Tutorial on Fitting of Potential Energy Surfaces

Guillermo Ochoa de Aspuru[1] and María Luz Hernández[2]

[1] Department of Chemistry, University College London, 20 Gordon Street; WC1H 0AJ London, UK
[2] Departamento de Fisica General y de la Atmosfera, Universidad de Salamanca, Plaza de la Merced 1, 37008 Salamanca, Spain

Summary. In this paper we present a tutorial on fitting of potential energy surfaces that was given at the I European Computational Chemistry School on Molecular and Reaction Dynamics. The tutorial consists in four exercises that involve the use of a computer running under UNIX operating system, programs and subroutines for assembling and analyzing potential energy surfaces, numerical methods for fitting and graphical tools for the visualization of results. This paper explains the background of the exercises and guides the reader to the succesful completion of the exercises.

The first exercise is devoted to fitting of potential energy surfaces (PES) by interpolation methods. The second exercise explains the use of the LEPS function to fit potential energy surfaces. In the third exercise the many-body expansion method is applied to fit a three-dimensional PES for a non-symmetric three-atom reaction. The last exercise illustrates the use of bond order coordinates to represent and develop potential energy functions. It also explains the concept of many-process expansion.

1. Introduction

A potential energy surface (PES) is an analytical function that provides the electronic potential energy of a set of atoms, depending on the relative configuration of the atoms in space. In most cases, it is necessary to have a potential energy surface to perform accurate theoretical studies about the dynamics and the spectroscopy of a chemical system. The procedure to obtain a potential energy surface is, first, to carry out electronic structure (*ab initio*) calculations on the system and, second, to fit a suitable analytical function so that it acceptably reproduces the *ab initio* results. However, the fitting of good quality potential energy surfaces is not an easy task.

In the tutorial four exercises illustrating well-known, powerful yet simple techniques for fitting potential energy surfaces are discussed. The paper explains these exercises in detail, and its purpose is to introduce the inexperienced reader to some techniques for fitting, checking and plotting potential energy surfaces.

To do the exercises, the reader needs a set of files to which we refer throughout the text. These are programs, scripts and data files. The files can be freely obtained from the authors in the Web (see Appendix). The files are grouped in four directories containing the files of each exercise. The programs used in the tutorial are written in FORTRAN and in the following we assume that they are used under UNIX operating system. However, the reader is required to know FORTRAN and UNIX only at basic level. Also, we refer to well-known plotting and visualization programs, such as `xmgr`, `gnuplot` and `ghostview`.

This applications are freely available on the Internet, but the reader may use any other suitable tool.

2. Potential energy surfaces by interpolation

One of the simplest ways of assembling an analytical potential energy surface is to interpolate the energy at the desired configuration to a set of energies and configurations obtained by means of *ab initio* calculations, i. e. a set of *ab initio* points. For this purpose one can use any of the several well-known interpolation methods, explained in any good book of numerical methods [1, 2].

The interpolation methods have been used extensively to obtain potential energy surfaces; for example, with cubic splines [3], with distributed approximating functionals (DAF) [4] or with Shepard interpolation [5].

The aim of this exercise is to build a potential energy surface for the H+ClI → HCl + I reaction at $\widehat{HClI}$ equal to 180° (i. e. collinear), using the cubic spline interpolation method. The interpolation will be done to a large set of *ab initio* points.

The files needed to carry out this exercise are in the directory **exercise1**. In the following, we provide instructions, step by step, on how to complete the exercise.

I   In the **abinitio.dat** file there is a set of points of *ab initio* electronic potential energy for the H + ClI → HCl + I reaction. For simplicity we label the atoms as A=Hydrogen, B=Iodine, C=Chlorine. The first field of the **abinitio.dat** file is the **rac** distance, the second field is the **rbc** distance and the third field is the potential energy for that geometry. The distances are in Å and the energies in kcal/mol referred to zero at the H+ClI configuration.

The **abinitio.dat** file contains a 20 × 20 grid of *ab initio* points for these sets of internuclear distances:

$$r_{\text{ICl}} = \quad \text{rbc} = \quad \{1.8, 1.9, 2.0, 2.1, 2.2, 2.32087, 2.4, 2.6, 2.8,$$
$$3.0, 3.2, 3.4, 3.6, 3.8, 4., 4.5, 5., 7., 10., 20.\}$$
$$r_{\text{HCl}} = \quad \text{rac} = \quad \{0.8, 0.9, 1.0, 1.1, 1.2, 1.27455, 1.4, 1.5, 1.6,$$
$$1.7, 1.8, 2.0, 2.2, 2.6, 2.8, 3., 4., 5., 7.5, 10.\}$$

The **abinitio.ps** file contains a graphical representation of the grid of *ab initio* points. It is possible to see it using the **ghostview** program to visualize postscript files, typing the following at the UNIX prompt:

% gv abinitio.ps

The grid of points is denser in the more interesting part of the potential energy surface, i. e. along the minimum energy path from reactants to products, and more sparse in the long range, dissociative region, where both **rac** and **rbc** are large. In the set of internuclear distances we have included the equilibrium distances of both diatoms: **rbc** = 2.32087 and **rac=1.27455** (see

Table 2.1), which are important to describe properly the asymptotic configurations H+ClI and HCl+I.

II   The program map.f calculates the potential energy for a regular grid of rac and rbc coordinates by means of spline interpolation. The grid has nac points along the rac coordinate and nbc points along the rbc coordinate. Essentially, the map.f program consists in a double loop along the rac and the rbc coordinates. Inside the double loop map.f calls the subroutine vspline.f, which actually performs the interpolation on the basis of data read from the filename file. After calculating the potential energy, map.f prints out rac, rbc and pot in the standard output.

III   Editing the map.f file one can set up racmin, racmax, rbcmin, rbcmax, nac and nbc. Good values to start with are: racmin=0.8, racmax=5.0, rbcmin=1.8, rbcmax=5.0, nac=50, nbc=50 (when filling in the numbers note that the first four are real and the last two are integers).

IV   Now the subroutine vspline.f must be prepared to work with the set of *ab initio* points, stored in the abinitio.dat file, editing the vspline.f file and setting the dimension of the grid of *ab initio* points: nx=20, ny=20, and the file where they are stored: filename='abinitio.dat' (lines 15 to 17).

V   The program map.f is run using the script map.scr, which compiles, links and runs the program all at once:

% map.scr

After executing the program there should be a file called map.res containing a 50 × 50 regular grid of points and potential energies, where the fields are r_{HCl}, r_{ClI} and potential energy.

VI   It is possible to see how the potential energy surface looks like doing a three-dimensional map of the interpolated grid. The map3d.mes file is a script that allows to do so; the name of the source file (map.res) must be given in the last line of the script. Then the gnuplot program is run using this script:

% gnuplot map3d.mes

It creates a map3d.ps postscript file with a pseudo-3D image of the potential energy surface, which can be seen using ghostview:

% gv map3d.ps

VII It is also possible to do a two-dimensional contour map of the surface. As above, the last line of map2d.mes is edited as appropriate, gnuplot is run and the output is visualized:

% gnuplot map2d.mes
% gv map2d.ps

VIII To have a close-up view of a specific region of the potential energy surface it is enough to change the range for rac and rbc in map.f (step III), and repeat the steps III to VII.

IX   Using the mep.f program it is possible to draw the minimum energy path from reactants to products in the potential energy surface. The mep.f program scans the whole potential energy surface, searchs the minimum potential energy along the reaction coordinate and plots it against such a coordi-

nate. The `mep.scr` script compiles the program, links it to the `vspline.f` subroutine and runs it:

`% mep.scr`

The results are in the `mep.res` file; the first field is the reaction coordinate and the second field is the minimum potential energy. The results can be plotted using the `xmgr` program as follows:

`% xmgr mep.res`

The reaction is exothermic and there is a barrier of 1.58 kcal/mol between reactants and products. Its geometry is `rac=2.351` and `rbc=2.330`.

The spline interpolation method is a powerful and easy technique that allows to fit potential energy surfaces without much effort. In this exercise, we have obtained a 2D potential energy surface that reproduces the input *ab initio* points within 1 kcal/mol. In a three-dimensional surface it is more difficult to obtain good accuracy, specially if there are structures like wells and barriers. The main drawback of the interpolation method is that it takes a large amount of *ab initio* points, of the order of thousands or more, to have an accurate, full-dimensional potential energy surface.

Table 2.1. Spectroscopic parameters for HI, ICl and HCl (Ref. 6).

	HI	ICl	HCl
r_{eq} (Å)	1.618	2.32087	1.27455
D_e (kcal/mol)	73.6185	50.1436	106.34809
ω_e (cm^{-1})	2309.01	384.29	2990.946
$\omega_e x_e$ (cm^{-1})	39.643	1.501	52.8186
β (Å^{-1})	1.7516	1.8496	1.8685

3. Semiempirical models: the LEPS function

Most of the times it is not possible to calculate a large amount of high-quality *ab initio* points for a chemical reaction, in particular if it involves more than three atoms or any of the atoms is heavy as is for several reactions of practical interest. For such cases it is necessary to have global methods for the representation of the potential energy surface. Global methods have far less adjustable parameters than local methods and, as a consequence, are less flexible. However, global methods can be successfully used with fewer *ab initio* points, and they give a reasonable representation of the potential energy surface even in regions for which no data exist.

LEPS potential energy surfaces are based on the Valence Bond theory. The starting point of this theory, and therefore of the LEPS method, is the Heitler and London treatment of the hydrogen molecule. The energies of the ground singlet state and the excited triplet state of two hydrogen atoms A and B are:

$$^1E_{AB} = \frac{Q_{AB} + J_{AB}}{1 + S_{AB}} \qquad ^3E_{AB} = \frac{Q_{AB} - J_{AB}}{1 - S_{AB}} \qquad (3.1)$$

where Q_{AB}, J_{AB} and S_{AB} are the Coulomb, exchange and overlap integrals, respectively.

$$\begin{aligned}
Q_{AB} &= \langle \phi_A(1)\phi_B(2)|\hat{H}|\phi_A(1)\phi_B(2)\rangle \\
J_{AB} &= \langle \phi_A(1)\phi_B(2)|\hat{H}|\phi_A(2)\phi_B(1)\rangle \\
S_{AB} &= \langle \phi_A(1)|\phi_B(2)\rangle\langle \phi_A(1)|\phi_B(2)\rangle
\end{aligned}$$

London extended the Valence Bond treatment to the H_3 molecule [7]. The LEP (London-Eyring-Polanyi) function that represents the potential energy surface of the H_3 molecule is [8]:

$$V(r_{AB}, r_{BC}, r_{AC}) = Q_{AB} + Q_{BC} + Q_{AC} \qquad (3.2)$$

$$- \left[\frac{(J_{AB} - J_{BC})^2}{2} + \frac{(J_{BC} - J_{AC})^2}{2} + \frac{(J_{AC} - J_{AB})^2}{2} \right]^{1/2}$$

where the overlap integrals are neglected (this greatly simplifies the equations) [9]. The first important conclusion is that the energy of a triatomic system does not equal the sum of the energies of the three diatomic couples. Sato proposed to mimic the effect of using the accurate overlap by introducing a parameter. The new LEPS (London-Eyring-Polanyi-Sato) function is [10, 11]:

$$V(r_{AB}, r_{BC}, r_{AC}) = Q_{AB} + Q_{BC} + Q_{AC} - \left[\frac{1}{2} \left(\frac{J_{AB}}{1 + S_{AB}} - \frac{J_{BC}}{1 + S_{BC}} \right)^2 \right.$$

$$\left. + \frac{1}{2} \left(\frac{J_{BC}}{1 + S_{BC}} - \frac{J_{AC}}{1 + S_{AC}} \right)^2 + \frac{1}{2} \left(\frac{J_{AC}}{1 + S_{AC}} - \frac{J_{AB}}{1 + S_{AB}} \right)^2 \right]^{1/2} \qquad (3.3)$$

where S_{AB}, S_{BC} and S_{AC} are the Sato parameters.

The LEPS formula was derived for the H_3 molecule but it is currently used to represent the potential energy surface of any triatomic system, provided it is regarded as an empirical function with adjustable parameters [12]. In fact, in such case the Sato parameters are not derived from theory, but fitted arbitrarily so that the potential energy surface mimics experimental or *ab initio* information.

The terms $Q_{ij}/(1 + S_{ij})$ and $J_{ij}/(1 + S_{ij})$ ($ij = AB, BC, AC$) in Eq. (3.3) are calculated using the energies of the singlet and triplet states of the related diatoms:

$$\frac{Q_{ij}}{1 + S_{ij}} = \frac{1}{2} \left({}^1E_{ij} + \frac{1 - S_{ij}}{1 + S_{ij}} {}^3E_{ij} \right) \qquad (3.4)$$

$$\frac{J_{ij}}{1 + S_{ij}} = \frac{1}{2} \left({}^1E_{ij} - \frac{1 - S_{ij}}{1 + S_{ij}} {}^3E_{ij} \right) \qquad (3.5)$$

The energies of the singlet and triplet states are obtained using empirical potential energy curves, such as the Morse function [13] for the singlet:

$$^1E = D_e \left[e^{-2\beta(r-r_e)} - 2e^{-\beta(r-r_e)} \right] \tag{3.6}$$

and the anti-Morse function for the triplet:

$$^3E = \frac{1}{2} D_e \left[e^{-2\beta(r-r_e)} + 2e^{-\beta(r-r_e)} \right] \tag{3.7}$$

where r_e, D_e and β are the equilibrium distance, the dissociation energy and the Morse parameter of the diatom, respectively.

Substituting Eqs. (3.6) and (3.7) in Eqs. (3.4) and (3.5) it is possible to evaluate easily the potential energy by means of Eq. (3.3).

The purpose of this exercise is to fit a potential energy surface for N+N_2 using the LEPS method. The parameters featured in the Morse and anti-Morse functions can be derived from spectroscopic data of the $N_2(X^1\Sigma_g^+)$ molecule. This well-known information is reported in Table 3.1 where the β parameter defined as [14]:

$$\beta = \omega_e \left(\frac{\pi\mu}{D_e} \right)^{1/2} \tag{3.8}$$

is used.

Table 3.1. Spectroscopic parameters for N_2 (Ref. 6).

	$N_2(X^1\Sigma_g^+)$
μ (a.m.u.)	7.00335
r_{eq} (Å)	1.0977
D_e (kcal/mol)	228.41
ω_e (cm^{-1})	2358.57
$\omega_e x_e$ (cm^{-1})	14.324
β (Å^{-1})	2.689

The potential energy surface of N+N_2 must have a barrier between reactants and products of 35.7 kcal/mol at collinear configuration [15]. Therefore, the only remaining thing is to find the Sato parameters that satisfy this condition. Of course, since N+N_2 is a fully symmetric system, only one parameter must be fitted because $S_{AB} = S_{BC} = S_{CA} = S_{NN}$.

The **leps.f** file in the **exercise2** directory contains the subroutine with the LEPS potential energy surface. The lines containing the data for the N+N_2 reaction are in the first part of the subroutine. All the data have been set up as in Table 3.1, except for the **SAB**, **SBC** and **SAC** variables (the Sato parameters), which are set up to 0.0.

The **mep.f** program calculates the fixed ϕ (variable **phi**) potential along the minimum energy path from reactants to products. The script **mep.scr** compiles **leps.f** and **mepe.f**, links them and runs the program. Executing the script
% **mep.scr**

generates a `mep.res` file containing the energy profile along the minimum energy path. The minimum energy path can be plotted with `xmgr`:

`% xmgr mep.res`

The potential is zero for reactants and products, and there is a barrier in between, but it does not have the right height.

To get the barrier height right, it is necessary to edit the `leps.f` file, to change the values of `SAB`, `SBC` and `SAC`, and to calculate again the minimum energy path. The Sato parameter is tuned until the desired barrier height is obtained. One can choose either positive or negative real numbers. Note that $S_{NN} = -1.0$ must not be used because it gives a singularity (i.e. division by zero) in Eq. (3.3).

Once the right Sato parameter has been found, one can calculate minimum energy paths at different, smaller angles by editing the `mep.f` program and changing the angle `phi` at pleasure ($0 \leq$ `phi` $\leq \pi$). It is worth finding out how the barrier height depends on `phi`. It is also interesting to find the geometry of the transition state between reactants and products and to devise a method to be able to tune also that geometry in the LEPS potential energy surface.

4. The many-body expansion

The many-body expansion (MBE), also known as Sorbie-Murrell method, is a global method to fit potential energy surfaces for reactions involving, in principle, any number of atoms [16, 17, 18]. The general expression of a many-body expansion for a three-atom system is given as a sum of one-body, two-body and three-body terms:

$$
\begin{aligned}
V(r_{AB}, r_{BC}, r_{AC}) \;=\; & V_A^{(1)} + V_B^{(1)} + V_C^{(1)} + \\
& V_{AB}^{(2)}(r_{AB}) + V_{BC}^{(2)}(r_{BC}) + V_{AC}^{(2)}(r_{AC}) + \\
& V_{ABC}^{(3)}(r_{AB}, r_{BC}, r_{AC}) \qquad (4.1)
\end{aligned}
$$

The one-body terms, $V_A^{(1)}$, $V_B^{(1)}$ and $V_C^{(1)}$, are the electronic energies of the atoms in the dissociation configurations. Since we are going to deal only with an electronic ground state potential energy surface we can set up these terms to zero:

$$
V_A^{(1)} = 0; \quad V_B^{(1)} = 0; \quad V_C^{(1)} = 0 \qquad (4.2)
$$

The two-body terms are the potential curves of each couple of atoms. In this case we choose Morse functions [13] for the diatomic potentials:

$$
\begin{aligned}
V_{ij}^{(2)}(r_{ij}) \;=\; & D_{ij}\left\{1 - \exp\left[-\beta_{ij}\left(r_{ij} - r_{eqij}\right)\right]\right\}^2 - D_{ij} \qquad (4.3) \\
ij \;=\; & AB, BC, AC
\end{aligned}
$$

where D_{ij} is the dissociation energy, β_{ij} is the Morse parameter (see Eq. (3.8)), r_{ij} is the internuclear distance and r_{eqij} is the equilibrium internuclear distance.

According to this expression the potential is zero at the dissociation points (i. e. A+B, B+C and A+C).

The three-body term accounts for the three-body interactions of the potential energy surface. It depends on the three internuclear coordinates.

$$V^{(3)}_{ABC}(r_{AB}, r_{BC}, r_{AC}) = \mathcal{S}(r_{AB}, r_{BC}, r_{AC})P_6(r_{AB}, r_{BC}, r_{AC}) \qquad (4.4)$$

In this case we give this term as a sixth-degree polynomial in r_{AB}, r_{BC} and r_{AC}, represented in short by $P_6(r_{AB}, r_{BC}, r_{AC})$.

$$P_6(r_{AB}, r_{BC}, r_{AC}) = \sum_{l=0}^{6}\sum_{m=0}^{6}\sum_{n=0}^{6} c_{lmn}\rho_{AB}^l\rho_{BC}^m\rho_{AC}^n \qquad (4.5)$$
$$l + m + n \le 6$$
$$\rho_{ij} = r_{ij} - r_{ij}^\circ \qquad ij = AB, BC, AC$$

where r_{ij}° are reference bond lengths. The three-body term must become zero when any of the internuclear distances becomes large; at such configurations the two-body terms take over. To make sure that the three-body term becomes zero at asymptotic configurations, it is multiplied by a damping function, $\mathcal{S}(r_{AB}, r_{BC}, r_{AC})$, which is one in the strong interaction region and becomes zero at the asymptotic configurations. The expression of the damping function is given in terms of three hyperbolic tangents, one for each internuclear coordinate.

$$\mathcal{S}(r_{AB}, r_{BC}, r_{AC}) = s(r_{AB})s(r_{BC})s(r_{AC}) \qquad (4.6)$$
$$s(r_{ij}) = \frac{1 - \tanh(\gamma_{ij}\rho_{ij})}{2} \qquad ij = AB, BC, AC$$

The MBE method has been used extensively over the years to assemble PES for three- and four-atom reactions [19, 20]. Also, there exists the double many-body expansion (DMBE) method, in which the interaction energy is partitioned into Hartree-Fock and dynamical correlation contributions [18, 21].

This exercise consists in assembling a global, three-dimensional potential energy surface for the $H + ICl \rightarrow HI + Cl, HCl + I$ reaction using the MBE method. The files needed for this purpose are in the directory exercise3.

The mbepot.f file contains the subroutine of the MBE for a three-atom system. The first data lines set up the dissociation energies, equilibrium distances and Morse parameters of the system. The data for HICl have been already filled in the file (see Table 2.1). In mbepot.f the couples HI, ICl and HCl are labelled as AB or 1, BC or 2 and AC or 3, respectively. The one-body terms (v1(i) i=1,2,3) are zero, as explained above. The two-body terms (v2(i) i=1,2,3) are given as Morse potentials [13], which use the information provided in the data lines. The next part deals with r_{ij}° (r0(i) i=1,2,3), γ_{ij} (gamma(i) i=1,2,3) and the damping functions (damping(i) i=1,2,3 and dampingtot). Since the optimization of the r_{ij}° distances is a lengthy task that requires some trial and error iterations, they have been already optimized. Their values are:

$$r^{\circ}_{AB} = 2.5\, r_{eqAB}$$
$$r^{\circ}_{BC} = 1.6\, r_{eqBC}$$
$$r^{\circ}_{AC} = 3.3\, r_{eqAC}$$

The coefficients of the damping functions have been set up as

$$\gamma_{AB} = \gamma_{BC} = \gamma_{AC} = 2.0$$

With these parameters, the damping functions $s(r_{ij})$ and $S(r_{AB}, r_{BC}, r_{AC})$ are straightforwardly calculated as in Eqs. (4.6).

The trickiest part of the potential is the three-body term (v3). To save computation time, this term is calculated only if the damping function has any significant value (i. e. larger than 10^{-6} in this case). Although the calculation of the polynomial could be packed up in a set of nested loops, it has been written explicitly to allow the reader to see easily the structure of the polynomial (also, writing in this way speeds up the calculation, but it takes many more lines of code).

Finally, all the contributions are added up to have the final potential energy (vpot).

The coefficients of the polynomial that gives the three-body term are c(i) i=1,2,...,84, and their values are given in a data sentence at the end of the file. The values of these coefficients are set up to zero and the objective is to work out the coefficients c that will give a good potential energy surface for HICl.

In the abinitiototal.dat there is a set of 768 *ab initio* points for HICl. The internuclear distances (three first fields: r_{HI}, r_{ICl} and r_{HCl}) are in Å and the potential energy (fourth field) is in kcal/mol referred to zero at the full dissociation configuration (H+I+Cl). The MBE surface must be fitted to these points.

Since the one- and two-body terms are already done, the remaining task is to obtain the three-body term. The procedure to follow is to obtain the three-body energies from the *ab initio* points and then fit the c coefficients to them. The to3body.f program allows to obtain the three-body component of the *ab initio* potential energies. This program reads nlines=768 lines from the standard input and, for every line it calculates first the two-body energies and then the three-body energy as follows:

$$\text{three-body energy} = \text{total energy} - \sum \text{two-body energy}$$

and writes the result in the standard output. The program is then compiled and executed:

```
% f77 -o to3body.e to3body.f
% to3body.e < abinitiototal.dat > abinitio3body.dat
```

Now the abinitio3body.dat file contains the 768 geometries and *ab initio* three-body energies.

The fitting of the three-body function is done with the least squares technique [1] in the lfit.f program. This program fits the coefficients of the three-body

function, which is given in the FUNCS subroutine of lfit.f. The program is run with the data just generated:

```
% f77 -o lfit.e lfit.f
% lfit.e < abinitio3body.dat > lfit.res
```

The lfit.res file contains the input data that it has read, the coefficients and the input data compared one by one to the result of the fitting. The coefficients must be retrieved and inserted appropriately in the data line of the mbepot.f file, so that the subroutine is complete and ready to be used.

The next step is to evaluate how good (or bad) is the MBE potential energy surface. For this purpose one has to write a small program that reads the 768 *ab initio* points from the file abinitiotot.dat (distances rab, rbc and rac and *ab initio* potential energy). Then the program calls the mbepot.f subroutine and calculates the MBE potential at such a configuration. The program calculates the difference (in absolute value) between the *ab initio* potential energy and the MBE potential energy at every point and the maximum and the average difference. With this information one has to decide if the potential energy surface is of acceptable quality or not. Should it not be good enough, one has to think of ways to further improve the agreement.

5. The bond order cordinates and the many-process expansion

Recently new methods to fit PES have been proposed. These methods use bond order coordinates instead of physical coordinates and are based on the many-process expansion. So far these methods have been used for reactions involving three and four atoms. The aim of this exercise is to get started with the bond order coordinates and with the many-process expansion.

The bond order (BO) coordinate for two atoms is related to the physical coordinate as follows:

$$n = \exp\left[-\beta(r - r_{\text{eq}})\right] \tag{5.1}$$

where n is the BO coordinate, r is the physical coordinate (internuclear distance), β is the Morse parameter (see Eq. (3.8)), and r_{eq} is the equilibrium internuclear distance.

If the potential energy curve of any diatomic molecule is plotted against the BO coordinate, it has a much simpler shape than if it is plotted against the physical coordinate. The most obvious case is the Morse potential. The Morse potential of a diatomic molecule in physical and bond order coordinates are:

$$V(r) = D\left\{1 - \exp\left[-\beta(r - r_{\text{eq}})\right]\right\}^2 \tag{5.2}$$

$$V(n) = D(1 - n)^2 \tag{5.3}$$

where D is the dissociation energy. When expressed in BO coordinates, the Morse potential is a very simple second degree polynomial in n.

To illustrate this point, it is interesting to plot the Morse potentials of three diatomic molecules as functions of the internuclear distances and also as functions of the BO coordinates. The program phys1D.f in the exercise4 directory calculates the Morse potential curves of the HI, ICl and HCl molecules. In the program the atoms are labelled as A=H, B=I and C=Cl. The first lines of the program provide the data of the molecules. After the data lines the program sets upper and lower limits for the AB internuclear distace and, by means of a loop, it calculates the Morse potential energy of AB at intervals of 0.01 Å. The same is done for BC and AC. The program sends the results to the standard output.

The program is compiled and executed and the three potential curves are plotted:

```
% f77 -o phys1D.e phys1D.f
% phys1D.e > phys1D.res
% xmgr phys1D.res
```

To compare with these Morse curves in physical coordinates one has to calculate the same curves but in BO coordinates. This is done with the bord1D.f program, which has an structure very similar to phys1D.f. However, in this case the lower and upper limits of the BO coordinate are the same for the three curves: boabmin=0.0 and boabmax=2.3. As before, one compiles and executes the program and plots the results:

```
% f77 -o bord1D.e bord1D.f
% bord1D.e > bord1D.res
% xmgr bord1D.res
```

The curves have four characteristics that are worth noticing:

– The curves in BO coordinates have a simpler shape than the ones in physical coordinates.
– All the curves have the minimum at the same point. In fact, the minimum of the Morse curve of *any* diatomic molecule is always at $n = 1$.
– In the curves plotted in physical coordinates the dissociation potential of A+B is reached asymptotically at large distances ($r \to \infty$) and the repulsive region is at short distances. On the contrary, in the curves plotted in BO coordinates the dissociation potential is at the origin of coordinates and the repulsive potential is at large values of n.
– The physical domain is infinite, $r \in [0, \infty)$, whereas the BO domain is finite, $n \in (0, \exp(\beta r_{\mathrm{eq}})]$.

These considerations lead to the first and clearest advantage of using BO coordinates for modelling potential energy functions: the representations in BO coordinates are simpler, and therefore simpler functions (with less adjustable parameters) are required to model potentials.

The BO coordinates have been used within the many-body expansion method explained in the previous section. When using BO coordinates, the two- and three-body terms are [22]:

$$V_{ij}^{(2)}(n_{ij}) \;=\; D_{ij} \sum_{k=0}^{N} a_{ij}^{(k)} n_{ij}^{k} \qquad ij = \mathrm{AB, BC, AC} \qquad (5.4)$$

$$V_{\text{ABC}}^{(3)}(n_{\text{AB}}, n_{\text{BC}}, n_{\text{AC}}) = \sum_{l=i}^{M} \sum_{m=j}^{M} \sum_{n=k}^{M} c_{lmn} n_{\text{AB}}^{l} n_{\text{BC}}^{m} n_{\text{AC}}^{n} \qquad (5.5)$$

$$l + m + n \leq M$$

In this formulation the damping function (Eq. (4.6)) is not needed because the three-body term is zero when either n_{AB} or n_{BC} or n_{AC} is zero (i. e. at atom+diatom and at full dissociation configurations). This formulation does not require the reference coordinates (r_{ij}°) either.

The next step of the exercise is to make representations of a three-atom PES in physical and in BO coordinates, and to compare both outcomes. The **phys2D.f** program calculates the potential energy of the HICl system in three different two-dimensional grids of configurations in the space of physical coordinates. In the program the atoms are labelled as before (A=H, B=I and C=Cl) so that the internuclear distances are **rab**, **rbc** and **rac**. Using a LEPS PES for this system, the program calculates three grids of 60×60 points: **rab** versus **rbc**, **rbc** versus **rac**, and **rac** versus **rab**. The calculations are done at a fixed internuclear angle **phi** (the same for the three grids), which is set up at the beginning of the program (note that $0 \leq$**phi**$\leq \pi$).

At this point, one sets the **phi** angle at a large value (e.g. **phi=3.14**) in **phis2d.f**, and then compiles and runs the program:

```
% f77 -o phys2D.e phys2D.f
% phys2D.e
```

The program generates three output files with the three sets of points: **rab_rbc.res**, **rbc_rac.res** and **rac_rab.res**.

To do a contour map of the potential energy of any of the files (e.g. **rbc_rac.res**), one has to edit the last line of the **map2d.mes** file, writing the name of the source file, and then run **gnuplot**:

```
% gnuplot map2d.mes
```

The postscript file generated, **map2d.ps**, can be seen with **ghostview**:

```
% gv map2d.ps
```

In the contour map one can notice the entrance channel of reagents, the intermediate barrier or well and the exit channel of products.

The **bord2D.f** program is similar to **phys2D.f**, but it calculates the potential energy in the space of physical coordinates. Once the **phi** angle is set up at the same value as before in **phys2D.f**, one compiles and runs the program.

```
% f77 -o bord2D.e bord2D.f
% bord2D.e
```

The output is in three files: **boab_bobc.res**, **bobc_boac.res** and **boac_boab.res**. To see the contents of the file, as above, one edits the last line of the **map2d.mes** and **map3d.mes** files, writing the name of the source file, and then plots the contour map in BO coordinates:

```
% gnuplot map2d.mes
% gv map2d.ps
% gnuplot map3d.mes
% gv map3d.ps
```

The BO contour map looks very different from the contour map in physical coordinates. The origin of coordinates corresponds to the full dissociation configuration H+I+Cl, and the minimum energy path pivots around this point. The repulsive zone of the potential is in the most external region. In fact, in the BO representation the whole potential energy surface can be regarded as a one-dimensional curve that revolves around the origin of coordinates. This important feature can be also noticed in the three-dimensional plots (`map3d.ps`).

One can repeat again the representations in BO coordinates but at smaller angles; for example, `phi=2.1`, `phi=1.5` and `phi=1.0`. The shape of the potential energy surface changes when the orientational angle gets smaller.

The BO coordinates can be transformed into polar BO coordinates or rotating BO (ROBO) coordinates. For example, for the couple of BO coordinates n_{ij} and n_{jk} we have that:

$$\rho_j = \left(n_{ij}^2 + n_{jk}^2\right)^{1/2} \tag{5.6}$$

$$\alpha_j = \arctan\left(\frac{n_{ij}}{n_{jk}}\right) \tag{5.7}$$

The ρ_j and α_j coordinates are perpendicular in the BO space [23]. The ρ_j coordinate corresponds to the vibrational coordinate of the system and the α_j coordinate corresponds to the reaction coordinate. The intervals of definition of these coordinates are:
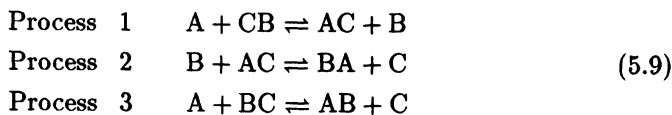
$$0 \leq \quad \rho_j \quad \leq \left[\exp\left(2\beta_{ij}r_{eqij}\right) + \exp\left(2\beta_{jk}r_{eqjk}\right)\right]^{1/2}$$
$$0 \leq \quad \alpha_j \quad \leq \pi/2$$

Using the ROBO coordinates, it is possible to fit the potential energy surface of the $A + BC \rightarrow AB + C$ process with a function as follows:

$$V_{B}^{ROBO}\left(\rho_B, \alpha_B; \phi_B\right) = \mathcal{D}_B\left(\alpha_B; \phi_B\right) \mathcal{F}_B\left(\rho_B; \alpha_B, \phi_B\right) \tag{5.8}$$

where ϕ_B is the angle $\widehat{ABC}$. The function $\mathcal{D}_B\left(\alpha_B; \phi_B\right)$ gives the fixed ϕ_B evolution of the minimum energy path along α_B when the system goes from the reactant to the product configuration. The function $\mathcal{F}_B\left(\rho_B; \alpha_B, \phi_B\right)$ represents the cut (normalised to -1 at the minimum) of the fixed ϕ_B potential energy surface along ρ_B at a given value of α_B.

In a three-atom system there are three possible rearrangement channels:

$$
\begin{array}{lll}
\text{Process} \quad 1 & A + CB \rightleftharpoons AC + B \\
\text{Process} \quad 2 & B + AC \rightleftharpoons BA + C & \tag{5.9}\\
\text{Process} \quad 3 & A + BC \rightleftharpoons AB + C
\end{array}
$$

Therefore the global potential energy surface of the system can be expressed as a linear combination of three ROBO functions, one for each rearrangement. This is called many-process expansion.

$$
\begin{aligned}
V\left(r_{AB}, r_{BC}, r_{CA}\right) &= w_A\left(\phi_A\right) V_A^{\text{ROBO}}\left(\rho_A, \alpha_A; \phi_A\right) \\
&+ w_B\left(\phi_B\right) V_B^{\text{ROBO}}\left(\rho_B, \alpha_B; \phi_B\right) \\
&+ w_C\left(\phi_C\right) V_C^{\text{ROBO}}\left(\rho_C, \alpha_C; \phi_C\right)
\end{aligned}
$$

The ROBO functions are multiplied by suitable weighting functions, in order to give more importance to the contribution that better represents the global potential energy surface at any given configuration of the atoms and to leave out the unsuitable contributions.

The result of combining the use of ROBO coordinates and the many-process expansion has led to the LAGROBO (Largest Angle Generalization of ROtating Bond Order) method to fit potential energy surfaces [25]. The LAGROBO method has been used for three- [24, 26, 27] and four-atom [28] reactions.

A. Appendix

The files of the tutorial can be freely obtained from the authors in the Web page `http://www3.usal.es/~dinmol/docencia.html`. Please note that the software is provided "as is", and without any further support or help from the authors. This software can be freely distributed, providing the source is acknowledged and reference to this paper is made.

The file supplied, `tutorial.tar.gz`, must be decompressed and "untarred":
`% gunzip tutorial.tar.gz`
`% tar xvf tutorial.tar`

Four directories will be created (`exercise1`, `exercise2`, `exercise3` and `exercise4`), containing the files of each exercise.

References

1. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery; *Numerical recipes in FORTRAN* 2nd Ed.; Cambridge University Press (1992). W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery; *Numerical recipes in C* 2nd Ed.; Cambridge University Press (1992).
2. J. H. Ahlberg, E. N. Nilson, J. L. Walsh; *The theory of splines and their applications*; Academic Press (1967).
3. N. Sathyamurthy, L. M. Raff, J. Chem.Phys. **63**, 464 (1975); L. Beneventi, P. Casavecchia, G. G. Volpi, C. C. K. Wong, F. R. W. McCourt, J. Chem. Phys. **98**, 7926 (1993); C. S. Maierle, G. C. Schatz, M. S. Gordon, P. McCabe, J. N. L. Connor, J. Chem. Soc. Faraday Trans. **93**, 709 (1997); D. H. Mordaunt, H. Flothmann, M. Stumpf, H. M. Keller, C. Beck, R. Schinke, K. Yamashita, J. Chem. Phys. **107**, 6603 (1997); C. Minichino, G. A. Voth, J. Phys. Chem. B **101**, 4544 (1997); D.

Oberlin, H. A. Scheraga, J. Comp. Chem. **19**, 71 (1998); D. Lauvergnat, D. C. Clary, J. Chem. Phys. **108**, 3566 (1998).

4. D. K. Hoffman, A. Frishman, D. J. Kouri, Chem. Phys. Lett. **262**, 393 (1996); A. Frishman, D. K. Hoffman, D. J. Kouri, J. Chem. Phys. **107**, 804 (1997); D. K. Hoffman, G. W. Wei, D. S. Zhang, D. J. Kouri, Phys. Rev. E **57**, 6152 (1998).

5. J. Ischtwan, M. A. Collins, J. Chem. Phys. **100**, 8080 (1994); M. J. T. Jordan, K. C. Thompson, M. A. Collins, J. Chem. Phys. **102**, 5647 (1995); M. J. T. Jordan, K. C. Thompson, M. A. Collins, J. Chem. Phys. **103**, 9669 (1995); M. J. T. Jordan, M. A. Collins, J. Chem. Phys. **104**, 4600 (1996). T. Ishida, G. C. Schatz, J. Chem. Phys. **107**, 3558 (1997).

6. K. P. Huber, G. Herzberg; *Molecular Spectra and Molecular Structure. IV Constants of Diatomic Molecules* (Van Nostrand, 1979).

7. F. London, Z. Elektrochem. **35**, 552 (1929).

8. H. Eyring, M. Polanyi, Z. Phys. Chem. Abt. B **12**, 279 (1931).

9. H. Eyring, S. H. Lin; in *Physical Chemistry: An Advanced Treatise, Vol. 6A, Kinetics of Gas Reactions*; H. Eyring, D. Henderson, W. Jost Eds. (Academic Press, 1974).

10. S. Sato, J. Chem. Phys. **23**, 592 (1955); S. Sato, J. Chem. Phys. **23**, 2465 (1955).

11. P. J. Kuntz, E. M. Nemeth, J. C. Polanyi, S. D. Rosner, C. E. Young, J. Chem. Phys. **44**, 1168 (1966).

12. M. Baer, J. Chem. Phys. **60**, 1057 (1974).

13. P. M. Morse, Phys. Rev. **34**, 57 (1929).

14. G. Herzberg; *Molecular Spectra and Molecular Structure. I Spectra of Diatomic Molecules*, 2nd ed. (Van Nostrand, 1950).

15. A. Laganà, E. García, L. Cicarelli, J. Phys. Chem. **91**, 312 (1987).

16. K. S. Sorbie, J. N. Murrell, Mol. Phys. **29**, 1387 (1975).

17. J. N. Murrell, S. Carter, S. C. Farantos, P. Huxley, A. J. C. Varandas; *Molecular potential energy functions* (Wiley, Chichester, 1984).

18. A. J. C. Varandas, Adv. Chem. Phys. **74**, 255 (1988).

19. J. N. Murrell, A. J. C. Varandas, Mol. Phys. **57**, 415 (1986); B. H. Wells, S. Wilson, Mol. Phys. **65**, 1363 (1988); J. N. Murrell, Int. J. Quant. Chem. **37**, 95 (1990); A. Laganà, M. Dini, E. Garcia, J. M. Alvariño, M. Paniagua, J. Phys. Chem. **95**, 8379 (1991); X. H. Liu, J. N. Murrell, J. Chem. Soc. Faraday Trans. **87**, 435 (1991); L. B. Harding, R. Guadagnini, G. C. Schatz, J. Phys. Chem. **97**, 5472 (1993); M. Simonson, N. Markovic, S. Nordholm, B. J. Persson, Chem. Phys. **200**, 141 (1995); M. Gonzalez, J. Hijazo, J. J. Novoa, R. Sayos, J. Chem. Phys. **108**, 3168 (1998).

20. G. C. Schatz, H. Elgersma, Chem. Phys. Lett. **73**, 21 (1980); T. G. Wei, R. E. Wyatt, J. Phys. Chem. **97**, 13580 (1993); A. Aguado, C. Suarez, M. Paniagua, J. Chem. Phys. **101**, 4004 (1994); M. A. Ter Horst, G. C. Schatz, L. B. Harding, J. Chem. Phys. **105**, 558 (1996); A. J. C. Varandas, W. Wang, Chem. Phys. **215**, 167 (1997); A. J. C. Varandas, H. G. Yu, Mol. Phys. **91**, 301 (1997); J. Sogas, M. Alberti, X. Gimenez, R. Sayos, A. Aguilar, J. Phys. Chem. A **101**, 8877 (1997).

21. A. J. C. Varandas, Mol. Phys. **53**, 1303 (1984); A. J. C. Varandas, Mol. Phys. **57**, 387 (1986); A. J. C. Varandas, A. A. C. C. Pais, Mol. Phys. **65**, 843 (1988); A. J. C. Varandas, J. Brandao, L. A. M. Quintales J. Phys. Chem. **92**, 3732 (1988); G. C. Lynch, R. Steckler, D. W. Schwenke, A. J. C. Varandas, D. G. Truhlar, B. C. Garrett, J. Chem. Phys. **94**, 7136 (1991); A. J. C. Varandas, A. I. Voronin, Mol. Phys. **85**, 497 (1995). A. J. C. Varandas, A. I. Voronin, J. Phys. Chem. **99**, 15846 (1995).

22. E. Garcia. A. Laganà, Mol. Phys. **56**, 629 (1985).

23. A. Laganà, J. Chem. Phys. **95**, 2216 (1991); A. Laganà, G. Ferraro, E. García, O. Gervasi, A. Ottavi, Chem. Phys. **168**, 341 (1992).

24. E. García, A. Laganà, J. Chem. Phys. **103**, 5410 (1995).

25. A. Laganà, G. Ochoa de Aspuru, E. García, J. Chem. Phys. **108**, 3886 (1998).

26. A. Laganà, G. Ochoa de Aspuru, E. García, J. Phys. Chem. **99**, 17139 (1995).
27. P. Casavecchia, A. Laganà, G. Ochoa de Aspuru, G. Lendvay, M. Alagia, N. Balucani, E. H. Van Kleef, G. G. Volpi, Chem. Phys. Lett. **258**, 323 (1996).
28. G. Ochoa de Aspuru, D. C. Clary, J. Chem. Phys. A **102**, 9631 (1998); G. Ochoa de Aspuru, D. C. Clary, "*Theoretical study of the* OH + HCl *reaction on a new potential energy surface*"; in preparation.