Computational and Theoretical Chemistry 963 (2011) 76–84

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

High accuracy interpolation of diatomic pseudo-potentials with high order Hermite splines

David Schmidling *

Department of Physics and Technology, Bronx Community College of the City University of New York, West 181 St & University Avenue, Bronx, NY 10453, USA

ARTICLE INFO

Article history:
Received 22 August 2010
Received in revised form 23 August 2010
Accepted 3 October 2010
Available online 12 October 2010

Keywords:
Hermite spline
Interpolation
Diatomic
Pseudo-potential

ABSTRACT

Pseudo-potentials for diatomic molecules can be interpolated to high accuracy with a high order Hermite spline via the function $Y = -[(E_S - E_U)/E_U]E_E\,[(1-z)/2]^2$ where E_S is the energy of the separated atoms, E_U the united atom energy, E_E the electronic energy and $z = (R - R_a)/(R + R_a)$, where R is the inter-nuclear distance and R_a is an adjustable parameter. Both Y and its derivative with respect to z are very smooth, which facilitates the interpolation process.

Limiting laws are used at each end to provide spline constraints and also to replace the spline in the regions beyond the data, giving a hybrid result. The low R limiting law is of the form $E_E = E_U + \Sigma_{i=2} A_i R^i$. The A coefficients are obtained by least-squares. The high R limiting law is of the form $E = E_S - \Sigma C_i R^{-i}$; the C coefficients can be obtained from perturbation theory, least-squares or in combination.

Test data with a pseudo-potential similar to that of most diatomic molecules with added noise was interpolated with an order 12 spline at data intervals of 0.05, 0.1 and 0.2 Bohr. The resulting maximum errors in E_E and vibrational energy levels were no more than several times a noise level of 10^{-17}, 10^{-15} and 10^{-10}, respectively. The lowest noise level at which the maximum errors were acceptable decreased with data range, spline order and smoothness of the function. The interpolation procedure was successfully applied to the H_2 X, C and a states, for which abundant data in the form of tables of energy and gradient are available.

1. Introduction

Interpolation is a process for obtaining an approximate value of a function $f(x)$ at arbitrary x intermediate between x values where the value of the function is given. This process must be applied to a pseudo-potential constructed from *ab initio* calculations of the energy of a molecule with fixed nuclei, since (a) the true analytic form of the pseudo-potential is not known and (b) the calculations at each point are expensive. The calculated points themselves are imprecise due to convergence cutoffs and finite precision arithmetic.

Progress has been made in solving the Schrödinger equation for diatomic molecules with nuclei and electrons treated together [1–4]. Currently though, the pseudo-potential concept is retained in most calculations, which start with a Born–Oppenheimer [5] approximation, and it is the interpolation of results from this first step that is addressed here.

Interpolating functions can be global or local in nature. Global functions seek to approximate the energy of a diatomic molecule over the entire range of inter-nuclear distances. Such functions have a very long history and include the Kratzer–Fues [6,7], Lennard–Jones [8] and Morse potentials [9]. Variations on these and other functions are in wide use today.

Spline functions in the form of piecewise polynomials with constraints at the knots connecting them are used in conjunction with diatomic molecule calculations [10,11]. The effect of cubic spline interpolation error on calculated energy levels has been analyzed [12]. Higher order Hermite splines, which are splines based on functions together with their first derivatives and will be used here, have been investigated and applied in other disciplines [13]. The effectiveness of an interpolation method depends on the spacing of the data, the noise level and the smoothness of the function to be fit. These obvious criteria are used to judge the value of the interpolation method introduced here and to define its limitations.

2. Theory

2.1. Energy function

The function of the energy to be used here is

$$Y(z) = -\left(\frac{E_S - E_U}{E_U}\right)E_E\left(\frac{1-z}{2}\right)^2$$

---

* E-mail address: david.schmidling@bcc.cuny.edu

2210-271X/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.comptc.2010.10.002

where E_S is the energy of the separated atoms, E_U is the energy of the united atom, E_E is the electronic energy and z is defined as $(R - R_a)/(R + R_a)$, where R is the inter-nuclear distance and R_a is an arbitrary distance greater than zero. The z definition differs from the usual one, which uses the equilibrium distance, R_e. The first term is a scaling parameter. The electronic energy is the total energy less the nuclear repulsion term and is used so as to give a function that is everywhere finite. The third term fixes a practical problem associated with the uncertainty in the gradient of the electronic energy and requires some explanation. The derivative of the electronic energy with respect to z is

$$\left(\frac{dE_E}{dz}\right) = \left(\frac{dE_E}{dR}\right)\left(\frac{dR}{dz}\right) = \left(\frac{dE_E}{dR}\right)\left[\frac{2R_a}{(1-z)^2}\right] \tag{1}$$

If dE/dR is given to a finite number of decimal places, for instance in a table of values, at least part of the uncertainty associated with dE_E/dR is a constant. The uncertainty associated with dE_E/dz then becomes large as R becomes large as z approaches one. The derivative of Y with respect to z can be written as

$$\frac{dY(z)}{dz} = \left(\frac{E_S - E_U}{E_U}\right)\left(\frac{1-z}{2}\right)E_E - \left(\frac{E_S - E_U}{E_U}\right)\left(\frac{R_a}{2}\right)\left(\frac{dE_E}{dR}\right) \tag{2}$$

The $(1 - z)^2$ term in the denominator of Eq. (1) is not present in Eq. (2), which fixes the problem. The function $Y(z)$, together with its derivative utilize all of the data available from *ab initio* calculations on diatomic molecules. The endpoints are obtained from atomic data, and no extrapolation is required.

Uncertainties in E_E and dE_E/dR are assumed to originate from imperfect integral convergence, which can be dependent on the inter-nuclear distance and from the number of decimal places used to report the final result. In the absence of a detailed analysis of the *ab initio* calculation, the combined uncertainty is assumed to be of the form

$$U = \left[(bR^c)^2 + \left(\frac{10^{-d}}{2\sqrt{3}}\right)^2\right]^{1/2} \tag{3}$$

where the parameters b and c can be determined from the noise residual of fits of closely spaced points to a low order polynomial, performed near various widely spaced R values, while d is the number of decimal places reported.

2.2. Low R limiting law (LRLL)

Ab initio calculations at low R are important for the study of scattering phenomena [14] and for dipole moment functions [15]. At low R values, the electronic energy takes the form [16]

$$E_E = E_U + \sum_{i=2} A_i R^i \tag{4}$$

Levine [17] obtained A_2 from second order perturbation theory for H_2^+. In general though, there does not appear to be a proper theoretical framework for obtaining the higher A coefficients at the present time. Accordingly, a least-squares fit to the electronic energy and its gradient is performed. The fitting procedure uses Jacobi polynomials.

$$E_E - E_U = \left(\frac{2R}{R_n}\right)^2 \sum_{j=1}^{m} a_j P_{j-1}^{(0,4)}\left(-1 + \frac{2R}{R_n}\right) \tag{5}$$

$$\frac{dE_E}{dR} = \sum_{j=1}^{m} a_j \left[\frac{8R}{R_n^2} P_{j-1}^{(0,4)}\left(-1 + \frac{2R}{R_n}\right) + \frac{4R^2}{R_n^3}(4+j) P_{j-2}^{(1,5)}\left(-1 + \frac{2R}{R_n}\right)\right] \quad P_{-1}^{(1,5)} = 0 \tag{6}$$

where n is the number of data set pairs, R_n is the highest R value used and m is the number of coefficients. No orthogonality exists, and the correlations can be quite high. The values and uncertainties at points not present in the data can be evaluated from the coefficients and the covariance matrix. The m and n are chosen to minimize the maximum uncertainty estimate in the range between $R = 0$ and R_n. The A coefficients and the derivatives of $Y(-1)$ are derived from the least-squares result.

$$A_k = \left[\frac{4}{R_n^k (k-2)!(k+2)!}\right] \sum_{i=k-1}^{m} a_i \frac{(-1)^{1+i-k}(1+i+k)!}{(1+i+k)!}$$

$$k = 2 \ to \ m+1$$

$$Y(-1) = -(E_S - E_U)$$

$$D^{(1)}Y(-1) = (E_S - E_U)$$

$$D^{(2)}Y(-1) = -\left(\frac{E_S - E_U}{E_U}\right)\left(\frac{E_U + R_a^2 A_2}{2}\right)$$

$$D^{(k)}Y(-1) = -\left(\frac{k!}{2^k}\right)\left(\frac{E_S - E_U}{E_U}\right)\sum_{j=3}^{k} A_j R_a^j \frac{(k-3)!}{(j-3)!(k-j)!}$$

Table 1

Test data results in a.u. with R range 0.1–40.0, interval $\Delta R = 0.1$, spline order 12 and noise level 1×10^{-11}. LRLL is the low R limiting law, HRLL is the high R limiting law, the A_i are the LRLL coefficients, the C_i are the HRLL coefficients, R_e is the equilibrium inter-nuclear distance, D_e is the equilibrium dissociation energy and $D_{v,J}$ is the dissociation energy for the vibrational state v, J.

Test data	1	2	3
LRLL			
Data points	11	11	11
Parameters	15	15	15
Variance	0.267	0.285	0.274
R range	0–1.1	0–1.1	0–1.1
Max. uncertainty	1.24×10^{-11}	1.25×10^{-11}	1.24×10^{-11}
Max. error	1.60×10^{-11}	1.60×10^{-11}	4.22×10^{-11}
HRLL			
Data points	118	118	118
Parameters	2	2	2
Variance	0.680	0.680	0.681
R range	28.3–∞	28.3–∞	28.3–∞
Max. uncertainty	1.04×10^{-11}	1.04×10^{-11}	1.04×10^{-11}
Max. error	2.16×10^{-11}	2.16×10^{-11}	
Added points	425	412	403
Errors			
Max. E_E	2.5×10^{-11}	2.5×10^{-11}	4.9×10^{-11}
A_2	4.6×10^{-8}	8.7×10^{-8}	1.6×10^{-8}
A_3	-4.6×10^{-6}	-3.9×10^{-6}	-5.7×10^{-7}
A_4	3.4×10^{-5}	7.4×10^{-5}	6.4×10^{-6}
A_5	-3.3×10^{-4}	-7.9×10^{-4}	-1.3×10^{-5}
A_6	2.1×10^{-3}	5.6×10^{-3}	-3.1×10^{-4}
A_7	-8.9×10^{-3}	-2.7×10^{-3}	-3.3×10^{-3}
A_8	2.7×10^{-2}	9.5×10^{-2}	-1.7×10^{-2}
A_9	-6.1×10^{-2}	2.4×10^{-1}	5.8×10^{2}
A_{10}	1.0×10^{-1}	4.6×10^{-1}	-1.3×10^{-1}
A_{11}	-1.2×10^{-1}	-6.2×10^{-1}	2.2×10^{-1}
A_{12}	1.1×10^{-1}	5.4×10^{-1}	-2.5×10^{-1}
A_{13}	-7.0×10^{-2}	-1.7×10^{-1}	2.0×10^{-1}
A_{14}	3.0×10^{-2}	-2.6×10^{-1}	-1.1×10^{-1}
A_{15}	-7.5×10^{-3}	4.4×10^{-1}	3.5×10^{-2}
A_{16}	7.4×10^{-4}	-2.3×10^{-1}	5.1×10^{-3}
C_6	6.3×10^{-3}	6.3×10^{-3}	5.0×10^{-3}
C_8	-5.6	-5.6	-3.4
R_e	8.0×10^{-10}	2.3×10^{-11}	6.8×10^{-13}
D_e	5.0×10^{-12}	4.2×10^{-12}	3.1×10^{-12}
Max. $D_{v,0}$	-3.4×10^{-12}	-3.4×10^{-12}	8.1×10^{-12}