

“Readme” file for the RAM36 code

(Rho-Axis Method for 3 and 6 fold barriers)

This is a “readme” file for a new program for fitting rotation–torsion energy levels in molecules in which the frame has C2v symmetry and the top has C3v symmetry (like toluene or nitromethane). Details of the theoretical approach are given in Ref. 1, where this V6 code is applied to the analysis of the toluene spectrum. One more successful application of the program (at the time of writing this readme file) may be found in Ref. 2.

Although the program was specifically designed for the V6 case of internal rotation, it also has a V3 mode suitable for molecules with a Cs frame and C3v top (like acetic acid, acetamide or methyl formate).

The V6 code uses a general expression of the form

$$H = (1/2) \sum_{knpqrs} B_{knpqrs0} [J_z^{2k} J_x^n J_y^p J_\alpha^q p_\alpha^r \cos(3s\alpha) + \cos(3s\alpha) p_\alpha^r J_y^q J_x^p J_z^n J^{2k}] \\ + (1/2) \sum_{knpqrt} B_{knpqrt} [J_z^{2k} J_x^n J_y^p J_\alpha^q p_\alpha^r \sin(3t\alpha) + \sin(3t\alpha) p_\alpha^r J_y^q J_x^p J_z^n J^{2k}], \quad (1)$$

which makes it possible to include almost any symmetry allowed torsion-rotation Hamiltonian term in the fitting Hamiltonian. By “almost any” we mean that we checked for overflows up to the 12th order of the parameters. It is possible that higher orders will be also ok, but we didn’t check this.

The program can treat the data both in MHz and in cm⁻¹, and provides the usual least squares fit, where the data are weighted by the inverse of their measurement uncertainties squared. There is a special treatment of blended transitions. If several transitions in the input file have the same frequency, then such transitions are considered to be blended, and the intensity weighted average of the calculated transition frequencies is put in correspondence to the measured line frequency in the fit. Two additional fitting options are built into the program. The first one is the robust fitting option, where the weighting scheme from Ref. 3 is implemented. The second one is the possibility to set a cut off for obs-calc in the fit, which excludes from fitting in the current iteration all transitions which have ratio of (obs.-cal.)/uncertainty larger then the selected cut-off.

In addition to the fitting option, the program has a prediction option which provides a calculation of the spectrum for a selected range of J quantum numbers and a selected frequency region. Since, particularly for the low-barrier case, normal asymmetric-rotor selection rules are not always a good guide for deciding which lines are the strongest, the program calculates all possible $\Delta K_a, \Delta K_c$ transition frequencies and intensities allowed by symmetry for a given pair of upper and lower-state J values, and displays all transitions with intensities above a user-set cutoff.

The rest of the “readme” file will be devoted to a brief description of the input file and the output generated by the program.

Input file

The input file should be present in the same directory as the *.exe file containing the V6 code. Its name should be “input.txt”.

The first 4 lines look as follows:

```
3, 6,           !isigma, idelm
10, 9, 0,       !ktronc, nvt, ispecRHO
251, 30000, 50, 0, !numpar, numlines, numcategories, norder
0.1E-04, 20.d0   !stopping criteria; maximum ratio (o-c)/uncertainty for fitted lines
```

The first line should contain 3, 6 for V6 mode of the program and 1, 3 for V3 mode of the program. This is the only key that switches the program between V6 and V3 modes.

The second line sets up: (1) the number of states treated in the first diagonalization step (ktronc =10 in the current example means that the number of torsional basis states is $2 \cdot \text{ktronc} + 1$); (2) the number of states treated in the second diagonalization step (nvt=9 in the current example means that the lowest 9 torsional energy levels of a given symmetry are included); and (3) a switch 0 or ± 1 between two rather different definitions for the parameter RHO. If this switch ispecRHO=0 then the F and RHO will appear in the fit as $F \cdot \text{pa}^2 + \text{RHO} \cdot \text{pa} \cdot \text{Jz}$, so that RHO has the same energy units as F. If ispecRHO=1 then F and RHO will appear as $F \cdot (\text{pa} + \text{RHO} \cdot \text{Jz})^2$, so that RHO is unitless. If ispecRHO= -1 then F and RHO will appear as $F \cdot (\text{pa} - \text{RHO} \cdot \text{Jz})^2$, with RHO is again unitless.

The third line gives: the number of allowed parameters (numpar=251), the number of lines (numlines=30000), the number of data categories grouped by uncertainties (numcategories=50) and the order of Hamiltonian parameters (norder=0). The first three may be any reasonable positive integer you want, but you should make sure that the actual number of parameters, lines, and different data group uncertainties in the input file will be less than these numbers. The last one - norder - determines the maximum order of terms which will be generated by the program to provide you a list of all possible symmetry allowed terms up to the selected order (again checked to work up to 12th order). This list will appear in a separate file “testparam.txt”.

The fourth line contains the stopping criterion and the maximum obs.-cal. which is allowed during fitting iterations. The fit stops if the change (compared to the previous iteration) in the sum of obs.-cal. residuals squared is less than 0.1E-04 in the current example. The meaning of maximum ratio (obs.-cal.)/uncertainty allowed in the fitting process is described above.

The next 6 lines should be kept as they are. They are skipped by the program and simply remind the user of the format of the Hamiltonian parameters. The operator expression below will be referred to as Exp. (1).

```
-----
H=Sum{1/2(J**2k*Jz**n*Jx**p*Jy**q*palfa**r*cos(3s*alfa)*sin(3*t*alfa)+
+sin(3*t*alfa)*cos(6s*alfa)*palfa**r*Jy**q*Jx**p*Jz**n*J**2k)}
-----
Name          k,      n,      p,      q,      r,      s,      t,          value[cm-1],      stage, float(Y=1/N=0)
-----
```

Then the list of parameters used in the fit is given. The line &&&END,,,,,,,,,,,,,,,,,,,,, signals the end of this list. The parameter values should be given in cm-1. A very brief example is:

```
muZ          ,      0,      1,      0,      0,      0,      0,      0.28599999999999999E+01 ,      0,      0,
0.5V6        ,      0,      0,      0,      0,      0,      0,      0.16149744155914532E+01 ,      1,      1,
0.5V6        ,      0,      0,      0,      0,      0,      2,      0.10000000000000000E+01 ,      1,      -1,
A-0.5(B+C)   ,      0,      2,      0,      0,      0,      0,      0.15939169657350583E+00 ,      2,      1,
0.5(B+C)     ,      1,      0,      0,      0,      0,      0,      0.29300182191213497E-01 ,      2,      1,
0.5(B-C)     ,      0,      0,      2,      0,      0,      0,      0.22615184835355637E-02 ,      2,      1,
0.5(B-C)     ,      0,      0,      0,      2,      0,      0,      -0.10000000000000000E+01 ,      2,      -1,
F            ,      0,      0,      0,      0,      2,      0,      0.37713516510541040E+00 ,      1,      1,
&&&END,,,,,,,,,,,,,,,,,,,,,
```

Each line with a Hamiltonian parameter contains the parameter name, a set of 7 integers that define the associated quantum mechanical operator, a value for the parameter (i.e., for the coefficient of this operator), the diagonalization stage in which the particular parameter is used, and a float/fix flag.

The name of the parameter may be anything you like, anything which helps you to remember what type of parameter you are using. The actual term is encoded by the set of 7 integers, which are the powers of operators from Exp. (1) above. Thus:

- k – (the first integer) corresponds to the power of J**2,
- n – (the second integer) corresponds to the power of Jz,
- p,q – are powers of Jx, Jy,
- r – is the power of pa,
- s – (the sixth integer) defines the argument of cos(3sa),
- t – (the seventh integer) defines the argument of sin(3ta).

On input, the program checks the set of these powers for conformity with symmetry requirements. The product of cos(3sa)*sin(3ta) is not allowed since this can be represented as a sum of cos(3s'a) and sin(3t'a) terms.

The value of the parameter is followed by a number that defines the stage in the diagonalization process in which the current parameter should be introduced (1 or 2). Usually the F, V6, and Rho terms are put in the first diagonalization stage. It is assumed, however, that at least Fpa**2 is put into the first diagonalization stage, in order to provide the eigenvectors which are necessary for the second diagonalization step. In the first stage only terms that do not contain Jx,Jy (i.e., only terms diagonal in the K quantum number) are allowed. Terms non diagonal in K should be put in the second diagonalization stage.

The last column of integers defines whether the particular parameter is fixed or floated in the fit (0 or 1) and also gives an opportunity to group Hamiltonian terms and fit a coefficient in front of a group of terms connected by a fixed linear relationship. Let's consider an example of such a "composite" Hamiltonian term.

```
0.5(B-C)      ,      0,      0,      2,      0,      0,      0,      0,      0.22615184835355637E-02 ,      2,      1,
0.5(B-C)      ,      0,      0,      0,      2,      0,      0,      0,      -0.10000000000000000E+01 ,      2,      -1,
```

These two lines together encode the $J_x^2 - J_y^2$ term. The first line encodes J_x^2 , the second line encodes J_y^2 . The parameter value in the second line fixes the relationship between J_x^2 and J_y^2 (in our case the second line is subtracted from the first line i.e. the value is -1.0). The parameter value in the first line is the actual value of the composite parameter which is fit. The last column in the second line contains the integer -1, which says that the parameter value by which the J_y^2 term should be multiplied is taken from previous line (so we take the parameter value from the previous line and multiply it by the "parameter value" from current line (in our case this is -1.0) and we get the parameter which will go with J_y^2 in the Hamiltonian matrix).

The composite term may contain more than two components like in the example below.

```
V3bc      ,      0,      0,      2,      0,      0,      0,      0,      -0.18859985631650738E-03 ,      2,      1,
V3bc      ,      0,      0,      0,      2,      0,      0,      0,      -0.10000000000000000E+01 ,      2,      -1,
V3bc      ,      0,      0,      2,      0,      0,      1,      0,      -0.10000000000000000E+01 ,      2,      -2,
V3bc      ,      0,      0,      0,      2,      0,      1,      0,      0.10000000000000000E+01 ,      2,      -3,
```

In this case the last integer increases its negative value (or to be more correct decreases it, since it is negative) each time by 1 with each new line, so that the parameter value (before modification) is always taken from the first line.

All components of the composite parameter should be put into the same diagonalization stage. Whether this composite parameter will be included or not is determined by the first line of the composite parameter. The composite parameter should go as one group.

A comment should be made about dipole moment components. The line which defines the dipole moment component should have zeros in last two columns.

```
muz      ,      0,      1,      0,      0,      0,      0,      0,      0.28599999999999999E+01 ,      0,      0,
```

Whether it is the μ_z , μ_x , or μ_y component of dipole moment is determined by a unity (integer 1) in the 2, 3, or 4 position, respectively (these positions correspond for the Hamiltonian terms to the J_z , J_x , and J_y operators). Therefore if you put unity in the second position like in the example above, then you work with the μ_z component of the dipole moment. If you put '1' in position 3 then you work with μ_x , and unity in the 4th position corresponds to μ_y . Certainly the permanent μ_x and μ_y components are not allowed by symmetry in the V6 mode of the program. But some $\cos(3s_a)$ or $\sin(3t_a)$ corrections to the dipole moment function containing μ_x or μ_y are allowed and also may be used in calculations by putting in addition to unity in the z,x,y position an appropriate integer in the 6 or 7 position to get a $\cos(3s_a)$ or $\sin(3t_a)$ variation of the component.

And the very last comment. It is the total number of lines (i.e. the lines in the composite parameters must also be counted) that should be less than 'numpar' specified in the third line of the input file.

The next 7 lines contain the following information.

```

20  !Jmax
9   !Number of iterations
8.d0, 0.0001, 1.E-2, ! Temperature, linestrength cutoff, "intensity" cutoff
5, 3, 3, 5,          ! V6 mode: stat.weights. for A1/A2, B1/B2, E1, E2 species; V3 mode - ignored
6000.d0, 80000.d0, 0, ! lower and upper frequency limits; 0/1 - prediction in MHz or in cm-1
0, 30, 0, 2,        ! Jlow, Jup, vtlow, vtup
0, 0,  !+/-1 prediction (0- no prediction, +1 only dvt=0, -1 dvt/=0 also); 0/1-calculate or not
impact of each parameter

```

The first line is the maximum J value for the transitions included in the fit, the second line defines the maximum number of iterations in the fit allowed if the stopping criterion will be not satisfied. If this number is negative, then the robust weighting scheme is used (see Ref. 3). If this number equals 0, then only a calculation will be done.

The next lines define parameters for the predictions. The third line contains the temperature, line strength cut-off, and intensity cut-off. If the temperature is less then 20K then a separate lowest energy is used in intensity calculations for each symmetry species. The 4-th line contains the nuclear statistical weights for torsion-rotation levels of the species indicated (the order should be as indicated in the comment, i.e. A1/A2, B1/B2, E1, E2). This line is important for intensity calculations in V6 mode. In V3 mode it is assumed that A:E lines have 1:1 ratio and this line is skipped by the program. The 5th line contains lower and upper frequency limits for the calculations and a MHz/cm⁻¹ switch. If the switch = 0 then predictions are made in MHz; if it equals 1 then predictions are made in cm-1. The frequency limits should be in the desired frequency units. The next line contains: Jlow, Jup, vtlow, vtup, i.e., contains the lower and upper limits for the J and vt quantum numbers. The ground vt=0, first vt=1 and up to second vt=2 excited states are allowed at the moment by the labeling procedure used in the program.

The last line contains two switches. The first switch has three positions. 0,1,-1. If it equals 0, then no prediction is made. If it equals 1, then only the transitions within each torsional state will be calculated. If it equals -1, then the transitions between different torsional states will be calculated also. The predictions will be put in separate files predictdvt0.txt and predictdvt1.txt. The first one will contain the transitions within torsional states and the second one will contain transitions between torsional states.

The last switch has two positions 0 and 1, and regulates whether the impact of different Hamiltonian parameters to the predicted transition frequencies will be calculated or not (0-not, 1-yes). The impact of the parameter is estimated as the expectation value of the corresponding operator, calculated using the eigenvectors of the upper and lower states of the transition. If the calculation is requested, then the result will be in separate files expectdvt0.txt and expectdvt1.txt, again for the transitions inside torsional states and between torsional states, respectively.

The next three lines are ignored by the program and simply remind the user of the format for the measured lines used in the fit

```

-----
 meas.freq  mut  Ju   Kau  Kcu      mlt  J1   Kal  Kcl   ifit  weight  comments
-----

```

The measured transition frequencies should be given as follows.

```

6997.7692    0    4    0    4          0    3    0    3    1    0.0030  comment
6999.5210    1    4    0    4          1    3    0    3    1    0.0030  comment
7005.4664    2    4    0    4          2    3    0    3    1    0.0030  comment
 123.7864   -2   31    8   23          1   31    9   23    1   -0.0012
 123.7864   -2    9    4    5          1   10    5    6    1   -0.0012
 123.7942   -2   19    8   12          1   19    9   10    1   -0.0012

```

Here it is crucial to follow the format of the line.

First a measured frequency is given. Then the assignment quantum numbers are given. They are followed by the include/exclude switch 'ifit' and by the measurement uncertainty. After that space for a comment is reserved. A negative measurement uncertainty means that the current measurement is in cm⁻¹, whereas a positive uncertainty assumes measured frequencies in MHz. If ifit=1 then the transition is used in the fit, if ifit=0 the transition is omitted from the fit.

The assignment consists of m, J, Ka, Kc quantum numbers. First the upper level is given then the lower level is given. The labelling procedure in the program determines from the eigenvector composition the levels which belong to each particular m-state and then assigns Ka,Kc labels within each m state according to the usual asymmetric top energy ordering scheme.

The *m* labeling scheme is as follows. Each symmetry species is characterized by $\sigma = 0, 1, 2$ or 3. For the four symmetry components of the $\nu_t=0$ ground state we choose $m = \sigma$. The four components of the first excited $\nu_t=1$ state corresponds to $m = \sigma - 6$, and the second excited state correspond to $\sigma + 6$. Therefore in the V6 mode we have the following m states: for $\sigma = 0$ we have $m=0, m=-6, m=6$; for $\sigma = 1$ we have $m=1, m=-5, m=7$, for $\sigma = 2$ we have $m=2, m=-4, m=8$ and for $\sigma = 3$ we have $m=3, m=-3, m=9$.

For the V3 mode we have following encoding for the torsional states of A and E symmetry species.

A-species: $\nu_t=0 \rightarrow m=0$; $\nu_t=1 \rightarrow m=-3$; $\nu_t=2 \rightarrow m=3$

E-species: $\nu_t=0 \rightarrow m=1$; $\nu_t=1 \rightarrow m=-2$; $\nu_t=2 \rightarrow m=4$

The list of transitions is read by the program up to the end of the file. If it is necessary to stop reading earlier it is necessary to put a negative Jup after the last line you would like to be read.

Output file

By default the output is directed to the screen. Certainly it is possible to redirect the output to a file with some desired name (RAM36.exe > output.txt).

The output file is organized as follows. First the initial values of parameters as well as different statistics on the experimental data (number of transitions, number of blended transitions, number of levels included in the fit etc.) are given. Then, for each iteration, the information on the parameter changes and current rms deviation for different groups of data are given. Then the list of the parameter values for the next iteration is presented in the format suitable for the input file (so it can be copied and pasted easily to the input file).

At each iteration the information on the fitted Hamiltonian parameters is given in the following format.

0.5V6	0	0	0	0	0	0	0	0.1614974468338617E+01	0.53E-07	0.16E-05	0.33E-01	0.10E+07	1
A-0.5(B+C)	0	2	0	0	0	0	0	0.1593916967236937E+00	0.15E-09	0.15E-07	0.10E-01	0.11E+08	2
0.5(B+C)	1	0	0	0	0	0	0	0.2930018217087762E-01	-.20E-10	0.77E-09	-0.26E-01	0.38E+08	2
0.5(B-C)	0	0	2	0	0	0	0	0.2261518449993083E-02	-.34E-10	0.55E-09	-0.61E-01	0.41E+07	2
F	0	0	0	0	2	0	0	0.3771351672264470E+00	0.21E-08	0.50E-07	0.43E-01	0.76E+07	1

The composite parameters are represented only by their first line in this table. First the name and integer number encoding is given, then the current value of the parameter (after addition of the correction from the current iteration) is given followed by the correction itself (from the current iteration of the least square fit) and the statistical uncertainty for the current parameter. Then two ratios are given. First is the ratio of the last change to the statistical uncertainty and the second is the ratio of the parameter value to the statistical uncertainty. Usually it is considered that if the last ratio is greater than 100 then the parameter is satisfactorily determined from the point of view of confidence intervals.

When the fit has converged or the maximum allowed number of iterations has passed the output of obtained results is given. The list of transitions with $J < J_{\max}$ from the input file is given three times. First it is given in order of frequency, second it is given in the order it appears in the input file, and third it is given sorted by spectroscopic branches. In these lists the assignments of the levels are augmented by the symmetry species labels from the G12 group (for V6) or from the G6 group (for V3). When the list of transitions is presented sorted by branch the expectation value of the (23)* operation is given and also if the transition is blended you will find in the end of the line the letter 'b' with the residual calculated from the special treatment of blends.

In between these three lists information on the rms deviations of different groups of data (grouped by uncertainty, by symmetry species etc.) as well as some statistical analysis (like the correlation table, or singular value decomposition analysis) can be found. One warning should be given, however. Whereas the weighted rms deviation as well as the rms deviation of the data groups grouped according to measurement uncertainty is calculated taking into account the special treatment of blends the rms deviation for transitions grouped by m-state and by symmetry species is calculated without this special treatment since if the components of a blended transition belong to different groups it is not clear what residual should be taken if you would like to calculate rms deviation for this particular group.

Also after the output of transitions in order of frequency the list of Hamiltonian parameters with confidence intervals is given both in MHz and cm^{-1} .

Tips inspired by user questions:

1) If you would like to get just a calculation of the transition frequencies using current set of parameters without fitting put zero as the number of allowed iterations.

2) Keep in mind that the statement above “the program uses a general expression” means that indeed one expression is coded inside the program and you only select the powers of different operators. In practice this means that attention should be paid on exact form of operator which multiplies fitted parameter. This is important when you would like to compare obtained parameter values with parameter values obtained using other programs (i.e. with BELGI program when V3 mode of RAM36 is used). As an example of peculiarities which may appear we may take the case of diagonal quartic centrifugal distortion constants. In commonly used definition it is assumed that Δ_J , Δ_{JK} , Δ_K are multiplied by $-J^4$, $-J^2J_z^2$, $-J_z^4$, respectively, whereas general expression used in RAM36 allows only J^4 , $J^2J_z^2$, J_z^4 expressions, i.e. they do not include minus sign. So, when you use in the list of parameters following three lines to encode Δ_J , Δ_{JK} , Δ_K :

```
-ΔJ      , 2, 0, 0, 0, 0, 0, 0, -0.30661003234599568E-06 , 2, 1,  
-ΔJK     , 1, 2, 0, 0, 0, 0, 0, 0.55979097051184501E-05 , 2, 1,  
-ΔK      , 0, 4, 0, 0, 0, 0, 0, -0.12945159093675880E-03 , 2, 1,
```

the minus sign will go to the values of obtained parameter values (that is why the names of diagonal centrifugal distortion parameters in the example above are given as $-\Delta_J$, $-\Delta_{JK}$, $-\Delta_K$ to attract attention to this fact). This fact should be taken into account when you try to compare obtained parameter values with ab initio results or fits obtained using other programs.

References.

1. V. Ilyushin, Z. Kisiel, L. Pszczółkowski, H. Mäder, J. T. Hougen “A New Torsion-Rotation Fitting Program for Molecules with a Six-Fold Barrier: Application to the Microwave Spectrum of Toluene” // *Journal of Molecular Spectroscopy* Vol. **259**, pp. 26-38 (2010).
2. Vadim V. Ilyushin, Laura B. Favero, Walther Caminati, and Jens-Uwe Grabow “Intertorsional Interactions Revealing Absolute Configurations: The V_6 Internal Rotation Heavy-Top Case of Benzotrifluoride” // *ChemPhysChem* 2010, 11, 2589 – 2593.
3. James K.G. Watson “Robust weighting in least-squares fits” of *Molecular Spectroscopy* Vol. 219, pp. 326-328 (2003).