

# Fitting Potential Energy Surfaces

George C. Schatz

Department of Chemistry, Northwestern University, Evanston IL 60208-3113 USA

**Summary.** This paper reviews several methods for determining analytical representations of potential energy surfaces for small molecule systems which are involved in unimolecular dissociation and/or bimolecular reaction. These methods may be categorized as “global” or “local” and they may involve either “fitting” or “interpolation” of ab initio data. The methods may be applied either directly to the full potential energy surface, or to individual terms in a many-body expansion of the full potential surface. In addition, most of the methods may be applied to the description of multiple coupled potential energy surfaces (typically using a diabatic representation) as well as to Born-Oppenheimer surfaces. Included in the global methods are least squares fitting and spline, Morse-spline, rotated Morse-spline, and reproducing kernel Hilbert space interpolation methods. The local methods include Shepard and moving least squares interpolation. Examples of the application of these methods to several triatomic reactive surfaces are discussed.

## 1. Introduction

This work is concerned with the development of analytical potential energy surfaces for chemical reactions that are suitable for molecular dynamics or quantum scattering calculations. For many applications to the description of chemical reaction dynamics, the development of accurate global surfaces is the most difficult part of the project, yet there are a variety of circumstances where it is desirable to put a lot of effort into this, despite the difficulties. However there are often ways to simplify the process of representing surfaces so that the effort is tolerable, and one of the goals of this paper is to show what works best under a variety of circumstances.

The typical situation that we shall consider is where the number of atoms in the reacting system is small enough (3-6), and the desired level of detail in the dynamics is high enough, that one can and wants to do high quality electronic structure calculations to generate the potential energy at a large number of geometries (say  $10^2$ - $10^5$ ). At the same time, the computational effort to determine the potential surface is too large to allow using a “direct dynamics” method, wherein the electronic structure calculations are done for every geometry needed in the dynamics calculation. The situation of interest is thus one where one wants to define an analytical function which can “fit” or “interpolate” the points (here we reserve the word “interpolate” to refer to methods that exactly match electronic structure data at each ab initio point, and “fit” to refer to methods that do not exactly match the input data). Evidently the function we seek should provide a realistic representation of the surface, and it must also be simple enough to evaluate that it can be used over and over again for dynamics calculations with acceptable computational effort. Although much of our discussion will be centered on the development of surfaces for chemical

reactions and other problems where a single valued potential energy surface is all that one needs to describe the problem of interest, we shall also be interested in electronically nonadiabatic processes, where one needs multiple potential energy surfaces and their couplings.

For certain classes of molecules, such as biomolecules, one can develop very general potential surfaces (force fields) by summing simple functions (harmonic stretches, bends, torsions, van der Waals, nonbonded interactions, etc) that contain empirically determined parameters. Often these parameters are highly transferable, which means that a CH stretch force constant can be assumed to have the same value for all CH's in all molecules. This is the idea behind the well known CHARMM, AMBER and GROMOS force fields that are used in biomolecular simulation studies. Such surfaces are good at describing dynamical processes that do not involve the making or breaking of bonds, such as conformational transitions that occur in protein folding. However the development of surfaces that describe bond breaking and forming is more difficult, and for the most part such surfaces have been developed only for reactions involving a few atoms, or for reactions involving a few atoms that are surrounded by "solvent" that is treated using empirical force fields.

If one can afford to do direct dynamics calculations, there are some additional strategies for developing potential surfaces that can be considered. Of course in the simplest version of direct dynamics, one solves the electronic structure problem for each time step or each grid point needed for the problem of interest, without regard to the proximity of this point to any previously calculated points. Obviously this procedure is very time consuming, and also very wasteful, but for applications where the time spent doing electronic structure calculations is small, this approach is often preferred due to its simplicity. For example, many kinds of semiempirical molecular orbital theory calculations (such as AM1) are simple enough to make this practical. However for reactions in which bonds are broken or formed, the accuracy of these methods is not sufficient to determine useful rate coefficient information reliably. Sometimes semiempirical methods can be reparametrized to circumvent this problem, but this by itself, introduces an element of "fitting" into the application. An alternative that may be of use in some situations is to develop methods that start with direct dynamics, but store away points on the potential surface as they are calculated, and then use this information to develop an interpolated surface that gradually replaces the direct dynamics calculation in the limit of enough points. We shall describe one method that does this "interpolation on the fly" later.

There are several approaches to developing analytical potential surfaces for chemical reactions, as reviewed in a variety of places [1, 2, 3, 4, 5]. This paper will primarily highlight those that have been developed or used in my group in the last few years. These methods have all been used for describing unimolecular or bimolecular reactions, and thus all use functions that allow for dissociation to produce fragments. The methods used may be categorized as follows:

A. Global Methods. These are methods in which the potential surface at each point is determined by all the ab initio points that are input to the fit.

Usually this means that the ab initio points must cover a broad range of geometries, corresponding to everything that is accessible to the dynamical process of interest. If more ab initio points are added to the data set, the fit must be redone. Within this group, there are several different methods, including: (1) least squares (LS) fitting methods, (2) Spline, Morse-spline and rotated Morse-spline interpolation methods, and (3) reproducing Kernel Hilbert Space (RKHS) interpolation methods.

B. Local Methods. These methods determine the potential surface at each point based only on ab initio data that is available for geometries close to that point. In these methods, if more ab initio points are added, only the fit near each new point must be redone. Examples of this method are: (1) Shepard Interpolation and (2) Moving Least Squares. A detailed description of each of these methods will be given later.

The methods described here are only a representative subset of all the methods that have been used for representing reactive potential energy surfaces for unimolecular and bimolecular potential energy surfaces. Among those which will not be discussed are the LAGROBO approach of Laganà and coworkers [6, 7], the diatomics-in-molecules (DIM) method [8] which has been implemented by several groups, such as by Kuntz, [9] and the double many-body expansion (DMBE) method of Varandas and coworkers [3]. LAGROBO is a global approach which is similar in spirit to the rotated Morse-spline method that is described in detail later. DIM is a semiempirical electronic structure approach derived from valence-body theory that uses empirical information about diatomic potential curves to determine parameters in the electronic structure calculation that describe bonding in a many-electron system. Often some of the diatomic curves are treated as adjustable parameters in optimizing potential energy surface properties so that they match high level ab initio results. An approximation to DIM is the commonly used LEPS (London-Eyring-Polanyi-Sato) potential surface, which was very popular in the 1960's and 1970's. DMBE uses a many-body expansion (described below) and expresses each term in terms of two contributions: that from extended Hartree-Fock (EHF) calculations, and that describing dynamical correlation(dc) effects that EHF doesn't include. Both parts have short and long range contributions, and the analytical form may be expressed in a variety of ways, but the long range components are typically inserted using expansions in inverse powers of the relevant separation coordinates. DMBE can be considered to be a "dual level" method [10] in that the EHF surface is developed using one level of ab initio calculation and the dc part involves a different level. There are many other examples of dual level methods, including methods that use semiempirical methods to define a low level surface, and then ab initio methods to correct it.

In the following sections, I first discuss several general strategies that must be considered in developing analytical representations of potential surfaces, then I describe each of the methods mentioned above, and finally I give examples of the application of these methods to specific chemical reactions. Among the reactions that will be considered are H + CO, H + N<sub>2</sub>, O(<sup>1</sup>D) + H<sub>2</sub>, H + H<sub>2</sub>, Cl + HCl.

## 2. Representing Potential Energy Surfaces: General Considerations

In developing global analytical representations of potential surfaces there are a number of general issues that must be considered that often determine what class of method is appropriate. These include the following:

- I What coordinate system does one want to use in describing the surface? For a diatomic molecule the surface only depends on the distance  $R$  between the nuclei. For a triatomic molecule it depends on the three internuclear distances ( $R_{12}, R_{23}, R_{13}$ ), but for linear geometries one of the distances is equal to the sum of the other two, and thus the three distances are no longer independent. More generally the potential depends on  $3N_{\text{nucl}} - 6$  distances, where  $N_{\text{nucl}}$  is the number of atoms. For  $N_{\text{nucl}} > 4$ , this number is smaller than the number of internuclear coordinates  $N_{\text{nucl}}(N_{\text{nucl}} - 1)/2$ , so it is not necessary to use all the internuclear coordinates to define the potential. For  $N_{\text{nucl}} > 4$ , there is a problem with distinguishing certain symmetry related configurations using  $3N_{\text{nucl}} - 6$  coordinates. See Collins and coworkers for a detailed discussion of this [11, 12].  
 Although internuclear distances are convenient for specifying the repulsive interactions between atoms when two are pushed together, they are often not the best choice for representing portions of the surface that relate to bond rearrangement. In this case it is often best to use other coordinates for defining the surface. For example for triatomic systems it is common to express the surface in terms of Jacobi coordinates. Also, even when internuclear distances are used, it is common to express the potential in terms of functions of these distances, such “bond order” variables, which are exponentials of the form  $n = e^{-\alpha R}$  where  $\alpha$  is a parameter, as this maps the physically meaningful range of coordinates from  $R = (0, \infty)$  to  $n = (0, 1)$ .
- II Is there any identical particle symmetry in the system being considered, and is it necessary to develop a fitting function that reflects this symmetry or perhaps a subgroup of the full symmetry group? While it may seem like an obvious point that one would want to insist on full group symmetry, this may be very difficult to invoke within the context of some fitting strategies, and moreover it may not be important if exchange of certain identical atoms is energetically or dynamically forbidden.
- III Does one want to apply the selected method directly to the full potential, or to individual terms in a many-body decomposition of the potential? In a many-body decomposition, one writes the full potential surface  $V$  in the following form:

$$V = V_1 + V_2 + V_3 + V_4 + \cdots + V_{N_{\text{nucl}}} \quad (2.1)$$

Here  $V_1$  refers to one-body terms that only depend on the positions of one atom at a time (generally these are taken to be zero if the atoms are all in their ground states when the system dissociates), while  $V_2$  refers to all the

two-body (diatomic) terms,  $V_3$  to the three-body (triatomic) terms, etc. The convention in these many-body terms is that  $V_i$  vanishes when the system has dissociated into fragments where the largest fragment molecule has fewer than  $N_i$  atoms. As a result, the  $V_i$ 's are of finite range in at least some sense, which can be convenient for applications to dynamical processes where molecules separate into one or more fragments. Another advantage is that once a given  $V_i$  has been determined, it can be used in any potential function requiring the same fragment in the same electronic state (i.e., the HCO three-body term can be used as one component of the HCO<sub>2</sub> potential surface). However a disadvantage of this approach is that several fitting functions must be determined when only one is needed or desired. Murrell and coworkers have used the many-body approach extensively in conjunction with least squares fitting methods [1].

- IV What are the geometries where the energy points are to be calculated using an electronic structure program, and can the point locations be optimized? Some methods, such as spline methods, require rectangular grids, while others, such as Shepard interpolation, can use points that are located randomly, or which are taken from trajectory calculations. For a triatomic potential surface, with only three internal degrees of freedom, it is often possible to determine a potential surface with a rectangular grid of roughly 10 points in each coordinate, thereby requiring 10<sup>3</sup> points. However, the same idea when applied to a four atom surface (six degrees of freedom) requires 10<sup>6</sup> points, which is usually not feasible. The energetically important region of the four atom surface may only occupy 10<sup>4</sup> of the 10<sup>6</sup> points, but the region covered may be highly irregular, so there may be advantages to using a method that can develop a surface based on an irregular region of the potential.
- V Does one have gradient (or even Hessian) information about the surface, and can this be used in refining the fit? Traditionally, the highest quality ab initio methods did not provide gradient or Hessian (second derivative) information, so the best surfaces were derived from energy information. However this situation is gradually changing, and it is becoming increasingly important to develop methods that can incorporate this additional information into the fitting function. The Shepard method [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23] is the one approach which makes this easy; in fact in most applications of this and related approaches, it is essential that this information be available if the method is to provide an acceptable surface. Other methods, such as spline methods, can be developed in such a way that gradient information and energy information are both fit, but such applications have not yet appeared.
- VI If multiple potential surfaces and couplings are to be fit, does the fitting function to be used have the correct functional form in regions of intersection between surfaces. Almost all potential surfaces that have been developed for studying nonadiabatic dynamics have been based on diabatic representations where the potential surfaces and couplings are smooth functions of the coordinates. Often these diabats intersect, and in fact an important

property of the fitted surfaces is that the intersection seam be described accurately, and that the coupling surface have the proper dependence on distance from the seam to make the intersection be conical. However there is no current method for generating global surfaces which makes it possible to constrain both the location and energy of the intersection seam in developing the fit. Sometimes it is sufficient to constrain only the symmetry of the intersection seam, and to force the intersection to be conical. These can be done with an appropriate choice of fitting function and coordinates.

- VII Does one need flexibility in the fitting function so that features of the surface (like a barrier height) can be adjusted if needed? This is often crucial to the application, as experiments are notoriously sensitive to the barrier heights, but electronic structure calculations are rarely of sufficient accuracy to predict the desired barrier without adjustment. Often the adjustment is a much smoother function of coordinates than is the potential function, so in this case it may be useful to use a “dual” fitting strategy, as described earlier. In addition, one sometimes finds that it is most convenient to use different fitting strategies for different parts of the potential, and then switch between one fitting form and another.

### 3. Fitting and Interpolation Methods

#### 3.1 Global Methods

**3.1.1 Least squares (LS) fitting methods.** Least squares is one of the oldest methods used to fit potential surfaces, and it is still a reasonable approach for some applications. The standard method involves expanding the potential surface in terms of basis functions  $b_k(\mathbf{R})$  (here we let  $\mathbf{R}$  denote the collection of coordinates used to describe the potential), with coefficients  $a_k$ , as in

$$V(\mathbf{R}) = \mathbf{a}^T \mathbf{b}(\mathbf{R}) = \sum_{k=1}^K a_k b_k(\mathbf{R}), \quad (3.1)$$

where the superscript  $T$  denotes “transpose” and  $K$  is the number of basis functions. We define the coordinates and energy values for the data points to be interpolated as  $\mathbf{R}(j), v(j)$  ( $j = 1, 2, \dots, J$ ), where  $J$  is the number of data points.

To determine the coefficients  $a_k$ , consider the functional  $E(V)$ :

$$E(V) = \sum_{j=1}^J w_j [V(\mathbf{R}) - v(j)]^2. \quad (3.2)$$

where  $w_j$  is a weight to be associated with the  $j$ th point (generally taken to be unity in fitting ab initio points). From the condition that  $E(V)$  be stationary with respect to variation in the parameters  $a_k$ , the following “normal” equations are obtained:

$$\mathbf{B} \mathbf{W} \mathbf{B}^T \mathbf{a} = \mathbf{B} \mathbf{W} \mathbf{v}, \quad (3.3)$$

where  $W$  is a diagonal matrix constructed from all the weights

$$W = \text{diag}[w_1, w_2, \dots, w_J], \quad (3.4)$$

and  $B$  is composed of values of the basis functions  $b(\mathbf{R})$  at  $\mathbf{R}(j)$ :

$$B = \begin{pmatrix} b_1(\mathbf{R}(1)) & b_1(\mathbf{R}(2)) & \dots & b_1(\mathbf{R}(J)) \\ b_2(\mathbf{R}(1)) & b_2(\mathbf{R}(2)) & \dots & b_2(\mathbf{R}(J)) \\ \dots & \dots & \dots & \dots \\ b_K(\mathbf{R}(1)) & b_K(\mathbf{R}(2)) & \dots & b_K(\mathbf{R}(J)) \end{pmatrix}. \quad (3.5)$$

The functions  $b_k$  can be chosen with great freedom, but often they involve polynomials in all the variables (a multinomial) multiplied by an exponential in the dissociating coordinate.

Solution of the normal equations only needs to be done once, and the simple linear form in Eq. (3.1) allows for very efficient evaluation of the potential for use in trajectory calculations. Because of this simplicity, least squares methods have been very popular [1, 24, 25]. Unfortunately the simplicity of this approach can be deceptive, as the quality of the fit is very strongly dependent on making a clever choice of the number and nature of the basis functions. In addition, it is easy for a least squares surface to have serious defects (such as spurious minima) in localized regions of the potential. Since potential surfaces often have an exponential dependence on coordinates, it is sometimes seductive to use expansions which depend on nonlinear parameters (parameters appearing in exponents) rather than the linear parameters presented above. However, nonlinear parameters can be difficult to optimize.

One of the more successful potentials that has been developed with least squares methods is the ground state water potential developed by Schinke and Lester [25]. This potential used the same many-body expansion form, an exponential times multinomial basis functions, as in the surfaces developed by Murrell, Carter and coworkers, including several that they developed for water [26], but unlike the MC surfaces, they allow the origin for expanding the multinomial to float so as to optimize the fit. In addition they fit ab initio calculations that were globally distributed. In fact this was the first water surface that was based on fits to globally defined ab initio points. Three fits were developed, representing slightly different parametrizations that influence the long range part of the potential. Each of these fits was of comparable quality in representing the ab initio points, but the long range potential was sufficiently different that some measurable properties, such as the  $O(^1D) + H_2$  thermal rate constant, were noticeably different. Only the first fit, denoted SL1, has been commonly used.

Based on comparisons with experiment [27], and comparisons with better surfaces for water that will be described later [28], the SL1 surface is now known to be qualitatively correct, but it has two serious flaws. One is that trajectories are difficult to integrate because the high order multinomial used was unphysically steep at geometries that are easy to access in  $O(^1D) + H_2$  reaction. The second is that the SL1 surface involves a fit to relatively low quality (by modern standards) electronic structure data, so a number of properties of the water molecule and of the dissociated fragments are inaccurate.

### 3.1.2 Spline and Morse-spline interpolation methods.

(a) Spline interpolation of all degrees of freedom. Spline methods start with the same basic equation as with linear least squares (Eq. (3.1)). However the basis functions are chosen in a special way (see below), and the coefficients are determined by interpolation of the ab initio data rather than by fitting the data. For a one-dimensional function, the basis functions are called B-splines. These functions are only nonzero over a finite range of grid points, and within this range they are polynomial functions (typically cubic) that are defined so that they vanish at the bounding grid points, and so also do as many derivatives as are compatible with the order of the polynomial. The interpolation equations that determine the coefficients are generated by demanding that Eq. (3.1) be exactly satisfied at each data point, where the number of data points equals the number of basis functions. The resulting set of linear equations is especially easy to solve if the points are taken from an equally spaced grid.

Multidimensional splines (2D, 3D) may be developed, as described in Sathyamurthy and Raff [29] (see also: Numerical Recipe's [30]), but the number of dimensions that may be represented using this approach is limited by the need for a relatively dense rectangular grid of ab initio data. One advantage of spline methods is that interpolation is not limited to just the energy function; gradients can also be incorporated provided that appropriate piecewise continuous polynomials are chosen.

There have been several applications of the multidimensional spline approach to the development of potential energy surfaces [31, 32], but this method has so-far been restricted to surfaces with 1-3 degrees of freedom. Although it is relatively easy to apply, and it is especially useful for surfaces where the topological features are complex or unknown, the method usually requires too high a point density to be useful for a large class of chemical problems. A good example of a three dimensional spline surface is one developed by Bowman et al(BBH) for the HCO potential surface [32]. This surface was based on 2000 high quality ab initio calculations, but the spline function wasn't applied directly to the ab initio data. Instead, the ab initio results were first fit to locally defined polynomial functions, and the polynomials were used to interpolate energies on a higher density grid (45000 points) for input to the three dimensional spline routine. In spite of the large number of points used to do the interpolation, the spline function was still very efficient to evaluate, as only a small subset of all the spline basis functions is nonzero at any point. The high quality of the BBH surface has been documented by a variety of experiment/theory comparisons [33].

(b) Morse-spline methods. Morse-spline methods are based on the idea that many potential functions look like Morse functions as a function of a selected internuclear separation. A Morse function has the form:

$$V_m = D_e [1 - e^{-\alpha(R - R_e)}]^2 \quad (3.6)$$

where the three parameters in the Morse function are  $D_e$ , the dissociation energy,  $R_e$ , the equilibrium distance, and  $\alpha$ , which determines the vibrational force constant. In fitting this function to the potential surface for one coordinate of a polyatomic molecule, it is necessary to add to this function with a fourth parameter that determines the potential at the bottom of the well. These four parameters are usually determined by nonlinear least squares fitting, so a minimum of four ab initio points is needed to do the fit for each set of values of the other coordinates.

The Morse-spline method [34] starts with fits of the Morse potential to describe one coordinate in the surface. The four fitting parameters are then determined as a function of the other coordinates, and then the dependence of the parameters on these other coordinates is represented using spline interpolation. Thus for a triatomic molecule, the Morse function describes one stretch coordinate, and then spline interpolation takes care of the other two (say stretch and bend) coordinates. Note that although spline interpolation is used, the method is really a fit rather than an interpolation, as the Morse parameters being interpolated are themselves obtained by fits to ab initio data. The primary advantage of the Morse-spline method compared to using splines for all degrees of freedom is that the degree of freedom that is represented using the Morse function only requires 4 ab initio points to represent that degree of freedom, whereas a typical spline fit would require significantly more. Also, the Morse parameters vary slowly with the other degrees of freedom, so spline fits of these parameters are usually well behaved.

A detailed application of the Morse-spline method was made by Koizumi et al [34] in studies of the  $\text{HN}_2$  molecule. In this project, 241 ab initio points were fit using the four parameter Morse function to describe the N-N stretch coordinate. A two dimensional spline was used to describe the dependence of the four parameters on two other coordinates (which were chosen to be the H –  $\text{N}_2$  Jacobi distance and angle). The rms error in the Morse fits was only about 0.01 kcal/mol, suggesting a very high quality fit to the Morse function, but no attempt was made to determine the global rms error of the surface. The potential was used to determine the lifetimes of metastable states at the bottom of the well using quantum scattering methods. These lifetimes are important in determining the mechanism of the thermal DeNOx process for removing nitric oxide from smokestacks. After correcting the surface for basis set incompleteness (which was found to be a slowly varying correction that could be described by a single exponential), the  $\text{HN}_2$  ground state lifetime was found to be a few nanoseconds. This result was controversial at the time, as combustion modellers thought that it should be five orders of magnitude longer, however the short value has withstood the test of time and the combustion models have been revised as a result.

- (c) Rotated Morse-spline methods. Rotated Morse-spline (RMS) methods [35, 36] involve a generalization of the Morse-spline fitting idea that allows one to describe potential energy surfaces for bimolecular chemical reactions, such as

the atom-transfer reaction  $A + BC \rightarrow AB + C$ . For such potential energy surfaces, the Morse-like function starts in the reagents looking like the  $BC$  diatomic potential, and ends up in the products looking like  $AB$  vibration. To smoothly switch from one limit to the other, one introduces polar coordinates that rotate smoothly from one diatomic to the other. If  $R_{AB}$  and  $R_{BC}$  are the two internuclear coordinates, then the polar coordinates that do this are defined by:

$$\rho = [(R_{AB}^* - R_{AB})^2 + (R_{BC}^* - R_{BC})^2]^{1/2} \quad (3.7)$$

$$\phi = \tan^{-1} \left[ \frac{(R_{BC}^* - R_{BC})}{(R_{AB}^* - R_{AB})} \right] \quad (3.8)$$

where  $R_{AB}^*$  and  $R_{BC}^*$  define the origin of the coordinate system, and are usually chosen to have values that are about three times the corresponding diatomic equilibrium distances. Note that when  $\phi=0$ ,  $R_{BC}$  is fixed at  $R_{BC}^*$ , and  $\rho$  measures the  $AB$  diatomic distance, whereas for  $\phi = 90^\circ$ ,  $R_{AB}$  is fixed at  $R_{AB}^*$ , and  $\rho$  varies linearly with the  $BC$  distance.

Provided that the coordinate system origin  $R_{AB}^*$ ,  $R_{BC}^*$  is chosen carefully, the Morse representation of the potential as a function of  $\rho$  is often very efficient. As a result, only four well chosen points are needed to define the Morse potential (for each value of the other coordinates). The disadvantage of RMS methods is that one needs to know a good deal about the shape of the potential surface in advance of the fit, in order to choose where the origin and other data points should be located.

Note that the RMS method as described above only describes reactions with a single reaction path, namely  $A + BC \rightarrow AB+C$ . However in many triatomic systems, one wants to be able to describe the three reaction paths that connect the three atom-diatom asymptotes. To describe this situation, one can do a weighted sum of three RMS potentials, one for each reaction path, with a weighting function that switches on whichever RMS potential is appropriate for the geometry being sampled.

A good example of application of the RMS method concerns the potential surface for the reaction  $Cl + HCl \rightarrow ClH + Cl$ . This surface has only one low energy reaction path and thus is ideal for this application. The first RMS surface for this reaction was developed by Garrett et al [37], and more recent work has been done by Maierle et al [38], and by Dobbyn et al [39] using the same polar coordinates but more extensive, and higher quality ab initio results. Both of the latter surfaces have been developed within the context of a diabatic representation fit to the three electronic states that correlate to the  $Cl + HCl$  ground state. Only 180 points were needed to generate the global surface for the Dobbyn et al application. A feature of both the Maierle and Dobbyn surfaces is that the RMS fits were grafted to an accurate long range potential due to Dubernet and Hutson [40] by using switching functions to turn on one and turn off the other over a range of geometries where the surfaces match up.

**3.1.3 Reproducing Kernel Hilbert Space (RKHS) interpolation methods.** The reproducing kernel Hilbert space method [28] provides a very general way to approach the issue of doing interpolation, one which includes multidimensional spline methods as a special case, but which also allows one to develop interpolation functions which satisfy asymptotic boundary conditions and other constraints on the interpolating basis functions which splines cannot.

In all interpolation methods, one uses known values of the potential  $V(R_i)$  to determine the global  $V(R)$  with the requirement that  $V(R)$  exactly match the ab initio data. To accomplish this, while keeping  $V(R)$  smooth, we require that  $V(R)$  satisfy the reproducing property (here written for a one-dimensional function)

$$V(R) = \int q(R, R') V(R') dR' \quad (3.9)$$

The kernel  $q(R, R')$  in this integral equation is determined by the order of smoothness desired, and by the boundary conditions associated with the chosen coordinates.

To define  $V(R)$ , we require that Eq. (3.9) be exactly satisfied at each ab initio point:

$$V(R_i) = \int q(R_i, R') V(R') dR' \quad i = 1, \dots, N \quad (3.10)$$

The solution of this equation is:

$$V(R) = \sum_i^N \alpha_i q(R_i, R) \quad (3.11)$$

where the coefficients  $\alpha_i$  are obtained by solving (either directly or by singular value decomposition):

$$\sum_i^N q(R_j, R_i) \alpha_i = V(R_j) \quad j = 1, \dots, N \quad (3.12)$$

For interpolation in one dimension, say for a diatomic molecule potential, an example [41] of a reproducing kernel is:

$$q_1^{n,m}(x, x') = n^2 B(m+1, n) x_{>}^{-m-1} {}_2F_1(-n+1, m+1; n+m-1; \frac{x_{<}}{x_{>}}) \quad (3.13)$$

where  $m$  and  $n$  are integers such that  $n > 0$  and  $m < 1$ ,  $x_{>}$  and  $x_{<}$  are, respectively, the larger and smaller of  $x$  and  $x'$ ,  $B(a, b)$  is the beta function, and  ${}_2F_1(a, b, c; z)$  is the Gauss hypergeometric function. This kernel is designed so that the first  $n - 1$  derivatives of it evaluated at  $x \rightarrow \infty$  vanish. (This makes the potential smooth and flat, as would be appropriate in the limit that the molecule dissociates in the  $x$  coordinate).

The choice of  $n$  and  $m$  is in principle arbitrary, but in reality it is appropriate to keep both as low in value as possible so that the interpolating function will be smooth. In actual applications,  $n = 2$  and  $m = 0$  has been commonly chosen. Other choices of reproducing kernels are possible, depending on the level of

smoothness of the interpolating function that is desired, and on what constraints on the interpolating function that one wishes to impose.

For applications to potential surfaces in three or more dimensions, the standard strategy has been to develop reproducing kernels separately for each degree of freedom, then take direct products to define multidimensional kernels.

A detailed application of the RKHS method was made to the H<sub>2</sub>O potential surface by Ho et al. [41] In this application, approximately 1280 high quality ab initio points were used for the final surface, but lower quality test surfaces with approximately 10 times as many points were first used to determine optimal grid spacings and kernel function parameters. A many-body representation of the surface was used, with the two-body terms expressed in terms of the Gauss hypergeometric function described above. The three-body term was expressed in terms of Jacobi coordinates ( $R, r, \theta$ ), and then  $R$  and  $r$  were converted to the bond-order coordinates  $x = \exp(-aR)$  and  $y = \exp(-br)$ , with  $a$  and  $b$  fixed to optimize the fit based on the test surfaces. The kernel used for the  $x$  and  $y$  coordinates involved a hypergeometric function similar to that described above, but which satisfies the requirement that the three body term vanish in the limit that  $x$  and  $y$  are either 0 or 1. The angular coordinate is represented by a kernel that is expanded in low order Legendre polynomials.

Tests of the quality of the RKHS surface were done by comparing the surface with a set of 514 ab initio points that were not part of the interpolating data, and which were randomly scattered around the surface at geometries that are accessible in the O(<sup>1</sup>D) + H<sub>2</sub> reaction. The resulting rms error was estimated to be 0.3 kcal/mol, which is remarkably small given that the total range of energies being fit is over 200 kcal/mol. Although this surface is not of spectroscopic quality (i.e., capable of describing the vibrational levels of water to within a wavenumber) it is quite adequate for describing the O(<sup>1</sup>D) + H<sub>2</sub> reaction dynamics. Schatz et al [42] found that although the agreement with experiment was excellent for low translational energy results (<2 kcal/mol), the agreement at higher energies was not. This pointed to the need to include excited potential surfaces in the dynamics. A RKHS surface for the lowest A" symmetry state was subsequently developed [42], and it was found that it contributes noticeably to the results at energies above 2 kcal/mol.

### 3.2 Local Methods

**3.2.1 Shepard Interpolation.** In the Shepard method, the potential energy surface  $V(\mathbf{R})$  is represented by a weighted sum of Taylor expansions  $T_i(\mathbf{R})$  about each ab initio point:

$$V(\mathbf{R}) = \sum_{i=1}^J w_i(\mathbf{R}) T_i(\mathbf{R}), \quad (3.14)$$

where  $\mathbf{R}$  is a vector of  $3N_{nucl} - 6$  internal coordinates (same notation as above), and  $J$  is the number of points. The basic idea underlying this equation is the assumption that we have ab initio data in sufficient density so that any geometry of interest is within the domain of convergence of Taylor expansions about at

least one point in the ab initio data set. The weight functions are therefore chosen so as to “turn on” whenever the ab initio points are suitably close to the geometry being considered, and the resulting potential is just the weighted average of Taylor expansion estimates from the nearby points. Of course one would like to have a method that smoothly interpolates between adjacent ab initio points even when they are not that close to one another. In the following paragraphs, we describe how this can be done.

Following Collins and coworkers, [11, 12, 13, 14, 15, 16, 17, 18, 19] we choose as internal coordinates the inverse coordinates  $\mathbf{z} = 1/\mathbf{R}$ . This has the advantage that a single *ab initio* point used in Eq. (3.14) describes the asymptotic behavior of the isolated diatomic potential quite accurately [21]. The reason for this is that Taylor expansions in inverse coordinates have a much larger domain of convergence for diatomic potentials than the corresponding expansions in the coordinates themselves.

The Taylor expansion  $T_i$  is

$$\begin{aligned} T_i(\mathbf{z}) &= V[\mathbf{z}(i)] + \sum_{\alpha=1}^{3N_{nuc}-6} [z_\alpha - z_\alpha(i)] \left. \frac{\partial V}{\partial z_\alpha} \right|_{\mathbf{z}(i)} \\ &+ \frac{1}{2!} \sum_{\alpha=1}^{3N_{nuc}-6} \sum_{\beta=1}^{3N_{nuc}-6} [z_\alpha - z_\alpha(i)][z_\beta - z_\beta(i)] \left. \frac{\partial^2 V}{\partial z_\alpha z_\beta} \right|_{\mathbf{z}(i)} + .(3.15) \end{aligned}$$

where  $V[\mathbf{z}(i)]$  is the value of the potential energy at  $\mathbf{z}(i)$  and the derivatives are with respect to inverse distances at  $\mathbf{z}(i)$ , evaluated at the  $i$ -th data point.

One might imagine that the higher the order of the Taylor expansion, the larger the domain of convergence, and the fewer number of points needed to achieve an acceptable surface. However there are practical considerations which limit the highest order expansion to quadratic. This is due to the fact that very few electronic structure codes provide derivatives higher than quadratic, and those that do tend to involve a lower quality electronic structure calculation than is needed for describing chemical reactions and other processes where bonds are made or broken. In fact, the very best electronic structure methods do not provide any derivative information, and this causes serious problems for the Shepard method, as Collins and coworkers have demonstrated that one needs to have gradient and hessian information in order to generate reliable surfaces. The reason for this is that if the Shepard method uses only the zeroth order term in the Taylor expansion, then the gradient function is entirely determined by the gradients of the weight functions, and these gradients generally vanish when evaluated exactly at each grid point (i.e., the weight function is usually chosen to be symmetrical about each data point). This means that derivatives of a given order cannot be described by Shepard interpolation based only on Taylor expansions that are of lower order.

Weight functions may be defined in many different ways, but one approach that is both simple and effective involves using unnormalized weight functions  $\nu_i(\mathbf{z})$  such that

$$\nu_i(\mathbf{z}) = \frac{1}{[\sum_{\alpha=1}^{3N_{nuc}-6} (z_\alpha - z_\alpha(i))^2 + a^2]^p}. \quad (3.16)$$

Here  $p$  is a parameter that determines how quickly the weight function drops off, and  $a$  is a parameter that is chosen to avoid letting the weight function become singular at the position of each data point, while at the same time keeping the weight function sufficiently peaked near each data point that the potential and its gradients are close to their true values. (If  $a$  is chosen to be zero, Ishida and Schatz found that [21] trajectories can be ill-behaved.)

Each normalized weight function  $w_i$  is obtained by dividing the corresponding unnormalized function  $\nu_i$  by the sum of unnormalized functions. The normalized weight functions  $w_i$  have the following desirable properties:

$$\begin{aligned} w_i(\mathbf{z}) &\rightarrow 1 & (|\mathbf{z} - \mathbf{z}(i)| \rightarrow 0), \\ w_i(\mathbf{z}) &\rightarrow 0 & (|\mathbf{z} - \mathbf{z}(i)| \rightarrow \infty) \end{aligned} \quad (3.17)$$

Evidently, the normalized weight functions are highly peaked around each ab initio data point. In fact the weight function associated with each point is close to unity as long as no other points are nearby. This means that the weight function in Eq. (3.17) allows for smooth interpolation between nearby points even when the points are too far away from the geometry of interest.

One of the great virtues of the Shepard method is that if one wants to improve the quality of an existing interpolated surface, one simply adds more ab initio points wherever one wants. The points need not be located on a uniform grid, and in fact Collins and coworkers have developed procedures for automatically selecting where new points should go based on where trajectories sample the surface, and on where the sampled surface shows the largest uncertainty in the current fit. Ishida and Schatz have developed an alternative to this which uses a constrained Monte Carlo sampling procedure to select new geometries.

An example of an application of Shepard interpolation to the well known H<sub>3</sub> surface is provided by Ishida and Schatz [21]. In this study, the effect of defining different values of the parameters  $p$  and  $a$  in the weight function were considered, with  $p = 3$  and  $a = 0.03 a_0$  giving the best results. Using the trajectory method for locating new points, it was found that a surface for the H + H<sub>2</sub> reaction with an rms error on the order of 0.1 kcal/mol could be obtained with on the order of 100 ab initio points (where each point includes energy, gradients and hessian).

**3.2.2 Moving Least Squares methods.** Moving least squares (MLS) is a method which mixes together local and global aspects to the surface fitting, although in the end it is a local method. The basic idea is to use a linear least square's algorithm that is similar to the global least squares method described earlier, but which involves basis functions which are only used locally to the geometry of interest. Because the basis functions are only used nearby each ab initio point, only low order polynomial functions are needed. However the coefficients of the basis functions are also functions of geometry, and as a result there is more computational effort involved in evaluation of the surface than in the Shepard method. In the following paragraphs, I first describe the basic MLS

method, and then some recently developed ideas for circumventing its shortcomings.

In the moving least squares scheme, [43] the value  $V$  at point  $\mathbf{R}$  is represented by a linear combination of linearly independent basis functions, say  $b_k(\mathbf{R})$  ( $j = 1, \dots, n$ ):

$$V(\mathbf{R}) = \mathbf{a}^T(\mathbf{R})\mathbf{b}(\mathbf{R}) = \sum_{k=1}^K a_k(\mathbf{R})b_k(\mathbf{R}), \quad (3.18)$$

where we have used the same notation as in Eq. (3.1). However unlike that equation, the coefficients  $a_1(\mathbf{R}), a_2(\mathbf{R}), \dots, a_K(\mathbf{R})$  are allowed to depend on the coordinates  $\mathbf{R}$  used to define the PES by using coordinate dependent weight functions that we define below.

Let us define the coordinates and energy values to be interpolated as  $\mathbf{R}(j), v(j)$  ( $j = 1, 2, \dots, J$ ), where  $J$  is the number of data points. This notation is the same as was used in Section 3.1.1, and the error functional is also similar:

$$E(V) = \sum_{j=1}^N w_j(\mathbf{R})[V(\mathbf{R}) - v(i)]^2. \quad (3.19)$$

but note that we allow the weights  $w_j$  to be functions of  $\mathbf{R}$ . From the condition that  $E(V)$  be stationary with respect to variation in the parameters  $a_k(\mathbf{R})$ , the following generalized normal equations are obtained:

$$BW(\mathbf{R})B^T\mathbf{a}(\mathbf{R}) = BW(\mathbf{R})\mathbf{f}, \quad (3.20)$$

where  $W$  and  $B$  are identical to what was discussed previously. This time the weights  $w_j$  are taken to vary with  $\mathbf{R}$ , and intuitively we want to take them to be larger for data points that are close to  $\mathbf{R}$  than for points far away. This can be accomplished using the same function as in Shepard interpolation, Eq. (3.16). Because the weights depend on position, the expansion coefficients  $a_j(\mathbf{R})$  must be obtained from a solution of Eq. (3.20) for each value of  $\mathbf{R}$  where one wants to determine the potential.

To calculate the gradient of the energy at any point (as would be needed in a trajectory calculation), take the first derivative of Eq. (3.18):

$$u_s(\mathbf{R}) = \mathbf{a}_s^T(\mathbf{R})\mathbf{b}(\mathbf{R}) + \mathbf{a}^T(\mathbf{R})\mathbf{b}_s(\mathbf{R}). \quad (3.21)$$

Here  $a_s(R)$ , the derivative of  $a(R)$  with respect to the  $s$ th coordinate, is obtained by solving the following equation

$$BW(\mathbf{R})B^T\mathbf{a}_s(\mathbf{R}) = BW_s(\mathbf{R})[\mathbf{f} - B^T\mathbf{a}(\mathbf{R})], \quad (3.22)$$

The interpolated energy and gradient values at any point can be obtained by solving Eqs. (3.20) and (3.22). Both of these equations have dimensions equal to the number of basis functions, which for a triatomic molecule with terms up to quadratic included is equal to 10. More generally there are  $(N_{nucl} + 1)(N_{nucl} + 2)/2$  terms up to quadratic for an  $N_{nucl}$  atom molecule. For trajectory calculations, this solution is fairly time consuming for large  $N_{nucl}$ , so as an alternative

Ishida and Schatz [44] developed a method which they call IMLS (interpolant moving least squares) in which equations (3.20) and (3.22) (and equivalent expressions for the second derivatives given below) are used to define gradients and hessians at the positions of the ab initio data points, and then the results are used to define a Shepard interpolation as described in the previous section. This means that the linear equations only need to be solved once to determine information needed for the Shepard interpolation. This also provides a method for doing Shepard interpolation without having to determine gradient and hessians directly from the ab initio calculations.

An additional feature of the IMLS method of Ishida and Schatz is that when evaluated at the ab initio points, Eq. (3.22) reduces to the simpler result:

$$u_s(\mathbf{R}) = \mathbf{a}^T(\mathbf{R})\mathbf{b}_s(\mathbf{R}). \quad (3.23)$$

because  $\mathbf{a}_s = 0$ . This avoids the need to solve Eq. (3.22).

For the second derivatives  $u_{st}(\mathbf{R})$  at the ab initio points, we have to solve:

$$u_{st}(\mathbf{R}) = \mathbf{a}_{st}(\mathbf{R})^T\mathbf{b}(\mathbf{R}) + \mathbf{a}(\mathbf{R})^T\mathbf{b}_{st}(\mathbf{R}), \quad (3.24)$$

and

$$BW(\mathbf{R})B^T\mathbf{a}_{st}(\mathbf{R}) = BW_{st}(\mathbf{R})[\mathbf{f} - B^T\mathbf{a}(\mathbf{R})]. \quad (3.25)$$

The method of Ishida and Schatz was applied to the H<sub>3</sub> potential surface, and it was found that a surface with accuracy comparable to that obtained in the Shepard application described in the previous section could be obtained using 300-400 ab initio points. Although this is more than were used in the Shepard application, the IMLS application did not require gradient or hessian input. Since many of the highest quality ab initio methods do not provide analytical gradients, and even fewer provide hessians, this approach should prove more useful than the original Shepard method. Of course one can determine gradients and hessians numerically in the ab initio calculations, but if the experience with H<sub>3</sub> generalizes, then it will be more efficient to use IMLS than to use Shepard with numerical gradients and hessians.

## 4. Conclusion

This paper has reviewed a number of strategies for fitting and/or interpolating molecular potential energy surfaces of relevance to unimolecular and bimolecular chemical reactions. Underlying each method is the assumption that one can calculate high quality potential surface information (energies, perhaps gradients and hessians) at hundreds to thousands of geometries. This means that the goal of the fitting method is to match these results as best one can, and also to provide an efficient function for evaluating the surface at any geometry one desires. Within these constraints we found that both global and local methods can be effective, usually with somewhat different domains of convergence.

Global methods are especially good for developing low dimensional potential surfaces (1D, 2D and 3D) where the shape of the surface varies smoothly with

coordinate values. Within this group of methods, spline methods require the least thought to implement but they are useless unless a fairly dense set of points is available (at least 10 points per degree of freedom), and they are useless for extrapolating outside the range of coordinates where points are available. Morse-spline and Rotating Morse-spline methods reduce the number of points needed for one degree of freedom to four, and they enable the determination of the complete range of coordinates for that degree of freedom, while also allowing for the change in shape of the potential that occurs when dissociation occurs. Global linear least squares methods have the advantage of not requiring points be on a grid, and in favorable cases one may be able to use only points located at low energy geometries, however the choice of basis functions requires great care, and the surface may be unphysical even for geometries that are inside the range sampled. RKHS is an alternative to splines that incorporates boundary conditions and smoothness constraints directly into the interpolating function.

As an alternative to global surfaces, local surfaces have the advantage that they are more immediately accessible for application to surfaces of higher dimensionality ( $> 3$ ) and they can be used to develop surfaces based on arbitrary point locations. Of the two global methods discussed, Shepard interpolation has received more attention, and it has already been used to describe surfaces for the reactions  $\text{OH} + \text{H}_2$  and  $\text{H} + \text{CH}_4$ . The disadvantage of the Shepard method is that all the successful implementations have required ab initio Hessians in addition to gradients and energies. The Hessian (and gradient) information is not generally available from the best ab initio codes, so the range of applications of this method is somewhat limited. In view of this, the IMLS method seems attractive, as it has the potential to avoid the gradient and Hessian determination. However the IMLS method has not yet been carefully studied so the point density needed to make it useful for describing higher dimensionality surfaces is unknown.

*Acknowledgement.* This work was supported by NSF Grant CHE-9873892. I thank Larry Harding, Timothy Hollebeek, Tak-San Ho, Herschel Rabitz, T. Ishida, Abigail Dobbyn, and Jonathan Connor for their help in projects that are cited in this work.

## References

1. J. N. Murrell, S. Carter, S. C. Farantos, P. Huxley, and A. Varandas, *Molecular Potential Energy Functions* (Wiley, Chichester, 1984).
2. G. C. Schatz, Rev. Mod. Phys. **61**, 669 (1989).
3. A. J. C. Varandas, Adv. Chem. Phys. **74**, 255 (1988).
4. A. J. C. Varandas, J. Molec. Str. (Theochem) **166**, 59 (1988).
5. D. G. Truhlar, R. Steckler, and M. S. Gordon, Chem. Rev. **87**, 217 (1987).
6. A. Laganà, G. O. de Aspuru, and E. Garcia, J. Chem. Phys. **108**, 3886 (1998).
7. G. O. de Aspuru and D. C. Clary, J. Phys. Chem. A **102**, 9631 (1998).
8. F. O. Ellison, J. Am. Chem. Soc. **85**, 3540 (1963).
9. P. J. Kuntz, Ber. Bunsenges. Phys. Chem. **86**, 367 (1982).

10. K. A. Nguyen, I. Rossi, and D. G. Truhlar, *J. Chem. Phys.* **103**, 5522 (1995).
11. K. C. Thompson, M. J. T. Jordan, and M. A. Collins, *J. Chem. Phys.* **108**, 564 (1998).
12. K. C. Thompson, M. J. T. Jordan, and M. A. Collins, *J. Chem. Phys.* **108**, 8302 (1998).
13. J. Ischtwan and M. A. Collins, *J. Chem. Phys.* **100**, 8080 (1994).
14. M. J. T. Jordan, K. C. Thompson, and M. A. Collins, *J. Chem. Phys.* **102**, 5647 (1995).
15. M. J. T. Jordan, K. C. Thompson, and M. A. Collins, *J. Chem. Phys.* **103**, 9669 (1995).
16. M. J. T. Jordan and M. A. Collins, *J. Chem. Phys.* **104**, 4600 (1996).
17. K. C. Thompson and M. A. Collins, *J. Chem. Soc. Faraday Trans.* **93**, 871 (1997).
18. R. P. A. Bettens and M. A. Collins, *J. Chem. Phys.* **108**, 2424 (1998).
19. R. P. A. Bettens and M. A. Collins, *J. Chem. Phys.* **109**, 9728 (1998).
20. Y. M. Rhee, T. G. Lee, S. C. Park, and M. S. Kim, *J. Chem. Phys.* **106**, 1003 (1997).
21. T. Ishida and G. C. Schatz, *J. Chem. Phys.* **107**, 3558 (1997).
22. T. Takata, T. Taketsugu, and K. Hirao, *J. Chem. Phys.* **109**, 4281 (1998).
23. T. Ishida and G. C. Schatz, *Chem. Phys. Lett.* **298**, 285 (1998).
24. K. S. Sorbie and J. N. Murrell, *Mol. Phys.* **31**, 905 (1976).
25. R. Schinke and W. A. Lester Jr., *J. Chem. Phys.* **72**, 3754 (1980).
26. J. N. Murrell and S. Carter, *J. Phys. Chem.* **88**, 4887 (1984).
27. M. S. Fitzcharles and G. C. Schatz, *J. Phys. Chem.* **90**, 3634 (1986).
28. T.-S. Ho and H. Rabitz, *J. Chem. Phys.* **104**, 2584 (1996).
29. N. Sathyamurthy and L. M. Raff, *J. Chem. Phys.* **63**, 464 (1975).
30. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge, Cambridge, UK, 1989).
31. D. R. McLaughlin and D. L. Thompson, *Chem. Phys.* **59**, 4393 (1973).
32. J. M. Bowman, J. S. Bittman, and L. B. Harding, *J. Chem. Phys.* **85**, 911 (1986).
33. G. K. Chawla, G. C. McBane, P. L. Houston, and G. C. Schatz, *J. Chem. Phys.* **88**, 5481 (1988).
34. H. Koizumi, G. C. Schatz, and S. P. Walch, *J. Chem. Phys.* **95**, 4130 (1992).
35. J. S. Wright and S. K. Gray, *J. Chem. Phys.* **69**, 67 (1978).
36. J. M. Bowman and A. Kuppermann, *Chem. Phys. Lett.* **34**, 523 (1975).
37. B. C. Garrett, D. G. Truhlar, A. F. Wagner, and T. H. Dunning Jr., *J. Chem. Phys.* **78**, 4400 (1983).
38. C. S. Maierle, G. C. Schatz, M. S. Gordon, P. McCabe, and J. N. L. Connor, *J. Chem. Soc. Faraday Trans.* **93**, 709 (1997).
39. A. J. Dobbyn, J. N. L. Connor, N. A. Besley, P. J. Knowles, and G. C. Schatz, *Phys. Chem. Chem. Phys.* **1**, 957 (1999).
40. M. L. Dubernet and J. M. Hutson, *J. Phys. Chem.* **98**, 5844 (1994).
41. T. S. Ho, T. Hollebeek, H. Rabitz, L. B. Harding, and G. C. Schatz, *J. Chem. Phys.* **105**, 10472 (1996).
42. G. C. Schatz, A. Papaioannou, L. A. Pederson, L. B. Harding, T. Hollebeek, T. S. Ho, and H. Rabitz, *J. Chem. Phys.* **107**, 2340 (1997).
43. P. Lancaster and K. Salkauskas, *Curve and Surface Fitting, An Introduction* (Academic, London, 1986).
44. T. Ishida and G. C. Schatz, *Chem. Phys. Lett.* **314**, 369 (1999) (1999).