Convex Hulls of Orbits and Orientations of a Moving Protein Domain

Marco Longinetti · Luca Sgheri · Frank Sottile

Abstract We study the facial structure and Carathéodory number of the convex hull of an orbit of the group of rotations in $\mathbb{R}^3$ acting on the space of pairs of anisotropic symmetric 3×3 tensors. This is motivated by the problem of determining the structure of some proteins in an aqueous solution.

1 Introduction

The most aesthetically appealing polytopes arise as convex hulls of orbits of finite groups acting on a vector space. These include the platonic and Archimedean solids and their higher-dimensional generalizations, such as the regular polytopes [7]. In contrast, the analogous objects for compact Lie groups have not attracted much

M. Longinetti
Dipartimento ingegneria agraria e forestale, Università degli Studi di Firenze, Piazzale delle Cascine,
15, 50144 Firenze, Italia
e-mail: longinetti@diaf.unfi.it
url: wwwnt.unifi.it/diaf/nuovosito/withframes/docenti/longinetti/sitolonginetti/index.htm

L. Sgheri
IAC–CNR Sede di Firenze, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italia
e-mail: luca@fi.iac.cnr.it
url: www.fi.iac.cnr.it/iaga/luca.html

F. Sottile (✉)
Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA
e-mail: sottile@math.tamu.edu
url: www.math.tamu.edu/~sottile

study. We investigate convex hulls of orbits of the group $SO(3)$ in a particular 10-dimensional representation. Our motivation comes from an algorithm to understand the fold of some proteins.

Certain proteins, such as calmodulin [2], consist of two rigid domains connected via a region that is flexible in an aqueous solution (i.e. under physiological conditions), and the problem is to determine the relative position and orientation of these two domains. Calmodulin, as do many other proteins, incorporates metal ions into its structure. When a paramagnetic ion is substituted, it interacts with the magnetic field of dipoles within the protein via its magnetic susceptibility tensor χ. Part of this interaction, the residual dipolar coupling, may be inferred from nuclear magnetic resonance data and depends solely upon the relative orientation of the two domains. When the relative orientation of the two domains is not constant, we infer the mean magnetic susceptibility tensor $\overline{\chi}$ from these data.

We model this relative orientation by a probability measure p on the group $SO(3)$ of rotations of $\mathbb{R}^3$. Then $\overline{\chi}$ is the average with respect to p of rotations of χ. Recovering p from $\overline{\chi}$ is an ill-posed inverse problem. Nevertheless, $\overline{\chi}$ contains useful information about p. Gardner, Longinetti, and Sgheri [10] gave an algorithm to determine the maximum probability of a given relative orientation of the two domains. Since $\overline{\chi}$ lies in the convex hull V of the orbit of χ under the group of rotations of $\mathbb{R}^3$, it admits a representation $\overline{\chi} = \sum_j p_j R_j . \chi$, where the sum is finite, $\sum_j p_j = 1$ with $p_j \geq 0$, R_j is a rotation in $\mathbb{R}^3$, and $R_j . \chi$ is the action of R_j on the tensor χ. The minimal number of summands needed to represent any point $\overline{\chi}$ in V is the Carathéodory number of V.

It is often possible to substitute a different metal ion into the protein with a different susceptibility tensor χ'. Repeating the measurements gives a second mean tensor $\overline{\chi}'$ which is the average of rotations of χ' with respect to the measure p. Combining this with $\overline{\chi}$ gives more information about p. Longinetti, Luchinat, Parigi, and Sgheri [13] adapted the algorithm of [10] when there are two or more metal ions and showed how this can be used to better understand the structure of calmodulin. Their algorithm uses some knowledge of the convex hull $V^{1,2}$ of the orbit of the pair (χ, χ') under the group of rotations.

We study the Carathéodory number and facets of $V^{1,2}$. When χ and χ' are linearly independent, $V^{1,2}$ has dimension 10. We call a subgroup of $SO(3)$ which stabilizes a line in $\mathbb{R}^3$ a *coaxial group*, and a face of $V^{1,2}$ which is stabilized by a such a subgroup a *coaxial face*. Our main result is the following.

Theorem 6.6 *Faces of $V^{1,2}$ have dimension at most* 6. *The coaxial faces of $V^{1,2}$ form a three-dimensional family whose union is a nine-dimensional subset of the boundary of $V^{1,2}$ if and only if χ and χ' have distinct eigenvectors. In that case, almost all coaxial faces have dimension* 6, *have Carathéodory number* 4, *and are facets of $V^{1,2}$.*

Our main result implies that the Carathéodory number of $V^{1,2}$ is at most 8. This is an advance over [10], where it was bounded between 4 and 10 inclusive.

We are unable to show that the boundary of $V^{1,2}$ is the union of coaxial faces when χ and χ' have distinct eigenvectors, but conjecture that this is the case. As a

consequence of our main result, we also conjecture that the Carathéodory number of $V^{1,2}$ is at most 5.

Given any number N of tensors $(\chi_1, \ldots, \chi_N)$ we may define the convex hull $V^{1,\ldots,N}$. In Sect. 4.2 we prove that $\dim V^{1,\ldots,N}$ is five times the dimension of the span of $(\chi_1, \ldots, \chi_N)$. In the text, we will omit the superscripts from our notation for the convex hull.

Magnetic susceptibility tensors are 3×3 symmetric trace zero matrices and form a five-dimensional irreducible representation of the group $SO(3)$ of rotations in $\mathbb{R}^3$. More generally, one could study the convex hulls of orbits of compact groups. We were surprised to find that very little is known about such convex bodies, particularly their Carathéodory numbers and facets. We hope that our work will stimulate a more thorough study of convex hulls of orbits of compact groups.

In Sect. 2, we describe the motivation for this work from protein structure. In Sect. 3 we discuss group actions and in Sect. 4 convex hulls of orbits. In Sect. 5 we complete the analysis of [10] in the case of one metal ion. In Sect. 6 we analyze the case of two metal ions and deduce Theorem 6.6.

2 Application to Protein Structure

Proteins are large biological molecules synthesized by living organisms. The genome of an organism contains the chemical formulae for its proteins. Currently, hundreds of organisms (including man) have had their genomes mapped, and such chemical formulae are readily available. An important step toward inferring the biological function of a protein from its chemical formula is to determine its three-dimensional structure, or its *fold*.

About one third of all proteins incorporate metal ions into their structures. The fold of these proteins may be inferred from nuclear magnetic resonance, which can measure the interactions between paramagnetic metal ions and dipoles within the protein. The main quantities that can be measured are the pseudo contact shifts (PCS) [1] and the residual dipolar coupling (RDC) [17]. In this paper we deal only with the RDC.

The residual dipolar coupling between a paramagnetic ion and a dipole formed by atoms a and b within the protein depends upon the vector displacement r from the atom a to the atom b and the *magnetic susceptibility tensor* χ of the metal ion, which is a 3×3 symmetric matrix. The RDC interaction has the following vector formula

$$\delta := \frac{C}{\|r\|^5} r^T \chi r - \frac{C}{3\|r\|^3} \text{Trace}(\chi). \tag{2.1}$$

Here, C is a constant and $\|r\|$ is the length of the vector r. This depends only upon the relative orientation of the dipole and metal ion, and so the RDC data may be used to infer this relative orientation.

Writing $\chi = \chi_0 + \frac{1}{3}\text{Trace}(\chi)I_3$, where I_3 is the 3×3 identity matrix and χ_0 is the trace-free or *anisotropic* part of χ, this formula becomes

$$\delta = \frac{C}{\|r\|^5} r^T \chi_0 r.$$
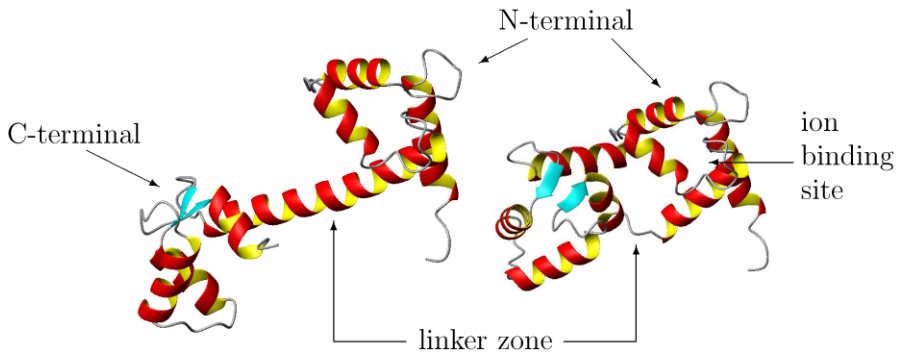
Fig. 1 Two orientations of calmodulin

We assume henceforth that $\chi = \chi_0$ is anisotropic.

The fold of the protein is usually unique, in the sense that small variations of the shape are allowed. There are proteins, however, that exhibit large variations of shape under particular conditions. A widely studied example is calmodulin, which has two rigid domains, called the *N-terminal* and *C-terminal* domains, connected by a short flexible linker. The N- and C-terminal domains are assumed to be rigid bodies with known structures. Figure 1, obtained with Molmol [11], shows calmodulin in two different orientations.

The binding site of the metal ion in calmodulin belongs to the N-terminal domain. The measured RDC of pairs of atoms belonging to the N-terminal domain can be used to obtain a good estimate of χ. The measured RDC of pairs of atoms belonging to the C-terminal domain can be used to study the relative orientation of the two domains.

Let us model the relative orientation of the N- and C-terminal domains with a rotation R. Then there is an unknown probability measure p on the set $SO(3)$ of rotations such that the *mean RDC $\bar{\delta}$* of the pair of atoms a, b in the C-terminal domain is given by

$$\bar{\delta} = \frac{C}{\|r\|^5} \int_{SO(3)} (Rr)^T \chi (Rr) \, dp(R) = \frac{C}{\|r\|^5} r^T \overline{\chi} r, \qquad (2.2)$$

where the *mean magnetic susceptibility tensor $\overline{\chi}$* is

$$\overline{\chi} = \int_{SO(3)} R^T \chi R \, dp(R). \qquad (2.3)$$

This tensor $\overline{\chi}$ can be estimated from the RDC of several dipole pairs in the C-terminal domain. The experimental measures show that in terms of difference of eigenvalues, $\overline{\chi}$ is between 5 and 20 times smaller than χ [3]. This indicates that p is not a point mass, that is, the C-terminal domain moves with respect to the N-terminal domain.

The availability of N distinct mean susceptibility tensors $\overline{\chi}_k$ with respect to different metal ions $k = 1, \ldots, N$ increases the information about p for N up to 5; see, for instance, [14] and [13, Theorem 3.2]. However, even the exact knowledge of five mean tensors $\overline{\chi}_k$ (i.e., 25 real numbers) does not allow the exact reconstruction of the probability measure p.

An approach to extract information from the mean tensors is to define $p_{\max}(R)$ as the maximal fraction of time that the C-terminal can stay in a particular orientation R, yet still produce the measured mean tensors. Orientations with a large $p_{\max}$ agree with what are thought to be the most-favored orientations of the C-terminal [13].

The calculation of $p_{\max}$ can be performed geometrically if only RDC is considered [13]. In the combined PCS+RDC case more information is added [4], however, the calculation of $p_{\max}$ can only be performed numerically. For the convergence and efficiency of the algorithm, the minimal number of orientations needed to reconstruct any admissible set of mean tensors $\overline{\chi}_k$ should be used. Experience suggests that adding the PCS data does not increase the actual number of orientations needed [4], so the Carathéodory number for the RDC case may be used as a basis for the numerical minimization.

3 Group Actions and Anisotropic Tensors

We first recall some basics about representations of compact groups, then consider the action of the group $SO(3)$ of rotations in $\mathbb{R}^3$ on the five-dimensional space of anisotropic 3×3 symmetric tensors, and finally investigate the coaxial subgroups of $SO(3)$.

3.1 Representations of Compact Groups

This material may be found in the book by Bröcker and tom Dieck [5, Chap. II]. Let G be a compact group, such as $SO(3)$. A *representation* of G is a finite-dimensional vector space W on which G acts by linear transformations. That is, we have a group homomorphism $\rho: G \to GL(W)$, where $GL(W)$ is the group of invertible linear transformations on W. For $g \in G$ and $w \in W$, write $g.w$ for $\rho(g)(w)$.

A representation W of G is *irreducible* if its only G-invariant subspaces are $\{0\}$ and W. Every representation of G decomposes as the direct sum of irreducible representations which is unique in the following way. Given a representation W of G and a positive integer l, let W^l be the l-fold direct sum of W,

$$W^l = \underbrace{W \oplus W \oplus \cdots \oplus W}_{l}.$$

Elements g of G act diagonally on elements $w = (w_1, \ldots, w_l)$ of W^l, $g.w = (g.w_1, \ldots, g.w_l)$. Suppose that $W_1, W_2, \ldots$ is the list of irreducible representations of G. If U is a representation of G, then there exist unique integers $l_1, l_2, \ldots$ such that

$$U \simeq W_1^{l_1} \oplus W_2^{l_2} \oplus W_3^{l_3} \oplus \cdots,$$

as representations of G. If U_i is the subrepresentation of U mapped to the summand $W_i^{l_i}$ under this isomorphism, then U_i does not depend on any choices and is called the *isotypical component* of U corresponding to W_i. If $l_i > 0$, then we say that U *contains* W_i. Furthermore, if $U' \subset U$ is a subrepresentation, then the ith isotypical component of U' is $U' \cap U_i$, which is also the image of U' under the projection to U_i.

Haar measure is a G-invariant measure μ on G with $1 = \int_G d\mu(g)$. Given a linear function $L: W \to \mathbb{R}$, where W is a representation of G, we may average L over orbits of G to get a new linear function L', defined by

$$L'(x) := \int_G L(g.x) \, d\mu(g).$$

Since L' is constant on orbits of G, it is G-invariant. This association $L \mapsto L'$ is called the *Reynolds operator*. It is an important tool for analyzing G-representations.

Another key tool is Schur's lemma. A linear map $\varphi: W \to U$ between representations of G is a *G-map* if for all $w \in W$ and $g \in G$, we have $g.\varphi(w) = \varphi(g.w)$. Let $\mathrm{Hom}_G(W, U)$ be the space of G-maps. A *division algebra* is a finite-dimensional associative algebra in which every nonzero element is invertible.

Schur's Lemma *If $W \not\cong U$ are irreducible representations of G, then $\mathrm{Hom}_G(W, U) = 0$ and $\mathrm{Hom}_G(W, W)$ is a division algebra that contains $\mathbb{R}$.*

Proof Let $\varphi: W \to U$ be a G-map. Then the kernel of φ is a subrepresentation of W and so it is either 0 or W, and the image of φ, which is a subrepresentation of U, is either 0 or U. Examining the possibilities leads to the conclusions. $\qquad\square$

There are exactly three division algebras which contain $\mathbb{R}$: The real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$, and the quaternions $\mathbb{H}$. An irreducible representation W of G has *real*, *complex*, or *quaternionic type*, depending on $\mathrm{End}_G(W) := \mathrm{Hom}_G(W, W)$.

Example 3.1 Consider the group $SO(2)$ of rotations of $\mathbb{R}^2$,

$$SO(2) = \left\{ R_\theta := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mid \theta \in [0, 2\pi) \right\}.$$

$\mathrm{End}_{SO(2)}(\mathbb{R}^2)$ consists of those 2×2 matrices M such that $M R_\theta = R_\theta M$, and so

$$\mathrm{End}_{SO(2)}(\mathbb{R}^2) = \mathbb{R} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This is isomorphic to $\mathbb{C}$ (we send $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ to $\sqrt{-1}$), so this representation of $SO(2)$ has complex type. It is the *defining representation* U_1 of $SO(2)$. If we identify $\mathbb{R}^2$ with $\mathbb{C}$ and $SO(2)$ with the circle group $S^1 := \{e^{i\theta} \mid 0 \le \theta < 2\pi\}$, then the action is scalar multiplication by elements of S^1. For any positive integer $k > 0$, let U_k be the representation of S^1 on $\mathbb{C}$ (identified with $\mathbb{R}^2$) where $z \in S^1$ acts as multiplication by z^k. These all have complex type.

Example 3.2 The orthogonal group $O(2)$ contains $SO(2)$ as well as the cosets of reflections

$$SO(2) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \left\{ R_\theta \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \mid \theta \in [0, 2\pi) \right\}.$$

The defining representation U_1 of $O(2)$ on $\mathbb{R}^2$ has real type, as

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

and so $\mathrm{End}_{O(2)}(\mathbb{R}^2) = \mathbb{R} \cdot \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \simeq \mathbb{R}$.

In the *trivial representation* $U_0 = \mathbb{R}$ of $O(2)$, elements act as multiplication by 1. For a positive integer k, define the map $\varphi_k \colon O(2) \to O(2) \subset GL(2, \mathbb{R})$ by $\varphi_k(R_\theta) = R_{k\theta}$ and $\varphi_k\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. This defines the representation U_k of $O(2)$, which has real type. Restricting to $SO(2)$ gives its representation U_k of complex type.

3.2 Rotations of Anisotropic Tensors

Let $e_1 = (1, 0, 0)^T$, $e_2 = (0, 1, 0)^T$, and $e_3 = (0, 0, 1)^T$ be the standard basis of $\mathbb{R}^3$.

The *special orthogonal group* $SO(3)$ is the group of rotations in $\mathbb{R}^3$. It consists of 3×3 real orthogonal matrices with determinant 1,

$$SO(3) := \left\{ R \in \mathbb{R}^{3 \times 3} \mid RR^T = 1 \text{ and } \det R = 1 \right\}.$$

Let $R \in SO(3)$, and let $T_R SO(3)$ be the tangent space to $SO(3)$ at the matrix R. Let I be the identity matrix, then $T_I SO(3)$ is the space $\mathfrak{so}_3$ of skew symmetric 3×3 matrices, which is the Lie algebra of $SO(3)$. That is,

$$T_I SO(3) = I + \mathfrak{so}_3.$$

Elements $R \in SO(3)$ act on 3×3 symmetric matrices (tensors) χ by conjugation, $R.\chi := R\chi R^T$.[1] This preserves the trace of χ, and so $SO(3)$ acts on the space W of *anisotropic* (trace-zero) tensors, a five-dimensional irreducible real representation. We introduce some useful coordinates for W. A point $(v, w, x, y, z) \in \mathbb{R}^5$ corresponds to

$$\chi(v, w, x, y, z) := \begin{pmatrix} v & 0 & 0 \\ 0 & -\frac{v}{2} & 0 \\ 0 & 0 & -\frac{v}{2} \end{pmatrix} + \begin{pmatrix} 0 & w & x \\ w & 0 & 0 \\ x & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & y & z \\ 0 & z & -y \end{pmatrix}. \quad (3.3)$$

Observe that e_1 is an eigenvector for $\chi(v, w, x, y, z)$ if and only if $w = x = 0$, e_2 is an eigenvector if and only if $w = z = 0$, and e_3 is an eigenvector if and only if $x = z = 0$.

3.3 Coaxial Subgroups

The *coaxial subgroup* Q_e is the set of rotations fixing a line in $\mathbb{R}^3$ with direction e. It is isomorphic to the orthogonal group $O(2)$. Its identity component Q_e^+ is isomorphic

---

[1] This left action (if $R, S \in SO(3)$, then $R(S.\chi) = RS.\chi$) is equivalent to the implied action in (2.3).

to $SO(2)$ and consists of rotations about the axis e, while the other component Q_e^- consists of reflections in axes orthogonal to e. For example, let

$$R_{e_1,\theta} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad \text{and} \quad R_{e_3,\pi} := \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad (3.4)$$

If we fix f perpendicular to e and let θ run over all angles, then $R_{e,\theta}$ and $R_{e,\theta} R_{f,\pi}$ give all elements of Q_e.

We consider the action of a coaxial subgroup Q_e on W. For this, suppose that $e = e_1$ and let $R_{e_1,\theta}$ act on $\chi(v, w, x, y, z)$. This gives the tensor $\chi(v', w', x', y', z')$, where

$$v' = v,$$

$$\begin{pmatrix} w' \\ x' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} w \\ x \end{pmatrix} = R_\theta \begin{pmatrix} w \\ x \end{pmatrix}, \quad \text{and}$$

$$\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = R_{2\theta} \begin{pmatrix} y \\ z \end{pmatrix}.$$

Thus $R_{e_1,\theta}$ acts trivially on the coordinate v, by rotation through the angle θ on the vector $(w, x)^T$, and by rotation through 2θ on the vector $(y, z)^T$. Note that $R_{e_3,\pi}$ sends $\chi(v, w, x, y, z)$ to $\chi(v, w, -x, y, -z)$. Thus, if we restrict the action of $SO(3)$ on W to $Q_{e_1} \simeq O(2)$, then it decomposes as a sum of irreducible representations

$$W = U_0 \oplus U_1 \oplus U_2. \qquad (3.5)$$

This decomposition corresponds to the coordinates (3.3). Projection to the trivial submodule $U_0 = \mathbb{R}$ is, up to a scalar multiple, the unique Q_e-invariant linear function $L \colon W \to \mathbb{R}$.

4 Convex Hulls of Orbits

Let $\mathcal{O}$ be an orbit of a compact group G in a representation W of G. The *convex hull* V of $\mathcal{O}$ is all points of W which are convex combinations of elements of $\mathcal{O}$,

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n,$$

where $v_1, \ldots, v_n \in \mathcal{O}$, and the nonnegative numbers λ_i have sum 1. The set V is a compact convex set, hence a convex body.

4.1 Faces and Carathéodory Number of V

A *face* F of V is the subset of its boundary where some linear function L achieves its maximum on V,

$$F := \{v \in V \mid L(v) \geq L(u) \text{ for all } u \in V\}.$$

We say that L *supports* F and also that the hyperplane $L(x) = L(F)$ *supports* F. (Here, $L(F)$ is a constant.) The tangent spaces to $\mathcal{O}$ of its points lying in F are contained in any hyperplane supporting F. Such a tangent space $T_v \mathcal{O}$ at the point v is

$$v + \mathfrak{g}.v,$$

where the action of the Lie algebra $\mathfrak{g}$ is the derivative of the action of G. The face F is *proper* if $F \neq V$. When V is full-dimensional so that $\dim V = \dim W$, this is equivalent to $L \neq 0$. A *facet* is a maximal face and a *vertex* is a minimal face. Vertices are not a convex combination of other points of V.

Lemma 4.1 *The vertices of V are exactly the points of $\mathcal{O}$.*

Proof The vertices of V are a subset of $\mathcal{O} := G.x$. Let $g.x \in \mathcal{O}$ and suppose that it is a convex combination of vertices,

$$g.x = \lambda_1 g_1.x + \lambda_2 g_2.x + \cdots + \lambda_n g_n.x.$$

Multiplying by $g_1 g^{-1}$ expresses the vertex $g_1.x$ as a convex combination of points of $\mathcal{O}$. Thus $n = 1$ and $g.x = g_1.x$ is a vertex. $\square$

When V has dimension d, Carathéodory's Theorem [6] (see, e.g., [16, Theorem 1.1.4]) implies that any point x of V is a convex combination of at most $d+1$ vertices. The *Carathéodory number* of V is the minimum number n such that any point $x \in V$ is a convex combination of at most n vertices. For example, a ball in $\mathbb{R}^d$ has Carathéodory number 2, while a d-simplex has Carathéodory number $d + 1$. Fenchel [9] (see, e.g., [15, Theorem 1.4]), showed that the Carathéodory number is at most d when the set of vertices is connected. More useful for us is a recursive bound, which is immediate from the observation that any point of V is the convex combination of any vertex and some boundary point.

Lemma 4.2 *The Carathéodory number of a convex body V is at most one more than the maximal Carathéodory number of its facets.*

Suppose that S is a closed (hence compact) subgroup of G that stabilizes a face F of V, that is $s.F = F$ for all $s \in S$.

Lemma 4.3 *When F is proper, there is a nonzero S-invariant linear function on W, and W contains the trivial representation of S.*

Proof Let $L \colon W \to \mathbb{R}$ be any linear function supporting F with $L(F) = \ell$. Let L' be the image of L under the Reynolds operator for S. For $u \in F$, we have $L'(u) = \ell$, as F is S-stable and $L'(u)$ is the average of L over the orbit of S through u.

Suppose that $w \in V \setminus F$. Then $L(w) < L(u) = \ell$ and $L(s.w)$ is bounded away from ℓ as $S.w$ is compact and disjoint from F. In particular, this implies that $L'(w) < L'(u)$, which shows that the S-invariant linear function L' supports F and that $L' \neq 0$. Such an S-invariant linear function must factor through the trivial isotypical component of W as a representation of S. This completes the proof. $\square$

4.2 The Dimension of V

If $w = (w_1, \ldots, w_l) \in W^l$, then we write $d(w)$ for the dimension of the linear span of the components $w_1, \ldots, w_l$ of w in W.

Lemma 4.4 *Suppose that W is an irreducible representation of a group G having real type. If $w \in W^l$, then the linear span U of the orbit $G.w$ in W^l is isomorphic to $W^{d(w)}$.*

Proof Write $k := d(w)$. We may assume that $w_1, \ldots, w_k$ form a basis for the linear span of $w_1, \ldots, w_l$. Let $A = (\alpha_{ij}) \in \mathrm{Mat}_{l \times k}(\mathbb{R})$ be the matrix which writes the components of w in terms of this basis, $w_i = \sum_{j=1}^{k} \alpha_{i,j} w_j$ for $i = 1, \ldots, l$. For each $i = 1, \ldots, l$, let $\varphi_i \colon U \to W$ be the projection to the ith coordinate. Since $\varphi_i(g.w) = g.w_i$, we have

$$\varphi_i = \sum_{j=1}^{k} \alpha_{i,j} \varphi_j \quad \text{for } i = 1, \ldots, l. \tag{4.5}$$

This matrix A defines a G-map $A \colon W^k \to W^l$ by

$$A \colon (w_1, \ldots, w_k) \longmapsto \left(\sum_j \alpha_{1,j} w_j, \sum_j \alpha_{2,j} w_j, \ldots, \sum_j \alpha_{l,j} w_j \right). \tag{4.6}$$

Composing the map $\psi := (\varphi_1, \ldots, \varphi_k) \colon U \to W^k$ with $A \colon W^k \to W^l$ gives the identity map on U: By (4.5), for $w \in U$, we have

$$w = \big(\varphi_1(w), \varphi_2(w), \ldots, \varphi_l(w) \big).$$

We show that the map (4.6) is injective and thus ψ is an isomorphism. A linear map $L \colon W \to \mathbb{R}$ induces maps $L^k \colon W^k \to \mathbb{R}^k$ and $L^l \colon W^l \to \mathbb{R}^l$, which commute with A. If $0 \neq w \in W^k$, then there is some linear map $L \colon W \to \mathbb{R}$ with $L^k(w) \neq 0$. Since A has full rank k, $A(L^k(w)) \neq 0$. But this implies that $A(w) \neq 0$, as $A(L^k(w)) = L^l(A(w))$. $\square$

Lemma 4.7 *Suppose that $W = W_1^{l_1} \oplus \cdots \oplus W_m^{l_m}$ is the decomposition of a representation W of G into isotypical pieces, each of which has real type. Let $w = (w_1, \ldots, w_m) \in W$, where w_i is the component of w in $W_i^{l_i}$. Then the dimension of the convex hull V of the orbit $G.w$ is*

$$\sum_{W_i \neq \mathbb{R}} d(w_i) \cdot \dim W_i.$$

If W does not contain the trivial representation, then 0 lies in V.

Proof If W contains the trivial representation, assume that it is W_1. Since $g.(w - w_1) = g.w - w_1$, we see that the orbits $G.w$ and $G.(w - w_1)$ are isomorphic,

and the same is true for their convex hulls. Thus it is no loss to suppose that $w_1 = 0$, which is equivalent to assuming that W does not contain the trivial representation.

The linear span U of the orbit $G.w$ is the direct sum of its projections to the isotypical components $W_i^{l_i}$ of W. Each projection is the linear span of $G.w_i$, which by Lemma 4.4 is isomorphic to $W_i^{d(w_i)}$. Thus U has dimension $\sum_i d(w_i) \dim W_i$. We may replace W by this linear span, and therefore assume that the orbit $G.w$ spans W.

The convex hull of $G.w$ lacks full dimension only if it lies in some hyperplane H not containing the origin. Suppose that this is not the case and let B be the convex hull of $G.w$ and the origin. Then V is a proper G-stable face of B and so by Lemma 4.3 W contains the trivial representation, which is a contradiction.

If $0 \notin V$, then there is some linear function L which is bounded above 0 on V. But then the image L' of L under the Reynolds operator is nonzero on V. This implies that $L' \neq 0$, and so W contains the trivial representation, a contradiction. $\square$

Example 4.8 Lemmas 4.4 and 4.7 do not hold if the representation W has complex type. For example, let $G = SO(2)$ and $W = U_k^l$ with $k, l \geq 1$. Identifying W with $\mathbb{C}^l$ and $SO(2)$ with the circle group S^1, elements $z \in S^1$ act on $\mathbb{C}^l$ as scalar multiplication by z^k. Thus the linear span $SO(2).w$ for $w \in U_k^l$ is a complex line, and therefore has real dimension 2, and not 4 as Lemma 4.4 predicts for general $w \in W$ when $l, k \geq 1$.

In particular, if $W = U_1 \oplus U_2^2$, and $w \in W$ is general, then the linear span of $G.w$ has complex dimension 2 and thus real dimension 4.

5 One Metal Ion

Let W be the space of symmetric 3×3 anisotropic tensors, a five-dimensional irreducible representation of $SO(3)$ of real type. For each unit vector $e \in \mathbb{R}^3$, there is a linear function

$$L_e \colon W \ni \chi \longmapsto \langle e, \chi e \rangle = e^T \chi e \in \mathbb{R}. \tag{5.1}$$

If e is an eigenvector of χ, then $L_e(\chi)$ is its eigenvalue. In general, $L_e(\chi)$ lies between the maximum and minimum eigenvalues of χ. Note that L_e is Q_e-invariant. By the decomposition (3.5) of W into irreducible Q_e representations, any Q_e-invariant linear function is a scalar multiple of L_e.

As a matrix, a tensor in W has real eigenvalues and its eigenvectors form an orthonormal basis for $\mathbb{R}^3$. Fix a nonzero anisotropic tensor $\chi \in W$ with maximum eigenvalue $M > 0$ and minimum eigenvalue $m < 0$. The intermediate eigenvalue of χ is $-M - m$, and we have $-\frac{M}{2} \geq m \geq -2M$. The orbit $\mathcal{O}$ of χ under $SO(3)$ consists of the anisotropic tensors with maximal eigenvalue M and minimal eigenvalue m. It is a manifold whose dimension we determine.

Proposition 5.2 *The orbit $\mathcal{O}_\chi$ is three-dimensional unless χ has an eigenvalue of multiplicity 2, in which case it is two-dimensional.*

Proof The dimension of $\mathcal{O}$ is equal to the dimension of any of its tangent spaces. Since

$$\mathcal{O} = \{ R\chi R^T \mid R \in SO(3) \},$$

the tangent space $T_\xi \mathcal{O}$ at a point $\xi \in \mathcal{O}$ is the affine space

$$\xi + \left\{ r\xi + \xi r^T \mid r \in \mathfrak{so}_3 \right\}. \tag{5.3}$$

Indeed, consider the action in ξ of an element $I + r$ of the tangent space $T_I SO(3) = I + \mathfrak{so}_3$:

$$(I + r)\xi(I + r)^T = \xi + r\xi + \xi r^T + r\xi r^T.$$

Discarding the term which is quadratic in $\mathfrak{so}_3$ gives (5.3).

It suffices to determine the tangent space to $\mathcal{O}$ at the point χ. We may suppose that χ is diagonal, and let r be a general element of $\mathfrak{so}_3$,

$$\chi = \begin{pmatrix} M & 0 & 0 \\ 0 & -M - m & 0 \\ 0 & 0 & m \end{pmatrix} \quad \text{and} \quad r = \begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix},$$

where $a, b, c \in \mathbb{R}$. Let $\alpha := M + \frac{m}{2}$, $\beta := M - m$, and $\gamma := -m - \frac{M}{2}$. Then $\beta > 0$ and $\alpha, \gamma \geq 0$ with $\alpha = 0$ only when the eigenvalue m has multiplicity 2 and $\gamma = 0$ only when the eigenvalue M has multiplicity 2. We see that $T_\chi \mathcal{O}$ is the affine subspace of W

$$\xi + \begin{pmatrix} 0 & 2a\alpha & b\beta \\ 2a\alpha & 0 & 2c\gamma \\ b\beta & 2c\gamma & 0 \end{pmatrix}, \tag{5.4}$$

where $a, b, c \in \mathbb{R}$. This is three-dimensional unless either $\alpha = 0$ or $\gamma = 0$. $\qquad \square$

Let V be the convex hull of the orbit $\mathcal{O}$ of $\chi \in W$. By Lemma 4.4, this is a five-dimensional convex body.

Lemma 5.5 *If $\overline{\chi} \in V$ and $e \in \mathbb{R}^3$ is a unit vector, then we have*

$$m \leq L_e(\overline{\chi}) \leq M. \tag{5.6}$$

In fact, V is the set of symmetric anisotropic tensors satisfying (5.6) [10, Theorem 3.3].

Proof $L_e(\overline{\chi})$ lies between the maximum and minimum eigenvalues of $\overline{\chi}$. Thus the inequality (5.6) holds for $\overline{\chi}$ in the orbit of χ. Since a general element of V is a convex combination of tensors in the orbit of χ, we deduce (5.6). $\qquad \square$

A *coaxial face* of V is a face that is stabilized by some coaxial subgroup Q_e. By Lemma 4.3, a coaxial face stabilized by Q_e is supported by a nontrivial Q_e-invariant linear function. As we noted earlier, this linear function is necessarily a scalar multiple of L_e. By the inequality (5.6), there are two possibilities for such a coaxial face,

$$\begin{aligned} F_e^M &:= \left\{ \overline{\chi} \in V \mid L_e(\overline{\chi}) = M \right\}, \quad \text{and} \\ F_e^m &:= \left\{ \overline{\chi} \in V \mid L_e(\overline{\chi}) = m \right\}. \end{aligned} \tag{5.7}$$

The coaxial face F_e^M consists of tensors $\overline{\chi} \in V$ having e as an eigenvector with eigenvalue M and tensors in F_e^m have e as an eigenvector with eigenvalue m.

We now describe the facets of V and determine its Carathéodory number. As in the proof of Proposition 5.2, set $\alpha := M + \frac{m}{2} \geq 0$ and $\gamma := -m - \frac{M}{2} \geq 0$.

Lemma 5.8 *The boundary of V is the union of coaxial faces* (5.7) *where e ranges over all unit vectors in $\mathbb{R}^3$. A nonempty intersection of two coaxial faces lies in the orbit of $SO(3)$. Each face F_e^M is a circle of radius γ and each face F_e^m is a circle of radius α. When χ has a repeated eigenvalue so that either α or γ vanishes, then the corresponding coaxial face degenerates to a point.*

A consequence of Lemma 5.8 is that the coaxial faces are maximal faces, and are therefore facets.

Proof By Theorem 3.3 of [10], V is the set of anisotropic tensors whose eigenvalues lie in the interval $[m, M]$, and so its boundary consists of tensors $\overline{\chi}$ either having maximal eigenvalue M or having minimal eigenvalue m. This shows that the boundary of V consists of coaxial faces, which are thus the facets of V.

We show that the intersection of two coaxial faces lies in the orbit of χ. Suppose that $\overline{\chi}$ lies on two different coaxial faces. If these are F_e^M and F_f^m, then e and f are eigenvectors of χ with eigenvalues M and m, respectively. The third eigenvalue of $\overline{\chi}$ is $-M - m$, and so $\overline{\chi}$ lies in the orbit of χ. If the two faces have the form F_e^M and F_f^M with e and f linearly independent, then the eigenvalue M of $\overline{\chi}$ has multiplicity 2 and its third (smallest) eigenvalue is $-2M$. Since $\overline{\chi} \in V$, this smallest eigenvalue is bounded below by m; as $m \geq -2M$, we see that $m = -2M$ and so again $\overline{\chi}$ lies in the orbit of χ. The argument is similar if the two faces are F_e^m and F_f^m.

The coaxial face F_e^M of V consists of tensors $\overline{\chi} \in V$ having e as an eigenvector with eigenvalue M. Since each point of the boundary of F_e^M lies in some other coaxial face, this boundary lies in the orbit of χ and is necessary an orbit of Q_e. We need only consider the case when $\chi \in F_e^M$ so that this boundary is $Q_e.\chi$. Suppose that $e = e_1$ and $\chi = \chi(M, 0, 0, \gamma, 0)$, in the coordinates (3.3). Here, $\gamma = -\frac{M}{2} - m \geq 0$. As in Sect. 3.3, elements of Q_e act on χ by rotation of the vector $(\gamma, 0)$ formed by the last two coordinates, and thus $F_{e_1}^M$ is a circle of radius γ, which degenerates to a point if $\gamma = 0$.

We omit the similar arguments for F_e^m. $\square$

Theorem 5.9 *If zero is not an eigenvalue of χ, then V has Carathéodory number 3, and when zero is an eigenvalue, V has Carathéodory number 2.*

When zero is not an eigenvalue of χ, this is the main result about V from [10].

Proof First, suppose that zero is not an eigenvalue of χ. Every facet has Carathéodory number 2, as it is a circle. So by Lemma 4.2, V has Carathéodory number either 2 or 3. By Lemma 4.7, 0 lies in V. If V has Carathéodory number 2, then there exist $\lambda \in [0, 1]$ and $R, S \in SO(3)$ with $0 = \lambda R.\chi + (1 - \lambda)S.\chi$. Multiplying by R^{-1}, this
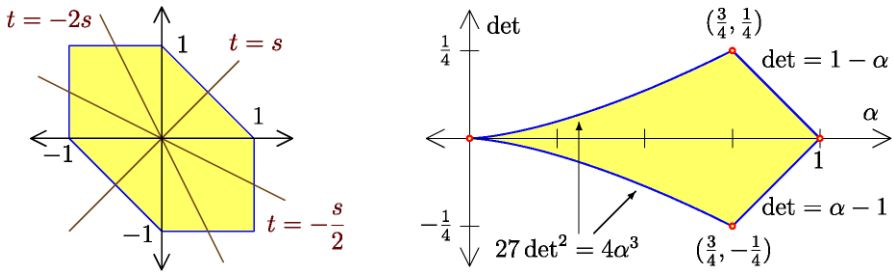
Fig. 2 Eigenvalues and invariants of tensors in V

becomes $0 = \lambda \chi + (1 - \lambda) R.\chi$, for a different rotation $R \in SO(3)$, and so

$$-\lambda \chi = (1 - \lambda) R.\chi. \tag{5.10}$$

Suppose that χ is diagonal. Then (5.10) implies that $R.\chi$ is also diagonal.

If $M \neq -m$ so that 0 is not an eigenvalue of χ, then one of the diagonal matrices $-\lambda \chi$ and $(1 - \lambda) R.\chi$ has two positive entries and the other has two negative entries, which is a contradiction. Thus if 0 is not an eigenvalue of χ, then V has Carathéodory number 3.

Now we assume that 0 is an eigenvalue of χ. We will show that the image of the map $[0, 1] \times SO(3) \times SO(3) \to V$ which takes (λ, R, S) to $\lambda R.\chi + (1 - \lambda) S.\chi$ meets each $SO(3)$-orbit in V and is therefore surjective. They key point is that two tensors are in the same orbit if and only if they have the same characteristic polynomial.

The characteristic polynomial of a trace-zero matrix $\overline{\chi}$ with eigenvalues s, t, $-s - t$ is

$$x^3 - x(st + t^2 + s^2) + (s^2 t + st^2).$$

The constant term is $- \det(\overline{\chi})$, while the coefficient $-\alpha$ of x is the sum of the pairwise products of eigenvalues, which is an invariant of the matrix.

Scaling χ, we may assume that its eigenvalues are 1, 0, and -1, so that V consists of tensors $\overline{\chi} \in W$ with eigenvalues $s, t, -s - t$ lying in the interval $[-1, 1]$. The set of such pairs (s, t) are the points of the hexagon of Fig. 2. The three lines through the origin $t = s$, $t = -s/2$, and $t = 2s$ divide the hexagon into six quadrilaterals and permutations of the eigenvalues permute these quadrilaterals. We leave it to the reader to check that $(s, t) \mapsto (\alpha, \det)$ is a one-to-one mapping of each quadrilateral onto the region shown in Fig. 2, which is

$$X := \left\{ (\alpha, \det) \in \mathbb{R}^2 \mid 27 \det^2 \leq 4\alpha^3 \text{ and } \alpha \leq 1 - |\det| \right\}, \tag{5.11}$$

and is bounded by the curves $\det = 1 - \alpha$, $\det = \alpha - 1$, and $27 \det^2 = 4\alpha^3$.

Consider matrices of the form $\chi(\lambda, \theta, \tau) := \lambda R(\theta).\chi + (1 - \lambda) S(\tau).\chi$, where

$$R(\theta) := \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad S(\tau) := \begin{bmatrix} \cos\tau & 0 & \sin\tau \\ 0 & 1 & 0 \\ -\sin\tau & 0 & \cos\tau \end{bmatrix}.$$
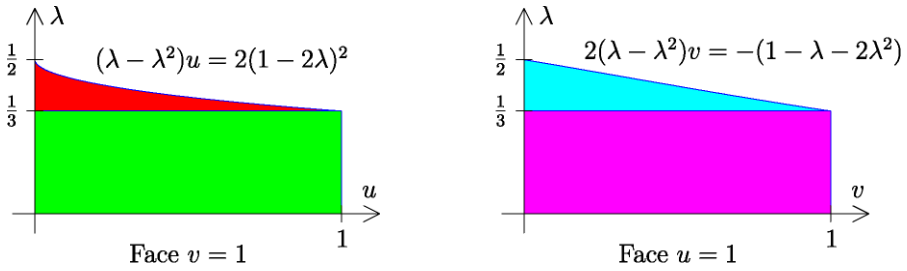
Fig. 3 Subsets of faces of the cube

The invariants $(\alpha, \det)$ of $\chi(\lambda, \theta, \tau)$ are

$$\left(1 - \lambda(1-\lambda)\left(4\sin^2\tau + \sin^2\theta - 2\sin^2\tau\sin^2\theta\right), \; \lambda(1-\lambda)\sin^2\theta\left(1 - 2\lambda\sin^2\tau\right)\right).$$

If we let $u = \sin^2\theta$ and $v = \sin^2\tau$, then the set of invariants of $\chi(\lambda, \theta, \tau)$ for all (λ, θ, τ) are the image of the unit cube $[0,1]^3$ under the map

$$f: (\lambda, u, v) \longmapsto \left(1 - \lambda(1-\lambda)(4v + u - 2uv), \; \lambda(1-\lambda)u(1 - 2\lambda v)\right).$$

We show that the image of f includes that part of X (5.11) where $\det \geq 0$. This will complete the proof, as replacing χ by $-\chi = S(\frac{\pi}{2}).\chi \in V$ in our definition of f changes the sign of the determinant and does not change the invariant α.

Figure 3 shows subsets of the faces $v = 1$ and $u = 1$ of the cube which include the segments $\lambda = 1/3$ and have boundaries the indicated curves. The map f is one-to-one on the interior of each region, and the images cover that part of X with $\det \geq 0$, meeting only along the curve $\det = \sqrt{\frac{4}{27}(\alpha - \frac{1}{4})(\alpha - 1)^2}$ for $\frac{1}{3} \leq \alpha \leq 1$, as shown in Fig. 4. The line $3\det = (\alpha - \frac{1}{9})$, which is tangent to the boundary curve $27\det^2 = 4\alpha^3$ at the point $(\frac{1}{3}, \frac{2}{27})$, is the image of the lines $\lambda = \frac{1}{3}$ in Fig. 3. This completes the proof.   □

Finally, we identify the hyperplanes supporting facets of V. By Lemma 5.8, the faces F_e^M and F_e^m are two-dimensional, unless χ has an eigenvalue with multiplicity 2, and in that case exactly one face is degenerate.

Proposition 5.12 *The facets F_e^M and F_e^m have a unique supporting linear function, unless they are degenerate.*

Proof It suffices to determine the hyperplanes supporting faces which contain χ. Any hyperplane supporting the vertex χ contains the tangent space $T_\chi \mathcal{O}$ at χ to the orbit $\mathcal{O}$ through χ. Choose coordinates so that $\chi = \chi(M, 0, 0, \gamma, 0)$ is diagonal $(\gamma = -\frac{M}{2} - m)$ so that $\chi \in F_{e_1}^M \cap F_{e_3}^m$. Recall that $T_\chi \mathcal{O}$ (5.4) is the affine 3-plane in W

$$\chi(M, *, *, \gamma, *),$$

Fig. 4 Image of subsets of Fig. 3

$$\det = \sqrt{\tfrac{4}{27}(\alpha - \tfrac{1}{4})(\alpha - 1)^2}$$
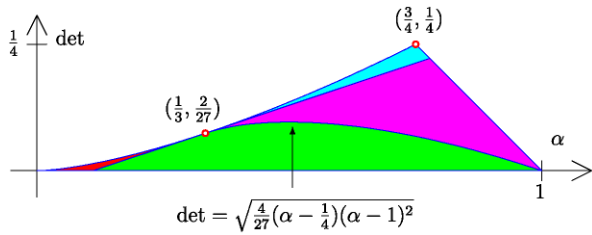
where $*$ represents an arbitrary real number. If $(x_1, \ldots, x_5)$ are the coordinates (3.3), then hyperplanes containing $T_\chi \mathcal{O}$ have equation $\sum_i c_i x_i = c$, where

$$c_2 = c_3 = c_5 = 0, \quad \text{and} \quad c_1 M + c_4 \gamma = c. \tag{5.13}$$

If $\gamma \neq 0$, then $F_{e_1}^M$ is nondegenerate and contains the additional point

$$R_{e_1, \frac{\pi}{2}} = \chi(M, 0, 0, 0, \gamma) = \begin{pmatrix} M & 0 & 0 \\ 0 & -\frac{M}{2} & \gamma \\ 0 & \gamma & -\frac{M}{2} \end{pmatrix},$$

which imposes the further condition $c_1 M = c$ on a support hyperplane to $F_{e_1}^M$. Thus $c_1 = c/M$ and $c_4 = 0$. Setting $c = M$ so that $c_1 = 1$, we see that the support hyperplane to $F_{e_1}^M$ is defined by $x_1 = M$, which is $L_{e_1}(\chi) = M$.

If $\nu = M + \frac{m}{2} \neq 0$, then $F_{e_3}^m$ is nondegenerate and it contains the point

$$\begin{pmatrix} -\frac{m}{2} & \nu & 0 \\ \nu & -\frac{m}{2} & 0 \\ 0 & 0 & m \end{pmatrix} = \chi\left(-\frac{m}{2}, \nu, 0, -\frac{3m}{4}, 0\right),$$

and so a support hyperplane to $F_{e_3}^m$ must satisfy (5.13) and also $-c_1 \frac{m}{2} - c_4 \frac{3m}{4} = d$. Subtracting these equations and dividing by ν, we see that $c_1 = 2c_4$, and so

$$x_1 + 2x_4 = -2m,$$

is the support hyperplane to $F_{e_3}^m$. Note that $x_1 + 2x_4$ is $-2L_{e_3}(\chi)$.   □

6 Two Metals

Let $\chi_1, \chi_2 \in W$ be linearly independent anisotropic tensors and set $\chi := (\chi_1, \chi_2) \in W^2$. By Lemma 4.7, the convex hull V of the orbit $\mathcal{O} := SO(3).\chi$ is a $2 \cdot 5 = 10$-dimensional convex body containing the origin. Its boundary is 9-dimensional.

We study the facial structure of $V \subset W^2$. One tool will be a family of $SO(3)$-equivariant maps $\pi_\alpha : W^2 \to W$. We first determine the dimension of the orbit, show that the maximum dimension of a facet is 6, and then define coaxial faces. Our main result is that coaxial faces are facets if χ_1 and χ_2 have distinct eigenvectors. In that case, almost all coaxial faces have dimension 6 and Carathéodory number 4. We are

unable to rule out the existence of other facets, but we conjecture that there are no other facets.

Let $\text{Span}(\chi) \subset W$ be the two-dimensional subspace of W spanned by χ_1 and χ_2. The structure of V depends only on $\text{Span}(\chi)$. Indeed, if $\chi_1', \chi_2' \in \text{Span}(\chi)$ are linearly independent, then there is a 2×2 invertible matrix $A = (a_{ij})$ such that

$$\chi_1' = a_{11}\chi_1 + a_{12}\chi_2, \qquad \chi_2' = a_{21}\chi_1 + a_{22}\chi_2.$$

This induces an $SO(3)$-isomorphism $W^2 \xrightarrow{\sim} W^2$:

$$W^2 \ni (w_1, w_2) \longmapsto (a_{11}w_1 + a_{12}w_2, a_{21}w_1 + a_{22}w_2) \in W^2$$

which sends V to the convex hull of the orbit of (χ_1', χ_2'). This is nothing more than a change of coordinates on W^2.

Any nonzero vector $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ gives an $SO(3)$-map

$$\pi_\alpha \colon W^2 \longrightarrow W \tag{6.1}$$

defined by $\pi_\alpha(w_1, w_2) := \alpha_1 w_1 + \alpha_2 w_2$. Write w_α for $\pi_\alpha(w)$. In particular, $\chi_\alpha := \pi_\alpha(\chi) \in \text{Span}(\chi)$. Set $V_\alpha \subset W$ to be the convex hull of the orbit $\mathcal{O}_\alpha := SO(3).\chi_\alpha$. Since $\pi_\alpha(SO(3).\chi) = SO(3).\chi_\alpha$, we have $\pi_\alpha(\mathcal{O}_\chi) = \mathcal{O}_\alpha$ and $V_\alpha = \pi_\alpha(V)$. We compute the dimension of the orbit $\mathcal{O}_\chi$.

Theorem 6.2 $\dim \mathcal{O}_\chi = 3$.

Proof We will show that $\dim \mathcal{O}_\alpha = 3$ for some $\alpha \in \mathbb{R}^2$. As $\pi_\alpha(\mathcal{O}_\chi) = \mathcal{O}_\alpha$, this implies $\dim \mathcal{O}_\chi \geq 3$. Since $\dim SO(3) = 3$, we have $\dim \mathcal{O}_\chi \leq 3$ and so $\dim \mathcal{O}_\chi = 3$.

By Proposition 5.2, the dimension of $\mathcal{O}_\alpha$ is 3 if and only if χ_α has distinct eigenvalues. If either χ_1 or χ_2, say χ_1, has distinct eigenvalues, then $\dim \mathcal{O}_{(1,0)} = 3$ and we are done. Suppose the contrary, that neither χ_1 nor χ_2 has distinct eigenvalues. That is, for each $i = 1, 2$, χ_i has a two-dimensional eigenspace with eigenvalue α_i. Since 0 cannot be a repeated eigenvalue, neither α_1 nor α_2 is zero. These eigenspaces must meet, so χ_1 and χ_2 share an eigenvector, which is an eigenvector for the nonzero tensor $\chi_\alpha := \alpha_2\chi_1 - \alpha_1\chi_2 \in \text{Span}(\chi)$ with eigenvalue 0. But then χ_α has distinct eigenvalues and so $\dim \mathcal{O}_\alpha = 3$. $\square$

Lemma 6.3 *The maximum dimension of a proper face of V is 6.*

Proof Let F be a proper face of V and let $S \subset SO(3)$ its stabilizer subgroup,

$$S = \{g \in SO(3) \mid g.F \subset F\}.$$

This is a closed, proper subgroup, and thus either has dimension 1 (in which case it is a coaxial subgroup Q_e or a rotation subgroup Q_e^+), or it is finite and has dimension zero.

Let F° be the relative interior of F, those points of F which do not lie in any other face of V of the same or smaller dimension. If $g.F^\circ \cap F^\circ \neq \emptyset$, then $g.F = F$, and so $g \in S$.

Let ∂V be the boundary of V and consider the map $f : SO(3) \times F^\circ \to \partial V$ defined by

$$f : (g, v) \longmapsto g.v.$$

This map is not 1-1: Suppose that $g.v = h.w$, for $g, h \in SO(3)$ and $v, w \in F^\circ$. Then $h^{-1}g.v = w$ and so $h^{-1}g.F^\circ \cap F^\circ \neq \emptyset$, which implies that $s := h^{-1}g \in S$. Then $s.v = w$.

This calculation shows that the fibers of f have the form

$$\left\{ \left(gs^{-1}, sv \right) \mid s \in S \right\}, \quad \text{for } g \in SO(3) \text{ and } v \in F^\circ.$$

Thus we have the dimension calculation

$$\dim \partial V \geq \dim SO(3) + \dim F^\circ - \dim S.$$

Since $\dim \partial V = 9$ and $\dim SO(3) = 3$, this gives

$$6 + \dim S \geq \dim F. \tag{6.4}$$

If S is finite, then $\dim F \leq 6$. If S has dimension 1 so that it is either Q_e or Q_e^+ for some e, then F could have dimension up to 7. By (3.5), $W^2 = \mathbb{R}^2 \oplus U_1^2 \oplus U_2^2$ as a representation of S. If $S = Q_e$, then Lemma 4.7 implies that F has even dimension, and if $S = Q_e^+$, then Example 4.8 implies that F has dimension 0, 2, or 4, which completes the proof. $\qquad \square$

6.1 Coaxial Faces

A *coaxial face* of V is a face that is stabilized by some coaxial subgroup, Q_e. By Lemma 4.3, such a face is supported by a Q_e-invariant linear function, which must factor through the projection to the trivial isotypical component of W^2, by Schur's Lemma. Since this component is $\mathbb{R}^2$ (3.5), L is the pullback of a linear map

$$\mathbb{R}^2 \ni (M_1, M_2) \longmapsto \alpha_1 M_1 + \alpha_2 M_2 \in \mathbb{R}.$$

Up to a scalar, this is the composition of the Q_e-invariant linear function L_e (5.1) on W with π_α, which is the map $L_{e,\alpha}$ defined by

$$L_{e,\alpha}(w) := L_e(w_\alpha) = \langle e, w_\alpha e \rangle.$$

Suppose now that e is a unit vector. For each nonzero $\alpha \in \mathbb{R}^2$, define

$$M_\alpha := \text{maximum eigenvalue of } \chi_\alpha, \quad \text{and}$$

$$m_\alpha := \text{minimum eigenvalue of } \chi_\alpha.$$

If $\overline{\chi} \in V$, then $\overline{\chi}_\alpha \in V_\alpha$, and so by Lemma 5.5 we have

$$M_\alpha \geq L_{e,\alpha}(\overline{\chi}) \geq m_\alpha, \tag{6.5}$$

with equality only when e is an eigenvector of $\overline{\chi}_\alpha$ having eigenvalue M_α or m_α. Thus coaxial faces are the faces of V defined by equality in (6.5).

For $e \in \mathbb{R}^3$ a unit vector and $0 \neq \alpha \in \mathbb{R}^2$, define the coaxial face

$$F_{e,\alpha} := \left\{ \overline{\chi} \in V \mid L_{e,\alpha}(\overline{\chi}) = M_\alpha \right\}.$$

If $\overline{\chi} \in F_{e,\alpha}$, then e is an eigenvector of $\overline{\chi}_\alpha$ with eigenvector M_α. As in Sect. 5, each coaxial face $F_{e,\alpha}$ is the convex hull of an orbit $Q_e.\chi'$, for some $\chi' \in \mathcal{O}$.

Theorem 6.6 *Faces of V have dimension at most* 6. *The coaxial faces of V form a three-dimensional family whose union is a nine-dimensional subset of the boundary of V if and only if χ_1 and χ_2 have distinct eigenvectors. When this happens, almost all coaxial faces have dimension* 6, *have Carathéodory number* 4, *and are facets of V.*

If the boundary of V is the union of the coaxial faces, then Lemma 4.2 implies that the Carathéodory number of V is at most 5, and we conjecture this is the case. If there are faces of dimension 6 that are not coaxial, then Carathéodory's Theorem implies that their Carathéodory number is at most 7. Then Lemma 4.2 implies the following corollary of Theorem 6.6.

Corollary 6.7 *The Carathéodory number of V is at most* 8.

By "almost all" in the statement of Theorem 6.6, we mean in the algebraic sense: Except for those $\alpha \in \mathbb{R}^2$ lying in finitely many half-rays in $\mathbb{R}^2$, $F_{e,\alpha}$ has dimension 6 when χ_1 and χ_2 have distinct eigenvectors. The proof of Theorem 6.6 is done in the series of lemmas below.

Remark 6.8 The condition that the magnetic susceptibility tensors χ_1 and χ_2 have distinct eigenvectors has already been considered in protein folding. It implies that RDC measurements from the two ions are sufficient to remove the symmetry property of the RDC [12].

Since $M_{-\alpha} = -m_\alpha$, there is no need for two types of coaxial faces as in Sect. 5. Since if $r > 0$, then $M_{r\alpha} = rM_\alpha$ and $L_e = L_{-e}$, we have

$$F_{e,\alpha} = F_{-e,\alpha} = F_{e,r\alpha}$$

if $r > 0$. Thus we may assume that α lies on the unit circle S^1 in $\mathbb{R}^2$. We also only need to consider the unit vector e up to multiplication by ± 1, that is, as a point in the real projective plane, $\mathbb{RP}^2$, which is a two-dimensional manifold.

Lemma 6.9 *The coaxial faces $F_{e,\alpha}$ form a three-dimensional family parameterized by $\mathbb{RP}^2 \times S^1$. For each $\alpha \in S^1$, any two coaxial faces $F_{e,\alpha}$ and $F_{e',\alpha}$ are isomorphic.*

Since the boundary of V is nine-dimensional and it has a three-dimensional family of coaxial faces, we see again that the maximum dimension of a coaxial face is 6.

Proof Suppose that $e \in \mathbb{R}^3$ is an eigenvector for χ_α with maximal eigenvalue M_α. Then $\chi \in F_{e,\alpha}$ and $F_{e,\alpha}$ is the convex hull of the orbit $Q_e.\chi$. If $R \in SO(3)$, then

$$R.Q_e.\chi = RQ_eR^T.R.\chi = Q_{Re}.(R.\chi).$$

But $(R.\chi)_\alpha$ is an anisotropic tensor having eigenvector Re with eigenvalue M_α. Therefore $F_{Re,\alpha}$ is the convex hull of $Q_{Re}.(R.\chi) = R.Q_e.\chi$, and thus equals $R.F_{e,\alpha}$. □

We now determine the dimension of the coaxial faces. By Lemma 6.9, we need only study one coaxial face $F_{e,\alpha}$ for each $\alpha \in S^1$. We compute the dimension of the affine span of an orbit $Q_e.\chi$, where e is an eigenvector of χ_α. This is the dimension of a coaxial face when the eigenvalue associated to e is a maximal eigenvalue of χ_α. Since this dimension is the rank of a matrix, those entries are algebraic functions of α. Thus, for all but finitely many α, this rank will be constant and it will be smaller for α in that finite set.

Let $\alpha \in S^1$ and suppose that $\chi_\alpha, \chi' \in \mathrm{Span}(\chi)$ are linearly independent, and let e be a unit eigenvector of χ_α. The decomposition (3.5) of W into Q_e-isotypical components induces a decomposition of the tensors χ_α and χ' into their components in $\mathbb{R} \oplus U_1 \oplus U_2$,

$$\chi_\alpha = M_\alpha \oplus 0 \oplus y_\alpha \quad \text{and} \quad \chi' = M' \oplus x' \oplus y'.$$

The U_1-component of χ_α is 0, because e is an eigenvalue of χ_α. Let $d_1 \in \{0, 1\}$ be the dimension of the linear span of x' in U_1 and $d_2 \in \{0, 1, 2\}$ be the dimension of the linear span of y_α, y' in U_2. By Lemma 4.7, the dimension of the convex hull of $Q_e.\chi$ is $d_1 \cdot \dim U_1 + d_2 \cdot \dim U_2$, which implies the following lemma.

Lemma 6.10 *The coaxial face $F_{e,\alpha}$ has dimension $2(d_1 + d_2)$.*

Thus again a coaxial face has dimension at most 6.

Lemma 6.11 *If χ_1 and χ_2 have a common eigenvector, then coaxial faces have dimension 2 or 4.*

Proof Fix $\alpha \in S^1$ and let f be a common eigenvector of χ_1 and χ_2. Then it is an eigenvector of any χ_α. Let χ' be another tensor in $\mathrm{Span}(\chi)$ which is not proportional to χ_α. Suppose that $e = e_1$ is an eigenvector of χ_α, that $f \in \{e_1, e_2, e_3\}$, and write χ_α and χ' in the coordinates (3.3),

$$\chi_\alpha = (M_\alpha, 0, 0, y, z) \quad \text{and} \quad \chi' = (M', w', x', y', z').$$

Note that $(w', x') \in U_1$ and $(y, z), (y', z') \in U_2$.

If $f = e_1$, then $(w', x') = (0, 0)$, and so $d_1 = 0$. If $f = e_2$ or e_3, then $z = z' = 0$ and so $d_2 = 1$. In either case, $d_1 + d_2 < 3$ and so the coaxial face $F_{e,\alpha}$ has dimension 2 or 4. □

Lemma 6.12 *If χ_1 and χ_2 do not have a common eigenvector, then there is a coaxial face with dimension* 6.

Proof Suppose that $e = e_1$ and e_1, e_2, e_3 is an ordered basis of eigenvectors of χ_1 with the eigenvalue of e maximal. Write χ_1 and χ_2 in the coordinates (3.3),

$$\chi_1 = \chi(M_1, 0, 0, \gamma, 0), \qquad \chi_2 = \chi(M_2, w, x, y, z).$$

The dimension of the coaxial face $F_{e_1,(1,0)}$ is 6 if and only if $(w, x) \neq (0, 0)$ and $(\gamma, 0), (y, z) \in U_2$ are linearly independent. Suppose that $\dim F < 6$. We cannot have $(w, x) = (0, 0)$ for then e_1 is a common eigenvector of χ_1 and χ_2, a contradiction. Thus the vectors $(\gamma, 0), (y, z)$ are dependent.

If $\gamma = 0$, then χ_1 has a repeated smallest eigenvalue with eigenspace spanned by e_2 and e_3. Changing the last two coordinates, we may assume that $z = 0$. We cannot also have either $w = 0$ or $x = 0$ for then χ_1 and χ_2 have either e_2 or e_3 as a common eigenvector. If $y = 0$, then $-xe_2 + we_3$ is a common eigenvector, so $y \neq 0$ and $F_{e_1,(1,0)}$ has dimension 4. In the coordinates (3.3) with respect to the ordered basis $e_3, e_2, e_1, -\chi_1$ and χ_2 are

$$-\chi_1 = \chi\left(\frac{M_1}{2}, 0, 0, 3\frac{M_1}{4}, 0\right) \quad \text{and} \quad \chi_2 = \chi\left(-\frac{M_2}{2} - y, 0, x, -3\frac{M_2}{4} + \frac{y}{2}, w\right).$$

Since $w, x \neq 0$, the affine span of $Q_{e_3}.\chi$ has dimension 6. Since $\frac{M_1}{2}$ is the maximal eigenvalue of $-\chi_1$ with eigenvector e_3, this shows that $F_{e_3,(-1,0)}$ has dimension 6.

The third possibility is that $z = 0$. But then the same arguments as in the previous paragraph show that coaxial face $F_{e_3,(-1,0)}$ has dimension 6. $\qquad \square$

6.2 Structure and Carathéodory Number of a Coaxial Facet

Suppose that F is a coaxial face of dimension 6. We may assume that F is the convex hull of the orbit $Q_e.\chi$ and that F spans the representation $U_1 \oplus U_2^2$. This Q_e-orbit is the union of two orbits of its identity component $Q_e^+ (\simeq SO(2))$. Call them $\mathcal{O}^+$ and $\mathcal{O}^-$. By Example 4.8, each orbit spans a subrepresentation of $U_1 \oplus U_2^2$ isomorphic to $U_1 \oplus U_2$. Set

$$F^{\pm} := \text{convex hull of } \mathcal{O}^{\pm} \quad \text{and} \quad W^{\pm} := \text{linear span of } \mathcal{O}^{\pm} \simeq U_1 \oplus U_2.$$

Proposition 6.13 *The faces $F^{\pm}$ each have dimension* 4 *and Carathéodory number* 3. *Points on their boundary are the convex hull of one or two vertices, while points in their relative interiors are the convex hull of three vertices.*

Proof As in Example 4.8, we identify $Q_e^+ \simeq SO(2)$ with the circle group S^1 and U_1, U_2 with $\mathbb{C}$. Then $z \in S^1$ acts on U_1 as scalar multiplication by z and on U_2 as scalar multiplication by z^2 and $F^{\pm}$ has dimension 4.

Let $(u, v)^T \in \mathbb{C}^2 \simeq U_1 \oplus U_2$ be the point corresponding to χ. Then

$$\mathcal{O}^+ = \left\{\left(e^{i\theta}u, e^{2i\theta}v\right) \mid 0 \leq \theta < 2\pi\right\},$$

and its convex hull is

$$\left\{ \left(\sum_{j=1}^{n} \lambda_j e^{i\theta_j} u, \sum_{j=1}^{n} \lambda_j e^{2i\theta_j} v \right) \, \Big| \, \sum_j \lambda_j = 1, 0 \le \theta_1, \ldots, \theta_n < 2\pi \right\}.$$

But this is $B.(u, v)^T$, where B is the set of 2×2 diagonal matrices whose entries are

$$\left(\sum_{j=1}^{n} \lambda_j e^{i\theta_j}, \sum_{j=1}^{n} \lambda_j e^{2i\theta_j} \right) \quad \text{where } \sum_j \lambda_j = 1 \text{ and } 0 \le \theta_1, \ldots, \theta_n < 2\pi. \quad (6.14)$$

Thus F^+ (and also F^-) is isomorphic to B.

Curto and Fialkow [8] characterized the points of B. Let (a, b) be a point of B (6.14) and p the corresponding measure on S^1,

$$p(e^{i\theta}) = \begin{cases} \lambda_j & \text{if } \theta = \theta_j, \\ 0 & \text{otherwise.} \end{cases}$$

Set $\gamma_{ij} := \int_{S^1} \overline{z}^i z^j \, dp(z)$ for $0 \le i, j$ with $i + j \le 2$ and form the moment matrix

$$M := \begin{pmatrix} \gamma_{00} & \gamma_{01} & \gamma_{10} \\ \gamma_{10} & \gamma_{11} & \gamma_{20} \\ \gamma_{01} & \gamma_{02} & \gamma_{11} \end{pmatrix} = \begin{pmatrix} 1 & a & \overline{a} \\ \overline{a} & 1 & \overline{b} \\ a & b & 1 \end{pmatrix}.$$

Proposition 6.15 (Curto and Fialkow [8]) *The points $(a, b) \in B$ are exactly the points $(a, b) \in \mathbb{C}^2$ such that M is positive semidefinite. The rank of M is the minimum number of summands needed to represent the point (a, b) (6.14).*

In particular, this implies that each body $F^\pm$ has Carathéodory number 3.
By Proposition 6.15,

$$B = \left\{ (a, b) \in \mathbb{C}^2 \mid 1 - |a|^2 \ge 0, 1 - |b|^2 \ge 0, 1 + \overline{a}^2 b + a^2 \overline{b} - 2|a|^2 - |b|^2 \ge 0 \right\}.$$

If $1 - |a|^2 = 0$, then $a \in S^1$ and so $n = 1$ in (6.14). If $1 - |b|^2 = 0$, then $b \in S^1$ and either $n = 1$ in (6.14) or $n = 2$ with $|\theta_1 - \theta_2| = \pi$. Thus if M has rank 3, then $|a| < 1$ and $|b| < 1$, and so (a, b) lies in the interior of B as the inequalities are strict. This implies that points on the boundary of B are the convex hull of one or two vertices and this completes the proof of Proposition 6.13. □

Now we complete the proof of Theorem 6.6, showing that the coaxial facet F has Carathéodory number 4. The coaxial facet F is the convex hull of F^+ and F^-. Let $v \in F$. We suppose that $v \notin F^+ \cup F^-$, for otherwise v is the convex combination of at most three vertices. Then there exist $v^\pm \in F^\pm$ and $\lambda \in (0, 1)$ such that

$$v = \lambda v^+ + (1 - \lambda)v^-.$$

If both v^+ and v^- lie on the boundary of their respective subfaces, then each is a convex combination of at most 2 vertices, and v is a convex combination of at most 4 vertices.

Suppose instead that v^+ lies in the relative interior of F^+. The linear span of v and W^- has dimension 5 in the six-dimensional space $U_1 \oplus U_2^2$ and therefore it meets W^+ in a three-dimensional affine subspace U^+. Similarly the span of v and W^+ meets W^- in a three-dimensional subspace U^-. Observe that both U^+ and U^- contain the two-dimensional linear subspace $W^+ \cap W^- = U_1$ so that their span has dimension 4.

Consider the cone over $U^- \cap F^-$ with vertex v. Removing v, this has two components. One meets F^-. Let C be the component that does not meet F^-, and let $C^+ := C \cap U^+$. This is a convex set which contains v^+ and thus meets the relative interior of $B^+ := U^+ \cap F^+$. Points $v' \in C^+ \cap B^+$ are exactly those points of F^+ for which there exists a point $v'' \in F^-$ such that v is a convex combination of v' and v''. There are two possibilities.

(1) The boundary of C^+ meets the boundary of B^+.
(2) Either the boundary of C^+ is a subset of B^+ or vice-versa.

In the first case, let v' be a point common to the two boundaries. Then v' lies on the boundary of B^+ and v'' lies on the boundary of B^-. But these are subsets of the boundaries of $F^\pm$, and so v is the convex combination of at most four vertices.

In the second case, suppose that the boundary of B^+ is a subset of C^+. Since B^+ is the intersection F^+ with a hyperplane, its boundary must contain a vertex of F^+, as the set of vertices of F^+ is a connected one-dimensional set whose convex hull is F^+. Suppose that $v' \in B^+$ is a vertex of F^+. Since $v'' \in F^-$ is a convex combination of three vertices of F^-, we see that v is a convex combination of $1 + 3 = 4$ vertices. If the boundary of C^+ is a subset of B^+, then we may choose the point v' in the boundary of C^+ so that the corresponding point v'' is a vertex of F^-. Again, v is the convex combination of $3 + 1 = 4$ vertices.

References

1. Banci, L., Bertini, I., Bren, K.L., Cremonini, M.A., Gray, H.B., Luchinat, C., Turano, P.: The use of pseudocontact shifts to refine solution structures of paramagnetic metalloproteins: Met80ala cyano-cytochrome c as an example. J. Biol. Inorg. Chem. **1**, 117–126 (1996)
2. Bertini, I., Gelis, I., Katsaros, N., Luchinat, C., Provenzani, A.: Tuning the affinity for lanthanides of calcium binding proteins. Biochemistry **42**, 8011–8021 (2003)
3. Bertini, I., Del Bianco, C., Gelis, I., Katsaros, N., Luchinat, C., Parigi, G., Peana, M., Provenzani, A., Zoroddu, M.A.: Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. Proc. Natl. Acad. Sci. USA **101**, 6841–6846 (2004)
4. Bertini, I., Gupta, K.J., Luchinat, C., Parigi, G., Peana, M., Sgheri, L., Yuan, J.: Paramagnetism-based nmr restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. J. Am. Chem. Soc. **129**, 12786–12794 (2007)
5. Bröcker, T., tom Dieck, T.: Representations of Compact Lie Groups. Graduate Texts in Mathematics, vol. 98. Springer, New York (1985)

6. Carathéodory, C.: Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen. Palermo Rend. **32**, 193–217 (1911). (German)
7. Coxeter, H.S.M.: Regular Polytopes, 3rd edn. Dover, New York (1973)
8. Curto, R., Fialkow, L.: The quadratic moment problem for the unit circle and unit disk. Integral Equ. Oper. Theory **38**(4), 377–409 (2000)
9. Fenchel, W.: Über Krümmung und Windung geschlossener Raumkurven. Math. Ann. **101**, 238–252 (1929). (German)
10. Gardner, R., Longinetti, M., Sgheri, L.: Reconstruction of orientations of a moving protein domain from paramagnetic data. Inverse Probl. **21**(3), 879–898 (2005)
11. Koradi, R., Billeter, M., Wüthrich, K.: Molmol: a program for display and analysis of macromolecular structures. J. Mol. Graph. **14**, 51–55 (1996)
12. Longinetti, M., Parigi, G., Sgheri, L.: Uniqueness and degeneracy in the localization of rigid structural elements in paramagnetic proteins. J. Phys. A **35**(39), 8153–8169 (2002)
13. Longinetti, M., Luchinat, C., Parigi, G., Sgheri, L.: Efficient determination of the most favoured orientations of protein domains from paramagnetic NMR data. Inverse Probl. **22**(4), 1485–1502 (2006)
14. Meiler, J., Prompers, J.J., Peti, W., Griesenger, C., Brüshweiler, R.: Model-free approach to the dynamic interpolation of residual dipolar coupling in globular proteins. J. Am. Chem. Soc. **123**, 6098–6107 (2001)
15. Reay, J.R.: Generalizations of a Theorem of Carathéodory. Memoirs of the Am. Math. Soc., vol. 54. AMS, Providence (1965)
16. Schneider, R.: The Brunn-Minkowski Theory. Cambridge University Press, Cambridge (1993)
17. Tolman, J.R., Flanagan, J.M., Kennedy, M.A., Prestegard, J.H.: Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. Proc. Natl. Acad. Sci. USA **92**, 9279–9283 (1995)