

А.Н. Павлов

МЕТОДЫ АНАЛИЗА
СЛОЖНЫХ СИГНАЛОВ

*Учебное пособие
для студентов физического факультета*

НАУЧНАЯ КНИГА

САРАТОВ

2008

УДК [621.391:519.2](075.8)

ББК 32.811 я73

П12

Павлов А.Н.

П12 Методы анализа сложных сигналов: Учеб. пособие для студ. физ. фак. / А.Н. Павлов. – Саратов: Научная книга, 2008. – 120 с.: ил. ISBN 978-5-9758-0811-0

Учебное пособие составлено в соответствии с программой спецкурса «Анализ временных рядов» для физического факультета Саратовского государственного университета им. Н.Г. Чернышевского. Содержит описание 9 лабораторных работ, выполняемых студентами в рамках специализированного практикума «Методы анализа сложных сигналов». Пособие включает описание необходимых теоретических сведений и порядка выполнения работ.

Для студентов, обучающихся по физическим специальностям, а также для аспирантов и научных работников, специализирующихся в области анализа структуры сигналов.

Рецензенты:

В.С. Анищенко, Заслуженный деятель науки РФ, д.ф.-м.н., профессор

Т.Е. Вадивасова, д.ф.-м.н., профессор

УДК [621.391:519.2](075.8)

ББК 32.811 я73

Работа издана в авторской редакции

ISBN 978-5-9758-0811-0

© Павлов А.Н., 2008

Предисловие

В настоящем учебном пособии рассматриваются как классические методы исследования структуры сигналов (корреляционный и спектральный анализ), так и целый ряд специальных методов, предназначенных для изучения процессов с меняющимися характеристиками (флуктуационный анализ, вейвлет-анализ, мультифрактальный анализ) или для перехода к более информативному представлению сигналов (преобразование Гильберта, реконструкция динамических систем по экспериментальным данным). Значительное внимание уделяется современным методам анализа нестационарных случайных процессов. Это связано с тем, что нестационарными являются очень многие сигналы в природе, в частности, практически все процессы, регистрируемые в динамике живых систем. В учебном пособии наряду с широко известными способами обработки экспериментальных данных рассматриваются перспективные подходы к исследованию структуры сложных сигналов, предложенные за последние два десятилетия.

Пособие составлено в соответствии с программой спецкурса «Анализ временных рядов», который читается студентам физического факультета Саратовского государственного университета им. Н.Г. Чернышевского и включает часть материала раздела «Классические методы» и основной материал раздела «Специальные методы» этого лекционного курса. Оно предназначено для студентов-физиков, выполняющих лабораторные работы в рамках специализированного практикума «Методы анализа сложных сигналов».

1 Корреляционный анализ

1.1 Ковариационные и корреляционные функции

Одним из классических методов исследования структуры сигналов является корреляционный анализ, который находит многочисленные применения в задачах, связанных с передачей информации, радарным обнаружением, системами управления, анализом вибраций, выделением слабого сигнала при наличии флуктуаций большой интенсивности, исследованием статистической взаимосвязи процессов в физике, биологии и т.д. Ковариационные (корреляционные) функции допускают разные варианты интерпретации. С одной стороны, они позволяют выявлять степень сходства (линейной зависимости) двух сигналов при варьировании сдвига по времени между ними. В этом случае говорят о взаимных ковариационных (взаимных корреляционных) функциях. С другой стороны, они применимы для изучения взаимосвязи значений одного и того же случайного процесса в разные моменты времени, что позволяет использовать терминологию автоковариационных (автокорреляционных) функций.

Предположим, что анализируется стационарный случайный процесс $X(t)$. Охарактеризовать статистическую взаимосвязь его значений $X_1 = X(t)$ и $X_2 = X(t + \tau)$ можно с помощью автоковариационной функции

$$K(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2, \tau) dx_1 dx_2 \quad (1.1)$$

или автокорреляционной функции

$$\Psi(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - \bar{X})(x_2 - \bar{X}) p(x_1, x_2, \tau) dx_1 dx_2, \quad (1.2)$$

где \bar{X} – математическое ожидание, $p(x_1, x_2, \tau)$ – двумерная плотность распределения вероятности. В задачах анализа структуры сигналов на этапе предварительной обработки данных часто используется приведение исходного процесса к нулевому среднему уровню, в этом случае характеристики $K(\tau)$ и $\Psi(\tau)$ совпадают.

Отметим, что проведение вычислений по формулам (1.1) и (1.2) предполагает наличие статистического ансамбля реализаций случайного процесса с заданной на нем функцией $p(x_1, x_2, \tau)$. Но при обработке экспериментальных данных исследователь имеет дело лишь с одной реализацией

случайного процесса $x^*(t)$, поэтому принципиальным моментом является свойство *эргодичности*, которое означает, что все статистические характеристики анализируемого процесса можно определить по сигналу $x^*(t)$, если заменить процедуру усреднения по ансамблю усреднением по времени. На практике (например, при исследовании физиологических процессов) доказать присутствие эргодических свойств невозможно, но полезно предположить наличие эргодичности, так как в противном случае мы лишимся возможности применять классические методы анализа структуры сигналов в исследованиях сложной динамики объектов живой природы. Обычно используется следующий подход – если процесс можно хотя бы приближенно считать стационарным и эргодическим, то это допущение принимают; если нет, то применяют специальные методы анализа.

1.2 Вычисление ковариационных и корреляционных функций

Ковариационная (или *автоковариационная*) функция эргодического случайного процесса $X(t)$, имеющего конечную длительность T , вычисляется по одной реализации следующим образом:

$$K(\tau) = \frac{1}{T - \tau} \int_0^{T-\tau} x(t)x(t + \tau)dt, \quad (1.3)$$

где $0 \leq \tau < T$. Она характеризует степень линейной зависимости между значениями случайного процесса, сдвинутыми друг относительно друга на интервал времени τ . Для периодической функции времени $K(\tau)$ также является периодической функцией с тем же самым периодом. Если $X(t)$ – случайный процесс, не содержащий периодических компонент, то с ростом τ ковариационная функция стремится к нулю, если $\bar{X} = 0$. Такое поведение $K(\tau)$ свидетельствует о потере статистической взаимосвязи: с увеличением τ предыдущие состояния «забываются».

Аналогичными свойствами обладает *корреляционная* (или *автокорреляционная*) функция (АКФ):

$$\Psi(\tau) = \frac{1}{T - \tau} \int_0^{T-\tau} (x(t) - \langle X \rangle)(x(t + \tau) - \langle X \rangle)dt, \quad (1.4)$$

где $\langle X \rangle$ – среднее по времени значение. С практической точки зрения отличие формул (1.3) и (1.4) состоит в том, что ковариационная функция содержит информацию об энергии процесса $X(t)$, тогда как корреляционная функция теряет часть информации при переходе к нулевому среднему уровню. Примеры расчета АКФ для различных сигналов приведены на рис. 1.1.

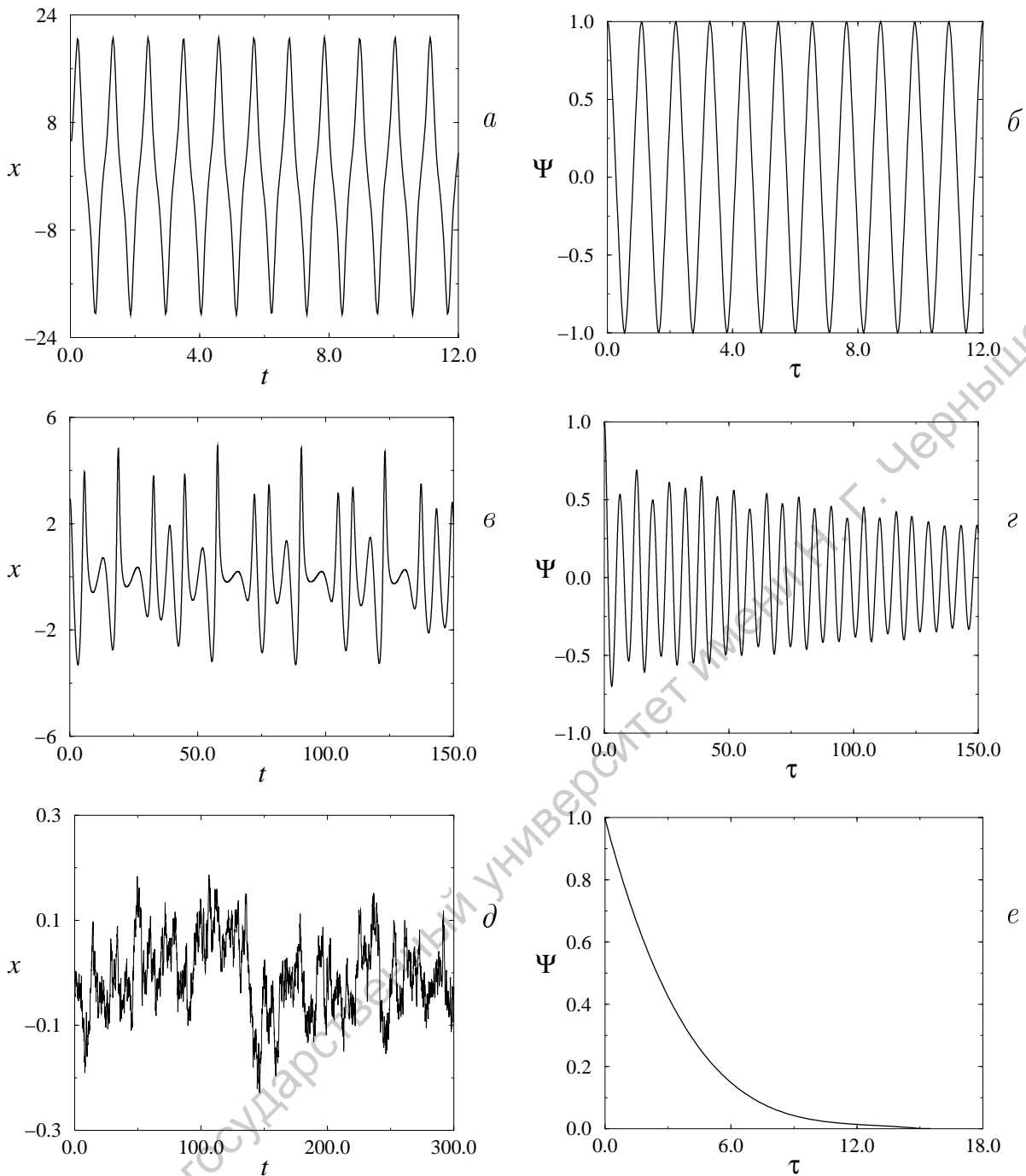


Рис. 1.1. Примеры сигналов и соответствующих им автокорреляционных функций: периодические колебания (а, б), хаотические колебания (в, г), цветной шум (д, е). Для удобства представления результатов проведена нормировка АКФ на значение $\Psi(0)$

Чтобы определить, как быстро теряется статистическая взаимосвязь между значениями случайного процесса $X(t)$ и $X(t + \tau)$, вводят понятие *времени корреляции*. Рассмотрим АКФ, которая стремится к нулю с ростом τ (такое поведение характерно для случайных процессов). Стремление АКФ к нулю не означает, что $\Psi(\tau)$ убывает монотонно; часто оно носит характер затухающих колебаний (рис. 1.2). Пусть $\gamma(\tau)$ – огибающая

функции $\Psi(\tau)$. Время корреляции τ_c случайного процесса $X(t)$ определяется как временной интервал, в течение которого огибающая $\gamma(\tau)$ спадает в h раз, где h – некоторое заданное значение (как правило, $h = e$ или $h = 10$).

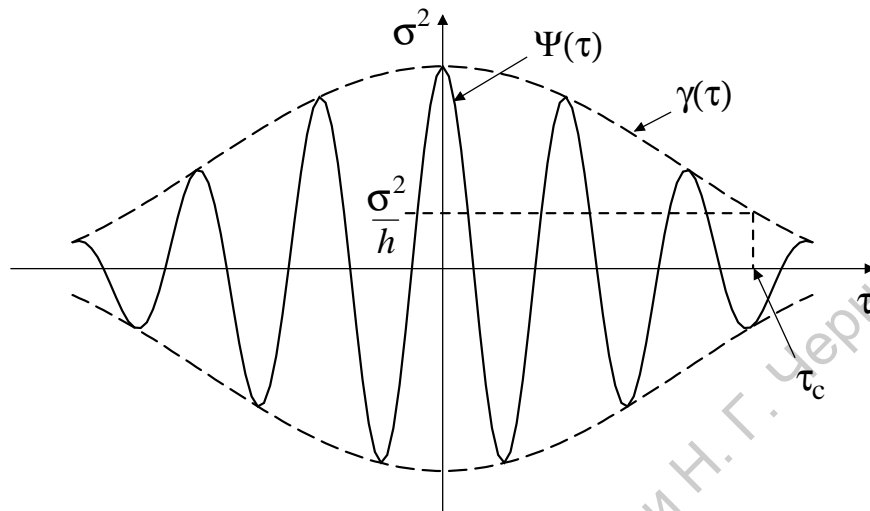


Рис. 1.2. Автокорреляционная функция и ее огибающая

Зависимость $\gamma(\tau)$ может иметь очень сложный вид, при котором ее не удастся аппроксимировать экспоненциальной или степенной функцией (рис. 1.3). В этом случае рассматривают другой способ определения τ_c :

$$\tau_c = \frac{1}{\sigma^2} \int_0^{\infty} \gamma(\tau) d\tau. \quad (1.5)$$

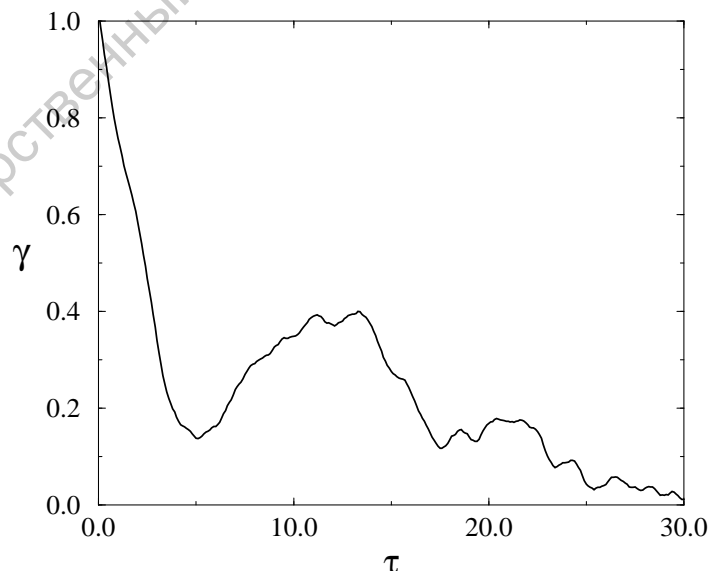


Рис. 1.3. Пример огибающей АКФ, вычисленной по сигналу малой длительности

Чем меньше время корреляции, тем быстрее происходит «забывание» предыдущих значений, то есть случайный процесс является более «хаотичным». Случай $\tau_c = 0$ соответствует белому шуму, для которого «память о прошлом» полностью отсутствует.

Формулы, по которым вычисляются ковариационные и корреляционные функции, были записаны для аналогового сигнала $x(t)$. Модифицируем их теперь для случая временных рядов. С этой целью рассмотрим временной ряд $x(i)$, представляющий собой значения непрерывной функции $x(t)$, дискретизованной с постоянным шагом Δt : $x(i) = x(i\Delta t)$, $i = 1, 2, \dots, N$. Процедура дискретизации всегда проводится при преобразовании аналоговых сигналов в цифровую форму для последующего анализа на компьютере. Длина временного ряда N и шаг дискретизации Δt определяются как практическими соображениями, например, приемлемой продолжительностью эксперимента $T = N\Delta t$, так и техническими возможностями используемой измерительной аппаратуры и характеристиками АЦП. Расчеты ковариационной и корреляционной функций в этом случае проводятся по формулам:

$$K(m) = K(m\Delta t) = \frac{1}{N-m} \sum_{i=1}^{N-m} x(i)x(i+m),$$

$$\Psi(m) = \Psi(m\Delta t) = \frac{1}{N-m} \sum_{i=1}^{N-m} (x(i) - \langle X \rangle)(x(i+m) - \langle X \rangle).$$

Здесь $m = 0, 1, \dots, M$ – число шагов дискретизации, характеризующее сдвиг по времени между значениями случайного процесса (обычно выбирают $M \ll N$). Заметим, что свое максимальное значение АКФ принимает в нуле $\Psi_{max} = \Psi(0)$, и оно совпадает с дисперсией: $\Psi(0) = \sigma^2$. По аналогии, ковариационная функция также принимает максимальное значение в нуле $K_{max} = K(0)$, оно соответствует средней энергии процесса $X(t)$.

Для того, чтобы охарактеризовать степень линейной зависимости между значениями анализируемых данных, удобно пользоваться нормированной корреляционной функцией – *коэффициентом корреляции*:

$$R(m) = \frac{\Psi(m)}{\sigma^2},$$

который может принимать значения от -1 до 1 и не зависит от единиц измерения переменной $x(t)$. Причем, как и сама АКФ, максимальное значение коэффициент корреляции принимает в нуле: $R_{max} = R(0) = 1$.

При проведении корреляционного анализа экспериментальных данных наряду с оценкой скорости спада АКФ (величины τ_c) в ряде задач значительный интерес представляет исследование эффектов длительных корреляций («длительной памяти» о предыдущих состояниях). Это означает, что важно оценивать закономерности поведения АКФ и аппроксимировать зависимость $\gamma(\tau)$ при больших значениях τ . Но в этом случае возникают сложности, связанные с тем, что с увеличением τ возрастают ошибки

вычислений, приводя к недостоверным оценкам. Избежать их возможно только при наличии очень большой выборки N , позволяющей сглаживать зависимость $\Psi(\tau)$ в ходе усреднения. Анализ длительных корреляций играет важную роль в исследованиях процессов медико-биологического происхождения. Например, известно, что при некоторых нарушениях режима работы сердца значительные изменения в структуре сердечного ритма связаны именно с областью длительных корреляций. Поэтому в рамках корреляционного анализа важным параметром является объем выборки – если N невелико, то можно проводить оценки $\Psi(\tau)$ лишь в области малых τ , рассматривать длительные корреляции в структуре экспериментальных данных имеет смысл лишь при больших N .

Порядок выполнения работы

Задание 1. Воспользовавшись программой «acf.x», провести расчеты автокорреляционной функции периодического процесса. Задав в файле параметров «acf.in» значения 100%, 20% и 5% от длины реализации, сопоставить результаты вычислений. Оценить погрешность расчета АКФ в зависимости от объема выборки.

Задание 2. Провести аналогичные вычисления для хаотического процесса. Определить время корреляции как время спадающей АКФ в e раз. Сопоставить оценки, проведенные при различных значениях длительности выборки. Считая, что огибающая АКФ спадает по экспоненциальному закону $\gamma(\tau) \sim e^{-a\tau}$, найти величину a .

№ варианта	динамическая система
1	система Ресслера
2	система Лоренца
3	генератор с инерционной нелинейностью
4	модель нефрона
5	модель нейрона (Хиндмарш-Розе)
6	модель бета-клетки

Задание 3. Вычислить АКФ по экспериментальным данным (сигналу электрокардиограммы). Провести оценки времени корреляции двумя способами: как время уменьшения огибающей АКФ в e раз и по формуле (1.5).

Контрольные вопросы

1. При решении каких задач применяются ковариационные и корреляционные функции?
2. Какие ограничения накладываются на случайный процесс, чтобы его можно было исследовать с помощью корреляционного анализа?
3. Как вычисляются ковариационные и корреляционные функции по временным рядам?
4. В чем состоит физический смысл значений $K(\tau)$ и $\Psi(\tau)$ при $\tau = 0$?
5. Как определить время корреляции случайного процесса?
6. Что характеризует коэффициент корреляции?
7. В чем заключается сложность анализа длительных корреляций?

2 Спектральный анализ

2.1 Преобразование Фурье

Очень многие задачи исследования структуры сигналов связаны со спектральными характеристиками и методами их эффективного вычисления. В отличие от вероятностных методов, описывающих свойства случайных процессов во временной области, спектральный анализ позволяет охарактеризовать *частотный состав* сигнала. Математической основой данного анализа является преобразование Фурье, которое играет важную роль не только при расчете спектров, но и как необходимый промежуточный этап при вычислении преобразования Гильберта, при проведении цифровой фильтрации экспериментальных данных, при определении передаточных и ковариационных функций и т.д.

2.1.1 Непрерывное преобразование Фурье

Сигнал $x(t)$, удовлетворяющий условию абсолютной интегрируемости

$$\int_{-\infty}^{\infty} x^2(t) dt < \infty,$$

может быть представлен интегралом Фурье

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df, \quad (2.1)$$

где

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt. \quad (2.2)$$

Функцию $X(f)$ называют спектральной плотностью амплитуды или спектральной функцией. Формула (2.2) определяет прямое преобразование Фурье, а формула (2.1) – обратное. Отметим, что преобразование Фурье не является локальным: чтобы вычислить $X(f)$, проводится интегрирование в диапазоне $(-\infty, \infty)$, то есть сигнал $x(t)$ должен быть задан на всей бесконечной прямой t .

2.1.2 Непрерывно-дискретное преобразование Фурье

При работе с временными рядами вместо непрерывных функций времени $x(t)$ анализируются дискретные выборки $x(i\Delta t)$, где Δt – шаг дискретизации. Это означает, что для проведения спектрального анализа экспериментальных данных формулы (2.1) и (2.2) должны быть модифицированы на случай дискретных процессов:

$$x(i\Delta t) = \int_{-F_H}^{F_H} X(f)e^{j2\pi fi\Delta t} df, \quad (2.3)$$

$$X(f) = \Delta t \sum_{i=-\infty}^{\infty} x(i\Delta t)e^{-j2\pi fi\Delta t}. \quad (2.4)$$

Записанные выражения (2.3) и (2.4) называют *непрерывно-дискретным преобразованием Фурье*. Обратим внимание на то, что помимо перехода от интегрирования к сумме для прямого преобразования (2.4) произошло еще одно изменение – пределы интегрирования в формуле обратного преобразования стали конечными. В соответствии с (2.3), спектр определен только для частот $|f| \leq F_H = 1/(2\Delta t)$, где величина F_H называется частотой Найквиста. Ее появление является следствием теоремы Котельникова, согласно которой сигнал $x(t)$ может быть однозначно восстановлен по дискретной выборке $x(i\Delta t)$, если максимальная частота в спектре не превышает частоту Найквиста. В противном случае при восстановлении сигнала $x(t)$ по временному ряду будут возникать неоднозначности. Ограничение частотного диапазона при анализе временных рядов отражает известный принцип неопределенности Гейзенберга: появление неопределенности во временной области, связанной с дискретизацией по времени с шагом Δt приводит к неопределенности в частотной области для $f > F_H$.

2.1.3 Дискретное преобразование Фурье

При расчете спектра на компьютере осуществляется дискретизация не только самого сигнала $x(t)$, но и вычисляемой спектральной функции. Кроме того, следует учитывать, что на практике анализируются не бесконечные во времени процессы, а временные ряды конечной длительности N . Поэтому более «практичной» формой записи преобразования Фурье является *дискретное преобразование*, которое имеет следующий вид:

$$x(i\Delta t) = \Delta f \sum_{k=0}^{N-1} X(k\Delta f)e^{j2\pi ik/N}, \quad (2.5)$$

$$X(k\Delta f) = \Delta t \sum_{i=0}^{N-1} x(i\Delta t) e^{-j2\pi ik/N}. \quad (2.6)$$

При записи формул (2.5) и (2.6) переход от интегрирования к суммированию проводился на основе очень простого способа численного интегрирования – метода прямоугольников. Расчеты спектра могут осуществляться непосредственно по формуле (2.6), однако для увеличения скорости вычислений обычно применяется быстрое преобразование Фурье (БПФ), позволяющее за счет специальных приемов получить искомую функцию $X(k\Delta f)$ значительно быстрее. Процедура БПФ входит во все пакеты прикладных программ, предназначенные для проведения спектрального анализа.

2.2 Влияние конечной длины реализации

2.2.1 Непрерывное преобразование Фурье

Формула непрерывного преобразования Фурье была записана для функции $x(t)$, заданной на всей бесконечной прямой t . Очевидно, что это математическая идеализация, которая неприменима для любых реальных сигналов. Регистрируемый в ходе эксперимента процесс $x_T(t)$ всегда имеет конечную длительность T , поэтому пределы интегрирования $(-\infty, \infty)$ на практике являются бессмысленными. Существует простой прием, позволяющий пользоваться формулой (2.2) для сигналов конечной длительности: в подынтегральном выражении рассматривается произведение бесконечного (и неизвестного нам целиком) процесса $x(t)$ и прямоугольного импульса, который «вырезает» лишь небольшой фрагмент $x_T(t)$ (рис. 2.1), то есть $x_T(t) = x(t) \cdot u(t)$. Наличие произведения двух функций в формуле Фурье-преобразования (2.2) приводит к тому, что спектральная плотность амплитуды $X(f)$ будет представлять собой свертку преобразований Фурье сигналов $x(t)$ и $u(t)$. Это означает, что результаты спектрального анализа зависят от Фурье-образа прямоугольного импульса $u(t)$, а следовательно, от длины реализации T . Эффекты конечной длительности сигнала $x_T(t)$ продемонстрированы на рис. 2.2 для случая гармонической функции $x(t) = \cos(2\pi f_0 t)$.

Если для бесконечного процесса (рис. 2.2, а) $X(f)$ представляет собой сумму двух функций Дирака, то для реализации конечной длительности спектр является суммой двух функций вида $(\sin x)/x$. Таким образом, информация о косинусоиде конечной длительности «размывается» по всей частотной области (рис. 2.2, б), в основном концентрируясь вблизи частот $\pm f_0$. Чем меньше T , тем более «размытой» является спектральная функция (рис. 2.2, в).

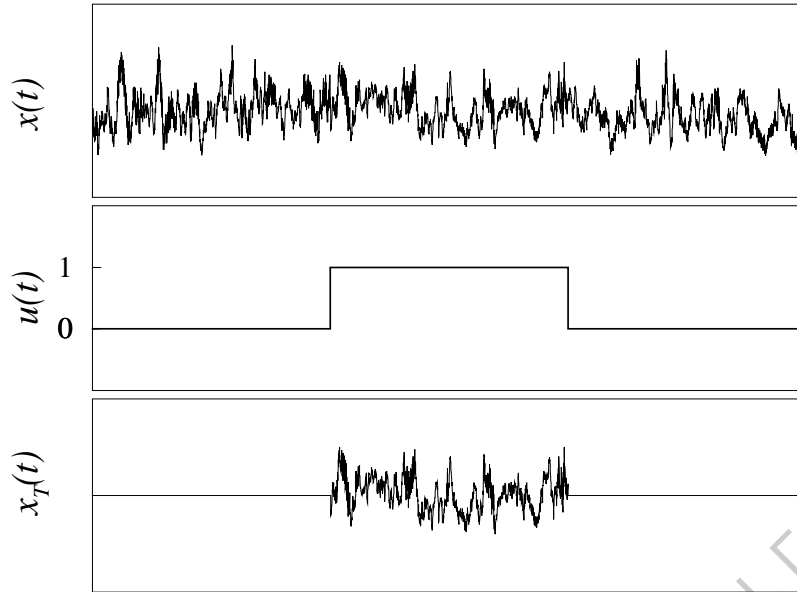


Рис. 2.1. Моделирование случайного процесса конечной длительности $x_T(t)$ путем умножения бесконечной функции $x(t)$ на прямоугольный импульс $u(t)$

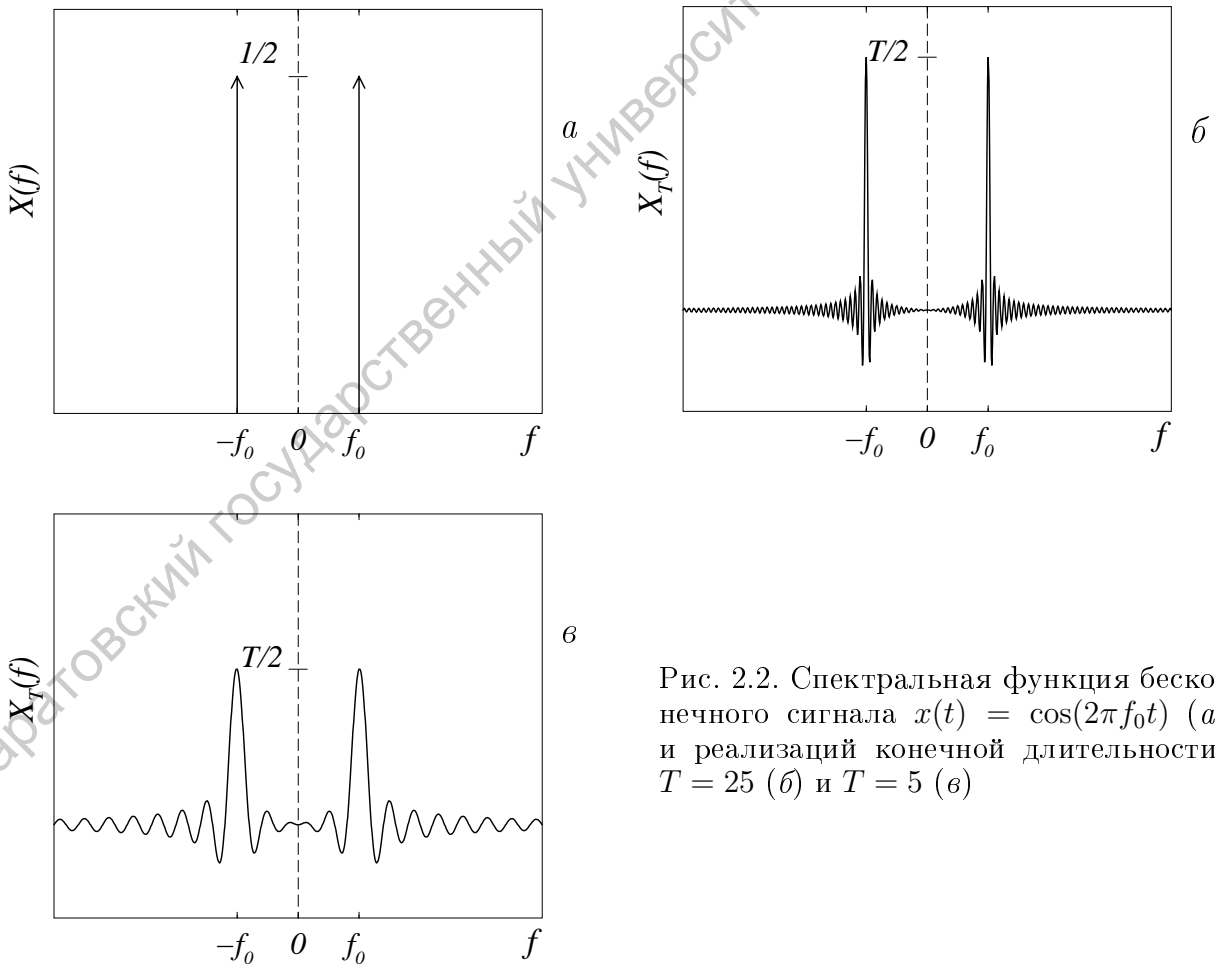


Рис. 2.2. Спектральная функция бесконечного сигнала $x(t) = \cos(2\pi f_0 t)$ (а) и реализаций конечной длительности: $T = 25$ (б) и $T = 5$ (в)

2.2.2 Дискретное преобразование Фурье

Эффекты конечной длины реализации в случае дискретного преобразования Фурье не ограничиваются только «размыванием» спектральной информации, а включают также зависимость от выбора шага дискретизации. В частности, если для гармонической функции подобрать Δt таким образом, что на один период колебаний будет приходиться целое число шагов дискретизации Δt , а длительность временного ряда N включает целое число периодов косинусоиды, то размывания спектральной информации не происходит (рис. 2.3, а). Но если временной ряд включает нецелое число периодов, или на один период функции $x(t) = \cos(2\pi f_0 t)$ приходится нецелое число шагов дискретизации Δt , то спектр усложняется (рис. 2.3, б, в): появляются дополнительные спектральные линии и уменьшается амплитуда спектральных составляющих на частотах $\pm f_0$. Соответствующие эффекты принято называть *эффектами утечки*, они означают, что происходит перетекание энергии из спектрального пика на частоте f_0 («главного лепестка») в соседние спектральные линии («боковые лепестки»).

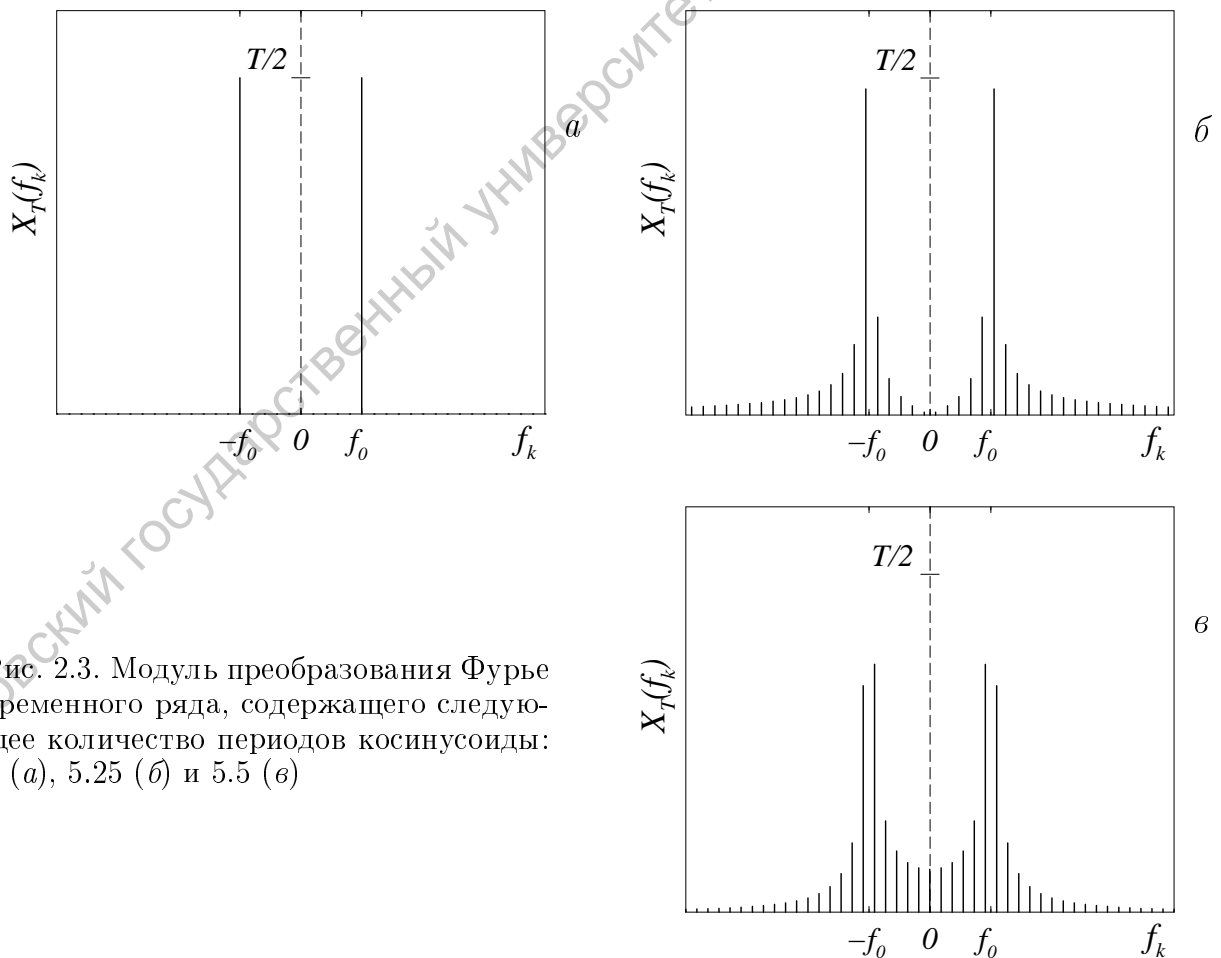


Рис. 2.3. Модуль преобразования Фурье временного ряда, содержащего следующее количество периодов косинусоиды: 5 (а), 5.25 (б) и 5.5 (в)

2.3 Оконный спектральный анализ

Эффекты утечки являются следствием резких разрывов на концах временного ряда, которые можно описать функцией Хевисайда («ступенькой»). Для того, чтобы их избежать, желательно каким-то образом сгладить эти разрывы. С этой целью используется прием, называемый *оконным спектральным анализом*: сигнал $x(t)$ домножают на специально подобранную оконную функцию, вырезающую из бесконечной реализации фрагмент $x_T(t)$. Прямоугольный импульс является одним из вариантов такой функции, но не единственным. На рис. 2.4 представлено несколько примеров часто используемых «окон».

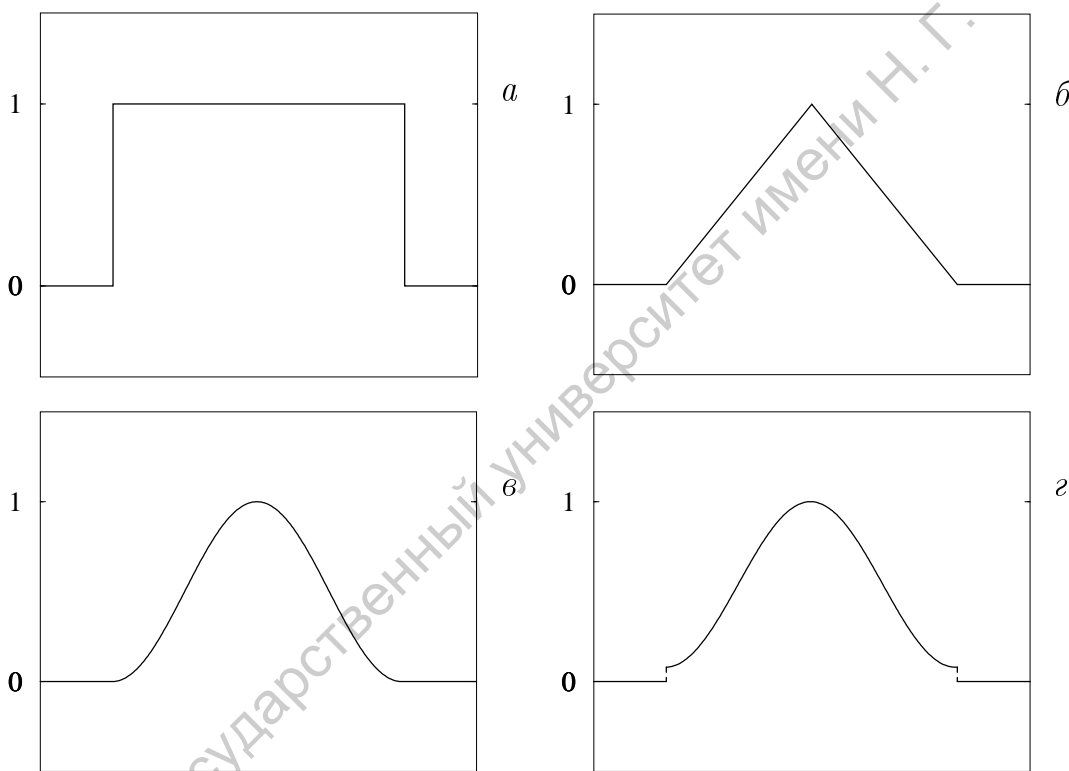


Рис. 2.4. Оконные функции: прямоугольное (а), Барлетта (б), Ханна (в) и Хемминга (г)

Выбор той или иной оконной функции зависит от целей исследования, поскольку каждая из них имеет свои характерные особенности. С точки зрения качества вычисляемого спектра желательно добиться двух вещей: улучшить спектральное разрешение (то есть получить более узкие спектральные пики) и устранить эффекты утечки (подавить боковые лепестки). Оказывается, что одновременно достичь и того, и другого невозможно – устранение эффектов утечки ухудшает спектральное разрешение, и наоборот. Поэтому при проведении вычислений важно определить, что в каждом конкретном случае является более важным. Сравнение характеристик различных оконных функций проводится в таблице 2.1.

Название окна	Ширина главного лепестка по уровню половинной энергии	Максимальный уровень боковых лепестков, дБ	Скорость спада боковых лепестков, дБ
прямоугольное	0.89	-13.3	-6
Барлетта	1.28	-26.5	-12
Ханна	1.44	-31.5	-18
Хемминга	1.30	-43.0	-6

Таблица 2.1. Характеристики оконных функций.

Если в спектре присутствуют две спектральные компоненты, расположенные рядом и имеющие близкие значения амплитуды (рис. 2.5, а), то спектральное разрешение оказывается важнее, и целесообразно применять прямоугольное окно. Для спектра, изображенного на рис. 2.5, б, принципиальным моментом является скорость спада боковых лепестков, чтобы они не наложились на слабую спектральную составляющую, расположенную на расстоянии от главного пика. В этом случае лучше всего выбрать окно Ханна.

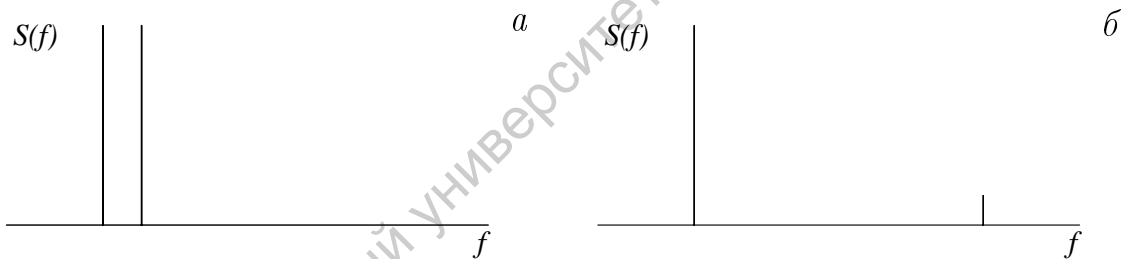


Рис. 2.5. Для изображенных на этом рисунке спектров целесообразно применять разные оконные функции

2.4 Расчет спектра мощности

Преобразование Фурье является характеристикой отдельной реализации $x(t)$. Чтобы охарактеризовать свойства всего случайного процесса (ансамбля реализаций), используется функция *спектральной плотности мощности* или просто *спектр мощности* $S(f)$. Существует несколько вариантов его определения, но чаще всего $S(f)$ вычисляется на основе преобразования Фурье. Рассмотрим реализацию конечной длительности $x_T(t)$:

$$x_T(t) = \begin{cases} x(t), & |t| \leq T/2 \\ 0, & |t| > T/2 \end{cases}$$

Функция $x_T(t)$ удовлетворяет условию абсолютной интегрируемости, и для нее можно записать преобразование Фурье:

$$x_T(t) = \int_{-\infty}^{\infty} X_T(f) e^{j2\pi ft} df.$$

$$X_T(f) = \int_{-T/2}^{T/2} x_T(t) e^{-j2\pi ft} dt,$$

Спектральная функция $X_T(f)$ характеризует спектральные свойства конечного отрезка конкретной реализации случайного процесса. Чтобы охарактеризовать исходный случайный процесс целиком, необходимо перейти к пределу $T \rightarrow \infty$. Этот предельный переход для функции $X_T(f)$ осуществить нельзя, поскольку преобразование Фурье не определено в этом пределе (из-за невыполнения условия абсолютной интегрируемости для реализации стационарного бесконечного во времени процесса). Оказывается, что переход к пределу $T \rightarrow \infty$ возможен, если вместо $X_T(f)$ рассмотреть величину $|X_T(f)|^2/T$.

Спектр мощности случайного процесса определяется следующим образом:

$$S(f) = \lim_{T \rightarrow \infty} \frac{\overline{|X_T(f)|^2}}{T},$$

где черта сверху означает процедуру усреднения по ансамблю реализаций. Таким образом, спектр мощности представляет собой среднее значение от квадрата модуля преобразования Фурье усеченной реализации.

На практике вычисление спектра мощности проводят по одной реализации, заменяя усреднение по ансамблю усреднением по различным участкам сигнала. Иными словами, временной ряд разбивается на неперекрывающиеся либо перекрывающиеся сегменты (которые называются периодограммами), и для каждого из них вычисляется квадрат модуля Фурье преобразования, затем результаты усредняются. Для получения сравнительно гладких функций $S(f)$ обычно рассматривают хотя бы 8–10 периодограмм.

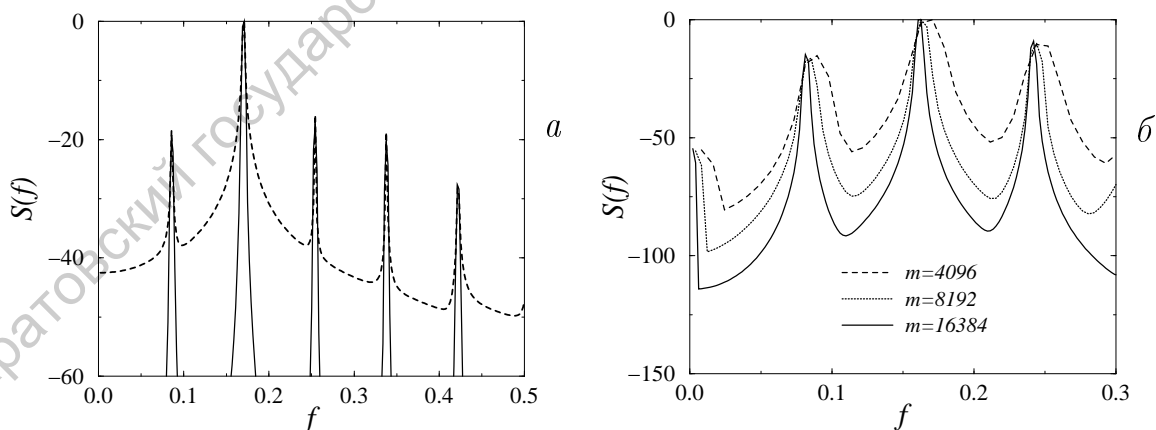


Рис. 2.6. (а) Использование окна Ханна (сплошная линия) вместо прямоугольного (пунктир) позволяет снизить «шумовой пьедестал» спектра для периодических процессов; (б) Спектральное разрешение зависит от длины периодограммы m .

Заметим, что расчет спектра мощности подразумевает поиск компромисса между длиной периодограммы (чем она больше, тем лучше спек-

тральное разрешение) и числом сегментов (с увеличением количества периодограмм спектр становится более гладким из-за лучшего усреднения – это особенно важно для зашумленных данных). Таким образом, чтобы качественно посчитать спектр мощности, целесообразно сопоставить результаты при различных параметрах счета и выбрать оптимальный (см., например, рис. 2.6).

Порядок выполнения работы

Задание 1. Воспользовавшись программой «spectr.x», провести расчеты спектра мощности периодического процесса. Путем перебора предлагаемых оконных функций (прямоугольное, Барлетта, Ханна и Хемминга), а также вариации длины (количества) периодограмм сделать вывод об оптимальных параметрах, позволяющих наиболее качественно посчитать спектр мощности.

Задание 2. Провести аналогичные вычисления для хаотического процесса. Сопоставить, в каком случае (периодическая или хаотическая динамика) оценки спектра мощности более чувствительны к выбору параметров счета.

№ варианта	динамическая система
1	система Ресслера
2	система Лоренца
3	генератор с инерционной нелинейностью
4	модель нефрона
5	модель нейрона (Хиндмарш-Розе)
6	модель бета-клетки

Задание 3. Вычислить спектр мощности по экспериментальным данным (электрокардиограмма, электроэнцефалограмма и электромиограмма). Сопоставить результаты при разных значениях параметров счета.

Контрольные вопросы

1. Чем непрерывно-дискретное преобразование Фурье отличается от непрерывного преобразования?
2. Что такое частота Найквиста, как она определяется?
3. В чем состоят эффекты конечной длины реализации для разных форм преобразования Фурье?
4. Для чего используются оконные функции?
5. По каким характеристикам выбираются оконные функции при проведении вычислений?
6. Как вычислить спектр мощности по экспериментальным данным?
7. Какие параметры влияют на качество вычисления $S(f)$?

3 Преобразование Гильберта

3.1 Преобразование Гильберта и его свойства

Понятия амплитуды, фазы и частоты являются базовыми в теории колебаний. Для гармонической функции $x(t) = A \cos(\omega_0 t + \varphi_0)$ амплитуда A определяет максимальное отклонение от среднего уровня $\langle x \rangle = 0$, фаза $\varphi = \omega_0 t + \varphi_0$ характеризует количество периодов, которые произошли с нулевого момента времени. Фаза линейно растет с коэффициентом пропорциональности, который представляет собой частоту колебаний, и связан с периодом T соотношением $\omega_0 = 2\pi/T$. Амплитуда и частота гармонической функции являются постоянными величинами, отражающими различную информацию: амплитуда описывает энергию сигнала $x(t)$, а частота – его повторяемость во времени. Поэтому амплитуду и частоту (или фазу) можно рассматривать как независимые характеристики.

Для негармонической функции $x(t)$ можно также ввести понятия амплитуды $a(t)$, фазы $\varphi(t)$ и частоты $\omega(t)$, формально представив ее в следующем виде:

$$\begin{aligned}x(t) &= a(t) \cos \varphi(t), \\ \varphi(t) &= \int_{t_0}^t \omega(\theta) d\theta + \varphi_0 = \omega_0 t + \Phi(t), \\ \omega(t) &= \omega_0 + \dot{\Phi}(t).\end{aligned}\tag{3.1}$$

Поскольку эти характеристики в общем случае зависят от времени, чтобы подчеркнуть их отличие от характеристик гармонической функции, часто используют термины *мгновенная* амплитуда, *мгновенная* частота и *мгновенная* фаза. Зависимость от времени функций $a(t)$ и $\omega(t)$ означает, что они могут использоваться для анализа как стационарных, так и нестационарных процессов.

Принципиальным отличием негармонических сигналов от гармонических колебаний является неоднозначность разложения функции $x(t)$ на амплитудную и фазовую составляющие. В частности, если рассмотреть «синусоидальное колебание с медленно меняющейся амплитудой», то данный термин внутренне противоречив подобно «слегка изогнутой прямой линии», и по наблюдаемой реализации такого процесса $x(t) = a(t) \cos \varphi(t)$ нельзя однозначно указать амплитуду и фазу. Это легко объяснить, если

перейти к комплексной форме записи сигнала

$$z(t) = x(t) + j\hat{x}(t), \quad (3.2)$$

то есть дополнить реальное колебание $x(t)$ произвольной мнимой частью $\hat{x}(t)$. В этом случае можно записать

$$z(t) = a(t)e^{j\varphi(t)},$$

и мгновенные амплитуда $a(t)$, фаза $\varphi(t)$ и частота $\omega(t) = d\varphi/dt = \dot{\varphi}$ определяются выражениями:

$$\begin{aligned} a(t) &= \sqrt{x^2(t) + \hat{x}^2(t)}, \\ \varphi(t) &= \arccos \frac{x(t)}{a(t)} = \arcsin \frac{\hat{x}(t)}{a(t)}, \\ \omega(t) &= \frac{\dot{\hat{x}}(t)x(t) - \dot{x}(t)\hat{x}(t)}{a^2(t)}. \end{aligned} \quad (3.3)$$

Следовательно, у сигнала $x(t)$ амплитуда, фаза и частота столь же произвольны как и мнимая часть. Для однозначного определения данных характеристик Габором был предложен подход, основанный на преобразовании Гильберта и позволяющий сконструировать *аналитический сигнал* $z(t)$ в соответствии с формулой (3.2), где функция $\hat{x}(t)$ есть преобразование Гильберта сигнала $x(t)$:

$$\hat{x}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau. \quad (3.4)$$

В формуле (3.4) функция $1/(t - \tau)$ называется ядром преобразования, а интеграл вычисляется в смысле главного значения Коши. Метод аналитического сигнала позволяет однозначно определить $a(t)$ и $\varphi(t)$ по формулам (3.3).

Отметим несколько особенностей данного подхода. С этой целью рассмотрим вначале его отличие от «интуитивного» введения амплитуды и фазы при представлении колебаний на фазовой плоскости. Предположим, что анализируются медленные изменения функций $a(t)$ и $\Phi(t)$ процесса (3.1). Рассмотрим движение изображающей точки на фазовой плоскости $(x, -\dot{x}/\omega_0)$:

$$\begin{aligned} x(t) &= a(t) \cos [\omega_0 t + \Phi(t)], \\ -\frac{\dot{x}(t)}{\omega_0} &= a(t) \sin [\omega_0 t + \Phi(t)]. \end{aligned} \quad (3.5)$$

В этом случае будут справедливы выражения (3.3), если выбрать $\hat{x}(t) = -\dot{x}(t)/\omega_0$. Казалось бы, что отслеживая поведение изображающей точки, можно легко ввести нужные характеристики. Однако при этом возникает

следующая проблема: если использовать определение (3.5), то для *любого* узкополосного колебания функции $a(t)$ и $\Phi(t)$ будут содержать быстрые компоненты по сравнению с частотой ω_0 . Это можно проиллюстрировать на примере амплитудной модуляции:

$$x(t) = (1 + m \cos \Omega t) \cos \omega_0 t, \quad \frac{\Omega}{\omega_0} = \varepsilon \ll 1.$$

В данном случае естественно рассматривать $a(t) = 1 + m \cos \Omega t$, $\varphi(t) = \omega_0 t$, и именно такие результаты позволяет получить метод аналитического сигнала на основе преобразования Гильберта. Но если воспользоваться методом фазовой плоскости (3.5), то будут получены другие выражения:

$$\begin{aligned} a(t) &= 1 + m \cos \Omega t + \frac{\varepsilon}{2} m \sin \Omega t \sin 2\omega_0 t, \\ \Phi(t) &= \frac{\varepsilon}{2} \left[\frac{m \sin \Omega t}{1 + m \cos \Omega t} \right] (1 + \cos 2\omega_0 t), \end{aligned}$$

которые записаны с точностью до членов порядка ε . Появление частот $2\omega_0$ не позволяет воспринимать функции $a(t)$ и $\Phi(t)$ медленными. Эти быстрые слагаемые пропадают при усреднении, и в рамках метода фазовой плоскости осмысленными являются только усредненные характеристики.

Другой особенностью метода преобразования Гильберта является то, что вычисление мгновенных характеристик сигнала $x(t)$ требует его знания на всем интервале времени $-\infty < t < \infty$. Исходя из интуитивных соображений, здесь содержится противоречие с привычными представлениями, ведь такие приборы как амплитудные детекторы, ограничители, системы автоподстройки частоты и т.д. основаны на измерении и подстройке характеристик в данный момент времени, без анализа процесса на всей бесконечной оси t . Несмотря на то, что формально интегрирование проводится в диапазоне $(-\infty, \infty)$, основной вклад в интеграл (3.4) вносит непосредственная окрестность выбранного момента времени (но не слишком малая). Таким образом, можно говорить о квазилокальном характере определяемых величин.

Отметим, что комплексная запись гармонического сигнала получается дополнением действительной функции мнимой частью, отличающейся поворотом фазы на $-\pi/2$. В случае негармонической функции $x(t)$, которая может быть разложена в ряд Фурье, преобразование Гильберта вносит фазовую задержку $-\pi/2$ для всех частотных компонент.

Продемонстрируем это. Согласно (3.4), преобразование Гильберта является сверткой функций $x(t)$ и $1/\pi t$

$$\hat{x}(t) = x(t) * \left(\frac{1}{\pi t} \right),$$

поэтому преобразование Фурье сигнала $\hat{x}(t)$ равно произведению преобразований Фурье функций $x(t)$ и $1/\pi t$, а именно:

$$\begin{aligned} F[\hat{x}(t)] &= F[x(t)] \cdot F[1/\pi t] \\ \hat{X}(f) &= X(f) \cdot [-j \operatorname{sgn}(f)], \end{aligned}$$

где F - оператор преобразования Фурье,

$$\operatorname{sgn}(f) = \begin{cases} 1, & f > 0, \\ 0, & f = 0, \\ -1, & f < 0. \end{cases}$$

Поскольку

$$-j \operatorname{sgn}(f) = \begin{cases} e^{-j\pi/2}, & f > 0, \\ 0, & f = 0, \\ e^{j\pi/2}, & f < 0, \end{cases}$$

а преобразование Фурье сигнала $x(t)$ можно записать в виде

$$X(f) = |X(f)|e^{-j\phi_x(f)},$$

то окончательно получаем:

$$\begin{aligned} \hat{X}(f) &= X(f)e^{-j\phi_b(f)}, \\ |\hat{X}(f)|e^{-j\hat{\phi}_x(f)} &= |X(f)|e^{-j[\phi_x(f)+\phi_b(f)]}, \\ \hat{X}(0) &= 0, \quad \phi_b(f) = \begin{cases} \pi/2, & f > 0, \\ -\pi/2, & f < 0. \end{cases} \end{aligned} \quad (3.6)$$

Отметим при этом, что комплексная функция $\hat{X}(f)$ не является преобразованием Гильберта функции $X(f)$.

Преобразование Гильберта не меняет энергетические характеристики, то есть энергии процессов $\hat{x}(t)$ и $x(t)$ одинаковы. Это означает, что выполняется равенство $|X(f)| = |\hat{X}(f)|$, и в результате преобразования происходит только сдвиг фазы для каждой спектральной компоненты. Более того, данное правило справедливо и для процессов, представляемых интегралом Фурье.

Формально преобразование Гильберта может быть вычислено для произвольного сигнала $x(t)$ (если интеграл (3.4) сходится), однако четкий физический смысл оно имеет только для узкополосных процессов. В этом случае мгновенная амплитуда совпадает с огибающей сигнала $x(t)$, а мгновенная частота соответствует частоте максимума спектра мощности, вычисленного в скользящем временном окне.

Преобразование Гильберта можно вычислить разными способами, в том числе и непосредственно по определению (3.4). Однако это достаточно длительная процедура по сравнению с подходом на основе быстрого преобразования Фурье (БПФ) исходного временного ряда. Как следует из (3.6),

преобразование Фурье аналитического сигнала $z(t)$ можно записать в виде:

$$Z(f) = X(f) + j\hat{X}(f) = \begin{cases} 2X(f), & f > 0, \\ 0, & f < 0, \\ X(0), & f = 0. \end{cases} \quad (3.7)$$

Учитывая это обстоятельство, вычисления целесообразно организовать следующим образом:

1) По временному ряду $x(i) = x(i\Delta t), i = 1, \dots, N$ вычислить $X(f_k)$ с помощью процедуры БПФ. При необходимости (если $N \neq 2^j$) к временному ряду можно добавить нули¹.

2) По формуле (3.7) вычислить Фурье-образ $Z(f_k)$ аналитического сигнала.

3) Вычислить обратное БПФ от $Z(f_k)$, получив в результате аналитический сигнал $z(i)$.

4) Выделить мнимую часть аналитического сигнала $z(i)$, которая является искомым преобразованием Гильберта: $\hat{x}(i) = \text{Im}[z(i)]$.

Для любых произвольных функций $x(t)$ и $y(t)$, имеющих Фурье-образы $X(f)$ и $Y(f)$ и преобразования Гильберта $\hat{x}(t) = H[x(t)]$ и $\hat{y}(t) = H[y(t)]$, справедливы следующие основные свойства:

- *Свойство линейности:*

$$H[ax(t) + by(t)] = a\hat{x}(t) + b\hat{y}(t)$$

- *Свойство сдвига:*

$$H[x(t - a)] = \hat{x}(t - a).$$

- *Свойство четности и нечетности.* Если функция $x(t)$ - четная, то $\hat{x}(t)$ - нечетная, и наоборот.

- *Двойное преобразование Гильберта* приводит к смене знака исходной функции $x(t)$:

$$H[\hat{x}(t)] = -x(t).$$

- *Обратное преобразование Гильберта (H^{-1}):*

$$x(t) = H^{-1}[\hat{x}(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{x}(\tau)}{t - \tau} d\tau.$$

¹На этапе предварительной обработки сигналов часто осуществляется их приведение к нулевому среднему уровню, в этом случае дополнение нулями не приводит к изменению среднего значения

- *Подобие при изменении масштаба:*

$$H[x(at)] = \hat{x}(at).$$

- *Энергетическая эквивалентность:*

$$\int_{-\infty}^{\infty} x^2(t) dt = \int_{-\infty}^{\infty} \hat{x}^2(t) dt.$$

Данное равенство следует из теоремы Парсеваля (энергия процесса равна сумме энергий всех его частотных составляющих) и равенства модулей спектров сигналов $x(t)$ и $\hat{x}(t)$.

- *Свойство ортогональности:*

$$\int_{-\infty}^{\infty} x(t)\hat{x}(t) dt = 0.$$

3.2 Примеры применения

Периодические колебания. В качестве первого примера рассмотрим периодический сигнал, являющийся решением уравнения Ван-дер-Поля:

$$\ddot{x} - \varepsilon(1 - bx^2)\dot{x} + \omega_0^2 x = 0. \quad (3.8)$$

При вариации параметров модель (3.8) демонстрирует различные режимы автоколебаний: как близкие по форме к гармоническому сигналу, так и носящие релаксационный характер. На рис. 3.1 представлены результаты анализа сигнала $x(t)$ данной модели для случая $\varepsilon = 1$, $b = 1$, $\omega_0 = 1$. Видно, что мгновенная амплитуда не является постоянной, а осциллирует во времени (рис. 3.1, *a*). Рост мгновенной фазы (рис. 3.1, *b*) также не является строго линейным: если рассмотреть данную зависимость при увеличении масштаба, то обнаруживаются малые колебания. Чтобы их лучше увидеть, можно вычесть из временной зависимости мгновенной фазы $\varphi(t)$, определенной на основе преобразования Гильберта, уравнение прямой линии, то есть перейти к разности $\varphi(t) - \omega_0 t$ (рис. 3.1, *в*).

Хаотические колебания. В качестве примера системы с хаотической динамикой выберем модифицированную модель генератора с инерционной нелинейностью:

$$\begin{aligned} \frac{dx}{dt} &= mx + y - xz, \\ \frac{dy}{dt} &= -y, \\ \frac{dz}{dt} &= g \left[\frac{x + |x|}{2} x - z \right]. \end{aligned} \quad (3.9)$$

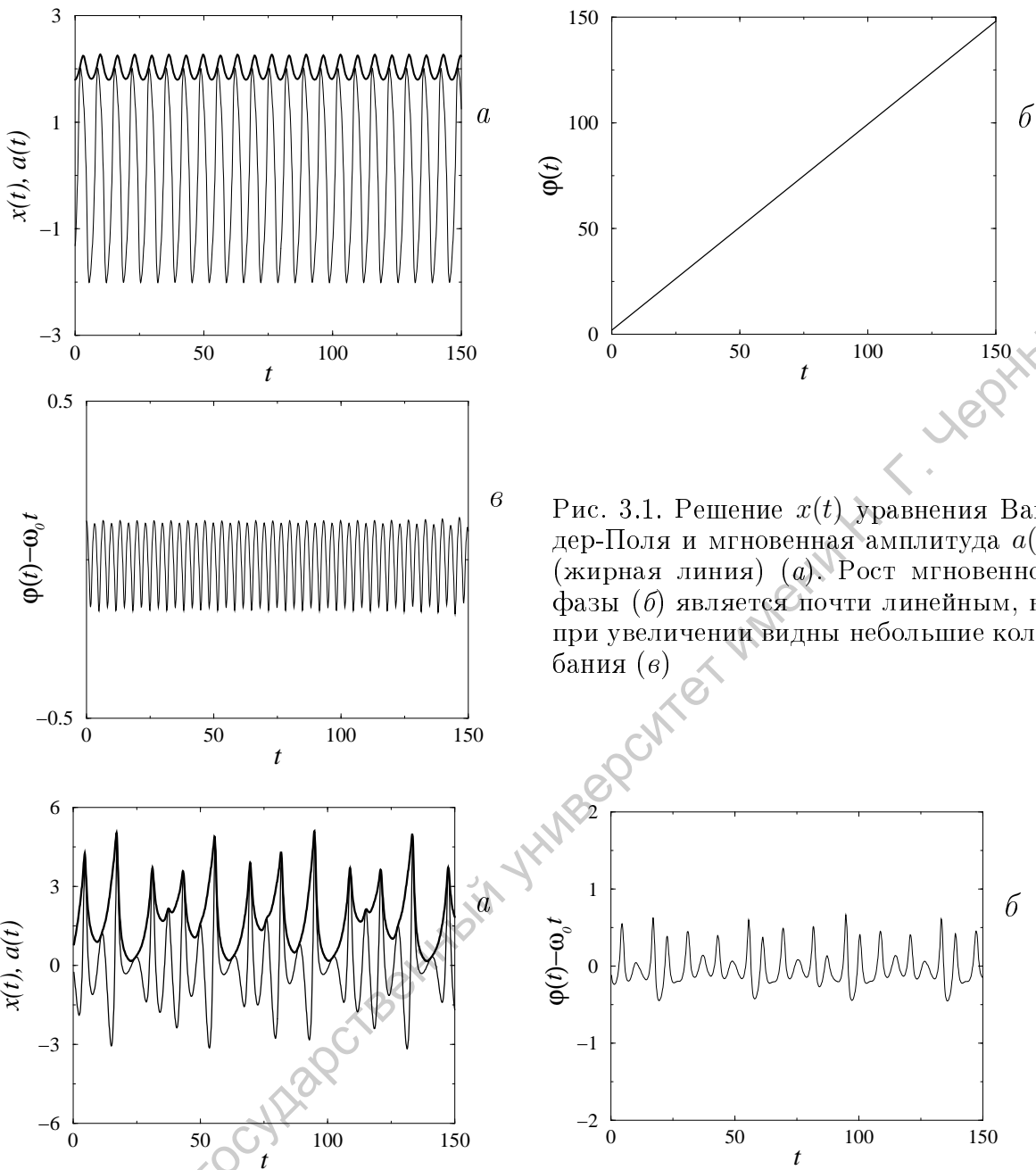


Рис. 3.1. Решение $x(t)$ уравнения Ван-дер-Поля и мгновенная амплитуда $a(t)$ (жирная линия) (а). Рост мгновенной фазы (б) является почти линейным, но при увеличении видны небольшие колебания (в)

Рис. 3.2. Решение $x(t)$ модифицированной модели генератора с инерционной нелинейностью (3.9) и его мгновенная амплитуда $a(t)$ (жирная линия) (а). В динамике мгновенной фазы при увеличении наблюдаются малые колебания (б)

Преобразование Гильберта сигнала $x(t)$ модели (3.9) позволяет получить мгновенную амплитуду и фазу, показанные на рис. 3.2. Хотя фаза растет почти линейно, заметны малые нерегулярные флуктуации этого роста. Такое поведение мгновенной фазы объясняется тем, что колебания в системе (3.9) при рассматриваемых значениях параметров ($m = 1.45$, $g = 0.2$) имеют хорошо выраженный базовый период (или базовую частоту, которая проявляется в виде четкого спектрального пика). Если проанализировать хаотическую динамику, в которой отсутствует характерный временной масштаб, то близкий к линейному рост мгновенной фазы будет

нарушаться.

Динамика биологических систем. В отличие от математических моделей, рассмотренных выше, большинство процессов в динамике живых систем являются нестационарными. Как уже отмечалось, преобразование Гильберта позволяет анализировать нестационарные процессы, однако оно является чувствительным к низкочастотному тренду. Для корректной интерпретации результатов необходимо, чтобы траектория на плоскости (x, \hat{x}) не проходила через начало координат (это приводит к неопределенности для фазы). Если при движении на плоскости (x, \hat{x}) траектория не включает точку $(0, 0)$, то преобразование Гильберта позволяет корректно определить искомые характеристики $a(t)$ и $\varphi(t)$.

В качестве примера биологического процесса рассмотрим сигнал давления жидкости в канальце нефрона – структурного элемента почки, размер которого составляет порядка 100 мкм. Из-за достаточно высокого уровня шума и нестационарности исходного сигнала временная зависимость мгновенной амплитуды является менее плавной, чем в случае моделей (рис. 3.3). Для большей наглядности, на рис. 3.3 представлен сигнал после цифровой фильтрации и удаления среднего уровня.

В настоящее время преобразование Гильберта широко используется для определения степени синхронности между временными рядами, которые порождаются взаимодействующими системами со сложной динамикой. С этой целью отдельно проводится исследование для амплитудных и фазовых составляющих анализируемых процессов. В частности, данный подход позволяет обнаружить взаимосвязь процессов, которая проявляется в синхронности динамики фаз (или фазовой синхронизации) в тех случаях, когда поведение мгновенных амплитуд остается несвязанным. Соответствующий пример представлен на рис. 3.4 для случая динамики двух соседних нефронов, взаимодействующих между собой.

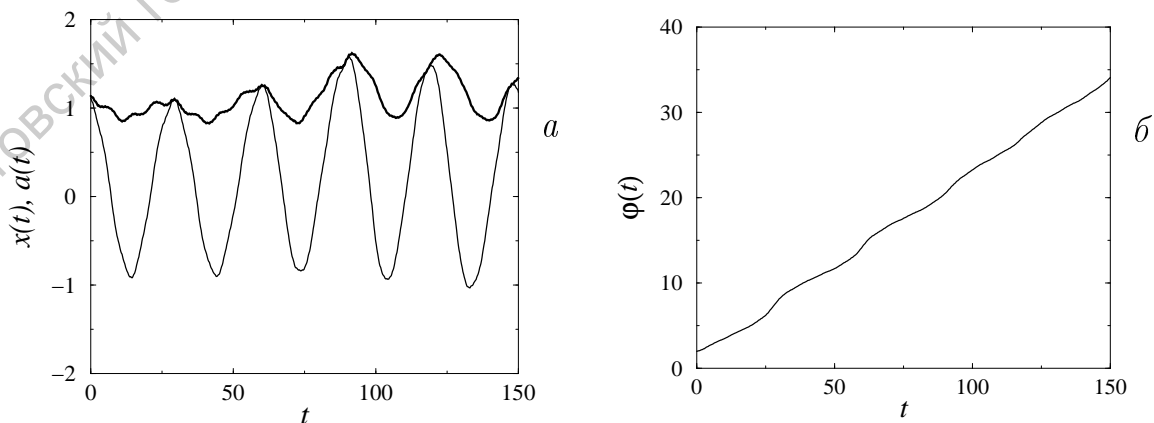


Рис. 3.3. Сигнал давления в канальце нефрона и его мгновенная амплитуда (жирная линия) (а). Близкий к линейному (из-за наличия сильно выраженного ритма) рост мгновенной фазы (б)

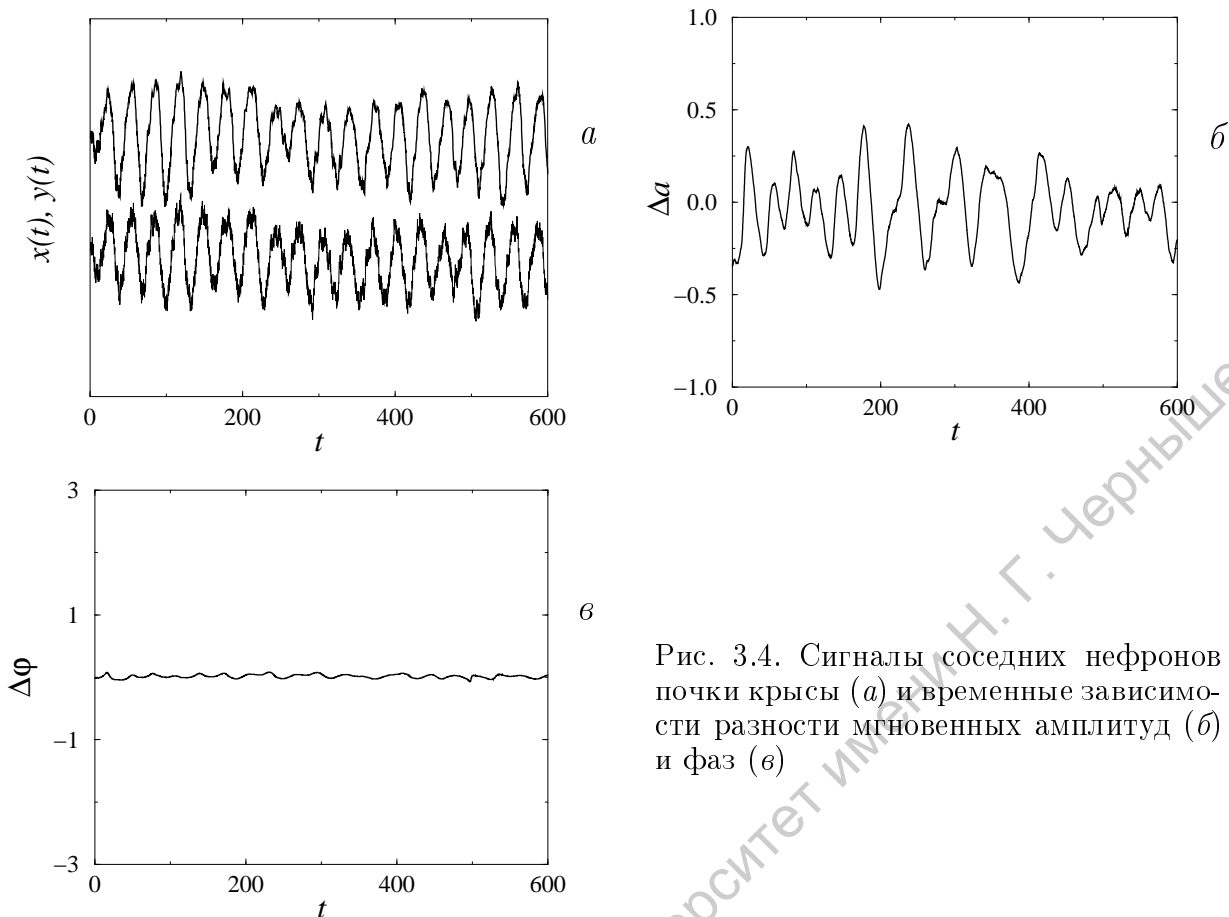


Рис. 3.4. Сигналы соседних нефронов почки крысы (а) и временные зависимости разности мгновенных амплитуд (б) и фаз (в)

Чтобы продемонстрировать наличие фазовой синхронизации и отсутствие эффекта подстройки мгновенных амплитуд анализируемых сигналов $x(t)$ и $y(t)$ (рис. 3.4, а), рассмотрим временные зависимости разностей $\Delta a = a_x(t) - a_y(t)$ и $\Delta\varphi = \varphi_x(t) - \varphi_y(t)$. Согласно рис. 3.4, б, первая зависимость демонстрирует сравнительно сильные вариации во времени, тогда как разность мгновенных фаз (рис. 3.4, в) остается практически неизменной. Таким образом, мы можем говорить о наличии фазовой синхронизации между взаимодействующими структурными элементами почки. Этот вывод можно сделать благодаря возможности разделить амплитудную и фазовую динамику с помощью метода аналитического сигнала.

Порядок выполнения работы

Задание 1. Воспользовавшись программами «hilbert1.x» и «hilbert2.x», провести расчеты преобразования Гильберта одного из предлагаемых процессов. Первая программа осуществляет вычисления непосредственно по определению, вторая использует алгоритм БПФ. Комментарии к программам и значения параметров содержатся в файлах «hilbert1.in» и «hilbert2.in» соответственно. Построить на одном графике исходный сигнал и временные зависимости мгновенной амплитуды колебаний, рассчитанные двумя спо-

собами. Сопоставить результаты. Охарактеризуйте преимущества (недостатки) каждого из рассмотренных подходов.

№ варианта	динамическая система
1	система Ресслера
2	система Лоренца
3	генератор с инерционной нелинейностью
4	модель нефрона
5	модель нейрона (Хиндмарш-Розе)

Задание 2. Провести расчеты мгновенных фаз колебаний в модели двух связанных хаотических систем. Построив временную зависимость разности между мгновенными фазами, сделать вывод о наличии (отсутствии) эффекта фазовой синхронизации. (*Небольшую программу для вычисления разности мгновенных фаз предлагается написать самостоятельно.*)

Задание 3. По аналогии с предыдущим заданием проанализировать динамику двух взаимодействующих нефронов по экспериментальным данным. Оценить влияние шума на расчеты мгновенных амплитуд.

Контрольные вопросы

1. Как можно определить амплитуду и фазу для непериодического процесса?
2. Применимо ли преобразование Гильберта для нестационарных данных?
3. Чем отличается подход на основе преобразования Гильберта от анализа колебаний на фазовой плоскости?
4. Как вычислить преобразование Гильберта с помощью процедуры БПФ?
5. Перечислите основные свойства преобразования Гильберта.
6. В каких задачах применяется метод аналитического сигнала?

4 Корреляционная размерность

4.1 Понятие фрактальной размерности

К настоящему времени разработаны различные методы, позволяющие проводить численное исследование сложных режимов колебаний. С этой целью может использоваться, например, спектральный анализ. Однако расчет спектров не дает возможности отличить хаотическую динамику в системах с малым числом степеней свободы от динамики многомерных систем. Визуально очень похожие спектры могут быть получены как для детерминированных хаотических колебаний, так и для случайных процессов. Определенные преимущества имеет исследование траекторий в фазовом пространстве с помощью сечений Пуанкаре. Но при этом удастся получать только качественную информацию, причем для наглядного представления о геометрии анализируемых объектов размерность фазового пространства не должна быть больше трех. Образом хаотического режима колебаний в фазовом пространстве является странный аттрактор – геометрически очень сложный объект. Особенности его геометрии можно количественно охарактеризовать с помощью *фрактальных размерностей*.

Существуют различные виды размерностей. *Размерность фазового пространства* (n) соответствует количеству переменных, определяющих состояние динамической системы (ДС). Если математическая модель ДС задана в виде системы обыкновенных дифференциальных уравнений первого порядка

$$\frac{d\vec{x}}{dt} = \vec{F}(\vec{x}, \vec{\mu}), \quad \vec{x} \in R^n, \quad \vec{\mu} \in R^k, \quad (4.1)$$

то n определяется числом данных уравнений. Чтобы охарактеризовать некоторое множество S в пространстве R^n , можно воспользоваться *топологической размерностью* d_T . Она равна минимальному количеству переменных, которое необходимо указать, чтобы обозначить положение точки на множестве S . Величина d_T (также как и n) принимает только целые положительные значения. Если речь идет о линии, то $d_T = 1$; для поверхности $d_T = 2$ и т.д.

Другое определение размерности было предложено Хаусдорфом. Пусть S – некоторое множество в пространстве R^n . Предположим, что мы покрываем данное множество кубиками $\{B_i\}$ с величиной ребра, не превышающей некоторое значение ε . При этом каждая точка множества S должна

обязательно попасть в тот или иной кубик. Тогда мера Хаусдорфа l_δ вводится следующим образом:

$$l_\delta(S) = \liminf_{\varepsilon \rightarrow 0} \sum_{B_i \in K(\varepsilon)} |B_i|^\delta.$$

Здесь \inf – минимальное значение (нижняя грань) по всем возможным покрытиям $K(\varepsilon)$ множества S кубиками $\{B_i\}$; $|B_i|$ – величина ребра кубика ($|B_i| \leq \varepsilon$). Указанный предел зависит от параметра δ . Размерность Хаусдорфа d_H представляет собой такое значение δ , при котором величина $l_\delta(S)$ является конечной:

$$\begin{cases} \delta > d_H(S) \implies l_\delta(S) = 0 \\ \delta < d_H(S) \implies l_\delta(S) = +\infty. \end{cases}$$

Согласно данному определению, d_H может принимать нецелые значения. В общем случае, если размерность является нецелой, ее называют *фрактальной*. Соответственно, объекты, характеризующиеся нецелой размерностью, называют *фракталами*. Наличие нецелой размерности является типичной особенностью хаотических аттракторов.

Понятие размерности Хаусдорфа хорошо определено с точки зрения математики, но ее чрезвычайно сложно вычислить. Поэтому обычно используют более «практичные» определения фрактальных размерностей. Одним из таких «практичных» определений является *емкость* (или емкостная размерность D_0). Пусть S – некоторое множество в пространстве R^n , которое покрывается кубиками размера ε . Если обозначить через $N(\varepsilon)$ число кубиков, необходимых для покрытия всего множества, то емкость представляет собой предел следующего вида:

$$D_0 = \lim_{\varepsilon \rightarrow 0} \frac{\ln N(\varepsilon)}{\ln (1/\varepsilon)}. \quad (4.2)$$

По сути, эта величина характеризует, как меняется число элементов покрытия при изменении ε :

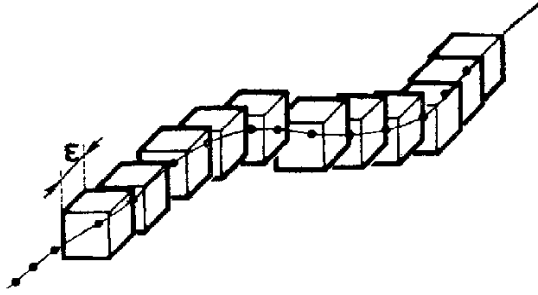
$$N \sim \varepsilon^{-D_0}.$$

Если в качестве S рассматривается единственная точка, то $N(\varepsilon) = 1$ и не зависит от ε :

$$N \sim \varepsilon^0 \implies D_0 = 0.$$

Если анализируется отрезок линии (рис. 4.1) длины L , то

$$N(\varepsilon) = \frac{L}{\varepsilon} \sim \varepsilon^{-1} \implies D_0 = 1.$$

Рис. 4.1. Покрытие отрезка линии кубиками размера ε

Для поверхности площади P :

$$N(\varepsilon) = \frac{P}{\varepsilon^2} \sim \varepsilon^{-2} \implies D_0 = 2.$$

Во всех этих случаях емкость D_0 совпадает с топологической размерностью d_T и является целым числом. В качестве примера объекта с дробной размерностью D_0 (фрактального объекта) рассмотрим канторово множество. Процедура его построения состоит в следующем. Берется отрезок единичной длины $[0, 1]$, разбивается на 3 равные части, и средняя из них выбрасывается. В результате на первом шаге процедуры построения канторова множества мы получаем два отрезка $[0, 1/3]$ и $[2/3, 1]$ длиной $\varepsilon = 1/3$ (рис. 4.2). На следующем шаге каждый из этих отрезков вновь разбивается на 3 равные части, и опять выбрасывается средняя часть. Такая процедура продолжается со всеми оставшимися отрезками. Если для покрытия множества на некотором шаге k используются кубики с величиной ребра $\varepsilon = 1/3^k$, то необходимое количество кубиков составит $N(\varepsilon) = 2^k$.

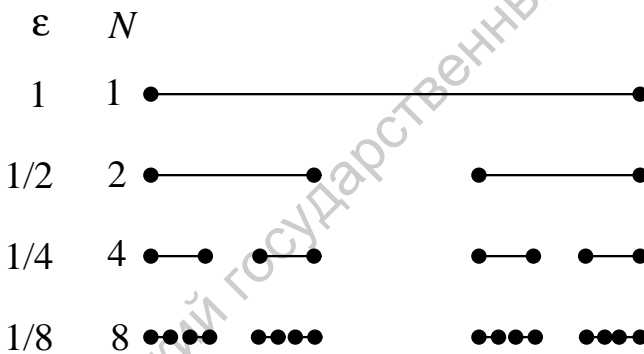


Рис. 4.2. Процедура построения канторова множества

Таким образом:

$$D_0 = \lim_{\varepsilon \rightarrow 0} \frac{\ln 2^k}{\ln 3^k} = \frac{\ln 2}{\ln 3} \approx 0.63.$$

Если говорить о геометрии данного объекта, то канторово множество есть нечто большее, чем точка (для которой $D_0 = 0$), но нечто меньшее, чем интервал ($D_0 = 1$). Для канторова множества емкость совпадает с размерностью Хаусдорфа $D_0 = d_H$, но не совпадает с топологической размерностью ($d_T = 0$). В общем случае справедливо неравенство

$$d_T \leq d_H \leq D_0.$$

Еще одним примером фрактала служит объект, называемый «снежинкой» (или кривой Кох). «Снежинка» имеет удивительное свойство: ее периметр является бесконечным, тогда как сама она занимает ограниченную площадь на плоскости. Процедура построения «снежинки» состоит в следующем. Рассматривается равносторонний треугольник, каждая сторона которого разбивается на 3 равные части, и к средней части «пририсовывается» равносторонний треугольник меньшего размера (рис. 4.3). Затем данная процедура повторяется с каждым отрезком полученной геометрической фигуры.

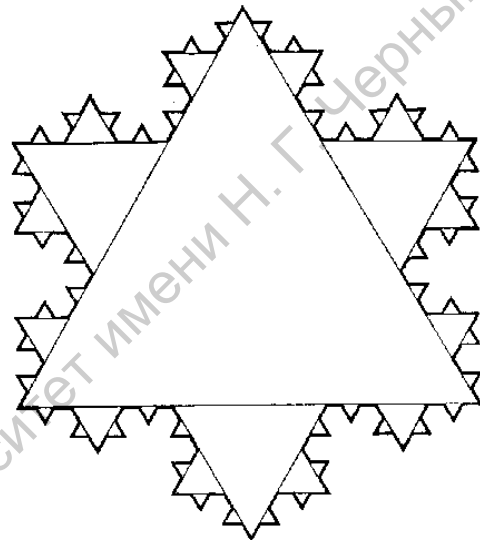


Рис. 4.3. «Снежинка»

На рисунках 4.4–4.6 изображены еще несколько примеров процедуры построения фрактальных объектов. Для получения «салфетки» Серпинского (рис. 4.4) из черного равностороннего треугольника вырезается треугольник меньшего размера (с длиной стороны, уменьшенной в 2 раза); та же процедура осуществляется с каждым из вновь образованных треугольников. С «салфеткой» Серпинского тесно связан другой пример – «ковер» Серпинского. В отличие от рис. 4.4 в качестве исходного объекта выбирается квадрат, из которого вырезаются квадраты меньшего размера (рис. 4.5). Если осуществлять данную операцию бесконечное число раз, то черные участки исходного квадрата исчезают, а полный периметр «дыр» в «ковре» Серпинского становится бесконечным.

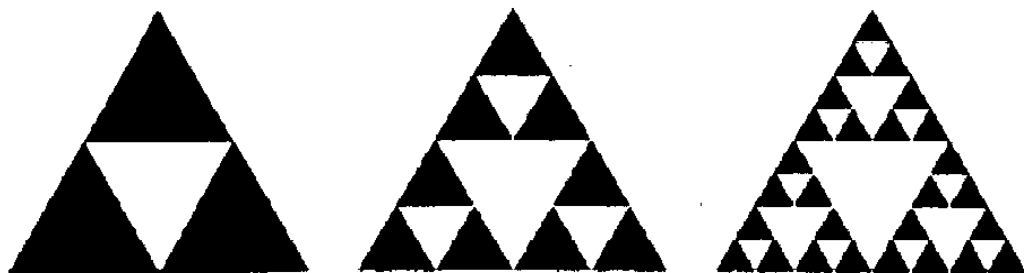


Рис. 4.4. «Салфетка» Серпинского

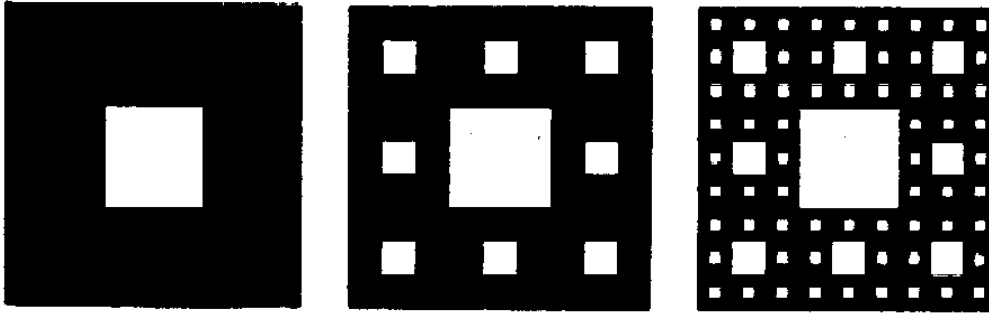


Рис. 4.5. «Ковер» Серпинского

На рис. 4.6 изображена процедура построения квадратной кривой Кох. В качестве исходного объекта выбирается квадрат, каждая сторона которого заменяется ломаной, состоящей из 8 отрезков равной длины. Во всех рассмотренных примерах значение D_0 легко можно вычислить аналитически.

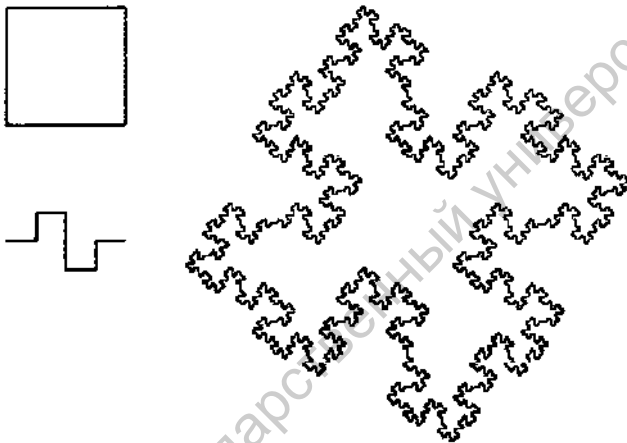


Рис. 4.6. Квадратная кривая Кох

К сожалению, во многих случаях, представляющих практический интерес (например, если проводится анализ сложной геометрии объектов фазового пространства), вычислить емкость можно только численно, и при этом определение D_0 непосредственно по формуле (4.2) зачастую осложняется очень медленной сходимостью отношения $\ln N(\varepsilon)/\ln(1/\varepsilon)$ к пределу $\varepsilon \rightarrow 0$ (если речь идет о расчете емкости объекта в фазовом пространстве размерности $n > 2$). Кроме того, D_0 не зависит от вероятности посещения тех или иных областей фазового пространства, то есть не учитывает статистические свойства потока, обусловленные динамикой системы (4.1). Поэтому на практике вместо емкости предпочитают вычислять корреляционную размерность ν , которую можно легче и быстрее оценить численно. Для объектов с целой размерностью $\nu = D_0$. В более общем случае $\nu \leq D_0$.

4.2 Расчет корреляционной размерности аттрактора

Рассмотрим ДС вида (4.1), демонстрирующую хаотический режим колебаний. Решением системы уравнений (4.1) является фазовая траектория $\vec{x}(t)$. При проведении численных исследований осуществляется анализ не непрерывных функций времени, а дискретных значений $\vec{x}(i\Delta t)$, где Δt – шаг дискретизации (он может быть выбран равным, например, шагу интегрирования). Таким образом, мы имеем дело с многомерными временными рядами $x_j(i\Delta t), i = 1, \dots, N, j = 1, \dots, n$, где N – число точек анализируемых данных. Предположим, что фазовая траектория $\vec{x}(t)$ n -мерной ДС (4.1) лежит на некотором аттракторе A и возвращается в любую сколь угодно малую окрестность произвольной точки данного аттрактора. Обозначим через $k_i(\varepsilon)$ число точек \vec{x}_j внутри шара радиуса ε с центром в точке \vec{x}_i . Вероятность того, что точка на аттракторе попадет в этот шар (меру шара) можно приближенно оценить как $P_i(\varepsilon, N) = k_i(\varepsilon)/N$. Эта вероятность, усредненная по всему аттрактору, называется корреляционным интегралом, и в качестве ее оценки можно использовать

$$C(\varepsilon, N) = \frac{1}{N} \sum_i P_i(\varepsilon, N).$$

Фактически $C(\varepsilon, N)$ – это просто отношение числа пар точек, расстояние между которыми $|\vec{x}_j - \vec{x}_i|$ меньше ε , к полному числу пар. Если обозначить среднее число ε -«соседей» точки через $k(\varepsilon) = \sum_i k_i(\varepsilon)/N$, то $C(\varepsilon, N) = k(\varepsilon)/N$. Определение корреляционного показателя ν основано на свойствах масштабной инвариантности корреляционного интеграла:

$$\nu = \lim_{\varepsilon \rightarrow 0} \left[\frac{\log[\lim_{N \rightarrow \infty} C(\varepsilon, N)]}{\log \varepsilon} \right],$$

то есть наличия степенной зависимости вида $C(\varepsilon, N) \sim \varepsilon^\nu$, справедливой при малых значениях параметра ε . Часто формулу для ν записывают иначе:

$$\nu = \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\ln C(\varepsilon, N)}{\ln \varepsilon}, \quad (4.3)$$

$$C(\varepsilon, N) = N^{-2} \sum_{i \neq j} \theta(\varepsilon - |\vec{x}_i - \vec{x}_j|), \quad (4.4)$$

где θ – функция Хевисайда, $\vec{x}_i = \vec{x}(i\Delta t)$, $\vec{x}_j = \vec{x}(j\Delta t)$. В численных экспериментах N всегда конечно, и оба предела $\varepsilon \rightarrow 0$ и $N \rightarrow \infty$ являются бессмысленными; поэтому отношение $\ln C(\varepsilon, N)/\ln \varepsilon$ рассматривается при не слишком малых значениях ε .

Расчет корреляционной размерности может производиться непосредственно по массиву векторов $\vec{x}(i\Delta t)$. Однако для оценки ν аттрактора A целесообразнее осуществлять переход к сечению Пуанкаре и вычислять корреляционную размерность точек в сечении (ν_c). Размерность аттрактора A в этом случае определяется путем увеличения значения ν_c на единицу: $\nu = \nu_c + 1$. Отметим одно важное обстоятельство: вычисление ν позволяет получать некоторую оценку снизу значения емкости D_0 либо размерности Хаусдорфа аттрактора A . Причем, *если корреляционная размерность рассматривается именно как оценка d_H , нужно принимать во внимание, что ν может существенно отличаться от размерности Хаусдорфа* (в частности, для хаотического аттрактора в системе Ресслера ν может принимать значение ≈ 1.6 , тогда как $d_H > 2$). В связи с этим привлечение алгоритма расчета корреляционной размерности является более привлекательным для исследования ДС по одномерным «проекциям» фазовой траектории (скалярным временным рядам). В этом случае вычисление корреляционной размерности позволяет сделать вывод о наличии (или отсутствии) маломерной динамики системы ($\nu < 4$). Метод расчета ν по скалярному временному ряду имеет лишь небольшое отличие от рассмотренного выше подхода: в формуле для корреляционного интеграла (4.4) вместо массива векторов $\vec{x}(i\Delta t)$, который предполагается неизвестным, рассматривается другой массив $\vec{z}(i\Delta t)$, полученный в результате реконструкции аттрактора A (задача реконструкции будет подробно рассмотрена далее). Сама процедура расчета корреляционного интеграла при этом не меняется.

4.3 Реконструкция аттрактора

Реконструкция представляет собой метод исследования динамических систем вида (4.1) по временным зависимостям одной или нескольких переменных x_j . Ранее считалось, что для изучения динамики автоколебательной системы в терминах фазового пространства необходимо знание всех координат, определяющих ее состояние. Однако в начале 80-х годов данное представление подверглось пересмотру. Было обосновано, что фазовый портрет динамической системы (4.1) может быть восстановлен по временному ряду c_i (представляющему собой дискретизованную с шагом Δt временную зависимость одной из «наблюдаемых» переменных ДС: $c_i = c(i\Delta t) = x_j(i\Delta t)$, $i = 1, \dots, N$), если в качестве недостающих координат вектора состояния используется тот же самый ряд c_i , взятый с некоторым запаздыванием. В 1981 году Такенсом была доказана теорема, утверждавшая следующее. Предположим, что фазовая траектория $\vec{x}(t)$ некоторой n -мерной ДС (4.1) лежит на аттракторе A , принадлежащем гладкому d -мерному многообразию. Тогда по одномерной «проекции» $c(t)$ этой траектории методом задержки можно получить m -мерную реконструкцию A_R

исходного аттрактора A как множество векторов $\vec{z}(t)$ в R^m при $m \geq 2d + 1$:

$$\vec{z}(t) = \Lambda_m(\vec{x}(t)) = \{c(t), c(t + \tau), \dots, c(t + (m - 1)\tau)\}. \quad (4.5)$$

Согласно теореме Такенса, отображение $\Lambda_m : A \rightarrow A_R$ является гладким и обратимым на A_R почти для любого τ (если $N \rightarrow \infty$). Это означает, что множество восстановленных \vec{z} -векторов может быть обработано вместо исходных (и часто неизвестных нам) векторов \vec{x} с целью вычисления метрических (фрактальные размерности) или динамических (ляпуновские показатели) характеристик аттрактора A системы (4.1).

В действительности на компьютере анализируется ряд значений переменной $c(t)$ в дискретные моменты времени $i\Delta t$, поэтому реконструируемое множество векторов также является дискретным $\vec{z}(i\Delta t)$, а величина τ , называемая задержкой, имеет вид $\tau = k\Delta t$, где k — целое число. Иными словами, на практике равенство (4.5) может быть переписано следующим образом:

$$\vec{z}_i = \Lambda_m(\vec{x}_i) = \{c_i, c_{i+k}, \dots, c_{i+(m-1)k}\},$$

где нижний индекс i соответствует моменту времени $i\Delta t$.

Техника реконструирования состоит в выборе значений задержки τ , размерности пространства вложения m и в формировании массива векторов $\vec{z}(i\Delta t)$. Предполагается, что полное время наблюдения $T_{obs} = N\Delta t$, и число точек N достаточно велики, чтобы по траектории можно было судить о важнейших свойствах интересующего нас аттрактора.

В качестве примера рассмотрим уравнения модели Лоренца:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= rx - y - xz, \\ \frac{dz}{dt} &= -bz + xy \end{aligned} \quad (4.6)$$

при значениях параметров $\sigma = 10$, $r = 28$, $b = 8/3$, соответствующих режиму динамического хаоса. Проекция фазового портрета хаотического режима на плоскость (x, y) изображена на рис. 4.7, *a*. По временной зависимости $x(t)$ уравнений (4.6) можно осуществить реконструкцию методом задержки (4.5) — рис. 4.7, *б, в*. В соответствии с теоремой Такенса мы ожидаем, что по восстановленному аттрактору (см. рис. 4.7, *в*) могут быть вычислены такие характеристики анализируемого режима динамики (см. рис. 4.7, *a*), как фрактальные размерности.

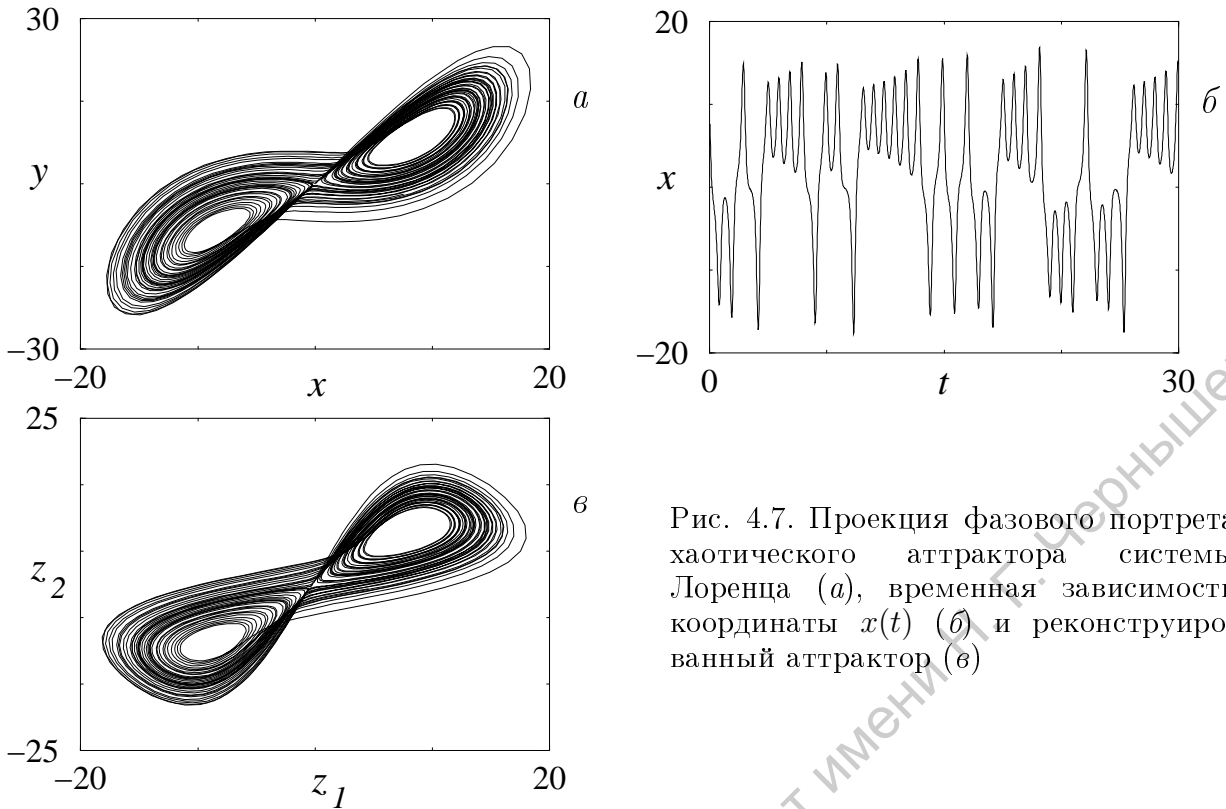


Рис. 4.7. Проекция фазового портрета хаотического аттрактора системы Лоренца (а), временная зависимость координаты $x(t)$ (б) и реконструированный аттрактор (в)

Из-за работы с конечным числом точек N необходимо тщательно выбирать параметр τ , поскольку качество реконструкции будет заметно отличаться при его вариации. Качество может оцениваться различным образом, например, с точки зрения близости численных значений тех или иных характеристик аттракторов A и A_R . Если τ слишком мало (рис. 4.8, а), то i -я и $(i+1)$ -я координаты точки в фазовом пространстве практически неразличимы. Реконструированный аттрактор в этом случае располагается вблизи главной диагонали пространства вложения, определяемой равенством координат $z_1 = z_2 = \dots = z_m$, то есть множество векторов \vec{z} концентрируется вблизи одной линии. Это приводит к заниженному значению размерности аттрактора A_R по сравнению с A . Слишком большое τ (рис. 4.8, б) также создает трудности: координаты оказываются некоррелированными, и реконструированный аттрактор не отражает истинной динамики.

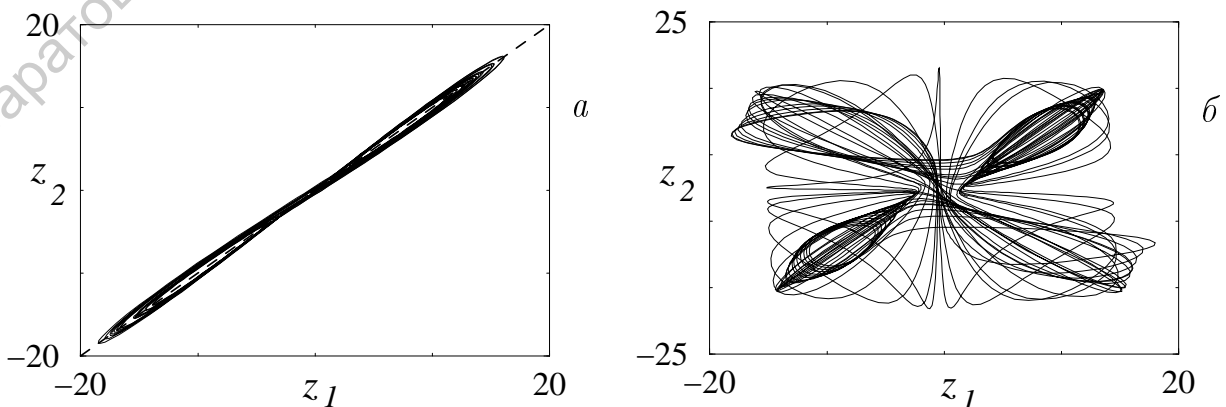


Рис. 4.8. Результат реконструкции при малой (а) и большой задержке (б)

Существует множество способов выбора оптимальной задержки. В качестве τ можно взять, например, время достижения первого нуля автокорреляционной функцией (АКФ). Можно пользоваться более сложными критериями, такими как минимум функции взаимной информации. Но для практических целей τ часто подбирают менее строгим образом, «на глаз», исходя из геометрии реконструируемого множества (чтобы аттрактор не был слишком вытянут ни в одном из направлений). Если анализируется сигнал, у которого наблюдается характерный период T , то лучше всего рассматривать значение $\tau = T/4$ (рис. 4.9).

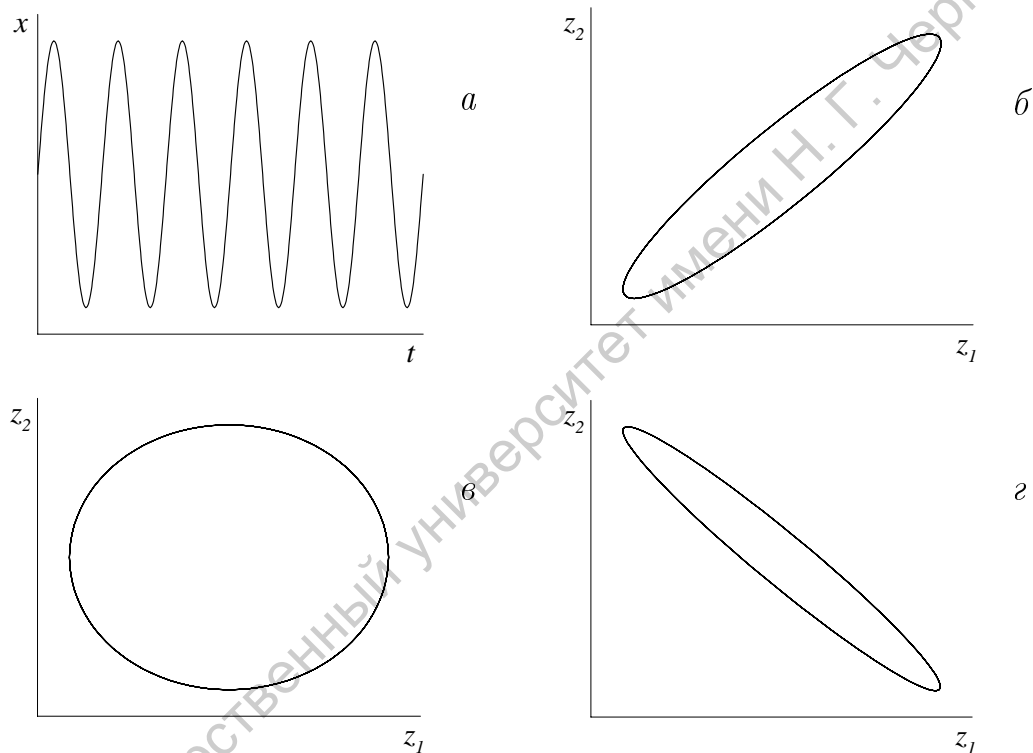


Рис. 4.9. Пример, иллюстрирующий, что при $\tau = T/4$ фазовый портрет имеет наименьшие геометрические искажения: исходный сигнал (а) и результаты реконструкции при $\tau = 0.1T$ (б), $\tau = 0.25T$ (в) и $\tau = 0.9T$ (г)

При выборе размерности пространства вложения обычно считается, что если хорошая реконструкция достигается при некотором m , то же справедливо и для $m + 1$. В частности, практически перестает меняться взаимное расположение точек в фазовом пространстве, то есть если расстояние между точками было мало, оно и останется малым; если велико – соответственно будет большим. Значение $m \geq 2d + 1$ определяет минимальную гарантированную размерность вложения. Иногда можно ограничиться меньшим значением $d < m < 2d + 1$, получив при этом хорошую реконструкцию (однако понижение параметра m не должно приводить к самопересечениям фазовой траектории). Существует мнение, что наиболее важным параметром при реконструкции оказывается длина окна, покрываемого \vec{z} -вектором

– величина $w = (m - 1)\tau$. Она не должна превышать время корреляции сигнала $c(t)$.

Метод задержки является наиболее популярным, но не единственным методом реконструкции аттрактора. Существуют и другие подходы, которые могут использоваться при решении данной задачи, например, метод последовательного дифференцирования координат:

$$\vec{z}(t) = \left\{ c(t), \frac{dc(t)}{dt}, \dots, \frac{d^{m-1}c(t)}{dt^{m-1}} \right\}.$$

Так как известны только значения наблюдаемой переменной $c(t)$ в дискретные моменты времени $i\Delta t$, координаты векторов \vec{z} определяются путем численного дифференцирования исходного временного ряда по приближенным математическим формулам. Очевидно, что точность вычисления производных будет определяться малостью величины шага дискретизации Δt . Недостатком метода является повышенная чувствительность к шуму при работе с экспериментальными данными, что ограничивает его применимость для пространств вложения большой размерности (по крайней мере, без проведения предварительной фильтрации). На настоящий момент разработан еще ряд способов задания вектора состояния, например, метод интегральной фильтрации:

$$z_j(t) = \int_0^t a(t_1) \exp\left(\frac{t - t_1}{\beta}\right) dt_1.$$

При его использовании обеспечивается сглаживание исходной реализации и фильтрация шума. Разновидностью данного метода является метод скользящего среднего

$$z_j(i, k) = \frac{1}{2k + 1} \sum_{l=-k}^k a(i + l),$$

где k – постоянный параметр. Данный метод также позволяет проводить сглаживание сигнала. Иногда для реконструкции аттрактора используется сразу несколько методов (для разных координат вектора состояния).

На рис. 4.10 изображены примеры реконструкции аттрактора в системе Ресслера

$$\begin{aligned} \frac{dx}{dt} &= -y - z, \\ \frac{dy}{dt} &= x + ay, \\ \frac{dz}{dt} &= b - cz + xz \end{aligned} \tag{4.7}$$

с помощью различных вариантов задания координат вектора состояния. Отметим, что при расчете размерности обычно ограничиваются методом

задержки, другие варианты восстановления траектории в фазовом пространстве применяются в более сложных задачах, например, при решении проблемы глобальной реконструкции, предусматривающей построение уравнений движения по временному ряду.

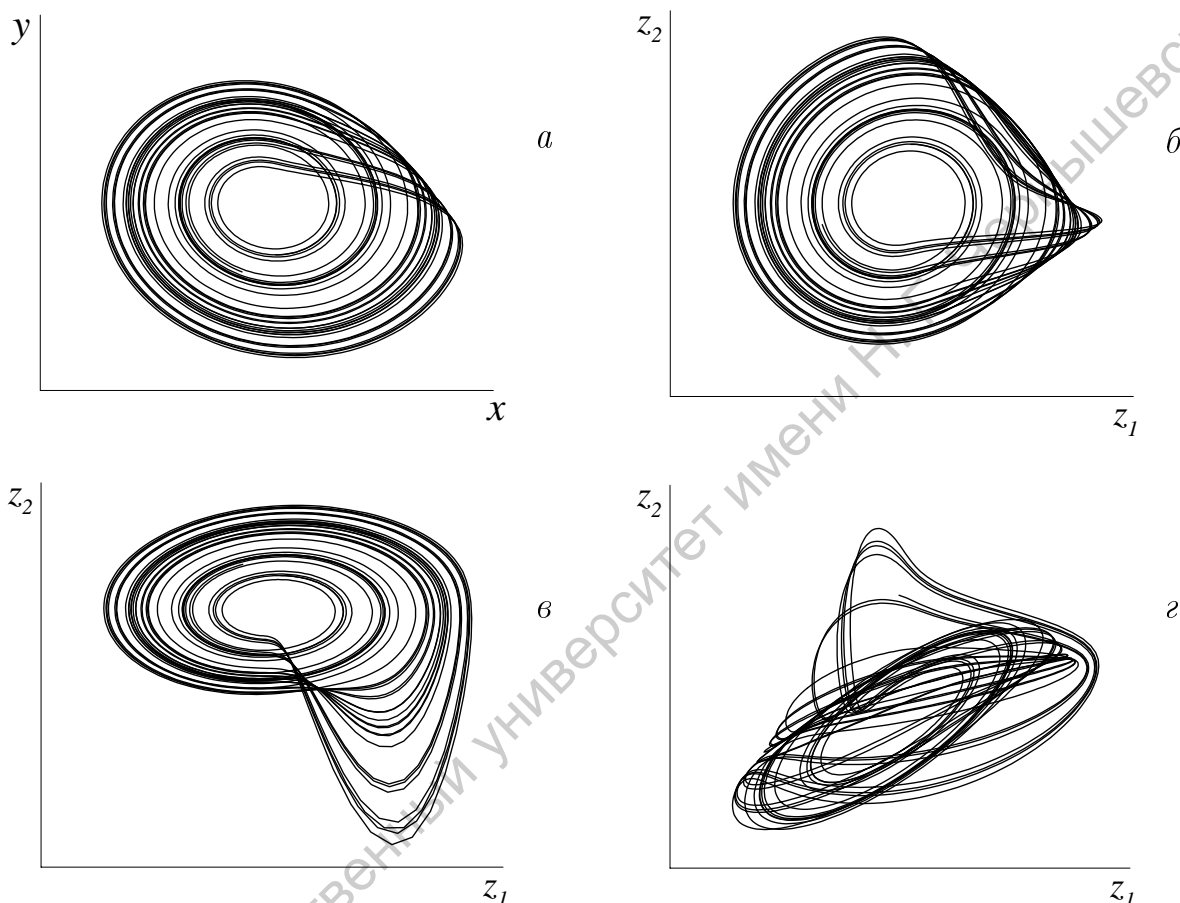
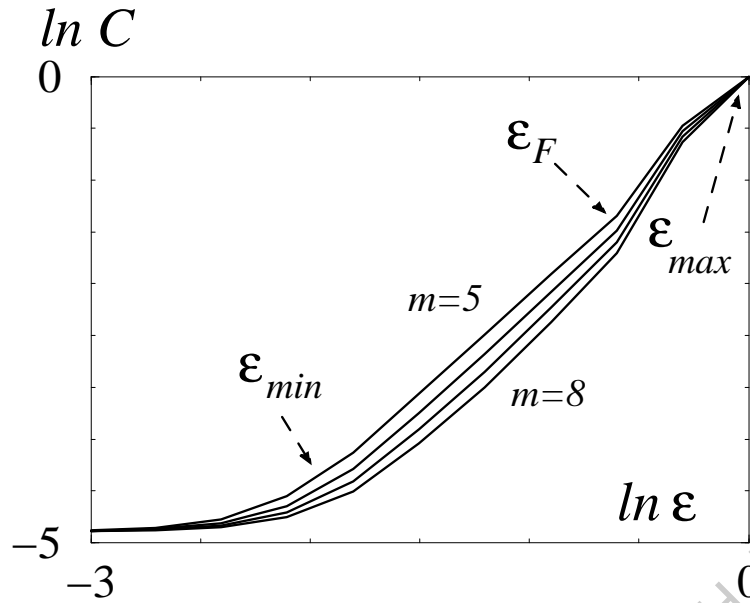


Рис. 4.10. Аттрактор в системе Ресслера (4.7) (а) и его реконструкция по переменной $x(t)$ методами задержки (б), дифференцирования (в) и скользящего среднего (з)

4.4 Особенности расчета корреляционной размерности

Как отмечалось выше, при вычислении ν по формуле (4.3) нужно учитывать, что в численных экспериментах N всегда конечно, и оба предела $\varepsilon \rightarrow 0$ и $N \rightarrow \infty$ не имеют реального смысла. Для определения ν строят зависимость $\ln C(\varepsilon, N)$ от $\ln \varepsilon$ в широком диапазоне по ε (рис. 4.11) и проводят поиск линейного участка, наклон которого определяет искомое значение размерности. Кроме того, при исследовании скалярных временных рядов анализируют зависимость ν от выбора m . Если размерность исходного аттрактора является конечной, то при увеличении m значение ν испытывает насыщение.

Рис. 4.11. Зависимости $\ln C(\varepsilon, N)$ от $\ln \varepsilon$ для $m = 5 \div 8$

При вычислении корреляционного интеграла существует ряд ограничений на величину ε . Если ε приближается к размерам аттрактора ε_{max} , то $C(\varepsilon, N) = 1$, а следовательно, $\rightarrow \ln C(\varepsilon, N) = 0$. На графике зависимости $\ln C(\varepsilon, N)$ от $\ln \varepsilon$ (рис. 4.11) обычно можно выделить еще два масштаба:

- верхнюю границу линейного участка ε_F (начало масштабной инвариантности $C \sim \varepsilon^\nu$). Часто отсутствие линейной зависимости при $\varepsilon > \varepsilon_F$ объясняют влиянием краев аттрактора, где среднее число ε -соседей обычно меньше, чем в «середине»;
- нижнюю границу линейного участка ε_{min} . Для $\varepsilon < \varepsilon_{min}$ структура аттрактора остается неразрешенной. Чем меньше ε , тем меньше число соседей внутри ячейки будет у точек в фазовом пространстве. Размерность не может быть измерена, если в среднем имеется меньше ν соседей, поэтому значение нижней границы ε_{min} зависит от времени наблюдения T_{obs} и числа точек N , используемых при расчете размерности.

Для удобства анализа вместо зависимости $\ln C(\varepsilon, N)$ от $\ln \varepsilon$ можно рассматривать локальные наклоны данного графика, соответствующие различным масштабам (рис. 4.12). Вариация размерности пространства вложения позволит наблюдать насыщение значения ν в области участка масштабной инвариантности (или фрактальности), который соответствует линейной зависимости $\ln C(\varepsilon, N)$ от $\ln \varepsilon$.

Точность расчета корреляционной размерности определяется длиной линейного участка, то есть величиной $p = \varepsilon_F / \varepsilon_{min}$. Считается, что при наличии свойств фрактальности p не может быть меньше $p_{min} \approx 2 \div 3$, а для надежного измерения ν необходимо $p \approx 10$.

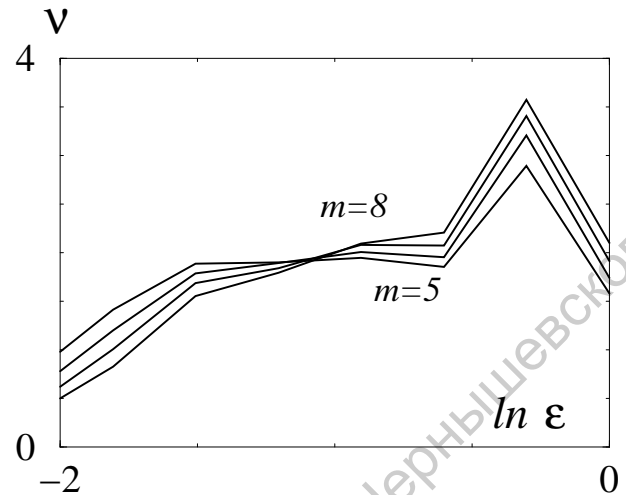


Рис. 4.12. Графики локальных наклонов

Существуют различные оценки минимального числа точек для правильного определения ν . Аргументы в пользу той или иной оценки слегка варьируются в разных исследованиях, но все они основаны на предположениях об «однородности» аттрактора. Аттракторы в динамических системах почти никогда однородными не бывают, но тем не менее эти оценки полезны и дают, по крайней мере, общее представление о возможности определения размерности по экспериментальным данным. При расчете ν можно выделить 3 важные характеристики массива данных: полное время наблюдения T_{obs} , число точек N и шаг между ними Δt . Они связаны соотношением $T_{obs} = N\Delta t$, и их следует рассматривать вместе, поскольку одной характеристики недостаточно: большое T_{obs} при малом N может быть столь же плохо, как и малое T_{obs} при большом N .

Возможность расчета корреляционной размерности по данным только об одной траектории основана на хорошо известном свойстве хаотических аттракторов: траектория, проходящая через любую точку, в течение некоторого времени блуждает по аттрактору, но затем (через интервал времени T_R) она возвращается в ε -окрестность этой точки. Чем меньше ε , тем больше должно быть T_R . При одном и том же ε величина T_R обычно меняется от одной точки к другой, но тем не менее, можно ввести некоторое среднее время возврата. Чтобы определить ν по интервалу масштабов $\varepsilon \in [\varepsilon_{min}, \varepsilon_F]$, точки аттрактора должны иметь достаточное число «соседей». Для различных точек аттрактора число ε -соседей будет отличаться, но предполагается, что ε_{min} соответствует ситуации, когда в среднем у точки аттрактора существует порядка одного соседа внутри шара радиуса ε_{min} , то есть когда $k(\varepsilon_{min}) \approx 1$. Соответствующее T_{obs} должно быть таким, чтобы для большинства точек траектория успела вернуться в их ε_{min} -окрестность 1–2 раза. Таким образом, если рассмотреть покрытие аттрактора шарами размера ε_{min} (число которых равно M), то большинство из них траектория посетит 1–2 раза.

Для данного T_{obs} существует некоторое оптимальное число точек N , необходимое для того, чтобы обеспечить разрешение на масштабе ε_{min} (это число должно быть близко к M). Так как скорость $|d\vec{x}/dt|$ вдоль траектории меняется, при однородном разбиении траектории по времени t_i плотность точек в фазовом пространстве однородной не будет, но этим эффектом можно пренебречь. Если $N \ll M$, то разрешаемый масштаб будет ограничен уже величиной N вместо T_{obs} , а в случае $N \gg M$ на малых масштабах ($\varepsilon < \varepsilon_{min}$) алгоритм будет измерять размерность не аттрактора, а самой траектории, если не принять специальных мер предосторожности. Следовательно, чтобы разрешить масштаб ε_{min} , необходимо $N \approx M$.

Обычно поступают следующим образом. Выбирают произвольное малое значение размерности пространства вложения m и вычисляют ν по наклону линейного участка графика $\ln C(\varepsilon, N)$ от $\ln \varepsilon$. Затем увеличивают m на единицу и вновь определяют ν . Таким образом, анализируется зависимость результатов расчета корреляционной размерности от выбора пространства вложения. Этот прием позволяет сделать вывод о существовании маломерной динамики ($\nu < 4$): при ее наличии зависимость $\nu(m)$ быстро достигает насыщения и при дальнейшем увеличении m не меняется в пределах точности вычислений. Если маломерная динамика отсутствует, то ν увеличивается с ростом m . В этом случае при достаточно больших m возможно насыщение, которое обусловлено фундаментальными ограничениями на значение корреляционной размерности, связанными с конечной длиной анализируемого временного ряда:

$$\nu_{max} = \frac{2 \ln N}{\ln(1/r)}, \quad r = \frac{\varepsilon}{\varepsilon_{max}}.$$

Данная формула означает, что алгоритм расчета размерности не может дать значение больше, чем ν_{max} при заданном числе точек N . Иными словами, если $r = 0.1$ и $N = 1000$, то $\nu_{max} \leq 6$; если $N = 100000$, то $\nu_{max} \leq 10$. Наличие фундаментальных ограничений создает серьезные проблемы, если проводится сравнение сложных, но детерминированных режимов динамики в системах с достаточно большим числом степеней свободы и случайных процессов. При изучении динамики маломерных систем таких проблем не возникает.

Практическая реализация алгоритма расчета размерности ν предполагает составление программы вычисления корреляционного интеграла в широком диапазоне по параметру ε и нахождение локальных наклонов зависимости $\ln C(\varepsilon, N)$ от $\ln \varepsilon$. Если уравнения ДС (4.1) известны, то есть можно проинтегрировать математическую модель, в формуле (4.4) рассматривается массив векторов \vec{x}_i . В случае одномерного временного ряда предварительно проводится реконструкция (выбирается значение задержки и размерности пространства вложения), после чего в формуле (4.4) рассматривается массив векторов \vec{z}_i . Если расчет размерности сопровождается

ется процедурой реконструкции аттрактора, появляются дополнительные параметры, которые также влияют на точность расчетов. Для того чтобы быть уверенным в результатах вычислений, экспериментаторы предпочитают немного поварьировать различные параметры и убедиться, что полученные оценки ν при этом будут меняться незначительно.

Меньше проблем возникает при исследовании дискретных отображений вида

$$\vec{x}_{n+1} = \vec{F}(\vec{x}_n, \vec{\mu}), \quad \vec{x} \in R^n, \vec{\mu} \in R^k,$$

поскольку в этом случае уменьшается число параметров, от которых зависит точность определения корреляционной размерности (в частности, отсутствует Δt , а величина временной задержки обычно принимается равной единице), в то время как сама программа расчета размерности остается прежней.

Порядок выполнения работы

Задание 1. Вычислить аналитически емкостную размерность одного из следующих фрактальных объектов.

Вариант (а): «снежинки» (рис. 4.3);

Вариант (б): канторова множества (рис. 4.2), если предположить, что в процессе его построения проводится разбиение отрезка $[0, 1]$ не на 3 части, а на 5. При этом на каждом шаге по-прежнему выбрасывается только одна – средняя часть;

Вариант (в): «салфетки» Серпинского (рис. 4.4);

Вариант (г): «ковра» Серпинского (рис. 4.5);

Вариант (д): квадратной кривой Кох (рис. 4.6).

Задание 2. Составить программу реконструкции фазового портрета по предоставленному временному ряду $x(i)$. Подбором задержки оценить оптимальное значение временной задержки (исходя из геометрии реконструированного множества – аттрактор не должен быть слишком вытянут ни в одном из направлений).

№ варианта	динамическая система
1	система Ресслера
2	система Лоренца
3	генератор с инерционной нелинейностью
4	модель нефрона
5	модель нейрона (Хиндмарш-Розе)
6	модель бета-клетки

Задание 3.

Вариант (а). Написать программу итерирования отображения Хенона, которая будет записывать в файл с названием «data» в две колонки соответственно значения x_i, y_i . Число точек для расчета ν выбрать равным $N = 10000$. Вычислить корреляционную размерность хаотического аттрактора в отображении Хенона с помощью программы «cordim.x», которая считывает данные из файла «data» и выводит результаты расчета в файл «dim.out» (в первую колонку данного файла записывается $\ln \varepsilon$, во вторую $\ln C(\varepsilon, N)$). Оценить величину ν по наклону линейного участка полученной зависимости.

Вариант (б). Определить корреляционную размерность хаотического аттрактора в двумерной дискретной динамической системе Лози. С этой целью необходимо выполнить те же операции, что и в варианте (а), только теперь в файл «data» будут записаны значения x_i, y_i другой модельной системы.

Вариант (в). Определить корреляционную размерность хаотического аттрактора в отображении Хенона по одной координате. Для этого нужно записать в файл «data» в две колонки значения $x_i, x_{i+1}, i = 1, \dots, 10000$. Далее воспользоваться программой расчета размерности «cordim.x» по аналогии с заданием 3(а).

Вариант (г). Определить корреляционную размерность хаотического аттрактора в отображении Лози по одной координате. Для этого нужно записать в файл «data» в две колонки значения $x_i, x_{i+1}, i = 1, \dots, 10000$. Далее воспользоваться программой расчета размерности «cordim.x» по аналогии с заданием 3(а).

Перечень модельных систем для выполнения задания 3**Отображение Хенона**

$$\begin{aligned}x_{n+1} &= 1 - 1.4x_n^2 + y_n \\y_{n+1} &= 0.3x_n\end{aligned}$$

Отображение Лози

$$\begin{aligned}x_{n+1} &= 1 - 1.5|x_n| + y_n \\y_{n+1} &= 0.3x_n\end{aligned}$$

Задание 4. С помощью программы «cordim2.x» провести расчеты корреляционной размерности для одного из физиологических сигналов при варьировании размерности вложения m в пределах от 3 до 10. Сделать вывод о наличии (или отсутствии) маломерной динамики. Если

маломерную динамику удастся обнаружить, оценить минимальное значение размерности пространства вложения, при котором целесообразно проводить расчеты.

№ варианта	анализируемый сигнал
1	последовательность RR-интервалов электрокардиограммы человека
2	последовательность межспайковых интервалов нейрона рыбы
3	последовательность межспайковых интервалов нейрона крысы
4	сигнал одиночного нефрона крысы
5	последовательность берстов, генерируемых нейроном крысы
6	сигнал электроэнцефалограммы человека

Контрольные вопросы

1. Что такое размерность Хаусдорфа?
2. Чем отличаются емкостная и корреляционная размерности?
3. Почему именно корреляционную размерность предпочитают вычислять на практике?
4. В чем состоит реконструкция динамических систем по экспериментальным данным?
5. Назовите основные методы реконструкции аттрактора по одномерной реализации.
6. Как подбираются параметры реконструкции?
7. Объясните, в чем состоят ограничения на величину ε при расчете корреляционного интеграла.
8. В чем состоят ограничения величины размерности при расчетах по сигналам малой длительности?

5 Ляпуновские показатели

5.1 Интерпретация ляпуновских показателей

Рассмотрим наиболее часто используемый на практике метод расчета ляпуновских характеристических показателей (ЛХП) по временному ряду. Этот метод немного напоминает стандартный подход к вычислению спектра ЛХП потоковых систем по известной математической модели. В целом техника расчета ляпуновских показателей базируется на следующей идеологии. Предположим, что задана n -мерная динамическая система с непрерывным временем. Чтобы охарактеризовать устойчивость ее решения, анализируется временная эволюция бесконечно малой n -мерной сферы начальных условий; с течением времени эта сфера преобразуется в эллипсоид (рис. 5.1).

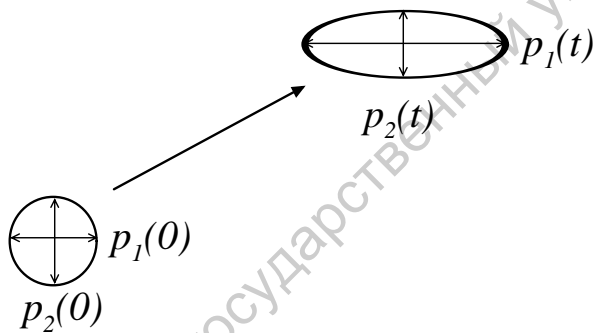


Рис. 5.1. Эволюция по времени бесконечно малой n -мерной сферы начальных условий

Если говорить о спектре ЛХП, то i -й показатель Ляпунова может быть определен в терминах длин осей эллипсоида $p_i(t)$:

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{p_i(t)}{p_i(0)},$$

где λ_i упорядочены от наибольшего к наименьшему. Таким образом, показатели Ляпунова определяются расширением либо сжатием сферы по различным направлениям в фазовом пространстве. Поскольку ориентация эллипсоида непрерывно меняется со временем, меняются и направления, ассоциирующиеся с тем или иным показателем. Поэтому нельзя говорить о каком-то одном направлении для каждой ляпуновской экспоненты.

Отметим, что если задается только одно возмущение (это значит, что мы следим лишь за главной осью эллипсоида), то оно будет в линейном приближении увеличиваться по закону $e^{\lambda_1 t}$. Для двух независимых возмущений площадь образуемого ими квадрата (или прямоугольника) эволюционирует по закону $e^{(\lambda_1 + \lambda_2)t}$; для трех – эволюция объема описывается законом $e^{(\lambda_1 + \lambda_2 + \lambda_3)t}$ и т.д. Данное свойство приводит к несколько иному определению спектра ляпуновских экспонент: сумма первых j показателей определяется скоростью экспоненциального роста j -мерного элемента объема. Такая интерпретация обеспечивает основу для техники анализа экспериментальных данных (когда требуется вычислить несколько показателей).

Любая потоковая ДС, если ее решение отлично от состояния равновесия, имеет по крайней мере один нулевой показатель, характеризующий практически не меняющуюся величину оси эллипсоида, касательной к потоку. Оси, размер которых будет увеличиваться (уменьшаться), соответствуют положительным (отрицательным) показателям.

Иногда ляпуновские показатели анализируют в терминах информации, в частности как скорость потери информации о начальном состоянии. В этих случаях их определяют в таких единицах, как бит/сек (количество бит информации в секунду), бит/период (для потоковых систем) или бит/итерация (для систем с дискретным временем). Соответственно вычисляют скорость разбегания соседних траекторий не по закону $e^{\lambda_1 t}$, а по закону $2^{\lambda_1 t}$. Например, если старший показатель Ляпунова хаотического аттрактора Лоренца равен 2.16 бит в единицу времени, а начальные условия определены с точностью до одной миллионной части (≈ 20 бит), то дальнейшее поведение не может быть предсказано спустя 9 единиц безразмерного времени ($20/2.16$), что соответствует примерно 20 оборотам вокруг одного из состояний равновесия. Спустя это время малая начальная неопределенность будет покрывать весь аттрактор, и фазовая точка может оказаться где угодно.

Поскольку расчет ляпуновских экспонент по экспериментальным данным имеет одну общую идеологию с расчетом показателей по системе обыкновенных дифференциальных уравнений (ОДУ), кратко рассмотрим эту процедуру.

5.2 Расчеты по системе уравнений

Как известно, для режима динамического хаоса характерно наличие экспоненциальной неустойчивости траекторий, количественной мерой которой является положительный ляпуновский показатель, характеризующий степень чувствительности системы к выбору начальных условий. Число положительных экспонент в спектре ЛХП определяется количеством неустойчивых направлений периодических орбит, встроенных в хаотиче-

ский аттрактор, хотя в принципе возможны и более сложные ситуации, состоящие в сосуществовании периодических орбит с различным числом неустойчивых направлений. Мы ограничимся рассмотрением динамических систем с одним положительным показателем Ляпунова λ_1 . При вычислении λ_1 делается предположение о «типичности» фазовой траектории, являющейся решением ДС при выбранных начальных условиях. В противном случае величина показателя, определенная на интервале времени T , может отличаться от предельного значения λ_1 , соответствующего $T \rightarrow \infty$. В частности, известны случаи, когда дискретные модели демонстрируют очень длительные переходные процессы (до 1500000 итераций динамика является «хаотической», после чего сменяется регулярной). С точки зрения вычисления ляпуновского показателя это соответствует тому, что величина λ_1 в течение переходного процесса сходится к некоторому положительному значению и только на очень длительных временах спадает до нуля. Чтобы охарактеризовать поведение типичной фазовой траектории, иногда используют понятие «ляпуновских показателей на конечном времени» (*finite time Lyapunov exponents*), которые характеризуют скорость разбегания или сжатия по различным направлениям в течение конечного временного интервала T .

Если уравнения ДС, генерирующие фазовую траекторию, известны, то определить величину максимального показателя (или полный спектр ЛХП) можно с помощью стандартного алгоритма расчета ляпуновских экспонент. Рассмотрим систему ОДУ:

$$\frac{d\vec{x}}{dt} = \vec{f}(\vec{x}, \vec{\mu}), \quad \vec{x} \in R^n, \vec{\mu} \in R^m,$$

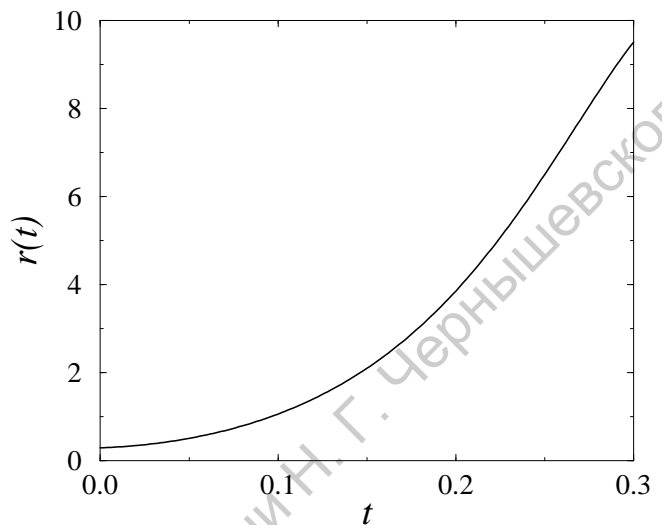
в которой \vec{x} – вектор состояния, \vec{f} – нелинейная вектор-функция, $\vec{\mu}$ – вектор управляющих параметров. Исследование данной системы на устойчивость ее частного решения $\vec{x}_p(t)$ сводится к анализу уравнений в вариациях и введению понятия k -мерного ляпуновского показателя. Но если ограничиться только вычислением старшего ЛХП, то этот алгоритм существенно упрощается. Искомая величина λ_1 будет определять эволюцию во времени вектора возмущения

$$r(t) = r_0 e^{\lambda_1 t},$$

где r_0 – величина начального возмущения (в момент времени $t_0 = 0$), $r_0 = |\vec{x}(t_0) - \vec{x}_p(t_0)|$. Данная формула является очень приближенной, поскольку скорость разбегания траекторий не является постоянной, а зависит от выбора точки на аттракторе. Строго говоря, ляпуновские показатели определяют путем решения уравнений в вариациях, рассматривая бесконечно малые возмущения. Однако экспоненциальный закон разбегания траекторий справедлив не только для бесконечно малых возмущений, но также и для малых возмущений конечной величины, поэтому последняя

формула вполне может быть использована при практическом анализе локальной неустойчивости (рис. 5.2).

Рис. 5.2. Пример экспоненциального разбегания близких в начальный момент траекторий для аттрактора в системе Лоренца



Поскольку большие значения $r(t)$ принимать не может (в противном случае не будет выполняться условие линейного приближения), при расчете старшего ЛХП проводят перенормировки, в ходе которых задаются новые возмущения — малые по величине, но выбранные в направлении, которое соответствует вектору возмущения непосредственно перед перенормировкой. Затем вновь оценивают скорость экспоненциального разбегания близлежащих траекторий. В результате вычисляется усредненная вдоль фазовой траектории $\bar{x}_p(t)$ количественная характеристика степени хаотичности. Схематично процедура перенормировки вектора изображена на рис. 5.3.

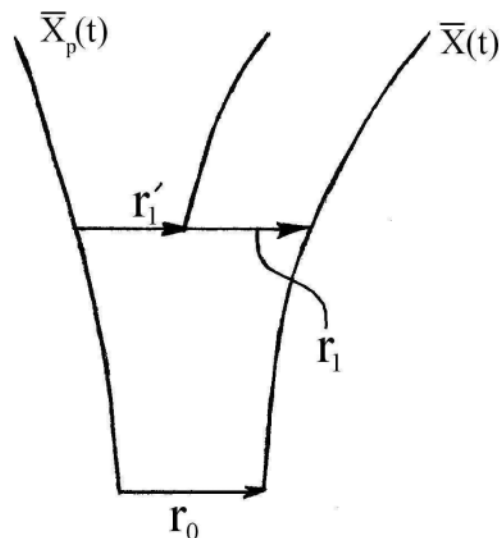


Рис. 5.3. Процедура перенормировок при расчете ЛХП по заданной системе уравнений. В данном случае r_1 — вектор возмущения до перенормировки, r'_1 — после

5.3 Расчеты по одномерной реализации

Метод расчета λ_1 по реализации использует ту же идеологию, но теперь возникает одна проблема: нет возможности произвольным образом задавать вектор возмущения. Поскольку анализируется экспоненциальная неустойчивость траекторий в фазовом пространстве, на первом этапе процедуры расчета показателя нужно осуществить реконструкцию траектории в фазовом пространстве, то есть сформировать множество векторов $\vec{z}(t)$ в соответствии с формулой (4.5).

Мы не можем выбирать вектор возмущения при перенормировках строго в заданном направлении и должны брать новую точку там, где ее удастся найти (в некотором конусе) (рис. 5.4). Невозможность выбора вектора возмущения в нужном направлении приводит к появлению ошибки ориентации (угол α между векторами r_1 и r'_1). Поскольку работа с одной единственной траекторией ограничивает возможности выбора вектора возмущения, на практике приходится искать компромисс между уменьшением ошибки ориентации в фазовом пространстве и минимизацией его длины. Ограничения на длину могут быть сформулированы следующим образом:

$$l_1 < r(t) < l_2.$$

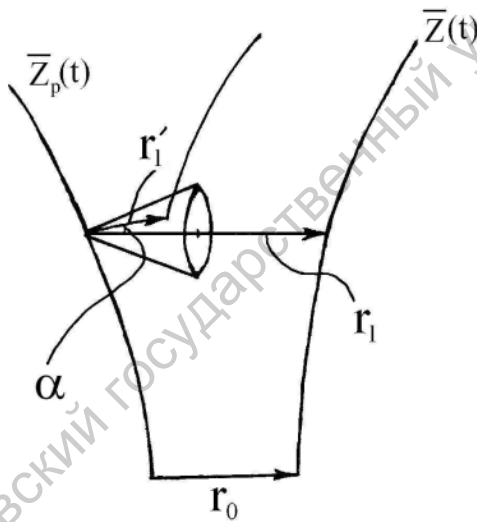


Рис. 5.4. Процедура перенормировок при расчете ЛХП по одномерной реализации.

Необходимо вводить минимальное значение l_1 , поскольку при $r(t) < l_1$ сказывается влияние шума, нарушающего экспоненциальный характер разбегания траекторий. Величина l_2 задает условие линейного приближения и часто может быть введена в процентном отношении от размера аттрактора (скажем, 5–10%). По аналогии со стандартным алгоритмом проводятся перенормировки вектора возмущения, если расстояние между траекториями перестает удовлетворять условию линейного приближения ($r(t) > l_2$). Поскольку вычисление λ_1 предполагает реконструкцию аттрактора, результат расчета данной величины будет зависеть от качества реконструкции,

что приводит к появлению дополнительных параметров численной схемы – размерности пространства вложения, задержки и т.п.

Несколько слов скажем о перенормировках вектора. Здесь возможны варианты:

- перенормировки проводятся после достижения определенного расстояния между траекториями, а именно, если $r(t) > l_2$;
- перенормировки проводятся через фиксированные интервалы времени (например, через характерный период колебаний).

Возможны также комбинации этих двух способов осуществления перенормировок. В любом случае выбором параметров алгоритма приходится искать компромисс: слишком частые перенормировки приводят к увеличению ошибок ориентации; слишком редкие – позволяют увеличиваться вектору возмущения, выводя за рамки линейного приближения. Метод перенормировок при достижении фиксированного расстояния между траекториями эффективен для достаточно однородных аттракторов, когда зависимость от выбора точки на нем не слишком сильная. Если анализируются неоднородные аттракторы, то предпочтительнее оказывается метод перенормировок через фиксированные промежутки времени.

Порядок выполнения работы

Задание 1. С помощью программы «lhp.x» провести расчеты старшего ляпуновского показателя периодического режима колебаний. Построить зависимости величины λ_1 от:

- времени задержки (при фиксированных остальных параметрах);
- размерности пространства вложения;
- минимального расстояния между точками в фазовом пространстве (l_1);
- максимального расстояния между точками в фазовом пространстве (l_2);
- времени между перенормировками.

Задание 2. Провести те же самые расчеты (как и в задании 1) для случая хаотического режима колебаний. Оценить погрешность вычислений.

№ варианта	динамическая система
1	система Ресслера
2	система Лоренца
3	генератор с инерционной нелинейностью
4	модель нефрона
5	модель нейрона (Хиндмарш-Розе)
6	модель бета-клетки

Задание 3. Вычислить старший ляпуновский показатель по экспериментальным данным (сигнал давления в канальце нефрона почки крысы). Сопоставить результаты при разных значениях параметров счета.

Контрольные вопросы

1. Что характеризуют ляпуновские показатели?
2. Что означает термин «типичная фазовая траектория»?
3. Как определить старший показатель Ляпунова по двум близлежащим траекториям в фазовом пространстве?
4. Как осуществляются перенормировки вектора начального возмущения при расчете λ_1 по известной системе уравнений?
5. В чем отличие перенормировок при расчете λ_1 по одномерной реализации?
6. В чем заключается процедура реконструкции фазового портрета?
7. Какие параметры влияют на точность вычисления старшего ляпуновского показателя?

6 Энтропия источника

6.1 Предварительные замечания

Понятие *мер сложности* относится к числу базовых представлений при проведении анализа динамики систем различной природы. В окружающем мире мы постоянно встречаемся с примерами сложного поведения самых разных объектов, и для того, чтобы описать наблюдаемые явления с точки зрения их упорядоченности (или предсказуемости) необходимо определить количественные характеристики, которые бы позволили это сделать. В настоящее время существует много самых разных подходов к определению сложности процессов. В рамках теории динамических систем они основываются на таких характеристиках как энтропия Колмогорова-Синая, показатели Ляпунова, обобщенные фрактальные размерности, взаимная информация и т.д.

При анализе последовательности символов часто применяется концепция *алгоритмической сложности*, которая характеризует минимальное число правил, необходимых для того, чтобы сгенерировать эту последовательность. Простые (периодические) последовательности требуют небольшого числа правил (команд для компьютерной программы, позволяющей их воспроизвести), тогда как для генерации сложных последовательностей необходимо значительно большее количество правил. Например, последовательность цифр иррациональных чисел (таких как π или e) отличается от случайного набора чисел, поскольку можно написать очень короткую программу, которая будет воспроизводить число π , но сделать это для совершенно случайной последовательности цифр не удастся. Аналогом алгоритмической сложности для различных текстов является *грамматическая сложность*, которая была предложена для описания возможности сжатия последовательности символов за счет введения специальных переобозначений.

К числу известных характеристик упорядоченности последовательностей символов относится *энтропия источника*, которую сравнительно легко вычислить на практике. В рамках теории динамических систем энтропия источника соответствует известной энтропии Колмогорова-Синая, которая характеризует предсказуемость поведения траектории в фазовом пространстве.

6.2 Определение энтропии источника

Исторически понятие энтропии было введено Больцманом в термодинамике. В 40-х годах 20 века концепция Больцмана была обобщена Шенноном для того, чтобы можно было охарактеризовать последовательности символов с точки зрения их упорядоченности.

Рассмотрим некоторую последовательность, составленную из алфавита, содержащего λ символов. Количество информации, доставляемое одним символом при условии, что все символы являются независимыми и равновероятными, или энтропия (в данном случае энтропия интерпретируется как информация о получении символа S с вероятностью $1/\lambda$) определяется следующим образом:

$$H = \log \lambda.$$

В общем случае, говоря об энтропии, нужно отметить следующие обстоятельства:

- Предположим, что мы получаем по буквам некоторый текст. Пока мы его еще не получили целиком, существует неопределенность о последующих символах. Получение каждого нового символа эту неопределенность уменьшает. При этом неопределенность, которая устраняется при получении символа S , будет тем меньше, чем больше вероятность символа. Редко встречающиеся (мало ожидаемые) символы несут больше информации, чем частые (в большей степени ожидаемые) символы. На языке математики для частного случая равновероятных символов это означает, что

$$H = f\left(\frac{1}{P}\right),$$

где f – функция, которая будет монотонно расти с ростом $1/P$.

- Количество информации, которое несут N независимых символов S_1, S_2, \dots, S_N , определяется следующим образом:

$$\begin{aligned} H(S_1, S_2, \dots, S_N) &= H(S_1) + H(S_2) + \dots + H(S_N), \\ f\left(\frac{1}{P^N}\right) &= Nf\left(\frac{1}{P}\right), \end{aligned}$$

а таким свойством обладает логарифмическая функция.

Поэтому H определяется следующим образом:

$$H = \log\left(\frac{1}{P}\right) = -\log P.$$

Если вероятность появления символов отличается, необходимо провести усреднение по всем λ символам, чтобы получить величину, которая характеризует количественную меру неопределенности в сообщении до его приема (энтропию Шеннона)

$$H = \langle -\log p_i \rangle = - \sum_{i=1}^{\lambda} p_i \log p_i,$$

при этом логарифм обычно рассматривается по основанию λ .

Следует отметить, что если речь идет о символьных последовательностях, то это не обязательно тексты, цепочки нуклеотидов ДНК и т.д. Порой и обычные сигналы преобразуют в последовательности символов, например, в бинарные последовательности: если значение случайного процесса в данный момент больше среднего, то это соответствует символу «1», если меньше среднего, то «0» и т.п. Изучение последовательностей символов вместо временного ряда осуществляется в рамках так называемой *символьной динамики* – отдельного направления в анализе сложных систем.

Что может сообщить нам энтропия о последовательности? Предположим, что анализируется бинарная последовательность символов «0» и «1», и вероятности данных символов соответственно равны p_0 и $p_1 = 1 - p_0$. Если логарифм рассматривается по основанию 2, то энтропия будет принимать значения в интервале от 0 до 1: $H = 1$, если вероятности одинаковы $p_0 = p_1 = 0.5$, и $H = 0$, если постоянно генерируется только один символ ($p_1 = 1$ или $p_0 = 1$). Значения $0 < H < 1$ характеризуют степень порядка в последовательности. Энтропия близка к нулю, когда один из символов доминирует (например, $H \approx 0.08$ для $p_0 = 0.01$ и $p_1 = 0.99$). Если же вероятности p_0 и p_1 почти одинаковые, то $H \approx 1$ (например, $H \approx 0.992$ для $p_0 = 0.45$ и $p_1 = 0.55$).

Однако такое определение энтропии на самом деле дает немного полезной информации. Например, принципиально различные процессы будут иметь одно и то же значение энтропии:

$$\begin{aligned} S_0 &= \dots 10101010101010\dots \\ S_1 &= \dots 1101110100101000\dots \end{aligned}$$

Если в обоих случаях $p_0 = p_1 = 0.5$, то как для периодической последовательности S_0 , так и для случайной последовательности S_1 энтропия будет принимать значение $H = 1$. Данный недостаток можно устранить, если ввести блочные энтропии (или энтропии высокого порядка). Очевидно, что уже вероятности встретить разные блоки из двух символов («00», «01», «10», «11») для S_0 и S_1 совпадать не будут. Поэтому более информативной характеристикой последовательности символов служит *n-блочная энтропия*

$$H_n = - \sum_i p_i^{(n)} \log p_i^{(n)}, \quad (6.1)$$

которая может интерпретироваться как среднее количество информации, содержащееся в блоке из n символов. В формуле (6.1) $p_i^{(n)}$ – вероятности блоков из n символов (S_1, \dots, S_n) , а суммирование проводится по всем встречающимся блокам.

Рассмотрим, как величина H_n меняется с ростом размера блока. В общем случае энтропии отдельных символов не позволяют охарактеризовать анализируемую последовательность, тогда как блочные энтропии H_1, H_2, H_3, \dots , взятые в совокупности, обеспечивают более информативное количественное описание структуры последовательностей. При этом могут наблюдаться различные особенности зависимости $H_n(n)$:

- Если рассматривается последовательность, в которой любые блоки из n символов имеют равные вероятности $p^{(n)} = 1/\lambda^{(n)}$, то

$$H_n = n \log \lambda. \quad (6.2)$$

Линейная зависимость (6.2) характерна для последовательности Бернулли.

- Для марковских процессов порядка k наблюдается линейный рост блочных энтропий при $n \geq k$:

$$H_n = H_k + (n - k)(H_{k+1} - H_k), \quad n \geq k.$$

- Для периодических последовательностей (с периодом q) достаточно знать только q символов. Остальные символы не несут новой информации, поэтому

$$H_n = H_q, \quad n \geq q.$$

- Символьные последовательности, несущие какой-то смысл, должны находиться между случаем полной неупорядоченности и полной периодичности. Это должно отражаться в уменьшении приращений $H_{n+1} - H_n$ при возрастании n , что является следствием эффектов «длительной памяти». Достаточно универсальная формула зависимости $H_n(n)$ имеет вид:

$$H_n = nh + gn^{\mu_0}(\log n)^{\mu_1} + c, \quad (6.3)$$

где c – константа, $0 < \mu_0 \leq 1$, $\mu_1 < 0$ – параметры, зависящие от конкретной последовательности, h – энтропия источника, определение которой будет рассмотрено далее. Логарифмическая зависимость (6.3) наблюдается, например, при анализе последовательностей, генерируемых логистическим отображением в критической точке.

Чтобы охарактеризовать приращения блочных энтропий, рассматривают так называемые *разностные блочные энтропии* (иногда называемые *динамическими энтропиями*)

$$h_n = H_{n+1} - H_n.$$

Их можно интерпретировать как среднюю неопределенность о символе, если известны n предшествующих символов. Поскольку чем больше символов мы знаем, тем меньше неопределенность о сообщении (в среднем), то значения h_n монотонно уменьшаются. Предел

$$h = \lim_{n \rightarrow \infty} \frac{H_n}{n} = \lim_{n \rightarrow \infty} h_n \quad (6.4)$$

называют *энтропией источника*, которая принимает максимальное значение $h = 1$ для полностью случайного процесса (белого шума) и $h = 0$ для периодических последовательностей. Для последовательностей, несущих полезную информацию, $h \in]0, 1[$.

При вычислении блочных энтропий и энтропии источника нужно обращать внимание на возможности достоверной оценки вероятности блока из n символов. Ее определяют как относительную частоту появления данного блока

$$p_i = \frac{k_i}{N},$$

где N – число блоков, k_i – число появлений блока с номером i . При работе с конечными последовательностями символов оценки вероятностей больших блоков будут очень неточными, так как возникает очень много вариантов блоков. В этом случае оценки энтропий являются недостоверными при $n > 6$ (для последовательности, содержащей 10000 символов) или $n > 9$ (для 100000 символов). Поэтому вместо рассмотрения характеристики h в зависимости от длины последовательности целесообразно выбирать одну из разностных блочных энтропий (например, h_5). Предел $n \rightarrow \infty$ в формуле (6.4) на практике оказывается бессмысленным.

Энтропию источника можно интерпретировать как характеристику сложности символьных последовательностей. В качестве примера рассмотрим определение сложности Лемпеля-Зива. Идея данного алгоритма состоит в следующем. Последовательность символов разбивается метками на фрагменты, которые ранее еще не были идентифицированы. Например, последовательность 100110... разбивается на фрагменты следующим образом: 1 · 0 · 01 · 10 ...

Число полученных фрагментов C представляет собой меру сложности последовательности. Естественно, что это число будет возрастать с ростом общего количества символов N . Для случайных бинарных последовательностей ($p_0 = p_1 = 1/2$) при больших N

$$C(N) = \frac{N}{\log_2 N}.$$

Для не совсем случайных последовательностей

$$C(N) = h \frac{N}{\log_2 N},$$

где коэффициентом пропорциональности служит энтропия источника. Чтобы характеристика сложности не зависела от N , часто рассматривают величину $C(N) \cdot (\log_2 N)/N$, что как раз и соответствует значению h . Возможность разбить последовательности на участки по определенным правилам играет важную роль в задачах сжатия данных. Исследование мер сложности в качестве одного из своих приложений предполагает создание алгоритмов оптимального сжатия. Например, идеология, рассмотренная в рамках метода Лемпеля-Зива, используется в архиваторе «gzip», а энтропию источника можно оценить, посчитав отношение размера бинарного файла после архивирования к первоначальному размеру.

Порядок выполнения работы

Задание 1. Написать программу, позволяющую с помощью логистического отображения

$$x_{n+1} = rx_n(1 - x_n)$$

сгенерировать бинарную последовательность (случай $r = 4$). Для этого нужно, итерируя данное отображение, записывать в файл символ «1», если $x \geq 0.5$, и «0» для $x < 0.5$. Число итераций взять равным 10000.

С помощью программы «block.x» провести расчеты блочных энтропий H_n для $1 \leq n \leq 10$. Программа «block.x» считывает символьную последовательность из файла данных и записывает в выходной файл «entropy.dat» значения n и H_n . Построить график $H_n(n)$ и определить, начиная с какого размера блока H_n заметно отличается от теоретического значения $H_n = n$. Построить зависимость погрешности вычислений от размера блока.

Задание 2. Провести сопоставление двух хаотических режимов динамики предложенной преподавателем модели с точки зрения их сложности. Воспользовавшись программой «block.x», по аналогии с предыдущим заданием построить зависимости $H_n(n)$. Оценить величину энтропии источника для каждого режима.

№ варианта	динамическая система
1	система Ресслера
2	система Лоренца
3	генератор с инерционной нелинейностью
4	модель нефрона
5	модель нейрона (Хиндмарш-Розе)
6	модель бета-клетки

Задание 3. Построить зависимость $H_n(n)$ по экспериментальным данным (последовательность кардиоинтервалов электрокардиограммы человека). Оценить динамическую энтропию h_5 .

Контрольные вопросы

1. Назовите методы исследования сложности, которые используются в рамках теории динамических систем.
2. Что такое алгоритмическая сложность?
3. Почему энтропия отдельных символов не может применяться в качестве информативной характеристики символьной последовательности?
4. Как блочная энтропия зависит от размера блока для разных последовательностей?
5. Как вычисляется энтропия источника? В чем состоят ограничения, обусловленные конечной длиной последовательности?
6. В чем состоит метод оценки сложности Лемпеля-Зива?

7 Флуктуационный анализ

7.1 Предварительные замечания

В функционировании многих систем в природе важную роль играют одновременно как детерминированные, так и случайные факторы, поэтому наблюдаемая сложная динамика может быть обусловлена сочетанием самых разных механизмов. Характерные временные масштабы этих систем часто варьируются в широком диапазоне значений, что может приводить к такому явлению как неэкспоненциальный закон спада автокорреляционной (или автоковариационной) функции. Для случайных процессов достаточно типичны степенные зависимости в поведении АКФ или функции спектральной плотности. Точная характеристика закономерностей потери корреляций важна при проведении анализа различных систем, так как она позволяет делать выводы о наличии и особенностях «длительной памяти» в их динамике. Например, такой анализ является математической основой для решения задач диагностики функциональных нарушений в структуре сердечного ритма.

Непосредственные расчеты АКФ по временному ряду ограничены достаточно малыми значениями τ по сравнению с самим исследуемым сигналом. Более того, для случайных процессов (в отсутствие характерного временного масштаба) автокорреляционная функция быстро спадает до нуля. Поэтому, начиная с некоторого τ , значения АКФ становятся настолько малы, что оказываются сопоставимыми с ошибками вычислений, возникающими при анализе сигналов конечной длительности. Как следствие, на больших временах возможность отслеживать закономерности спада корреляций просто отсутствует (в рамках классического корреляционного анализа).

За последние двадцать лет было предложено несколько методов, позволяющих решить такие проблемы и изучать особенности длительных корреляций в структуре сигналов. Эти методы базируются на идеологии одномерных случайных блужданий («*random walk*»). Прежде чем перейти к непосредственному изложению данных подходов, введем некоторые обозначения. Предположим, что автоковариационная функция $K(\tau)$ демонстрирует степенную зависимость для значений временного аргумента $\tau > \tau_1$

$$K(\tau) \sim \tau^{-\gamma}.$$

Тогда для спектральной плотности мощности в частотном диапазоне $\omega < 2\pi/\tau_1$ будет справедливо следующее выражение

$$S(\omega) \sim \omega^{-\beta}. \quad (7.1)$$

Величины γ и β принято называть *скейлинговыми показателями* (в зарубежной литературе иногда используют термин *скейлинговые экспоненты*). Для процессов с конечной энергией

$$\int_0^{\infty} S(\omega) d\omega < \infty$$

значение β ограничено диапазоном $0 < \beta < 1$. Если скейлинговый показатель γ на больших временах также принимает значение в диапазоне $0 < \gamma < 1$ (что типично для спадающей до нуля функции), то между β и γ существует простое соотношение

$$\beta = 1 - \gamma.$$

Если автоковариационная функция спадает быстрее чем τ^{-1} , то в пределе $\omega \rightarrow 0$ спектральная плотность мощности будет приближаться к постоянному значению $\beta = 0$. При этом на больших частотах спектральная плотность может вести себя в соответствии с зависимостью (7.1). Как уже отмечалось, при больших τ АКФ принимает малые значения (вблизи нуля), что приводит к росту статистических ошибок. В результате интересующая нас характеристика γ не может быть определена, так как АКФ просто «растворяется» в шуме. Основная идея флуктуационного анализа (от англ. *fluctuation analysis*) состоит в том, чтобы преобразовать спадающую АКФ в некоторую возрастающую функцию, которая будет менее чувствительна к статистическим ошибкам. Данный анализ применяется именно с целью выявления эффектов *длительных корреляций* в исследуемом процессе (анализ корреляций при малых τ целесообразно проводить с помощью классического метода расчета АКФ). Существуют различные варианты методов флуктуационного анализа. Во всех вариантах на первом шаге алгоритма проводится суммирование значений временного ряда $x(i)$, $i = 1, \dots, N$

$$y(k) = \sum_{i=1}^k x(i), \quad (7.2)$$

где $k = 1, \dots, N$. Сумму (7.2) называют *профилем* переменной x , ее иногда рассматривают как *одномерное случайное блуждание*, при котором изменение состояния в момент времени с номером i происходит на величину $x(i)$. Далее мы проанализируем различные характеристики, базирующиеся на построенной таким образом функции $y(k)$.

7.2 Простой флуктуационный анализ

В рамках данного метода последовательность $y(k)$ делится на непрерывающиеся по времени сегменты длиной s , и рассматриваются приращения значений y в конечных точках этих сегментов:

$$\Delta y_j(s) = y(js + 1) - y((j - 1)s + 1),$$

которые вычисляются для $j = 1, \dots, J$, где $J = \left[\frac{N-1}{s} \right]$, а квадратные скобки обозначают целую часть числа. Далее вводят следующую характеристику

$$F_{fa}^2(s) = \frac{1}{J} \sum_{j=1}^J \Delta y_j(s)^2 = \overline{\Delta y(s)^2},$$

где черта сверху означает усреднение по всем сегментам.

Для автоковариационной функции процесса $x(i)$, спадающей по закону $K(\tau) \sim \tau^{-\gamma}$, $\tau > t_0$, существует взаимосвязь между величиной τ и скейлинговым показателем, описывающим степенной характер зависимости функции $F_{fa}^2(s) \sim s^\alpha$. В частности, при $0 < \gamma < 1$ справедливо следующее выражение:

$$F_{fa}^2(s) \sim s^{2-\gamma}, \quad \alpha = 2 - \gamma.$$

Аналогично, если наблюдается степенной характер частотной зависимости функции спектральной плотности мощности $S(\omega) \sim \omega^{-\beta}$ на низких частотах, то в соответствии с соотношением $\beta = 1 - \gamma$ можно записать

$$F_{fa}^2(s) \sim s^{\beta+1}, \quad \alpha = \beta + 1.$$

Характеристика $F_{fa}^2(s)$ слабо чувствительна к быстрым флуктуациям и не позволяет обеспечивать хорошее разрешение различных степенных законов в высокочастотной области. С ее помощью также не удастся обеспечить хорошее разрешение промежуточных скейлинговых характеристик, если анализируются сложные сигналы, для которых поведение $K(\tau)$ нельзя описать одним показателем γ .

7.3 Анализ флуктуаций относительно тренда

В рамках данного метода¹ сигнал вначале приводится к нулевому среднему значению

$$x^*(i) = x(i) - \langle x(i) \rangle.$$

¹В зарубежной литературе название метода звучит как *detrended fluctuation analysis (DFA)*.

По аналогии с предыдущим подходом вычисляется сумма

$$y(k) = \sum_{i=1}^k x(i),$$

после чего последовательность $y(k)$ разбивается на участки (сегменты) равной длины n .

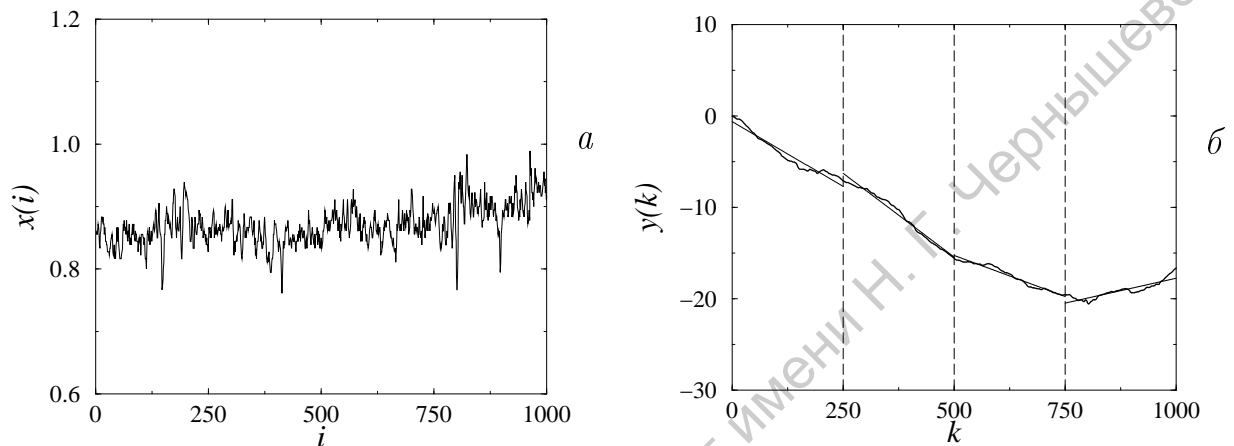


Рис. 7.1. Анализируемый сигнал (а) и последовательность $y(k)$ (б). Во всех сегментах (выделенных пунктирными линиями) проведена линейная аппроксимация локального тренда.

Для каждого участка проводится линейная аппроксимация данных – методом наименьших квадратов определяется уравнение прямой $y_n(k)$, представляющей локальный тренд в пределах одного сегмента. Далее он устраняется в каждом сегменте путем перехода к разности $y(k) - y_n(k)$, и проводится анализ среднеквадратического отклонения от локального тренда, то есть вычисляется функция

$$F_{dfa}(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2}.$$

Такие вычисления повторяются для разных масштабов (то есть различных размеров участков), чтобы получить зависимость $F_{dfa}(n)$ в широком диапазоне значений параметра n . Обычно с ростом n функция $F_{dfa}(n)$ также возрастает, и линейная зависимость $\lg F_{dfa}$ от $\lg n$ свидетельствует о наличии свойства масштабной инвариантности (скейлинга), которое проявляется в степенной зависимости

$$F_{dfa}(n) \sim n^\alpha. \quad (7.3)$$

В этом случае флуктуации можно охарактеризовать скейлинговым показателем α , представляющим собой наклон соответствующего линейного

участка при представлении зависимости (7.3) в двойном логарифмическом масштабе.

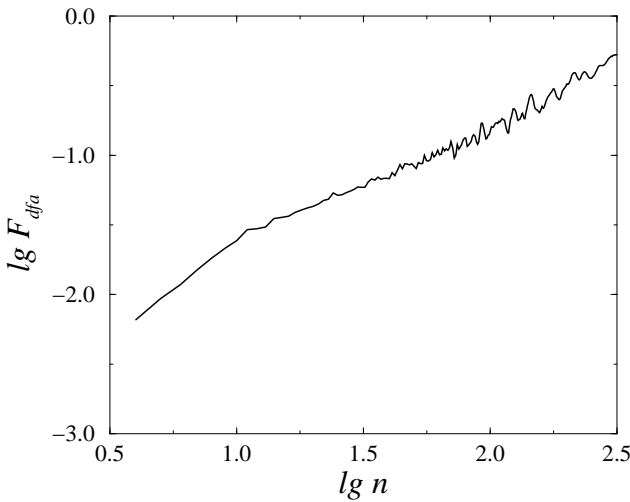


Рис. 7.2. Зависимость $F_{dfa}(n)$, изображенная в двойном логарифмическом масштабе, демонстрирует наличие участка, близкого к линейному ($\lg n > 1$).

Рассмотрим вначале процесс, для которого значения переменных $x^*(i)$ в соседние моменты времени являются полностью некоррелированными. Потери корреляций можно достичь, например, если случайным образом перемешать порядок точек временного ряда (в этом случае часто говорят о «суррогатных» данных). В отсутствие корреляций $\alpha = 0.5$. Если в сигнале существуют только корреляции небольшой длительности, начальный наклон может отличаться от 0.5, но α будет приближаться к 0.5 для больших n . Значение $0.5 < \alpha < 1$ указывает на наличие длительных корреляций, при которых за большим (по сравнению со средним) значением переменной x чаще следует большое, а за малым – малое. Диапазон $0 < \alpha < 0.5$ характеризует отличный тип корреляций – чередование больших и малых значений (антикоррелированная динамика). Особый случай $\alpha = 1$ соответствует $1/f$ -шуму, часто встречающемуся в динамике многих биологических систем. При $\alpha > 1$ корреляции существуют, но они не обязательно носят степенной характер. Например, $\alpha = 1.5$ характеризует винеровский случайный процесс. Скейлинговый показатель α может служить индикатором, который описывает «гладкость» исходного временного ряда: чем больше α , тем более гладким является анализируемый процесс. С этой точки зрения $1/f$ -шум может интерпретироваться как компромисс между полностью некоррелированным белым шумом и более «гладким» винеровским процессом².

Для бесконечного временного ряда $\alpha = (\beta + 1)/2$, то есть скейлинговый показатель метода анализа флуктуаций относительно тренда связан со скейлинговым показателем функции спектральной плотности мощности, а, следовательно, и с характеристикой спада АКФ. Преимущество данного метода по сравнению с простым флуктуационным анализом состоит в

² В данном случае о степени «гладкости» можно говорить лишь в кавычках, характеризуя внешний вид сигнала, так как винеровский процесс не является дифференцируемым.

том, что появляется возможность анализировать нестационарные процессы. Хотя мы пытаемся аппроксимировать тренд прямой линией (что на практике не всегда эффективно), можно рассматривать и другие варианты аппроксимации (полиномами, сплайнами и т.п.), также рассматривая флуктуации относительно локального тренда. Если же при анализе нестационарных процессов тренд вовсе не устранять, то можно сделать ошибочные выводы при расчете скейлинговых характеристик. Рассматриваемый вариант флуктуационного анализа в настоящее время является одним из наиболее широко используемых методов исследования длительных корреляций в нестационарных процессах. Он очень часто применяется в исследованиях сложной динамики биологических систем по экспериментальным данным. В частности, он активно используется в исследованиях структуры символьных последовательностей, таких как цепочки нуклеотидов ДНК. В последнем случае в зависимости от того, какой нуклеотид встречается, значения $x(i)$ выбирают равными либо $+1$, либо -1 , и проводится поиск корреляций различной длительности в полученных бинарных последовательностях. Известно много других применений данного метода, изначально предложенного для исследования эффектов длительных корреляций в структуре сердечного ритма и решения задач медицинской диагностики.

Порядок выполнения работы

Задание 1. Провести анализ корреляционных свойств тестовых сигналов (белый шум и винеровский случайный процесс). С этой целью запустить программу «dfa.x» и указать файлы данных «noise1.dat» и «noise2.dat». Проанализировать погрешность расчета величины α в зависимости от длительности анализируемого временного ряда (выбирая N в диапазоне от 100 до 10000).

Задание 2. Для тестовых примеров, используемых в предыдущем задании, провести сравнение погрешности расчета α в области длительных корреляций ($\lg n > 2$) и корреляций малой длительности ($\lg n < 1$). Выбрать $N = 10000$. Сделать вывод об эффективности используемого метода в разных диапазонах значений n .

Задание 3. Провести анализ корреляционных свойств физиологических процессов (последовательность кардиоинтервалов электрокардиограммы человека в обычном состоянии и при физической нагрузке). Сопоставить полученные результаты.

Контрольные вопросы

1. Чем вызваны сложности исследования эффектов длительных корреляций в рамках классического корреляционного анализа?
2. Как связаны между собой скейлинговые экспоненты, характеризующие поведение автоковариационной функции и спектра мощности?
3. В чем состоит основная идея метода флуктуационного анализа?
4. Почему метод анализа флуктуаций относительно тренда применим в исследованиях структуры нестационарных процессов?
5. Как связаны характеристики флуктуационного анализа со скейлинговыми экспонентами классического спектрально-корреляционного анализа?
6. Объясните, каким образом по величине α можно сделать вывод об особенностях корреляций в анализируемом процессе?
7. При решении каких задач применяется флуктуационный анализ?

8 Вейвлет-анализ

8.1 Предварительные замечания

Вейвлеты¹ появились в восьмидесятых годах двадцатого века как альтернатива классическому спектральному анализу, основанному на преобразовании Фурье. Возникновение теории вейвлетов считается одним из важнейших событий в математике за последние десятилетия, поскольку это, пожалуй, единственная новая математическая концепция, которая сразу же после ее появления стала восприниматься в качестве инструмента *прикладных* исследований практически во всех естественных науках и многих областях техники. В настоящее время вейвлеты широко применяются при анализе и синтезе различных сигналов, при обработке изображений, для сжатия больших объемов информации, для распознавания образов и т.д. В данном пособии рассматривается одна из задач, для решения которой используется теория вейвлетов – анализ нестационарной динамики ритмов в сигналах различной природы.

Очень многие процессы в природе являются нестационарными и демонстрируют изменения во времени своих статистических свойств. Примерами могут служить переходные процессы в радиофизических устройствах, атмосферная и гидродинамическая турбулентность, нестационарные волны в океане, нестационарные геофизические и физиологические сигналы. Классические методы обработки экспериментальных данных (спектрально-корреляционный анализ, вычисление моментных функций и т.д.) представляют собой инструменты исследования стационарных случайных процессов; их применение для анализа нестационарных данных зачастую приводит к различным проблемам в интерпретации полученных результатов. В частности, наличие двух пиков в спектре мощности с некрратными частотами может соответствовать принципиально разным ситуациям: в динамике изучаемой системы могут одновременно присутствовать два независимых ритма или может наблюдаться процесс переключения частоты, и в каждый момент времени удается зафиксировать только один ритмический процесс.

Довольно часто при исследовании экспериментальных данных используется идеология анализа систем с медленно-меняющимися параметрами: предполагается, что в течение небольших интервалов времени свойства процесса меняются незначительно, и его можно рассматривать как стацио-

¹ Термин «вейвлет» в дословном переводе означает «маленькая волна».

нарный, применяя классический аппарат статистической обработки. Такой подход следует признать эффективным лишь в том случае, если нестационарность ассоциируется с низкочастотной областью спектра по отношению к динамике, представляющей интерес для исследователя. Если же свойства процесса даже в пределах сравнительно малых временных интервалов успевают существенно поменяться, то есть два варианта дальнейших действий – либо отказываться от классических методов анализа временных рядов и ориентироваться на специальные методики, либо тщательно проводить предварительную обработку экспериментальных данных, выбирая только те участки сигналов, на которых их можно считать приближенно стационарными. Но даже при условии осуществления предварительной обработки данных целесообразно проводить последующий анализ на основе наиболее универсальных методов, которые могут применяться независимо от свойства стационарности случайных процессов. Таких универсальных инструментов существует немного, и наиболее известным среди них является вейвлет-анализ.

К настоящему времени вейвлеты продемонстрировали свою эффективность при решении очень широкого круга задач, связанных с подавлением шумов, сжатием больших объемов информации, анализом изображений, синтезом сигналов и т.д. Вероятно, не было ни одной другой математической концепции, которая бы быстро приобрела столь широкое прикладное значение в самых разных областях науки. Возможности теории вейвлетов очень широки. Иногда говорят, что вейвлет-анализ – это математический аппарат для решения задач, которые оказались «не по зубам» классическому спектральному анализу. Вейвлеты появились изначально в качестве альтернативы оконным преобразованиям Фурье для исследования структуры сигналов. Однако вскоре стало ясно, что они представляют собой намного большее чем просто другой вариант спектрального анализа. Перечислим лишь некоторые направления, где в настоящее время широко применяются вейвлеты:

1) Анализ структуры нестационарных процессов. Вейвлет-преобразование позволяет исследовать изменение во времени мгновенных амплитуд и частот ритмов. Это свойство особенно хорошо подходит для изучения нестационарных данных, например, процессов, возникающих в динамике живых систем.

2) Сжатие больших объемов информации. Вейвлет-преобразование является основой стандарта сжатия видеоизображений MPEG4. При допустимых потерях качества изображения (несущественных мелких деталей) существует возможность провести сжатие данных почти в 300 раз.

3) Компьютерная графика. Вейвлеты позволяют осуществлять редактирование изображения с переменным разрешением, поклеточное представление поверхностей по контуру, формировать запрос для поиска изображения по образцу.

4) Системы передачи данных. Обладая высокой устойчивостью к воздействию помех, вейвлет-преобразование все активнее используется в тех областях, где раньше применялись другие методы (например, преобразование Фурье), в частности, для передачи информации.

По аналогии с различными формами преобразования Фурье существует непрерывное и дискретное вейвлет-преобразование. Во многих практических приложениях используются дискретные вейвлеты, которые позволяют осуществлять более точное представление сигнала, особенно в задачах, связанных с его сжатием и последующим восстановлением. Они обеспечивают возможность осуществления процедуры быстрого преобразования, что является актуальным не только при передаче информации, но и при анализе случайных процессов. Дискретное преобразование наиболее естественно в случае дискретных оцифрованных данных (временных рядов) и применяется на практике, если требуется быстро вычислить те или иные характеристики. Непрерывное преобразование содержит избыточную информацию. Тем не менее, при анализе структуры сложных сигналов возникают ситуации, когда данная избыточность может быть полезным свойством, позволяющим получать более наглядные и ясные результаты. Информацию, которую можно извлечь из непрерывного вейвлет-преобразования, например, об изменении характерных частот ритмических процессов и их взаимодействии, легче анализировать – в этом случае можно визуально отслеживать поведение во времени мгновенных частот и амплитуд на разных масштабах. Именно эта возможность и будет рассмотрена нами далее.

8.2 Непрерывное вейвлет-преобразование и его свойства

В отличие от бесконечно осциллирующих гармонических функций, которые применяются в классическом спектральном анализе, вейвлет-преобразование использует «локальные» функции или «маленькие волны» – вейвлеты. Несколько примеров вейвлетов, часто применяемых на практике, изображены на рис. 8.1. Данные функции могут быть подвергнуты двум математическим преобразованиям, а именно изменению масштаба (сжатию или растяжению) и смещению вдоль анализируемого сигнала – рис. 8.2. Для чего это нужно? Во-первых, появляется возможность изучать локальные свойства сигнала на разных масштабах наблюдения. Можно образно представить, что мы вначале смотрим на сигнал невооруженным глазом, затем через увеличительное стекло и, наконец, с помощью микроскопа рассматриваем совсем мелкие детали (а при необходимости, наоборот, отойдем на несколько шагов и взглянем на сигнал издали). Во-вторых, смещение вейвлета вдоль анализируемого процесса позволяет отслеживать, как раз-

личные характеристики сигнала меняются во времени. Очень популярна и наглядна интерпретация вейвлет-преобразования как метода математического микроскопа: путем перемещения вейвлета вдоль сигнала меняется точка фокусировки микроскопа, изменение масштаба функций за счет сжатия или растяжения аналогично выбору разрешения, а использование различных вейвлетов по сути соответствует выбору объектива микроскопа (его оптических свойств).

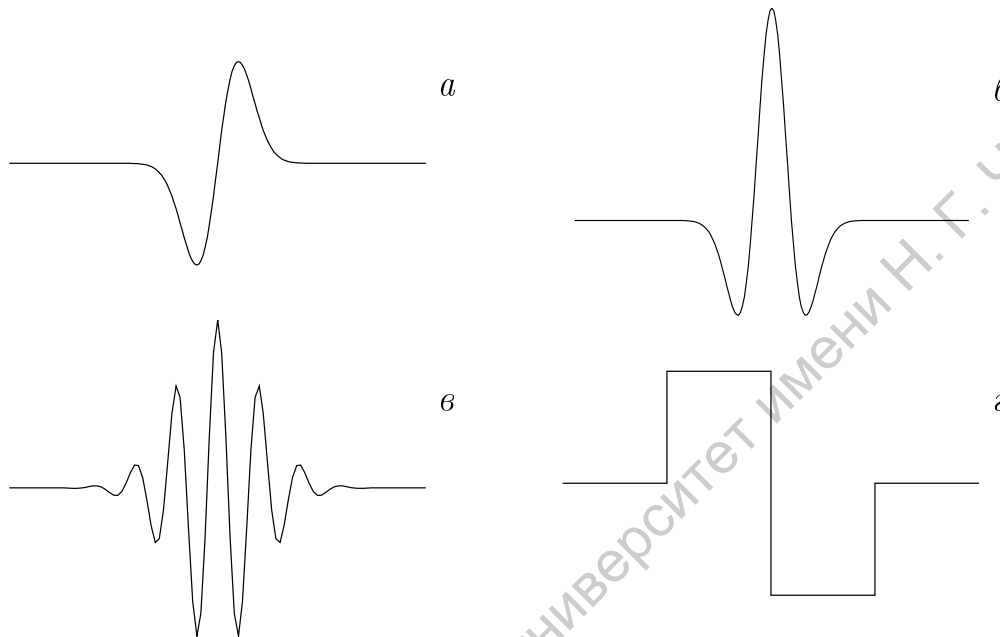


Рис. 8.1. Примеры вейвлетов: (а) WAVE, (б) МНАТ, (в) Морле, (г) Хаар

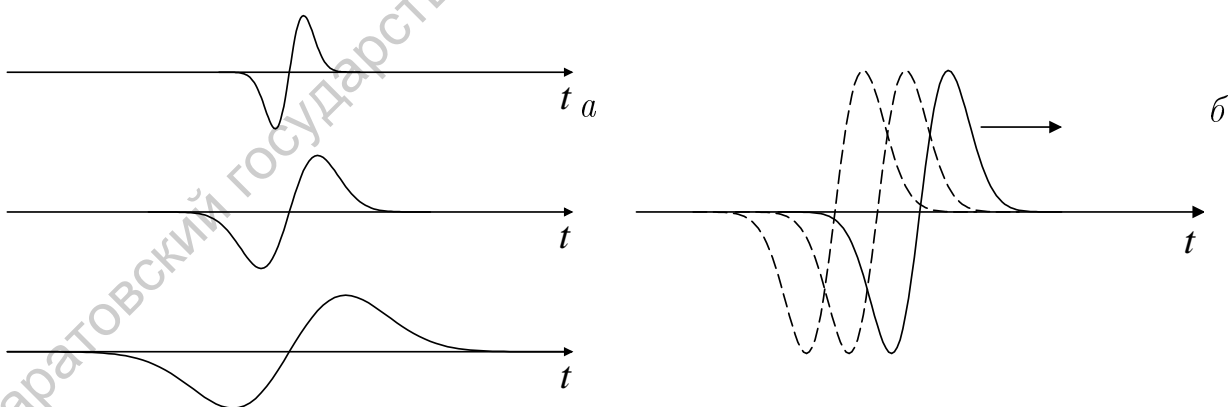


Рис. 8.2. Математические преобразования функций – изменение масштаба (а) и смещение вдоль сигнала (б)

Если в рамках классического Фурье-анализа спектр сигнала $x(t)$ представляет собой функцию одной переменной (частоты), то вейвлет-преобразование оперирует уже с двумя переменными: 1) масштаб или частота, 2)

локализация во времени. Одной из принципиальных особенностей теории вейвлетов является то, что частота и время рассматриваются как независимые переменные, а именно это и позволяет говорить о «локальном» спектральном анализе. Поместив вейвлет в выбранной точке сигнала и проведя его перемасштабирование, можно получить информацию о частотных составляющих исследуемого процесса в окрестности фиксированного момента времени. С этой целью фактически осуществляется простое перемножение вейвлета и сигнала – если они коррелируют, то при перемножении будет получено большое значение, если не коррелируют – то малое. Затем аналогичные математические операции повторяются для других масштабов и в другие моменты времени. В результате будет получена двумерная развертка одномерного сигнала $x(t)$ (рис. 8.3). Выбор масштаба a и смещения b может быть произвольным (при проведении непрерывного вейвлет-преобразования), или данные величины могут принимать фиксированный набор значений, например, $a = 2^k$ (в случае дискретного преобразования). В дальнейшем мы будем рассматривать только непрерывное преобразование, которое допускает более наглядное представление результатов.

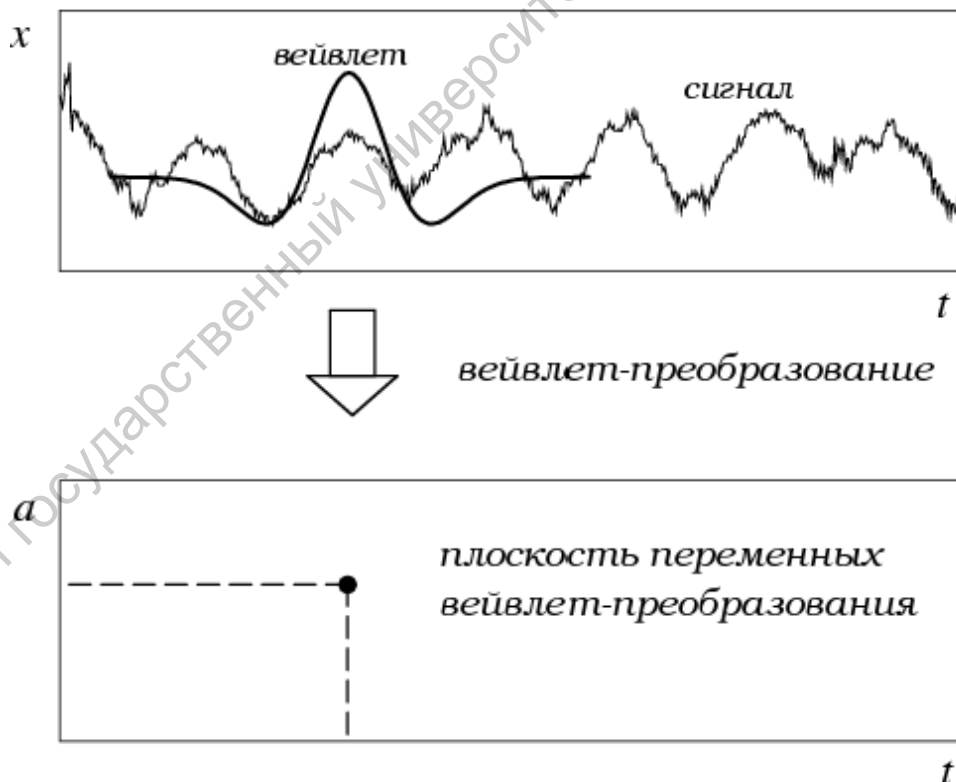


Рис. 8.3. Анализируемый сигнал, вейвлет и двумерная развертка сигнала при вейвлет-преобразовании. Пунктирными линиями обозначены текущие значения параметров масштаба и смещения

Одним из наиболее важных моментов при проведении анализа является выбор конкретного вейвлета $\psi(t)$ (его называют базисным или «материнским»). Этот выбор зависит от решаемой задачи и от самого сигнала.

ла. Например, чтобы определить фазу колебательного процесса на основе вейвлет-преобразования, используются комплексные функции. Для анализа свойств самоподобия процесса на разных масштабах целесообразно применять вещественные функции. При выборе вейвлета большую роль играет интуиция и практический опыт исследователя, а также существующие традиции использования тех или иных функций при решении определенного круга задач. Это вовсе не означает, что с другим вейвлетом нельзя получить правильные результаты. Выбор разных базисных функций можно интерпретировать как выбор разных объективов для микроскопа: если данный объектив позволяет рассмотреть нужные детали, то он подходит для решаемой задачи (более того, если взять другой объектив, то при этом мы не увидим ничего нового). Иногда также используют аналогию с разными системами исчисления. Применение десятичной или двоичной системы определяется удобством и традициями. В зависимости от выбора системы исчисления решение конкретной задачи может потребовать больших или меньших усилий, и интуиция при этом играет важную роль. В то же время существуют системы исчисления, применение которых просто исключено при проведении расчетов (например, римские цифры не подходят, если нужно осуществлять операции умножения). Ситуация с вейвлетами действительно в чем-то аналогичная.

Рассмотрим теперь, какими признаками должна обладать функция $\psi(t)$, чтобы ее можно было рассматривать в качестве вейвлета:

- **Локализация.** Вейвлет $\psi(t)$ должен быть локализован и во временном пространстве, и по частоте. Это означает, что функция спадает до нуля за пределами некоторой области, и для практических целей чем быстрее она спадает, тем лучше.
- **Нулевое среднее.** Необходимым признаком вейвлета является условие

$$\int_{-\infty}^{\infty} \psi(t) dt = 0.$$

Чтобы игнорировать медленную нестационарность (полиномиальный тренд) бывает необходимо выполнение более сильного требования:

$$\int_{-\infty}^{\infty} t^m \psi(t) dt = 0,$$

которое означает равенство нулю первых m моментов функции $\psi(t)$.

- **Конечная энергия.** Должно выполняться условие:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty.$$

Очень популярны на практике вейвлеты, представляющие собой производные функции Гаусса, в частности, WAVE-вейвлет

$$\psi(t) = te^{-t^2/2}$$

и МНАТ-вейвлет

$$\psi(t) = (1 - t^2)e^{-t^2/2}.$$

Среди комплексных функций чаще всего применяется хорошо локализованный и по времени, и по частоте вейвлет Морле

$$\psi(t) = \pi^{-1/4} \left(e^{j2\pi f_0 t} - e^{-(2\pi f_0)^2/2} \right) e^{-t^2/2},$$

где f_0 – параметр, называемый центральной частотой. Второе слагаемое в скобках осуществляет корректировку при проведении вейвлет-преобразования сигналов с ненулевым средним значением. В случае $f_0 \gg 0$ им можно пренебречь, получив упрощенную и наиболее часто применяемую на практике формулу вейвлета Морле:

$$\psi(t) = \pi^{-1/4} e^{j2\pi f_0 t} e^{-t^2/2}. \quad (8.1)$$

Здесь множитель $e^{j2\pi f_0 t}$ представляет собой комплексную форму записи гармонической функции, которая промодулирована по амплитуде функцией Гаусса $e^{-t^2/2}$, а нормировочный коэффициент $\pi^{-1/4}$ необходим для того, чтобы величина энергии вейвлета была равна единице.

После выбора «материнского» вейвлета $\psi(t)$ на его основе строится базис

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right),$$

где a – масштабный коэффициент (он определяет сжатие или растяжение функции), b – параметр сдвига, с помощью которого вейвлет перемещается вдоль анализируемого сигнала $x(t)$.

Непрерывное вейвлет-преобразование имеет вид:

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-b}{a} \right) dt = \int_{-\infty}^{\infty} x(t) \psi_{a,b}^*(t) dt, \quad (8.2)$$

где звездочка обозначает комплексно сопряженную величину. Результат преобразования (8.2) представляет собой поверхность в трехмерном пространстве. Для наглядного представления об особенностях исследуемого сигнала существуют различные варианты визуализации поверхности коэффициентов $W(a, b)$. Обычно рассматривают ее проекцию на плоскость ab и различными оттенками цвета изображают интенсивность амплитуд вейвлет-преобразования (по аналогии с географическими картами). Такой

способ позволяет проследить, как на разных масштабах происходит изменение во времени свойств сигнала. Другой способ визуализации состоит в построении картин линий локальных экстремумов поверхности $W(a, b)$ или так называемого «скелетона». Считается, что данный скелетон фактически содержит всю основную информацию о вейвлет-преобразовании, не «загроможденную» лишними деталями. Для иллюстрации рассмотрим тестовый пример – сигнал, обладающий свойствами самоподобия (фрактальной структурой) в широком диапазоне масштабов (рис. 8.4, *а*). Проекция поверхности модулей коэффициентов $W(a, b)$ на плоскость ab и выделенный скелетон представлены на рис. 8.4, *б, в*. В вычислениях использовался МНАТ-вейвлет.

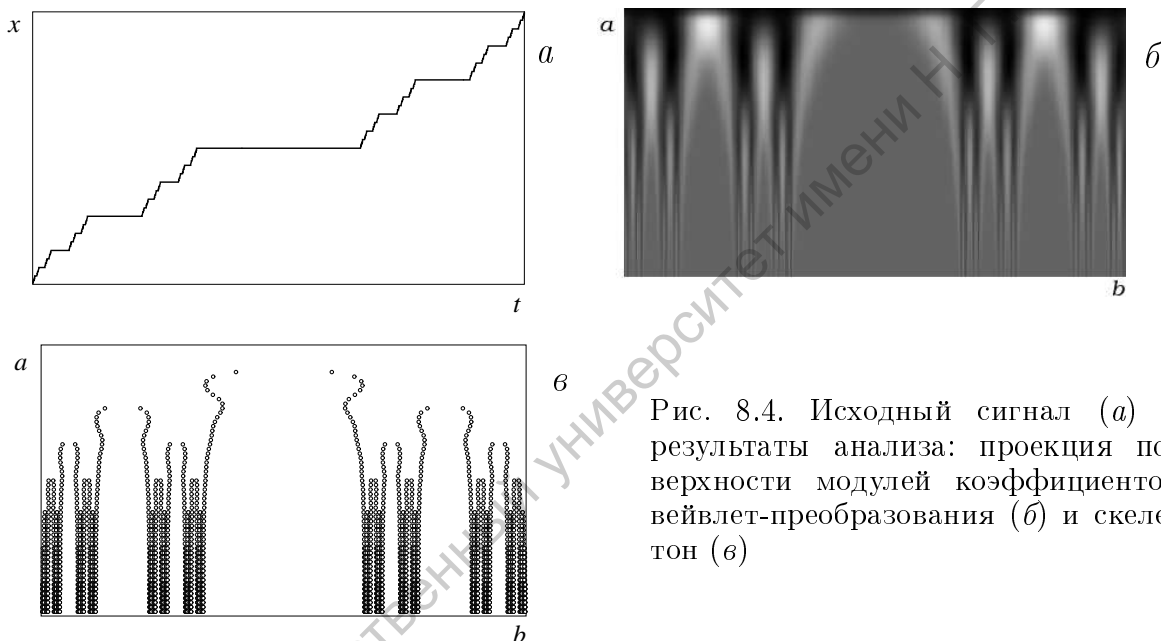


Рис. 8.4. Исходный сигнал (*а*) и результаты анализа: проекция поверхности модулей коэффициентов вейвлет-преобразования (*б*) и скелетон (*в*)

По аналогии с классическим спектром мощности в рамках вейвлет-анализа также оперируют с энергетическими характеристиками. Плотность энергии сигнала можно определить следующим образом:

$$E(a, b) = W^2(a, b).$$

Но при этом у вейвлет-преобразования существует ряд принципиальных особенностей:

1. Локальный и глобальный энергетические спектры. Использование ограниченных по времени функций позволяет изучать локализованные характеристики процессов. Локальный спектр энергии можно интерпретировать как сечение поверхности $E(a, b)$ в фиксированный момент времени t_0

$$E(a, t_0) = W^2(a, t_0).$$

Существование локального энергетического спектра представляет собой одну из основных особенностей вейвлет-преобразования, которая позволяет изучать временную динамику мгновенных характеристик нестационарного процесса.

Распределение по масштабам полной энергии характеризуется глобальным спектром энергии вейвлет-преобразования

$$E(a) = \int_{-\infty}^{\infty} E(a, b) db,$$

который часто называют скалограммой. Глобальный энергетический спектр является аналогом классического спектра мощности. Отличие состоит лишь в том, что он является более гладким – $E(a)$ фактически представляет собой спектр мощности, сглаженный на каждом масштабе спектром Фурье анализирующего вейвлета.

2. Проблема нормировки. В рамках классического спектрального анализа мощность колебаний равна площади под кривой функции спектральной плотности $S(f)$, а величина спектрального пика позволяет определить амплитуду колебаний на соответствующей частоте. В случае вейвлет-анализа ситуация несколько сложнее – при корректном оценивании энергии по полученному спектру не удастся правильно вычислить амплитуду, и наоборот. В зависимости от того, какую характеристику необходимо вычислить, проводятся различные нормировки плотности энергии $E(a, b)$. Для иллюстрации рассмотрим, как меняется спектр гармонической функции при разных значениях частоты f (рис. 8.5). Воспользуемся вейвлетом Морле с $f_0 = 1$, в этом случае $f \approx 1/a$. С увеличением частоты происходит «расплывание» спектрального пика (причины этого явления будут рассмотрены позднее). Если необходимо вычислить энергию колебаний в фиксированной полосе частот, то данный эффект не является принципиальным – увеличение ширины спектрального пика сопровождается уменьшением его амплитуды, и энергия колебаний (площадь под кривой $E(f)$ на рис. 8.5б) сохраняется неизменной. Однако оценка амплитуды для каждой синусоиды (как корень из величины энергии на данной частоте) дает разные результаты – амплитуда будет уменьшаться с ростом f .

Для корректной оценки соотношений между амплитудами колебательных процессов необходимо провести специальную нормировку (с этой целью можно заменить множитель $1/\sqrt{a}$ перед знаком интеграла в формуле (8.2) на множитель $1/a$). Это позволит правильно определить амплитуды ритмических процессов с разными периодами (с точностью до некоторой константы, если сравнивать с Фурье-анализом), но приведет к нарушению энергетических характеристик. Суммируя вышесказанное, можно говорить о том, что в случае вейвлет-преобразования значения энергии колебаний в некотором частотном диапазоне и амплитуды характерных ритмов целесообразно рассматривать отдельно, выбирая тот вариант нормировки, при

котором удобнее вычислить требуемую характеристику.

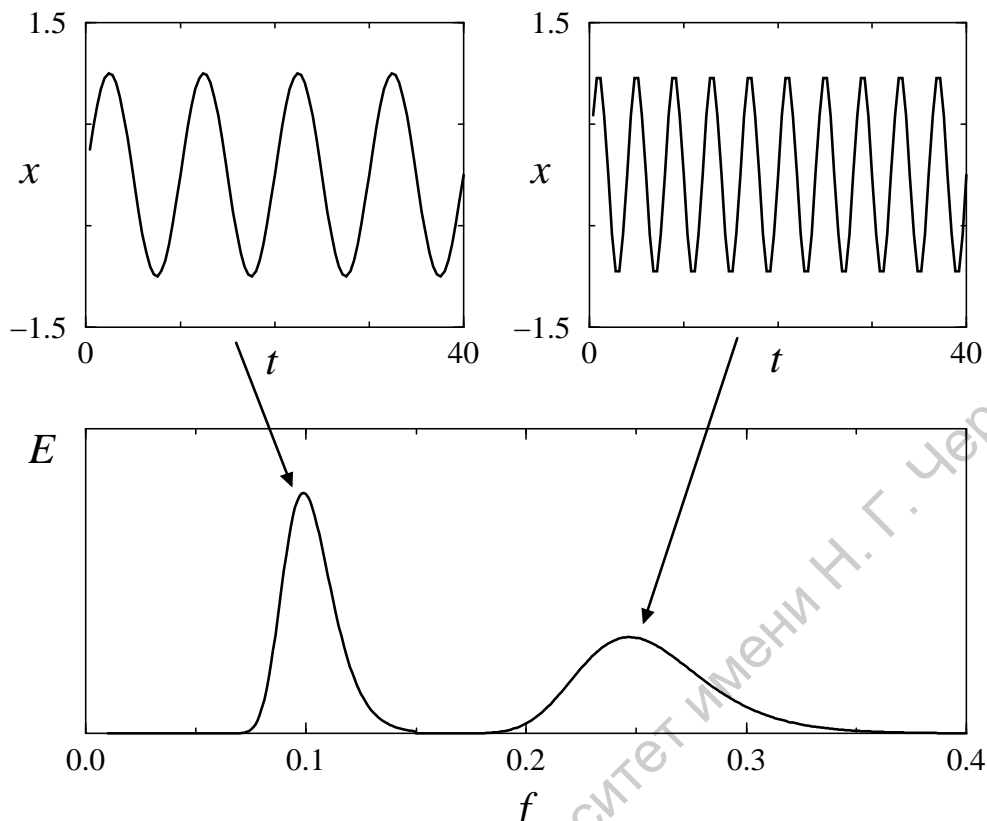


Рис. 8.5. Спектр, вычисленный на основе вейвлет-преобразования двух гармонических функций с единичной амплитудой и разной частотой

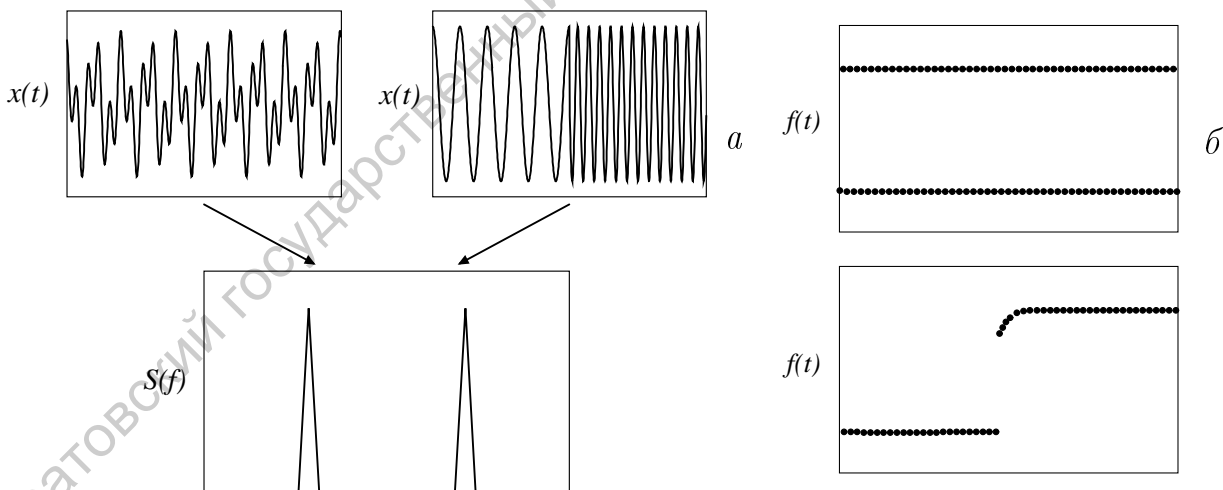


Рис. 8.6. Классический спектральный анализ не отличает два разных сигнала (а), тогда как расчет мгновенных частот на основе локальных энергетических спектров вейвлет-анализа (б) позволяет эти сигналы четко идентифицировать

Одним из наиболее важных моментов при сопоставлении вейвлет- и Фурье-преобразований является различная частотно-временная локализация. Преобразование Фурье оперирует с бесконечными во времени функциями \sin и \cos , и для того, чтобы получить спектральную информацию на

выбранной частоте, необходимо вычислить интеграл в диапазоне $(-\infty, \infty)$. Это означает, что необходимо располагать и прошлой, и будущей временной информацией. Кроме того, преобразование Фурье не учитывает то обстоятельство, что частота может меняться во времени. В частности, преобразование Фурье не позволяет отличить сигнал, представляющий собой сумму гармонических функций с разными частотами, от сигнала, демонстрирующего «переключение» частоты – рис. 8.6, а. Использование локализованных функций в рамках вейвлет-анализа эту проблему решает (рис. 8.6, б).

Вейвлеты могут быть рассмотрены в качестве «оконных» функций, имеющих свою ширину во временной (σ_t) и в частотной (σ_f) областях. Ширина определяется как стандартное отклонение (квадратный корень из дисперсии) функций $|\psi_{a,b}(t)|$ и $|\hat{\psi}_{a,b}(f)|$, где $\hat{\psi}_{a,b}(f)$ - Фурье-образ вейвлета. Частотно-временное разрешение вейвлет-преобразования можно охарактеризовать прямоугольником на плоскости tf со сторонами σ_t и σ_f . Этот прямоугольник иногда называют *прямоугольником Гейзенберга*, отражающим известный принцип неопределенности, который характеризует минимальную площадь, занимаемую данным прямоугольником. В случае вейвлет-преобразования эта площадь не может быть меньше $1/4\pi$.

Существование принципа неопределенности означает, что при анализе сигналов невозможно одновременно улучшить и частотное, и временное разрешение: уменьшение ширины вейвлета во временной области сопровождается увеличением ширины спектральных пиков, как это было показано на рис. 8.5. Иными словами, улучшение временного разрешения (возможность ограничиваться меньшими участками сигналов) приводит к ухудшению частотного разрешения (более размытым спектральным пикам). Наоборот, стремление получить узкие линии в спектре неизбежно будет приводить к необходимости рассматривать более длительные фрагменты анализируемого процесса. Это приводит к тому, что частотно-временное окно вейвлет-преобразования является подвижным: прямоугольник Гейзенберга сужается по частоте и расширяется по времени на низких частотах, расширяется по частоте и сужается по времени – на высоких (рис. 8.7).

Данное свойство вейвлет-преобразования является очень важным. Поскольку частота сигнала обратно пропорциональна его продолжительности (периоду колебаний), то для получения высокочастотной информации с хорошей точностью требуется извлекать ее из сравнительно малых временных интервалов, тогда как для извлечения низкочастотной спектральной информации необходимо рассматривать более длительные участки. Отметим, что и в классическом спектральном анализе, основанном на преобразовании Фурье, информация может извлекаться из отдельных фрагментов сигнала (при использовании оконных функций), поэтому целесообразно рассмотреть отличие вейвлет-преобразования от оконного преобразования Фурье.

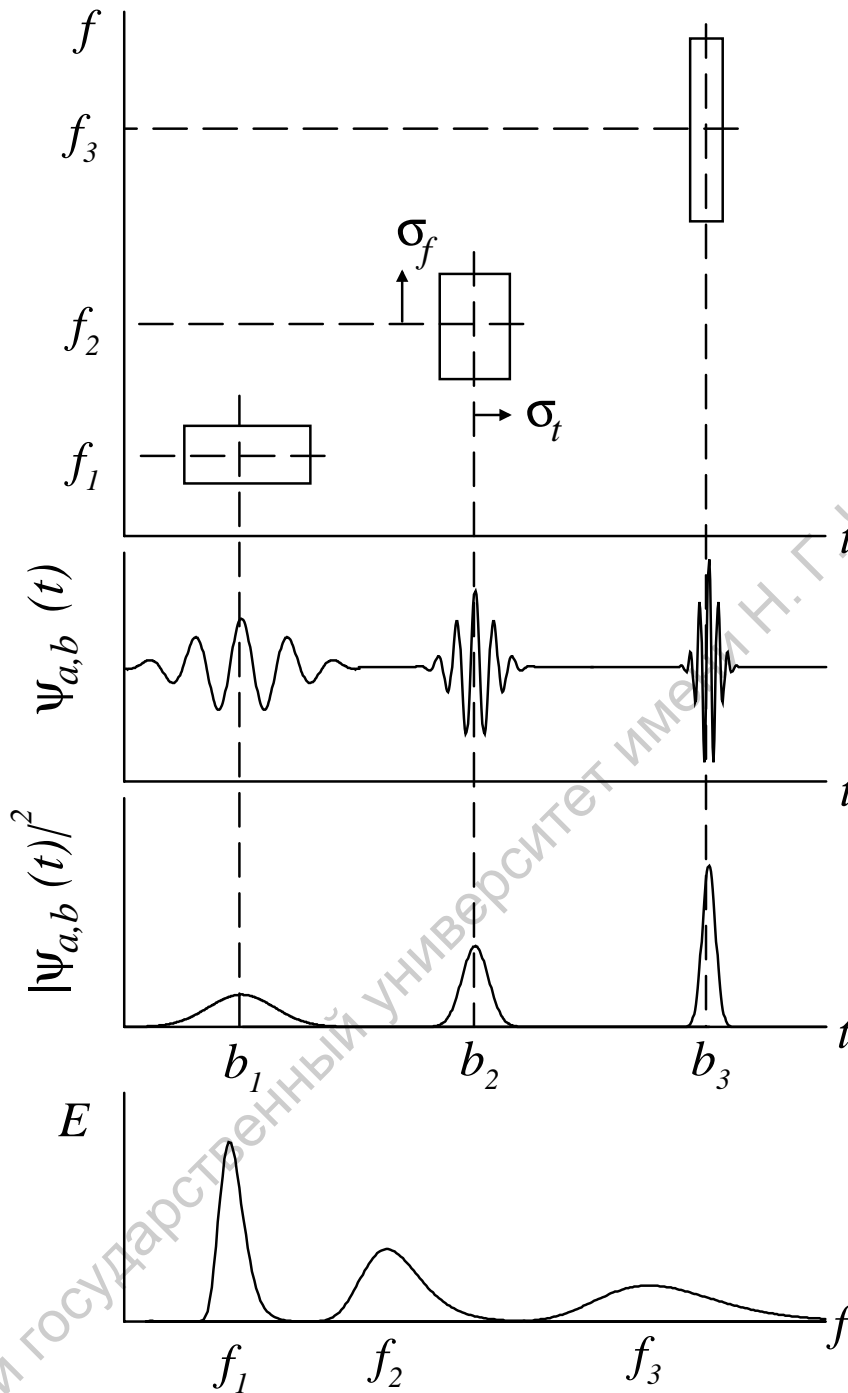


Рис. 8.7. Частотно-временное разрешение вейвлет-анализа. Длина и ширина прямоугольников Гейзенберга (верхний рисунок) зависят от масштаба, но их площадь остается постоянной. С увеличением частоты улучшается временное разрешение, но спектральные пики (нижний рисунок) «расплываются»

Интуитивно представляется, что у этих методов есть нечто общее – вейвлет Морле представляет собой гармоническую функцию, промодулированную по амплитуде функцией Гаусса. Предположим, что мы проводим оконный спектральный анализ на основе преобразования Фурье, также выбрав в качестве временного «окна» функцию Гаусса. Будут ли при этом ре-

зультаты совпадать с результатами вейвлет-анализа? Использование оконной функции во временной области приведет к ограничению спектрального разрешения в области частот и появлению прямоугольника Гейзенберга на плоскости tf (рис. 8.8).

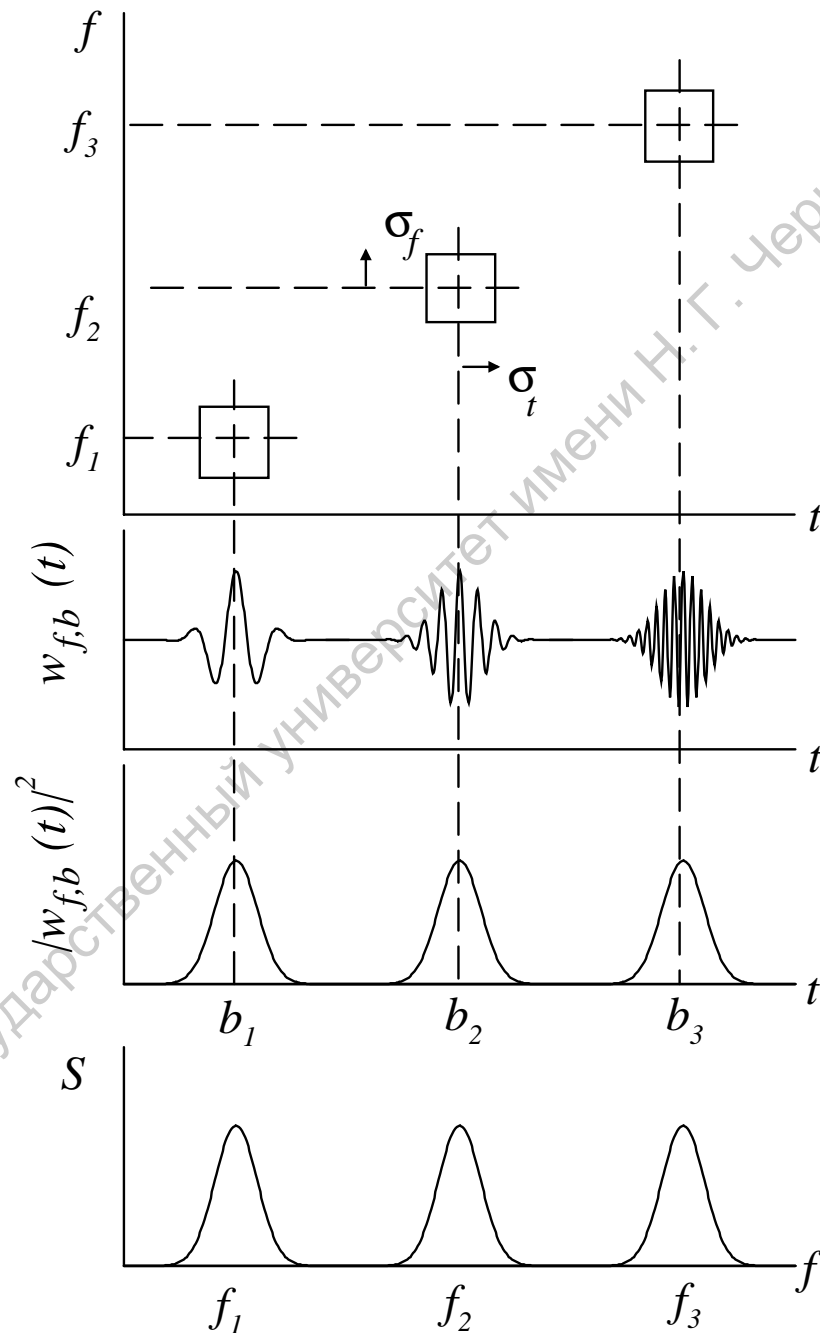


Рис. 8.8. Частотно-временное разрешение оконного преобразования Фурье. Длина и ширина прямоугольников Гейзенберга (верхний рисунок) в данном случае сохраняются постоянными на всех масштабах. Разрешение не зависит от частоты

Но обратим внимание на то, что этот прямоугольник не меняется при изменении частоты – частотное и временное разрешение остаются постоянными

ными на всех частотах. В этом, с одной стороны, есть свои положительные моменты – не происходит «расплывания» спектральных пиков с ростом частоты, не меняется их амплитуда (не требуется никаких нормировок для корректного вычисления одновременно и амплитуд ритмов, и энергии колебаний). С другой стороны, в этом есть и свои недостатки, поскольку на разных частотах в пределы прямоугольника Гейзенберга попадает разное число колебаний. Если их оказывается много, то пропадает возможность проводить по-настоящему локальный анализ, и не удастся отслеживать быстрые изменения во времени свойств сигнала. Вместо выявления быстрых вариаций его характеристик будет получена лишь усредненная в пределах окна информация. Например, вейвлет Морле при $f_0 = 1$ содержит примерно 5 периодов колебаний – этого достаточно для проведения локального спектрального анализа (при умножении вейвлета на сигнал $x(t)$ не учитывается вся информация за пределами этого участка сигнала). Исследование динамики на другой частоте при перемасштабировании вейвлета означает, что мы вновь ограничиваемся участком сигнала, содержащим около 5 колебаний. Это позволяет говорить о локальности проводимого анализа вне зависимости от выбора масштаба наблюдения. Оконный спектральный анализ, основанный на преобразовании Фурье, такую локальность не обеспечивает – при изменении масштаба внутрь окна будет попадать произвольное число периодов (от единиц до сотен).

8.3 Примеры применения

Чтобы нагляднее представить, как вейвлеты позволяют изучать локальную структуру сигналов, рассмотрим гармоническую функцию и проведем расчеты на основе вейвлета WAVE. Согласно формуле (8.2), коэффициенты вейвлет-преобразования зависят от расположения вейвлета (параметра сдвига b) и от масштабного коэффициента a , определяющего сжатие или растяжение базисной функции $\psi_{a,b}(t)$.

На рис. 8.9,а вейвлет демонстрирует ту же «периодичность», что и сам сигнал (синусоида), а параметр b выбран таким образом, чтобы достичь максимального совпадения «накладываемых» друг на друга функций. Очевидно, что это приведет к большой корреляции между сигналом и вейвлетом при выбранных значениях a и b , и при вычислении интеграла (8.2) будет получено большое положительное значение $W(a, b)$. При увеличении b на половину периода синусоиды вейвлет и сигнал оказываются в противофазе. Как следствие, коэффициент $W(a, b)$ принимает большое отрицательное значение. Любые промежуточные значения b приводят к уменьшению коэффициентов вейвлет-преобразования от максимума (рис. 8.9,а) к минимуму (рис. 8.9,б). В частности, рис. 8.9,в иллюстрирует случай, когда $W(a, b) = 0$. На рис. 8.9,а,б,в вейвлет имеет приблизительно ту же «форму»

и «размер», что и сам сигнал в окрестности выбранного значения b . Для очень малого масштаба a (рис. 8.9, *г*) участкам положительных и отрицательных значений вейвлета соответствует почти одна и та же часть периода синусоиды, что приводит к коэффициентам $W(a, b)$, близким к нулю. Таким образом, если a стремится к нулю, то и $W(a, b)$ также уменьшается до нуля. При очень больших a вейвлет-коэффициенты также стремятся к нулю (рис. 8.9, *д*), так как в пределах функции $\psi_{a,b}(t)$ оказывается много участков положительных и отрицательных значений синусоиды, и перемножение вейвлета и сигнала приведет к малым коэффициентам $W(a, b)$.

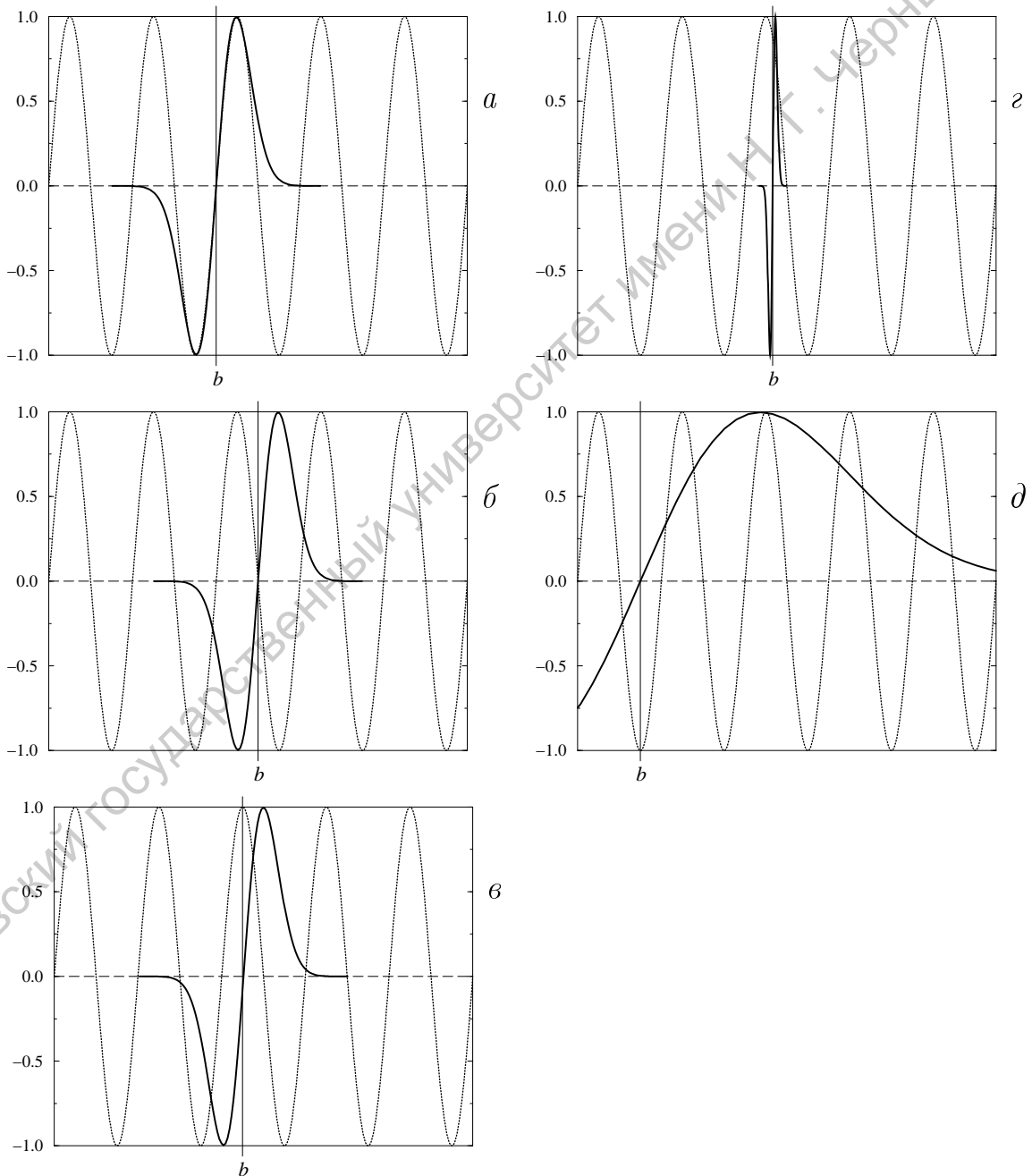


Рис. 8.9. Анализируемый гармонический сигнал и WAVE-вейвлет при различных значениях параметра сдвига и масштабного коэффициента.

Вейвлет-преобразование обычно вычисляется в широком диапазоне зна-

чений параметров a и b . Для рассмотренного примера гармонической функции и вейвлета WAVE проекция поверхности $W(a, b)$ изображена на рис. 8.10, где оттенками цвета показаны изменения коэффициентов от максимального значения (белый) до минимального (черный). Визуальный анализ поверхности вейвлет-коэффициентов в целях извлечения информации об особенностях структуры сигнала в данном случае провести нетрудно – мы видим лишь один характерный масштаб, который не меняется во времени. Для многочастотных и нестационарных процессов анализ соответствующих проекций поверхности $W(a, b)$ является более сложным и неудобным. Чередование светлых и темных «эллипсов» в окрестности каждого ее локального максимума или минимума загромождает рисунок большим количеством деталей, которые на самом деле не очень то нужны для того, чтобы разобраться в структуре анализируемого процесса.

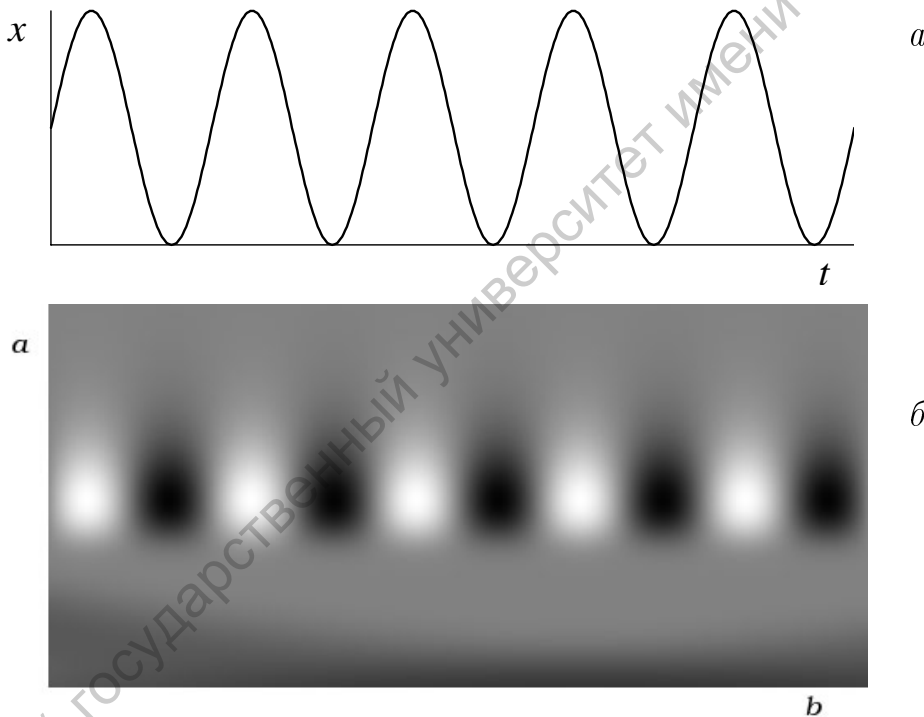


Рис. 8.10. Результаты анализа гармонической функции (а) на основе WAVE-вейвлета, представленные в виде проекции поверхности вейвлет-коэффициентов (б)

Результаты, аналогичные рис. 8.10, будут получены и с применением других вещественных вейвлетов. Чтобы упростить визуальный анализ, целесообразно исключить информацию о фазе за счет рассмотрения модулей вейвлет-коэффициентов и выбора комплексной базисной функции. В частности, для анализа нестационарных многочастотных колебаний и изучения особенностей динамики ритмов удобно пользоваться функцией Морле (8.1). Если ограничиться только изображением скелетона², то в этом слу-

²При анализе ритмов скелетон строится путем поиска локальных экстремумов $E(a, b)$ для фикси-

чае результаты будут более простыми и наглядными. На рис. 8.11 показана зависимость от времени мгновенной частоты синусоиды, вычисленная на основе вейвлет-преобразования. Таким образом, рассматриваемый метод позволяет перейти от изучения исходных сигналов к анализу временных зависимостей мгновенных частот (или амплитуд) ритмических процессов.

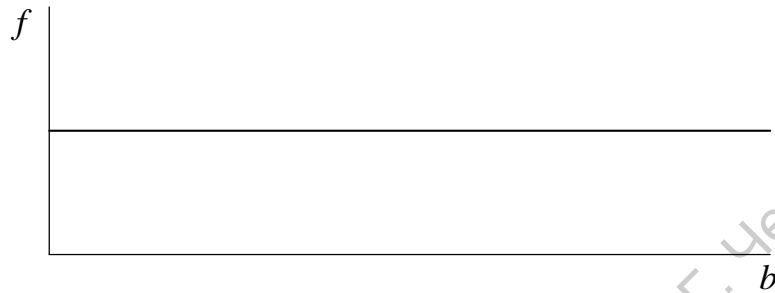


Рис. 8.11. Результат анализа гармонической функции на основе вейвлета Морле (приводится только выделенный скелетон)

Мы выбрали очень простой случай, который не позволяет в полной мере оценить возможности, которые предоставляет вейвлет-анализ. Чтобы лучше понять преимущества вейвлетов по сравнению с классическим спектральным анализом, рассмотрим несколько примеров нестационарных процессов.

1. «Переключение» частоты. Классический спектральный анализ, основанный на преобразовании Фурье, позволяет обнаружить сам факт существования различных ритмов, но не дает возможности проследить эволюцию во времени мгновенных характеристик ритмических процессов. Вейвлет-анализ обладает более широкими возможностями и является эффективным инструментом исследования локальных свойств сигналов даже в случае быстрых изменений мгновенных частот ритмических составляющих (сильной нестационарности). На рис. 8.12, *a* приведен тестовый пример сигнала, для которого вначале частота колебаний принимает одно значение, затем на некотором участке колебания полностью прекращаются, после чего они возникают уже на другой частоте. Вейвлет-анализ с применением базисной функции Морле позволяет с хорошей точностью идентифицировать моменты времени, в которые происходят изменения структуры сигнала. Правильно определить моменты «переключения» частоты удается и в случае сравнительно быстрых изменений. На рис. 8.12, *б* показан случай, когда частота дважды переключается в пределах одного периода. Возможность анализа колебаний с частотой порядка 0.01 Гц важна в исследованиях сложной динамики живых систем, в частности, различных физиологических процессов. Очень многие ритмы в биологии и медицине соответствуют диапазону $10^{-2} \div 10^{-1}$ Гц, и применение специальных ме-

рованных b и варьировании параметра a , то есть путем поиска пиков локальных спектров.

тодов, позволяющих проводить детальный анализ их структуры, играет важную роль в задачах диагностики состояния системы по экспериментальным данным.

2. «Чирпы». На практике более типичны ситуации, когда частота колебаний меняется не скачкообразно, а плавно. Примером может служить чирп-сигнал (от англ. «chirp» – чирикание) – процесс с частотой, монотонно меняющейся во времени. Рис. 8.13, *a* иллюстрирует случай сигнала, включающего в себя два параллельных линейных чирпа. После проведения вейвлет-преобразования можно идентифицировать мгновенные частоты обоих ритмических процессов, меняющиеся по линейному закону, и, таким образом, вейвлет-анализ дает правильную информацию о структуре исследуемого сигнала.

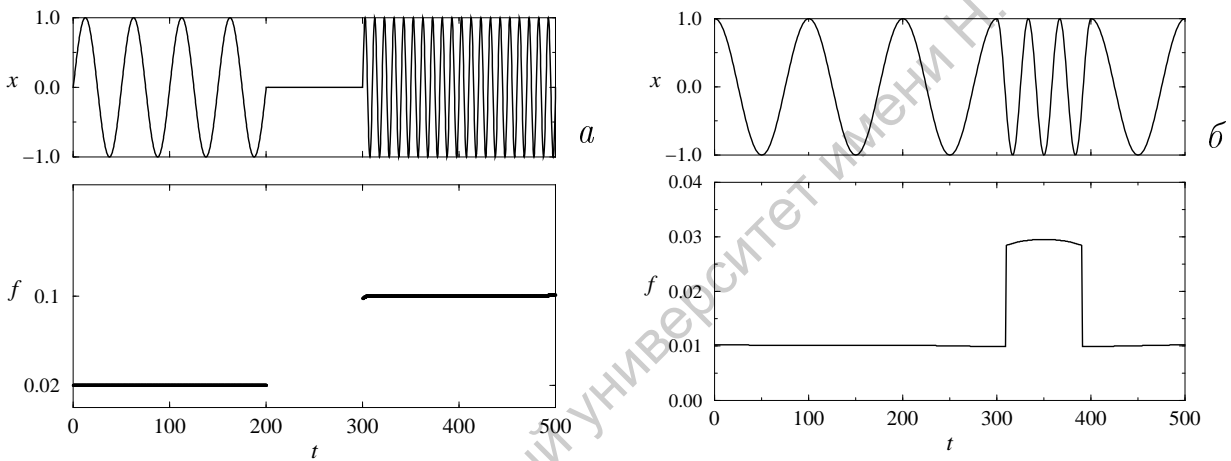


Рис. 8.12. Анализ сигналов, содержащих «переключения» частоты

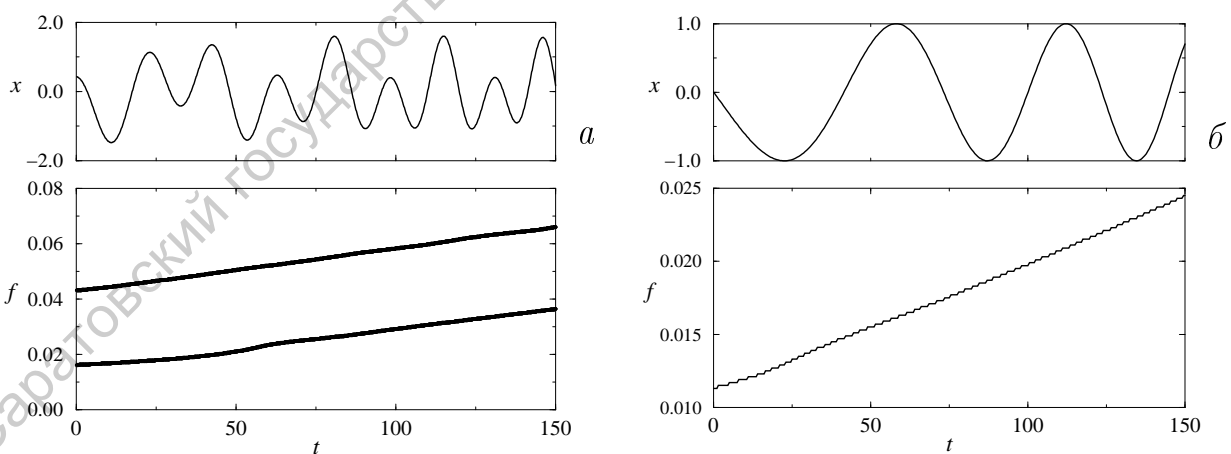


Рис. 8.13. Анализ «чирпов»

Отметим, что несмотря на то, что мы можем проводить локальный спектральный анализ, тем не менее все характеристики оцениваются в пределах некоторого временного окна, соответствующего функции $\psi_{a,b}(t)$ (или в пределах вышеупомянутого прямоугольника Гейзенберга). Это означает, что

характеристики вычисляются не совсем локально, а на основе некоторого усреднения по короткому участку сигнала. Эффект усреднения приводит к увеличению погрешности вычисления мгновенных характеристик с ростом скорости их изменения. Тем не менее, даже при быстрых изменениях свойств сигнала метод позволяет получать корректные результаты. Для иллюстрации рассмотрим случай линейного чирпа, частота которого меняется примерно в два раза за один период колебаний (рис. 8.13, б). Вейвлет-анализ позволяет идентифицировать в этом случае линейную зависимость мгновенной частоты с высокой точностью.

8.4 Вейвлет-анализ сердечного ритма

Вейвлет-анализ представляет несомненный интерес в задачах исследования динамики живых систем. Эти системы демонстрируют сложное нерегулярное поведение, характеристики которого непрерывно меняются во времени. Привлечение для анализа соответствующей динамики классических вероятностных и спектральных методов означает априорное предположение о том, что рассматриваемые процессы можно приближенно считать эргодическими, а справедливость этого допущения довольно сложно обосновать, если живой организм демонстрирует процесс адаптации к изменению внешних условий функционирования. Порой возникают проблемы с интерпретацией результатов анализа биологических данных. Как уже отмечалось (рис. 8.6), наличие двух пиков в спектре мощности с некрратными частотами может соответствовать принципиально разным ситуациям: в динамике изучаемой системы могут одновременно присутствовать два независимых ритма или может наблюдаться процесс изменения частоты, и в каждый момент времени удастся зафиксировать только один ритмический процесс. Такие ситуации нередко встречаются в функционировании объектов живой природы, они выявляют ограничения классических подходов к анализу случайных процессов и определяют важность разработки новых, более эффективных инструментов анализа структуры сигналов. Современные биологические исследования уже невозможны без самого широкого использования специальных физических методов. Развитие техники привело в настоящее время к высочайшему уровню экспериментальных исследований, когда сигналы биологических систем можно измерять на микроскопическом уровне отдельных клеток и внутриклеточной динамики. В то же время анализ таких сигналов зачастую ограничивается простой статистической обработкой экспериментальных данных. Создание более точных инструментов исследования сигналов, позволяющих выявить детали их сложной структуры, является в этой связи очень актуальной задачей: под высокоточные эксперименты, выполняемые в настоящее время в биологии, нужны соответствующие методы анализа. Биологические приложения

физических подходов и методов обогащают и саму физику. В частности, разработанные специальные методы, для которых нестационарность динамики не является препятствием, не только существенно расширяют возможности экспериментальных исследований, но и в значительной степени определяют дальнейший прогресс в развитии теории анализа структуры сигналов. Вейвлет-анализ как раз относится к числу тех инструментов, которые в ближайшие годы будут служить основой исследования сложной динамики объектов живой природы. Далее мы рассмотрим применение вейвлет-анализа при изучении структуры сердечного ритма.

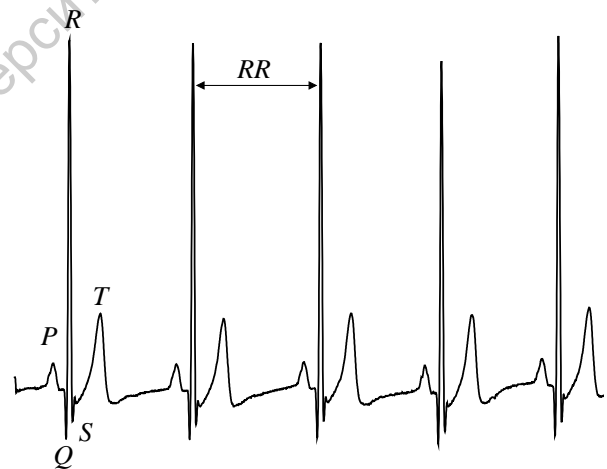
В настоящее время в экспериментальной и клинической кардиологии широко используются статистические методы анализа динамических рядов кардиоинтервалов. Задачи математического анализа сердечного ритма заключаются в том, чтобы на основании изучения активности синусного узла по вариациям длительности последовательности кардиоинтервалов сделать заключение о состоянии системы управления динамикой сердца и ее отдельных уровней. При оценке результатов математического анализа ритма сердца речь идет о характеристиках состояний регуляторных систем, обусловленных приспособительными (адаптационными) реакциями организма. Возможность судить о сложных явлениях, происходящих на разных регуляторных уровнях, путем обработки доступных физиологических сигналов, содержащих информацию о состоянии систем организма, имеет при этом принципиальное значение.

Измерение средней частоты пульса позволяет судить лишь о конечных результатах деятельности механизмов регуляции (о суммарном эффекте воздействия многих факторов). Сама же эта деятельность проявляется в «разбросе» значений кардиоинтервалов относительно среднего уровня (в вариабельности ритма). Поэтому только исследование вариабельности позволяет выяснить степень активности различных звеньев регуляторного механизма и составить представление о выраженности общей адаптационной реакции организма. Известно, что в ритме сердца отражена информация о процессах, протекающих не только в пределах сердечно-сосудистой системы, но и в различных звеньях системы регуляции: высших нервных центрах, железах внутренней секреции, различных рецепторах и т.д. Сложная иерархия структур системы управления и их взаимодействие приводит к тому, что на ритм сердца воздействуют многочисленные нервные и гуморальные факторы, которые обуславливают вариабельность сердечного ритма. Даже в условиях постоянного уровня физической активности человека сердечный ритм не является постоянным, и сердечно-сосудистая система может реализовать различные режимы динамики. Отметим, что сердечный ритм – это сложный, в общем случае непериодический процесс, в котором можно выделить несколько характерных частот: частоту сердечных сокращений (ЧСС), частоту дыхания (кардиологи предпочитают говорить о дыхательных волнах) и т.д. Поэтому, хотя мы пользуемся понятием сер-

дечный «ритм», речь, естественно, не идет о периодической динамике.

Типичный вид электрокардиограммы здорового человека представлен на рис. 8.14. Один сердечный цикл включает $PQRST$ -комплекс, отражающий процессы электрической активности, протекающие в предсердиях и желудочках при каждом сокращении сердца, и период паузы перед следующим сокращением. В медицине, однако, предпочитают работать не с сигналом электрокардиограммы, а с последовательностью RR -интервалов – расстояний между соседними R -пиками. Данное обстоятельство объясняется следующим образом. Во-первых, при создании больших банков данных в клиниках хранение электрокардиограмм требует резервирования значительного объема памяти компьютера (особенно при проведении Холтеровского мониторинга – суточных записей ЭКГ), что создает проблемы не только при хранении, но и при обработке сигналов. Во-вторых, считается, что именно RR -интервалы несут в себе основную информацию о функциональном состоянии организма, тогда как $PQRST$ -комплекс повторяется с минимальными искажениями при каждом сердечном сокращении. Важной характеристикой является вариабельность сердечных сокращений, т.е. изменение величины RR от удара к удару (рис. 8.14).

Рис. 8.14. Электрокардиограмма здорового человека



Известные методы анализа RR -интервалов включают:

- Временные методы, подразделяемые на статистические и геометрические. К первым относятся такие характеристики, как среднее, стандартные отклонения и т.д.; ко вторым – различные методы интерполяции либо аппроксимации кривой плотности распределения вероятности, рассчитанной по последовательности RR -интервалов.
- Спектральные методы. Для коротких реализаций (порядка 5 минут) осуществляется деление спектра мощности на следующие диапазоны:
 - 1) < 0.04 Гц – область очень низких частот (VLF)
 - 2) $0.04 \div 0.15$ Гц – область низких частот (LF)
 - 3) $0.15 \div 0.4$ Гц – область высоких частот (HF).

Для суточных ЭКГ выделяют еще область ультра-низких частот (ULF), < 0.003 Гц. В качестве диагностических критериев используется величина энергии в указанных частотных диапазонах, либо отношение энергий колебаний в диапазонах LF и HF .

При проведении спектрального анализа важным является вопрос о представлении последовательности RR -интервалов. Традиционным методом представления является зависимость величины каждого RR -интервала от номера сердечного сокращения i (рис. 8.15, *a*). Такая зависимость позволяет сразу же оценить диапазон изменения кардиоинтервалов, построить плотность распределения вероятности и вычислить моментные функции (среднее \overline{RR} и дисперсию σ_{RR}^2). Для определения частот ритмических процессов необходим переход к временной оси, для чего иногда просто домножают номер i на величину \overline{RR} . Заметим, что такой подход является не очень физическим. Чтобы сохранить информацию о времени каждого сердечного сокращения, целесообразно использовать другой вариант представления результатов, изображенный на рис. 8.15, *б*, где $RR_i = t(R_{i+1}) - t(R_i)$, $t_i = t(R_i)$. Для удобства дальнейшего анализа полученную последовательность значений можно интерполировать гладкой функцией (например, сплайном), что позволяет далее перейти к постоянному шагу по времени.

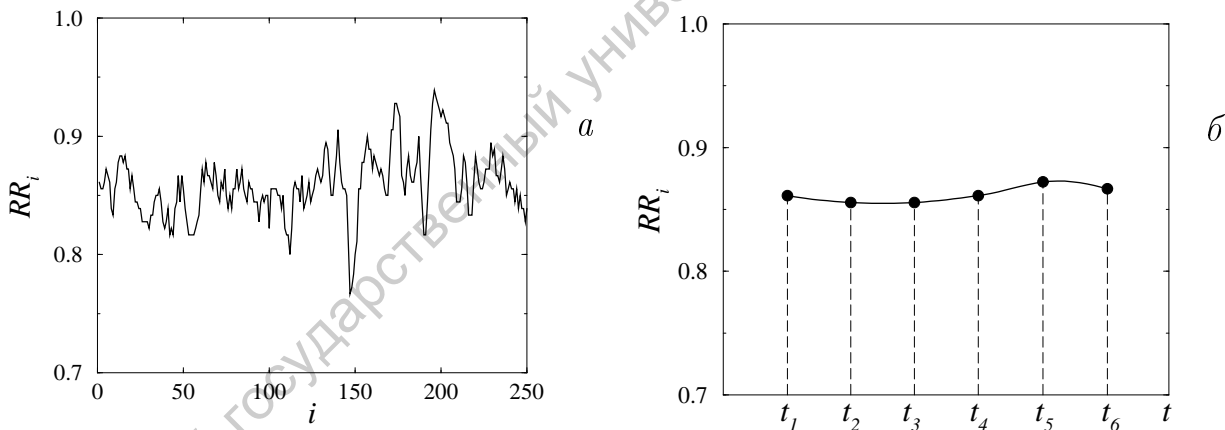


Рис. 8.15. Различные варианты представления последовательностей RR -интервалов

Полученный сигнал отражает вариабельность сердечного ритма, вызванную различными механизмами регуляции. Применение вейвлет-анализа позволяет исследовать динамику мгновенных частот ритмических процессов, приводящих к вариабельности последовательности RR -интервалов. По коротким последовательностям (300-500 отсчетов) обычно хорошо идентифицируется динамика на частоте ≈ 0.1 Гц, отражающая барорецепторную регуляцию, и на частоте ≈ 0.25 Гц (ритм дыхания) – рис. 8.16.

Мгновенные частоты ритмов могут демонстрировать как незначительные, так и сильные изменения, отражая степень нестационарности анализируемого сигнала. Чтобы количественно охарактеризовать их динамику,

можно воспользоваться мерой динамической стабильности ритмов

$$s = \frac{1}{\sigma_f},$$

где σ_f – стандартное отклонение мгновенной частоты от среднего значения. Чем больше величина s , тем более стабильным является исследуемый ритм.

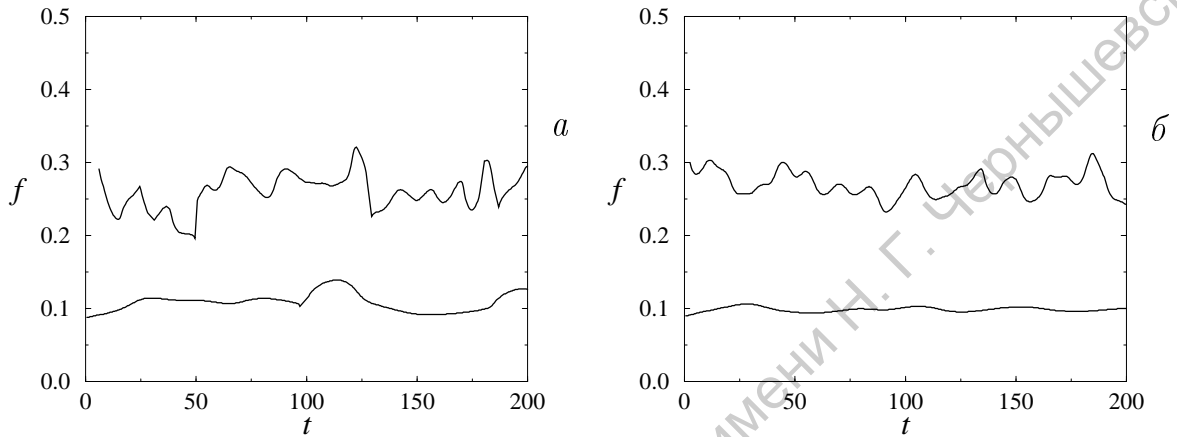


Рис. 8.16. Динамика мгновенных частот ритмов в LF и HF диапазонах для одного человека в разных функциональных состояниях: при слабой (a) и более сильной (b) физической нагрузке. Во втором случае наблюдаются более стабильные ритмы

Переход к рассмотрению мгновенной динамики соответствующих ритмов позволяет исследовать наличие синхронизации между различными процессами в организме. Например, хорошо известно явление синхронизации между процессом дыхания и частотой сердцебиений. Чтобы анализировать подстройку ритмов и взаимодействие между ними, можно построить распределение отношения мгновенных частот. Более узкие распределения характеризуют более высокую степень когерентности динамики рассматриваемых ритмических процессов. Примеры распределений приведены на рис. 8.17.

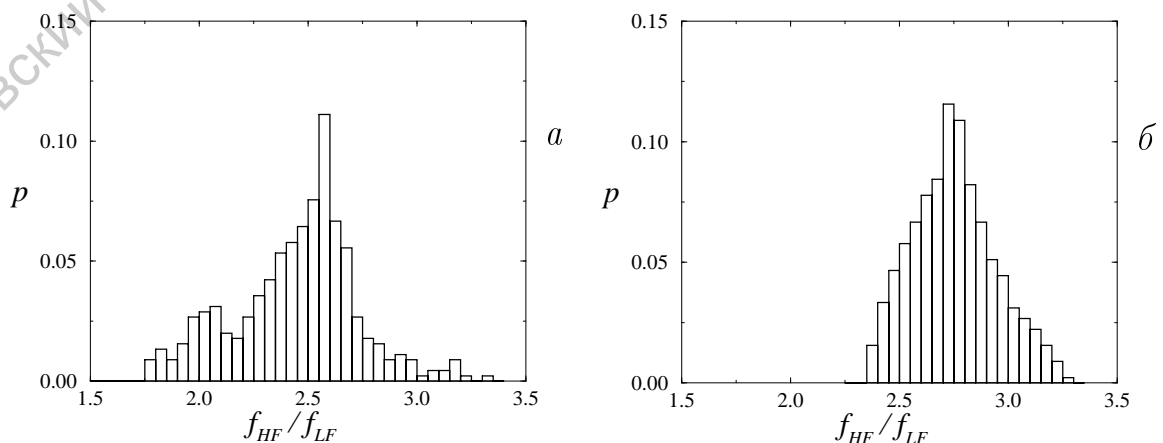


Рис. 8.17. Распределения отношения частот ритмов в HF и LF диапазонах для примеров, представленных на рис. 8.16

Порядок выполнения работы

Задание 1. Для предложенного нестационарного временного ряда провести расчет вейвлет-преобразования на основе базисной функции Морле (воспользоваться программой «morle.x»). Проанализировать частотно-временную динамику, т.е. выявить характерные частоты и проследить их эволюцию во времени. Комментарии к работе программы «morle.x» даны в файле «morle.in».

№ варианта	название файла
1	test1.dat
2	test2.dat
3	test3.dat
4	test4.dat
5	test5.dat
6	test6.dat

Задание 2. По аналогии с предыдущим заданием проанализировать частотно-временную динамику последовательностей RR-интервалов человека в двух состояниях (норма и стресс).

Задание 3. Провести анализ частотно-временной динамики другого процесса медико-биологического происхождения. Выявить характерные ритмы и проследить их эволюцию во времени. Оценить (приблизительно) динамическую стабильность наблюдаемых ритмов.

№ варианта	анализируемый сигнал
1	послед. межспайковых интервалов нейрона рыбы
2	сигнал нефрона (периодическая динамика)
3	сигнал нефрона (хаотическая динамика)
4	сигнал артериального давления крысы (норма)
5	сигнал артериального давления крысы (стресс)

Контрольные вопросы

1. Какими свойствами должна обладать функция, чтобы ее можно было выбрать в качестве вейвлета?
2. В чем состоят основные отличия Фурье- и вейвлет-преобразований?
3. Чем отличаются локальные и глобальные энергетические спектры?
4. В чем состоят особенности частотно-временного разрешения вейвлет-преобразования?
5. Объясните основные отличия вейвлет-преобразования от оконного преобразования Фурье.

6. Какие существуют варианты визуализации результатов вейвлет-преобразования?
7. Чем вызвана переменчивость сердечного ритма?
8. В чем состоят различные варианты представления последовательностей кардиоинтервалов?

9 Мультифрактальный анализ

9.1 Предварительные замечания

Теория фракталов и мультифракталов в настоящее время широко используется для описания свойств самоподобия и сложного скейлинга, наблюдаемых в самых разных физических ситуациях. К числу фракталов относят геометрические объекты – линии, поверхности, пространственные тела, которые имеют сильно изрезанную форму и демонстрируют некоторую повторяемость в широком диапазоне масштабов. Повторяемость может быть полной (в этом случае говорят о регулярных фракталах), либо может наблюдаться некоторый элемент случайности (такие фракталы называют случайными). Структура случайных фракталов на малых масштабах не является идентичной всему объекту, но их статистические характеристики совпадают, и свойства самоподобия справедливы после усреднения по статистически независимым реализациям объекта.

Для количественного описания фракталов достаточно одной величины – размерности Хаусдорфа или показателя скейлинга, описывающего сохраняемость геометрии или статистических характеристик при изменении масштаба. Однако в физике, химии, биологии и других науках встречаются многие явления, которые требуют расширения понятия фрактала на сложные структуры с более чем одним показателем скейлинга. Такие структуры часто характеризуются целым спектром показателей, и размерность Хаусдорфа представляет собой лишь один из них. Сложные фракталы, называемые *мультифракталами*, важны прежде всего потому, что именно они, как правило, и встречаются в природе, тогда как простые самоподобные объекты являются идеализацией реальных явлений. Фактически, мультифрактальный подход означает, что изучаемый объект каким-то образом можно разделить на части, для каждой из которых наблюдаются свои свойства самоподобия. Данный подход изначально был предложен для статистического анализа особенностей скейлинга сингулярных мер и с успехом применялся в разных областях науки – при изучении агрегационных свойств клеточных элементов крови в биологии и диффузионного роста кластеров, для характеристики разрушения материалов в физике металлов, в теории развитой гидродинамической турбулентности, при исследовании несоразмерных структур и квазикристаллов в физике твердого тела, для анализа структуры молекул ДНК, в задачах об одномер-

ных случайных блужданиях и при исследовании броуновского движения, для описания инвариантной вероятностной меры странных аттракторов и т.д. Самые разные объекты природы могут быть отнесены к специальному классу «мультифракталов», и, пожалуй, довольно сложно найти область науки, где бы мы не встретились с представителями этого класса.

Важными представителями мультифракталов служат многие сигналы, регистрируемые в натуральных экспериментах, и для целого ряда практических задач несомненную ценность представляет наличие строгого математического подхода к анализу сложной структуры процессов различной природы. Простые или *монофрактальные* сигналы (например, $1/f$ -шум, винеровский случайный процесс и т.д.) являются однородными в том смысле, что их скейлинговые характеристики остаются неизменными в любом диапазоне масштабов. Спектр таких сигналов имеет вид $S(f) \sim f^{-\beta}$ и не меняется в широком частотном диапазоне, то есть β представляет собой постоянную величину. Мультифрактальные процессы допускают разложение на участки с различными локальными свойствами скейлинга. Соответственно, для их количественного описания требуется большое число характеристик. В частности, спектр таких процессов не может быть описан степенным законом с единственным показателем β .

Известно несколько попыток расширить концепцию мультифракталов на случай функциональных зависимостей (сигналов). Одна из них основана на методе структурных функций, который достаточно часто использовался разными исследователями. Возможно наиболее широкое распространение данный подход приобрел в задачах изучения сильно развитой турбулентности. В начале 90-х годов в работах Мьюзи, Бакри и Арнеодо был предложен более совершенный метод «максимумов модулей вейвлет-преобразования» (ММВП), имеющий ряд существенных преимуществ: анализ широкого класса сингулярностей – не только самих сигналов, но и их производных, меньшая погрешность вычисления скейлинговых характеристик и т.д. Техника ММВП может успешно применяться в исследованиях структуры неоднородных процессов различной природы. Она базируется на вейвлет-анализе, который называют математическим «микроскопом» из-за способности сохранять хорошее разрешение на разных масштабах.

Метод ММВП основан на непрерывном вейвлет-преобразовании. Привлекательность данного метода состоит в том, что с его помощью можно анализировать как сингулярные меры, так и сингулярные функции; он является более универсальным аппаратом для исследования мультискейлинговых свойств объектов по сравнению с ранее разработанными подходами. Вейвлеты часто интерпретируют как обобщение классических алгоритмов покрытия множества сферами, кубиками и т.п. Поскольку базисные функции вейвлет-преобразования являются хорошо локализованными (солитоноподобными), они представляют собой эффективный математический аппарат для анализа нестационарных процессов.

В настоящее время метод максимумов модулей вейвлет-преобразования становится одним из популярных инструментов анализа нестационарных данных, который широко используется в прикладных исследованиях, в частности, для обработки медико-биологических процессов в случаях, когда нестационарность ограничивает применение классических методов анализа экспериментальных данных. За последние несколько лет наличие мультифрактальной структуры было обнаружено и численно охарактеризовано в динамике очень многих систем различной природы. Существует мнение, что метод ММВП является одним из наиболее эффективных в настоящее время инструментов статистического описания неоднородных процессов.

Отметим, что методов обработки нестационарных данных существует совсем немного. Если свойства процесса даже на коротких интервалах времени успевают существенно измениться, то использование классических алгоритмов анализа случайных процессов способно приводить к ошибкам в интерпретации полученных результатов. Арсенал универсальных инструментов, применимых независимо от свойства стационарности, включает концепцию аналитического сигнала, метод анализа флуктуаций относительно тренда, вейвлет-анализ. Мультифрактальный анализ может с полным основанием интерпретироваться как еще один инструмент, входящий в этот список. Несмотря на то, что он использует вейвлет-преобразование на промежуточных этапах вычислений, было бы неправильно воспринимать метод ММВП в качестве составной части вейвлет-анализа: он представляет собой сочетание двух различных теорий – теории вейвлетов и теории мультифракталов, имеющей глубокую аналогию со статистической термодинамикой. Такое сочетание двух различных теорий открывает новые возможности решения прикладных задач.

9.2 От сингулярных мер к сингулярным функциям

9.2.1 Фрактальные меры

Фракталы в природе могут встречаться не только в виде сложных геометрических объектов. Различные процессы на фракталах (физические, химические и т.д.) могут генерировать стационарные распределения, называемые *фрактальными мерами*. Примерами таких мер служат инвариантная вероятностная мера на странном аттракторе, распределение вероятностей роста для процессов ограниченной диффузией агрегации или распределение массы на некотором фрактальном множестве. Если в качестве количественной характеристики сложной геометрии фракталов используется размерность D_0 , то математическое описание фрактальных мер осуществляется в терминах *спектра сингулярностей* $f(\alpha)$. Смысл этой функции состоит в следующем. Предположим, что задано распределение меры μ на

некотором множестве, например, распределение заряда или массы. Если это множество покрывать шарами диаметра ε , то мера шара с центром в точке x_i зависит от ε по степенному закону вида:

$$\mu_{x_i}(\varepsilon) \sim \varepsilon^{\alpha(x_i)}, \quad (9.1)$$

где $\alpha(x_i)$ - называется *экспонентой сингулярности*. Зависимость (9.1) удобнее переписать в виде

$$\alpha(x_i) = \lim_{\varepsilon \rightarrow 0} \frac{\ln \mu_{x_i}(\varepsilon)}{\ln \varepsilon}. \quad (9.2)$$

Чем меньше $\alpha(x_i)$, тем более сингулярным является распределение меры в этой точке. Значение $\alpha(x_i) = 1$ будет приниматься при равномерном распределении. Предел $\alpha(x_i) = 0$ соответствует распределению меры, подобному функции Дирака, и означает, что заряд или масса сосредоточены вблизи одной точки. Спектр сингулярностей $f(\alpha)$ характеризует зависимость от ε числа элементов покрытия N_α , соответствующих точкам с экспонентой сингулярности, равной некоторому значению α :

$$N_\alpha(\varepsilon) \sim \varepsilon^{-f(\alpha)}. \quad (9.3)$$

По смыслу величина f соответствует размерности Хаусдорфа. В случае равномерного распределения меры на множестве $\alpha = \text{const}$, и спектр сингулярностей представляет собой единственную точку на плоскости (α, f) . При неравномерном распределении меры функция $f(\alpha)$ имеет более сложный («колоколообразный») вид.

Вышесказанное можно проиллюстрировать на примере канторова множества. Предположим, что на этом множестве задано равномерное распределение меры μ , например, массы, и для покрытия множества рассматриваются 2^n элементов (окружностей) диаметра $\varepsilon = 3^{-n}$. Мера каждого элемента покрытия будет равна $\mu_{x_i}(\varepsilon) = 2^{-n}$, где x_i - центр окружности. Согласно формуле (9.2), экспонента сингулярности $\alpha(x_i)$, определяемая наклоном зависимости $\ln \mu_{x_i}(\varepsilon)$ от $\ln \varepsilon$, принимает значение $\alpha(x_i) = \ln 2 / \ln 3$. В пределе $\varepsilon \rightarrow 0$ это значение экспоненты соответствует каждой точке канторова множества. В рассматриваемом примере экспонента α совпадает с размерностью Хаусдорфа d_H , а спектр сингулярностей $f(\alpha)$ состоит из одной точки $f(\alpha) = \alpha \approx 0.63$. Таким образом, мы наблюдаем только один «сорт» сингулярностей, и соответствующая мера μ называется *однородной*.

Если мера распределена неравномерно на множестве, спектр сингулярностей усложняется. Для иллюстрации проанализируем случай биномиального распределения: предположим, что отрезок $[0, 1]$ вновь делится на три равные части; средняя часть выбрасывается, но теперь мы приписываем разные весовые коэффициенты p_1 и $p_2 = (1 - p_1) \neq p_1$ двум оставшимся интервалам $[0, 1/3]$ и $[2/3, 1]$. Если вначале ($n = 0$) мы примем

для всего интервала $[0, 1]$ $\mu_0 = 1$, то на первом шаге процедуры построения канторова множества двум отрезкам соответствуют меры $\mu_1 = p_1\mu_0$ и $\mu_2 = (1 - p_1)\mu_0$. При последующих шагах будем использовать те же самые весовые коэффициенты p_1 и p_2 , осуществляя деление на части каждого из отрезков (рис. 9.1).

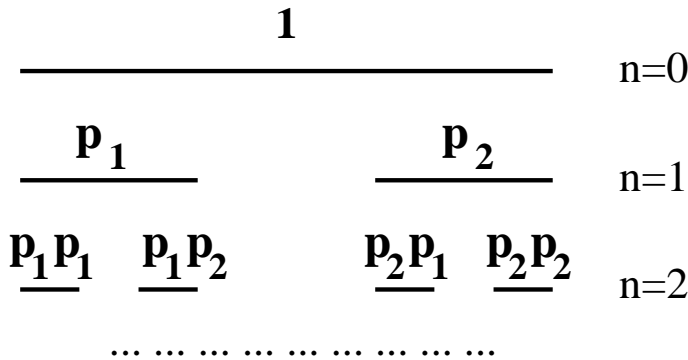


Рис. 9.1. Канторово множество с неравномерным распределением меры

Покрывая полученное фрактальное множество окружностями радиуса $\varepsilon = 3^{-n}$, рассмотрим крайний левый и крайний правый элементы покрытия. Для первого из них, содержащего точку $x_0 = 0$, мера равна $\mu_0 = p_1^n \mu_0 = p_1^n$. Поэтому, согласно (9.2), $\alpha(0) = \ln p_1 / \ln(1/3)$. Аналогично, для крайнего правого элемента покрытия, содержащего точку $x_0 = 1$, можно записать $\alpha(1) = \ln p_2 / \ln(1/3)$. Поскольку изначально $p_1 \neq p_2$, то и $\alpha(0) \neq \alpha(1)$. Соответственно, спектр сингулярностей $f(\alpha)$ уже не будет состоять только из одной точки. Типичный вид этой функции для случая неоднородной меры представлен на рис. 9.2.

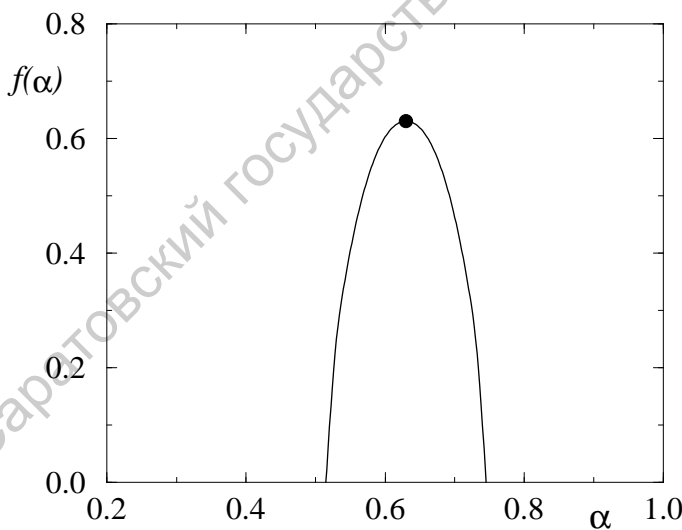


Рис. 9.2. Спектры сингулярностей для монофрактального объекта (канторова множество с равномерным распределением меры) – точка, и мультифрактального объекта (канторова множество с неравномерным распределением меры) – сплошная линия.

Наличие неравномерного распределения меры на множестве является свойством мультифрактальных объектов: чем более неоднородной является мера, тем шире спектр сингулярностей. Отметим, что максимум $f(\alpha)$

совпадает со спектром сингулярностей для случая равномерного распределения меры на канторовом множестве, характеризуя тем самым наиболее часто встречающуюся сингулярность.

На практике вычислить функцию $f(\alpha)$ на основе формулы (9.3) весьма проблематично из-за очень медленной сходимости при $\varepsilon \rightarrow 0$. Кроме того, значения оцениваемых характеристик могут заметно варьироваться для разных выбранных точек. Поэтому в теории мультифракталов предпочитают использовать специальный подход, основанный на расчете обобщенных фрактальных размерностей как глобальных характеристик, зная которые, можно вычислить спектр сингулярностей $f(\alpha)$. В рамках данного подхода вводятся в рассмотрение так называемые *частичные функции* (или *обобщенные статистические суммы*):

$$Z(q, \varepsilon) = \sum_{i=1}^{N(\varepsilon)} \mu_i^q(\varepsilon), \quad (9.4)$$

где $N(\varepsilon)$ - число элементов покрытия размера ε , μ_i - мера элемента покрытия с номером i , $q \in \mathbb{R}$. Зависимость функций Z от ε , как правило, носит степенной характер вида:

$$Z(q, \varepsilon) \sim \varepsilon^{(q-1)D_q}, \quad (9.5)$$

где величины D_q представляют собой характеристики, называемые *обобщенными фрактальными размерностями*. Коэффициент $(q - 1)$ включен в показатель степени для выполнения равенства $Z(1, \varepsilon) = 1$, означающего условие нормировки для меры μ . Обычно используется следующее обозначение

$$\tau(q) = (q - 1)D_q,$$

а величины $\tau(q)$ называются скейлинговыми экспонентами. Как видно из формул (9.4) и (9.5), при $q = 0$ мы приходим к ранее введенному определению фрактальной размерности или емкости.

Для простых фракталов, называемых также *монофракталами*, $D_q = \text{const}$. В общем случае мультифрактальных объектов значения D_q монотонно уменьшаются с ростом q (рис. 9.3, а). Это уменьшение можно рассматривать в качестве диагностического критерия того, что анализируемый объект является неоднородным фракталом. Как следствие, зависимость $\tau(q)$ представляет собой прямую линию для однородных фрактальных мер и нелинейную функцию для неоднородных (рис. 9.3, б).

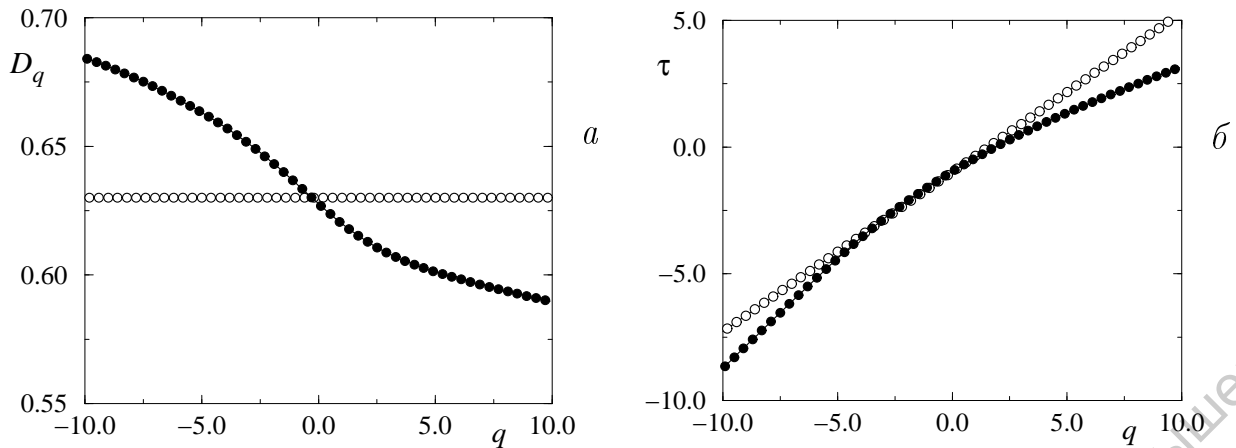


Рис. 9.3. Спектры обобщенных фрактальных размерностей (а) и скейлинговых экспонент (б) для монофрактала (белые кружочки) и мультифрактала (черные кружочки).

Метод расчета спектра сингулярностей на основе скейлинговых экспонент $\tau(q)$ или обобщенных фрактальных размерностей является более устойчивым и надежным чем непосредственно по определению (9.3). В рамках этого метода нахождение функций $f(\alpha)$ осуществляется с помощью преобразования Лежандра:

$$\begin{cases} \alpha = d\tau/dq, \\ f(\alpha) = q\alpha - \tau(q). \end{cases}$$

Таким образом, знание спектра скейлинговых экспонент $\tau(q)$ позволяет сразу же определить искомую функцию $f(\alpha)$. Мультифрактальный анализ часто называют *мультифрактальным формализмом*, подразумевая под этим термином подход, в рамках которого спектр сингулярностей $f(\alpha)$ рассматривается как преобразование Лежандра спектра $\tau(q)$. Отметим, что существует глубокая аналогия между мультифрактальным формализмом и статистической термодинамикой. Переменные q и $\tau(q)$ играют ту же роль, что и величина, обратная температуре, и свободная энергия в термодинамике, а преобразование Лежандра указывает на то, что вместо энергии и энтропии рассматриваются α и $f(\alpha)$. Ряд строгих математических результатов, относящихся к мультифрактальному формализму, был получен в рамках теории динамических систем. В последние годы данный подход приобретает большую популярность в различных экспериментальных исследованиях. Следует однако отметить, что сам термин мультифрактальный анализ (хотя он активно используется) является, возможно, не вполне корректным с той точки зрения, что получить «колоколообразную» структуру зависимости $f(\alpha)$, изображенную на рис. 9.2, можно и для объектов, не являющихся мультифракталами. Поэтому в общем случае должны быть дополнительно какие-то физические причины, чтобы с полным основанием можно было использовать терминологию мультифракталов. На ограниче-

ниях мультифрактального анализа мы еще остановимся позднее.

9.2.2 Фрактальные функции

Если бы теория мультифракталов ограничивалась только сингулярными мерами, ее практическая ценность была бы не столь велика. Однако фрактальные объекты встречаются в природе еще и в виде сингулярных функций, что позволяет рассматривать аппарат мультифрактального анализа в качестве инструмента исследования сигналов сложной структуры, необходимо было только адаптировать этот инструмент на случай нерегулярных процессов.

Возможно, наиболее интенсивные исследования случайных фрактальных функций были связаны с изучением процессов диффузии или броуновского движения. В одномерном случае смещение броуновской частицы $x(t)$ вдоль прямой линии представляет собой случайный процесс с нулевым средним и дисперсией

$$\langle x^2(t) \rangle \sim t, \quad (9.6)$$

где скобки обозначают процедуру усреднения по ансамблю реализаций. Линейная зависимость дисперсии от времени (9.6) характеризует случай нормальной диффузии. Функция $x(t)$ может интерпретироваться как «фрактальная» в том смысле, что ее график на плоскости «время – координата» представляет собой фрактальное множество точек с размерностью $D_0 = 1.5$ – рис. 9.4.

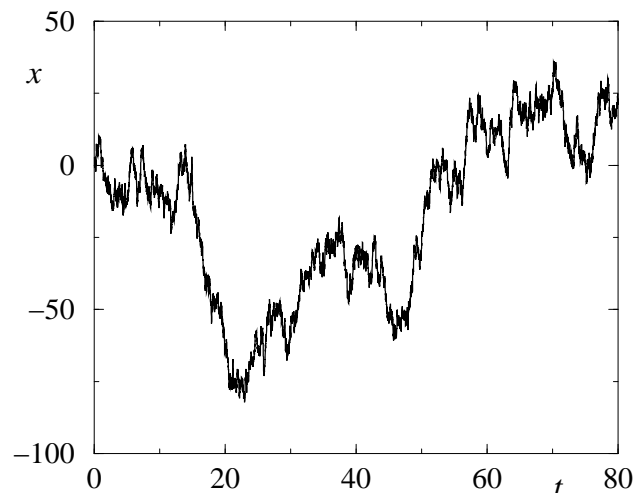


Рис. 9.4. Броуновское движение (случай нормальной диффузии)

Обобщением концепции обычного броуновского движения является фрактальное броуновское движение, при котором смещение частицы $x(t)$ в одномерном пространстве является случайным процессом с дисперсией

$$\langle x^2(t) \rangle \sim t^{2H},$$

где $0 < H < 1$. При $H \neq 0.5$ такая дисперсия соответствует аномальной диффузии с коррелированными приращениями. Корреляции приращений оказывают влияние на график зависимости $x(t)$, делая его более изрезанным ($H < 0.5$) или более гладким ($H > 0.5$). Соответствующие изменения влияют на величину размерности, которая связана с показателем степени H следующим образом:

$$D_0 = 2 - H.$$

Несмотря на то, что при изучении фрактальных функций может проводиться расчет размерности графика, этот классический подход дает мало информации по сравнению с мультифрактальным анализом, позволяющим получать непрерывный спектр фрактальных размерностей и статистические сведения о наличии различных сингулярностей. Вычисление одной единственной величины D_0 позволяет охарактеризовать функцию при условии, что она демонстрирует только один «сорт» сингулярного поведения, которое является неизменным во времени (случай однородной фрактальной функции). Такая ситуация, конечно же, представляет собой идеализацию реальных процессов, наблюдаемых в природе.

Броуновское движение относится к числу случайных фракталов: увеличенный фрагмент функции $x(t)$ не является идентичным всему графику. В теории фракталов часто используется понятие само-афинности для характеристики множеств, инвариантных при аффинных преобразованиях. В частности, на плоскости это означает, что перемасштабирование отдельного фрагмента графика функции с различными коэффициентами увеличения по абсциссе и по ординате будет приводить к графику, похожему на первоначальный. Самоподобие, о котором упоминалось применительно к фрактальным объектам, строго говоря, является частным случаем само-афинности и означает наличие сходства отдельного фрагмента всему объекту при перемасштабировании с одинаковыми коэффициентами увеличения по обеим осям. Для само-афинной функции можно записать следующее выражение:

$$g(x_0 + lx) - g(x_0) \simeq l^H (g(x_0 + x) - g(x_0)).$$

При рассмотрении случайных процессов, таких как броуновское движение, равенство может выполняться только при фиксированных значениях l и x_0 . Показатель степени H в последней формуле называется *экспонентой Хёрста* и характеризует, насколько нерегулярна функция g в точке x_0 . При $H < 1$ эта функция является недифференцируемой и, по аналогии с ранее введенной экспонентой сингулярности α , чем меньше H , тем более сингулярна $g(x)$. В рамках мультифрактального анализа само-афинные функции рассматриваются не с точки зрения сложной геометрии их графика на плоскости «время – координата», а как функции с локальным нерегулярным поведением. Учитывая то, что сингулярности функции могут

различаться в разных точках, их описание в общем случае не может осуществляться с помощью одного параметра H . В связи с этим используют понятие локальных экспонент Хёрста или *экспонент Гельдера* $h(x_0)$, которые вводятся в рассмотрение путем некоторого изменения определения H , а именно:

$$|g(x_0 + l) - g(x_0)| \sim l^{h(x_0)}$$

и характеризуют локальное сингулярное поведение функции g в точке x_0 . Мультифрактальный подход для сигналов потенциально способен характеризовать широкий класс процессов, являющихся более сложными по сравнению с процессами, для описания которых достаточно одного числа (единственного значения фрактальной размерности либо одной скейлинговой характеристики, описывающей, например, частотную зависимость спектральной плотности мощности).

Анализ нерегулярных функций, также как и анализ фрактальных мер, осуществляется в терминах спектра сингулярностей. Однако, при рассмотрении функциональных зависимостей (нерегулярных сигналов) проводится смена используемых обозначений. Вместо спектра $f(\alpha)$ рассматривается аналогичная функция $D(h)$: экспонента Гельдера по смыслу соответствует α , а $D(h_0)$ представляет собой размерность подмножества анализируемых данных, которое характеризуется локальной экспонентой h_0 .

Существует несколько способов вычисления спектра сингулярностей. Также как и в случае фрактальных мер проведение вычислений непосредственно по определению осложняется медленной сходимостью и тем обстоятельством, что при рассмотрении локальной окрестности произвольной точки с нерегулярным поведением x_0 в эту окрестность могут попадать другие сингулярности. Эффект «наложения» нескольких сингулярностей будет приводить к значительным погрешностям расчета скейлинговых характеристик и нестабильности результатов при вариации параметров алгоритма в ходе проведения численного анализа. Для повышения надежности расчетов $D(h)$ проводится статистический анализ на основе структурных или частичных функций. Последний метод является предпочтительным, так как он позволяет проводить исследования более широкого класса сингулярностей. В следующем разделе изложены основы теории мультифрактального анализа, в рамках которого вычисление частичных функций базируется на вейвлет-преобразовании.

9.3 Анализ на основе вейвлет-преобразования

9.3.1 Вейвлет-анализ сингулярных функций

В начале 90-х годов Мьюзи, Бакри и Арнеодо разработали новый подход к исследованию мультифрактальных свойств сигналов сложной структуры

– метод «*максимумов модулей вейвлет-преобразования*». Как следует из названия метода, он базируется на вейвлет-преобразовании, состоящем в разложении сигнала по базису, сконструированному из обладающей определенными свойствами солитоноподобной функции (вейвлета) посредством масштабных изменений и переносов. Каждая из функций этого базиса характеризует как определенную пространственную или временную частоту, так и ее локализацию в физическом пространстве или времени.

Непрерывное вейвлет-преобразование функции $g(x)$ определяется следующей формулой:

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} g(x) \psi \left(\frac{x - b}{a} \right) dx, \quad (9.7)$$

где a представляет собой параметр масштаба, b – пространственная координата или момент времени, ψ – солитоноподобная функция (вейвлет), обладающая определенными свойствами и сконструированная, например, на основе производных функции Гаусса:

$$\psi^{(m)} = (-1)^m \frac{\partial^m}{\partial x^m} \left[\exp\left(-\frac{x^2}{2}\right) \right]. \quad (9.8)$$

При изучении локальных сингулярностей функции $g(x)$ обычно рассматривают $m = 1$ («WAVE»-вейвлет) или $m = 2$ («МНАТ»-вейвлет). Более высокие производные применяются редко. Выбор базисной функции ψ является очень важным моментом при проведении исследований – нужно выбрать такой вейвлет, который позволяет разглядеть нужную нам информацию. Сделать это можно при помощи разных функций, и продолжая аналогию вейвлет-анализа с математическим микроскопом, использование разных базисных функций можно интерпретировать как выбор разрешения объектива для микроскопа – если выбранное разрешение позволяет увидеть нужные детали, то вейвлет подходит для решаемых задач. Более того, дальше уже выбор объектива с лучшим разрешением не дает ничего нового.

Как отмечалось в разделе 9.2.2, степень сингулярности функции g может быть описана с помощью экспоненты Гельдера. Ее более строгое математическое определение можно ввести следующим образом. Рассмотрим случай, когда функция g в точке x_0 является n раз дифференцируемой, а $(n + 1)$ -я производная не существует. Тогда функцию $g(x)$ можно разложить в ряд Тейлора до степени n включительно. Обозначим полученный ряд (полином степени n) символом $P_n(x)$. Экспонентой Гельдера в точке x_0 называется наибольшее значение h , при котором выполняется следующее неравенство:

$$|g(x) - P_n(x - x_0)| \leq C |x - x_0|^h.$$

Чем больше h , тем более регулярной (более «гладкой») является функция g . Интегрирование этой функции увеличивает значение h на 1, а дифференцирование – уменьшает. Преположим, что анализируемую сингулярную функцию в окрестности точки $x = x_0$ можно представить в виде двух слагаемых:

$$g(x) = P_n(x) + C |x - x_0|^{h(x_0)},$$

то есть в виде суммы регулярной компоненты (полинома P_n) и слагаемого, которое определяет нерегулярное поведение и характеризуется нецелым значением $h(x_0)$. Одной из особенностей вейвлет-преобразования, широко используемой для устранения из сигнала полиномиальных составляющих, является то, что вейвлеты, имеющие первые m нулевых моментов

$$\int_{-\infty}^{\infty} x^m \psi(x) dx = 0,$$

будут ортогональны полиномам до степени m включительно, и при $m \geq n$:

$$\int_{-\infty}^{\infty} P_n(x) \psi(x) dx = 0. \quad (9.9)$$

Поэтому вейвлет-преобразование функции $g(x)$ примет вид:

$$W(a, x_0) = Ca^{1/2} \int_{-\infty}^{\infty} \psi(x) |ax|^{h(x_0)} dx.$$

В целях упрощения анализа, предлагается немного изменить определение вейвлет-преобразования, умножив выражение (9.7) на величину $1/\sqrt{a}$:

$$W(a, x_0) = \frac{1}{a} \int_{-\infty}^{\infty} \psi\left(\frac{x-b}{a}\right) g(x) dx. \quad (9.10)$$

В этом случае при $a \rightarrow 0$ наблюдается простая степенная зависимость:

$$W(a, x_0) \sim a^{h(x_0)}. \quad (9.11)$$

Таким образом, если функция $g(x)$ имеет особенность в точке $x = x_0$, то ее локальное сингулярное поведение характеризуется степенным законом вида (9.11). Если же эта функция является непрерывно дифференцируемой в точке x_0 вплоть до производной порядка m , то при $a \rightarrow 0$

$$W(a, x_0) \leq a^m.$$

В задачах исследования структуры сигналов обычно анализируют особенности самой функции $g(x)$, а ее производные рассматривают редко. В этом случае значения гельдеровских экспонент $h(x_0) < 1$, и наличие нерегулярности в точке x_0 будет приводить к разной скорости уменьшения вейвлет-коэффициентов с масштабом a . Чем быстрее коэффициенты уменьшаются при $a \rightarrow 0$, тем более регулярна функция в этой точке. Это приводит к тому, что значения $W(a, x)$ будут расходиться в окрестности точки с локальной особенностью. Такое поведение вейвлет-коэффициентов позволяет проводить детальный анализ сингулярностей.

9.3.2 Метод максимумов модулей вейвлет-преобразования

Алгоритм ММВП предполагает исследование нерегулярного поведения функции $g(x)$ в два этапа. На первом этапе осуществляется вейвлет-преобразование по формуле (9.10). Результат вейвлет-преобразования можно интерпретировать как поверхность в 3-мерном пространстве. Наиболее важная информация о ней содержится в «скелетоне» или линиях локальных экстремумов поверхности коэффициентов $W(a, x)$, поиск которых проводится на каждом масштабе a – рис. 9.5.

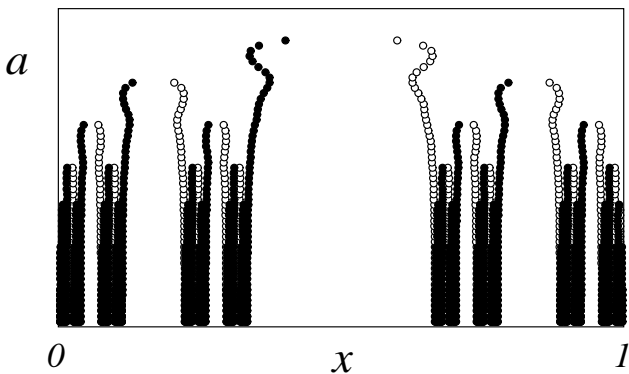


Рис. 9.5. Скелетон или линии локальных экстремумов поверхности коэффициентов $W(a, x)$: локальных минимумов (белые кружочки) и максимумов (черные)

Выбор базисной функции определяется тем, какую информацию необходимо извлечь из $g(x)$. Необходимым условием является то, что выбранный вейвлет должен быть не менее гладким чем сам анализируемый сигнал. При выборе параметра m в (9.8) следует отметить, что, с одной стороны, увеличение m позволяет проигнорировать крупномасштабные полиномиальные составляющие (устранить тренд) и анализировать мелкомасштабные вариации данной функции. С другой стороны, многократное дифференцирование приводит к увеличению числа линий локальных экстремумов вейвлет-коэффициентов и появлению большого числа дополнительных линий, обрывающихся на малых масштабах. Такие линии являются короткими для того, чтобы проводить по ним оценки степенных зависимостей вида (9.11) и служат помехами при проведении численного анализа сингулярностей. Вейвлет-преобразование устроено таким образом, что $W(a, x_0)$ является регулярной функцией даже при нерегулярной $g(x)$. Вся информация о возможной особенности $g(x)$, включая ее локализацию x_0 и показатель $h(x_0)$, заключена в асимптотическом поведении коэффициентов $W(a, x_0)$ при малых a . Если коэффициенты на малых масштабах расходятся, то g имеет особенность в x_0 , и экспонента Гельдера может быть определена путем представления зависимости (9.11) в двойном логарифмическом масштабе и вычисления наклона $\ln W(a, x_0)$ от $\ln a$. Если коэффициенты $W(a, x_0)$ близки к нулю в окрестности x_0 на малых масштабах, то g является регулярной в этой точке. Важное обстоятельство при расчете гильдеровских экспонент состоит в том, что искомые характеристики теоретически

не зависят от выбора базисных функций вейвлет-преобразования, что позволяет говорить о том, что анализ локальных нерегулярностей в некотором смысле универсален (хотя разложение по вейвлетам от выбранного базиса, конечно же, зависит).

Выделением «скелетона» заканчивается первый шаг алгоритма ММВП. Теоретически, анализ выделенных линий локальных экстремумов или локальных максимумов модулей вейвлет-преобразования позволяет вычислять гильдеровские экспоненты и анализировать сингулярности функции $g(x)$. Однако такой подход является неточным – при увеличении масштаба сказывается влияние соседних нерегулярностей, что приводит к различным ошибкам. В теории мультифракталов предпочитают проводить расчеты на основе статистических функций, позволяющих получать более надежные оценки вычисляемых характеристик. Поэтому второй шаг метода ММВП состоит в построении функций $Z(q, a)$, называемых статистическими функциями, по формуле:

$$Z(q, a) = \sum_{l \in L(a)} |W(a, x_l(a))|^q, \quad (9.12)$$

где $L(a)$ - множество всех линий (l) локальных максимумов модулей вейвлет-коэффициентов, существующих на масштабе a ; $x_l(a)$ характеризует расположение на этом масштабе максимума, относящегося к линии l . В данном случае принципиальным моментом, обеспечивающим устойчивость метода, является то, что рассматриваются модули вейвлет-коэффициентов. Без этого (с учетом информации о фазах) метод не позволил бы получить устойчивое решение. Заметим, что в общем случае рассмотрение максимумов может создавать проблемы с точки зрения устойчивости метода (лучше, когда оперируют средними величинами). Тем не менее, сама процедура вейвлет-преобразования подразумевает вычисление коэффициентов в пределах частотно-временного окна, что автоматически обеспечивает проведение усреднения. Определение (9.12) не подходит для отрицательных значений q , поскольку возможна ситуация, когда $W(a, x_l(a)) = 0$. В связи с этим на практике используется другая формула:

$$Z(q, a) = \sum_{l \in L(a)} (\sup_{a' \leq a} |W(a', x_l(a'))|)^q, \quad (9.13)$$

то есть выбирается максимальное значение модуля коэффициентов вейвлет-преобразования вдоль каждой линии на масштабах, меньших заданного значения a . Обычно выполняется следующая зависимость:

$$Z(q, a) \sim a^{\tau(q)},$$

где величину $\tau(q)$, определяемую для некоторого значения q путем вычисления наклона $\ln Z(q, a)$ от $\ln a$, называют *скейлинговой экспонентой* по

аналогии со случаем сингулярных мер. Вариация степеней q при построении частичных функций (9.13) позволяет получить линейную зависимость $\tau(q)$ для монофрактальных объектов ($H = d\tau/dq = \text{const}$) и нелинейную зависимость $\tau(q) = qh - D(h)$ с большим числом гельдеровских экспонент $h(q) = d\tau/dq \neq \text{const}$ в случае мультифракталов.

При некоторых значениях q скейлинговые экспоненты $\tau(q)$ имеют простую интерпретацию. Так, существует взаимосвязь между $\tau(2)$ и показателем степени β функции спектральной плотности мощности $S(f) \sim 1/f^\beta$:

$$\beta = 2 + \tau(2).$$

В свою очередь, учитывая то обстоятельство, что спектральная плотность мощности связана с корреляционной функцией преобразованием Фурье, знание β позволяет определить скорость спада корреляций $\Psi(\tau) \sim \tau^{-\gamma}$, то есть показатель γ . Взаимосвязь между основными величинами, рассматриваемыми в рамках алгоритма ММВП, определяется преобразованием Лежандра:

$$\begin{cases} h = d\tau/dq, \\ D(h) = qh - \tau(q). \end{cases}$$

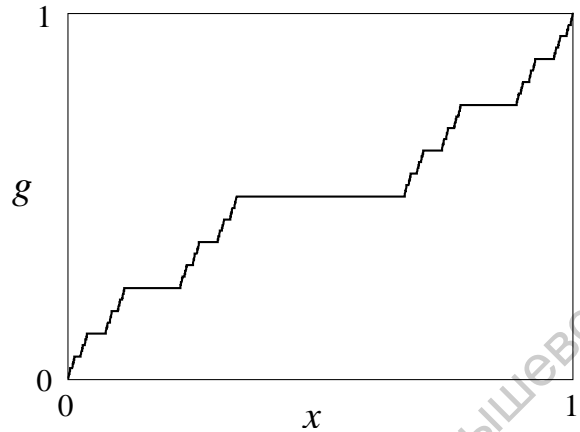
В отличие от метода структурных функций, основанный на вейвлет-преобразовании мультифрактальный анализ позволяет исследовать сингулярности при отрицательных значениях q . Статистические функции $Z(q, a)$ при $q < 0$ характеризуют особенности скейлинга для слабых сингулярностей (малых флуктуаций), а при $q > 0$ – для сильных сингулярностей (больших флуктуаций).

Проиллюстрируем алгоритм ММВП на примере канторова множества. Для проведения численного анализа канторову множеству может быть поставлена в соответствие бинарная последовательность, состоящая из нулей и единиц (нули обозначают выбрасываемые при построении ряда части отрезка). Таким образом, данная последовательность вначале ($n = 0$) состоит из одного элемента «1», на первом этапе из трех: «101», на втором – из девяти: «101000101» и т.д. Рассмотрим в качестве фрактальной функции $g(x)$ функцию распределения однородной меры μ на канторовом множестве. Нормируя меру ($\mu([0, 1]) = 1$), определим $g(x)$ следующим образом:

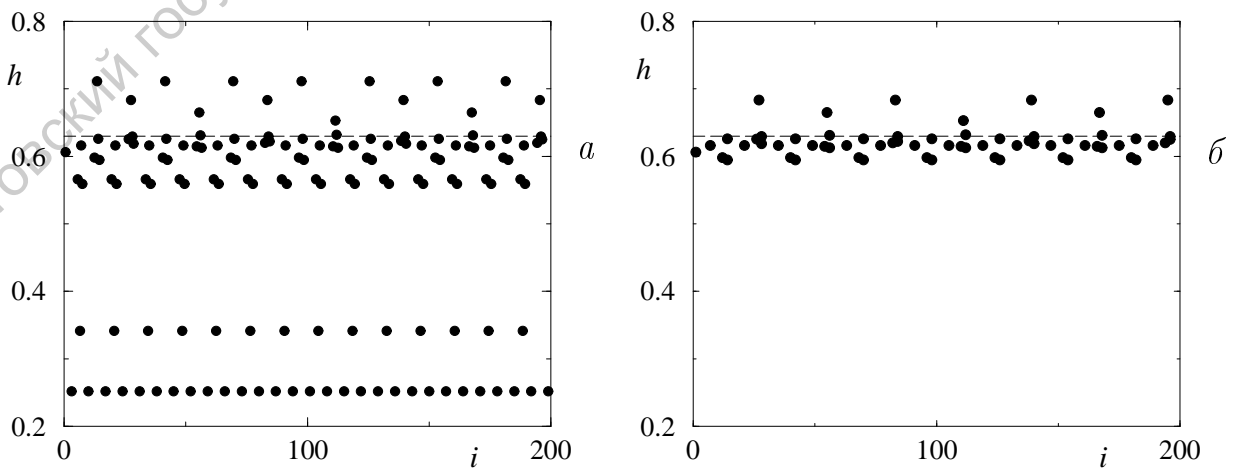
$$g(x) = \mu([0, x]) = \int_0^x d\mu.$$

Полученная при этом функция изображена на рис. 9.6.

Рис. 9.6. Функция распределения однородной меры на канторовом множестве



Выбрав в качестве базисного вейвлета $\psi^{(m)}$ «МНАТ»-вейвлет и следуя алгоритму, изложенному в предыдущем разделе, на первом этапе проведем вейвлет-преобразование функции $g(x)$ и выделение линий локальных максимумов модулей вейвлет-коэффициентов. Как видно из рис. 9.5, число локальных максимумов уменьшается при увеличении масштаба a . В соответствии с (9.11), локальное сингулярное поведение $g(x)$ может анализироваться путем построения зависимости $\ln W(a, x_0)$ от $\ln a$ и вычисления наклона аппроксимирующей прямой. Данные расчеты, проведенные для различных линий локальных максимумов, представлены на рис. 9.7, *а*. Наряду с ожидаемой величиной $h = \ln 2 / \ln 3$, для коротких линий получены значения гельдеровских экспонент, варьирующиеся в широких пределах. Картина меняется, если исключить из рассмотрения линии, «обрывающиеся» на малых масштабах и рассматривать степенные зависимости вейвлет-коэффициентов, например, при $\ln a > 1.5$ (рис. 9.7, *б*). Чтобы избежать проблем, связанных с недостаточно хорошим разрешением на малых масштабах, изучение скейлинговых особенностей проводится при не слишком малых масштабных коэффициентах a .

Рис. 9.7. Значения экспонент Гельдера, вычисленных по наклону степенной зависимости вида (9.12) для всех линий локальных экстремумов (*а*) и без учета линий, «обрывающихся» на малых масштабах ($\ln a < 1.5$) — (*б*)

Непосредственные оценки гельдеровских экспонент h из степенных зависимостей для вейвлет-коэффициентов менее точны, чем определение спектра $\tau(q)$ статистических функций и вычисление локальных наклонов $h = d\tau/dq$ (второй этап метода ММВП). При большом числе шагов n процедуры построения канторова множества численно найденные значения h и спектра сингулярностей $D(h)$ практически не отличаются от теоретических. Например, в расчетах для $n = 10$ отклонение от теоретического значения $\ln 2 / \ln 3$ не превышает 1%.

Уже отмечалась одна особенность метода анализа локальной регулярности сигнала на основе коэффициентов вейвлет-преобразования, которая состоит в том, что вычисляемые экспоненты Гельдера теоретически не зависят от выбора базисного вейвлета. На практике, однако, такая зависимость все-таки появляется. В большинстве известных работ в качестве базисных выбирались функции «WAVE» или «МНАТ» (1-я и 2-я производные функции Гаусса, соответственно). Если в качестве анализирующего вейвлета используются производные высокого порядка, то число дополнительных коротких линий локальных максимумов $W(a, x_0)$ возрастает, что обуславливается осциллирующими «хвостами» солитоноподобных функций $\psi^{(m)}$. В данной ситуации приходится искать компромисс между игнорированием полиномиальных составляющих (тренда), для чего желательно увеличивать параметр m , и значительным ростом при этом числа коротких линий локальных максимумов, которые усложняют вид скелетона и приводят к отклонениям от степенных зависимостей на малых масштабах. При проведении мультифрактального анализа на практике целесообразно ограничиваться малым значением m (обычно $m \leq 2$).

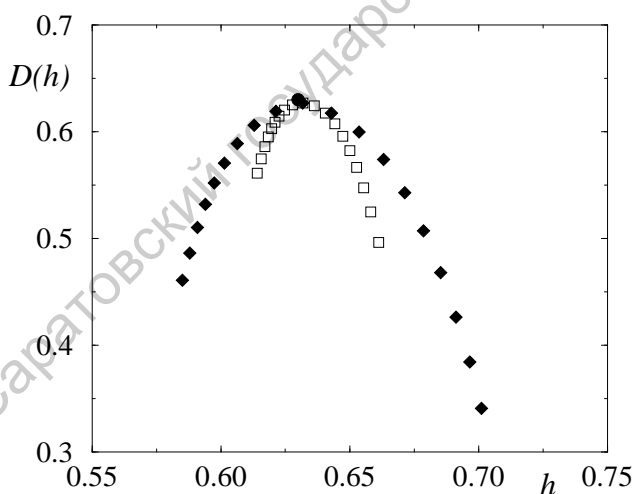


Рис. 9.8. Спектры сингулярностей для канторова множества со случайным распределением меры. Увеличение ширины спектра $D(h)$ соответствует увеличению неравномерности распределения меры

Мы рассмотрели случай равномерного распределения меры μ на фрактальном множестве. В численном анализе можно реализовать более сложный пример - канторово множество со случайным распределением меры.

При построении такого множества весовые коэффициенты p_1 и $p_2 = (1 - p_1)$ на каждом шаге выбираются равными $0.5 \pm \xi$, где ξ – случайная величина. При малой дисперсии случайной величины созданное множество будет иметь структуру, близкую к монофрактальной. При увеличении дисперсии динамика становится мультифрактальной: точка на плоскости $D(h)$ преобразуется в «колоколообразную» зависимость, типичную для мультифракталов (рис. 9.8). Таким образом, при случайном распределении меры на канторовом множестве наиболее вероятная сингулярность соответствует значению $\ln 2 / \ln 3$, а степень мультифрактальности или ширина спектра $D(h)$ зависит от дисперсии случайной величины ξ , то есть от неравномерности распределения меры.

Мультифрактальный анализ может проводиться применительно к сигналам различной природы. Помимо нерегулярных реализаций непрерывнозначных случайных процессов на основе техники ММВП могут анализироваться, например, символьные последовательности, такие как цепочки нуклеотидов молекул ДНК. С этой целью их преобразуют в бинарные последовательности $u(i) = 1$ или $u(i) = 0$ в зависимости от того, какой нуклеотид встречается под номером i , после чего по аналогии с задачами об одномерных случайных блужданиях рассматривается сумма:

$$g(k) = \sum_{i=1}^k u(i). \quad (9.14)$$

Проводя мультифрактальный анализ зависимости $g(k)$ методом ММВП, можно говорить об отсутствии корреляций в последовательности символов, если $h = 0.5$, и о наличии корреляций при $h \neq 0.5$. Иногда говорят также о корреляциях ($h > 0.5$) и анти-корреляциях ($h < 0.5$). Анतिकоррелированная динамика бинарной последовательности означает, что вероятность чередования нулей и единиц (пар «01» и «10») выше вероятности пар «00» или «11». В случае коррелированной динамики за нулями чаще следуют нули, за единицами – единицы, то есть вероятность обнаружения пар «00» и «11» увеличивается. Для реализаций непрерывнозначных случайных процессов значения гельдеровских экспонент также позволяют делать вывод о наличии анти-коррелированной ($h < 0.5$) и коррелированной динамики ($h > 0.5$). В первом случае наблюдается чередование больших и малых значений случайного процесса (вслед за большим значением с большей вероятностью следует малое и наоборот). Во втором случае за большим значением чаще следует большое, за малым – малое: процесс является более «гладким». Частный случай $h = 1$ соответствует $1/f$ -шуму, $h = 1.5$ – винеровскому случайному процессу.

Заметим, что процедура анализа мультискейлинговых явлений для последовательностей вида (9.14) приводит к тем же самым результатам, что и рассмотрение в качестве функции g исходной бинарной последователь-

ности $g(i) = u(i)$, если в последнем случае исключить множитель $1/a$ перед знаком интеграла в (9.10). Это объясняется тем, что интегрирование увеличивает значения h на 1. Если учесть это обстоятельство в формуле (9.11), можно проводить расчеты без вычисления сумм (9.14), аналогичных по смыслу взятию интеграла с переменным верхним пределом. Такое свойство метода полезно при применении мультифрактального анализа на практике.

9.4 Возможности и ограничения мультифрактального анализа

Основанный на вейвлет-преобразовании мультифрактальный анализ может рассматриваться в качестве метода исследования спектрально-корреляционных свойств различных процессов, в том числе нестационарных. Как следует, в частности, из формулы (9.9), медленная нестационарность (полиномиальный тренд) не оказывает влияния на результат, если выбирается базисная функция ψ , имеющая несколько первых нулевых моментов. Помимо проблемы нестационарности метод ММВП может демонстрировать еще одно преимущество по сравнению с классическим спектральным анализом. Являясь инструментом «локального» исследования структуры функций времени, мультифрактальный анализ позволяет проводить оценку корреляционных свойств случайных процессов по сравнительно коротким сигналам.

Чтобы оценить возможности и ограничения мультифрактального формализма, проанализируем неоднородные сигналы, содержащие различные типы сингулярного поведения.

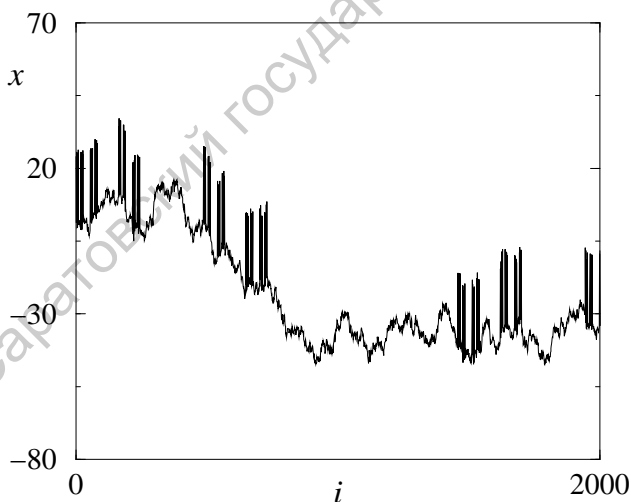
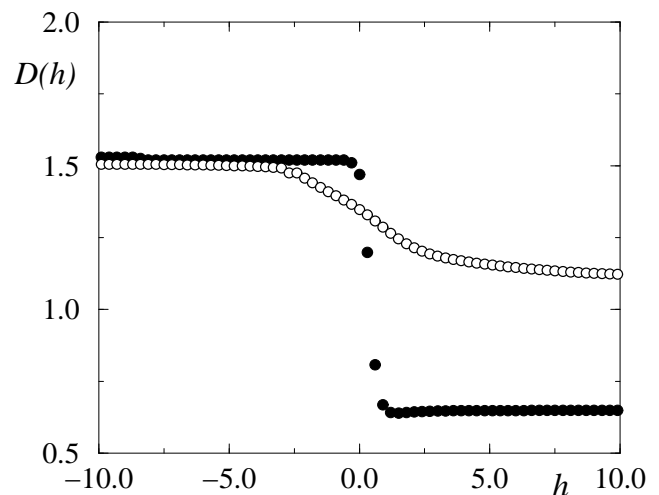


Рис. 9.9. Тестовый сигнал, представляющий собой сумму реализации винеровского случайного процесса и последовательности импульсов, имеющей структуру канторова множества

Если рассматривается сумма процессов с различными статистическими свойствами, и если данные процессы отличаются по амплитудам, то можно ожидать, что на малых масштабах (при рассмотрении слабых син-

гулярностей) проявится один из них, а на больших масштабах (сильные сингулярности) обнаружит себя другой процесс. Для наглядности рассмотрим искусственный сигнал, представляющий собой сумму реализации винеровского случайного процесса и последовательности импульсов, имеющей структуру канторова множества. За счет такого сложения (рис. 9.9) ожидается, что при рассмотрении больших флуктуаций проявится структура канторова множества, на меньших масштабах будет доминировать винеровский случайный процесс. Проведенные расчеты это подтверждают (рис. 9.10). При отрицательных q (слабые сингулярности) значения гильдеровских экспонент приближаются к $h \approx 1.5$, тогда как при положительных q (сильные сингулярности) $h \approx 0.63$. Таким образом, метод ММВП четко фиксирует различия в структуре анализируемого сигнала при изменении масштаба наблюдения. При наличии нескольких типов сингулярностей метод ММВП позволяет их четко фиксировать, если данные особенности относятся к разным масштабам. Если различия в масштабах несущественны, то четкого разделения сингулярностей не произойдет. В частности, в рассматриваемом примере суммы импульсов со структурой канторова множества и винеровского процесса увеличение интенсивности последнего будет приводить к тому, что при $q > 0$ значения гильдеровских экспонент будут смещаться к $h \approx 1.5$ (рис. 9.10). Второй тип сингулярностей ($h \approx 0.63$) может проявиться лишь при очень больших q , но при увеличении q снижается надежность оценок гильдеровских экспонент. При наличии нескольких типов сингулярностей, относящихся к близким масштабам, метод ММВП приводит к усредненным значениям h .

Рис. 9.10. Зависимость $h(q)$ для сигнала, изображенного на рис. 9.9 (черные кружочки). Белыми кружочками отмечен случай незначительных различий в масштабах наблюдения сингулярностей



Одним из ограничений мультифрактального анализа является то, что он оценивает верхнюю огибающую истинного мультифрактального спектра. Это обстоятельство может приводить к ошибочным интерпретациям результатов численного анализа. Во-первых, если истинный спектр сингулярностей $D(h)$ является дискретным, то есть h принимает только дискретное множество значений, огибающая будет включать «ложные» точки,

представляя собой интерполяцию дискретного спектра $D(h)$. Во-вторых, огибающая не позволяет идентифицировать «внутренние» точки, не принадлежащие верхней огибающей спектра $D(h)$, если они существуют – рис. 9.11. «Колоколообразная» форма спектра сингулярностей может быть получена даже в том случае, когда метод ММВП применяется для обработки сигналов, не обладающих мультифрактальными свойствами. Если процесс характеризуется не сплошным спектром $D(h)$, а небольшим дискретным набором гельдеровских экспонент, то получив огибающую, подобную изображенной на рис. 9.11, в общем случае нельзя достоверно сказать, существуют ли сингулярности, характеризующиеся значением h , равным, например, 0.63 ± 0.005 ? Чтобы избежать сложностей в интерпретации результатов, целесообразнее не анализировать, конечно либо бесконечно число скейлинговых характеристик, а рассматривать метод ММВП как аппарат численного анализа, позволяющий оценивать диапазон изменения гельдеровских экспонент (степень мультифрактальности) и характеризовать наличие корреляций различного типа в нестационарных случайных процессах. Тем самым, мы получаем возможность осуществлять корреляционный анализ по нестационарным сигналам малой длительности и проводить оценку количественной меры однородности случайного процесса.

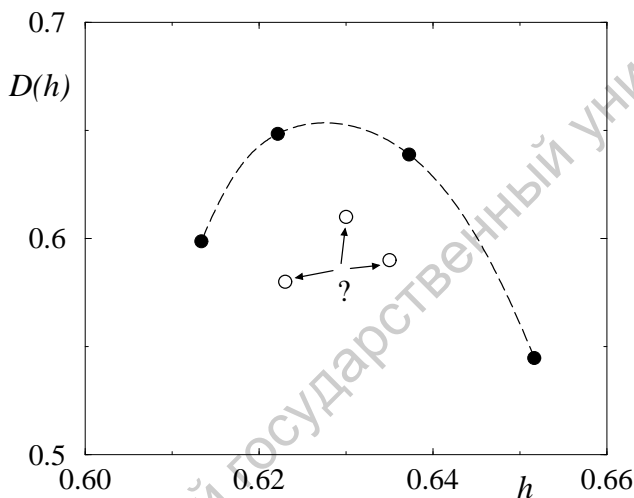


Рис. 9.11. Иллюстрация ограничений метода мультифрактального формализма. Кружочки обозначают истинный мультифрактальный спектр, пунктирная линия – результат применения мультифрактального анализа. Внутренние точки (белые кружочки) не идентифицируются на основе метода ММВП

Порядок выполнения работы

Задание 1. Провести мультифрактальный анализ канторова множества с равномерным распределением меры. С этой целью в программе «multi.x» указать файл данных «cantor.dat». Проанализировать погрешность расчета спектра сингулярностей для разной длительности последовательности (выбирая N в диапазоне от 300 до 3000).

Задание 2. Провести мультифрактальный анализ канторова множества со случайным распределением меры. По аналогии с заданием 1 вос-

пользоваться программой «multi.x», указав файл данных «cantor2.dat». Проанализировать погрешность расчета спектра сингулярностей при варьировании N в диапазоне от 1000 до 5000.

Задание 3. Провести мультифрактальный анализ по экспериментальным данным (сигнал артериального кровяного давления). Сделать вывод о наличии и о характере коррелированной динамики.

Контрольные вопросы

1. В чем состоит отличие фракталов и мультифракталов?
2. Как определяется фрактальная размерность множества?
3. Что характеризует спектр сингулярностей?
4. Как определяются частичные функции и спектр скейлинговых экспонент?
5. Что характеризуют экспоненты Гельдера?
6. Почему вейвлеты позволяют проводить анализ мультифрактальных свойств?
7. Перечислите возможности и ограничения метода мультифрактального анализа.

Список литературы

Анищенко В.С., Вадивасова Т.Е. Лекции по статистической радиофизике. – Саратов: Изд-во Саратов. ун-та, 1992.

Хованова Н.А., Хованов И.А. Методы анализа временных рядов. – Саратов: Изд-во ГосУНЦ «Колледж», 1996.

Бендат Дж., Пирсол А. Прикладной анализ случайных данных. – М.: Мир, 1989.

Отнес Р., Эноксон Л. Прикладной анализ временных рядов. – М.: Мир, 1982.

Дженкинс Г., Ваттс Д. Спектральный анализ и его приложения. – М.: Мир, 1972.

Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. – М.: Мир, 1978.

Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения. – М.: Мир, 1990.

Розанов Ю.А. Теория вероятностей, случайные процессы и математическая статистика. – М.: Наука, 1989.

Купер Дж., Макгиллем К. Вероятностные методы анализа сигналов и систем. – М.: Мир, 1988.

Вайнштейн Л.А., Вакман Д.Е. Разделение частот в теории колебаний и волн. – М.: Наука, 1983.

Пиковский А., Розенблюм М., Куртс Ю. Синхронизация: фундаментальное нелинейное явление. – М.: Техносфера, 2003.

Безручко Б.П., Смирнов Д.А. Математическое моделирование и хаотические временные ряды. – Саратов. Изд-во ГосУНЦ «Колледж», 2005.

Мандельброт Б. Фрактальная геометрия природы. – М.: Институт компьютерных исследований, 2002.

Шредер М. Фракталы, хаос, степенные законы. – Ижевск.: Регулярная и хаотическая динамика, 2001.

Божокин С.В., Паршин Д.А. Фракталы и мультифракталы. – Ижевск.: Регулярная и хаотическая динамика, 2001.

Малла С. Вэйвлеты в обработке сигналов. – М.: Мир, 2005.

Короновский А., Храмов А. Непрерывный вейвлетный анализ. – Саратов. Изд-во ГосУНЦ «Колледж», 2002.

Игошева Н.Б., Павлов А.Н., Анищенко Т.Г. Методы анализа сердечного ритма. – Саратов. Изд-во ГосУНЦ «Колледж», 2001.

Астафьева Н.М. Вейвлет–анализ: основы теории и примеры применения // *Успехи физических наук.* – 1996. – Т. 166. – №. 11. – С. 1145.

Дремин И.М., Иванов О.В., Нечитайло В.А. Вейвлеты и их применение // *Успехи физических наук.* – 2001. – Т. 171. – №. 5. – С. 465.

Павлов А.Н., Анищенко В.С. Мультифрактальный анализ сложных сигналов // *Успехи физических наук.* – 2007. – Т. 177. – №. 8. – С. 859.

Саратовский государственный университет имени Н. Г. Чернышевского

ОГЛАВЛЕНИЕ

Предисловие	3
1. Корреляционный анализ	4
1.1 Ковариационные и корреляционные функции	4
1.2 Вычисление ковариационных и корреляционных функций	5
Порядок выполнения работы	9
Контрольные вопросы	10
2. Спектральный анализ	11
2.1 Преобразование Фурье	11
2.1.1 Непрерывное преобразование Фурье	11
2.1.2 Непрерывно-дискретное преобразование Фурье	12
2.1.3 Дискретное преобразование Фурье	12
2.2 Влияние конечной длины реализации	13
2.2.1 Непрерывное преобразование Фурье	13
2.2.2 Дискретное преобразование Фурье	15
2.3 Оконный спектральный анализ	16
2.4 Расчет спектра мощности	17
Порядок выполнения работы	19
Контрольные вопросы	19
3. Преобразование Гильберта	20
3.1 Преобразование Гильберта и его свойства	20
3.2 Примеры применения	25
Порядок выполнения работы	28
Контрольные вопросы	29
4. Корреляционная размерность	30
4.1 Понятие фрактальной размерности	30
4.2 Расчет корреляционной размерности аттрактора	35
4.3 Реконструкция аттрактора	36
4.4 Особенности расчета корреляционной размерности	41
Порядок выполнения работы	45
Контрольные вопросы	47
5. Ляпуновские показатели	48
5.1 Интерпретация ляпуновских показателей	48
5.2 Расчеты по системе уравнений	49
5.3 Расчеты по одномерной реализации	52
Порядок выполнения работы	53
Контрольные вопросы	54
6. Энтропия источника	55
6.1 Предварительные замечания	55
6.2 Определение энтропии источника	56
Порядок выполнения работы	60

Контрольные вопросы	61
7. Флуктуационный анализ	62
7.1 Предварительные замечания	62
7.2 Простой флуктуационный анализ	64
7.3 Анализ флуктуаций относительно тренда	64
Порядок выполнения работы	67
Контрольные вопросы	68
8. Вейвлет-анализ	69
8.1 Предварительные замечания	69
8.2 Непрерывное вейвлет-преобразование и его свойства	71
8.3 Примеры применения	82
8.4 Вейвлет-анализ сердечного ритма	87
Порядок выполнения работы	92
Контрольные вопросы	92
9. Мультифрактальный анализ	94
9.1 Предварительные замечания	94
9.2 От сингулярных мер к сингулярным функциям	96
9.2.1 Фрактальные меры	96
9.2.2 Фрактальные функции	101
9.3 Анализ на основе вейвлет-преобразования	103
9.3.1 Вейвлет-анализ сингулярных функций	103
9.3.2 Метод максимумов модулей вейвлет-преобразования	106
9.4 Возможности и ограничения мультифрактального анализа	112
Порядок выполнения работы	114
Контрольные вопросы	115
<i>Список литературы</i>	116

Учебное издание

Павлов Алексей Николаевич

МЕТОДЫ АНАЛИЗА СЛОЖНЫХ СИГНАЛОВ

*Учебное пособие
для студентов физического факультета*

Оригинал-макет подготовил Павлов А.Н.

Подписано в печать 08.05.2008. Формат 60x84 1/16.
Бумага офсетная. Гарнитура Times. Печать трафаретная.
Усл. печ. л. 7.5 Уч.-изд. л. 7.18 Тираж 300 Заказ 0110

Издательство “Научная книга”
410054, Саратов, Б. Садовая, 127, тел.: 27-40-08

Отпечатано с готового оригинал-макета
410005, Саратов, Пугачевская, 161, офис 320, тел.: 27-26-93