
MolSOC version 0.1

User's Manual



Author: Sandro Giuseppe Chiodo

1 Introduction

The basic theoretical approach implemented in MolSOC provides a procedure that takes explicitly into account monodeterminantal wave functions of a pair of pre-optimized states, coupled through a Spin-Orbit Coupling (SOC) operator. More explicitly, a full Breit-Pauli operator[1],

$$\begin{aligned} \hat{H}_{BP}^{SO} = & \frac{e^2}{2m_e^2 c^2} \sum_{i=1}^{N_{el}} \left\{ \sum_{\alpha=1}^{N_A} Z_{\alpha} \left(\frac{\mathbf{r}_{i\alpha}}{r_{i\alpha}^3} \times \mathbf{p}_i \right) \cdot \mathbf{s}_i \right. \\ & \left. - \sum_{j \neq i}^{N_{el}} \left(\frac{\mathbf{r}_{ij}}{r_{ij}^3} \times \mathbf{p}_i \right) \cdot (\mathbf{s}_i + 2\mathbf{s}_j) \right\}, \end{aligned} \quad (1)$$

and a reduced screened-nuclear charge one,

$$\hat{H}_{SO} = \frac{e^2}{2m_e^2 c^2} \sum_{i=1}^{N_{el}} \sum_{\alpha=1}^{N_A} Z_{eff}(\alpha) \left(\frac{\mathbf{r}_{i\alpha}}{r_{i\alpha}^3} \times \mathbf{p}_i \right) \cdot \mathbf{s}_i, \quad (2)$$

are considered. The explicit expressions of the derived matrix elements and more informations about this method can be found in earlier detailed works[2, 3].

To run MolSOC are needed the Self-Consistent Field (SCF) Molecular Orbitals (MO) linear combination coefficients of the two states which, in principle, can be supplied by means of many quantum mechanical codes. However, the present version can support only formatted checkpoint files of the **GAUSSIAN 03** program package, the formatted *mos* files of the **TURBOMOLE V6.0** code, and the unformatted *deMon.rst* files of the **deMon V1.02** code.

The sets of MOs, due to the separate optimization of the two states, can be not quite orthogonal to each other. To make these MOs pairwise orthogonals, biorthogonalization is performed by using transformation matrices, which, in turns, are obtained by Singular Value Decomposition (SVD) of the two-state MO overlap matrix.

Notice that the most tedious work is in the preliminary procedure for preparing MOs of the two states. We strongly recommend to pay close attention to this procedure, to look at the MOs of the two states, to compare them before to be coupled. One can be interested to specific MO configurations, even belonging to excited states. It is not instantly obvious that, sometime, these are not properly preserved during the SCF iterations. However, one can refine particular strategies to avoid this drawback[2, 3], or one can easily find with **GAUSSIAN 03** typical therapeutic keywords fit for this purpose, like the *symm* keyword.

2 MolSOC installation

2.1 MolSOC installation

To compile MolSOC is very simple, just you need to open the **Makefile**, located in the source code directory (**source**), and uncomment the lines where are defined own fortran compiler, flags, and linker. Note that MolSOC is supplied with LAPACK and BLAS math libraries which must be compiled before the source code. The math library **Makefile** is located in the **lib** directory. These **Makefiles** contain the targets specific to the following compilers:

- (i) OSF/1 Tru64 compiler;
- (ii) Intel Fortran compiler;
- (iii) Portland Group Fortran-90 compiler.

MolSOC needs a working Fortran 90 compiler. The following installation procedure can be used:

1. Unpack the MolSOC distribution files in subdirectory **molSOC0.1** of own desired directory;
2. Go to the **molSOC0.1/lib** subdirectory and uncomment or add own fortran compiler, flags, and linker in the **Makefile**;
3. Issue in this subdirectory the command **make** to compile the MolSOC libraries;
4. Go to the **molSOC0.1/source** subdirectory and uncomment or add own fortran compiler, flags, and linker in the **Makefile**;
5. Issue in this subdirectory the command **make** to compile the MolSOC source code. The **molSOC0.1.exe** executable will be created and placed in the directory **molSOC0.1/bin**;
6. To make the MolSOC executable permanently recognizable by the operating system by simply typing **molSOC0.1.exe** from the linux shell prompt, add its location path to own shell initialization file (**.cshrc** or **.profile** or ...).

2.2 Contents of the distribution package

MolSOC distribution package contains the following directories and files:

bin	Executable
doc	Documentation files
examples	Example and test files
include	Include files
lib	Library files
source	Source code files

3 How to run MolSOC

3.1 Description of the input files

To actually run MolSOC four input files are necessary, containing the following information:

- (i) File **molSOC.inp** must contain the geometry common to both the electronic states, together with keywords specifying desired calculation type;
- (ii) The basis set file, **basis**, must contain gaussian contraction coefficients and exponents;
- (iii) The files **mos1** and **mos2** must contain the MO linear combination coefficients of the first and second state, respectively. These can be obtained simply by changing the name of the file **deMon.rst** (**deMon**) or **Test.FChk** (**Gaussian**) or **mos** (**TURBOMOLE**) to **mos1** or **mos2** depending on which is the first or the second state.

All these files must be located in the working directory.

3.2 The file molSOC.inp

All keywords in **molSOC.inp** ASCII text file are not case sensitive. They can be assembled only in one keyword block line containing not more than 80 characters, including the blank ones. The ordering of these keywords is free. Comment lines are turned on by the symbol **#**, inserted as first character of each comment line. No blank line are allowed. The structure of the **molSOC.inp** file includes the following sections:

- (i) *Route section*: Specify keywords for the calculation of the SOC contributions and other options. Note that this section can be omitted if the default keyword list is used;
- (ii) *Charge and multiplicities section*: Specify the charge, multiplicities, and MOs coupling mode (α or β orbitals);
- (iii) *Molecular Geometry section*: Specify the molecular geometry in cartesian coordinates.

The end of the input is indicated by the **END** statement or by a blank line.

Route section

In this section the following keywords can be specified with at least three characters for each of them:

BOHR	The input geometry is given in Bohr (default) (a)
ANGSTROM	The input geometry is given in Angstrom (a)
ORTHO	Cram-Schmidt MOs orthonormalization of each state is performed before SOC calculations. Note that this keyword should not be used since MOs coming from SCF calculations should be already orthonormals (the default is NO-ORTHO)
UKS	SOC calculations are performed starting from unrestricted kohn-Sham MOs (b)
RKS	SOC calculations are performed starting from restricted kohn-Sham MOs (default) (b)
SPHERICAL	SOC calculations are performed using spherical MOs (c)
CARTESIAN	SOC calculations are performed using cartesian MOs (c)
ONE	One-electron SOC calculations are performed (d)
ZEFF	One-electron SOC calculations with the screened-nuclear charge method are performed[3] (d)
TWO	SOC matrix elements are calculated using the full Breit-Pauli operator (default) (d)
GAUSSIAN	SOC matrix elements are evaluated using MOs computed with the Gaussian code (default) (e)
DEMON	SOC matrix elements are evaluated using MOs computed with the deMon code (e)
TURBOMOLE	SOC matrix elements are evaluated using MOs computed with the TURBOMOLE code (e)
DIPOLE	Activates the calculation of the dipole moment components (f)
QUADRUPOLE	Activates the calculation of the dipole and quadropole moment components (f)
OCTUPOLE	Dipole, quadropole, and octupole moment components are calculated (f)
ALTER	The MOs occupation is altered (see the subsection reserved to this keyword)
NOBIORTHO	Inactivates the biorthogonalization procedure

Notice that it is not needed to specify default keywords. If more than one of the keywords denoted with the same letter (a, b, c, d, e, and f) are given, the last one will override the previous one.

Charge and multiplicities section

This section consists of a single line given as follow:

C ^IM ^IO ^{II}M ^{II}O

C	Charge of the system (integer value). This is common to both the states
$^I\mathbf{M}$	Multiplicity (2S+1) of the first state (integer value)
$^I\mathbf{O}$	Identify the discoincent orbital of the first state with respect to the second state
$^{II}\mathbf{M}$	Multiplicity (2S+1) of the second state (integer value)
$^{II}\mathbf{O}$	Identify the discoincent orbital of the second state with respect to the first state

For the spin discoinidence case, **MolSOC** works only if $^I\mathbf{M}$ is the multiplicity of the low spin state and $^{II}\mathbf{M}$ those of the high spin state ($^{II}\mathbf{M} - ^I\mathbf{M} = 2$). Both $^I\mathbf{M}$ and $^{II}\mathbf{M}$ must be integer values. The **O** identifier can be **D** (direct) or **R** (reverse). It depends on which discoincent orbital of each state is involved in the coupling mechanism. As an example, for better understand how to set this identifier let's consider the diagrams in Figure 1 showing a graphical representation of this mechanism. Particularly, these are relative to the coupling between singlet and $\alpha\alpha$ -triplet determinants (Figure 1a), α -doublet and α -doublet determinants with an α -electron transition (Figure 1b), and α -doublet and α -doublet determinants with a β -electron transition (Figure 1c). The case shown in Figure 1a requires that both $^I\mathbf{O}$ and $^{II}\mathbf{O}$ must be **D**, since here a spin discoinidence occurs. The spin discoincent orbitals are $^I\phi_{j\beta}$ and $^{II}\phi_{k\alpha}$. Even for the case shown in Figure 1b it is needed to use the character **D** for both $^I\mathbf{O}$ and $^{II}\mathbf{O}$. This is identified as orbital discoinidence, where the discoincent orbitals are $^I\phi_{j\alpha}$ and $^{II}\phi_{k\alpha}$. Instead, the diagram in Figure 1c concerns still an orbital discoinidence, but since the coupling is between β -orbitals, $^I\phi_{j\beta}$ and $^{II}\phi_{k\beta}$, the **O** identifier for both the states must be **R**.

Molecular Geometry section

The geometry is read in cartesian coordinates. The atomic symbol, the x, y, and z coordinates of each atom of the system, and a scaling factor of the nuclear charge ($\lambda(\alpha)$) of this atom have to be specified in each line. Usually, the latter must be set equal 1.0 except if one is interested to carry out calculations with the one-electron SOC operator scaling the nuclear charges using certain aptly chosen λ factors ($Z_{eff}(\alpha) = \lambda(\alpha)Z(\alpha)$). Remind that in this case it is needed the keyword **ONE**.

Keyword ZEFF

As hinted, this keyword activates the calculation of SOC matrix elements using the screened-nuclear charge method. The Z_{eff} used in **MolSOC** have been derived by fitting procedure performed over computed fine-structure splittings (FSS) in Π states of diatomic hydrides[3]. More precisely, the initial λ_n factor of the A atom (in the AH specie) has been obtained varying it manually so as to reach as closely as possible the corresponding value of the matrix elements computed using the full SOC operator and the MOs of the optimizad Π states. On doing so, it has been found for the λ_n factors to follow a straight-line model along the entire series of elements that belong to

the same row of the periodic table. A further fitting has given the following expressions:

$$\begin{aligned}
Z_{eff}(\alpha) &= \lambda_n(\alpha)Z(\alpha) \\
\lambda_1(\alpha) &= 0.2517 + 0.0626N_1^{val}(\alpha) && \text{B - F} \\
\lambda_2(\alpha) &= 0.7213 + 0.0144N_2^{val}(\alpha) && \text{Al - Cl} \\
\lambda_3(\alpha) &= 0.8791 + 0.0039N_3^{val}(\alpha) && \text{Ga - Br} \\
\lambda_4(\alpha) &= 0.9228 + 0.0017N_4^{val}(\alpha) && \text{In - I}
\end{aligned}$$

Where $N_n^{val}(\alpha)$ is the number of valence electrons of the α -th element. Together with the derived value of the Fe element, the Z_{eff} available in MolSOC are those of abovementioned elements.

Keyword *ALTER*

Particularly, this keyword is useful for the calculation of the matrix elements between singlet S_i and triplet T_j excited states, which, in turn, can be evaluated by this formula[4, 5]:

$$\langle S_i | H_{SO} | T_j \rangle = \sum_l^{N_{S_i}} \sum_m^{N_{T_j}} C_{il}^S C_{jm}^T \langle \Psi_{il}^S | H_{SO} | \Psi_{jm}^T \rangle \quad (3)$$

Here Ψ^S and Ψ^T are the singlet and triplet state wave functions, respectively, generated from one-electron vertical excitations performed over the ground state (S_0) electronic configuration, C_{il}^S and C_{jm}^T are the weighted coefficients of the l -th one electron singlet transition and the m -th one-electron triplet transition belonging to the S_i and T_j states, respectively. These coefficients come from time-dependent density functional (TD-DFT) calculations. N_{S_i} and N_{T_j} are the number of transitions defining the i -th and j -th singlet and triplet states, respectively. H_{SO} is the Breit-Pauli operator[1].

It is possible to find a good approximation by simply truncating the above expression to a few terms, those belonging to larger values of the weighted coefficients. Once the keyword *alter* is defined, alteration sections are required, specified at the end of the *molSOC.inp* file, spaced from the *Molecular Geometry section* by the **END** statement or by a black line. These sections, consist, in the case of the $\langle S_i | H_{SO} | T_j \rangle$ matrix elements, of a set of traspositions indicating that occupied MOs of the S_0 state are replaced by its virtual MOs. In this way, one can systematically (give rise) generate the electronic configuration of the Ψ^S and Ψ^T singlet and triplet state wave functions according to the TD-DFT results. Notice that, using this approach, the *mos1* and the *mos2* files must be the same, both containing the linear combination MOs coefficients of the S_0 state. Moreover, since it is a workable procedure in MolSOC to biorthogonalize the MOs of states coming from separate optimization, here occasionally it is needed to switched off the biorthogonalization since the MOs of the singlet, Ψ^S , and triplet, Ψ^T , wave functions, being unaltered from the ground state, S_0 , wave function, are already biorthogonals. To better understand the alteration sections is useful to take into account the following example:

1	2
162	163

162	165
162	163

1	1
162	163

162	163
-----	-----

Here the 163th MO is the HOMO orbital. The transpositions of the first section belong to the first state, 1 transposition for α orbitals (the 162th orbital is exchanged with the 163th one) and 2 for the β orbitals (the 162th orbital is exchanged with the 165th one and the 165th with the 163th). Note that, as for the transposition sections, the α set of transpositions is separated from the β set by a black line. The second section, relative to the second state, specifies two transposition, 1 α and 1 β . Notice that multiple α or β transpositions are needed to reorder the MOs in such a way that the discoinident orbitals, the orbitals involved in the coupling mechanism, must be shifted to external ones, as HOMO orbitals, as required in the implementation of this code[2]. But due to this operation the antysimmetrization principle requires to assign the correct sign for the wave function. To end the transposition sections it is needed a black line.

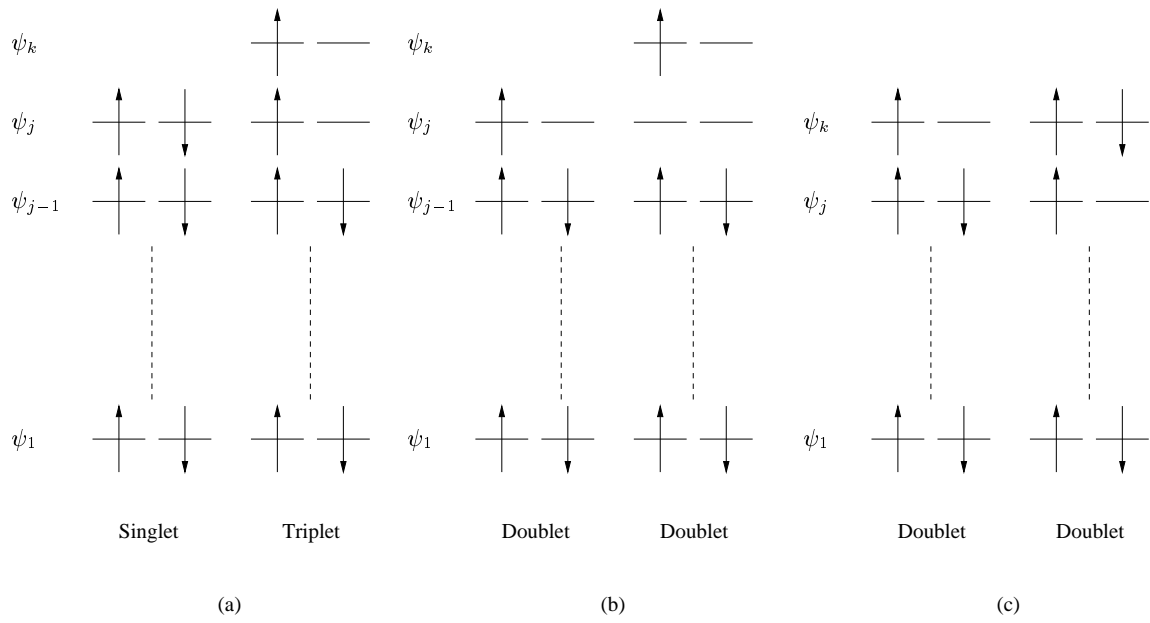


Figure 1

On the basis of the above guidelines, below is an example of the **molSOC.inp** file concerning the calculation of the SOC matrix elements of RbI^+ specie in $^2\Pi$ states:

```
#                                     Comment line
Zeff                                Route section
#                                     Comment line
+1 2 R 2 R                          Charge and multiplicities section
Rb  0.00000000E+00  0.00000000E+00 -3.53537734E+00  1.0 Molecular Geometry section
I   0.00000000E+00  0.00000000E+00  2.46809361E+00  1.0
End
```

In this job, the route section consists of a single three characters (the minimum allowed) keyword which activates the calculation of SOC matrix elements using the screened-nuclear charge method[3]. Moreover, the remaining computational options are accepted as default. The charge and multiplicities section specify that SOC contributions are calculated by coupling between a pair of doublet states and that β orbitals are involved in the coupling mechanism (see Figure 1c), both the identifier **O** are **R**.

3.3 The file basis

In the **basis** ASCII text file, basis set are specified by center definitions blocks separated to each other by the *Separator* *******. Each block has to be composed of a:

1. *Center identifier line*, specifying the atomic symbol (X) or one or more numbers belonging to the atomic order given in the *Molecular Geometry section* of the **molSOC.inp** file. This line must be ended by a **0**;
2. *Shell definition lines*, each one must specify the shell type (SH which can be **s**, **p**, **d**, **f**), the number of gaussian contraction coefficients and exponents (N_{GF}), and a scale factor λ (all gaussian exponents are scaled by λ^2);
3. *Gaussian contraction coefficients and exponents* ($d_{\mu i}$ and $\zeta_{\mu i}$), the first column must contain the $\zeta_{\mu i}$ values and the second one the corresponding $d_{\mu i}$.

The end of the basis set input must be indicated by the **END** statement or by a black line. Notice that it is enough to specify desired basis set of atomic centers of the same element once. Below is an example showing the structure of this file:

O	0		<i>Center identifier line</i>
s	6	1.00	<i>Shell definition line</i>
	5222.90220	-0.193640000E-02	<i>Exponents and contraction coefficients</i>
	782.539940	-0.148507000E-01	
	177.267430	-0.733187000E-01	
	49.5166880	-0.245116200	
	15.6664400	-0.480284700	
	5.17935990	-0.335942700	
s	2	1.00	<i>Shell definition line</i>
	10.6014410	0.788058000E-01	<i>Exponents and contraction coefficients</i>
	0.94231700	-0.567695200	
s	1	1.00	<i>Shell definition line</i>
	0.27747460	1.000000000	<i>Exponents and contraction coefficients</i>
p	4	1.00	<i>Shell definition line</i>
	33.4241260	0.175603000E-01	<i>Exponents and contraction coefficients</i>
	7.62217140	0.107630000	
	2.23820930	0.323525600	
	0.68673000	0.483222900	
p	1	1.00	<i>Shell definition line</i>
	0.19381350	1.000000000	<i>Exponents and contraction coefficients</i>
d	1	1.00	<i>Shell definition line</i>
	0.80000000	1.000000000	<i>Exponents and contraction coefficients</i>
***			<i>Separator</i>
H	0		<i>Center identifier line</i>
s	4	1.00	<i>Shell definition line</i>
	50.9991780	0.966050000E-02	<i>Exponents and contraction coefficients</i>
	7.48321810	0.737289000E-01	
	1.77746760	0.295858100	
	0.51932950	0.715905300	
s	1	1.00	<i>Shell definition line</i>
	0.15411000	1.000000000	<i>Exponents and contraction coefficients</i>
End			<i>End statement</i>

The format of the basis set is quite similar to that of the basis set given in the GAUSSIAN 03 input.

3.4 Description of the output files

After run MolSOC code it will generate five ASCII output files:

- (i) **molSOC.out** is the output containing the results;
- (ii) **coef** contains spherical and cartesian MO coefficients of the first and second state, respectively;
- (iii) **aover** and **bover** contain the two-state α and β molecular overlap integrals between the first and second state, respectively, after biorthogonalization;

- (iv) **soint** contains the one-electron SOC integrals over contracted cartesian gaussian functions.

3.5 Driving steps in some of the output files

It is important to pay attention to the number of α and β electrons belonging to each state (**molSOC.out**). These come from the integration of the α and β density and, since MolSOC is a support code of GAUSSIAN 03, they can be an important point verifying if the calculation is going correctly. Below is an example concerning still the RbI^+ specie relative to the calculation to the number of the electrons before of the Singular Value decomposition and after:

*** BEFORE SVD (BIORTHOGONALIZATION) ***

```
NUMBER OF ELECTRONS =    43.999999987
NUMBER OF ELECTRONS =    44.999999987
NUMBER OF ELECTRONS =    43.999999983
NUMBER OF ELECTRONS =    44.999999984
```

⋮

*** AFTER SVD (BIORTHOGONALIZATION) ***

```
NUMBER OF ELECTRONS =    43.999999987
NUMBER OF ELECTRONS =    44.999999987
NUMBER OF ELECTRONS =    43.999999983
NUMBER OF ELECTRONS =    44.999999984
```

```
SPIN DISCOINCIDENCE      =    0
ORBITAL DISCOINCIDENCE   =    1
COUPLING DETERMINANTS    =  BETA BETA
```

Here the spin discoincedence is zero and the orbital discoincedence is one (the multiplicities of the two states are the same). β orbitals are responsible of the coupling between these $^2\Pi$ states (see Figure 1c).

To verify if the biorthogonalization is performed successful, one needs to look the contents of the files **aover** and **bover** where diagonal elements of the two-state MO overlap integrals are printed:

aover

⋮

```
40 40    0.999999594
41 41    0.999995468
42 42    0.999992586
43 43    0.999726186
44 44    0.000000000
```

bover

⋮

```
41 41    0.999995468
42 42    0.999992586
43 43    0.999726186
44 44    0.999550567
45 45    0.999550567
```

Note that in this case (an orbital dis coincidence occurs) the 44th β orbitals (file **aover**), which are treated as α orbitals, are quite orthogonals after SVD.

Matrix elements are printed at the end of the file **molSOC.out** in this format:

```
*****
*** MATRIX ELEMENTS (CM-1) ***
*****
```

(I-STATE) MS	(II-STATE) MS	1-el		2-el		TOTAL	
		R	I	R	I	R	I
0.50	0.50	2430.7287	0.0000	0.0000	0.0000	2430.7287	0.0000
-0.50	-0.50	-2430.7287	0.0000	0.0000	0.0000	-2430.7287	0.0000
0.50	-0.50	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-0.50	0.50	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

References

- [1] Bethe, H. A.; Salpeter, E. E. *Quantum Mechanics of the One and Two Electron Atoms* (Plenum, New York, 1977).
- [2] Chiodo, S.; Russo, N. J Comput Chem, **29**, 912 (2008)
- [3] Chiodo, S. G.; Russo, N. J Comput Chem, **30**, 832 (2009)
- [4] Chiodo, S. G.; Russa, N. Chem Phys Lett, **490**, 90 (2010)
- [5] Quartarolo, A. D.; Chiodo, S. G.; Russo, N. J Chem Theory Comput, **6**, 3176 (2010)