Internal or shape coordinates in the n-body problem

Robert G. Littlejohn and Matthias Reinsch
Department of Physics, University of California, Berkeley, California 94720
(Received 17 January 1995)

The construction of global shape coordinates for the n-body problem is considered. Special attention is given to the three- and four-body problems. Quantities, including candidates for coordinates, are organized according to their transformation properties under so-called democracy transformations (orthogonal transformations of Jacobi vectors). Important submanifolds of shape space are identified and their topology studied, including the manifolds upon which shapes are coplanar or collinear, and the manifolds upon which the moment of inertia tensor is degenerate.

PACS number(s): 31.15.−p, 03.20.+i, 02.40.−k

I. INTRODUCTION

In recent years the n-body problem has been studied from a new perspective, one that involves certain gauge fields which arise in the separation of rotations and internal motions. By the n-body problem we mean a multiparticle system in which the interactions are invariant under overall rotations of the system. The role of gauge fields in such systems has been realized and developed in a number of papers, including those by Guichardet [1], Tachibana and Iwai [2], Iwai [3], and Shapere and Wilczek [4,5]. We envision the primary applications of this new approach to occur in atomic, nuclear, and molecular physics, as well as in classical problems such as celestial mechanics.

It has long been recognized that coordinates for an n-body system naturally break up into internal or shape coordinates and orientational coordinates. This decomposition has a natural description in the language of fiber bundles, which is the proper mathematical framework for gauge theories. In this language, the internal or shape coordinates are coordinates on the base space (essentially the space we call "shape space") and the orientational coordinates (usually chosen to be Euler angles) are coordinates on the rotation fibers.

In applications it has not always been clear how to construct internal or shape coordinates explicitly. This difficulty, and the related problem of understanding the topology of shape space, has been an obstacle in certain problems in molecular dynamics. One possible system of shape coordinates was given by Eckart [6], but his coordinates are primarily useful for small amplitude vibrations about the equilibrium positions.

In the case of the four-body problem, the construction of explicit, global shape coordinates has been dealt with by Keating and Mead [7], who analyzed this problem in the context of a study of conical intersections between Born-Oppenheimer potential energy surfaces. These authors defined a coordinate system based on the interparticle distances, and used the theory of the permutation group to find linear combinations of the squares of the interparticle distances that resulted in useful coordinates for their study. They worked out the ranges of these coordinates and also described a way of modifying this coordinate system in order to distinguish shapes which differ only by parity. In our work we make use of techniques very similar to those developed by Keating and Mead; one difference is that we employ a continuous group (the "democracy group") rather than the permutation group.

A technical treatment of the four-body problem has been given by Narasimhan and Ramadas [8]. These authors proved that the space of all shapes in the four-body problem is homeomorphic (topologically equivalent) to $\mathbb{R}^6$, and that the space of all shapes with dimensionality greater than or equal to 2 is homeomorphic to the Cartesian product of the real line $\mathbb{R}$ and a certain five-dimensional manifold. We will discuss these results from a less abstract perspective in Sec. IV.

In this paper we consider the general n-body problem (for arbitrary n), devoting special attention to the three- and four-body problems. There is a vast amount of literature on this subject, much of which focuses on the three-body problem. Works in this area include papers by Breit [9], Smith [10], Bhatia and Temkin [11], Zickendraht [12], De Celles and Darling [13], Mead and Truhlar [14], Johnson [15], Sutcliffe [16], Pack and Parker [17], and Mezey [18]. The subject of hyperspherical coordinates has been considered by Smith [10,19], Smirnov and Shitikova [20], Johnson [15], and Pack and Parker [17]. The construction of quantities which are invariant under permutations of particle labels or under the more general democracy transformations has been considered by authors such as Keating and Mead [7], Smith [19], Dragt [21], Levý-Leblond and Lévy-Nahas [22], and Louck and Galbraith [23]. We make use of similar techniques below.

In Sec. II of this paper, we describe the general features of the n-body problem and introduce quantities that facilitate the study of shape space. This section also includes a discussion of the action of various groups on shape space. In Sec. III, the methods developed in the previous section are applied to the three-body problem. This is done to illustrate the theory and to explain the significance of some of the standard coordinate systems that are used in the literature. In Sec. IV, we study the shape space of the four-body problem and some of its submanifolds. We also discuss the use of democracy transformations for visualizing four-body shape space.

1050-2947/95/52(3)/2035(17)/$06.00

II. GENERAL CONSIDERATIONS

In this section we develop several results which are of use in the n-body problem for any value of $n \geq 3$. One of our main accomplishments is to define a mapping from shape space onto the space of real, symmetric, $(n-1) \times (n-1)$ non-negative definite matrices with rank $r \leq 3$. This mapping plays a central role in the construction of coordinate systems on shape space.

A. Translational degrees of freedom and Jacobi coordinates

Consider a system of n distinguishable particles with lab (inertial frame) positions $(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ and masses $(m_1, \ldots, m_n)$. The total kinetic energy is

$$T_{\rm tot} = \frac{1}{2}\sum_{\alpha=1}^{n} m_\alpha |\dot{\mathbf{r}}_\alpha|^2. \tag{2.1}$$

To separate out the translational degrees of freedom, we perform a linear coordinate transformation $(\mathbf{r}_1, \ldots, \mathbf{r}_n) \to (\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_{n-1}, \mathbf{R})$, where $\mathbf{R}$ is the center of mass,

$$\mathbf{R} = \frac{1}{M}\sum_{\alpha=1}^{n} m_\alpha \mathbf{r}_\alpha, \tag{2.2}$$

and where the $n-1$ vectors $(\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_{n-1})$ are mass-weighted Jacobi vectors. In Eq. (2.2), M is the total mass,

$$M = \sum_{\alpha=1}^{n} m_\alpha. \tag{2.3}$$

The Jacobi vectors $\boldsymbol{\rho}_\alpha$ are chosen so that the kinetic energy about the center of mass has the form

$$T = \frac{1}{2}\sum_{\alpha=1}^{n-1} |\dot{\boldsymbol{\rho}}_\alpha|^2, \tag{2.4}$$

to which the kinetic energy of the center of mass, $(M/2)|\dot{\mathbf{R}}|^2$, must be added to obtain the total kinetic energy $T_{\rm tot}$ as in Eq. (2.1). We will henceforth ignore the kinetic energy of the center of mass, and simply refer to T as the "kinetic energy."

The $n-1$ Jacobi vectors $\boldsymbol{\rho}_\alpha$ are coordinates on the "translation-reduced configuration space," which is $\mathbb{R}^{3n-3}$. We will henceforth refer to this space simply as "configuration space," and to the translation-reduced configurations in it simply as "configurations." The kinetic energy (2.4) defines a Euclidean metric on this configuration space. Linear transformations among the $3n-3$ coordinates on this space which preserve the Euclidean form of the metric belong to the group $O(3n-3)$, which is the overall symmetry group of the kinetic energy.

The usual procedure for defining Jacobi vectors involves organizing the particles into a hierarchy of clusters, in which each cluster consists of one or more particles, and where each Jacobi vector joins the centers of mass of two clusters, thereby creating a larger cluster. Each Jacobi vector must be weighted by the square root of the reduced mass of the two clusters it joins in order to achieve the Euclidean form of the kinetic energy shown in Eq. (2.4). For example, in the four-body problem, we can use the clustering of particles illustrated in Fig. 1 to define

$$\begin{aligned}\boldsymbol{\rho}_1 &= \sqrt{\mu_1}(\mathbf{r}_2 - \mathbf{r}_1),\\ \boldsymbol{\rho}_2 &= \sqrt{\mu_2}(\mathbf{r}_4 - \mathbf{r}_3),\\ \boldsymbol{\rho}_3 &= \sqrt{\mu_3}\left(\frac{m_3\mathbf{r}_3 + m_4\mathbf{r}_4}{m_3+m_4} - \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1+m_2}\right),\end{aligned} \tag{2.5}$$

where the reduced masses μ_α are defined by

$$\begin{aligned}&\frac{1}{\mu_1} = \frac{1}{m_1} + \frac{1}{m_2}, \quad \frac{1}{\mu_2} = \frac{1}{m_3} + \frac{1}{m_4},\\ &\frac{1}{\mu_3} = \frac{1}{m_1+m_2} + \frac{1}{m_3+m_4}.\end{aligned} \tag{2.6}$$

B. Three groups acting on configuration space

Different clusterings lead to different definitions of Jacobi vectors, but all choices are connected by linear transformations of the form,

$$\boldsymbol{\rho}'_\alpha = \sum_{\beta=1}^{n-1} D_{\alpha\beta}\,\boldsymbol{\rho}_\beta, \tag{2.7}$$

where D (with components $D_{\alpha\beta}$) is an $(n-1) \times (n-1)$ matrix depending on the masses. Since all choices of mass-weighted Jacobi vectors lead to the same Euclidean form (2.4) of the kinetic energy, the matrix D is orthogonal, i.e., an element of the group $O(n-1)$. We introduce this group as a continuous group which interpolates between all the discrete choices of Jacobi coordinates. We call this group the "democracy group," because the transformations (2.7) arise in questions of particle democracy, such as how to construct coordinates or functions which are invariant under permutations of particle labels. Often it is convenient to restrict the democracy group to the connected group $SO(n-1)$. This entails little loss

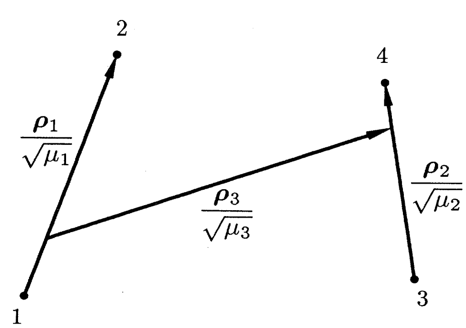

FIG. 1. An example of Jacobi vectors for the four-body problem. The figure is drawn for $m_1 \approx 3m_2$, $m_4 \approx 3m_3$.

of generality, since by changing the sign of one Jacobi vector, if necessary, it is always possible to force the matrix D to have determinant +1. The democracy group $O(n-1)$ or $SO(n-1)$ is only a subgroup of the overall symmetry group $O(3n-3)$ of the kinetic energy, because the democracy transformations (2.7) do not mix the individual components of the Jacobi vectors $\boldsymbol{\rho}_\alpha$.

Another group of importance acting on configuration space consists of ordinary rotations, i.e., the transformations

$$\boldsymbol{\rho}'_\alpha = \mathsf{R}\boldsymbol{\rho}_\alpha, \quad \alpha = 1, \ldots, n-1, \tag{2.8}$$

where $\mathsf{R} \in SO(3)$. The rotation group is a symmetry group, not only of the kinetic energy, but also (as we shall assume) of the potential energy, and so has a deeper significance than the larger group $O(3n-3)$ which is the symmetry group of the kinetic energy alone.

We will also be interested in the parity operator P, whose action on configuration space is given by

$$\boldsymbol{\rho}'_\alpha = P\boldsymbol{\rho}_\alpha = -\boldsymbol{\rho}_\alpha. \tag{2.9}$$

The rotation group combined with parity gives the group $O(3)$ of proper and improper rotations, also a subgroup of $O(3n-3)$.

The democracy group has a significance which goes beyond the motivations of its definition, and it arises in many places in the analysis presented below. One reason for the importance of the democracy group is that it is the largest subgroup of $O(3n-3)$ which commutes with all rotations $\mathsf{R} \in SO(3)$. (The proof involves a simple application of Schur's lemma.) Only those elements of $O(3n-3)$ which commute with rotations have an action on shape space (defined momentarily); thus, the democracy group acts on shape space and is the symmetry group of the metric tensor defined on that space. It is also the symmetry group of the kinetic energy in the reduced (shape space) description of the dynamics.

C. Vectors for visualization

Sometimes for purposes of visualization it is useful to have some set of vectors which describe the geometry of an n-particle configuration in a more immediate manner than the Jacobi vectors. For example, we may wish to use the displacement vectors $\mathbf{d}_\alpha$ taking us from one of the particles, say, the nth, to the other $n-1$,

$$\mathbf{d}_\alpha = \mathbf{r}_\alpha - \mathbf{r}_n, \qquad \alpha = 1, \ldots, n-1, \tag{2.10}$$

as illustrated in Fig. 2. Such vectors are related to the Jacobi vectors by a linear transformation,

$$\mathbf{d}_\alpha = \sum_{\beta=1}^{n-1} U_{\alpha\beta}\, \boldsymbol{\rho}_\beta. \tag{2.11}$$

For our purposes, all we ever need to know about this transformation is that it is invertible, i.e., $\det \mathsf{U} \neq 0$.

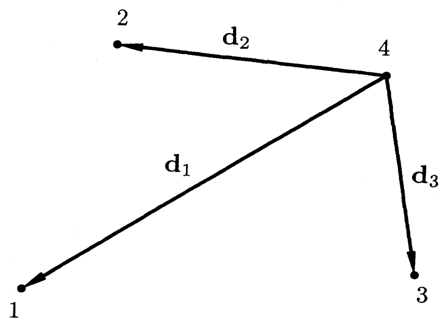

FIG. 2. It is easier to visualize the geometry of the n-particle system in terms of the $n-1$ vectors $\mathbf{d}_\alpha$, rather than the $n-1$ Jacobi vectors $\boldsymbol{\rho}_\alpha$. But the two sets of vectors are related by an invertible linear transformation.

D. Shape space and the action of parity on it

Now we turn to shape space, which is the space in which a single point represents a family of configurations of a given shape. Two configurations $\{\boldsymbol{\rho}_\alpha\}$ and $\{\boldsymbol{\rho}'_\alpha\}$ are defined to have the same shape if $\boldsymbol{\rho}'_\alpha = \mathsf{R}\boldsymbol{\rho}_\alpha$ for some proper rotation $\mathsf{R} \in SO(3)$; this defines an equivalence relation between configurations, and shape space $\mathcal{S}$ is the quotient space

$$\mathcal{S} = \frac{\mathbb{R}^{3n-3}}{SO(3)}. \tag{2.12}$$

For example, in the three-body problem, shape space is the space of triangles formed by three distinguishable bodies. For $n \geq 3$, shape space is a manifold of dimensionality $3n-6$, and internal or shape coordinates are coordinates on this manifold. By a coordinate system on shape space we mean a one-to-one mapping from some region of shape space onto some region of $\mathbb{R}^{3n-6}$; as is the general case with coordinates on manifolds, we expect to find that the whole manifold cannot be covered with a single coordinate patch, and multiple, overlapping patches must be used. On the other hand, it turns out that in the cases $n=3$ and $n=4$, shape space can be covered with a single coordinate patch. Finding a coordinate system on shape space is closely related to finding the topology of shape space; for example, we shall prove below that shape space for the four-body problem is homeomorphic (topologically equivalent) to $\mathbb{R}^6$ by finding a single coordinate patch which covers all of shape space, in which each of the six coordinates ranges from $-\infty$ to $+\infty$.

We now consider the action of parity on shape space, in preparation for later analysis. The action of parity on configuration space, given in Eq. (2.9), commutes with the action of every element of the rotation group, so parity has a well-defined action on shape space. We can just as well think in terms of reflections as parity, since an arbitrary reflection acting on configuration space induces the same mapping on shape space as parity (a reflection is the same as parity composed with a 180° proper rotation about the normal to the plane of the reflection).

Thus, if an n-particle configuration is planar, i.e., if the Jacobi vectors $\boldsymbol{\rho}_\alpha$ or, equivalently, the displacement vectors $\mathbf{d}_\alpha$ lie in a plane, then the shape is invariant under reflections in that plane, and therefore also under parity. Conversely, if a configuration is three dimensional, i.e., if the span of the $\boldsymbol{\rho}_\alpha$ or $\mathbf{d}_\alpha$ is all of three-dimensional space, then parity must map the given shape into a distinct shape, i.e., one which cannot be reached by a proper rotation. This is clear geometrically, but to present a proof we select three linearly independent Jacobi vectors, say, $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \boldsymbol{\rho}_3$, and form the nonzero triple product $\boldsymbol{\rho}_1 \cdot (\boldsymbol{\rho}_2 \times \boldsymbol{\rho}_3)$; since this changes sign under parity but not under proper rotations, parity maps the old shape into a distinct new shape. As simple corollaries of these facts, we note that all shapes in the three-body problem (triangles with distinguishable vertices) are planar, and are therefore invariant under parity [18]; but for $n = 4$, only the planar configurations, a subset of measure zero, are invariant under parity, while all other configurations (tetrahedra of nonzero volume and with distinguishable vertices) are mapped into distinct shapes by parity.

E. Motivation for study of matrices F, K, and Q

Below we present an analysis of three matrices, F, K, and Q, which prepare us for the establishment of coordinates on shape space. The general motivation for this analysis is as follows. First, for the case $n = 3$, it is obvious that the lengths of the three sides of the triangle can be used as a coordinate system on the three-dimensional shape space. Similarly, for the four-body problem, there are six interparticle distances which almost form a coordinate system on the six-dimensional shape space, failing only to distinguish two shapes (tetrahedra) related by parity. Of course, any invertible functions of the interparticle distances will serve as well for coordinates; interesting choices are the squares of the interparticle distances (which are somewhat better behaved than the interparticle distances themselves), or the dot products of the Jacobi vectors, $\boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$. [In the n-body problem, the number of interparticle distances and Jacobi dot products is the same, namely, $n(n-1)/2$, and they are simple functions of one another; thus, if the interparticle distances can be used as coordinates, then so can the Jacobi dot products.] In many respects, the Jacobi dot products are nicer candidates as coordinates than the interparticle distances or their squares, because with the Jacobi dot products, all the mass dependencies in the problem become localized in the definition of the Jacobi coordinates. Unfortunately, for $n \geq 5$, the number of Jacobi dot products, $n(n-1)/2$, exceeds the dimensionality of shape space, $3n-6$, so the Jacobi dot products can no longer be used as coordinates, unless we arbitrarily throw some of them out, an unsymmetrical and unattractive solution.

So the two defects of the Jacobi dot products, $\boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$, as coordinates on shape space are that for $n \geq 4$ they do not distinguish shapes related by parity, and that for $n \geq 5$ there are too many of them. As far as the parity ambiguity is concerned, we may wish to consider the triple products $\boldsymbol{\rho}_\alpha \cdot (\boldsymbol{\rho}_\beta \times \boldsymbol{\rho}_\gamma)$, of which there are $(n-1)(n-2)(n-3)/6$ in the n-body problem, because these do distinguish shapes related by parity. But the square of any triple product can always be expressed as a cubic polynomial in the dot products $\boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$, as shown by the vector identity

$$[\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})]^2 = A^2B^2C^2 + 2(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{C})(\mathbf{C}\cdot\mathbf{A}) - A^2(\mathbf{B}\cdot\mathbf{C})^2 - B^2(\mathbf{A}\cdot\mathbf{C})^2 - C^2(\mathbf{A}\cdot\mathbf{B})^2. \tag{2.13}$$

Thus we see that, apart from a sign ambiguity arising on taking the square root (which is really the same as the parity ambiguity), the triple products are functions of the dot products. Therefore it seems best first to study the dot products and worry about their surplus when $n \geq 5$, and then to deal with the parity ambiguity. To this end we introduce the matrices F, K, and Q and develop their properties. The following discussion is a variation of the standard theory of the singular value decomposition.

F. Definitions of matrices F, K, and Q

The matrix F is an $(n-1) \times 3$ matrix, defined by

$$\mathsf{F} = \begin{pmatrix} \rho_{1x} & \rho_{1y} & \rho_{1z} \\ \rho_{2x} & \rho_{2y} & \rho_{2z} \\ \vdots & \vdots & \vdots \\ \rho_{n-1,x} & \rho_{n-1,y} & \rho_{n-1,z} \end{pmatrix}, \tag{2.14}$$

i.e., $F_{\alpha i} = \rho_{\alpha i}$, $\alpha = 1, \ldots, n-1$, $i = 1, 2, 3$. Since this matrix has three columns, its rank r satisfies

$$r = \operatorname{rank} \mathsf{F} \leq 3. \tag{2.15}$$

The rank r is the dimensionality of the span of the vectors $\boldsymbol{\rho}_\alpha$ or $\mathbf{d}_\alpha$, so it has a direct geometrical meaning, as follows:

$r = 0$, n-particle collision;

$r \leq 1$, collinear configuration;

$r \leq 2$, planar configuration;

$r = 3$, three-dimensional configuration. (2.16)

As indicated by the inequalities, we consider the collinear configurations to include the n-body collision, and the planar configurations to include the collinear; if we wish to indicate configurations for which $r = 1$ or $r = 2$ exclusively, we will speak of one-dimensional or two-dimensional configurations, respectively.

From F we construct two more matrices,

$$\mathsf{K} = \mathsf{F}^t\mathsf{F}, \tag{2.17}$$

$$\mathsf{Q} = \mathsf{F}\mathsf{F}^t, \tag{2.18}$$

where the t indicates the transpose, which are obviously both symmetric and non-negative definite. We will call the 3×3 matrix K the "moment tensor;" its definition can also be written

$$K_{ij} = \sum_{\alpha=1}^{n-1} \rho_{\alpha i}\, \rho_{\alpha j}. \tag{2.19}$$

The moment tensor K is closely related to the moment of inertia tensor M, defined by

$$M_{ij} = \sum_{\alpha=1}^{n-1} \left[|\boldsymbol{\rho}_\alpha|^2 \delta_{ij} - \rho_{\alpha i}\, \rho_{\alpha j} \right], \tag{2.20}$$

or

$$\mathsf{M} = (\mathrm{tr}\, \mathsf{K})\mathsf{I} - \mathsf{K}. \tag{2.21}$$

As for Q, it is an $(n-1) \times (n-1)$ matrix containing the $n(n-1)/2$ Jacobi dot products,

$$Q_{\alpha\beta} = \boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta. \tag{2.22}$$

G. Interrelations among matrices F, K, and Q

It turns out that matrices K and Q have the same rank r as F. We prove this first for K. Let κ_i and $\mathbf{e}_i$, $i = 1, 2, 3$, be the eigenvalues and normalized eigenvectors of K, so that

$$\mathsf{K}\mathbf{e}_i = \kappa_i \mathbf{e}_i. \tag{2.23}$$

Since K is symmetric, the eigenvectors can be chosen to be orthonormal and are complete on $\mathbb{R}^3$; and since K is non-negative definite, we have $\kappa_i \geq 0$. For definiteness we order the eigenvalues and eigenvectors according to

$$\kappa_1 \geq \kappa_2 \geq \kappa_3 \geq 0. \tag{2.24}$$

Then by Eq. (2.17) we have

$$|\mathsf{F}\mathbf{e}_i|^2 = \mathbf{e}_i^t \cdot \mathsf{F}^t \mathsf{F} \cdot \mathbf{e}_i = \kappa_i, \tag{2.25}$$

so that $\mathsf{F}\mathbf{e}_i = \mathbf{0}$ if and only if $\kappa_i = 0$. Therefore the null spaces or kernels of K and F (as subspaces of $\mathbb{R}^3$) coincide, and K and F have the same rank. A similar argument applies to Q, which has the same null space as F^t (as subspaces of $\mathbb{R}^{n-1}$), and therefore the same rank as F.

There is a close relationship between the non-null eigenvectors of K and those of Q. (We call an eigenvector null if the corresponding eigenvalue is zero.) As for Q, let the eigenvectors and eigenvalues be ξ_α^k and λ_k, respectively, where $k = 1, \ldots, n-1$ labels the eigenvectors and eigenvalues, so that

$$\sum_{\beta=1}^{n-1} Q_{\alpha\beta}\, \xi_\beta^k = \lambda_k\, \xi_\alpha^k. \tag{2.26}$$

These eigenvectors are complete on $\mathbb{R}^{n-1}$ and we choose them to be orthonormal, so that

$$\sum_{k=1}^{n-1} \xi_\alpha^k\, \xi_\beta^k = \delta_{\alpha\beta}, \tag{2.27}$$

and

$$\sum_{\alpha=1}^{n-1} \xi_\alpha^k\, \xi_\alpha^\ell = \delta_{k\ell}. \tag{2.28}$$

For definiteness, we order the eigenvalues of Q according to

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1} \geq 0. \tag{2.29}$$

Since the rank of both K and Q is $r \leq 3$, only the first r eigenvalues of each matrix are nonzero.

Now let us consider the vectors (in $\mathbb{R}^3$),

$$\mathbf{X}_k = \sum_{\alpha=1}^{n-1} \xi_\alpha^k\, \boldsymbol{\rho}_\alpha, \tag{2.30}$$

for $k = 1, \ldots, n-1$. We claim that $\mathbf{X}_k$ vanishes if $\lambda_k = 0$, and otherwise it is an unnormalized eigenvector of K with eigenvalue $\kappa_k = \lambda_k$. To prove this, we first note that

$$\mathbf{X}_k \cdot \mathbf{X}_\ell = \sum_{\alpha,\beta} \xi_\alpha^k\, Q_{\alpha\beta}\, \xi_\beta^\ell = \lambda_k\, \delta_{k\ell}. \tag{2.31}$$

Thus $\mathbf{X}_k = \mathbf{0}$ if $\lambda_k = 0$, and the vectors

$$\hat{\mathbf{e}}_k = \frac{1}{\sqrt{\lambda_k}} \sum_{\alpha=1}^{n-1} \xi_\alpha^k\, \boldsymbol{\rho}_\alpha \tag{2.32}$$

for $k = 1, \ldots, r$ are orthonormal. They are also eigenvectors of K with eigenvalues $\kappa_k = \lambda_k$, according to the direct calculation

$$\mathsf{K}\hat{\mathbf{e}}_k = \frac{1}{\sqrt{\lambda_k}} \sum_{\alpha,\beta} \boldsymbol{\rho}_\beta\, Q_{\alpha\beta}\, \xi_\alpha^k = \lambda_k \hat{\mathbf{e}}_k. \tag{2.33}$$

We see that the nonvanishing eigenvalues of K and Q are the same, $\kappa_k = \lambda_k$, $k = 1, \ldots, r$.

In a similar manner, we can show that the vectors (in $\mathbb{R}^{n-1}$)

$$\eta_\alpha^i = \mathbf{e}_i \cdot \boldsymbol{\rho}_\alpha, \tag{2.34}$$

for $i = 1, 2, 3$, vanish if $\kappa_i = 0$, and otherwise are unnormalized eigenvectors of Q with eigenvalues $\lambda_i = \kappa_i$. If $\lambda_i > 0$, we can normalize these vectors and write

$$\xi_\alpha^i = \frac{1}{\sqrt{\lambda_i}} \hat{\mathbf{e}}_i \cdot \boldsymbol{\rho}_\alpha, \tag{2.35}$$

where $\hat{\xi}_\alpha^i$ is a normalized eigenvector of Q as in Eq. (2.26).

As a simple corollary of these facts, we note that the non-null eigenvectors of K, i.e., $\hat{\mathbf{e}}_i$, $i = 1, \ldots, r$, span the same subspace of $\mathbb{R}^3$ as the Jacobi vectors $\boldsymbol{\rho}_\alpha$, $\alpha = 1, \ldots, n-1$, and form an orthonormal basis in this subspace. This follows because the two subspaces have the same dimensionality (namely, r) and because $\mathbf{e}_i \cdot \boldsymbol{\rho}_\alpha = 0$ for $i > r$, so the orthogonal subspaces are the same. Therefore the $\boldsymbol{\rho}_\alpha$ can be written as linear combinations of the $\hat{\mathbf{e}}_i$, $i = 1, \ldots, r$. To obtain this relation explicitly, we transform Eq. (2.32), which is valid for $k = 1, \ldots, r$, into

$$\sum_{k=1}^{r} \sqrt{\lambda_k}\, \xi_\beta^k \, \hat{\mathbf{e}}_k = \sum_{\alpha=1}^{n-1} \sum_{k=1}^{r} \xi_\beta^k \, \xi_\alpha^k \, \boldsymbol{\rho}_\alpha. \tag{2.36}$$

But since $\mathbf{X}_k = \mathbf{0}$ for $k > r$, we can extend the k limit on the right hand side of Eq. (2.36) from $k = r$ to $k = n-1$, and then use the completeness relation (2.27) to obtain

$$\boldsymbol{\rho}_\alpha = \sum_{k=1}^{r} \sqrt{\lambda_k}\, \xi_\alpha^k \, \hat{\mathbf{e}}_k. \tag{2.37}$$

Finally, we note the simple relation between the eigenvalues κ_i of K (all of them, the null and non-null) and the eigenvalues μ_i of the moment of inertia tensor M; it is

$$\mu_1 = \kappa_2 + \kappa_3, \quad \mu_2 = \kappa_1 + \kappa_3, \quad \mu_3 = \kappa_1 + \kappa_2. \tag{2.38}$$

The moment of inertia tensor M does not generally have the same rank as K and Q; instead, if $r = 0$, then $\operatorname{rank}\mathsf{M} = 0$; if $r = 1$, then $\operatorname{rank}\mathsf{M} = 2$, and the two nonzero eigenvalues of M are equal; if $r = 2$, then $\operatorname{rank}\mathsf{M} = 3$ and one eigenvalue of M equals the sum of the other two; and if $r = 3$, then $\operatorname{rank}\mathsf{M} = 3$ with no special condition on the eigenvalues of M. These are elementary results with elementary interpretations. As for the eigenvectors of K and M, they are the same,

$$\mathsf{M}\hat{\mathbf{e}}_i = \mu_i \hat{\mathbf{e}}_i. \tag{2.39}$$

H. Group actions on quantities of interest

In later sections we will be interested in the action of rotations and democracy transformations on various quantities of interest. The general idea is that a given quantity will belong to some irreducible representation of the rotation group and some other irreducible representation of the democracy group, i.e., it will be a member of two sets of quantities which span carrier spaces of irreducible representations of the two groups. The quantities of interest are usually polynomials in the components of the Jacobi vectors, $\rho_{\alpha i}$. Quantities which are invariant under one or the other group (belonging to the trivial representation) are especially interesting; for example, coordinates on shape space must be invariant under rotations (they must belong to the representation $j = 0$), and democratic invariants (such as the hyperradius) are interesting because they treat all particles in a democratic manner, being invariant under permutations of particle labels. Democratic invariants are also important because the democracy group is a symmetry group of the kinetic energy (both its "horizontal" and "vertical" components), and therefore the construction of democratic invariants is involved in finding operators which commute with the kinetic energy. In the following discussion we will speak in terms of arbitrary n, but in fact we will mainly be interested in the cases $n = 3$ and $n = 4$, for which the democracy group is SO(2) and SO(3), respectively.

Under rotations $\mathsf{R} \in \mathrm{SO}(3)$, the Jacobi vectors transform according to Eq. (2.8), and the moment tensor according to

$$\mathsf{K}' = \mathsf{R}\mathsf{K}\mathsf{R}^t. \tag{2.40}$$

But the matrix Q is invariant under rotations, which is why we are interested in the components of Q for coordinates on shape space.

Under democracy transformations $\mathsf{D} \in \mathrm{SO}(n-1)$, the moment tensor K and the moment of inertia tensor M are invariant, but the Jacobi vectors transform according to Eq. (2.7). Thus first degree polynomials in the components of the Jacobi vectors $\rho_{\alpha i}$ transform according to the fundamental representation of $\mathrm{SO}(n-1)$ (actually the direct sum of three copies of the fundamental representation, one each for $i = 1, 2, 3$). None of these first degree polynomials is rotationally invariant.

The general second degree polynomial in the components of the Jacobi vectors is a linear combination of the monomials $\rho_{\alpha i}\, \rho_{\beta j}$; usually we are only interested in the quantities which are rotationally invariant, so we contract on i, j to obtain the Jacobi dot products $\boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$, the components of Q. These transform according to

$$Q'_{\alpha\beta} = \sum_{\mu,\nu=1}^{n-1} D_{\alpha\mu}\, D_{\beta\nu}\, Q_{\mu\nu}, \tag{2.41}$$

i.e., according to the symmetric part of the tensor product of the fundamental representation of $\mathrm{SO}(n-1)$ with itself (only the symmetric part because Q is a symmetric matrix). This tensor product is generally reducible, and yields a Clebsch-Gordan series of other irreducible representations of $\mathrm{SO}(n-1)$, which always includes the trivial representation (the scalar or democratic invariant). This is clear because the trace of Q, essentially the hyperradius, is a democratic invariant.

If we ask for rotational invariants among cubic polynomials in the components of the Jacobi vectors, we come up with the triple products $\boldsymbol{\rho}_\alpha \cdot (\boldsymbol{\rho}_\beta \times \boldsymbol{\rho}_\gamma)$, which transform according to the totally antisymmetric part of the threefold tensor product of the fundamental representation of $\mathrm{SO}(n-1)$ with itself. The Clebsch-Gordan series for this tensor product may or may not yield the trivial representation (i.e., a democratic invariant); it does so for $n = 4$.

Higher order polynomials which are rotational invariants can be formed by taking products of Jacobi dot products or triple products. For example, rotationally invariant polynomials which are quartic in the Jacobi vectors can be formed by taking products of the Jacobi dot products with themselves, i.e., they are quantities of the form $Q_{\alpha\beta}Q_{\mu\nu}$. It is straightforward to work out the Clebsch-Gordan series for these and to pick out the various irreducible representations of the democracy group. In later sections we will illustrate these operations with explicit examples.

I. A useful mapping and its properties

We now return to the Jacobi dot products and attempt to remedy their defects as coordinates on shape space.

These dot products are well-defined functions on shape space and are the $n(n-1)/2$ components of the matrix Q, which we think of as an element of the space of $(n-1)\times(n-1)$ symmetric matrices, or "symmetric matrix space" ($\mathcal{M}$) for short. The independent components of such a matrix can be used as coordinates on symmetric matrix space, so symmetric matrix space can be identified with $\mathbb{R}^{n(n-1)/2}$. A coordinate system on shape space ($\mathcal{S}$) will be a mapping from some region of shape space to some region of $\mathbb{R}^{3n-6}$; we now construct a related mapping by regarding the definition (2.22) of Q as a mapping f from shape space to symmetric matrix space,

$$f:\mathcal{S}\to\mathcal{M},\qquad f:s\mapsto \mathsf{Q}(s), \tag{2.42}$$

where s represents a shape, i.e., an equivalence class of configurations related by proper rotations, and where

$$Q_{\alpha\beta}(s)=\boldsymbol{\rho}_\alpha\cdot\boldsymbol{\rho}_\beta. \tag{2.43}$$

Here $\boldsymbol{\rho}_\alpha$ and $\boldsymbol{\rho}_\beta$ are Jacobi vectors chosen from any configuration in the equivalence class represented by s.

The mapping f has three important properties. First, if shapes s and s' are related by parity, $s'=Ps$, then $\mathsf{Q}(s)=\mathsf{Q}(s')$. This trivial property follows immediately from the invariance of the dot product under parity, and it implies that whenever $Ps\neq s$ the mapping f is at least two to one. As discussed above, for $n\geq 4$, only the planar configurations, a subset of measure zero, satisfy $Ps=s$; for all others, $Ps\neq s$. The two-to-one nature of the mapping f is what we have called the "parity ambiguity." Of course, we need a one-to-one mapping for a coordinate system.

According to the second property, however, the mapping f is never worse than two to one. That is, if $\mathsf{Q}(s)=\mathsf{Q}(s')$ for two shapes s and s', then either $s'=s$ or $s'=Ps$. To prove this, suppose we have two shapes s and s', let $\mathsf{Q}=\mathsf{Q}(s)$ and $\mathsf{Q}'=\mathsf{Q}(s')$, and suppose $\mathsf{Q}=\mathsf{Q}'$. We let λ_k and ξ^k_α be the eigenvalues and orthonormal eigenvectors of $\mathsf{Q}=\mathsf{Q}'$, as above. From the equivalence classes of configurations represented by s and s', we arbitrarily choose configurations $\{\boldsymbol{\rho}_\alpha\}$ and $\{\boldsymbol{\rho}'_\alpha\}$, and from these we construct the matrices K and K' as in Eq. (2.19), which are not in general equal. But since the nonzero eigenvalues of Q and those of K are the same, the eigenvalues of K and K' are the same, $\kappa_i=\kappa'_i$, $i=1,2,3$. As for the non-null eigenvectors of K and K', we use Eq. (2.32) to write

$$\hat{\mathbf{e}}_k=\frac{1}{\sqrt{\lambda_k}}\sum_{\alpha=1}^{n-1}\xi^k_\alpha\,\boldsymbol{\rho}_\alpha,\qquad \hat{\mathbf{e}}'_k=\frac{1}{\sqrt{\lambda_k}}\sum_{\alpha=1}^{n-1}\xi^k_\alpha\,\boldsymbol{\rho}'_\alpha, \tag{2.44}$$

for $k=1,\ldots,r$. These sets of vectors, $\{\hat{\mathbf{e}}_k\}$ and $\{\hat{\mathbf{e}}'_k\}$, form r-dimensional, orthonormal frames in the spaces spanned by $\{\boldsymbol{\rho}_\alpha\}$ and $\{\boldsymbol{\rho}'_\alpha\}$, respectively; these spaces have the same dimensionality. Now if $0\leq r\leq 2$, then any r-dimensional frame in $\mathbb{R}^3$ is related to any other such frame by some proper rotation, i.e., $\hat{\mathbf{e}}'_k=\mathsf{S}\hat{\mathbf{e}}_k$, for some $\mathsf{S}\in SO(3)$; and if $r=3$, then any three-frame in $\mathbb{R}^3$ is related to any other such three-frame by some (possibly improper) $\mathsf{S}\in O(3)$. But by Eq. (2.37) this implies that

$$\boldsymbol{\rho}'_\alpha=\mathsf{S}\boldsymbol{\rho}_\alpha. \tag{2.45}$$

Thus either configurations $\{\boldsymbol{\rho}_\alpha\}$ and $\{\boldsymbol{\rho}'_\alpha\}$ are related by some proper rotation (always the case when $0\leq r\leq 2$), meaning that $s'=s$, or else they are related by some improper rotation (which sometimes happens when $r=3$), meaning that $s'=Ps$.

The mapping f is certainly not onto symmetric matrix space, because the matrices $Q_{\alpha\beta}(s)$ constructed by Eq. (2.43) are always non-negative definite and have rank $r\leq 3$, and symmetric matrix space, as we have defined it, includes all symmetric matrices. On the other hand, it turns out that f is onto the subset of symmetric matrix space consisting of non-negative definite matrices of rank $r\leq 3$. That is, if Q is an $(n-1)\times(n-1)$ symmetric, non-negative definite matrix with rank $r\leq 3$, then there exists some shape s such that $\mathsf{Q}=\mathsf{Q}(s)$. This is the third property of the mapping f, and its significance is that if we can find coordinates on the subset of symmetric matrix space consisting of non-negative definite matrices with rank $r\leq 3$, then, apart from the parity ambiguity, we will have found coordinates on shape space. As we shall show, this subset of symmetric matrix space is a manifold of dimensionality $3n-6$.

To prove this property, we let Q be any $(n-1)\times(n-1)$ symmetric, non-negative definite matrix with rank $r\leq 3$. We construct its eigenvectors ξ^k_α and eigenvalues λ_k as in Eqs. (2.26), (2.27), and (2.28). Since Q is non-negative definite, we have $\lambda_k\geq 0$, and only the first $r\leq 3$ eigenvalues λ_k are nonzero. We order the eigenvalues as in Eq. (2.29). Thus we can write Q in terms of its eigenvalues and eigenvectors by

$$Q_{\alpha\beta}=\sum_{k=1}^{r}\xi^k_\alpha\,\lambda_k\,\xi^k_\beta. \tag{2.46}$$

Now we arbitrarily choose an r-dimensional, orthonormal frame $\{\hat{\mathbf{e}}_k\}$, $k=1,\ldots,r$ in $\mathbb{R}^3$, and use Eq. (2.37) to define Jacobi vectors $\boldsymbol{\rho}_\alpha$. Then by direct computation we find

$$\boldsymbol{\rho}_\alpha\cdot\boldsymbol{\rho}_\beta=\sum_{k,\ell=1}^{r}\sqrt{\lambda_k\lambda_\ell}\,\xi^k_\alpha\,\xi^\ell_\beta\,\hat{\mathbf{e}}_k\cdot\hat{\mathbf{e}}_\ell=\sum_{k=1}^{r}\xi^k_\alpha\,\lambda_k\,\xi^k_\beta=Q_{\alpha\beta}. \tag{2.47}$$

Thus, if s is the shape of configuration $\{\boldsymbol{\rho}_\alpha\}$, then $\mathsf{Q}=\mathsf{Q}(s)$, as claimed.

We will henceforth refer to the subset of symmetric matrix space for which Q is non-negative definite and has rank $r\leq 3$ as the "physical region," and the rest of symmetric matrix space as the "nonphysical region."

III. SHAPE COORDINATES FOR THE THREE-BODY PROBLEM

There is a large literature on the three-body problem, and many different coordinate systems have been used on the corresponding three-dimensional shape space. In this section we will pretend that we do not know any of this, but instead we will apply the general theory of Sec. II to the case $n = 3$. In the process we will illustrate the general theory, we will "discover" some coordinate systems in use, and we will prepare ourselves for the case $n = 4$ which we take up in the next section.

A. Coordinates on symmetric matrix space for $n = 3$

In the case $n = 3$ there are only two Jacobi vectors, $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$, and symmetric matrix space is the three-dimensional space of real, 2×2 symmetric matrices. We let Q be such a matrix (whether or not it can be written in the form $Q_{\alpha\beta} = \boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$, i.e., whether or not Q lies in the physical region). We write Q as a linear combination of the identity matrix and the Pauli matrices,

$$\mathsf{Q} = \frac{1}{2}(w\mathsf{I} + w_1\boldsymbol{\sigma}_z + w_2\boldsymbol{\sigma}_x), \tag{3.1}$$

where the expansion coefficients are given by

$$\begin{aligned} w &= \mathrm{tr}\,\mathsf{Q} = Q_{11} + Q_{22}, \\ w_1 &= \mathrm{tr}(\mathsf{Q}\boldsymbol{\sigma}_z) = Q_{11} - Q_{22}, \\ w_2 &= \mathrm{tr}(\mathsf{Q}\boldsymbol{\sigma}_x) = 2Q_{12}. \end{aligned} \tag{3.2}$$

The Pauli matrix $\boldsymbol{\sigma}_y$ is omitted because Q is symmetric. The coefficients (w, w_1, w_2) are convenient coordinates on symmetric matrix space; each ranges from $-\infty$ to $+\infty$, and symmetric matrix space itself is $\mathbb{R}^3$.

B. The physical region for $n = 3$

The physical region is the subset of symmetric matrix space for which Q is non-negative definite. It is not necessary to add the qualification $\mathrm{rank}\,\mathsf{Q} \le 3$, because Q is only a 2×2 matrix. The physical region is easily characterized in terms of the eigenvalues of Q; these are

$$\begin{aligned} \lambda_1 &= \frac{1}{2}\left(w + \sqrt{w_1^2 + w_2^2}\right), \\ \lambda_2 &= \frac{1}{2}\left(w - \sqrt{w_1^2 + w_2^2}\right), \end{aligned} \tag{3.3}$$

which are arranged so that $\lambda_1 \ge \lambda_2$. Specifically, the physical region is $\lambda_2 \ge 0$. Thus, in the coordinates (w, w_1, w_2), the boundary separating the physical from the nonphysical region is the surface $\lambda_2 = 0$, i.e.,

$$w = \sqrt{w_1^2 + w_2^2}. \tag{3.4}$$

This surface is a cone, and the physical region itself is the union of the interior of this cone with the boundary, as illustrated in Fig. 3.

On the boundary of the physical region we have $\lambda_1 \ge \lambda_2 = 0$ and $\mathrm{rank}\,\mathsf{Q} \le 1$, so the boundary consists of the collinear shapes. Most of the boundary points represent one-dimensional shapes, but at the apex of the cone we have $\lambda_1 = \lambda_2 = 0$ and $\mathrm{rank}\,\mathsf{Q} = 0$, so the apex represents the three-particle collision. In the interior of the cone we have $\lambda_2 \ge \lambda_1 > 0$ and $\mathrm{rank}\,\mathsf{Q} = 2$, so the interior shapes are two dimensional, representing triangles of nonvanishing area.

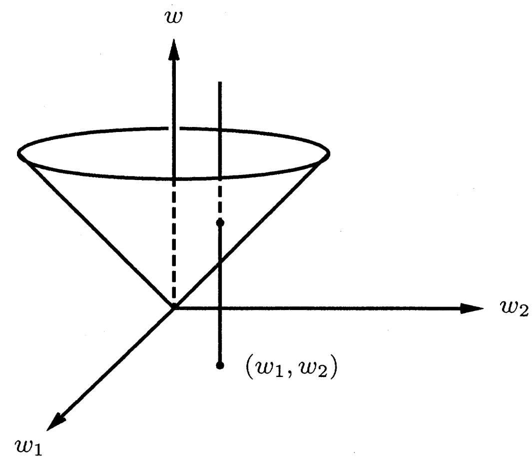

FIG. 3. Symmetric matrix space in the case $n = 3$ is $\mathbb{R}^3$, on which (w, w_1, w_2) are convenient coordinates. The surface of the cone (only the upper sheet, as illustrated) separates the physical from the nonphysical region. The surface itself contains collinear shapes, and the interior contains two-dimensional shapes.

In the physical region we are allowed to write $Q_{\alpha\beta} = \boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$, so that

$$w = \rho_1^2 + \rho_2^2, \quad w_1 = \rho_1^2 - \rho_2^2, \quad w_2 = 2\boldsymbol{\rho}_1 \cdot \boldsymbol{\rho}_2. \tag{3.5}$$

We note that w is the square of the "hyperradius" (see Smith [10]). To these we add

$$w_3 = 2|\boldsymbol{\rho}_1 \times \boldsymbol{\rho}_2| \ge 0, \tag{3.6}$$

to be used only in the physical region, and we note the identity

$$w^2 = w_1^2 + w_2^2 + w_3^2. \tag{3.7}$$

Thus, the surface of the cone is also characterized by $w_3 = 0$, or $\boldsymbol{\rho}_1 \times \boldsymbol{\rho}_2 = \mathbf{0}$ ($\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ are parallel). We note that w_3 is proportional to the area of the parallelogram spanned by $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$, and thus, by Eq. (2.11), to the area of the triangle formed by the three bodies. This area is considered to be strictly non-negative.

Topologically speaking, the boundary of the physical region is homeomorphic to $\mathbb{R}^2$, as is easily seen by projecting the surface of the cone onto the w_1-w_2 plane. That is, Eq. (3.4) specifies a one-to-one mapping between the surface of the cone and the w_1-w_2 plane. We will see an analog of this feature in the four-body problem.

There is no parity ambiguity in the three-body problem, because triangular shapes are invariant under parity, as discussed in Sec. II D, and the mapping f from shape space onto the physical region of symmetric matrix space is one to one. Therefore coordinates on the physical region of symmetric matrix space are also coordinates on shape space. But coordinates (w, w_1, w_2) are not convenient for the physical region, since the range of w has nonconstant bounds ($\sqrt{w_1^2 + w_2^2} \le w < \infty$). In this respect, w_3 is better than w. If we move from some point (w_1, w_2) in the w_1-w_2 plane parallel to the w axis, as illustrated in Fig. 3, then at a certain point we will puncture the surface of the cone and pass into the physical region. At the point of puncture, $w_3 = 0$, and beyond this point, w_3 increases monotonically. Therefore if we use (w_1, w_2, w_3) as coordinates in the physical region, we see that the physical region is characterized by $w_3 \ge 0$. These coordinates make it obvious that shape space itself is homeomorphic to half of $\mathbb{R}^3$, including the boundary $w_3 = 0$.

The transformation from w to w_3 is not differentiable at the apex of the cone, as is indicated geometrically by the transformation of the boundary of the physical region, which is the surface of a cone in the coordinates (w, w_1, w_2), and a plane ($w_3 = 0$) in the coordinates (w_1, w_2, w_3). Therefore scalar or tensor fields of interest on shape space which appear smooth in one coordinate system will appear nonsmooth in the other. In the three-body problem this lack of differentiability occurs only at the three-particle collision, which is a highly singular configuration, but it is a warning to beware of notions of smoothness unless they are connected with an invariant physical or geometrical meaning.

The coordinates (w_1, w_2, w_3), or variations of them, have been used by a number of authors, including Smith [19], Dragt [21], Mead and Truhlar [14], Iwai [3], and Pack and Parker [17]. One simple variation, suggested by the identity (3.7), is to transform to spherical coordinates in (w_1, w_2, w_3) space; this gives the coordinates (w, χ, ψ), defined by

$$\begin{aligned} w_1 &= w \cos\chi \cos\psi, \\ w_2 &= w \cos\chi \sin\psi, \\ w_3 &= w \sin\chi. \end{aligned} \tag{3.8}$$

The angles (χ, ψ) are "hyperspherical" angles (see Smith [10] or Pack and Parker [17]).

C. Group theoretical significance of the coordinates

In the following we present a group theoretical analysis of various quantities of interest in the three-body problem, and show in a different way how one might "discover" the coordinates we have introduced above. This analysis is overly formalistic for the purposes of the three-body problem, but it is useful practice for our later treatment of the four-body problem.

Shape coordinates are automatically invariant under rotations $\mathsf{R} \in SO(3)$, as in Eq. (2.8), but they may have different transformation properties under democracy transformations. In the three-body problem, the democracy group is simply SO(2), whose one-dimensional irreducible representations can be labeled by the "magnetic quantum number" $m = \ldots -1, 0, 1, 2, \ldots$, with characters $\exp(im\alpha)$. The fundamental representation of SO(2) consists of the matrices

$$\mathsf{D}(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \tag{3.9}$$

for $0 \le \alpha < 2\pi$. This representation is reducible and contains the irreducible representations $m = -1$ and $m = +1$,

$$SO(2) = -1 \oplus +1, \tag{3.10}$$

although the reduction requires complex transformations.

Under democracy transformations, the Jacobi dot products transform according to the symmetric part of $SO(2) \otimes SO(2)$, or

$$[(-1 \oplus +1) \otimes (-1 \oplus +1)]_{\text{sym}} = -2 \oplus 0 \oplus +2. \tag{3.11}$$

(The antisymmetric part gives a single copy of $m = 0$.) Thus, of the three independent components of $Q_{\alpha\beta}$ in the three-body problem, there is one linear combination which is a democratic invariant ($m = 0$); this is just the quantity w introduced in Eq. (3.5). As for the representations $m = \pm 2$, these are complex, but can be combined into a single, two-dimensional real representation. The two quantities transforming under this representation are w_1 and w_2, as shown by

$$\begin{pmatrix} w_1' \\ w_2' \end{pmatrix} = \begin{pmatrix} \cos 2\alpha & -\sin 2\alpha \\ \sin 2\alpha & \cos 2\alpha \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \tag{3.12}$$

Alternatively, the quantities $w_1 \pm i w_2$ transform according to the complex but one-dimensional $m = \pm 2$ representations. Geometrically, Eq. (3.12) shows that democracy transformations are rotations about the w axis in the coordinates (w, w_1, w_2), or about the w_3 axis in the coordinates (w_1, w_2, w_3), by twice the angle α; thus 2α can be identified with the azimuthal hyperspherical angle ψ as in Eq. (3.8). This fact is widely recognized in the literature on the three-body problem (e.g., Pack and Parker [17]).

Three-body shape space is three dimensional, and the democracy group is one dimensional. Therefore we expect there to be at least two democratic invariants on shape space, but so far we have only one, namely, w. We have not seen any others yet because there are no others which can be formed from quadratic polynomials in the Jacobi vectors. To see the other invariant, we must go to higher order polynomials. Cubic polynomials will not do, because the threefold tensor product of SO(2) with itself does not contain the representation $m = 0$, and in any case there are no rotational invariants among the cubic polynomials because we cannot construct a triple product with only two Jacobi vectors. In fact, the new invariant turns up among the quartic polynomials in $\boldsymbol{\rho}_\alpha$. The rotationally invariant, quartic polynomials in the Jacobi vectors are quadratic polynomials in the quantities (w, w_1, w_2); but since w already belongs to the representation $m = 0$ we might as well discard it, since it can never create a new $m = 0$ quantity in combination with w_1 and w_2. Thus we consider quadratic polynomials which are linear combinations of $w_i w_j$, $i, j = 1, 2$, where the w_i transform according to Eq. (3.12). There are three independent such polynomials, which transform according to

$$[(-2 \oplus +2) \otimes (-2 \oplus +2)]_{\text{sym}} = -4 \oplus 0 \oplus +4. \quad (3.13)$$

The $m = 0$ quantity in this decomposition is just the obvious invariant of the transformation (3.12); we write a for this invariant, which is

$$a = w_1^2 + w_2^2. \quad (3.14)$$

Thus w and a are the two democratic invariants in the three-body problem. Finally, the quantity w_3 introduced in Eq. (3.6) is just a simple function of the two invariants,

$$w_3 = \sqrt{w^2 - a}. \quad (3.15)$$

A more direct way to construct the democratic invariants is to use the obvious invariants of the matrix Q, i.e., its trace and determinant. Indeed, we find

$$w = \operatorname{tr} \mathsf{Q}, \qquad w_3^2 = 4 \det \mathsf{Q}. \quad (3.16)$$

Alternatively, we can use the eigenvalues λ_1 and λ_2, which are also democratic invariants.

The action of the democracy group on three-body shape space suggests that the natural coordinates to use on shape space would be (w, w_3, α), that is, two democratic invariants plus the angle α of the democracy transformation as in Eq. (3.9). Of course, any two functions of the democratic invariants could be used instead; it is hard to argue that one invariant is better than another. In practice, however, it seems that either coordinates (w_1, w_2, w_3) or (w, χ, ψ) are more convenient for the three-body problem. The reason is that the various tensor fields of interest on the three-body shape space (such as the metric tensor or the Coriolis tensor) are invariant under a larger group than the democracy group, namely, under a certain SO(3) group which contains the democracy group SO(2) as a subgroup. This SO(3) symmetry group consists of rotations in the usual sense in the (w_1, w_2, w_3) coordinates, and its presence is related to the fact that the holonomy group for the three-body problem is only SO(2) and that the planar three-body problem has a larger symmetry group than just the democracy group. There is nothing analogous in the four-body problem, for which the various tensor fields of interest are invariant under the democracy group but not a larger group.

IV. SHAPE COORDINATES FOR THE FOUR-BODY PROBLEM

We now apply the theory of Sec. II to the case $n = 4$. This case is the first one for which the generic features occurring at large n occur; the three-body problem is a special case (just as is the two-body problem). Our main accomplishments in this section are to show that four-body shape space is homeomorphic to $\mathbb{R}^6$ (a fact first proven by Narasimhan and Ramadas [8]), to find convenient coordinates on this space, to identify various submanifolds of interest, and to explore the action of the democracy group on shape space.

A. Coordinates on symmetric matrix space for $n = 4$

In the case $n = 4$ there are three Jacobi vectors, $\boldsymbol{\rho}_1$, $\boldsymbol{\rho}_2$, $\boldsymbol{\rho}_3$, and symmetric matrix space is the six-dimensional space of real, symmetric, 3×3 matrices Q. The democracy group for $n = 4$ is SO(3); of course the democracy group is not to be confused with the rotation group, also SO(3) as an abstract group. [The rotation group acts on configuration space (not shape space), according to Eq. (2.8).] We will denote the irreducible representations of the democracy group by $\ell = 0, 1, 2, \ldots$; the fundamental representation SO(3) is the irreducible representation $\ell = 1$.

A convenient basis of matrices for symmetric matrix space consists of the identity I and the following five matrices:

$$\mathsf{B}_1 = \sqrt{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathsf{B}_2 = \sqrt{3} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\mathsf{B}_3 = \sqrt{3} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathsf{B}_4 = \sqrt{3} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$
$$\mathsf{B}_5 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \quad (4.1)$$

These matrices obey

$$\operatorname{tr} \mathsf{B}_i = 0, \quad (4.2)$$
$$\operatorname{tr}(\mathsf{B}_i \mathsf{B}_j) = 6\, \delta_{ij}, \quad (4.3)$$

for $i, j = 1, \ldots, 5$. The matrices B_i are the analogs of the Pauli matrices for the case $n = 3$. Their definition is motivated by the action of the democracy group on symmetric matrix space, shown in Eq. (2.41); this action consists of the symmetric part of SO(3)×SO(3), which decomposes according to the ℓ values

$$(1 \otimes 1)_{\text{sym}} = 0 \oplus 2. \quad (4.4)$$

Thus the identity matrix spans the one-dimensional $\ell = 0$ subspace of symmetric matrix space, and the matrices $\{\mathsf{B}_i\}$ span the five-dimensional $\ell = 2$ subspace of symmetric and traceless matrices.

We write an arbitrary symmetric 3×3 matrix Q (not necessarily in the physical region) as a linear combination of the basis matrices, according to

$$\mathsf{Q} = \frac{1}{3} \left(w \mathsf{I} + \sum_{i=1}^{5} w_i \mathsf{B}_i \right), \quad (4.5)$$

so that $(w; w_1, \ldots, w_5)$ are coordinates on symmetric matrix space. Thus we have

$$w = \operatorname{tr} \mathsf{Q}, \qquad w_i = \frac{1}{2} \operatorname{tr}(\mathsf{Q} \mathsf{B}_i), \quad (4.6)$$

or

$$
\begin{aligned}
w &= Q_{11} + Q_{22} + Q_{33}, \quad w_3 = \sqrt{3}\, Q_{23}, \\
w_1 &= \frac{\sqrt{3}}{2}(Q_{11} - Q_{22}), \quad w_4 = \sqrt{3}\, Q_{31}, \\
w_2 &= \sqrt{3}\, Q_{12}, \quad w_5 = \frac{1}{2}(-Q_{11} - Q_{22} + 2Q_{33}). \qquad (4.7)
\end{aligned}
$$

The coordinates $(w; w_1, \ldots, w_5)$ transform under democracy transformations just like the corresponding basis matrices; w is the $\ell = 0$ democratic invariant, and the w_i transform according to $\ell = 2$. All six coordinates range from $-\infty$ to $+\infty$, and symmetric matrix space itself is $\mathbb{R}^6$. Keating and Mead [7] used a similar set of coordinates, making reference to interparticle distances rather than mass-weighted Jacobi vectors.

B. The physical region for $n = 4$

The physical region is the subset of symmetric matrix space for which Q is non-negative definite. It is not necessary to add the qualification that rank Q ≤ 3, since Q is only a 3×3 matrix. Inside the physical region we are allowed to write $Q_{\alpha\beta} = \boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$, so the coordinates become

$$
\begin{aligned}
w &= \rho_1^2 + \rho_2^2 + \rho_3^2, & w_3 &= \sqrt{3}\, \boldsymbol{\rho}_2 \cdot \boldsymbol{\rho}_3, \\
w_1 &= \frac{\sqrt{3}}{2}(\rho_1^2 - \rho_2^2), & w_4 &= \sqrt{3}\, \boldsymbol{\rho}_3 \cdot \boldsymbol{\rho}_1, \\
w_2 &= \sqrt{3}\, \boldsymbol{\rho}_1 \cdot \boldsymbol{\rho}_2, & w_5 &= \frac{1}{2}(-\rho_1^2 - \rho_2^2 + 2\rho_3^2).
\end{aligned} \qquad (4.8)
$$

We note that w is again the square of the hyperradius. To these equations we add the definition

$$
w_6 = \sqrt{3}\left(|\boldsymbol{\rho}_1 \times \boldsymbol{\rho}_2|^2 + |\boldsymbol{\rho}_2 \times \boldsymbol{\rho}_3|^2 + |\boldsymbol{\rho}_3 \times \boldsymbol{\rho}_1|^2\right)^{1/2} \geq 0, \qquad (4.9)
$$

to be used only in the physical region, and we note the identity

$$
w^2 = \left(\sum_{i=1}^{5} w_i^2\right) + w_6^2. \qquad (4.10)
$$

In terms of the three eigenvalues of Q, ordered according to $\lambda_1 \geq \lambda_2 \geq \lambda_3$, the physical region is characterized by $\lambda_3 \geq 0$. In Sec. III we used explicit formulas for the eigenvalues, Eq. (3.3), to study the physical region. But in the four-body problem the eigenvalues are roots of a cubic, and it is not easy to work with their explicit expressions. Therefore we take another approach.

We imagine the $(w; w_1, \ldots, w_5)$ coordinates in symmetric matrix space as illustrated in Fig. 4, in which the five-dimensional hyperplane $w = 0$ is illustrated schematically as if it were two dimensional. The eigenvalues λ_k are functions of the six coordinates $\lambda_k = \lambda_k(w; w_1, \ldots, w_5)$; we denote the eigenvalues evaluated on the five-dimensional hypersurface $w = 0$ by λ_{k0}, so that

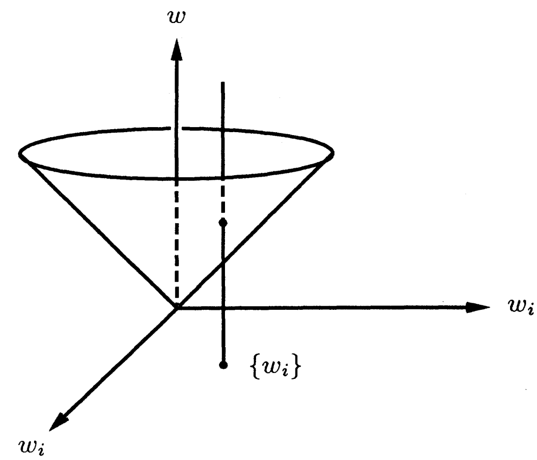

FIG. 4. Symmetric matrix space in the case $n = 4$ is $\mathbb{R}^6$, on which $(w; w_1, \ldots, w_5)$ are convenient coordinates. On the five-dimensional hyperplane $w = 0$, there is only one point which lies in the physical region; this is the origin $w_i = 0$, representing the four-body collision. At all other points on this hyperplane, at least one eigenvalue λ_{k0} is negative. The surface dividing the physical from the nonphysical region is a higher dimensional analog of a cone, and is homeomorphic to $\mathbb{R}^5$.

$$
\lambda_{k0} = \lambda_{k0}(w_1, \ldots, w_5) = \lambda_k(0; w_1, \ldots, w_5). \qquad (4.11)
$$

Then we claim that

$$
\lambda_k(w; w_1, \ldots, w_5) = \frac{w}{3} + \lambda_{k0}(w_1, \ldots, w_5). \qquad (4.12)
$$

The proof follows immediately from the secular equation for Q, which is

$$
\det(\mathsf{Q} - \lambda_k \mathsf{I}) = \det\left[\frac{1}{3}\sum_{i=1}^{5} w_i \mathsf{B}_i - \left(\lambda_k - \frac{w}{3}\right)\mathsf{I}\right] = 0. \qquad (4.13)
$$

Thus $\lambda_{k0} = \lambda_k - w/3$.

Now since $w = \operatorname{tr} \mathsf{Q}$ on the hyperplane $w = 0$ we have

$$
\lambda_{10} + \lambda_{20} + \lambda_{30} = 0. \qquad (4.14)
$$

Therefore either all three λ_{k0} vanish, or else some are positive and some are negative. The vanishing of all three λ_{k0} implies rank Q $= 0$ or equivalently $w_i = 0$, $i = 1, \ldots, 5$; this is the four-body collision, and it is the only point of the hyperplane $w = 0$ which is contained in the physical region. It is the analog of the apex of the cone in the three-body problem. At all other points on the hyperplane $w = 0$, at least one eigenvalue λ_{k0} must be negative (certainly the last one, λ_{30}), so these points lie outside the physical region. A typical arrangement for the three eigenvalues on the λ axis is illustrated in Fig. 5.

Now let us choose a point $\{w_i\}$ on the hyperplane $w = 0$, and move parallel to the w axis, as illustrated in Fig. 4. According to Eq. (4.12), as we move along this line in the direction of increasing w, all eigenvalues λ_k

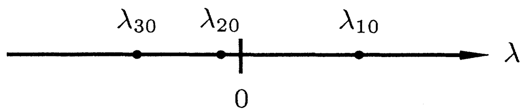

FIG. 5. Typical appearance of the three eigenvalues λ_{k0} of Q on the hyperplane $w = 0$. Except at the origin, at least one of the eigenvalues must be negative.

increase monotonically, their differences remaining constant. When we reach the point

$$w = -3\lambda_{30}(w_1, \ldots, w_5) \geq 0, \tag{4.15}$$

the eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \lambda_3 = 0$, and we are at the boundary of the physical region. On this boundary, $\operatorname{rank} \mathsf{Q} \leq 2$, so the boundary configurations are planar. At all larger w values, we have $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$ and $\operatorname{rank} \mathsf{Q} = 3$, so such matrices lie inside the physical region and the corresponding shapes are three dimensional.

Equation (4.15) can be thought of as a one-to-one mapping from points of the hyperplane $w = 0$ to the points on the boundary of the physical region, and shows that this boundary is homeomorphic to $\mathbb{R}^5$. In other words, Eq. (4.15) is the equation for a single-valued surface over the hyperplane $w = 0$. The value of w occurring in Eq. (4.15) is always non-negative, and takes on the value $w = 0$ only at the origin of the coordinates. The surface itself is a higher dimensional analog of the cone we found in the three-body problem, which is why it is illustrated as a cone in Fig. 4. But the illustration is misleading in certain respects; in particular, when the boundary surface is represented in the coordinates $(w; w_1, \ldots, w_5)$, it is not smooth at places where the lowest two eigenvalues are equal, $\lambda_2 = \lambda_3 = 0$. The vanishing of two eigenvalues implies $\operatorname{rank} \mathsf{Q} \leq 1$, which implies in turn the collinear shapes. The boundary surface [as viewed in the coordinates $(w; w_1, \ldots, w_5)$] is continuous at such points, but it is not differentiable; in a moment we will introduce a coordinate transformation which irons out these singularities. Again, we have a warning to beware of notions of smoothness, this time in the neighborhood of collinear shapes. It is well known (e.g., Watson [24]) that the dynamics of an n-body system has singular features in the neighborhood of collinear configurations, but there is much more that can be said about this subject.

C. Coordinates on shape space for $n = 4$

Unlike the case $n = 3$, for $n = 4$ the mapping f from shape space onto the physical region is not one to one. It is one to one for planar shapes, which we now recognize as the boundary of the physical region; but it is two to one for all points inside the physical region. Therefore to find coordinates on shape space we need coordinates which cover the physical region twice, except for the boundary points, which are to be covered once.

To this end, we introduce the quantity V,

$$V = \boldsymbol{\rho}_1 \cdot (\boldsymbol{\rho}_2 \times \boldsymbol{\rho}_3), \tag{4.16}$$

which is a rotational invariant and therefore a single-valued function on shape space. The geometrical meaning of V is that it is the signed volume of the parallelepiped spanned by the three Jacobi vectors, which to within a constant factor is the volume of the tetrahedron formed by the four particles. To express V as a function on symmetric matrix space, we note the identity

$$\det \mathsf{Q} = (\det \mathsf{F})^2 = V^2, \tag{4.17}$$

which is meaningful because in the four-body problem the matrix F is square. Therefore we have

$$V = \pm\sqrt{\det \mathsf{Q}} = \pm\sqrt{\lambda_1\lambda_2\lambda_3}, \tag{4.18}$$

or

$$V^2 = \left(\frac{w}{3} + \lambda_{10}\right)\left(\frac{w}{3} + \lambda_{20}\right)\left(\frac{w}{3} + \lambda_{30}\right). \tag{4.19}$$

We see that V, regarded as a function on symmetric matrix space, is real in the physical region, its value is zero on the boundary (the planar shapes), and it is double valued in the interior of the physical region. The $\pm$ sign resolves the parity ambiguity for shapes of nonzero volume. Furthermore, V^2 increases monotonically with w as we move past the boundary parallel to the w axis, as in Fig. 4, so the range of V, for all values of w_i, is $-\infty$ to $+\infty$.

Therefore suitable coordinates on shape space are $(V; w_1, \ldots, w_5)$, each of which ranges from $-\infty$ to $+\infty$. These coordinates provide the proof that shape space in the four-body problem is homeomorphic to $\mathbb{R}^6$. If we identify shape space with $\mathbb{R}^6$ through these coordinates, then we see that the five-dimensional hyperplane $V = 0$ contains the planar shapes; in these coordinates, all the lack of smoothness that was present in the boundary surface of planar shapes in symmetric matrix space has been ironed out.

D. Democratic invariants for $n = 4$

Since shape space for $n = 4$ is six dimensional and the democracy group SO(3) is three dimensional, we expect there to be at least three independent democratic invariants, of which $w = \operatorname{tr} \mathsf{Q}$ is one. In fact, there are precisely three independent democratic invariants, since the manifold swept out by allowing the democracy group to act on a typical point of shape space is three dimensional. For a typical point of shape space, the matrix Q has distinct eigenvalues, so the subgroup of the democracy group which leaves this point invariant is a discrete (0-dimensional) subgroup. There are exceptional points of shape space constituting a subset of measure zero where these statements are not true, but for now we will concentrate on the typical points.

The easy way to find the three democratic invariants is to write down the invariants of the matrix Q, which can be identified with the coefficients of the secular polynomial. We write this polynomial in the form

$$P(\lambda) = -\det(\mathsf{Q} - \lambda \mathsf{I}) = \lambda^3 - c_2\lambda^2 + c_1\lambda - c_0. \tag{4.20}$$

Of the coefficients, c_2 is the previously found invariant, $\operatorname{tr} \mathsf{Q} = w$. The coefficient c_1 is new, and can be written in several forms,

$$\begin{aligned} c_1 &= Q_{11}Q_{22} + Q_{22}Q_{33} + Q_{33}Q_{11} - Q_{12}^2 - Q_{23}^2 - Q_{31}^2 \\ &= \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 = \frac{1}{3}(w^2 - a) = \frac{1}{3}w_6^2, \end{aligned} \quad (4.21)$$

where

$$a = \sum_{i=1}^{5} w_i^2 \geq 0. \quad (4.22)$$

By Eq. (4.10), we have

$$w \geq \sqrt{a}, \quad (4.23)$$

in the physical region. The final coefficient is

$$c_0 = \det \mathsf{Q} = \lambda_1\lambda_2\lambda_3 = V^2 = \frac{1}{27}(w^3 - 3aw + b), \quad (4.24)$$

where

$$\begin{aligned} b = 2w_5^3 &- 6w_5(w_1^2 + w_2^2) + 3w_5(w_3^2 + w_4^2) \\ &-3\sqrt{3}\, w_1(w_3^2 - w_4^2) + 6\sqrt{3}\, w_2w_3w_4. \end{aligned} \quad (4.25)$$

In deriving these equations, we have solved Eqs. (4.7) for the components of Q, and used the results to express things in terms of the coordinates $(w; w_1, \ldots, w_5)$.

The three democratic invariants for the four-body problem can be taken to be (c_2, c_1, c_0) or (w, a, b) or $(\lambda_1, \lambda_2, \lambda_3)$ or any invertible functions of these; note that a and b depend only on $(w_1, \ldots, w_5)$, and are independent of w. Note also that V^2 is a democratic invariant; so is V itself, if we restrict the democracy group to SO(3) [instead of O(3)].

The secular polynomial simplifies somewhat when expressed in terms of $\lambda_0 = \lambda - w/3$. If we write $P_0(\lambda_0) = P(\lambda)$, then by Eq. (4.12) we obtain $P_0(\lambda_0)$ by setting $w = 0$ and $\lambda = \lambda_0$ in $P(\lambda)$. Thus,

$$P_0(\lambda_0) = \lambda_0^3 - \frac{1}{3}a\lambda_0 - \frac{1}{27}b. \quad (4.26)$$

If one has to find the eigenvalues λ of Q, the easiest way is to solve first for the roots λ_0 of $P_0(\lambda_0)$, which depend only on the invariants a and b, and then to use Eq. (4.12).

The democratic invariants can also be obtained from a group theoretical approach. As noted earlier, the coordinates $(w; w_1, \ldots, w_5)$ are rotationally invariant quadratic polynomials in the Jacobi vectors, transforming according to $\ell = 0$ and $\ell = 2$. To examine quartic polynomials in the Jacobi vectors, we discard the invariant w and form the 15 independent monomials w_iw_j, $i, j = 1, \ldots, 5$. These transform according to

$$(2 \otimes 2)_{\mathrm{sym}} = 0 \oplus 2 \oplus 4, \quad (4.27)$$

so they contain precisely one democratic invariant; this is simply a. To obtain a third invariant, we move on to sixth degree polynomials in the Jacobi vectors, i.e., cubic polynomials in the w_i. These transform according to

$$(2 \otimes 2 \otimes 2)_{\mathrm{sym}} = 0 \oplus 2 \oplus 3 \oplus 4 \oplus 6, \quad (4.28)$$

which again contains a single invariant; this turns out to be b, so that the funny coefficients seen in Eq. (4.25) are essentially vector coupling coefficients for SO(3). We can also examine cubic polynomials in the Jacobi vectors. The only one of these which is rotationally invariant is the triple product $V = \boldsymbol{\rho}_1 \cdot (\boldsymbol{\rho}_2 \times \boldsymbol{\rho}_3)$, which is the totally antisymmetric part of $1 \otimes 1 \otimes 1$, which is just a single scalar ($\ell = 0$).

It is also convenient to express the invariants of the moment of inertia tensor in terms of (w, a, b). From Eq. (2.38) we have $\operatorname{tr} \mathsf{M} = 2w$, and

$$\begin{aligned} \det \mathsf{M} &= (\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3)(\lambda_3 + \lambda_1) \\ &= c_2c_1 - c_0 \\ &= (\lambda_1 + \lambda_2 + \lambda_3)(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) - \lambda_1\lambda_2\lambda_3 \\ &= \frac{1}{27}(8w^3 - 6aw - b). \end{aligned} \quad (4.29)$$

E. Interesting submanifolds of shape space

We now study certain interesting submanifolds of shape space. Technically, these are not proper differentiable manifolds, because they contain points where they are not smooth.

The polynomial $P_0(\lambda_0)$ has a local maximum at $\lambda_0 = -\sqrt{a}/3$ and a local minimum at $\lambda_0 = +\sqrt{a}/3$; a graph of this polynomial for typical values of a and b is shown in Fig. 6. Since the roots of P_0 are real, the maximum must be non-negative and the minimum nonpositive. This leads to the inequality

$$-2a^{3/2} \leq b \leq +2a^{3/2}, \quad (4.30)$$

which is not at all easy to prove from the definitions Eqs. (4.22) and (4.25). When the locally maximum value of P_0 vanishes, then $b = +2a^{3/2}$ and the two lowest roots are equal, $\lambda_{20} = \lambda_{30} = -\sqrt{a}/3$, with $\lambda_{10} = +2\sqrt{a}/3$. Likewise, when the locally minimum value of P_0 vanishes, then $b = -2a^{3/2}$ and the two largest roots are equal, $\lambda_{10} = \lambda_{20} = +\sqrt{a}/3$, with $\lambda_{30} = -2\sqrt{a}/3$. An alternate

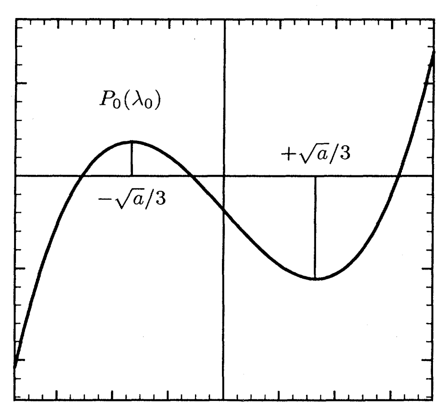

FIG. 6. Typical graph of the polynomial $P_0(\lambda_0)$. Local extrema occur at $\lambda = \pm\sqrt{a}/3$. The roots are real and sum to zero.

way to prove inequality (4.30) is to use Eqs. (4.21) and (4.24) to express a and b in terms of the eigenvalues of Q. One then finds that

$$4a^3 - b^2 = 27(\lambda_1 - \lambda_2)^2(\lambda_2 - \lambda_3)^2(\lambda_3 - \lambda_1)^2, \quad (4.31)$$

which shows that $4a^3 - b^2$ is non-negative.

These algebraic relations connecting a, b, and the three eigenvalues λ_{0k}, all of which are functions of the five w_i, can be viewed geometrically either in the five-dimensional hyperplanes w = const in symmetric matrix space, or in the five-dimensional hyperplanes V = const in shape space. For example, if we take a point $(w_1, \ldots, w_5)$ in the hyperplane $w = 0$ in symmetric matrix space such that $b = +2a^{3/2}$ and project it onto the boundary of the physical region ($V = 0$) according to Eq. (4.15) or as illustrated in Fig. 4, then $w = \sqrt{a} = \lambda_1$ and $\lambda_2 = \lambda_3 = 0$. At such a point of matrix space rank $\mathsf{Q} \leq 1$ and the corresponding shapes are collinear. Conversely, every collinear shape can be obtained in this manner. Therefore the two equations $b = +2a^{3/2}$ and $V = 0$ characterize the collinear shapes; such shapes are properly one dimensional, unless $a = 0$, in which case we have the four-body collision. It would appear that the manifold of collinear shapes is four dimensional, since we have two equations connecting the six coordinates $(V; w_1, \ldots, w_5)$ specifying it; but actually the manifold of collinear shapes is three dimensional (the functions in the two equations are not smooth at the points in question). It is easy to see this from another point of view, because any collinear shape is specified by the coordinates of the three Jacobi vectors along a single direction. Similarly, the submanifold of zero-dimensional shapes is zero dimensional (the single shape representing the four-body collision).

Another submanifold of interest in shape space is that upon which the moment of inertia tensor is degenerate, since at such points the principal axis frame is not uniquely defined. We will call this the "degeneracy submanifold." The principal axis frame, assumed to be right handed, is defined at all other points of shape space and is unique modulo proper (right handed) inversions of some of the eigenvectors. [We can exclude permutations of the eigenvectors if we sequence the eigenvalues of K according to Eq. (2.24).] In any contractible region of shape space, excluding the degeneracy submanifold, the principal axis frame can be defined as a smooth function of shape. The principal axis frame does not, however, approach a unique limit as the degeneracy manifold is approached, so the shape derivatives of the eigenvectors become infinite as the degeneracy manifold is approached. This in turn implies a divergence of the gauge potential $\mathbf{A}_\mu$, so the manifold of degeneracy of the moment of inertia tensor is also the singular manifold of the gauge potential (the "string singularity") in the principal axis gauge. These are all reasons for being interested in the degeneracy manifold of the moment of inertia tensor.

The moment of inertia tensor is degenerate if and only if $\lambda_2 = \lambda_3$ or $\lambda_1 = \lambda_2$. In view of Eq. (4.12), these conditions are true for any w (or for any V) if and only if $\lambda_{20} = \lambda_{30}$ or $\lambda_{10} = \lambda_{20}$, i.e, $b = \pm 2a^{3/2}$. The manifold $b = -2a^{3/2}$ inside the hyperplane $V = 0$ looks exactly like the collinear manifold $b = +2a^{3/2}$ inside the same hyperplane and joins with it at the four-body collision. In particular, both manifolds are three dimensional. Therefore the degeneracy manifold is formed by taking the union of these two three-dimensional manifolds in the hyperplane $V = 0$ and translating them parallel to the V axis. Thus the degeneracy manifold is four dimensional.

Narasimhan and Ramadas [8] also considered submanifolds of shape space and drew some topological conclusions about them. The treatment given in the present paper can be used to derive some of their results by alternate means. As an example of the utility of the $(V; w_1, \ldots, w_5)$ coordinate system, we present a simple proof of the homeomorphism

$$B = \mathbb{R} \times (S^5 - P), \quad (4.32)$$

where B is the set of shapes with dimensionality greater than or equal to 2, and P is a submanifold of the five-sphere S^5 homeomorphic to the projective plane $\mathbb{R}P^2$. The manifold B is of interest because it is the subset of shape space consisting of noncollinear shapes, and such shapes are acted upon freely by the rotation group. In other words, B is the base space of the principal SO(3) fiber bundle obtained from the set of noncollinear configurations.

B is homeomorphic to the subset of the $(V; w_1, \ldots, w_5)$ parameter space obtained by removing the points which represent collinear shapes. As discussed above, points representing collinear shapes satisfy $b = +2a^{3/2}$ and $V = 0$. These conditions are homogeneous in the coordinates, so we are removing a set of rays from $\mathbb{R}^6$. The set which remains is $\mathbb{R} \times (S^5 - P)$, where P is a submanifold of S^5 to be determined. Because of the condition $V = 0$, P is homeomorphic to the set of points that are on a sphere centered at the origin in $(w_1, \ldots, w_5)$ space and that represent collinear shapes. For concreteness, let the radius of this sphere be $\sqrt{a_0}$ where a_0 is an arbitrary positive constant. If we define C to be the set of all shapes that are collinear and have $a = a_0$, then P is homeomorphic to C. We define a mapping from a two-sphere of radius $(a_0)^{1/4}$ imbedded in $\mathbb{R}^3$ to C by mapping (x_1, x_2, x_3) to the shape of the collinear configuration given by $\{\boldsymbol{\rho}_\alpha = x_\alpha \mathbf{u}, \alpha = 1, 2, 3\}$, where $\mathbf{u}$ is an arbitrary constant unit vector in physical space. It is clear that this mapping is onto C and that it is everywhere two to one since (x_1, x_2, x_3) and $-(x_1, x_2, x_3)$ get mapped to the same shape. Thus P is homeomorphic to S^2 modulo the equivalence relation of defining antipodal points to be equivalent. This is $\mathbb{R}P^2$.

F. Foliation of shape space by the democracy group

Visualization of shape space for $n = 4$ or the submanifolds within it is not easy because of the high dimensionalities involved. This task is made much easier by foliating shape space under the action of the democracy group, i.e., dividing shape space into the family of "orbits" of the democracy group. The same foliation arises in the construction of basis wave functions which transform according to irreducible representations of the democracy group. In this section we will consider the democracy group to be SO(3) (excluding the improper transformations).

If we identify a point of shape space by the symmetric matrix $\mathsf{Q}_{\alpha\beta} = \boldsymbol{\rho}_\alpha \cdot \boldsymbol{\rho}_\beta$, then the action of the democracy group on this point is given by Eq. (2.41). Thus the orbit of a matrix Q under the action of the democracy group is the set $\{\mathsf{DQD}^t | \mathsf{D} \in \mathrm{SO}(3)\}$. The symmetric matrix Q does not uniquely identify a point of shape space due to the parity ambiguity, but, since this ambiguity is resolved by the sign of the coordinate V and since V itself is a democratic invariant, the orbit of a point of shape space can be identified with the orbit of the corresponding matrix Q.

Since every symmetric matrix Q can be diagonalized by some proper orthogonal D, every orbit contains a diagonal matrix. In fact, unless all the eigenvalues are equal, it contains several diagonal matrices, corresponding to the permutations of the eigenvalues. In any case, every orbit contains a unique diagonal matrix satisfying the eigenvalue ordering (2.29); we will call this the "principal diagonal matrix."

If all three eigenvalues are distinct, then the orbit of the democracy group is three dimensional, as we can see by considering the actions of infinitesimal democracy transformations taken about the three eigenvectors of the principal diagonal matrix (these actions produce three linearly independent matrices). The orbit is not a copy of the group SO(3), because there is a discrete subgroup of SO(3) which leaves the principal diagonal matrix invariant; this is the viergruppe of proper diagonal orthogonal matrices, or the group D_2 in the Schoenflies notation [25], which has the effect of changing the directions of either zero or two of the eigenvectors of Q. Therefore the orbit is the three-dimensional space of cosets of SO(3) with respect to this discrete subgroup.

If two eigenvalues of Q are equal and the third distinct, then the orbit is two dimensional, being the space of cosets of SO(3) with respect to a certain one-dimensional subgroup. Such matrices correspond to the degeneracy manifold discussed above. Finally, if all three eigenvalues of Q are equal, then the orbit is just a (zero dimensional) point (because a multiple of the identity matrix is invariant under democracy transformations).

Coordinates $(V; w_1, \ldots, w_5)$ are convenient for examining the action of the democracy group. Since V is democratic invariant, we can work in any five-dimensional hyperplane $V =$ const, and study the action of the group on the five w_i. This action is independent of V and looks the same in any hyperplane $V =$ const. Furthermore, diagonal matrices are represented in the coordinates $(w_1, \ldots, w_5)$ by $w_2 = w_3 = w_4 = 0$, i.e., they lie in the two-dimensional w_1-w_5 coordinate plane in the five-dimensional hyperplane $V =$ const, as can be seen by Eq. (4.1). But since every orbit includes diagonal matrices, every orbit passes through this two-dimensional plane.

The w_1-w_5 plane is illustrated in Fig. 7. In this plane, Eqs. (4.22) and (4.25) simplify, and we have

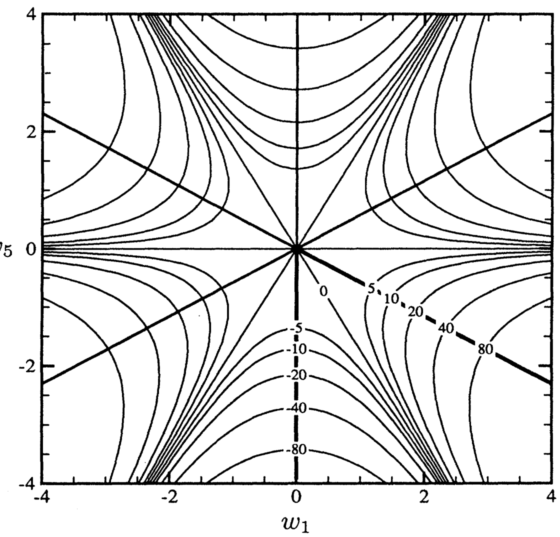

FIG. 7. The w_1-w_5 plane is a section through the orbits of the democracy group in the subspace $V =$ const. The matrix Q is diagonal on this plane. The heavy lines bound the "fundamental sector," within which the eigenvalues are in descending order. Contours of b are shown.

$$a = w_1^2 + w_5^2, \tag{4.33}$$

$$b = 2w_5(w_5^2 - 3w_1^2). \tag{4.34}$$

There is a mathematical similarity between the present analysis for the four-body problem and the Bohr-Mottelson treatment of quadrupolar deformations of a sphere [26]. In both cases one has an $\ell = 2$ representation of the group SO(3) (in Ref. [26] SO(3) is the rotation group; here it is the democracy group) and one uses the group theoretical methods indicated in Eqs. (4.27) and (4.28) to find two invariants.

From Eqs. (4.33) and (4.34) follows

$$4a^3 - b^2 = w_1^2(3w_5^2 - w_1^2)^2 \geq 0. \tag{4.35}$$

Thus we have another proof of the inequality (4.30), valid not only in the w_1-w_5 plane but everywhere in shape space since a and b are constant along the democracy orbits. One can easily show that b actually takes on all values between and including the limits shown in Eq. (4.30), as is also indicated by the contours of b in the w_1-w_5 plane, illustrated in Fig. 7. (The contours of a are just circles, and are not shown.)

Equation (4.35) also shows that the set $b = \pm 2a^{3/2}$, which is the degeneracy set of the moment of inertia tensor, appears in the w_1-w_5 plane as the three lines $w_1 = 0$, $w_1 = \pm\sqrt{3}w_5$ which divide the w_1-w_5 plane into six sectors. In fact, the rays at polar angles $\theta = -150°$, $-30°$, and $+90°$ correspond to $b = +2a^{3/2}$, while those at $\theta = -90°$, $+30°$, and $+150°$ correspond to $b = -2a^{3/2}$. The former set of rays on the hyperplane $V = 0$ is the set of collinear shapes (including the origin, which is the four-body collision).

The Q matrices on the w_1-w_5 plane are diagonal but do not necessarily satisfy the eigenvalue ordering $Q_{11} \geq Q_{22} \geq Q_{33}$. But we can solve Eqs. (4.7) on this plane, to obtain

$$Q_{11} - Q_{22} = \frac{2}{\sqrt{3}} w_1,$$
$$Q_{22} - Q_{33} = -\frac{1}{\sqrt{3}} w_1 - w_5,$$
$$Q_{33} - Q_{11} = -\frac{1}{\sqrt{3}} w_1 + w_5, \quad (4.36)$$

from which we conclude that the principal diagonal matrices lie in the sector $-90^\circ \leq \theta \leq -30^\circ$, which we call the "principal sector." This sector is indicated by heavy lines in Fig. 7. When a diagonal matrix in the interior of the principal sector is acted upon by the democracy group, the resulting orbit intersects the w_1-w_5 plane five more times, corresponding to the six proper orthogonal D matrices which permute the eigenvalues.

Therefore if we wish to represent all shapes of the four-body problem once and only once (except for a set of measure zero), we can range over the principal sector, say, by using coordinates a and b in the range $0 \leq a < \infty$, $-2a^{3/2} \leq b \leq +2a^{3/2}$, and by allowing the Euler angles of the democracy group to cover the space of cosets $SO(3)/D_2$. Alternatively, we can range over all of the w_1-w_5 plane, as is conveniently done in coordinates (w_1, w_5) themselves, and allow the Euler angles of the democracy group to cover the space of cosets $SO(3)/T$, where T is the 24-element point group consisting of the four elements of D_2 times the six proper permutations of the axes.

V. CONCLUSION

In conclusion, we will mention several applications and extensions of the results of this paper, some of which will appear in future publications. First there is the problem of coordinates on shape space for $n \geq 5$. For $n \geq 5$, it turns out that shape space can be represented as a union of a $(3n-12)$-dimensional family of copies of the (six-dimensional) four-body shape space, which gives a $(3n-6)$-dimensional manifold, as expected. Thus four-body shape space can be regarded as the building block out of which shape space for all higher values of n can be constructed. We will present a more detailed analysis of the case $n \geq 5$ in the future.

In another application, we have succeeded in transforming the four-body kinetic energy operator on shape space into its horizontal and vertical components under the action of the democracy group. That is, we have carried out the same transformation on shape space with the democracy group as has previously been carried out on configuration space under the action of the rotation group. Surprising results of this calculation are that the metric on the new base space (shape space divided by the democracy group) is Euclidean, and that the new curvature of the connection vanishes (as does the Riemann curvature).

We have also studied the Chern classes of the rotational SO(3) fiber bundle which appears in the four-body problem and found that they are all trivial. We will report on these and other applications in the future.

ACKNOWLEDGMENTS

We would like to thank Greg Ezra, Jerry Marsden, and Alden Mead for many useful conversations. This work was partially supported by the National Science Foundation through a grant for the Institute for Theoretical Atomic and Molecular Physics at Harvard University and Smithsonian Astrophysical Laboratory. This work was also supported by the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

[1] A. Guichardet, Ann. Inst. H. Poincaré **40**, 329 (1984).
[2] Akitomo Tachibana and Toshihiro Iwai, Phys. Rev. A **33**, 2262 (1986).
[3] Toshihiro Iwai, J. Math. Phys. **28**, 1315 (1987).
[4] *Geometric Phases in Physics*, edited by Alfred Shapere and Frank Wilczek (World Scientific, Singapore, 1989).
[5] Alfred Shapere and Frank Wilczek, Am. J. Phys. **57**, 514 (1989).
[6] Carl Eckart, Phys. Rev. **47**, 552 (1935).
[7] Sean P. Keating and C. Alden Mead, J. Chem. Phys. **82**, 5102 (1985).
[8] M. S. Narasimhan and T. R. Ramadas, Commun. Math. Phys. **67**, 121 (1979).
[9] G. Breit, Phys. Rev. **35**, 569 (1930).
[10] Felix T. Smith, Phys. Rev. **120**, 1058 (1960).
[11] A. K. Bhatia and A. Temkin, Rev. Mod. Phys. **36**, 1050 (1964).
[12] Werner Zickendraht, Ann. Phys. (N.Y.) **35**, 18 (1965).
[13] Michel De Celles and Byron T. Darling, J. Mol. Spectrosc. **29**, 66 (1969).
[14] C. Alden Mead and Donald G. Truhlar, J. Chem. Phys. **70**, 2284 (1979).
[15] B. R. Johnson, J. Chem. Phys. **73**, 5051 (1980).
[16] B. T. Sutcliffe, in *Quantum Dynamics of Molecules*, edited by R. G. Wooley (Plenum, New York, 1980), p. 1.
[17] Russell T. Pack and Gregory A. Parker, J. Chem. Phys. **87**, 3888 (1987).
[18] Paul G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987); *Shape in Chemistry* (VCH Publishers, New York, 1993).
[19] Felix T. Smith, J. Math. Phys. **3**, 735 (1962).
[20] Yu. F. Smirnov and K. V. Shitikova, Fiz. Elem. Chastits At. Yadra **8**, 847 (1977) [Sov. J. Part. Nucl. **8**, 344 (1977)].
[21] A. J. Dragt, J. Math. Phys. **6**, 533 (1965).
[22] Jean-Marc Lévy-Leblond and Monique Lévy-Nahas, J. Math. Phys. **6**, 1571 (1965).
[23] James D. Louck and Harold W. Galbraith, Rev. Mod.

Phys. **44**, 540 (1972).
[24] James K. G. Watson, Mol. Phys. **19**, 465 (1970).
[25] Morton Hamermesh, *Group Theory and its Applications to Physical Problems* (Dover, New York, 1962).
[26] Aage Bohr and Ben R. Mottelson, *Nuclear Structure, Vol. II, Nuclear Deformations* (W. A. Benjamin, Reading, MA, 1975).