

## Алгоритм аппроксимации базовой линии экспериментальных спектров

Подосенова Татьяна Борисовна, кандидат физико-математических наук,  
старший научный сотрудник, факультет ВМК  
МГУ имени М.В. Ломоносова (г. Москва)

На основе непрерывного вейвлет-преобразования (НВП) спектров с использованием базисных гауссовых вейвлетов рассмотрены задачи поиска и локализации резонансных линий в спектре. Предложен алгоритм аппроксимации базовой компоненты спектра.

**Ключевые слова:** спектр излучения, резонансная линия, базовая линия, локализация линии, гауссов вейвлет, непрерывное вейвлет-преобразование.

1. Исходные данные измерений - экспериментальные спектры излучения - задаются обычно в виде суммы трех компонент - полезной, базовой и шумовой:

$$y = \{y_i = y(x_i) = Y(x_i) + B(x_i) + \varepsilon_i \geq 0, i = 1, \dots, N\}, Y(x) = \sum_j g_j(x).$$

Здесь  $y_i$  - значение интеграла по  $i$ -му каналу измерения,  $N$  - число каналов. Величины  $y_i$  предполагаются случайными и статистически независимыми, распределенными по закону Пуассона или нормальному, т.е.  $E(\varepsilon_i) = 0$ ,  $D(\varepsilon_i) = \sigma^2(\varepsilon_i)$ ,  $\sigma(\varepsilon_i) = \sqrt{\max(1, y_i)}$ . Базовую компоненту  $B(x)$  рассматривают как достаточно гладко меняющуюся функцию. Кроме того, во многих случаях достаточно ограничиться гауссовой моделью формы пиков:  $g_j(x) = A_j g_0(x; \mu_j, \beta_j)$ ,  $A_j \geq 0$  - амплитуды пиков,  $g_0(x; \mu, \beta) = e^{-(x-\mu)^2/(2\beta^2)}$ ,  $\int_{-\infty}^{\infty} g_0(x; \mu, \beta) dx = \sqrt{2\pi} \beta$ ,  $\beta > 0$ . Обработка данных сводится к определению параметров суммы резонансных унимодальных функций - кривой  $Y(x)$ .

2. Одной из значимых при обработке данных является задача распознавания фрагментов локализации резонансных линий в спектре, как синглетов - одиночных линий, - так и мультиплетов - совокупностей плохо разрешенных линий. При поиске границ локализации пиков мы будем использовать метод непрерывного вейвлет-преобразования (НВП) на основе гауссовых вейвлетов [1, с. 93; 1, с. 151].

Метод НВП сводится к вычислению корреляции между фрагментами анализируемого спектра и смещаемыми вдоль оси аргумента спектра растянутыми/сжатými копиями анализирующей базовой вейвлет-функции [2].

3. Базисные гауссовы вейвлеты  $\psi_n(x)$  и  $\varphi_n(x)$ , применяемые в данной работе, основаны на вычислении производных гауссовых функций [1, с. 151]:

$$\psi_n(x) = (-1)^{n+1} \frac{d^n}{dx^n} e^{-x^2/2}, \varphi_n(x) = \lambda_n C_n^{-1} \frac{d^n}{dx^n} e^{-x^2}, C_n = \|\varphi_n(x)\|_{L_2}, n \geq 1. \text{ Здесь } \lambda_n = -1 \text{ для значений } n = 2, 3, \text{ и}$$

$\lambda_n = 1$  для остальных значений  $n \geq 1$ . В работе используются гауссовы вейвлеты 2-го и 4-го порядка:

$$\psi_2(x) = (1 - x^2) e^{-x^2/2}, \psi_4(x) = (-x^4 + 6x^2 - 3) e^{-x^2/2}, \varphi_2(x) = 2C_2^{-1} (1 - 2x^2) e^{-x^2},$$

$$\varphi_4(x) = 4C_4^{-1} (4x^4 - 12x^2 + 3) e^{-x^2}. \text{ Авторами аналитически проверено и доказано, что } C_1 = (\pi/2)^{1/4},$$

$$C_2 = \sqrt{3} \cdot (\pi/2)^{1/4}, C_3 = \sqrt{15} \cdot (\pi/2)^{1/4}, C_4 = \sqrt{105} \cdot (\pi/2)^{1/4} [3].$$

Графики функций  $\psi_n(x)$  и  $\varphi_n(x)$  при четных значениях порядков  $n$  напоминают формой резонансные кривые, а их аналитические представления допускают запись через ортогональные полиномы Эрмита [4, с. 772]:  $\psi_n(x) = -He_n(x) e^{-x^2/2}$ ,  $\varphi_n(x) = C_n^{-1} H_n(x) e^{-x^2}$ . Т.к. функции  $He_n(x)$  и  $H_n(x)$  различны по масштабу:  $H_n(x) = 2^{n/2} He_n(x\sqrt{2})$  [4, с. 584; 5, с. 772], то и вейвлеты  $\varphi_n(x)$  и  $\psi_n(x)$  отличаются масштабом (полушириной):  $\varphi_n(x) = (-1)^{n+1} 2^{n/2} \lambda_n C_n^{-1} \psi_n(x\sqrt{2})$ .

4. При использовании метода НВП коэффициенты вейвлет-разложения спектра записываются в виде свертки:  $W_\psi(a, b; y) = a^{-1/2} \int_{-\infty}^{\infty} y(t) \psi\left(\frac{t-b}{a}\right) dt$  [1, с. 93], где  $a, b$  - параметры масштаба и сдвига. В силу коммутирования операции дифференцирования и НВП [2, с. 1155], для вейвлет-преобразований справедливы соотношения:

$$W_\psi\left(a, b; \frac{d^n}{dx^n} y(x)\right) = (-1)^n \int_{-\infty}^{\infty} y(t) \cdot \frac{d^n}{dx^n} \Psi_{a,b}(t) \cdot dt, \Psi_{a,b}(t) = a^{-1/2} \psi\left(\frac{t-b}{a}\right), \text{ что позволяет перейти от диффе-}$$

ренцирования заданного численно исходного спектра  $y(x)$  к дифференцированию выписываемого в виде аналитической функции вейвлета  $\Psi_{a,b}(t)$ .

В случае, когда исходные данные представляют собой одиночный резонанс:  $y(x) = A g_0(x; \mu, \beta)$ , для базисных гауссовых вейвлетов 2-го и 4-го порядков выражения вейвлет-коэффициентов выписываются аналитически [6]:

$$W_{\psi_n}(a, b; y) = \frac{A\sqrt{2\pi}\beta}{\sqrt{a}} \left( 1 + \left( \frac{\beta}{a} \right)^2 \right)^{-(2[n/2]+1)/2} \cdot \psi_n \left( \frac{b - \mu}{\sqrt{a^2 + \beta^2}} \right), \quad n = 2, 4,$$

$$W_{\varphi_n}(a, b; y) = (-1)^{n+1} 2^{(n-1)/2} \lambda_n C_n^{-1} \cdot W_{\psi_n}(a, \sqrt{2}b; \bar{y}),$$

$$\bar{y}(x) = y(x / \sqrt{2}) = A \cdot g_0(x; \sqrt{2}\mu, \sqrt{2}\beta),$$

$$W_{\varphi_n}(a, \sqrt{2}b; \bar{y}) = \frac{2A\sqrt{\pi}\beta}{\sqrt{a}} \left( 1 + \left( \frac{\sqrt{2}\beta}{a} \right)^2 \right)^{-(2[n/2]+1)/2} \cdot \psi_n \left( \frac{\sqrt{2}(b - \mu)}{\sqrt{a^2 + 2\beta^2}} \right).$$

Из последних соотношений следует, что указанные вейвлет-преобразования сохраняют положения одиночных резонансов на оси аргумента.

Заметим, что крайние нули  $\bar{b}_{1,4}$  вейвлет-коэффициента  $W_{\varphi_4}(a, b; y)$ ,  $\bar{b}_{1,4} \approx \pm 1.6507 \cdot \sqrt{a^2 + 2\beta^2}$ , при параметре масштаба  $a = \beta$  равны  $\bar{b}_{1,4} \approx \pm 1.6507 \cdot \beta \sqrt{3} \approx 2.86\beta$  [6]. Если обозначить  $a = k\beta > 0$ , то решением уравнения  $1.6507 \cdot \sqrt{k^2 + 2} = 3$  будет значение  $k \approx \sqrt{(1.8174)^2 - 2} \approx 1.1415$ , т.е. при выборе масштаба  $a = k\beta$ ,  $k = 1.1415$ , область  $(\mu - 3\beta, \mu + 3\beta)$  локализации синглета  $y(x)$  будет практически совпадать с отрезком  $(\bar{b}_1, \bar{b}_4)$  оси аргумента  $b$ .

5. Вейвлет-образы спектров на базе гауссовых вейвлетов даже в случае зашумленных данных имеют вид достаточно гладких кривых и позволяют анализировать резонансные пики спектра, не учитывая при этом влияние характеризующейся достаточно медленной динамикой своего изменения базовой компоненты исследуемого спектра.

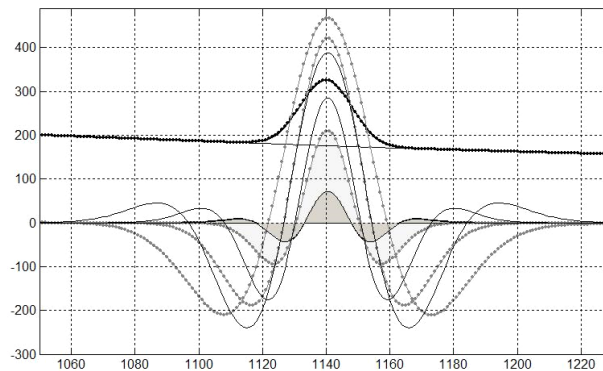


Рис.1. Графики кривых  $Y_B(x)$ ,  $B(x)$  и вейвлет-образов  $W_{\varphi_n}(a, x; Y_B)$  в случае одиночного резонанса,  $a \in \{\beta, 2\beta, 3\beta\}$ ,  $\beta = 8$ ,  $n = 2, 4$

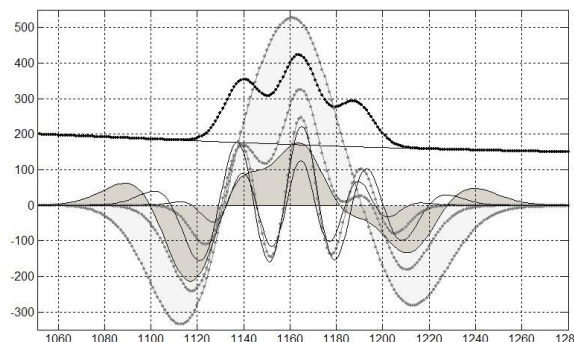


Рис.2. Графики кривых  $Y_B(x)$ ,  $B(x)$  и вейвлет-образов  $W_{\varphi_n}(a, x; Y_B)$  в случае мультиплета (триплета),  $a \in \{\beta, 2\beta, 3\beta\}$ ,  $\beta = 8$ ,  $n = 2, 4$

В силу свойств линейности для вейвлет-преобразований можно записать:  $W_{\psi}(a, b; Y_B) = W_{\psi}(a, b; B) + \sum_j W_{\psi}(a, b; g_j)$ , где  $Y_B(x) = Y(x) + B(x)$ . В области локализации резонансов  $g_j(x)$  значения  $W_{\psi}(a, b; B)$  малы, что объясняется как достаточно гладким поведением базовой компоненты спектра  $B(x)$ , так и наличием нескольких первых нулевых моментов у гауссовых вейвлетов [2].

На рис.1-2 показаны графики спектра  $Y_B(x)$  и базовой линии  $B(x)$ , а также графики вейвлет-коэффициентов  $W_{\varphi_n}(a, x; Y_B)$  для значений масштаба  $a \in \{\beta, 2\beta, 3\beta\}$ ,  $\beta = 8$ , для порядков вейвлета  $n = 2, 4$ . Точками выделены графики вейвлет-коэффициентов 2-го порядка. Рис.1 иллюстрирует случай, когда  $Y(x)$  есть синглет, а рис.2 - случай, когда  $Y(x)$  представляет собой плохо разрешенный мультиплет (триплет). Закрашены графики, соответствующие масштабам (полуширинам):  $a = \beta = 8$  (рис.1),  $a = 3\beta = 24$  (рис.2).

Проведенные численные расчеты показывают, что при расстановке границ локализации мультиплетов в спектре, для восстановления базовой линии спектра, можно использовать масштабы вейвлетов из диапазона  $\beta \leq a \leq 3\beta$ .

6. Область задания  $D = [x_1, x_N]$  исследуемого одномерного спектра излучения  $y(x)$  представим в виде объединения перемежающихся областей  $D_j$  локализации базовой компоненты спектра и областей  $P_j$  локализации резонансных линий:  $D = D_1 \cup P_1 \cup D_2 \cup \dots \cup D_K \cup P_K \cup D_{K+1}$ ,  $j = 1, \dots, K$ . Границы областей  $P_j$  первоначально будем задавать граничными точками фрагментов спектра, на которых вейвлет-коэффициенты и 2-го, и 4-го порядка превышают заданный уровень ошибок данных. Границы локализации резонансных компонент, вычисляемые по аргументам точек локальных минимумов вейвлет-образов, могут не совпадать с границами  $P_j$ .

Заметим, что гауссовы вейвлеты позволяют достаточно надежно выявить структуру спектра, но зачастую при использовании метода НВП требуется предварительное сглаживание обрабатываемых спектров, поскольку при малых масштабах вейвлеты выявляют не только резонансные линии в спектрах, но и случайные флуктуации в данных.

7. Предлагаемый в работе алгоритм разметки границ резонансов в спектре для аппроксимации базовой линии спектра включает в себя несколько этапов. Выпишем основные этапы алгоритма.

1) Сглаживаем исходный спектр фильтром Савицкого-Голея (SG-фильтром) [7; 8; 9], скользящим окном ширины  $2w+1$ ,  $5 \leq w \leq 20$ . Значение  $w > 0$  есть оценка полуширины  $\beta$  одиночного синглета. Сглаженный SG-фильтром спектр обозначим  $\tilde{y}$ , или  $\tilde{y}(x)$ .

2) Вычисляем вейвлет-коэффициенты  $W_{\varphi_n}(a, x; y)$  для входного спектра, при значении масштаба вейвлета  $a = p \cdot w$ ,  $n = 2, 4$ . Величина  $p > 0$  есть параметр алгоритма. Вычисляем аргументы (положения) точек локальных экстремумов графиков полученных кривых  $W_{\varphi_n}(a, x; y)$ .

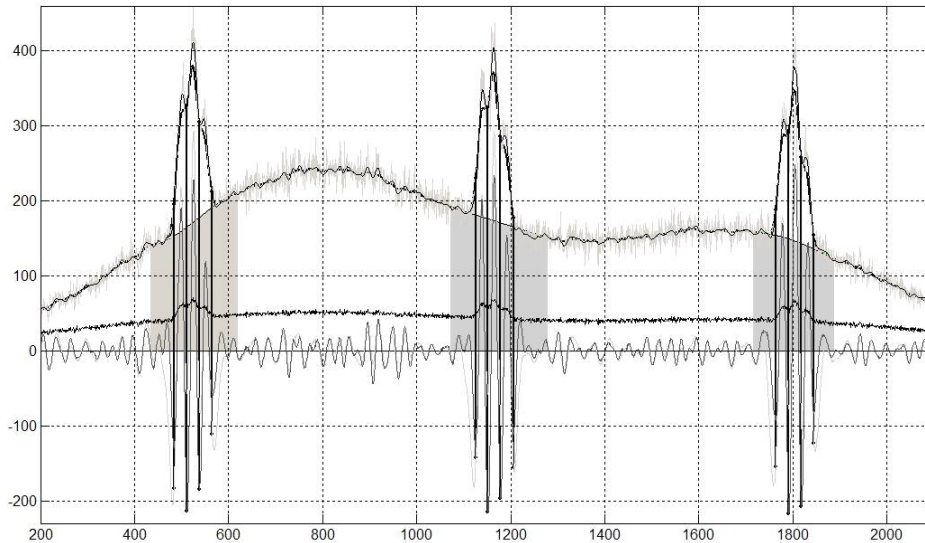
3) Находим границы  $(x_{j_1}, x_{j_2})$  фрагментов  $\Omega_j = \{x_i, i = j_1, \dots, j_2\}$  оси аргумента таких, что для всех точек  $x_i \in \Omega_j$  значения вейвлет-коэффициентов превышают уровень ошибки данных:  $\max(W_{\varphi_2}(a, x_i; y), W_{\varphi_4}(a, x_i; y)) \geq h \cdot \sigma_i$ ,  $\sigma_i = \sigma(\varepsilon_i) = \sqrt{\max(1, y_i)}$ ,  $j = 1, \dots, K_0$ . В алгоритме используется значение  $h = 3.3$ .

4) Используя сглаженный спектр  $\tilde{y}$ , проверяем эти фрагменты спектра на значимость (наличие резонансов), по критерию:  $(\max_{x_i \in \Omega_j} \tilde{y}_i - \min_{x_i \in \Omega_j} \tilde{y}_i) / (\max_{x_i \in \Omega_j} \sigma_i) \geq 1$ . Проверяем критерий мультиплетности спектра на выбран-

ных фрагментах: как правило, если участок спектра представляет собой мультиплет, то локальные минимумы и максимумы графика кривой  $W_{\varphi_4}(a, x; y)$  чередуются между собой.

5) Для каждого значимого фрагмента  $\Omega_j$  ищем ближайшие слева и справа к границам фрагмента значения локальных минимумов вейвлет-образа  $W_{\varphi_4}(a, x; y)$ , а затем, опять-таки расширяя область локализации резонансов, вычисляем ближайшие к найденным значениям аргументы нулей функции  $W_{\varphi_2}(a, x; y)$ . Полученные точки оси аргумента будем считать границами областей  $D_j$  и  $P_j$  локализации базовой компоненты спектра и резонансных линий соответственно.

На фрагментах  $D_j$  будем аппроксимировать базовую линию  $B(x)$ , используя сглаженный спектр  $\tilde{y}(x)$ , а на областях  $P_j$  локализации резонансов будем моделировать  $B(x)$  с помощью кубических сплайнов [9].



**Рис.3. Результаты работы алгоритма обработки квазиреального возмущенного модельного спектра, с использованием вейвлет-образов  $W_{\varphi_n}(a, x; y)$ ,  $n = 2, 4$  и при задании параметров  $a = 1.5\beta$ ,  $\beta = 8$**

Как показывает практика, преобразованные SG-фильтрами спектры  $\tilde{y}(x)$  не обладают достаточной степенью гладкости на областях  $D_j$  локализации базовой компоненты спектра, и потому по спектрам  $\tilde{y}(x)$  сложно смоделировать близкую к истине базовую компоненту спектра. Поэтому на областях  $D_j$  будем использовать регуляризирующие алгоритмы.

8. Известно, что задача сглаживания зашумленной функции  $\tilde{y}(x) \in L_2[c, d]$ ,  $c \leq x \leq d$ , сводится к минимизации функционала Тихонова [10, с. 140 - 141]:  $M^\alpha(f) = \|f - \tilde{y}\|_{L_2}^2 + \alpha \|f\|_{W_2^1}^2$ , где  $f(x) \in W_2^1[c, d]$ ,  $\|f\|_{W_2^1}^2 = \int_c^d \{f^2(t) + \{f'(t)\}^2\} dt$ . У функционала  $M^\alpha(f)$  существует единственный минимум, а функция  $f(x)$ , на которой достигается минимум, удовлетворяет уравнению Эйлера [11, с. 21 - 22; 11, с. 46]:  $f'' - q^2 f = -\theta \tilde{y}$ ,  $\theta = \alpha^{-1}$ ,  $q^2 = 1 + \theta$ . Уравнение Эйлера есть неоднородное линейное обыкновенное дифференциальное уравнение (ОДУ) 2-го порядка  $lf(x) = -\theta \tilde{y}$ , где дифференциальный оператор есть  $lf(x) = f'' - q^2 f$ .

В работе авторов [12] получены аналитические решения краевых задач для указанного уравнения Эйлера для ряда способов задания краевых условий. Численное решение используемой в данной работе краевой задачи, с условиями на значения 2-й производной функции,  $f''(c) = f''(d) = 0$ , имеет вид [13]:

$$\begin{aligned}
 f(x_i) = & \frac{\theta}{q^2 \cdot sh(q\{d-c\})} \times \{ sh(qh\{n-i\}) \cdot \tilde{y}(c) + sh(qhi) \cdot \tilde{y}(d) \} + \\
 & + \frac{2\theta \cdot sh(0.5qh)}{q^2 \cdot sh(q\{d-c\})} \times \{ sh(qh\{n-i\}) \cdot \sum_{k=1}^i \tilde{y}(x_k) \cdot sh(qh\{k-0.5\}) + \\
 & + sh(qhi) \cdot \sum_{k=i+1}^n \tilde{y}(x_k) \cdot sh(qh\{n-k+0.5\}) \} .
 \end{aligned}$$

Функция  $f(x)$  гораздо более гладкая, чем  $\tilde{y}(x)$ , и поэтому значения  $f(c_j)$ ,  $f(d_j)$ ,  $f'(c_j)$ ,  $f'(d_j)$ , где  $c_j, d_j$  - границы, например, областей  $D_j$ , могут быть использованы для моделирования базовой линии спектра кубическими сплайнами на заданных фрагментах локализации полезного сигнала [9].

9. На рис.3 показаны графики исходного и сглаженного спектров, графики восстановленной базовой линии  $\tilde{B}(x)$  и вейвлет-коэффициентов  $W_{\varphi_n}(a, x; y)$ ,  $n = 2, 4$ , для значения масштаба  $a = 1.5\beta$ ,  $\beta = 8$ , а также график линии отсечки  $h \cdot \sigma(x) = h \cdot \sqrt{\max(1, y(x))}$ . В приведенном примере  $\gamma(x)$  представляет собой совокупность трех плохо разрешенных мультиплетов (триплетов). Области  $P_j$  локализации резонансов закрашены цветом.

#### Литература:

1. Дьяконов, В. П. Вейвлеты. От теории к практике / В. П. Дьяконов. - М.: Солон-Р, 2002. - 448 с.

2. Астафьева, Н.М. Вейвлет-анализ: основы теории и примеры применения / Н.М. Астафьева // УФН. - 1966. - Т. 166. - № 11. - С. 1145-1170.
3. Подосенова, Т. Б. Вопросы нормировки при проведении CWT преобразований спектров / Т. Б. Подосенова // Перспективы развития науки и образования: сб. тр. науч.-практич. конф. - М.: АР-Консалт, 2018. С. 16-22.
4. Справочник по специальным функциям с формулами, графиками и таблицами / Под ред. М. Абрамовица, И. Стиган - М.: Наука, 1979. - 832 с.
5. Корн, Г. Справочник по математике. Для научных работников и инженеров / Г. Корн, Т. Корн. - М.: Наука, 1973. - 832 с.
6. Подосенова, Т. Б. О вейвлет-образах плохоразрешенных мультиплетов [Электронный ресурс] / Т. Б. Подосенова // Электронный научный журнал. - 2019. - № 3 (25). - С. 8-15.
7. Savitzky, A. Smoothing and differentiation of data by simplified least squares procedures / A. Savitzky, M.J.E. Golay // Analytical Chemistry. - 1964. - Vol. 36. - No 8. - P. 1627-1639.
8. Подосенова, Т. Б. О сглаживающих SG фильтрах для задач обработки спектров [Электронный ресурс] / Т. Б. Подосенова // Электронный научный журнал. - 2019. - № 3 (25). - С. 16-24.
9. Подосенова, Т. Б. Комбинированный алгоритм подавления шума в спектрах излучения [Электронный ресурс] / Т. Б. Подосенова // Электронный научный журнал. - 2020. - № 4 (33). - С. 12-17.
10. Тихонов, А. Н. Методы решения некорректных задач: учеб. пособие для вузов / А. Н. Тихонов, В. Я. Арсенин. - М.: Наука. Гл. ред. физ.-мат. лит., 1986. - 288 с.
11. Гельфанд, И. М. Вариационное исчисление: Учебник для ун-тов / И. М. Гельфанд, С. В. Фомин. - М.: Гос. изд-во физ.-мат. лит., 1961. - 228 с.
12. Подосенова, Т. Б. Аналитическое решение уравнения Эйлера для задачи сглаживания экспериментальных данных / Т. Б. Подосенова // Science and education: problems and innovations: сб. ст. III Междунар. науч.-практич. конф. - Пенза: МЦНС «Наука и Просвещение», 2020. - С. 19-24.
13. Подосенова, Т. Б. Численное решение уравнения Эйлера для задачи сглаживания экспериментальных данных / Т. Б. Подосенова // Science and education: problems and innovations: сб. ст. III Междунар. науч.-практич. конф. - Пенза: МЦНС «Наука и Просвещение», 2020. - С. 14-18.