

01

Сравнение различных методов разделения непрерывных перекрывающихся спектральных линий

© В.С. Сизиков, А.В. Лавров

Университет ИТМО,
197101 Санкт-Петербург, Россия

e-mail: sizikov2000@mail.ru, lavrov@corp.ifmo.ru

Поступила в редакцию 23.01.2018 г.

Сравниваются несколько методов разделения (сепарации) непрерывных перекрывающихся спектральных линий. В большинстве методов профиль каждой линии моделируется гауссианом или лоренцианом и обрабатывается суммарный измеренный спектр z . Количество N и параметры линий обычно оценивают методом производных, однако дифференцирование зашумленного спектра z связано с большими погрешностями. Для повышения точности дифференцирования предлагается использовать сглаживающие сплайны. В методе Фурье-самодеконволюции для разрешения перекрывающихся линий используется аподизация (искусственное усечение интерферограммы), позволяющая разрешить линии, однако за счет значительного уменьшения их ширины. В настоящей работе уменьшение ширины линий ради их разрешения не используется, а восстанавливаются истинные профили линий путем минимизации функционала невязки модифицированным методом координатного спуска с применением способа сужающихся ограничений, а также для сравнения методом Нелдера—Мида. В методе Манойлова параметры линий-пиков определяются по сверткам производных от спектра с отдельными пиками. В этом методе введено также понятие „степень наложения“. В настоящей работе вводится обобщенная степень наложения для случая разных амплитуд, ширины и расстояний соседних линий. Приведены численные иллюстрации.

DOI: 10.21883/OS.2018.06.46071.28-18

Введение

Одной из актуальных задач спектроскопии является задача о разделении (сепарации) непрерывных перекрывающихся спектральных линий [1–7]. Перекрывание линий может возникнуть из-за их близости или вследствие воздействия на излучающее вещество магнитного или электрического поля, в результате чего каждая линия расщепляется (эффект Зеемана или Штарка) на ряд близких пиков (компонент) и возникает сверхтонкая структура линии. Перекрывание спектральных линий может возникнуть также из-за взаимодействия молекул вещества, что ведет к движению молекул и к доплеровскому уширению линий и, как следствие, к их (частичному) наложению.

В этих случаях необходимо разделить перекрывающиеся линии в спектре некоторым способом. Особенностью данной задачи является то, что линии невозможно разделить технически, даже используя спектрометр с повышенной разрешающей способностью. Технически можно уточнить суммарный профиль линий, но сами линии по отдельности не разрешить (если не менять внешние условия — вещество, его температуру, электромагнитные поля и т.д.). Линии (их профили) можно разделить только математическим (и компьютерным) путем. Данная задача относится к так называемым обратным задачам второго типа (когда в принципе невозможно восстановить искомую функцию без математической обработки) [8,9]

Постановка задачи

Пусть в спектре присутствуют N непрерывных (протяженных) линий с профилями интенсивности $z_j(\lambda)$, $j = \overline{1, N}$, где j — номер линии, а λ — длина волны. Считаем, что линии, вообще говоря, налагаются друг на друга в той или иной степени. Полагаем, что измеряется суммарный спектр-профиль $z(\lambda)$, равный сумме N линий-компонент:

$$z(\lambda) = \sum_{j=1}^N z_j(\lambda) + \delta z, \quad a \leq \lambda \leq b, \quad (1)$$

где $[a, b]$ — пределы суммарного спектра, а δz — шум.¹

Требуется на основе измеренного суммарного, обычно зашумленного спектра $z(\lambda)$ определить количество N линий-компонент и их профили $z_j(\lambda)$. При этом допускается моделирование профилей линий некоторыми функциями (гауссианами, лоренцианами и т.д.) и использование дополнительной (априорной) информации о линиях (начальные приближения параметров линий, диапазоны изменения параметров и т.д.).

Моделирование линий и оценка их количества

Задача восстановления профилей линий-компонент $z_j(\lambda)$ по измеренному суммарному профилю $z(\lambda)$

¹ В настоящей работе не рассматривается сглаживающее влияние аппаратной функции спектрометра на суммарный спектр. Полагается, что такое влияние устранено [6].

(см. (1)) в общем случае не имеет однозначного решения. Поэтому обычно используют моделирование линий $z_j(\lambda)$ некоторыми функциями, адекватно представляющими реальные линии. Мы воспользуемся моделированием линий гауссианами и лоренцианами.

При этом для оценки количества линий N используем известный метод производных суммарного спектра $d^l z(\lambda)/d\lambda^l$, где обычно $l = \overline{1, 4}$, т.е. используются производные от 1-го до 4-го порядков [1,3,4,6,7].

Моделирование линий гауссианами

В работе [1] впервые предложен метод производных с использованием производной лишь 1-го порядка, а в работах [3,4,6,7] и др. использованы производные до 4-го порядка. Моделируем каждую линию $z_j(\lambda)$ гауссианом (гауссовой функцией):

$$z_j(\lambda) = A_j \exp\left(-\frac{(\lambda - \bar{\lambda}_j)^2}{2\sigma_j^2}\right), \quad j = \overline{1, N}, \quad (2)$$

где A_j — амплитуда линии, $\bar{\lambda}_j$ — координата максимума, σ_j^2 — дисперсия, при этом $\tau_j = \sqrt{2 \ln 2} \sigma_j = 1.1773 \sigma_j$ — полуширина гауссиана по уровню 0.5 [10,11].

Производные порядков $l = 1, 2, 3, 4$ от $z_j(\lambda)$ по λ равны (ср. [4,6,7])

$$z_j'(\lambda) = -\frac{A_j}{\sigma_j} \exp\left(-\frac{(\lambda - \bar{\lambda}_j)^2}{2\sigma_j^2}\right) \left(\frac{\lambda - \bar{\lambda}_j}{\sigma_j}\right), \quad (3)$$

$$z_j''(\lambda) = -\frac{A_j}{\sigma_j^2} \exp\left(-\frac{(\lambda - \bar{\lambda}_j)^2}{2\sigma_j^2}\right) \left(-\frac{(\lambda - \bar{\lambda}_j)^2}{\sigma_j^2} + 1\right), \quad (4)$$

$$z_j'''(\lambda) = \frac{A_j}{\sigma_j^3} \exp\left(-\frac{(\lambda - \bar{\lambda}_j)^2}{2\sigma_j^2}\right) \left(-\frac{(\lambda - \bar{\lambda}_j)^2}{\sigma_j^2} + 3\right) \times \left(\frac{\lambda - \bar{\lambda}_j}{\sigma_j}\right), \quad (5)$$

$$z_j^{IV}(\lambda) = \frac{A_j}{\sigma_j^4} \exp\left(-\frac{(\lambda - \bar{\lambda}_j)^2}{2\sigma_j^2}\right) \left(\frac{(\lambda - \bar{\lambda}_j)^4}{\sigma_j^4} - 6 \frac{(\lambda - \bar{\lambda}_j)^2}{\sigma_j^2} + 3\right). \quad (6)$$

Из формул (3)–(6) следует, что

$$\begin{aligned} z_j'(\bar{\lambda}_j) &= 0, & z_j''(\bar{\lambda}_j) &= -\frac{A_j}{\sigma_j^2} < 0, \\ z_j'''(\bar{\lambda}_j) &= 0, & z_j^{IV}(\bar{\lambda}_j) &= 3 \frac{A_j}{\sigma_j^4} > 0, \end{aligned} \quad (7)$$

т.е. отрицательный экстремум 2-й производной и положительный экстремум 4-й производной соответствуют максимуму j -й линии, а нулевое значение 1-й и 3-й производных (при условии, что $z_j''(\bar{\lambda}_j) < 0$, $z_j^{IV}(\bar{\lambda}_j) > 0$) также указывают на j -ю линию.

Из выражений для $z_j''(\bar{\lambda}_j)$ и $z_j^{IV}(\bar{\lambda}_j)$ в (7) можно определить A_j и σ_j из системы двух уравнений:

$$\left. \begin{aligned} \frac{A_j}{\sigma_j^2} &= -z_j''(\bar{\lambda}_j), \\ \frac{A_j}{\sigma_j^4} &= \frac{1}{3} z_j^{IV}(\bar{\lambda}_j), \end{aligned} \right\}$$

откуда

$$A_j = 3 \frac{[z_j''(\bar{\lambda}_j)]^2}{z_j^{IV}(\bar{\lambda}_j)}, \quad \sigma_j = \sqrt{3 \frac{-z_j''(\bar{\lambda}_j)}{z_j^{IV}(\bar{\lambda}_j)}}, \quad j = \overline{1, N}. \quad (8)$$

В работе [4] также использованы производные 2-го и 4-го порядков и изложена другая оригинальная методика, согласно которой вычисляются свертки 2-й и 4-й производных с отдельно взятым пиком (линией) типа (2). Это также дает возможность оценить A_j и σ_j .

Однако реальный (не модельный) измеренный спектр обычно зашумлен и его численное (не по аналитическим формулам (3)–(6)) дифференцирование выполняется со значительными погрешностями [7]. Можно отфильтровать шумы, например фильтром Савицкого–Голея с помощью m -функции `sgolayfilt.m` в системе MatLab [12]. Можно также аппроксимировать спектр сглаживающим сплайном [13,14] и дифференцировать сплайн, как это успешно сделано в работах [7,15,16].

При этом заметим, что на практике невозможно дифференцировать непосредственно линии $z_j(\lambda)$, а можно дифференцировать только суммарный спектр $z(\lambda)$ (1). В результате параметры отдельных линий будут определяться по производным суммарного спектра с погрешностями, а ряд линий может вообще не проявиться по некоторым производным согласно критериям (7). Поэтому нужно использовать все производные 1–4-го порядков, а также использовать более точные методы сепарации линий (модифицированный метод координатного спуска [7] и др.), полагая в качестве начальных приближений значения N , A_j , $\bar{\lambda}_j$ и σ_j , оцененные по методу производных.

Моделирование линий лоренцианами

Если каждая линия $z_j(\lambda)$, $j = \overline{1, N}$, — лоренциан (дисперсионный контур, лоренцева линия) [6,15,16] (ср. (2)):

$$z_j(\lambda) = A_j \frac{\tau_j^2}{(\lambda - \bar{\lambda}_j)^2 + \tau_j^2}, \quad (9)$$

где τ_j — полуширина лоренциана по уровню 0.5, то производные будут равны (ср. (3)–(6))

$$z_j'(\lambda) = -2 \frac{A_j}{\tau_j} \frac{(\lambda - \bar{\lambda}_j) \tau_j^3}{[(\lambda - \bar{\lambda}_j)^2 + \tau_j^2]^2}, \quad (10)$$

$$z_j''(\lambda) = -2 \frac{A_j}{\tau_j^2} \frac{[\tau_j^2 - 3(\lambda - \bar{\lambda}_j)^2] \tau_j^4}{[(\lambda - \bar{\lambda}_j)^2 + \tau_j^2]^3}, \quad (11)$$

$$z_j'''(\lambda) = 2 \frac{A_j}{\tau_j^3} \frac{12(\lambda - \bar{\lambda}_j)[\tau_j^2 - (\lambda - \bar{\lambda}_j)^2]\tau_j^5}{[(\lambda - \bar{\lambda}_j)^2 + \tau_j^2]^4}, \quad (12)$$

$$z_j^{IV}(\lambda) = 2 \frac{A_j}{\tau_j^4} \frac{12[\tau_j^4 - 10\tau_j^2(\lambda - \bar{\lambda}_j)^2 + 5(\lambda - \bar{\lambda}_j)^4]\tau_j^6}{[(\lambda - \bar{\lambda}_j)^2 + \tau_j^2]^5}. \quad (13)$$

Из формул (10)–(13) следует

$$\begin{aligned} z_j'(\bar{\lambda}_j) &= 0, & z_j''(\bar{\lambda}_j) &= -2 \frac{A_j}{\tau_j^2} < 0, \\ z_j'''(\bar{\lambda}_j) &= 0, & z_j^{IV}(\bar{\lambda}_j) &= 24 \frac{A_j}{\tau_j^4} > 0. \end{aligned} \quad (14)$$

Можно найти A_j и τ_j из системы двух уравнений

$$\left. \begin{aligned} \frac{A_j}{\tau_j^2} &= -\frac{1}{2} z_j''(\bar{\lambda}_j), \\ \frac{A_j}{\tau_j^4} &= \frac{1}{24} z_j^{IV}(\bar{\lambda}_j), \end{aligned} \right\}$$

откуда (ср. (8))

$$A_j = 6 \frac{[z_j''(\bar{\lambda}_j)]^2}{z_j^{IV}(\bar{\lambda}_j)}, \quad \tau_j = \sqrt{12 \frac{-z_j''(\bar{\lambda}_j)}{z_j^{IV}(\bar{\lambda}_j)}}, \quad j = \overline{1, N}. \quad (15)$$

Сепарация спектральных линий

Следующий этап — выполнение собственно сепарации (разделения, разрешения) линий-компонент, входящих в суммарный измеренный спектр $z(\lambda)$ (см. (1)). Рассмотрим и сравним несколько методов сепарации.

Метод фурье-самодеконволюции

В методе фурье-самодеконволюции спектра (Fourier self-deconvolution) [2] для разрешения перекрывающихся линий используется аподизация (про аподизацию см. [17]) — искусственное усечение интерферограммы, по которой с помощью преобразования Фурье вычисляется спектр в фурье-спектрометрах [9]. За счет аподизации ширины линий искусственно уменьшаются (до 5 раз). В результате истинные профили линий-компонент искажаются ради их разрешения.

Далее в настоящей работе приводится пример из [2] с двумя перекрывающимися линиями, обработанный методом фурье-самодеконволюции и для сравнения методами координатного спуска [7] и Нелдера–Мида [18,19]. Показывается, что линии можно разрешить, не искажая их профили, например, методами координатного спуска и Нелдера–Мида.

Степень наложения линий

В работе [4] введен эффективный параметр, характеризующий ситуацию с наложением двух пиков, — степень наложения пиков (в наших обозначениях):

$$D = \frac{\tau}{\Delta}, \quad (16)$$

где τ — полуширина пиков-линий по уровню 0.5, а Δ — расстояние между пиками. Однако формула (16) применима для случая, когда полуширины обоих пиков одинаковы. А для случая, вообще говоря, разных полуширин предлагаем следующую формулу:

$$D = \frac{\bar{\tau}}{\Delta}, \quad (17)$$

где $\bar{\tau} = (\tau_1 + \tau_2)/2$ — средняя полуширина 1-го и 2-го пиков. Однако формула (17) не учитывает возможного различия амплитуд A пиков и случая, когда число пиков N больше двух. Предлагается следующая формула для случая, вообще говоря, разных амплитуд A_j , $j = \overline{1, N}$, пиков и $N \geq 2$:

$$D_{j,j+1} = \frac{\bar{\tau}_{j,j+1}}{\Delta_{j,j+1}} \left(1 + \frac{|A_j - A_{j+1}|}{A_j + A_{j+1}} \right), \quad j = \overline{1, N-1}, \quad (18)$$

где $\bar{\tau}_{j,j+1} = (\tau_j + \tau_{j+1})/2$ — средняя полуширина соседних (j и $j+1$) пиков, а $\Delta_{j,j+1} = \bar{\lambda}_{j+1} - \bar{\lambda}_j$ — расстояние между соседними пиками.

Сделаем некоторый анализ формул (16)–(18). Если амплитуды линий одинаковы: $A_j = A_{j+1}$, то формула (18) переходит в формулу $D_{j,j+1} = \bar{\tau}_{j,j+1}/\Delta_{j,j+1}$, а при $N = 2$ — в формулу (17). А если $A_{j+1} \ll A_j$, то $D_{j,j+1} = 2\bar{\tau}_{j,j+1}/\Delta_{j,j+1}$. Это говорит о том (и формула (18) это подтверждает), что чем шире пики, а также чем они ближе друг к другу и чем больше отличаются их амплитуды, тем больше степень наложения пиков D , а значит, тем сложнее разделить пики при решении обратной задачи (см. далее).

Будем использовать также следующий параметр:

$$\bar{D} = \frac{1}{N-1} \sum_{j=1}^{N-1} D_{j,j+1} \quad (19)$$

— среднее значение степени наложения всех N пиков-линий.

Следует отметить также метод генетических алгоритмов [19], метод „полуслепой“ деконволюции [5], способ моделирования (вычислительных экспериментов) [6] — довольно сложные методы разложения суммарного спектра на составляющие. В настоящей работе предлагается более простая методика, включающая дифференцирование суммарного спектра, использование сглаживающего сплайна и минимизацию функционала.

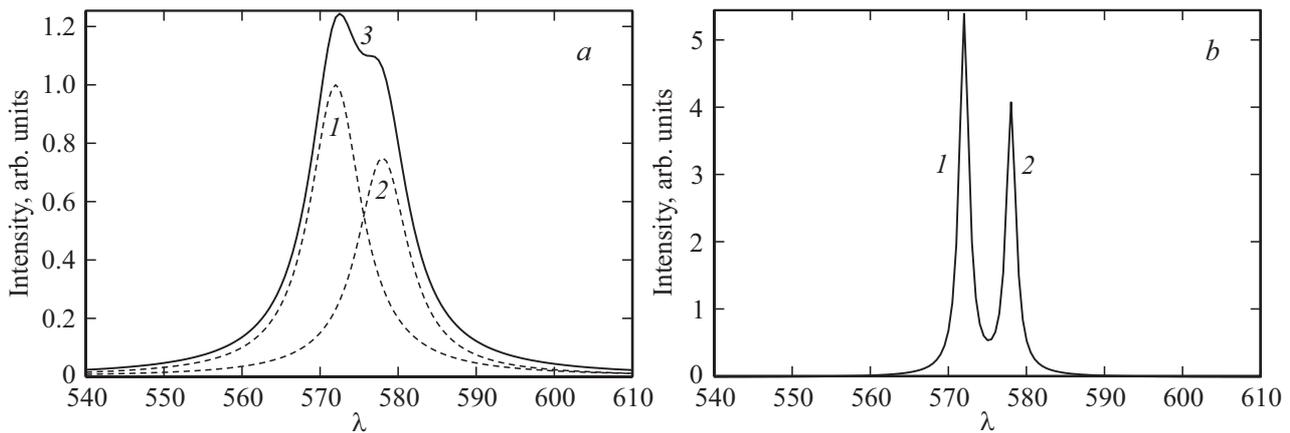


Рис. 1. (a) 1, 2 — исходные профили линий $z_1(\lambda)$ и $z_2(\lambda)$; 3 — суммарный спектр $z(\lambda) = z_1(\lambda) + z_2(\lambda)$; (b) суммарный спектр $z(\lambda)$, полученный в результате применения метода ФСД, включая аподизацию.

Численные иллюстрации

В работе [6] нами приведены результаты обработки трех близких линий-гауссиан со средней степенью наложения $\bar{D} \approx 0.52-0.72$. Суммарный спектр $z(\lambda)$ был практически незашумлен, поэтому вычисление 1-й и 2-й производных по $z(\lambda)$ не потребовало использования сплайна. Уточненные параметры трех гауссиан A_j , $\bar{\lambda}_j$ и σ_j были найдены путем минимизации функционала невязки модифицированным методом координатного спуска (КС) с ограничениями на параметры, а также (для сравнения) симплекс-методом Нелдера–Мида (НМ) с использованием начальных приближений для параметров. Среднеквадратическая погрешность восстановления линий методом КС получилась равной $\delta p \approx 0.011$, а методом НМ — $\delta p \approx 0.028$ [6].

С ростом степени наложения \bar{D} производные $z'(\lambda)$ и $z''(\lambda)$ выделяют не все три линии, но методы КС и НМ их восстанавливают.

В работе [7] приведены результаты обработки также трех линий-гауссиан с $\bar{D} \approx 0.55-0.77$, причем с зашумлением суммарного спектра $z(\lambda)$ 5-процентным шумом. Из-за шума 1-я и 2-я производные от $z(\lambda)$ дали очень неустойчивые результаты, а аппроксимация зашумленного спектра $z(\lambda)$ сглаживающим кубическим сплайном и последующее дифференцирование сплайна позволили оценить число линий N и их координаты $\bar{\lambda}_j$. Затем параметры линий A_j , $\bar{\lambda}_j$ и σ_j были найдены путем минимизации функционала методами КС (погрешность $\delta p \approx 0.0075-0.043$) и НМ ($\delta p \approx 0.045-0.072$).

Эти результаты продемонстрировали эффективность применения сплайн-сглаживания для устойчивого вычисления производных от суммарного спектра $z(\lambda)$ и эффективность определения параметров линий путем минимизации функционала.

В настоящей работе мы сравним результаты сепарации (разделения) линий-компонент методом фурье-самодеконволюции (ФСД) [2], алгоритмом свертки с

производными (АСП) [4] и методом минимизации функционала невязки (МФН) [6,7], развиваемым в настоящей работе.

Рассмотрим один из примеров, приведенных в работе [2]. Даны две перекрывающиеся линии-лоренциана вида (9). Амплитуды линий равны $A_1 = 1$, $A_2 = 0.75$, координаты максимумов $\bar{\lambda}_1 = 572$, $\bar{\lambda}_2 = 578$, полуширины линий $\tau_1 = \tau_2 = 8/2 = 4$, а расстояние между пиками $\Delta = \bar{\lambda}_2 - \bar{\lambda}_1 = 6$. Степень наложения пиков, согласно (19), равна $\bar{D} = 0.76$. Спектр задан численно с шагом дискретизации $h = 0.5$ в пределах по λ : левый предел $a = 540$, правый предел $b = 610$. Число узлов по λ равно $n = (b - a)/h + 1 = 141$ (рис. 1, a). Все величины в произвольных единицах, отличных от единиц в [2], где вместо λ использовано $1/\lambda$ — волновое число в см^{-1} , но графически пример как в [2].

Из рис. 1, b видно, что метод ФСД разделил линии, определил их координаты $\bar{\lambda}_1$ и $\bar{\lambda}_2$ (точнее, расстояние между ними $\Delta = \bar{\lambda}_2 - \bar{\lambda}_1$), но занижил (в 5.3 раза) их полуширины τ (с 4 до 0.75), т.е. исказил профили линий.

Используем для сравнения метод МФН. Добавим к суммарному спектру 1-процентный шум ($\delta z = 0.005$), меньше, чем в работе [7] (рис. 2, a).

Используем метод производных, а именно вычислим численно (с помощью m -функции diff.m системы MatLab) производные $z'(\lambda)$, $z''(\lambda)$, $z'''(\lambda)$ и $z^{IV}(\lambda)$ от незашумленного и зашумленного суммарного спектра $z(\lambda)$.

На рис. 2, b представлены производные $z'(\lambda)$, $z''(\lambda)$, $z'''(\lambda)$ и $z^{IV}(\lambda)$ от незашумленного суммарного спектра $z(\lambda)$. Видим, что кривые в общем соответствуют соотношениям (14). Кривая $z'(\lambda)$ имеет два нуля, кривая $z''(\lambda)$ имеет два отрицательных экстремума, кривая $z^{IV}(\lambda)$ имеет два положительных экстремума, причем нули и экстремумы дают значения $\bar{\lambda}_1 = 571.5$ и $\bar{\lambda}_2 = 577.5$, близкие точным значениям $\bar{\lambda}_1 = 572$ и $\bar{\lambda}_2 = 578$. Кривая $z'''(\lambda)$ имеет слишком много нулей (> 5), но надо учитывать лишь те нули, где функция $z'''(\lambda)$ возрастает,

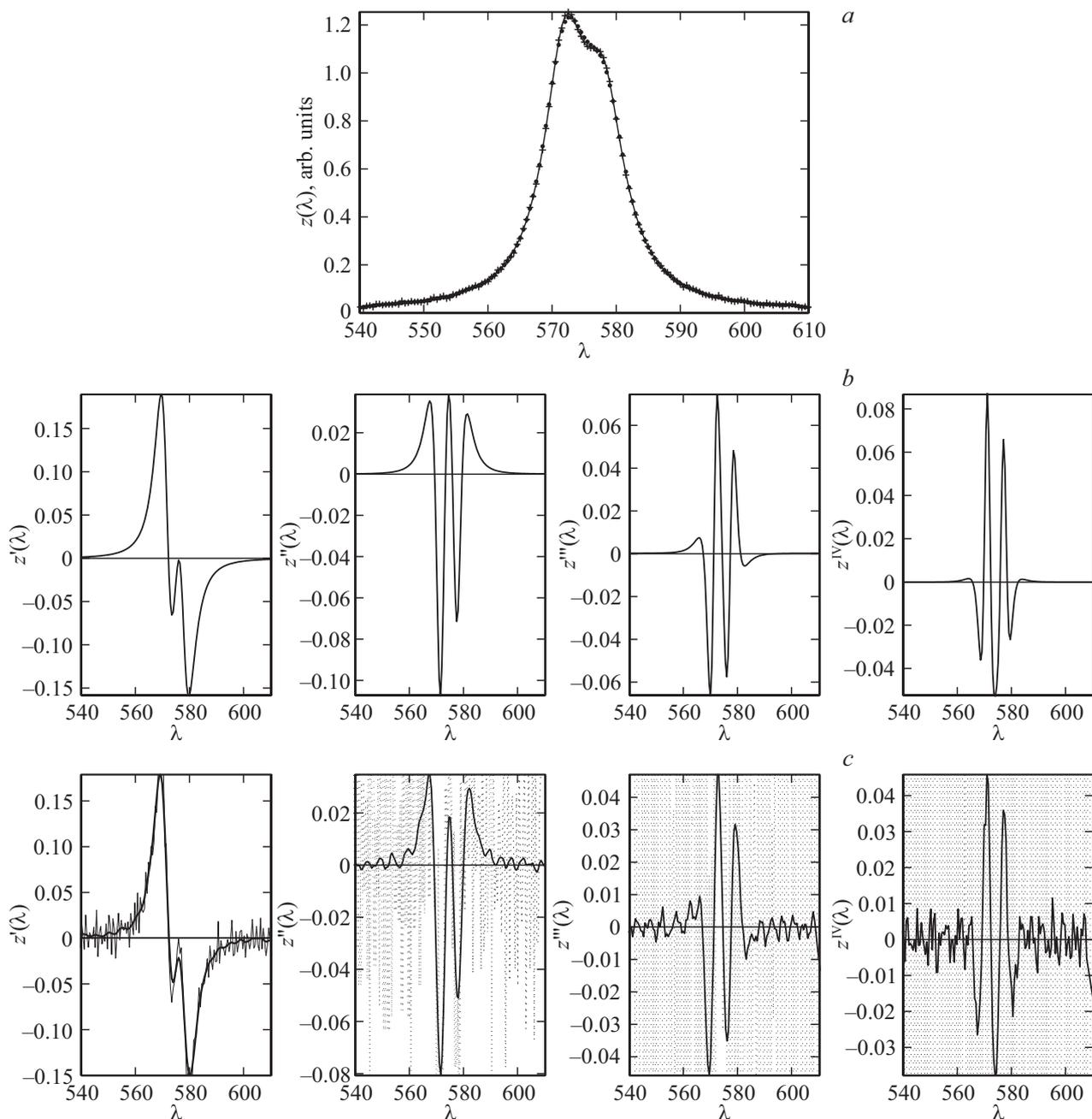


Рис. 2. (a) Суммарный спектр $z(\lambda)$ незашумленный (непрерывная линия), зашумленный (плюсы) и сглаживающий сплайн (точки); (b) производные $z'(\lambda)$, $z''(\lambda)$, $z'''(\lambda)$ и $z^{IV}(\lambda)$ от незашумленного суммарного спектра $z(\lambda)$; (c) производные от зашумленного $z(\lambda)$ (тонкие линии, „пилообразные“ решения) и от сплайна (непрерывные линии).

другими словами, $z'''(\bar{\lambda}) < 0$ и $z^{IV}(\bar{\lambda}) > 0$ (см. (14)). Таких нулей два.

По рис. 2, b можно в соответствии с формулами (15) оценить значения A_1 , τ_1 , A_2 и τ_2 . Из кривых $z''(\lambda)$ и $z^{IV}(\lambda)$ при $\bar{\lambda}_1 = 571.5$, $\bar{\lambda}_2 = 577.5$ получаем $z''(\bar{\lambda}_1) = -0.107$, $z^{IV}(\bar{\lambda}_1) = 0.0689$, $z''(\bar{\lambda}_2) = -0.0713$, $z^{IV}(\bar{\lambda}_2) = 0.0531$. Используя (15), вычисляем

$$A_1 = 1.0039, \quad \tau_1 = 4.3246, \quad A_2 = 0.5752, \quad \tau_2 = 4.0154. \quad (20)$$

Значения (20) заметно отличаются (особенно A_2) от точных значений

$$A_1 = 1, \quad \tau_1 = 4, \quad A_2 = 0.75, \quad \tau_2 = 4,$$

что объясняется (как было сказано выше) тем, что дифференцируется только суммарный спектр $z(\lambda)$, а непосредственно линии $z_j(\lambda)$ невозможно дифференцировать в силу их недоступности. Это порождает погрешности в определении параметров линий даже в отсутствие

шумов. Чтобы уточнить параметры линий, необходимо использовать дополнительно некоторый метод.

Уточнение параметров линий. Для уточнения параметров линий используем модификацию метода координатного спуска (КС) [6,7] (в работе [7] приведен псевдокод этого метода, однако в настоящей работе внесены дополнения в метод КС). Согласно методу КС, компоненты $z_j(\lambda)$, $j = \overline{1, N}$, суммарного спектра $z(\lambda)$ моделируем гауссианами (см. (2)) или лоренцианами (см. (9)), у каждой линии по три искомым параметра: амплитуда A_j , координата максимума $\bar{\lambda}_j$ и полуширина τ_j (или СКО σ_j).

Далее $3N$ параметров (где N — число линий-компонент) объединяем в один вектор

$$\mathbf{p} = [A_1, \bar{\lambda}_1, \tau_1, \dots, A_N, \bar{\lambda}_N, \tau_N] \quad (21)$$

длиной $3N$ (в нашем примере $N = 2$). Параметры p_J , $J = \overline{1, 3N}$, находим путем минимизации функционала невязки между результатами измерений и расчета:

$$F = \sum_{i=1}^n (\tilde{z}_i - z_i)^2, \quad (22)$$

где $\tilde{z}_i = \tilde{z}(\lambda_i)$ — измеренные значения суммарного спектра, $z_i = z_i(\mathbf{p})$ — рассчитанные значения суммарного спектра, n — число дискретных отсчетов λ , \mathbf{p} — единый вектор (21).

Вводятся также ограничения на параметры p_J в виде

$$p_{\min_j} \leq p_J \leq p_{\max_j}, \quad J = \overline{1, 3N}. \quad (23)$$

Ограничения вида (23) не позволяют выходить решению \mathbf{p} за пределы „коридора“, даваемого неравенствами (23), обеспечивая тем самым устойчивость и сходимость решения. При этом предполагается, что ограничения (23) вначале могут быть широкими, а затем могут сужаться — способ сужающихся ограничений.

Для сравнения предлагается выполнить уточнение параметров линий также путем минимизации функционала (22) симплекс-методом Нелдера–Мида (НМ) [18,19], реализованным в m -функции `fminsearch.m` [20]. Специфика метода НМ состоит в том, что он не использует ограничения, а использует начальные приближения параметров.

Погрешность определения параметров линий. Средне-квадратическая относительная погрешность определения параметров линий-компонент методами КС и НМ может быть вычислена по формуле (ср. [6,7]):

$$\delta p = \left[\frac{1}{3N} \sum_{J=1}^{3N} w_J (p_J - \bar{p}_J)^2 \right]^{1/2}, \quad (24)$$

где p_J — вычисленные значения параметров линий, а \bar{p}_J — точные значения параметров (известные лишь при обработке модельных спектров). Здесь w_J — вес, равные, согласно [6,7], $w_J = 1/p_{\text{mid } J}^2$, причем

$p_{\text{mid } J} = (p_{\min J} + p_{\max J})/2$. Однако такое выражение w_J годится лишь для метода КС. В настоящей работе предлагается использовать следующее выражение, пригодное как для метода КС, так и для метода НМ: $w_J = 1/\bar{p}_J^2$.

Результаты расчетов. Вначале рассмотрим случай отсутствия шумов (см. рис. 2, *b* и оценки (20)). Обозначим

$$\mathbf{p} = [1.0039, 571.5, 4.3246, 0.5752, 577.5, 4.0154] \quad (25)$$

вектор вида (21), полученный по производным (см. (20)).

Далее выполняем уточнение вектора \mathbf{p} путем минимизации функционала (22) методом КС с использованием сначала довольно широких ограничений вида (23) вокруг значений (20). При этом широкие ограничения \mathbf{p}_{\min} и \mathbf{p}_{\max} вида (23) задаем на основе вектора \mathbf{p} следующим образом (в кодах):

$$\text{width} = \mathbf{p}; \quad \text{width}(2) = \text{width}(5) = 10 * \mathbf{h}; \quad \mathbf{p}_{\min} = \mathbf{p} - \text{width};$$

$$\mathbf{p}_{\max} = \mathbf{p} + \text{width};$$

или

$$\left. \begin{aligned} p_{\min}(1) &= p_{\min}(3) = p_{\min}(4) = p_{\min}(6) = 0, \\ p_{\min}(2) &= p(2) - 10h, \quad p_{\min}(5) = p(5) - 10h, \\ p_{\max}(1) &= 2p(1), \quad p_{\max}(2) = p(2) + 10h, \\ p_{\max}(3) &= 2p(3), \quad p_{\max}(4) = 2p(4), \\ p_{\max}(5) &= p(5) + 10h, \quad p_{\max}(6) = 2p(6), \end{aligned} \right\} \quad (26)$$

т.е. (см. (25) и (26))

$$\left. \begin{aligned} \mathbf{p}_{\min} &= [0, 566.5, 0, 0, 572.5, 0], \\ \mathbf{p}_{\max} &= [2.0078, 576.5, 8.6492, 1.1504, 582.5, 8.0308] \end{aligned} \right\} \quad (27)$$

и получаем методом КС уточненный вектор:

$$\mathbf{p} = [1.0637, 572.37, 4.4238, 0.6374, 578.36, 3.6686], \quad (28)$$

$F = 0.0119$, $\delta p = 0.0863$, $N_f = 1715$ — общее число вызовов функционала.

Видим, что широкие ограничения (27) породили параметры линий \mathbf{p} (28), заметно отличные от точных значений (погрешность 8.6%):

$$\bar{\mathbf{p}} = [1, 572, 4, 0.75, 578, 4]. \quad (29)$$

Для дальнейшего уточнения вектора \mathbf{p} применим способ сужающихся ограничений. Уменьшим ширину ограничений (27) в 1.1 раза:

$$\text{width} = \text{width}/1.1; \quad \mathbf{p}_{\min} = \mathbf{p} - \text{width}; \quad \mathbf{p}_{\max} = \mathbf{p} + \text{width}$$

(где \mathbf{p} — текущий вектор), после чего вновь выполним минимизацию функционала (22) методом КС. Повторим эти действия в цикле по $i = 1, \dots, k$, получаем

$$\mathbf{p} = [1.0003, 572.00, 4.0023, 0.7495, 578.00, 3.9986], \quad (30)$$

$F = 3.11 \cdot 10^{-7}$, $\delta p = 0.00041$, $N_f = 23370$, $k = 29$. Цикл по i выполняем, пока

$$F_{i+1} < F_i/1.001, \quad (31)$$

где F_i — значение функционала F в i -й итерации, т.е. пока новое значение функционала меньше предыдущего хотя бы на 0.1%.

Метод КС со способом сужающихся ограничений дает параметры линий \mathbf{p} , согласно (30), почти совпадающие с точными значениями (29) (погрешность 0.04%), несмотря на то что используемые исходные значения параметров (20) (или (25)) были весьма далеки от истинных значений.

Выполняем минимизацию функционала (22) также методом НМ, используя в качестве начального приближения вектора \mathbf{p} (21) значения (20) (или (25)). Получаем методом НМ (без шумов):

$$\mathbf{p} = [1.0000, 572.00, 4.0000, 0.7500, 578.00, 4.0000], \quad (32)$$

$F = 4.13 \cdot 10^{-14}$, $\delta p = 1.06 \cdot 10^{-7}$, $k = 362$. Метод НМ от начальных приближений (25) привел к параметрам (32) — практически точным значениям (29).

Чтобы из полученных решений сделать выбор, помимо критерия (31) предлагается следующий.

Критерий выбора варианта решения:

— если обрабатывается модельный спектр с известными параметрами линий $\bar{\mathbf{p}}$, то отдаем предпочтение варианту с минимальной погрешностью δp (согласно (24));

— если обрабатывается реальный спектр с неизвестным $\bar{\mathbf{p}}$, то отдаем предпочтение варианту с минимальным значением функционала невязки F (согласно (22)).

Исходя из этого критерия в данном примере выбираем решения с минимальной погрешностью δp .

Теперь рассмотрим случай зашумления суммарного спектра $z(\lambda)$. Из рис. 2, *c* видно, что зашумление суммарного спектра даже умеренным 1-процентным шумом (рис. 2, *a*) ведет к большому погрешностям производных $z'(\lambda)$, $z''(\lambda)$, $z'''(\lambda)$ и $z^{IV}(\lambda)$ — к так называемым „пилообразным“ решениям (ср. [21,8,9]), по которым невозможно выделить ограниченное количество нулей производных $z'(\lambda)$ и $z'''(\lambda)$, а также экстремумов производных $z''(\lambda)$ и $z^{IV}(\lambda)$. Такое поведение производных связано с тем, что дифференцирование (тем более высокого порядка) зашумленных функций является некорректной (сильно неустойчивой) задачей [21].

Чтобы сделать вычисление производных устойчивым и умеренно гладким, используем аппроксимацию зашумленного суммарного спектра $z(\lambda)$ сглаживающим кубическим сплайном [13,14] с помощью m -функции `csaps.m` (с параметром сглаживания $P = 0.5$) (ср. [15,16]). После сглаживания были вычислены с помощью m -функции `diff.m` производные $z'(\lambda)$, $z''(\lambda)$, $z'''(\lambda)$ и $z^{IV}(\lambda)$ (рис. 2, *c*). Сравнение рис. 2, *b* и 2, *c* показывает, что сплайн-производные дают значения $\bar{\lambda}_1 = 571.5$ и $\bar{\lambda}_2 = 578$,

почти такие же, как и незашумленные производные (рис. 2, *b*), но значения производных $z''(\bar{\lambda}_{1,2})$ и $z^{IV}(\bar{\lambda}_{1,2})$ несколько различные.

Со сплайн-кривых $z''(\lambda)$ и $z^{IV}(\lambda)$ на рис. 2, *c* снимаем при $\bar{\lambda}_1 = 571.5$, $\bar{\lambda}_2 = 578$ значения: $z''(\bar{\lambda}_1) = -0.0809$, $z^{IV}(\bar{\lambda}_1) = 0.04336$, $z''(\bar{\lambda}_2) = -0.05114$, $z^{IV}(\bar{\lambda}_2) = 0.02740$. Используя (15), вычисляем (ср. (20))

$$A_1 = 0.9049, \quad \tau_1 = 4.7307, \quad A_2 = 0.5728, \quad \tau_2 = 4.7329. \quad (33)$$

Параметры (33) заметно отличаются от точных значений и от „бесшумных“ значений (20). Поэтому требуется их уточнение.

Выполняем минимизацию функционала (22) методом КС с использованием „широких“ ограничений вокруг значений (33) по правилу (26):

$$\left. \begin{aligned} \mathbf{p}_{\min} &= [0, 566.5, 0, 0, 573, 0], \\ \mathbf{p}_{\max} &= [1.8099, 576.5, 9.4615, 1.1457, 583, 9.4659]. \end{aligned} \right\} \quad (34)$$

Получаем на основе (34)

$$\mathbf{p} = [1.0193, 572.29, 4.4051, 0.6501, 578.19, 3.9534], \quad (35)$$

$F = 0.0171$, $\delta p = 0.0689$, $N_f = 1422$. Далее используем способ сужающихся ограничений, в результате получаем

$$\mathbf{p} = [1.0129, 572.04, 4.0332, 0.740, 578.06, 3.9353], \quad (36)$$

$F = 0.00263$, $\delta p = 0.0106$, $N_f = 14400$, $k = 17$. Параметры линий (36) получились более точными (в $0.0689/0.0106 = 6.5$ раз), чем (35), и близкими к точным значениям (29). Это говорит об эффективности способа сужающихся ограничений.

Выполняем минимизацию функционала (22) также методом НМ, используя в качестве начального приближения вектора \mathbf{p} (21) значения (33), а именно

$$\mathbf{p} = [0.9049, 571.5, 4.7307, 0.5728, 578, 4.7329]. \quad (37)$$

Получаем методом НМ (с шумами)

$$\mathbf{p} = [1.0115, 572.0382, 4.0262, 0.7417, 578.0555, 3.9427], \quad (38)$$

$F = 0.00260$, $\delta p = 0.0092$, число итераций $k = 414$. Начальное приближение (37) породило решение (38), дающее значения параметров линий, довольно близкие к точным значениям (29). При этом значения функционала F и погрешности параметров δp в методах КС и НМ получились практически одинаковыми.

На рис. 3 представлены линии-лоренцианы, восстановленные методом КС (рис. 3, *a*) и методом НМ (рис. 3, *b*) по зашумленным данным, но сглаженным сплайном.

Погрешность δp решения методом КС с сужающимися ограничениями получилась равной 1.06%, а методом НМ равной 0.92%, что сопоставимо. Но применение метода КС с многократным сужением ограничений потребовало 14 400 оценок целевой функции, а метод НМ использовал 414 оценок целевой функции.

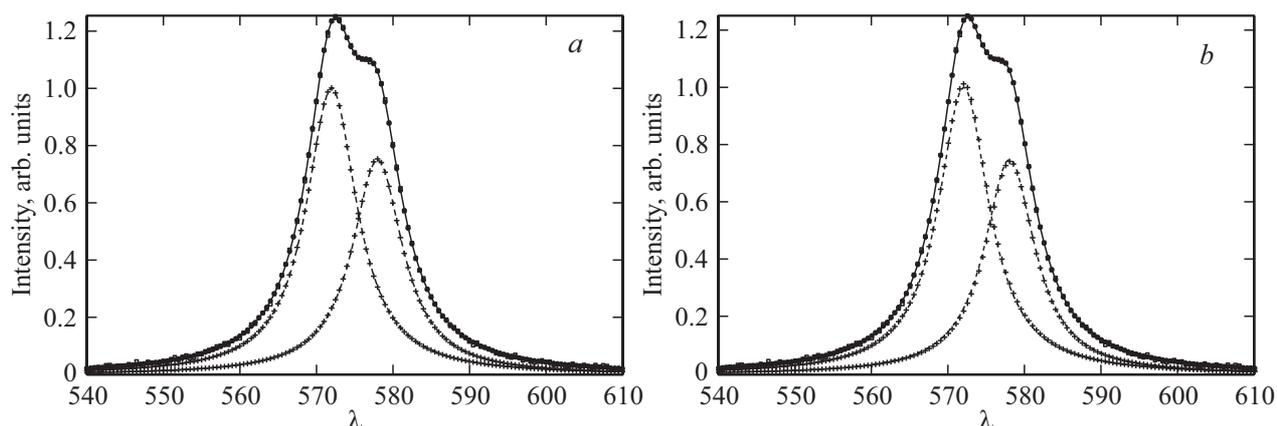


Рис. 3. (а) Восстановление линий-лоренцианов методом КС по данным $z(\lambda)$, зашумленным, но сглаженным сплайном, с использованием способа сужающихся ограничений: непрерывная линия — точный суммарный спектр $z(\lambda)$, две пунктирные линии — точные профили линий с параметрами (29), квадратики — зашумленный суммарный спектр $z(\lambda)$, плюсы — профили двух линий, рассчитанные методом КС; (б) восстановление линий методом НМ, обозначения такие же, как в (а).

Точность результатов метода НМ (и зависящее от этого время выполнения) можно регулировать, задавая порог сходимости целевой функции через параметр `optimset` `m`-функции `fminsearch.m` (в данной работе результаты получены при задании порога сходимости 10^{-6}). А точность результатов и время работы метода КС регулируется заданием порога сходимости для каждой из переменных минимизируемой функции (в данной работе использовался порог 10^{-4} для всех параметров вектора p), а также условиями сужения ограничений (здесь использовался коэффициент сужения 1.1 и пороговый коэффициент сходимости функционала 1.001, см. (31)).

Заключение

1. В работе сопоставлены некоторые методы разделения (сепарации) перекрывающихся спектральных линий, а именно метод производных суммарного спектра, метод фурье-самодеконволюции (ФСД), метод сверток производных с пиками, а также методы координатного спуска (КС) и Нелдера–Мида (НМ) минимизации функционала невязки.

2. Построена следующая цепочка действий в предлагаемой методике разделения линий. Вычисляются численно производные 1-, 2-, 3- и 4-го порядков от, вообще говоря, зашумленного суммарного спектра $z(\lambda)$ с использованием сглаживающего сплайна (рис. 2). По производным делается оценка количества линий N и их координат $\bar{\lambda}_j$. Выполняются оценки других параметров линий: амплитуд A_j и полуширин τ_j гауссианов согласно (8) или лоренцианов согласно (15) по производным 2-го и 4-го порядков согласно новым выведенным формулам. Далее эти оценки могут быть начальными приближениями при более точном определении параметров линий методами КС и НМ.

3. Предложены новые формулы (18), (19) для степени наложения пиков для случая нескольких пиков, имеющих, вообще говоря, разные ширины и амплитуды.

4. В методе ФСД за счет аподизации удастся разделить линии, но при этом ширины линий значительно уменьшаются. Предлагаемая в настоящей работе и в работах [4,6,7] и др. методика восстанавливает истинные профили без их заужения.

5. Рассмотрен численный пример разделения линий-лоренцианов методом ФСД и предлагаемой методикой.

Список литературы

- [1] Giese A.T., French C.S. // Appl. Spectrosc. 1955. V. 9. N 2. P. 78. doi 10.1366/000370255774634089
- [2] Kauppinen J.K., Moffatt D.J., Mantsch H.H., Cameron D.G. // Appl. Spectrosc. 1981. V. 35. N 3. P. 271. doi 10.1366/0003702814732634
- [3] Михайленко В.И., Михальчук В.В. // ЖПС. 1987. Т. 46. № 4. С. 535; Mikhailenko V.I., Mikhal'chuk V.V. // J. Appl. Spectrosc. 1987. V. 46. N 4. P. 327. doi 10.1007/BF00660037
- [4] Манойлов В.В., Заруцкий И.В. // Научное приборостроение. 2009. Т. 19. № 4. С. 103.
- [5] Yan L., Liu H., Zhong S., Fang H. // Appl. Spectrosc. 2012. V. 66. N 11. P. 1334. doi 10.1366/11-06256
- [6] Сизиков В.С., Лавров А.В. // Опт. и спектр. 2017. Т. 123. № 5. С. 678. doi 10.7868/S0030403417110216; Sizikov V.S., Lavrov A.V. // Opt. Spectrosc. 2017. V. 123. N 5. P. 682. doi 10.1134/S0030400X17110200
- [7] Сизиков В.С., Лавров А.В. // Научно-техн. вестник ИТМО. 2017. Т. 17. № 5. С. 879. doi 10.17586/2226-1494-2017-17-5-879-889; Sizikov V.S., Lavrov A.V. // Sci. Techn. J. ITMO. 2017. V. 17. N 5. P. 879. doi 10.17586/2226-1494-2017-17-5-879-889
- [8] Сизиков В.С. Обратные прикладные задачи и MatLab. СПб.: Лань, 2011. 256 с.
- [9] Сизиков В.С. Прямые и обратные задачи восстановления изображений, спектроскопии и томографии. СПб.: Лань, 2017. 412 с.

- [10] Сизиков В.С., Кривых А.В. // Опт. и спектр. 2014. Т. 117. № 6. С. 1040. doi 10.7868/S0030403414110166; Sizikov V.S., Krivykh A.V. // Opt. Spectrosc. 2014. V. 117. N 6. P. 1010. doi 10.1134/S0030400X14110162
- [11] Sizikov V., Sidorov D. // Appl. Spectrosc. 2017. V. 71. N 7. P. 1640. doi 10.1177/0003702817694181
- [12] Дьяконов В., Абраменкова И. MATLAB. Обработка сигналов и изображений. СПб.: Питер, 2002. 608 с.
- [13] Воскобойников Ю.Е., Преображенский Н.Г., Седельников А.И. Математическая обработка эксперимента в молекулярной газодинамике. Новосибирск: Наука, 1984. 240 с.
- [14] Сизиков В.С. Математические методы обработки результатов измерений. СПб.: Политехника, 2001. 240 с.
- [15] Сизиков В.С. // Научно-техн. вестник ИТМО. 2013. № 6(88). С. 1.
- [16] Sizikov V.S., Evseev V., Fateev A., Clausen S. // Appl. Opt. 2016. V. 55. N 1. P. 208. doi 10.1364/AO.55.000208
- [17] Bell R.J. Introductory Fourier Transform Spectroscopy. N.Y.–London: Academic Press, 1972. 382 p.
- [18] Nelder J.A., Mead R. // Computer J. 1965. V. 7. N 4. P. 308. doi 10.1093/comjnl/7.4.308
- [19] Kincaid D., Cheney W. Numerical Analysis: Mathematics of Scientific Computing. 3rd ed. Providence, RI: AMS, 2009. 788 p.
- [20] Дьяконов В. MATLAB 6 / Учебный курс. СПб.: Питер, 2001. 592 с.
- [21] Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. 3-е изд. М.: Наука, 1986. 288 с.