

Приборостроение, метрология и информационно-измерительные приборы и системы

УДК 681.391:543/545

СГЛАЖИВАНИЕ ХРОМАТОГРАФИЧЕСКИХ СИГНАЛОВ ПУТЕМ ИХ АППРОКСИМАЦИИ В БАЗИСЕ ФУНКЦИЙ ЧЕБЫШЕВА – ЭРМИТА

*А.В. Бочкарев, Р.Т. Сайфуллин**

Самарский государственный технический университет
443100, г. Самара, ул. Молодогвардейская, 244

E-mail: bochkarevpubliciit@gmail.com

Аннотация. Целью настоящей работы является оценка применимости аппроксимационного алгоритма сглаживания хроматографических сигналов в базисе функций Чебышева – Эрмита. Для аппроксимации используется алгоритм, подразумевающий вычисление коэффициентов сдвига и масштаба для функций базиса, а также деление сигнала на фрагменты. Рассматривается погрешность сглаживания модельного сигнала при его аппроксимации в выбранном базисе в сравнении с цифровым фильтром скользящего среднего. Также исследуется погрешность сглаживания производной сигнала и положения на производной экстремумов, производится сравнение, как и в прошлом случае, с цифровым фильтром скользящего среднего. В рамках экспериментального исследования в работе подвергаются обработке девять наиболее характерных хроматографических пиков, извлеченных из реальных хроматограмм. Для оценки качества сглаживания сравниваются среднеквадратичные отклонения шума и остатка, а также определяется закон распределения остатка. По результатам исследования устанавливаются ограничения по максимально допустимому значению коэффициента асимметрии обрабатываемых хроматографических пиков. Благодаря использованию выбранного аппроксимационного подхода удается решить задачу сглаживания хроматографических сигналов без ослабления полезной составляющей. Для вычислений и графического представления результатов моделирования использована система компьютерной алгебры Wolfram Mathematica 11.3.

Ключевые слова: функции Чебышева – Эрмита, сглаживание, случайная помеха, СКО, хроматографический сигнал, преобразование сигналов, аппроксимационный метод, хроматография.

* Андрей Владимирович Бочкарев, аспирант, старший преподаватель кафедры «Информационно-измерительная техника».

Раухат Талгатович Сайфуллин, профессор кафедры «Информационно-измерительная техника», доктор технических наук, профессор.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-31-90014.

Введение

Одной из задач обработки хроматографических сигналов является сглаживание или подавление помехи. На практике сглаживание чаще всего реализуют различными цифровыми фильтрами. Проблемой цифровой фильтрации является тот факт, что частотный спектр пиков хроматограммы, как правило, частично совпадает со спектром помехи и вместе с подавлением помехи происходит ослабление полезной составляющей, приходится либо снижать сглаживающие свойства фильтра, либо пытаться восстановить потерянную при сглаживании информацию [1]. Фильтр Савицкого – Голея решает проблему ослабления полезной составляющей, но осуществляет сглаживание менее эффективно, причем степень сглаживания зависит от подхода к расчету параметров фильтра [2]. Фильтр Калмана имеет лучшие сглаживающие свойства, чем фильтр Савицкого – Голея, но подразумевает существенную зависимость от формы шума, выходной сигнал фильтра может иметь фазовые и прочие нелинейные искажения [3, 4]. Сплайн-фильтры эффективно подавляют помеху с учетом гладкости полезного сигнала и при сравнении превосходят фильтр Калмана, но имеют значительную зависимость погрешности от метода расчета коэффициентов сплайна [4–6]. Фильтр Винера также эффективно сглаживает пики хроматограммы, но вызывает осцилляции, обусловленные краевыми эффектами; трудностью являются одновременно хорошее сглаживание и низкие осцилляции [3]. Таким образом, актуальным является вопрос разработки нового математического метода сглаживания хроматографических сигналов.

В настоящей работе для создания такого метода предлагается воспользоваться аппроксимацией сигналов в базисе функций Чебышева – Эрмита (за рубежом просто Эрмита). При этом из коэффициентов аппроксимации в данном базисе возможно восстановить по другим, предварительно рассчитанным базисам не только сам сигнал [7], но и его производную различных порядков [7, 8], вейвлет-преобразование [9, 10] и т. п. Функции Чебышева – Эрмита находят широкое применение в науке и технике. Разложение по данным функциям используется для аппроксимации двумерных [11] и одномерных [12] сигналов и аналитически заданных функций, обработки файлов мультимедиа форматов, в частности для выделения границ зашумленного изображения [13], биометрической идентификации по изображению радужки глаза человека [14], фильтрации и поиска изображения по базе данных, индексации говорящего по записи голоса [15], извлечения полезной медицинской информации из изображений компьютерной томографии [16] и распознавания речи [17]. Помимо этого, данные функции используются для построения банка фильтров [18], диагностики состояния оборудования [19] и здоровья человека [20], обнаружения полезных составляющих в сигнале [21], а также для решения дифференциальных уравнений [22, 23].

Аппроксимация сигналов в базисе функций Чебышева – Эрмита

Функции Чебышева – Эрмита определяются следующим образом:

$$\varphi_n(x) = \frac{1}{\alpha_n} \cdot e^{-\frac{x^2}{2}} \cdot H_n(x), \quad (1)$$

где $\alpha_n = \sqrt{2^n n! \sqrt{\pi}}$ – нормирующая константа; $H_n(x)$ – полином Эрмита n -го порядка;

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}. \quad (2)$$

Набор функций $\varphi_n(x)$ (1) задает базис для кодирования и декодирования сигналов. Представление сигнала $f(x)$ с помощью $N+1$ функций разложения (алгоритм кодирования) заключается в нахождении коэффициентов c_n . В силу локализации функций Чебышева – Эрмита нахождение данных коэффициентов можно свести к формуле:

$$c_n = \int_{-\tau_n}^{\tau_n} f(x) \cdot \varphi_n(x) dx, \quad n = 0, 1, \dots, N, \quad (3)$$

где $[-\tau_n, \tau_n]$ – отрезок, на котором локализована функция $\varphi_n(x)$.

Алгоритм декодирования с помощью $N+1$ коэффициентов, в результате которого восстанавливается аппроксимированный сигнал, выражается формулой:

$$\hat{f}(x) = \sum_{n=0}^N c_n \cdot \varphi_n(x). \quad (4)$$

Описанный алгоритм кодирования-декодирования обладает сглаживающим свойством [18] и позволяет получить не только сглаженный сигнал, но также и его сглаженную производную k порядка при наличии соответствующего базиса. Формируя различные базисы перехода от коэффициентов разложения (4) к различным преобразованиям, можно получить универсальный аппарат для быстрого получения также коэффициентов непрерывного вейвлет-преобразования, устранения аппаратной функции прибора и т. д. Для восстановления необходимого преобразования сигнала требуется заменить в (4) $\varphi_n(x)$ на соответствующее преобразование данной функции и вычислить полученную формулу относительно рассчитанных для исходного сигнала коэффициентов разложения (4) [7].

Одной из отличительных черт функций Чебышева – Эрмита является их локализация на временной оси. Данное обстоятельство может создать трудности при кодировании сигналов большой длительности – даже при простой их конфигурации (6 и менее точек перегиба) для кодирования в рассматриваемом базисе потребуются функции высоких порядков, вычисление которых затруднительно; трудность вычисления только нарастает с ростом порядка функции.

Очевидным решением является масштабирование функций Чебышева – Эрмита на некоторый коэффициент γ . Помимо этого наилучших результатов кодирования можно достичь, если положение максимума в сигнале будет совпадать с началом координат базисных функций, что обусловлено наибольшим перекрытием областей локализации исследуемого сигнала и базисных функций. Следовательно, требуется ввести и коэффициент сдвига для данных функций, в настоящей работе обозначенный как x_0 . Тем самым формируется новый базис функций

$\varphi_n\left(\frac{x-x_0}{\gamma}\right)$. Для восстановления нормировки базиса был рассчитан коэффициент

$\sqrt{1/\gamma}$, на который умножается и результат кодирования, и результат декодирования. Тем самым можно вынести множитель $1/\gamma$ в выражение для вычисления коэффициентов кодирования, которое при этом следует записать:

$$c_n = \frac{1}{\gamma} \int_{-\tau_n}^{\tau_n} f(x) \varphi_n \left(\frac{x-x_0}{\gamma} \right) dx. \quad (5)$$

С учетом описанных масштабных коэффициентов был разработан алгоритм аппроксимации сигналов в базисе функций Чебышева – Эрмита (рис. 1). Помимо вычисления коэффициентов масштаба и сдвига в ходе данного алгоритма происходит деление сигнала на фрагменты, что снижает погрешность аппроксимации [24].

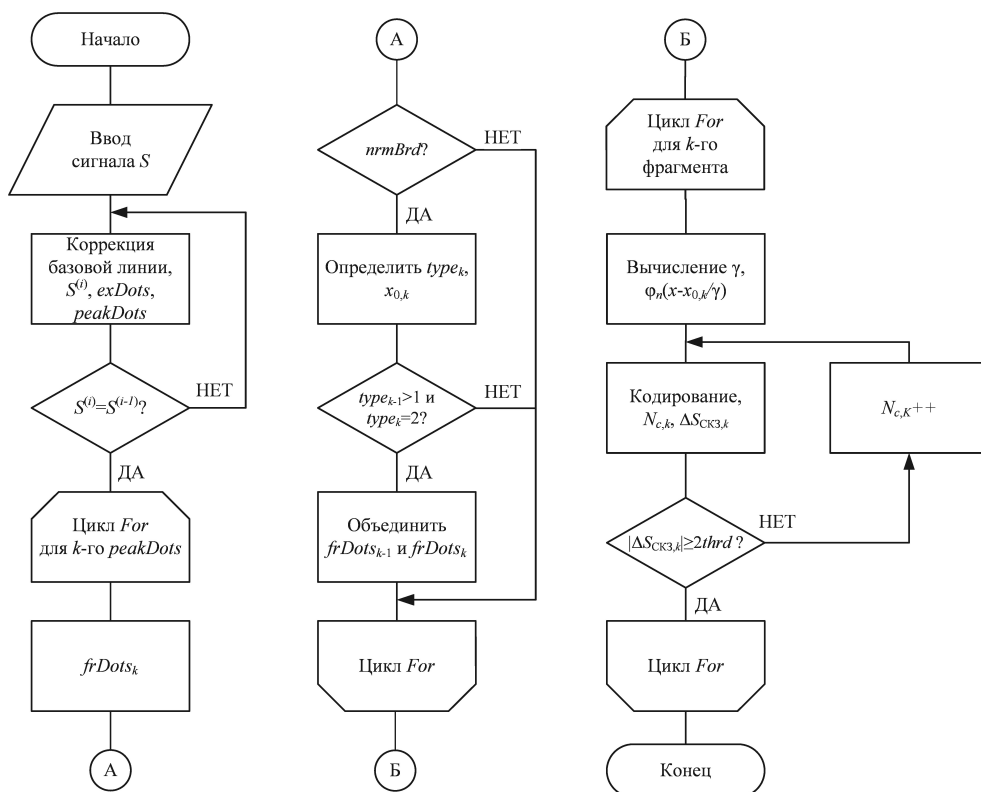


Рис. 1. Алгоритм аппроксимации

Для решения различных задач первичной обработки в хроматографии используются сглаженные производные хроматографических сигналов. Как было сказано выше, при аппроксимации сигналов в базисе функций Чебышева – Эрмита можно аппроксимировать также преобразования этого сигнала; для получения производной требуется вычислить производную функций Чебышева – Эрмита, получив выражение

$$\frac{d^k \varphi_n \left(\frac{x-x_0}{\gamma} \right)}{dx^k} = \frac{1}{\gamma^k} \sum_{m=0}^{\min(k,n)} \left[C_k^m (-1)^{k-m} \sqrt{\frac{2^m n!}{(n-m)!}} He_{k-m} \left(\frac{x-x_0}{\gamma} \right) \varphi_{n-m} \left(\frac{x-x_0}{\gamma} \right) \right], \quad (6)$$

которое задает базис производной k порядка n -й функции Чебышева – Эрмита с учетом масштаба γ и сдвига x_0 и может быть использовано для восстановления k -й сглаженной производной исходного сигнала по рассчитанным заранее

коэффициентам разложения в базисе Чебышева – Эрмита, если вынести за знак суммы нормировочный коэффициент:

$$\frac{d^k \hat{f}(x)}{dx^k} = \frac{1}{\gamma^k} \sum_{n=0}^N c_n \cdot \frac{d^k \varphi_n \left(\frac{x-x_0}{\gamma} \right)}{d \left(\frac{x-x_0}{\gamma} \right)^k}, \quad (7)$$

где $\frac{d^k \hat{f}(x)}{dx^k}$ – оценка k -й сглаженной производной сигнала $f(x)$; c_n – коэффициенты кодирования (5) сигнала $f(x)$ в базисе функций Чебышева – Эрмита.

Подавление случайной помехи

Аппроксимация сигналов в базисе функций Чебышева – Эрмита обладает сглаживающими свойствами, что позволяет устранять случайную помеху. На рис. 2 представлен пример сглаживания реального хроматографического сигнала с предварительным его делением на фрагменты по алгоритму рис. 1. Начальный и конечный фрагменты классифицируются как шум, поэтому на рис. 2 данные области обнулены.

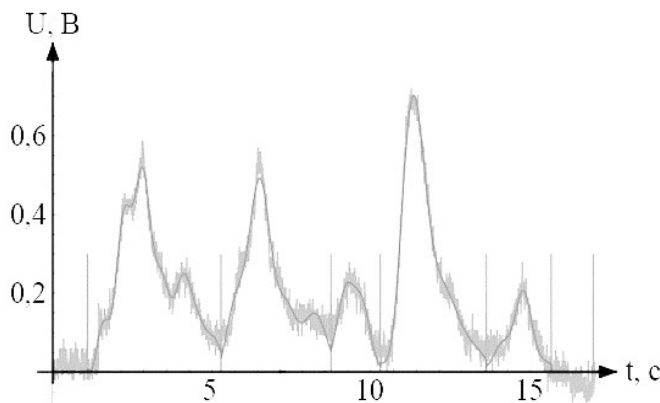


Рис. 2. Пример сглаживания хроматографического сигнала

Для оценки эффективности устранения помехи относительно описанных в первой главе методов было произведено сравнение результата декодирования сглаженного сигнала по коэффициентам базиса функций Чебышева – Эрмита со сглаженным сигналом, полученным при использовании наиболее распространенного на практике метода сглаживания – фильтра скользящего среднего:

$$L_k(N) = \frac{1}{N} \sum_{k=-0,5(N-1)}^{0,5(N-1)} \varsigma(i, k) S_{i-k}, \quad (8)$$

где

$$\varsigma(i, k) = \begin{cases} 1, & \text{если } (i - k > 0) \vee (i - k \leq M), \\ \text{иначе } 0, \end{cases} \quad (9)$$

N – порядок фильтра (только нечетное число); Δx – интервал дискретизации сигнала.

При исследованиях был задан сигнал S , полученный дискретизацией пика, заданного выражением

$$f(t) = Ae^{-\frac{(t-\mu)^2}{2\sigma^2}}, \quad (10)$$

где A – амплитуда пика; μ – время удерживания пика; σ – среднеквадратичная ширина пика [25].

На полученный сигнал накладывалась сгенерированная псевдослучайная помеха с нормальным законом распределения и различным СКО помехи σ . При разных σ достигалось разное отношение сигнал/шум (ОСШ):

$$\text{ОСШ} = 20 \lg \left(\frac{S_{\text{СКЗ}}}{n_{\text{СКЗ}}} \right), \quad (11)$$

где $S_{\text{СКЗ}}$ – среднеквадратическое значение сигнала; $n_{\text{СКЗ}}$ – среднеквадратическое значение помехи.

Оценка погрешности сглаженного сигнала производилась по выражению

$$\delta = \max \left(\frac{\hat{S}_i - S_i}{\max(S_i)} \right) \cdot 100\%, \quad (12)$$

где \hat{S} – сглаженный сигнал; S – исходный сигнал до наложения помехи; $\hat{S}_i - S_i$ – разность i -х значений сглаженного и исходного сигналов.

На рис. 2 данное сравнение представлено в виде графиков, где a – сглаживание с применением цифрового фильтра N порядка, b – путем восстановления сигнала по коэффициентам кодирования N порядка (максимальный порядок функции) в базисе функций Чебышева – Эрмита.

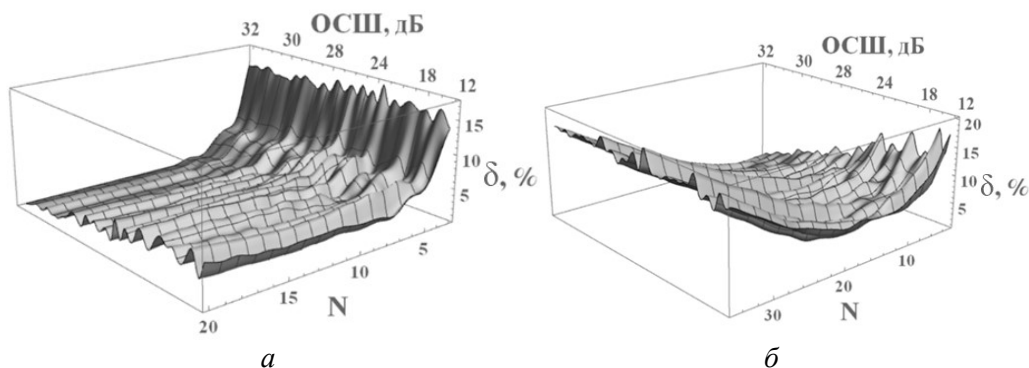


Рис. 3. Сравнение сглаживания с цифровым фильтром

По рис. 3, b можно видеть, что при больших значениях ОСШ цифровой фильтр малого порядка обеспечивает низкую погрешность ($\delta < 1\%$), но со снижением ОСШ повышение порядка не дает снижения погрешности, поскольку, как было сказано в первой главе, вместе с помехой происходит ослабление полезной составляющей. Для сглаживания в базисе функций Чебышева – Эрмита видна иная картина – по достижении минимально необходимого порядка (в данном случае – $N=8$) обеспечивается $\delta < 1\%$ и со снижением ОСШ погрешность растет значительно медленнее, чем при использовании цифрового фильтра.

Поскольку помимо самих сигналов сглаживание требуется (и в большей степени) при вычислении их производных, для того же S была оценена погрешность вычисления производных первого и второго порядка относительно численно определенной производной [26] по сглаженному цифровым фильтром сигналу. Для оценки погрешности использовалось выражение (12) с заменой сигнала на его производную (как в случае исходного, так и в случае сглаженного сигнала). Результаты представлены на рис. 4, где a и b – соответственно первая и вторая производные, полученные численно по сглаженному цифровому сигналу, \bar{b} и \bar{c} – восстановленные по базису (7) сглаженные производные.

По данному рисунку можно видеть, что применение цифрового фильтра не позволяет обеспечить погрешность $\delta < 50\%$ для обеих производных даже при больших значениях ОСШ, а с его уменьшением ситуация меняется только в сторону ухудшения. Увеличение порядка фильтра не дает существенных результатов, позволяя вернуться лишь к начальному значению погрешности $\delta \approx 50\%$. Применение сглаживания в базисе функций Чебышева – Эрмита дает иные результаты: при достижении минимально необходимого порядка (в данном случае – $N=8$) обеспечивается $\delta < 1\%$ для обеих производных, но при снижении ОСШ погрешность возрастает до $\delta \approx 10\%$.

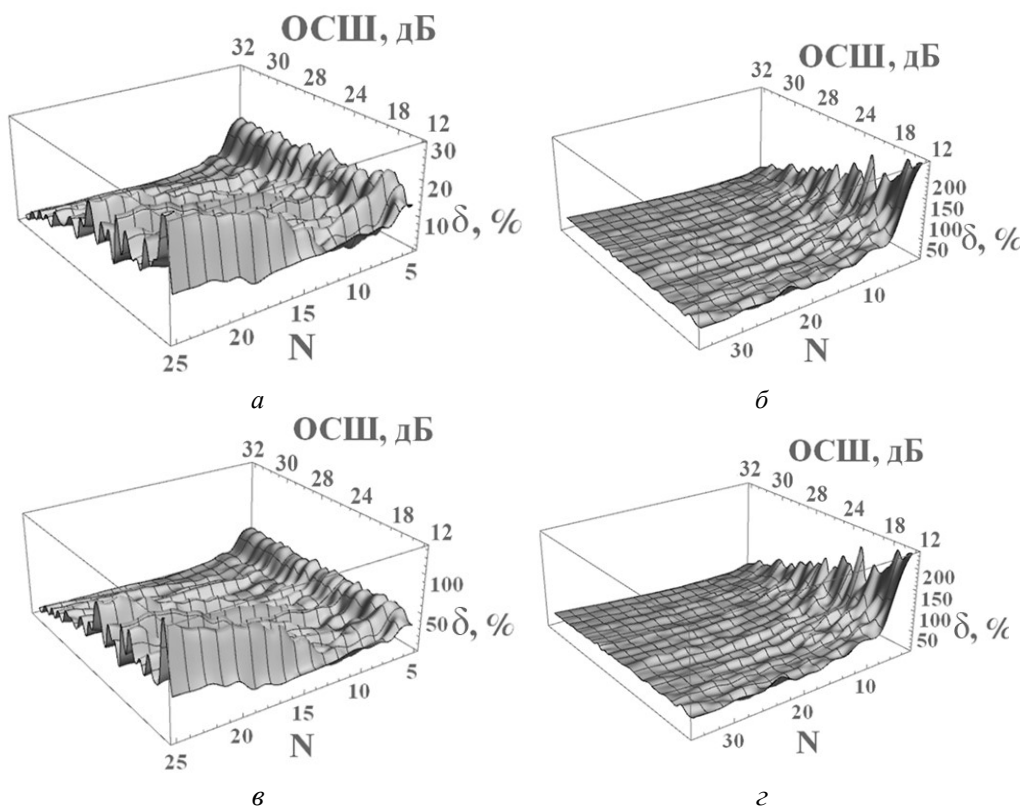


Рис. 4. Сравнение погрешности восстановления производной

Тем не менее для производной имеет значение не только погрешность аппроксимации (см. рис. 4), но и погрешность положения экстремумов, по которым производится определение положения вершины и границ пиков. Поэтому ниже рассматривается погрешность восстановления экстремумов

производных первых двух порядков с применением цифрового фильтра и базиса функций Чебышева – Эрмита, для чего используется выражение

$$\delta = \frac{\left| ex\left[\hat{S}^{(n)}\right] - ex\left[S^{(n)}\right] \right|}{ex\left[S^{(n)}\right]} \cdot 100\%,$$

где $ex[\cdot]$ – функция для вычисления положения экстремума; $\hat{S}^{(n)}$ – восстановленная по функциям Чебышева – Эрмита производная n порядка; $S^{(n)}$ – полученная при цифровой фильтрации производная n порядка.

В качестве экстремумов рассматриваются наиболее применяемые на практике точки: положение прохода через ноль для первой производной и положение минимума второй производной [1]. Результаты представлены на рис. 5.

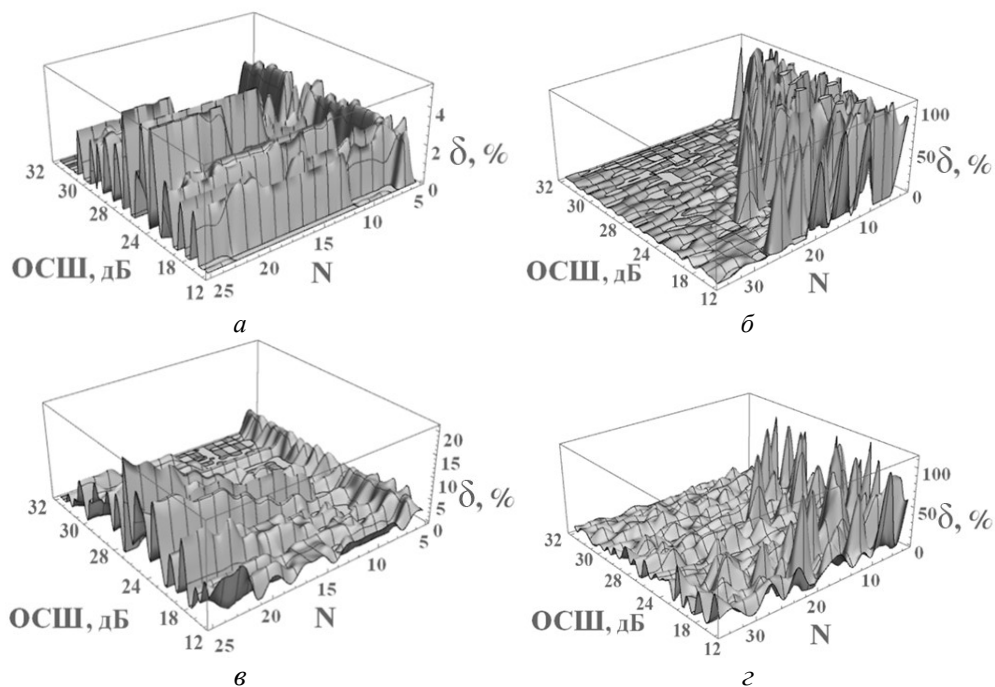


Рис. 5. Сравнение погрешности восстановления экстремумов производных

Можно видеть, что применение цифровой фильтрации позволяет сохранить положение экстремумов при достижении определенного N (при котором согласно рис. 5 возрастает погрешность аппроксимации производной) с $\delta < 5\%$ для первой производной и с $\delta < 10\%$ для второй. Восстановление сглаженной производной в базисе функций Чебышева – Эрмита при достижении $N=8$ обеспечивает для большинства значений ОСШ близкую к нулю погрешность, которая на малых его значениях растет до $\delta < 3\%$ для первой производной и

$\delta < 5\%$ для второй. Таким образом, подавление помехи в используемом базисе обеспечивает снижение погрешности относительно традиционного метода.

Погрешность восстановления одиночных пиков

Для оценки погрешности представленного метода были аппроксимированы пики реальных хроматограмм. Хроматограммы были извлечены из программы Хроматэк Аналитик 3.1. Все они были получены на хроматографе Кристалл-5000 с пламенно-ионизационным детектором.

Из полученных хроматограмм были извлечены 9 наиболее характерных пиков, конфигурация которых наиболее часто встречается в реальных хроматограммах (рис. 6). Многие пики на практике близки по форме к гауссиане, но с небольшой разницей в крутизне фронтов; при выборе пиков для обработки были собраны наиболее часто встречаемые асимметричные конфигурации, которые, как правило, являются одиночными. Для оценки зависимости погрешности отклонения пика от симметричной гауссианы (как наиболее простой для обработки модели) был использован метод моментов, на основе которого вычислялись коэффициенты асимметрии γ_1 и эксцесса γ_2 [27]:

$$m_0 = \Delta x \sum_{i=1}^M S_i, \quad m_1 = \frac{\Delta x^2}{m_0} \sum_{i=1}^M S_i (i-1), \quad m_n = \frac{\Delta x}{m_0} \sum_{i=1}^M S_i [(i-1)\Delta x - m_1]^n, \quad n > 1,$$

$$\gamma_1 = \frac{m_3}{m_2^{3/2}}, \quad \gamma_2 = \frac{m_4}{m_2^2} - 3.$$

Для оценки эффективности сглаживания проведена проверка закона распределения шума остатка, являющегося разностью исходного и аппроксимированного в базисе функций Чебышева – Эрмита сигналов (при отсутствии существенных отклонений аппроксимированного сигнала от действительного закон распределения остатка должен совпадать с законом распределения помехи и быть нормальным). Эффективность сглаживания также оценена численно посредством сравнения СКО шума в пустой области исходного сигнала и СКО озвученного остатка, а также по закону распределения остатка.

Результаты приведены в таблице. Закон распределения помехи в пустых областях исходных сигналов совпадает с законом распределения остатка. Математическое ожидание равно нулю во всех случаях, поэтому в таблице не приводится. Почти для всех рассматриваемых сигналов помеха описывается нормальным законом, для других – законом Коши или логистическим законом, которые отличны от нормального плавностью фронтов, но приближенно являются сопоставимыми с нормальным. Гистограммы распределения помехи для некоторых сигналов приведены на рис. 7.

Разделение симметричных пиков

Пик	γ_1	γ_2	СКО шума сигнала	СКО остатка	Закон распределения шума сигнала
№ 1	0,10	0,42	$1,72 \cdot 10^{-2}$	$1,47 \cdot 10^{-2}$	Нормальный
№ 2	-0,69	-3,59	$1,54 \cdot 10^{-2}$	$1,76 \cdot 10^{-2}$	Нормальный
№ 3	0,79	1,02	$1,50 \cdot 10^{-2}$	$1,59 \cdot 10^{-2}$	Нормальный
№ 4	0,96	0,41	$1,97 \cdot 10^{-2}$	$1,57 \cdot 10^{-2}$	Нормальный
№ 5	0,97	0,78	$1,71 \cdot 10^{-2}$	$1,44 \cdot 10^{-2}$	Нормальный
№ 6	1,34	3,83	$0,91 \cdot 10^{-1}$	$1,14 \cdot 10^{-1}$	Нормальный
№ 7	1,53	3,38	$1,39 \cdot 10^{-2}$	$1,79 \cdot 10^{-2}$	Нормальный
№ 8	1,86	3,85	$1,55 \cdot 10^{-2}$	3,18	Коши
№ 9	2,37	8,78	$4,70 \cdot 10^{-2}$	$6,35 \cdot 10^{-1}$	Логистический

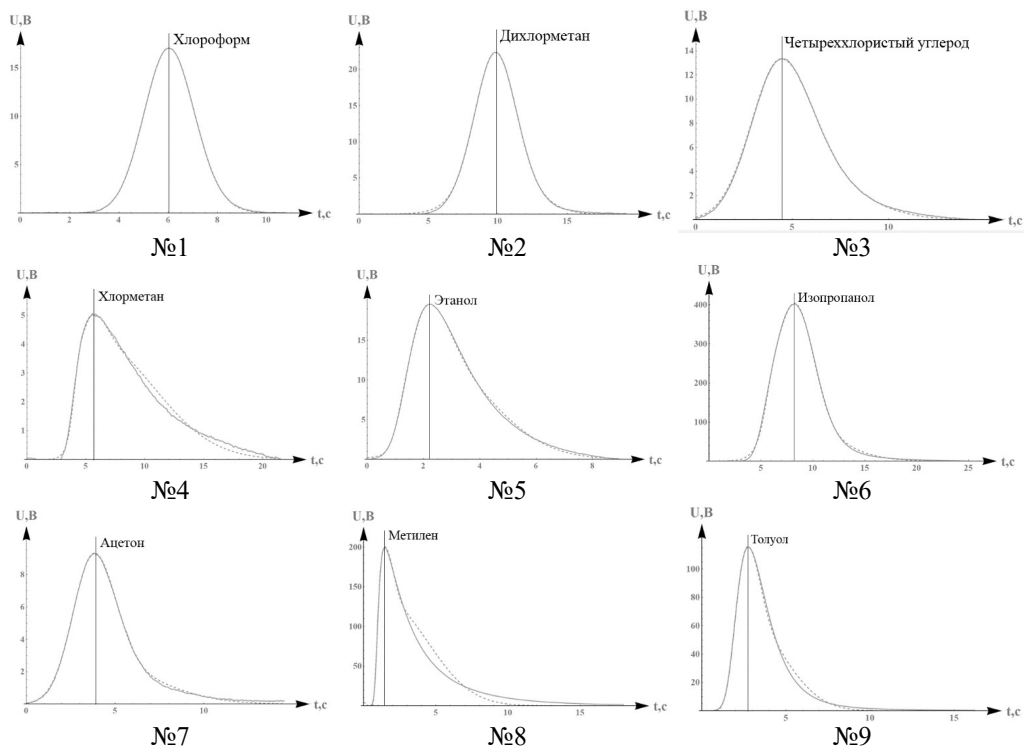


Рис. 6. Исследуемые реальные пики

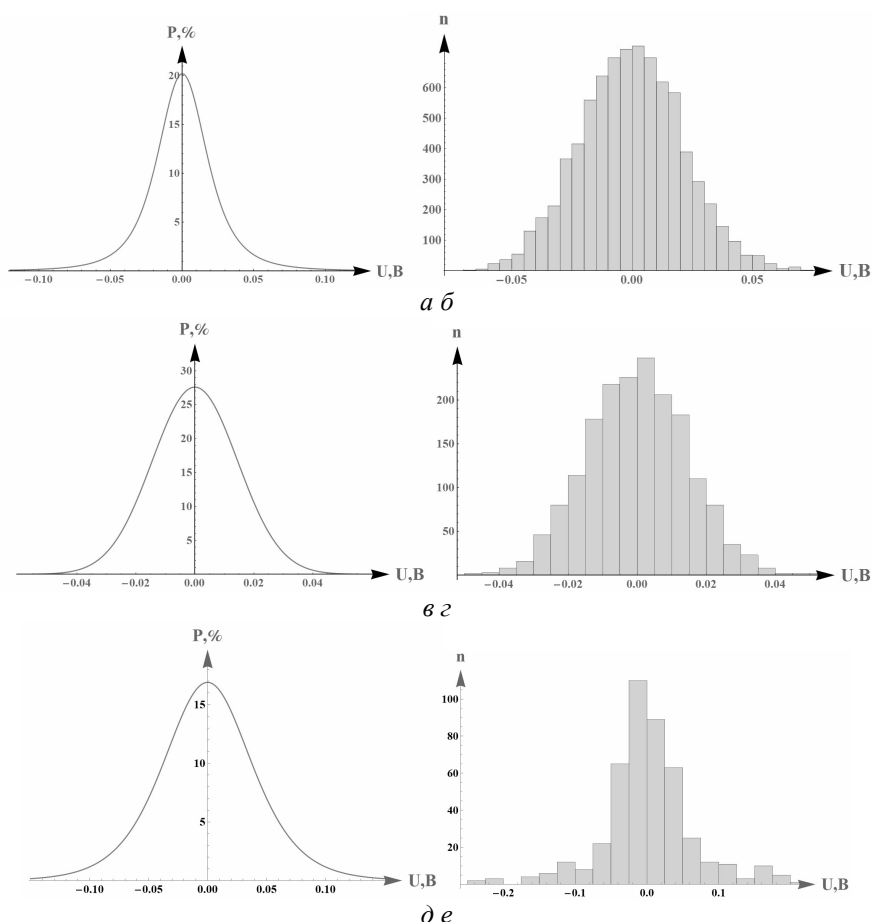


Рис. 7. Гистограммы распределения помехи: а, б – для пика № 8, распределение Коши; в, г – для остатка хроматограммы, нормальное распределение; д, е – для пика № 9, логистическое распределение

Выводы

Анализируя результаты, приведенные в таблице, можно видеть, что СКО шума совпадает порядком с СКО остатка во всех случаях, кроме пиков 8 и 9. Поскольку согласно рис. 6 помеха имеет существенно меньшую амплитуду по сравнению с полезным сигналом, можно положить, что совпадение порядков СКО шума и остатка говорит о низкой погрешности сглаживания. Следовательно, конфигурации пиков, подобных номерам 8 и 9, не могут быть аппроксимированы с достаточно низкой погрешностью, что является ограничением представленного метода. Соответственно предельным значением асимметрии для аппроксимированных одиночных пиков можно считать наибольшую асимметрию в таблице, не считая параметров пиков 9 и 10, $\gamma_{1\max} = 1,53$, а влияния γ_2 на качество аппроксимации не выявлено. С учетом сравнения качества сглаживания сигнала при использовании предложенного аппроксимационного метода и цифрового фильтра можно заключить, что аппроксимационный подход позволяет совместить высокую степень сглаживания и низкую погрешность.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Felinger A.* Data Analysis and Signal Processing in Chromatography. Amsterdam, Netherlands: Elsevier Science B.V., 1998. 413 p.
2. *Savitzky A., Golay M.J.E.* Smoothing and Differentiation of Data by Simplified Least Squares Procedures // *Anal. Chem.* 1964. Vol. 36. P. 1627–1639.
3. *Kalman R.E.* A New Approach to linear Filtering and Prediction Problems // *Trans ASME, Series D, J. BasicEng.* 1960. Vol. 82. P. 35–45.
4. *Гуторов А.С.* Математическое моделирование и исследование алгоритмов фильтрации при траекторной обработке данных по целям // *Автоматизация процессов управления.* 2015. № 1 (39). С. 34–40.
5. *Ланге П.К.* Сплайн-аппроксимация дискретных значений сигналов с применением методов цифровой фильтрации // *Вестн. Самар. госуд. технич. ун-та. Серия: Физ.-матем. науки.* 2003. Вып. 18. С. 87–93.
6. *Спиридонов А.В., Тиме И.В.* Применение сглаживающих сплайнов для фильтрации сильно зашумленных сигналов // *Автомат. и телемех.* 1998. Вып. 7. С. 75–82.
7. *Сайфуллин Р.Т., Бочкарев А.В.* Использование функций Чебышева – Эрмита в обработке сигналов аналитических приборов // *Вестн. Самар. госуд. технич. ун-та. Серия: Технич. науки.* 2019. 1(61). С. 68–81.
8. *Сайфуллин Р.Т., Бочкарев А.В.* Вычисление производных аналитического сигнала в базисе функций Чебышева – Эрмита // *Материалы XI Всеросс. научн. конф. с междунар. участ. «Матем. моделир. и краевые задачи», 27-30 мая 2019. Самара, 2019. Т. 2.* С. 137–139.
9. *Сайфуллин Р.Т., Бочкарев А.В.* Вычисление непрерывного вейвлет-преобразования сигналов в базисе функций Чебышева – Эрмита // *Вестн. Самар. госуд. технич. ун-та. Серия: Технич. науки.* 2019. 2(62). С. 99–113.
10. *Сайфуллин Р.Т., Бочкарев А.В.* Алгоритм вычисления коэффициентов вейвлет-преобразования сигналов с использованием базиса функций Чебышева – Эрмита // *Вестн. Самар. госуд. технич. ун-та. Серия: Технич. науки.* 2019. 4(64). С. 113–124.
11. *Beliczynski B.* A method of multivariable Hermite basis function approximation // *Neurocomputing.* 2012. Vol. 96. P. 12–18.
12. *Jung S.-M.* Approximation of analytic functions by Hermite functions // *Bulletin des Sciences Mathématiques.* 2009. Vol. 133. Issue 7. P. 756–764.
13. *Наджафи М.М.* Проекционный метод решения некоторых задач обработки и анализа изображений: специальность 05.13.18 «Математическое моделирование, численные методы и комплексы программ»: дисс. ... канд. физ.-мат. наук. Московский государственный университет им. М.В. Ломоносова. М., 2004. 110 с.
14. *Павельева Е.А.* Анализ и восстановление изображений проекционными методами, использующими функции Эрмита: специальность 05.13.18 «Математическое моделирование, численные методы и комплексы программ»: дисс. ... канд. физ.-мат. наук. Московский государственный университет им. М.В. Ломоносова. М., 2015. 130 с.
15. *Корчагин Д.Н.* Быстрый проекционный метод обработки мультимедиа информации: специальность 05.13.18 «Математическое моделирование, численные методы и комплексы программ»: дисс. ... канд. физ.-мат. наук. Московский государственный университет им. М.В. Ломоносова. М., 2005. 180 с.
16. *Мамаев Н.В., Лукин А.С., Юрин Д.В., Глазкова М.А., Сеницин В.Е.* Алгоритм нелокального среднего на основе разложения по функциям Эрмита в задачах компьютерной томографии // *ГРАФИКОН'2013. Сб. трудов 23-й междунар. конф. по компьютерной графике и зрению, 16-20 сентября 2013. Владивосток, 2013. С. 254–258.*
17. *Баяковский Ю.М., Жирков А.О., Корчагин Д.Н., Крылов А.С., Лукин А.С.* Нейросетевой анализ и сопоставление частотно-временных векторов на основе краткосрочного спектрального представления и адаптивного преобразования Эрмита // *Препринты ИПМ им. М.В. Келдыша.* 087. 2001.
18. *Балакин Д.А., Штыков В.В.* Построение ортогонального банка фильтров на основе преобразований Эрмита для обработки сигналов // *Журнал радиоэлектроники.* 2014. № 9. С. 1–15.
19. *Балакин Д.А., Штыков В.В.* Диагностика циклических систем с помощью алгоритма, основанного на функциях Гаусса – Эрмита // *Цифровая обработка сигналов.* 2018. № 2. С. 59–62.

20. Балакин Д.А., Штыков В.В., Чуркин С.С. Использование фазометрического комплекса диапазона КВЧ для диагностики жизнедеятельности человека // Электроника и микроэлектроника СВЧ. 2017. Т. 1. С. 127–132.
21. Балакин Д.А., Чуркин С.С., Штыков В.В. Обнаружение сигналов с использованием ортогонального базиса функций Гаусса – Эрмита // Инфокоммуникационные и радиоэлектронные технологии. 2018. Т. 1, № 1. С. 48–61.
22. Яремко Н.Н. Функции Эрмита с разрывными коэффициентами и их применения для решения обратных задач теплопроводности // Известия Пензен. госуд. педагогич. ун-та им. В.Г. Белинского. 2011. № 26. С. 326–330.
23. Paranda K., Dehghanb M., Rezaei A.R., Ghaderi S.M. An approximation algorithm for the solution of the nonlinear Lane–Emden type equations arising in astrophysics using Hermite functions collocation method // Computer Physics Communications. 2010. Vol. 181. Issue 6. P. 1096–1108.
24. Сайфуллин Р.Т., Бочкарев А.В. Алгоритм определения положения границ хроматографических пиков в сигналах аналитических приборов // Молодежь и наука: шаг к успеху: сб. научн. статей 5-й Всероссийской научной конференции перспективных разработок молодых ученых (22-23 марта 2021 года), в 4 т. 2021. Т. 3. С. 200–204.
25. Гуревич А.Л., Русинов Л.А., Сягаев Н.А. Автоматический хроматографический анализ. Л.: Химия, 1980. 192 с.
26. Слабнов В.Д. Численные методы. СПб.: Лань, 2020. 392 с.
27. Misra S., Wahab M., Patel D. The utility of statistical moments in chromatography using trapezoidal and Simpson's rules of peak integration // Journal of Separation Science. 2019. Vol. 42. DOI: 10.1002/jssc.201801131

Статья поступила в редакцию 22 ноября 2021 г.

SMOOTHING OF CHROMATOGRAPHIC SIGNALS BY THEIR APPROXIMATION IN THE BASIS OF CHEBYSHEV-HERMITE FUNCTIONS

*A.V. Bochkarev., R.T. Sayfullin **

Samara State Technical University
244, Molodogvardeyskaya st., Samara, 443100, Russian Federation

E-mail: bochkarevpubliciit@gmail.com

Abstract. *The paper deals with assessing the scope of the approximation algorithm for smoothing chromatographic signals. The algorithm is based on Chebyshev-Hermite functions. For approximation, an algorithm is used that implies the calculation of the shift and scale factors for the basis functions, as well as the division of the signal into fragments. The smoothing error of the model signal is considered when it is approximated in the selected basis in comparison with the digital moving average filter. The error of smoothing the derivative of the signal and the position on the derivative of extremums is also investigated, a comparison is made, as in the previous case, with a digital moving average filter. As part of an experimental study, the nine most characteristic chromatographic peaks extracted from real chromatograms are processed in the work. To assess the quality of smoothing, the standard deviation of the noise and the residual are compared, and the distribution law of the residual is determined. According to the results of the study, restrictions are set on the maximum allowable value of the asymmetry coefficient of the processed chromatographic peaks. Thanks to the use of the chosen approximation approach, it is possible to solve the problem of smoothing chromatographic signals without attenuating the useful component. The computer algebra system Wolfram Mathematica 11.3 was used for calculations and graphical presentation of the simulation results.*

Keywords: *Chebyshev-Hermite functions, smoothing, random noise, standard deviation, chromatographic signal, signal conversion, approximation method, chromatography.*

REFERENCES

1. *Felinger A.* Data Analysis and Signal Processing in Chromatography / A. Felinger. – Amsterdam, Netherlands: Elsevier Science B.V., 1998. 413 p.
2. *Savitzky A.* Smoothing and Differentiation of Data by Simplified Least Squares Procedures / A. Savitzky, M.J.E. Golay // Anal. Chem. 1964. Vol. 36. P. 1627–1639.
3. *Kalman R.E.* A New Approach to linear Filtering and Prediction Problems // Trans ASME, Series D, J. BasicEng. 1960. Vol. 82. P. 35–45.
4. *Gutorov A.S.* Mathematical modeling and research of filtering algorithms for trajectory processing of data by goals // Automation of management processes. 2015. No. 1(39). P. 34–40.
5. *Lange P.K.* Spline approximation of discrete values of signals using digital filtering methods // Bulletin of the Samara State Technical University. Series "Physical and mathematical sciences". 2003. Issue 18. P. 87–93.
6. *Spiridonov A.V., Time I.V.* Application of smoothing splines for filtering highly noisy signals // Avtomat. and telemekh. 1998. 7. P. 75–82.
7. *Saifullin R.T., Bochkarev A.V.* Use of Chebyshev-Hermite functions in signal processing of analytical instruments // Bulletin of the Samara State Technical University. Series "Technical Sciences". 2019. 1(61). P. 68–81.
8. *Saifullin R.T., Bochkarev A.V.* Calculation of derivatives of an analytical signal in the basis of Chebyshev-Hermite functions // Proceedings of the XI All-Russian Scientific Conference

* *Andrey V. Bochkarev, Postgraduate Student.
Raukhat T. Sayfullin, Professor.*

- with international participation "Mathematical modeling and boundary value problems", May 27-30, 2019, Samara. Samara, 2019. V. 2. P. 137–139.
9. *Saifullin R.T., Bochkarev A.V.* Calculation of continuous wavelet transform of signals in the basis of Chebyshev-Hermite functions // Bulletin of the Samara State Technical University. Series "Technical Sciences". 2019. 2 (62). P. 99–113.
 10. *Saifullin R.T., Bochkarev A.V.* Algorithm for calculating the coefficients of the wavelet transform of signals using the basis of the Chebyshev-Hermite functions // Bulletin of the Samara State Technical University. Series "Technical Sciences". 2019. 4(64). P. 113–124.
 11. *Beliczynski B.* A method of multivariable Hermite basis function approximation // Neurocomputing. 2012. Vol. 96. P. 12–18.
 12. *Jung S.-M.* Approximation of analytic functions by Hermite functions // Bulletin des Sciences Mathématiques. 2009. Vol. 133, 7. P. 756–764.
 13. *Najafi M.M.* Projection method for solving some problems of image processing and analysis: specialty 05.13.18 "Mathematical modeling, numerical methods and software packages": thesis for the degree of candidate of physical and mathematical sciences. Moscow State University M.V. Lomonosov. Moscow, 2004. 110 p.
 14. *Paveleva E.A.* Analysis and restoration of images by projection methods using Hermite functions: specialty 05.13.18 "Mathematical modeling, numerical methods and software packages": dissertation for the degree of Candidate of Physical and Mathematical. Moscow State University M.V. Lomonosov. Moscow, 2015. 130 p.
 15. *Korchagin D.N.* Fast projection method for processing multimedia information: specialty 05.13.18 "Mathematical modeling, numerical methods and software packages": dissertation for the degree of candidate of physical and mathematical sciences. Moscow State University M.V. Lomonosov. Moscow, 2005. 180 p.
 16. *Mamaev N.V., Lukin A.S., Yurin D.V., Glazkova M.A., Sinitsin V.E.* Algorithm of non-local average based on expansion in terms of Ermit functions in problems of computed tomography // GRAFIKON'2013. Proceedings of the 23rd International Conference on Computer Graphics and Vision, September 16-20, 2013, Vladivostok. Vladivostok, 2013. P. 254–258.
 17. *Bayakovskiy Yu.M., Zhirkov A.O., Korchagin D.N., Krylov A.S., Lukin A.S.* Neural network analysis and comparison of time-frequency vectors based on short-term spectral representation and adaptive Hermite transform // Preprints IPM them. M. V. Keldysh. 087. 2001.
 18. *Balakin D.A., Shtykov V.V.* Construction of an orthogonal filter bank based on Ermit's transformations for signal processing // Journal of radio electronics. 2014. 9. P. 1–15.
 19. *Balakin D.A., Shtykov V.V.* Diagnosis of cyclic systems using an algorithm based on Gauss-Hermite functions // Digital signal processing. 2018. 2. P. 59–62.
 20. *Balakin D.A., Shtykov V.V., Churkin S.S.* The use of the phase-metric complex of the EHF range for the diagnosis of human life // Electronics and microwave microelectronics. 2017. V. 1. P. 127–132.
 21. *Balakin D.A., Churkin S.S., Shtykov V.V.* Detection of signals using the orthogonal basis of Gauss-Hermite functions // Infocommunication and radio-electronic technologies. 2018. V. 1(1). P. 48–61.
 22. *Yaremko N.N.* Hermite functions with discontinuous coefficients and their applications for solving inverse problems of heat conduction // Proceedings of the Penza State Pedagogical University. V.G. Belinsky. 2011. 26. P. 326–330.
 23. *Paranda K., Dehghan M., Rezaei A.R., Ghaderi S.M.* An approximation algorithm for the solution of the nonlinear Lane–Emden type equations arising in astrophysics using Hermite functions collocation method // Computer Physics Communications. 2010. Vol. 181(6). P. 1096–1108.
 24. *Saifullin R.T., Bochkarev A.V.* Algorithm for determining the position of the boundaries of chromatographic peaks in the signals of analytical devices // Youth and science: a step to success: collection of scientific articles of the 5th All-Russian scientific conference of promising developments of young scientists (March 22-23, 2021), in 4 volumes. 2021. V. 3. P. 200–204.
 25. *Gurevich A.L., Rusinov L.A., Syagaev N.A.* Automatic chromatographic analysis. L.: Chemistry, 1980. 192 p.
 26. *Slabnov V.D.* Numerical methods. St. Petersburg: Lan, 2020. 392 p.
 27. *Misra S., Wahab M., Patel D.* The utility of statistical moments in chromatography using trapezoidal and Simpson's rules of peak integration // Journal of Separation Science. 2019. Vol. 42. DOI: 10.1002/jssc.201801131