

+++++

STRFIT - GENERAL STRUCTURE FITTING PROGRAM, FITTING EFFECTIVE STRUCTURAL  
PARAMETERS DEFINED IN THE 'CART' SCHEME TO ANY COMBINATION  
OF ROTATIONAL CONSTANTS

+++++

This program was originally written in 1990 for determination of effective  $r_0$  structures of hydrogen bonded complexes, but currently it can also be used to fit also  $r_z$ ,  $rm(2L)$ ,  $re^{(SE)}$  and other geometries

Citation and additional information on this program:  
Z.Kisiel, J. Mol. Spectrosc. 218, 58-67 (2003).

This STRFIT was written independently of, but seems to embody similar features to Schwendeman's STRFIT. Once I found this out it was too late to change the name.

See: R.H.Schwendeman, "Structural parameters from rotational spectra", in "Critical evaluation of chemical and physical structural information", D.R.Lide and M.A.Paul, Eds., National Academy of Sciences, Washington, D.C. (1974).

For a similar least-squares fitting approach see also:

P.Nosberger,A.Bauder, and Hs.H.Gunthard, "A versatile method for molecular structure determinations from ground state rotational constants", Chem.Phys. 1,418-425(1973).

The least-squares engine for fitting  $r_0$  internal parameters to ground state moments of inertia can now be used to evaluate also:

- 1/  $r.z$ ,  $re^{SE}$  or any other geometry fitted to moments of inertia modified by user supplied corrective contributions. For example, in the  $re^{SE}$  geometry (= semi-experimental equilibrium) the Be-B $\theta$  corrections from a CFOUR anharmonic force field calculation can be used directly
- 2/  $r.m(1r)$  and  $rm(1rL)$  geometry
- 3/  $r.m(2r)$  and  $rm(2rL)$  geometry

For  $r.m(1)$ ,  $r.m(2)$  and their subvariants see:

J.K.G.Watson, A.Roytburg, W.Ulrich, J.Mol.Spectrosc. 196,102-119(1999)

Declaration scheme of internal structural coordinates (colloquially called CART since it was adapted from a Wilson group program of that name):

H.R.Thompson, J.Chem.Phys. 47,3407-3410(1967)

Main features of STRFIT:

- fit of up to MAXPAR structural parameters to up to MAXCON measured rotational constants in up to MAXISO isotopic species
- each structural parameter can have up to twenty repetitions in the molecule
- up to MAXCHG structural parameters can be different between the parent and the substituted species
- non-linear Levenberg-Marquardt least squares fitting algorithm is used (Numerical Recipes Chapt.14)

- the fit is to (equally weighted) moments of inertia or planar moments and not to rotational constants, although when rotational constants are to be fitted they are declared conventionally
- the spectroscopic constants that can be fitted to are A,B,C, B+C, and A+B, and the three planar moments Pa,Pb,Pc (in any combination)
- the data set for STRFIT is compatible with PMIFST so the latter can be used to check the 'CART' definitions
- The user only needs to specify the generic name MOLNAM on startup, and the input/output files are then:

MOLNAM.STF = the input file

MOLNAM.OUT = the main output file with complete results of the fit

MOLNAM.PMI = the abbreviated output file containing at the top the connectivity declaration of the fitted structure for viewing with PMIFST

MOLNAM.XYZ = the fitted geometry in Cartesians for viewing with packages such as Chimera, JMol, MacMolPlot...

ver. 22.V.2015

----- Zbigniew KISIEL -----

Institute of Physics, Polish Academy of Sciences Al.Lotnikow 32/46, Warszawa, POLAND kisiel@ifpan.edu.pl <a href="http://info.ifpan.edu.pl/~kisiel/prospe.htm">http://info.ifpan.edu.pl/~kisiel/prospe.htm</a>
---

Modification history:

- 1990: Creation
- 24.06.95: PARAMETER dimensioning
- 10.12.99: Fitting of B+C and A+B
- 30.08.00: Echo of fitted structure to STRFIT.PMI
- 8.01.01: Various incremental modifications
- 8.04.02: Fitting also of planar moments
- 19.04.02: Elimination of bug in writing STRFIT.PMI
- 4.06.02: rm(1) and rm(2) model + fixed parameters + overhaul
- 29.10.02: increase in MAXPAR and associated corrections
- 30.01.03: fitting equal c or d constants
- 8.08.05: correction of 'final principals' bug spotted by Stew Novick
- 14.11.08: correction to mass used for d terms, spotted by Jean Demaison
- 25.05.09: elimination of erroneous temporary code that crept in about 2005 in Laurie parameter + improvement of symmetric top performance
- 15.02.10: allowance also for dB=B\*-Bexp input and better input debugging
- 15.12.10: optional user parameter descriptors and updated input description
- 17.02.12: uncertainties in principal coordinates
- 12.12.12: increased dimensioning
- 6.03.13: allowance for exclusion from fit of some of the declared constants
- 6.05.15: parameter repetitions with changes
- 22.05.15: generic i/o file names and addition of .xyz output

+++++

STRFIT DATA FILE STRUCTURE - specimen data file is reproduced below, with lines numbered to facilitate description:

lno.	first column of the data file	32nd column of data file (beginning of numerical input for all lines below the CART deck)
1.	buteneyne...HF, HF at the TRIPLE bond, in the bey plane	
2.	11 0	
3.	1 0 0 0 .000000 .000000 .000000 12.000000	
4.	2 1 0 0 1.208600 .000000 .000000 12.000000	
5.	3 2 1 0 1.431000 177.900000 .000000 12.000000	
6.	4 3 2 1 1.341600 123.100000 180.000000 12.000000	
7.	5 1 2 4 1.062000 182.300000 180.000000 1.007825	
8.	6 4 3 2 1.087000 121.600000 .000000 1.007825	
9.	7 4 3 2 1.087000 118.700000 180.000000 1.007825	
10.	8 3 4 7 1.087000 121.700000 0.000000 1.007825	
11.	9 2 3 8 0.604300 177.900000 .000000 .000001	
12.	10 9 2 3 3.080000 90.000000 180.000000 18.998404	
13.	11 10 9 2 0.925700 360.000000 .000000 1.007825	
14.	TOTAL NUMBER OF PARAMETERS: 2	
15.	atom no., parameter no. 10 1 0	
16.	atom no., parameter no. 9 1 0	
17.	NO OF CONSTANTS TO FIT TO: 6	
18.	constant,species,value 1 1 5182.82917	
19.	constant,species,value 2 1 3312.90406	
20.	constant,species,value 3 1 2021.475817	
21.	constant,species,value 1 2 5182.55955	
22.	constant,species,value 2 2 3271.83826	
23.	constant,species,value 3 2 2005.91671	
24.	NO OF CHANGES FROM PARENT SP.: 2	
25.	atom no.,parameter no.,value 11 4 2.0141022	
26.	atom no.,parameter no.,value 11 1 -0.0023	

The various lines are:

line 1: descriptive comment (truncated to 78 characters if longer)

-----  
 lines 2-13 - block containing the CART definition of the molecular geometry  
 -----

line 2: The number of atoms in the molecule, read in with format=(I5)  
 Optionally a second number (also in an I5 field) can be included,  
 and if equal to 1 specifies output of debugging information.

line 3+: CART DEFINITION LINE: N,NA,NB,NC,R,A,D,W

this is read in free format, but N,NA,NB,NC are to be integers, and  
 R,A,D,W are to be floating point values. There should be as many  
 such lines as the number of atoms declared in Line 2.

The structure of the CART line is defined in a separate set of  
 comments further below.

CART lines declare either real or dummy atoms, and the use of dummy  
 atoms is encouraged since it may considerably simplify the declaration  
 of the molecule. In the example above a dummy atom  
 (carrying an infinitesimally small but non-zero mass) defines  
 the centre of the triple bond to which the HF molecule  
 is notionally bonded. The program will actually use the mass in  
 the calculation so specify as low as possible (but non-zero mass)  
 and mass of 0.000001u is recommended.

Dummy atoms should not be counted in the evaluation of the  $rm^2$   
 d-parameters and are identified as such if their mass is less



- V - the optional initial value for the declared parameter, which will override the value in the preceding CART deck
- AD - the optional alphanumeric descriptor of this parameter (up to 20 characters long), which will be echoed in the output. This has to start in column 72.
- FIX - the command to fix this parameter in the fit, details at the end of 15b below

line 15b: DEFINITION OF AN ANCILLARY PARAMETER OF FIT = N,M,I,V

this is read in with format=(31x,3i2,f14.0) so that the numbers have to be EXACTLY in the columns indicated below:

```

column: 32      38              51
         |      |              |
         ---+---+-----+-----+

```

Examples:

```

rm(1) parameter c_a.....-4 1 0   0.001
rm(2) parameter_d_b.....-5 2 0
rm(2) parameter_d_b.....FIX-5 2 0   0.001

```

N - the code (always negative) defining the type of the ancillary parameter, and also the meaning of M,I,V

N = -4 Watson's rm(1) c parameter, in which case  
M = 1,2,3,4,5,6 for c\_a,c\_b,c\_c,c\_ab,c\_ac,c\_bc  
= 7 for fitting of c\_a = c\_b  
= 8 for fitting of c\_a = c\_c  
= 9 for fitting of c\_b = c\_c  
= 10 for fitting c\_a = c\_b = c\_c

I = 0

V = optional initial value for this parameter

N = -5 Watson's rm(2) d parameter, in which case

M = 1,2,3 for d\_a,d\_b,d\_c  
= 7 for fitting of d\_a = d\_b  
= 8 for fitting of d\_a = d\_c  
= 9 for fitting of d\_b = d\_c  
= 10 for fitting d\_a = d\_b = d\_c

I = 0

V = optional initial value for this parameter, which will override the value in the CART definitions

N = -6 Laurie correction term for an XH bond, in which case

M = the atom for which this bond is defined

I = the number of additional bonds subject to the same correction: if this is greater than 0 then an additional line, format=(a31,20i2) should be inserted immediately following, with numbers of atoms for which these internals are defined

V = optional initial value for this parameter

N = -7 experimental Laurie term defined in terms of a projection angle

M = the atom for which this angle is defined

I = the number of additional angles subject to this correction, and I>0 requires an additional line, format=(a31,20i2), specifying the numbers of atoms carrying such angles

V = optional initial value for this parameter

N = -8 isotopic bond length change relative to the value in the parent - this is only for use with Deuterium substitution.

M = the atom for which this bond is defined  
 I = the number of additional bonds subject to this correction,  
 and I>0 requires an additional line, format=(a31,20i2), specifying  
 the numbers of atoms carrying such bonds  
 V = optional initial value for this parameter

FIX = Any of the declared internal or ancillary parameters can be fixed  
 instead of fitted by writing FIX in columns 29-31 of the  
 declaration line.

This is one of only two cases for which the text in  
 in columns 1-31 is relevant (the other is XXX for an excluded  
 spectroscopic constant). Otherwise this field is ignored and  
 can be used for arbitrary explanatory comments.

-----  
 lines17-23 - block defining the spectroscopic constants to be fitted to  
 -----

line 17: The total number of available spectroscopic constants.

line 18+: DEFINITION OF THE SPECTROSCOPIC CONSTANT TO BE FITTED TO: N,M,B,dI,dB

read in with format=(31x,2i2,3f14.0) so that the numbers have  
 to be EXACTLY in the columns indicated below:

```

column: 32 36          50          64          78
         |  |          |          |          |
         +-+-----+-----+-----+-----+

```

Examples:

```

Constant,species,value..... 1 1 312142.390          NH2CN
Constant,species,value..... 1 1 312142.390          1.0          NH2CN
Constant,species,value..... 1 1 312142.390          30000.          NH2CN
Constant,species,value....XXX 1 1 312142.390          NH2CN

```

N - the code defining the type of spectroscopic constant  
 = 1,2,3 for A,B,C  
 = 4,5 for B+C and A+B  
 = 6,7,8 for P.a,P.b,P.c

When B+C and A+B are declared, the fit is actually to  
 $4*505379.01/(B+C) = 4*I.b.I.c/(I.b+I.c) = ca. I.b+I.c$  or to  
 $4*505379.01/(A+B) = 4*I.a.I.b/(I.a+I.b) = ca. I.a+I.b$

M = the number of the isotopic species  
 B = the value of the constant, in MHz for  
 rotational constants and uA\*\*2 for planar moments.  
 The constants should be declared in  
 increasing order of isotopic species number, all constants for  
 a given species are not required.

It is optionally possible to specify EITHER fourth OR fifth numbers,  
 dI or dB (only one allowed), such that:

dI - vibration-rotation correction  
 dI=Icorr-Iobs (in uA\*\*2)  
 to the experimental moment of inertia defined by N and B (even if  
 experimental rotational constants are specified). In this case  
 the dB field should be left blank.

dB - vibration-rotation correction  
 dB=Bcorr-Bobs (in MHz)  
 to the experimental rotational constant defined by N and C. In  
 this case the space corresponding to dI is to be left blank.

The structure will then be fitted to corrected quantities  
 $I_{corr} = I_{obs} + dI$ ,  $I_{obs} = conv / B_{obs}$  or to  
 $I_{corr} = conv / (B_{obs} + dB)$   
 where  $I_{obs}$  is the moment of inertia derived from the preceding rotational  
 constant, or to  $P = P_{obs} + dI$  if planar moment has been specified

The main purpose of the dI option is evaluation of  $r^*$  (r.z)  
 structures by using dI's from a harmonic force field  
 The main purpose of the dB option is evaluation of  
 semi-experimental equilibrium ( $r_e^{SE}$ ) structures by using  
 dB's from an anharmonic force field calculation

XXX = Any of the input rotational constants can be excluded from the fit  
 by writing XXX in columns 29-31 of the declaration line.  
 This is one of two only two cases for which the text in  
 in columns 1-31 is relevant. Otherwise this field is ignored and  
 can be used for arbitrary explanatory comments.

-----  
 lines24-26 - Block defining the isotopic changes  
 -----

There is to be one such block for each substituted isotopic species  
 represented in the spectroscopic constants. Species 1 is taken to  
 be the parent species, the first isotopic block is to be for  
 species number 2, and successive such blocks are to be for  
 successively numbered species.  
 Each block begins with a repeat count line (line 24), which is  
 followed by the declaring lines.

line 24: The number of changes relative to the parent for a given  
 isotopic species

line 25+ - DEFINITION OF ISOTOPIIC CHANGE = N,M,V

read in with format=(31x,2i2,3f14.0) so that the numbers have  
 to be EXACTLY in the columns indicated below:

```

column:  32  36          49
         |  |          |
         +-+-----+
  
```

N = the number of the isotopic atom  
 M = the number of the changed parameter (1,2,3,4 for  
 bond length, bond angle, dihedral angle and atomic mass resp.),  
 V = the value of this parameter in the substituted species.  
 When the parameter is not a mass (N not equal to 4) then V declares  
 the change relative to the value in the parent species

-----  
 GENERAL NOTES:

- As noted in several places above the input is in the  
 old-fashioned fixed format, so it is necessary to keep input  
 numbers within the specified columns.
- Columns 1-31 can be used for any clarifying comments, which are  
 up to the user and are not echoed to the output
- Columns to the right of the input fields can be also be used for  
 comments, which are not echoed to the output.  
 Usually columns 78 and higher are safe for this purpose except in  
 the definitions of parameters of fit where the descriptive  
 alphanumeric parameter extends up to and inclusive of column 91

- The input file is not read past the last set of isotopic changes, so the rest of the file can be used for any purpose - usually storage of alternative data and of results.
- STRFIT does not write to the input file so it will not corrupt it
- Additional blocks of comments are allowed before all lines with a repeat count: i.e. those that specify:
  - the number of atoms,
  - the number of parameters,
  - the number of spectroscopic constants
  - the number of isotopic changes.These comment lines have to have a '!' character in the first column, and there is no limit on their number. Such lines are echoed at the top of the output file.

WARNING: There is a known problem with symmetric top molecules which break symmetry on isotopic substitution. This is associated with the way the program implements parallel axes for the rm(1r) scheme and the resulting possibility of perpendicular axis switching. Steps have been taken to stabilise such behaviour, and those are marked by the string May2009 in the code

-----  
CART = Thompson J.Chem.Phys. 47,3407(1967) definitions:  
-----

1. Atom N in the molecule is defined in terms of three other neighbouring atoms: NA, NB, and NC, and a bond length  $R(N-NA)$ , a bond angle  $A(N-NA-NB)$  and a dihedral angle  $D(N-NA-NB-NC)$
- 1a. Atoms are declared in the sequence of increasing N (starting from N=1) and the numbers of the defining atoms NA, NB, NC all have to be smaller than N (and all different)
- 1b. For N=1 set NA,NB,NC, R,A,D to zero  
For N=2 set NB,NC, A,D to zero  
For N=3 set NC, D to zero
2. For a right handed coordinate system dihedral angles are defined by viewing the configuration
  - i/ from the NC direction along the NB-NA axis if NC is bonded to NB
  - ii/ from the NC direction along the NA-NB axis if NC is bonded to NAThe angle is defined positive for a clockwise rotation of atom NC into the plane N-NA-NB.
3. In the calculation of initial Cartesian coordinates the atom N=1 defines the origin, the line N=1 to N=2 defines the X axis and atoms N=1, N=2, N=3 define the XY plane.

-----  
C O M P I L A T I O N:  
-----

If you want to compile this program then please note that it is crucially dependent on the use of 'static memory allocation'. Please read the following notes copied from the PROSPE webpage (and also use the data files from the webpage for testing):

Most contemporary compilers no longer use default static memory allocation that preserves values of variables previously set in some subroutine. Many programs assume this to be the case so that you need to use an appropriate keyword to enforce static allocation. For example on the f77 compiler on SGI machines this is the -static option, whereas with Intel Visual Fortran you have to use option -Qsave. Note that with some compilers optimization options used not to be safe. With good compilers this is no longer the case, but if problems crop up it is still good to check by disabling optimization.

Command line example for gfortran:  
gfortran -fno-automatic strfit.for -o strfit

Command line example for Intel Visual Fortran 9:  
ifort -nopdbfile -nodebug -traceback -arch=pn1 -tune=pn1 -O2 -Qsave  
-ccdefault:fortran -fpscomp:filesfromcmd strfit.for

Command line example for Intel Visual Fortran 11:  
ifort -nopdbfile -nodebug -traceback -O3 -Qsave -ccdefault:fortran  
-fpscomp:filesfromcmd strfit.for

---