the sphere. Let $(x_1, \ldots, x_r) \in \mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_r}$ be a point such that $z = \ell(x_1, \ldots, x_r)$ has the maximum norm and let $y = z/\|z\|$. Then

$$\begin{aligned} &|\widehat{\ell}(x_1, \ldots, x_r, y)| = |\langle z, y \rangle| = \frac{\langle z, z \rangle}{\|z\|} = \|z\| = \|\ell(x_1, \ldots, x_r)\| \Longrightarrow \\ &\max_{\|x_1\|=\cdots=\|x_r\|=\|y\|=1} |\widehat{\ell}(x_1, \ldots, x_r, y)| \geq \max_{\|x_1\|=\cdots=\|x_r\|=1} \|\ell(x_1, \ldots, x_r)\|. \end{aligned}$$

Analogously, let $(x_1, \ldots, x_r, y) \in \mathbb{S}^{n_1} \times \cdots \times \mathbb{S}^{n_r} \times \mathbb{S}^s$ be a point such that $|\widehat{\ell}(x_1, \ldots, x_r, y)|$ is maximum. Let $z = \ell(x_1, \ldots, x_r)$. Then

$$\begin{aligned} &|\widehat{\ell}(x_1, \ldots, x_r, y)| = |\langle z, y \rangle| \leq \|z\| \, \|y\| = \|\ell(x_1, \ldots, x_r)\| \Longrightarrow \\ &\max_{\|x_1\|=\cdots=\|x_r\|=\|y\|=1} |\widehat{\ell}(x_1, \ldots, x_r, y)| \leq \max_{\|x_1\|=\cdots=\|x_r\|=1} \|\ell(x_1, \ldots, x_r)\|. \quad \square \end{aligned}$$

As a corollary of the previous lemma, we will work with multilinear forms. Specifically, to make the notation easiest, we will work with $\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \times \mathbb{R}^{s+1} \longrightarrow \mathbb{R}$ a trilinear form. Our goal is to find the maximum of ℓ over a product of three spheres.

Using Lagrange's method of multipliers, ([1, Section 13.7]), we know that the extreme points of ℓ, over $\mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$, satisfy

$$\begin{cases} \partial\ell/\partial x_i(x_0, \ldots, x_n, y_0, \ldots, y_m, z_0, \ldots, z_s) = 2\alpha x_i, & 0 \leq i \leq n, \\ \partial\ell/\partial y_j(x_0, \ldots, x_n, y_0, \ldots, y_m, z_0, \ldots, z_s) = 2\beta y_j, & 0 \leq j \leq m, \\ \partial\ell/\partial z_k(x_0, \ldots, x_n, y_0, \ldots, y_m, z_0, \ldots, z_s) = 2\lambda z_k, & 0 \leq k \leq s, \end{cases}$$
$$\alpha, \beta, \lambda \in \mathbb{R}, \qquad \|x\| = \|y\| = \|z\| = 1.$$

Let us use a better notation,

$$\begin{aligned} &x = (x_0, \ldots, x_n), \qquad y = (y_0, \ldots, y_m), \qquad z = (z_0, \ldots, z_s), \\ &\frac{\partial\ell}{\partial x}(x, y, z) = \left(\frac{\partial\ell}{\partial x_0}(x, y, z), \ldots, \frac{\partial\ell}{\partial x_n}(x, y, z) \right), \\ &\frac{\partial\ell}{\partial y}(x, y, z) = \left(\frac{\partial\ell}{\partial y_0}(x, y, z), \ldots, \frac{\partial\ell}{\partial y_m}(x, y, z) \right), \end{aligned}$$

$$\frac{\partial \ell}{\partial z}(x, y, z) = \left(\frac{\partial \ell}{\partial z_0}(x, y, z), \dots, \frac{\partial \ell}{\partial z_s}(x, y, z) \right),$$
$$\nabla \ell(x, y, z) = \left(\frac{\partial \ell}{\partial x}(x, y, z), \frac{\partial \ell}{\partial y}(x, y, z), \frac{\partial \ell}{\partial z}(x, y, z) \right).$$

Definition 5. A point $(x, y, z) \in \mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$ is called an *extreme point* of ℓ if it satisfies the system of equations

$$\nabla \ell(x, y, z) = (2\alpha x, 2\beta y, 2\lambda z),$$

for some $\alpha, \beta, \lambda \in \mathbb{R}$. Note that if (x, y, z) is an extreme point, then $(\pm x, \pm y, \pm z)$ is also an extreme point. We say that they belong to the same *class*.

Proposition 6. *There is a bijection between classes of extreme points of ℓ and fixed points of the map*

$$\nabla \ell : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s,$$
$$([x], [y], [z]) \to \left(\frac{\partial \ell}{\partial x}([x], [y], [z]), \frac{\partial \ell}{\partial y}([x], [y], [z]), \frac{\partial \ell}{\partial z}([x], [y], [z]) \right).$$

In the multilinear case, we get a similar result.

Proof. Given an extreme point (x, y, z), consider $([x], [y], [z])$. This assignment is independent of the class of (x, y, z). By definition, it gives a fixed point of $\nabla \ell$.

Given a fixed point $([x], [y], [z]) \in \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$ of $\nabla \ell$, consider representatives x, y, z such that $\|x\| = \|y\| = \|z\| = 1$. Then (x, y, z) is an extreme point of ℓ. $\square$

Remark 7. The map $\nabla \ell : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$ from Proposition 6 is not defined over the closed subset

$$\left\{ (x, y, z) \middle| \frac{\partial \ell}{\partial x}(x, y, z) = 0 \text{ or } \frac{\partial \ell}{\partial y}(x, y, z) = 0 \text{ or } \frac{\partial \ell}{\partial z}(x, y, z) = 0 \right\} \subseteq \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s.$$

This set is empty if and only if the *hyperdeterminant* of ℓ is zero. The hyperdeterminant is a polynomial in the coefficient of ℓ; for the definition and some properties see [13, Section 14].

By a result in [13, Section 14, 1.3], if

$$2n, 2m, 2s \leq n + m + s$$

then a generic choice of ℓ, makes $\nabla \ell$ defined everywhere.

In [17], there is a definition of singular values and singular vectors for a multilinear form. For example, for a trilinear form ℓ, the author defined the singular vectors of ℓ as the solutions of the system $\nabla \ell(x, y, z) = (2\alpha x, 2\beta y, 2\lambda z)$. It is the same as our definition of extreme points. It is of interest to know the number of singular values/vectors of ℓ, and in Section 3, we count them. In the same article, the author proved that the first singular value is the maximum of ℓ over a product of spheres. Also, under the hypothesis $2n, 2m, 2s \leq n + m + s$, he proved that the hyperdeterminant of ℓ is zero if and only if 0 is a singular value of ℓ. Given that the hyperdeterminant is a polynomial in the coefficients of ℓ, if ℓ is generic, then the number 0 is not a singular value of ℓ.

There exists another article to mention [8]. In it, the authors gave a different definition of singular values and proposed a multidimensional singular value decomposition. Their decomposition does not preserve the properties that we need, for example, the first singular value of ℓ, does not correspond to the maximum of ℓ over a product of spheres.

3. Number of extreme points of a multilinear form

In this section we use intersection theory ([12, 8.4]) to count the number of fixed points of a generic map $\mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$ over $\mathbb{C}$. Recall from Proposition 6 that there is a bijection between fixed points of

$$\nabla\ell : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$$

and classes of extreme points of the trilinear form ℓ over $\mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$. It is known that if $F : \mathbb{P}^N \to \mathbb{P}^N$ is a generic map of degree d, then F has $1 + d^2 + \cdots + d^N$ fixed points, [11, 1.3]. Here we generalize this result to a generic map between products of projective spaces.

Before we continue with this section, let us make a survey of some related concepts that are in the literature.

In [6,21,20,5,16,3] there is a notion of eigenvectors and eigenvalues assigned to a multilinear form ℓ. There are a lot of applications and in [5], the authors counted the number of eigenvalues of ℓ as the number of roots of a characteristic polynomial assigned to ℓ. The idea is to look at $\ell : \mathbb{C}^n \times \cdots \times \mathbb{C}^n \to \mathbb{C}$ as a polynomial map $P : \mathbb{C}^n \to (\mathbb{C}^n)^\vee \cong \mathbb{C}^n$, $P(x) = \ell(x, \ldots, x, -)$, and then an eigenvector of ℓ is a vector $x \in \mathbb{C}^n$ such that $P(x) = \lambda x$. If ℓ is m-multilinear, P has degree $m-1$ and as a map, $\mathbb{P}^{n-1} \to \mathbb{P}^{n-1}$, it has $(m-1)^{n-1} + (m-1)^{n-2} + \cdots + 1$ fixed points, i.e. eigenvectors of P. They arrived at this number using toric varieties and Newton polytopes.

In [10, 7.1.4] and [19, 3.1] there is a theory of multihomogeneous Bézout number, or m-Bézout. The m-Bézout gives an upper bound on the cardinality of the intersection of multihomogeneous polynomials in $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_k}$. Given that we are counting the fixed points of a map $F : \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_k} \to \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_k}$, in order to apply this formula, we need to realize the fixed points of F as an intersection in some product of projective spaces. Concretely, the intersection of the graph of F and the diagonal. Let us make an explicit example. Assume for simplicity that F is linear, $F : \mathbb{P}^n \to \mathbb{P}^n$, we will see that the m-Bézout formula gives a very bad bound. Recall that the number of fixed points in this case is the number of eigenvectors, that is, $n+1$. Let us apply the formula to the equations of the graph $\Gamma = \{(x, F(x))\}$ and the diagonal $\Delta = \{(x, x)\}$. The points in the intersection satisfy the following equations,

$$y_i F_j(x) = y_j F_i(x), \qquad x_i y_j = x_j y_i, \quad 0 \le i,\ j \le n,$$
$$((x_0 : \ldots : x_n), (y_0 : \ldots : y_n)) \in \mathbb{P}^n \times \mathbb{P}^n.$$

The equations imply that the following matrices have rank one,

$$\begin{pmatrix} y_0 & \cdots & y_n \\ F_0(x) & \cdots & F_n(x) \end{pmatrix}, \qquad \begin{pmatrix} x_0 & \cdots & x_n \\ y_0 & \cdots & y_n \end{pmatrix}.$$

By abuse of notation, we denote the equations,

$$y = F(x), \qquad x = y, \quad (x, y) \in \mathbb{P}^n \times \mathbb{P}^n.$$

Given that the equations have bidegree $(1, 1)$, the m-Bézout number is the coefficient of α_1^{n+1} α_2^{n+1} in the polynomial $(\alpha_1 + \alpha_2)^{2n+2}$. It is the binomial $\binom{2n+2}{n+1} \neq n+1$.

Bernstein proved in [4] that the number of solutions of a *sparse system* equals the *mixed volume* of the corresponding *Newton polytopes*. A sparse system is a collection of Laurent polynomials,

$$f_i = \sum_{(v_1,\dots,v_n)\in\mathcal{A}_i} c_{i,v_1,\dots,v_n} x_1^{v_1}\cdots x_n^{v_n}, \quad 1 \le i \le n,$$

where $\mathcal{A}_i$ are fixed finite subsets of $\mathbb{Z}^n$. The convex hull Q_i of $\mathcal{A}_i$, $Q_i = \operatorname{conv}(\mathcal{A}_i) \subseteq \mathbb{R}^n$, is called the Newton polytope of f_i. Consider the function

$$R(\lambda_1,\dots,\lambda_n) := \operatorname{vol}(\lambda_1 Q_1 + \cdots + \lambda_n Q_n), \quad \lambda_i \ge 0,\ 1 \le i \le n$$

where *vol* is the usual Euclidean volume in $\mathbb{R}^n$ and $Q + Q'$ denotes the Minkowski sum of polytopes. It is a fact that R is a homogeneous polynomial and the coefficient of the monomial $\lambda_1\cdots\lambda_n$ is called the mixed volume of $Q_1,\dots,Q_n$. The mixed volume (i.e the number of solutions of a sparse system) is a very difficult number to compute; see [7, p. 363]. In some situations, this is possible and in the general case, there are a lot of algorithms to compute it. In our situation, we are working with a multihomogeneous polynomial system, and using Bernstein's theorem, in [18], the author gives a recursive formula to compute this number. In fact, it is proved that, under some hypothesis, if the system is over $\mathbb{R}$ and the functions are generic, then all the solutions are reals. Here, we present a different and more direct method using intersection theory.

Let us make an introduction to intersection theory. The germ of intersection theory is the Fundamental Theorem of Algebra. It implies that given a generic homogeneous polynomial in two variables F of degree d, the set of zeros $\{x \in \mathbb{P}^1,\ F(x) = 0\}$ has d points. Generalizing this result, Bézout's theorem says that given two generic homogeneous polynomials in three variables of degrees d and e, the set of zeros $\{x \in \mathbb{P}^2,\ F_1(x) = F_2(x) = 0\}$ consists of de points. In $\mathbb{P}^r$ the situation is similar, if $F_1,\dots,F_r$ are generic homogeneous polynomials of degree $d_1,\dots,d_r$ respectively, the set $\{x \in \mathbb{P}^r,\ F_1(x) = \cdots = F_r(x) = 0\}$ has $d_1 d_2 \dots d_r$ points.

To formalize these ideas, let us introduce the Chow ring of $\mathbb{P}^r$, [12, proof of Proposition 8.4]

$$A(\mathbb{P}^r) = \mathbb{Z}[\alpha]/(\alpha^{r+1}).$$

Every variety $X \subseteq \mathbb{P}^r$ has a class, $[X] \in A(\mathbb{P}^r)$. The intersection of two generic varieties $X \cap Y$ corresponds to the product of the classes $[X]\cdot[Y] = [X \cap Y]$. Two different varieties may correspond to the same class, for example, every hypersurface of degree d corresponds to the same class, $d\alpha$, where α is the class of a hyperplane. For example, α^r corresponds to the intersection of r generic hyperplanes, i.e. a point. The product

$$(d_1\alpha)\cdot(d_2\alpha)\cdots(d_r\alpha) = d_1\cdots d_r\alpha^r$$

corresponds to the intersection of r generic hypersurfaces of degree $d_1,\dots,d_r$ respectively. We get $d_1\cdots d_r$ points in the intersection as mentioned. The class of a variety of codimension c is a homogeneous polynomial of degree c in $\mathbb{Z}[\alpha]/(\alpha^{r+1})$.

The Chow ring is very useful to solve problems in enumerative geometry. For example, to count the number of fixed points of a generic map $F : \mathbb{P}^r \to \mathbb{P}^r$, the procedure is the following. Let $A(\mathbb{P}^r \times \mathbb{P}^r)$ be the Chow ring of $\mathbb{P}^r \times \mathbb{P}^r$, defined as $A(\mathbb{P}^r \times \mathbb{P}^r) = \mathbb{Z}[a,\alpha]/(a^{r+1},\alpha^{r+1})$, [12, Example 8.4.2]. Let $[\Delta] \in A(\mathbb{P}^r \times \mathbb{P}^r)$ be the class of the diagonal, $\Delta = \{(x,x)\}$, and let $[\Gamma] \in A(\mathbb{P}^r \times \mathbb{P}^r)$ be the class of the graph of F, $\Gamma = \{(x, F(x))\}$. Given that

$$\dim\Delta + \dim\Gamma = \dim(\mathbb{P}^r \times \mathbb{P}^r),$$

the product $[\Delta]\cdot[\Gamma]$ is a multiple of the class of a point, $da^r\alpha^r$, [12, Section 8.3]. The coefficient d is the number of fixed points of F.

The Chow ring of a product of projective spaces, [12, Example 8.3.7], is

$$\begin{aligned} A(\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_k}) &= A(\mathbb{P}^{n_1})\otimes_{\mathbb{Z}}\cdots\otimes_{\mathbb{Z}} A(\mathbb{P}^{n_k}) \\ &= \mathbb{Z}[\alpha_1,\ldots,\alpha_k]/(\alpha_1^{n_1+1},\ldots,\alpha_k^{n_k+1}). \end{aligned}$$

Note that in $A(\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_k})$ there is only one class of a point, $\alpha_1^{n_1}\cdots\alpha_k^{n_k}$, so there is a well defined map called *degree*. The degree of a class is the coefficient of $\alpha_1^{n_1}\cdots\alpha_k^{n_k}$. It may be zero (or negative).

The last thing to mention is that every map $F:\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_k}\to\mathbb{P}^{m_1}\times\cdots\times\mathbb{P}^{m_l}$ induces a morphism of rings, [12, Proposition 8.3(a)],

$$F^\star : A(\mathbb{P}^{m_1}\times\cdots\times\mathbb{P}^{m_l})\to A(\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_k}),\qquad F^\star([X])=[F^{-1}(X)].$$

For a more extensive treatment of intersection theory, see [15, Section A], [12].

Let us use the previous introduction. First, we will compute the number of fixed points of a generic map $F:\mathbb{P}^r\to\mathbb{P}^r$ of degree d. Then we will adapt the proof to a generic map $F:\mathbb{P}^n\times\mathbb{P}^m\times\mathbb{P}^s\to\mathbb{P}^n\times\mathbb{P}^m\times\mathbb{P}^s$.

Proposition 8. *The number of fixed points of a generic map $F:\mathbb{P}^r\to\mathbb{P}^r$ of degree d is*

$$1+d+\cdots+d^r.$$

Proof. The following proof is standard in intersection theory. The fixed points of a map $F:\mathbb{P}^r\to\mathbb{P}^r$ may be computed in $A(\mathbb{P}^r\times\mathbb{P}^r)$ as the degree of the product of the class of the graph of F, $[\Gamma]$, and the class of the diagonal, $[\Delta]$. First, let us find out the class of the diagonal,

$$[\Delta]\in A^r(\mathbb{P}^r\times\mathbb{P}^r)=\mathbb{Z}[a,\alpha]/(a^{r+1},\alpha^{r+1}).$$

Being of codimension r, the class is a homogeneous polynomial of degree r,

$$[\Delta]=t_0\alpha^r+t_1a\alpha^{r-1}+\cdots+t_{r-1}a^{r-1}\alpha+t_ra^r,\quad t_i\in\mathbb{Z}.$$

Here, a represents a class of a hyperplane in $\mathbb{P}^r$ and a^i represents the intersection of i of these generic hyperplanes, in other words, a^i is a generic linear space of dimension $r-i$ inside $\mathbb{P}^r$. Same for α and α^j. Viewed in $\mathbb{P}^r\times\mathbb{P}^r$, a^i is the class of $U\times\mathbb{P}^r$, where U is a generic linear space of dimension $r-i$, and α^j is the class of $\mathbb{P}^r\times V$, $\alpha^j=[\mathbb{P}^r\times V]$, where $\dim V=r-j$. The class $a^i\alpha^j$, represents a product of general linear spaces $U\times V\subseteq\mathbb{P}^r\times\mathbb{P}^r$, where $\dim U=r-i$ and $\dim V=r-j$.

The class of the diagonal is determined by the coefficients $t_0,\ldots,t_r$. Note that $t_i=[\Delta]\cdot a^{r-i}\alpha^i$. Then we need to count the number of points in $(U\times V)\cap\Delta$,

$$\begin{aligned} &(U\times V)\cap\Delta\cong U\cap V=\{p\}\Longrightarrow \\ &t_0=\cdots=t_r=1\Longrightarrow[\Delta]=\sum_{i=0}^r a^i\alpha^{r-i}. \end{aligned}$$

Now, let us compute the class of the graph of a map, $\Gamma=\{(x,F(x))\}\subseteq\mathbb{P}^r\times\mathbb{P}^r$,

$$[\Gamma]\in A^r(\mathbb{P}^r\times\mathbb{P}^r)=\mathbb{Z}[a,\alpha]/(a^{r+1},\alpha^{r+1}),$$

it is also a homogeneous polynomial of degree r,

$$[\Gamma]=\tau_0\alpha^r+\tau_1a\alpha^{r-1}+\cdots+\tau_{r-1}a^{r-1}\alpha+\tau_ra^r,\quad\tau_i\in\mathbb{Z}.$$

Again, we have $\tau_i = [\Gamma] \cdot a^{r-i}\alpha^i$, so we need to count the points in $\Gamma \cap (U \times V)$, where $\dim U = i$ and $\dim V = r - i$,

$$\Gamma \cap (U \times V) \cong \{x \in U \mid F(x) \in V\} = U \cap F^{-1}(V) \subseteq \mathbb{P}^r.$$

If F is formed by homogeneous polynomials of degree d, the pull-back of a hyperplane is a hypersurface of degree d, then

$$[U \cap F^{-1}(V)] = \alpha^i.F^\star(\alpha^{r-i}) = \alpha^i.F^\star(\alpha)^{r-i} = \alpha^i(d\alpha)^{r-i} = d^{r-i}\alpha^r.$$

Then $U \cap F^{-1}(V)$ has d^{r-i} points, i.e. $\tau_i = d^{r-i}$,

$$\begin{aligned}
&[\Gamma] = d^r\alpha^r + d^{r-1}a\alpha^{r-1} + \cdots + da^{r-1}\alpha + a^r \Longrightarrow \\
&[\Delta] \cdot [\Gamma] = \left(\sum_{i=0}^{r} a^i\alpha^{r-i}\right)\left(\sum_{j=0}^{r} d^{r-j}a^j\alpha^{r-j}\right) \\
&\qquad = \sum_{i,j=0}^{r} d^{r-j}a^{i+j}\alpha^{2r-(i+j)} = \sum_{j=0}^{r} d^{r-j} = 1 + d + \cdots + d^r.
\end{aligned}$$

Given that a constant map has one fixed point, we use the convention $d^0 = 1$ for $d = 0$. □

Let us adapt the previous calculation to $\mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$.

Theorem 9. *The number of fixed points of a map $F = (F_1, F_2, F_3) : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$ is the coefficient of $\alpha^n\beta^m\gamma^s$ in the following polynomial in $\mathbb{Z}[\alpha, \beta, \gamma]$,*

$$\begin{aligned}
&\sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{s}(d_1\alpha + d_2\beta + d_3\gamma)^{n-i}(e_1\alpha + e_2\beta + e_3\gamma)^{m-j} \\
&\quad \times (f_1\alpha + f_2\beta + f_3\gamma)^{s-k}\alpha^i\beta^j\gamma^k,
\end{aligned}$$

where $(d_1, d_2, d_3), (e_1, e_2, e_3)$ and (f_1, f_2, f_3) are the multidegrees of F_1, F_2 and F_3 respectively.

For a generic map $F : \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_k} \to \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_k}$ the result is similar.

Proof. The class of the diagonal $\Delta = \{(x, y, z, x, y, z)\} \in \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \times \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$, is a homogeneous polynomial of degree $n + m + s$,

$$\begin{aligned}
&[\Delta] \in A(\mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \times \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s) \\
&\quad = \mathbb{Z}[\alpha, \beta, \gamma, a, b, c]/(\alpha^{n+1}, \beta^{m+1}, \gamma^{s+1}, a^{n+1}, b^{m+1}, c^{s+1}).
\end{aligned}$$

Instead of doing the same computation as before, let

$$\pi_{1,4} : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \times \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^n,$$

be the projection in the first and the fourth factor (same for $\pi_{2,5}$ and $\pi_{3,6}$) and let $\Delta_n \subseteq \mathbb{P}^n \times \mathbb{P}^n$ be the diagonal of $\mathbb{P}^n$ (same for Δ_m and Δ_s). Then we have

$$[\Delta] = \pi_{1,3}^\star([\Delta_n]) \cdot \pi_{2,5}^\star([\Delta_m]) \cdot \pi_{3,6}^\star([\Delta_s]) = \sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{s} a^i\alpha^{n-i}b^j\beta^{m-j}c^k\gamma^{s-k}.$$

The class of $\Gamma = \{(x, y, z, F(x, y, z))\} \subseteq \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \times \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$ is a homogeneous polynomial of degree $n + m + s$,

$$[\Gamma] = \sum_{i+j+k+i'+j'+k'=n+m+s} \tau_{ijki'j'k'} a^i \alpha^{i'} b^j \beta^{j'} c^k \gamma^{k'}, \quad \tau_{ijki'j'k'} \in \mathbb{Z} \Longrightarrow$$
$$\deg([\Delta] \cdot [\Gamma]) = \sum_{i=0}^{n} \sum_{j=0}^{m} \sum_{k=0}^{s} \tau_{ijkijk},$$

where deg is the coefficient of $a^n \alpha^n b^m \beta^m c^s \gamma^s$, the number of points in the intersection $\Delta \cap \Gamma$. Note that the integer τ_{ijkijk} may be computed as the degree of $[\Gamma] \cdot a^{n-i} \alpha^i b^{m-j} \beta^j c^{s-k} \gamma^k$, it is the number of points in

$$\begin{aligned}
&\Gamma \cap (U_1 \times U_2 \times U_3 \times V_1 \times V_2 \times V_3),\\
&U_1,\ V_1 \subseteq \mathbb{P}^n, \qquad U_2,\ V_2 \subseteq \mathbb{P}^m, \qquad U_3,\ V_3 \subseteq \mathbb{P}^s,\\
&\dim U_1 + \dim V_1 = n, \qquad \dim U_2 + \dim V_2 = m, \qquad \dim U_3 + \dim V_3 = s \Longrightarrow\\
&\Gamma \cap (U_1 \times U_2 \times U_3 \times V_1 \times V_2 \times V_3)\\
&\quad \cong (U_1 \times U_2 \times U_3) \cap F^{-1}(V_1 \times V_2 \times V_3) \subseteq \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s.
\end{aligned}$$

Let us use the fact that F is equal to $(F_1, F_2, F_3) : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$,

$$\begin{aligned}
&F_1 : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n, \qquad F_2 : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^m,\\
&F_3 : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^s
\end{aligned}$$

where (d_1, d_2, d_3), (e_1, e_2, e_3) and (f_1, f_2, f_3) are the multidegrees of F_1, F_2 and F_3 respectively. Then

$$F^{-1}(V_1 \times V_2 \times V_3) = F_1^{-1}(V_1) \cap F_2^{-1}(V_2) \cap F_3^{-1}(V_3).$$

Thus, the class of the intersection that defines τ_{ijkijk} in the Chow ring $A(\mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s)$, is

$$\begin{aligned}
\tau_{ijkijk} &= \alpha^i \beta^j \gamma^k F^\star(\alpha^{n-i} \beta^{m-j} \gamma^{s-k}) = \alpha^i \beta^j \gamma^k F_1^\star(\alpha^{n-i}) F_2^\star(\beta^{m-j}) F_3^\star(\gamma^{s-k})\\
&= \alpha^i \beta^j \gamma^k (d_1 \alpha + d_2 \beta + d_3 \gamma)^{n-i} (e_1 \alpha + e_2 \beta + e_3 \gamma)^{m-j}\\
&\quad \times (f_1 \alpha + f_2 \beta + f_3 \gamma)^{s-k}. \quad \square
\end{aligned}$$

Example 10. Let us apply the previous formula to $\nabla \ell$ where $\ell : \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$ is a generic trilinear form. The multidegrees of $\partial \ell / \partial x$, $\partial \ell / \partial y$ and $\partial \ell / \partial z$ are $(0, 1, 1)$, $(1, 0, 1)$ and $(1, 1, 0)$ respectively. Then the number of fixed points of this map (over $\mathbb{C}$) is equal to 37. According to [17, 3], the number of singular values of ℓ is 37.

Example 11. Let us apply the formula to count the number of eigenvectors of a generic linear map $L : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$. The map L induces a map $\mathbb{P}^n \to \mathbb{P}^n$ of degree 1. Then

$$\sum_{i=0}^{n} \alpha^{n-i} \alpha^i = \sum_{i=0}^{n} \alpha^n = (n + 1) \alpha^n.$$

The map $\mathbb{P}^n \to \mathbb{P}^n$ has $n + 1$ fixed points over $\mathbb{C}$, that is, L has $n + 1$ eigenvectors over $\mathbb{C}$.

Example 12. Finally, let us apply the formula to find the number of singular values of a generic linear map $L : \mathbb{R}^{n+1} \to \mathbb{R}^{m+1}$. The map L induces a bilinear form $\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \to \mathbb{R}$ and the bidegrees of $\partial\ell/\partial x$ and $\partial\ell/\partial y$ are (0, 1) and (1, 0) respectively (assume $n \geq m$).

$$\sum_{i=0}^{n}\sum_{j=0}^{m}\beta^{n-i}\alpha^{m-j}\alpha^{i}\beta^{j} = \sum_{i=0}^{n}\sum_{j=0}^{m}\beta^{n-i+j}\alpha^{m-j+i}.$$

The coefficient of $\alpha^n\beta^m$, appears when $i - j = n - m$,

$$i = n - m + j \Longrightarrow \sum_{j=0}^{m}\beta^{n-(n-m+j)+j}\alpha^{m-j+(n-m+j)} = (m+1)\alpha^n\beta^m.$$

Then $\nabla\ell : \mathbb{P}^n \times \mathbb{P}^m \to \mathbb{P}^n \times \mathbb{P}^m$ has $m+1$ fixed points. That is L has $m+1$ singular values over $\mathbb{C}$. We used the variational definition of singular values; see [17]. In the case $n < m$ we can use the fact that the number of non-zero singular values of L and $L^t : \mathbb{R}^{m+1} \to \mathbb{R}^{n+1}$ is the same.

4. Theory for a bilinear form

In this section we present a method to find the maximum of a bilinear form, ℓ, over a product of spheres, $\mathbb{S}^n \times \mathbb{S}^m$. This case is very special and the method presented here does not work for a general multilinear form.

The key point of this method is the fact that the partial derivatives of $\ell = \sum a_{ij}x_i y_j$ are linear,

$$\frac{\partial\ell}{\partial x_i}(x, y) = \ell(e_i, y), \qquad \frac{\partial\ell}{\partial y_j}(x, y) = \ell(x, e_j), \qquad (x, y) \in \mathbb{R}^{n+1} \times \mathbb{R}^{m+1},$$
$$0 \leq i \leq n,\ 0 \leq j \leq m,$$

where $\{e_0, \ldots, e_n\}$ is the standard basis of $\mathbb{R}^{n+1}$. Same for $\{e_0, \ldots, e_m\} \subseteq \mathbb{R}^{m+1}$. The map $\nabla\ell$ induces a linear map $L : \mathbb{P}^{n+m+1} \to \mathbb{P}^{n+m+1}$. Let $(x_0 : \ldots : x_n : y_0 : \ldots : y_m)$ be a point in $\mathbb{P}^{n+m+1}$. Then

$$L(x_0 : \ldots : x_n : y_0 : \ldots : y_m) = (\ell(e_0, y) : \ldots : \ell(e_n, y) : \ell(x, e_0) : \ldots : \ell(x, e_m)).$$

This map is well-defined. Let $\lambda \in \mathbb{R}$, $\lambda \neq 0$,

$$\begin{aligned} &L(\lambda x_0 : \ldots : \lambda x_n : \lambda y_0 : \ldots : \lambda y_m) \\ &\quad = (\ell(e_0, \lambda y) : \ldots : \ell(e_n, \lambda y) : \ell(\lambda x, e_0) : \ldots : \ell(\lambda x, e_m)) \\ &\quad = (\lambda\ell(e_0, y) : \ldots : \lambda\ell(e_n, y) : \lambda\ell(x, e_0) : \ldots : \lambda\ell(x, e_m)) \\ &\quad = L(x_0 : \ldots : x_n : y_0 : \ldots : y_m). \end{aligned}$$

Theorem 13. *Let $p = (x, y) \in \mathbb{S}^n \times \mathbb{S}^m$ be an absolute maximum of ℓ. Then*

$$\lim_{r\to\infty} L^r(q) = [p]$$

for a generic $q \in \mathbb{P}^{n+m+1}$.

Proof. Let $A \in \mathbb{R}^{n+m+2\times n+m+2}$ be a matrix representing the linear map L. Given that L is linear, the differential of L at any point, q, is equal to L,

$$dL_q = L, \quad \forall q \in \mathbb{P}^{n+m+1}.$$

In particular, the matrix A, also represents the differential of L at p,

$$A = \widehat{dL_p}.$$

Let $\{v_0, \ldots, v_{n+m+1}\}$ be a basis of $\mathbb{R}^{n+m+2}$ formed by eigenvectors of A. Let λ_i be the eigenvalue of v_i, $0 \le i \le n+m+1$. By Remark 3 we know that p is an eigenvector of A with eigenvalue $\ell(p)$. In particular, if the magnitude of λ_0 is maximum, then $|\lambda_0| = |\ell(p)|$ and $[p] = [v_0]$.

Let $z \in \mathbb{R}^{n+m+2}$ be a vector representing q such that $z = a_0 v_0 + \cdots + a_{n+m+1} v_{n+m+1}$, $a_0 \neq 0$. Then

$$L^r(q) = \left[A^r \cdot \sum_{i=0}^{n+m+1} a_i v_i \right] = \left[a_0 \lambda_0^r v_0 + \sum_{i=1}^{n+m+1} a_i \lambda_i^r v_i \right]$$
$$= \left[a_0 v_0 + \sum_{i=1}^{n+m+1} a_i \frac{\lambda_i^r}{\lambda_0^r} v_i \right] \to [a_0 v_0] = p. \quad \square$$

In the proof of the previous theorem, we saw that the iterations of a linear map in the projective space converge to an eigenvector with eigenvalue of maximum magnitude. In particular, given a square matrix $A \in \mathbb{R}^{n+1 \times n+1}$ and a generic vector $w \in \mathbb{R}^{n+1}$, the sequence $\{[w], [Aw], [A^2 w] \cdots\} \subseteq \mathbb{P}^n$, converges to a point $[v]$. The vector v satisfies $Av = \lambda v$, where $|\lambda|$ is the spectral radius of A.

The rate of convergence of this method is linear.

Remark 14. Based on Theorem 13, let us give an algorithm to find the absolute maximum of a generic bilinear form,

$$\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \longrightarrow \mathbb{R}.$$

Let $\nabla \ell = (\partial \ell / \partial x, \partial \ell / \partial y)$ be the gradient of ℓ and let $q = (x, y)$ be a vector such that $x \in \mathbb{R}^{n+1}$, $x \neq 0$ and $y \in \mathbb{R}^{m+1}$, $y \neq 0$.

`Input:`	`A bilinear form` $\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \to \mathbb{R}$.
`Output:`	`The absolute maximum` $(x, y) \in \mathbb{S}^n \times \mathbb{S}^m$.
1.	`Let` $q = q/\|q\|$ `and` $aux = (1, 0, \ldots, 0)$.
2.	`While` $\|\langle q, aux \rangle\|$ `is different from` 1, `do`
	2.1 $aux = q$
	2.2 $q = \nabla \ell(q)$
	2.3 $q = q/\|q\|$
3.	`Let` $x = (q_0, \ldots, q_n)$, $y = (q_{n+1}, \ldots, q_{n+m+1})$.
4.	`Return` $(x/\|x\|, y/\|y\|)$.

The iterations stop when the points in the projective space are equal, in other words, when the cosine of the angle between q and aux is 1 or -1 (when they are aligned). Given that the absolute maximum is attractive (see Theorem 13), the program ends. The maximum value is $|\ell(x, y)|$.

Remark 15. We may adapt the previous algorithm to a multilinear form, but in the multilinear case, in general, the absolute maximum is not an attractive fixed point. For example, the trilinear form $\ell : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$\ell(x, y, z) = 6x_0 y_0 z_0 + 3x_1 y_0 z_0 - 6x_0 y_1 z_0 + 16 x_1 y_1 z_0 - 14 x_0 y_0 z_1$$
$$- 15 x_1 y_0 z_1 - 11 x_0 y_1 z_1 + 8 x_1 y_1 z_1,$$

induces a map $\mathbb{P}^5 \to \mathbb{P}^5$ of degree 2 without attractive fixed points. Even more, the 4-multilinear form $\ell : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$\ell(x, y, z, t) = 4x_0 y_0 z_0 t_0 + 6x_1 y_0 z_0 t_0 + x_0 y_1 z_0 t_0 - 6x_1 y_0 z_1 t_0 t_0 - 5x_0 y_0 z_1 t_0$$

$$+ 7x_1y_1z_0t_0 - 5x_0y_1z_1t_0 + 2x_0y_0z_0t_1 - 3x_1y_0z_0t_1 - 7x_0y_1z_0t_1$$
$$+ 9x_1y_1z_0t_1 - 9x_0y_0z_1t_1 - 9x_1y_0z_1t_1 - 6x_0y_1z_1t_1 + 8x_1y_1z_1t_1,$$

induces a map $\mathbb{P}^7 \to \mathbb{P}^7$ of degree 3 with two attractive fixed points. One is the absolute maximum.

5. Presentation of the general algorithm

In this section, we present an algorithm to find the maximum of a multilinear form over a product of spheres. First, we reduce the problem to solve a system of multilinear equations. In the first part of the section, we present an algorithm to find the absolute maximum of a multilinear form. In the second, we give an algorithm to find the point where the maximum occurs.

Proposition 16. *Let $\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \times \mathbb{R}^{s+1} \to \mathbb{R}$ be a generic trilinear form. There exists a bijection between classes of extreme points of ℓ and solutions of the following system of trilinear equations in $\mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$,*

$$\begin{cases} \ell(x_j e_i - x_i e_j, y, z) = 0, & 0 \le i < j \le n, \\ \ell(x, y_j e_i - y_i e_j, z) = 0, & 0 \le i < j \le m, \\ \ell(x, y, z_j e_i - z_i e_j) = 0, & 0 \le i < j \le s. \end{cases}$$

The vector e_k satisfies $(e_k)_l = 0$ if $l \neq k$ and $(e_k)_k = 1$.

In the multilinear case, we obtain a similar result, a system of multilinear equations.

Proof. From Proposition 6, we know that every class of an extreme point of ℓ, is a fixed point of $\nabla\ell : \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \to \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$. If ℓ is a generic trilinear form, we know that the number of fixed points is finite (see Section 3).

A fixed point of $\nabla\ell$, $([x], [y], [z])$, satisfies

$$\begin{cases} \partial\ell/\partial x(x, y, z) = 2\alpha x \\ \partial\ell/\partial y(x, y, z) = 2\beta y \\ \partial\ell/\partial z(x, y, z) = 2\lambda z \end{cases}$$

where α, β and λ are three non-zero real numbers. In $\mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s$, the equations are

$$\begin{cases} x_j \partial\ell/\partial x_i(x, y, z) = x_i \partial\ell/\partial x_j(x, y, z), & 0 \le i < j \le n, \\ y_j \partial\ell/\partial y_i(x, y, z) = y_i \partial\ell/\partial y_j(x, y, z), & 0 \le i < j \le m, \\ z_j \partial\ell/\partial z_i(x, y, z) = z_i \partial\ell/\partial z_j(x, y, z), & 0 \le i < j \le s. \end{cases}$$

The result follows from the equalities,

$$\partial\ell/\partial x_i(x, y, z) = \ell(e_i, y, z), \qquad \partial\ell/\partial y_j(x, y, z) = \ell(x, e_j, z),$$
$$\partial\ell/\partial z_k(x, y, z) = \ell(x, y, e_k). \quad \square$$

Let us present the algorithm to find the absolute maximum of a generic multilinear form. The algorithm is based on *Eigenvalue Theorem*. Let us recall it. Consider a system of polynomial equations with finitely many solutions in $\mathbb{C}^n$,

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ \vdots \\ f_m(x_1, \dots, x_n) = 0 \end{cases}$$

where $f_1, \ldots, f_m$ are polynomials in $\mathbb{C}[x_1, \ldots, x_n]$. The quotient ring,

$$\mathcal{A} = \mathbb{C}[x_1, \ldots, x_n]/\langle f_1, \ldots, f_m \rangle,$$

is a finite-dimensional vector space, [10, Theorem 2.1.2]. The dimension of $\mathcal{A}$ is the number of solutions of the system.

Every polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$, determines a linear map $M : \mathcal{A} \to \mathcal{A}$,

$$M(\overline{g}) = \overline{fg}, \quad g \in \mathbb{C}[x_1, \ldots, x_n],$$

where $\overline{g}$ denotes the class of the polynomial g in the quotient ring $\mathcal{A}$. The matrix of M is called the *multiplication matrix* assigned to the polynomial f.

Theorem (*Eigenvalue Theorem*)**.** *The eigenvalues of M are $\{f(p_1), \ldots, f(p_r)\}$, where $\{p_1, \ldots, p_r\}$ are the solutions of the system of polynomial equations. See* [10, *Theorem* 2.1.4] *for a proof.*

The algorithm in Appendix, first generates the following system of polynomial equations,

$$\begin{cases} \ell(x_j e_i - x_i e_j, y, z) = 0, & 0 \le i < j \le n, \\ \ell(x, y_j e_i - y_i e_j, z) = 0, & 0 \le i < j \le m, \\ \ell(x, y, z_j e_i - z_i e_j) = 0, & 0 \le i < j \le s, \\ \|x\|^2 = 1, \qquad \|y\|^2 = 1, & \|z\|^2 = 1, \end{cases}$$

where the vector e_k satisfies $(e_k)_l = 0$, if $l \neq k$, and $(e_k)_k = 1$. Then, computes the real eigenvalues, $\{\lambda_0, \ldots, \lambda_r\}$, of the multiplication matrix assigned to ℓ. Finally, it returns λ_i such that $|\lambda_i| \ge |\lambda_j|$ for all $0 \le j \le r$. This number is the maximum of ℓ over $\mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$.

For the algorithm and an implementation in Maple, see Appendix.

Now, let us give an algorithm to find the point $(x, y, z) \in \mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$ such that $|\ell(x, y, z)|$ is maximum. We need to use the following result (same notation as Eigenvalue Theorem),

Theorem. *Let $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$ be a generic linear form and let M be its multiplication matrix. Assume that $B = \{1, x_1, \ldots, x_n, \ldots\}$ is a finite basis of $\mathcal{A}$ formed by monomials. Then the eigenvectors of M determine all the solutions of the system of polynomial equations. Specifically, if $v = (v_0, \ldots, v_n, \ldots)$ is an eigenvector of M such that $v_0 = 1$, then $(v_1, \ldots, v_n)$ is a solution of the system of polynomial equations. Even more, every solution is of this form. See* [10, *Section* 2.1.3] *for a proof.*

Note that the theorem requires that the variables $\{x_1, \ldots, x_n\}$ are elements of the basis B. It could be the case that some variables are missing from B. For example, if $x_1, \ldots, x_i \in B$, and $x_{i+1}, \ldots, x_n \notin B$, then every missing variable, say x_j, is a linear combination of $\{x_1, \ldots, x_i\}$,

$$x_j = a_{j1} x_1 + \cdots + a_{ji} x_i, \quad i + 1 \le j \le n.$$

If $v = (1, v_1, v_2 \ldots)$ is an eigenvector of M, the j-coordinate of the solution corresponding to v, is $a_{j1} v_1 + \cdots + a_{ji} v_i$. See [10, Section 2.1.3].

Let us call the *affine system* to the following system of polynomial equations,

$$\begin{cases} \ell(x_j e_i - x_i e_j, y, z) = 0, & 0 \le i < j \le n, \\ \ell(x, y_j e_i - y_i e_j, z) = 0, & 0 \le i < j \le m, \\ \ell(x, y, z_j e_i - z_i e_j) = 0, & 0 \le i < j \le s, \\ x_0 = 1, \qquad y_0 = 1, & z_0 = 1. \end{cases}$$

The vector e_k satisfies $(e_k)_l = 0$ if $l \neq k$ and $(e_k)_k = 1$. The solutions of this system determine classes of extreme points of ℓ. The genericity of ℓ implies that all the extreme points of ℓ, (x, y, z), satisfy $x_0 \neq 0, y_0 \neq 0, z_0 \neq 0$. Then all the classes of extreme points appear as the solutions of the affine system. Even more, the cardinal of the basis B of the affine system is equal to the number of classes of extreme points.

Theorem 17. *Assume that $\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \times \mathbb{R}^{s+1} \to \mathbb{R}$ is a generic trilinear form and that $2n, 2m, 2s \leq n + m + s$; see Remark 7. Then the basis B of the affine system contains all the variables.*

In the multilinear case, we obtain a similar result.

Proof. Given that the equations in Proposition 16 are multilinear, the quotient ring, $\mathcal{A}$, is multi-graded. Let us denote $\mathcal{A}_{(d_1,d_2,d_3)}$ the multidegree part (d_1, d_2, d_3), where $d_1, d_2, d_3 \geq 0$. The hypothesis $2n, 2m, 2s \leq n + m + s$, implies that the following set is empty,

$$\left\{(x, y, z) \in \mathbb{P}^n \times \mathbb{P}^m \times \mathbb{P}^s \,\middle|\, \frac{\partial \ell}{\partial x}(x, y, z) = 0 \text{ or } \frac{\partial \ell}{\partial y}(x, y, z) = 0 \right.$$
$$\left. \text{or } \frac{\partial \ell}{\partial z}(x, y, z) = 0\right\} = \emptyset.$$

Then the equations $\{\partial \ell/\partial x_i\}_{i=0}^n$ are linearly independent. Same for $\{\partial \ell/\partial y_j\}_{j=0}^m$ and $\{\partial \ell/\partial z_k\}_{k=0}^s$. In the quotient ring, $\mathcal{A}$, the partial derivatives are proportional to the variables, thus, the variables are linearly independent too. For example, a basis for the multidegree part $(0, 0, 0)$ is $\{1\}$, and a basis for the multidegree part $(1, 0, 0)$ is $\{x_0, \ldots, x_n\}$. Even more, a basis for

$$\mathcal{A}_{(1,0,0)} \oplus \mathcal{A}_{(0,1,0)} \oplus \mathcal{A}_{(0,0,1)}$$

is $\{x_0, \ldots, x_n, y_0, \ldots, y_m, z_0, \ldots, z_s\}$.

Let us add the equations $x_0 = y_0 = z_0 = 1$ to the system of polynomial equations. The equations are not multilinear, so the corresponding quotient ring is not multi-graded,

$$\widehat{\mathcal{A}} = \mathcal{A}/\langle x_0 - 1, y_0 - 1, z_0 - 1\rangle.$$

Notice that the variables $\{x_1, \ldots, x_n, y_1, \ldots, y_m, z_1, \ldots, z_s\}$ are linearly independent in $\widehat{\mathcal{A}}$. This implies that the basis B of $\widehat{\mathcal{A}}$, formed by monomials, contains all the variables.

$$\sum_{i=1}^n \alpha_i x_i + \sum_{j=1}^m \beta_j y_j + \sum_{k=1}^s \lambda_k z_k = 0 \in \widehat{\mathcal{A}}, \quad \alpha_i, \beta_j, \lambda_k \in \mathbb{C} \Longrightarrow$$
$$\sum_{i=1}^n \alpha_i x_i + \sum_{j=1}^m \beta_j y_j + \sum_{k=1}^s \lambda_k z_k = P \in \mathcal{A}, \quad P \in \langle x_0 - 1, y_0 - 1, z_0 - 1\rangle,$$

where P is a polynomial combination of $x_0 - 1$, $y_0 - 1$ and $z_0 - 1$.

Denote $P_{(d_1,d_2,d_3)}$ as the multidegree part (d_1, d_2, d_3) of P. Given that $\mathcal{A}$ is multi-graded, we get the following equalities in $\mathcal{A}$,

$$P = P_{(1,0,0)} + P_{(0,1,0)} + P_{(0,0,1)}, \qquad \sum_{i=1}^n \alpha_i x_i = P_{(1,0,0)},$$
$$\sum_{j=1}^m \beta_j y_i = P_{(0,1,0)}, \qquad \sum_{k=1}^s \lambda_k z_k = P_{(0,0,1)}.$$

Using the fact that the variables $\{x_0, \ldots, x_n\}$ are linearly independent in $\mathcal{A}$, we find that x_0 is not a variable in P. Same for y_0 and z_0. Given that P is a polynomial combination of $x_0 - 1$, $y_0 - 1$ and $z_0 - 1$, it must be 0. Then $\alpha_1 = \cdots = \alpha_n = 0$, $\beta_1 = \cdots = \beta_m = 0$ and $\lambda_1 = \cdots = \lambda_s = 0$. □

Remark 18. Let us give the algorithm to find the point $(x, y, z) \in \mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$ such that $|\ell(x, y, z)|$ is maximum.

We choose to work with Gröbner bases because they are implemented in most computer algebra systems (Maple, Macaulay2, Singular). In [2], the authors proposed an algorithm without the need to use Gröbner bases. See also [10, 2.3.1].

Input: A generic trilinear form $\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \times \mathbb{R}^{s+1} \to \mathbb{R}$,
where $2n + 2m + 2s \leq n + m + s$.
Output: The absolute maximum $(x, y, z) \in \mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$.

1. Compute the system of trilinear equations of Proposition 16.
2. Add the equations $x_0 = y_0 = z_0 = 1$.
3. Compute a Gröbner basis for the resulting system, $\mathcal{I}$.
4. Find a basis B of $\mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_m, z_1, \ldots, z_s]/\mathcal{I}$.
5. Compute the multiplication matrix of x_1.
6. Compute the eigenvectors of the multiplication matrix.
7. For each eigenvector v, do
   - 7.1 Normalize v such that $v = (1, v_1, \ldots)$.
   - 7.2 Let $x = (x_1, \ldots, x_n)$ be such that $x_i = v_{\sigma_i}$ where σ_i is the coordinate of x_i in B, $1 \leq i \leq n$.
   - 7.3 Same for y and z.
   - 7.4 Normalize the points, $x = x/\|x\|$, $y = y/\|y\|$, $z = z/\|z\|$.
   - 7.5 Evaluate ℓ at (x, y, z) if the coordinates are real.
   - 7.6 Save the maximum.
8. Return the maximum, (x, y, z).

In Step 5 of the algorithm we used the linear form x_1 as a generic linear form. This fact is not restrictive. Given that the trilinear form is generic, we may suppose that the first coordinates of the classes of extreme points of ℓ are all different. In other words, the eigenvalues of the multiplication matrix of x_1 have multiplicity one. See [10, Section 2.1.3]. We added the hypothesis $2n, 2m, 2s \leq n + m + s$ to imply that B contains all the variables. The reader may adapt the algorithm to the general case.

For a multilinear form, the algorithm is similar.

6. Applications and examples

Let us start with some applications. First, we give applications of the iterative algorithm to find the maximum of a bilinear form. Then, we give applications of the general algorithm.

Remark 19. Given a real matrix A, its first singular value (the 2-norm) is given by

$$\max_{\|x\|=\|y\|=1} x^t A y.$$

In other words, it is the maximum over a product of spheres of the bilinear form $(x, y) \to x^t A y$.

An interesting aspect of Theorem 13 is that we can find the first singular vectors and the first singular value, $|\ell(x,y)|$, of ℓ without using the *spectral radius formula*. Recall that the 2-norm of a matrix A is computed using the spectral radius formula,

$$\|A\|_2 = \sqrt{\lim_{k\to+\infty} \|(AA^t)^k\|^{\frac{1}{k}}}.$$

Example 20. Let $A \in \mathbb{R}^{4\times 3}$ be the matrix

$$A = \begin{pmatrix} 3 & 2 & 32 \\ 2 & 1 & 36 \\ -3 & 25 & 2 \\ 0 & -1 & 1 \end{pmatrix}.$$

Then with the algorithm in Remark 14, we get that the 2-norm is 48.46054603. Computing the spectral radius of AA^t, we get the same number 48.46054603.

Example 21. An interesting example is to apply the algorithm to a bilinear form over $\mathbb{S}^1 \times \mathbb{S}^0$. Note that the domain is a cylinder in $\mathbb{R}^3$, so we can draw the whole situation. Take, for example, the bilinear form

$$\ell : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}, \qquad \ell(x,y) = 4x_0y + 2x_1y.$$

The maximum of ℓ over $\mathbb{S}^1 \times \mathbb{S}^0$ is the 2-norm of the vector (4, 2), that is,

$$\|(4,2)\| = \sqrt{20} \cong 4.472135954.$$

Let us compute this using the algorithm in Remark 14. First of all, note that the gradient of ℓ determines a vector field over the cylinder, and the iteration follows the arrow. Over the ending point of the iteration, the flow is orthogonal to the surface. This means that we have reached an extreme,

$$\max_{\|x\|=|y|=1} \ell(x,y) \cong 4.472135953.$$

Let us give now some applications of the general algorithm; see Appendix.

Remark 22. The first interesting application of the algorithm in Appendix is to the theory of entanglement. It is of interest to find the maximum of the form $\langle\rho,-\rangle$ over the space of *separable states*. The matrix ρ is called a *state* if it is Hermitian, $\rho \geq 0$ and $\mathrm{tr}(\rho) = 1$. It is easy to see that the space of states is a convex set and is generated by the matrices of the form $vv^\dagger$ where v is a column vector of norm one in a finite dimensional vector space $\mathcal{H}$, $v \in \mathcal{H}$, $\|v\| = 1$. The general theory says that when we work with two particles, we need to consider the space of states over the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. In this situation, a state is called *separable* if it is a convex combination of the form $\sum a_i v_i \otimes w_i$, where v_i is a state of $\mathcal{H}_1$ and w_i a state of $\mathcal{H}_2$. Let us call $\mathrm{Sep}(\mathcal{H})$ the convex space of separable states. It is true that the space of separable states is a convex set generated by the matrices of the form $xx^\dagger \otimes yy^\dagger$, where $x \in \mathcal{H}_1$, $y \in \mathcal{H}_2$ and $\|x\| = \|y\| = 1$. Then

$$\max_{\mathrm{Sep}(\mathcal{H})} \langle\rho,-\rangle = \max_{\|x\|=\|y\|=1} \langle\rho, xx^\dagger \otimes yy^\dagger\rangle.$$

Note that the form is not bilinear in x nor in y. Rewriting the state ρ in the form $\rho = \sum \lambda_i \rho_i \rho_i^\dagger$ with $\langle \rho_i, \rho_j \rangle = 0$, $\|\rho_i\| = 1$, and using the equality $\langle \rho_i, x \otimes y \rangle^2 = \langle \rho_i \rho_i^\dagger, xx^\dagger \otimes yy^\dagger \rangle$, we get

$$\max_{\mathrm{Sep}(\mathcal{H})} \langle \rho, - \rangle = \max_{\|x\|=\|y\|=1} \left\| \sum \sqrt{\lambda_i} \langle \rho_i, x \otimes y \rangle \rho_i \right\|^2 .$$

The resulting map $\sum \sqrt{\lambda_i} \langle \rho_i, x \otimes y \rangle \rho_i$ is bilinear in x and in y and our algorithm is capable of maximizing it. See the next example.

Example 23. Let ρ be following state in $\mathbb{R}^4 = \mathbb{R}^2 \otimes \mathbb{R}^2$,

$$\begin{pmatrix} 0.242894940524649938 & -0.123994312358229969 & -0.0712215842649899789 & 0.219784373378769966 \\ -0.123994312358229969 & 0.0888784895376599772 & 0.111143109132249979 & -0.0627261109839499926 \\ -0.0712215842649899789 & 0.111143109132249979 & 0.361255602168969903 & 0.0603142605185699871 \\ 0.219784373378769966 & -0.0627261109839499926 & 0.0603142605185699871 & 0.306970967813849916 \end{pmatrix}.$$

We choose to work over the real numbers to make the exposition clearer, but all the results can be adapted to work with Hermitian matrices instead of symmetric matrices. Using the Cholesky algorithm and making a singular value decomposition, we have $\rho = \sum \lambda_i \rho_i \rho_i^\dagger$,

$$\lambda_1 = 0.5435016101, \qquad \lambda_2 = 0.4146107959,$$
$$\lambda_3 = 0.04113792919, \qquad \lambda_4 = 0.0007496649711, \qquad \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1.$$
$$\rho_1 = \begin{pmatrix} -0.656481390369177854 \\ 0.326643787198963642 \\ 0.245965753592146592 \\ -0.633906040705653040 \end{pmatrix}, \qquad \rho_2 = \begin{pmatrix} -0.0253829550629408562 \\ -0.209402292907094082 \\ -0.881013881627254691 \\ -0.423463015737623016 \end{pmatrix}$$
$$\rho_3 = \begin{pmatrix} -0.444223726945872255 \\ 0.546710519336902400 \\ -0.399219012690601172 \\ 0.586853532298337144 \end{pmatrix}, \qquad \rho_4 = \begin{pmatrix} -0.609141338368644146 \\ -0.741998735885678107 \\ 0.0627659808038323054 \\ 0.272846362424434330 \end{pmatrix}.$$

Applying the algorithm in Appendix to the trilinear form

$$\ell(x, y, z) = \sum \sqrt{\lambda_i} \langle \rho_i, x \otimes y \rangle \langle \rho_i, z \rangle, \quad (x, y, z) \in \mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{S}^3,$$

we get that the maximum is 0.7228016991. Finally,

$$\max_{\mathrm{Sep}(\mathcal{H})} \langle \rho, - \rangle = \max_{\|x\|=\|y\|=1} \left\| \sum \sqrt{\lambda_i} \langle \rho_i, x \otimes y \rangle \rho_i \right\|^2 = \max_{\|x\|=\|y\|=\|z\|=1} |\ell(x, y, z)|^2$$
$$\cong 0.7228016991^2 \cong 0.5224422962.$$

Note that if ρ is separable, then $\langle \rho, \rho \rangle \leq \max_{\mathrm{Sep}(\mathcal{H})} \langle \rho, - \rangle$. This is not the case, but for example, the following state is not separable (it is called *entangled state*),

$$\phi = \begin{pmatrix} 0.168106937369559950 & -0.190509527669719958 & -0.200004375511779936 & -0.0690454833860399825 \\ -0.190509527669719958 & 0.257651665981429912 & 0.267759084652009926 & 0.0985801483325399742 \\ -0.200004375511779936 & 0.267759084652009926 & 0.320790216378169901 & 0.194053687463299957 \\ -0.0690454833860399825 & 0.0985801483325399742 & 0.194053687463299957 & 0.253451180300149959 \end{pmatrix}.$$

We have $\langle \phi, \phi \rangle \cong 0.6620536187 \not\leq 0.4862909489 \cong \max_{\mathrm{Sep}(\mathcal{H})} \langle \phi, - \rangle$.

Remark 24. Our final application is the ability to find numerically a closest rank-one tensor of a given tensor. In [9], the authors considered the problem of finding the best rank-r approximation of a given tensor. They proved that for $r > 1$ the problem is ill-posed, but when $r = 1$, the problem has a solution, [9, 4.5]. Here we find the solution. Let us prove that a computation of the absolute maximum of ℓ over a product of spheres gives the closest *rank-one multilinear form* to ℓ. A rank-one multilinear form is a product of linear forms, $\ell_1 \cdots \ell_s$, where $\ell_i : \mathbb{R}^{n_i+1} \to \mathbb{R}$, $1 \le i \le s$. We choose to do this remark about multilinear forms, but dually, the same is true for tensors.

For simplicity, we do the proof for a trilinear form. The proof is similar in the multilinear case. Consider the affine Segre map (it is not an isometry)

$$\mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \times \mathbb{R}^{s+1} \longrightarrow \mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1}, \quad (x, y, z) \longrightarrow x \otimes y \otimes z.$$

Using the usual inner product in the tensor product, we identify

$$\mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1} \cong (\mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1})^\vee,$$
$$x \otimes y \otimes z \longrightarrow \langle x \otimes y \otimes z, -\rangle, \qquad \langle x \otimes y \otimes z, a \otimes b \otimes c\rangle = \langle x, a\rangle\langle y, b\rangle\langle z, c\rangle.$$

We can identify the following three different notations

$$\ell(x, y, z) = \ell(x \otimes y \otimes z) = \langle \ell, x \otimes y \otimes z\rangle.$$

The first equality identifies a trilinear form with a linear map $\ell : \mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1} \to \mathbb{R}$. The second equality identifies, under the isometry $(\mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1})^\vee \cong \mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1}$, the linear form ℓ with the tensor $\ell \in \mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1}$.

Let $\mathbb{S}$ be the immersion of $\mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$ under the Segre map,

$$\mathbb{S} = \{\langle x \otimes y \otimes z, -\rangle \colon \|x\| = \|y\| = \|z\| = 1\} \subseteq (\mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1})^\vee.$$

Then, for all $\phi = \langle x \otimes y \otimes z, -\rangle \in \mathbb{S}$, we have

$$\|\ell - \phi\|^2 = \Big(\|\ell\|^2 + \|\phi\|^2 - 2\langle \ell, \phi\rangle\Big) = \|\ell\|^2 + 1 - 2\ell(x, y, z).$$

In other words, a local maximum of ℓ is a local minimum of the distance function, $\|\ell - \phi\|$.

Let $\mathbb{B}$ be the image, under the Segre map, of a product of balls,

$$\mathbb{B} = \{\langle x \otimes y \otimes z, -\rangle \colon \|x\|, \|y\|, \|z\| \le 1\} \subseteq (\mathbb{R}^{n+1} \otimes \mathbb{R}^{m+1} \otimes \mathbb{R}^{s+1})^\vee.$$

Note that the elements of $\mathbb{B}$ are rank-one multilinear forms. It is easy to see that $\mathbb{B}$ is compact and convex, so the distance from ℓ to $\mathbb{B}$ is achieved in $\mathbb{S}$ (the border). In other words, a closest rank-one multilinear form to ℓ is an element of $\mathbb{S}$. Summing up, a computation with the algorithm in Remark 18 of the absolute maximum of ℓ, gives a rank one multilinear form such that the distance to ℓ is minimum. Note that there can be more than one closest rank-one approximation. Every closest rank-one multilinear form, determines an absolute maximum and vice versa.

Example 25. Let $\ell : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ be the multilinear form

$$\begin{aligned}\ell(x, y, z, t) = {} & 4x_0y_0z_0t_0 + 6x_1y_0z_0t_0 + x_0y_1z_0t_0 + 7x_1y_1z_0t_0 - 5x_0y_0z_1t_0 \\ & - 6x_1y_0z_1t_0 - 5x_0y_1z_1t_0 + 2x_0y_0z_0t_1 - 3x_1y_0z_0t_1 \\ & - 7x_0y_1z_0t_1 + 9x_1y_1z_0t_1 \\ & - 9x_0y_0z_1t_1 - 9x_1y_0z_1t_1 - 6x_0y_1z_1t_1 + 8x_1y_1z_1t_1.\end{aligned}$$

Using the algorithm in Remark 18, we get that the closest rank one multilinear form is

$$\begin{aligned}
&\ell_1(x)\ell_2(y)\ell_3(z)\ell_4(t),\\
&\ell_1(x) = 0.4799354720x_0 - 0.8773037918x_1\\
&\ell_2(y) = 0.2732019392y_0 - 0.9619567040y_1\\
&\ell_3(z) = 0.7563638894z_0 + 0.6541511043z_1\\
&\ell_4(t) = 0.3260948315t_0 + 0.9453370622t_1.
\end{aligned}$$

The value of the absolute maximum of ℓ is 16.71262553.

Example 26. Let $v \in \mathbb{R}^2 \otimes \mathbb{R}^3$ be the following tensor

$$v = 4x_0 \otimes y_0 - 9x_1 \otimes y_0 + 2x_0 \otimes y_1 + x_1 \otimes y_1 - 5x_0 \otimes y_2 - 7x_1 \otimes y_2.$$

Using the algorithm in Remark 18 we get that the closest rank one tensor is

$$\begin{aligned}
&(0.01162554952x_0 + 0.9999324213x_1) \otimes (-0.7821828869y_0\\
&\quad + 0.08939199251y_1 - 0.6166027924y_2).
\end{aligned}$$

In this case, we can check this result. The first singular vectors of the matrix

$$\begin{pmatrix} 4 & -9 \\ 2 & 1 \\ -5 & -7 \end{pmatrix},$$

are

$$\begin{aligned}
&(0.01162554952, 0.99993242102),\\
&(-0.7821828866, 0.08939199251, -0.6166027924).
\end{aligned}$$

Acknowledgments

The author was fully supported by CONICET, Argentina. The author would like to thank Federico Holik for very fruitful discussions and for presenting him with this subject. Thanks are also due to the reviewer for his detailed and useful suggestions.

Appendix. A general algorithm for a multilinear form

Let us give the algorithm to find the maximum value of a generic multilinear map over a product of spheres,

$$\ell : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r} \to \mathbb{R}^{n_{r+1}}, g \qquad \max_{\|x_1\|=\cdots=\|x_r\|=1} \|\ell(x_1, \ldots, x_r)\|.$$

Recall from Section 2, that we may assume that ℓ is a multilinear form,

$$\widehat{\ell} : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r} \times \mathbb{R}^{n_{r+1}} \to \mathbb{R}, \qquad \max_{\|x_1\|=\cdots=\|x_{r+1}\|=1} |\widehat{\ell}(x_1, \ldots, x_{r+1})|.$$

The following is a pseudocode in the trilinear case.

Input: A generic trilinear form $\ell : \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \times \mathbb{R}^{s+1} \to \mathbb{R}$.
Output: The maximum value $|\ell(x, y, z)|$, where $(x, y, z) \in \mathbb{S}^n \times \mathbb{S}^m \times \mathbb{S}^s$.

1. Compute the system of trilinear equations of Proposition 16.
2. Add the polynomial equations $\|x\|^2 = \|y\|^2 = \|z\|^2 = 1$.
3. Compute a Gröbner basis for the resulting system, $\mathcal{I}$.
4. Find a basis B of $\mathbb{C}[x_0, \ldots, x_n, y_0, \ldots, y_m, z_0, \ldots, z_s]/\mathcal{I}$.
5. Compute the multiplication matrix of ℓ.
6. Return the magnitude of the maximum real eigenvalue.

Let us give an implementation of the algorithm in Maple. The code computes the maximum of a trilinear form over $\mathbb{S}^{n-1} \times \mathbb{S}^{m-1} \times \mathbb{S}^{s-1}$. The reader may change the values of n, m and s and the trilinear form, to get different examples.

```
> restart;with(Groebner):with(linalg):
> n:=2:m:=2:s:=2:
> L:=6*x[1]*y[1]*z[1]+3*x[2]*y[1]*z[1]-6*x[1]*y[2]*z[1]+16*x[2]*y[2]*z[1]-
  14*x[1]*y[1]*z[2]-15*x[2]*y[1]*z[2]-11*x[1]*y[2]*z[2]+8*x[2]*y[2]*z[2];
> #Step 1 and 2
> J:={add(x[i]^2,i=1..n)-1,add(y[j]^2,j=1..m)-1,add(z[k]^2,k=1..s)-1,
  seq(seq(x[i]*diff(L,x[j])-x[j]*diff(L,x[i]),j=1..i-1),i=1..n),
  seq(seq(y[i]*diff(L,y[j])-y[j]*diff(L,y[i]),j=1..i-1),i=1..m),
  seq(seq(z[i]*diff(L,z[j])-z[j]*diff(L,z[i]),j=1..i-1),i=1..s)}:
> #Step 3
> G:=Basis(J,'tord'):
> #Step 4
> ns,rv:=NormalSet(G, tord):
> #Step 5
> mulMat:=evalm(evalf(MultiplicationMatrix(L,ns,rv,G,tord))):
> #Step 6
> max(op(map(abs,map(Re,{eigenvalues(mulMat)}))));
```

The following is a table that shows the time, in seconds, used to compute the maximum. In the first column appears different values of (n, m, s), in the second, the time used to compute the Steps 1 through 4, and in the third, the total time of the algorithm. We ran a Maple 11 session on a 2.1 GHz CPU, with 2 GB of memory. We used the previous code to find the maximum value of a generic trilinear form over $\mathbb{S}^{n-1} \times \mathbb{S}^{m-1} \times \mathbb{S}^{s-1}$,

(n, m, s)	Steps 1–4	Total time
(2, 2, 2)	0.03	0.33
(2, 2, 3)	0.05	0.79
(2, 2, 4)	0.09	0.99
(2, 2, 5)	0.14	1.20
(2, 3, 3)	0.31	7.13
(2, 3, 4)	0.89	30.03
(3, 3, 3)	5.06	397.28

Note that the computation of the multiplication matrix using Gröbner bases requires most of the time. A method to compute the multiplication matrix of ℓ without the need to use Gröbner bases is required.

References

[1] T.M. Apostol, Mathematical Analysis, second ed., Addison-Wesley Publishing Co., Reading, Mass., London, Don Mills, Ont, 1974.

[2] W. Auzinger, H.J. Stetter, An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations, in: Numerical Mathematics, Singapore 1988, in: Internat. Schriftenreihe Numer. Math., vol. 86, Birkhäuser, Basel, 1988, pp. 11–30.

[3] G. Ballard, T. Kolda, T. Plantenga, Efficiently computing tensor eigenvalues on a GPU, 2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and Ph.D. Forum, 2011, pp. 1340–1348.

[4] D.N. Bernstein, The number of roots of a system of equations, Funkcional. Anal. Priložen. 9 (1975) 1–4.

[5] D. Cartwright, B. Sturmfels, The number of eigenvalues of a tensor, Linear Algebra Appl. (2011).

[6] P. Comon, Tensor decompositions: state of the art and applications, in: Mathematics in Signal Processing, V (Coventry, 2000), in: Inst. Math. Appl. Conf. Ser. New Ser., vol. 71, Oxford Univ. Press, Oxford, 2002, pp. 1–24.

[7] D.A. Cox, J. Little, D. O'Shea, Using Algebraic Geometry, second ed., in: Graduate Texts in Mathematics, vol. 185, Springer, New York, 2005.

[8] L. De Lathauwer, B. De Moor, J. Vandewalle, A multilinear singular value decomposition, SIAM J. Matrix Anal. Appl. 21 (2000) 1253–1278 (electronic).

[9] V. de Silva, L.-H. Lim, Tensor rank and the ill-posedness of the best low-rank approximation problem, SIAM J. Matrix Anal. Appl. 30 (2008) 1084–1127.

[10] A. Dickenstein, I.Z. Emiris (Eds.), Solving Polynomial Equations: Foundations, Algorithms, and Applications, in: Algorithms and Computation in Mathematics, vol. 14, Springer-Verlag, Berlin, 2005.

[11] T.-C. Dinh, N. Sibony, Dynamics in several complex variables: endomorphisms of projective spaces and polynomial-like mappings, in: Holomorphic Dynamical Systems, in: Lecture Notes in Math., vol. 1998, Springer, Berlin, 2010, pp. 165–294.

[12] W. Fulton, Intersection Theory, second ed., in: Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics (Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics), vol. 2, Springer-Verlag, Berlin, 1998.

[13] I.M. Gel'fand, M.M. Kapranov, A.V. Zelevinsky, Discriminants, Resultants, and Multidimensional Determinants, in: Mathematics: Theory & Applications, Birkhäuser Boston Inc., Boston, MA, 1994.

[14] J. Harris, Algebraic Geometry, in: Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1992, A first course.

[15] R. Hartshorne, Algebraic Geometry, in: Graduate Texts in Mathematics, vol. 52, Springer-Verlag, New York, 1977.

[16] T.G. Kolda, J.R. Mayo, Shifted power method for computing tensor eigenpairs, SIAM J. Matrix Anal. Appl. 32 (2011) 1095–1124.

[17] L.-H. Lim, Singular values and eigenvalues of tensors: a variational approach, in: 2005 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, December 2005, pp. 129–132.

[18] A. McLennan, The maximum number of real roots of a multihomogeneous system of polynomial equations, Beiträge Algebra Geom. 40 (1999) 343–350.

[19] A.P. Morgan, A.J. Sommese, C.W. Wampler, A product-decomposition bound for Bezout numbers, SIAM J. Numer. Anal. 32 (1995) 1308–1325.

[20] G. Ni, L. Qi, F. Wang, Y. Wang, The degree of the E-characteristic polynomial of an even order tensor, J. Math. Anal. Appl. 329 (2007) 1218–1229.

[21] L. Qi, Eigenvalues and invariants of tensors, J. Math. Anal. Appl. 325 (2007) 1363–1377.