

УДК 681.3.01 : 621.384.8

Определение точных масс ионов в масс-спектре высокого разрешения. Е. В. Л а н и н — В кн.: Научное приборостроение. Теоретические и экспериментальные исследования. Л.: Наука, 1984, с. 27—33.

Рассматривается проблема высокоточного определения масс ионов в масс-спектре высокого разрешения путем экстраполяции и интерполяции закона развертки. Приводятся результаты исследования развертки масс-спектрометра МХ1310. Обосновывается выбор вида и степени аппроксимирующего полинома. Приводятся результаты обработки спектров и вычисления точных масс по прилагаемому алгоритму. Рассматриваются пути повышения точности в процедуре экстраполяции посредством прогнозирования ошибки вычисления, обусловленной случайными шумами, на основе модели авторегрессии. Параметры авторегрессии выражаются через значения автокорреляционной функции, вычисленной по значениям ряда ошибок. Лит. — 5 назв., ил. — 4, табл. — 2.

ОПРЕДЕЛЕНИЕ ТОЧНЫХ МАСС ИОНОВ В МАСС-СПЕКТРЕ ВЫСОКОГО РАЗРЕШЕНИЯ

Электронные устройства масс-спектрометров высокого разрешения обычно позволяют автономно определить массу с относительной погрешностью не больше $(1 \div 3) \cdot 10^{-6}$ лишь для небольшого числа ионов в спектре, так как измерения производятся для дублета на участке, где $\Delta M/M \leq 10\%$, причем эти измерения требуют значительной затраты времени.

Использование вычислительных средств открывает возможность высокоточного определения масс всех ионов в масс-спектре высокого разрешения путем экстраполяции и интерполяции закона развертки [1].

Методика измерения массовых чисел состоит в напуске в масс-спектрометр вместе с исследуемым веществом калибровочного соединения (фтороуглеродов), масс-спектр которого известен заранее. В начале эксперимента одновременно с запуском развертки масс-спектрометра включается подача сигналов от стабильного генератора импульсов на вход счетчика времени. Во время развертки масс-спектра ЭВМ решает задачу определения центров масс-спектральных пиков и привязывает к ним соответствующие временные значения. В результате обработки в реальном масштабе времени и предварительной информации о составе калибровочного вещества и последовательности масс в спектре имеем параметры M_i и t_i , позволяющие построить функциональную зависимость $M=f(t)$ и определить массы путем экстраполяции и интерполяции по спектру. При этом M можно представить в виде аддитивной смеси полезного сигнала и помехи, т. е.

$$M(\varphi, t) = m(\varphi, t) + S(t), \quad (1)$$

где φ — набор параметров, определяющих закон развертки.

Под помехой $S(t)$ понимаются, во-первых, все те причины, которые искажают развертку масс-спектра (нестабильность источников питания электромагнита, неоднородность магнитного поля, гистерезис, помехи от внешних магнитных полей) и во-вторых, те, которые связаны с погрешностью вычисления центров масс-спектральных пиков и обусловлены ионной статистикой, нестабильностью генератора импульсов, погрешностями счетчика ионов и вычислений [2].

Дополнительный вклад в $S(t)$ дают пропуски реперных пиков, т. е. нарушения шага между опорными пиками, и ложные реперные пики, обусловленные

наложением пика исследуемого вещества на реперный пик. В первом приближении можно предположить, что случайные колебания развертки представляют собой стационарный случайный процесс с выборочным нормальным законом распределения вероятностей и математическим ожиданием процесса, равным нулю. Ошибки измерения центров пиков также обусловлены влиянием большого числа независимых факторов и поэтому можно считать, что они подчинены нормальному закону распределения. Ложные пики и пропуски пиков случайны и независимы по всему спектру. Таким образом, зависимость $M = f(t)$ можно отнести к процессу со случайно изменяющимися во времени параметрами.

Представим $m(\varphi, t)$ в виде

$$m(\varphi, t) = \varphi_0 + \varphi_1(t) + \varphi_2 t^2 + \dots + \varphi_k t^k = \sum_{r=0}^k \varphi_r t^r. \quad (2)$$

Найдем оценку математического ожидания функции $m(\varphi, t)$ по результатам одной реализации на дискретном множестве t_1, t_2, \dots, t_n . Поскольку случайная функция $M(f, t)$ распределена нормально, то и $M(\varphi, t)$ подчинена нормальному закону распределения. В соответствии с методом максимального правдоподобия n -мерная функция правдоподобия, соответствующая нормальному закону, имеет вид [3]

$$L(m_1, m_2, \dots, m_n/\varphi) = \frac{1}{\sqrt{(2\pi)^n B_{nn}}} \times \\ \times \exp \left\{ -\frac{1}{2|B_{nn}|} \sum_{j=1}^n \sum_{i=1}^n K_{ij} [m_i - M(m_i)][m_j - M(m_j)] \right\}, \quad (3)$$

где (B_{nn}) — определитель корреляционной матрицы.

$$B_{nn} = \begin{vmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & R_{2n} \\ \dots & \dots & \dots & \dots \\ R_{n1} & R_{n2} & \dots & R_{nn} \end{vmatrix}, \quad (4)$$

где R_{ij} — корреляционные моменты; K_{ij} — адьюнкты элементов матрицы; $M(m_i)$ — математическое ожидание величины m_i .

Обозначим отклонение приближения $[m_j - M(m_j)]$ через w_j и с учетом выражения (2) запишем в виде

$$w_j = m_j - \sum_{r=0}^k \varphi_r t_j^r, \quad (j = 1, 2, \dots, n). \quad (5)$$

Как известно, для нахождения максимума функции $L(m_1, m_2, \dots, m_n/\varphi)$ необходимо прологарифмировать выражение (3); полученный результат продифференцировать по искомым параметрам, а конечное выражение подставить в (5) [3]. В результате система уравнений максимального правдоподобия будет иметь вид

$$\sum_{i=1}^n \sum_{j=1}^n \frac{K_{ij}}{|B_{nn}|} w_j \frac{\partial}{\partial \varphi_p} \sum_{l=0}^k \varphi_l t_i^l = 0. \quad (6)$$

Решение этой системы уравнений дает выражение для оптимальной оценки коэффициентов полинома. При $R_s(t, t+1) = 0$ и $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$ корреляционная матрица превращается в диагональную $\frac{K_{ij}}{|B_{nn}|} = \frac{1}{\sigma^2}$ при $i = j$.

Тогда система (6) принимает вид

$$\sum_{i=1}^n \left(m_j - \sum_{r=0}^k \varphi_r t_j^r \right) \frac{\partial}{\partial \varphi_p} \sum_{l=0}^k \varphi_l t_i^l = 0. \quad (7)$$

Для решения системы уравнений вида (7) необходимо задаться предположением о характере зависимости $M=f(t)$ с целью выбора аппроксимирующего полинома. Для приближенной оценки зависимости $M=f(t)$ на масс-спектрометре высокого разрешения МХ1310 был снят спектр перфторкеросина. Использовалась развертка вида $M=M_0 e^{-\alpha t}$, полученные данные приведены на рис. 1. Коэффициент пропорциональности A был найден путем аппроксимации зависимости $\ln M=f(t)$ линейной функции. Из построенных зависимостей видно, что для аппроксимации закона развертки целесообразно использовать полином, содержащий члены первого и второго порядка. Для полинома второй степени вида

$$M_k = M_i e^{a(t_k - t_i) + b(t_k - t_i)^2} \quad (8)$$

система уравнений запишется как

$$\left. \begin{aligned} \sum_{k=1}^n [\ln(M_k/M_i) - a(t_k - t_i) - b(t_k - t_i)^2](t_k - t_i) &= 0, \\ \sum_{k=1}^n [\ln(M_k/M_i) - a(t_k - t_i) - b(t_k - t_i)^2](t_k - t_i)^2 &= 0. \end{aligned} \right\} \quad (9)$$

Нахождение оптимальных оценок коэффициентов временного полинома необходимо связать с числом реперных точек, по которым строится интерполяционный полином, т. е. в окончательном виде задача вычисления точных масс может быть сформулирована следующим образом. Необходимо найти такие r и n , где n — число реперных пиков, по которым строится интерполяционный полином, аппроксимирующий зависимость $\ln M=f(t)$, а r — порядок полинома, при котором для любых подряд следующих реперных пиков имеет место

$$\frac{1}{n-r-1} \sum_{i=1}^{l+n} [\ln M_i - \varphi_r(t_i)]^2 \approx \sigma^2, \quad (10)$$

где σ^2 — требуемая среднеквадратичная погрешность определения массы; M_i — масса i -го реперного пика; t_i — время центра i -го пика; $\varphi_r(t_i)$ — значение интерполяционного полинома в точке i .

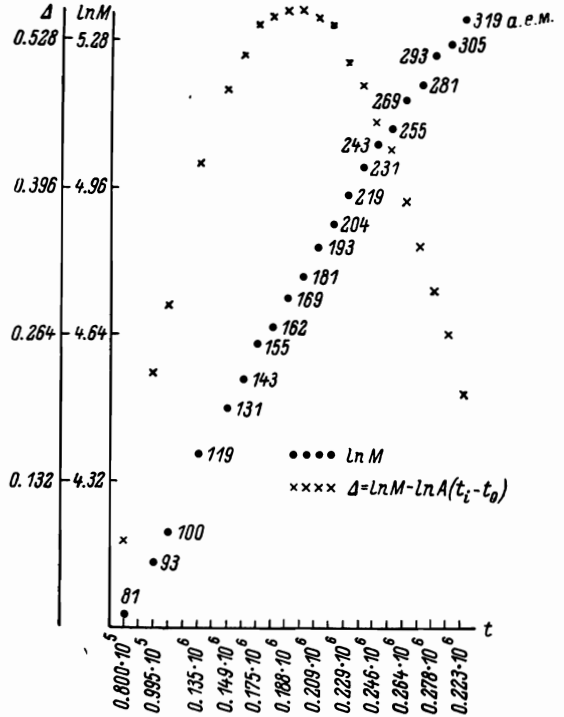


Рис. 1. Зависимость логарифма массы $\ln M$ от времени $t = \frac{1}{f} n_i$ (f — частота опроса, n — число импульсов).

Проверка выбора r и n производилась на ЭВМ. Задавался список точных масс и времен, полученный в результате регистрации спектра калибровочного соединения в цифровом виде. Проверялись процедуры экстраполяции и интерполяции.

В первом случае программа по первым n пикам $M_i, M_{i+1}, \dots, M_{i+n}$ вычисляла коэффициенты полинома, массу M_{i+n+1}^s , $\Delta = M_{(i+n+1)}^s - M_{(i+n+1)}^t$; $\delta M = \frac{\Delta}{M_{(i+n+1)}^t}$. Далее из списка опорных пиков исключался первый пик M_i , коэффициенты вычислялись по $M_{i+1}, M_{i+2}, \dots, M_{i+n+1}$ и проводилась экстраполяция в M_{i+n+2} пик и т. д. (Здесь M_i^s — экстраполированное, M^t — теоретическое значение массы i -го пика).

Исследовались полиномы вида (8) до 5-й степени включительно, число опорных точек менялось от 3 до 6 для процедуры экстраполяции, проводилась также оценка влияния пропуска репера на погрешность δM . В процедуре интерполяции исследовались те же полиномы и влияние длины участка на точность вычисления масс.

На основании сравнения δM для процедур экстраполяции и интерполяции были выбраны $n=5$, $r=2$. Причем пики репера следовали подряд, а вычисление точных масс на каждом шаге программы интерполяции производилось на центральном участке интервала. Для полинома вида (8) средняя относительная погрешность вычисления масс (интерполяция) равна примерно $4 \cdot 10^{-6}$.

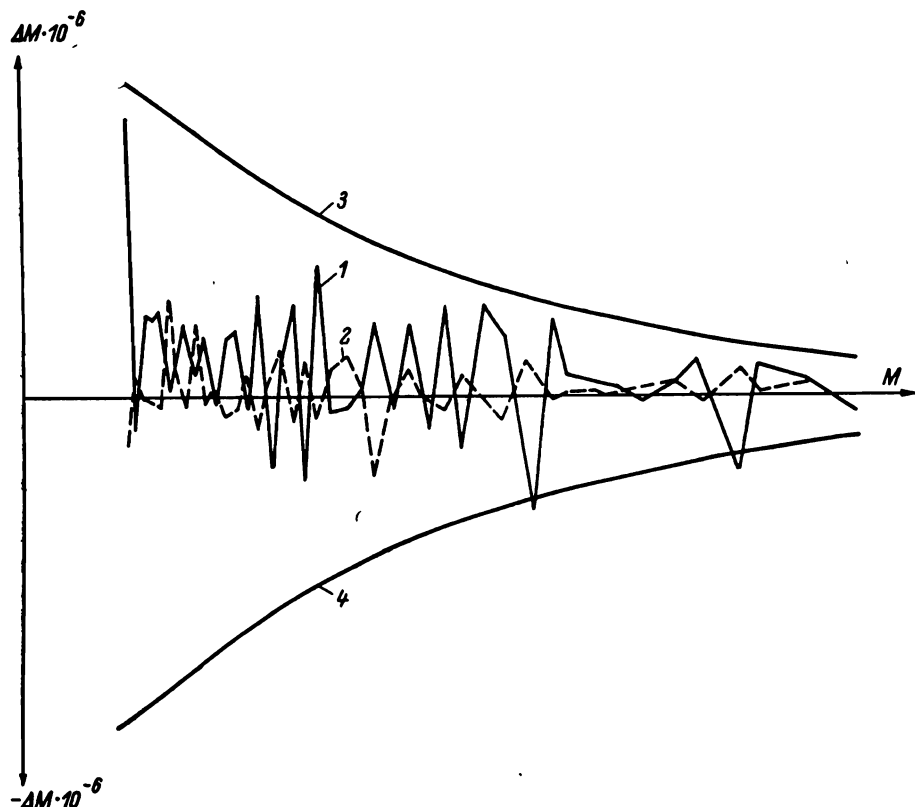


Рис. 2. Характеристика погрешности вычисления масс в процедуре экстраполяции (1) и интерполяции (2). Кривые 3 и 4 ограничивают область допустимой ошибки.

Результаты экстраполяции и интерполяции (рис. 2) показывают, что при $n=5$ и $r=2$ обеспечивается необходимая точность определения масс ионов. При снятии спектров смеси калибровочного и исследуемого вещества в систему уравнений вида (9), решаемую методом наименьших квадратов, подставляются кандидаты на очередной репер, лежащие в стробе «грубой» массы $\Delta M = \pm(0.5 \div 1)$ а. е. м. С этой целью в алгоритме используются показания от датчика измерения массовых чисел прибора. В результате отбирается кандидат, давший наименьшее среднеквадратичное отклонение при решении системы вида (9).

С целью повышения точности определения масс была исследована возможность прогнозирования ошибки вычисления при процедуре экстраполяции. Ошибку вычисления массы реперного пика (рис. 2, кривая 1) можно рассматривать в качестве временного ряда значений $\Delta M_i = x_i$, оценить который можно с помощью математического ожидания

$$x = \frac{\sum_{i=1}^N x_i}{N}; \quad i = 1, 2, \dots, N, \quad (11)$$

дисперсия D_1 . По значениям ряда x_i с использованием выражения (13) были рассчитаны значения корреляционной функции, представленной на рис. 4. Точные значения первых четырех K ($K_0 \div K_3$) использовались для расчета Δ , $\Delta_1 \div \Delta_3$

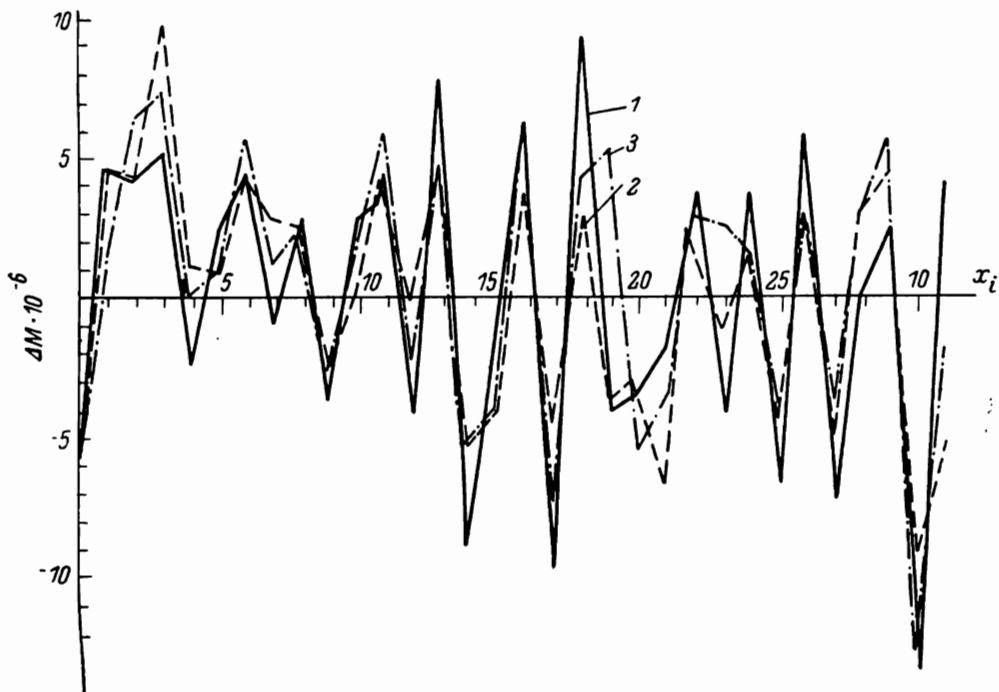


Рис. 3. Погрешность вычисления масс при процедуре экстраполяции (1), то же с учетом прогноза, т. е. процесса авторегрессии, (2) и с учетом прогноза $x'_{i+1} = x_{i+1} - K_x x_i$ (3).

системы (18) и параметров φ ($\varphi_1 \div \varphi_3$). Полученные значения введены в уравнение (20) и рассчитаны 29 значений Δx_{i+3} .

Результирующая ошибка в определении массы с учетом прогноза приведена на рис. 3 (кривая 2). Дисперсия в ряде ошибок с учетом прогноза в 2 раза ме-

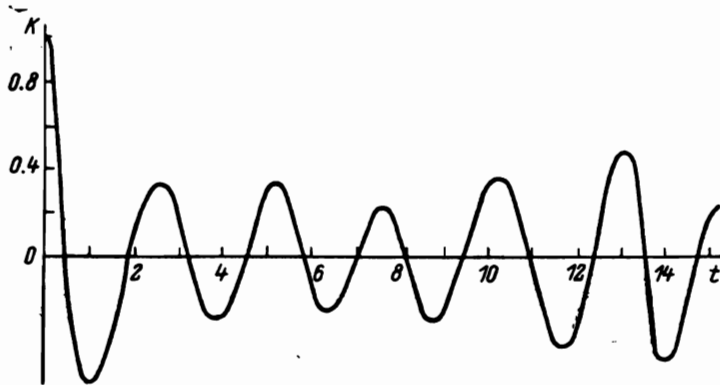


Рис. 4. Корреляционная функция K временного ряда ошибок вычисления массы (процедура экстраполяции).

ньше D_1 . В 21 случае из 29 прогноз дал уменьшение ошибки вычисления массы, в 2 случаях ошибки оказались равными, в 6 случаях (из них 3 относятся к первым значениям) ошибка увеличилась. Учет прогноза с использованием выражения $x'_{i+1} = x_{i+1} - K_x x_i$, где $K_x = K_1 = 0.5$, дает дисперсию полученного ряда в 1.4 раза меньше D_1 (рис. 3, кривая 3).

Таким образом, применение прогноза по временному ряду ошибок в определении M дает дополнительные возможности в уменьшении погрешности вычисления масс в процессе экстраполяции.

Полученные результаты положены в основу алгоритма обработки в СОИ к масс-спектрометру МХ1310. Результаты снятия спектров на комплексе МХ1310—СОИ в ряде институтов АН СССР показали, что обеспечивается определение масс ионов в диапазоне $12 \div 1000$ а. е. м. с относительной погрешностью не более $5 \cdot 10^{-6}$.

Достижение указанной точности вычисления масс ионов позволило производить автоматическое определение масс (табл. 1) и элементного состава для отдельных ионов в спектре (табл. 2). Спектр был снят на комплексе, установленном в Институте химии растительных веществ АН УзССР.

Т а б л и ц а 1
Точные массы ионов спектра $C_{32}H_{44}N_2O_8$

Состав иона	Масса		$\Delta M \cdot 10^{-5}$	$\left(\frac{\Delta M}{M}\right) \cdot 10^{-6}$
	теоретическая	вычисленная		
$C_8H_9O_3$	137.06025	137.06072	47	3
$C_8H_4NO_2$	146.02420	146.02462	42	3
$C_9H_7O_2N$	161.04767	161.04816	49	3
$C_{11}H_{16}N_2O_4$	178.12318	178.12392	74	4
$C_{16}H_{32}NO_3$	276.15996	276.16000	4	0.1
$C_{21}H_{31}NO_3$	345.23038	345.22899	139	4
$C_{22}H_{32}NO_4$	374.23312	374.23463	151	4
$C_{22}H_{23}NO_5$	390.22803	390.22786	17	1
$C_{23}H_{34}NO_5$	392.24368	392.24234	130	3
$C_{23}H_{35}NO_5$	405.25158	405.25258	100	3
$C_{30}H_{39}N_2O_6$	523.28076	523.28038	38	1
$C_{31}H_{41}N_2O_7$	553.29130	553.28980	150	3

Т а б л и ц а 2
Элементный состав ионов в спектре $C_{32}H_{44}N_2O_8$

M	C	H	N	O	^{13}C	Примечание
137	6	7	3	1	0	
	2	9	1	4	2	
	8	9	0	2	0	!
178	11	16	1	1	0	!
	1	17	4	5	1	
	3	19	1	6	1	
	8	14	3	0	2	
390	4	15	5	2	1	
	14	30	5	6	2	
	20	30	4	4	0	
	22	32	1	5	0	!
	17	28	6	3	2	
	23	28	5	1	0	

В графе «Примечание» восклицательным знаком отмечен наиболее вероятный состав.

ЛИТЕРАТУРА

1. *Biemann K., Bommerly P., Desiderio D., McMurray W. J.* — In: *Advances in Mass Spectrometry*. London, 1968, v. 4, p. 139.
2. *Raznikov V. V., Dodonov A. F., Lanin E. V.* — *Int. J. Mass Spectrom. Ion Phys.*, 1977, N 25, p. 295.
3. *Левин Б. Р.* Теоретические основы статистической радиотехники. М., 1969.
4. *Бокс Дж., Дженкинс Г.* Анализ временных рядов. М., 1974, т. 1.
5. *Хеннан Э.* Многомерные временные ряды. М., 1974.