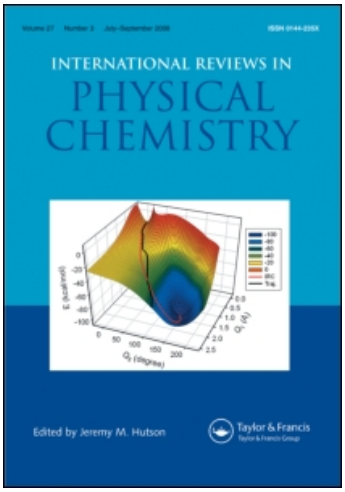

International Reviews in Physical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713724383

Permutationally invariant potential energy surfaces in high dimensionality

Bastiaan J. Braams[a]; Joel M. Bowman[a]
[a] Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, GA 30322, USA

To cite this Article Braams, Bastiaan J. and Bowman, Joel M.(2009) 'Permutationally invariant potential energy surfaces in high dimensionality', International Reviews in Physical Chemistry, 28: 4, 577 — 606
To link to this Article: DOI: 10.1080/01442350903234923
URL: http://dx.doi.org/10.1080/01442350903234923

International Reviews in Physical Chemistry
Vol. 28, No. 4, October–December 2009, 577–606

Taylor & Francis
Taylor & Francis Group

Permutationally invariant potential energy surfaces in high dimensionality

Bastiaan J. Braams[†] and Joel M. Bowman*

Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, GA 30322, USA

(*Received 16 June 2009; final version received 29 July 2009*)

We review recent progress in developing potential energy and dipole moment surfaces for polyatomic systems with up to 10 atoms. The emphasis is on global linear least squares fitting of tens of thousands of scattered *ab initio* energies using a special, compact fitting basis of permutationally invariant polynomials in Morse-type variables of all the internuclear distances. The computational mathematics underlying this approach is reviewed first, followed by a review of the practical approaches used to obtain the data for the fits. A straightforward symmetrization approach is also given, mainly for pedagogical purposes. The methods are illustrated for potential energy surfaces for CH_5^+, $(H_2O)_2$ and CH_3CHO. The relationship of this approach to other approaches is also briefly reviewed.

Keywords: potential energy surfaces; invariant fitting; CH_5^+; CH_3CHO; water dimer

Contents

*Corresponding author. Email: jmbowma@emory.edu
[†]Current Address: International Atomic Energy Agency, Division of Physical and Chemical Sciences, P.O. Box 100, Wagramerstrasse 5, A-1400 Vienna, Austria.

ISSN 0144–235X print/ISSN 1366–591X online

DOI: 10.1080/01442350903234923
http://www.informaworld.com

1. Introduction

The potential energy surface (PES) plays a central role in concepts and theories of molecular interactions spanning so-called non-bonding and bonding interactions that govern molecular structures and dynamics. The PES also plays a central role in the computation and prediction of structure and dynamics. In the Born–Oppenheimer (BO)approximation there exists a PES for each electronic state of the molecular system. Thus, as is well known, there are many PESs and their isolation from one another is an approximation, albeit often an excellent one.

It is important to note that there are two ways to think about a PES. One is as the eigenvalue of the electronic Schrödinger equation plus the nuclear repulsion energy, which in the BO approximation is a function of the nuclear coordinates. Thus this energy can be determined 'on the fly', that is directly at the nuclear coordinates of interest by solving the electronic Schrödinger equation at each value of these coordinates. The other way to think about the PES is as an analytical function of the nuclear coordinates. Certainly historically the latter way of thinking has dominated the development of the PES in chemical physics and computational chemistry. Ubiquitous examples include the harmonic, Morse and the Lennard-Jones potentials. These are two-body potentials and they can be used to describe most diatomic molecule interactions. The usefulness of having such functional forms for these simple cases is clear and not worth belabouring.

The extension of functional representations of the potential to larger molecules has by and large not been done. Of course if the many-body interactions were just a sum of pairwise ones then the extension would be straightforward and mathematically trivial. Representations of the many-body PES based on two-body interactions are in fact widespread. However, such representations do not describe most chemical reactions, which are inherently not two-body additive. Perhaps the $H + H_2$ reaction is the simplest example of this. This reaction has a barrier to exchange of the H atom and a simple pairwise interaction based on, say, the Morse potential for H_2 cannot describe this and instead would predict a bound H_3 molecule. A functional form that semi-quantitatively describes this barrier has been derived based on simple valence bond theory and is known by the acronym LEPS for London–Eyring–Polanyi–Sato. This functional form was used extensively to study triatomic reactions in the 1960–1980s. The function contains several parameters than can be optimized/selected for a specific reaction. If high-level *ab initio* information is available for, say, the barrier height these parameters can be adjusted to optimize agreement with this information. Otherwise the parameters were often adjusted empirically to optimize agreement with experiment.

As *ab initio* electronic energy calculations became more efficient and accurate it soon became clear that the LEPS functional form was not flexible enough to represent this high-quality data precisely. More mathematical forms to represent the potential were needed. For triatomic systems the Rotated Morse Potential was one such form, which however still

had an underlying model as its basis, as is clear from the name. This model was subsequently superseded for triatomic systems by strictly mathematical fitting approaches such as 3D splines, and related methods, nonlinear fitting based on physically based functional forms, many-body expansions, etc. Most applications of these numerical methods, until quite recently, were restricted to three-atom systems. Some of these techniques were applied to four-atom reactions but by mainly adding non-reactive degrees of freedom onto the three-atom methods. Excellent reviews of these methods along with applications up to 1989 are given by Schatz [1] and more recently by Ho and Rabitz [2].

The fitting techniques for triatomics do not generalize favourably to larger molecules. This is easily seen even for tetratomics, which have six internal degrees of freedom. A direct-product representation of the potential for, say, a 6D spline would require of the order of 10^6 electronic energies, compared to of the order of 10^3 electronic energies for triatomics. For N atoms this scaling is of the order of 10^{3N-6} and obviously is impractical beyond triatomics.

This exponential scaling has stimulated many groups to avoid the issue of fitting and use the 'direct-dynamics' approach, also known as '*ab initio* molecular dynamics' (AIMD) [3–11], where the citations are but a small sample of the work in this field. The attractiveness of the AIMD approach is, however, counterbalanced by its well-known limitations, i.e., the level of *ab initio* theory, the number of propagation time steps, etc. due to its large computational overhead. The focus of this article is the global, analytical representation of PESs in high dimensionality, i.e., for molecules with more than four atoms. However, as will be discussed in detail below, AIMD is used as one major means to generate the database of electronic energies for this representation.

Specifically, we review an approach we have developed to represent global PESs for up to 10 atoms of the order of 10^4 or 10^5 *ab initio* electronic energies. This approach has been applied to the following molecules and reactions: CH_5^+ [12,13], $(H_2O)_2$ [14,15], $(H_2O)_3$ [16], $C_2H_3^+$ [17], $H_3O_2^-$ [18], $H_5O_2^+$ [19], $CHOCH_2CHO$ (malonaldehyde) [20], C_2H_3 [21], $H + CH_4$ [22], $F + CH_4$ [23], FCH_4^- [24], $C + C_2H_2$ [25], $NO_2 + OH$ [26].

That one can obtain accurate global PESs for such large systems with only 10^4–10^5 energies is at first glance quite surprising given the high dimensionality of the space spanned. For six atoms the space is 12d and for nine atoms it is 21d. The key to the approach described in detail in the following sections is to explicitly make use of the fact that many molecules of interest contain identical atoms, e.g., H atoms. Since the PES must be invariant with respect to all permutations of identical atoms much of the configuration space is redundant and it is this fact that is exploited to reduce the size of the electronic energy database. Even before this reduction was realized, it was clear that permutational invariance can be essential in a variety of dynamics applications. Examples of this are the isomerization of acetylene to vinylidene and the unimolecular dissociation of formaldehyde. Our earliest approach to enforce permutational symmetry was done by simply replicating electronic energies for permutationally equivalent configurations for both C_2H_2 [27] and H_2CO [28] and then to fit the enlarged dataset. This approach did produce a PES that is numerically invariant with respect to permutations, however, the number of terms in the representation had to be large in order to represent the additional data. Clearly such an approach cannot be used for a molecule such as CH_5^+, where the order of the permutation group is $5! = 120$, and thus the size of the dataset to fit would be of the order of 10^6.

Clearly the ideal approach is to incorporate permutational symmetry directly into the PES representation by using a fitting basis that is permutationally invariant. The implementation of this approach is by no means trivial. This invariance property of the PES was clearly stated by Murrell *et al*. in their classic book published in 1984 [29]. They noted that an invariant fitting basis with this property can be represented in terms of what they termed an integrity basis. Explicit expressions for such a basis for triatomic molecules and for the special case of four identical atoms were given. The variables chosen for this basis were the internuclear distances; this is the obvious choice since they form a closed set under permutations. The techniques they used to generate this basis are in their words rather tedious even for the case of three identical atoms. To the best of our knowledge the use of this basis has not been developed beyond three and four identical atoms by this group [30].

This review describes the progress we have made developing PESs that are manifestly permutationally invariant. This is described in the next section. A brief presentation of related techniques to represent the dipole moment will also be given in that section. Following that, illustrations of the method are given for global PESs for CH_5^+, $(H_2O)_2$, and CH_3CHO. A summary and conclusions are given in the final section.

Before concluding this Introduction it is important to take note of other effective strategies to deal with high-dimensional potential energy surfaces. One is the modified Shepard approach, which has been advanced and developed by Collins and co-workers [31–35]. This approach represents the PES as a weighted sum of force fields, centred at numerous reference geometries. These force fields are Taylor series representations of the potential, typically truncated at fairly low order, using inverse internuclear distances. Other strategies include the *n*-mode representation of the potential [36–40], the related cut-high dimensional model representation [41,42] and the potfit method [43]. These approaches will be discussed in more detail at the end of the next section. It is also important to note that numerous groups are investigating a variety of strategies to represent PESs from *ab initio* electronic energies [44–56]. However, to date, these other approaches do not directly incorporate permutational symmetry into the fitting. The modified Shepard approach does enforce permutational invariance by replicating force fields, as discussed in more detail in the next section.

2. Methods

In this section we give the details of two approaches to explicitly incorporate permutational invariance into the global representation of potential energy surfaces. These approaches use basis functions that are invariant with respect to permutations of like atoms. In general any approach that incorporates this invariance property explicitly into the representation of potential must express the potential in terms of a set of variables that is closed under all permutations. Cartesian coordinates are an obvious but clearly non-optimum choice. Another choice, and the one we have made, is the set of all internuclear distances (in fact functions of these distances as described in detail below). The first approach, which we denote as monomial symmetrization, is given largely for pedagogical purposes. In this approach, which was briefly mentioned in [19], direct symmetrization of monomial basis functions is done. Details of this approach along with five examples are given in the next section. The relationship of this approach to other

methods to represent high-dimensional PESs will be briefly described. Following that in Sections 2.2 and 2.3 we give an account of results from invariant polynomial theory, where invariant bases are obtained in a very compact fashion in terms of *primary* and *secondary* polynomials. Some of the examples given for the monomial symmetrization approach are reconsidered using this more compact approach. That will be followed by a brief description of permutationally invariant fitting of the dipole moment.

2.1. *Monomial symmetrization*

To begin we motivate the invariant basis approach by recalling earlier work done on the acetylene/vinylidene PES [27]. To describe the isomerization to vinylidene it was necessary to account for the permutational symmetry of the HH and CC atoms. This was done by replicating electronic energies representing the potential given by the following expansion in multinomials:

$$V = \sum_{m=0}^{M} C_{abcdef} \left[y_{12}^{a} y_{13}^{b} y_{14}^{c} y_{23}^{d} y_{24}^{e} y_{34}^{f} \right]; \quad (m = a + b + c + d + e + f), \tag{1}$$

where y_{ij} is a Morse variable, which for the C_2H_2 PES was given by $y_{ij} = 1 - \exp(-r_{ij}/a)$. Currently we use $y_{ij} = \exp(-r_{ij}/a)$ and remarks about the choice of Morse variables instead of the r_{ij} will be given in the next section. The summation is over all powers of the y_{ij} subject to the constraint that the total degree m is at most M. (For simplicity we denote the sum over powers by the single index m.) In the above expression the H atoms are labelled 1, 2, and the C atoms are labelled 3 and 4. Internuclear distances (and thus the y_{ij}) are given using the normal lexical order $i < j$, and N is the number of atoms, which is four in this example. The convention in the monomial follows the obvious pattern. The linear coefficients were determined using a standard linear least squares method. The dataset in the fit consisted of electronic energies, three-quarters of which were obtained by replicating energies upon interchanging the two groups of identical atoms. The resulting dataset was still small enough that the least squares fitting effort was trivial. However, a large number of coefficients were numerically equal. This was a direct consequence of taking account of the permutational symmetry only by replicating the data.

Clearly as the number of variables in V increases the number of terms in the above representation grows nonlinearly for a fixed M and replicating data ceases to be a viable option. It is also clear from the above example that monomials with identical coefficients could be grouped into a single polynomial given by the sum of monomials multiplied by a single coefficient. Obviously this implies that symmetrization of the monomial basis should be done at the outset, thus eliminating the redundancy in coefficients and eliminating the need to replicate energies. It is quite straightforward to proceed with this symmetrization, which we illustrate below for five molecule types. We do this mainly for pedagogical purposes as a more efficient and mathematically elegant approach, based on invariant polynomial theory, is the one we have implemented. The details of that approach will be given after presentation of the straightforward symmetrization approach.

To proceed, we replace Equation (1) with the symbolic expression

$$V = \sum_{m=0}^{M} D_{abcdef}\, \mathcal{S}\Big[\, y_{12}^{a} y_{13}^{b} y_{14}^{c} y_{23}^{d} y_{24}^{e} y_{34}^{f} \,\Big], \tag{2}$$

where '$\mathcal{S}$' is the operator that symmetrizes monomials. To implement the symmetrization all that is needed is the mapping of atom permutations to permutations of the internuclear distances. This is straightforward to do as illustrated for A_3 molecules. Let 1, 2, 3 denote the initial arrangement of the three atoms. Consider the permutation 2, 3, 1. Initial internuclear distances r_{12}, r_{13}, r_{23} map onto r_{23}, r_{21}, r_{31}, which we rewrite in normal order as r_{23}, r_{12}, r_{13}. Thus, the monomial $r_{12}^{a} r_{13}^{b} r_{14}^{c}$ maps onto $r_{23}^{a} r_{12}^{b} r_{13}^{c}$ which can be rewritten as the original monomial with permuted powers as $r_{12}^{b} r_{13}^{c} r_{23}^{a}$. Adding to these two monomials the four others corresponding to the four additional permutations gives a fully symmetrized sum of monomials. Replacing r_{ij} with y_{ij} then yields the symmetrized basis function for the representation of V. Different from Equation (1), in Equation (2) it is understood that only a single term, and only a single unknown coefficient D_{abcdef}, appears for every set of tuples (a, b, c, d, e, f) that are permutationally equivalent.

The results of this explicit procedure are given for A_2B, A_3, A_2B_2, A_3B and A_3B_2 molecules in the tables below. For pedagogical purposes we use internuclear distances r_{ij} in these tables instead of y_{ij}, with the understanding that the latter variables are the ones used in the representation of V. The first row of each table lists the numerical labels of all atoms starting with the largest group of identical atoms, followed by the next group of identical atoms. The next column in this row gives the monomial corresponding to this original order of atoms and the third column gives the monomial in normal order. The subsequent rows in these tables gives the atom labels for all permutations of like atoms followed by the permutation of the original monomial (permuting the indicies), followed by the monomial in normal order, i.e, with the internuclear distances arranged as r_{ij} and with permuted powers.

First consider triatomic molecules A_2B and A_3. Table 1 gives these results for the simplest example A_2B which has a single permutation to consider. As seen, symmetrizing the single monomial gives two monomials and thus replacing the single monomial $y_{12}^{a} y_{13}^{b} y_{23}^{c}$ with $y_{12}^{a}(y_{13}^{b} y_{23}^{c} + y_{13}^{c} y_{23}^{b})$ leads to a invariant basis function.

The results for A_3 molecules are given in Table 2. As seen, the symmetrized monomial consists of $3! = 6$ terms and the symmetrized basis function consists of the sum of the six monomials (in the y_{ij}) in Table 2.

The results for A_2B_2 molecules, of which C_2H_2 is an example, are given in Table 3. In this case there are $2!2! = 4$ permutations and the symmetrized basis function consists of four terms.

Table 1. Symmetrized monomials for A_2B molecules.

Atom labels	Monomial	Normal order
1 2 3	$r_{12}^{a} r_{13}^{b} r_{23}^{c}$	$r_{12}^{a} r_{13}^{b} r_{23}^{c}$
2 1 3	$r_{12}^{a} r_{23}^{b} r_{13}^{c}$	$r_{12}^{a} r_{13}^{c} r_{23}^{b}$
Symmetrized term	$r_{12}^{a} r_{13}^{b} r_{23}^{c} + r_{12}^{a} r_{13}^{c} r_{23}^{b}$	

The next example is A_3B molecules with results shown in Table 4 and the final example we consider is A_3B_2 molecules with results shown in Table 5. Here there are 12 permtuations of 10 internuclear distances.

We have noted above that the use of symmetrized basis functions reduces the number of independent terms in the representation of V. That is there are fewer terms in

Table 2. Symmetrized monomials for A_3 molecules.

Atom labels	Monomial	Normal order
1 2 3	$r_{12}^a r_{13}^b r_{23}^c$	$r_{12}^a r_{13}^b r_{23}^c$
2 1 3	$r_{12}^a r_{23}^b r_{13}^c$	$r_{12}^a r_{13}^c r_{23}^b$
3 2 1	$r_{23}^a r_{13}^b r_{12}^c$	$r_{12}^c r_{13}^b r_{23}^a$
1 3 2	$r_{13}^a r_{12}^b r_{23}^c$	$r_{12}^b r_{13}^a r_{23}^c$
3 1 2	$r_{13}^a r_{23}^b r_{12}^c$	$r_{12}^c r_{13}^a r_{23}^b$
2 3 1	$r_{23}^a r_{12}^b r_{13}^c$	$r_{12}^b r_{13}^c r_{23}^a$
Symmetrized term	$r_{12}^a r_{13}^b r_{23}^c + r_{12}^a r_{13}^c r_{23}^b + \cdots$	$+r_{12}^c r_{13}^a r_{23}^b + r_{12}^b r_{13}^c r_{23}^a$

Table 3. Symmetrized monomials for A_2B_2 molecules.

Atom labels	Monomial	Normal order
1 2 3 4	$r_{12}^a r_{13}^b r_{14}^c r_{23}^d r_{24}^e r_{34}^f$	$r_{12}^a r_{13}^b r_{14}^c r_{23}^d r_{24}^e r_{34}^f$
2 1 3 4	$r_{12}^a r_{23}^b r_{24}^c r_{13}^d r_{14}^e r_{34}^f$	$r_{12}^a r_{13}^d r_{14}^e r_{23}^b r_{24}^c r_{34}^f$
1 2 4 3	$r_{12}^a r_{14}^b r_{13}^c r_{24}^d r_{23}^e r_{34}^f$	$r_{12}^a r_{13}^c r_{14}^b r_{23}^e r_{24}^d r_{34}^f$
2 1 4 3	$r_{12}^a r_{24}^b r_{23}^c r_{14}^d r_{13}^e r_{34}^f$	$r_{12}^a r_{13}^e r_{14}^d r_{23}^c r_{24}^b r_{34}^f$
Symmetrized term	$r_{12}^a r_{13}^b r_{14}^c r_{23}^d r_{24}^e r_{34}^f + \cdots$	$+r_{12}^a r_{13}^e r_{14}^d r_{23}^c r_{24}^b r_{34}^f$

Table 4. Symmetrized monomials for A_3B molecules.

Atom labels	Monomial	Normal order
1 2 3 4	$r_{12}^a r_{13}^b r_{14}^c r_{23}^d r_{24}^e r_{34}^f$	$r_{12}^a r_{13}^b r_{14}^c r_{23}^d r_{24}^e r_{34}^f$
2 1 3 4	$r_{12}^a r_{23}^b r_{24}^c r_{13}^d r_{14}^e r_{34}^f$	$r_{12}^a r_{13}^d r_{14}^e r_{23}^b r_{24}^c r_{34}^f$
3 2 1 4	$r_{23}^a r_{13}^b r_{34}^c r_{12}^d r_{24}^e r_{14}^f$	$r_{12}^d r_{13}^b r_{14}^f r_{23}^a r_{24}^e r_{34}^c$
1 3 2 4	$r_{13}^a r_{12}^b r_{14}^c r_{23}^d r_{34}^e r_{24}^f$	$r_{12}^b r_{13}^a r_{14}^c r_{23}^d r_{24}^f r_{34}^e$
3 1 2 4	$r_{13}^a r_{23}^b r_{34}^c r_{12}^d r_{14}^e r_{24}^f$	$r_{12}^d r_{13}^a r_{14}^e r_{23}^b r_{24}^f r_{34}^c$
2 3 1 4	$r_{12}^a r_{13}^b r_{14}^c r_{23}^d r_{24}^e r_{34}^f$	$r_{12}^b r_{13}^d r_{14}^f r_{23}^a r_{24}^c r_{34}^e$
Symmetrized term	$r_{12}^a r_{13}^b r_{14}^c r_{23}^d r_{24}^e r_{34}^f + \cdots$	$+r_{12}^b r_{13}^d r_{14}^f r_{23}^a r_{24}^c r_{34}^e$

Table 5. Symmetrized monomials for A_3B_2 molecules.

Atom labels	Monomial	Normal order
1 2 3 4 5	$r_{12}^a r_{13}^b r_{14}^c r_{15}^d r_{23}^e r_{24}^f r_{25}^g r_{34}^h r_{35}^i r_{45}^j$	$r_{12}^a r_{13}^b r_{14}^c r_{15}^d r_{23}^e r_{24}^f r_{25}^g r_{34}^h r_{35}^i r_{45}^j$
2 1 3 4 5	$r_{12}^a r_{23}^b r_{24}^c r_{25}^d r_{13}^e r_{14}^f r_{15}^g r_{34}^h r_{35}^i r_{45}^j$	$r_{12}^a r_{13}^e r_{14}^f r_{15}^g r_{23}^b r_{24}^c r_{25}^d r_{34}^h r_{35}^i r_{45}^j$
3 2 1 4 5	$r_{23}^a r_{13}^b r_{34}^c r_{35}^d r_{12}^e r_{24}^f r_{25}^g r_{14}^h r_{15}^i r_{45}^j$	$r_{12}^e r_{13}^b r_{14}^h r_{15}^i r_{23}^a r_{24}^f r_{25}^g r_{34}^c r_{35}^d r_{45}^j$
1 3 2 4 5	$r_{13}^a r_{12}^b r_{14}^c r_{15}^d r_{23}^e r_{34}^f r_{35}^g r_{24}^h r_{25}^i r_{45}^j$	$r_{12}^b r_{13}^a r_{14}^c r_{15}^d r_{23}^e r_{24}^h r_{25}^i r_{34}^f r_{35}^g r_{45}^j$
3 1 2 4 5	$r_{13}^a r_{23}^b r_{34}^c r_{35}^d r_{12}^e r_{14}^f r_{15}^g r_{24}^h r_{25}^i r_{45}^j$	$r_{12}^e r_{13}^a r_{14}^f r_{15}^g r_{23}^b r_{24}^h r_{25}^i r_{34}^c r_{35}^d r_{45}^j$
2 3 1 4 5	$r_{23}^a r_{12}^b r_{24}^c r_{25}^d r_{13}^e r_{34}^f r_{35}^g r_{14}^h r_{15}^i r_{45}^j$	$r_{12}^b r_{13}^e r_{14}^h r_{15}^i r_{23}^a r_{24}^c r_{25}^d r_{34}^f r_{35}^g r_{45}^j$
1 2 3 5 4	$r_{12}^a r_{13}^b r_{15}^c r_{14}^d r_{23}^e r_{25}^f r_{24}^g r_{35}^h r_{34}^i r_{45}^j$	$r_{12}^a r_{13}^b r_{14}^d r_{15}^c r_{23}^e r_{24}^g r_{25}^f r_{34}^i r_{35}^h r_{45}^j$
2 1 3 5 4	$r_{12}^a r_{23}^b r_{25}^c r_{24}^d r_{13}^e r_{15}^f r_{14}^g r_{35}^h r_{34}^i r_{45}^j$	$r_{12}^a r_{13}^e r_{14}^g r_{15}^f r_{23}^b r_{24}^d r_{25}^c r_{34}^i r_{35}^h r_{45}^j$
3 2 1 5 4	$r_{23}^a r_{13}^b r_{35}^c r_{34}^d r_{12}^e r_{25}^f r_{24}^g r_{15}^h r_{14}^i r_{45}^j$	$r_{12}^e r_{13}^b r_{14}^i r_{15}^h r_{23}^a r_{24}^g r_{25}^f r_{34}^d r_{35}^c r_{45}^j$
1 3 2 5 4	$r_{13}^a r_{12}^b r_{15}^c r_{14}^d r_{23}^e r_{35}^f r_{34}^g r_{25}^h r_{24}^i r_{45}^j$	$r_{12}^b r_{13}^a r_{14}^d r_{15}^c r_{23}^e r_{24}^i r_{25}^h r_{34}^g r_{35}^f r_{45}^j$
3 1 2 5 4	$r_{13}^a r_{23}^b r_{35}^c r_{34}^d r_{12}^e r_{15}^f r_{14}^g r_{25}^h r_{24}^i r_{45}^j$	$r_{12}^e r_{13}^a r_{14}^g r_{15}^f r_{23}^b r_{24}^i r_{25}^h r_{34}^d r_{35}^c r_{45}^j$
2 3 1 5 4	$r_{23}^a r_{12}^b r_{25}^c r_{24}^d r_{13}^e r_{35}^f r_{34}^g r_{15}^h r_{35}^i r_{45}^j$	$r_{12}^b r_{13}^e r_{14}^i r_{15}^h r_{23}^a r_{24}^d r_{25}^c r_{34}^g r_{35}^f r_{45}^j$
Symmetrized term	$r_{12}^a r_{13}^b r_{14}^c r_{15}^d r_{23}^e r_{24}^f r_{25}^g r_{34}^h r_{35}^i r_{45}^j + \cdots$	$+r_{12}^b r_{13}^e r_{14}^i r_{15}^h r_{23}^a r_{24}^d r_{25}^c r_{34}^g r_{35}^f r_{45}^j$

Table 6. Number of terms for indicated molecules versus total order indicated.

Molecule	5	6	7	8
A_3	16	23	31	41
A_2B	34	50	70	95
ABC	56	84	120	165
A_4	40	72	120	195
A_3B	103	196	348	590
A_2B_2	153	291	519	882
A_2BC	256	502	918	1589
ABCD	462	924	1716	3003
A_5	64	140	289	580
A_4B	208	495	1101	2327
A_3B_2	364	889	2022	4343
A_3BC	636	1603	3737	8163
A_2B_2C	904	2304	5416	11910
A_2BCD	1632	4264	10208	22734
ABCDE	3003	8008	19448	43758

Equation (1) than in Equation (2). The reduction is bounded above by the order of the symmetric group. The number of terms in each representation can be obtained explicitly using the Molien series, which is given below. For now we use the results of that calculation to present in Table 6 the number of terms for a series of three, four, and five atom molecules and for total polynomial orders 5–8. Note the number of terms for a PES representation with no symmetrization, Equation (1), is given by the number of terms for

molecule types with no identical atoms. As seen, there is a large reduction in the number of terms using the symmetrized representation of V and also note that as the total degree increases the reduction factor approaches the order of the relevant symmetric group.

Before proceeding to the next subsection we make some observations about the representation of the potential given by the unsymmetrized restricted sum of monomials, Equation (1), and n-mode [36–40] and cut-high dimensional model [42,57,58] representations of the potential. Both of these representations decompose the full-dimensional potential, in, say f degrees of freedom, into a sum of lower dimensional terms. In the n-mode representation specifically, the f-mode potential in normal coordinates, denoted collectively by $\mathbf{Q}$, is given by

$$V(\mathbf{Q}) = V^{(0)} + \sum_i V_i^{(1)}(Q_i) + \sum_{ij} V_{ij}^{(2)}(Q_i, Q_j) + \sum_{ijk} V_{ijk}^{(3)}(Q_i, Q_j, Q_k) + \cdots, \tag{3}$$

where $V^{(0)}$ is a constant, the one-mode terms, $V_i^{(1)}(Q_i)$, are the cuts through the hyperspace of normal coordinates with just one coordinate varying at a time, the two-mode terms, $V_{ij}^{(2)}(Q_i, Q_j)$, are potentials in all pairwise variations of the coordinates, etc. Each of the terms in each summation is an *intrinsic p*-mode potential, because they each vanish if one of the variables is zero. The efficiency of this representation is realized if it can be accurately truncated at much less than full f-mode terms. (This is explicitly made use of in the vibrational analysis code Multimode [36–38] and similar codes [39,40].)

Now consider the monomial representation of the potential, Equation (1), and note that formally that expression can be rewritten as

$$V = V^{(0)} + V^{(1)} + V^{(2)} + V^{(3)} + V^{(4)} + \cdots, \tag{4}$$

where $V^{(0)}$ is a constant and

$$\begin{aligned} V^{(1)} &= \sum_i V^{(1)}(y_i), \\ V^{(1)}(y_i) &= \sum_n C_n^{(i)} y_i^n, \\ V^{(2)} &= \sum_{i,j} V^{(2)}(y_i, y_j), \\ V^{(2)}(y_i, y_j) &= \sum_{m,n} C_{m,n}^{(i,j)} y_i^n y_j^m, \end{aligned} \tag{5}$$

and so forth, in principle up to a full f-mode term (where f equals the number of internuclear distances). However, if the maximum sum of powers in Equation (1), M, is less than the number of variables in V, which it typically is for molecules of more than four atoms, then an n-mode truncation is realized. Clearly each of these terms can be symmetrized and the resulting representation is then both truncated and permutationally invariant.

As one example of this effective mode reduction, consider the potential for the Zundel cation, $H_5O_2^+$, where the number of Morse variables is 21. The accurate PES for this cation was based on fitting tens of thousands of CCSD(T) electronic energies using full permutational symmetry, containing terms that at most involve seven variables [19].

Thus, from the perspective of the n-mode representation, the $H_5O_2^+$ potential is truncated at the seven-mode term. It is of interest to note that this PES has been used in quantum calculations of the IR spectrum of $H_5O_2^+$ [59,60], as well as several classical MD calculations of this spectrum [61,62]. In the impressive Multiconfiguration Time Dependent Hartree (MCTDH) quantum calculations by Meyer and co-workers [59] the n-mode representation of the PES was an essential component of the calculation to make it feasible. In this case the so-called potfit representation of the potential i.e., a product of one-mode potentials, developed by Meyer and co-workers specifically for use in MCTDH calculations [43], was not able to be used effectively for the $H_5O_2^+$ application.

Note the n-mode, cut-high dimensional model and potfit representations are not (at least currently) permutationally invariant. Post symmetrization of these representations is possible, perhaps by following the approach taken in the modified Shepard approach. This is done by replacing the non-symmetrized representation

$$V(Z) = \sum_{i=1}^{N} w_i(Z) T_i(Z)$$

with one replicating the force fields in variables Z (inverse internuclear distances), $T_i(Z)$, by [63]

$$V(Z) = \sum_{g \in G} \sum_{i=1}^{N} w_i(Z; g) T_i(Z; g),$$

where the outer sum is over the elements of the relevant symmetric group. Thus, in the application to CH_5^+ [64], N local force fields were evidently replicated 120 times.

Next we present the essentials of the theory of polynomial invariants which leads to more efficient representation of permutationally invariant bases. These are the bases used to obtain the numerous PESs mentioned in the Introduction.

2.2. *Theory of polynomial invariants*

We begin with elements of invariant polynomial theory of relevance to a compact representation of potential energy surfaces that are invariant under the action of a finite group. The mathematical background is classical, and can found in the monograph by Derksen and Kemper [65], which emphasizes the constructive computer-algebra aspects of the theory.

Given an n-dimensional vector space V over the field K (the real or the complex numbers), a polynomial on V means a polynomial of n real or complex variables, say $p(x)$, $x = (x_1, \ldots, x_n)$. The set of all polynomials on V is denoted $K[V]$. The polynomials form themselves a vector space over K: if p and q are polynomials and α and β are scalars (elements of K) then $\alpha p + \beta q$ is a polynomial, and the addition of polynomials and multiplication by scalars satisfy the familiar rules of arithmetic. The polynomials also form a (commutative) *ring*: if p and q are polynomials then so are $p + q$ and pq, the commutative, associative, and distributive laws are satisfied (commutative also for multiplication), there is a zero element ($p(x) = 0$) and a unit element ($p(x) = 1$), and there is an inverse for addition ($p + (-p) = 0$), but not in general for multiplication (only the non-zero constant polynomials have a multiplicative inverse). Finally, combining the vector

space and the ring properties, $K[V]$ is said to form a (commutative) *algebra*; i.e., a vector space in which elements may be multiplied and again familiar rules of arithmetic are satisfied. As a vector space $K[V]$ is infinite-dimensional, but it may be written as a direct sum of finite-dimensional vector spaces:

$$K[V] = \sum_{k=0}^{\infty} K[V]_k \tag{6}$$

in which $K[V]_k$ represents the vector space of homogeneous polynomials of precise degree k together with the zero polynomial. The degrees of homogeneous polynomials satisfy the rule $\deg(pq) = \deg(p)\deg(q)$. As a vector space or algebra $K[V]$ is said to be *graded by total degree*.

We next consider the case where V carries a representation of a group G. Thus, an element $g \in G$ defines an invertible linear mapping $g : V \to V$ and we have the group composition law: for $g_1, g_2 \in G$ and $x \in V$, $g_2(g_1 x) = (g_2 g_1)x$. Associated with the action of G on V is an action of G on $K[V]$; if $g \in G$ and $p \in K[V]$ then $(gp)(x) = p(g^{-1}x)$. (It has to be g^{-1} on the right-hand side in order to make the group composition law come out right.) A polynomial p is invariant under G if $gp = p$ for all $g \in G$. The set of invariant polynomials is denoted by $K[V]^G$ and, like $K[V]$, it forms a vector space and a ring and an algebra: invariant polynomials may be added, multiplied by scalars, and multiplied among themselves and the result is again an invariant polynomial, and all the rules of arithmetic are inherited from $K[V]$.

If $p_1, \ldots, p_k$ are polynomials then we can form further polynomials by taking sums of products of powers of the p_i and the set of all polynomials formed that way may be denoted $K[p_1, \ldots, p_k]$. It again forms a ring and an algebra and is called the ring or algebra *generated* by the family $\{p_1, \ldots, p_k\}$, which are called the *generators*. For example, if we let $k = \dim(V)$ and $p_i(x) = x_i$ for $i \in \{1, \ldots, k\}$ then $K[p_1, \ldots, p_k] = K[V]$. Thus, $K[V]$ is generated by just $\dim(V)$ basic polynomials, although as a vector space it is infinite-dimensional. It is of interest, and not obvious, that this finite generation property is also true of rings of invariant polynomials. Limiting ourselves to the case of interest to us, if V is a finite-dimensional real or complex vector space that carries a representation of a finite group G then $K[V]^G$ is finitely generated. (We will not prove this statement.)

The finite generation property says nothing about uniqueness of the representation in terms of generators. For example, consider the real plane, $V = R^2$, and let G be the two-element group that is generated by reflection into the origin; so the non-identity element of G maps (x, y) to $(-x, -y)$. A general polynomial on V may be written as

$$p(x, y) = \sum_{i,j=0}^{\infty} c(i, j) x^i y^j \tag{7}$$

and in order for p to be invariant under G it must be that $c(i, j) \neq 0$ only if $i + j$ is even: all non-zero terms must be of even total degree. In this case, three polynomials suffice to generate the invariant ring: $K[V]^G = K[x^2, y^2, xy]$, and also the reader can convince him or herself that none of the generators is redundant; for example, xy is invariant but it is not generated by the set of polynomials $\{x^2, y^2\}$. On the other hand, the representation of a higher-degree invariant polynomial by these generators is not normally unique; for

example, the polynomial x^2y^2 is represented as the product of generators in the form $(x^2)(y^2)$ and also in the form $(xy)^2$.

For the case of interest to us, a finite group acting on a finite-dimensional real or complex vector space, the finite generation property may be strengthened in a manner that makes the representation unique.

Theorem: *Let V be a finite-dimensional vector space over the real or complex field K and carrying a representation of the finite group G. Then there exists a family of $n = \dim(V)$* primary generators $\{p_i \in K[V]^G : 1 \le i \le n\}$ *and a finite family of* secondary generators $\{q_\alpha \in K[V]^G : 1 \le \alpha \le M\}$ *such that every $f \in K[V]^G$ has a unique representation in the form*

$$f(x) = \sum_{\alpha=1}^{M} h_\alpha(p_1(x), \ldots, p_n(x)) q_\alpha(x) \tag{8}$$

where h_α is an arbitrary polynomial.

We refer to [65], Chapter 3 for a proof, and for a statement of the theorem in a more general context. This representation expresses that $K[V]^G$ has the *Cohen–Macaulay property*. The primary invariants generate a polynomial ring, $K[p_1, \ldots, p_n]$. A polynomial $h_\alpha(p_1, \ldots, p_n)$ is just an element of that ring, and so it is seen that the secondary invariants q_α form a basis for a representation of $K[V]^G$ as a *free module* with coefficients h_α in the ring $K[p_1, \ldots, p_n]$. (A free module is essentially a vector space, but with coefficients in a ring instead of in a field.)

Returning to the example of $V = R^2$ and the two-element group G generated by reflections, we recognized that polynomial p is invariant if every non-zero monomial term $c(i,j)x^iy^j$ is of even total degree. But if $i+j$ is even then either i and j are both even or they are both odd. The terms in which i and j are both even may be collected into a polynomial of x^2 and y^2 and those with i and j both odd may be collected into xy times such a polynomial. We see therefore that $K[V]^G$ is generated in the Cohen–Macaulay form by primary generators $p_1(x,y) = x^2$ and $p_2(x,y) = y^2$ and secondary generators $q_1(x,y) = 1$ and $q_2(x,y) = xy$.

2.3. *Application to potential energy surfaces*

Here we give the essentials of the formal theory of invariant polynomials as applied to the representation of potential energy surfaces. The independent variables are functions of the internuclear distances, Morse variables in our work, denoted y_{ij} or collectively $\mathbf{y}$, a vector that has $N(N-1)/2$ components where N is the number of atoms. The potential is expressed as

$$V(\mathbf{y}) = \sum_{\alpha=1}^{M} h_\alpha(\mathbf{p}(\mathbf{y})) q_\alpha(\mathbf{y}), \tag{9}$$

where $\mathbf{p}(\mathbf{y})$ is the vector formed by the $N(N-1)/2$ *primary invariant polynomials* and $q_\alpha(\mathbf{y})$ (for $1 \le \alpha \le M$) are the *secondary invariant polynomials*. The primary and secondary invariant polynomials are invariant under all permutations of identical atoms, so under the direct product of symmetric groups, one symmetric group for each kind of atoms.

The number of primary invariant polynomials is equal to the number of variables, i.e., the number of internuclear distances, $N(N-1)/2$ for an N-atom system. A finite number of secondary invariant polynomials suffices to express all possible invariant polynomials, to arbitrarily high degree, through Equation (9), although the number of secondary invariant polynomials is not easily expressed in terms of the number of variables or the order of the group. Finding these polynomials for direct product groups $S_n \times S_m \times \cdots \times S_p$ (for $A_nB_m \ldots X_p$ molecules) acting on the space of internuclear distances is non-trivial and computational algebra software MAGMA [66] is therefore called on to perform the task. Below we give examples of these polynomials for molecular systems of up to five atoms. We make use of the Molien series which tells us the number of terms in the above summation at each order. This series is the power series defined by

$$M(t) = \left(\sum_{\alpha=1}^{M} t^{e_\alpha} \right) \Pi_{i=1}^{n} (1 - t^{d_i})^{-1} , \tag{10}$$

where p_i has degree d_i and q_α has degree e_α.

Next we consider a few examples where explicit expressions for primary and secondary invariant polynomials are given.

2.3.1. *Two-atom and three-atom molecules*

For the case of two atoms, equal or distinct, there is just one internuclear distance, r_{12}. The natural choice for the single primary invariant is the polynomial $p_1(r) = r_{12}$, and then for the single secondary invariant the constant polynomial $q_1(r) = 1$. The Molien series is generated by $M(t) = 1/(1-t)$. Once again we stress that even in this simple example the primary and secondary invariants are not uniquely defined. The polynomial $p(r) = r_{12}^2$ could have been chosen as the primary invariant, with, for example, $q_1(r) = 1$ and $q_2(r) = r_{12}$ as secondary invariants.

For the case of three atoms there are three internuclear distances and therefore three primary invariants, whatever the symmetry. For three unlike atoms there is no symmetry operation other than the identity. As primary invariants we choose the three functions $p_1(r) = r_{12}$, $p_2(r) = r_{13}$ and $p_3(r) = r_{23}$, and the single secondary invariant is $q_1(r) = 1$. The Molien series is generated by $M(t) = (1-t)^{-3}$.

The first non-trivial case is A_2B. According to our procedures there is a single primary invariant associated with the AA system, polynomial $p_1(r) = r_{12}$ and two primary invariants associated with the set of AB distances: $p_2(r) = (r_{13} + r_{23})/2$ and $p_3(r) = (r_{13}^2 + r_{23}^2)/2$. The single secondary invariant is $q_1(r) = 1$. The Molien series is generated by $M(t) = (1-t)^{-2}(1-t^2)^{-1}$.

For A_3 there are three primary invariants, $p_k(r) = (r_{12}^k + r_{13}^k + r_{23}^k)/3$ for $k = 1$, $k = 2$, and $k = 3$. The single secondary invariant is $q_1(r) = 1$. The Molien series is generated by $M(t) = (1-t)^{-1}(1-t^2)^{-1}(1-t^3)^{-1}$.

As an aside we mention that for these small cases it is not too difficult to provide the transformations between the vector space basis of invariant polynomials that is generated by explicit symmetrization of a monomial basis and the vector space basis that is generated by the primary and secondary invariants. However, we shall not carry out that exercise here.

2.3.2. *A_2B_2 molecules*

For the A_2B_2 system label the atoms A(1), A(2), B(3), B(4). There are six internuclear distances, thus six primary invariants. One is a primary invariant that involves only the A_2 system; this subsystem has only a single variable associated with it, r_{12}, and that variable is a primary invariant. Likewise there is one primary invariant associated with the B_2 system, r_{34}. Finally the four AB distances permute among themselves; thus there are four primary invariants that involve only these distances. Of these four, there is one of degree 1 and three of degree 2. (More precisely, the primary invariants can be chosen so that one is of degree 1 and three are of degree 2, and such a choice minimizes the product of the degrees.)

A natural choice for the degree-1 primary invariant for the AB subsystem is the average AB distance. For the three degree-2 primary invariants there are several reasonable choices. In the codes we use auxiliary variables (they are covariants) $e_0 = (r_{13} + r_{14})/2$, $e_1 = (r_{23} + r_{24})/2$, $f_0 = (r_{13} + r_{23})/2$, and $f_1 = (r_{14} + r_{24})/2$, and in terms of these the three degree-2 primary invariants are $(e_0^2 + e_1^2)/2$, $(f_0^2 + f_1^2)/2$, and $(r_{13}^2 + r_{23}^2 + r_{14}^2 + r_{24}^2)/4$.

Given that the primary invariants for the A_2B_2 system are chosen to be three of degree 1 and three of degree 2 then there are two secondary invariants, one of degree 0 and one of degree 3. The secondary invariant of degree 0 is the polynomial that has the constant value 1. In the codes, the secondary invariant of degree 3 is $(r_{13}^3 + r_{23}^3 + r_{14}^3 + r_{24}^3)/4$. Other choices for the degree-3 secondary invariant are possible too.

2.3.3. *A_3B_2 molecules*

For A_3B_2 the atoms are labelled as A(1), A(2), A(3), B(4), B(5), there are 10 internuclear distances, three AA distances, six AB distances, and a single BB distance. Again choosing the degrees of the primaries in such a way that the product of degrees is minimized, there should be three primaries for the AA system and they have degrees 1, 2, and 3; six primaries for the AB system and they have degrees 1, 2, 2, 2, 3, and 6; and a single primary for the BB system of degree 1.

For the three primary invariants for the A_3 subsystem we chose polynomials $(r_{12}^k + r_{13}^k + r_{23}^k)/3$ for $k \in \{1, 2, 3\}$. The primary invariant for the BB subsystem is the single BB distance, r_{45}. For the primary invariants associated with AB distances there are again several reasonable choices, and the choice that was made for the codes is obtained as follows. Let $e_0 = (r_{14} + r_{15})/2$, $e_1 = (r_{24} + r_{25})/2$, $e_2 = (r_{34} + r_{35})/2$, $f_0 = (r_{14} + r_{24} + r_{34})/3$, and $f_1 = (r_{15} + r_{25} + r_{35})/3$; then the six primary invariants for the AB system are:

$$\begin{aligned} p_1 &= (r_{14} + r_{24} + r_{34} + r_{15} + r_{25} + r_{35})/6 \\ p_2 &= (e_0^2 + e_1^2 + e_2^2)/3 \\ p_3 &= (f_0^2 + f_1^2)/2 \\ p_4 &= (r_{14}^2 + r_{24}^2 + r_{34}^2 + r_{15}^2 + r_{25}^2 + r_{35}^2)/6 \\ p_5 &= (e_0^3 + e_1^3 + e_2^3)/3 \\ p_6 &= (r_{14}^6 + r_{24}^6 + r_{34}^6 + r_{15}^6 + r_{25}^6 + r_{35}^6)/6. \end{aligned} \tag{11}$$

Given that the primary invariants are three of degree 1, four of degree 2, two of degree 3, and one of degree 6, then the number of secondary invariants at each degree starting at degree 0 is given by this sequence: [1, 0, 1, 6, 8, 6, 12, 14, 9, 8, 5, 2, 0, . . .].

(There are no secondary invariants of degree 12 or higher.) The secondary invariant at degree 0 is the constant 1. In order to describe the higher-degree secondaries it is useful to introduce some conventions. Indices i_0, i_1, i_2, will run over the A-nuclei and indices j_0, j_1 will run over the B-nuclei. Whenever two or more i-indices appear in a product then they are all distinct; likewise for the j indices. Expressions such as $r(i_0, i_1)$ or $r(i_0, j_0)$ represent a corresponding internuclear distance (or Morse variable). The $\mathcal{S}$ in what follows indicates that the subsequent expression must be averaged over all choices of the i and j indices (it is the same symmetrization operator defined in Section 2.1). With those conventions the single secondary invariant of degree 2 is $\mathcal{S}[r(i_0, i_1)r(i_0, j_0)]$. The six secondary invariants of degree 3 are:

$$\begin{aligned}&\mathcal{S}[r(i_0,i_1)^2 r(i_0,j_0)],\ \mathcal{S}[r(i_0,i_1)r(i_0,j_0)^2],\ \mathcal{S}[r(i_0,j_0)^3],\ \mathcal{S}[r(i_0,i_1)r(i_0,j_0)r(i_1,j_0)],\\ &\mathcal{S}[r(i_0,j_0)^2 r(i_1,j_0)],\ \mathcal{S}[r(i_0,i_1)r(i_0,j_0)r(i_0,j_1)].\end{aligned} \tag{12}$$

Of the eight secondary invariants of degree 4, one is obtained by taking the square of the single secondary invariant of degree 2. The other seven are expressed thus:

$$\begin{aligned}&\mathcal{S}[r(i_0,i_1)^2 r(i_0,j_0)^2],\ \mathcal{S}[r(i_0,i_1)r(i_0,j_0)^3],\ \mathcal{S}[r(i_0,j_0)^4],\ \mathcal{S}[r(i_0,i_1)^2 r(i_0,j_0)r(i_1,j_0)],\\ &\mathcal{S}[r(i_0,i_1)r(i_0,j_0)^2 r(i_1,j_0)],\ \mathcal{S}[r(i_0,j_0)^3 r(i_1,j_0)],\ \mathcal{S}[r(i_0,j_0)^3 r(i_0,j_1)].\end{aligned} \tag{13}$$

We do not list the higher degree secondary invariants here.

2.4. *Dipole moment*

The dipole moment is a vector and thus its representation in terms of invariant polynomials must reflect this. A straightforward way to address this is to represent the dipole moment $\vec{\mu}$ as a vector sum of effective charges:

$$\vec{\mu} = \sum_i w_i(X)\vec{x}(i), \tag{14}$$

where X denotes the molecular configuration, $w_i(X)$ is the effective charge on the i-th nucleus (this effective charge depends on the entire configuration) and $\vec{x}(i)$ is the position vector of the i-th nucleus. The effective charges are scalar quantities, which *can* be expressed in terms of internuclear distances alone. They are expanded in a set of basis functions, $w_i(X) = \sum_k c_k b_k(i, X)$, where b_k is the k-th basis function, depending on the nuclear index i as well as on the configuration X, and the coefficients c_k are obtained by solving a weighted least squares system with the *ab initio* dipole moment data on the right-hand side. Similar to the expansion of the potential energy the basis functions are polynomials in Morse variables.

It appears that the expansion of the dipole moment is essentially the same as that of the potential, but there is a twist when permutations of identical nuclei are considered. Like the potential energy the dipole moment is invariant under permutation of identical nuclei, but in order to make it invariant the functions $w_i(X)$ in Equation (14) must transform like a *covariant*: if configuration X is transformed into X' by interchange of identical nuclei i and j then pair $(w_i(X), w_j(X))$ is equal to the pair $(w_j(X'), w_i(X'))$. Ideally we would have a compact library of appropriate covariant basis functions for each molecular symmetry group, but to date we have used a somewhat indirect and not quite optimal approach.

Consider an A_mB_n molecule and let the index i belong to an A-atom. Then the function w_i, which is a component of a covariant for the group $S_m \times S_n$, is also an *invariant* for the smaller group $S_{m-1} \times S_n \times S_1$, so for the $A_{m-1}B_nC_1$ molecule. Relying on that, we can directly use our comprehensive invariants library to represent these covariants as well and fit a dipole moment. A drawback to this approach enters when the behaviour of the dipole moment under translations is considered. If Z the total charge of the molecule and if $X + \vec{y}$ is the result of applying a uniform displacement $\vec{y}$ to all atoms in the configuration X then it must be that $\vec{\mu}(X + \vec{y}) = \vec{\mu}(X) + Z\vec{y}$. The fitted dipole moment will satisfy this condition if $\sum_i w_i(X) = Z$, but in our present work this condition is not built into the basis, rather it is imposed as an additional constraint in the least squares system, and is therefore not satisfied exactly.

2.5. *Library of invariant polynomial basis functions*

A Fortran-95 library has been written that contains code to compute primary and secondary invariants and a polynomial basis, up to some maximum degree, for almost all molecular permutational symmetry groups of at most 10 atoms total. (The exceptions at this time are the A_6B_4 system and systems containing eight or more identical atoms; these exceptions will be filled in.) The code for the computation of the secondary invariants is largely computer-generated. Magma computer-algebra procedures were used to print out representations of the invariant polynomials that could be translated into Fortran (or C, if we would have chosen that) in a quite direct manner. The principal kind of routine from the user point of view is a subroutine with a name such as mg421-base; the digits indicate the permutation symmetry group, here that of the $A_4B_2C_1$ molecule, and the subroutine takes as argument a vector of internuclear distances and an integer that is the maximum polynomial degree; it returns a vector of which each component is the value of an invariant basis function evaluated for the given vector of internuclear distances or Morse variables. When constructing the fitted potential energy surface that output vector becomes a row in the matrix that defines the least squares system, and when evaluating the potential energy that output vector is multiplied in an inner product with the coefficient vector that solved the least squares system. Additional procedures in the library make it easy to apply the invariant polynomials in the context of a many-body expansion. Further procedures are provided to set up and solve the least squares system. The data can be function values, function values and gradients, or function values, gradients and Hessians.

The compiled library is available for download at the iOpenShell website [http://iopenshell.usc.edu/downloads/ezpes/] in connection with potential energy surfaces for specific molecules.

3. Applications

In this section we give details of several applications of the fitting methods described in the previous section including practical aspects of generating the database of electronic energies, etc.

First, we comment on the use of Morse variables, $y_{ij} = \exp(-r_{ij}/a)$ instead of the r_{ij} for global fitting. It is clear that representing V in terms of the latter variables will not yield a physically correct description in the limit of asymptotically large values of r_{ij}, since V will

diverge as these variables go to infinity. By contrast the Morse variables go to a constant (zero) in this limit and this permits a correct physical description of the potential in these fragmentation regions. The terminology 'Morse' comes directly from the observation that the well-known Morse potential $V_M(r)$ is given simply by the three-term expression $D(y^0 - 2y^1 + y^2)$ in contrast to the infinite order expansion in the variable r.

The observant reader will note that if a single, global expression is used for V that this representation does not rigorously become separable in fragment coordinates, except in the case of atom + molecule fragments. Our extensive experience with fitting using a single global representation is that the artificial dependency is generally quite small, i.e., of the order of $10\,cm^{-1}$ (0.03 kcal/mol) or less. We have used other approaches to remove this small dependency, e.g., switching functions [67,68], fitting many-body representations that contain damping functions that give the correct asymptotic behaviour [25,26]. We direct the interested reader to these cited references for more details. It remains to make some remarks about the nonlinear parameter, a, in the definition of y_{ij}. In principle this range parameter could be optimized for each fit and could even be made dependent on i and j. However, our experience with fitting roughly 20 potentials, indicates that the fits are largely insensitive to the precise value of a provided it is in the range 1.5–3 bohr; a typical value is 2.0 bohr. The physical condition is that the potential in, say, a particular bond ij, should be tending to a constant for r_{ij} in the range 4–8 bohr. As long as the basis functions in y_{ij} are still varying in this range we rely on the numerical fitting procedure to produce the correct results.

The least squares fitting is done using standard library routines including singular value decomposition (from the LAPACK library) and employs simple weighting of the data, the details of which depend on the particular application.

Next we give some details on how the database of electronic energies is obtained. Generally there is a trade-off to be made between local accuracy and global coverage, depending on the intended application. If the interest is in ro-vibrational spectroscopy then one requires the highest possible accuracy, but only in the vicinity of a reference configuration, where the nuclear wavefunction is significantly different from zero. In this extreme limit where fragmentation is not of interest fragment data are not needed for the database.

At the opposite end of the trade-off between local accuracy and global coverage are applications to bimolecular reaction dynamics. Clearly this is the most demanding limit as reactants and products can have any relative orientation and impact parameter, and in addition multiple isomers can be formed. An example of a potential surface where this applies is one we developed for the $C(^3P) + C_2H_2$ reaction to give *linear*-C_3H and *cyclic*-$C_3H + H$ [25].

An intermediate case is illustrated by unimolecular reactions, where an energetic molecule can fragment to a single channel. Examples of this are PESs describing the CH_5^+ cation and the water dimer, both of which are described below. If multiple reaction channels are possible and also where 'roaming' dynamics can play a role, then the PES has the full complexity of bimolecular reaction dynamics PES. An example of this is the unimolecular dissociation dynamics of acetaldehyde for which a PES has been developed [69].

To conclude this preamble we note a recent application of the invariant fitting method for the vinyl radical, where two PESs were reported [21]. One PES was fit to a subset of *ab initio* electronic energies that spanned the two equivalent global minima and the

saddle-point separating them. This PES was used in accurate calculations of the vibrational energies. A second PES was fit to these energies plus $H + C_2H_2$ fragment data, the dissociation saddle point, and the high energy 1,2 H-atom migration regions of the PES.

In the following discussion of techniques for sampling configuration space we have in mind mainly applications to dynamics, where barrier crossings and reaction processes are all of interest.

For a three-atom system the geometries for the database can be obtained from a regular mesh in suitable internal coordinates, but already for a four-atom system (six degrees of freedom) this is not really feasible for global potentials. One can then consider placement of configurations on 'sparse grids', and for some applications, e.g., Multimode vibrational calculations, this may be appropriate. (The sparse grids do not need to be associated with normal modes of the system as they are in the Multimode and related codes [36–40], as was discussed briefly in the previous section. These grids can be defined with respect to any system of internal coordinates and this is the basis of the so-called cut-high dimensional model representation [42,57,58], also briefly reviewed in the previous section.) However, in our work for systems of three or more atoms we have relied almost exclusively on sampling procedures that produce scattered data. And as noted already the PESs have generally required of the order of a few tens of thousands of energies, although the PES for CH_3CHO is a fit to roughly 135,000 energies, owing to its great complexity.

Generally we aim to construct highly accurate surfaces based on the best affordable *ab initio* method. However, the procedure to develop such PESs generally starts with a reasonable PES based on a low-level method and basis. At the earliest stage of constructing a PES we may use direct *ab initio* molecular dynamics (AIMD) calculations to sample the configuration space; we also sample the space by making random displacements from known geometries. The AIMD calculations are done at several total energies using the efficient density functional theory, generally with the B3LYP functional, and using a small basis, e.g., VDZ, or even Hartree–Fock with a minimal basis, depending on the molecular system. Once we have obtained an initial sample of geometries then we carry out higher-level calculations, construct an initial fitted surface, and proceed to improve that initial surface through further sampling and further *ab initio* calculations. We use the Molpro program package [70] almost exclusively, and for the case of DFT methods we thereby obtain analytical gradients and also a dipole and quadrupole moment.

The iterative improvement of a PES involves a combination of techniques. We carry out a quasi-Newton search for all stationary configurations on the surface; additional *ab initio* electronic energies are then obtained in the vicinity of these configurations. We always perform further molecular dynamics calculations to sample the configuration space, and often diffusion Monte Carlo calculations of the ground state wavefunction on the preliminary PES as well. The database of configurations from such sampling is pruned by removing near-duplicate configurations. (For this purpose we have designed tests for similarity that involve only permutationally invariant functions of internuclear distances, so we recognize the similarity between configurations in arbitrary relative position and with arbitrary relabelling of nuclei.) A new fit is then done and the process proceeds iteratively until a satisfactory PES is obtained.

We have often used a two-stage procedure in which at first we construct a reasonable PES based on DFT calculations and then we proceed to construct a surface based on the

higher quality coupled cluster method. In cases where the final surface is based on very expensive *ab initio* calculations we have used the earlier dataset to identify a relatively small subset of the database that is adequate to reproduce the surface, and for which the higher-quality and higher-cost *ab initio* calculations are to be carried out and a new final PES obtained. This strategy was used very effectively to obtain a full-dimensional very high-level PES for malonaldehyde that described the H-atom transfer and was used in essentially exact calculations of the tunnelling splitting of H- and D-atom transfer [20].

Next we present the salient aspects of three PESs that illustrate the methods described above. They are PESs for CH_5^+, the water dimer, and CH_3CHO unimolecular dissociation.

3.1. CH_5^+

Research on the CH_5^+ cation has a rich history and we refer the reader to references [67] and [71] for background material. CH_5^+ is also of fundamental interest owing to its highly fluxional nature, even in the zero-point state. Thus it is a molecule that truly requires a proper treatment of the permutational symmetry of the five H atoms. We obtained such a PES [67] by fitting 36,173 CCSD(T) energies obtained with the aug-cc-pVTZ basis using the invariant polynomial methods described in the preceding section. This PES, given in terms of the 15 internuclear distances, mapped to Morse variables, was the successor to an earlier one based on fitting MP2/cc-pVTZ energies [72]. The database of configurations was obtained using a variety of strategies; however, the major one was based on direct dynamics. AIMD trajectories were propagated using the efficient MP2/cc-pVTZ method and basis. These were done at three total energies and the distribution of potential values for a total energy of 3000 cm^{-1} is shown in Figure 1. As seen, the distributions are

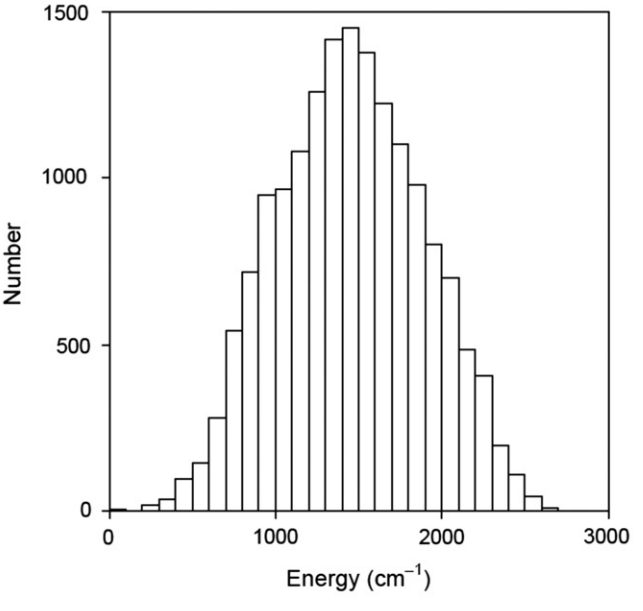

Figure 1. Distribution of *ab initio* molecular dynamics electronic energies for CH_5^+ run at the total energy of 3000 cm^{-1} relative to the global minimum.

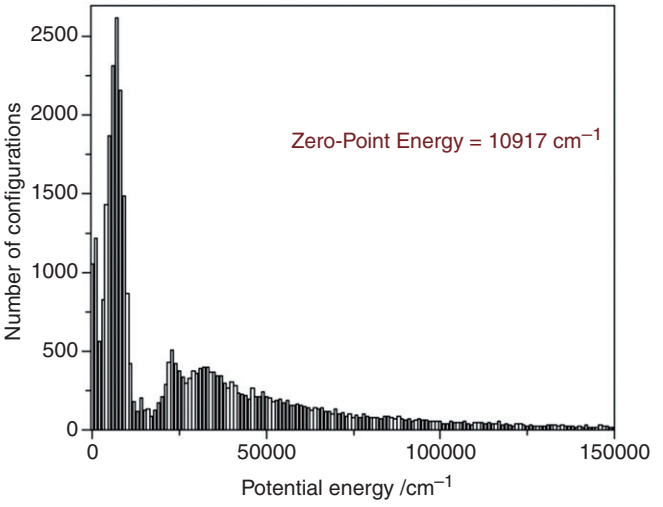

Figure 2. [Colour online] Distribution of electronic energies for CH_5^+ PES.

Gaussian-shaped with the maximum occurring at roughly half the total energy. The most recent CCSD(T)-based PES configurations from these distributions were supplemented with additional high energy ones from direct-dynamics calculations (done at the MP2/cc-pVDZ level of theory and basis). In addition, a dissociation coordinate, defined as the distance between the H_2 centre of mass and the C atom of the CH_3^+, was used to obtain many additional configurations.

The energy distribution of the *ab initio* data used in the fit (relative to the global minimum energy) is shown in Figure 2. As seen, most of the dataset is at energies below 50,000 cm^{-1}, with a large concentration of energies below the dissociation energy of roughly 16,000 cm^{-1}. Energies above that value are from configurations where bonds are highly compressed.

The PES obtained from these data was fit using permutationally invariant polynomials up to total degree 7. Thus there are roughly 2300 terms in the fit. The fitting error (root mean square deviation between the fit and the data) is shown in Figure 3 as a function of the maximum energy in the sample. As seen, it is very small, indicating a very precise fit. An indication of the precision is shown in Table 7 where the geometry and normal mode frequencies are given for the global minimum, denoted C_s(I), and two low-lying first-order saddle points, denoted, C_s(II) and C_{2v}. As seen, the PES describes these properties very well and schematic representations of the structures of these stationary points are shown in Figure 4. In addition the energies of the C_s(II) and C_{2v} saddle points on the PES, 29.1 and 340.7 cm^{-1}, respectively, are in excellent agreement corresponding to CCSD(T)/aug-cc-pVTZ values of 37.7 and 331.2 cm^{-1}. The geometry and harmonic normal mode frequencies at these stationary points on the PES are given in Table 7 where agreement with direct *ab initio* results, also given in the table, is very good.

The potential surface dissociates accurately to fragments $CH_3^+ + H_2$ with an accurate D_e. The harmonic normal frequencies of these fragments are given in Table 8 where the comparison with direct *ab initio* results shows good agreement.

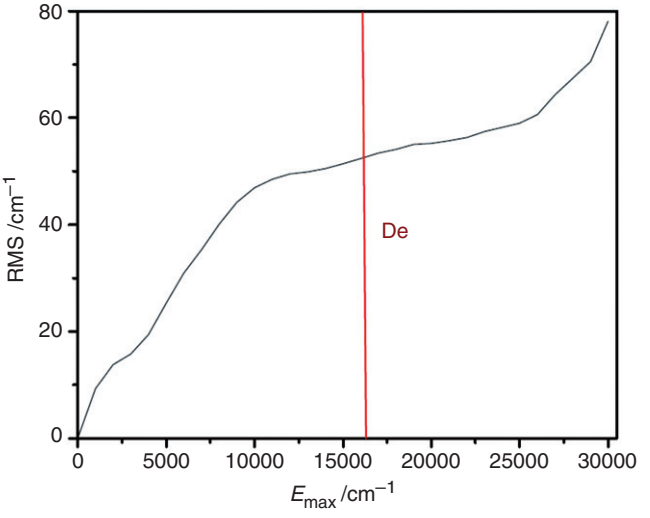

Figure 3. [Colour online] Root mean square fitting error (RMS) of the CH_5^+ PES vs. E_{max}, defined in the text. D_e is the electronic dissociation energy.

Table 7. Geometry (Ångstroms) and normal mode frequencies (cm^{-1}) from the CH_5^+ potential energy surface at the minimum and two saddle points, denoted C_s(I), C_s(II) and C_{2v}, respectively. Results from *ab initio* calculations are given in ref. [67].

	C_s(I)	C_s(II)	C_{2v}
Bond			
C-Ha	1.108 (1.106)a	1.084 (1.084)	1.163 (1.162)
C-Hb	1.197 (1.198)	1.200 (1.199)	1.087 (1.087)
C-Hc	1.197 (1.197)	1.100 (1.098)	1.142 (1.140)
C-Hd	1.088 (1.088)		
H2	0.952 (0.945)	0.945 (0.940)	
Mode			
1	199.9 (227.4)	188.6i	698.5i
2	839.4 (831.3)	995.7	497.7
3	1296.7 (1284.3)	1151.2	1261.1
4	1303.5 (1291.2)	1340.3	1332.6
5	1477.7 (1464.1)	1482.1	1393.2
6	1499.9 (1492.9)	1507.1	1452.1
7	1586.7 (1595.8)	1603.2	1476.3
8	2417.9 (2430.1)	2401.1	2639.9
9	2707.7 (2714.9)	2728.8	2685.0
10	3001.0 (3011.0)	3037.8	2848.0
11	3132.9 (3130.2)	3090.5	3122.0
12	3224.4 (3224.4)	3236.1	3239.1

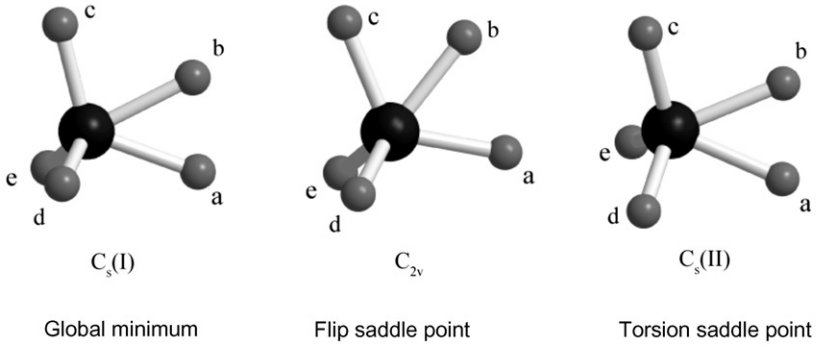

Figure 4. Structure of CH_5^+ global minimum and two low-lying saddle-points from the PES.

Table 8. Normal mode frequencies of CH_3^+ and H_2 from the CH_5^+ PES. Numbers in parentheses are from CCSD(T)/aug-cc-pVTZ calculations of [67].

Mode	CH_3^+	H_2
1	1390.2 (1421.6)a	4377.0 (4400.2)
2	1412.1 (1429.0)	
3	1412.1 (1429.0)	
4	3031.2 (3039.3)	
5	3195.7 (3236.9)	
6	3195.7 (3236.9)	

This PES and in some instances the fitted dipole moment surface, which was also obtained [72], has been used in a number of quantum calculations of the vibrational energies of CH_5^+ [71,73–76] and quasiclassical trajectory calculations of the exchange reaction $CH_3^+ + HD \rightarrow CH_2D^+ + H_2$ [77].

It should be noted that a PES for CH_5^+ has also been reported [64] using the post-symmetrized modified Shepard method, which was briefly reviewed in the previous section, based on CCSD(T)/aug′-cc-pVTZ *ab initio* data, where aug′ indicates no diffuse functions for the H atoms.

3.2. *$(H_2O)_2$*

Next consider the PES for the water dimer. This is a recent application with several versions of the PES reported. In the first report [14] the PES was a fit to roughly 20,000 CCSD(T)/aug-cc-pVTZ electronic energies in 15 Morse variables. A second version of this PES, denoted HBB1, was reported recently and used in essentially exact quantum calculations of the vibration/rotation tunnelling splittings and rotation constants of $(H_2O)_2$ and $(D_2O)_2$ [78], where the agreement with experiment was shown to be of unprecedented quality. HBB1 was a fit to roughly 30,000 CCSD(T)/aug-cc-pVTZ electronic energies (obtained by adding roughly 10,000 energies to the database of

Table 9. Relative energies (cm^{-1}) of 10 stationary points on indicated PES and benchmark value [79]. N_i is the order of the stationary point.

Str. No.	Sym.	N_i	HBB1	HBB2	Benchmark
1 (Min)	Cs	0	0	0	0
2	C1	1	164.1	185.3	181.0
3	Cs	2	196.7	210.2	199.4
4	Ci	1	244.7	260.7	242.7
5	C2	1	328.6	354.4	331.1
6	C2h	3	347.0	375.5	347.0
7	Cs	2	601.4	640.9	627.3
8	C2h	3	1176.4	1230.4	1236.2
9	C2v	1	588.9	629.2	618.1
10	C2v	2	897.1	946.1	939.2

Table 10. Harmonic frequencies (cm^{-1}) at the global minimum on indicated PES and benchmark (CCSD(T)/avqz) values.

Mode	HBB1	HBB2	Benchmark
1	120	126	130
2	132	141	146
3	142	149	152
4	180	181	186
5	351	351	354
6	601	610	619
7	1646	1653	1651
8	1661	1673	1671
9	3734	3757	3748
10	3805	3827	3825
11	3894	3915	3912
12	3911	3934	3931

20,000 used in the first fit). This fit is a combination of a global expression in all the internuclear distance Morse variables, with range parameter equal to 3.0 bohr, of total polynomial order of 7 with 5227 terms plus damped two-body terms. This potential, like the one for CH_5^+, dissociates to fragments, in this case two water monomers.

An even more recent PES, denoted HBB2, was reported. The major difference in this PES is that it gives the essentially exact electronic dissociation energy, D_e (4.98 kcal/mol) [79]. (The previous PESs gave a D_e that is below the "exact" D_e by 0.26 kcal/mol.) This was achieved by noting that the counterpoise corrected CCSD(T)/aug-cc-pVTZ electronic energies slightly underestimate D_e (whereas the uncorrected energies slight overestimate D_e, as usual). Thus we took a simple average of the two sets of energies to get a database of energies with the correct D_e. The high precision of the fit gave a D_e in virtually perfect agreement with that correct D_e.

The energies of 10 low-lying stationary points, including the global minimum, and harmonic frequencies of the minimum for HBB1 and HBB2 are given in Tables 9 and 10,

respectively, along with benchmark results. As seen, there is excellent agreement with benchmarks results.

The structure of the stationary points is shown in Figure 5 and Figure 6 shows the behaviour of HBB2 along a cut where the OO distance varies and the monomer geometries are fixed at their values at the global minimum (str01). As seen, the curve is quite smooth. In addition we have verified that it does have the proper dipole–dipole distance dependence, at least over the range of the plot. At very long range the potential decays

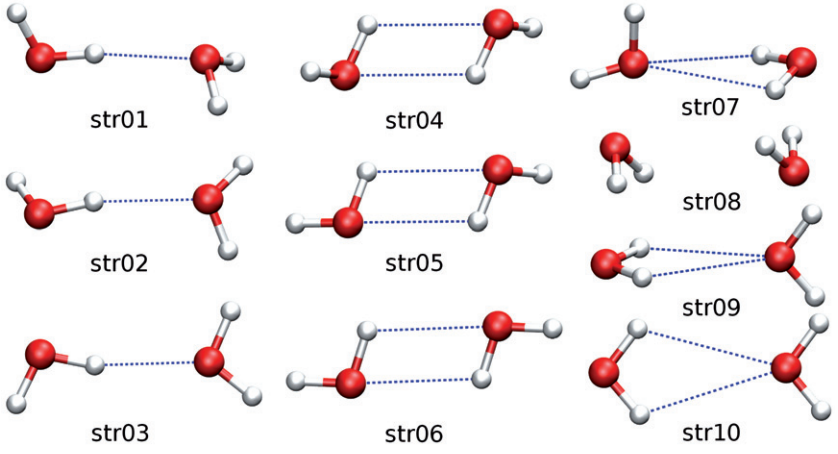

Figure 5. [Colour online] Stationary configurations of the water dimer; str01 is the global minimum.

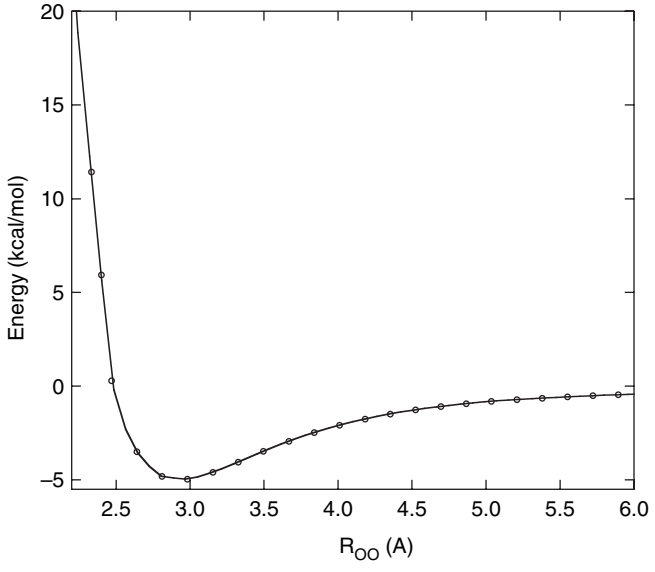

Figure 6. Potential cut of the water dimer PES.

exponentially. This potential along with a soon-to-be-published three-body potential, also fit using invariant polynomials, is planned to be the major component of a flexible, many-body potential for arbitrary numbers of water monomers.

3.3. *CH_3CHO*

We conclude this section with the most complex PES we have considered to date, namely one describing the unimolecular dissociation of $^1CH_3CHO$, where our interest centred on the molecular and radical products, $CH_4 + CO$ and $CH_3 + HCO$, respectively.

The potential landscape of this system is shown in Figure 7, where the complexity of this system can be appreciated. This is a challenging system, not only due to the large number of atoms and many isomers and fragment channels but technically as well because there are multi-reference regions of configuration space, e.g., the doublet + doublet radicals $CH_3 + HCO$. We dealt with this essentially by interpolating through the multi-reference regions. That is, the database of roughly 135,000 CCSD(T)/cc-pVTZ energies was obtained in the single reference regions of configuration space and also from separate calculations for a number of fragments, including $CH_3 + HCO$. The least squares fitting was done with invariant polynomial bases of total order 5 and 6. There are 2655 terms in the fifth degree fit and 9953 in the sixth degree fit. Two fits were done in this case because of the high complexity of the problem and also to use in dynamics calculations, where the lower order fit is more efficient to evaluate. The least squares procedure smoothly interpolates between the single reference region, e.g., the CH_3CHO well region and, say, the radical channel $CH_3 + HCO$. A simple example of how this works is the H_2 molecule,

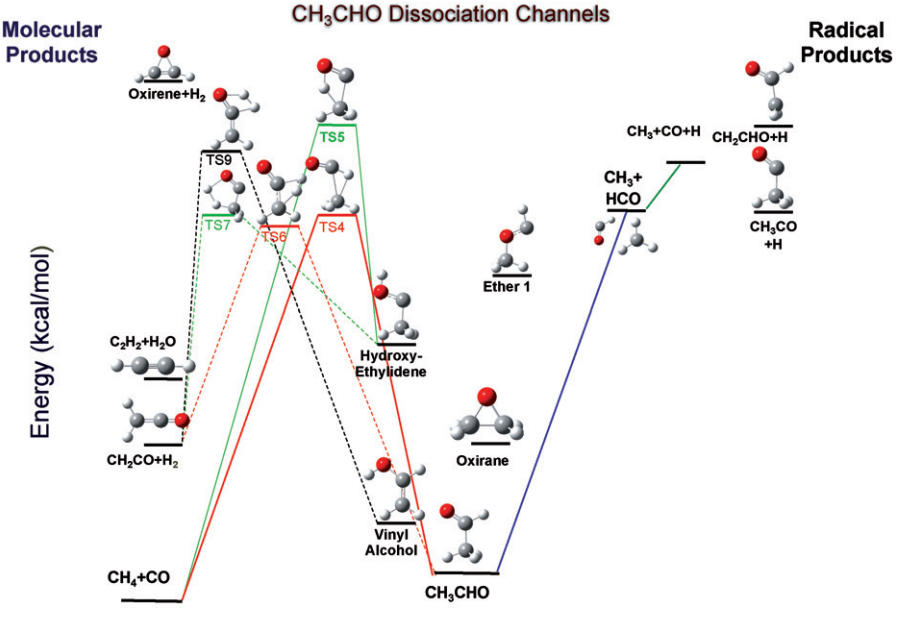

Figure 7. [Colour online] Energy landscape of CH_3CHO PES.

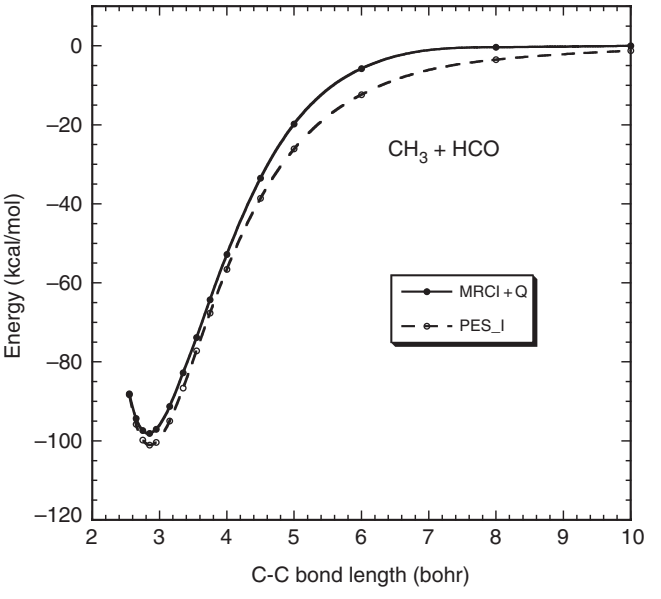

Figure 8. Potential for dissociation of CH_3CHO to $CH_3 + CHO$.

which is well described by a single-reference wavefunction except in the region of dissociation where that description breaks down. However, we know that asymptotically the correct limit is two H atoms with a total electronic energy of -1 h. Thus, if that data is included in a database of energies that contains energies in the single-reference region a least squares fit will smoothly interpolate between the two limits. An example of this interpolation for the $CH_3 + HCO$ channel is shown in Figure 8 from the global minimum. This figure also contains multi-reference, *ab initio* calculations (MRCI + Q) of the electronic energy to test the reality of the PES interpolation. As seen, the PES does produce a smooth result in qualitative agreement with the MRCI + Q energies, as expected. However, as seen, the PES cut decays more gradually than the correct result. Further (unpublished) investigations indicate that the main cause of this behaviour is the placement of the $CH_3 + HCO$ fragment data. That is, the choice of what CC distance to centre that data will obviously determine the rate at which the interpolation becomes asymptotic. With the multi-reference energies as a guide we now know better where to place the fragment data, and new fits based on roughly 30,000 MRCI + Q energies are being developed presently.

The energetics of these two PESs are compared to benchmark results for a variety of stationary points (which are single reference in character) and to fragment channels in Table 11. Note 'TS1 (1–2)', etc. refers to the saddle-point transition state connecting minima '1' and '2'. As seen, the PESs do faithfully represent the *ab initio* energies. Quasiclassical trajectory calculations were performed on both PESs and substantial 'roaming' dynamics was found, i.e., molecular products were formed via a pathway that by-passes the conventional molecular products saddle-point transition state, TS4 (1–10). The experimental and theoretical signature of this other pathway is very highly excited CH_4, following photolysis at 308 nm [80].

Table 11. Comparison of the fitted potential energy surfaces and *ab initio* benchmark calculations for selected stationary points on the H_4C_2O surface (kcal/mol).

Species	CCSD(T)/CBS+CV	Fifth order PES	Sixth order PES
1. Acetaldehyde	0.0	0.0	0.0
2. Vinyl alcohol	9.1	13.5	15.0
3. Hydroxyethylidene	50.8	54.0	53.4
4. TS1 (1-2)	70.8	72.4	72.5
5. TS2 (1-3)	83.0	87.2	85.6
6. TS3 (2-3)	75.5	87.6	81.6
7. TS4 (1-10)	87.5	84.8	96.1
8. TS5 (3-10)	110.9	102.6	108.0
9. TS6 (1-11)	86.0	91.8	95.1
10. $CH_4 + CO$	−2.0	−2.5	−1.6
11. $CH_2CO + H_2$	35.2	43.5	37.1
12. $CH_3 + HCO$	90.8	91.0	94.0
13. $CH_3CO + H$	95.6	100.9	100.5
14. $CH_2CHO + H$	102.5	103.7	103.9
15. $CH_3 + CO + H$	110.5	104.7	106.0
16. $CH_2CH + OH$	124.5	141.7	142.0

4. Summary

We reviewed methods to incorporate permutational symmetry into global representations of potential energy surfaces. A direct symmetrization method to do this was presented, mainly for pedagogical purposes, where basis functions in Morse variables were explicitly symmetrized starting with simple monomial functions. More sophisticated and compact basis functions were described making use of theorems and computational algorithms from the field of invariant polynomial theory. The dipole moment was also represented using these invariant basis functions. Currently basis functions for PESs with up to 10 atoms have been developed.

By construction these PESs contain all permutationally equivalent stationary points and so they can be used in complex dynamics calculations where isomerization, etc. can occur. The other major benefit in using permutationally invariant basis functions is that complex PESs for fairly large molecules can be constructed with only of the order of tens of thousands of electronic energies. We described how the electronic energies are selected for the fitting, and low-level *ab initio* molecular dynamics are often the first step in creating the database of energies. The potentials are linear least squares fit to these energies and so the fitting is quite straightforward.

Roughly 20 PESs have been fit with these methods and these can be downloaded from the website iOpenshell (http://iopenshell.usc.edu/downloads/ezpes/). Here we illustrated the methods, including their practical implementation, for three challenging systems, CH_5^+, $(H_2O)_2$ and CH_3CHO. For CH_5^+ and $(H_2O)_2$ roughly 30,000 electronic energies were sufficient to obtain smooth precisely-fit PESs that dissociate to a single set of fragments. For the much more complex CH_3CHO system the PES was a fit to roughly 135,000 electronic energies. Very recently, we have fit the intrinsic three-body interaction for the water trimer [81]. The goal of this work is to combine that potential with the water dimer PES (described briefly in this review) to have a general many-body PES for water.

Acknowledgements

We thank the Department of Energy and the Office of Naval Research for supporting this work. BJB thanks Prof. Gregor Kemper for advice about invariant computations with Magma.

References

[1] G. C. Schatz, Rev. Mod. Phys. **61**, 669 (1989).
[2] T. Hollebeek, T. S. Ho, and H. Rabitz, Annu. Rev. Phys. Chem. **50**, 537 (1999).
[3] K. Bolton, W. L. Hase, and C. Doubleday, J. Phys. Chem. B **103**, 3691 (1999).
[4] K. Bolton, W. L. Hase, H. B. Schlegel, and K. Song, Chem. Phys. Lett. **288**, 621 (1998).
[5] L. P. Sun, K. Y. Song, and W. L. Hase, Science **296**, 875 (2002).
[6] L. P. Sun and G. C. Schatz, J. Phys. Chem. B **109**, 8431 (2005).
[7] D. Troya, R. Z. Pascual, and G. C. Schatz, J. Phys. Chem. A **107**, 10497 (2003).
[8] J. Z. Pu and D. G. Truhlar, J. Chem. Phys. **116**, 1468 (2002).
[9] D. Troya and E. Garcia-Molina, J. Phys. Chem. A **109**, 3015 (2005).
[10] A. Gonzalezlafont, T. N. Truong, and D. G. Truhlar, J. Phys. Chem. **95**, 4618 (1991).
[11] Y. Qi, K. L. Han, and A. J. C. Varandas, Chin. J. Chem. Phys. **20**, 109 (2007).
[12] A. Brown, B. J. Braams, K. Christoffel, Z. Jin, and J. M. Bowman, J. Chem. Phys. **119**, 8790 (2003).
[13] Z. Jin, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **110**, 1569 (2006).
[14] X. Huang, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **110**, 445 (2006).
[15] A. Shank, Y. M. Wang, A. Kaledin, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **130**, 144314 (2009).
[16] Y. Wang, S. Carter, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **128**, 071101 (2008).
[17] A. R. Sharma, J. Y. Wu, B. J. Braams, S. Carter, R. Schneider, B. Shepler, and J. M. Bowman, J. Chem. Phys. **125**, 224306 (2006).
[18] X. Huang, B. J. Braams, S. Carter, and J. M. Bowman, J. Am. Chem. Soc. **126**, 5042 (2004).
[19] X. Huang, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **122**, 044308 (2005).
[20] Y. Wang, B. J. Braams, J. M. Bowman, S. Carter, and D. P. Tew, J. Chem. Phys. **128**, 224314 (2008).
[21] A. R. Sharma, B. J. Braams, S. Carter, B. C. Shepler, and J. M. Bowman, J. Chem. Phys. **130**, 174301 (2009).
[22] X. B. Zhang, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **124**, 021104 (2006).
[23] G. Czako, B. C. Shepler, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **130**, 084301 (2009).
[24] G. Czako, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **112**, 7466 (2008).
[25] W. K. Park, J. Park, S. C. Park, B. J. Braams, C. Chen, and J. M. Bowman, J. Chem. Phys. **125**, 081101 (2006).
[26] C. Chen, B. C. Shepler, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **127**, 104310 (2007).
[27] S. L. Zou and J. M. Bowman, Chem. Phys. Lett. **368**, 421 (2003).
[28] X. B. Zhang, S. L. Zou, L. B. Harding, and J. M. Bowman, J. Phys. Chem. A **108**, 8980 (2004).
[29] J. N. Murrell, S. Carter, S. C. Farantos, P. Huxley, and A. J. C. Varandos, *Molecular Potential Energy Functions* (John Wiley, Chichester, 1984).
[30] A. Schmelzer and J. N. Murrell, Int. J. Quant. Chem. **2**, 287 (1985).
[31] M. A. Collins, Theor. Chem. Acc. **108**, 313 (2002).
[32] G. E. Moyano and M. A. Collins, Theor. Chem. Acc. **113**, 225 (2005).
[33] M. A. Collins, Comput. Sci. – ICCS 2003, Pt IV, Proc. **2660**, 159 (2003).
[34] M. A. Collins, J. Chem. Phys. **127**, 024104 (2007).
[35] K. C. Thompson, M. J. T. Jordan, and M. A. Collins, J. Chem. Phys. **108**, 8302 (1998).
[36] S. Carter, S. J. Culik, and J. M. Bowman, J. Chem. Phy. **107**, 10458 (1997).
[37] J. M. Bowman, S. Carter, and X. Huang, Int. Rev. Phys. Chem. **22**, 533 (2003).

[38] J. M. Bowman, T. Carrington, and H. D. Meyer, Molec. Phys. **106**, 2145 (2008).
[39] G. Rauhut, V. Barone, and P. Schwerdtfeger, J. Chem. Phys. **125**, 054308 (2006).
[40] K. Yagi, C. Oyanagi, T. Taketsugu, and K. Hirao, J. Chem. Phys. **118**, 1653 (2003).
[41] H. Rabitz and O. F. Alis, J. Math. Chem. **25**, 197 (1999).
[42] G. Y. Li, C. Rosenthal, and H. Rabitz, J. Phys. Chem. A **105**, 7765 (2001).
[43] A. Jäckle and H.-D. Meyer, J. Chem. Phys. **109**, 3772 (1998).
[44] G. G. Maisuradze, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **119**, 10002 (2003).
[45] Y. Guo, A. Kawano, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **121**, 5091 (2004).
[46] I. V. Tokmakov, A. F. Wagner, M. Minkoff, and D. L. Thompson, Theor. Chem. Acc. **118**, 755 (2007).
[47] Y. Guo, I. Tokmakov, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **127**, 214106 (2007).
[48] G. G. Maisuradze, A. Kawano, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **121**, 10329 (2004).
[49] A. Kawano, Y. Guo, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **120**, 6414 (2004).
[50] A. Kawano, I. V. Tokmakov, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **124**, 054105 (2006).
[51] R. Dawes, D. L. Thompson, Y. Guo, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **126**, 184108 (2007).
[52] R. Dawes, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **128**, 084107 (2008).
[53] R. Dawes, A. Passalacqua, A. F. Wagner, T. D. Sewell, M. Minkoff, and D. L. Thompson, J. Chem. Phys. **130**, 144107 (2009).
[54] L. Zhang, P. Luo, Z. Huang, H. Chen, and A. J. C. Varandas, J. Phys. Chem. A **112**, 7238 (2008).
[55] S. Manzhos and T. Carrington, J. Chem. Phys. **127**, 12707 (2007).
[56] S. Manzhos, X. G. Wang, R. Dawes, and T. Carrington, J. Phys. Chem. A **110**, 5295 (2006).
[57] O. F. Alis and H. Rabitz, J. Math. Chem. **29**, 127 (2001).
[58] G. Y. Li, S. W. Wang, C. Rosenthal, and H. Rabitz, J. Math. Chem. **30**, 1 (2001).
[59] O. Vendrell, F. Gatti, and H. D. Meyer, Angew Chem. – Int. Ed. **48**, 352 (2009).
[60] N. I. Hammer, E. G. Diken, J. R. Roscioli, M. A. Johnson, E. M. Myshakin, K. D. Jordan, A. B. McCoy, X. Huang, J. M. Bowman, and S. Carter, J. Chem. Phys. **122**, 244301 (2005).
[61] M. Kaledin, A. L. Kaledin, and J. M. Bowman, J. Phys. Chem. A **110**, 2933 (2006).
[62] M. Kaledin, A. L. Kaledin, J. M. Bowman, J. Ding, and K. D. Jordan, J. Phys. Chem. A **ASAP 117**, 7671 (2009).
[63] C. R. Evenhuis and M. A. Collins, J. Chem. Phys. **121**, 2515 (2004).
[64] K. C. Thompson, D. L. Crittenden, and M. J. T. Jordan, J. Amer. Chem. Soc. **127**, 4954 (2005).
[65] H. Derksen and G. Kemper, *Computational Invariant Theory* (Springer-Verlag, Berlin, 2002).
[66] W. Bosma, J. Cannon, and C. Playoust, J. Comp. Symb. Comp. **24**, 235 (1997).
[67] Z. Jin, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **110**, 1569 (2006).
[68] Z. Xie, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **122**, 224307 (2005).
[69] B. C. Shepler, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **111**, 8282 (2007).
[70] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, A. Mitrushenkov, G. Rauhut, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, G. Hetzer, T. Hrenar, G. Knizia, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, and A. Wolf, *MOLPRO* version 2009.1. http://www.molpro.net

[71] X. Huang, A. B. McCoy, J. M. Bowman, L. M. Johnson, C. Savage, F. Dong, and D. J. Nesbitt, Science **311**, 60 (2006).
[72] A. Brown, A. B. McCoy, B. J. Braams, Z. Jin, and J. M. Bowman, J. Chem. Phys. **121**, 4105 (2004).
[73] A. B. McCoy, B. J. Braams, A. Brown, X. Huang, Z. Jin, and J. M. Bowman, J. Phys. Chem. A **108**, 4991 (2004).
[74] X. Huang, L. M. Johnson, J. M. Bowman, and A. B. McCoy, J. Amer. Chem. Soc. **128**, 3478 (2006).
[75] M. P. Deskevich, A. B. McCoy, J. M. Hutson, and D. J. Nesbitt, J. Chem. Phys. **128**, 094306 (2008).
[76] X. G. Wang and T. Carrington, J. Chem. Phys. **129**, 234102 (2008).
[77] K. M. Christoffel, Z. Jin, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **111**, 6658 (2007).
[78] X. Huang, B. J. Braams, J. M. Bowman, R. E. A. Kelly, J. Tennyson, G. C. Groenenboom, and A. van Der Avoird, J. Chem. Phys. **128**, 034312 (2008).
[79] G. S. Tschumper, M. L. Leininger, B. C. Hoffman, E. F. Valeev, H. F. Schaefer, and M. Quack, J. Chem. Phys. **116**, 690 (2002).
[80] B. R. Heazlewood, M. J. T. Jordan, S. H. Kable, T. M. Selby, D. L. Osborn, B. C. Shepler, B. J. Braams, and J. M. Bowman, Proc. Nat. Acad. Sci. **105**, 12719 (2008).
[81] Y. Wang, B. C. Shepler, B. J. Braams, and J. M. Bowman, J. Chem. Phys. 131, 054511 (2009).