

H2SOLV: Fortran solver for diatomic molecules in explicitly correlated exponential basis

K. Pachucki^{a,*}, M. Zientkiewicz^a, V. A. Yerokhin^b

^a*Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland*

^b*Center for Advanced Studies, Peter the Great St. Petersburg Polytechnic University, Polytekhnicheskaya 29, St. Petersburg 195251, Russia*

Abstract

We present the Fortran package H2SOLV for an efficient computation of the nonrelativistic energy levels and the wave functions of diatomic two-electron molecules within the Born-Oppenheimer approximation. The wave function is obtained as a linear combination of the explicitly correlated exponential (Kołos-Wolniewicz) functions. The computations of H2SOLV are performed within the arbitrary-precision arithmetics, where the number of working digits can be adjusted by the user. The key part of H2SOLV is the implementation of the algorithm of an efficient computation of the two-center two-electron integrals for arbitrary values of internuclear distances developed by one of us (K. Pachucki, Phys. Rev. A **88** (2013) 022507). This has been one of the long-standing problems of quantum chemistry. The code is parallelized, suitable for large-scale computations limited only by the computer resources available and can produce highly accurate results. As an example, we report several benchmark results obtained with H2SOLV, including the energy value accurate to 18 decimal digits.

Keywords: Schrödinger equation; explicitly correlated basis set; hydrogen molecule

PROGRAM SUMMARY

Manuscript Title: H2SOLV: Fortran code for diatomic molecules in explicitly correlated exponential basis.

Authors: K. Pachucki, M. Zientkiewicz and V. A. Yerokhin.

*Corresponding author.

E-mail address: krp@fuw.edu.pl

Program Title: H2SOLV

Journal Reference:

Catalogue identifier:

Licensing provisions: None.

Programming language: Fortran 90.

Computer: PCs or higher performance computers.

Operating system: Linux, Unix.

Has the code been vectorized or parallelized?: Yes

RAM: From several Mbytes to 512 Gbytes, depending on the size of the basis.

External routines: MPFR library [2] and GMP library [3] need to be pre-installed on the computer.

Keywords: Schrödinger equation; explicitly correlated basis set; hydrogen molecule.

Classification: 2.1 Structure and Properties.

Nature of problem: Numerical solution of the two-center two-electron Schrödinger equation within the Born-Oppenheimer approximation using the explicitly correlated basis set of Kolos-Wolniewicz functions.

Solution method: The method of solution is based on the algorithm developed in Ref. [1].

Running time: From seconds to days, depending on the size of the basis.

References:

- [1] K. Pachucki, Efficient approach to two-center exponential integrals with application to excited states of molecular hydrogen, Phys. Rev. A **88** (2013) 022507.
- [2] The GNU Multiple Precision Floating-Point Library, <http://www.mpfr.org/>.
- [3] The GNU Multiple Precision Arithmetic Library, <https://gmplib.org/>.

1. Introduction

Rovibrational spectroscopy of simplest diatomic molecules, in particular, H₂, HD and D₂, demonstrated spectacular progress during the recent years, achieving the fractional accuracy of energy level determinations of few parts in 10⁻⁸ [1, 2, 3]. In order to match such accuracy in theoretical calculations, one needs not only to accurately determine the nonrelativistic energy levels within the Born-Oppenheimer approximation, but also to account for the relativistic, quantum electrodynamics (QED), adiabatic and non-adiabatic effects. The systematic treatment of these effects has recently been performed within the nonrelativistic quantum electrodynamics

(NRQED) expansion [4, 5] and the nonadiabatic perturbation theory [6]. Within these approaches, all corrections to the nonrelativistic energy are obtained by perturbation expansion in the fine-structure constant α and the electron-to-nucleus mass ratio m/M , as expectation values of various effective Hamiltonians with the nonrelativistic wave function. An accurate and efficient numerical procedure for solving the nonrelativistic electronic Hamiltonian in the Born-Oppenheimer approximation is the basis underlying the high-precision calculations [6, 4, 5].

Solution of the Schrödinger equation for the diatomic molecule has a long history. It started in 1927 by the study of Heitler and London [9], who were the first to prove theoretically that the hydrogen molecule is stable against the dissociation into two hydrogen atoms. Several years later, James and Coolidge obtained the first accurate theoretical value for the dissociation energy of H_2 with help of the exponential basis set that explicitly included the interelectronic distance r_{12} [10]. Even more accurate results were obtained by Kołos and Wolniewicz [11] who used a more general form of the exponential basis set. The outstanding accuracy of their results remained unsurpassed until very recently.

The Kołos and Wolniewicz basis, despite the obvious fact that it is the most general and thus most powerful among the exponential basis sets used in the literature, did not become popular up to now. The reason is the notorious difficulty of the evaluation of the two-center radial integrals in the Hamiltonian matrix with such basis functions. In their pioneering work, Kołos and Wolniewicz [11, 12] calculated the integrals by using the Neumann expansion of $1/r_{12}$ in the spherical oblate coordinates. This approach, later used also by other authors [14], is plagued by serious technical problems among which are the complexity of the expansion for high powers of r_{12} and numerical instabilities.

A way towards an effective numerical evaluation of the two-center radial integral of the general form was paved by one of us in Ref. [13]. This approach calculates the radial integrals by using the Taylor expansion in the internuclear distance r . The expansion was shown to be absolutely convergent for all positive values of r , thus making the method applicable even for large internuclear distances.

In the present work, we develop the approach further and present the results in the form of the H2SOLV code, which solves the two-electron two-center Schrödinger equation by using an explicitly correlated basis set of exponential Kołos-Wolniewicz functions. H2SOLV is the only code reported

so far in the literature that is capable of computing two-center integrals with arbitrary powers of interparticle distances, with the efficiency sufficient for practical calculations. We report several benchmark results obtained with H2SOLV, which are by several orders of magnitude more accurate than any values published so far. The H2SOLV code thus solves the long standing problem in quantum chemistry, that of the evaluation of the two-center exchange integrals with exponential functions with Coulomb interaction.

It should be noted that the other exponential basis sets used in the literature, such as the symmetric and asymmetric James-Coolidge basis [10, 17, 18], the symmetric H_2 basis [18] and the symmetric $H-H^+$ basis [19], are just particular cases of the Kołos-Wolniewicz functions (with some restrictions on the nonlinear parameters), so that all of them can be handled by H2SOLV.

2. Formulation of the problem

We consider the stationary Schrödinger equation for the two electrons in the field of two infinitely heavy nuclei,

$$[H_{\text{el}}(r) - E(r)] \Psi(r_{1A}, r_{1B}, r_{2A}, r_{2B}, r_{12}) = 0, \quad (1)$$

where H_{el} is the Hamiltonian of a diatomic molecule in the Born-Oppenheimer approximation,

$$H_{\text{el}}(r) = -\frac{\nabla_1^2 + \nabla_2^2}{2} - \frac{Z_A}{r_{1A}} - \frac{Z_A}{r_{2A}} - \frac{Z_B}{r_{1B}} - \frac{Z_B}{r_{2B}} + \frac{1}{r_{12}} + \frac{Z_A Z_B}{r}, \quad (2)$$

where the indices A and B label the nuclei, the indices 1 and 2 numerate the electrons, and $r \equiv r_{AB}$ is the distance between the nuclei.

Considering the particular case of a Σ ($J = 0$) electronic state, the wave function can be represented as

$$\Psi_{I_{12} I_{AB}}(r_{1A}, r_{1B}, r_{2A}, r_{2B}, r_{12}) = \hat{S}_{I_{12}} \hat{S}_{I_{AB}} \sum_{n_0 \dots n_4} c_{n_0 \dots n_4} \phi_{n_0 \dots n_4}(r_{1A}, r_{1B}, r_{2A}, r_{2B}, r_{12}), \quad (3)$$

where I_{12} is the parity with respect to the permutation of the two electrons ($r_1 \leftrightarrow r_2$), I_{AB} is the parity with respect to the permutation of the two nuclei ($r_A \leftrightarrow r_B$), $\phi_{n_0 \dots n_4}$ are the basis radial functions and $c_{n_0 \dots n_4}$ are linear coefficients. The operators $\hat{S}_{I_{12}}$ and $\hat{S}_{I_{AB}}$ enforce the desired symmetry of

the electronic wave function with respect to the permutation $r_1 \leftrightarrow r_2$ and $r_A \leftrightarrow r_B$,

$$\hat{S}_I = \frac{1}{2} (\hat{1} \pm \hat{P}), \quad \text{for } I = \pm 1, \quad (4)$$

$$\hat{S}_I = \hat{1}, \quad \text{for } I = 0, \quad (5)$$

where $\hat{1}$ is the unity operator and $\hat{P}$ is the permutation operator ($1 \leftrightarrow 2$ or $A \leftrightarrow B$).

In the present work we use the basis radial functions $\phi_{n_0 \dots n_4}$ in the form of the explicitly correlated exponential Kołos-Wolniewicz functions [11]

$$\begin{aligned} \phi_{n_0 n_1 n_2 n_3 n_4}(r_{1A}, r_{1B}, r_{2A}, r_{2B}, r_{12}) &= e^{-y(r_{1A}-r_{1B})-x(r_{2A}-r_{2B})-u(r_{1A}+r_{1B})-w(r_{2A}+r_{2B})} \\ &\times r_{12}^{n_0} (r_{1A}-r_{1B})^{n_1} (r_{2A}-r_{2B})^{n_2} (r_{1A}+r_{1B})^{n_3} (r_{2A}+r_{2B})^{n_4} \\ &\times r^{-n_0-n_1-n_2-n_3-n_4-3}, \end{aligned} \quad (6)$$

where u, w, x and y are real nonlinear parameters.

The energy eigenvalue and the linear coefficients $c_{n_0 \dots n_4}$ of the eigenfunction are determined by solving the secular equation

$$\det \left[\langle n_0 \dots n_4 | H_{\text{el}} | n'_0 \dots n'_4 \rangle - E \langle n_0 \dots n_4 | n'_0 \dots n'_4 \rangle \right] = 0. \quad (7)$$

The matrix elements in this equation can be expressed as a linear combination of the basis integrals $f_{k_0 k_1 k_2 k_3 k_4}(r)$ defined as follows

$$\begin{aligned} f_{k_0 k_1 k_2 k_3 k_4}(r) &= r \int \frac{d^3 r_1}{4\pi} \int \frac{d^3 r_2}{4\pi} r_{12}^{k_0} (r_{1A}-r_{1B})^{k_1} (r_{2A}-r_{2B})^{k_2} \\ &\times (r_{1A}+r_{1B})^{k_3} (r_{2A}+r_{2B})^{k_4} \frac{e^{-u(r_{1A}+r_{1B})-w(r_{2A}+r_{2B})-y(r_{1A}-r_{1B})-x(r_{2A}-r_{2B})}}{r_{12} r_{1A} r_{1B} r_{2A} r_{2B}}. \end{aligned} \quad (8)$$

The explicit expression of the secular equation in terms of the basis integrals $f_{k_0 k_1 k_2 k_3 k_4}(r)$ is given in Appendix A.

In practical calculations, the expansion of the wave function over $n_0 \dots n_4$ in Eq. (3) is truncated. It is a common approach to introduce the rank of the basis set Ω and restrict the expansion by the condition

$$n_0 + n_1 + n_2 + n_3 + n_4 \leq \Omega. \quad (9)$$

The energy value and the wave function are thus becoming functions of the rank of the basis, $E = E(\Omega)$ and $\Psi = \Psi(\Omega)$. By studying the convergence of $E(\Omega)$ as Ω is increased, one obtains the estimation of the uncertainty of the obtained energy value. It should be mentioned that the energy of a bound state $E(\Omega)$ is negative and always decreases when Ω is increased. Thus, $E(\Omega)$ always represents an upper bound for the exact energy level.

3. Calculation of basis integrals

In the previous section, we reduced the problem of the evaluation of the secular matrix to the calculation of the basis integrals $f(n_0 n_1 n_2 n_3 n_4)$. This is the most nontrivial part of the problem. It can be deduced from Eqs. (A.1)-(A.5) that for the evaluation of the secular matrix we need only the basis integrals with non-negative indices $n_i \geq 0$. Such integrals can be obtained from the single master integral $f_{00000}(r) \equiv f(r)$,

$$f(r) = r \int \frac{d^3 r_1}{4\pi} \int \frac{d^3 r_2}{4\pi} \frac{e^{-w_1 r_{12} - u(r_{1A} + r_{1B}) - w(r_{2A} + r_{2B}) - y(r_{1A} - r_{1B}) - x(r_{2A} - r_{2B})}}{r_{12} r_{1A} r_{1B} r_{2A} r_{2B}} \quad (10)$$

by differentiation with respect to the nonlinear parameters,

$$f_{n_0 n_1 n_2 n_3 n_4}(r) = \left(-\frac{\partial}{\partial w_1} \right)^{n_0} \bigg|_{w_1=0} \left(-\frac{\partial}{\partial y} \right)^{n_1} \left(-\frac{\partial}{\partial x} \right)^{n_2} \left(-\frac{\partial}{\partial u} \right)^{n_3} \left(-\frac{\partial}{\partial w} \right)^{n_4} f(r). \quad (11)$$

The above equation reduces the problem of the evaluation of basis integrals to the calculation of the master integral and its partial derivatives over the nonlinear parameters. It may be mentioned that the nonlinear parameter w_1 plays only a supplementary role and is used to derive recursion relations in powers of r_{12} .

The calculation of the master two-center integral $f(r)$ is a difficult and long-standing problem. In the present work, we calculate the master two-center integral and its derivatives by Taylor expansion in the internuclear distance r , using the approach developed by one of us in Ref. [13]. With explicit numerical calculations in Ref. [13] it was demonstrated that the Taylor expansion method is not restricted to the region of small values of r but works well also when r is significantly larger than 1 and for arbitrary large powers of r_{12} .

The Taylor expansion of the basis integrals in r has the following general form [13]:

$$f_{n_0 n_1 n_2 n_3 n_4}(r) = \sum_{n=1}^{\infty} r^n \left[(\ln r + \gamma_E) f_{n n_0 n_1 n_2 n_3 n_4}^{(1)} + f_{n n_0 n_1 n_2 n_3 n_4}^{(2)} \right], \quad (12)$$

where γ_E is the Euler constant. The coefficients of the Taylor expansion $f_{n n_0 n_1 n_2 n_3 n_4}^{(1,2)}$ can be calculated by the recurrence formulas derived from the differential equations for $f(r)$; the method is described in details in Ref. [13]. The differential equations for the master integral $f(r)$ are

$$\begin{aligned} (w_1^2 - 4w^2) \frac{\partial f'(r)}{\partial w} &= -\frac{r f(r)}{2} \frac{\partial \sigma_{02}}{\partial w} - 2r w f''(r) + \frac{-F_1 - F_2 + F_3 + F_4}{2}, \\ (w_1^2 - 4u^2) \frac{\partial f'(r)}{\partial u} &= -\frac{r f(r)}{2} \frac{\partial \sigma_{02}}{\partial u} - 2r u f''(r) + \frac{F_1 + F_2 + F_3 + F_4}{2}, \\ (w_1^2 - 4x^2) \frac{\partial f'(r)}{\partial x} &= -\frac{r f(r)}{2} \frac{\partial \sigma_{02}}{\partial x} - 2r x f''(r) + \frac{F_1 - F_2 + F_3 - F_4}{2}, \\ (w_1^2 - 4y^2) \frac{\partial f'(r)}{\partial y} &= -\frac{r f(r)}{2} \frac{\partial \sigma_{02}}{\partial y} - 2r y f''(r) + \frac{-F_1 + F_2 + F_3 - F_4}{2}, \end{aligned} \quad (13)$$

where $f'(r) = df(r)/dr$, $f''(r) = d^2 f(r)/dr^2$,

$$\begin{aligned} F_1 &= \text{Ei}[-r(w_1 + 2u)] \exp[r(u - w + x - y)] - \text{Ei}[-r(w_1 + 2w)] \exp[-r(u - w + x - y)], \\ F_2 &= \text{Ei}[-r(w_1 + 2u)] \exp[r(u - w - x + y)] - \text{Ei}[-r(w_1 + 2w)] \exp[-r(u - w - x + y)], \\ F_3 &= \text{Ei}[-2r(u + w)] \exp[r(u + w + x + y)] + \left\{ \text{Ei}[2r(x + y)] - \text{Ei}[-r(w_1 - 2x)] \right. \\ &\quad \left. - \text{Ei}[-r(w_1 - 2y)] - \ln \left[\frac{(w_1 + 2u)(w_1 + 2w)(x + y)}{(u + w)(w_1 - 2x)(w_1 - 2y)} \right] \right\} \exp[-r(u + w + x + y)], \\ F_4 &= \text{Ei}[-2r(u + w)] \exp[r(u + w - x - y)] + \left\{ \text{Ei}[-2r(x + y)] - \text{Ei}[-r(w_1 + 2x)] \right. \\ &\quad \left. - \text{Ei}[-r(w_1 + 2y)] - \ln \left[\frac{(w_1 + 2u)(w_1 + 2w)(x + y)}{(u + w)(w_1 + 2x)(w_1 + 2y)} \right] \right\} \exp[-r(u + w - x - y)], \end{aligned} \quad (14)$$

and

$$\sigma_{02} = (u + w - x - y)(u - w + x - y)(u - w - x + y)(u + w + x + y). \quad (15)$$

From the dimensional analysis, one can derive that the function f satisfies the following identity

$$w_1 \frac{\partial f(r)}{\partial w_1} = -w \frac{\partial f(r)}{\partial w} - u \frac{\partial f(r)}{\partial u} - x \frac{\partial f(r)}{\partial x} - y \frac{\partial f(r)}{\partial y} + r f'(r) - 2 f(r). \quad (16)$$

The explicit expressions for the recurrence relations obtained from the above differential equations are too long to be presented in the printed form. They were derived by the symbolic computation in the Mathematica package and stored directly in the Fortran form.

Despite the fact that the recurrence relations are quite long, the time required for their computation grows only linearly with the increase of the length of the Taylor series. Because of this it is possible to use these expressions for computing basis integrals even in the cases when r is not small and and the length of the Taylor expansion is large.

General recurrence relations derived from above differential equations contain $1/x$, $1/y$, $1/w$, $1/u$ and thus become numerically unstable for small values of nonlinear parameters. Small values of w and u are not needed in actual calculations, so the corresponding instability is not important. The parameters x and y , however, might happen to be small. We were not able to find reasonably compact recurrence relations that are stable for small values of x and y . Instead, we obtain separate sets of recurrence relations for 3 special cases: $(x = 0, y \neq 0)$, $(x \neq 0, y = 0)$, and $x = y = 0$. The special-case recurrences can be derived from the same set of differential equations (13). In the actual calculations one has to avoid cases where x or y are nonzero but small (typically < 0.01), this being the only drawback of the present computational approach.

4. Description of the code

The H2SOLV code solves the two-electron two-center Schrödinger equation by using an explicitly correlated basis set of exponential Kołos-Wolniewicz functions (6). The user should provide the nuclear charges Z_A and Z_B of the two nuclei, the internuclear distance R (in a.u.), the parity with respect to the permutation of the two electrons I_{12} , the parity with respect to the permutation the two nuclei I_{AB} , and the nonlinear parameters of the wave function y , x , u , and w . The code currently supports the Σ^+ states (with zero total angular momentum, $J = 0$). Generalizations to the higher electronic angular

momentum Π and Δ states are possible but not implemented. Normalizability of the wave function requires u and w to be positive. Parameters x and y are in principle arbitrary, but in practice their absolute value should not exceed $\max(u, w)$. The cases where x and(or) y are small but nonzero have to be excluded because of numerical instabilities. Typically, the condition $|x|, |y| \notin (\epsilon, 0.01)$ is sufficient, where ϵ is the smallest positive machine value. The cases of $x = 0$ and(or) $y = 0$ are handled by special sets of recurrence relations and are numerically safe.

In the beginning of the calculation, the code checks the nonlinear parameter for the special cases, in order to avoid the linear dependence of the basis set. The special cases are

- 1: $x = y = 0$ and $u = w$, the symmetric James-Coolidge basis,
- 2: $x = y = 0$ and $u \neq w$, the asymmetric James-Coolidge basis,
- 3: $x = -y$ and $u = w$, the symmetric H_2 basis,
- 4: $x = y$ and $u = w$, the symmetric H^-H^+ basis.

If none of the above applies, the most general basis type (0: the Kołos-Wolniewicz basis) is assumed.

Depending on the basis type, the restrictions on powers of interparticle distances should be imposed. The first restriction is $\text{mod}_2(n_2 + n_3) = 0, 1$ and the second restriction is $n_2 > n_3$ or $n_2 = n_3$ and $n_4 \geq n_5$. For example, if the symmetric James-Coolidge ($u = w, y = x = 0$) basis is chosen, then the gerade symmetry leads to the condition that $n_2 + n_3$ is even. Once the restrictions are imposed, the code generates the basis with the powers of interparticle distances restricted by $\sum_i n_i \leq \Omega$, see Eq. (9). The complete list of all restrictions applied in each particular case can be found in the code. The user has the choice to let the code assign the restrictions automatically or to specify them explicitly in the input parameters.

Once the set of all allowed combinations of n_i is generated, the code allocates memory and calculates the Taylor expansion coefficients by using the recursion relations. After the Taylor expansion coefficients are computed and stored, the code calculates the radial integrals by summing up the expansion coefficients. After that the code calculates the Hamiltonian and the overlap matrices. The linear coefficients of the wave function and the energy eigenvalue are determined by solving the secular equation by the inverse iteration method using the user-supplied initial approximation to the energy. The resulting eigenfunction is stored in the external file on the disk.

5. Numerical issues

The common feature of nearly all explicitly correlated basis sets is that calculations with them should be done in an extended precision. The main reason is that the overlap matrices of such basis sets are almost singular, and become increasingly singular as the size of the basis set is increased. This is the case also for the basis set considered in the present work. We found that in order to get reliable results with several thousands of basis functions, we need to perform calculations with working precision of at least 60 decimal digits.

In H2SOLV code, we perform all calculations with help of the arbitrary precision arithmetic library MPFUN2015 written in Fortran 90 by David Bailey [15]. MPFUN2015 is a thread-safe library, which allowed us to greatly speed up calculations by using shared memory parallelization with OPENMP. In our calculations we use the version v04 of MPFUN2015 based on the GNU MPFR library [16], since it is several times faster than the pure Fortran version. In fact this is the fastest thread-safe arbitrary precision arithmetic library publicly available at present. The parallelization of the calculation of integrals in H2SOLV is almost 100% efficient, so that the evaluation time is inversely proportional to the number of available cores (threads). This makes the evaluation of integrals very efficient, and for large basis sets the code spends most of the time on the linear algebra, rather than on integrals.

An additional important feature of MPFUN2015 is that the number of digits of working precision can be changed on the fly. While the maximal number of digits of working precision is fixed during the compilation, the actual number of digits can be decreased during the execution. In H2SOLV, we take advantage of this opportunity, by using different number of digits in the evaluation of integrals and in the linear algebra. For large basis sets, the most time consuming part is the LDLT decomposition, so decreasing the number of digits in this part may speed up the computations considerably.

6. Usage of the code

The input parameters for H2SOLV can be supplied either in the interactive mode or, more conveniently, by creating an input text file and redirecting the standard input (typically, `./h2solv <input_file`). Below is an example of the input file for a calculation of the energy and the wave function of the first excited electronic state of $^1\Sigma_g^+$ symmetry of the hydrogen molecule, for the internuclear distance $R = 6.0$ a.u.

```

1      ZA nuclear charge
1      ZB nuclear charge
1      IAB, A <-> B symmetry (=-1,0,1)
1      IE, r1 <-> r2 symmetry (=-1,1)
6.0    internuclear distance R
0      RESTRICT1, basis restriction type 1 (=-1..2; -1 for auto)
0      RESTRICT2, basis restriction type 2 (=-1..2; -1 for auto)
7      OMEGA, basis rank (=1..16)
140    NMAX, length of Taylor series (0 for the default)
60     NDIGITS, working precision for the linear algebra in digits
-0.694 approximate energy EN
-0.103 nonlinear parameter Y
-0.303 nonlinear parameter X
0.444 nonlinear parameter U
0.684 nonlinear parameter W

```

The left column of the input file contains the numerical values that are read by the code, whereas the right column contains commentaries that are ignored. The first two parameters ZA and ZB are the nuclear charges (positive, integer). The next two parameters IAB and IE describe the symmetries with respect to the permutation of nuclei and of electrons, respectively (1, symmetric; -1, antisymmetric; 0, no symmetry imposed). In the example, both symmetries are set to 1, which corresponds to the gerade (IAB = 1) and singlet (IE = 1) state.

R is the internuclear distance in atomic units (positive, real). RESTRICT1 and RESTRICT2 are the basis restrictions (integer, -1...2). Their meaning is explained by the following excerpt from `h2solv.f`:

```

c      RESTRICT1 = -1 ! auto
c      RESTRICT1 = 0 ! no restriction
c      RESTRICT1 = 1 ! MOD(n2+n3,2) = 0
c      RESTRICT1 = 2 ! MOD(n2+n3,2) = 1

c      RESTRICT2 = -1 ! auto
c      RESTRICT2 = 0 ! no restriction
c      RESTRICT2 = 1 ! (n2.gt.n3).OR.((n2.eq.n3).AND.(n4.ge.n5))
c      RESTRICT2 = 2 ! (n2.gt.n3).OR.((n2.eq.n3).AND.(n4.gt.n5))

```

If the restrictions are set to -1, the code will automatically impose minimal restrictions needed in order to avoid linear dependencies of the basis functions.

The rank of the basis `Omega` is the maximal value of the sum of powers of all interparticle distances Ω , see Eq. (9), which determines the size of the basis set. NMAX is the length of the Taylor series. If set to 0, the code determines this parameter automatically as $NMAX = 60 + 10 \cdot R$. This choice should be adequate for most practical applications, but it is recommended to check the final results by varying NMAX. NDIGITS is the working precision

for the linear algebra, in decimal digits (60 is sufficient for most cases). The working precision for the evaluation of integrals (80 by default) is set up by the parameter `mpipl` in the MPFUN library. (Naturally, `NDIGITS` should not exceed `mpipl`.) An extension to a larger working precision is straightforward; one needs just to increase the parameter `mpipl` in `mpfunf.f90` and recompile the code.

EN is the initial approximation to the energy. The code will compute the eigenvalue of the Hamiltonian matrix which is closest to the initial approximation provided. In the example, the initial approximation `EN = -0.694` is closest to the first excited state (i.e., the second lowest-lying state) of this symmetry, so this state is calculated.

The last four inputs variables `Y`, `X`, `U`, `W` are the nonlinear parameters of the basis set, which can be chosen differently. For small internuclear distances $R \leq 6$ a.u., the James-Coolidge basis $x = y = 0$ works well [18]. For large distances $R > 12$, one may use generalized Heitler-London basis (the electrons on the corresponding centers $y = u$ and $x = -w$ or both electrons on the same center $y = u$ and $x = w$, see Ref. [18] for details). In the intermediate region of R , a reasonable starting choice of parameters is $y = 0$ and $x = w = 1/2$. In order to achieve the best possible results for a given size of the basis set (value of Ω), one needs to minimize the energy with respect to these nonlinear parameters. However, in many situations it is sufficient just to enlarge the size of the basis set with the recommended values of nonlinear parameters.

The computation time of H2SOLV with the input example specified above is about a minute on a dual processor Intel Xeon CPU E5-2680 v3 2.50GHz. Below is the control output of this exemplary calculation.

```
=====
Nonrelativistic energy of H2-like molecule

General KOLOS-WOLNIEWICZ basis

Nuclear charge A:          1
Nuclear charge B:          1
A <-> B symmetry:          GERADE
r1 <-> r2 symmetry:          SYM
Internuclear distance:      6.0000

Work. prec. for integr.:    80
Work. prec. for lin. alg.:  60

Basis parameters:
Rank (OMEGA):              7
```

```

Type 1 of restriction:      0
Type 2 of restriction:      0
Length of Taylor series:    140
Basis size:                 792

Non-linear parameters:
Y =      -1.0300000000000000e-1
X =      -3.0300000000000000e-1
U =       4.4400000000000000e-1
W =       6.8400000000000000e-1

Resulting energy:
EN =      -6.942666307257072980e-1

```

Beside the output in the standard output channel, the code creates the file `wf_7.dat` ("7" corresponds to the value of `Omega`), which contains the wave function. The computed wave function can be used for calculating expectation values of various operators, as demonstrated by the exemplary code `example.f`. The energy value obtained with 792 basis functions in the above example, $E = -0.694\,266\,6$, can be compared with the benchmark result of $E = -0.694\,267\,038\,209$ [17], see Table 1.

We provide also a number of other input and output files of exemplary calculations which illustrate the usage of the code with different variants of the James-Coolidge basis and the generalized Heitler-London basis. The results of these computational exercises can be compared with the benchmark values collected in Table 1. The examples were chosen so as to keep the computation time low, so the precision of the results obtained is not very high, typically about 6 digits. Their accuracy can be enhanced significantly by increasing the rank of the basis Ω . The typical precision of the energy determination for basis sets with the rank $\Omega = 20$ is 14-15 decimal digits [18], which is by far the best accuracy reported in the literature so far for diatomic molecules.

Table 1: Literature benchmark values for BO potential for the comparison with exemplary inputs for different types of the basis set.

Basis	System	State	R	Energy	Reference
general	H ₂	EF	6	-0.694 267 038 209	[17]
jc_sym	H ₂	Σ_g^+	3	-1.057 326 268 873	[18]
jc_asym	H ₂	EF	3	-0.690 747 056 393	[17]
h2_sym	H ₂	Σ_g^+	6	-1.000 835 707 655	[18]
h-h+	HeH ⁺	Σ^+	3	-2.904 314 561 859	[19]

7. Benchmark results

As a demonstration of the potential of the H2SOLV code, we present large-scale calculations of the long-range asymptotics of the exchange energy of the hydrogen molecule ΔE , which is the difference between the lowest-lying singlet and triplet levels. The exchange energy ΔE is exponentially decreasing with the increase of the internuclear distance R and thus its large- R asymptotics is difficult to be accessed in numerical calculations. In Table 2 we present our numerical results for the energy of the $^1\Sigma_g^+$ state of the H_2 molecule and the exchange energy ΔE multiplied by the exponential prefactor, at $R = 25$ a.u. These results were obtained by H2SOLV with the following input parameters: $y = u = 0.52$, $-x = w = 0.52$, MPIPL= 200, NDIGITS= 100, NMAX= 450. The detailed numerical study of the large- R asymptotics of the exchange energy of the H_2 molecule is an important physical problem and will be presented elsewhere.

Table 2: Numerical values for the Born-Oppenheimer and exchange energies obtained with the generalized Heitler-London functions at $R = 25$ a.u.

Ω	N	$E(^1\Sigma_g^+)$	$e^{2R} \Delta E$
17	13332	-1.000 000 027 470 268 507 696 150	4815.918
18	17017	-1.000 000 027 470 268 507 701 647	4816.266
19	21472	-1.000 000 027 470 268 507 703 947	4816.275
20	26818	-1.000 000 027 470 268 507 704 945	4816.490
21	33176	-1.000 000 027 470 268 507 705 394	4816.516
22	40690	-1.000 000 027 470 268 507 705 603	4816.527
∞			

The advantage of the explicitly correlated exponential functions is that the consecutive increments of the series of energies as a function of Ω converge exponentially (as $e^{-\beta\Omega}$), thus allowing for a very accurate extrapolation to the limit of $\Omega \rightarrow \infty$. As an example, by using a slightly adapted H2SOLV code (the linear algebra performed with the proprietary HSL_MP54 routine and a double set of the non-linear parameters as in Ref. [18]), we were able to obtain results for Ω as large as 25 (basis size 61 425). The extrapolated result for the Bohr-Oppenheimer energy at $R = 1.4011$ is

$$E = -1.174\,475\,931\,400\,217\,165(2), \quad (17)$$

which is two orders of magnitude more accurate than previous best literature value.

8. Summary

The Fortran package H2SOLV provides an efficient computation of the nonrelativistic energies and the wave functions of diatomic two-electron molecules in the Born-Oppenheimer approximation. The calculation is performed with the explicitly correlated basis set of exponential (Kołos-Wolniewicz) functions. This is one of the most general basis sets for the two-center two-electron problem suitable for an accurate and compact representation of the two-center wave function both for small as well as large internuclear distances. The calculation is performed in the arbitrary-precision arithmetics, which allows for large-scale computations limited only by the computer resources available. The typical precision of the energy determination for basis sets with the rank $\Omega = 20$ ($\approx 25,000$ functions) is 14-15 decimal digits, which is by far the best accuracy reported in the literature for diatomic molecules. Already with the basis set of $\approx 1,000$ functions, one typically gets energies accurate to 6-7 digits, which is better than most of literature results.

Acknowledgements

The authors acknowledge support from NCN Grant 2012/04/A/ST2/00105. V.A.Y. acknowledges support by the Russian Federation program for organizing and carrying out scientific investigations and RFBR (grant No. 16-02-00538).

Appendix A. Matrix elements of the secular equation

The overlap integral of the two wave functions in Eq. (7) is obtained as

$$\begin{aligned} \langle k_0 \dots k_4 | l_0 \dots l_4 \rangle = & \frac{1}{16} \left\{ f(n_0 + 1, n_1, n_2, n_3 + 2, n_4 + 2) \right. \\ & - f(n_0 + 1, n_1, n_2 + 2, n_3 + 2, n_4) - f(n_0 + 1, n_1 + 2, n_2, n_3, n_4 + 2) \\ & \left. + f(n_0 + 1, n_1 + 2, n_2 + 2, n_3, n_4) \right\}, \end{aligned} \quad (\text{A.1})$$

where

$$f(n_0, n_1, n_2, n_3, n_4) = f_{n_0 n_1 n_2 n_3 n_4}(r) / r^{2+n_0+n_1+n_2+n_3+n_4}, \quad (\text{A.2})$$

and $n_i \equiv k_i + l_i$. The matrix element of the potential energy part of the Hamiltonian (all terms on the right-hand side of Eq. (2) except for the gradient terms) is given by

$$\begin{aligned}
\langle k_0 \dots k_4 | V | l_0 \dots l_4 \rangle = & \frac{1}{16r} \left\{ f(n_0, n_1, n_2, n_3 + 2, n_4 + 2) \right. \\
& - f(n_0, n_1, n_2 + 2, n_3 + 2, n_4) \\
& - f(n_0, n_1 + 2, n_2, n_3, n_4 + 2) + f(n_0, n_1 + 2, n_2 + 2, n_3, n_4) \\
& + 2(Z_A + Z_B) \left[-f(n_0 + 1, n_1, n_2, n_3 + 1, n_4 + 2) \right. \\
& - f(n_0 + 1, n_1, n_2, n_3 + 2, n_4 + 1) + f(n_0 + 1, n_1, n_2 + 2, n_3 + 1, n_4) \\
& \left. + f(n_0 + 1, n_1 + 2, n_2, n_3, n_4 + 1) \right] \\
& + 2(Z_A - Z_B) \left[f(n_0 + 1, n_1, n_2 + 1, n_3 + 2, n_4) \right. \\
& + f(n_0 + 1, n_1 + 1, n_2, n_3, n_4 + 2) - f(n_0 + 1, n_1 + 1, n_2 + 2, n_3, n_4) \\
& \left. - f(n_0 + 1, n_1 + 2, n_2 + 1, n_3, n_4) \right] \\
& + Z_A Z_B \left[f(n_0 + 1, n_1, n_2, n_3 + 2, n_4 + 2) \right. \\
& - f(n_0 + 1, n_1, n_2 + 2, n_3 + 2, n_4) - f(n_0 + 1, n_1 + 2, n_2, n_3, n_4 + 2) \\
& \left. + f(n_0 + 1, n_1 + 2, n_2 + 2, n_3, n_4) \right] \Big\}. \tag{A.3}
\end{aligned}$$

The matrix element of the kinetic energy part of the Hamiltonian (the terms with gradient in Eq. (2)) is given by

$$\begin{aligned}
\langle k_0 \dots k_4 | T | l_0 \dots l_4 \rangle = & \frac{1}{16 r^2} \{ k_0 l_0 [f(n_0 - 1, n_1, n_2, n_3 + 2, n_4 + 2) \quad (\text{A.4}) \\
& - f(n_0 - 1, n_1, n_2 + 2, n_3 + 2, n_4) - f(n_0 - 1, n_1 + 2, n_2, n_3, n_4 + 2) \\
& + f(n_0 - 1, n_1 + 2, n_2 + 2, n_3, n_4)] \\
& - 2 (\delta_k (l_1 - 1) l_1 + (k_1 - 1) k_1 \delta_l) (f(n_0 + 1, n_2 - 2, n_2, n_3, n_4 + 2) \\
& - f(n_0 + 1, n_2 - 2, n_2 + 2, n_3, n_4)) + 4 (y_l \delta_k l_1 + y_k k_1 \delta_l) \\
& \times (f(n_0 + 1, n_1 - 1, n_2, n_3, n_4 + 2) - f(n_0 + 1, n_1 - 1, n_2 + 2, n_3, n_4)) \\
& - 2 ((u_l^2 + w_l^2) \delta_k + (u_k^2 + w_k^2) \delta_l) f(n_0 + 1, n_1, n_2, n_3 + 2, n_4 + 2) \\
& + 2 (\delta_k (l_3 - 1) l_3 + (k_3 - 1) k_3 \delta_l) (f(n_0 + 1, n_1, n_2, n_3 - 2, n_4 + 2) \\
& - f(n_0 + 1, n_1, n_2 + 2, n_3 - 2, n_4)) - 4 (u_l \delta_k l_3 + u_k k_3 \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2, n_3 - 1, n_4 + 2) - f(n_0 + 1, n_1, n_2 + 2, n_3 - 1, n_4)) \\
& + 2 [\delta_k (u_l^2 - y_l^2 + l_1 (l_1 + 1) - l_3 (l_3 + 1)) \\
& + \delta_l (u_k^2 - y_k^2 + k_1 (k_1 + 1) - k_3 (k_3 + 1))] \\
& \times (f(n_0 + 1, n_1, n_2, n_3, n_4 + 2) - f(n_0 + 1, n_1, n_2 + 2, n_3, n_4)) \\
& + 4 (u_l \delta_k (l_3 + 1) + u_k (k_3 + 1) \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2, n_3 + 1, n_4 + 2) - f(n_0 + 1, n_1, n_2 + 2, n_3 + 1, n_4)) \\
& + 2 ((u_l^2 + x_l^2) \delta_k + (u_k^2 + x_k^2) \delta_l) f(n_0 + 1, n_1, n_2 + 2, n_3 + 2, n_4) \\
& - 4 (y_l \delta_k (l_1 + 1) + y_k (k_1 + 1) \delta_l) \\
& \times (f(n_0 + 1, n_1 + 1, n_2, n_3, n_4 + 2) - f(n_0 + 1, n_1 + 1, n_2 + 2, n_3, n_4)) \\
& - 2 (\delta_k (l_2 - 1) l_2 + (k_2 - 1) k_2 \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2 - 2, n_3 + 2, n_4) - f(n_0 + 1, n_1 + 2, n_2 - 2, n_3, n_4)) \\
& + 4 (x_l \delta_k l_2 + x_k k_2 \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2 - 1, n_3 + 2, n_4) - f(n_0 + 1, n_1 + 2, n_2 - 1, n_3, n_4)) \\
& + 2 (\delta_k (l_4 - 1) l_4 + (k_4 - 1) k_4 \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2, n_3 + 2, n_4 - 2) - f(n_0 + 1, n_1 + 2, n_2, n_3, n_4 - 2)) \\
& - 4 (w_l \delta_k l_4 + w_k k_4 \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2, n_3 + 2, n_4 - 1) - f(n_0 + 1, n_1 + 2, n_2, n_3, n_4 - 1)) \\
& + 2 (\delta_k (w_l^2 - x_l^2 + l_2 (1 + l_2) - l_4 (1 + l_4)) \\
& + \delta_l (w_k^2 - x_k^2 + k_2 (1 + k_2) - k_4 (1 + k_4))) \\
& \times (f(n_0 + 1, n_1, n_2, n_3 + 2, n_4) - f(n_0 + 1, n_1 + 2, n_2, n_3, n_4)) \\
& + 4 (w_l \delta_k (1 + l_4) + w_k (1 + k_4) \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2, n_3 + 2, n_4 + 1) - f(n_0 + 1, n_1 + 2, n_2, n_3, n_4 + 1)) \\
& + 2 ((w_l^2 + y_l^2) \delta_k + (w_k^2 + y_k^2) \delta_l) f(n_0 + 1, n_1 + 2, n_2, n_3, n_4 + 2) \\
& - 4 (x_l \delta_k (1 + l_2) + x_k (1 + k_2) \delta_l) \\
& \times (f(n_0 + 1, n_1, n_2 + 1, n_3 + 2, n_4) - f(n_0 + 1, n_1 + 2, n_2 + 1, n_3, n_4)) \\
& - 2 ((y_l^2 + x_l^2) \delta_k + (y_k^2 + x_k^2) \delta_l) f(n_0 + 1, n_1 + 2, n_2 + 2, n_3, n_4) \} , \quad (\text{A.5})
\end{aligned}$$

where $\delta_k = 1/2$ for $k_0 + l_0 = 0$ and $\delta_k = k_0/(k_0 + l_0)$ otherwise; $\delta_l = 1/2$ for $k_0 + l_0 = 0$ and $\delta_l = l_0/(k_0 + l_0)$ otherwise; y_k, x_k, u_k, w_k are the nonlinear parameters of the left-hand side wave function, y_l, x_l, u_l, w_l are the nonlinear parameters of the right-hand side wave function.

References

- [1] J. Liu et al., J. Chem. Phys. **130**, 174306 (2009).
- [2] J. Liu et al., J. Chem. Phys. **132**, 154301 (2010).
- [3] G.D. Dickenson *et al.*, Phys. Rev. Lett. **110**, 193601 (2013).
- [4] K. Piszczatowski, G. Łach, M. Przybytek, J. Komasa, K. Pachucki, B. Jeziorski, J. Chem. Theor. Comp. **5**, 3039 (2009).
- [5] J. Komasa, K. Piszczatowski, G. Łach, M. Przybytek, B. Jeziorski, and K. Pachucki, J. Chem. Theor. Comp. **7**, 3105 (2011).
- [6] K. Pachucki and J. Komasa, J. Chem. Phys. **130**, 164113 (2009).
- [7] C. Herring and M. Flicker, Phys. Rev. **134**, A362 (1964)
- [8] B.L. Burrows, A. Dalgarno, and M. Cohen, Phys. Rev. A **86**, 052525 (2012).
- [9] W. Heitler and F. London, Z. Phys. **44** (1927) 455.
- [10] H. M. James and A. S. Coolidge, J. Chem. Phys. **1** (1933) 825.
- [11] W. Kołos and L. Wolniewicz, Phys. Rev. Lett. **20**, 243, (1968).
- [12] W. Kołos and L. Wolniewicz, J. Chem. Phys. **49**, 404 (1968).
- [13] K. Pachucki, Phys. Rev. A **88**, 022507 (2013).
- [14] F.E. Harris, Int. J. Quant. Chem. **88**, 701 (2002).
- [15] MPFUN2015: A thread-safe arbitrary precision package, D. Bailey, <http://www.davidhbailey.com/dhbssoftware/>
- [16] The GNU MPFR library: www.mpfr.org

- [17] M. Zientkiewicz, Master thesis, University of Warsaw (2014).
- [18] K. Pachucki, Phys. Rev. A **82**, 032509 (2010).
- [19] K. Pachucki, Phys. Rev. A **85**, 042511 (2012).