

УДК 621.391

ДЕКОМПОЗИЦИЯ НА ЭМПИРИЧЕСКИЕ МОДЫ И ЕЕ ИСПОЛЬЗОВАНИЕ ПРИ АНАЛИЗЕ ДРОБНОГО БРОУНОВСКОГО ДВИЖЕНИЯ

Клионский Д.М., Орешко Н.И., Геппенер В.В.

Введение

Большинство сигналов, рассматриваемых в рамках решения различных прикладных задач, являются по своей природе нестационарными. Стационарность сигналов, обычно рассматриваемая на практике в широком смысле, означает, что некоторые характеристики не зависят от времени, а некоторые зависят только от величины интервала между конкретной парой временных отсчетов. Среди упомянутых характеристик одними из наиболее информативных являются начальные и центральные моменты различных порядков, значения автокорреляционной функции (АКФ), а также значения мгновенной частоты, обычно вычисляемой на основе преобразования Гильберта.

Широко используемый классический анализ Фурье на основе аппроксимации сигнала конечным набором дискретных гармоник предназначен для обработки стационарных сигналов в силу очевидных особенностей используемых базисных функций (гармонических функций). Но у них при этом отсутствует точная временная локализация отдельных особенностей (при вычислении коэффициентов разложения сигнала по Фурье-базису используется весь сигнал, что негативно сказывается на временной локализации) и, кроме того, для нестационарного сигнала может потребоваться довольно большое число гармоник для обеспечения требуемой точности представления. Энергия сигнала при этом рассеивается по широкой полосе частот, что затрудняет выявление основных диапазонов ее концентрации. Для повышения точности и достоверности анализа нестационарных сигналов необходим особый подход, обладающий свойством *адаптивности* к конкретному рассматриваемому сигналу. Другими словами, разложение по базису с целью дальнейшего анализа должно проводиться с учетом *локальных особенностей* (таких как, например, экстремумы и нули) и *внутренней структуры* (наличия амплитудной, частотной и фазовой модуляций, шумов и помех) каждого конкретного сигнала. Этим важнейшим требованиям отвечает алгоритм *декомпозиции на эмпирические моды* (ДЭМ, *EMD – Empirical Mode Decomposition*) [1-4], являющийся весьма перспективным при решении таких задач, как очистка сигналов от шумов различной природы, выделение тренда среднего значения, анализ степени хаотичности и регулярности сигналов на основе оценивания показателя Херста, экстраполяция сигналов, частотно-временной анализ на основе *преобразования Гильберта-Гуанга* (*HHT – Hilbert-Huang Transform*). Наряду с этим, не менее важным является возможность проведения мультиразрешающего анализа сигналов (представление в виде набора компонент с различным временным разрешением), что позволяет выявить отдельные компо-

Представлен принципиально новый подход к анализу нестационарных сигналов сложной внутренней структуры – алгоритм декомпозиции на эмпирические моды. Описаны идеи, заложенные в его основу, а также основные преимущества по сравнению с существующими методами анализа нестационарных сигналов. В качестве примера рассмотрена задача анализа дробного броуновского движения, в частности, оценивание его показателя Херста.

ненты, входящие в состав сложного сигнала, классифицировать их в частотной области и сконструировать на их основе банк фильтров. Основным преимуществом подхода ДЭМ является адаптивность всех используемых методов анализа на его основе, что дает возможность обеспечить более высокую точность решения вышеперечисленных задач. Следует также отметить, что в силу адаптивной природы появляются дополнительные возможности для физической интерпретации полученных базисных функций в различных предметных областях [1].

Выше упомянутые свойства ДЭМ позволяют также применить эту стремительно развивающуюся технологию для «интеллектуального» анализа сигналов (сегментации, кластер-анализа, классификации на основе разложения и спектра Гильберта) [4]. Безусловно, требуют дальнейшего исследования новые возможности ДЭМ при работе с различными классами нестационарных сигналов.

Понятие эмпирической и характеристической мод. Основные свойства

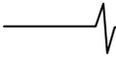
Из названия алгоритма следует, что *эмпирические моды* (*ЭМ, Empirical mode, Intrinsic mode function*) представляют собой те компоненты, на которые осуществляется декомпозиция (разложение) исходного сигнала. В общем случае данные функции могут иметь произвольную форму, быть заданы аналитически (точной формулой) или численно (в виде набора дискретных отсчетов), однако они непременно должны удовлетворять 2-м необходимым условиям [1-4]:

1) На всей длительности функции общее число экстремумов должно быть равно общему числу нулей с точностью до 1:

$$N_{\max} + N_{\min} = N_{\text{zero}} \pm 1, \quad (1)$$

где N_{\max} , N_{\min} , N_{zero} - соответственно, общее число максимумов, минимумов и нулей на всей длительности сигнала, не считая начальный и конечный отсчеты (которые в редких случаях могут быть единственными экстремумами сигнала).

2) Полусумма 2-х огибающих функции – верхней, интерполирующей весь найденный набор локальных макси-



мумов и нижней, интерполирующей весь найденный набор локальных минимумов, не должна превышать некоторого малого значения η , определяемого машинной точностью \mathcal{E} и погрешностями, связанными с получением, преобразованием и передачей данных. При интерполяции обеих огибающих используются в основном *кубические сплайны* [11,12], причина выбора которых будет пояснена далее. Точного равенства нулю данной полусуммы на практике добиться невозможно в силу ряда объективных причин: точность определения экстремумов, вычислительные погрешности (связанные с особенностями машинной арифметики с плавающей точкой), плохая обусловленность систем уравнений, на основе которых определяются коэффициенты сплайнов, а также влияние самого исходного сигнала (например, краевые эффекты – сильные осцилляции интерполирующей функции на краях). Аналитически данное условие записывается как

$$\frac{U(k) + L(k)}{2} \leq \eta, \quad k = \overline{1, N}, \quad (2)$$

где $U(k)$ и $L(k)$ - значения верхней и нижней огибающих сигнала, сформированных на основе интерполяции кубическими сплайнами (k - номер отсчета сигнала), η - выбранное пороговое значение, N - общее количество отсчетов сигнала. Данная полусумма есть не что иное, как *локальное среднее* сигнала, которое должно быть меньше либо равно пороговому значению η в каждый дискретный момент времени.

Данные условия требуют некоторых комментариев. Первое из них, по сути, является требованием узкополосности ЭМ. Данная мера узкополосности ν выражается следующим образом [10]:

$$\nu^2 = \frac{(N_{\max} + N_{\min})^2 - N_{\text{zero}}^2}{N_{\text{zero}}^2}. \quad (3)$$

Очевидно, эта величина равна нулю, если число экстремумов равняется числу нулей, и она близка к нулю, если между ними есть отличие на 1. Именно это и утверждается в первой части определения ЭМ.

Второе условие также имеет физическую интерпретацию. Из него следует, что ЭМ является стационарной функцией по отношению к вышеопределенному локальному среднему значению, которое неизменно и меньше порогового значения η в любой момент времени. Кроме того, для достижения этого условия ЭМ должна иметь положительные значения в точках максимумов и отрицательные значения в точках минимумов (должна соблюдаться симметрия относительно оси времени), т.к. иначе для отдель-

ных моментов времени условие может нарушиться. Наконец, важно отметить, что ЭМ в общем случае обладает одновременно и амплитудной и частотной модуляциями. Закон амплитудной модуляции может быть установлен из огибающих, полученных интерполяцией экстремумов или преобразованием Гильберта, а закон частотной модуляции устанавливается путем вычисления мгновенной частоты.

Понятие огибающей, введенное для определения ЭМ, рассматривается с учетом выбранного вида интерполяции. Как было указано, в данном алгоритме используется кубическая сплайн-интерполяция. У сплайнов в общем и кубических сплайнов в частности есть ряд преимуществ [11] по сравнению с другими функциями, обладающими высокой степенью гладкости, наличие которой существенно повышает точность при извлечении ЭМ из сигнала.

Во-первых, по сравнению с полиномами у сплайнов нет эффекта "раскачивания" [11] – сильных осцилляций интерполирующей функции в случае, если данные заведомо имеют не полиномиальную природу. Кроме того, кубические сплайны являются непрерывными (подразумевается равенство значений в узлах стыков) и дважды дифференцируемыми. Последнее означает, что сам сплайн не имеет острых углов (определенность 1-й производной), и в любой точке определен радиус его кривизны (определенность 2-й производной), что дает возможность оценивать интенсивность осцилляций самой моды. Но самым важным является то, что кубический сплайн минимизирует осциллирующее поведение функции – т.е. из всего многообразия дважды дифференцируемых, непрерывных на некотором интервале $[a; b]$ функций $f(k)$, интерполирующих заданную совокупность точек $\{(x_k, y_k)\}_{k=1}^N$, кубический сплайн меньше всего осциллирует. Математически данное утверждение можно записать в следующем виде:

$$\int_a^b [S''(x)]^2 dx \leq \int_a^b [f''(x)]^2 dx, \quad (4)$$

при условии $S'(a) = f'(a), S'(b) = f'(b)$ где $S(x)$ - кубический сплайн, $f(x)$ - произвольная дважды дифференцируемая функция. Условие записано для аналоговых (непрерывных) функций, но оно также верно и для дискретных. Аналогичное условие справедливо и для сплайнов более высоких порядков (там рассматривается множество n раз дифференцируемых функций, где $n+1$ - порядок сплайна). Но их применение сопряжено с неоправданно большими для данного алгоритма временными затратами на вычисление коэффициентов.

Ниже (рис. 1) приведены примеры ЭМ с интерполяцией огибающих кубическими сплайнами:

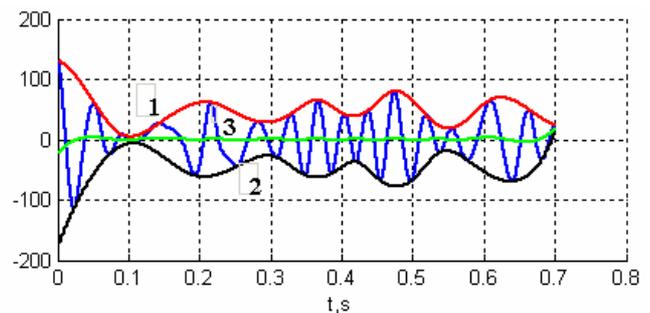
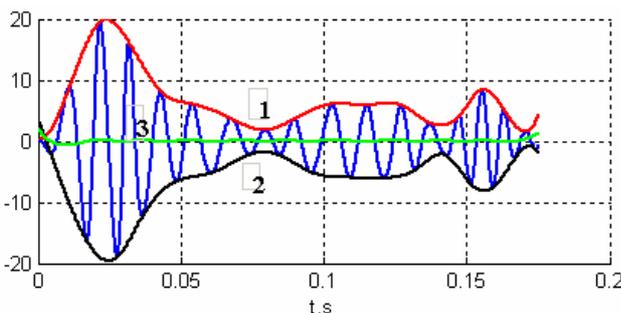


Рис. 1. Примеры ЭМ. 1 - верхняя огибающая, 2 - нижняя огибающая, 3 - их полусумма

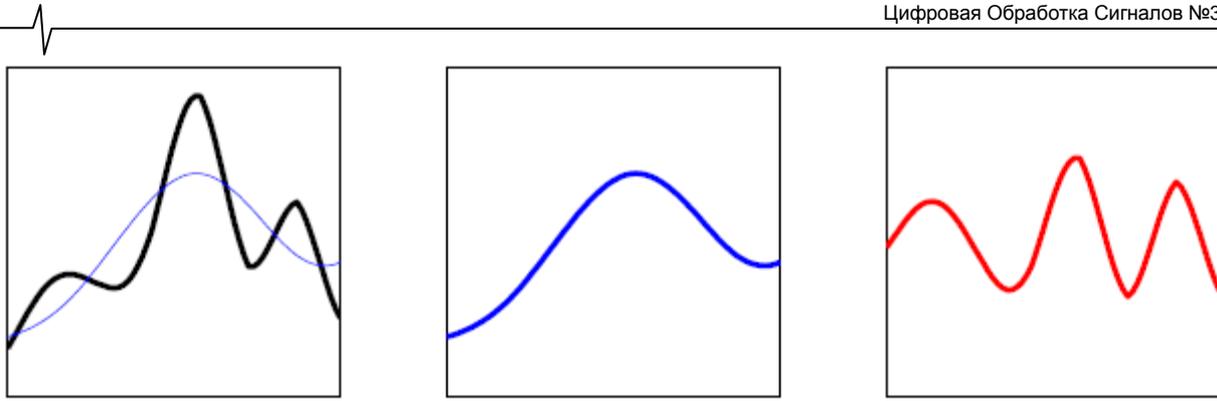


Рис. 2. Иллюстрация представления сигнала (слева) в виде суммы низкочастотной, слабо осциллирующей компоненты (в центре) и высокочастотной, заметно осциллирующей (справа)

Характеристические моды (ХМ, *Characteristic mode*), в отличие от ЭМ, всегда имеют строгое и точное аналитическое выражение, но они довольно редко встречаются на практике. При этом оба необходимых условия, сформулированные для ЭМ, остаются в силе. Подобные модели можно использовать при аппроксимации ЭМ с помощью ХМ, что иногда облегчает дальнейший анализ и обработку. Некоторые примеры ХМ приведены ниже:

- 1) $s_1(k) = A \cos(w_0 k + \varphi)$ - гармонический сигнал с известными A, w_0, φ .
- 2) $s_2(k) = A \cos(w_0 k + 0.5 \mu k^2 + \varphi)$ - ЛЧМ-сигнал с известными A, w_0, μ, φ .
- 3) $s_3(k) = e^{-ak^2} \cos(w_0 k + \varphi)$ - Гауссов радиоимпульс с известными a, w_0, φ .

Алгоритм ДЭМ. Процесс отсеивания

Алгоритм ДЭМ представляет любой сигнал с конечным числом отсчетов (конечной энергией) в виде набора ЭМ. В основе алгоритма лежит важная и фундаментальная идея, основанная на том, что на каждом этапе, т.е. при извлечении очередной моды, сигнал (вернее, текущий остаток) представляется в виде суммы быстро осциллирующей компоненты, которой является сама извлеченная ЭМ, и медленно осциллирующей, которой является результирующий остаток. Последний затем подвергается дальнейшему разложению, если у него имеется хотя бы один минимум и хотя бы один максимум.

Сказанное иллюстрируется графически на рис. 2.

Основные этапы алгоритма, в соответствии со схемой на рис. 3, описаны далее.

Шаг 1. Рассматривается текущий остаток $r_p(k)$ ($r_1(k)$ - первый остаток, который и есть сам исходный сигнал $s(k)$). Определяются его экстремумы и формируются 2 набора:

$$\{M_i\}, i = 1, 2, 3, \dots; \{m_i\}, i = 1, 2, 3, \dots, \tag{5}$$

где $\{M_i\}$ и $\{m_i\}$ - набор максимумов и минимумов соответственно;

Если экстремумы отсутствуют в текущем остатке, то возможно применение операции численного дифференцирования с помощью конечноразностных формул для их порождения. По окончании работы всего алгоритма производится возврат к начальному

множеству значений путем интегрирования каждой ЭМ на основе численных методов [11,12].

Далее по найденным экстремумам строятся две огибающие с помощью интерполяции кубическими сплайнами:

$$U_j(k) = f_U(M_i, k), L_j(k) = f_L(m_i, k), \tag{6}$$

где $U_j(k)$ и $L_j(k)$ - верхняя и нижняя огибающие, построенные, соответственно, по найденным локальным максимумам и минимумам, j - номер итерации отсеивания (см. ниже).

Необходимо отметить, что построенные огибающие должны заключать в себе весь текущий остаток, т.е. любой отсчет должен удовлетворять условию:

$$L_j(k) \leq r_p(k) \leq U_j(k), \forall k. \tag{7}$$

Однако в некоторых случаях наблюдаются так называемые “всплески” - явление, при котором какой-либо отсчет превышает значения верхней огибающей или же уступает значению нижней огибающей. Также необходимо учитывать влияние начального и конечного отсчетов на интенсивность осцилляций на краях (краевые эффекты). Для уменьшения этих осцилляций используются методы зеркального отображения экстремумов, ближайших к краевым отсчетам [1], что позволяет в некоторой степени подавить или сгладить колебания на краях.

После построения огибающих вычисляется их полусумма - локальное (мгновенное) среднее значение, зависящее от времени:

$$e_j(k) = 0.5 \cdot [U_j(k) + L_j(k)]. \tag{8}$$

Затем выполняется переход к шагу 2.

Шаг 2. Найденное среднее значение вычитается из текущего остатка, и полученный результат $h_j(k)$ оказывается “кандидатом” на то, чтобы стать очередной модой:

$$h_j(k) = r_p(k) \text{ при } j = 1, \tag{9}$$

$$h_j(k) = h_{j-1}(k) - e_j(k) \text{ при } j > 1, \tag{10}$$

где p - номер извлекаемой ЭМ.

При этом требуется проверить два необходимых условия отнесения эмпирической функции $h(k)$ к классу ЭМ. Если оба условия выполняются, то осуществляется переход к шагу 3.

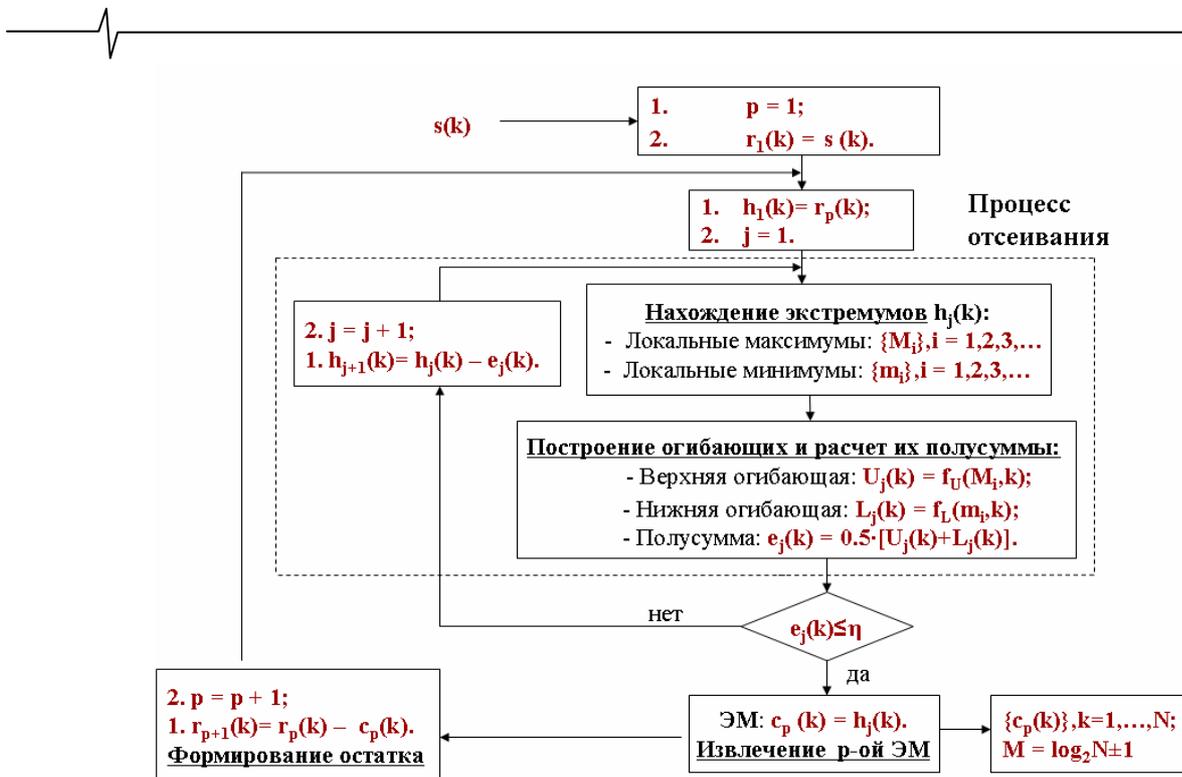


Рис. 3. Графическое представление основных этапов алгоритма ДЭМ

Если какое-либо из них нарушено, то осуществляется возврат к шагу 1, но теперь уже в качестве текущего остатка выступает полученный на 2-м шаге результат. Тем самым начинается так называемый процесс отсеивания (*sifting process*) [1-4], который может быть записан в следующем виде:

$$\begin{aligned} h_1(k) &= r_p(k), h_2(k) = h_1(k) - e_1(k), \\ h_{j+1}(k) &= h_j(k) - e_j(k), \dots c_p(k) = h_{iter}(k), \end{aligned} \quad (11)$$

где $e_j(k)$ - среднее значение функции на j -ой итерации процесса отсеивания, $h_j(k)$ - текущий результат на j -ой итерации отсеивания, $iter$ - общее число итераций для данной ЭМ.

На итерации с номером $iter$ процесс отсеивания для извлечения очередной моды прекращается и выполняется переход к шагу 3.

Шаг 3. После извлечения ЭМ в ее окончательном виде осуществляется ее вычитание из текущего остатка для формирования нового (для обновления остатка):

$$r_{p+1}(k) = r_p(k) - c_p(k), \quad (12)$$

где $c_p(k)$ - полученная ЭМ; $r_p(k)$ - текущий остаток, $r_{p+1}(k)$ - новый остаток.

Шаг 4. Далее осуществляется переход к шагу 1, где в качестве функции, из которой будут извлекаться ЭМ с более высокими номерами, выступит тот остаток, который был получен на 3-м шаге, т.е. $r_{p+1}(k)$.

Основные этапы представленного алгоритма требуют некоторых комментариев. В настоящее время он еще не имеет такой стройной и глубоко изученной теоретической базы, которой обладает классический анализ Фурье или вейвлет-анализ. Алгоритм имеет эмпирическую природу. Например, строго не доказана его сходимость при произвольном способе интерполяции огибающих. Однако, во-первых, она вытекает из того обстоятельства, что на каждом этапе происходит выделе-

ние компонент таким образом, что удаляется локальное среднее значение, и в итоге остается результирующий низкочастотный монотонный остаток, отражающий общую тенденцию в сигнале (тренд среднего). Он не будет иметь достаточного числа экстремумов, и понятие огибающих (либо верхней, либо нижней, либо обеих) для него потеряет смысл - это будет означать завершение работы алгоритма. Кроме того, возможно введение ограничения на число извлекаемых ЭМ, например, с использованием *информационных критериев* (Акаике, Риссанена и др.) [5]. Последние, как известно, находят оптимальное значение порядка модели, доставляющее экстремум (максимум или минимум в зависимости от конкретного вида) выражению, определяемому выбранным критерием. Он обычно содержит два слагаемых, одно из которых возрастает с увеличением порядка модели, а другое, наоборот, убывает. Применительно к данной задаче в качестве порядка модели будет выступать число ЭМ в разложении, а в качестве дисперсии ошибки - дисперсия сигнала ошибки приближения исходного сигнала выбранным числом ЭМ. Данный подход направлен на то, чтобы сохранить лишь наиболее информативные компоненты, относящиеся к полезному сигналу, и исключить из рассмотрения те, что являются шумовыми. Наиболее часто рассматриваемые информационные критерии имеют вид:

$$F(q, D) = N \ln D + g(q, N), \quad (13)$$

где N - число отсчетов в сигнале, q - порядок модели (число ЭМ), D - дисперсия ошибки восстановления сигнала числом ЭМ, равным q . Широко используемыми являются

$$ICA(q, D) = N \ln D + 2q, \quad (14)$$

критерий Акаике, где $g(q, N) = g(q) = 2q$, и

$$MDL(q, D) = N \ln D + q \ln N, \quad (15)$$

критерий Риссанена (критерий минимума длины описания), где $g(q, N) = q \ln N$.

При отыскании оптимального числа ЭМ разумно начать с тех мод, что имеют самые большие номера, т.е. являются самыми низкочастотными, и наращивать это число в направлении высокочастотных мод, которые являются шумовыми и нежелательными с точки зрения результирующего отношения сигнал-шум.

Исходя из практических результатов, была выведена эмпирическая формула для оценки общего числа ЭМ в разложении произвольного сигнала:

$$M = \log_2 N \pm 1, \quad (16)$$

где M - общее число ЭМ, N - общее число отсчетов сигнала.

Ключевая роль во всем алгоритме отводится уже упоминавшемуся процессу отсеивания (на рис. 3 он обозначен пунктиром). Он имеет итерационную природу и предназначен для извлечения конкретной ЭМ, удовлетворяющей двум необходимым условиям из ее определения. Смысл названия "отсеивание" состоит в том, что отсеиваются (отбраковываются) все те "кандидаты" на роль ЭМ, которые ими на самом деле не являются по причине невыполнения одного из двух требуемых условий.

Всю процедуру разложения можно рассматривать как два вложенных друг в друга цикла: внешний отвечает за извлечение очередной моды, а внутренний – за процесс отсеивания для очередной моды. Отсеивание применяется для каждой ЭМ в отдельности, однако число итераций может отличаться, причем порой весьма существенно. В силу итерационного характера невозможно заранее предсказать, сколько именно вычислений потребуется для каждой ЭМ. Для более точного выполнения отсеивания были специально разработаны различные критерии его остановки, наиболее широко используемые из которых приведены ниже.

Процесс отсеивания для очередной моды прекращается, если величина показателя нормированной квадратичной разности между значениями эмпирической функции $h(k)$ на двух последовательных итерациях отсеивания SD выходит за границы интервала $[0.2; 0.3]$. SD определяется как

$$SD = \sum_{k=1}^N \frac{[h_j(k) - h_{j-1}(k)]^2}{h_{j-1}^2(k) + \mu}, \quad (17)$$

где j и $j-1$ – номера двух последовательных итераций процесса отсеивания; $h_j(k)$ и $h_{j-1}(k)$ – значения эмпирической функции на двух последовательных итерациях отсеивания, μ - маленькое число (обычно $\mu = 0.00001$), введенное на случай равенства $h_{j-1}(k)$ нулю в отдельные моменты времени. Указанный интервал $[0.2; 0.3]$ получен в работе [1]. В общем случае пороговое значение показателя SD может быть изменено. Главным недостатком данного критерия является то, что он никак не учитывает определение ЭМ, а именно ни условие равенства числа нулей и экстремумов, ни требование симметричности огибающих. Следовательно, вполне возможен вариант, когда результатом процесса отсеивания окажется функция, не удовлетворяющая одному из необходимых условий, что даст неверный результат в целом. Однако в большинстве случаев,

как показывают эксперименты, этого не происходит, и полученные по результатам процесса отсеивания функции являются истинными ЭМ.

Процесс отсеивания для очередной моды прекращается, если число экстремумов и нулей на текущей и предыдущей итерациях совпадают, и эти числа между собой одинаковы либо отличаются на 1. Если при этом полусумма двух огибающих функции превосходит пороговое значение η (т.е. не выполняется второе условие из определения ЭМ), то число итераций может быть увеличено.

После завершения алгоритма можно провести реконструкцию (восстановление) исходного сигнала с помощью выделенных мод. Для этого используются следующие формулы:

$$s(k) = \sum_{i=1}^{M-1} c_i(k) + r_M(k) \quad (18)$$

при использовании всех ЭМ (как высокочастотных, так и низкочастотных) и

$$s(k) = \sum_{i \in I} c_i(k) + r_M(k), \quad (19)$$

в случае исключения некоторых ЭМ из рассмотрения, например, в результате применения информационных критериев. В этом случае верхняя граница в суммировании в (19) не указывается, а индекс суммирования i принадлежат индексному множеству I , которое упорядочивается в зависимости от включения той или иной ЭМ.

Алгоритм ДЭМ может быть применен для решения широкого класса задач, часть из которых уже была упомянута в начале данной статьи. Далее внимание будет уделено одному из важных и наименее освещенных в литературе практических приложений – анализу на основе ДЭМ дробного броуновского движения и, в частности, оцениванию показателя Херста.

Показатель Херста дробного броуновского движения

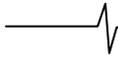
Показатель Херста [8,9] является важной характеристикой сигналов, используемой для оценки степени их хаотичности (близости по своим свойствам к шумовым сигналам) и регулярности. Согласно определению

$$R/S = (N/2)^H, \quad (20)$$

где R - размах сигнала (разница между минимальным и максимальным значениями), S - стандартное отклонение сигнала, N - число отсчетов в сигнале, H - значение показателя Херста.

Существует несколько способов оценивания значения показателя Херста [6-9], однако все они сильно зависят от выбора начальных параметров соответствующих алгоритмов. В данной статье предлагается новый метод оценивания показателя Херста на основе ДЭМ.

Дробное броуновское движение (ДБД) является распространенной моделью для многих прикладных задач. Например, оно широко используется в экономике (при анализе рынков, колебаний курсов валют на биржах) [8], гидрологии [9] и др. научных направлениях. Смысл названия "дробное" состоит в том, что энергетический спектр такого сигнала описывается обратным степенным законом вида $f^{-\nu}$, где ν - целое число. Согласно



определению, ДБД с показателем Херста H - гауссовский сигнал $x_H(k)$ с нулевым средним значением, для которого выполняются следующие условия:

$$x_H(0) = 0, \quad (21)$$

$$E\{[x_H(k) - x_H(k - \tau)]^2\} = \sigma^2 \tau^{2H}, \quad (22)$$

где E - оператор математического ожидания, σ^2 - дисперсия $x_H(k)$. ДБД является обобщением обычного броуновского движения (гауссовского сигнала с независимыми приращениями) со значением $H = 0.5$, производная которого есть белый гауссовский шум.

Определим АКФ сигнала, удовлетворяющего данным свойствам. Для этого необходимо учесть тот факт, что при задержке на $\tau = k$ справедливо соотношение

$$E\{[x_H(k)]^2\} = \sigma^2 k^{2H}. \quad (23)$$

Подставив в выражение для АКФ $\tau = k - u$ и используя вышеприведенное соотношение и свойство линейности оператора математического ожидания E , можно получить следующее соотношение для $R_1(k, \tau)$ (АКФ):

$$R_1(k, \tau = k - u) = E\{x_H(k) \cdot x_H(u)\} = 0.5 \cdot \sigma^2 [k^{2H} + u^{2H} - (k - u)^{2H}] \quad (24)$$

Можно показать, используя приведенные выше формулы и физический смысл величины H , что ее значения лежат в пределах $[0; 1]$. При этом диапазон значений $[0; 0.5]$ соответствует так называемым *антиперсистентным* сигналам или сигналам с *неустойчивым трендом* (называемым также *сигналами с кратковременной памятью*). Последнее означает, что сигнал стремится вернуться к своему среднему значению, и, следовательно, на всей длительности наблюдается чередование участков монотонного возрастания и убывания тренда. Чем ближе значение H к нулю, тем устойчивее общая тенденция к чередованию. Сигналы со значениями H из диапазона $[0.5; 1]$ соответствуют так называемым *персистентным* сигналам или сигналам с *устойчивым трендом*, (называемым также *сигналами с долговременной памятью*) для которых общая тенденция является долговременной и сохраняется на более длительном временном интервале, чем у антиперсистентных сигналов. Чем ближе значение H к 1, тем более устойчивой эта тенденция является.

Данные предположения основываются на том, что сигналы с кратко- и долговременной памятью имеют разные особенности АКФ. Сигнал с кратковременной памятью должен иметь абсолютно суммируемую АКФ (строгое доказательство имеется для реализации бесконечной длины, в практических задачах рассматривается ее конечный фрагмент) [9]:

$$\sum_{u=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} |R(k, u)| < \infty, \quad (25)$$

в то время как для сигнала с долговременной памятью вышеприведенный ряд является расходящимся, т.е. АКФ не суммируема абсолютно. Зная выражение для $R_1(k, \tau)$, непосредственной подстановкой можно убедиться, что для $H \in (0; 0.5]$ АКФ суммируема абсолютно (сигнал с кратковременной памятью), а для $H \in [0.5; 1]$ данное свойство не выполняется (сигнал с долговременной памятью).

Функция $R_1(k, \tau)$ зависит не только от относительного сдвига τ , но и от выбранного отсчета k , и поэтому ДБД является нестационарным сигналом. При этом данный тип сигналов обладает свойством *статистического самоподобия* [7], проявляющимся при масштабировании:

$$E\{x_H(nk)x_H(nu)\} = E\{n^H x_H(k)n^H x_H(u)\}, \quad (26)$$

при $n > 0$, что может быть выведено аналогично АКФ ДБД. Поскольку оба сигнала $x_H(nk)$ и $n^H x_H(k)$ являются гауссовскими, имеют одинаковые среднее значение и ковариацию (n^H - константа, не влияющая на закон распределения), то они обладают свойством самоподобия.

В силу нестационарности ДБД для него нельзя определить энергетический спектр (который будет использоваться в дальнейшем при оценке значений H). Поэтому далее переходят к рассмотрению приращений данного сигнала, которые являются стационарным сигналом. Разностный сигнал ДБД (РДБД) первого порядка определяется как

$$z_H(k) = x_H(k) - x_H(k - 1). \quad (27)$$

Для такого сигнала можно аналогично определить АКФ $R_2(\tau)$, зависящую на этот раз только от величины сдвига τ :

$$R_2(\tau) = 0.5 \cdot \sigma^2 [(\tau + 1)^{2H} + (\tau - 1)^{2H} - 2\tau^{2H}]. \quad (28)$$

Согласно теореме Винера-Хинчина, энергетический спектр $S_H(f)$ РДБД определяется как ДПФ от АКФ $R_2(\tau)$:

$$S_H(f) = C\sigma^2 |e^{i2\pi f} - 1|^2 \sum_{k=-\infty}^{\infty} \frac{1}{|f + k|^{2H+1}} = C\sigma^2 4 \sin^2(\pi f) \sum_{k=-\infty}^{\infty} \frac{1}{|f + k|^{2H+1}}, \quad (29)$$

где $|f| \leq 0.5$ - частота, нормированная к частоте Найквиста. При $H = 0.5$ (что соответствует белому гауссовскому шуму, т.к. $R_2(\tau) = 0$ при $\forall \tau$) сумма приобретает вид

$$\sum_{k=-\infty}^{\infty} \frac{1}{|f + k|^{2H+1}} = \sum_{k=-\infty}^{\infty} \frac{1}{(f + k)^2} = \frac{1}{4 \sin^2(\pi f)}, \quad (30)$$

$$S_H(f) = C\sigma^2 = const. \quad (31)$$

Исследуем значения и характер энергетического спектра для трех различных диапазонов значений показателя Херста H :

1) $H = 0.5$; $S_H(f) = C\sigma^2 = const$, что соответствует равномерному энергетическому спектру в полосе частот $|f| \leq 0.5$.

2) $0 < H < 0.5$; $S_H(0) = 0$, т.е. значение энергетического спектра равно нулю на нулевой частоте и увеличивается по мере приближения к частоте Найквиста. Таким образом, энергетический спектр имеет высокочастотный характер.

3) $0.5 < H < 1$; $S_H(0) = \infty$, т.е. значение энергетического спектра равно бесконечности на нулевой частоте и уменьшается по мере приближения к частоте Найквиста. Таким образом, энергетический спектр имеет низкочастотный характер.

На рис. 4 показано типичное поведение энергетических спектров ЭМ, соответствующих РДБД с разными

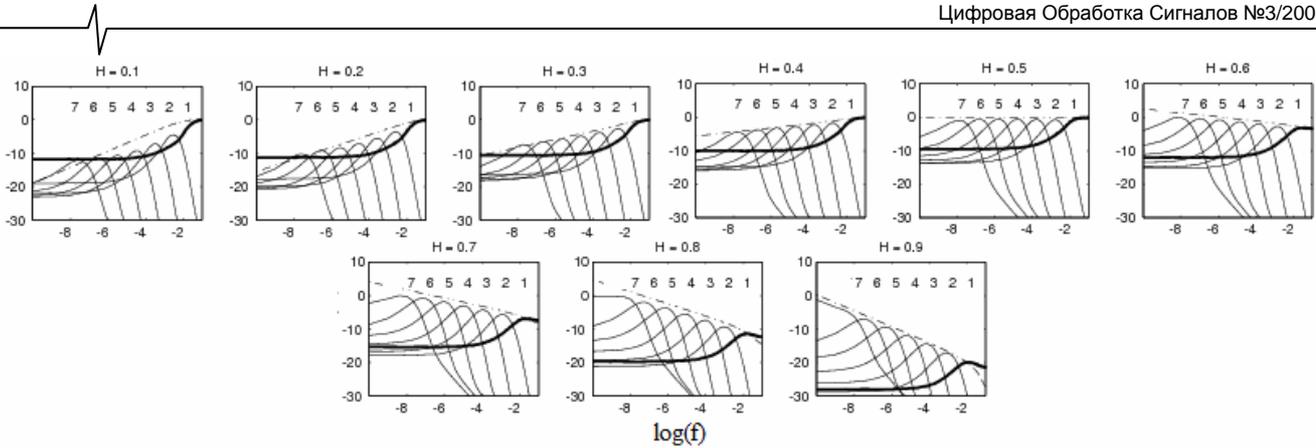


Рис. 4. Энергетические спектры ЭМ, полученных в результате декомпозиции РДБД с разными значениями H (на основе 5000 независимых реализаций длиной по 512 отсчетов)

значениями H . Из данной картины можно сделать следующие выводы:

Во-первых, первая мода РДБД, показанная на рис. 4 жирной линией, независимо от значения H , имеет ярко выраженный высокочастотный характер. При этом максимальное затухание в области низких частот составляет 10 дБ (по отношению к затуханию на частоте Найквиста) для $H < 0.6$, и увеличивается при больших значениях показателя Херста (при $H \geq 0.6$). Все остальные моды по структуре частотного спектра носят полосовой характер.

Во-вторых, с увеличением значения H от 0.1 до 0.9 энергетический спектр последней моды, представляющей собой либо константу, либо тренд среднего значения, постепенно меняет свой характер с полосового на низкочастотный.

В-третьих, огибающая семейства энергетических спектров также имеет определенную тенденцию. Для значений H , находящихся в диапазоне $0 < H < 0.5$, она представляет собой прямую с наклоном, уменьшающимся при приближении H к 0.5. Значению $H = 0.5$ соответствует нулевой наклон. Наконец, при изменении H в пределах $0.5 < H < 1$ огибающая имеет отрицательный наклон, увеличивающийся по абсолютной величине при стремлении H к 1.

В-четвертых, при рассмотрении ЭМ, имеющих полосовой характер спектра (всех, кроме первой и последней), можно проверить, что их ширина полосы частот уменьшается в 2 раза с увеличением номера. На этом основана интерпретация всего разложения как *диадического банка фильтров* [3].

Построение данных семейств спектров ЭМ, соответствующих РДБД, может использоваться для приближенной оценки показателя Херста в силу вышеизложенных свойств и соотношений, а именно на основе поведения огибающей семейства энергетических спектров.

Оценивание показателя Херста на основе метода ДЭМ

Основываясь на том факте, что ДЭМ можно интерпретировать как диадический банк фильтров, можно сделать вывод, что эффективная ширина Фурье-спектра каждой моды уменьшается в число раз, равное степени 2. Исходя из этого, энергетический спектр очередной моды определяется как

$$S_{p',H}(f) = 2^{\alpha(p'-p)} S_{p,H}(2^{p'-p} f), \tag{32}$$

где $\alpha = 2H + 1$ - экспериментально определенная константа, p, p' - номера ЭМ, причем $p' > p \geq 2$. Эта рекуррентная формула определяет энергетический спектр очередной ЭМ на основе известных спектров предыдущих ЭМ.

Зная энергетические спектры всех ЭМ, можно разделить их дисперсии, причем для удобства дальнейших вычислений лучше представить выражение для дисперсий ЭМ в рекуррентном виде:

$$V_H[p'] = \int_{-1/2}^{1/2} S_{p',H}(f) df = 2^{\alpha(p'-p)} \int_{-1/2}^{1/2} S_{p,H}(2^{p'-p} f) df = 2^{\alpha(p'-p)} V_H[p]. \tag{33}$$

В вышеприведенной формуле использовалась теорема Винера-Хинчина, а также то, что интегрирование ведется по симметричному промежутку частот, нормированных к частоте Найквиста. Из нее можно получить в явном виде выражение для дисперсии p -ой моды как функции ее номера:

$$V_H[p] = C \cdot 2^{2Hp}, \tag{34}$$

где C - некоторая константа. Справедливость формулы устанавливается путем проверки рекуррентного соотношения (33). После логарифмирования данного выражения по основанию 2 получается окончательный результат:

$$\log_2 V_H[p] = \log_2 C + 2Hp. \tag{35}$$

Из этого выражения непосредственно вытекает способ экспериментальной оценки показателя Херста. Выполняется ДЭМ для данной реализации ДБД, рассчитываются значения дисперсий каждой моды и их двоичные логарифмы откладываются на графике. После этого на основе регрессии методом наименьших квадратов (МНК) строится линейная зависимость и определяется ее угловой коэффициент. Тогда само значение H будет найдено следующим образом:

$$H = \frac{|k_H|}{2}, \tag{36}$$

где k_H - угловой коэффициент прямой регрессии. Знак модуля нужен потому, что в отдельных случаях линеаризованная зависимость логарифма дисперсии ЭМ от ее номера может иметь отрицательный наклон, а показатель Херста - априорно положительная величина.

Альтернативный способ оценивания показателя Херста с использованием среднего периода ЭМ иллюстрируется следующей формулой:

$$V_H[p] = C' \cdot (T_H[p])^{2H}, \quad (37)$$

где C' - константа, $T_H[p'] = 2^{p'-p} T_H[p]$ для $p' > p \geq 2$, $T_H[k]$ - средний период p -ой ЭМ, определяемый как расстояние между первым и последним нулем (или максимумом, или минимумом) функции, деленное на число нулей (или максимумов, или минимумов). Рекуррентная формула, связывающая средние периоды двух ЭМ, основывается на интерпретации ДЭМ как диадического банка фильтров, и, следовательно, средний период ЭМ связан с ее номером как

$$T_H[p] = 2^p. \quad (38)$$

Из описанных процедур следует, что для оценки значения H необходимо определение дисперсии каждой из мод. Это делается, в основном, по двум формулам, дающим оценку по методу наименьших квадратов и по медиане абсолютных отклонений [7]. Первая имеет вид:

$$V_H[p] = \frac{1}{N-1} \sum_{k=1}^N c_p^2(k). \quad (39)$$

Выражение записано в предположении о центрированности ЭМ, которое обычно выполняется с высокой точностью. Главным преимуществом оценки является ее несмещенность, состоятельность и эффективность. Вторая оценка (40) существует в двух незначительно отличающихся модификациях, приведенных выше. Ее использование объясняется широкой распространенностью гауссовых сигналов, к числу которых относятся ДБД (данная оценка строго применима только к гауссовым сигналам) и высокой робастностью. Применение этой оценки имеет смысл только для начальных уровней разложения, которые являются высокочастотными, сильно осциллирующими и содержат значительную часть шумовой составляющей.

$$\sigma_{1H}[p] = \frac{\text{median}\{c_p - \text{median}(c_p)\}}{0.6745};$$

$$\sigma_{2H}[p] = \frac{\text{median}\{c_p\}}{0.6745}; \quad V_{1,2H}[p] = \sigma_{1,2H}^2[p] \quad (40)$$

Генерации ДБД с заданным значением H может осуществляться несколькими возможными способами. Один из наиболее известных - алгоритм Абри-Селлана-Мейера, использующий для этого биортогональные вейвлеты [6]. В качестве примера работы алгоритма рассмотрим две реализации сигналов ДБД (рис. 5), имеющих показатели Херста, соответственно, $H_1 = 0.3$ и $H_2 = 0.6$. Из рис. 5 видно, что реализация ДБД с $H_2 = 0.6$ является более регулярной и с более устойчивым трендом, чем та, у которой $H_1 = 0.3$.

После применения алгоритма ДЭМ к данным сигналам необходимо выбрать те ЭМ, на основании которых будет осуществляться оценка значения H . В качестве критерия отбора целесообразно использовать *выборочные средние значения* c_i всех выделенных ЭМ, так называемые компоненты с большими номерами, относящиеся к числу трендовых, могут существенно исказить экспериментально построенные зависимости (за счет большого разброса значений в случае монотонных трендовых компонент и нулевого разброса значений в случае констант), что в итоге приведет к неточной оценке показателя Херста. В данном случае к трендовым модам будут относиться те моды, среднее значение которых по модулю больше 0.04 (при таком пороге достигается наибольшая точность оценки). Не следует также забывать, что первая ЭМ не используется при оценивании в силу особенностей ее спектра, который имеет ярко выраженный высокочастотный характер (рис. 4) в отличие от остальных компонент, имеющих полосовой спектр. В таблицах 1 и 2 жирным выделены те ЭМ, которые участвуют в оценивании.

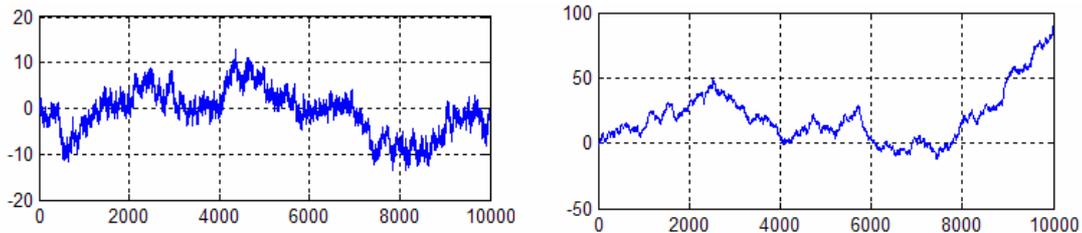


Рис. 5. Две реализации ДБД с показателями Херста $H_1 = 0.3$ (слева) и $H_2 = 0.6$ (справа)

Таблица 1.

Средние значения ЭМ для ДБД с $H_1 = 0.3$.

№ ЭМ (i)	1	2	3	4	5	6	7	8	9	10	11	12
$\bar{c}_i \cdot 10^3$	-0.9	0.8	0.6	-5.2	7.2	15.8	-41.9	-16.1	0.9	6.4	-580.8	-312.2

Таблица 2.

Средние значения ЭМ для ДБД с $H_2 = 0.6$.

№ ЭМ (i)	1	2	3	4	5	6	7	8	9	10
$\bar{c}_i \cdot 10^3$	-0.1	1.1	4.3	-0.4	3.1	-6.4	-17.9	0.5	-2200.7	22038.4

Таблица 3.

Сравнения оценок для сигнала ДБД с $H_1 = 0.3$

Название оценки	На основе ДЭМ-1	На основе ДЭМ-2	На основе биортогональных вейвлетов	На основе R/S анализа
Величина	0.3038	0.2958	0.3010	0.3146

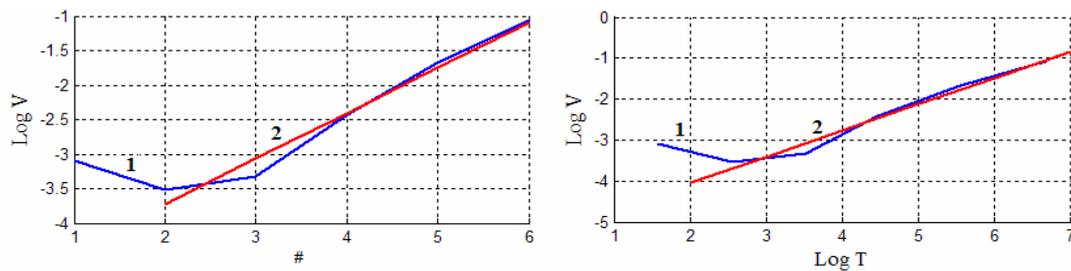


Рис. 6. Зависимость логарифма дисперсии от номера ЭМ (слева) и от среднего периода ЭМ (справа) для ДБД с $H_1 = 0.3$ 1 – экспериментальные значения, 2 – регрессия по МНК

Таблица 4.

Сравнения оценок для сигнала ДБД с $H_2 = 0.6$

Название оценки	На основе ДЭМ-1	На основе ДЭМ-2	На основе биортогональных вейвлетов	На основе R/S анализа
Величина	0.6996	0.6021	0.6454	0.5932

Далее, после определения соответствующих ЭМ (со 2-й по 6-ю для ДБД с $H_1 = 0.3$ и со 2-й по 8-ю для ДБД с $H_2 = 0.6$) строится зависимость логарифма дисперсии ЭМ от их номера (первая оценка) и логарифма дисперсии ЭМ от среднего периода (вторая оценка), как было описано выше. Выше (рис. 6) приведены графики, на основании которых определяются две оценки показателя Херста для сигнала ДБД с $H_1 = 0.3$ (для сигнала с $H_2 = 0.6$ все делается аналогично).

В таблицах 3 и 4 приведены найденные значения оценок H для 2-х рассмотренных реализаций ДБД и здесь же приведены результаты сравнения двух оценок на основе ДЭМ с другими оценками, полученными в результате использования аналогичного ДЭМ разложения сигнала с помощью биортогональных вейвлетов и R/S анализа – наиболее распространенных на сегодняшний день альтернативных способов оценки показателя Херста.

Как видно из полученных результатов, оценка по методу ДЭМ в первом случае незначительно уступает по точности оценке на основе вейвлет-преобразования, однако это справедливо лишь при верном выборе типа вейвлет-базиса, что довольно затруднено на практике в связи с их большим многообразием и необходимостью тщательного и порой затратного перебора. К тому же, правильный выбор соответствующего вейвлет-базиса для обеспечения высокой точности во многом зависит от алгоритма генерации самого ДБД, который зачастую неизвестен. В данных примерах использовался алгоритм генерации на основе биортогональных вейвлетов, что и объясняет столь высокую точность результатов на их основе. Другие варианты (вейвлет-базис Хаара, Добеши и пр.) дают неудовлетворительные результаты (абсолютная погрешность оценивания может достигать 0.1, что неприемлемо). Кстати, даже на основе биортогональных вейвлетов оценка показателя Херста для реализации с $H_2 = 0.6$ проигрывает по точности оценке на основе ДЭМ. Имеет смысл также сравнивать представленные результаты с результатами, полученными на основании R/S анализа [8], который применяется непосредственно к самой реализации во временной области, что часто обеспечивает высокую точность оценивания.

Заключение

В статье представлен принципиально новый подход к анализу нестационарных сигналов сложной внутренней структуры – алгоритм декомпозиции на эмпирические моды. Описаны идеи, заложенные в его основу, а также основные

преимущества по сравнению с существующими методами анализа нестационарных сигналов. В качестве примера рассмотрена задача анализа дробного броуновского движения, в частности, оценивание его показателя Херста. В результате проведенных экспериментов на модельных сигналах достигнута очень высокая точность результатов. Основное достоинство предлагаемого подхода в том, что по сравнению с существующими альтернативными методами (вейвлет-анализ, R/S - анализ) устранен весьма существенный недостаток – необходимость выбора начальных параметров соответствующих алгоритмов оценивания (таких как вид вейвлет-базиса, длина окна во временной области при использовании R/S - анализа). Полученные результаты позволят в дальнейшем применять ДЭМ для оценки показателя Херста реальных сигналов – в медицине (кардиологии, неврологии), экономике, гидрологии и других областях.

Литература

- Huang N., Shen S. Hilbert-Huang Transform and Its Applications, World Scientific, 2005.
- Huang N.E., etc. The Empirical Mode Decomposition and The Hilbert Spectrum for Non-linear and Non-stationary Time Series Analysis, Proc. Royal Soc. London, Vol. 454, pp. 903-995, 1998.
- Flandrin P., Rilling G., Gonsalves P. Empirical mode decomposition as a filter bank, IEEE Signal Process. Lett., Vol. 11, pp. 112-114, 2004.
- Klionsky D.M., Geppener V.V. Empirical mode decomposition in segmentation of slowly and fast changing non-stationary signals, Proceedings of PRIA Conference, Yoshkar-Ola, Russia, 2007.
- Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ., изд. "Мир", Москва, 1990.
- Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в Matlab, изд. "ДМК", Москва, 2008.
- Малла С. Вейвлеты в обработке сигналов, изд. "МИР", Москва, 2005.
- Петерс Эдгар Э. Фрактальный анализ финансовых рынков, изд. "Интернет-Трейдинг", Москва, 2004.
- McLeod H. Time Series Modelling of Water Resources and Environmental Systems, Elsevier, Amsterdam, 1994.
- Longuet-Higgins M.S. The statistical analysis of random moving surface, Proc. R. Soc. Lond., 1957.
- Мэтьюз Д. Г., Финк К. Д. Численные методы. Использование Matlab, изд. дом "Вильямс", Москва-Санкт-Петербург-Киев, 2001.
- Вержбицкий В.М. Основы численных методов, изд. "Высшая школа", Москва, 2002.