УДК539.67:669.293

В.Е. Бахрушин

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА АНАЛИЗА СЛОЖНЫХ СПЕКТРОВ НА ЯЗЫКЕ R

Аннотация. Предложена программная реализация алгоритма анализа сложных спектров, основанная на использовании нелинейной минимизации суммы квадратов остатков и комплекса критериев, проверяющих их статистические свойства.

Ключевые слова: сложный спектр, разложение, программная реализация, язык программирования R, нелинейная минимизация, критерии адекватности.

Постановка проблемы. Проблема разложения сложных спектров на элементарные компоненты актуальна для широкого круга задач исследования материалов, технической диагностики и т. п. В частности она возникает при исследованиях механической релаксации [1], фотолюминесценции [2], нестационарной спектроскопии глубокоуровневых центров в полупроводниках [3] и др. Характерной особенностью этой задачи является то, что каждая компонента имеет определенный физический смысл. В частности, они могут соответствовать различным процессам в кристаллах, различным состояниям примесей и дефектов и т. п. Поэтому адекватная модель спектра не должна содержать «лишних» компонент. В связи с этим важной задачей является не только правильное приближение формы спектра, но и правильное определение числа компоненти их параметров.

Анализ последних исследований и публикаций. Существуют разные подходы к решению задачи разложения сложного спектра на составляющие [4, 5]. При этом одной из основных проблем является необходимость правильного выбора числа компонент. В [1, 6] для ее решения было предложено использовать систему критериев адекватности, которые оценивают соответствие остатков модели нормальному распределению, равенство нулю их среднего значения и их статистическую независимость. В этом случае может быть предложен цикл, который предполагает выполнение следующих основных этапов:

© Бахрушин В.Е., 2015

– задание числа компонент (его можно принять равным единице);

– задание начальных значений параметров компонент;

– подбор параметров компонент с использованием методов нелинейной минимизации суммы квадратов остатков модели;

– проверка адекватности модели по набору критериев;

– в случае если модель оказывается неадекватной – увеличение числа компонент на единицу и повторение последующей процедуры.

Постановка задачи. Разработать программную реализацию указанного алгоритма на языке R.

Изложение основного материала исследования. Язык R получил широкое распространение в сфере прикладной статистики и анализа данных [7]. Это обусловлено наличием большого числа специализированных функций и библиотек, предназначенных для решения таких задач, а также тем, что R является свободно распространяемым продуктом с открытым программным кодом. Это дает возможность неограниченной разработки программ для реализации новых алгоритмов и методов.

Модель сложного спектра, представляющего собой сумму n дебаевских пиков, можно записать [1] в виде:

$$Q(T) = \sum_{j=1}^{n} Q0_j \cosh^{-1}\left[\frac{E_j}{R}\left(\frac{1}{T} - \frac{1}{T0_j}\right)\right], \quad (1)$$

$$E_j = k_b T0_j \ln\frac{k_b T0_j}{hf}, \quad (2)$$

где $Q0_j, E_j, T0_j$ – параметры компонент, R, k_b, h, f – константы.

Для формирования спектра была написана такая функция:

```
Q = function(T, Q0, T0) {
E = R*T0*log(kb*T0/h/f)
QQ = matrix(nrow = length(T), ncol = length(T0))
for (j in1:length(T0))
QQ[,j] = Q0[j]/(cosh(E[j]/R*(1/T - 1/T0[j])))
rowSums(QQ)
}
Q(T, Q0, T0)
```

 ISSN 1562-9945

Она использует числовые векторы $T, Q0, T0$ и возвращает значения числового вектора Q, соответствующие значениям T.

Модель строится с помощью функции nls(). Для ее корректной работы необходимо задать массив данных Qempи начальные значения параметров:

```
Data = data.frame(T, Qemp)
names(Data) = c("T", "IFr")
mod = nls(IFr~Q(T, Q1, T01), Data, start = list(Q1 = c(1, 1, 1), T01 = c(450, 550, 650)))
```

Функция nls() в данном случае оценивает параметры модели минимизацией суммы квадратов ее остатков с помощью ньютоновского алгоритма. Однако при необходимости могут быть заданы аргументы, устанавливающие другие методы оценивания.

Следующим шагом является оценка адекватности модели. Для проверки соответствия остатков нормальному распределению используется критерий Андерсона – Дарлинга, который реализует функция ad.test() из библиотеки nortest. Равенство среднего арифметического остатков нулю проверяем по одновыборочному критерию Стьюдента с помощью функции t.test(). Наличие автокорреляции остатков первых пяти порядков проверяем с помощью функции durbinWatsonTest() из библиотеки car:

```
library(nortest)
ad.test(resid
t.test(resid)
library("car")
durbinWatsonTest(lm(resid~T), max.lag = 5)
```

Важным показателем адекватности модели является также различие дисперсий ее остатков и погрешностей эмпирических данных. Ее можно оценить по критерию Фишера. Однако стандартная функция var.test() в данном случае неприменима, т. к. не учитывает числа параметров модели спектра. Поэтому для этой цели может быть использован другой подход:

```
var_res = sum(((resid - mean(resid))^2)/(length(resid)-
length(Q1)))
FF = max(var_eps, var_res)/min(var_eps, var_res)
p_value = 1-pf(FF, length(T) - length(Q1), length(T) -
length(Q1))
```

Примеры результатов разложения спектра для случаевправильного и ошибочного выбора числа компонент показаны на рис. 1, 2.

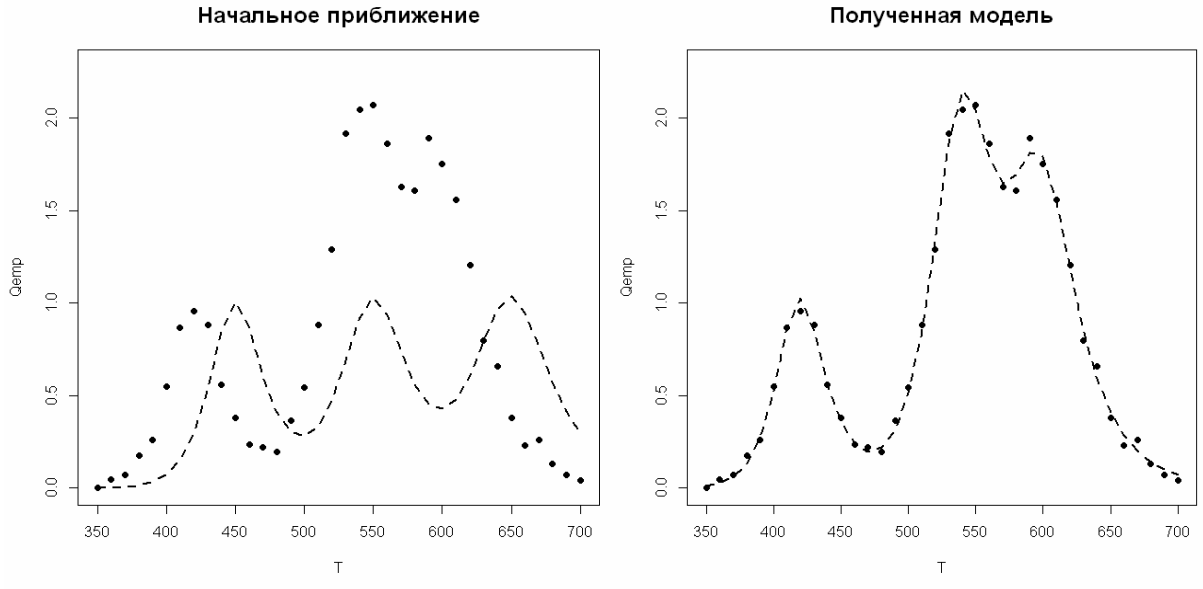

Рисунок 1 – Результаты разложения спектра
при правильном выборе числа компонент

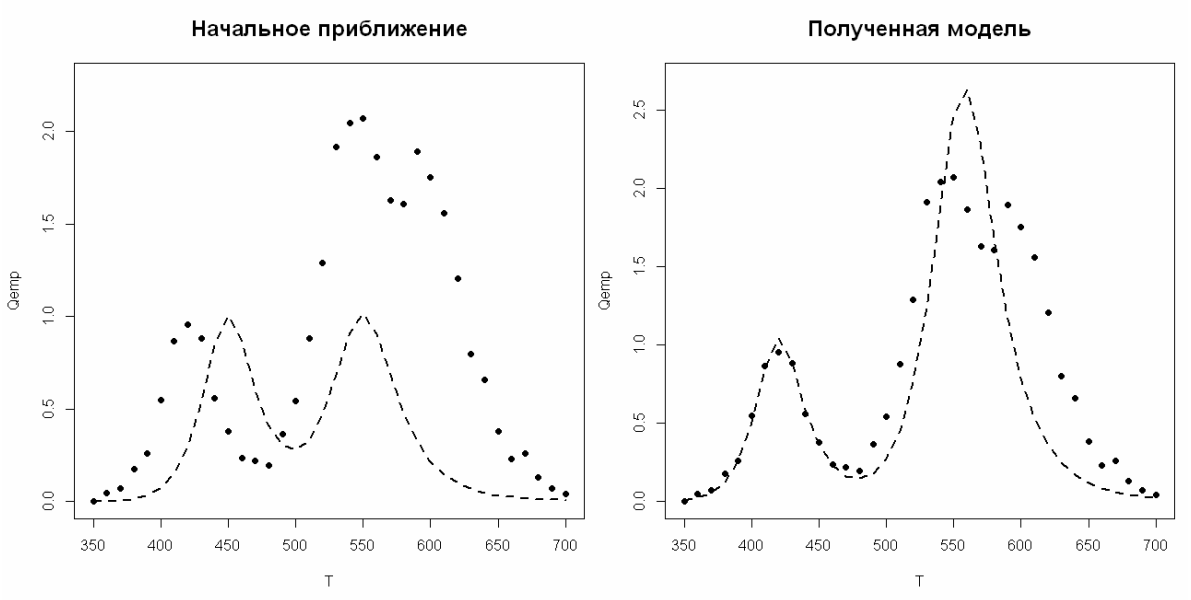

Рисунок 2 – Результаты разложения спектра
при ошибочном выборе числа компонент

 ISSN 1562-9945

Выводы. Предложена программная реализация алгоритма разложения сложного спектра на заранее неизвестное числодебаевских компонент. Тестирование показало корректность работы программы для широкого диапазона условий, соответствующего практически важным ситуациям.

ЛИТЕРАТУРА

1. Бахрушин В.Є., Чиріков О.Ю. Моделі та механізми механічної релаксації, пов'язаної з перебудовою домішково-дефектної підсистеми кристалів. ГУ "ЗІДМУ", Запоріжжя. – 2004. – 140 с.
2. Новосад С. С., Новосад И. С. Влияние примеси самария на спектральные характеристики кристаллов йодистого кальция // Журнал прикладной спектроскопии. – 2013. – Т. 80, № 1. – С. 78 – 84.
3. Qin-ShengZhu, H. Nagai, Y. Kawaguchi, K. Hiramatsu, N. Sawaki. EffectofthermalannealingonholetraplevelsinMg-dopedGaNgrownbymetalorganicvaporphaseepitaxy // J. Vac. Sci. Technol. – 2000. – V. A 18, N1. – P. 261–267.
4. Литвинов В. С. Ядерная гамма-резонансная спектроскопия сплавов / В. С. Литвинов, С. Д. Каракишев, В. В. Овчинников. – М.: Металлургия, 1982 . – 142 с.
5. Фок М. В. Разделение сложных спектров на индивидуальные полосы при помощи обобщённого метода Аленцева // Труды ФИАН СССР. – 1972. – Т. 59. – С. 3 – 24.
6. Бахрушин В. Е. Анализ сложных релаксационных спектров внутреннего трения твердых растворов на основе ниобия / В. Е. Бахрушин, А.Ю. Чириков // Высокочистые металлические и полупроводниковые материалы. Сборник докладов 9 Международного симпозиума / Под ред. В. М. Ажажи, В. И. Лапшина, И. М. Неклюдова, В. М. Шулаева. – Харьков: ННЦХФТИ, 2003. – С. 77 – 82.
7. Буховец А. Г. Статистический анализ данных в системе R. / А. Г. Буховец, П. В. Москалев, В. П. Богатова, Т. Я. Бирючинская – Воронеж: ВГАУ, 2010. – 124 с.