

Г.Е. Слепнев, канд.хим.наук,
В.Н. Купреев, Л.В. Вилков,
докт.хим.наук, Г.И. Новиков,
докт.хим.наук

ПОСТАНОВКА ПРОГРАММЫ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ ДЛЯ ОБРАБОТКИ ЭЛЕКТРОНОГРАФИЧЕСКИХ ДАННЫХ НА ЭВМ ЕС-1020

В связи с дальнейшим совершенствованием методики обработки электронографических данных, а также повышением требований к получаемым структурным параметрам нами была поставлена на ЭВМ ЕС-1020 норвежская программа КСЕД26, которая первоначально была написана на языке FORTRAN USASI (MASTER) для ЭВМ СДС 6600. Программа КСЕД26 реализует метод наименьших квадратов [1-4] в применении к экспериментальной кривой молекулярной составляющей интенсивности рассеяния, заданной в интервале (S_{\min}, S_{\max}) . Программа состоит приблизительно из 1000 выполняемых операторов, что создало известные трудности при ее постановке на ЭВМ ЕС-1020. КСЕД26 является одной из программ норвежского комплекса [9], позволяющего определить параметры молекул, используя минимизацию функционала метода наименьших квадратов.

В данном случае последний имеет вид

$$F = \sum_s W_s \Delta_s^2 = \sum_s W_s [sM^E(s) - k \cdot sM^T(s)]^2, \quad (1)$$

где суммирование проводится по всем s , W_s - вес молекулярной составляющей интенсивности рассеяния в точке s ; k - масштабный множитель.

Теоретическое выражение для $sM(s)$, реализованное в программе, определяется в квазикинематическом приближении [6,7]. При этом предусмотрена ангармоничность колебаний. $sM(s)$ вычисляется по формуле

$$sM^T(s) = \sum_{i < j} \frac{A_{ij}}{r_{ij}} g_{ij}(s) \alpha_{ij} e^{-\frac{1}{2}u_{ij}^2 s^2} \sin((r_{ij} - B_{ij}s^2)s), \quad (2)$$

где величины α_{ij} , r_{ij} , u_{ij} и B_{ij} могут варьироваться; A_{ij} — кратность расстояния r_{ij} .

Функция $g_{ij}(s)$ рассчитывается по формуле

$$g_{ij}(s) = -\frac{f_i(s)|f_j(s)|}{I_{at}(s)} \cos(\eta_i(s) - \eta_j(s)). \quad (3)$$

Здесь амплитуда рассеяния на i -ом атоме является комплексной величиной:

$$f_i(s) = |f_i(s)| \exp(i\eta_i(s)), \quad (4)$$

где r_{ij} — расстояние между атомами "i" и "j"; u_{ij} — соответствующая среднеквадратичная амплитуда колебания; α_{ij} — вес данного расстояния. Предполагается, что молекула может существовать в нескольких конформациях, причем содержание конформеров может варьироваться B_{ij} — константа, определяемая агармоничностью колебания [8].

Минимизация функционала в данной программе осуществляется по схеме Ньютона—Гаусса.

Обозначим варьируемые параметры через p_1, p_2, \dots, p_n . Параметрами p могут быть α , r , u , B и k . Дифференцируя функционал $F = \sum_s W_s \Delta_s^2$ по параметрам, получим систему уравнений, где число уравнений равно числу варьируемых параметров.

$$\frac{\partial E}{\partial p_i} = 2 \sum_s W_s \Delta_s \frac{\partial \Delta_s}{\partial p_i} = 0. \quad (5)$$

Разлагая функцию Δ_s в ряд Тейлора и при этом ограничиваясь линейными членами, приходим к системе уравнений:

$$D \Delta p = H, \quad (6)$$

где Δp — вектор искомых изменений параметров,

$$D_{ij} = \sum_s W_s \left(\frac{\partial \Delta_s}{\partial p_i} \right) \left(\frac{\partial \Delta_s}{\partial p_j} \right); \quad H_i = - \sum_s W_s \Delta_s \frac{\partial \Delta_s}{\partial p_i}. \quad (7)$$

Матрица D в программе обращается методом квадратного корня.

В молекуле обычно только часть расстояний r_i варьируется как независимые параметры, когда известен эффект сокращения [5] или им пренебрегают. Программа KCED26 обращается подпрограмме SPESR (LINK), которая вычисляет зависящие расстояния как функции независимых расстояний и углов. Подпрограмма также рассчитывает производные от независимых расстояний по независимым для каждой конкретной молекулы. Производные вычисляются по разностной схеме. Возможно и аналитическое задание производных. Для пользователя нами написан стандартный вариант подпрограммы SPESR (LINK), в который нужно лишь вставлять блок вычисления координат атомов конкретной молекулы, исходя из ее параметризации и принятой системы координат.

Данная программа позволяет также вычислять некоторые статистические характеристики параметров.

Обозначим

$$\sigma_0^2 = \frac{1}{m-n} \sum_S W_S \Delta_S^2, \quad (8)$$

где m - число точек наблюдения, n - независимых варьируемых параметров. Ковариационная матрица M определяется выражением

$$M = \sigma_0^2 D^{-1}. \quad (9)$$

Стандартные отклонения параметров вычисляются из диагональных элементов матрицы M , т.е. стандартное отклонение для параметра r_i рассчитывается по формуле

$$\sigma_i = \sqrt{M_{ii}} = \sigma_0 \sqrt{(D^{-1})_{ii}}. \quad (10)$$

Недиагональные элементы матрицы M дают коэффициенты корреляции ρ ($-1 \leq \rho \leq 1$). Величина ρ , близкая ± 1 , показывает, что корреляция между параметрами велика.

$$\rho_{ij} = \frac{M_{ij}}{\sqrt{M_{ii} M_{jj}}} = \frac{M_{ij}}{\sigma_i \sigma_j}. \quad (11)$$

В программе предусмотрено также вычисление стандартных отклонений зависимых расстояний по формуле

$$\sigma_m^2 = \sum_{i=1}^k \sum_{j=1}^k \left(\frac{\partial r_m}{\partial r_i} \right) \left(\frac{\partial r_m}{\partial r_j} \right) M_{ij}, \quad (12)$$

где k - число независимых расстояний и углов.

Для более успешного проведения структурного анализа к программе КСЕД26 нами было добавлено несколько блоков.

Первый блок позволяет рассчитывать теоретическую и экспериментальную кривые радиального распределения, а также разностную кривую. Кривые радиального распределения в этом случае вычисляются по формуле:

$$f(r) = \frac{1}{\sqrt{2\pi}} \sum_{s_{\min}}^{s_{\max}} sM(s) e^{-bs^2} \sin(sr) \Delta s, \quad (13)$$

где b -- искусственный температурный фактор.

Второй блок рассчитывает обостренную теоретическую и экспериментальную кривые радиального распределения и их разностную кривую [10,11]. Обостренная кривая вычисляется в соответствии с выражением:

$$f(r) = \sqrt{2\alpha} \sum_{s_{\min}}^{s_{\max}} sM(s) e^{-\alpha(s_{\max}-s)} \sin(sr) \Delta s. \quad (14)$$

Величина α обычно выбирается равной 0,3-0,4. Этим блоком следует пользоваться в тех случаях, когда необходимо разрешить близко лежащие пики.

Третий блок предназначен для получения теоретической и экспериментальной кривых радиального распределения с учетом поправки на внеядерное рассеяние и с экстраполяцией экспериментальной $sM(s)$ в начальной области значениями теоретической $sM(s)$ для последней модели. Как правило, необходимость в таких кривых радиального распределения возникает на последнем этапе уточнения параметров молекулы. Получившиеся кривые не имеют отрицательных областей и могут рассматриваться как пропорциональные $\sum_{i < j} c_{ij} P_{ij}(r | b) / r$,

где P_{ij} -- плотность вероятности распределения расстояний между атомами "i" и "j", а c_{ij} -- константа.

Четвертый блок позволяет вводить новые начальные значения параметров, изменять коды этих параметров. Блоком удобно пользоваться в тех случаях, когда бывает необходимо просчитывать модели, различающиеся значениями тех или иных параметров, а также при варьировании от разных начальных приближений для проверки единственности найденного минимума.

Введение дополнительных блоков сделало программу более удобной при проведении структурного анализа.

В ы в о д ы

1. Программа написана для достаточно быстродействующей машины, что позволяет за относительно короткое время проводить большой объем вычислений; при этом процессорное время для среднего варианта не превышает 10 мин.

2. Расчет теоретической $sM(s)$ производится в квазиклассическом приближении. Амплитуды рассеяния на атомах берутся из таблиц Шефера [12]. Все это позволяет значительно повысить точность расчета функций $sM(s)$.

3. Введена весовая функция для экспериментальной $sM(s)$, что отражает неравноточность измерений в интервале (s_{\min} , s_{\max}).

4. Методически последовательно применен метод наименьших квадратов к задаче электронографии. Рассчитываются стандартные отклонения, корреляционная матрица независимых параметров, стандартные отклонения зависимых параметров.

5. Предусмотрено наличие различных конформаций в паре, возможность варьирования процентного содержания конформеров.

6. С целью уменьшения числа определяемых параметров возможно варьирование амплитуд группами, когда ряд амплитуд получает одинаковые изменения в процессе минимизации.

7. Блочная структура программы, наличие ключевых управляющих слов позволяют планировать схему минимизации, по-разному образуя группы параметров с учетом корреляции между ними.

Л и т е р а т у р а

1. Линник Ю.В. Метод наименьших квадратов и основы теории обработки наблюдений. М., 1958.
2. Худсон Д. Статистика для физиков. М., 1970.
3. Вилков Л.В. К использованию метода наименьших квадратов в газовой электронографии. Ж. структ. хим., 1964, 5, №6, с. 809.
4. Вилков Л.В., Садов Н.И. Об ошибках определения параметров в газовой электронографии. - Ж. структ.хим., 1967, 8, №2, с. 344.
5. Сивин С. Колебания молекул и среднеквадратичные амплитуды. М., 1971.
6. Рамбиди Н.Г., Спиридонов В.П. Исследование строения молекул высокотемпературного пара по рассеянию быстрых электронов. - "Теплофизика высоких температур", 1964, 2, №2, с. 280.
7. Спиридонов В.П., Рамбиди Н.Г., Алексеев Н.В. Современное состояние газовой электронографии. Теория рассеяния электронов молекулами. - Ж. структ.хим., 1963, 4, №5.

c. 779. 8. Kuchitsu K., Bartell L.S. Effect of anharmonicity of molecular vibrations on the diffraction of electrons. - J. Chem. Phys., 1961, 35, №6, p. 1945. 9. Andersen B. et al. Procedure and computer programs for the structure determination of gaseous molecules from electron diffraction data. - Acta Chem. Scand., 1969, 23, №9, p. 3224. 10. Traetteberg M. The molecular structure of 1,3,5-cycloheptatrien in the vapor phase as determined by sector electron diffraction method. - J. Am. Chem. Soc., 1964, 86, №20, p. 4265. 11. Traetteberg M., Bonham R.A. Auto- and cross-correlation functions used as tools for information retrieval in electron diffraction structure studies. - J. Chem. Phys., 1965, 42, №2, p. 587. 12. Schafer L., Yates A.C., Bonham R.A. New values for the partial wave electron scattering factor for the elements $1 \leq Z \leq 57$ and $72 \leq Z \leq 90$ for incident electron energies of 10, 40, 70, and 100 kev. - J. Chem. Phys., 1971, 55, №6, p. 3055.