Описание работы с программами решения обратных задач спектроскопии (Disp) и электронографии (ElDiff)

Игорь Кочиков (igor@kochikov.ru)
Юрий Тарасов (tarasov@phys.chem.msu.ru)

В результате долгих объединительных усилий обе программы, а также и программа приведения по симметрии Symm, уже в значительной мере имеют общий код и общий интерфейс. Упомянутые программы позволяют решать задачу восстановления матрицы силовых постоянных молекулы по заданным колебательным частотам (Disp) и одновременного восстановления этой матрицы с определением равновесной геометрии молекулы по данным газовой электронографии (ElDiff). Применяется подход, основанный на методе регуляризации А.Н.Тихонова, позволяющий отыскивать приближения к нормальному (относительно априорно заданной матрицы F^0) псевдо- (квази-) решению. Более подробно описание алгоритма и составных частей программы дано в статьях и книге.

1. Основные возможности программы.

Предусмотрен расчет в произвольной системе естественных координат, а также в декартовых координатах. Система координат может быть зависимой. Ввод матрицы силовых постоянных и ее вывод может производиться в различных формах (в декартовых координатах, в естественных координатах, в координатах симметрии) и различных единицах (10^6 см$^{-2}$, мдин/Å, Хартри/Бор^2), что позволяет преобразовывать матрицы.

Расчет силовых постоянных предусмотрен в обычном режиме (когда определению подлежат элементы матрицы силовых постоянных) и в режиме, когда определяется набор масштабирующих множителей для заранее заданной матрицы. Предусмотрена возможность фиксации определенных силовых постоянных на заранее заданных значениях.

Предусмотрена возможность совместного расчета нескольких изотопных модификаций молекулы с целью определения единого силового поля. Возможен также расчет нескольких различных молекул одновременно при наложенных условиях равенства некоторых силовых постоянных в одной и в другой молекуле. Возможно проведение расчета в ситуациях, когда известны не все колебательные частоты.

Более подробно возможности программы можно узнать из формата входного файла.

2. Формат входных файлов

Предыдущая версия программы использовала формат входных данных, принятый еще в 1985 году, что в настоящее время приводит к ненужным ограничениям и неудобствам. Новый формат более гибок и имеет возможности расширения без изменения структуры файла. Для конвертирования старых файлов данных в новый формат существует особая программа.

В отличие от старой версии входного файла, часть информации, не относящаяся собственно к молекуле (а определяющая некоторые общие параметры расчета) вынесена в отдельный файл DISP.INI. Данные для каждой молекулы (либо для набора изотопных разновидностей одной и той же молекулы одинаковой симметрии) размещаются в файле с расширением .MOL.

Все входные файлы состоят из разделов, порядок следования которых безразличен. Некоторые разделы необходимы, в то время как другие могут отсутствовать. Каждый раздел начинается с его заголовка (в квадратных скобках) и продолжается до следующего заголовка раздела или конца файла. Ввод основан на использовании ключевых слов, то есть многие строки файла содержат ключевые слова и данные. Ключевые слова могут присутствовать в любом регистре (большие и малые буквы считаются одинаковыми). Во всех случаях разделителями данных являются пробелы или знаки табуляции. Каждая строка данных не должна быть длиннее 1000 символов.

Пустые строки игнорируются; игнорируется также содержимое каждой строки после комбинации символов // (как это принято в программах на C++). Это удобно использовать для комментариев в файле. Для временного отключения большого куска входного файла надо перед ним вставить строку #if 0, а после него - #endif.

2.1. Формат файла DISP.INI

Файл DISP.INI содержит данные, относящиеся ко всему расчету в целом. Он включает в себя на настоящий момент три раздела. Опции внутри каждого раздела идентифицируются по ключевым словам и могут следовать в любом порядке.

2.1.1. Содержание раздела [Job]

Type: тип решаемой задачи. Может принимать значения *Inverse* – обратная или *Direct* - прямая.
AlphaMax, AlphaMin: начальное и конечное значения параметра регуляризации. Эти параметры зависят от нормировки данных и менять их, вероятно, придется редко. Обычно разумные значения параметра составляют 1 (максимальное) и 10^{-6} (минимальное).
AlphaStepFactor: величина, на которую умножается значение параметра от шага к шагу. По умолчанию стоит 1/10.
MinGrad: критерий окончания оптимизации по величине нормы градиента. Его разумное значение также 10^{-6}.
MaxIter: максимальное число итераций в процессе оптимизации. Разумное значение – от сотен до тысяч. В электронографической версии должен задаваться небольшим (10 или около того), поскольку там есть еще внешний цикл итераций.
Method: режим решения задачи. Для обычного режима задается как *Standard*; для поиска масштабирующих множителей должен быть задан как *Pulay*.
OptiMethod: выбор метода оптимизации. Пока единственным работающим является градиентный (Gradient); он и используется по умолчанию.
ToTheEnd: режим, отключающий выбор параметра регуляризации. В этом режиме параметр регуляризации будет выбран равным минимальному значению (см. выше). Бывает полезен, если ваша задача на самом деле корректна (или чтобы определить меру несовместности).

2.1.2. Содержание раздела [Units]

В этом разделе указываются единицы измерения, в которых производится расчет и выдаются результаты. Входные данные не обязаны быть заданы в тех же единицах.

Length указывает на единицу измерения длин. Она может быть задана равной *Angstrom* или *Bohr*. Вторая из этих единиц длины часто используется в квантовомеханических расчетах. По умолчанию единицей длины является ангстрем.

Force задает единицы измерения для матрицы силовых постоянных и может иметь три различных значения:
1) *Cm* – матрица задается в единицах 10^6 см$^{-2}$, как это принято в советской спектроскопической литературе.
2) *MDyn* – матрица задается в единицах мдин/Å, мдин, мдин·Å (в зависимости от размерностей координат), как это принято в зарубежной спектроскопической литературе.
3) *Hartree* – матрица задается в единицах Хартри/Бор^2, Хартри/Бор, Хартри (в зависимости от размерностей координат). В этих единицах обычно приводятся результаты квантовомеханических расчетов.

2.1.3. Содержание раздела [Options]

UseIntFile: эта опция определяет, как используется файл промежуточных результатов, создаваемый в процессе работы программы (и позволяющий, в частности, продолжить ранее прерванную оптимизацию). Может принимать значения *YES* (будет использован, если имеется); *NO* (не будет использован); *ASK* (программа запросит пользователя, если обнаружит подходящий файл).

Все прочие опции этого раздела могут принимать значения *TRUE* или *FALSE*.

UseSymmetry: Включает или отключает использование координат симметрии, если информация о симметрии молекулы содержится в файле данных.
UseConstraints: Включает или исключает учет ограничений на элементы матрицы силовых постоянных, если эти ограничения содержатся в файле данных.
UseEquivalencies: Включает или исключает учет ограничений на равенство элементов матрицы силовых постоянных при совместной обработке молекул, если эти ограничения содержатся в файле данных (.pck) - см. раздел о совместной обработке.
MinimizeOffDiagonal: Матрица силовых постоянных в зависимой системе естественных координат определяется неоднозначно. Когда исходная матрица задана в декартовых координатах, пересчет ее в естественные координаты может быть сделан различными способами. При значении *FALSE* эта опция позволяет получить каноническую матрицу силовых постоянных (т.е. матрицу с рангом, равным числу независимых координат). Если же задано *TRUE*, выбирается матрица с минимальной недиагональной нормой.
SkipDirectInternal: В обычном режиме работа программы завершается выдачей результатов в координатах симметрии и в естественных (или декартовых) координатах. При больших размерностях матриц вторая из этих задач требует много времени и объема, поэтому ее можно отключить.
VerifyCoordinates: Во входном файле описание естественных координат может сопровождаться указание их значений в равновесной конфигурации. Если эта опция включена, программа проверит заданные декартовы координаты атомов на предмет соответствия указанным значениям естественных координат и, если потребуется, исправит положения атомов. Тем самым появляется возможность автоматически исправлять приближенные значения координат атомов.

2.1.4. Содержание раздела [Print Options]

Здесь собраны опции вывода на печать.

PrintInput: Распечатывать введенные данные перед решением задачи.
PrintCartesian: Если расчет ведется в естественных координатах, можно заказать также выдачу окончательной матрицы силовых постоянных в декартовых координатах.
PrintInternal: Для больших разреженных (содержащих много нулей) матриц бывает удобен ввод/вывод матриц в виде списка значений и набора элементов матрицы, соответствующих каждому значению. Более подробно об этом формате сказано позже.
PrintGMatrix: Печатать матрицы кинематических коэффициентов.
PrintGMatrixSymm: Печатать матрицы кинематических коэффициентов по блокам симметрии.
PrintModes: Печатать векторы форм колебаний.
PrintModeSymms: Печатать векторы форм колебаний по блокам симметрии.
PrintPED: Печатать распределение потенциальной энергии по нормальным колебаниям.
PrintPEDSymm: То же самое для координат симметрии.
PrintCubic: Печатать матрицы кубических силовых постоянных (только ElDiff).
PrintGFactors: Печатать таблицу атомных факторов рассеяния (только ElDiff).
PrintTime: Время от времени печатать текущее время выполнения программы.
DumpBinary: Используется для сохранения результирующей матрицы силовых постоянных в файле в двоичном формате, который может быть впоследствии загружен в программу (более подробно см. позже). Это бывает полезно при больших размерах матриц (с сотнями координат), когда их выдача в текстовом виде становится неудобной.
DumpCartesian: Используется для сохранения результирующей матрицы силовых постоянных в декартовых координатах файле в двоичном формате.

2.1.5. Содержание раздела [ED Options]

Этот раздел используется только при решении обратной задачи с данными по электронной дифракции. Он содержит следующие основные группы:
а) Задание параметров для оптимизации (каждый из параметров может быть *TRUE* либо *FALSE*:
Geometry: проводить оптимизацию геометрии;
Matrix: проводить оптимизацию силового поля.
б) Задание входных экспериментальных данных и весов для них.
Spectroscopy=1.0
Electronography=1.0
Microwave=1.0
Если данные соответствующего типа присутствуют во входном файле, они будут использованы при оптимизации. Задавая нулевой вес, можно исключить какой-либо вид данных (или понизить его влияние, задавая сравнительно меньший вес).
в) Управление стабилизацией данных. При решении неустойчивых задач оптимизируемый функционал включает добавки – так называемые стабилизаторы. Если задача устойчива, то они не являются необходимыми и могут быть исключены путем задания нулевых весов в следующих строках:
StabGeom=0.0
StabForce=0.1
(здесь исключен стабилизатор по расстояниям, а по силовой матрице оставлен, но имеет меньший вес).
г) *RfactorErr*=1.0

Эта величина показывает, какое отклонение *R*-фактора от его минимального значения мы считаем допустимым. При решении устойчивых задач можно минимизировать *R*-фактор до самого минимума. Если же задача неустойчива, то мы жертвуем некоторой величиной *R*-фактора для получения стабилизированного решения. Разумно ставить значение этого параметра несколько меньшим, чем вклад в *R*-фактор от случайных погрешностей.

д) *UseNonlinearCoordinates=(TRUE,FALSE)*
Ghb htitybb ufhvjybxtcrb[pflfx

2.1.6. Пример файла DISP.INI

```
// This is a sample INI file for Eldiff program.
// It now includes ED options as well as general options.
// Disp program will not read sections it does not need.

[Job]
Type=inverse            // для обратной задачи
AlphaMax=1.000000e+04
AlphaMin=1.000000e-04
MinGrad=1.000000e-6
MaxIter=100
Method=Standard         // может также быть Pulay
ToTheEnd=FALSE

[Units]
Length=Angstroms        // по умолчанию
Force=MDyn

[Options]
UseIntFile=ASK              // YES, NO or ASK
UseSymmetry=TRUE            // использовать координаты симметрии
UseConstraints=TRUE         // использовать ограничения
MinimizeOffDiagonal=TRUE    // брать матрицу с мин. недиагональной нормой
SkipDirectInternal=FALSE     // распечатать результаты не только в
                  // координатах симметрии, но и в естественных координатах
VerifyCoordinates=TRUE      // проверить и исправить координаты атомов

[Print Options]
PrintInput=TRUE             // распечатать введенные данные
PrintGMatrix=TRUE           // печатать матрицу G для молекулы
PrintModes=TRUE             // печатать собственные векторы колебаний
PrintGMatrixSymm=TRUE       // печатать матрицу G по блокам симметрии
PrintModesSymm=TRUE         // печатать собственные векторы по блокам
PrintPED=TRUE               // печатать распределение потенциальной энергии
PrintCubic=TRUE             // печатать кубические постоянные
                             // (для электронографии)
PrintCartesian=TRUE         // распечатать матрицу в декартовых координатах
PrintInternal=FALSE         // не выдавать матрицу в виде списка элементов
PrintTime=FALSE             // временами печатать время работы (в секундах)
DumpCartesian=FALSE          // не записывать декартову матрицу в
                             // двоичный файл
DumpBinary=FALSE            // не записывать матрицу в двоичный файл

[ED Options]
// What to optimize
Geometry = TRUE             // оптимизировать геометрию
Matrix = TRUE               // оптимизировать матрицу силовых постоянных
// What data to use
Spectroscopy = TRUE
```

```
Electronography = TRUE
Microwave = FALSE

// Anharmonicity section
UseNonlinearCoords = TRUE
AnharmonicConstants = MODEL   // Can be set to MODEL, FILE or NONE
// Scaling has sense for file input only
ScaleMode = NONE
// The following lines have sense only for model
Stretch = TRUE
StretchStretch = TRUE
StretchBend = TRUE
Bend = TRUE

Rotation = TRUE          // If to include rotation correction to dr
UseSecondOrder = FALSE   // Don't enable: does not work now

UseC3 = TRUE
UseC4 = TRUE            // With 1nd order perturbations, is zero anyway
UseRA = TRUE            // Use averaging with P(r)/r rather than P(r)
UseM = FALSE            // True if input is M(s) rather than sM(s)

Errors = TRUE           // Calculate errors (correlation matrix)
```

2.2. Формат файла данных (.mol)*

Файл данных (с расширением .MOL) содержит данные, относящиеся к одной молекуле или к набору изотопных разновидностей молекулы с одной и той же симметрией. Он содержит множество разделов, не все из которых является обязательными. Разделы могут следовать в любом порядке.

Многие разделы содержат списки, которые устроены одинаковым образом. *Список* начинается с указания его длины (*Count* = ...), за которым следует набор перенумерованных строк. Содержание строк (за исключением первого числа – номера строки – может быть разным в различных списках.

2.2.1. Содержание раздела [Isotopes]

Основное назначение этого раздела – задать имя молекулы. Кроме того, здесь же задается число изотопных разновидностей молекул. Раздел по сути является одним списком. В отличие от других мест, имя молекулы может содержать пробелы. При наличии изотопных модификаций, каждая из них может иметь свое имя.

Пример для одной молекулы:

```
[Isotopes]
Count=1
1 C60-4 linear chain
```

Пример для нескольких разновидностей:

```
[Isotopes]
Count=3
1 H2O ordinary water
2 D2O heavy water
3 T2O if you can get it
```

2.2.2. Содержание раздела [Atoms]

В этом разделе описывается равновесная конфигурация молекулы. Он начинается с указания единицы измерения длин (*Length*), которая может иметь значения *Angstrom* или *Bohr* (см. также 2.1.2). Если этой строки нет, координаты измеряются в ангстремах.

Остальная часть раздела представляет собой список атомов. Для каждого из атомов указываются его декартовы координаты (*x,y,z*) в произвольной системе координат, химический символ атома и масса (в а.е.м.). Имена (символы) атомов используются для поиска некоторых табличных значений, а потому должны соответствовать настоящим. При выдаче значений, относящихся к отдельным атомам, программа будет ссылаться на них, используя химический символ и номер атома в списке (например, C1, C2, C3 в следующем примере).

Пример задания информации об атомах:

```
[Atoms]
Length=Angstrom   // unnecessary line
Count=180
//No       x         y         z      Name  Mass(amu)
  1   0.000000  3.501544  -7.901932   C    12.011
  2   3.501544  0.701156  -8.603088   C    12.011
  3   0.785000  0.000000  -5.101544   C    12.011
... more atoms
```

Если в файле содержится информация о нескольких изотопных модификациях молекулы, имена и массы атомов должны быть заданы для каждой модификации (см. пример далее). Координаты при этом могут быть опущены, так как они предполагаются одинаковыми для всех модификаций молекулы.

Пример данных для трех изотопных модификаций:

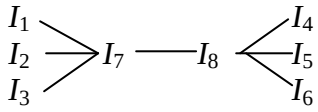
```
[Atoms]
Length=Angstrom
Count=3
//No      x         y          z      Name   Mass(amu)
  1   0.000000  3.501544  -7.901932   H         1.0
                                      D         2.0
                                      T         3.0
  2   3.501544  0.701156  -8.603088   C        12.0
                                      C        12.0
                                      C        12.0
... more atoms
```

2.2.3. Содержание раздела [Coordinates]

Программа может работать в декартовых или естественных координатах. Если выбрать тип *Type = Cartesian*, то никакой дополнительной информации не требуется. Программа сама введет 3*N* декартовых координат и даст им соответствующие имена. Если же выбрать естественные координаты (*Type = Internal*), то надо их описать. Описание координат представляет собой список. В этом списке для каждой координаты указывается ее имя, вид и список атомов, участвующих в определении данной координаты. Имя, как обычно, не может содержать пробелов. Единственное, что требует

комментариев – это вид координаты. В настоящее время реализованы следующие виды (индексы I_1, I_2, ... соответствуют номерам атомов в порядке их следования в описании координаты):

Вид 1: Растяжение связи (I_1-I_2).
Вид 2: Угол между связями (I_1-I_2) и (I_1-I_3).
Вид 3: Угол между связью (I_1-I_2) и плоскостью тройки атомов (I_2-I_3-I_4).
Вид 4: Двугранный угол между плоскостями троек атомов (I_1-I_2-I_3) и (I_2-I_3-I_4). Подразумевается, что плоскости пересекаются по связи (I_2-I_3).
Вид 5: Угол между плоскостями двух любых троек атомов (I_1-I_2-I_3) и (I_4-I_5-I_6).
Вид 6: Угол взаимного вращения двух волчков (см. рис.).



Если какой-либо волчок содержит менее 3 атомов, вместо номера отсутствующего атома вводится 0. Хотя бы один атом из каждой тройки присутствовать должен.

Порядок следования координат разных видов не имеет значения.

Пример:

```
[Coordinates]
Type=Internal
Count=838
//No Name      Type     Atoms
   1 R1-10      1       1   10            // связь определяется 2 атомами
   2 R2-19      1       2   19
   3 R3-34      1       3   34
      ...
 118 A28        2       28   30   11      // угол определяется 3 атомами
 119 A29        2       29   31   19
      ... more coordinates
```

Особый случай представляет собой угловая координата (тип 2) в линейных молекулах. В этом случае деформация угла должна описываться двумя углами, описывающими деформации в двух взаимно перпендикулярных направлениях. Эти две координаты должны стоять подряд с идентичными описаниями наборов атомов. Направления деформаций для двух углов программа выбирает автоматически. Если же желательно задать эти направления определённым образом, то декартовы компоненты векторов нормалей должны быть указаны вслед за описаниями атомов. Пример:

```
//No Name      Type     Atoms
 3 A123x        2       2   1   3   1.0 0.0 0.0
 4 A123y        2       2   1   3   0.0 1.0 0.0
```

Здесь предполагается, что ось молекулы направлена вдоль оси z.

Наконец, в этом же разделе могут быть представлены данные для уточнения координат атомов. Если мы не уверены в точности декартовых координат атомов, можно потребовать уточнения конфигурации для того, чтобы определённые координаты приняли заданное значение в равновесном состоянии. Значения координат указываются после описаний атомов (значения углов – в градусах, длин – в ангстремах). Пример:

```
  1 R1-10      1        1   10            Value=1.490
  2 R2-19      1        2   19            Value=1.490
  3 R3-34      1        3   34            Value=1.490
     ...
118 A28         2       28   30   11      Value=120.0
119 A29         2       29   31   19      Value=120.0
```

Если в файле DISP.INI заказано уточнение координат (*VerifyCoordinates = TRUE*), то перед началом расчёта декартовы координаты атомов будут изменены так, чтобы соответствовать указанным параметрам. Если параметры несовместны, рассчитывается наиболее близкая к заданной конфигурация (в смысле среднеквадратичного отклонения параметров от заданных значений).

Дополнительная возможность предоставляется опцией *Type = Auto*. В этом случае программа сама построит систему внутренних координат, состоящую из связей и углов. При этом связи определяются пользователем заданием максимального расстояния (оно может быть своим для каждой пары химических элементов), а углы вводятся все, какие есть между полученными связями. Поскольку этих координат может оказаться недостаточно, имеет смысл использовать данную опцию в программе SYMM, так как полученный набор координат будет вписан в модифицированный MOL файл. Пример:

```
[Coordinates]
Type=Auto
Count=2
//No Pair   Max Distance
  1  C-C      2.0
  2  C-H      1.4
```

В данном случае будут все пары атомов C, находящиеся на расстояниях меньше 2 Å друг от друга, будут считаться связанными. Для расстояний CH этот критерий ниже. После создания связей будут введены все валентные углы.

2.2.4. Содержание раздела [Symmetry]

Этот раздел не является абсолютно необходимым. Если он отсутствует, программа выдает предупреждение и продолжает работать, приняв симметрию C_1. Даже если раздел присутствует, использование координат симметрии можно отключить, установив опцию *UseSymmetry = FALSE* в файле DISP.INI.

В настоящее время программа анализа симметрии (SYMM.EXE) способна, получив файл данных без раздела [Symmetry], определить симметрию и вставить правильно построенный раздел. Создавать его вручную поэтому приходится редко. Для полноты все же опишем его содержимое.

Первое, что нужно задать в этом разделе – это матрицу перехода к координатам симметрии. Матрица должна быть квадратной и такого же размера, что и число естественных (или декартовых) координат. Возможные варианты:

Matrix =Unit – означает, что матрица – единичная. Встречается либо в случае отсутствия симметрии, либо в редких случаях, когда случайно все координаты и без того совпадают с координатами симметрии.
Matrix =Text – означает, что элементы матрицы приведены в последующих строках файла данных в текстовом виде. Если присутствуют комплексные коэффициенты (об их вводе –

см. ниже), необходимо добавить слово *Complex* вслед за *Text* (например, через запятую или пробел).
Matrix =File – означает, что матрица должна быть прочитана из двоичного файла. Имя файла в этом случае печатается после двоеточия (см. пример ниже). Такой способ ввода удобен для больших матриц, создаваемых программой автоматического анализа симметрии.

Матрица в текстовом виде может быть ненормированной (она будет отнормирована при вводе), и вводится с любым количеством чисел в каждой строке (но так, чтобы не превзойти максимальную ее длину из 1000 символов). Можно переносить строки. Вместо чисел допустимы следующие условные обозначения:

c и *d* (соответствуют комплексным кубическим корням из 1);
i (соответствует квадратному корню из –1);
и т. п. (полный список приводится в отчете программы SYMM).

Эти обозначения могут сопровождаться знаком “-”.

После матрицы приведения по симметрии задается блочная структура всех прочих матриц. По сути дела это список типов симметрии присутствующих в молекуле колебаний. Для каждого типа симметрии указывается его имя, количество независимых координат (или частот колебаний) и общее число координат (включая зависимые).

Имена блоков (за исключением первой буквы) могут быть произвольными, а первая буква используется для определения кратности блока. Для 2-х, 3-х, 4-х и 5-кратных блоков первая буква имени блока должна быть соответственно E, F, G и H.

Пример раздела с матрицей, содержащейся непосредственно в файле данных:

```
[Symmetry]
Matrix=Text
   1   0   0   0   0   0   0   0   0   0   0   0   0
   0   1   1   1   1   0   0   0   0   0   0   0   0
   0   0   0   0   0   1   1   1   1   0   0   0   0
   0   0   0   0   0   0   0   0   0   1   1   1   1
   0   1  -1   1  -1   0   0   0   0   0   0   0   0
   0   0   0   0   0   1  -1   1  -1   0   0   0   0
   0   0   0   0   0   0   0   0   0   1  -1   1  -1
   0   1   0  -1   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   1   0  -1   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   1  -1  -1   1
   0   0   1   0  -1   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   1   0  -1   0   0   0   0
   0   0   0   0   0   0   0   0   0   1   1  -1  -1
BlocksCount=4
//Name  Frequencies Coordinates
1   A1          3           4
2   B1          2           2
3   B2          1           1
4   E           3           3
```

Другой пример, когда матрица читается из двоичного файла:

```
[Symmetry]
Matrix=File:c60-3.cmx
BlocksCount=8
//Name  Frequencies Coordinates
1   Ag         71         118
2   Au         64          94
```

```
3   B1g       64          96
4   B1u       69         114
5   B2g       66         101
6   B2u       67         107
7   B3g       66         104
8   B3u       67         104
```

В данном случае размерность матрицы – 838 x 838, и представлять ее в текстовом виде было бы крайне неудобно.

2.2.5. Содержание раздела [Frequencies]

Этот раздел содержит описание экспериментальных частот колебаний и необходим только для решения обратных задач. Он представляет собой список частот с указанием экспериментальных ошибок и типа симметрии. Тип симметрии должен совпадать с одним из определенных в разделе [Symmetry], если таковой присутствует в файле данных. Все частоты и экспериментальные ошибки измеряются в обратных сантиметрах.

Пример раздела:

```
[Frequencies]
Count=20//  Symm.  Freq.   Error
1  A1g  3074.00   3.00
2  A1g   993.00   3.00
3  A2g  1350.00   3.00
4  A2u   674.00   3.00
... more frequencies
```

Если в файле содержится информация о нескольких изотопных модификациях молекулы, данные об их частотах записываются в последующих столбцах того же списка. Напомним, что один файл содержит информацию только об изотопных модификациях одной и той же симметрии. Пример файла для двух изотопных разновидностей:

```
[Frequencies]Count=20
//        Simple benzene   Deuterated benzene
//  Symm.  Freq.   Error      Freq.   Error
1  A1g  3074.00   3.00    2303.00   3.00
2  A1g   993.00   3.00     946.00   3.00
3  A2g  1350.00   3.00    1059.00   3.00
4  A2u   674.00   3.00     496.00   3.00
5  B1u  3057.00   3.00    2285.00   3.00
... more frequencies
```

2.2.6. Содержание раздела [Matrix Z]

Для начала работы программы необходимо задать какую-либо матрицу силовых постоянных, которая будет использована в качестве начального приближения для оптимизации (или для решения прямой задачи). Иногда эта же матрица может быть использована в качестве стабилизатора функционала Тихонова (см. 2.2.7).

Для ввода следует задать тип используемых координат и единицы измерения. Что касается единиц измерения (*Units*=), то они могут быть те же, что описаны в 2.1.2.

Тип координат определяет форму представления матрицы и может принимать следующие значения:
Coordinates=Cartesian означает ввод матрицы в декартовых координатах. При этом вся матрица вводится в треугольном виде, в каком в программе принято осуществлять вывод (это позволяет результаты одного расчета легко использовать в качестве входных данных для другого). Фактически нижняя треугольная матрица разрезается на столбцы определенной ширины, и эти столбцы вводятся один за другим. Приведенный ниже пример иллюстрирует этот способ представления матрицы.

Число столбцов матрицы опять-таки произвольно, но длина строки не должна превышать 1000 символов. Пример ввода декартовой матрицы:

```
[Matrix Z]
Coordinates=Cartesian
Units=Hartree

    1-x   0.7744
    1-y  -0.0056  0.7357
    1-z  -0.0000  0.0000  0.1530
    2-x  -0.2814 -0.1256 -0.0000  0.8052
    2-y  -0.0382 -0.2088 -0.0000 -0.0537  0.7674
    2-z  -0.0000 -0.0000 -0.0747  0.0000  0.0000  0.1708
    3-x   0.0744 -0.0138 -0.0000 -0.1688 -0.0345 -0.0000  0.7976
    3-y  -0.0945 -0.0434  0.0000  0.0703 -0.3823 -0.0000  0.0598  0.8022
    3-z  -0.0000  0.0000  0.0071  0.0000 -0.0000 -0.0708 -0.0000 -0.0000
    ... (пропущено сколько-то строк)...
   14-z  -0.0000 -0.0000  0.0210  0.0000  0.0000  0.0024 -0.0000 -0.0000
          1-x     1-y     1-z     2-x     2-y     2-z     3-x     3-y
    3-z   0.1502
    4-x  -0.0000  0.7707
    4-y   0.0000  0.0058  0.7853
    4-z  -0.0799  0.0000  0.0000  0.1904
    5-x   0.0000 -0.3172 -0.1464 -0.0000  0.8331
    5-y   0.0000 -0.0201 -0.1988  0.0000 -0.0481  0.7672
    5-z   0.0089 -0.0000 -0.0000 -0.0797 -0.0000 -0.0000  0.1476
    6-x   0.0000  0.0860 -0.0214 -0.0000 -0.1671 -0.0242  0.0000  0.7907
    ... (пропущено сколько-то строк)...
   14-z  -0.0010  0.0000  0.0000  0.0003  0.0000 -0.0000 -0.0005  0.0000
           3-z     4-x     4-y     4-z     5-x     5-y     5-z     6-x
    ... (пропущено много строк)...
   14-y   0.1494
   14-z  -0.0000  0.0273
           14-y    14-z
```

При этом обозначения координат (присутствующие в примере) могут быть опущены (но только все сразу).

Coordinates=Internal означает ввод матрицы в естественных координатах. За исключением этого, ввод полностью аналогичен вводу матрицы в декартовых координатах. Как и в предыдущем случае, можно сопровождать матрицу указанием координат (слева и снизу от каждого блока), но их функция – только облегчить контроль за правильностью ввода (при чтении эта информация игнорируется). Подразумевается, однако, что введенные координаты соответствуют (по типу, числу и порядку следования) тем, что описаны в разделе [Coordinates].

И в декартовых, и в естественных координатах можно (что особенно полезно при больших размерностях) вводить матрицу силовых постоянных из двоичного файла. Такие файлы должны быть сгенерированы другими программами, либо их вывод должен быть запрошен из программы DISP (см. раздел 2.1.3). Пример ввода матрицы из файла:

```
[Matrix Z]
Coordinates=Cartesian
Units=Hartree
```

```
File=car-02.dmp
```

Coordinates=Symmetry означает ввод матрицы в координатах симметрии. Этот тип ввода аналогичен предыдущим, но матрица вводится для каждого типа симметрии отдельно. Число блоков и размерность матриц должны соответствовать описанию в разделе [Symmetry]. Пример ввода матрицы в координатах симметрии:

```
[Matrix Z]
Coordinates=Symmetry
Units=Cm
Block=A1
 8.4620
 -0.0497 11.9070
-0.1476 -0.6315 5.3850
 0.0000  0.0000 0.0000 0.0000
Block=E
 7.3750
 0.3624  1.1702
 0.0921 -0.4248 2.3871
```

Матрицы в декартовых и естественных координатах, если они содержат много нулей, могут быть введены в виде списка элементов с указанными значениями, а оставшиеся элементы будут положены равными нулю. При этом после указания типа координат должно быть указано слово *List*, а дальнейшие данные содержат в каждой строке значение силовой постоянной и набор пар индексов элементов матрицы. Пример ввода списка значений:

```
[Matrix Z]
Coordinates=Internal(List)
Units=Cm
  7.0000    1,1
  0.2039    1,2   1,3   1,4   1,11  1,12  1,13
  0.9604    11,13 12,13
 12.0000    2,2   3,3   4,4   11,11 12,12 13,13
  0.6178    1,14  1,15  1,16
  0.9604    2,3   2,4   3,4   11,12
  0.6178    1,5   1,6   1,7
  0.4632    2,5   3,6   4,7   11,14 12,15 13,16
```

При наличии симметрии не требуется указывать все элементы матрицы, соответствующие какому-либо значению. При заполнении заданное значение будет присвоено всем эквивалентным по симметрии элементам матрицы. Список подобного рода можно запросить в качестве выходного файла программы (см. 2.1.3); он там упорядочен по убыванию элементов матрицы.

При вводе проверяется, удовлетворяет ли матрица условиям симметрии. Если она не соответствует заданной симметрии, то элементы матрицы соответствующим образом исправляются, а пользователь получает сообщение о произведенных изменениях.

2.2.7. Содержание раздела [Matrix Z0]

Для решения обратной задачи необходима некоторая априорно заданная матрица F^0, по отношению к которой будет производиться поиск нормального решения (псевдорешения). Она задается в этом разделе в точности в том же виде, что и в предыдущем.

В большинстве случаев удобно использовать в качестве F^0 ту же матрицу, что и заданная в предыдущем разделе. В этом случае раздел [Matrix Z0] может отсутствовать в файле данных. Напротив, если отсутствует раздел [Matrix Z], вместо него используется матрица, заданная в разделе [Matrix Z0].

2.2.8. Содержание раздела [Constraints]

Этот раздел требуется, если решение обратной (или прямой) задачи требуется провести на множестве с ограничениями. В настоящее время работают только ограничения типа равенства. Фактически можно положить некоторый набор силовых постоянных равными нулю либо зафиксировать их на начальных значениях. Набор элементов можно указать явно и неявно.Задание ограничений состоит из двух частей: задания типа ограничений (*Action=*) и задания набора элементов (*Set=*). Можно последовательно задать несколько видов ограничений со своими наборами элементов.

Action=Fix соответствует фиксации значений силовых постоянных. Это означает, что величины, заданные в исходной матрице, в дальнейшем изменены не будут.
Action=Fix-At-0 означает, что перечисленные элементы матрицы будут положены равными нулю и таковыми останутся в процессе минимизации.
Action=Release означает, что для перечисленных элементов ограничения будут сняты. Иногда бывает удобно наложить ограничения на более широкий набор элементов, а затем для некоторого их подмножества эти ограничения отменить.

Набор элементов, на который распространяются указанные ограничения, может быть задан после ключевого слова *Set=*. Возможны следующие варианты:

All - означает распространение указанных ограничений на все элементы матрицы. Этот вид введен только для полноты и вряд ли кому-либо может понадобиться.
Diag – означает наложение ограничений только на диагональные элементы матрицы. Этот вид данных чаще нужен для снятия ограничений (*Action=Release*), чем для их наложения.
OffDiag - распространяет ограничения на недиагональные элементы матрицы.
Distant – накладывает ограничения на элементы матрицы, соответствующие удаленным друг от друга структурным фрагментам. Фактически элемент матрицы участвует в этих ограничениях, если пара координат, соответствующая данному элементу, не имеет общих атомов, участвующих в определении координат (в частности, при этом равны нулю и соответствующие элементы матрицы *G*).
Far – то же, что и *Distant*, но пара координат не должна иметь не только общих, но и соседних атомов, участвующих в их определении. Это более узкий набор элементов, нежели в *Distant*.
Zeroes - ограничения налагаются на элементы матрицы, заданные равными нулю (это означает, что таковыми они и останутся в процессе оптимизации).
List – означает, что непосредственно после этой строки явно задан список элементов матрицы, подлежащих указанным ограничениям. Этот список имеет вид пар номеров координат, разделенных дефисом или запятой. Остальные виды *Set=* не требуют задания дополнительной информации. { Вместо номера координаты возможно задание символа *, который означает, что данное ограничение применяется ко всем элементам данного столбца (строки) матрицы. Помимо того, можно указывать не один номер, а интервал, например, 2-(5...10) означает набор из 6 элементов от (2,5) до (2,10). }[1].

Пример введения ограничений:

[1] Сейчас не работает

```
[Constraints]
Action=Fix        Set=OffDiag
Action=Fix-at-0   Set=List
10,1  10,2  10,3  10,4  10,5  10,6  10,7  11,1  11,2
11,3  11,4  11,5  11,6  11,7
Action=Release    Set=List
5,4   6,4
```

В данном примере сначала зафиксированы начальные значения всех недиагональных силовых постоянных, затем для некоторого их количества (по списку) заданы нулевые значения (которые также не будут меняться в процессе оптимизации), а затем включены в процесс минимизации значения некоторых других (ранее зафиксированных) недиагональных элементов.

2.2.9. Содержание раздела [Amplitudes]

Этот раздел используется для вывода интересующих значений среднеквадратичных амплитуд колебаний. Он начинается со списка пар атомов, для которых нужно рассчитать амплитуды, и содержит набор температур, для которых они рассчитываются. Пример задания раздела:

```
[Amplitudes]
Count=3
  1-3         2-3         1-2
Temperatures
0.000000  300.000000  400.000000
```

2.2.10. Содержание раздела [Pulay Factors]

Когда расчет ведется в режиме поиска масштабирующих множителей, этот раздел позволяет задать начальные значения множителей и набор множителей, подлежащих оптимизации (в противном случае все они будут изначально положены равными 1 и участвовать в процессе оптимизации). Часть множителей может быть положена равными друг другу из соображений, не следующих непосредственно из симметрии задачи (например, для всех пар одинаковых атомов). Структура раздела полностью видна из следующего примера:

```
[Pulay Factors]Count=17
// attribute = Fix for fixed factor, equal for equal factors
// may be absent if it is the only one
//No Name Value Attribute
   1 Q    1.10      Fix
   2 Q1   1.00      Q
   3 Q2   1.00      Q
   4 q1   1.00      q
   5 q2   1.00      q
   6 q3   1.00      q
   7 G1   1.00      Fix
   8 G2   0.90      Fix
   9 G3   1.00      Fix
  10 B1   1.00      B
  11 B2   1.00      B
  12 B3   1.00      B
```

```
 13 A12  1.00      A
 14 A13  1.00      A
 15 A23  1.00      A
 16 X    1.00
 17 T    1.00
```

Множители с равным значением *Attribute* будут поддерживаться равными в процессе оптимизации. Если *Attribute* не указан, данный множитель считается независимым. Если же *Attribute = Fix*, этот множитель не изменяется в процессе оптимизации. Имена координат должны соответствовать указанным в разделе [Coordinates]. Имена атрибутов могут быть любыми (кроме зарезервированного слова *Fix*).

В будущем (по мере развития программного комплекса) предполагается появление новых разделов. В настоящее время в электронографической версии уже имеются разделы

[Geometry Parameters] - для описания уточняемых параметров равновесной геометрии;
[Matrix Z3] - для ввода кубических силовых постоянных;
[Rotational Constants] - для ввода соответствующих экспериментальных значений;
[ED Data] - для задания экспериментальных кривых рассеяния.

2.3. Совместная обработка нескольких молекул.

Для совместной обработки двух и более молекул необходимо составить дополнительный входной файл, содержащий информацию о молекулах, подлежащих обработке, и о характере связей между ними. Этот файл (с расширением .PCK) должен стоять в командной строке вместо файла .MOL, который используется в обычном режиме. Структура файла .PCK описана ниже; она аналогична структуре остальных входных файлов. Выдача программы направляется в файл, имеющий то же имя, что и .PCK файл, и расширение .RES.

2.3.1. Содержание раздела [Molecules]

Этот раздел является обязательным. Он содержит список .MOL файлов; способ задания понятен из следующего примера.

```
[Molecules]
Count=4
1. adam-td.mol
2. adam-c3v.mol
3. adam-c2v.mol
4. adam-cs.mol
```

В настоящее время общее число молекул ограничено десятью.

2.3.2. Содержание раздела [Links]

Совместная обработка приобретает смысл, только если указаны связи между силовыми матрицами различных молекул (в противном случае совместный расчет должен дать тот же результат, что и последовательная обработка молекул).

Задание соотношений аналогично заданию ограничений. Также задается действие (*Action=*), которое может быть равным *Fix* или *Release*, а за ним в той же строке следует, к какому набору элементов его следует применить (*Set=*).

Возможные значения для *Set=*:

All - когда положены равными все силовые постоянные для всех участвующих молекул. Естественно, этот тип соотношений допустим только для молекул с одинаковыми по размеру силовыми матрицами; его имеет смысл использовать для обработки нескольких изотопных разновидностей молекулы с разной симметрией (если симметрия одинакова, лучше работать в обычном варианте программы).

!!! В настоящее время работает только этот вариант: остальные будут добавлены в скором времени. !!!

List - когда перечислены эквивалентные постоянные в разных молекулах. Пример задания списка:

```
Action=Fix Set=List
1(2,2) = 2(3,3) = 3(3,3)
```

В данном случае зафиксировано равенство каких-то диагональных силовых постоянных.

```
Action=Fix Set=All
Action=Release Set=List
1(1,*)
2(2,*)
```

Здесь сначала зафиксировано равенство всех элементов матриц, а затем из этих условий исключены элементы первого столбца и строки первой молекулы и элементы второго столбца и строки второй молекулы.

3. Выходные файлы

Программа формирует выходной файл с именем, совпадающим с именем входного файла, но с расширением .RES. Его структура достаточно проста: содержит много комментариев и ее лучше изучить на примере любой реальной выдачи.

Помимо этого файла, формируется файл с тем же именем и с расширением .VIB, содержащий информацию о частотах и формах колебаний молекулы. Этот файл используется в качестве входного программой MOLGR, которая демонстрирует колебания. Демонстрационная программа может воспринять и файл типа .MOL, но в этом случае она будет неспособна показать нормальные колебания молекулы.