

Министерство науки и высшего образования Российской Федерации

---

САНКТ-ПЕТЕРБУРГСКИЙ  
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО

---

Физико-механический институт  
Высшая школа прикладной математики и вычислительной физики

*А. Н. Баженов*

МУЛЬТИИНТРЕВАЛЫ  
В ИНТЕРВАЛЬНОМ АНАЛИЗЕ.  
ИНТЕРВАЛЬНАЯ МОДА И МУЛЬТИМОДА

Учебное пособие

Санкт-Петербург  
2024

УДК 519.9 Р32

А в т о р ы :

А.Н. Баженов

Мультиинтервалы в интервальном анализе. Интервальная мода и мультимода. 2024. – 60 с.

Учебное пособие посвящено применению интервального анализа к анализу данных. Рассматривается ситуация, когда оценка измеряемой или расчётной величины представляет собой неоднозначное множество. В таком случае в интервальном анализе можно использовать специальный объект — мультиинтервал.

Рассмотрены примеры мультиинтервалов в науке и технике. Особое внимание уделено спектроскопии, где мультиинтервалы возникают естественным образом.

В пособии конкретизированы идеи обработки мультимодальных распределений, намеченные в книге А.Н. Баженов, С.И. Жилин, С.И. Кумков, С.П. Шарый «Обработка и анализ данных с интервальной неопределённостью». В дополнение к алгоритму вычисления интервальной моды разработан ряд алгоритмов для обработки зашумлённых данных сложной структуры.

Разработанный аппарат проиллюстрирован на наборе примеров из спектроскопии видимого излучения и гамма-спектроскопии.

Книга предназначена для студентов и аспирантов естественнонаучных специальностей, а также для учёных и инженеров, интересующихся вопросами математической обработки экспериментов.

# Оглавление

Введение	4
<b>1 Интервальный анализ</b>	<b>9</b>
1.1 Базовые понятия интервального анализа . . . . .	9
1.2 Мультиинтервалы. . . . .	12
<b>2 Примеры мультиинтервалов в науке и технике</b>	<b>14</b>
2.1 Спектроскопия электромагнитного излучения . . . . .	14
2.1.1 Спектр неба на Земле . . . . .	15
2.1.2 Элементо-определение состава смеси веществ . . . . .	17
2.1.3 Красное смещение и лес Лаймана . . . . .	23
2.2 Масс-спектрометрия . . . . .	28
<b>3 Мультиинтервалы в обработке данных</b>	<b>33</b>
3.1 Мода интервальной выборки . . . . .	33
3.2 Выборки унимодальные и мультимодальные . . . . .	37
3.2.1 Алгоритм вычисления мод мультимодального распределения . . . . .	37
<b>4 Примеры</b>	<b>43</b>
4.1 Модельная выборка — спектральный триплет . . . . .	43
4.2 Спектроскопия гамма-квантов . . . . .	47
4.3 Спектроскопия атомного и молекулярного излучения . . . . .	52
Литература	55
Предметный указатель	59

# Введение

Учебное пособие посвящено применению интервальной статистики, иначе — анализу данных с интервальной неопределённостью.

Понятия и методы интервальной статистики представлены в коллективной монографии А.Н. Баженова, С.И. Жилина, С.И. Кумкова и С.П. Шарого «Обработка и анализ данных с интервальной неопределённостью» [1]. Математической базой интервальной статистики служит современная ветвь математики — интервальный анализ. Наиболее основательное изложение этой дисциплины можно найти в фундаментальной монографии С.П. Шарого «Конечномерный интервальный анализ» [2]. Краткое введение в интервальный анализ содержится в [3].

План изложения следующий. В первой главе вводятся базовые определения и обозначения, соответствующие современной нотации интервального анализа. На основе базового объекта, интервала, вводится понятие мультиинтервала.

Во второй главе приведены примеры мультиинтервалов в науке и технике. Для каждого примера формируются мультиинтервальные объекты, осуждается их специфика в рассмотренном примере.

В третьей главе обсуждается применение мультиинтервалов в обработке данных. При анализе данных часто нужно дать оценку информационного множества задачи. Естественным способом является использование моды выборки, как наиболее её наиболее представительной характеристики. В связи с этим, сначала обсуждается понятие интервальной моды, рассмотренное в [1]. В многообразии практических задач обработки данных нередко встречаются ситуации, когда мода является набором непересекающихся интервалов. В книге [1] эта ситуация рассмотрена на идейном уровне. В настоящем пособии представлен набор алгоритмов, позволяющих обрабатывать данные с набором мод.

Алгоритмы иллюстрируются на конкретных примерах. В первую очередь рассмотрен простой модельный пример. Далее рассматриваются два примера, находящиеся на разных полюсах по характеру данных, из ядерной гамма-спектроскопии и спектроскопии видимой области электромагнитного излучения.

Вычисления подробно иллюстрированы.

# Список иллюстраций

1.1	Мультиинтервал в $\mathbb{R}$ . Рис. 1.11 из [2]. . . . .	12
2.1	Спектр голубого неба, Википедия . . . . .	15
2.2	Спектр неба, С.-Петербург, 2017-01-04 . . . . .	16
2.3	Спектр паров меди в дуге . . . . .	17
2.4	Спектр стандарта РУС-2 . . . . .	18
2.5	Участок спектра стандарта РУС-2 . . . . .	19
2.6	Участок спектра образца . . . . .	19
2.7	Участок спектров образца стандарта и стандарта . . . . .	20
2.8	Участок неполного покрытия стандартом спектра образца . . . . .	20
2.9	Покрытие стандартом спектра образца СГХМ-2 . . . . .	21
2.10	Перегрузка области спектра . . . . .	22
2.11	Типичный спектр квазара . . . . .	23
2.12	Спектральная серия Лаймана . . . . .	24
2.13	Варианты геометрии регистрации излучения [15] . . . . .	26
2.14	Примеры спектров для различных случаев геометрии (рис. 2.13) регистрации излучения квазара [15] . . . . .	27
2.15	Спектры для различных случаев красного смещения [15] . . . . .	27
2.16	Масс-спектры GaInAsP в масс-спектрометрии с вторичными ионами . . . . .	28
2.17	Масс-спектры GaInAsP, область масс Ga и In . . . . .	30
2.18	Масс-спектры GaInAsP, область масс GaIn . . . . .	30
2.19	Масс-спектры GaInAsP, область масс 39-40 а.е. . . . .	31
3.1	Диаграмма рассеяния интервальной выборки (3.1) . . . . .	35
3.2	Значения частот $\mu_i$ и интервальная мода mode $\mathbf{X}$ . . . . .	36
4.1	Шаг 0 — основная мода. . . . .	44

4.2	Шаг 1 — расширение основной моды . . . . .	45
4.3	Шаг 2 — вторая мода . . . . .	45
4.4	Шаг 3 — третья мода . . . . .	45
4.5	Общая картина — все моды . . . . .	46
4.6	Массив частот $\mu$ при регистрации гамма-квантов . . . . .	47
4.7	Гистограмма разностей $\Delta\mu$ . . . . .	48
4.8	Пример работы алгоритма Табл. 3.2 . . . . .	51
4.9	Тепловая карта интенсивностей для разряда токамака Глобус-М2 [25] . . . . .	52
4.10	Спектр излучения вещества плазмы токамака Глобус-М2	53
4.11	Главные моды спектра рис. 4.10 . . . . .	54

# Список таблиц

1.1	Интервальное умножение в полной интервальной арифметике . . . . .	11
2.1	Длины волн серии Лаймана . . . . .	25
2.2	Изотопный состав индия и галлия . . . . .	29
2.3	Изотопный состав пар атомов GaInAsP . . . . .	31
3.1	Алгоритм для нахождения моды интервальной выборки и её частоты [1] . . . . .	34
3.2	Алгоритм для нахождения частот мод мультимодальной интервальной выборки . . . . .	38
3.3	Алгоритм для расширения границ интервальной моды . . . . .	41
4.1	Алгоритм вычисления значения параметра <i>noise</i> алгоритма Табл. 3.3 . . . . .	49



# Глава 1

## Интервальный анализ

### 1.1 Базовые понятия интервального анализа

Забегая вперёд, введём понятие интервала. *Интервалом* вещественной оси  $[a, b]$  называется множество всех чисел, расположенных между заданными числами  $a$  и  $b$  включая их самих, т.е.

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$

При этом  $a$  и  $b$  называются концами интервала.

Интервальную величину принято обозначать жирным шрифтом, например,  $\mathbf{x}$ . Левую границу интервала подчёркивают снизу, а правую — сверху. Границы берутся в квадратные скобки, что передаёт идею интервала как отрезка вещественной оси.

$$\mathbf{x} = [\underline{\mathbf{x}}, \overline{\mathbf{x}}].$$

Важнейшими характеристиками интервала являются его *середина* (центр)

$$\text{mid } \mathbf{a} = \frac{1}{2}(\overline{\mathbf{a}} + \underline{\mathbf{a}}),$$

и его *радиус*

$$\text{rad } \mathbf{a} = \frac{1}{2}(\overline{\mathbf{a}} - \underline{\mathbf{a}}).$$

Радиус интервала является мерой абсолютного рассеяния точек интервала. При описании относительной погрешности в интервальном анализе приходится использовать разные меры.

Полезной характеристикой интервала является так называемый функционал Рачека  $\chi$ :

$$\chi(\mathbf{a}) = \begin{cases} \underline{\mathbf{a}}/\overline{\mathbf{a}}, & \text{если } \underline{\mathbf{a}} \leq \overline{\mathbf{a}}, \\ \overline{\mathbf{a}}/\underline{\mathbf{a}}, & \text{иначе.} \end{cases}$$

Он характеризует «относительную узость» интервала.

Множество всех интервалов из  $\mathbb{R}$  обозначается символом  $\mathbb{IR}$ . Используемая система обозначений следует неформальному международному стандарту на обозначения в интервальном анализе, выработанному в 2002-2010 годах [4].

Неформально можно сказать так: интервалы предназначены для величин, для которых существуют двусторонние ограничения. Можно также говорить об интервальных оценках.

Важной особенностью интервальной арифметики является учёт неопределённости выполнения арифметических операций. В частности, при последовательном выполнении сложения и вычитания получается не точно 0, а величина, содержащая 0:

$$[1, 2] - [1, 2] = [-1, 1] \ni 0.$$

Таким образом, производится двусторонняя оценка величины результата прямого и обратного действия. Многократное повторение этой операции приводит к увеличению границ результата

$$\sum_{i=1}^n ([1, 2] - [1, 2]) = n \cdot [-1, 1].$$

То есть, имеет место эффект нарастания «снежного кома», или «обёртывания». Такое свойство классической интервальной арифметики отражает факт «внешнего» оценивания множества решений задачи.

Помимо наиболее естественного понимания интервала как отрезка вещественной оси, существуют и более сложные конструкции. В частности, очень важна полная интервальная арифметика или арифметика Каухера. Она обобщает обычную интервальную арифметику на случай, когда у интервала есть «направление». Именно, в этом случае концы интервала не обязательно упорядочены от меньшего к большему. Такое свойство даёт дополнительные возможности, которые мы обсудим далее. Обозначается такая арифметика символом  $\mathbb{KR}$ .

Символически, можно представить соотношение арифметик следующим образом

$$\mathbb{R} \subseteq \mathbb{IR} \subseteq \mathbb{KR}.$$

Если концы интервалов совпадают, имеем обычную вещественную арифметику.

В арифметике Каухера содержательный смысл имеет операция замены порядка следования концов интервала, при которой получается интервал, дуальный исходному:

$$\text{dual } \mathbf{a} := [\bar{\mathbf{a}}, \underline{\mathbf{a}}].$$

В частности, дуализация даёт возможность получать при интервальных операциях точечные значения, или *внутреннюю оценку*:

$$[1, 2] \ominus [1, 2] = 0.$$

Символ  $\ominus$  соответствует так называемому *алгебраическому вычитанию*.

Аналогично для деления имеем внешние и внутренние оценки:

$$[1, 2]/[1, 2] = [0.5, 4], \quad [1, 2] \oslash [1, 2] = 1$$

Таким образом, в арифметике Каухера имеются гибкие арифметические возможности оценок: от самых строгих, рассчитанных на наилучший вариант, до локализирующих результат. Последнее обстоятельство особенно важно при многократных проведениях однотипных операций и построения итерационных алгоритмов.

Приведём правила умножения в  $\mathbb{KR}$  в виде так называемой таблицы Кэли.

	$\mathbf{b} \in P$	$\mathbf{b} \in Z$	$\mathbf{b} \in -P$	$\mathbf{b} \in \text{dual } Z$
$\mathbf{a} \in P$	$[\underline{ab}, \bar{ab}]$	$[\bar{ab}, \underline{ab}]$	$[\bar{ab}, ab]$	$[\underline{ab}, \bar{ab}]$
$\mathbf{a} \in Z$	$[\underline{ab}, \bar{ab}]$	$[\min\{\underline{ab}, \bar{ab}\}, \max\{\underline{ab}, \bar{ab}\}]$	$[\bar{ab}, ab]$	0
$\mathbf{a} \in -P$	$[\underline{ab}, \bar{ab}]$	$[\bar{ab}, \underline{ab}]$	$[\bar{ab}, ab]$	$[\bar{ab}, \underline{ab}]$
$\mathbf{a} \in \text{dual } Z$	$[\underline{ab}, \bar{ab}]$	0	$[\bar{ab}, ab]$	$[\max\{\underline{ab}, \bar{ab}\}, \min\{\underline{ab}, \bar{ab}\}]$

Таблица 1.1. Интервальное умножение в полной интервальной арифметике

Замечательно, что данные в таблице 1.1 правила верны и для классической интервальной арифметики. Её область ограничена квадратом

$3 \times 3$  и не включает неправильных интервалов самой нижней строки и самого правого столбца.

Помимо уже упомянутого, в полной интервальной арифметике всегда имеет содержательный смысл обобщение операции пересечения интервалов, взятие *минимума по включению*, обозначаемому как  $\wedge$ .

Продemonстрируем этот факт на примере. Найдем пересечение двух пересекающихся интервалов  $[1, 3]$  и  $[2, 4]$ :

$$[1, 3] \cap [2, 4] = \{\max \min\{1, 2\}, \min \max\{3, 4\}\} = [2, 3].$$

Поступим аналогичным образом с непересекающимися интервалами  $[1, 2]$  и  $[3, 4]$ , взяв минимум по включению:

$$[1, 2] \wedge [3, 4] = \{\max \min\{1, 3\}, \min \max\{2, 4\}\} = [3, 2].$$

В классической интервальной арифметике этот результат не имеет смысла, а полной имеет: это минимальный интервал, имеющий общие элементы с исходными.

Такая возможность даёт гибкость при неизбежной в экспериментальной практике работе с несовместными данными. *Минимаксный подход*, свойственный полной интервальной арифметике, также обеспечивает и другие важные свойства, см. [2].

## 1.2 Мультиинтервалы.

### Неодносвязные интервальные величины (мультиинтервалы).

В ряде разделов науки и техники имеют место ситуации, когда исследуемая величина содержится в неодносвязной области.

Согласно [2], мультиинтервал — это объединение конечного числа несвязных интервалов числовой оси (Рис. 1.1).



Рис. 1.1. Мультиинтервал в  $\mathbb{R}$ . Рис. 1.11 из [2].

Между мультиинтервалами также могут быть определены арифметические операции «по представителям», аналогично тому, как это делается на множестве интервалов [5]. Мультиинтервальная арифметика

применяется редко ввиду серьёзных ограничений, которые возникают при алгебраических операциях с мультиинтервальными величинами и вычислительных сложностей. Тем не менее, сама по себе идея мультиинтервалов содержательна и полностью отменить её не стоит.

В естественных науках возникновение неодносвязных областей часто связано с наличием периодичности в уравнениях или граничных условиях. Спектр таких явлений достаточно широк. Некоторые примеры приводятся в [6].

## Глава 2

# Примеры мультиинтервалов в науке и технике

В главе §2 рассмотрены примеры возникновения мультиинтервалов в науке и технике. В каждом из разделов главы приводятся сведения о физике явлений. Далее выделяются величины, представляющие группы с неоднозначными параметрами. Работа с полученными объектами рассматривается в главе §3 «Мультиинтервалы в обработке данных».

### 2.1 Спектроскопия электромагнитного излучения

Спектральный анализ — совокупность методов качественного и количественного определения состава объекта, основанная на изучении спектров взаимодействия материи с излучением, включая спектры электромагнитного излучения, акустических волн, распределения по массам и энергиям элементарных частиц и др.

В настоящей главе мы рассматриваем спектры электромагнитного излучения. В зависимости от целей анализа и типов спектров выделяют несколько методов спектрального анализа. Атомный и молекулярный спектральные анализы [10] позволяют определять элементарный и мо-

лекулярный состав вещества. Соответственно, в эмиссионном (§2.1.2) и абсорбционном (§2.1.1, §2.1.3) методах состав определяется по спектрам испускания и поглощения.

Приводимые примеры возникновения мультиинтервальных объектов в данных совершенно не являются исключительными. Скорее наоборот — они составляют совершенно незначительное подмножество различных вариантов возникновения мультиинтервальных объектов в спектроскопии.

## 2.1.1 Спектр неба на Земле

Тёмные линии на спектральных полосках были замечены давно (1802 г. — Волластон), но первое серьёзное исследование этих линий было предпринято только в 1814 году Фраунгофером. В его честь эффект получил название «Фраунгоферовы линии». Фраунгофер установил стабильность положения линий, составил таблицу 574 линии. Он присвоил каждой линии буквенно-цифровой код. Не менее важным стало его заключение, что линии не связаны ни с оптическим материалом, ни с земной атмосферой, но являются природной характеристикой солнечного света. Аналогичные линии он обнаружил у искусственных источников света, а также в спектрах Венеры и Сириуса.

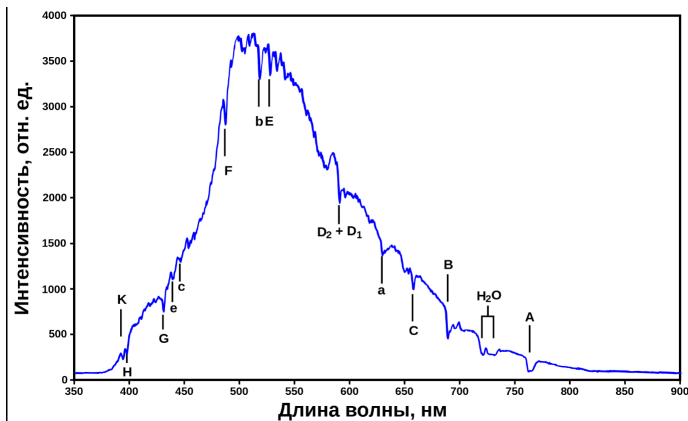


Рис. 2.1. Спектр голубого неба, Википедия

На рисунке 2.1 показан спектр голубого неба. Для некоторых линий поглощения приведена легенда, объясняющая их происхождение.

Линии  $D_1$  и  $D_2$  — широко известный «натриевый дублет», пара хорошо различимых солнечных линий.

Вскоре выяснилось, что одна из самых отчётливых линий всегда появляется в присутствии натрия. В 1859 году Г. Кирхгоф и Р. Бунзен после серии экспериментов заключили: каждый химический элемент имеет свой уникальный линейчатый спектр, и по спектру небесных светил можно сделать выводы о составе их вещества. С этого момента в науке появился спектральный анализ, мощный метод дистанционного определения химического состава.

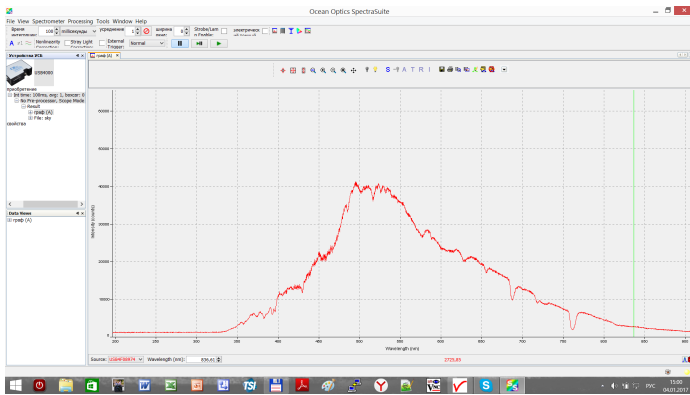


Рис. 2.2. Спектр неба, С.-Петербург, 2017-01-04

На рисунке 2.2 показан спектр неба, С.-Петербург, 2017-01-04. Видно качественное соответствие спектров рисунков 2.1 и 2.2.

Для классификации причин возникновения различных линий поглощения необходимо выделить из полного набора те факты, которые относятся к свойствам атмосферы, и те, что относятся к источнику излучения — Солнцу.

Математически это можно выразить следующим образом. Построим, используя приёмы §3, мультиинтервалы, обязанные Солнцу и атмосфере Земли

$$M_S = M_{Sun} = \{\lambda_1, \dots, \lambda_{n1}\}, \quad (2.1)$$

$$M_E = M_{Earth} = \{\lambda_1, \dots, \lambda_{n2}\}, \quad (2.2)$$

$$M_s = M_{sample} = \{\lambda_1, \dots, \lambda_{ns}\}. \quad (2.3)$$



Здесь  $n1, n2, ns$  — количество интервалов в соответствующих мультиинтервалах  $M_S, M_E, M_s$ .

Далее можно строить различные отношения между мультиинтервалами стандартов и образца  $M_S, M_E, M_s$  и вычислять меры совместности, характеризующие и сходство [7].

## 2.1.2 Элементо-определение состава смеси веществ

Рассмотрим пример из геологии. Медно-цинковые руды являются сложным комплексом сульфидов меди, цинка и железа а так же минералов вмещающих пород.

Атомарные спектры (поглощения или испускания) получают переводением вещества в парообразное состояние путём нагревания пробы до 1000—10000 °С. В качестве источников возбуждения атомов при эмиссионном анализе токопроводящих материалов применяют искру, дугу переменного тока; при этом пробу помещают в кратер одного из угольных электродов.

Регистрация излучения может производиться различными способами. Например, в качестве детектора моуг использоваться линейки ПЗС (приборы с зарядовой связью, ССD). Системы регистрации на основе линейных ПЗС были названы фотоэлектронными кассетами (ФЭК).

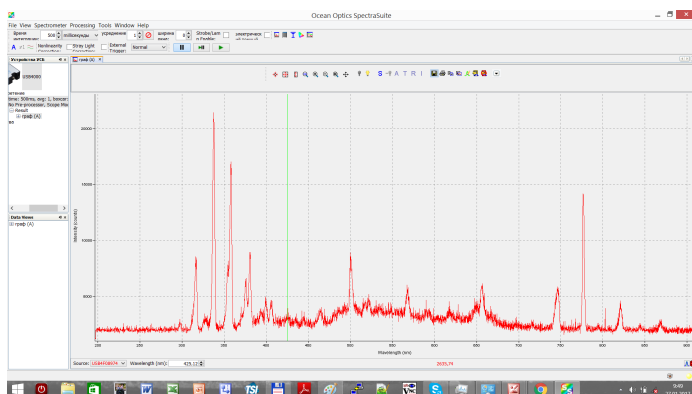


Рис. 2.3. Спектр паров меди в дуге

Для качественных измерений используют приборы низкого разрешения. На рисунке 2.3 показан спектр паров меди в дуге, малогаба-

ритный прибор Ocean Optics (<https://oceanoptics.ru/spectrometers/369-oceansr.html>).

Для определения состава руд используются различные стандарты, внесённые в Государственный реестр. На рисунке 2.4 показан спектр стандартного образца РУС-2, снятый в ВНИИОкеангеологии им. И.С.Грамберга на спектральном приборе спектрограф со скрещенной дисперсии СТЭ-1 [11] с ФЭК МОРС (<http://www.oomors.ru/kasseti.html>).

Стандартный образец РУС-2 содержит 0.7% меди, 0.05% цинка, 0.02% свинца и ряд других элементов [12].

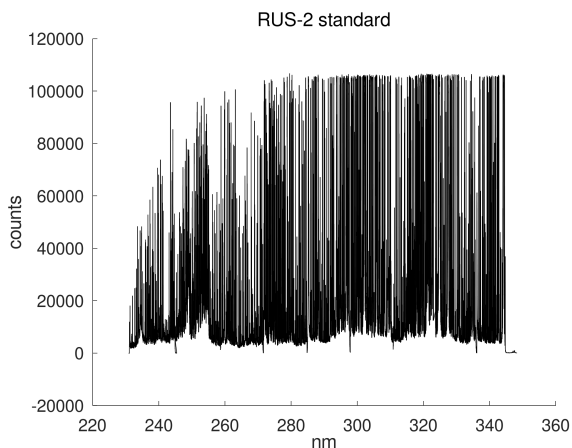


Рис. 2.4. Спектр стандарта РУС-2

Спектр содержит большое количество спектральных линий — см. рисунок 2.4. Их идентификация проводится на основе табличных баз данных.

### Сравнение спектров образца и стандарта

Проведём сравнение спектров стандарта и образца. На рисунке 2.5 показан участок спектра стандарта РУС-2 в диапазоне 261-263 нм.

Стандарты подбираются таким образом, чтобы их состав был сходен с исследуемым объектом. ФГУП ВНИИОкеангеологии им. И.С.Грамберга традиционно занимается изучением донных осадков. На

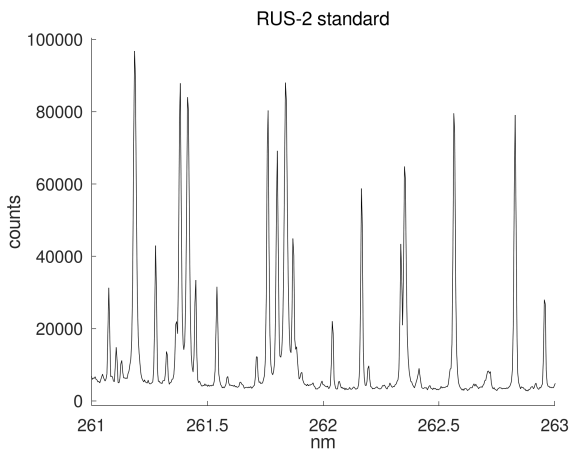


Рис. 2.5. Участок спектра стандарта РУС-2

рисунке 2.6 показан участок спектра исследуемого образца 257 в том же диапазоне 261-263 нм, что и на рисунке 2.5 для спектра стандарта РУС-2.

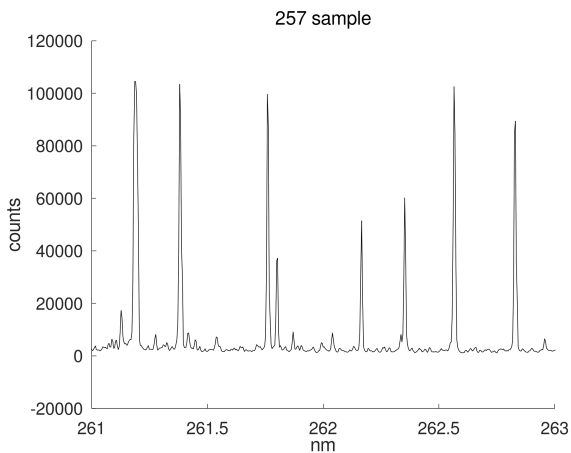


Рис. 2.6. Участок спектра образца

Проведём непосредственное сравнение спектров образца и стандар-

та. На рисунке 2.7 показан участок спектров исследуемого образца и стандарта в диапазоне 261-263 нм.

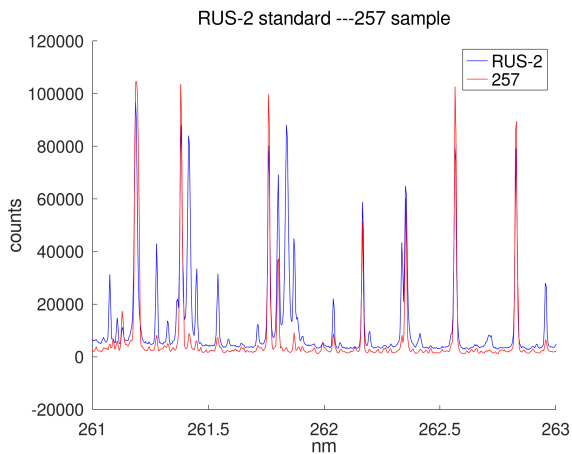


Рис. 2.7. Участок спектров образца стандарта и стандарта

Из рисунка 2.7 видно, что значительная часть линий спектра образца покрывается данными стандарта. Однако покрытие неполное.

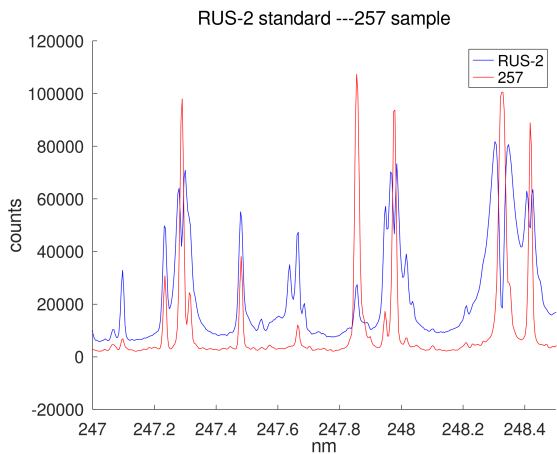


Рис. 2.8. Участок неполного покрытия стандартом спектра образца

Из рисунка 2.8 видно, что ряд линий спектра образца расположен между линиями в спектре стандарта. Вопрос о том, является наблюдаемый эффект приборным, или же в спектре имеются линии веществ, не присутствующие в стандарте, требует отдельного изучения. Например, можно попытаться обнаружить неопознанные линии в других стандартах. Например, пробуем стандартный образец СГХМ-2 «Стандартный образец состава алюмосиликатных рыхлых отложений» [13].

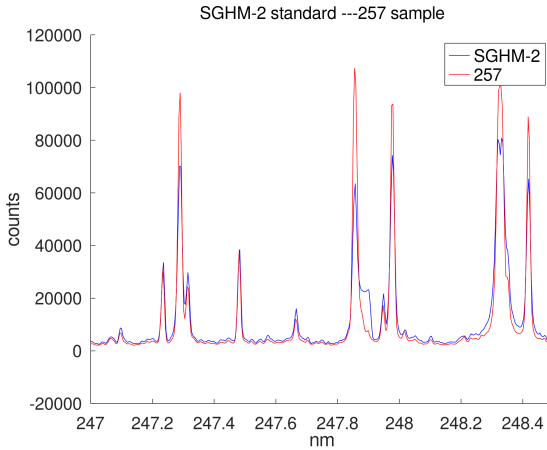


Рис. 2.9. Покрытие стандартом спектра образца СГХМ-2

Из рисунка 2.9 видно, что для стандарта СГХМ-2 покрытие линий спектра образца стандартом хорошее.

Совокупность этих фактов (рис. 2.8 и 2.9) говорит о необходимости использования для идентификации спектральных линий образца спектры нескольких стандартных образцов.

Математически это можно выразить следующим образом. Постороним, используя приёмы §3, мультиинтервалы образцов и пробы

$$M_{s1} = M_{standard1} = \{\lambda_1, \dots, \lambda_{n1}\}, \quad (2.4)$$

$$M_{s2} = M_{standard2} = \{\lambda_1, \dots, \lambda_{n2}\}, \quad (2.5)$$

⋮

$$M_s = M_{sample} = \{\lambda_1, \dots, \lambda_{ns}\}. \quad (2.6)$$

Здесь  $n_1, n_2, \dots, n_s$  — количество интервалов в соответствующих мультиинтервалах  $M_{s1}, M_{s2}, \dots, M_s$ .

Далее можно строить различные отношения между мультиинтервалами стандартов и образца  $M_{s1}, M_{s2}, \dots, M_s$  и вычислять меры совместности, характеризующие и сходство.

Следует заметить, что реальные данные часто содержат артефакты, которые усложняют интерпретацию спектров и связанных с ними математических приёмов. К одному из типов артефактов относится перегрузка приёмника излучения при очень интенсивном сигнале.

Например, на рисунке 2.10 наблюдаются два примера перегрузки, которые находятся в областях  $\approx 285$  и  $\approx 288$  нм. Спектральные линии имеют неправильную форму и большую ширину.

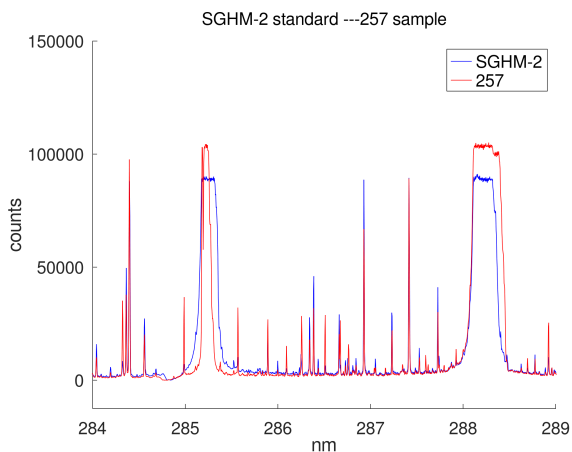


Рис. 2.10. Перегрузка области спектра

Идентификация спектральной линии в таком случае затрудняется возможным расположением в её области нескольких табличных линий элементов. Для выяснения природы сигнала в таких случаях требуется дополнительная информация и обработка элементов-кандидатов на наличие и других линий в спектре.

### 2.1.3 Красное смещение и лес Лаймана

В астрофизике большое количество информации о межзвёздной среде получают из линий поглощения [15] объектов Вселенной. Важным объектом изучения являются квазары. Квазар — класс астрономических объектов, являющихся одними из самых ярких в видимой Вселенной. Английский термин *quasar* образован от слов *quasi-stellar* («квазизвёздный» или «похожий на звезду») и *radiosource* («радиоисточник») и дословно означает «похожий на звезду радиоисточник».

Квазары видны с огромных расстояний, по ним исследуют структуру и эволюцию Вселенной, определяют распределение вещества на луче зрения: сильные спектральные линии поглощения водорода разворачиваются в лес линий по красному смещению поглощающих облаков. Ввиду большой удалённости квазары, в отличие от звёзд, выглядят практически неподвижными (не имеют параллакса), поэтому радиоизлучение квазара используется для высокоточного определения с Земли параметров траектории автоматической межпланетной станции.

Каждая порция газа на линии прямой видимости от квазара будет избирательно поглощать определенные длины волн континуального света квазара из-за наличия различных химических элементов в газе. Благодаря анализу этих линии поглощения можно изучать различные эффекты.

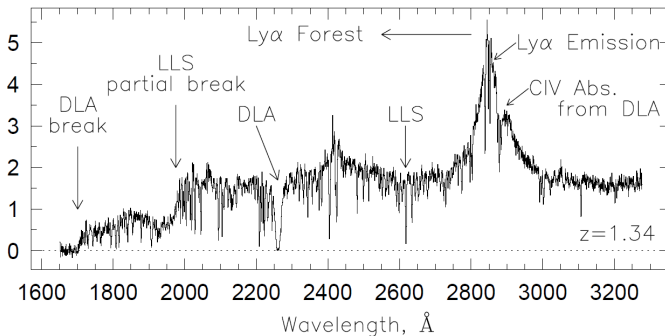


Рис. 2.11. Типичный спектр квазара

На рисунке 2.11 представлен типичный спектр квазара, показывающий квазарный континуум, эмиссионные линии и линии поглощения, которые производятся галактиками и межгалактическим материалом,

находящимся между квазаром и наблюдателем. Спектр  $z = 1.34$  от квазара PKS0454+039 был получен с помощью спектрографа слабых объектов на космическом телескопе Хаббл. Эмиссия линии на  $2400 \text{ \AA}$  и  $2850 \text{ \AA}$  — это  $\text{Ly}\beta$  и  $\text{Ly}\alpha$ . Лес  $\text{Ly}\alpha$ , поглощение, создаваемое различными межгалактическими облаками, очевиден на длинах волн, расположенных ближе к синему свету от линии излучения  $\text{Ly}\alpha$ . Два самых сильных поглотителя, обусловленные галактиками, представляют собой затухающую Поглотитель  $\text{Ly}\alpha$  при  $z = 0.86$  и предельная система Лаймана при  $z = 1.15$ . Первый вызывает нарушение предела Лаймана при температуре  $1700 \text{ \AA}$ . и последний является частичным нарушением предела Лаймана при  $1950 \text{ \AA}$ , поскольку плотность столба нейтрального водорода недостаточно велика для этого, поглощать все ионизирующие фотоны. Многие линии поглощения создаются DLA при  $z = 0,86$  (например,  $\text{Civ}$   $\lambda\lambda$  1548, 1550, смещено в красную область на красное крыло эмиссионной линии  $\text{Ly}\alpha$  квазара).

Линия Лайман-альфа (англ. Lyman-alpha line,  $\text{Ly}\alpha$ ) — спектральная линия водорода (или, в общем случае, водородоподобного атома) в серии Лаймана. Излучается в случае, когда электрон в атоме переходит с уровня  $n = 2$  в основное состояние ( $n = 1$ ), здесь  $n$  — главное квантовое число. Для атома водорода длина волны линии составляет  $1215.67$  ангстремов ( $121.567 \text{ нм}$ ), то есть линия находится в ультрафиолетовой области спектра электромагнитного излучения.

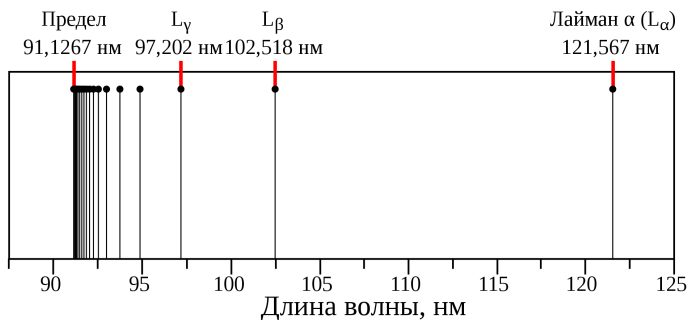


Рис. 2.12. Спектральная серия Лаймана

На рисунке 2.12 показана спектральная серия Лаймана.

Формула Ридберга для серии Лаймана выглядит следующим обра-



зом:

$$\frac{1}{\lambda} = R_H \left( \frac{1}{1^2} - \frac{1}{n^2} \right), \quad \left( R_H \approx 1,0968 \times 10^7 \text{ м}^{-1} \approx \frac{13,6 \text{ eV}}{hc} \right), \quad (2.7)$$

где  $n$  — главное квантовое число — натуральное число большее либо равное 2;  $R_H$  — постоянная Ридберга для атома водорода.

Самая коротковолновая граница серии Лаймана 91.175 нм, что соответствует энергии 13.6 эВ — минимальной энергии ионизации невозбуждённого атома водорода. С этой стороны в наблюдаемом спектре водорода к серии Лаймана примыкает сплошной спектр, соответствующей энергиям превышающим энергию связи электрона и протона в атоме водорода.

Все линии серии Лаймана находятся в ультрафиолетовом диапазоне. В таблице 2.1,  $n$  — главное квантовое число. Длина волны 91.15

$n$	2	3	4	5	6	7	8	9	10	11	$\infty$
$\lambda$ , нм	121.6	102.5	97.2	94.9	93.7	93.0	92.6	92.3	92.1	91.9	91.15

Таблица 2.1. Длины волн серии Лаймана

нм называется *пределом серии Лаймана*.

*Лес Лайман-альфа* (Ly- $\alpha$ -лес) — многократное повторение абсорбционной линии Лайман-альфа в спектрах далеких астрономических объектов. Для очень далёких объектов это явление может быть настолько сильным, что вызывает значительный спад интенсивности в некотором интервале длин волн; это называется *эффектом Ганна—Петерсона*.

Ly- $\alpha$ -лес возникает из-за облаков нейтрального водорода, через которые проходит свет от наблюдаемого объекта. Эти облака имеют разные красные смещения  $z$ . Длины волн линий, которые добавляет к спектру объекта каждое такое облако, зависят от его красного смещения. В результате густота и интенсивность этих линий несет в себе информацию о состоянии межгалактического газа, находящегося по пути следования принимаемого света.

На рисунке 2.13 показаны различные варианты геометрии регистрации излучения квазара. Слева модель описывает радиальное падение облаков к центру сферы с постоянной скоростью. Луч зрения проходит через пять облаков, что приводит к пяти различным особенностям поглощения (для одного перехода) в спектре квазара. В двух случаях

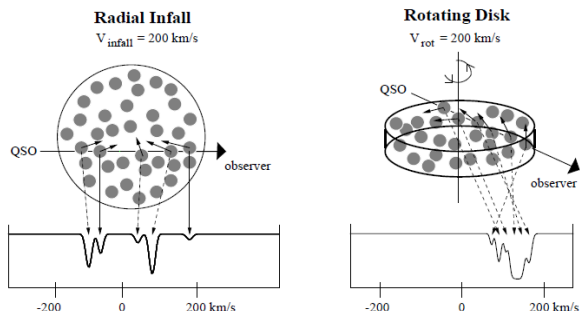


Рис. 2.13. Варианты геометрии регистрации излучения [15]  
 Слева — движение вещества к центру, справа — вращение

спектры смещены в синюю область относительно остальной поглощающей галактики, а остальные три имеют красное смещение. Такая модель поглощения от радиальной модель падения может быть распространена на скорости 100–200 км/с, типичные для дисперсии скоростей гало галактики.

Справа изображена модель вращающегося диска. В этом случае все «облака» имеют тангенциальную составляющую движения. Они смещены в красную область и имеют тенденцию группироваться вместе в пространстве скоростей с типичным разбросом 20–60 км/с.

На рисунке 2.14 показан пример профилей поглощения  $0.4 < z < 1.4$  для ионов Mg-ii, наблюдаемых с помощью спектрографа Keck/HIRES при разрешении по длине волны  $R = 45\,000$ , что соответствует разрешению в поле скоростей равному 6 км/с. Сплошные линии на этих данных соответствуют *профилю Фойгта* [16], а отметки представляют скорости облаков. Некоторые из этих профилей соответствуют кинематике вращающегося диска, а другие соотносятся с радиальной кинематикой падения. Однако для объяснения всего ансамбля профилей необходима модель, объединяющая обе базовые виды кинематики.

На рисунке 2.15 показаны спектры для различных случаев величины красного смещения.

Иллюстрация эволюции структуры межгалактического газа от высокого красного смещения к низкому. Верхний спектр квазара с  $z = 3.6$  — это наблюдение Keck/HIRES, а нижний спектр — это наблюдения FOS/HST квазара с  $z = 1.3$ . На верхнем рисунке лес линий Ly гораздо более густой.

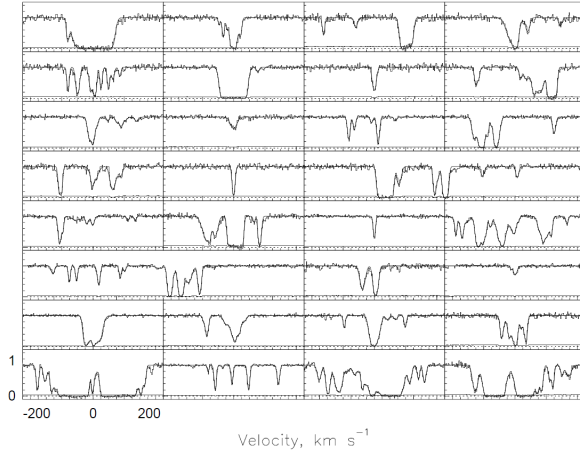


Рис. 2.14. Примеры спектров для различных случаев геометрии (рис. 2.13) регистрации излучения квазара [15]

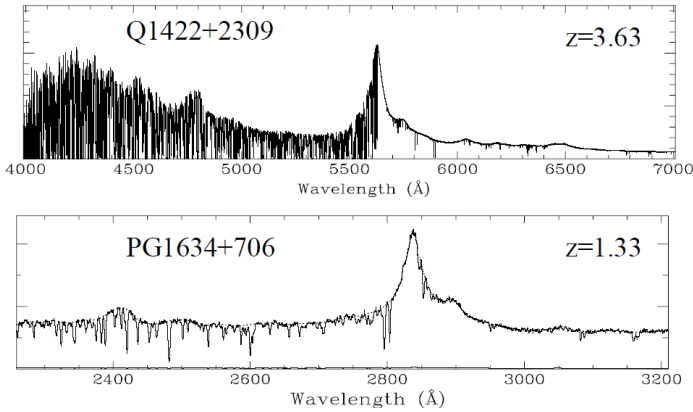


Рис. 2.15. Спектры для различных случаев красного смещения [15]

Рассмотренный выше примеры различных типов движений межзвёздной среды (рис. 2.13 и 2.14) подталкивает к описанию спектральных данных мультиинтервалами с разными свойствами.

Математически это можно выразить следующим образом. Построим, используя приёмы §3, мультиинтервалы двух теоретических типов

движения (th1, th2) и реальных спектров (s) как

$$M_{th1} = M_{Infall1} = \{\lambda_1, \dots, \lambda_{n1}\}, \quad (2.8)$$

$$M_{th2} = M_{Radial2} = \{\lambda_1, \dots, \lambda_{n2}\}, \quad (2.9)$$

$$M_s = M_{sample} = \{\lambda_1, \dots, \lambda_{ns}\}. \quad (2.10)$$

Здесь  $n1, n2, ns$  — количество интервалов в соответствующих мультиинтервалах  $M_{th1}, M_{th2}, M_s$ .

Далее можно строить различные отношения между мультиинтервалами стандартов и образца  $M_{s1}, M_{s2}, \dots, M_s$  и вычислять меры совместности, характеризующие и сходство. В данном случае количество интервалов в  $M_{th1}$  и  $M_{th2}$  ожидается разное. Также может существенно отличаться ширина интервалов.

## 2.2 Масс-спектрометрия

Масс-спектрометрический анализ осуществляется по спектрам масс атомарных или молекулярных ионов. Данные масс-спектров можно использовать в широком классе задач. В частности, можно определять изотопный состав объекта.

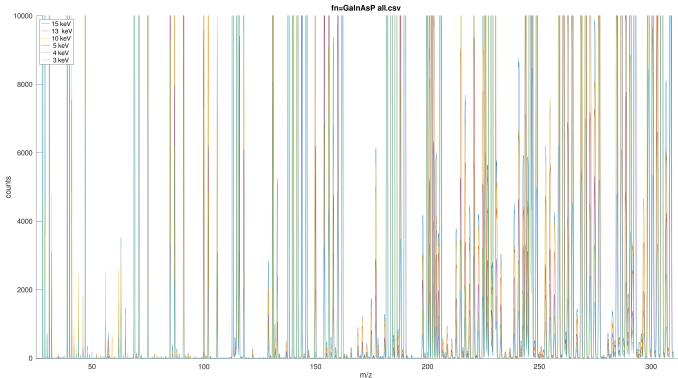


Рис. 2.16. Масс-спектры GaInAsP в масс-спектрометрии с вторичными ионами

Сплав  $Ga_xIn_{1-x}As_yP_{1-y}$  (Gallium Indium Arsenide Phosphide) является весьма перспективным материалом для преобразователей солнечной энергии. В частности, с использованием материала

$Ga_{0.93}In_{0.07}As_{0.87}P_{0.13}$  достигнута рекордная эффективность преобразования солнечной энергии, порядка 40% [17].

На рисунке 2.16 показаны масс-спектры GaInAsP в масс-спектрометрии с вторичными ионами (Secondary-ion mass spectrometry, SIMS) (лаб. Брункова П.Н., ФТИ им. А.Ф. Иоффе) [18]. Энергия первичных ионов составляла от 3 до 15 кэВ, что отражено в легенде. Вертикальный масштаб выбран так, чтобы наблюдать как линии с большой интенсивностью, так и с малой.

В спектр вносят вклад различные комбинации изотопов элементов. Фосфор и мышьяк имеют по одному стабильному изотопу (31 и 75 соответственно), галлий и индий — по 2 (69 и 71, 113 и 115) [20].

Рассмотрим сначала «чистые» пики галлия и индия.

Элемент	Масса-распространённость	Масса-распространённость
Ga	69 — 60%	71 — 40%
In	113 — 4%	115 — 96%

Таблица 2.2. Изотопный состав индия и галлия

В таблице 2.2 представлены приближённые данные распространённости изотопов индия и галлия (<https://applets.kcvs.ca/IPTEI/IPTEI.html>).

На рисунке 2.18 показаны линии спектра, соответствующие различным изотопам галлия и индия.

Заметим, что сумма интенсивностей линий изотопов галлия и такой же суммы галлия различны. Это может быть связано с разной эффективностью регистрации ионов этих элементов [18].

Рассмотрим теперь, какие варианты возможны в линиях, содержащих различные атомы соединения GaInAsP. Каждая комбинация, содержащая индий или галлий, имеет 2 представителя. Если индий и галлий входят в соединение вместе, то всего имеется 4 элемента такой группы. Ввиду того, что разница масс изотопов каждого элемента равна 2, имеем не 4 пика, а три.

Минимальная масса такого иона составляет  $69 + 113 = 182$  атомной массы, максимальная —  $71 + 115 = 186$ . Исходя из этих данных, выберем диапазон масс 182-186. На рисунке 2.18 показаны линии спектра, соответствующие различным комбинациям изотопов галлия и индия:

$$M_{InGa} = \{182, 184, 186\}. \quad (2.11)$$

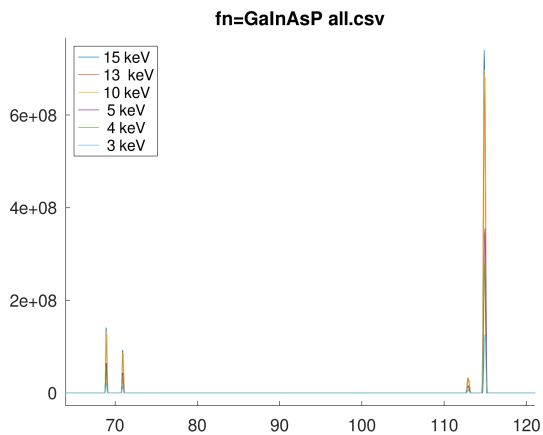


Рис. 2.17. Масс-спектры GaInAsP, область масс Ga и In

Интенсивность линий существенно различна, помимо разной эффективности регистрации ионов. Объясним этот факт, исходя из различной распространённости изотопов. В таблице 2.2 представлены приближённые данные по распространённости изотопов.

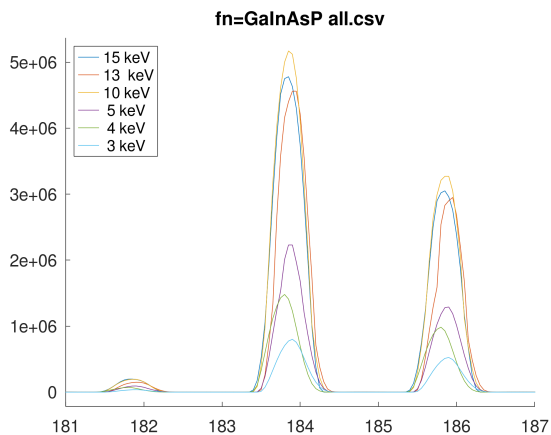


Рис. 2.18. Масс-спектры GaInAsP, область масс GaIn

Наибольшую интенсивность имеет пик с массой 184, являющийся суммой *изотопологов* [21]  $^{71}\text{Ga}^{113}\text{In}$  и  $^{69}\text{Ga}^{115}\text{In}$ . Наименьшую интенсивность имеет пик с массой 182 —  $^{69}\text{Ga}^{113}\text{In}$ .

Также в спектре наблюдаются различные ионы, состоящие из пары атомов: GaP, InP, GaAs, InAs.

Элемент	Масса-распространённость	Масса-распространённость
GaP	100 — 60%	102 — 40%
InP	144 — 4%	146 — 96%
GaAs	144 — 60%	146 — 40%
InAs	188 — 4%	190 — 96%

Таблица 2.3. Изотопный состав пар атомов GaInAsP

В таблице 2.3 представлены приближённые данные распространённости изотопов пары атомов: GaP, InP, GaAs, InAs.

Помимо уже рассмотренных масс-пиков в спектре есть и иные их комбинации, а также ионы от остаточной атмосферы, что даёт различные комбинации C, O, N. Также возможно загрязнение поверхности образца.

Приведём пример двух массовых пиков, не относящихся к исследуемому веществу, а возможно связаны с загрязнением калием и кальцием (39 и 40 масса).

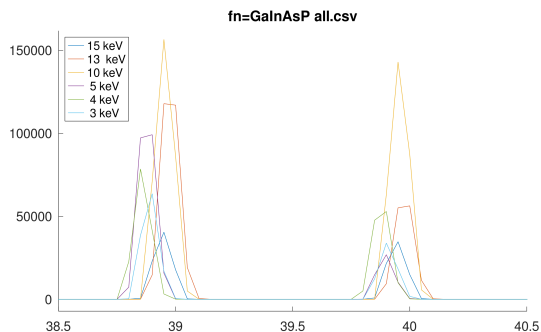


Рис. 2.19. Масс-спектры GaInAsP, область масс 39-40 а.е.

Совокупность этих фактов (рис. 2.8 и 2.9 ) говорит о необходимости использования для идентификации спектральных линий образца

спектры нескольких стандартных образцов.

Математически это можно выразить следующим образом. Построим, используя приёмы §3, мультиинтервалы теоретических масс-спектров и пробы

$$M_{Ga} = \{\lambda_1, \dots, \lambda_{n1}\}, \quad (2.12)$$

$$M_{In} = \{\lambda_1, \dots, \lambda_{n2}\}, \quad (2.13)$$

$$M_{GaIn} = \{\lambda_1, \dots, \lambda_{n3}\}, \quad (2.14)$$

$$M_{bg} = \{\lambda_1, \dots, \lambda_{n4}\}, \quad (2.15)$$

$$M_s = \{\lambda_1, \dots, \lambda_{ns}\}. \quad (2.16)$$

Здесь  $n1, n2, n3, ns$  — количество интервалов в соответствующих мультиинтервалах  $M_{Ga}, M_{In}, M_s$ .

$M_{Ga}$  — масс-спектр ионов, содержащих Ga;

$M_{In}$  — масс-спектр ионов, содержащих In;

$M_{GaIn}$  — масс-спектр ионов, содержащих Ga и In;

$M_{bg}$  — масс-спектр ионов, содержащих фон.

Далее можно строить различные отношения между мультиинтервалами стандартов и образца  $M_{Ga}, M_{In}, M_{GaIn}, M_s$  и вычислять меры совместности, характеризующие их сходство.

Численно, в конкретном случае, величина загрязнения мала и оценивается авторами работ на уровне 0.1% [19]. Вместе с тем, при проведении исследований весьма важно контролировать этот уровень и, по-возможности, понимать происхождение примесей.



## Глава 3

# Мультиинтервалы в обработке данных

### 3.1 Мода интервальной выборки

Для оценки выборки её важной характеристикой является *мода* — значение из выборки, которое встречается наиболее часто. Выборки и распределения с одной модой называются *унимодальными*, а с двумя и более модами — *мультимодальными* (бимодальными т.д.). Мода лучше, чем среднее значение (математическое ожидание) характеризует выборки с большим разбросом значений. Кроме того, мода, как характеристика «средней величины», может применяться при обработке данных, имеющих нечисловую природу.

В книге [1], понятие моды расширено на интервальные выборки и предложено

**Определение 3.1.1** *Модой интервальной выборки назовём совокупность интервалов пересечения наибольших совместных подвыборок рассматриваемой выборки. Наибольшая длина совместных подвыборок данной выборки называется частотой моды.*

Вычисление моды связано с введением важного понятия *элементарных подинтервалов измерений*. Это упорядоченное по линейному порядку множество интервалов, образованных из вершин интервалов выборки.

Таблица 3.1. Алгоритм для нахождения моды интервальной выборки и её частоты [1]

Вход

Интервальная выборка  $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^n$  длины  $n$ .

Выход

Мода  $\text{mode } \mathbf{X}$  выборки  $\mathbf{X}$  и её частота  $\mu$ .

Алгоритм

$\mathbf{I} \leftarrow \bigcap_{i=1}^n \mathbf{x}_i$ ;

IF  $\mathbf{I} \neq \emptyset$  THEN

$\text{mode } \mathbf{X} \leftarrow \mathbf{I}$ ;

$\mu \leftarrow n$

ELSE

    помещаем все концы  $\underline{\mathbf{x}}_1, \overline{\mathbf{x}}_1, \underline{\mathbf{x}}_2, \overline{\mathbf{x}}_2, \dots, \underline{\mathbf{x}}_n, \overline{\mathbf{x}}_n$

    интервалов рассматриваемой выборки  $\mathbf{X}$  в один

    массив  $Y = (y_1, y_2, \dots, y_{2n})$ ;

    упорядочиваем элементы в  $Y$  по возрастанию значений;

    порождаем интервалы  $\mathbf{z}_i = [y_i, y_{i+1}]$ ,  $i = 1, 2, \dots, 2n - 1$

    (назовём их *элементарными подинтервалами измерений*);

    для каждого  $\mathbf{z}_i$  подсчитываем число  $\mu_i$  интервалов

    из выборки  $\mathbf{X}$ , включающих интервал  $\mathbf{z}_i$ ;

    вычисляем  $\mu \leftarrow \max_{1 \leq i \leq 2n-1} \mu_i$ ;

    выбираем номера  $k$  интервалов  $\mathbf{z}_k$ , для которых  $\mu_k$

    равно максимальному, т. е.  $\mu_k = \mu$ , и формируем

    из таких  $k$  множество  $K = \{k\} \subseteq \{1, 2, \dots, 2n - 1\}$ ;

$\text{mode } \mathbf{X} \leftarrow \bigcup_{k \in K} \mathbf{z}_k$

END IF

Так как сама выборка, очевидно, является своей подвыборкой, то в случае её совместности мода совпадает с пересечением всех интервалов выборки. Если же выборка несовместна, то в ней может найтись несколько совместных подвыборок максимальной длины, и их пересечения нужно рассматривать в совокупности друг с другом. Как следствие, мода может быть мультиинтервалом (см. §1.2).

Псевдокод алгоритма для нахождения моды выборки интервальных измерений и её частоты из [1] приведён в Табл. 3.1. Число  $N$ , превосходящее общего числа  $2n$  концов интервалов выборки, необходимо для работы алгоритма потому, что некоторые концы интервалов обрабатываемой выборки могут совпадать друг с другом. Отметим также, что мода интервальной выборки — это интервал или мультиинтервал, который не обязан совпадать с каким-либо из интервалов обрабатываемой выборки.

**Пример 3.1.1.** Рассмотрим пример вычисления моды интервальной выборки.

Пусть имеется интервальная выборка из 4 элементов

$$\mathbf{X} = \{ [1, 4], [5, 9], [1.5, 4.5], [6, 9] \}. \quad (3.1)$$

Диаграмма рассеяния выборки  $\mathbf{X}$  приведена на Рис. 3.1.

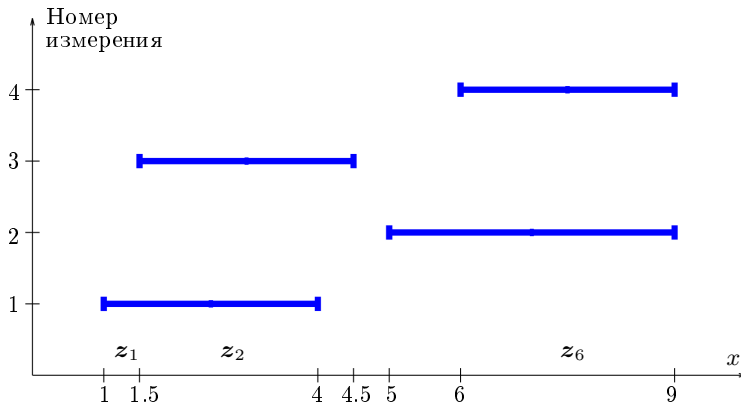


Рис. 3.1. Диаграмма рассеяния интервальной выборки (3.1) и элементы массива  $\mathbf{z}$ .

В соответствии с алгоритмом Табл. 3.1, проверим совместность  $\mathbf{X}$ .

Пересечение элементов выборки пусто

$$I = \bigcap_{i=1}^n x_i = \emptyset.$$

Таким образом, необходимо выполнить шаги алгоритма после ключевого слова ELSE.

Сформируем массив интервалов  $\mathbf{z}$  из концов интервалов  $\mathbf{X}$

$$\mathbf{z} = \{ [1, 1.5], [1.5, 4], [4, 4.5], [4.5, 5], [5, 6], [6, 9], [9, 9] \}. \quad (3.2)$$

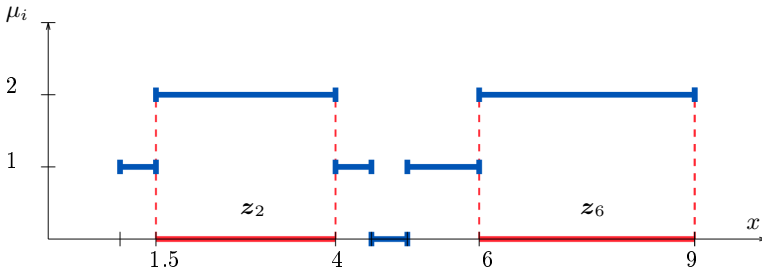


Рис. 3.2. Значения частот  $\mu_i$  и интервальная мода  $\text{mode } \mathbf{X}$  выборки (3.1), а также элементы  $\mathbf{z}_k$ ,  $k \in K$ , массива  $\mathbf{z}$ .

Для каждого интервала  $\mathbf{z}_i$  подсчитываем число  $\mu_i$  интервалов из выборки  $\mathbf{X}$ , включающих  $\mathbf{z}_i$ , получаем массив  $\mu_i$  в виде

$$\{1, 2, 1, 0, 1, 2, 2\}. \quad (3.3)$$

Максимальные  $\mu_i$ , равные 2, достигаются для индексного множества

$$K = \{2, 6, 7\},$$

так что частота моды равна  $\mu = 2$ . Как итог, мода является мультиинтервалом

$$\text{mode } \mathbf{X} = \bigcup_{k \in K} \mathbf{z}_k = [1.5, 4] \cup [6, 9]. \quad (3.4)$$

На Рис. 3.2 значения частот  $\mu_i$  (3.3) показаны синим цветом, а интервальная мода  $\text{mode } \mathbf{X}$  (3.4) — красным цветом. ■

## 3.2 Выборки унимодальные и мультимодальные

Тот факт, что выборка не является унимодальной, указывает на её неоднородность и может служить признаком сложной внутренней структуры описываемого ею явления. Получается, что из всего диапазона охватываемых выборкой значений выделяются тогда два или более изолированных друг от друга участка, одинаково доминирующих над остальными значениями по частоте. Если это доминирование велико, то исследуемая величина может, к примеру, не быть постоянной, а является «смесью» нескольких близких постоянных величин.

В предыдущих рассуждениях возможное различие частот подвыборок, на которые распадается исходная выборка, никак не учитывалось, и это является их существенным недостатком. Соответственно, необходимо распространить наши рассуждения на ситуации, когда интервальная выборка имеет несколько пиков частоты с разными значениями.

Дополнительное замечание. Здесь и далее предполагается, что множество элементарных подинтервалов измерений выборки уже вычислено.

### 3.2.1 Алгоритм вычисления мод мультимодального распределения

Представим основные идеи возможного алгоритма вычисления мод многомодального распределения.

Неформально процесс можно представить как последовательное исключение из массива частот  $\mu_i$  областей пиков. Границы этих областей трудно определить формально в силу их неизвестной формы и возможного наличия участков немонотонности «склонов» пиков. Мы предлагаем приём, похожий на алгоритм сегментации областей методом «водораздела» (watershed) в компьютерной графике. Отличием от поиска водоразделов является использование максимумов, а не минимумов функции распределения.

В книге [1] предложенный алгоритм был только намечен. Здесь же будет дано развёрнутое изложение.

Предполагается, что все моды достаточно «компактны» и не содержат «тонкой структуры». Имеется в виду, что если текущая мода является мультиинтервалом, то его интервальная оболочка не имеет

Таблица 3.2. Алгоритм для нахождения частот мод мультимодальной интервальной выборки

Вход

Интервальная выборка  $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^n$  длины  $n$ .

Массив частот  $\{\mu\}$ . Первая мода  $\text{mode } \mathbf{X}^{(1)}$ .

Выход

Массив областей  $\mathbf{MX} = \{\mathbf{IH}\}$  выборки  $\mathbf{X}$  и массив частот  $M\mu(\mathbf{MX})$ .

Алгоритм

$\mathbf{MX} \leftarrow \text{mode } \mathbf{X}^{(1)}; M\mu \leftarrow \max \mu$

Установка признака выполнения задания  $\text{G0}=1$

WHILE  $\text{G0}$

    Обработка текущей моды  $k$ .

Вычисление  $\text{mode } \mathbf{X}^{(k)}$  — Алгоритм Табл. 3.1 §3.1

        Вычисление границ текущей моды  $k$ :

            Установка границ интервала  $\mathbf{IH}^{(k)} = \text{mode } \mathbf{X}^{(k)}$ .

Расширение границ  $\mathbf{IH}^{(k)}$  — Алгоритм Табл. 3.3.

        Зануление значений массива частот  $\mu$  с индексами  $\mathbf{IH}^{(k)}$ :

$\mu(\mathbf{IH}^{(k)}) := 0$ .

$\mathbf{MX} \leftarrow \mathbf{MX} \cup \mathbf{IH}^{(k)}, M\mu \leftarrow M\mu \cup \mu^{(k)}$ .

        Вычисление распределения частот  $\mu$ .

    IF  $\max \mu \leq 1$  THEN

        Окончание работы  $\text{G0}=0$

    ELSE

        Переход к следующей моде

    ENDIF

END WHILE

непустых пересечений с областями других мод. Такое непустое пересечение может иметь место в спектральном анализе атомов и особенно молекул, где мультиплетные распределения случаются часто [10].

Основой алгоритма является «Алгоритм для нахождения моды интервальной выборки», представленный в Табл. 3.1 §3.1. Этот алгоритм используется для вычисления первой («старшей») моды  $\text{mode } \mathbf{X} = \bigcup_{k \in K} z_k$ . Здесь множество  $K$  состоит из номеров  $k$  интервалов  $z_k$ , входящих в первую моду. Также на этом этапе вычисляется массив частот  $\mu$ .

Результатом работы алгоритма является вычисление массива мод  $\mathbf{MX} = \{\text{mode } \mathbf{X}^{(k)}\}_{k=1}^n$  выборки  $\mathbf{X}$  и соответствующего массива частот  $M\mu(\mathbf{MX})$ . Ввиду того, что массив мод полностью определяется массивом частот  $M\mu(\mathbf{MX})$ , дальнейшее изложение будет посвящено нахождению этого массива. Здесь  $n$  — полное число найденных мод.

Вычисления носят итерационный характер. Пусть найдена  $k - 1$  мода. Для следующих («младших») мод из интервальной выборки следует удалить последовательно области  $\mathbf{IH}^{(k)} \supseteq \text{mode } \mathbf{X}^{(k)}$ , в которых содержатся уже найденные моды  $\text{mode } \mathbf{X}^{(k)}$  с номерами  $1, 2, \dots, n - 1$ . Обозначение  $\mathbf{IH}$  указывает на *interval hall* — «расширенную» интервальную оболочку моды. Области  $\mathbf{IH}^{(k)}$  расширяют интервалы мод  $\text{mode } \mathbf{X}^{(k)}$ .

Критерием останова расширения (окончанием сегментации в алгоритме водораздела) является тот факт, что значения массива частот  $\mu$  на границах  $\underline{\mathbf{IH}}^{(k)}$  и  $\overline{\mathbf{IH}}^{(k)}$  становятся меньше, чем значение частоты следующей моды

$$\max\{ \mu(\underline{\mathbf{IH}}^{(k)}), \mu(\overline{\mathbf{IH}}^{(k)}) \} < \max \mu. \quad (3.5)$$

При выполнении условия (3.5) полагаем, что текущая мода распределения грубо околонулена.

Исключаем область расширенной области текущей моды  $\mathbf{IH}^{(k)}$  из общей выборки частот, следующая мода становится текущей и для неё вычисляется расширенная область.

Множество интервалов

$$\mathbf{IH} = \{ \mathbf{IH}^{(1)}, \mathbf{IH}^{(2)}, \dots \} \quad (3.6)$$

грубо описывает структуру интервальной выборки  $\mathbf{X}$ , а множество ин-

тервалов

$$\bigcup_{1 \leq k \leq n} \text{mode } \mathbf{X}^{(k)} = \left\{ \text{mode } \mathbf{X}^{(1)}, \text{mode } \mathbf{X}^{(2)}, \dots \right\} \quad (3.7)$$

даёт набор соответствующих интервальных мод со значениями частот  $\mu_n = \max_{k \in K^{(n)}} \mu_k$ . Множества  $K^{(k)}$  — суть множества номеров индексов  $\underline{\mathbf{IH}}^{(k)}$ .

Наиболее проблемным является этап алгоритма Табл. 3.2 «Расширение границ  $\underline{\mathbf{IH}}^{(k)} \supseteq \text{mode } \mathbf{X}^{(k)}$ ». Он, как и основной алгоритм, носит итерационный характер. В графике частот могут содержаться участки немонотонности «склонов» графика  $\mu$  или очень пологие склоны. В большинстве случаев для устойчивой работы алгоритма необходимы дополнительные параметры настройки. Пример реализации алгоритма приведён в Табл. 3.3.

В случае унимодальной выборки конструкции (3.6) и (3.7) соответствуют определению интервальной моды 3.1.1.

В алгоритме Табл. 3.3 выполняется проверка монотонности  $\mu$  на границах  $\underline{\mathbf{IH}}^{(k)}$ ,  $\overline{\mathbf{IH}}^{(k)}$ . В обозначениях алгоритма это означает проверку условий

$$\mu(\underline{\mathbf{IH}}^{(k)}) \geq \mu(\underline{\mathbf{IH}}^{(k)} - \text{step}), \quad (3.8)$$

$$\mu(\overline{\mathbf{IH}}^{(k)}) \geq \mu(\overline{\mathbf{IH}}^{(k)} + \text{step}). \quad (3.9)$$

Условия (3.8) и (3.9) используются независимо. Например, на Рис. 4.6 для первой моды левый и правый склоны пика не симметричны.

Критерием останова расширения (окончанием сегментации в алгоритме водораздела) является тот факт, что значения массива частот  $\mu$  на границах  $\underline{\mathbf{IH}}^{(k)}$  и  $\overline{\mathbf{IH}}^{(k)}$  становятся меньше, чем значение частоты следующей моды

$$\max\{ \mu(\underline{\mathbf{IH}}^{(k)}), \mu(\overline{\mathbf{IH}}^{(k)}) \} < \max \mu. \quad (3.10)$$

Значение  $\max \mu$  вычисляется для выборки с исключёнными из неё элементами расширенной моды:  $\mu(\underline{\mathbf{IH}}^{(k)}) = 0$ .

При выполнении условия (3.10) полагаем, что текущая мода распределения грубо оконтурена.



Таблица 3.3. Алгоритм для расширения границ интервальной моды

<p>Вход</p> <p>Мода <math>k</math> интервальной выборки <math>\mathbf{X} = \text{mode } \mathbf{X}^{(k)}</math></p> <p>Выход</p> <p>Интервал <math>\mathbf{IH}^{(k)}</math></p> <p>Алгоритм</p> <p><math>\mathbf{IH}^{(k)} \leftarrow [ \min \text{mode } \mathbf{X}^{(k)}, \max \text{mode } \mathbf{X}^{(k)} ]</math></p> <p><u>Вычисление начальных значений параметров алгоритма:</u></p> <p><math>step</math> — шаг по индексу, <math>noise</math> — уровень шума</p> <p><math>addstep</math> — увеличение шага по индексу</p> <p>Проверяем монотонность <math>\mu</math> на границах <math>\underline{\mathbf{IH}}^{(k)}, \overline{\mathbf{IH}}^{(k)}</math></p> <p>Проверяем условие COND валидности расширения - формула (4.11)</p> <p>WHILE COND</p> <p style="padding-left: 2em;">IF COND</p> <p style="padding-left: 4em;"><math>\underline{\mathbf{IH}}^{(k)} \leftarrow \underline{\mathbf{IH}}^{(k)} - step,</math></p> <p style="padding-left: 4em;"><math>\overline{\mathbf{IH}}^{(k)} \leftarrow \overline{\mathbf{IH}}^{(k)} + step</math></p> <p style="padding-left: 2em;">ELSE</p> <p style="padding-left: 4em;"><math>step \leftarrow step + addstep</math></p> <p style="padding-left: 2em;">END IF</p> <p style="padding-left: 2em;">Проверяем условие COND (4.11)</p> <p>END WHILE</p> <p><math>\mathbf{IH}^{(k)} \leftarrow [ \underline{\mathbf{IH}}^{(k)}, \overline{\mathbf{IH}}^{(k)} ].</math></p>
--

Начальный этап Вычисление начальных значений параметров алгоритма требует отдельного рассмотрения и довольно существенно зависит от характера зависимости  $\mu$  и задач ее интерпретации. В разделе §4.2 представлен практический пример.

# Глава 4

## Примеры

В настоящей главе рассмотрим несколько примеров вычислений. Первый пример §4.1, носит модельный характер. Далее следуют реальные примеры §4.2 и §4.3.

### 4.1 Модельная выборка — спектральный триплет

В спектральном анализе удобным способом калибровки прибора является использование *триплетов*. В таком случае по крайним линиям можно задать цену шкалы прибора по длине волн, а по центральной линии — проверить линейность.

**Пример 4.1.1.** Рассмотрим численный пример вычисления частот мультимоды по алгоритму Табл. 3.2.

Пусть уже рассчитан массив частот с помощью «Алгоритма для нахождения моды интервальной выборки», представленного в Табл. 3.1 §3.1.

$$\mu = \{1, 2, 3, 3, 4, 4, 3, 2, 1, 3, 4, 3, 5, 5, 4, 3, 0, 1, 3, 3, 1, 1, 0\}. \quad (4.1)$$

Идеализации в примере:

- Иерархия мод строгая: все моды имеют различные значения максимумов

- Наклоны на различных склонах мод сопоставимы по величине  
Проследим работу алгоритма по шагам.

**Шаг 0 — выделение основной моды.** На первом шаге локализуется основная мода. Основная мода  $K$  для выборки с частотами элементарных подинтервалов  $\mu$  (4.1) является интервалом и имеет индексы

$$K = \{ [13, 14] \}. \quad (4.2)$$

$$IH_{ini} = \{ [4, 5], [13, 14], [19, 20] \}. \quad (4.3)$$

На рисунке 4.1 показан график массива  $\mu$  (4.1). Красным цветом на рисунке 4.1 выделен интервал  $K$  (4.2).

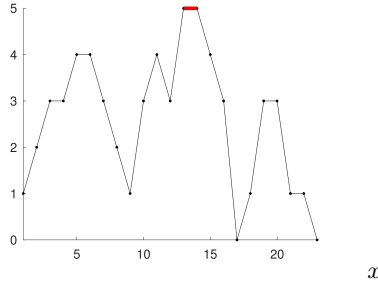


Рис. 4.1. Шаг 0 — основная мода.

**Шаг 1 — расширение основной моды.** На следующем шаге необходимо расширить интервал основной моды до тех пор, пока на его границах значение  $\mu$  графика не станет ниже «фона». Здесь работает алгоритм Табл. 3.3.

На рисунке 4.2 показан интервал расширения основной моды

$$IH^{(1)} = [10, 16] \supseteq IH_{ini}^{(1)} = [13, 14]. \quad (4.4)$$

**Шаг 2 — вторая мода.** Для вычисления минорных мод необходимо воспользоваться алгоритмом Табл. 3.2 и Табл. 3.3 §3.1. При этом данные предшествующей моды «зануляются».

$$IH^{(2)} = [2, 9] \supseteq IH_{ini}^{(2)} = [5, 6]. \quad (4.5)$$

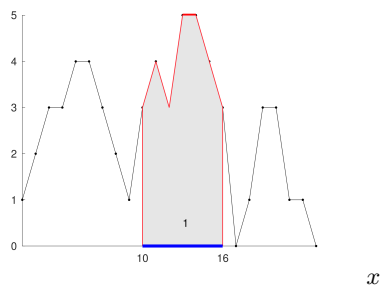


Рис. 4.2. Шаг 1 — расширение основной моды

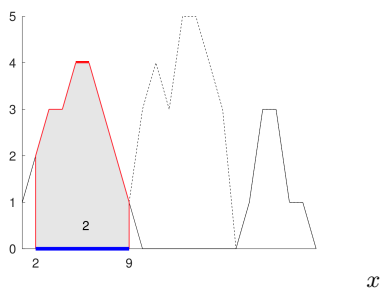


Рис. 4.3. Шаг 2 — вторая мода

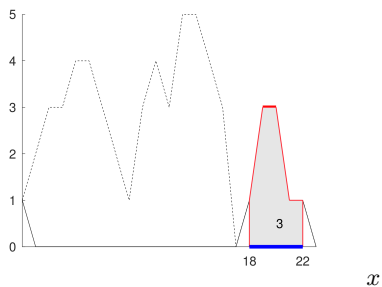


Рис. 4.4. Шаг 3 — третья мода

**Шаг 3 — третья мода.** Как и на предыдущем шаге, используем алгоритмы Табл. 3.2 и Табл. 3.3 §3.1. При этом данные предшествующей моды «зануляются».

$$\mathbf{IH}^{(3)} = [18, 22] \supseteq \mathbf{IH}_{ini}^{(3)} = [19, 20]. \quad (4.6)$$

### Общая картина — все моды.

Соберём результаты предыдущей работы вместе. На рисунке 4.5 показан график массива  $\mu$  (4.1).

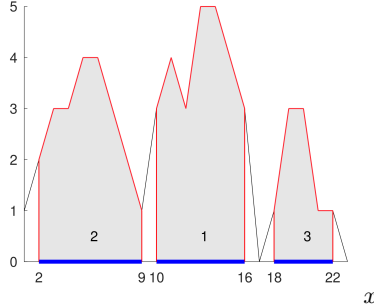


Рис. 4.5. Общая картина — все моды

Массив значений интервалов расширенных областей мод

$$\mathbf{IH} = \{[2, 9], [10, 16], [18, 22]\}, \quad (4.7)$$

$$\mathbf{IH}_{ini} = \{[4, 5], [13, 14], [19, 20]\}. \quad (4.8)$$

Для ориентировки даны начальные интервалы областей мод. В соответствии с алгоритмом расширения

$$\mathbf{IH}^{(i)} \supseteq \mathbf{IH}_{ini}^{(i)}. \quad (4.9)$$

Значения индексов частот мод в массиве  $\mu$

$$M\mu = \{4, 5, 3\}. \quad (4.10)$$

■

Заметим, что в приведённом примере не отмечено использование алгоритма вычисления параметра *noise* в алгоритме Табл. 3.3. Он разобран ниже — см. Табл. 4.1. Без правильной установки этого параметра алгоритму расширения интервала моды не удалось бы преодолеть локального минимума для индекса 12 и точки «застоя» 3.

## 4.2 Спектроскопия гамма-квантов

Приведём практический пример для иллюстрации работы алгоритма Табл. 3.3 Вычисление начальных значений параметров алгоритма и проблем, возникающих при работе с реальными данными.

Пример относится к спектроскопии электромагнитных излучений при высокой энергии. Речь идёт о детектировании гамма-квантов с энергией около 1 МэВ, что является типичной величиной для ядерных реакций, идущих в промышленных и исследовательских ядерных реакторах.

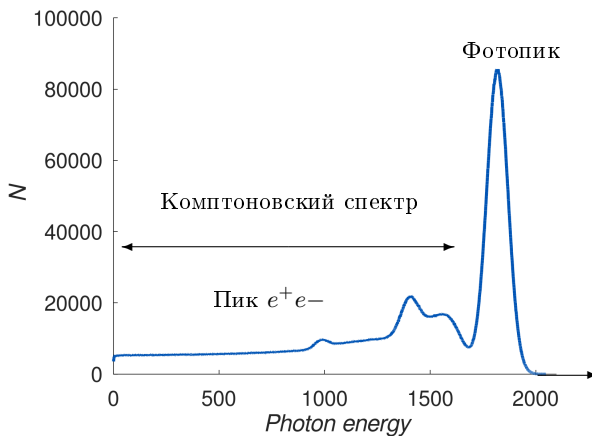


Рис. 4.6. Массив частот  $\mu$  при регистрации гамма-квантов

На Рис. 4.6 приведён типичный спектр регистрации такого кванта в сцинтилляционном детекторе. По оси абсцисс представлены энергии вторичных частиц (интервалы энергии), возникающих при регистрации. По оси ординат — абсолютное количество соответствующих событий. Спектр получен при помощи хорошо апробированной программы МСС3d [23]. Пик с наибольшей интенсивностью соответствует наиболее полному поглощению энергии гамма-кванта в детекторе — *фотопик*.

Между тем, в детекторе при попадании в него гамма-кванта возможны различные физические процессы, которые идут с различными вероятностями и приводят к сложному виду спектра, простирающегося влево вплоть до нулевых энергий, что соответствует прохождению

гамма-кванта через детектор практически без взаимодействия. Кроме этого, в спектре присутствуют дополнительные локальные максимумы, обязанные комптоновскому рассеянию гамма-квантов и образованию электрон-позитронных пар в электрическом поле атомных ядер. Диапазон комптоновского спектра на Рис. 4.6 показан двунаправленным вектором, его край в области высоких энергий имеет два локальных максимума из-за сложного состава детектора. Пик аннигиляции электрон-позитронных пар имеет энергию, равную удвоенной массе покоя электрона, 1022 кэВ.

Для того, чтобы придать процессу расширения границ объективный характер, необходимо учесть типичную разность между соседними величинами массива частот  $\mu$ . Характерное значение шума опоставимо с наиболее вероятным значением гистограммы рис. 4.7.

### Вычисление параметра *noise* алгоритма Табл. 3.3.

Способ вычисления параметра *noise* представлен в Табл. 4.1.

Выполним численное дифференцирование этого массива и построим гистограмму для массива  $N_z$  результата вычислений  $\text{HIST}(\Delta\mu)$  с числом бинов гистограммы  $M = 10$ .

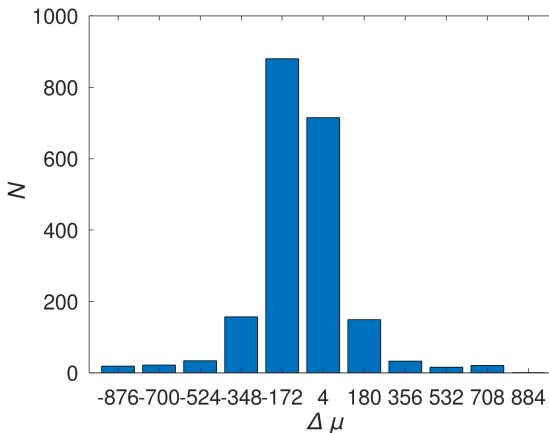


Рис. 4.7. Гистограмма разностей  $\Delta\mu$ .

Интервал гистограммы 5 имеет наибольшее значение  $N_5 \simeq 900$ . Этот интервал, как видно из Рис.4.7, имеет середину  $\text{mid } z_5 = -172$ .



Таблица 4.1. Алгоритм вычисления значения параметра *noise* алгоритма Табл. 3.3

<p><b>Вход</b></p> <p>Массив частот <math>\{\mu\}</math>, значение уровня представительности <math>\beta &lt; 1</math></p>
<p><b>Выход</b></p> <p><i>noise</i> — уровень шума</p>
<p><b>Алгоритм</b></p> <p><math>\{\Delta\mu\}_{k=1}^{n-1}</math> — численное дифференцирование массива <math>\{\mu\}_{k=1}^n</math></p> <p>Зададим число бинов гистограммы <math>M</math></p> <p>Установим условие NOK=TRUE</p> <p>WHILE NOK</p> <p style="padding-left: 2em;">Построение гистограммы <math>[z, N_z] = \text{HIST}(\Delta\mu)</math>,</p> <p style="padding-left: 4em;"><math>z_k</math> — массив интервалов, <math>N_z</math> — сумма <math>\Delta\mu</math> в бине <math>k</math></p> <p style="padding-left: 2em;">Отбрасываем выбросы гистограммы <math>\Delta\mu</math> — (1)</p> <p style="padding-left: 2em;">Находим интервал <math>z_i = \arg \max_i N_z</math></p> <p style="padding-left: 4em;">IF <math>N_z &gt; \beta \cdot \sum_{i=1}^n \mu_k</math></p> <p style="padding-left: 6em;"><math>noise \leftarrow \text{mid } z, \text{ NOK} = \text{FALSE}</math></p> <p style="padding-left: 4em;">ELSE</p> <p style="padding-left: 6em;"><u>изменить значение <math>M</math></u> — (2)</p> <p style="padding-left: 2em;">END IF</p> <p>END WHILE</p> <p>Возвращаем <i>noise</i></p> <p>(1) — выбросы гистограммы возникают в результате зануления элементов массива <math>\{\mu\}</math> при последовательном нахождении мод</p> <p>(2) — отдельный алгоритм</p>

Величина  $N_z$  немного меньше половины  $\sum_{i=1}^n \mu_k$ , так что при параметре  $\beta < 0.4$  работа алгоритма 2 даёт рабочее значение параметра  $noise \simeq 200$  алгоритма Табл. 4.1.

Уточним условие останова алгоритма (3.5):

$$\max \mu > \max\{ \mu(\underline{IH}^{(k)}), \mu(\overline{IH}^{(k)}) \} + noise. \quad (4.11)$$

По мере выявления мод, из распределения  $\mu$  удаляются соответствующие участки распределения, так что и параметр  $noise$  в (4.11) тоже изменяется.

**Выбор параметра  $step$  алгоритма Табл. 3.3.** Параметр  $step$ , который используются для расширения границ  $I\mathbf{H}^{(n)}$  задает степень подробности определения мод исходной выборки. Их величина не фиксирована чётко, и служит для исследований. Например, при обзорном рассмотрении распределений можно оконтурить мультиплетные области и далее разрешить их структуру с использованием меньшего шага.

**Выбор параметра  $addstep$  алгоритма Табл. 3.3.** Параметр  $addstep$  необходим для работы алгоритма в областях распределения, где шумы приводят к немонотонности.

Конкретно, на Рис. 4.8 показаны результаты работы алгоритма при величинах шага  $step = 100$  и  $50$ . Параметр  $addstep = 50$ .

**Результаты вычислений.** Приведем численные результаты. Примем значение шага, равное 100. Получаем массив расширенных областей мод

$$I\mathbf{H} = \{ [1618, 2018], [1156, 1617], [744, 1155], \quad (4.12)$$

$$[481, 743], [217, 480], [1, 216] \}. \quad (4.13)$$

Более грубый шаг, равный 100, даёт и более грубое выявление структуры. Конкретно, область  $I\mathbf{H}^{(2)} = [1156, 1617]$ , на Рис. 4.8 а) содержит два локальных максимума. Кроме того, область  $I\mathbf{H}^{(1)}$  чрезмерно широка и содержит часть соседней области слева.

В случае шага, равного 50, удастся разрешить структуру  $I\mathbf{H}^{(2)}$ , и выделить две моды на краю комптоновского распределения,  $[1305, 1506]$  и  $[1506, 1702]$ , что видно из Рис. 4.8 б).

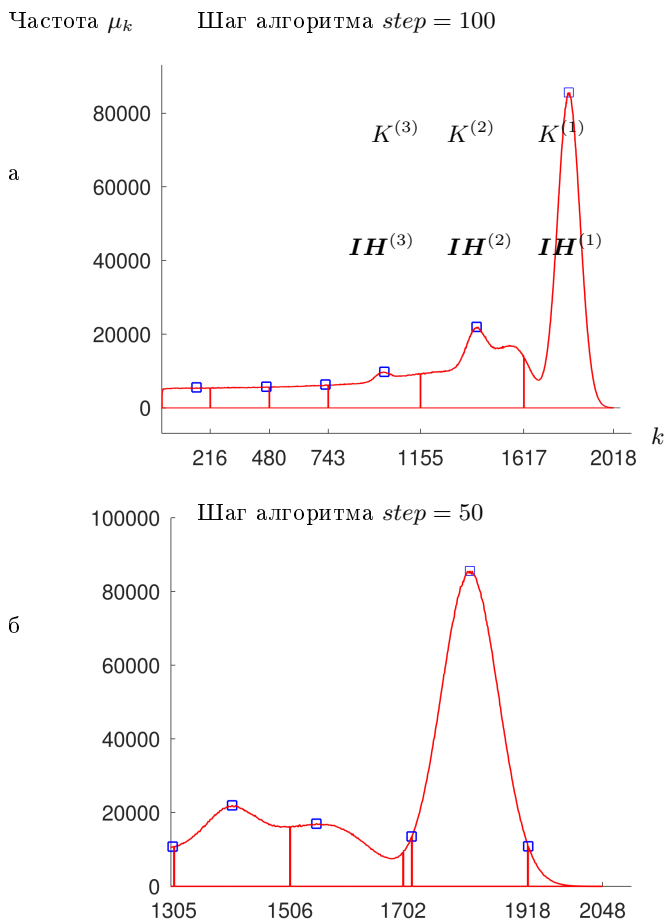


Рис. 4.8. Пример работы алгоритма Табл. 3.2  
 а — грубое распознавание структуры, б — уточнённое распознавание структуры распределения мод мультимодальной выборки

Область фотопика [1702, 1918] на Рис. 4.8 б) также определена более корректно. Обратной стороной более тонкого вычисления является появление артефактов малой ширины, например [ 1918, 2048 ]. Для их отбрасывания нужны дополнительные алгоритмы.

Ширины расширенных областей мод (4.12) составляют

$$\text{wid } \mathbf{IH} = \{ 400, 461, 411, 262, \dots \}. \quad (4.14)$$

В §4.3 мы сравним эти данные с результатами обработки атомарных спектров излучения плазмы.

### 4.3 Спектроскопия атомного и молекулярного излучения

Приведём ещё один практический пример для иллюстрации работы алгоритма Табл. 3.3 Вычисление начальных значений параметров алгоритма и проблем, возникающих при работе с реальными данными.

Пример относится к спектроскопии атомного и молекулярного излучения — см. §2.1. Речь идёт о детектировании световых квантов с энергией 1-10 эВ, что на 5-6 порядков меньше, чем в случае §4.2.

#### Входные данные.

Предметная область относится к физике плазмы.

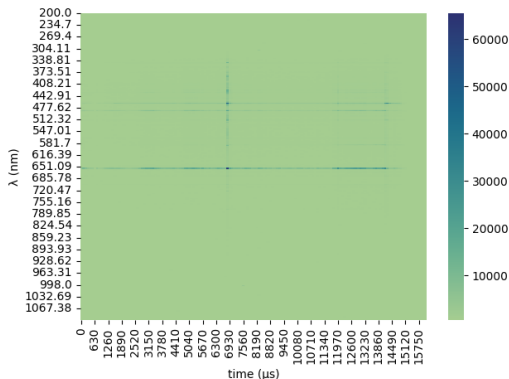


Рис. 4.9. Тепловая карта интенсивностей для разряда токамака Глобус-М2 [25]

В качестве входных данных использовалась таблица измерений интенсивности УФ-излучения, полученная с токамака Глобус-М2 ФТИ

им. А. Ф. Иоффе [24]. Наглядное представление спектра в виде двумерного графика приведено на рис. 4.9 [25].

На рис. 4.9 каждый вертикальный срез соответствует спектру излучения вещества плазмы. Пример такого спектра приведён на рис. 4.10. Квадратом выделена главная мода.

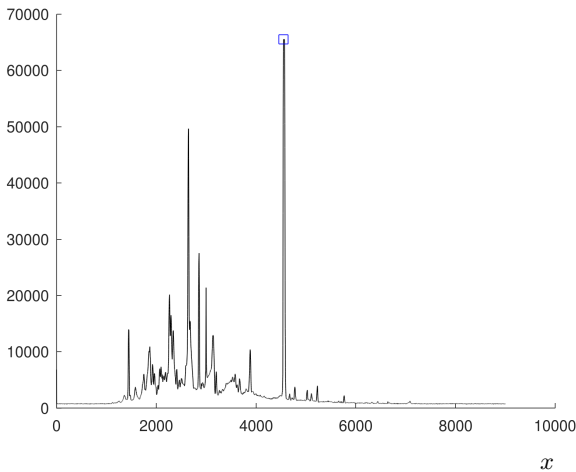


Рис. 4.10. Спектр излучения вещества плазмы токамака Глобус-М2

В зависимости от времени и течения разряда токамака получаются спектры, сходные с тем, который приведён на рис. 4.10, но с разными интенсивностями для различных длин волн.

**Работа алгоритма по выделению мод.** На рис. 4.11 показан результат работы алгоритма по выделению мод.

**Значения расширенных областей мод.**

Значения расширенных областей мод рис. 4.11:

$$\mathbf{IH} = \{ [4515, 4585], [2627, 2667], [2837, 2877], [2978, 3018], \dots \}. \quad (4.15)$$

Ширины расширенных областей мод (4.15) составляют

$$\text{wid } \mathbf{IH} = \{ 70, 40, 40, 40, \dots \}. \quad (4.16)$$

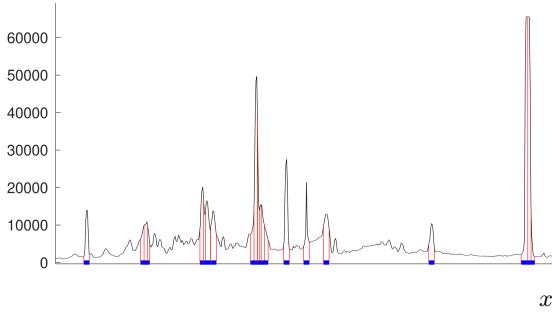


Рис. 4.11. Главные моды спектра рис. 4.10

Сравним ширины расширенных областей мод (4.16) с ширинами (4.14) для расширенных областей мод гамма-спектров.

$$\text{wid } \mathbf{IH} = \{ 400, 461, 411, 262, \dots \}. \quad (4.14)$$

Видно, что для оптических спектров ширины мод существенно более узкие. В связи с этим, параметры работы алгоритмов тоже иные. В частности, параметр *step* брался равным 10, а не 50-100, как в §4.2.

Значения частот для (4.15)

$$M\mu = \{ 65535, 49659, 27536, 21394, \dots \}. \quad (4.17)$$

Значение частоты для главной моды, равное  $65535 = 2^{16} - 1$ , указывает на факт «насыщения» регистрирующей системы. Таким образом, информация об интенсивности излучения для соответствующей линии спектра искажена.

# Заключение

Автор признателен участникам Всероссийского вебинара по интервальному анализу (<http://interval.ict.nsc.ru/WebSeminar/index.html>) С. И. Жилину, С. И. Кумкову, А. В. Пролубникову, Е. В. Чаусовой, С. П. Шарому за конструктивное сотрудничество в области интервального анализа.

Выражаю благодарность за экспериментальные данные, представленные коллегами по ФТИ им. А.Ф.Иоффе и ВНИООкеангеологии им. И.С.Грамберга.

Часть материала пособия была представлена на Всероссийской конференции «Неделя науки ФизМех» (1-5 апреля 2024 года) [25], сборник готовится к публикации.

# Литература

- [1] А.Н. БАЖЕНОВ, С.И. Жилин, С.И. Кумков, С.П. ШАРЫЙ. «Обработка и анализ данных с интервальной неопределённостью». – Москва–Ижевск: НИЦ «Регулярная и хаотическая динамика» РХД, 2024, - 356 с.
- [2] ШАРЫЙ С.П. Конечномерный интервальный анализ. – ФИЦ ИВТ: Новосибирск, 2022. Электронная книга, доступная на <http://www.nsc.ru/interval/Library/InteBooks/SharyBook.pdf>
- [3] А.Н.БАЖЕНОВ Интервальный анализ. Основы теории и учебные примеры: учебное пособие. – СПб. 2020 <https://elib.spbstu.ru/dl/2/s20-76.pdf/info>
- [4] KEARFOTT R.B., NAKAO M., NEUMAIER A., RUMP S., SHARY S.P., VAN HENTENRYCK P. Standardized notation in interval analysis // Вычислительные технологии. – 2010. – Т. 15, №1. – С. 7–13.
- [5] ЯКОВЛЕВ А.Г. Машинная арифметика мультиинтервалов // Вопросы кибернетики (Научный Совет по компл. проблеме «Кибернетика»: АН СССР). – 1986. – Вып. 125. – С. 66–81.
- [6] А.Н.БАЖЕНОВ Естественнаучные и технические мотивации интервального анализа: учебное пособие. – СПб., 2021. URL:<https://elib.spbstu.ru/dl/5/tr/2021/tr21-169.pdf>. — DOI 10.18720/SPBPU/5/tr21-169
- [7] А.Н.БАЖЕНОВ, А.Ю.Тельнова Обобщение коэффициента Жаккара для анализа данных с интервальной неопределённостью // Измерительная техника. 2022. № 12. С. 15–22.



- [8] БАЖЕНОВ А.Н., С.И.Жилин, А.Ю.Тельнова. Приведенные меры совместности в анализе данных с интервальной неопределенностью // Измерительная техника 2023 № 11, 2023, стр.17-25.
- [9] БАЖЕНОВ А. Н., Тельнова А. Ю. Комбинированный метод коррекции интервальных систем линейных алгебраических уравнений // Метрология. 2021. № 4. С. 17–39. [https:// doi.org/10.32446/0132-4713.2021-4-17-39](https://doi.org/10.32446/0132-4713.2021-4-17-39)
- [10] Ельяшевич М.А. Атомная и молекулярная спектроскопия. — М.: Физматгиз, 1962; М.: Эдиториал УРСС, 2001.
- [11] ТАРАСОВ К.И. Спектральные приборы. Изд-во «Машиностроение», 1968, стр. Табл. 22. Илл. 241. Библ. 183 назв.
- [12] Свидетельство на комплект стандартных образцов медноцинковых сульфидных руд (РУС-1, РУС-2, РУС-3, РУС-4). г.Свердловск, г.Полевской. 1976
- [13] СГХМ-2 (ГСО 3486-84). Стандартный образец состава алюмосиликатных рыхлых отложений Федеральное государственное бюджетное учреждение науки Институт геохимии им. А.П. Виноградова Сибирского отделения Российской академии наук. Иркутск, 1986 <http://www.igc.irk.ru/ru/item/2420-sgkxm-2-gso-3486-86>
- [14] СГХМ-4 (ГСО 3486-86). Стандартный образец состава алюмосиликатных рыхлых отложений Федеральное государственное бюджетное учреждение науки Институт геохимии им. А.П. Виноградова Сибирского отделения Российской академии наук. Иркутск, 1986 <http://www.igc.irk.ru/ru/item/2420-sgkxm-4-gso-3486-86>
- [15] J. CHARLTON, C. CHURCHILL. Quasistellar Objects: Intervening Absorption Lines. Encyclopedia of Astronomy and Astrophysics, Edited by Paul Murdin, article id. 2366. Bristol: Institute of Physics Publishing, 2001. doi:10.1888/0333750888/2366
- [16] ТЕЛИКОВА КСЕНИЯ НИКОЛАЕВНА. Изучение теплового состояния ранней Вселенной методами абсорбционной спектроскопии квазаров. Диссертация на соискание ученой степени канд. физ.-мат. наук. С.-Петербург, 2023. [https://www.ioffe.ru/serve/theses/avtoref/Thes\\_0595.pdf](https://www.ioffe.ru/serve/theses/avtoref/Thes_0595.pdf)

- [17] Wafer bonded four-junction GaInP/GaAs//GaInAsP/GaInAs concentrator solar cells with 44.7% efficiency. PROGRESS IN PHOTOVOLTAICS: RESEARCH AND APPLICATIONS Prog. Photovolt: Res. Appl. 2014; 22:277–282
- [18] Методическое пособие к лабораторной работе «вторично-ионная масс-спектрометрия». ФТИ им. А.Ф.Иоффе, СПб. 2024 - 12 с.
- [19] Sergey N. Bocharov, Boris E. Burakov, Anton I. Isakov, Ksenia N. Orekhova, Ekaterina V. Dementeva, Maria V. Zamoryanskaya, Peter A. Dementev, Boris Y. Ber, Dmitry Yu. Kasantsev, Mikhail V. Tokarev, Yuri Yu. Petrov, Self-glowing single crystal diamond activated with carbon-14: First synthesis and characterization, Diamond and Related Materials, Volume 141, 2024, 110650, ISSN 0925-9635, <https://doi.org/10.1016/j.diamond.2023.110650>. (<https://www.sciencedirect.com/science/article/pii/S0925963523009755>).
- [20] А.Н.БАЖЕНОВ, А.Ю.ТЕЛЬНОВА. Изотопы и таблица Менделеева: учебное пособие. – СПб., 2024. <https://elib.spbstu.ru/dl/5/tr/2024/tr24-29.pdf/info>. DOI:10.18720/SPBPU/5/tr24-29
- [21] А.Н.БАЖЕНОВ. Интервальные арифметики и прослеживаемость изотопной подписи: учебное пособие. – СПб., 2023. <https://elib.spbstu.ru/dl/5/tr/2023/tr23-167.pdf/info>. DOI:10.18720/SPBPU/5/tr23-167
- [22] Мода интервальной выборки и алгоритм для мультимоды. Методический материал. 22.10.2022
- [23] БАГАЕВ К.А., КОЗЛОВСКИЙ С.С. Применение компьютерного моделирования для калибровки детекторов в водной среде // Журнал «Научно-технические ведомости СПбГПУ. Физико-математические науки» 2011. № 2 (122). с. 106 - 111.
- [24] ТЮХМЕНЕВА Е.А. Исследование процессов переноса в плазме токамака Глобус-М2 с использованием комплексного подхода на основе данных измерений оптических диагностик: науч. док.; специальность: 01.04.08 – физика плазмы. – СПб.: ФГБУН ФТИ им. А. Ф. Иоффе, 2022. – 44 с.

- [25] А.И.ИВАНОВ, А.Н.БАЖЕНОВ. Применение мультиинтервалов для описания спектральных данных. Всероссийская конференция «Неделя науки ФизМех» (1-5 апреля 2024 года). Сборник аннотаций докладов. С.-Петербург. 2024 [https://amcp.spbstu.ru/userfiles/files/Annotations\\_Science\\_Week\\_PhysMech\\_2024.pdf](https://amcp.spbstu.ru/userfiles/files/Annotations_Science_Week_PhysMech_2024.pdf)

# Index

- «расширенная» интервальная оболочка моды, 39
- Gallium Indium Arsenide Phosphide, 28
- Г.Кирхгоф и Р.Бунзен, 16
- алгебраическое вычитание, 11
- алгоритм вычисления значения параметра *noise*, 48
- алгоритм для нахождения моды интервальной выборки и её частоты, 33
- алгоритм для расширения границ интервальной моды, 40
- алгоритм нахождения частот мод мультимодальной интервальной выборки, 37
- атомарные спектры поглощения или испускания, 17, 52
- внешнее оценивание, 10
- внутреннее оценивание, 11
- донные осадки, 18
- дуальный интервал, 11
- изотопологи, 31
- интервал, 9
- интервальная арифметика Каухера, 10
- интервальная оболочка, 39
- интервальное умножение в полной интервальной арифметике, 12
- квazar, 23
- квazar PKS0454+039, 24
- классическая интервальная арифметика, 10
- комптоновское рассеяние гамма-квантов, 48
- космический телескоп Хаббл, 24
- лес Лайман-альфа, 25
- линия Лайман-альфа, 24
- масс-спектрометрия, 28
- масс-спектрометрия с вторичными ионами (Secondary-ion mass spectrometry, SIMS), 29
- медно-цинковые руды, 17
- метод «водораздела» (watershed), 37
- минимаксный подход, 12
- минимум по включению, 12
- мода выборки, 33
- модель вращающегося диска, 26
- модель радиального падения облаков, 25
- мультиинтервал, 12
- мультимодальная выборка, 33
- относительная ширина интервала, 9
- пик аннигиляции электрон-позитронных пар, 48

постоянная Ридберга, 25  
предел Лаймана, 24  
преобразователи солнечной  
    энергии, 28  
приборы с зарядовой связью,  
    ПЗС, CCD, 17  
профиль Фойгта, 26

радиус интервала, 9  
середина интервала, 9  
спектр неба на Земле,  
    «Фраунгоферовы  
    линии», 15  
спектральная серия Лаймана, 24  
спектральный анализ, 14  
спектрограф Кекк/HIRES, 26  
спектрограф со скрещенной  
    дисперсии СТЭ-1, 18  
спектроскопия гамма-квантов, 47  
стабильные изотопы, 29  
стандартный образец РУС-2, 18  
стандартный образец СГХМ-2, 21  
триплет, 43  
унимодальная выборка, 33  
формула Ридберга, 24  
фотошпик, пик полного  
    поглощения, 47  
фотоэлектронные кассеты, ФЭК,  
    17  
функционал Рачека, 10  
частота моды, 33  
электрон-позитронная пара, 48  
элементарные подинтервалы  
    измерений, 33  
эффект «обёртывания», 10  
эффект Ганна—Петерсона, 25