

CATCO Group Internal Use ONLY

User Manual for pURVA program

by

Yunwen Tao

Ver 0.1.1

May 2017

Contents

List of Figures	iii
List of Tables	iv
Abbreviations	v
Symbols	vi
1 Before we start	1
1.1 Copyright	1
1.2 History	1
1.3 Execution of pURVA	2
1.4 Theory as Unified Reaction Valley Approach	2
2 Input description	3
2.1 Keyword input	3
2.1.1 @DATAFILETYPE keyword	3
2.1.2 @DATAFILEPATH keyword	4
2.1.3 @PARM keyword	4
2.1.4 @VIBRATION keyword	4
2.1.5 @DIRCURV keyword	4
2.1.6 @AVAM keyword	5
2.1.7 @CURVCPL keyword	5
2.1.8 @CORIOLIS keyword	5
2.2 Section input	5
2.2.1 TITLE section	6
2.2.2 PARAMETER section	6
2.2.3 CURVCOR section	7
2.2.4 AUTOSMTH section	8
2.2.5 RMSPK section	8
2.2.6 DMO section	9
3 Output description	10
3.1 Energy and derivatives	10
3.2 Internal coordinates	11
3.3 Decomposition of reaction path direction and curvature into internal co-ordinates	11

3.4	Generalized vibrational frequency	11
3.5	Scalar curvature	12
4	Examples	13
4.1	Example 1. $\text{HCN} \rightarrow \text{HNC}$ isomerization	14
4.2	Example 2. $\text{CH}_3 + \text{H}_2 \rightarrow \text{CH}_4 + \text{H}$	15
4.3	Example 3. Gold catalysis Step-1	15
4.3.1	First run	16
4.3.2	Second run	17
4.3.3	Third run	17
	Bibliography	19

List of Figures

List of Tables

Abbreviations

pURVA standalone **p**rogram of **U**nified **R**eaction **V**alley **A**pproach
IRC **I**ntrinsic **R**eaction **C**oordinates

Symbols

\AA	distance	Angstrom
ω	harmonic frequency	cm^{-1}

Chapter 1

Before we start

It is not a good idea to have no idea about the basic copyright, history, requirement, functionality as well as theory of the program before we use it.

1.1 Copyright

The program pURVA as well as this manual MUST NOT be released out of CATCO group at Southern Methodist University. According to policy §12.1 and §12.2 of Southern Methodist University, the leakage of the intellectual property may face legal charge.

1.2 History

The original URVA method was implemented by Zoran Konkoli in the link L716 of Gaussian package. However, this part is never incorporated into the public version of Gaussian. Since then, many other contributors including Dr. Wenli Zou added functionalities into this part and migrated this part from older version of Gaussian into newer version of Gaussian for several times. As Gaussian package was written in Fortran 77, the corresponding URVA part was written in the same language.

Later on from 2011, Dr. Dieter Cremer and Dr. Elfi Kraka wanted to have an independent version of URVA program. They asked Yunwen Tao in the group to this job. He started with programming in Fortran 90 language which is an extension to Fortran 77. Then he switched the whole project into Python language which is more flexible and easy to use. The new version of the URVA program was then named as **pURVA**.

1.3 Execution of pURVA

The proper execution of pURVA requires Python interpreter with the version 2.7.x. Versions lower than this might lead to trouble.

Here are the list of Python modules needed to run pURVA: (1) NumPy, (2) SciPy, (3) SymPy, (4) sys, (5) os, (6) copy, (7) gc, (8) math and (9) time.

Make sure that all these modules have been installed properly.

pURVA expects and then reads in an external text file as the user input file. After this file is prepared, in the terminal, type in

```
$ python main.py myinputfile
```

From the standard output, we could monitor how the calculation goes. The calculation results will be dumped into external text files on the disk.

To make life easier, running pURVA on ManeFrame cluster is recommended as pURVA has been developed and tested on the same machine. Before running pURVA, remember to load the Python interpreter by using

```
$ module load python
```

1.4 Theory as Unified Reaction Valley Approach

The name of “Unified Reaction Valley Approach” firstly appeared on scientific journals in 1997 when Konkoli, Kraka and Cremer published their comprehensive studies of $\text{CH}_3 + \text{H}_2 \rightarrow \text{CH}_4 + \text{H}$ on *J. Phys. Chem. A*[1]. In that paper, one of the highlights is to introduce the approach that calculates the adiabatic mode coupling coefficient that is decomposition of reaction path curvature into adiabatic local modes which was a novel approach dealing with vibrational spectroscopy. URVA is based on the Reaction Path Hamiltonian(RPH) that was intensively developed by Miller, Handy[2], Page and McIver[3]. In the year of 2011, Dr. Kraka published a well-written review on the relationship between RPH and URVA[4]. Most recently, Dr. Zou proposed a new approach to decompose the reaction path direction and curvature into internal coordinates which opens the possibility to study chemical reactions in large systems, e.g. organometallic compounds and enzymes[5].

One of the most important papers involved in URVA is the introduction of Diabatic Mode Ordering(DMO) procedure which has now been widely used in several projects within CATCO group[6].

Chapter 2

Input description

There are two major types of input that are allowed in the input file.

- Keyword input
- Section input

2.1 Keyword input

Just for convenience, keyword input is often written before section input. The format of keyword input line is:

```
@keyword_name = [keyword_value]
```

The @ symbol should be in the first column. No space is allowed after it. On both sides of = sign, it should be space. There might be several optional keyword values available, however, only one option is accepted.

2.1.1 @DATAFILETYPE keyword

This keywords specifies the format of input data source file for URVA analysis.

```
@DATAFILETYPE = old/new/xyz
```

old: The input data source file is generated by Gaussian package by setting corresponding IOp(1/45). This type of data contains most complete information.

new: The input data source file is generated by a modified version of Gaussian package. This type of data has no Hessian and gradient stored.

`xyz`: XYZ file containing the Cartesian coordinates of multiple snapshots.

2.1.2 @DATAFILEPATH keyword

This keyword specifies the path of the input data file.

```
@DATAFILEPATH = "../path/to/data/file"
```

The quotation marks should be included.

2.1.3 @PARM keyword

This keyword specifies the way to deal with internal coordinates parameters provided by user.

```
@PARM = No/GeomOnly/All
```

`No`: Do nothing with regard to these internal coordinates specifications.

`GeomOnly`: Only calculate the value of these internal coordinates.

`All`: Besides the value of internal coordinates, other properties related to these internal coordinates will be calculated.

2.1.4 @VIBRATION keyword

This keyword specifies whether or not to do normal mode analysis.

```
@VIBRATION = on/off
```

If the keyword value is set to `on`, the `@DATAFILETYPE` must be set to `old` and `@DIRCURV` must be set to `on`.

2.1.5 @DIRCURV keyword

This keyword decides whether or not to calculate reaction path direction $\eta(s)$ and curvature $\kappa(s)$.

```
@DIRCURV = on/off
```

If the keyword value is set to `on`, the `@DATAFILETYPE` must be set to `old` or `new`.

2.1.6 @AVAM keyword

This keyword specifies whether or not to calculate the adiabatic mode coupling coefficient $A_{n,s}(s)$.

@AVAM = on/off

If the keyword value is set to on, the @DATAFILETYPE must be set to old, the @PARM must be set to All, the @VIBRATION must be set to on and the @DIRCURV must be set to on

2.1.7 @CURVCPL keyword

This keyword specifies whether or not to calculate the curvature coupling coefficient $B_{\mu,s}(s)$.

@CURVCPL = on/off

If the keyword value is set to on, the @DATAFILETYPE must be set to old, the @VIBRATION must be set to on, and the @DIRCURV must be set to on.

2.1.8 @CORIOLIS keyword

This keyword specifies whether or not to calculate the Coriolis mode-mode coupling coefficient $B_{\mu,\nu}(s)$.

@CORIOLIS = on/off

If the keyword value is set to on, the @DATAFILETYPE must be set to old and the @VIBRATION must be set to on.

2.2 Section input

Section input is used when multiple parameters need to be read in, the format of the section input is:

SECTION_NAME

parameter line_1

parameter line_2

...

END SECTION_NAME

2.2.1 TITLE section

This section accepts remarks provided by user. The content will be displayed in standard output.

TITLE

Please put remarks here.

Multiple lines are accepted.

END TITLE

This section is quite useful to take note of the parameters we use for URVA calculations.

2.2.2 PARAMETER section

This section contains the internal coordinates specifications provided by the user. Different types of internal coordinates including ring coordinates are acceptable.

Bond length, bond angle, dihedral angle, out-of-plane angle, pyramidalization angle, ring puckering amplitude, ring puckering phase angle, ring deformation amplitude and ring deformation phase angle are supported.

PARAMETER

Internal coordinate specification

END PARAMETER

Bond length:

std N_1 N_2 : "bond_name"

Bond angle:

std N_1 N_2 N_3 : "angle_name"

Dihedral angle:

std N_1 N_2 N_3 N_4 : "dihedral_name"

Out of plane angle(the angle between the bond length N_1-N_2 and the plane $N_2-N_3-N_4$):

```
oop N1 N2 N3 N4 : "out_of_plane_name"
```

Pyramidalization angle(the angle θ_P is related to the three bond angles $N_2-N_1-N_3$, $N_3-N_1-N_4$, $N_4-N_1-N_2$):

```
pyr N1 N2 N3 N4 : "pyramidalization_angle_name"
```

Radius of planar reference ring(R)(N_{ring} : number of ring atoms):

```
ring Nring - ( N1 N2 ... Natoms ) -[0 0]: "ring_breathing_name"
```

Planar deformation amplitude(t_n)($n=1\sim N_{ring}-2$):

```
ring Nring - ( N1 N2 ... Natoms ) -[1 n]: "deformation_amplitude_name"
```

Planar deformation phase angle(τ_n)($n=1\sim N_{ring}-2$):

```
ring Nring - ( N1 N2 ... Natoms ) -[2 n]: "deformation_phase_angle_name"
```

Puckering amplitude(q_n)($n=2\sim(N_{ring}-1)/2$ for odd N_{ring} or $2\sim N_{ring}/2$ for even N_{ring}):

```
ring Nring - ( N1 N2 ... Natoms ) -[3 n]: "puckering_amplitude_name"
```

Puckering phase angle(ϕ_n)($n=2\sim(N_{ring}-1)/2$ for odd N_{ring} or $2\sim N_{ring}/2-1$ for even N_{ring}):

```
ring Nring - ( N1 N2 ... Natoms ) -[4 n]: "puckering_phase_angle_name"
```

2.2.3 CURVCOR section

The CURVCOR interface will be activated if this section is found.

For most situations, it is usually enough for N_l and N_r to take the value of 25.

```
CURVCOR
```

```
Ln = Nl
```

```
Rn = Nr
```

```
END CURVCOR
```

2.2.4 AUTOSMTH section

The AUTOSMTH interface will be activated if this section is found.

AUTOSMTH interface requires the activation of CURVCOR interface.

δs is the stepsize of mass-weighted IRC with the unit of amu^{1/2}-Bohr.

Using the value of 3 is usually enough for N_l and N_r .

t is a cut-off for second derivative of smoothened curve. Increase it when necessary.

Recommended value: 2.5.

```
AUTOSMTH
```

```
StepSize =  $\delta s$ 
```

```
Ln =  $N_l$ 
```

```
Rn =  $N_r$ 
```

```
d2ythresh =  $t$ 
```

```
END AUTOSMTH
```

2.2.5 RMSPK section

The RMSPK interface will be activated if this section is found.

RMSPK interface requires the activation of AUTOSMTH.

Any points in the curvature plot having the value larger than k will be left out as spike.

The value of p ranges from 0.5 to 1.0 as a percentage number. Any points leading to consecutive difference larger than the percentile of p will be labeled as spike candidates.

Recommended value: 0.85.

Gradient check threshold g is used to filter out normal points from spike candidates.

Recommended value: 1.2.

```
RMSPK
```

```
CutHigh =  $k$ 
```

```
Percentage =  $p$ 
```

```
GradRatio =  $g$ 
```

END RMSPK

2.2.6 DMO section

If this section input is not found, default parameter values will be used.

s_{max} is an overlap threshold after each mode reordering step. If the overlap criteria of s_{max} could not be reached, the criteria will be reduced to s_{min} gradually. Recommend values for s_{max} and s_{min} : 0.990 and 0.890.

If local difficulty is encountered, linear interpolation will be adopted, space between two consecutive points will be divided into N_{min} pieces. If the difficulty is still not solved, N_{min} will be increased up to N_{max} . Recommended values for N_{min} and N_{max} : 30 and 200.

If the DMO could not get through for a specific point due to the following reasons:

- Change of symmetry of reaction complex, e.g. linear \rightarrow non-linear
- Discontinuity of reaction path

one solution to circumvent this problem is to calculate and re-order the vibrational frequencies for a specific region of reaction path. This function could be activated by setting IO_{cut} to 1. In this way, the reaction path with its s value ranging from s_{start} to s_{end} will have vibrational frequencies calculated.

DMO

$S_{thresh} = s_{max}$

$S_{lowest} = s_{min}$

$N_p = N_{min}$

$N_{Max} = N_{max}$

$Cut = IO_{cut}$

$CutA = s_{start}$

$CutB = s_{end}$

END DMO

Chapter 3

Output description

In pURVA, results are all written to external files instead of standard output. All output files have the suffix of “.csv” or “.dat”. The execution of pURVA will abort if a result file with a duplicated name is found in the current folder. Make sure that current directory is cleaned up before execution.

3.1 Energy and derivatives

Usually the Self-Consistent Field(SCF) energy is calculated and used to construct the potential energy surface along the reaction path.

In order to check this value, the @DATAFILETYPE must be set to `old` or `new`.

The unit of SCF energy is Hartree as one of the atomic units(a.u.). In order to calculate first and second derivatives of SCF energy against reaction coordinate/parameter `s`, cubic spline fitting is used. For the second derivative of SCF energy, the region between `s = -0.1` and `s = +0.1` is predicted via cubic spline fitting from the information outside this region.

NOTE: 1 Hartree = 627.509 474 kcal/mol

Output files:

- `energy.csv`

SCF energy vs. `s`

- `energy_1_d.csv`

First derivative of SCF energy vs. `s`

- [energy_2_d.csv](#)

Second derivative of SCF energy vs. **s**

3.2 Internal coordinates

The value of user-defined internal coordinates could be calculated. All types of internal coordinates described in section 2.2.2 are supported.

In order to have this result, **@PARM** must be set to **GeomOnly** or **All**.

The unit of printed internal coordinates is atomic unit with bohr for distance and radian for angles.

NOTE: 1 Bohr = 0.529177 Å; 1 rad = 57.295 8 °

Output file:

- [q_n.csv](#)

3.3 Decomposition of reaction path direction and curvature into internal coordinates

In order to have this result, **@PARM** must be set to **All**.

Output files:

- [eta-q_n.csv](#)

Decomposition of reaction path direction into internal coordinates

- [kappa-q_n.csv](#)

Decomposition of reaction path curvature into internal coordinates

3.4 Generalized vibrational frequency

For any point on reaction path, we could have $3N - K - 1$ vibrations, in which K is the total number of translations and rotations. In pURVA, K takes the value of 6 which excludes the possibility of analysis of reactions like $\text{H}_2 + \text{H} \rightarrow \text{H} + \text{H}_2$ where the whole reaction complex stays in a linear geometry.

In order to have this result, `@VIBRATION` must be set to `on`.

Output file:

- `freq_dmo.csv`

Generalized vibrational frequencies vs. `s`

NOTE: Unit of frequencies is cm^{-1} .

3.5 Scalar curvature

The original scalar curvature calculated without correction around the TS region and spike removal will be written to file `originalkappa.dat`.

If the CURVCOR and AUTOSMTH modules are used, the corrected curvature data will be written to `merged.dat`.

If RMSPK module is also used, the curvature data after spike removal will be written to `merged-nospk.dat`.

Chapter 4

Examples

4.1 Example 1. HCN \rightarrow HNC isomerization

```
@DATAFILETYPE = old
@PARM = All
@VIBRATION = on
@DIRCURV = on
@AVAM = off
@CURVCPL = off
@CORIOLIS = off

@DATAFILEPATH = "../examples/hcn/IRC.browse"

TITLE
  HCN Reaction test job
  Frequency calculation from -3.09985216909 to 3.97984648135
END TITLE

PARAMETER
  std 2 3 : bond-NH
  std 1 3 : bond-CH
END PARAMETER

DMO
  Sthresh = 0.990
  Slowest = 0.890
  Np = 30
  NMax = 200
  Cut = 1
  CutA = -3.09985216909
  CutB = 3.97984648135
END DMO
```

4.2 Example 2. $\text{CH}_3 + \text{H}_2 \rightarrow \text{CH}_4 + \text{H}$

The whole calculation may take up to 15 minutes.

```
@DATAFILETYPE = old
@PARM = All
@VIBRATION = on
@DIRCURV = on
@AVAM = off
@CURVCPL = off
@CORIOLIS = off

@DATAFILEPATH = "./examples/ch3h2/IRC.forward.2"

TITLE
  CH3+H2 Reaction test job
END TITLE

PARAMETER
  std 2 3 : bond-NH
  std 1 3 : bond-CH
END PARAMETER

DMO
  Sthresh = 0.990
  Slowest = 0.890
  Np = 30
  NMax = 200
  Cut = 0
END DMO
```

4.3 Example 3. Gold catalysis Step-1

For this reaction complex, we do three URVA analysis runs. In the first run, basic information including energy, internal coordinates are calculated. In the second run, scalar curvature is corrected in its TS region and spikes are also removed. In the last run, vibrational frequencies are calculated for three segments of the reaction path due to several local difficulties around TS point.

4.3.1 First run

```
@DATAFILETYPE = old
@PARM = GeomOnly
@VIBRATION = off
@DIRCURV = off
@AVAM = off
@CURVCPL = off
@CORIOLIS = off

@DATAFILEPATH = "../examples/gold/IRC.browse"

TITLE
  Gold catalysis step 1 test job - first run
END TITLE

PARAMETER
std 1 16 : bond-C1C16
std 5 16 : bond-O5C16
std 1 10 : bond-C1C10
std 2 5  : bond-C2O5
std 2 4  : bond-C2O4
END PARAMETER
```

4.3.2 Second run

```
@DATAFILETYPE = old
@PARM = All
@VIBRATION = off
@DIRCURV = on
@AVAM = off
@CURVCPL = off
@CORIOLIS = off

@DATAFILEPATH = "../examples/gold/IRC.browse"

TITLE
  Gold catalysis step 1 test job - second run
END TITLE

PARAMETER
  std 1 16 : bond-C1C16
  std 5 16 : bond-O5C16
  std 1 10 : bond-C1C10
  std 2 5  : bond-C2O5
  std 2 4  : bond-C2O4
END PARAMETER

CURVCOR
  Ln = 25
  Rn = 25
END CURVCOR

AUTOSMTH
  StepSize = 0.03
  Ln = 3
  Rn = 3
  d2ythresh = 2.4
END AUTOSMTH

RMSPK
  CutHigh = 20.0
  Percentage = 0.85
  GradRatio = 1.2
END RMSPK
```

4.3.3 Third run

This run may take up to 25 minutes.


```
@DATAFILETYPE = old
@PARM = No
@VIBRATION = on
@DIRCURV = off
@AVAM = off
@CURVCPL = off
@CORIOLIS = off

@DATAFILEPATH = "./examples/gold/IRC.browse"

TITLE
  Gold catalysis step 1 test job - third run
END TITLE

DMO
  Sthresh = 0.980
  Slowest = 0.880
  Np = 5
  NMax = 80
  Cut = 0
END DMO
```

Bibliography

- [1] Zoran Konkoli, Elfi Kraka, and Dieter Cremer. Unified reaction valley approach mechanism of the reaction $\text{CH}_3 + \text{H}_2 \rightarrow \text{CH}_4 + \text{H}$. *The Journal of Physical Chemistry A*, 101(9):1742–1757, 1997.
- [2] William H. Miller, Nicholas C. Handy, and John E. Adams. Reaction path hamiltonian for polyatomic molecules. *The Journal of Chemical Physics*, 72(1):99–112, 1980.
- [3] Michael Page and James W. McIver. On evaluating the reaction path hamiltonian. *The Journal of Chemical Physics*, 88(2):922–935, 1988.
- [4] Elfi Kraka. Reaction path hamiltonian and the unified reaction valley approach. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(4):531–556, 2011.
- [5] Wenli Zou, Thomas Sexton, Elfi Kraka, Marek Freindorf, and Dieter Cremer. A new method for describing the mechanism of a chemical reaction based on the unified reaction valley approach. *Journal of Chemical Theory and Computation*, 12(2):650–663, 2016.
- [6] Zoran Konkoli, Dieter Cremer, and Elfi Kraka. Diabatic ordering of vibrational normal modes in reaction valley studies. *Journal of Computational Chemistry*, 18(10):1282–1294, 1997.