

УДК 621.372

ВЕЙВЛЕТ-РЕГУЛЯРИЗАЦИЯ ОПЕРАЦИИ ДИФФЕРЕНЦИРОВАНИЯ СИГНАЛОВ С ШУМОМ

И. А. Патрикеев¹, Р. А. Степанов¹, П. Г. Фрик¹

Рассматриваются различные алгоритмы приближенного вычисления производной функции, заданной неточно. Алгоритм дифференцирования с использованием вейвлет-преобразования сравнивается с алгоритмом на основе преобразования Фурье и алгоритмами вычисления производной в физическом пространстве. Проведена численная оценка погрешности вычислений различных алгоритмов на модельных примерах.

Ключевые слова: вейвлет, дифференцирование, фильтрация шума, вейвлет-преобразования, регуляризация, преобразование Фурье.

1. Введение. Задача о численном дифференцировании функции, известной приближенно, является классическим примером некорректно поставленной задачи, приводящей к неустойчивости решения [1]. Для обеспечения устойчивости по Тихонову точное решение заменяют приближенным, которое управляется параметром регуляризации α и стремится к точному при отсутствии погрешности измерений (подробнее о различных вариантах стабилизации решения см. в [1–3]). На практике регуляризация обычно сводится либо к сглаживанию исходных данных в физическом пространстве, либо к подавлению высоких частот в спектре измеренных данных. При этом оптимальная ширина сглаживающего окна или соответствующая ему полоса пропускания фильтра связывается с ожидаемым уровнем шума.

Представляется целесообразным сформулировать задачу регуляризации операции дифференцирования зашумленных данных на языке вейвлет-представления сигналов, которое позволяет естественным образом совместить преимущества работы в физическом пространстве и пространстве Фурье. Вейвлет-анализ, превратившийся за последнее десятилетие в хорошо развитый раздел функционального анализа (см., например, [4]), показал свою эффективность в задачах, связанных с обработкой всевозможных многомасштабных сигналов и полей.

Первые попытки использования аппарата вейвлет-анализа при нахождении производной зашумленных данных были выполнены в работах [5–8]. В данной работе методика вейвлет-дифференцирования описывается в рамках общей проблемы без привязки к специфике сигнала. При этом проводится систематическое сравнение вейвлет-регуляризации в задаче дифференцирования зашумленных сигналов с другими общеизвестными подходами. Эффективность использования демонстрируется на конкретных примерах.

2. Методы численного дифференцирования. Пусть функция $f(x)$ имеет первую производную $g(x)$, так что $f(x) = \int_0^x g(x) dx$, и определена на наборе точек x_n с точностью до некоторой случайной ошибки ξ : $\tilde{f}_n = f(x_n) + \xi$. Производная $g(x)$ в простейшем случае может быть приближенно вычислена по формуле

$$\tilde{g}_n = \frac{\tilde{f}_n - \tilde{f}_{n-1}}{\Delta x}, \quad (1)$$

где $\Delta x = x_n - x_{n-1}$.

В условиях шума формула (1) может стать неустойчивой. Принимая для случайной ошибки оценку $|\xi| \leq A$, можно записать

$$|g_n| \leq \left| \frac{f(x_n) - f(x_{n-1})}{\Delta x} \right| + \frac{2A}{\Delta x}. \quad (2)$$

При $\Delta x \rightarrow 0$ первое слагаемое в правой части (2) стремится к $g(x)$, а второе может быть сколь угодно большим. Так, в частности, если функция задана на конечном интервале, то увеличение числа точек приводит к увеличению второго слагаемого в (2).

¹ Институт механики сплошных сред, ул. Королева, 1, 614013, г. Пермь; e-mail: pat@icmm.ru

2.1. Дифференцирование в физическом пространстве. В теории обобщенных функций [9] вводится δ -функция, задаваемая неявно как

$$\int_{-\infty}^{\infty} f(x)\delta(x) dx = \int_{0-}^{0+} f(x)\delta(x) dx = f(0),$$

где $f(x)$ — гладкая функция.

Известно, что δ -функция дифференцируема. Ее первая производная, обозначаемая $\dot{\delta}$ и называемая “дублет”, обладает свойством

$$g(x) = \int_{-\infty}^{\infty} f(x')\dot{\delta}(x-x') dx' = \int_{0-}^{0+} f(x-x')\dot{\delta}(x') dx'. \quad (3)$$

Отметим, что функция $\dot{\delta}$ нечетна и сосредоточена в бесконечно малой окрестности точки $x = x'$. Функцию $\dot{\delta}$ называют иногда *бесконечно малым диполем*.

2.2. Дифференцирование в пространстве Фурье. По теореме о производной [10] дифференцирование в физическом пространстве сводится к умножению в частотном пространстве. Фурье-образы функций g и f связаны соотношением $\hat{g}(k) = ik\hat{f}(k)$, где $\hat{f}(k)$ — результат преобразования Фурье:

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{-ikx} dx.$$

Таким образом, находя Фурье-образ сигнала $\hat{f}(k)$, умножая его в частотном пространстве на ik и выполняя обратное преобразование Фурье, можно найти производную $g(x)$:

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} ik \left(\int_{-\infty}^{\infty} f(x)e^{-ikx} dx \right) e^{ikx} dk. \quad (4)$$

Отметим, что фильтр ik имеет импульсную характеристику $\dot{\delta}$: $ik = \int_{-\infty}^{\infty} \dot{\delta}(x)e^{-ikx} dx$.

Принципиальное отличие алгоритма (4) от прямого дифференцирования в физическом пространстве заключается в том, что при вычислении преобразования Фурье используется информация о сигнале во всех точках числовой оси, в то время как дифференцирование является по определению операцией локальной.

2.3. Дифференцирование с использованием вейвлет-преобразования. С точки зрения локальности, метод дифференцирования на основе вейвлет-анализа занимает промежуточное положение между дифференцированием в физическом пространстве и в пространстве Фурье. Вейвлет-образ исходной функции $f(x)$ имеет вид

$$W_{a,b}\{f\} = \int \psi^*\left(\frac{x-b}{a}\right)f(x) dx,$$

где в качестве анализирующего вейвлета $\psi(x)$ используется комплексная или действительная функция, локализованная как в физическом пространстве, так и в пространстве Фурье, а параметры a и b определяют, соответственно, масштаб и положение функции $\psi_{a,b} = \psi\left(\frac{x-b}{a}\right)$.

Записывая аналогичным образом вейвлет-образ функции $g(x)$ и проводя дифференцирование по частям, легко получить

$$W_{a,b}\{g\} = \int_{-\infty}^{\infty} \psi_{a,b}^*(x)g(x) dx = - \int_{-\infty}^{\infty} \frac{\partial \psi_{a,b}^*(x)}{\partial x} f(x) dx.$$

Итак, выполнив вейвлет-разложение функции $f(x)$ по семейству $-\frac{\partial \psi_{a,b}(x)}{\partial x}$ и выполнив затем обратное вейвлет-преобразование с помощью вейвлет-семейства $\psi_{a,b}(x)$, можно получить искомую функцию $g(x)$:

$$g(x) = \frac{1}{C} \int_{0+}^{\infty} \frac{da}{a^3} \int_{-\infty}^{\infty} db \psi_{a,x}(b) \left(\int_{-\infty}^{\infty} dx \chi_{a,b}^*(x) f(x) \right), \quad (5)$$

где χ — анализирующий вейвлет, ψ — синтезирующий вейвлет и C — константа, определяемая выражением

$$C = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(k)| |\widehat{\chi}(k)|}{|k|} dk = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\widehat{\psi}(k)|^2 dk < \infty.$$

Выбор пары χ и ψ (для анализа и синтеза соответственно) из условия

$$\chi(x) = -\frac{\partial \psi(x)}{\partial x} \quad (6)$$

предполагает, что производная от ψ существует и является вейвлетом. Например, пара функций $\chi = (1-x^2) \exp(-x^2/2)$ и $\psi = -x \exp(-x^2/2)$ удовлетворяет условию (6). В принципе, любой вейвлет, имеющий первую производную, можно использовать в качестве ψ . На практике выбор конкретной пары вейвлетов осуществляется с учетом специфики постановки задачи.

В качестве предельного случая можно использовать пару функций $\chi = \dot{\delta}$ и $\psi = \delta$. Подставляя их в уравнение (5) и рассматривая δ -функцию как сингулярный вейвлет [4], получаем

$$g(x) = \int_{-\infty}^{\infty} dx' g(x') \delta(x-x') = \int_{-\infty}^{\infty} dx' f(x') \dot{\delta}(x-x').$$

3. Регуляризация алгоритмов дифференцирования. Рассмотренные выше алгоритмы не являются устойчивыми и для использования на практике требуют регуляризации.

3.1. Приближенное вычисление производной в физическом пространстве. Для практического использования необходимо провести регуляризацию уравнения свертки (3). Свертка с аппроксимирующей дублет-функцией, ширина которой управляется параметром регуляризации α , дает приближенное значение производной. Обобщенная функция $\dot{\delta}$ аппроксимируется регулярной нечетной функцией с нулевым средним значением, область локализации которой определяется ожидаемым уровнем шума. Простейшей аппроксимацией дублета на дискретном множестве является вычисление разности в соседних точках. Устойчивость можно обеспечить, увеличив расстояние между точками, в которых вычисляется разность. Такой метод известен со времен Ньютона [1].

Одним из практических методов оценки производной в физическом пространстве является свертка измеренного сигнала со сглаживающим окном (Хэмминга, Винера и др. [2, 11]) с последующим вычислением производной в виде конечной разности (1). Сглаженный сигнал находится по формуле

$$\int_{-\infty}^{\infty} f(x') u_{\alpha}(x-x') dx',$$

где u — сглаживающее окно, ширина которого управляется регуляризирующим параметром α .

Применяемое на практике усреднение по соседним точкам эквивалентно сглаживанию прямоугольным окном. Численная оценка погрешности вычисления производной в физическом пространстве (при использовании различных сглаживающих окон) будет дана ниже.

3.2. Приближенное вычисление производной в пространстве Фурье. Как альтернативу свертке со сглаживающим окном в физическом пространстве рассмотрим умножение на фильтр низких частот (ФНЧ) в частотной области. Идеальный ФНЧ отсекает все частоты, превосходящие $k_0(\alpha)$. Сигнал, пропущенный через идеальный ФНЧ, не будет иметь высоких частот, т.е. будет сглаженным. После выполнения обратного преобразования Фурье дифференцирование производится методом конечных разностей.

Можно объединить в частотной области обе операции — обеспечение устойчивости путем подавления высоких частот и вычисление производной путем умножения на ik . Алгоритм (4) в условиях шума является неустойчивым, так как умножение на ik приводит к неограниченному усилению высоких частот. Для обеспечения устойчивости используется ФНЧ с шириной полосы пропускания, управляемой регуляризирующим параметром (при $\alpha \rightarrow 0$ полоса пропускания стремится к бесконечности и приближенное решение стремится к точному). Форма фильтра может выбираться достаточно произвольно, исходя из специфики задачи. Ниже будут рассмотрены два фильтра: дифференцирующий фильтр, представляющий собой произведение ik и идеального ФНЧ, и дифференцирующий гауссов фильтр — произведение ik и гауссова фильтра.

3.3. Вейвлет-регуляризация. При использовании вейвлет-алгоритма вместо свертки сигнала с дублетом выполняется анализ с использованием вейвлета χ и последующий синтез с использованием вейвлета ψ .

Интегрирование по a на этапе синтеза на практике осуществляется в конечных пределах от $a_{\min}$ до $a_{\max}$. В случае высокочастотного шума выбор $a_{\min}$ необходимо связать с ожидаемым уровнем шума. В простейшем случае пределы интегрирования по a выбираются на основании интегральных критериев и не учитывают поведение функции и особенности шума в различные моменты времени. Преимущество вычисления с помощью вейвлетов состоит в том, что легко реализовать локальную регуляризацию, когда пределы интегрирования адаптируются под локальные свойства вейвлет-спектра, т.е. $a_{\min}$ зависит от x .

4. Примеры. Выбор оптимального способа дифференцирования зависит, безусловно, от вида сигнала и характеристик шума. Мы ограничимся рассмотрением двух характерных примеров. Первый пример представляет сигнал, основная энергия которого сосредоточена в высокочастотной части спектра (периоды характерных колебаний существенно меньше длины интервала, на котором задан сигнал). В этом случае можно считать, что сигнал задан на неограниченном отрезке или является периодическим, что позволяет избежать проблем, связанных с границами области определения. В качестве второго пример выбрана кусочно-гладкая функция.

В первом примере рассматривается заданный на интервале от 0 до 1 с шагом 0.001 осциллирующий сигнал, модулированный по частоте и амплитуде, искомая производная которого определяется по формуле

$$g(x) = \sin(2\pi x(10x + 10))(1 - 1/2 \cos(6\pi x)).$$

Производная второго модельного сигнала определяется выражением

$$g(x) = \begin{cases} -2, & x \leq 0, \\ 1, & x > 0. \end{cases}$$

Второй сигнал более гладкий и не имеет высокочастотной составляющей за исключением области разрыва производной. Этот сигнал задан только в интервале от -1 до 1 ; ограниченность области определения оказывает прямое влияние на дифференцирование.

Для каждого из модельных сигналов были сформированы наборы точек f_n по формуле

$$f_n = \int_{x_0}^{x_n} g(x) dx + \xi(\mu),$$

где $\xi(\mu)$ — белый шум с уровнем $\mu * 100\%$ от среднего абсолютного значения функции.

Полученные распределения представлены на рис. 1. Уровень шума составлял 30% для первого сигнала (рис. 1, слева) и 10% для кусочно-гладкого примера (рис. 1, справа). В высокочастотной части соответствующих вейвлет-образов (рис. 1, внизу) наблюдаются нерегулярные структуры, вызванные добавлением шума.

Массивы значений f_n использовались для численного определения производной g_n различными алгоритмами и сравнения с аналитическими значениями производной $g(x_n)$. Среднеквадратичная ошибка (СКО) дифференцирования определялась по формуле [2]

$$\sigma = \sqrt{\frac{\sum_n (g_n - g(x_n))^2}{\sum_n g(x_n)^2}},$$

где $g(x_n)$ — значения производной модельного сигнала в точках x_n , g_n — результат численного дифференцирования.

В данной работе сравнивались следующие алгоритмы.

Прямоугольное окно в физическом пространстве + конечная разность. Сигнал усреднялся по соседним точкам и затем применялся метод конечных разностей для оценки производной. Оптимальное значение ширины окна, минимизирующее σ , определялось путем варьирования количества точек, по которым проводится усреднение, и последующим сравнением с аналитическим решением. Оптимальное значение ширины окна имеет значение 0.018 для первого сигнала и 0.14 — для второго сигнала. Минимальные значения σ приведены в таблице. Графически зависимость σ от ширины окна показана на рис. 2.

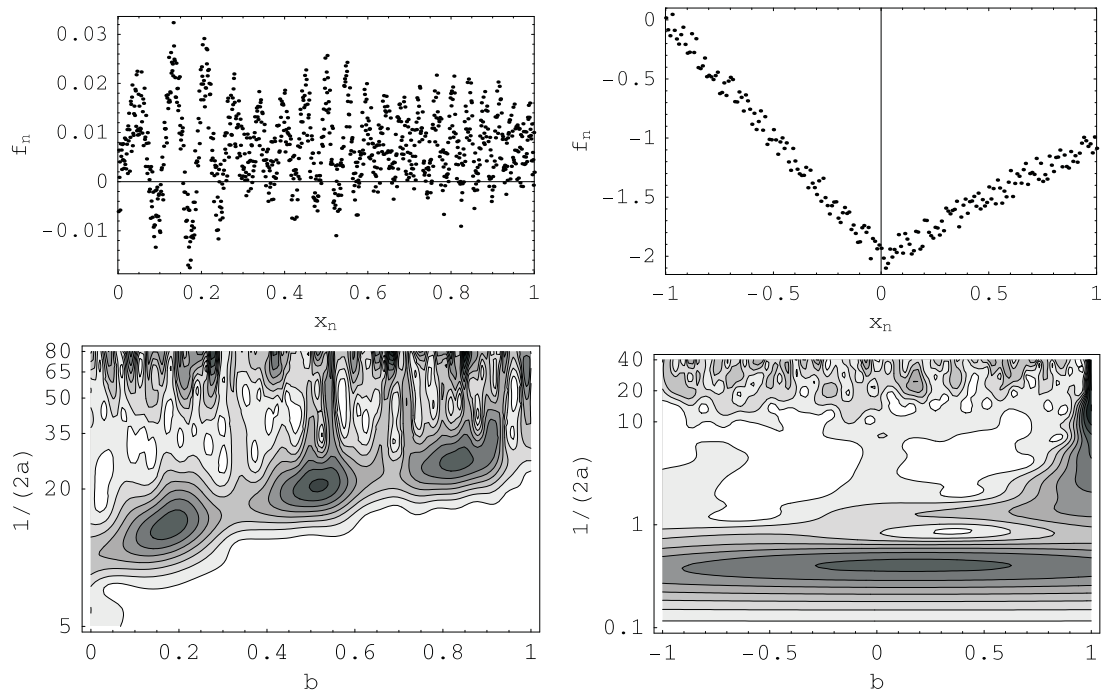


Рис. 1. Модельные сигналы с аддитивным белым шумом (*вверху*) и соответствующие вейвлет-образы (*внизу*)

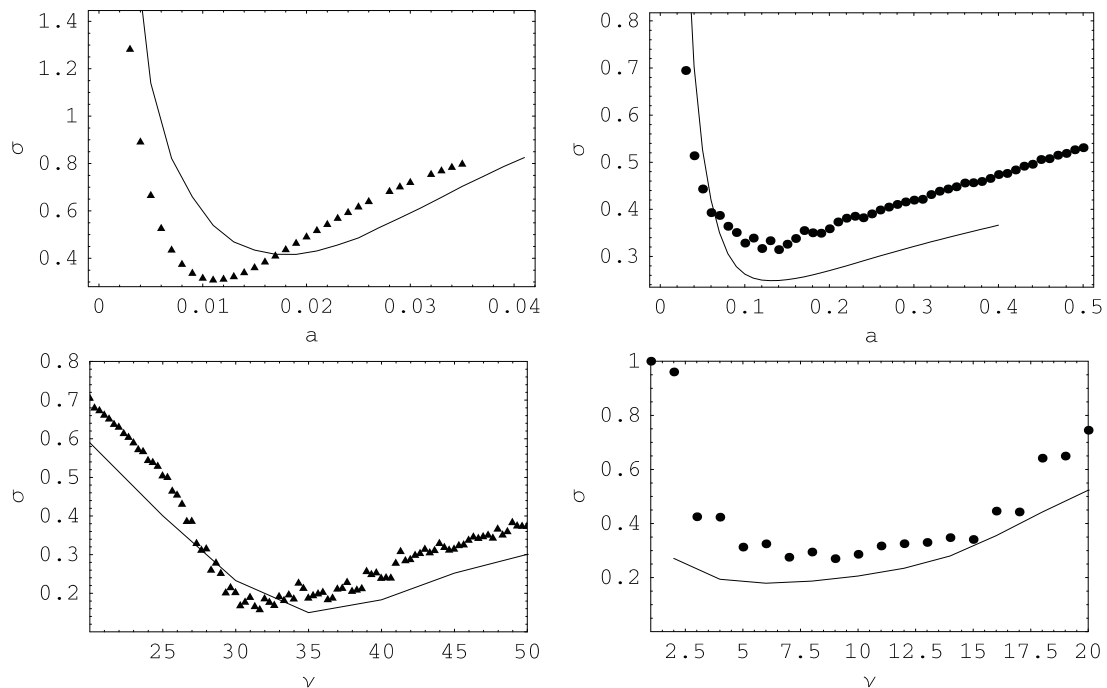


Рис. 2. Зависимость среднеквадратичной ошибки дифференцирования (пример 1 — слева, пример 2 — справа) от ширины окна в физическом пространстве: прямоугольное окно — линия, гауссово окно — точки (*вверху*); в частотном пространстве: прямоугольный ФНЧ — точки, вейвлет-дифференцирование — линия (*внизу*)

Гауссово окно в физическом пространстве + конечная разность. Аналогичным образом определялось оптимальное значение ширины гауссова окна (на уровне половины от максимального значения). Оптимальное значение ширины окна равно 0.011 для первого сигнала и 0.31 — для второго сигнала. Минимальные значения σ представлены в таблице. Сравнивая графики зависимости σ от ширины окна (рис. 2), можно видеть, что гауссово окно дает более точное восстановление производной, чем прямоугольное.

Сравнение эффективности фильтрации шума при вычислении
производной

Метод	Минимальное значение σ	
	Пример 1	Пример 2
Прямоугольное окно	0.42	0.31
Гауссово окно	0.31	0.25
Фурье (ФНЧ) + кон. разн.	0.16	0.8
Фурье гаусс + кон. разн.	0.16	0.7
Дифференцирующий фильтр	0.18	0.27
Диф. гауссов фильтр	0.3	0.32
Вейвлет-регуляризация (Морле)	0.15	0.2
Вейвлет $a_{\min} = (20 * x + c)^{-1}$	0.12	—

ФНЧ в частотном пространстве + конечная разность. Оптимальное значение частоты отсечки k_0 , минимизирующее σ , определялось следующим образом: вычислялось FFT сигнала, отбрасывались коэффициенты ряда Фурье, соответствующие частотам $k > k_0$, а затем выполнялось обратное FFT и применялся метод конечных разностей. Зависимость σ от k_0 представлена на рис. 2. Минимальное значение σ , достигаемое при использовании идеального фильтра низких частот, указано в таблице.

Гауссов фильтр в частотном пространстве + конечная разность. Фильтр получается заменой прямоугольного ФНЧ на гауссов. Зависимость σ от ширины фильтра в частотном пространстве находилась следующим образом: вычислялось FFT сигнала, умножались коэффициенты ряда Фурье на гауссов фильтр, а затем выполнялось обратное FFT и применялся метод конечных разностей. Оптимальное значение ширины фильтра равно 32 для первого примера и 1 — для второго примера. Минимальные значения σ указаны в таблице.

Дифференцирующий фильтр в частотном пространстве. Умножение идеального ФНЧ на ik дает фильтр специальной формы, осуществляющий в частотной области вычисление производной одновременно со сглаживанием. Такой фильтр, а также его ближайшие родственники используются в томографии для вычисления обратного преобразования Радона [2, 12]. Оптимальное значение частоты отсечки k_0 равно 32 для первого примера и 8 — для второго примера. Минимальные значения σ указаны в таблице.

Дифференцирующий гауссов фильтр в частотном пространстве. Заменяв идеальный ФНЧ на гауссов фильтр, модифицируем предыдущий фильтр. Умножив в частотном пространстве гауссов фильтр на ik , получим фильтр $ik \exp(-k^2)$, импульсная характеристика которого равна первой производной от гауссова окна. Варьируя ширину фильтра в частотной области, найдем значение параметра масштаба, минимизирующее σ . Оптимальное значение ширины фильтра равно 37 для первого примера и 10 — для второго примера. Минимальные значения σ указаны в таблице.

Вейвлет Морле в физическом пространстве. Разложим сигнал по масштабам, используя производную от вейвлета Морле в качестве анализирующего вейвлета. Диапазон масштабов зададим от 0.1 до 0.006, за пределами которого сигнал практически отсутствует. Затем для каждого масштаба вычислим свертку полученных вейвлет-коэффициентов с вейвлетом Морле и проинтегрируем по масштабам от $a_{\max} = 0.1$ до $a_{\min}$. Варьируя $a_{\min}$, найдем соответствующие значения σ . Зависимость σ от $a_{\min}$ (переведенного в частотный аналог по формуле $\nu = 1/(2a_{\min})$) представлена на рис. 2. Оптимальное значение $\nu = 35$ для первого примера и 7 — для второго примера. Минимальное значение σ указано в таблице. Анализ вейвлет-образа квазигармонического сигнала (рис. 1, слева) показывает, что существенная часть информации сосредоточена в трех компактных областях, расположенных вдоль наклонной прямой. Это дает возможность задать минимальный масштаб в виде функции $a_{\min} = (10 * x + c)^{-1}$. Тогда перебор значений c дополнительно минимизирует σ . В этом случае для значения $c = 42$ получается результат $\sigma = 0.12$, что заметно лучше, чем результат, полученный с использованием преобразования Фурье. Для кусочно-гладкого сигнала (пример 2) дополнительная оптимизация не проводилась, так как его вейвлет-образ (рис. 1, справа) существенно искажен влиянием шума.

На рис. 3 представлены графики аналитических производных модельных сигналов и результаты численного дифференцирования методом на основе вейвлета Морле при оптимальном значении $a_{\min}$ (без использования дополнительной оптимизации). Вейвлет-образы производных сигналов, приведенные на рис. 3, внизу, показывают распределения спектральных свойств в заданных интервалах.

5. Обсуждение. Алгоритм с использованием вейвлет-преобразования позволяет проводить устой-

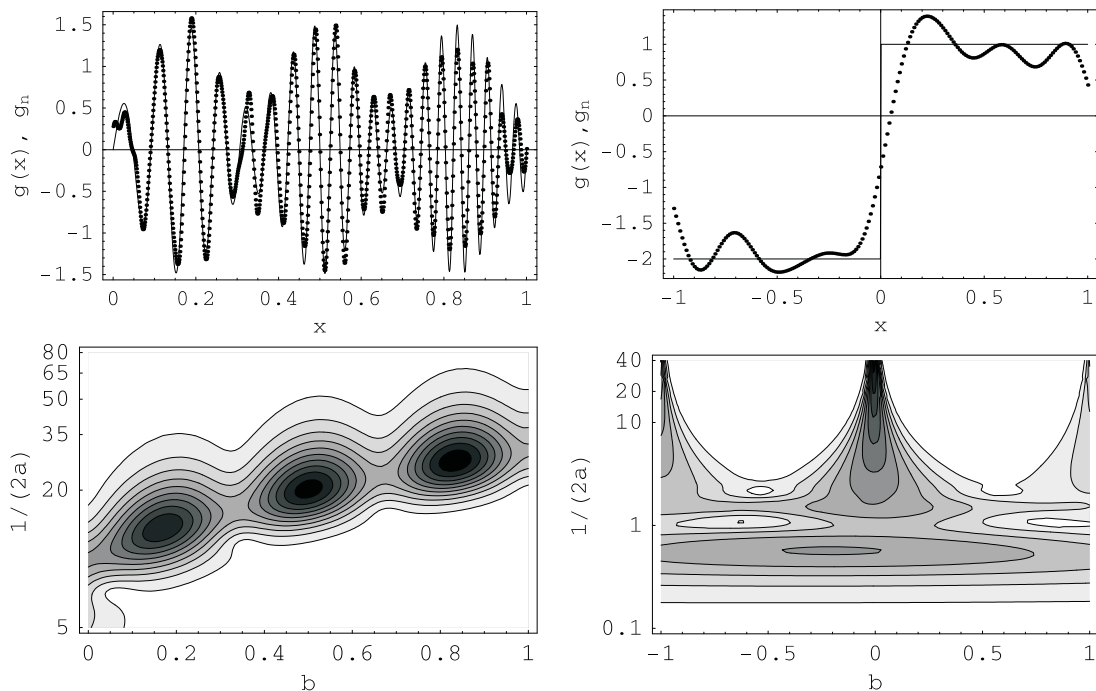


Рис. 3. Производные модельных сигналов, вычисленные аналитически — линии; численная оценка с использованием вейвлета Морле при оптимальном значении $a_{\min}$ — точки (вверху). Вейвлет-образы производных модельных сигналов (внизу)

чивое дифференцирование в условиях шума. Локальность базисных функций позволяет более точно учитывать свойства сигнала по сравнению с методами фильтрации в частотном пространстве.

Особый интерес предложенные методы могут представлять при решении обратных задач (в медицине, астрономии, гидродинамике и т.д.). Например, интегральное уравнение Абеля, к которому сводится задача осесимметричной томографии, имеет решение [12]

$$g(r) = -\frac{1}{\pi} \int_r^{\infty} \frac{dp}{\sqrt{p^2 - r^2}} \frac{\partial f(p)}{\partial p},$$

где $f(p)$ — измеренные с некоторой погрешностью проекционные данные.

В общем случае решение двумерной задачи томографии сводится к обратному преобразованию Радона, которое также может быть выражено через первую производную измеренного сигнала [2]:

$$g(x, y) = -\frac{1}{2\pi^2} \int_0^{\pi} d\phi \int_{-\infty}^{\infty} \frac{dp}{p - x \sin \phi + y \cos \phi} \frac{\partial f(p, \phi)}{\partial p},$$

где $f(p, \phi)$ — измеренные с некоторой погрешностью проекционные данные.

В работе [14] приведен ряд примеров, показывающих, что метод позволяет эффективно подавить шумы, возникающие в спектре из-за пробелов, и повысить точность восстановления спектра исходного сигнала. Этот алгоритм легко переносится и на рассматриваемую задачу о вычислении производной и состоит в этом случае в следующем. На первом шаге с помощью метода “дырявых” вейвлетов вычисляется вейвлет-образ $W_{a,b}$ исходной функции $\tilde{f}(x)$, а на втором шаге — по формуле (5) восстанавливается искомая производная $g(x) = f'(x)$. При этом на втором шаге используются *полные* вейвлет-функции $\chi(x)$.

Работа выполнена при финансовой поддержке РФФИ-ННИО (код проекта 03-02-04031), РФФИ (код проекта 03-02-16384) и Научно-образовательного центра (код проекта РЕ-009-0). Второй автор также благодарен Уральскому отделению РАН (грант молодым ученым).

СПИСОК ЛИТЕРАТУРЫ

1. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1986.

2. Пикалов В.В., Мельникова Т.С. Томография плазмы. Новосибирск: Наука, 1995.
3. Троицкий И.Н. Статистическая теория томографии. М.: Радио и связь, 1989.
4. Holschneider M. Wavelets: tool of analysis. Oxford: Oxford University Press, 1995.
5. Патрикеев И.А., Фрик П.Г. Вейвлет-томография в условиях шума // Математическое моделирование систем и процессов. Вып. 5. Пермь: Изд-во ПГТУ, 1997. 86–92.
6. Patrickeyev I., Frick P. Lymphocyte nucleus reconstruction via wavelet tomography // Journal of Biomedical Optics. 1999. N 7. 376–380.
7. Степанов Р.А. Двумерная вейвлет-томография галактических полей // Математическое моделирование систем и процессов. Вып. 7. Пермь: Изд-во ПГТУ, 1999. 86–91.
8. Stepanov R., Frick P., Shukurov A., Sokoloff D. Wavelet-tomography of the Galactic magnetic field. I. The method // Astronomy and astrophysics. 2002. **391**. 361–368.
9. Гельфанд И.М., Шилов Г.Е. Обобщенные функции и действия над ними. М.: Наука, 1959.
10. Bracewell R. Fourier transform and its application. New York: McGraw-Hill, Inc., 1965.
11. Хемминг Р. Цифровые фильтры. М.: Советское радио, 1980.
12. Levin G.G. (Ed.) Analytical methods for optical tomography, Bellingham: SPIE, 1992.
13. Frick P., Baliunas S., Galyagin D., Sokoloff D., Soon W. Wavelet analysis of stellar chromospheric activity variations // Astrophysical Journal. 1997. **483**. 426–434.
14. Frick P., Grossmann A., Tchamichian Ph. Wavelet analysis of signals with gaps // Journal of Mathematical Physics. 1998. **39**, N 8. 4091–4107.

Поступила в редакцию
28.12.2004
