

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

C+++++
C
C This program was originally written in 1990 for determination of
C effective r_0 structures of hydrogen bonded complexes, but it can also
C be used to fit r_z , $r_m^{(1L)}$, $r_m^{(2L)}$, $r_e^{(SE)}$ and other geometries.
C

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

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

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

C For a similar least-squares fitting approach see also:

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

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

C 1/ r_z , r_e^{SE} or any other geometry fitted to moments of inertia
C modified by user supplied corrective contributions. For example,
C in the r_e^{SE} geometry (= semi-experimental equilibrium) the Be-B0
C corrections from a CFOUR anharmonic force field calculation can
C be used directly. The much smaller rotational g-tensor contributions
C can also be used.

C 2/ $r_m^{(1r)}$ and $r_m^{(1rL)}$ geometry

C 3/ $r_m^{(2r)}$ and $r_m^{(2rL)}$ geometry
C

C For $r_m^{(1)}$, $r_m^{(2)}$ and their subvariants see:

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

C Reviews dealing (among other topics) with r_e^{SE} evaluations:

C J.Demaison, Molec.Phys. 105,3109-3138(2007)

C C.Puzzarini et al., Int.Rev.Pjys.Chem. 29,273-367(2010)
C

C Demonstration of the usefulness of rotational g-tensor corrections
C in r_e^{SE} geometry evaluations:

C A.G.Csaszar, J.Demaison, H.D.Rudolph, "Equilibrium structures of three-
C four-, five-, six-, and seven-membered unsaturated N-containing
C heterocycles", J.Phys.Chem.A 119,1731-1746(2015)
C

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

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

C Main features of STRFIT:
C

C - fit of up to MAXPAR structural parameters to up to MAXCON measured
 C rotational constants in up to MAXISO isotopic species
 C
 C - each structural parameter can have up to twenty repetitions in the
 C molecule
 C
 C - up to MAXCHG structural parameters can be different between the parent
 C and the substituted species
 C
 C - non-linear Levenberg-Marquardt least squares fitting algorithm is used
 C (Numerical Recipes Chapt.14)
 C
 C - the fit is to (equally weighted) moments of inertia or planar moments
 C and not to rotational constants, although when rotational constants
 C are to be fitted they are declared conventionally
 C
 C - the spectroscopic constants that can be fitted to are A,B,C, B+C,
 C and A+B, and the three planar moments Pa,Pb,Pc (in any combination)
 C
 C - the data set for STRFIT is compatible with PMIFST so the latter can
 C be used to check the 'CART' definitions
 C
 C - The user only needs to specify the generic name MOLNAM on startup,
 C and the input/output files are then:
 C
 C MOLNAM.STF = the input file
 C
 C MOLNAM.OUT = the main output file with complete results of the fit
 C MOLNAM.PMI = the abbreviated output file containing at the top
 C the connectivity declaration of the fitted structure
 C for viewing with PMIFST
 C MOLNAM.XYZ = the fitted geometry in Cartesians for viewing with packages
 C such as Chimera, JMol, MacMolPlot...

ver. 8.IV.2020

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

Institute of Physics, Polish Academy of Sciences Al.Lotnikow 32/46, Warszawa, POLAND kisiel@ifpan.edu.pl http://info.ifpan.edu.pl/~kisiel/prospe.htm

Modification history:

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

C 8.10.16: improved diagnostics of errors in the input data file
C 25.07.19: electronic correction in r_e^SE evaluation + misc. updates
C 8.04.20: some output and documentation changes

C
C+++++

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

C first column of the data file
C 32nd column of data file (beginning of
C numerical input for all lines
C below the CART deck)

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

C
C The various lines are:
C line 1: descriptive comment (truncated to 78 characters if longer)

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

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

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

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

C The structure of the CART line is defined in a separate set of
C comments further below.
C CART lines declare either real or dummy atoms, and the use of dummy
C atoms is encouraged since it may considerably simplify the declaration
C of the molecule. In the example above a dummy atom

C (carrying an infinitesimally small but non-zero mass) defines
 C the centre of the triple bond to which the HF molecule
 C is notionally bonded. The program will actually use the mass in
 C the calculation so specify as low as possible (but non-zero mass)
 C and mass of 0.000001u is recommended.
 C Dummy atoms should not be counted in the evaluation of the rm^2
 C d-parameters and are identified as such if their mass is less
 C than 0.00001u.
 C
 C

C -----
 C lines14-16 - block defining the parameters of fit
 C -----
 C

C line 14: The total number of declared parameters. These parameters can either
 C be parameters of fit or are to be held fixed at declared values.
 C The parameters can be either internal structural coordinates, or
 C ancillary parameters such as c,d parameters of the $rm(1)$ and
 C $rm(2)$ models, Laurie δ_H parameters, or empirical bond length
 C change (mostly for use with deuteration)
 C

C Parameter of fit declaration lines 15+ can be in two different forms, 15a or 15b:
 C

C line 15a: DEFINITION OF AN INTERNAL COORDINATE:
 C N,I,J,V,AD
 C

C this is read in with input format=(31x,3i2,f14.0,20x,A20) so that
 C the numbers have to be EXACTLY in the columns indicated below:
 C

column:	32	38	51	72	91
	-----			cccccccccccccccccccc	

C Examples:
 C

atom no., parameter no.	7 1 0	r(C-C)
atom no., parameter no.	7 2 2	CCH (in CH3)
	8 9	
atom no., parameter no.	FIX 8 2 0 15.346	fixed at lit.value

C N - the number of atom containing the required internal coordinate in
 C its CART definition
 C I - the CART parameter for that atom corresponding to the required internal
 C (1=bond length, 2=bond angle, 3=dihedral angle)
 C
 C J - the optional repeat count declaring whether this parameter
 C has any additional occurrences:
 C J = 0 means that the parameter is unique
 C J > 0 makes it possible to fit 'degenerate' parameters,
 C occurring for more than one atom (eg. CX bond in a CX3 pyramidal
 C segment) and the value of J defines the number
 C of additional occurrences of the parameter (e.g. J=2 for bond
 C lengths in a methyl group). An additional line
 C in format =(a31,20i2) should then be inserted immediately
 C following the current definition line and containing
 C the numbers of atoms also carrying this parameter.
 C Up to twenty such additional occurrences are allowed.
 C NOTE: symmetric +/- dihedral angle values can be fitted by
 C specifying the values which have opposite sign to that of
 C the parent value by using a negative number for the respective
 C atom. Regrettably, owing to current formatting restrictions,
 C the numbers of such atoms have to be smaller
 C than 10 (single digit)
 C J < 0 makes it possible to fit 'degenerate' parameters with changes.
 C This feature is designed primarily for dihedral angles in
 C order to allow implementation of local rotational symmetry based

C on a non-zero rotation angle for atom N.
C This line is to be followed by |J| lines, each containing
C two numbers: integer number of the atom carrying the repeated value,
C and the difference of this value relative to that for atom N.
C The entries are to start from column 32, but are free-format
C otherwise. Negative atom numbers are allowed, and these will
C specify a reversed sign value of V (see below). With this method
C there is no limitation on the atom number that is present for J>0.
C

C V - the optional initial value for the declared parameter,
C which will override the value in the preceding CART deck

C AD - the optional alphanumeric descriptor of this parameter
C (up to 20 characters long), which will be echoed in the output.
C This has to start in column 72.

C FIX - the command to fix this parameter in the fit, details at the end
C of 15b below
C

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

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

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

C Examples:

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

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

C N = -4 Watson's rm(1) c parameter, in which case

C M = 1,2,3,4,5,6 for c_a,c_b,c_c,c_ab,c_ac,c_bc

C = 7 for fitting of c_a = c_b

C = 8 for fitting of c_a = c_c

C = 9 for fitting of c_b = c_c

C = 10 for fitting c_a = c_b = c_c

C I = 0

C V = optional initial value for this parameter
C

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

C M = 1,2,3 for d_a,d_b,d_c

C = 7 for fitting of d_a = d_b

C = 8 for fitting of d_a = d_c

C = 9 for fitting of d_b = d_c

C = 10 for fitting d_a = d_b = d_c

C I = 0

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

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

C M = the atom for which this bond is defined

C I = the number of additional bonds subject to the same correction: if

C this is greater than 0 then an additional

C line, format=(a31,20i2) should be inserted immediately

C following, with numbers of atoms for which these internals

C are defined
C

C V = optional initial value for this parameter
C

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

```

C      M = the atom for which this angle is defined
C      I = the number of additional angles subject to this correction,
C          and I>0 requires an additional line, format=(a31,20i2), specifying
C          the numbers of atoms carrying such angles
C      V = optional initial value for this parameter
C
C      N = -8 isotopic bond length change relative to the value in
C          the parent - this is only for use with Deuterium
C          substitution.
C      M = the atom for which this bond is defined
C      I = the number of additional bonds subject to this correction,
C          and I>0 requires an additional line, format=(a31,20i2), specifying
C          the numbers of atoms carrying such bonds
C      V = optional initial value for this parameter
C
C      FIX = Any of the declared internal or ancillary parameters can be fixed
C            instead of fitted by writing FIX in columns 29-31 of the
C            declaration line.
C            This is one of only two cases for which the text in
C            in columns 1-31 is relevant (the other is XXX for an excluded
C            spectroscopic constant). Otherwise this field is ignored and
C            can be used for arbitrary explanatory comments.
C
C-----
C lines17-23 - block defining the spectroscopic constants to be fitted to
C-----
C
C   line 17:  The total number of available spectroscopic constants.
C
C   line 18+:  DEFINITION OF THE SPECTROSCOPIC CONSTANT TO BE FITTED TO:
C              N,M,B,dI,dB
C
C              read in with format=(31x,2i2,3f14.0) so that the numbers have
C              to be EXACTLY in the columns indicated below:
C
C              column:  32  36              50              64              78
C                      |  |              |              |              |
C                      +-+-----+-----+-----+-----+
C
C      Examples:
C
C Constant,species,value..... 1 1 312142.390              NH2CN
C Constant,species,value..... 1 1 312142.390          1.0              NH2CN
C Constant,species,value..... 1 1 312142.390              30000.              NH2CN
C Constant,species,value....XXX 1 1 312142.390              NH2CN
C Constant,species,value..... 1 1 312142.390          0.1          30000.              NH2CN
C
C      N - the code defining the type of spectroscopic constant
C          = 1,2,3 for A,B,C
C          = 4,5 for B+C and A+B
C          = 6,7,8 for P.a,P.b,P.c where P.a=(1/2)[Ib+Ic-Ia] etc.
C
C      When B+C and A+B are declared, the fit is actually to
C       $4*505379.01/(B+C) = 4*I.bI.c/(I.b+I.c) = ca. I.b+I.c$  or to
C       $4*505379.01/(A+B) = 4*I.aI.b/(I.a+I.b) = ca. I.a+I.b$ 
C
C      M = the number of the isotopic species
C      B = the value Bobs of the experimentally determined constant, in MHz for
C          rotational constants and uA**2 for planar moments.
C          The constants should be declared in increasing order of
C          isotopic species number, although all constants for
C          a given species are not required.
C
C      It is optionally possible to specify EITHER fourth OR fifth numbers,
C      dI or dB (usually only one allowed), such that:

```

C
 C dI - vibration-rotation correction $dI = I_{\text{corr}} - I_{\text{obs}}$ (in uA^{**2})
 C to the experimental moment of inertia defined by N and B (even if
 C experimental rotational constants are specified). In this case
 C the dB field should be left blank.
 C
 C dB - vibration-rotation correction $dB = B_{\text{corr}} - B_{\text{obs}}$ (in MHz)
 C to the experimental rotational constant defined by N and M.
 C
 C --> this is most often (Be-B0) from anharmonic force field calculations
 C and used for r_e^{SE} structure evaluation
 C
 C In this case the space corresponding to dI is normally to be left blank.
 C BUT if a value is placed in this position then it is assumed to be
 C the (dimensionless) rotational g-factor, g_{bb} , for rotational constant B.
 C STRFIT will then subtract from the pertinent rotational constant
 C an electronic correction dB_{el} (in MHz) = $0.000544617 * g_{\text{bb}} * B$,
 C where B is taken to be $B + dB$.
 C
 C The structure will then be fitted to corrected quantities
 C $I_{\text{corr}} = I_{\text{obs}} + dI$, $I_{\text{obs}} = \text{conv} / B_{\text{obs}}$ or to
 C $I_{\text{corr}} = \text{conv} / (B_{\text{obs}} + dB)$ or to
 C $I_{\text{corr}} = \text{conv} / (B_{\text{obs}} + dB - dB_{\text{el}})$
 C where I_{obs} is the moment of inertia derived from the preceding
 C rotational constant, or to $P = P_{\text{obs}} + dI$ if planar moment has been
 C specified
 C
 C The main purpose of the dI option is evaluation of r^* (r_z)
 C structures by using dI's from a harmonic force field
 C
 C The main purpose of the dB option is evaluation of
 C semi-experimental equilibrium (r_e^{SE}) structures to either
 C the specified $B = B_{0\text{exp}} + (B_e - B_0)_{\text{calc}}$
 C or to $B = B_{0\text{exp}} + (B_e - B_0)_{\text{calc}} - dB_{\text{el}}$
 C
 C XXX = Any of the input rotational constants can be excluded from the fit
 C by writing XXX in columns 29-31 of the declaration line.
 C This is one of two only two cases for which the text in
 C in columns 1-31 is relevant. Otherwise this field is ignored and
 C can be used for arbitrary explanatory comments.
 C
 C
 C -----
 C lines24-26 - Block defining the isotopic changes
 C -----
 C
 C There is to be one such block for each substituted isotopic species
 C represented in the spectroscopic constants. Species 1 is taken to
 C be the parent species, the first isotopic block is to be for
 C species number 2, and successive such blocks are to be for
 C successively numbered species.
 C Each block begins with a repeat count line (line 24), which is
 C followed by the declaring lines.
 C
 C line 24: The number of changes relative to the parent for a given
 C isotopic species
 C
 C line 25+ - DEFINITION OF ISOTOPIC CHANGE:
 C N,M,V
 C
 C read in with format=(31x,2i2,3f14.0) so that the numbers have
 C to be EXACTLY in the columns indicated below:
 C
 C column: 32 36 49
 C | | |
 C +-+-----+

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

C-----
C
C GENERAL NOTES:

- C - As noted in several places above the input is in the
C old-fashioned fixed format, so it is necessary to keep input
C numbers within the specified columns.
- C - Columns 1-31 can be used for any clarifying comments, which are
C up to the user and are not echoed to the output
- C - Columns to the right of the input fields can be also be used for
C comments, which are not echoed to the output.
C Usually columns 78 and higher are safe for this purpose except in
C the definitions of parameters of fit where the descriptive
C alphanumeric parameter extends up to and inclusive of column 91
- C - The input file is not read past the last set of isotopic changes,
C so the rest of the file can be used for any purpose - usually
C storage of alternative data and of results.
- C - STRFIT does not write to the input file so it will not corrupt it
- C - Additional blocks of comments are allowed before all lines
C with a repeat count: i.e. those that specify:
C the number of atoms,
C the number of parameters,
C the number of spectroscopic constants
C the number of isotopic changes.
C These comment lines have to have a '!' character in the first
C column, and there is no limit on their number. Such lines are
C echoed at the top of the output file.

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

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

- C 1. Atom N in the molecule is defined in terms of three other neighbouring
C atoms: NA, NB, and NC, and a bond length R(N-NA), a bond angle
C A(N-NA-NB) and a dihedral angle D(N-NA-NB-NC)
C
- C 1a. Atoms are declared in the sequence of increasing N (starting from N=1)
C and the numbers of the defining atoms NA, NB, NC all have to be smaller
C than N (and all different)
C
- C 1b. For N=1 set NA,NB,NC, R,A,D to zero
C For N=2 set NB,NC, A,D to zero
C For N=3 set NC, D to zero
C
- C 2. For a right handed coordinate system dihedral angles are defined

C by viewing the configuration
C i/ from the NC direction along the NB-NA axis if NC is bonded to NB
C ii/ from the NC direction along the NA-NB axis if NC is bonded to NA
C The angle is defined positive for a clockwise rotation
C of atom NC into the plane N-NA-NB.
C

C 3. In the calculation of initial Cartesian coordinates the atom N=1
C defines the origin, the line N=1 to N=2 defines the X axis and
C atoms N=1, N=2, N=3 define the XY plane.
C
C

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

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

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

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

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

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