

ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

ДУБНА



10350

ЭКЗ. ЧИТ. ЗАЛА

Р10 - 10350

В.Б.Злоказов

**UREAK - СПЕКТРООРИЕНТИРОВАННАЯ ПРОГРАММА
РАЗЛОЖЕНИЯ СМЕСЕЙ**

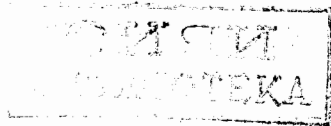
1976

P10 - 10350

В.Б.Злоказов

**UREAK- СПЕКТРООРИЕНТИРОВАННАЯ ПРОГРАММА
РАЗЛОЖЕНИЯ СМЕСЕЙ**

Направлено в "Computer Phys. Communicat."



UREAK - спектроориентированная программа разложения смесей

В работе описана программа разложения смеси регрессий, ориентированная на анализ спектров ядерного или молекулярного излучения (поглощения). Особенности программы: 1) общность в трактовке понятия компоненты, т.е. вклада в спектр отдельного излучателя (поглотителя); 2) независимость от возможности описать форму компоненты аналитически, т.к. в качестве модели могут быть взяты экспериментально измеренные линии (полосы); 3) минимальность числа параметров для описания компонент произвольной формы; 4) возможность налагать связи на параметры; 5) возможность контроля качества и достоверности получаемых оценок параметров.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1976

UREAK - Spectrooriented Routine for Mixture Decomposition

A computer routine for mixture decomposition is described, oriented to be applied to the analysis of spectra of nuclear or molecular radiation (absorption). Unusual features of the routine are the following: 1) generality in treatment of component notion, i.e., by a component a spectrum fraction is meant, which is produced by a single radiator (absorber); 2) independence of the possibility of describing the component shape analytically, because experimentally measured lines (bands) may be taken as shape models; 3) minimum of parameter number for the description of the components of arbitrary shape is required; 4) ability to impose the constraints on the parameters; 5) ability to check the quality and the likelihood of the obtained parameter estimates.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Preprint of the Joint Institute for Nuclear Research, Dubna 1976

ВСТУПЛЕНИЕ

Данные физического эксперимента, в частности спектры ядерного излучения, могут формально рассматриваться как линейная смесь полезных и фоновых компонент, на которые наложена статистическая ошибка. Математически это выражается так:

$$s(x) = \sum_{i=1}^n \phi_i(x) + b(x) + e(x),$$

где $s(x)$ - анализируемые данные, $\phi_i(x)$ - i -я полезная компонента, $b(x)$ - сумма всех фоновых компонент, $e(x)$ - статистическая ошибка с нулевым средним и дисперсией $D(x)$. Цель программы UREAK - разложение $s(x)$ на $\phi_i(x)$ и $b(x)$ оптимальным в определенном смысле способом для возможно более широкого класса аппаратных спектров или других спектроподобных смесей. Метод, реализованный в UREAK, был создан после тщательного анализа имеющегося опыта в области обработки спектров и может рассматриваться как обобщение обычно используемых методов в этой области. Его главные особенности состоят в следующем.

1) Спектр формально рассматривается как смесь пикоподобных и (или) непрерывных компонент (имеющих, возможно, пикоподобные субкомпоненты); под одиночной компонентой понимается либо одиночный пик, либо непрерывная линия (α -, β -спектры), либо их фиксированная комбинация (спектр изотопа в γ -или X-спектрах).

2) Для каждого типа компонент спектров должна быть построена модель компоненты $m_i(x)$ (для фона $b(x)$): это может быть аналитическая функция или экспериментально измеренная гистограмма соответствующей одиночной компоненты (во 2-м случае недостающие значения $m_i(x)$ создаются интерполяцией).

3) На основе общих свойств логических отношений между моделями $m(x)$ и соответствующими реальными объектами $\Gamma(x)$ вводятся параметры в эти отношения так, что

$$\Gamma(x, A, P, W, K) = Am \left(\frac{x-P}{Kx+W} \right); \quad (1)$$

эти параметры, будучи общими характеристиками вышеупомянутых отношений, имеют в то же самое время ясный физический смысл: A – амплитуда, P – положение, W – ширина, K – коэффициент линейной зависимости ширины от канала x .

4) Далее для получения методом наименьших квадратов оценок параметров A_i, P_i, W_i, K_i i -й компоненты и их ошибок используется обычная техника регрессионного анализа; частные производные $\Gamma(x, A, P, W, K)$ по параметрам вычисляются на основе (1):

$$\frac{\partial \Gamma}{\partial A} = m(z), \quad \frac{\partial \Gamma}{\partial P} = -\frac{A}{Kx+W} m'_x(z), \quad \frac{\partial \Gamma}{\partial W} = -\frac{Az}{Kx+W} m'_x(z), \quad (2)$$

$$\frac{\partial \Gamma}{\partial K} = -\frac{Azx}{Kx+W} m'_x(z),$$

где $z = (x-P)/(Kx+W)$.

Аналогично площадь i -й компоненты вычисляется при использовании (1):

$$S = \int \Gamma(x, A, P, W, K) dx = (A+PK)W \int m(z) \frac{1}{(1-Kz)^2} dz. \quad (3)$$

Математическое обоснование и подробное описание метода дано в работах^{/1-3/}. Хотя метод был применен к анализу различных аппаратных спектров, фактически он может быть использован в большем классе задач разложения спектроподобных смесей.

1. ФОРМАЛЬНОЕ ОПИСАНИЕ ЗАДАЧИ

Задачи разложения спектров могут быть разделены на 3 подгруппы согласно трем способам выбора величин,

которые рассматриваются как дискриминаторы одиночных полезных компонент:

1) 1 подгруппа: положение – дискриминатор (например, анализ спектров, состоящих из пикообразных полезных компонент: α - и γ -спектры, спектры электронов конверсии, нейтронных резонансов и т.д.);

2) 2 подгруппа: ширина – дискриминатор (например, анализ спектров, состоящих из непрерывных компонент: спектры радиоактивного распада нескольких источников, β -спектры и т.д.);

3) 3-я подгруппа: и положение, и ширина – дискриминаторы (например, спектр нейтронной дифракции, когда пик упругого рассеяния расположен на пике неупругого рассеяния и оба – на фоне).

Возможны и дальнейшие разделения задач согласно:

1) количеству параметров для описания полезных компонент; A, P, W, K (одиночная компонента – сумма $K\alpha$ – , $K\beta \rightarrow L\alpha$ – , $L\beta$ – линий в спектре X -лучей); A, P, W (одиночная компонента-пик в γ -спектре); A, W (одиночная компонента – кривая радиоактивного распада);

2) знаку амплитуды (плюс – для спектров реакций рассеяния или эмиссии, минус – для спектров поглощения).

Итак, мы ставим задачу разложения спектра формально:

$$\min \sum_{j=1}^M \frac{1}{D(x_j)} \left(s(x_j) - \sum_{i=1}^n A_i \phi_i \left(\frac{x_j - P_i}{K_i \phi_i + W_i} \right) \right)^2, \quad (4)$$

где одна из $A_i \phi_i$ – функция фона.

В работе^{/1/} показано, что для единственности решения задачи минимизации (4) необходимы следующие ограничения на параметры:

$$|A_i| \geq s_i, \quad P_{i+1} - P_i \geq r_i \quad (\text{для задач 1-й подгруппы}); \quad (5)$$

$$|A_i| \geq s_i, \quad W_{i+1} - W_i \geq r_i \quad (\text{для задач 2-й подгруппы}),$$

где s_i, r_i должны быть положительными, согласованными с дисперсией помехи $e(x)$ и предварительно определенными (т.е. перед минимизацией) и $K_i x_j + W_i \geq c_i$ должны выполняться.

Далее могут быть иногда добавлены следующие ограничения, основанные на физических соображениях:

$$\phi_\ell = \phi_k, \ell, k \in \{i\} \quad (\text{одинаковая форма некоторых компонент}); \quad (6)$$

$$K_j = 0 \text{ или } P_j = 0 \quad (\text{для анализа пиков или экспонент});$$

$$W_j = k_j W_i + d_j \pm \epsilon_j, j > i \quad (\text{для 2-й подгруппы, } P_j = k_j P_i + d_j \pm \epsilon_j);$$

например, для γ -спектров $k_j = 1$, $d_j = \epsilon_j = 0$ в каждом коротком интервале.

Для улучшения качества оценок можно наложить ограничения на амплитуды и положения (и на каждый параметр):

$$A_j = b_j A_i \pm \epsilon_j, P_j = P_i + k_j \pm \delta_j, j > i, p_{i\ell} \leq P_i \leq p_{iu} \quad (7)$$

при условии, что область, ограничиваемая (5)-(7), содержит истинные значения p_0 .

Чтобы минимизировать (4) при связях (5)-(7), применяется следующая двухстадийная процедура:

1) на первой стадии используется демпфированный процесс Гаусса-Ньютона с проектированием всех промежуточных оценок на область (5)-(7);

2) на второй - процесс Ньютона при условии положительной определенности его матрицы.

Такая процедура делает обрыв процесса минимизации более надежным; с другой стороны, положительная определенность конечной матрицы есть вернейший признак того, что решение соответствует минимуму (4) (а не другим нулям градиента (4)).

Байесовские оценки параметров могут быть получены при условии, что начальные значения p_0 и их матрица Грамма R_{ij} дают априорное распределение оценок. В этом случае величина $(R_{ij}(p_i - p_{i0}), (p_j - p_{j0}))$ прибавляется к (4).

Лучший способ сформировать модели (как полезных компонент, так и фона) - строить их на основе теоретических соображений, но если это невозможно, можно использовать для этих целей экспериментально измеренные компоненты, прибегая к байесовой рекурсивной

процедуре. Такие компоненты должны иметь статистику по меньшей мере на порядок выше, чем анализируемые компоненты. Можно использовать сглаживающую процедуру для экспериментально измеренных пиков, выбранных для построения модели, следующим образом: предположим, что интервал спектра (α_0, α_1) содержит пик и линейный фон, которые подгоняются функцией:

$$f(x, p) = A \exp \left\{ - \left(\frac{x - P}{P(x)} \right)^2 \right\} + c + az,$$

где $P(x)$ - полином Чебышева 3-го порядка по x , $z = \frac{2x - \alpha_0 - \alpha_1}{\alpha_1 - \alpha_0}$; выражение $A \exp \left\{ - \left(\frac{x - P}{P(x)} \right)^2 \right\}$ берется в качестве модели пика. Если результаты подгонки недостаточно хороши, модель формируется простым способом: фон вычитается и остаток берется в качестве модели.

2. СТРУКТУРА ПРОГРАММЫ

Спектр делится на интервалы, которые могут содержать n компонент, или изолированных, или перекрывающихся, и фон. Каждая изолированная компонента может быть взята для формирования модели. Не более двух различных моделей полезных компонент и одна модель фона (если таковой нет, тогда полином Чебышева порядка не выше третьего) могут использоваться в каждом интервале спектра. В качестве чисел s_i и r_i программой берутся максимальные значения амплитуд и разностей положений, гарантирующие однозначную статистическую интерпретацию результатов минимизации (4) (см. /1/).

UREAK представляет собой фортранную подпрограмму, которая должна вызываться главной программой (MP). Возможны 2 режима взаимодействия подпрограммы UREAK и MP:

1) IM1 - нет обмена информацией (простой вызов); в этом случае входные данные читаются с карт, а выходные печатаются; оценки основных параметров и их ошибки могут по желанию пользователя быть записаны в специальный массив TAB);

2) IM2 - активный обмен информацией (например, в случае, если UPEAK должен быть лишь процедурой внутри более сложной программы); тогда входные данные читаются из памяти, выходные пишутся в память (как фактические параметры) и по желанию могут быть распечатаны и (или) накоплены в TAB.

Структура UPEAK следующая:

а) основная подпрограмма UPEAK, организующая работу всего комплекса и вызываемая MP;

б) сегменты (фактически, а формально - подпрограммы; они готовы к оформлению в сегменты, например, для сокращения памяти):

PREPR - подготавливает все необходимые данные для подгонки;

FITPR - осуществляет подгонку и анализ ее результатов;

OUTPR - осуществляет вывод результатов подгонки;

SIMPR - строит (по желанию) простые оценки полезных параметров одиночных компонент;

SAMOD - строит (по желанию) модели на базе реальных компонент;

FPEAK - "заглушка" (для нестандартных целей);

в) подпрограммы для целей общих вычислений:

LSM76 - находит минимум (4) при связях между параметрами; включает в себя вспомогательные подпрограммы OUTIN и PODEM;

LAPIN - осуществляет линейную интерполяцию;

SMXIN - обращает симметричную матрицу;

г) вспомогательные подпрограммы, используемые другими подпрограммами: AMXNS, APPRO, ARIT2, ARITH, ARRAM, CALF, MOPR, OPOSP, POLCA, PRANA, INIPR, PRIPR, PROGM, PROO2, PROOP, RESCA, RINC0, SEMAT, STAR, UPDAM, WAREC, WHDEF, WIPRO; графики печатаются подпрограммой GRAPH; каждая из этих подпрограмм содержит комментарии, описывающие их назначение, и другую необходимую информацию;

д) помеченные общие блоки: C0BL, M0D1, M0D2, M0DB, SBL0CK, C0NAR, RBUCH, LIC0N; большинство переменных этих блоков описано в секции входных данных или комментариях за исключением:

в C0BL: PRI - текущий указатель режима печати;

KM - количество моделей полезных компонент;

в M0D1, M0D2, M0DB: RL1, RL2, RLB - те же величины, что и NLI, NL2, NLB, только в вещественном формате; CS1, WII..., CO - вспомогательные переменные для операций с моделями;

в SBL0CK: СВ и RB описаны в комментариях к LSM76;

P - массив параметров, их нижних и верхних границ;

F, Y, OP - вспомогательные массивы для LSM76;

в RBUCH: AMX, AMN, KX, KN - максимум и минимум и их индексы:

WH - вес i-го значения спектра;

в LIC0N: IAZ, IPZ - указатели точных связей для

A_i и P_i;

RESA - массив g_i (разрешений);

Общая схема работы программы следующая:

а) IM1: все входные данные делятся на группы; каждая из них снабжается номером группы, который пробивается на отдельной карте и предшествует группе; UPEAK вводит эти группы в последовательности, не фиксированной и определенной пользователем, и реагирует на них, в основном запасая их в соответствующих ячейках памяти; однако некоторые группы вызывают большие действия (подгонку, формирование и нормализацию модели, остановку работы).

б) IM2: те же группы читаются из специального массива 0NL, результаты анализа записываются в 0NL и TAB, и возврат из UPEAK делается после обработки каждой группы; т.е. в этом режиме UPEAK следует вызывать в цикле.

Пользователь определяет последовательность групп в обоих режимах, следуя логике обработки спектра; введенная группа остается в своих ячейках, пока новая группа с таким же номером не будет введена, т.е. требуется лишь ввод групп, обновляющих данные.

UPEAK вызывается так:

CALL UPEAK (SP, DISP, 0NL, ANP, TAB, AM1, AMD, RM, NM), где SP(LS) - массив спектра длиной LS;

DISP(LS) - дисперсия спектра; если отсчеты имеют распределение Пуассона, UPEAK формирует дисперсию автоматически (байесова оценка):

$DISP(K) = SP(K) + 1, K=1, LS; \quad (7)$

иначе (например, если был вычтен измеренный фон) она должна быть вычислена перед вызовом UPEAK и помещена в DISP;

$\emptyset NL(MS+3)$ - массив для взаимодействия;

$\emptyset NL(1) = A\emptyset 1 = 0(1)$ - режим взаимодействия IM1 (IM2);

$\emptyset NL(2) = DEM = 0(1,2)$ - решаются задачи 1-й (2-й, 3-й) подгрупп;

остальные ячейки $\emptyset NL$ содержат в IM1 произвольную информацию;

в IM2 значения остальных ячеек $\emptyset NL$ следующие:

$\emptyset NL(3) = RNG$ - номер группы;

$\emptyset NL(4-...)$ - данные группы с номером RNG (максимальный размер = MS);

$ANP(3 * MM)$ - вспомогательный массив; MM - максимальная длина интервалов, на которые делится спектр.

$TAB(6 * NP + 1)$ - массив результатов, NP - общее количество обработанных компонент; содержимое этого массива после анализа:

$TAB(1) = NP$ (тип - real) (не включая компоненты для моделей);

$TAB(2) - TAB(7)$ - положение, ошибка положения, площадь, ошибка площади, ширина, ошибка ширины 1-ой компоненты;

$TAB(8) - TAB(13)$ - положение, ошибка положения и т.д.

2-й компоненты; и так далее;

$AMI(NM, NM), AMD(NM), RM(NM, NM)$ - вспомогательные массивы;

NM - максимальное число параметров на интервал спектра;

Число параметров вычисляется по формуле: $NP = 2N + KM + LI + 4$, где N - число компонент в интервале,

KM - количество моделей полезных компонент на интервал,

$LI = 1(0)$ - параметр K (не) должен использоваться;

4 параметра используются для описания фона.

Если дисперсия формируется по (7) или байесовы оценки не используются, можно опустить резервирование DISP или RM и писать вместо них произвольные идентификаторы.

3. ВХОДНЫЕ ДАННЫЕ

UPEAK использует для ввода карт единственный формат: 10E8.0.

Описание данных состоит из описания групп данных.

DATA I. RKM, RNN, RIDI, RMP, DEM, RBFM - 1 карта;

RKM (в программе KM) - количество моделей на интервал; если RKM=0, то автоматически: модель 1 = GAUSSIAN, модель 2 = LORENTZIAN (для 2-й подгруппы: модель 1 = модель 2 = экспонента);

RNN=0(1) - одиночные компоненты должны подгоняться (вычисляться по простому алгоритму: сложение отсчетов и вычитание фона);

RIDI=0(1)-DISP(K) = SP(K) + 1, K - канал (DISP формируется пользователем);

RMP=1(0) - гауссова и лоренцева модели (не)распечатываются;

DEM=0(1,2) - решается задача 1(2,3) подгруппы;

RBFM=0 - фон имеет полиномиальную форму;

RBFM=1,3 - модель фона дается пользователем;

RBFM=2,3 - полезные компоненты 1-й модели имеют 4

параметра: A, P, W, K; PKM=2, полезные компоненты 2-й модели имеют 3 параметра: A, P, W; W - одинакова для компонент одинаковых моделей; K - общий параметр в каждом интервале спектра;

DATA 2.	R1, RMP1	-	1 карта;
	F1	-	колода карт;
	(R2, RMP2)	-	(1 карта);
	(F2)	-	колода карт;
	(RB, RMPB)	-	(1 карта);
	(FB)	-	колода карт;

F1(F2) содержит таблицы модели 1 (модели 2); FB содержит модель фона; каждая таблица содержит R1(R2, RB) чисел перед максимумом, 1 число в максимуме и

R1(R2, RB) чисел после максимума; если асимметрия имеет место, недостающие значения заполняются нулями; RMPi=1(0) - модель i (не) печатается;

если RKM=1, карты с R2, RMP2 и F2 отсутствуют;

если RBFM=0 или 2, карты с RB, RMPB и FB отсутствуют; если некоторые Ri=0, это означает, что ввод i-й модели не требуется и старая модель остается; в этом

случае F_i не нужен. Если вводятся все 3 модели, R_i не должны превосходить 99,49,49 (если вводятся только F_i , F_B , - соответственно 148, 49).

Введенные модели автоматически нормализуются.

DATA 3. ACH, B, RMPS - 1 карта;
SPIN - колода карт.

ACH - длина, B - начало вводимого интервала спектра;
SPIN - вводимый интервал спектра (ACH каналов);
RMPS=I(0) - введенный интервал (не)печатаются;
если $B=0, 1$ присваивается B; если спектр был введен в массив SP перед вызовом UPEAK, DATA 3 не требуется.

DATA 4. CB(4)-CB(13) - 1 карта;
DATA 4 служит для управления процессом минимизации.

CB(4) - шаг, с которым берутся каналы для вычисления (4) (в том случае, если (4) следует вычислять лишь на части множества каналов);

CB(5)=NG - верхний предел количества итераций Гаусса-Ньютона;

CB(6) - желаемая точность минимизации;

CB(7) - желаемое отношение χ^2 к числу степеней свободы (если CB(7)=-1, тогда оно определяется программой);

CB(8), CB(9) - нижний и верхний пределы количества итераций Ньютона;

CB(10)=1(0) - матрица (не)проверяется на положительную определенность;

CB(11)= λ - длина шага приращений параметров ($0 < \lambda \leq 1$);

CB(12)= a - признак регуляризации задачи;

$a=0$ - нет регуляризации;

$0 < a < 1$ - итерации демпфируются (заменой диагональных элементов матрицы a_{ij} на $(1+a)a_{ij}$); $a=1$ - байесовские оценки используются;

CB(13)=PRI=1(0) - результаты каждой итерации (не)печатаются; рекомендуемые значения CB(4-13) : 1; 5; 1; -1; 0; 1; 1; 1; 0,0001; 0; в случае слабых или перекрывающихся компонент $a \approx 0,01-0,1$; $\lambda \approx 0,1-0,5$; NG $\approx 10-15$.

DATA 5. PRS, FIS, \emptyset US - 1 карта
PRS=1(0) -- информация о подготовке (не)печатаются;

FIS=1(0) - факторы корреляции (не)печатаются; если i -й фактор много больше 1, имеют место трудности при вычислении i -го параметра;

\emptyset US=0 - выходные результаты не печатаются;

\emptyset US=1(2) - графики (численные значения) спектра и результатов разложения печатаются;

\emptyset US=3 - печатаются и графики, и числа; .

если \emptyset US отрицательно, IABS(IFIX(\emptyset US)) используются, как это описано выше, но результаты анализа не посылаются в TAB.

DATA 6 - нет информации;

карта с номером 6 означает возврат из UPEAK к MP.

DATA 7 - A0, A1, RN, RE, RILC, PIR - 1 карта;

A0, A1 - начало и конец анализируемого интервала спектра;

RN - число компонент в [A0, A1]; если N=0, подгоняется лишь фон; $N = \text{FIX}(RN)$;

RE=0(-1) - есть (нет) фон(a) в интервале; границы параметров формируются автоматически;

RE=1 - все начальные значения параметров, их нижние и верхние границы вводятся;

RILC=1(0) - линейные зависимости A_i и P_i (не)используются (см. DATA 14);

PIR=1(0) - чувствительности s_i и разрешения r_i используются мягким (жестким) способом; жесткий - абсолютные пределы положений фиксированы; мягкий - фиксированы только минимальные разности положений.

DATA 8. DP \emptyset S, DBACK, CAM, CRE, W1, DW1, W2, DW2 - 1 карта;
(DBK, DWB, DLWB, B \emptyset W, RL \emptyset CF, DR \emptyset LCF) - (1 карта);

DP \emptyset S - точность (в каналах) начальных значений положений (1-я подгруппа) или ширины (2-я подгруппа);

DBACK - степень фоновой полинома (0, 1, 2, 3); если введенное DBACK < 0, то DBACK = min(N, 3); N - число компонент в интервале;

CAM - каждое s_i умножается на CAM;

CPE - каждое r_i умножается на CRE; таким образом, можно контролировать эти величины s_i и r_i и менять их значения;

W1(W2) - начальные значения ширины для 1-й подгруппы или положений для 2-й подгруппы компонент 1(2) модели;

DW1(DW2) - точность (в каналах) W1(W2); если DW2 < 0, то на W1, W2 налагается связь: W1 = W2 ;

DBK, DWB, DLWB - точность фоновых параметров, P, W, K (если RBFM=1,3);

BOW - нижний предел $Kx_j + W$ (если RBFM=1,3);

RICF - начальное значение параметра K для полезных компонент (если 4 параметра A, P, W, K описывают их форму, т.е. если RBFM=2,3);

DRLCF - точность RLCF;

2-я карта DATA 8 требуется, если RBFM=1,2,3; иначе она отсутствует;

для обработки спектров поглощения CAM должно быть отрицательным; если только 1 модель используется, W2 и DW2 могут отсутствовать; используя DATA 8, UPEAK формирует границы параметров.

DATA 9.	TAM	-	колода карт;
	ACC	-	колода карт;
	BCC	-	колода карт.

TAM: если TAM(i)=k, i -ая компонента относится к k-й модели; ACC, BCC - ширина i -й компоненты W_i связана с общей шириной W соотношением $W_i = ACC(i) * W + BCC(i)$; ввод DATA 9 следует за вводом DATA 7; в начале работы: TAM(i)=1; ACC(i)=1; BCC(i)=0; i=1, N; если такие значения подходят, ввод DATA 9 может быть опущен

DATA 10. AP - колода карт;

AP есть массив или начальных значений положений-ширин (RE=0), или начальных значений фоновых параметров (отдельно) и начальных значений положений-ширин (RE=0, RBFM=1,3), или всех начальных параметров, их нижних и верхних границ (RE=1); для второй подгруппы ширины играют роль положений; если RE=1 и IM1, колоды карт параметров, их нижних и верхних границ должны быть отдельными; последовательность параметров такова:

4 фоновых параметра (A, P, W, K или 4 коэффициента полинома);

N амплитуд; N положений (ширин); KM ширин (положений); параметр K - общий для всех полезных компонент 1-й модели; K присутствует, если RBFM=2,3; ввод DATA 10 инициирует подгонку; если RNN=1 и N=1, AP не нужен (лишь число группы 10).

DATA 11. RICM, BEM, FSM, A0, A1	-	1 карта;
FAR(i), i=1,8	-	1 карта;

ввод DATA 11 начинает формирование моделей на основе некоторых одиночных полезных компонент;

RICM - номер модели (1 или 2); для фона - 0;

BEM=1(0) - старая модель (не)учитывается с помощью метода Байеса; старая модель может быть, в частности, введена с карт, новая модель может быть отперфорирована на карты;

FSM=0 - нет ввода старой модели; нет вывода новой модели;

FSM=1 - ввод старой модели; нет вывода новой модели;

FSM=2 - нет ввода старой модели; вывод новой модели;

FSM=3 - ввод старой модели; вывод новой модели.

A0, A1 - начало и конец интервала, содержащего одиночную компоненту (и, возможно, линейный фон) для модели; FAR(1) - минимальное число каналов на каждом конце, где зарегистрирован только фон; если фон отсутствует, FAR(1)=0; если FAR(1) отрицательно, модель формируется просто: вычитанием фона и упорядочением остатка; FAR(2)-FAR(8) - если FAR(i)=0, i -й параметр в функции, подгоняющей модель, варьируется, иначе фиксируется и равняется FAR(i) (за исключением 1-го параметра, который не управляется FAR(1)); последовательность параметров: A, P, W0, W1, W2, W3, C, L; где W_i - коэффициент полинома ширины, C и L - постоянный и линейный параметры фона.

Вся остальная информация для подгонки модели берется из DATA 2, DATA 3, DATA 4, DATA 5, введенных ранее. Если подгонка не удастся, программа формирует модель простым способом (описанным выше).

DATA 13. DATA 13=DATA 7 +DATA 10;

номера DATA 7 и DATA 10 не требуются, только их массивы (отдельно);

оперирование с DATA 13 возможно лишь в IM1;

DATA 14.	TMA	-	колода карт;
	ACA	-	колода карт;
	EAA	-	колода карт;
	PCA	-	колода карт;
	EPA	-	колода карт.

Каждый из этих массивов содержит N чисел;

TMA: если $TMA(j)=i$, ($j>i$) амплитуда и положение j -й компоненты зависят от амплитуды и положения i -й; ACA, EAA: $A(j)=ACA(j) * A(i) \pm EAA(j)$;

если $EAA(k)=C$, то $EAA(k)=SQRT(A(i)) * ABS(C)$;

PCA, EPA: $P(j)=P(i)=PCA(j) \pm EPA(j)$;

ввод DATA 7 должен предшествовать вводу DATA 14; в начале работы (в IM1) автоматически;

$TMA(i)=i$, $ACA(i)=1$, $PCA(i)=0$, $EAA(i)=EPA(i)=0$, $i=1, N$;
если $RILC=0$ (в DATA 7), DATA 14 отсутствует вообще.

Рекомендуемая последовательность ввода DATA 1-14: DATA 1, DATA 2, DATA 4, DATA 5, затем в цикле DATA 3, DATA 8, DATA 7, DATA 10 (или DATA 13); DATA 3 и DATA 8 могут быть введены до цикла; DATA 9, DATA 11, DATA 14 могут встречаться как в цикле, так и вне его.

4. ВЫХОДНЫЕ ДАННЫЕ И ТЕСТ ПРОВЕРКИ

Результаты анализа таковы:

- 1) оценки величин, представляющих интерес: положение, площадь, ширина и их ошибки;
- 2) различные сообщения (например, диагнозы ошибок) о разложении и его результатах:
 - а) есть ли параметры на границе симплекса (5)-(7);
 - б) χ^2 на степень свободы (CHIFD), достигнутая точность (EPS), число использованных итераций Гаусса-Ньютона и Ньютона, уклонение χ^2 (DCH), значение байесовского члена в (4) (CHIR);

- в) факторы корреляции, которым предшествуют их сумма и среднее;
- г) оценки, их ошибки, графики и числа, иллюстрирующие разложение (если $CHIFD < CB(7)$, графики и числа не печатаются).

В IM2 $\emptyset NL$ после подгонки (и формирования модели)

содержит:

- $\emptyset NL(4)=1(0)$ - процесс минимизации (не)сошелся;
- $\emptyset NL(5)=P\emptyset BI=0(1)$ есть (нет) параметры на границе;
- $\emptyset NL(6)=AFD$ - число степеней свободы;
- $\emptyset NL(7)=EPS$ - достигнутое;
- $\emptyset NL(8)=N1$ - число использованных итераций Гаусса-Ньютона;
- $\emptyset NL(9)=N2$ - число использованных итераций Ньютона;
- $\emptyset NL(10)=1(0)$ - матрица (не)обращена;
- $\emptyset NL(11)=CHIFD$;
- $\emptyset NL(12)=DCH$;
- $\emptyset NL(13)=CB(7)$;
- $\emptyset NL(14)$ =число каналов с ненулевыми весами;
- $\emptyset NL(15)$ = значение байесова члена в (4).

Полученные оценки, их ошибки и матрица ковариаций могут быть также получены из массивов P, F и AM1 в общем блоке SBL \emptyset CK;

Замечания по использованию подпрограммы UPEAK.

- 1) сбои во время вычислений (переполнение, деление на 0 при обращении матрицы) могут быть устранены, если положить $CB(12) \neq 0$;
- 2) если процесс не сходится, могут быть предприняты следующие меры: взять меньшую длину шага ($CB(11)$), большее CAM и $CRE(DATA 8)$, большее a ($CB(12)$), попытаться зафиксировать положения или ширины ($DP\emptyset S = 0$ или $DW1=DW2=0$), наложить связи на параметры.

Тест проверки состоит в разложении смесей, содержащих:

- 1) 6 гауссианов и линейный фон;
- 2) 2 лоренцевых пика абсорбции и постоянный фон;
- 3) 2 пика (1 гауссиан и 1 лоренциан) и сложный фон;
- 4) 3 пика с переменной шириной и постоянный фон;
- 5) 2 негауссовских пика (перекрывающихся) и 1 пик (одиночный) для построения модели;

Все пики, упомянутые выше, перекрываются. В качестве МР служит следующая программа:

```
PROGRAM MAIPR (INPUT, OUTPUT)
DIMENSION SP(4096), ANP(1000), TAB(500), AM1(20,20), AMD(20), ONL(50)
TAB(1)=0
ONL(1)=0
ONL(2)=0
CALL UPEAK (SP, DUMMI, ONL, ANP, TAB, AM1, AMD, DUMMI, 20)
END
```

В заключение автор выражает глубокую благодарность проф. Н.Н.Говоруно за поддержку данной работы и полезные обсуждения.

ЛИТЕРАТУРА

1. Злоказов В.Б. Препринт ОИЯИ, Р11-10186, Дубна, 1976.
2. Злоказов В.Б. Nucl. Instr. and Meth., 1975, 130, p.543.
3. Злоказов В.Б. Препринт ОИЯИ, Е10-10192, Дубна, 1976.

Рукопись поступила в издательский отдел
31 декабря 1976 года.