

О. В. Чумак

**ЭНТРОПИИ И ФРАКТАЛЫ
В АНАЛИЗЕ ДАННЫХ**



Москва ♦ Ижевск

2011

УДК 53 + 519.2 + 519.6
ББК 22.3 + 22.171 + 22.19
Ч 90

Чумак О. В.

Энтропии и фракталы в анализе данных. — М.–Ижевск: НИЦ «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2011. — 164 с.

Энтропия как мера хаоса и мультифрактальность как мера порядка рассматриваются как важнейшие универсальные взаимосвязанные и взаимодополняющие понятия, характеризующие сложные системы самой разнообразной природы. Прослежены основные этапы развития этих понятий. Обсуждаются различные варианты построения вероятностных мер и соответствующих им энтропий на примерах реальных временных рядов и видеоизображений. Также на конкретных примерах показаны методы расчета мультифрактальных параметров через расчет энтропий. Обсуждаются некоторые специальные методы расчета фрактальных размерностей.

Книга ориентирована на студентов, аспирантов и специалистов, занимающихся анализом данных в самых разных областях. В связи с этим изложение ведется на *предельно* доступном уровне, текст снабжен большим количеством примеров. Так примеры расчета выборочных энтропий содержатся разделах 3.3 и 3.4, примеры конструктивных фракталов — в разделе 4.1, примеры природных фракталов представлены на рис. 3.6 и 3.7. Кроме того, все обсуждаемые алгоритмы расчета энтропий, фрактальных размерностей и мультифрактальных спектров иллюстрируются конкретными примерами.

ISBN 978-5-93972-852-2

© О. В. Чумак, 2011

© НИЦ «Регулярная и хаотическая динамика», 2011

<http://shop.rcd.ru>

<http://ics.org.ru>

ОГЛАВЛЕНИЕ

Предисловие	5
Глава 1. Введение	9
1.1. Фракталы и энтропии	9
1.2. Становление понятия «энтропия»	13
1.2.1. Сади Карно	14
1.2.2. Бенуа П. Клапейрон.....	15
1.2.3. Рудольф Ю. Клаузиус.....	16
1.2.4. Людвиг Больцман	18
1.2.5. Джозайя У. Гиббс	19
1.2.6. Клод Э. Шеннон.....	20
1.2.7. Андрей Н. Колмогоров.....	22
1.2.8. Альфред Реньи	23
1.2.9. Константино Тсаллес	25
1.2.10. Джон фон Нейман.....	28
1.2.11. Заключительные замечания	29
1.3. Список литературы	29
Глава 2. Энтропия и информация	33
2.1. Информация как первичное понятие	37
2.2. Мера Хартли. Теорема Хартли	41
2.3. Энтропия систем с не равновероятными состояниями	44
2.4. О принципе максимума энтропии	46
2.5. Список литературы	46
Глава 3. Энтропии временных рядов и видеоизображений	49
3.1. Общие замечания о временных рядах.....	52
3.2. О параметризации	56
3.3. Выборочные энтропии.....	57
3.3.1. Первая выборочная энтропия.....	62
3.3.2. Вторая выборочная энтропия.....	65
3.3.3. Перестановочная энтропия.....	67
3.3.4. Энтропия Реньи	69
3.4. Энтропии видеоизображений	79
3.5. Список литературы	79

Глава 4. Фракталы и их свойства

4.1. Конструктивные фракталы	81
4.1.1. Парадокс Лебега и производные Гельдера	81
4.1.2. H – фрактал	83
4.1.3. Канторова пыль	84
4.1.4. Фигуры Коха	85
4.1.5. Фрактал Минковского	86
4.1.6. Фракталы Серпинского	88
4.1.7. Круговые фракталы	
4.2. Природные фракталы	90
4.2.1. Размерность Хаусдорфа, бокс алгоритм, скейлинг	91
4.2.2. Соотношения «масса-радиус», «периметр-площадь» и «площадь-объем»	94
4.2.3. Информационная размерность	97
4.2.4. Корреляционная размерность	97
4.3. Мультифракталы	99
4.3.1. Понятие мультифрактала	99
4.3.2. Спектр размерностей Реньи	101
4.3.3. Расчет мультифрактальных спектров размерностей	102
4.3.4. Бокс алгоритм для CCD изображений	105
4.4. Список литературы	111

Глава 5. Специальные методы расчета фрактальных размерностей временных рядов

5.1. Элементы R/S анализа	113
5.1.1. Эмпирический закон Хэрста	113
5.1.2. Оценка показателя Хэрста	116
5.2. Алгоритм Хигучи	117
5.3. Алгоритм Грасбергера – Прокачча	119
5.4. Перестановочная размерность	122
5.5. Список литературы	123

Заключение	125
Список литературы	135

Приложение. Ральф Нельсон Эллиотт — автор оригинальной концепции изменчивости сложных систем	137
---	-----

Предисловие

Такие, на первый взгляд, противоположные понятия как энтропии и фракталы дают в распоряжение исследователей один из универсальных методов анализа данных, продуцируемых сложными системами самой разной природы. Последние десятилетия фракталы стали очень популярны, и уровень интереса к ним не падает, скорее, наоборот. Большую роль в этом сыграла прекрасно написанная книга франко-американского математика Бенуа Манадельброта «*Фрактальная геометрия природы*». Не меньшую роль сыграли простота и красота этого нового раздела математики, а также многочисленные и очень эффективные приложения фрактального анализа в самых разных областях — от физики, до социологии.

Общепризнано, что фрактальная геометрия является полезным инструментом для количественного описания самых разных структур. Это могут быть как идеализированные абстрактно-математические построения, временные ряды, возникающие в результате функционирования какой-либо сложной системы (например, ряд чисел Вольфа, движение цены товара на бирже и др.), так и природные объекты астрономии, геофизики, физики, химии, биологии и медицины. В последние несколько лет техника фрактального анализа привлекает все большее внимание в области обработки сигналов и изображений в астрофизике и геофизике. Есть положительный опыт применения фрактального анализа в космологии, космогонии звездных и планетных систем, физике Солнца.

Эта небольшая книга ставит своей целью дать первые представления о фракталах, а также познакомить читателя с некоторыми практическими приемами работы с техникой фрактального анализа и ее приложениями в методах обработки экспериментальных и наблюдательных данных. В силу профессиональных интересов автора в качестве примеров для иллюстрации методов анализа были взяты астрофизические временные ряды и ряды изображений, получаемых в результате космических и наземных наблюдений, но рассматриваемые подходы и методы достаточно универсальны и могут быть применимы к анализу данных любой природы.

Одним из примеров приложения фрактального анализа к нестационарным временным рядам является анализ рядов биржевых индексов. Эти ряды со сложным временным поведением очень интересны как в плане теории, так и в плане приложений. Следует отметить, что ряд фундаментальных положений фрактального анализа временных рядов был впервые сформулирован именно в этой области (см. Приложение).

В технике анализа природных фрактальных структур важное место занимают методы, основанные на расчете энтропий К. Шеннона, А. Колмогорова, А. Реньи. Энтропийный подход дает универсальный ключ как для расчетов мультифрактальных спектров размерностей систем, так и для осмысления и интерпретации результатов их фрактального анализа. Это стало возможным, когда была выявлена глубокая в фундаментальном плане и достаточно простая по форме связь между этими понятиями.

Мы привыкли к тому, что энтропия ассоциируется у нас с мерой хаотичности системы, в то время как фрактал — с предельно упорядоченной, самоподобной и масштабно инвариантной структурой. Но еще Л. Больцман поставил аккуратный знак равенства между энтропией как мерой хаотичности — с одной стороны энтропией как мерой недостающей информации о состоянии системы — с другой. К. Шеннон с других исходных позиций пришел к аналогичному выводу.

Дальнейшее развитие теории информации вскрыло глубокую связь между информацией и структурой. Стало понятно, что эти понятия практически неразрывны — каждая структура несет в себе информацию, и структура без информации немыслима, равно как и наоборот — невозможна информация сама по себе — без носителя, структуры. Таким образом, информация является универсальной, первичной категорией и везде присутствует в виде структур. Стало также понятно, что энтропия как мера хаотичности и мультифрактальность как мера структурной упорядоченности являются взаимно дополняющими понятиями. Наконец, А. Рени нашел изящную формулу связи между энтропией и континуумом мультифрактальных размерностей, открыв тем самым один из простых и достаточно универсальных путей к решению проблемы фрактальной параметризации как конструктивных фракталов, так и природных структур.

Кроме того, энтропии сами по себе являются очень информативными параметрами, количественно характеризующими состояние систем, и в этой связи в книге значительное внимание уделено изложению смысла понятий различных энтропий, а также технике расчета тех из них, которые связаны с тематикой фрактального анализа.

Поскольку книга ориентирована на читателей разных специальностей, занимающихся анализом данных в самых разных областях, автор

стремился к предельной доступности изложения и основное внимание уделял приложениям, что неизбежно привело к потере строгости. В вопросах строгости определений и доказательств автор полностью доверяет математикам, которые профессионально занимаются этими вопросами. Здесь строгость по возможности компенсировалась рассмотрением примеров, иллюстрирующих основные определения, утверждения и алгоритмы, а также ссылками на оригинальные работы математиков.

Книга начинается с краткого обзора становления и эволюции понятия энтропии — от Сади Карно (1824 г.) до Константина Тсаллеса (1988 г.). При этом обзор не претендует на охват всех обобщений энтропий (в настоящее время известно несколько десятков) и областей их применения. Цель обзора узкоспециальная: представить ключевые сведения, необходимые для понимания последующего изложения энтропийного подхода в методах фрактального анализа. Обсуждение этих вопросов составляет содержание первой главы.

Во второй главе — «Энтропии и информация» — обсуждаются смысл энтропии Шеннона и идеи о связи энтропии и информации, а также связь между энтропией Больцмана и Шеннона.

В третьей главе — «Энтропии временных рядов и видеоизображений» — приведены примеры расчета выборочных энтропий временных рядов и энтропий различных изображений. Рассмотрено применение энтропии Шеннона для анализа временных рядов и параметризации состояний сложных систем на основе такого анализа. Обсуждается алгоритм расчета выборочной энтропии Шеннона и рассматриваются примеры вычисления выборочных энтропий, реализующих этот алгоритм. Рассматриваются вопросы зависимости величины и смысла получаемых значений энтропии от выбора вероятностной меры. Приведены примеры, демонстрирующие эти различия. Приводится определение и обсуждается алгоритм расчета перестановочной энтропии временного ряда. Далее, рассмотрен пример вычисления структурной энтропии CCD-изображения. В качестве примера рассчитана структурная энтропия Шеннона для магнитного поля солнечной активной области. В этой же главе рассмотрено важное обобщение понятия энтропии, предложенное Реньи. Рассмотрена структурная, информационная и корреляционная энтропии. Обсуждаются примеры вычисления первых трех энтропий Реньи для одномерных и двумерных покрытий.

В четвертой главе — «Фракталы и их свойства» — дано краткое введение во фрактальный анализ. Рассмотрены примеры математических фракталов. Показана связь между фрактальными размерностями и выборочными энтропиями. Далее рассмотрены различные способы вычисления Хаусдорфовой размерности по данным наблюдений. Рассмотрен

алгоритм расчета фрактальных размерностей методом бокс-алгоритма. Приведены примеры расчета размерностей методом бокс-алгоритма для различных временных рядов. Далее метод бокс-алгоритма обобщается на случай расчета спектра размерностей Реньи. Обсуждается понятие мультифрактала и мультифрактальных спектров.

В пятой главе — «Специальные методы расчета фрактальных размерностей временных рядов» — рассмотрены некоторые специальные, не энтропийные, методы фрактального анализа временных рядов. Отдельный раздел целиком посвящен так называемому «методу нормированного размаха Хэрста», или, по другому, R/S анализу, временных рядов (важному в плане приложений). Рассмотрена классификация временных рядов с длинной памятью по показателю Хэрста. Показана связь показателя Хэрста с Хауздорфовой размерностью. Обсуждается метод Хигучи и алгоритм Грасбергера–Прокачча.

В Заключении подводятся краткий итог и обсуждаются некоторые общие вопросы фрактального анализа; показано место фрактального анализа в современных методах исследования сложных систем.

В конце каждой главы имеется список литературы, где можно найти более строгое или более развернутое изложение вопросов, обсуждаемых в соответствующей главе. В Приложении приведена биография (на русском языке публикуется впервые) и основные идеи одного из основоположников фрактального анализа временных рядов Р. Н. Элиота.

В заключение выражаю искреннюю благодарность дирекции Государственного астрономического института им. П. К. Штернберга МГУ за внимание и поддержку, а также Академику РАН Л. М. Зеленому и проф. А. С. Расторгуеву, прочитавшим рукопись и сделавшим ряд ценных замечаний.

ГЛАВА 1 ВВЕДЕНИЕ

1.1. Фракталы и энтропии

Понятие «фрактал» настолько общее, что до сих пор нет его общепринятого определения. Термин «фрактал» был предложен Бенуа Мандельбротом в 1975 г., получив широкую известность после опубликования в 1977 г. книги «Фрактальная геометрия природы» [1]. Словом «фрактал» — дробный, разбитый, ломаный — Мандельброт называл самоподобные геометрические фигуры, каждый фрагмент которых повторяется при уменьшении масштаба, то есть фигура обладает свойством масштабной инвариантности. Это свойство принципиально отличает фракталы от графиков гладких функций, у которых при уменьшении масштаба фрагмент стремится к отрезку прямой или плоскости, то есть к объектам с целочисленной размерностью. Специальные скейлинговые процедуры позволяют определить размерность фракталов, и эта размерность оказывается дробной. Таким образом, фрактал имеет вид ломанного или разбитого на фрагменты объекта, а его размерность — дробная. Термин «фрактал», предложенный Мандельбротом, оказался очень емким.

Тот факт, что для большинства масштабно инвариантных объектов результат не сходится к целочисленным показателям степени, которые обычно играют роль размерности (0 — точка (нульмерье), 1 — линия (одномерье), 2 — поверхность (двумерье), 3 — объем (трехмерье)... N — n-мерье), вначале казался удивительным. На самом деле ничего удивительного здесь нет. В природе нет нульмерных точек, двумерных линий, трехмерных объемов, гладких функций. Все это лишь абстрактные модели, которые, отбрасывая все «лишнее», помогают нам приближенно представить реальные объекты и вводить простую меру для их количественного сравнения.

Но таким ли уж «лишним» и «несущественным» является то, что мы отбрасываем? Иногда да, но чаще — нет. Дело в том, что скейлинговые процедуры играют фундаментальную роль в идентификации природных процессов и структур, так как скейлинг (масштабная самофинность) яв-

ляется важным свойством окружающего нас мира. Поэтому если мы, обновившись на данных наблюдений, обнаружили некую новую масштабную самоафинность, и она однозначно не является эффектами, связанными с регистрирующей аппаратурой, или процедурами обработки, то мы, скорее всего, имеем дело с новой, не открытой еще закономерностью. Так вот, как правило, показатель степени найденной нами скейлинговой степенной зависимости для этой закономерности будет дробным. А величина этой дроби — размерность нашего фрактала — является важнейшей количественной характеристикой обнаруженной закономерности. Именно эта характеристика, фрактальная размерность, помогает отличить данную закономерность от других законов или, напротив, обнаружить сходные законы в других областях. Кроме того, фрактальная размерность вместе с другими аналогичными параметрами спектра размерностей часто позволяет найти простую модель исследуемого процесса или структуры.

Мир масштабно инвариантных объектов, очень разнообразен. Это и многочисленные математические объекты самого разного вида, это и множество природных объектов. Не менее многочисленны и разнообразны методы построения, обнаружения и исследования таких объектов. Разумеется, надо иметь в виду, что математические фракталы имеют фундаментальное отличие от природных объектов: математические фракталы обладают бесконечной иерархией масштабов, в то время как природные объекты всегда имеют конечную и часто очень небольшую глубину иерархии. Подобно тому как не существует в природе бесконечных прямых линий, так нет в природе и инвариантности на бесконечном числе масштабов. Бесконечная масштабная инвариантность, подобно одномерной бесконечной прямой или бесконечно тонкой двумерной плоскости, есть лишь удобное в определенных случаях модельное представление реальных объектов, имеющее ограниченную область применения.

Этот факт налагает определенные ограничения на применение к природным объектам некоторых математических методов, прекрасно работающих на абстрактных фракталах. Более того, небольшая глубина иерархии природных фракталов часто затрудняет процедуру скейлинга и налагает ограничения на точность определения размерности. Это обстоятельство потребовало разработки таких робастных методов оценки размерности по данным наблюдений, которые бы корректно работали на конечной глубине иерархии масштабов и на, как правило, ограниченном объеме данных. Кроме того, желательно, чтобы метод не был эксклюзивным, то есть чтобы он хорошо работал не только в какой-либо одной области или с каким-либо одним классом фрактальных объектов, но имел бы достаточно широкую область применения. Это крайне важно при ана-

лизе данных, так как заранее неизвестно, есть скейлинг или нет, и, если есть, то к какому типу фракталов относится исследуемый объект.

На сегодняшний день, на наш взгляд, наиболее продвинутое решение этой проблемы основано на учете глубокой связи фрактальной геометрии с фундаментальными положениями теории образования структур при необратимых процессах. Важную роль здесь играют такие базовые понятия как хаос и структура, а также различные энтропии, количественно характеризующие эти понятия в системах разной природы. Мы здесь будем знакомиться, в первую очередь, с этим подходом. Энтропийный подход дает достаточно универсальный ключ для расчетов фрактальных размерностей и мультифрактальных спектров размерностей и применим к широкому классу природных объектов. Он полезен также для осмысления и интерпретации результатов фрактального анализа. Кроме того, энтропии сами по себе являются очень информативными параметрами, количественно характеризующими состояние систем.

Фракталы и энтропия — два диаметрально противоположных, но взаимодополняющих понятия, как свет и тьма, порядок и хаос, как китайские янь и инь. В самом деле, фракталы — это всегда структуры, причем часто высокоорганизованные структуры. Энтропия — это мера хаоса. Ее величина дает нам представление о том, как далеко система находится от упорядоченного, структурированного состояния и как близко — к полностью хаотичному, бесструктурному, однородному виду. Но что же такое структура?

Есть разные определения понятия «структура». Здесь под структурой будем понимать характер организации элементов и совокупность отношений между элементами системы. Понятно, что структура задается природой связей между элементами системы. Такое определение структуры не налагает никаких ограничений на природу самой системы и ее элементов. Это могут быть системы атомов, буквенные сообщения, космические, биологические или социальные системы. Каждая система может быть либо высокоорганизованной, то есть обладать развитой и сложной структурой, либо менее организованной, имеющей достаточно простую структуру, либо и вовсе хаотической, то есть когда ее элементы распределены случайно и, в среднем, однородно.

Как уже отмечалось, мерой хаоса является энтропия. Но в такой же степени энтропия является и мерой структурной организованности систем, так как порядок и хаос — это не только противоположные, но, как отмечалось, и взаимодополняющие понятия. Это то самое единство противоположностей, равновесие или неравновесие между которыми определяет направление и темп развития или деградации структур в рассматриваемой системе.

Энтропия как мера хаоса или порядка обнаруживалась в системах разной природы: это и энтропия Клаузиуса в термодинамике, и энтропия Больцмана в статистической физике, и энтропия Шеннона в теории информации, энтропия Колмогорова в теории динамических систем, и энтропия фон Неймана в квантовой механике. Понимание универсальности энтропии как меры хаотичности независимо от природы системы приходило постепенно. Вот, например, что писал в 1987 г. А. Н. Колмогоров [2] по поводу энтропии Шеннона: «Общеизвестно и то обстоятельство, что выражение $H(\zeta)$ формально тождественно с выражением энтропии в физике. Это совпадение я считаю вполне достаточным для оправдания наименования величины $H(\zeta)$ и в теории информации «энтропией»: такие математические аналогии следует всегда подчеркивать, так как сосредоточение внимания на них содействует прогрессу науки». С развитием физики открытых систем стало понятно, что «среди различных макроскопических функций только энтропия S обладает совокупностью свойств, позволяющих использовать ее в качестве меры неопределенности (хаотичности) при статистическом описании процессов в макроскопических системах», отмечал в 2002 г. Ю. Л. Климонтович [3].

Максимальная энтропия заданной системы, соответствует низшей степени ее структурной организованности, то есть наибольшей хаотичности и неупорядоченности. Низкая энтропия, напротив, соответствует высокой структурной упорядоченности. Энтропия нам здесь будет интересна, прежде всего, именно в этом ее качестве — как мера структурной организованности системы. Чтобы подчеркнуть это обстоятельство, далее мы будем иногда называть ее структурной энтропией.

В реальных природных системах обнаруживаются не только пространственные, но и временные структуры. Как правило, пространственная и временная организации настолько взаимосвязаны, что обычно говорят о пространственно-временных структурах. Временные структуры обычно изучаются путем анализа временных рядов данных наблюдений, пространственные, как правило, — путем анализа двумерных изображений. Поэтому в последующих главах определенное внимание будет уделено методам расчета структурных энтропий как рядов, так и изображений.

Всякая ли природная структура является фракталом? Вот что думал по этому поводу основоположник фрактальной геометрии Б. Мандельброт [1] (с. 13): «Почему геометрию так часто называют «холодной» и «сухой»? Одна из причин — ее неспособность описать форму облака, горы, дерева или береговой линии. Облака не являются сферами, горы — конусами, береговые линии нельзя изобразить с помощью окружностей, кору деревьев не назовешь гладкой, а путь молнии — прямолинейным.

В более общем виде я заявляю, что многие формы природы настолько неправильны и фрагментированы, что в сравнении с евклидовыми фигу-

рами (евклидовыми в данной работе мы будем называть все, что относится к обычной геометрии) природа демонстрирует не просто более высокую степень, но совершенно иной уровень сложности. Количество масштабов длины в естественных формах можно считать бесконечным для каких угодно практических задач».

Таким образом, Б. Мандельброт полагал, что модельные представления евклидовой геометрии слишком грубы и примитивны для описания природных структур, а фрактальные представления являются более адекватными.

Выше мы уже упоминали о принципиальных отличиях абстрактных фракталов от природных структур. Важнейшим из них является отсутствие в природных объектах бесконечной масштабной иерархии самоподобных фигур. Да и сами фигуры на разных масштабах, как правило, не до конца самоподобны. Строгое самоподобие — это тоже лишь модельное, абстрактное понятие. Формальная строгость, необходимая в математике, не всегда бывает востребованной в практических приложениях. «Фрактальная геометрия оказывается во многих случаях излишне формальной, и возникает необходимость развития “Физики фракталов”» — отмечает Ю. Л. Климонтович [3] (с. 87). Энтропийный подход к исследованию природных структур, на наш взгляд, — шаг в этом направлении.

1.2. Становление понятия «энтропия»

1.2.1. Сади Карно

История энтропии началось с любознательного молодого человека, военного инженера в отставке **Сади Карно**. С. Карно поразили и увлекли только недавно появившиеся паровые машины. Пытливый и ясный ум инженера по образованию и физика по призванию сумел проникнуть в самую суть таинственного *превращения* (греч. *εὐτροπὴ*) тепла в механическое движение. После длительных размышлений, наблюдений за работой паровых машин он опубликовал в Париже 45-страничную брошюру, или, как тогда принято было говорить, мемуар, название которого можно перевести как «Размышления о движущей силе огня и о машинах, способных развить эту силу». В этом труде С. Карно объяснил, как и почему теплота может превращаться в работу. Размышляя о «получении движения из теп-



С. Карно
(1796–1832)

ла» и анализируя полный рабочий цикл идеальной тепловой машины (известный сейчас как цикл Карно), он впервые пришёл к выводу о том, что работа производится только при переходе тепла от нагретого тела к более холодному. Карно высказал также положение, что величина работы обусловлена разностью температур нагревателя и холодильника и не зависит от природы вещества, работающего в тепловой машине (теорема Карно).

В рассматриваемой работе С. Карно было сформулировано, как выяснилось спустя полтора столетия, одно из важнейших научных положений современного естествознания. Закон называется *Принцип Карно*, но сейчас он больше известен как *второе начало термодинамики* или *принцип энтропии*. Таким образом, второй закон термодинамики, принцип энтропии был открыт *раньше* первого закона термодинамики — принципа сохранения энергии.

С. Карно было 28 лет, когда он опубликовал свой мемуар в 1824 г., а через восемь лет, в 1832 г., Сади Карно умер во время эпидемии холеры в Париже. При жизни его труд остался неизвестным. Но после публикации «Размышлений...» Карно продолжил изучение тепловых машин. Свои наблюдения, размышления и выводы он записывал в дневнике, намереваясь со временем их систематизировать и опубликовать. Однако судьба не предоставила ему такой возможности... Дневники С. Карно были опубликованы лишь через семьдесят лет после его смерти — в 1902 г. Как выяснилось, в этих заметках Карно сформулировал основные положения кинетической теории, подробно обосновал принцип сохранения энергии (первое начало термодинамики), вычислил механический эквивалент теплоты. Полстолетия понадобилось, чтобы все это открыть заново!

1.2.2. Бенуа Поль Эмиль Клапейрон



Б. П. Э. Клапейрон
(1799–1864)

Спустя два года после смерти С. Карно, в 1834 г., **Бенуа Поль Эмиль Клапейрон** прочитал его «Размышления...» и по достоинству оценил открытия, содержащиеся в этом труде. Клапейрон познакомил широкую научную общественность своего времени с идеями Карно и далее развивал их в своих работах. Он формализовал принцип Карно, придав ему графическую и аналитическую форму, и вывел КПД тепловой машины. Работа W в цикле Карно равна:

$$W = Q_1 - Q_2, \quad (1.1)$$

где Q_1 — количество тепла в рабочем теле до выполнения работы, Q_2 — количество тепла в

рабочем теле после выполнения работы. Клапейрон также выяснил, что количество тепла в конкретном тепловом цикле пропорционально температуре, то есть $Q = \xi T$, где ξ — константа. Коэффициент полезного действия (κ) машины (КПД) Клапейрон определил как отношение работы к затраченной теплоте:

$$\kappa = W/Q_1 = (Q_1 - Q_2)/Q_1 = (T_1 - T_2)/T_1. \quad (1.2)$$

Здесь T_1 и T_2 — температура рабочего тела до и после выполнения работы. Клапейрон вывел уравнение состояния идеального газа (уравнение Клапейрона), которое объединило закон Бойля–Мариотта, закон Гей-Люссака и закон Авогадро.

1.2.3. Рудольф Юлиус Эммануэль Клаузиус

Однако все, что было сказано выше, можно считать лишь предысторией понятия энтропии. Но вот, в 1865 г. **Рудольф Юлиус Эммануэль Клаузиус** ввел представление о новой термодинамической величине, которую он определил из следующих соображений. Работа в цикле Карно (см. уравнения (1.2)) может быть записана так:

$$W = Q_1 - Q_2 = Q_1(T_1 - T_2)/T_1 \quad (1.3)$$

Откуда видим, что в цикле Карно $Q_1/T_1 = Q_2/T_2$. То есть отношение Q/T сохраняется. Этой сохраняющейся величине он дает звучное название **энтропов** — **превращение**, имея в виду, что эта величина как раз и является мерой превращения тепловой энергии в механическую, и наоборот.

Далее Клаузиус вводит дифференциал $dS = dQ/T$, в котором приращение энтропии определяет как изменение тепловой энергии dQ , отнесенное к абсолютной температуре T , и показывает, что интеграл $S_1(V_1, T_1) - S_2(V_2, T_2) = \int dQ/T = Q_1/T_1 - Q_2/T_2 = 0$, то есть не зависит от пути интегрирования в пространстве переменных (V, T) . Таким образом, интеграл Клаузиуса S — оказывается функцией состояния системы.



Р. Ю. Э. Клаузиус
(1822–1888)

Размерность энтропии Клаузиуса — Дж/К. К безразмерной энтропии можно перейти, если разделить ее на некоторую постоянную, которую в последствии стали называть постоянной Больцмана:

$$S = (1/k) \int dQ / T, \quad (1.4)$$

$k = 1.38 \cdot 10^{-23}$ Дж/К. Величина k , как видим, очень мала. Тот факт, что она стоит в знаменателе, говорит о том, что изменения энтропии, даже в самых простых термодинамических системах, очень велики.

Так и остался бы в теплофизике известный лишь специалистам интеграл Клаузиуса, если бы сам Клаузиус не заварил всю кашу, — он эту свою идею о *превращениях* энергии распространил на все процессы и даже на всю Вселенную и пришел к выводу о неизбежной тепловой смерти Вселенной. Малопонятное, но звучное название «энтропия» плюс «триллерный» сюжет со Вселенной сделали свое дело: об энтропии заговорили все — от философов до продавцов на рынке, при этом каждый вкладывал в это слово свой смысл. Из чисто научного термин постепенно превратился (ох, уж этот энтропон!) в бытовую [4]¹. В сознании широкой публики второе начало термодинамики и по сей день окружено странным магическим ореолом. О нем слагаются стихи и песни, о нем читают лекции. Считается, что прямо и доступно объяснить это понятие нельзя.

1.2.4. Людвиг Больцман

Крупнейший физик XIX и начала XX века **Людвиг Больцман** потратил большую часть жизни, стараясь объяснить смысл энтропии. Современный известный итальянский физик-теоретик Карло Черчичьяни², написавший замечательную книгу о Л. Больцмане [5], подчеркивая важность его работ в этой области, пишет: «Только хорошо поняв второе

¹ Это ссылка на прекрасную книгу П. Шамбодала, в которой ее автору удалось удачно сочетать научную строгость с популярностью изложения основ классической термодинамики. Особое место занимает история становления и развития понятия энтропии в теплофизике. Несмотря на 1967 г. издания, книга не потеряла своей ценности в качестве хорошего введения в предмет.

² Великий научный подвиг Л. Больцмана и его полная драматизма научная судьба привлекали внимание многих писателей, философов, историков науки. Один из ведущих современных специалистов в области кинетической теории и статистической физики Карло Черчичьяни, изучавший научное наследие Л. Больцмана в силу своей профессиональной деятельности, в полной мере осознал и оценил тот огромный прорыв в области теоретической физики, который был сделан благодаря работам Больцмана. Черчичьяни проанализировал огромное количество документов, связанных с жизнью и научной деятельностью Л. Больцмана. Живой стиль изложения, большой фактический материал, положенный в основу, делают книгу Черчичьяни интересной как ученым, работающим в области теоретической физики, так и всем, кто интересуется историей науки.

начало термодинамики, можно ответить на вопрос, почему вообще возможна жизнь». Дело в том, что живые существа должны поддерживать свою форму и все внутренние структуры, постоянно борясь со всеобщей тенденцией к рассеянию и хаосу. Именно эту тенденцию, утверждал Клаузиус, описывает второе начало.

Больцману удалось прояснить кинетический смысл энтропии. Он показал, что логарифм числа P равновероятных микросостояний термодинамической системы обладает всеми свойствами термодинамической энтропии Клаузиуса. А если его умножить на специальную константу $k = 1.38 \cdot 10^{-23}$ Дж/К, то такая функция становится полностью тождественной в физическом смысле размерной энтропии Клаузиуса

$$S = k \log P. \quad (1.5)$$

Так в физике появилась константа, которую потом стали называть постоянной Больцмана.

Л. Больцман, давая статистико-физическое определение энтропии, в 1877 году заметил, что *энтропия характеризует недостающую информацию* о системе. То есть, энтропия есть мера нашего незнания о системе. Таким образом, Больцман был первым, кто разглядел информационный смысл энтропии. И только спустя 80 лет теоретико-информационная трактовка энтропии стала общепринятой.

На самом деле формула (1.5) является удивительным научным прозрением Больцмана. Научная общественность, встретившая по началу в штыки его кинетические идеи, в конце концов по достоинству оценила научный подвиг Л. Больцмана в выяснении кинетической и информационной природы энтропии. Формула (1.5) выбита на надгробии могилы Больцмана на Центральном кладбище в Вене.

Остановимся и мы немного на этой формуле, чтобы приоткрыть и немного прочувствовать ее смысл. Здесь P называется статистическим весом макросостояния системы. Он равен числу способов, которым может быть реализовано данное макросостояние. Рассчитывают P обычно следующим образом. Разобьем фазовый объем системы на n шестимерных кубиков. Число частиц в таком i -том кубике пусть равно N_i , а общее число частиц N , соответственно, равно:

$$N = \sum_{i=1}^n N_i.$$



Л. Больцман
(1844–1906)

Полное число перестановок всех частиц равно $N!$, однако перестановки частиц внутри каждого кубика не различаются в рамках данного разбиения. Таким образом, число способов, которым может быть реализовано данное микросостояние, равно:

$$P = \frac{N!}{N_1! N_2! \dots N_n!}. \quad (1.6)$$

Подставляя (1.6) в (1.5) и воспользовавшись формулой Стирлинга, получаем:

$$S \approx k(N \log N - \sum_{i=1}^n N_i \log N_i). \quad (1.7)$$

Учитывая, что $N_i \log N_i = N(f_i \log f_i + f_i \log N)$, где $f_i = N_i/N$, имеем:

$$S \approx -Nk \sum_{i=1}^n f_i \log f_i. \quad (1.8)$$

Разделив левую и правую части равенства (1.8) на N , получаем удельную (приходящуюся на одну частицу) энтропию. Переходя затем к так называемому термодинамическому пределу $\lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} f_i = f(x, v) dx dv$, получаем, что удельная энтропия Больцмана S_B равна:

$$S_B = -k \int_{\mu} f(x, v) \log f(x, v) dx dv + S_0, \quad (1.9)$$

где $dx dv$ — элемент шестимерного фазового μ -пространства, $f(x, v)$ — функция распределения, S_0 — произвольная постоянная. Поэтому в классической термодинамике можно говорить только о разности энтропий. Таким образом, формула (1.9) дает определение удельной энтропии Больцмана, выраженное в терминах функции распределения элементов системы в шестимерном фазовом μ -пространстве.

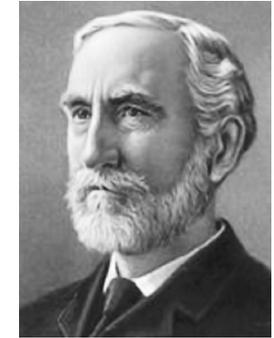
1.2.5. Джозайя Уиллард Гиббс

В своей фундаментальной работе «Основные принципы статистической механики» Джозайя Уиллард Гиббс [6] пошел несколько иным путем.

Он ввел энтропию через распределение плотности $F(X)$ вероятности состояний по фазовому $6N$ -мерному Γ -пространству статистико-физической системы:

$$S_G = -k \int_{\Gamma} F(X) \log F(X) d\Gamma + S_0, \quad (1.10)$$

где X — $6N$ координата этого пространства. Гиббс пришел к выводу, что его путь, в частности, приводит к результату Больцмана, если пренебречь корреляциями между частицами, то есть если положить вероятность $F(X)$ равной произведению вероятностей $f(x, v)$. Таким образом, энтропия Гиббса S_G носит более общий характер по сравнению с энтропией Больцмана, так как формально учитывает корреляции состояний. По сути, энтропия Гиббса стала первым существенным обобщением этого понятия со времени Клаузиуса [4].



Д. У. Гиббс
(1839–1903)

1.2.6. Клод Элвуд Шеннон

В работе «Математическая теория связи» (1948 г.) К. Шеннон [7], исследуя передачу сообщений по шумящим линиям, ввел меру дискретного распределения вероятности P_i на множестве альтернативных состояний передатчика (transmitter) и приемника (receiver) сообщений и вывел формулу, ставшую основой количественной теории информации:

$$H = -\sum_{i=1}^n p_i \log_2 p_i \quad (1.11)$$

Здесь n — число символов, из которых может быть составлено сообщение (алфавит), H — мера неопределенности. Если N — число всех переданных и принятых символов в сообщении, то $P_i = m_i/N$ — вероятность появления i -го символа в сообщении, m_i — число встречаемости i -го символа в сообщении.

По аналогии с формулой Больцмана, он назвал величину H энтропией. Шеннон говорил, что назвать величину H энтропией ему посоветовал



К. Э. Шеннон
(1916–2001)

товал знаменитый математик Джон фон Нейман, который, шутя, обосновывал это тем, что мало кто из математиков и инженеров знает об энтропии, и это даст Шеннону большое преимущество в неизбежных дискуссиях. Но, скорее всего, это была не просто шутка, основанная на формальной аналогии. И Дж. Фон Нейман и К. Шеннон, вероятно, были хорошо осведомлены об информационном смысле энтропии Больцмана как о величине, характеризующей недостающую информацию о системе.

Но что же такое тогда сама информация? На этот вопрос Шеннон дает гениально простой ответ: это убыль энтропии в результате приема сообщений. Итак, информацию I Шеннон определил как уменьшение энтропии при получении сообщений:

$$I = H_1 - H_2. \quad (1.12)$$

На этих идеях и определениях ниже мы остановимся более подробно. А здесь упомянем еще о четырех важных обобщениях понятия энтропии, которые нам понадобятся в дальнейшем.

1.2.7. Андрей Николаевич Колмогоров



А. Н. Колмогоров
(1903–1987)

А. Н. Колмогорову принадлежит не только существенная роль в превращении теории информации, сформулированной К. Э. Шенноном в виде, скорее, технической дисциплины, в строгую математическую науку, но и построение основ теории информации на принципиально ином, отличном от шенноновского, фундаменте. В процессе этих работ, а также работ в области теории динамических систем, Ф. Н. Колмогоров [8] обобщает понятие энтропии на эргодические случайные процессы через предельное распределение вероятности, имеющее плотность $f(x)$.

Чтобы понять смысл этого обобщения, напомним некоторые определения из теории случайных процессов. Пусть имеется некоторый стационарный случайный процесс (например, броуновское движение) и пусть u — некоторый измеряемый параметр этого процесса (например, вектор положения броуновской частицы). *Реализацией* случайного процесса назовем временную последовательность значений параметра $u(t)$. *Выборкой* i из данной реализации назовем некий конечный фрагмент данной реализации, или

последовательность значений $u(t)$, полученную путем применения к исходной реализации определенного правила отбора значений, то есть некую конечную последовательность значений параметра $u_i(t)$.

Случайный процесс называется эргодическим, если среднее по времени от значений выборок процесса

$$M \left(\frac{1}{n} \sum_{i=1}^n u_i(t) \right) \quad (1.13)$$

стремится к его математическому ожиданию реализации $Mu(t)$, которое не зависит от t в силу стационарности процесса. То есть это такие процессы, для которых

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t) dt = Mu(t) = \text{const}. \quad (1.14)$$

Разобьем пространство параметра u на ε ячейки, а время t пусть тоже изменяется дискретно так, что $t_n = ndt$. Пусть теперь $f(u_i(t_n))$ — вероятность данной выборки $u_i(t_n)$, то есть совместная вероятность того, что в момент времени t_0 броуновская частица находилась в ε ячейке с номером 0, в ячейке 1 — в момент времени $t_1 = t_0 + dt, \dots, t_n = t_0 + ndt$. Тогда мера априорной неопределенности положения броуновской частицы дается формулой:

$$K(n) = - \sum_{i=0}^n f(u_i(t_n)) \ln f(u_i(t_n)). \quad (1.15)$$

Разность энтропий $(K(n+1) - K(n))$ есть потерянная информация о положении частицы на интервале dt . Скорость потери информации на всем интервале ndt равна $(K(0) - K(n)) / ndt$. Энтропия Колмогорова определяется как предельное значение этой оценки при заданном начальном значении ($K(0) = 0$):

$$K = \lim_{dt \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{K(n)}{ndt}. \quad (1.16)$$

Таким образом, энтропия Колмогорова (или K -энтропия) дает нам оценку скорости потери информации и может рассматриваться как мера «памяти» системы, или мера скорости «забывания» начальных условий

[9]. Ее можно также рассматривать как меру хаотичности системы. Так, для чисто случайного процесса прошлые значения u_i никак не связаны и не несут никакой информации о настоящем и будущем (даже ближайшем) в значениях u . В этом случае все значения u_i равновероятны и $f(u_i) = 1/n$. Откуда $K(n) = \ln(n)$. Подставляя это значение в (1.16) и учитывая, что $ndt = \text{const}$ при предельном переходе, получаем, что для процесса «без памяти», $K \rightarrow \infty$. Аналогично можно показать, что, в случае регулярного движения, когда $K \rightarrow 0$, можно строго предсказать любые будущие значения $u(t)$. Таким образом, чем меньше K , тем более детерминированной является система. Напротив, высокие значения K говорят о высокой стохастичности системы, короткой «памяти» и, соответственно, плохой прогнозируемости ее будущих состояний. Роль K -энтропии для нелинейных систем в чем-то похожа на ту роль, которую играет автокорреляционная функция для линейных систем. K -энтропия обладает и другими интересными свойствами. В частности, она связана с показателями Ляпунова, характеризующими устойчивость динамической системы.

1.2.8. Альфред Реньи



А. Реньи
(1921–1970)

В 1960 г. венгерский математик Альфред Реньи предлагает свое обобщение энтропии [10]. Он вводит энтропию как q — момент меры ε -разбиения (покрытия)

$$R(q) = \frac{1}{1-q} \ln \left(\sum_{i=1}^{N(\varepsilon)} p_i^q \right), \quad \sum_{i=1}^{N(\varepsilon)} p_i = 1. \quad (1.17)$$

Здесь $p_i = N_i/N(\varepsilon)$, N_i — число элементов системы, приходящихся на i -элемент ε -разбиения, $N(\varepsilon)$ -полное число элементов заданного ε покрытия. Константа q может принимать любые значения, однако смысл энтропии Реньи при этом, соответственно, меняется.

Далее мы будем более подробно рассматривать энтропию Реньи, поэтому здесь ограничимся только этим определением. А сейчас кратко остановимся еще на одном обобщении энтропии, которое, возможно,

в ближайшем будущем будет играть важную роль в теоретической физике и астрофизике. Имеется в виду так называемая энтропия Тсаллеса.

1.2.9. Константино Тсаллес

В 1988 г. Константино Тсаллес предлагает новое обобщение энтропии [11] — так называемую неэкстенсивную (неаддитивную) энтропию T_q , которая на дискретном числе микросостояний N определяется следующим выражением:

$$T_q = k \frac{1 - \sum_{i=1}^N p_i^q}{q-1} \quad \left(\sum_{i=1}^N p_i = 1; q \in R; k > 0 \right). \quad (1.18)$$

Здесь p_i определены так же, как и в (1.17), k — размерная константа, если размерность существует для понимания задачи.

Тсаллес и его последователи предлагают «Неэкстенсивную статистическую механику и термодинамику» как обобщение этих классических дисциплин на случай систем с длинной памятью и/или далекодействующими силами.

Перечислим основные свойства энтропии Тсаллеса.

1. При q , стремящемся к единице, она также, как и энтропия Реньи, переходит в энтропию Шеннона:

$$\begin{aligned} \lim_{q \rightarrow 1} T_q &= k \frac{1 - \sum_{i=1}^N p_i p_i^{q-1}}{q-1} = k \frac{1 - \sum_{i=1}^N p_i e^{(q-1) \ln p_i}}{q-1} \approx \\ &\approx k \frac{1 - \sum_{i=1}^N p_i [1 + (q-1) \ln p_i]}{q-1} = -k \sum_{i=1}^N p_i \ln p_i. \end{aligned} \quad (1.19)$$

2. Если $q > 1$, а все $p_i \sim 0$, кроме одного $p_i \sim 1$ (заполнен один уровень — система полностью детерминирована), то $T_q > \sim 0$.

3. T_q достигает экстремума (максимума — при $q > 1$ и минимума — при $q < 1$) в случае равновероятностного распределения (то есть $p_i = 1/N$).

В этом случае

$$T_q = k (N1 - q - 1) / (1 - q). \quad (1.20)$$



К. Тсаллес
(род. 1943)

При $q = 1$, $S = k \ln N$, то есть получаем известную формулу Больцмана–Хартли–Шеннона для равновероятных распределений.

4. T_q является вогнутой при $q > 0$ и выпуклой при $q < 0$, что является основной термодинамической устойчивости.

5. Важнейшим свойством энтропии Тсаллеса является ее неаддитивность, в отличие от всех рассмотренных нами энтропий, в которых была принята аддитивная мера Хартли. То есть если две системы A и B независимы, то

$$\frac{T_q(A+B)}{k} = \frac{T_q(A)}{k} + \frac{T_q(B)}{k} + (1-q) \frac{T_q(A)T_q(B)}{k^2}. \quad (1.21)$$

При $q < 1$ получаем сверхэкстенсивный случай, $q = 1$ — экстенсивный и при $q > 1$ — субэкстенсивный.

6. Если N состояний системы разбить на две подсистемы L и M ($N = L + M$) так, что

$$\sum_L p_l = p_L \text{ и } \sum p_m = p_M; (p_L + p_M) = 1,$$

тогда

$$T_q(\{p_i\}) = T_q(p_L, p_M) + p_L^q T_q(\{p_l / p_L\}) + p_M^q T_q(\{p_m / p_M\}). \quad (1.22)$$

Эта формула обобщает известную формулу Шеннона.

От всех энтропий в том числе и от энтропии Реньи, энтропия Тсаллеса отличается тем, что не является аддитивной. Это принципиальное и важное отличие. По мнению Тсаллеса и его многочисленных последователей, это отличие позволяет построить новую термодинамику и новую статистическую теорию, которые способны просто и корректно описывать системы с длинной памятью, а также и такие системы, в которых каждый элемент взаимодействует не только с ближайшим соседом или несколькими ближайшими соседями, но и со всей системой в целом или ее частями. Типичным примером таких систем, а следовательно, перспективным объектом возможных приложений новой теории являются гравитирующие системы: звездные скопления, газопылевые комплексы, галактики, скопления галактик и др.

Новая формулировка термодинамики аномальных (т. е. с длиной памяти или с далекодействующими силами) систем на основе энтропии (1.18) была предложена, как уже отмечалось, Константино Тсаллесом в 1988 г.

С тех пор появилось значительное число приложений этой теории, в том числе и в области термодинамики гравитирующих систем. Заинтересованных читателей отсылаю к обзору (Тсаллес, 1999) [12]. На сайте <http://tsallis.cat.cbpf.br/biblio.htm> можно найти полную и весьма обширную по данной проблеме библиографию, которая постоянно обновляется.

1.2.10. Квантовая энтропия фон Неймана

В заключение этой главы упомянем об энтропии фон Неймана, которая играет важную роль в квантовой физике и в некоторых астрофизических приложениях. В этой связи напомним некоторые понятия из квантовой статистики.

Среднее значение физической величины A , оператором которой является \hat{A} , а квантовое состояние описывается волновой функцией ϕ , в x представлении равно:

$$\langle A \rangle = (\phi^*, \hat{A} \phi) = \int \phi^*(x) \hat{A} \phi(x) dx, \quad (1.23)$$

где (*) означает комплексное сопряжение. Состояния квантовой системы, которые можно описать волновой функцией, называются *чистыми состояниями*. Вся классическая квантовая механика основана на применении таких чистых ансамблей.

Теперь представим себе систему, которая с разной вероятностью находится в некотором дискретном наборе n состояний (дискретный базис), каждое из которых описывается своей волновой функцией ϕ_i . Такое состояние системы называется *смешанным*. Среднее значение физической величины A для системы в смешанном состоянии равно:

$$\langle A \rangle = \sum_i p_i (\phi_i^*, \hat{A} \phi_i),$$

$$\sum_i p_i = 1, \quad p_i = \frac{N_i}{N}, \quad \sum_i N_i = N, \quad (1.24)$$

где p_i — вес состояния ϕ_i .

Таким образом, *оператором плотности вероятности*, соответствующим чистому состоянию, является (см. 1.23):

$$\rho = (\phi, \phi^*), \quad (1.25)$$



Д. фон Нейман
(1903–1957)

а оператором плотности, соответствующим плотности вероятности смешанного состояния, является (см. 1.24):

$$\rho_{i^*i} = p_i(\phi_i^*, \phi_i). \quad (1.26)$$

ρ_{i^*i} называется *матрицей плотности вероятности* смешанного состояния или просто матрицей плотности. Формулу (1.24) для среднего значения физической величины A в этих обозначениях можно записать в виде³:

$$\langle A \rangle = \text{Tr}(\rho_{i^*i} \hat{A}_{ii^*}). \quad (1.27)$$

Матрица плотности является фундаментальным понятием квантовой статистической физики. Она представляет собой *оператор плотности вероятности* состояний системы и описывает как замкнутые системы (*чистые состояния*), так и открытые системы (*смешанные состояния*), то есть системы, взаимодействующие как со своими частями, так и с окружением. Формализм квантовой физики, основанный на матрице плотности, был предложен Дж. фон Нейманом и независимо от него Л. Д. Ландау и Ф. Блохом в 1927 г. Необходимость такого подхода для описания квантовых систем была связана с тем, что формализм волновой функции описывает замкнутые системы, то есть только чистые состояния. Пространство же состояний для матрицы плотности это не только набор всех дискретных (базисных) состояний, это и все возможные корреляции между ними. Таким образом, матрица плотности в квантовой статистике играет в некотором смысле роль, подобную функции распределения Гиббса в классической статистической физике.

Оператор плотности — это неотрицательный *эрмитов* оператор с *единичным следом* в пространстве состояний:

$$\text{Tr}(\rho) = 1. \quad (1.29)$$

Равенство следа единице соответствует единичной нормировке полной вероятности на данном пространстве состояний.

След квадрата матрицы плотности:

$$\text{Tr}(\rho^2) = \sum_{i^*i} |\rho_{i^*i}|^2 \leq 1. \quad (1.30)$$

Его можно рассматривать в качестве *меры чистоты состояния* системы. В самом деле, для чистых состояний только одно значение — $p = 1$, все остальные $p_i = 0$ (см. (1.25) и (1.26)), и, соответственно, след квадрата

³ $\text{Tr}(A)$ — след матрицы (от англ. tract — след); иногда пишут $\text{Sp}(A)$ (от нем. Spur — след).

матрицы состояний равен единице. Для смешанных состояний выражение (1.30) всегда будет меньше единицы.

Энтропия, как известно, является важной характеристикой любой системы. Для случая квантовых систем в смешанных состояниях фон Нейман обобщил шенноновское определение энтропии (1.11), приняв в качестве вероятностной меры элементы матрицы плотности. В диагональном представлении энтропия фон Неймана может быть записана в виде:

$$S = -\sum_{i=1}^n \rho_i \log \rho_i. \quad (1.31)$$

Так как, согласно определению, $\text{Tr}A = \sum_i a_{ii}$, то эту формулу можно записать в виде:

$$S = -\text{Tr}(\rho_i \log \rho_i). \quad (1.32)$$

а поскольку для эрмитовых матриц от диагонального представления всегда можно перейти к произвольному и обратно, то в общем виде энтропия фон Неймана обычно записывается без индексов:

$$S = -\text{Tr}(\rho \log \rho). \quad (1.33)$$

Мы уже упоминали, чем отличаются чистые и смешанные состояния. Здесь необходимо упомянуть еще один чисто квантовый эффект, присущий смешанным состояниям. Речь идет о такой специфической черте квантовых систем как *нелокальность*. Дело в том, что волновые функции по своей природе нелокальны, и расстояние между частями системы, описываемой этими функциями, существенной роли не играет. Пусть у нас имеется квантовая система в смешанном состоянии, описываемом матрицей плотности (1.26). Такая система разными своими частями может одновременно находиться в разных точках обычного пространства. Причем расстояние между этими точками пространства, в принципе, может быть любым. Таким образом, воздействуя на одну часть системы, мы изменяем вероятность ее состояния и тем самым мгновенно изменяем распределение вероятностей в других ее частях, находящихся в том числе и в произвольно удаленных местах. Этот чисто квантовый эффект нелокальности получил название квантовой телепортации.

Эффект квантовой нелокальности широко обсуждался в первой половине прошлого века в связи с формулировкой парадокса Эйнштейна–Подольского–Розена. Эти физики полагали, что эффект нелокальности не имеет никакого отношения к реальности и является недостатком (неполнотой, неадекватностью) математического аппарата квантовой механики.

Однако развитие теории и многочисленные эксперименты, выполняемые с 1997 г., показали, что нелокальность является реальным квантовым эффектом. К настоящему времени обсуждаются различные возможные приложения этого эффекта, например, в связи с разработкой квантовых компьютеров. Эффекты квантовой нелокальности играют определенную роль в астрофизике, в том числе в космологии и в физике черных дыр.

Энтропия фон Неймана, как всякая энтропия, связана с информацией; в данном случае — с информацией о квантовой системе, описываемой матрицей плотности (1.26). И в этом плане она играет роль фундаментального параметра, количественно характеризующего состояние и направление эволюции квантовой системы.

1.2.11. Заключительные замечания

Возникает важный вопрос: имеют ли между собой все эти и другие многочисленные определения энтропии нечто общее или кроме общего названия их ничто не объединяет? Есть учёные, которые отвечают на этот вопрос утвердительно, есть те, кто отвечает отрицательно, остальные — сомневаются...

Клаузиус, по-видимому, был первым, кто был убежден в универсальном характере энтропии и полагал, что во всех процессах, происходящих во Вселенной, она играет важную роль, определяя их направление развития во времени.

В определенной мере, но уже с других позиций, к подобному выводу пришел И. Пригожин [13]. Пригожин полагает, что принцип энтропии ответственен за необратимость времени во Вселенной и, возможно, играет важную роль в понимании смысла времени как физического феномена. Имеется множество различных обобщений энтропии на случай неравновесных сред. Это целая большая область исследований. Читателям, интересующимся такого рода приложениями энтропии, можно порекомендовать книги И. Л. Пригожина [13–14], Р. Балеску [15–16], Г. Хакена [17], П. Резибуа и ДеЛенера [18], Ю. Л. Климонтовича [19–20], В. Эбеллинга [21] и др.

Имеется множество других исследований и обобщений энтропии, в том числе и в плане строгого математического подхода. Однако большая активность математиков в этой области пока не востребована в приложениях, за исключением, пожалуй, работ Колмогорова, Реньи и Тсаллеса. Математическая теория энтропии хорошо представлена в книге (Мартин и Ингленд, 1988) [23], которую можно рекомендовать заинтересованным читателям.

Говоря о прикладной математике, нельзя не упомянуть приложения принципа максимума энтропии. Хороший обзор применения различных вариационных принципов максимальной производительности энтропии в физике, геофизике, биологии и других областях можно найти в работе (Матюшев, Селезнев, 2006) [25].

Как уже отмечалось, важными областями приложений энтропии являются квантово-механические и релятивистские объекты. В квантовой физике и астрофизике такие приложения энтропии представляют большой интерес. Достаточно вспомнить один из красивых результатов термодинамики черных дыр: энтропия черной дыры равна четверти площади ее поверхности (площади горизонта событий) (Хокинг, 1978) [22]. В космологии энтропия Вселенной равна числу квантов реликтового излучения, приходящихся на один нуклон (Зельдович и Новиков, 1975) [24].

На наш взгляд, широкая популярность формул типа $S \sim k \sum p \log(p)$ в разных областях связана, возможно, с особой ролью степенных законов в природе [26]. Для таких случаев выражение $p \log(p)$ является удобной аддитивной мерой. Удобной в плане нашего восприятия количества чего бы то ни было — информации, хаоса, нелокальности, интенсивности света звезды и так далее. Ведь наши органы чувств работают по логарифмическому закону Вебера–Фехнера: интенсивность ощущения пропорциональна логарифму интенсивности стимула. С этой точки зрения, энтропия, выражаемая формулой типа $S \sim p \sum \log(p)$ есть всего лишь удобная для нашего восприятия абстрактная аддитивная мера. Ее физический смысл зависит от того, что именно мы оцениваем, то есть тем, на каких данных и как мы определяем аргумент p — плотность вероятности. Основание логарифмов несущественно и выбирается из соображений удобства для той или иной задачи.

Несколько слов о непрерывности, дискретности и предельных переходах. Переход к термодинамическому пределу в случае больших систем является необходимой и полезной процедурой. Но не нужно забывать, что, как пишет фон Нейман [27], «...непрерывность возникает лишь в результате процесса усреднения в прерывном, по существу, своему миру... и истинная природа единичного процесса оказывается полностью завуалированной все нивелирующим законом больших чисел». В реальном мире предельные переходы от дискретности к непрерывности не всегда возможны и часто даже противоестественны. Далее мы будем иметь дело с конечными наборами данных, и в этом плане основной интерес для нас будут представлять энтропии, не связанные с необходимостью предельных переходов.

Список литературы

1. Мандельброт Б. Б. Фрактальная геометрия природы / Под ред. В. А. Садовниченко и др. М.: Институт компьютерных исследований, 2002 г. 666 с.
(Книга прекрасно написана и хорошо иллюстрирована. Это лучшая книга для тех, кто хочет познакомиться с фракталами, что называется, из первых рук).
2. Колмогоров А. Н. Теория информации и теория алгоритмов. М.: «Наука», 1987. С. 39.
(Это третий том избранных трудов Академика А. Н. Колмогорова. Книга очень полезна для тех, кого профессионально интересует теорией информации).
3. Климонтович Ю. Л. Введение в физику открытых систем. М.: «Янус-К», 2002. С. 31.
(В книге кратко и в доступной форме изложены основные идеи, методы и результаты физики открытых систем. При всей краткости книга уникальна по широте охвата всех основных проблем самоорганизации и систематичности изложения. Основная ее идея, вынесенная в эпиграф, — фраза, приписываемая А. Н. Колмогорову: «Важно не то, что строго, а то, что верно»).
4. Шамбодаль П. Развитие и приложения понятия энтропии. М.: Наука, 1967.
5. Cercignani C. Ludwig Boltzmann: The Man Who Trusted Atoms. Oxford, 1998.
6. Гиббс Дж. В. Термодинамика. Статистическая механика: Избранные труды. М.: Наука, 1982. 584 с.
7. Шеннон К. Э. Работы по теории информации и кибернетике. М.: ИЛ, 1963. 829 с.
8. Колмогоров А. Н. Новый метрический инвариант транзитивных динамических систем и автоморфизмов пространства Лебега. ДАН СССР. Т. 119. 1958. С. 861–864.
9. Колмогоров А. Н. Об энтропии на единицу времени как метрическом инварианте автоморфизмов. ДАН СССР. Т. 124. 1959. С. 754–755.
10. Renyi A. Proc. Fourth Berkeley Symposium. Vol. 1. Berkeley, Los-Angeles: University of California Press, 1961. P. 547.
11. Tsallis C. Possible Generalization of Boltzmann-Gibbs-Statistics // J. Stat. Phys, 1988. Vol. 52. N 1/2. P. 479–487.

12. Tsallis C. Nonextensive Statistics: Theoretical, Experimental and Computational Evidences and Connections // Brazilian Journal of Physics, vol. 29, No. 1. March. 1999.
13. Пригожин И. От существующего к возникающему. М.: Наука, 1985. 328 с.
14. Пригожин И. Неравновесная статистическая механика. М.: Мир, 1964. 314 с.
15. Балеску Р. Равновесная и неравновесная статистическая механика. Т. 1. М.: Мир, 1978. 408 с.
16. Балеску Р. Равновесная и неравновесная статистическая механика. Т. 2. М.: Мир, 1978. 400 с.
17. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам. М.: Мир, 1991. 240 с.
18. Резебуа П., ДеЛенер М. Классическая кинетическая теория жидкостей и газов. М., Мир, 1980. 423 с.
19. Климонтович Ю. Л. Турбулентное движение и структура хаоса. Новый подход к статистической теории открытых систем. М.: Наука, 1990. 320 с.
20. Климонтович Ю. Л. Статистическая теория открытых систем. М.: ТОО «Янус», 1995. 624 с.
21. Эбелинг В. Образование структур при необратимых процессах. М.: Мир, 1979. 279 с.
22. Хокинг С. В. Черные дыры и термодинамика // сб. статей «Черные дыры». Новости фундаментальной физики. Вып. 9. 1978 г. С. 204–221. (там же другие его статьи).
23. Мартин Н., Ингленд Дж. Математическая теория энтропий. М.: Мир, 1988 г. 387 с.
24. Зельдович Я. Б., Новиков И. Д. Строение и эволюция Вселенной. М.: «Наука», 1975 г. 735 с.
25. Мартюшев Л. М., Селезнев В. Д. Принцип максимальности производства энтропии в физике и смежных областях. Екатеринбург: ГОУ ВПО УГТУ-УПИ, 2006. 83 с.
26. Шредер М. Фракталы, хаос, степенные законы. М.-И.: R&C Dynamics, 2001. 527 с.
27. Фон Нейман И. Математические основы квантовой механики. М.: «Наука», 1964. 367 с.

ГЛАВА 2

ЭНТРОПИЯ И ИНФОРМАЦИЯ

В различных прикладных задачах энтропия фигурирует в основном в структурном и информационном смысле. Это касается анализа временных рядов разной природы, области обработки спутниковой информации, в том числе астрофизических и геофизических изображений. Основной акцент в данной книге сделан именно на эту область приложений. Поэтому данную главу мы начнем с введения некоторых понятий теории информации, которые нам потребуются в дальнейшем. Далее, через теорему Хартли, введем понятие энтропии Шеннона. Рассмотрим без детальных доказательств ее основные свойства, знание которых потребуется нам в практических приложениях.

2.1. Информация как первичное понятие

Мы помним слова Л. Больцмана о том, что *энтропия* характеризует *недостающую информацию* о системе. То есть энтропия есть мера нашего незнания о системе. Но что имел в виду Л. Больцман, говоря о недостающей информации или о незнании системы? Ответ на этот вопрос лежит на поверхности. Вспомним физический смысл величины P , фигурирующей в формуле Больцмана $S = k \log(P)$. Параметр P , как мы помним, называется статистическим весом макросостояния системы и равен числу способов, которыми может быть реализовано данное микросостояние. Иными словами, мы не знаем, наблюдая данную макросистему, какой из этих P способов в ней реализован. Это и есть недостающая информация о системе. А что же тогда полная информация о системе? В данном случае это знание координат и скоростей всех частиц. В самом деле, если мы все это знаем, то мы знаем тот единственный способ, которым реализована наша макросистема на микроуровне. В этом случае $P = 1$, и энтропия $S = 0$.

А что означает на макроуровне знание координат и скоростей всех частиц? Это означает, что мы знаем, как распределены частицы системы в заданном 6 -мерном фазовом пространстве. Иными словами, мы знаем

структуру системы. Таким образом, Больцман был первым, кто неявно указал на тесную *связь энтропии, информации и структуры*.

В предыдущей главе был кратко описан путь, эволюции понятия «энтропия» почти за два столетия. Мы отметили некоторые важные этапы становления и развития этого фундаментального понятия современной науки. Энтропия играет важную роль в теплофизике, в статистической физике, в физической и химической кинетике, в биофизике, астрофизике, космологии и теории информации. Можно сказать, что понятие энтропии пронизывает все современное естествознание. В чем причина такой универсальности энтропии? Ведь по степени важности энтропию ставят в один ряд с такими фундаментальными понятиями как материя, энергия, время, пространство. Определенный ответ на этот вопрос дал великий американский инженер и ученый К. Шеннон. Он раскрыл тесную связь энтропии с другой фундаментальной сущностью нашего мира — информацией.

Информация, наряду с материей, энергией, пространством и временем, является первичным понятием нашего мира и поэтому в строгом смысле не может быть определена. Согласно современным представлениям, Вселенная возникла в результате Большого взрыва. Большой взрыв породил излучение, вещество, темную материю и космический вакуум. А откуда возникли законы, управляющие всем этим? Откуда первичная энергия узнала, как ей разделиться на кванты и породить излучение, вещество, галактики, звезды, планеты, весь растительный и животный мир и нас с вами? Была ли заложена в первичную энергию некая первичная информация, некий Код Вселенной, нечто подобное генетическому коду, по которому надлежало далее развиваться Вселенной? Иными совами, был ли Большой взрыв также и Большим информационным взрывом, или физические законы — правила игры полей и вещества — стали появляться по ходу самой игры?

Что представляют собой физические законы? Это наши собственные постулаты, нами самими придуманные с целью упорядочить наши представления об окружающем нас мире, а Вселенная о них ничего не знает? Или физические законы есть некий информационный атрибут Вселенной, столь же объективный, как вещество или энергия, и мы лишь открываем их для себя? Для ответа на эти вопросы надо бы больше узнать о происхождении, роли и способах функционирования информации во Вселенной. Иными словами, теория информации, по-видимому, должна быть тесно связана с космологией, астрофизикой, квантовой механикой, кинетической теорией и биофизикой.

Существующая теория информации является очень важным шагом на этом пути, и эта современная теория информации тесно связана с понятием энтропии. Чтобы обсудить эту связь, перечислим вначале основные свойства информации как первичного понятия.

- 1) **Дискретность.** Информация состоит из сообщений. Сообщение — это дискретный сигнал или последовательность дискретных сигналов, которые приходят из окружающего мира в рассматриваемую открытую систему и на которые эта система способна реагировать, то есть так или иначе изменять свое состояние. Например, квант света от звезды попадает в телескоп и изменяет состояние пикселя светоприемной матрицы. Другой пример: мы смотрим новости по телевидению — последовательность фреймов, попадая на сетчатку глаза, изменяет состояние световых рецепторов.
- 2) **Нематериальность.** Информация нематериальна, но она проявляется в связи с материальными носителями самой разной природы: букв, написанных на бумаге, клинописи на глиняных дощечках, иероглифах на папирусе, книг на деревянных дощечках, отпечатках древних растений и животных на камне, модуляции электромагнитной волны, последовательности аминокислот в генетическом коде и так далее. Иными словами, информация сама по себе без носителя, то есть без наличия той или иной структуры, невозможна. С другой стороны, как бессмыслица информация без структуры, так и всякая структура всегда несет в себе информацию. Более того, структура без информации невозможна.
- 3) **Закодированность.** Информация всегда закодирована: сообщения в виде знаков, символов, сигналов могут быть восприняты только получателем, способным их распознать. На этом свойстве информации основана вся криптография. Любой язык также есть код. Лишь люди, владеющие одним и тем же кодом, способны обмениваться информацией и понимать друг друга. Протоколы передачи данных по компьютерным сетям позволяют осуществлять обмен информацией между компьютерами. Такие примеры можно приводить бесконечно.
- 4) **Универсальность и объективность.** Это важнейшие свойства информации. Универсальность информации заключается в том, что это понятие одинаково применимо как к живому, так и к неживому миру. Примеры: рост кристаллов — зародыши кристаллов несут информацию молекулам, как им выстраиваться в опре-

деленном порядке; молекулы ДНК несут генетическую информацию о том, как строить организм. Информация объективна, то есть необязательно связана с наличием воспринимающего субъекта — она присутствует и функционирует в природе как объективный фактор.

- 5) **Креативность.** В основе любой структуры лежит информация, которая эту структуру порождает. Весь фрактальный анализ является иллюстрацией этого свойства — элемент и итеративное правило способны порождать сложные структуры. Мы знаем, что структуры возникают из хаоса вследствие фазовых переходов разного рода. Возможно, и информация, по которой строится структура, также возникает из хаоса? Опыт говорит, что это не так. Важно, из каких элементов состоит хаос. Конкретный хаос в конкретных условиях порождает конкретные структуры. Правила построения структур изначально заложены в элементах, из которых возникают структуры, а хаос лишь одно из возможных состояний среды.
- 6) **Аддитивность.** Это еще одно важное свойство информации. Это свойство попросту означает, что если к имеющейся информации мы добавляем новую, то оказавшаяся в нашем распоряжении информация равна сумме старой и новой.

Перечисленный перечень свойств информации как первичного понятия наверняка неполон. Есть, например, работы, в которых рассматривается такое свойство информации как ее ценность или значимость, обсуждаются и другие свойства. Но перечисленных свойств пока достаточно, чтобы двигаться дальше и, следуя за К. Шенноном, выяснить связь информации и энтропии.

Мы уже говорили, что *информация* наряду, с *массой*, *энергией*, *пространством* и *временем*, является первичным понятием нашего мира. Но равноправны ли эти понятия? Эквивалентность массы и энергии была показана Эйнштейном. Более того, стало понятно, что масса есть одна из форм состояния энергии. Была понята также относительность пространства и времени и их связь с массой и энергией. А как в этот ряд вписывается информация? Квантовая физика, похоже, установила эту связь, и она оказалась неожиданной: похоже, информация играет первичную и основную роль в современной физической картине мира.

Вот что по этому поводу писал Джон Арчибалт Уиллер. «Моя жизнь в физике представляется мне разделенной на три периода. В первый из них я был захвачен идеей, что „всё — это частицы“. Я искал способы вы-

строить все базовые элементы материи из наиболее фундаментальных частиц — электронов и фотонов.

Второй период я называю «всё — это поля». С тех пор как я влюбился в общую теорию относительности и гравитацию, я придерживался взгляда, согласно которому мир состоит из полей. Мир, в котором то, что представляется нам частицами, — это в действительности проявления электрических, магнитных, гравитационных полей и самого пространства–времени.

Теперь же я захвачен новой идеей: «Всё — это информация». Чем больше я размышляю о квантовых тайнах и о нашей странной способности постигать этот мир, в котором мы живем, тем очевиднее фундаментальное значение информации как основы физической теории».

2.2. Мера Хартли. Теорема Хартли

В основе шенноновской теории информации лежит представление о четырех исходных элементах: *передатчика*, *линии*, *приемника* и собственно *информации*. Информация по Шеннону дискретна. Атомом информации является сообщение. Сообщение — это некий сигнал, который (будучи эмитирован передатчиком), пройдя по линии, способен изменить состояние приемника. Например, звезда — передатчик, линия — космическое пространство, атмосфера, оптика телескопа, приемник — CCD матрица. Конечно, понятия «передатчик», «приемник», «линия», «сигнал» в теории информации не являются чисто техническими. Несмотря на то, что в их основе лежат представления о конкретных технических устройствах и процессах, они представляют собой абстрактно математические и даже в некотором роде философские категории.

Вообще говоря, чтобы найти решение какой-либо проблемы, очень важно суметь найти переход от конкретного к абстрактному. В результате такого перехода натянутый канат становится бесконечной и бестелесной прямой, кирпич или грубо обработанный камень — параллелепипедом, Юпитер со всеми его спутниками — материальной тяготеющей точкой. Абстрактные понятия, удерживая самые существенные для рассматриваемой проблемы свойства, дают возможность построить абстрактно математическую модель, позволяющую описать простыми средствами широкий круг явлений. Шеннону удалось это проделать с информацией.

Рассмотрим простые примеры. Разговаривают два человека. Каждый из них поочередно становится источником сообщений — передатчиком и приемником. Сообщения здесь — отдельные звуки (сигналы), «линия» — воздушная среда. Мы говорим, что эти люди обмениваются информацией. Обмениваться между собой информацией могут два или больше компьютеров, объединенных линиями связи в сети. Песня соловья для его самки тоже является очень важной информацией. Мы принимаем и анализируем свет звезды. Его интенсивность, спектральный состав, поляризация приносят нам важную информацию о массе и радиусе звезды, о ее химическом составе, температуре, наличию на ней магнитных полей. Таким образом, передатчиком информации потенциально является любая открытая система, а приемником — такая открытая система, которая, получив сигнал от первой, способна на него реагировать, то есть изменять свое состояние.

Если речь идет о какой-то конкретной системе, то с ростом информации об этой системе уменьшается наше незнание о ней. Мерой же нашего незнания системы является ее энтропия. Это, как мы помним, показал еще Больцман. Таким образом, энтропия и информация являются как бы двумя противоположностями, которые дополняют друг друга. Энтропия, это показал еще Клаузиус, также является аддитивной величиной. Вспомним интеграл Клаузиуса. Энтропия по Клаузиусу определяется через интеграл, то есть, по сути, через сумму. Аддитивность энтропии и информации, в частности, означает, что мера как той, так и другой, может быть только экстенсивной, а не интенсивной переменной.

Например, масса и энергия — экстенсивные переменные, их величины растут с ростом размеров системы. Температура и давление — примеры интенсивных величин. Они инвариантны относительно размеров системы. Температура одного кубического сантиметра воздуха в помещении такая же, как и температура одного кубического метра. Энергия же и масса 1 см^3 и 1 м^3 воздуха различны. Соответственно, различны и энтропии этих объемов воздуха.

Именно такую аддитивную меру для информации впервые предложил и достаточно строго обосновал Хартли в 1928 г. (Хартли, 1959) [1]. Рассмотрим этот вопрос более подробно. Пусть имеются два источника сообщений с числами *равновероятных* состояний m и n . Эти два источника можно рассматривать как один, с числом состояний, равным, соответственно, $m \cdot n$. Желательно, чтобы мера неопределенности H такого источника равнялась сумме неопределенностей двух исходных, то есть $H(m \cdot n) = H(m) + H(n)$. Можно строго показать, что единственной такой функцией

является логарифм. Если $N = m \cdot n$, то $H(N) = \log(N) = \log(m) + \log(n)$, при этом основание логарифмов особой роли не играет.

Хартли сформулировал и доказал следующую важную теорему. «Если в заданном множестве $\{M\}$, состоящем из N элементов, содержится элемент x , о котором известно только, что он принадлежит этому множеству $\{M\}$, то, чтобы найти x , необходимо получить об этом множестве количество информации, равное $\log_2 N$ бит⁴».

Теорема доказывается с помощью так называемого алгоритма Бар-Кохбы⁵ и метода индукции. Полное доказательство было дано А. Реньи (Renyi, 1970) [2]. Здесь мы продемонстрируем лишь интуитивно понятную общую идею этого доказательства. Рассмотрим ситуацию с двумя равновероятными исходами. Пусть, например, нас интересует, где у упавшей монеты будет находиться орел — сверху или снизу. До испытания равновероятны оба исхода, однако одно испытание полностью устраняет эту неопределенность и отвечает на наш вопрос.

Пусть теперь мы рассматриваем некоторую ситуацию с четырьмя равновероятными исходами. Четыре исхода делим на две равные части и задаем первый вопрос: где тот исход, который нас интересует? Пусть он в правой части. Ее снова делим пополам и задаем второй точно такой же вопрос. Ответ на этот второй вопрос полностью устраняет исходную неопределенность. Аналогично рассматриваем ситуацию с 8-ю равновероятными исходами и приходим к выводу, что достаточно получить ответы на три вопроса, чтобы устранить неопределенность. Принимаем, что количество информации в одном ответе на вопрос типа «да – нет» равно одному биту. Согласно методу индукции заключаем, что число ответов H и число равновероятных исходов N связаны друг с другом как $N = 2^H$, или $H = \log_2 N$ бит. Теорема Хартли доказана. Как уже отмеча-

⁴ Название БИТ (BIT — аббревиатура от Binary digiT) было предложено американским математиком Джоном Тьюки в 1946 г. Тьюки избрал бит для обозначения одного разряда в двоичной системе исчисления. Хартли и Шеннон использовали бит как единицу измерения информации. Теперь, как мы знаем, это одно из главных понятий нашего времени.

⁵ Бар Кохба — один из руководителей израильских повстанцев против римлян (132–135 гг.). По преданию Бар Кохба послал своего лазутчика в стан римлян, чтобы выяснить, около какого горного прохода римляне сосредоточили свои войска. Лазутчик все выяснил, но был схвачен. Для устрашения римляне отрезали ему язык и руки и отправили обратно к Бар Кохбе. Все же Бар Кохба получил необходимую ему информацию. Он нарисовал ему карту местности и, поделив ее пополам, попросил указать кивком головы ту половинку, где сосредоточены войска. Затем эту половинку карты он снова делил пополам, и снова та же процедура. Таким образом Бар Кохба точно выяснил, где именно стоят войска противника.

лось, это лишь общая идея доказательства. Фактически же, теорема доказана для произвольных N . Полное доказательство можно найти в цитированных книгах Р. Хартли и А. Реньи.

Величину

$$H = \log_2 N \quad (2.1)$$

называют иногда информацией Хартли. На наш взгляд, это не совсем правильно. Информацией она становится только после проведения всех H экспериментов. Более правильно ее называть энтропией, потому что она выражается формулой, аналогичной формуле для энтропии Больцмана, от которой она отличается только размерным множителем k и основанием логарифмов. А также потому, что дает меру нашего незнания о состоянии системы до того, как мы начинаем экспериментировать с этой системой.

Еще раз отметим этот важный момент. Информация, которую необходимо получить, чтобы устранить неопределенность ситуации, равна двоичному логарифму числа *равновероятных* состояний системы. Но эта информация появляется только после необходимого числа испытаний. А до опыта эта величина является мерой неопределенности нашей системы и называется ее энтропией. Убыль этой величины в результате испытаний и есть та информация о системе, которой мы владеем на данный момент. Иными словами, информация

$$I = H_1 - H_2, \quad (2.2)$$

где H_1 и H_2 — энтропии до и после испытаний соответственно. Это и есть знаменитое Шенноновское определение количества информации.

На бытовом уровне все мы пользуемся словом «информация», связывая его с определенными представлениями. В обиходе под информацией обычно понимают совокупность сведений об окружающем мире. Информация является объектом хранения, передачи и преобразования. Заслуга Шеннона заключалась в том, что он сумел в соответствии бытовому термину поставить четкое определение, позволившее количественно оценивать объемы информации. Ключевым при этом, как уже было отмечено во введении, является понятие энтропии. Можно сказать, что Шеннон осуществил энтропийную параметризацию информации. Как уже отмечалось, еще Л. Больцман, давая статистико-физическое определение энтропии в 1877 году, заметил, что энтропия характеризует недостающую информацию о системе. Тогда эта фраза осталась без внимания.

Однако теперь, благодаря Шеннону, мы можем уже количественно оценить как имеющуюся, так и недостающую информацию о системе.

Следует отметить, что шенноновское определение информации удовлетворяет, или, по крайней мере, не противоречит, шести перечисленным свойствам информации как первичного понятия. Она дискретна, аддитивна, универсальна и объективна, нематериальна. Определение информации по Шеннону не содержит свойства креативности, но и не противоречит ему, поскольку по Шеннону информация определена через энтропию. Энтропия же характеризует состояние системы, а не ее «элементную базу» — те дискреты, из которых эта система построена. Так же и информация Шеннона характеризует меру нашего знания устройства системы и не содержит сведений о первичной информации, заключенной в элементах, из которых эта система построена.

2.3. Энтропия систем с неравновероятными состояниями

Системы, для которых Хартли ввел свою меру неопределенности, являются системами с равновероятными состояниями. Однако в природе такие системы представляют очень узкий класс и являются, скорее, исключением, чем правилом. Игральные кости, монеты, различные лототроны, игровые автоматы, датчики случайных чисел — примеры таких систем. Все они в основном искусственного происхождения. Это не случайно. Природа *не любит* равновероятные системы и практически никогда их не реализует.

Обобщение понятий энтропии и информации на случай систем с неравновероятными состояниями, а именно такие системы и представляют основной интерес, получается достаточно просто. Пусть i -е состояние системы (отсчет ряда, символ в сообщении) имеет вероятность $p_i = N_i/N$, где N — объем выборки (объем или число членов ансамбля в термодинамике, число отсчетов в ряду, число символов в сообщении), N_i — число заполнения (частота встречаемости в сообщении) i -го уровня. Так что

$$N = \sum_{i=1}^n N_i, \quad \sum_{i=1}^n p_i = 1, \quad (2.3)$$

n — число возможных уровней состояния системы (число символов в алфавите).

Для произвольно временного ряда $f(t)$ n задается искусственно и равно числу уровней амплитуды ряда (величины сигнала) в данной выборке: $n = (\max(f(t)) - \min(f(t))) / df$, где df — искусственный «квант» — обычно величина, не меньше точности измерения амплитуды $f(t)$.

Для k -мерного объекта n равно числу минимальных k -мерных кубиков размера e , целиком покрывающих данный объект. В этой связи множество k -мерных кубиков будем называть покрытием, e — элементом данного покрытия, а n — числом элементов покрытия. Здесь и далее термин «покрытие» употребляется в узкоспециальном смысле некоторых задач комбинаторной геометрии (см., например, [3], [4]).

Пусть P — статистический вес системы при заданном e покрытии, то есть полное число *равновероятных* способов, которыми может быть реализовано рассматриваемое состояние системы (эту величину называют еще абсолютной или термодинамической вероятностью состояния системы). Эта величина равна числу всех значимых перестановок в системе, то есть таких, когда ячейки покрытия обмениваются между собой элементами. Число перестановок всех элементов ансамбля (отсчеты выборки из ряда) равно $N!$. Однако перестановки внутри элемента покрытия (внутри одного уровня) очевидно, не дают новых ситуаций и не учитываются. Таким образом, система может реализовать данное состояние числом способов (число перестановок с повторениями), равным

$$P = N! / (N_1! N_2! \dots N_n!). \quad (2.4)$$

Отсюда определяем так называемую абсолютную энтропию Ha . Учитывая теорему Хартли, получаем:

$$Ha = \log_2 P = \log_2 (N! / (N_1! N_2! \dots N_n!)), \quad (2.5)$$

воспользовавшись формулой Стирлинга, получим

$$Ha \sim N \log_2 N - \sum N_i \log_2 N_i = -N \sum p_i \log_2 p_i \text{ бит.} \quad (2.6)$$

Удельная энтропия, то есть в расчете на одно состояние (на удельный объем в термодинамике, один отсчет в ряде), равна:

$$H = - \sum_{i=1}^n p_i \log_2 p_i. \quad (2.7)$$

Это и есть известная формула Шеннона, или энтропия Шеннона для систем с неравновероятными состояниями.

Важно отметить, что порядок элементов покрытия, как видим из (2.7), не влияет на величину H , то есть мы можем поменять местами любое число элементов покрытия, но значение H останется прежним.

Проясним и некоторые другие важные для приложений следствия формулы (2.7). Представим себе систему, которая все время находится в одном и том же состоянии, то есть все p_i , кроме одного, равны нулю. Это означает, что система полностью детерминирована, и всякая неопределенность в ней отсутствует. В этом случае $p = 1$ и $H = 0$. Таким образом, чем меньше неопределенности в системе, чем меньше состояний она способна реализовать, тем меньше энтропия. Для случая временного ряда $f(t)$ это соответствует $f(t) = \text{const}$.

Из формулы Шеннона также видно, что максимально возможное значение энтропия принимает в случае, когда все возможные состояния равновероятны (белый шум), то есть все уровни заполнены равномерно. Для случая временного ряда $f(t)$ это соответствует белому шуму, пилообразной функции или, самое простое, линейной зависимости: $f(t) = at$ на интервале выборки. В этом случае из формулы следует, что $H = \log_2 n$, где n — число элементов покрытия (в случае ряда — уровней). В этих конкретных границах и может изменяться энтропия в данной конкретной системе с данным фиксированным числом состояний n .

Большое значение имеет знак изменения энтропии со временем. Если $dH/dt > 0$, то система релаксирует, то есть стремится к своему наиболее вероятному состоянию с максимумом энтропии. В случае временного ряда это стремление к белому шуму. Если же $dH/dt < 0$, то система находится на фазе самоорганизации, в ней возникает и развивается некая упорядоченность. В случае временного ряда это стремление к $f(t) = \text{const}$. Если же $dH/dt \sim 0$, то система находится в фазе стагнации или/и, возможно, бифуркации. Иными словами, определив знак и величину изменения энтропии, мы можем судить о темпах, фазе и направлении эволюции системы в текущий момент.

Рассмотрим простой пример расчета величины энтропии. Пусть наш временной ряд представляет собой функцию Радемахера: $Ra(t) = \text{sign}(\sin(t))$. Чему равна энтропия такого ряда? Легко видеть, что ряд имеет лишь два равновероятных уровня. Следовательно, $p_i = 1/2$. Подставляя это значение p_i в формулу (2.7), получаем $H = \log_2 2 = 1$.

2.4. О принципе максимума энтропии

В литературе часто можно встретить утверждение, что энтропия Шеннона это совсем не то же самое, что термодинамическая энтропия или энтропия Больцмана. Более того, энтропия Шеннона ведет себя прямо противоположным образом, стремится со временем не к максимуму, а к минимуму. При этом обычно авторы ссылаются на 7-ю теорему Шеннона [5], утверждающую, что энтропия выходного сигнала не возрастает при любом преобразовании. Таким образом, информационная (2.7) и термодинамическая (1.8) энтропии при внешней похожести формул ведут себя противоположным образом, что при применении аппарата теории информации к процессам в физических системах может приводить к вышеупомянутому утверждению [6].

Вопрос важен для прояснения смысла энтропии вообще, поэтому рассмотрим его более подробно. Для этого нам необходимо вспомнить принцип максимума производства энтропии. *Принцип максимума энтропии* можно сформулировать так. *При отсутствии внешних воздействий, когда множество возможных микросостояний n зафиксировано, система переходит в такое макросостояние, при котором все микросостояния равновероятны.* Математически это означает выход на наибольшую однородность распределения вероятностей микросостояний при фиксированной системе физических связей [7].

Так ведут себя все замкнутые термодинамические системы. Именно этот факт, как факт эмпирический, составляет смысл второго начала термодинамики. В наших обозначениях мы запишем второе начало так:

$$S = -\sum_{i=1}^n p_i \log_2 p_i \Rightarrow \log_2 n, \quad (2.8)$$

где S — энтропия Больцмана.

Однако 7-я теорема Шеннона утверждает, и это тоже эмпирический факт, что энтропия выходного сигнала не возрастает при любом преобразовании, более того, стремится к минимальному значению. То есть 7-я теорема Шеннона в наших обозначениях может быть записана так:

$$H = -\sum_{i=1}^n p_i \log_2 p_i \Rightarrow 0. \quad (2.9)$$

Таким образом, при идентичности формул термодинамической и информационной энтропий постулируется прямо противоположный вывод. Суть дела заключается в разном определении вероятностной меры p_i в физике и теории информации. Во всех случаях p_i соответствует относительному числу дискретов системы, попавших в i -тый элемент покрытия. Однако в физических системах, как правило, дискретами являются изображающие точки фазового пространства, и покрытие производится на этом пространстве. В теории информации дискретами являются сигналы («буквы алфавита»), и распределения «букв» рассматриваются по частоте их встречаемости в сообщении. Здесь p_i имеет смысл относительной частоты встречаемости в сообщении данной «буквы». Это существенно разные определения вероятностных пространств. Рассмотрим эту разницу на примере временного ряда.

В случае временного ряда физический подход соответствовал бы распределению величин амплитуд $f(t)$ ряда по *временным интервалам*. В этом случае $f(t) = \text{const}$ дает нам максимальную энтропию, что полностью соответствует второму началу. Если же на выделенном временном интервале выборки имеем только один δ -импульс, то энтропия такой выборки равна нулю.

В теории информации p_i имеют смысл распределений отсчетов ряда по *амплитудным уровням*. В этом случае $f(t) = \text{const}$ соответствует нулевой энтропии, а δ -импульс, так же, как и рассмотренный ранее линейный тренд или пила, дает максимальное значение энтропии. Таким образом, при стремлении $f(t) \rightarrow \text{const}$ поведение энтропии «противоположно» второму началу термодинамики. Во избежание путаницы этот факт нужно всегда иметь в виду при применении понятий теории информации к физическим системам. Важно отметить, что белый шум во всех случаях соответствует максимальной энтропии.

И еще одно важное замечание. Следует всегда помнить, что формула Шеннона (2.7) всего лишь алгоритм для расчета величины H . Смысл и значения этой величины целиком определяются выбором пространства состояний и характеристиками его покрытия.

Подведем краткий итог. Число равновероятных состояний системы N можно рассматривать как число возможных реализаций одного и того же состояния системы. Тогда N тождественно больцмановскому статистиче-

скому весу P , то есть числу таких реализаций. Такое переосмысливание параметра N позволяет обобщить энтропию Хартли на случай не равновероятных состояний системы. Как обычно подсчитывают число возможных реализаций состояния системы? Для этого вводится понятие e -покрытия системы (вспомним алгоритм Бар Кохбы). Минимально допустимой величиной e -элемента покрытия обычно является предел точности измерений параметров. Каждой такой e -ячейке ставим в соответствие число $p_i = N_i/N$, где N_i — число встречаемости значений параметра в пределах i -той e -ячейки покрытия, p_i — относительное число значений параметра, попавших в i -ю ячейку, которую мы интерпретируем как вероятность обнаружения значения параметра в пределах i -того элемента покрытия. Распределение $p_i(i)$ может быть любым. Сумма всех p_i равна 1 по определению.

Для неравных p_i число $P = N!/N_1!N_2!\dots N_n!$, где n — число элементов покрытия. Подставляя P в выражение (2.1) вместо N , получаем после несложных преобразований выражение для удельной энтропии Шеннона (2.7). Из этого выражения следует, что максимальное значение энтропии Шеннона достигается, когда все p_i равны, то есть все состояния системы равновероятны. При этом значение $H_{\max} = \log(n)$, где n — число элементов покрытия. Если же все p_i , кроме одного, равны 0, то есть система полностью детерминирована, то энтропия Шеннона достигает минимального значения $H_{\min} = 0$.

Список литературы

1. Хартли Р. В. Л. Передача информации / Теория информации и ее приложения / пер. с англ. / под ред. А. А. Харкевича. М.: Физматгиз, 1959 г.
2. Renyi A. Probability Theory. Amsterdam: North-Holland Publ. Co. 1970. 573 p.
3. Болтянский В. Г., Гогберг И. Ц. Теоремы и задачи комбинаторной геометрии. М. Наука, 1965. 109 с.
4. Болтянский В. Г., Гогберг И. Ц. Разбиение фигур на меньшие части. М.: Наука, 1971.
5. Шеннон К. Э. Работы по теории информации и кибернетике. М.: Ил., 1963–829 с.

6. Коротаев С. М. Энтропия и информация — универсальные естественнонаучные понятия. <http://www.chronos.msu.ru>
7. Мартюшев Л. М., Селезнев В. Д. Принцип максимальности производства энтропии в физике и смежных областях. Екатеринбург: ГОУ ВПО УГТУ-УПИ, 2006. 83 с.

ГЛАВА 3

ЭНТРОПИИ ВРЕМЕННЫХ РЯДОВ И ВИДЕОИЗОБРАЖЕНИЙ

3.1. Общие замечания о временных рядах

В практике обработки данных *под временным рядом* обычно понимают упорядоченную последовательность пар измеренных величин, одна из которых является временем. Вторая величина может иметь самую разную природу и размерность.

По характеру природы этой второй величины различают *скалярные, обобщенно-скалярные, векторные, тензорные* и *логические* ряды. Например, ежедневные средние значения радиопотока от Солнца на волне 10.7 см — пример *скалярного* временного ряда. Все ряды гелио- и геофизических индексов тоже скалярные ряды. Если каждому моменту времени ставится в соответствие два, три и более параметров, то такую последовательность можно рассматривать как *обобщенно-скалярный* ряд. Биржевые ряды являются примерами рядов такого типа. Электрокардиограммы и энцефалограммы — другие примеры подобных рядов. Набор этих скалярных параметров, являющихся компонентами вектора, называется *векторным* рядом: например, три компоненты вектора скорости тела как функции времени. Если каждому моменту времени ставится в соответствие тензор (матрица значений), то это *тензорный* ряд. Чернобелые и цветные фильмы в цифровых форматах можно рассматривать как примеры тензорных рядов. Если имеем упорядоченный по времени ряд логических значений, то это, соответственно, *логический* ряд. Примером таких рядов могут быть ряды событий, амплитуда которых количественно не определена или не имеет особого значения, а важным является лишь сам факт наличия или отсутствия события. Например, последовательность дорожно-транспортных происшествий, смертельный травматизм на производстве и так далее.

Понятно, что ряд любой размерности можно рассматривать как определенную композицию скалярных рядов. По этой причине мы в основном будем говорить о скалярных рядах. Обобщение на случай большей размерности достаточно тривиально, если не приходится решать какие-либо специальные задачи.

По характеру измерения времени, можно говорить об эквидистантных, неэквидистантных и автономных рядах.

Эквидистантным называется ряд измерений с фиксированным временным шагом. Это самые комфортные с точки зрения обработки данных ряды. В последние годы в общем объеме астрофизических и геофизических данных такие ряды приобретают все больший вес в связи с появившейся возможностью вести непрерывные и регулярные наблюдения с борта космических аппаратов. Некоторые разделы астрофизики, например, гелиосейсмология, вообще не могут развиваться без получения и анализа таких рядов. В связи с необходимостью получения длинных эквидистантных рядов строятся глобальные сети станций наблюдения, обсерватории в Арктике и Антарктике, разрабатывается и выводится в космическое пространство соответствующая аппаратура. Важную роль такие ряды данных играют в геофизике — метеорологии, геомагнетизме и других областях.

Если временной шаг не фиксирован, то ряд, соответственно, называется *неэквидистантным*. Такие ряды возникают, когда по тем или иным причинам процедуру измерений не имеет смысла или невозможно проводить с фиксированным временным шагом. В астрофизике в таких рядах до начала ширококомасштабных космических наблюдений содержался практически весь объем данных. Эти ряды более трудны в обработке, так как часто требуют дополнительных процедур, таких как, например, интерполяция, которые вносят дополнительные ошибки.

Необходимо отличать неэквидистантные ряды от *автономных*. Неэквидистантный ряд порождается спецификой измерительного процесса, в то время как между измерениями процесс продолжается, и измеряемые параметры могут изменяться на этих ненаблюдаемых интервалах как угодно. Можно сказать, что эквидистантные и неэквидистантные ряды есть лишь некие выборки значений параметров изучаемых процессов.

Другое дело *автономные* ряды. Для таких рядов само понятие времени приобретает иной, специфический смысл. Как правило, это ряды некоторых событий или специфических процессов в квазизамкнутых системах, когда фиксируются лишь изменения параметров, а промежутки времени без изменений невозможно определить, или они просто теряют смысл и не рассматриваются. С точки зрения внешнего наблюдателя, пользующегося астрономическим временем, эти процессы могут происходить крайне неравномерно, но в пределах самой системы точка зрения внешнего наблюдателя и специфика его времени не играют особой роли.

Например, в биржевых рядах, как правило, важно изменение цены и не очень важны величины промежутков времени между этими изменениями: цена не меняется, значит, на рынке ничего не происходит. *Геоло-*

гические ряды, являются другим примером автономных рядов. Вообще, любая квазизамкнутая система живет как бы в своем собственном специфическом времени, слабо связанно с тикающими часами, отсчитывающими наше астрономическое время. В этом плане говорят о *геологическом* времени, *биологическом* времени, *психологическом* времени, *космологическом* времени и так далее. Можно сказать, что каждая квазизамкнутая система сама определяет для себя собственное время и живет по собственным часам, то есть образует как бы свой собственный темп-мир [5]. Время системы обусловлено последовательностью изменений, происходящих в ней самой⁶. Автономный ряд — это такой ряд, в котором принимаются во внимание и измеряются лишь изменения параметров, а внешнее, астрономическое время при этом по тем или иным причинам не фиксируется. Иными словами, *автономный ряд* — это ряд упорядоченных значений изменения переменной. Сама переменная при этом может иметь любую природу.

Безусловно, деление всех временных рядов на регулярные и автономные в определенном смысле является условным. Так, если возможно значениям автономного ряда поставить в соответствие моменты астрономического времени, то такой ряд становится обычным неэквидистантным рядом. А если ввести фиксированный шаг по времени и соответствующим образом дополнить множество значений, то автономный ряд превращается в эквидистантный. Процедуру установки соответствия астрономического времени значениям автономного ряда обычно называют *привязкой* или *датировкой*. Эти часто очень сложные, неоднознач-

⁶ В философии идея смысла времени как последовательности событий восходит, по видимому, к Лейбницу, который полагал, что само восприятие времени порождается изменениями, и если нигде никакие изменения не происходят, то говорить о времени бессмысленно — времени нет. Намеки на такую природу времени можно найти и у античных авторов. Похожей точки зрения на природу времени придерживался В. И. Вернадский (Вернадский, 1975) [6], который полагал, что характеристики пространства–времени тесно связаны с рассматриваемой системой и процессами, происходящими в ней. В этом плане пространство–время Общей теории относительности Вернадский рассматривал лишь как частный случай такого общего подхода, когда объектом рассмотрения является такая система как Вселенная. В качестве противоположной точки зрения на природу времени обычно рассматривают представления Ньютона о времени. Здесь время рассматривается как первичный атрибут мира, который представляет собою вместилище событий, подобно тому, как пространство является вместилищем объектов. Известно в этом плане высказывание А. Пуанкаре о том, что наука не изучает время, она изучает проявление природных процессов в ходе времени, которое от этих явлений абсолютно не зависимо. Мы не станем здесь углубляться в эти вопросы. Отметим лишь, что на наш взгляд, эти представления есть лишь модели одного, но многоликого феномена времени, и подобно корпускулярно-волновому дуализму в оптике, в большей или меньшей степени применимы в той или иной конкретной ситуации (см., в том числе, Приложение).

ные и трудоемкие процедуры широко распространены в геологии, палеонтологии, археологии и других областях.

Для интерпретации результатов выборочных энтропий важно отличать неэквидистантные ряды от автономных. Как уже отмечалось, истинные неэквидистантные ряды порождаются процессом измерения и являются, по сути, некоторой случайной выборкой значений параметра исследуемого процесса. Это, в частности, означает, что такой ряд лишь частично, не полно, представляет рассматриваемый процесс. В то время как автономные ряды являются результатом фиксации реально происходящих изменений и не являются произвольной выборкой значений параметра. Можно считать, что процессы, формирующие такие ряды, обладают нерегулярным и, часто, *фрактальным* временем. Соответственно, значения энтропий таких рядов должны интерпретироваться по-разному.

Еще временные ряды делят **по характеру изменения амплитуды** на *регулярные, стохастические и смешанные*, а также на *стационарные и нестационарные*. Всякий ряд порождается некоторым процессом. Регулярные механические движения порождают регулярные ряды. Например, измерение расстояния между центрами Земли и Луны дает регулярный ряд. Случайные стохастические процессы порождают, соответственно, стохастические ряды, (например, изменение скорости течения в турбулентном потоке, броуновское движение и др.). Смешанные ряды порождаются суперпозицией регулярных и случайных процессов. Например, вариации освещенности участка поверхности Земли порождаются регулярными процессами суточного вращения и движения Земли по орбите вокруг Солнца и случайными изменениями прозрачности атмосферы над этим участком.

Резюмируя все вышесказанное, стоит отметить, что временные ряды измеряемых величин разной природы и разных размерностей, с которыми имеют дело исследователи разных предметных областей, как правило, порождаются нестационарными процессами. В таких рядах обычно присутствуют как квазирегулярные, так и стохастические компоненты. Кроме того, в силу условий и/или специфики измерений такие ряды часто имеют разрывы, поведение измеряемых параметров внутри которых может быть, в принципе, каким угодно.

3.2. О параметризации

Вначале несколько слов о процедуре параметризации. *Параметризация* — это извлечение из данных наблюдений минимального набора наи-

более существенных параметров, характеризующих изучаемую систему или процесс. Фактически, параметризация — это редукция (свертка) массивов данных наблюдений к набору небольшого количества параметров и правил. Правильно проделанная параметризация позволяет восстановить исходные данные с заданной точностью по полученному набору этих правил и параметров.

Наиболее популярными методами параметризации являются интегральные преобразования данных с помощью полных систем ортонормированных тригонометрических, специальных, обобщенных функций или специально подобранных локализованных функций — вейвлетов. Это преобразования Фурье, Лежандра, Уолша, другие подобные преобразования и вейвлет-анализ. Результатом таких процедур является получение некоторого конечного набора амплитуд базовых функций, *суперпозиция* которых с заданной точностью представляет исходные данные.

Параметризация временного ряда — это, вообще говоря, процедура выбора такого набора параметров, которые позволяют нам количественно охарактеризовать изменение системы со временем, понять происходящие в ней процессы и, в конечном итоге, построить ее теоретическую модель. Параметризация играет очень важную роль в получении новой информации о любых системах или процессах. Она тесно связана с общетеоретическими представлениями об изучаемых процессах в данной системе и часто диктует основные положения экспериментальных и наблюдательных программ.

Исходным материалом для исследования физических процессов являются временные ряды. Примерно до середины прошлого века господствовала достаточно упрощенная, «наивная», точка зрения на временные ряды данных. Считалось, что любой ряд содержит лишь общий тренд, регулярную и случайную компоненты, параметры которых не меняются. Причем тренд можно было рассматривать как фрагмент регулярного движения с очень большим периодом. Методы анализа рядов сводились к выделению регулярных составляющих из ряда таким образом, чтобы в остатке оказался «белый шум» (так называемое «выбеливание» ряда). Далее определялись параметры регулярной составляющей (периоды, амплитуды) и стохастической (моменты распределений). После этого считалось, что параметризация выполнена, и вся значимая информация из ряда извлечена.

Такая точка зрения молчаливо предполагала, что любой временной ряд является непрерывным, стационарным, линейным, то есть может быть представлен с заданной точностью суперпозицией тех или иных гладких функций, а случайная компонента (ошибки измерений) является стационарным гауссовым шумом.

Сейчас понятно, что линейными и стационарными, в приемлемом приближении, можно считать лишь небольшую часть всех временных рядов. Как правило, это ряды, порождаемые динамическими процессами в простых системах. Например, изменение расстояния от Земли до Солнца, смена фаз Луны, фотометрические кривые широких затменно-переменных звезд и другие. Большинство же анализируемых временных рядов, как правило, нестационарно и нелинейно. То есть параметры регулярных и стохастических компонент могут быть произвольными функциями времени. Кроме того, ряды часто могут содержать и непериодические компоненты. Как ведет себя ряд в интервалах между отсчетами, не всегда понятно. Также далеко не всегда очевидно, представим ли он в частях и в целом гладкими функциями. Нестационарность, стохастичность, возможная негладкость накладывают серьезные ограничения на многие традиционные методы анализа и стимулируют поиск новых подходов. Одним из таких подходов, применимых к слабо нестационарным рядам, является выборочная фрактальная параметризация, то есть параметризация выборок с помощью негладких масштабно инвариантных функций. Этот подход будет обсуждаться в следующей главе.

А сейчас несколько слов о «философии эксперимента», который тесно связан с проблемой параметризации. Обычно состояние системы характеризуется набором параметров ее состояния. Для простой системы, например, для газа, близкого по своим характеристикам к идеальному, такими параметрами могут быть объем, давление и температура. В данном случае набор из этих трех составляющих является «полным» набором, так как полностью определяет состояние идеального газа. Если исследуемая система сложная, то характер и число этих параметров существенно зависит от того, что мы хотим узнать о системе, то есть от постановки задачи.

Рассмотрим примеры, когда одни и те же объекты можно характеризовать различным набором параметров в зависимости от поставленной задачи. Георгий Гамов в своей биографической повести «Моя мировая линия» выбрал в качестве параметров своей биографии географические координаты и время своего пребывания в тех или иных местах земного шара. Конечно, это только литературный прием, но для физика подход вполне естественный – параметризовать процесс, упорядочить сложные перипетии судьбы с помощью всех трех параметров. Конечно же, это пример параметризации сложного процесса, хотя и шуточный.

Продолжая шуточный подход Г. Гамова, заметим, что его параметризация продиктована чисто *механическим* подходом к проблеме. Но можно подойти к этому вопросу с физиологической точки зрения и характеризовать объект «Гамов» набором таких, например, параметров как

вес, рост, артериальное давление, температура тела, процент C_2H_5OH в крови и так далее. Набор таких биометрических и физиологических параметров также будет описывать наш объект, только уже в рамках другой задачи; они могут быть интересны его лечащему врачу.

Но если Г. Гамов интересен нам прежде всего как ученый, то все эти параметры не годятся. В этом плане наибольший интерес представляют такие параметры как, например, число его научных публикаций за год, число ссылок на его публикации за год, число публикаций о нем и его работах. Условно этот подход назовем *информационным*. Именно такой информационный подход представляется наиболее существенным и интересным, так как именно способность Г. Гамова к переработке и генерированию новой научной информации составляет основную его ценность с точки зрения социума.

Рассмотрим другой объект — звезду. Мы можем в качестве параметров состояния звезды рассматривать ее положение в галактике и ее полный вектор скорости. Это очень важные параметры звезды. С их помощью мы можем найти траекторию движения звезды, узнать, откуда пришла эта звезда и где она окажется в тот или иной момент в будущем. Как и в случае мировой линии Г. Гамова, это чисто механический подход.

Если же нас интересует звезда как астрофизический объект, то больший интерес будут представлять такие ее параметры как светимость, спектр, масса. Нас будет интересовать, в первую очередь, положение звезды на диаграмме Герцшпрунга–Рассела, а не ее положение в галактике. Конечно, эти два подхода дополняют друг друга, как и механический и физиологический подходы в случае объекта «Гамов», но тем не менее, это два разных научных подхода к одному и тому же объекту.

Возможен ли в этом примере со звездой информационный подход? На данном этапе развития науки (а возможно, и вообще) информационный подход к звезде неприменим. Это связано с тем, что все известные нам устройства приема, передачи и обработки информации обычно являются сложными высокоструктурированными системами, начиная от самых примитивных форм жизни и заканчивая человеческим мозгом, состоящим из миллионов нейронов, или компьютерами, состоящими из множества интегральных схем. Чтобы работать с информацией, звезда должна обладать, как минимум, очень сложной внутренней структурой, способной принимать, обрабатывать и генерировать информацию. Насколько сейчас известно, ни чем подобным Солнце и звезды не располагают. Увы, выражение «звезда с звездой говорит» остается только поэтической метафорой.

К вопросу о параметризации в данной работе мы будем периодически возвращаться, рассматривая проблему на разных уровнях и с разных точек зрения. Сейчас несколько слов об *идентификации* сложных систем. Вопросы идентификации возникают в том случае, когда ряд продуцируется процессом не известной или не до конца изученной природы. Например, ряды интегральных индексов солнечной активности продуцируются некоторыми не до конца понятыми магнитогидродинамическими (мгд) процессами.

Идентифицировать систему или процесс по данным измерений — значит построить математическую модель временного поведения системы и, по возможности, выяснить, какие физические процессы продуцируют данный ряд. С математической точки зрения это задача аппроксимации данных некоторым уравнением, или системой уравнений. В процессе идентификации возникают вопросы *полноты* набора параметров:

- сколько параметров необходимо и достаточно для идентификации объекта;
- насколько информативен ряд данных для идентификации системы;
- проблемы связи и перехода между измеряемыми переменными и переменными математической модели и другие вопросы.

3.3. Выборочные энтропии

К анализу временных рядов произвольной природы, кроме известных традиционных методов, можно подойти путем их энтропийной и/или фрактальной параметризации. Примером энтропийной параметризации является расчет энтропии Колмогорова, о которой мы уже упоминали в первой главе. Энтропия Колмогорова изначально вводилась для анализа стохастических временных рядов, порождаемых динамическими системами. Такие ряды, как правило, получают из решений нелинейных динамических уравнений, и, как следствие, они являются гладкими, непрерывными и, соответственно, могут быть практически бесконечными; в идеале они избавлены от случайных шумов, то есть они представляют динамический хаос в чистом виде. Ряды наблюдений всегда конечны, как правило, не являются гладкими иотяжены случайными шумами разной природы. Это обстоятельство налагает серьезные ограничения на использование энтропии Колмогорова для анализа реальных рядов данных наблюдений, так как соответствующие предельные переходы не всегда возможны, а если и возможны в приемлемом приближении, то часто пре-

делы не всегда существуют. Иными словами, энтропия Колмогорова для реальных временных рядов может вести себя неустойчиво.

Поиск робастных подходов к энтропийно-фрактальной параметризации природных временных рядов велся все последние годы, и к настоящему времени предложено несколько перспективных методов. В основе всех этих методов лежит понятие элементарной энтропии выборки, или выборочной энтропии. *Под выборкой обычно понимают конечную последовательность значений временного ряда $\{x_k\}$, выбранных по определенному правилу.* Фрагмент ряда является простым примером выборки. В общем случае, в зависимости от наших целей мы можем выбирать из ряда некоторое число значений и составлять из них ту или иную последовательность. Каждая из таких последовательностей будет выборкой.

3.3.1. Первая выборочная энтропия

Чтобы определить элементарную энтропию Шеннона выборки из N значений рассматриваемого ряда $\{x\}$, находим максимальную (x_{\max}) и минимальную (x_{\min}) величины в анализируемом ряде. Разбиваем этот интервал ($x_{\max} - x_{\min}$) на n подинтервалов (уровней) таким образом, чтобы величина интервала Δx была не меньше доверительного интервала данных наблюдений. Рассматриваем выборку как «сообщение», а подинтервалы i как «алфавит». Далее находим число значений выборки $\{x_k\}$, попавших в каждый из подинтервалов, и определяем относительную населенность уровня p_i (вероятность попадания значения из выборки в подинтервал i , то есть относительную частоту встречаемости «буквы» в «сообщении»).

$$p_i = \Delta N_i / N. \text{ Очевидно, что при этом } \sum_{i=1}^n \Delta N_i = N, \sum_{k=1}^n p_i = 1. \text{ Элементарная энтропия выборки определяется как энтропия Шеннона (2.7) на данном наборе } p_i:$$

Элементарная энтропия выборки определяется как энтропия Шеннона (2.7) на данном наборе p_i :

$$H = - \sum_{i=1}^n p_i \log_2 p_i. \quad (3.1)$$

Продемонстрируем алгоритм расчета выборочной энтропии Шеннона на конкретном примере временного ряда вариаций рентгеновского потока от Солнца, регистрируемого спутниками GOES. На рис. 3.1 показан небольшой фрагмент этого ряда. Данные взяты с сайта: http://www.sec.noaa.gov/ftpd/warehouse/2003/2003_plots/. Этот непрерывный ряд наблюдений охватывает период свыше 20 лет, то есть около 2-х циклов солнечной активности.

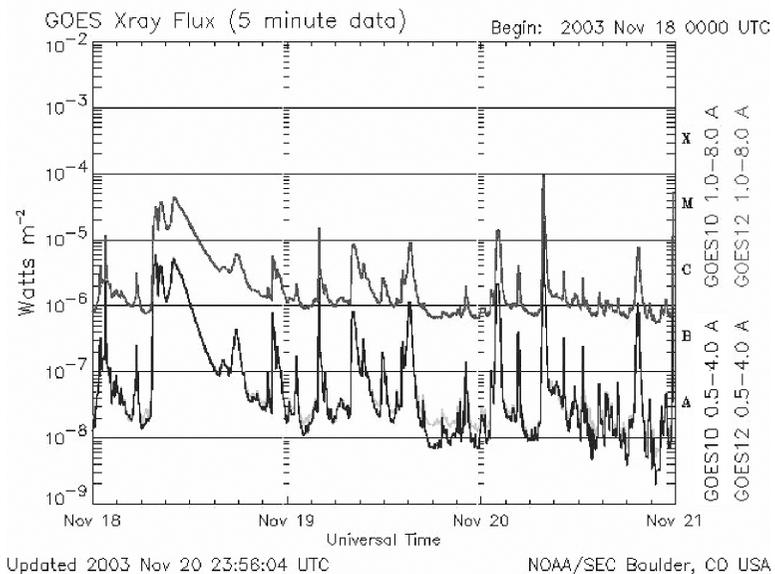


Рис. 3.1. Фрагмент временного ряда вариаций интегрального рентгеновского потока от Солнца (NOAA/SEC Boulder, USA) в двух диапазонах 0.5–4.0 А и 1.0–8.0 А. По оси абсцисс отложено Универсальное время (UT) с шагом 5 мин. Данный фрагмент ряда относится к ноябрю 2003 года и охватывает период с 18 по 20 ноября включительно. По оси ординат дана шкала значений потока в Вт/м²

Известно, что уровень и характер солнечной активности количественно оценивается с помощью специальных индексов. Наиболее популярным индексом является индекс Вольфа, определяемый уже в течение более 260 лет с 1749 г. Полный радиопоток на волне 10.7 см является другим примером такого индекса. Полная площадь солнечных пятен также рассматривается в качестве индекса, количественно характеризующего солнечную активность. Имеются и другие индексы.

Все эти параметры на больших интервалах времени в целом неплохо коррелируют друг с другом. Вместе с тем они отражают различные стороны многопланового явления солнечной активности и часто могут существенно различаться, отражая специфику различных активных процессов на Солнце в тот или иной момент времени. Индексы, как правило,

количественно представляют уровень тех или иных процессов солнечной активности с фиксированным временным шагом за определенный промежуток времени. Наиболее популярными индексами солнечной активности являются дневные, месячные и годовые индексы. Их обычно сопоставляют с соответствующими геофизическими, метеорологическими, биофизическими и другими индексами при рассмотрении различных проблем солнечно-земной физики.

В качестве примера энтропийной параметризации сложных процессов введем новый индекс солнечной активности. Назовем его Шенноновская Энтропия Рентгеновского Потoka первого типа (или Шэрп1 индекс). Если мы будем его вычислять из данных за сутки, это будет суточный Шэрп1 индекс, если из данных за неделю — недельный Шэрп1 индекс и так далее.

Рассмотрим процедуру расчета суточного Шэрп1 индекса. Из ряда интегрального рентгеновского потока от Солнца выделим суточный фрагмент (см. рис. 3.2). Такой фрагмент будем называть также суточной выборкой. Определим шенноновскую энтропию этой выборки (далее — суточная Шенноновская энтропия). Делается это следующим образом.

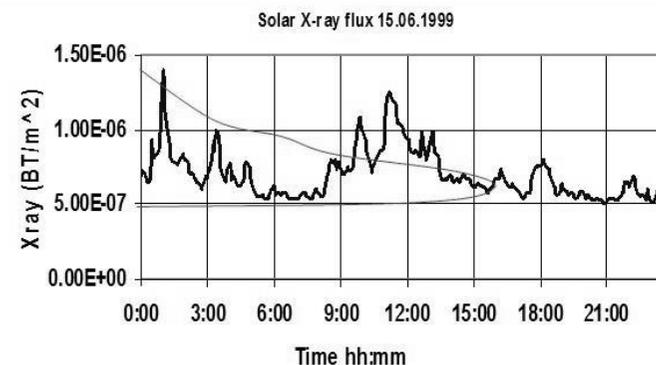


Рис. 3.2. Суточная выборка из данных солнечного рентгеновского потока в диапазоне 1.0–8.0 ангстрем. Плавная кривая — набросок распределения потока по амплитудам за этот день

Определяем минимальное и максимальные значения рентгеновского потока за анализируемый интервал наблюдений, скажем, за один цикл солнечной активности. Пусть это будет F_{\max} и F_{\min} . Этот интервал ампли-

туд разбиваем на конечное число равных интервалов n так, что один элемент d этого покрытия равен $d = (F_{\max} - F_{\min})/n$. Пусть N — полное число измерений амплитуды в выборке. В нашем случае $N = 288$ — число пятиминуток в сутках.

Сумма амплитуд по выборке дает нам норму A выборки:

$$A = \sum_{i=0}^N F_i, \quad (3.2)$$

а $p_i = F_i/A$ дает относительную заселенность i -го уровня или вероятность нахождения значения рентгеновского потока на этом уровне. Понятно, что

$$\sum_{i=1}^N p_i = 1. \quad (3.3)$$

Тогда выборочная энтропия Шеннона равна

$$H = -\sum_{i=1}^N p_i \log_2 p_i = -\sum p_i \log_{10} p_i / \log_{10} 2. \quad (3.4)$$

Как уже отмечалось, минимальное значение $H_{\min} = 0$, а максимальное H_{\max} достигается при равномерном заполнении всех уровней, то есть когда все $p_i = 1/n$. Тогда $H_{\max} = \log_2 n$. Если, например, мы положили $n = 300$, то $H_{\max} = 8.229$. Таким образом, при $n = 300$ наш дневной Шерп индекс может принимать значения в интервале от 0 до 8.229.

По данному алгоритму нетрудно написать простую компьютерную программу для расчета Шерп1 индекса. Расчет энтропии Шеннона для суточной выборки, приведенной на рис. 3.2, дает значение $H = 2.299$, то есть наш Шерп1 индекс равен 2.299.

На рис. 3.3 представлена суточная выборка рентгеновского потока от Солнца для другой даты — на 17.06.1999 г. Сравнивая рис. 3.2 и рис. 3.3, мы видим существенную разницу в характере вариабельности рентгеновского потока в эти дни. Наш Шерп1 индекс мгновенно реагирует на эту ситуацию. Его значение становится равным 3.377. Смысл этих двух значений индекса, на первый взгляд, понятен: в первой выборке амплитуда меняется в пределах 1.5 порядков, во втором — в пределах примерно 3.6 порядков. Но Шерп1 индекс характеризует не только размах амплитуды.

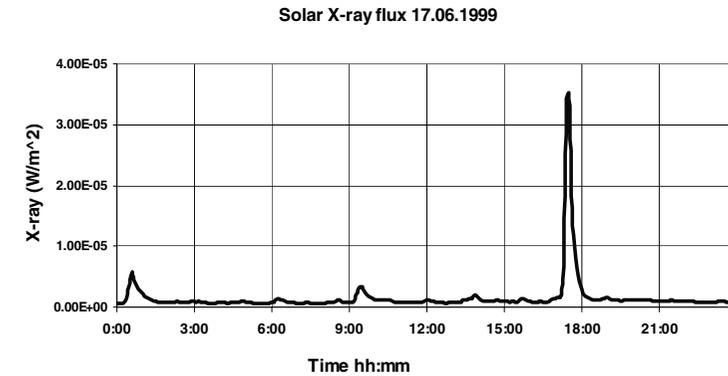


Рис. 3.3. Дневная выборка (17.06.1999 г.) из данных солнечного рентгеновского потока в диапазоне 1.0 – 8.0 ангстрем

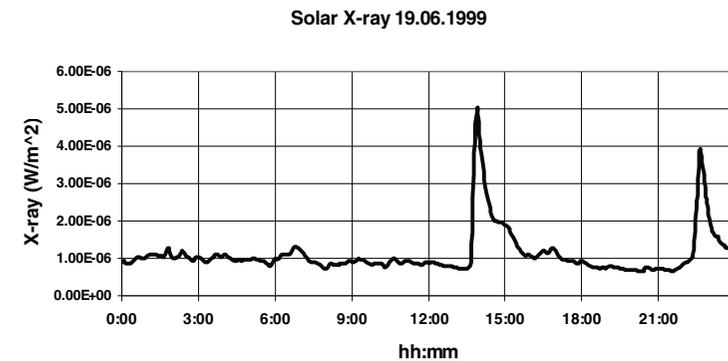


Рис. 3.4. Дневная выборка (19.06.1999 г.) из данных солнечного рентгеновского потока в диапазоне 1.0 – 8.0 ангстрем

На рис. 3.4 представлена дневная выборка из данных солнечного рентгеновского потока за 19.06.1999 г. Для этой выборки Шерп1 = 3.179. Интересно отметить, что в этом случае амплитуда потока изменяется в больших пределах, чем в выборке за 17.06.1999 г., однако значение Шерп1 индекса меньше. Это связано с тем, что энтропия Шеннона является специфической характеристикой. Она дает количественную оценку отклонения данного распределения амплитуды по уровням от однородного, то есть когда все уровни равномерно заполнены. Таким обра-

зом, наш Шэрип1 индекс характеризует степень вариабельности процесса. При нулевой вариабельности (амплитуда не меняется со временем) Шэрип1 = 0, при максимальной вариабельности выборки (все уровни равномерно заполнены) Шэрип1 достигает своего максимального значения, равного $\log_2 n$, где n — число уровней; в рассматриваемом примере максимальное значение Шэрип1 = 8.229.

О чем говорят нам значения Шэрип1 индекса в плане интенсивности и направления эволюции физических процессов на Солнце, генерирующих этот рентгеновский поток? Собственно говоря, это один из тех вопросов, ради ответа на который мы и делали нашу энтропийную параметризацию временного ряда. Ответ простой. Большая вариабельность, то есть ситуация, когда Шэрип1 индекс принимает значения, близкие к максимальным, говорит о чрезвычайной интенсивности и нестабильности процессов генерации рентгеновского потока. Напротив, небольшая вариабельность говорит об относительно спокойной фазе. Знак тренда Шэрип1 индекса (если тренд надежно просматривается) указывает на направление эволюции процессов, порождающих рентгеновский поток. В этом качестве тренд может быть использован для краткосрочных прогнозов.

Энтропия Шеннона по определению количественно характеризует произвольное распределение какого-либо параметра системы — энергии, скорости, числа атомов в разных состояниях возбуждения и др. Если в системе происходят какие-либо изменения, то меняются распределения ее параметров. Соответственно, изменяется и величина энтропии. В этом смысле энтропия Шеннона является функцией состояния системы, так как количественно характеризует меру неопределенности значений параметров, характеризующих систему.

3.3.2. Вторая выборочная энтропия

Для формулы Шеннона (3.1) неважно, каким физическим параметром мы характеризуем состояние системы. Более того, несущественно, как мы выбираем наше ε -покрытие. Все это наш выбор. Но как только мы определили p_i для нашей задачи, мы можем по формуле (3.1) вычислить энтропию данного параметра на выбранном покрытии. Если p_i — вероятность появления i -го сигнала в сообщении, то это будет энтропия Шеннона. Если p_i -элементы матрицы плотности квантово-механической системы, то это энтропия Фон Неймана (см. раздел 1.2.10), и так далее. В этом смысле формулу (3.1) мы будем называть обобщенной энтропией Шеннона.

Если параметром является относительное число атомов в ε -ячейке фазового объема, то обобщенная энтропия Шеннона будет тождественна энтропии Больцмана (см. раздел 2.4). В этом смысле энтропия Больцмана является частным случаем обобщенной энтропии Шеннона. Это утверждение можно найти в работах (Климонтвич, 1982, 2002) [1, 2] и (Эбелинг, 2004) [3]. Эбелинг утверждает, что это положение является одним из основных постулатов статистической теории. В этой связи Эбелинг пишет: «Аналогия между термодинамической энтропией, статистической энтропией Больцмана и информационной энтропией наводит на мысль об отождествлении этих величин». И далее: эти «равенства рассматриваются сегодня как основные постулаты статистической теории вещества».

Однако во избежание путаницы и для правильной интерпретации результатов расчета энтропии под *первой энтропией Шеннона* мы, как и ранее (см. раздел 2.4), будем понимать энтропию на амплитудном покрытии, а под *второй энтропией Шеннона* — на пространственном или временном.

Рассмотрим несколько примеров. Ранее мы рассчитывали первую энтропию Шеннона временных рядов. Сейчас мы сделаем то же самое для второй энтропии Шеннона. Для данных, представленных на рис. 3.2–3.4, получаем следующие значения второй энтропии Шеннона: для 15.06.99 — $H_2 = 2.703$ ($H_1 = 2.485$); для 17.06.99 — $H_2 = 1.918$ ($H_1 = 4.037$); для 19.06.99 — $H_2 = 2.518$ ($H_1 = 3.768$), при том же диапазоне ее изменения — от 0 до 8.229. В скобках, рядом с энтропиями Больцмана, выписаны соответствующие значения энтропии Шеннона H_1 (то есть нашего Шэрип1 индекса).

Сравнивая эти значения, видим, что эти величины существенно различаются. Это естественно, поскольку они получены для разных определенных вероятностных пространств, хотя оба эти пространства генерируются одной и той же выборкой. Это два ортогональных пространства и, следовательно, первая и вторая энтропии Шеннона и линейно независимы⁷. Это значит, что H_2 можно рассматривать как еще один независимый индекс, количественно характеризующий особенности суточных вариаций рентгеновского потока от Солнца (далее будем называть его Шенно-

⁷ Линейная независимость энтропий, определенных на ортогональных вероятностных пространствах, обсуждается в работе (Коротеев, 2003) [4]. В этой работе рассматриваются также некоторые обобщения понятия энтропии, однако они не представляют прямого интереса для нашей темы и здесь не обсуждаются

новская Энтропия Рентгеновского Потока второго типа, или, сокращенно, Шэрп2 индекс).

Для постоянного по времени потока Шэрп2 индекс принимает свое максимальное значение, для δ -образного распределения стремится к нулю. Иначе говоря, ведет себя в определенной мере противоположно Шэрп1 индексу: большие значения Шэрп2 индекса соответствуют относительно спокойным (равновесным) фазам процессов, генерирующих рентгеновские потоки. Малые значения указывают на их взрывообразное течение. Таким образом, рост Шэрп1 индекса и падение Шэрп2 индекса указывают на нестабильность и тенденцию к возможным резким вариациям потока.

Оба индекса, Шэрп1 и Шэрп2, являются двумя независимыми и дополняющими друг друга характеристиками временных вариаций амплитуды потока за сутки. Этих двух индексов, разумеется, не достаточно, для более или менее полного в плане геофизических приложений описания рентгеновского потока от Солнца. Для этих целей полезно дополнить эти два индекса третьим, характеризующим полную энергию рентгеновского потока за соответствующий промежуток времени. Такая ортогональная система индексов может оказаться полезной для различных геофизических и гелиофизических задач, в том числе и в области прогнозов космической погоды.

Рассмотренную выше энтропию H_1 произвольной выборки из временного ряда можно назвать элементарной выборочной энтропией. Это связано с тем, что она рассчитывается на основе элементарного обобщения исходной энтропии Шеннона: сама выборка рассматривается как «сообщение», а введенное одномерное пронумерованное покрытие на оси амплитуд играет роль «алфавита». Фактически такое обобщение аналогично тому, которое проделал фон Нейман: квантовые состояния системы он рассматривал как «алфавит», а распределение вероятностей этих состояний — как «сообщение». Но в нашем случае важно отметить, что если мы предполагаем сравнивать энтропии разных выборок рассматриваемого ряда, то «алфавит» должен быть один для всех выборок, то есть наше покрытие должно быть одним и тем же для всего исследуемого ряда.

Менее тривиально обобщение энтропии Шеннона во втором рассмотренном нами случае H_2 . Здесь «алфавит» задается самой выборкой, а относительное значение амплитуды p_i выступает в качестве частоты встречаемости (вероятности) «символа» в данном «сообщении». Нормировка p_i , разумеется, должна быть единой для всех сравниваемых выборок.

Понятно, что процедура нахождения двух таких выборочных энтропий применима к любому временному экспериментально-наблюдательному ряду, исключая, пожалуй, ряды логического типа. Для последних возможен только расчет выборочных энтропий Шеннона второго типа.

3.3.3. Перестановочная энтропия

Теперь мы познакомимся с еще одним определением вероятностной меры p_i на выборке и, соответственно, с еще одним вариантом энтропии — с так называемой *перестановочной энтропией* временного ряда. Рассмотрим снова выборку $\{x_k\}$ из исследуемого временного ряда $\{x\}$. *Перестановочная энтропия выборки порядка n определяется как Шенноновская энтропия $n!$ перестановок, определяемых отношениями n последовательных значений выборки.*

Так, например, перестановочная энтропия порядка $n = 2$, определяется так. Пусть в выборке $\{x_k\}$ содержится N значений ряда, то есть $1 \leq k \leq N$. Сравниваем два соседних значения x_k и x_{k+1} . Находим число m_1 случаев в выборке, когда $x_k < x_{k+1}$, и число случаев в выборке m_2 , когда $x_{k+1} < x_k$. Если все x_k различимы, то есть нет равных значений x_k , тогда $m_1 + m_2 = N$, $p_{2,1} = m_1/N$, $p_{2,2} = m_2/N$. Перестановочная энтропия порядка $n = 2$, равна: $H_2 = -p_{2,1} \log_2 p_{2,1} - p_{2,2} \log_2 p_{2,2}$.

Перестановочная энтропия порядка $n = 3$, соответственно, определяется следующим образом. Находим число m_1 случаев в выборке, когда $x_k < x_{k+1} < x_{k+2}$, m_2 — когда $x_{k+1} < x_k < x_{k+2}$, и так далее. Число таких перестановок в данном случае будет $3! = 6$. Тогда

$$\sum_{j=1}^{3!} m_j = N, \quad p_j = m_j / N \quad \text{и} \quad H_3 = -\sum_{j=1}^{3!} p_j \log_2 p_j. \quad (3.5)$$

Соответственно, перестановочная энтропия выборки порядка n определяется так:

$$\sum_{j=1}^{n!} m_j = N, \quad p_j = m_j / N \quad \text{и} \quad H_n = -\sum_{j=1}^{n!} p_j \log_2 p_j. \quad (3.6)$$

Здесь, как и в случае классической энтропии Шеннона, перестановочная энтропия выражается в битах. Разумеется, ее можно выражать и в «натах», то есть использовать натуральные логарифмы, и в «датах», то есть пользоваться десятичными логарифмами, но нужно учитывать,

что «биты» все же являются общепринятой размерностью для шенноновских энтропий.

Максимально возможное значение перестановочной энтропии любого порядка $\max H_n$ достигается, как обычно, в равновероятном случае, то есть когда в выборке все случаи m_j равны. Из формул (3.6) видим, что $\max H_n = \log_2 n!$. Например, такие значения перестановочной энтропии может иметь бесконечный ряд случайных чисел. Минимальное значение энтропии $\min H_n = 0$ достигается в случаях, когда во всей выборке реализуется только одна из $n!$ перестановок. Например, это может быть ряд монотонно растущих или монотонно убывающих значений.

Каждая из этих энтропий и все вместе дают оценку меры сложности ряда, то есть их обычно рассматривают как параметры сложности наряду с вышерассмотренными энтропиями, фрактальными размерностями, о которых будет идти речь ниже, а также энтропией Колмогорова и показателями Ляпунова. Эти последние параметры, как известно, рассматривались в контексте теории типичных, предположительно эргодических, динамических систем и имеют между собой глубокие связи, позволяющие вычислить один из них по-другому [7, 8].

Однако, как уже отмечалось, прямое применение результатов теории динамических систем имеет существенные ограничения в случае работы с реальными рядами наблюдений, продуцируемых природными системами со сложным поведением и имеющих стохастические компоненты разной природы. Как было показано в ряде работ (см., например, [9, 10]), перестановочная энтропия также тесно связана с показателями Ляпунова и энтропией Колмогорова, но ее существенным преимуществом является то обстоятельство, что гладкость ряда для нее несущественна в силу ее определения. Метод ее расчета является достаточно робастным, и она вполне применима к реальным рядам наблюдений.

При практической работе с перестановочной энтропией возникает вопрос о ее необходимом и достаточном числе порядков n при работе с конкретными рядами и выборками. Очевидным ограничением является соотношение $n! < N$, где N — объем выборки. Однако практика показывает, что приемлемые по точности значения получаются только при $N > 2n!$. Компьютерные программы, построенные по вышеописанному алгоритму с учетом этих ограничений, дают устойчивые и достаточно точные значения перестановочной энтропии разных порядков. Так, например, для случайного ряда, а также для выборок, представленных на рис. 3.2, рис. 3.3 и рис. 3.4, были получены следующие значения перестановочной энтропии порядков 2, 3, 4, 5 и 6, представленные в Таблице 3.1.

Таблица 3.1.

Перестановочные энтропии до шестого порядка

Order	$2n!$	$\log_2 n!$	Random	15.06.1999	17.06.1999	19.06.1999
H2	4	1.00000	0.99947	0.98523	0.96930	0.94744
H3	12	2.58522	2.58195	2.27745	2.16185	2.08709
H4	48	4.58541	4.56407	3.75564	3.48265	3.34729
H5	240	6.90759	6.80836	5.22848	4.85203	4.67028
H6	1440	9.49280	8.88883	6.35545	5.92288	5.78027

В первом столбце таблицы указаны порядки энтропий. Во втором — необходимый объем выборки N согласно практическому критерию $N > 2n!$. В третьем столбце — максимально возможные значения энтропий соответствующих порядков. В четвертом — энтропии случайного ряда из 1000 значений. В пятом, шестом и седьмом рядах — энтропии выборок, представленных на рис. 3.2, рис. 3.3 и рис. 3.4.

Сравнивая третий и четвертый столбцы, можно оценить, как работает практический критерий $N > 2n!$. Мы видим, что энтропия случайного ряда, как и положено, близка к максимально возможным значениям. До энтропий пятого порядка, то есть до тех пор, пока критерий выполняется, относительная ошибка не превышает 1.5%. Для энтропии шестого порядка, когда критерий уже не выполняется, относительная ошибка превышает 6%.

Объемы выборок, по которым рассчитывались энтропии, содержащиеся в пятом, шестом и седьмом столбцах, составляли по 288 значений. Для этих выборок требование $N > 2n!$ выполняется также до энтропий пятого порядка. Энтропии шестого порядка уже могут содержать значительные ошибки.

Перестановочные энтропии разных порядков являются полезным инструментом анализа для большого количества временных рядов самой разной природы. Они позволяют получить важную информацию о динамических процессах в сложных системах, продуцирующих нестационарные ряды со значительной стохастической компонентой.

3.3.4. Энтропия Реньи

Выше мы уже упоминали об обобщении Реньи энтропии Шеннона. Здесь мы рассмотрим это обобщение более подробно.

Пусть некий объект (фрактал, аттрактор) вложен в n -мерное евклидово пространство. Часть пространства, занятого объектом, покроем равными гиперкубами с ребром ε . Пусть m_i — элемент «массы» или «значе-

ние» (число точек, поток вектора, почернение, амплитуда процесса и т. д.) в i -том гиперкубе данного покрытия. Тогда соответствующая вероятность равна: $p_i = m_i/m$, где

$$m = \sum_{i=1}^{N(a)} m_i \quad (3.7)$$

— полная «масса» исследуемого объекта (ряда, магнитограммы, ССД изображения, аттрактора и т. д.), $N(a)$ — общее число ячеек, покрывающих объект. Ясно, что

$$\sum_{i=1}^{N(a)} p_i = 1. \quad (3.8)$$

Обобщенная энтропия Реньи порядка q определяется так:

$$R_q(a) = \frac{1}{1-q} \ln \left(\sum_{i=1}^{N(a)} p_i^q \right), \quad \text{где } q \in (-\infty, \infty). \quad (3.9)$$

Частные случаи энтропии Реньи

1. Если наш объект *однороден*, то есть все m_i равны между собой, то $R_q(a) = \ln N(a)$. То есть энтропия Реньи *любого* порядка q в случае однородного объекта равна логарифму числа ячеек разбиения. Покажем это. Из (3.7) получаем, $m = N(a) \cdot m_i$, откуда $p_i = 1/N(a)$. Подставляя в (3.9), получаем

$$\begin{aligned} R_q(a) &= \frac{1}{1-q} \ln \left(\sum_{i=1}^{N(a)} \left(\frac{1}{N(a)} \right)^q \right) = \frac{1}{1-q} \ln \left(N(a) \left(\frac{1}{N(a)} \right)^q \right) = \\ &= \frac{1}{1-q} \ln \left(N(a)^{1-q} \right) = \ln N(a). \end{aligned}$$

2. Если $q = 0$ и все $p_i \neq 0$, то из формулы (3.9) получаем тот же результат:

$$R_0(a) = \ln N(a). \quad (3.10)$$

3. Пусть $q = 1$. Подставляя $q = 1$ в (3.9) и раскрывая неопределенность по правилу Лопитала, получаем:

$$R_1 = \lim_{q \rightarrow 1} \frac{\sum_{i=1}^{N(a)} p_i^q \ln p_i}{-\sum_{i=1}^{N(a)} p_i^q} = -\sum_{i=1}^{N(a)} p_i \ln p_i, \quad (3.11)$$

что с точностью до основания логарифмов (то есть постоянного множителя) совпадает с энтропией (информацией) Шеннона. В этой связи энтропию Реньи первого порядка называют *информационной энтропией*.

4. Пусть $q = 2$. Квадрат вероятности $(p_i)^2$ можно рассматривать как вероятность корреляции «значений» нашего объекта на расстоянии ε . Тогда сумма

$$\sum p_i^2 = c(a) \quad (3.12)$$

— аналог корреляционного интеграла. Поэтому энтропию Реньи второго порядка

$$R_2(a) = \ln \left(\sum_{i=1}^{N(\varepsilon)} p_i^2 \right) \quad (3.13)$$

называют также *корреляционной энтропией*.

В заключение этого раздела отметим следующее. Все рассмотренные выше выборочные энтропии временных рядов при надлежащем определении выборок нечувствительны к разрывам в данных измерений и вообще к вариациям шага по времени. В этом плане энтропийная параметризация временного ряда выгодно отличается от любого типа фильтрации, которая к разрывам в данных обычно очень чувствительна. Такая параметризация особо предпочтительна для стохастических и автономных временных рядов, поскольку не требует предположений о гладкости и непрерывности. Данные экспериментов и наблюдений, вообще говоря, представляют собой достаточно произвольную выборку из вариаций параметра некоторого стохастического процесса. Сама вариация может быть не гладкой и даже разрывной. Процедура измерений вносит дополнительно свои разрывы и свою стохастичность, связанную с ошибками и спецификой измерений. Как уже отмечалось, применение методов, основанных на теории гладких функций, к анализу рядов такого рода сопряжено с определенными трудностями. Расчет выборочных энтропий H_1 , H_2 и перестановочной энтропии разных порядков и энтропий Реньи открывает новые возможности для анализа данных такого типа.

3.4. Энтропии видеоизображений

Видеоизображения, наряду с другими временными рядами, являются важнейшим источником информации во многих естественных науках,

в том числе в физике, астрофизике и геофизике. Раньше это были фотографические снимки, аналоговая информация которых (распределение по снимку величины почернения) переводилась в цифровые массивы. Далее производились специальные процедуры трансформации (фильтрации, коррекции, редукации, калибровки) этих числовых значений величин почернения в прямые интенсивности световых потоков, вызвавших данные почернения (абсолютная фотометрия), и затем эти данные анализировались теми или иными численными методами.

С внедрением в технику наблюдений цифровой регистрации изображений все необходимые трансформации первичных сигналов осуществляются, как правило, аппаратно или программно в реальном времени, и на выходе мы имеем блоки данных в определенном формате⁸, содержащие числовые массивы в той или иной физической размерности, а также всю сопроводительную документацию (метаданные). В метаданных указываются назначение и характер данных, дата и время регистрации, тип и характеристики регистрирующего прибора, точность измерений, условия получения данных, информация о фотометрической и масштабной калибровке, и ряд других параметров. Далее под видеоизображением мы будем понимать блоки именно таких данных.

Чтобы количественно охарактеризовать структурные и статистические особенности изображений, различных фрагментов одного изображения или разные кадры одного временного ряда изображений, необходимо иметь удобные и легко вычисляемые количественные параметры, отражающие эти особенности. Энтропии Шеннона, аналогичные рассмотренным выше H_1 , H_2 , а также перестановочная энтропия и энтропии Реньи, вычисляемые на двумерном покрытии, хорошо подходят для этих целей. Они позволяют количественно оценить изменения в структурах и статистике, определить направление и темп эволюции наблюдаемой системы, могут служить количественными критериями при разработке моделей физических процессов, видимой манифестацией которых являются наблюдаемые структуры.

⁸ В последние годы для регистрации и передачи научных астрономических данных используется fits формат (Flexible Image Transport System). В июле 2008 года версия 3.0 FITS официально принята МАС в качестве стандарта для хранения, передачи и редактирования изображений и их метаданных. Полное описание fits-формата можно найти здесь: http://fits.gsfc.nasa.gov/standard30/fits_standard30.pdf

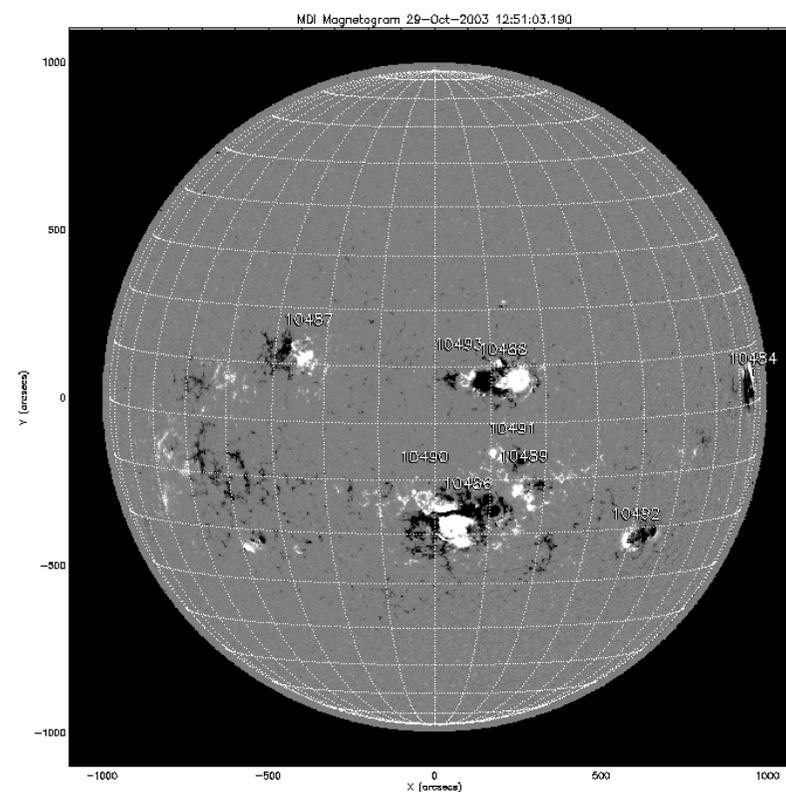


Рис. 3.5. Пример MDI магнитограммы солнечного диска. Серый цвет соответствует нулевому значению продольной компоненты магнитного поля. Области светлее фона соответствуют северной полярности, области темнее фона — южной полярности. Номера активных областей по каталогу NOAA указаны выше каждой АО

Анализируемое изображение мы рассматриваем как выборку из временного ряда аналогичных изображений (2D ряд). В отличие от выборки из рассмотренных выше скалярных временных рядов (1D ряд), ряды 2D и более высоких размерностей, применяемые в астрономии⁹, позволяют

⁹ В качестве дополнительных координат в астрономических данных обычно фигурируют цветовые индексы. Но анализируются также данные и больших размерностей. Так, при анализе снимков звездного неба каждой звезде ставится в соответствие ее пространственное положение, вектор скорости в той или иной системе координат, а также ряд фотометрических и спектральных характеристик.

делать более широкий выбор вероятностной меры p_i в зависимости от решаемой задачи. Здесь мы ограничимся рассмотрением выборок из 2D ряда и на конкретном примере рассмотрим введение некоторых вариантов покрытий и вероятностных мер. Формальное обобщение на данные с большим числом размерностей в основном является технической проблемой и может решаться аналогичным образом.

В качестве примера выборки из 2B ряда рассмотрим несколько последовательных магнитограмм двух солнечных активных областей (далее АО): NOAA 10486 и NOAA 10488. Магнитограммы были получены на космической научной платформе Solar and Heliospheric Observatories (SOHO) с помощью Michelson Doppler Imager (MDI) [11]. Пример магнитограммы SOHO/MDI показан на рис. 3.5.

Активная область NOAA 10486 (далее АО 10486) появилась на видимой полусфере Солнца 23 октября 2003 г. как сложная $\beta\gamma\delta$ -конфигурация. В течение всего периода наблюдений вплоть до ее исчезновения за западным лимбом 6 ноября АО 10486 оставалась сложной $\beta\gamma\delta$ -структурой. Область изменялась очень быстро и продуцировала высокий уровень вспышечной активности. За время прохождения по видимой части диска в этой АО произошло восемь вспышек X класса и множество вспышек класса M. АО 10486, таким образом, являлась одной из наиболее мощных АО 23 солнечного цикла. Общее представление об эволюции магнитного поля этой АО можно получить из рис. 3.6.

Активная область NOAA 10488 (далее АО 10488) появилась в центре солнечного диска 26 октября примерно в 9-10 часов как новый быстро возникающий и развивающийся магнитный поток в северной полусфере, всего на 5 градусов восточнее АО 20486. К 27 октября 10488 становится уже сложной $\beta\gamma\delta$ -областью, и вплоть до 31 октября (конец анализируемых данных) она остается растущей группой. За анализируемый в этой работе период в 10488 наблюдались только три события M класса и одно класса X. Общий характер эволюции магнитного поля этой АО можно видеть на рис. 3.7.

Выбор этих двух объектов неслучаен. АО 10486 находилась на высоком уровне структурной сложности, выражающейся в сильной перепутанности магнитных потоков. Фаза структурной эволюции соответствовала стагнации с элементами депрессии и относительного упрощения, особенно перед заходом за солнечный лимб. Напротив, АО 10488 находилась на фазе роста и бурного усложнения структуры. Таким образом, эти две АО позволяют продемонстрировать возможности энтропии как удобного параметра для количественной характеристики фазы и направления эволюции сложной системы.

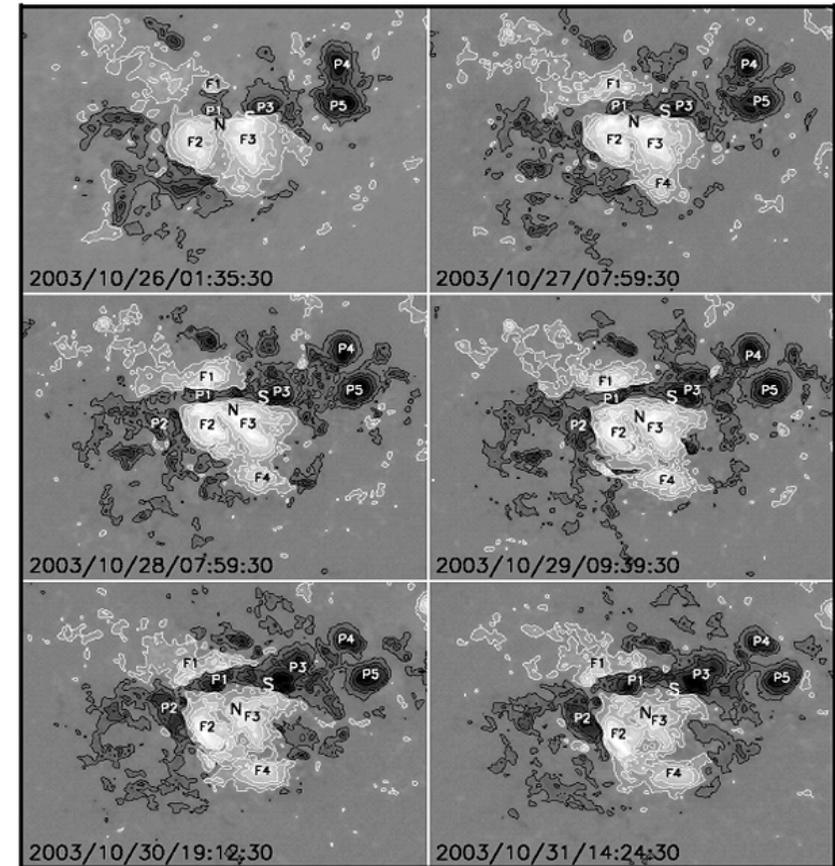


Рис. 3.6. Пример шести MDI магнитограмм NOAA10486 с временными интервалами порядка суток. Показана эволюция магнитного поля этой АО с 26 по 31 октября 2003 г. (данные SOHO/MDI). Изолинии поля проведены через 200 гаусс. Литерой F с индексом обозначены пятна N полярности, литерой P — пятна S полярности. Дата и время (UT) представлены в левом нижнем углу каждой магнитограммы

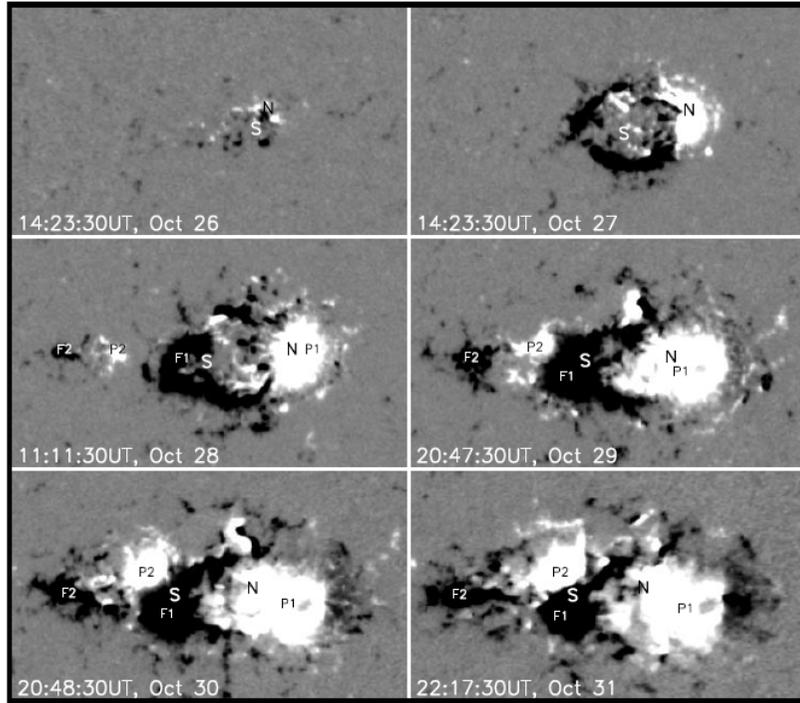


Рис. 3.7. Пример шести MDI магнитограмм АО 10488 с временными интервалами порядка суток. Показана временная последовательность продольной компоненты магнитного поля этой АО в период с 26–31 октября 2003 г. (данные SOHO/MDI). Изолинии поля не показаны. Литерой F с индексом обозначены пятна S полярности, литерой P — пятна N полярности. Дата и время (UT) представлены в левом нижнем углу каждой магнитограммы

В наших примерах весь солнечный диск проектируется на матрицу 1024×1024 пикселя, (рис. 3.5). Пусть у нас имеется K таких изображений. С помощью специальных процедур «вырезаем» из диска нужную нам АО размером m на n пикселей (рис. 3.6 и рис. 3.7).

Пусть B_{\max} — максимальное значение величины магнитного поля в данном ряду магнитных карт, B_{\min} — уровень дискриминации (например, точность измерений магнитометра), ΔB — интервал выбранной нами шкалы покрытия по полю. Тогда $L = (B_{\max} - B_{\min})/\Delta B$ — число «квантовых» уровней величины поля для данной шкалы. Пусть $M_{ns} = m \cdot n$ — полное число отсчетов поля на магнитной карте. Если оценивается энтропия полярности, то M_n или M_s — полное число отсчетов в со-

ответствующей полярности. Для северной полярности $p_i = n_i/M_n$, где n_i — число отсчетов, в которых значение поля этой полярности соответствует i -му уровню (населенность i -го уровня); p_i — относительная населенность i -го уровня, или вероятность обнаружить значение поля в интервале $B_{i+1} - B_i$ на данной магнитной карте. Тогда амплитудная энтропия Шеннона для северной полярности, выраженная в битах, равна:

$$H1_n = -\sum_{i=1}^L p_i \log(p_i) / \log(2). \quad (3.13)$$

Из этой формулы понятно, что значения энтропии заключены в пределах: $0 < H1 < \ln(L) / \ln 2$. Нуль соответствует случаю, когда все поле на карте имеет везде одно и то же значение (то есть случай вырожденного распределения — представлен только один уровень), $\ln(L)$ — случаю, когда на карте представлена только одна полярность и каждому уровню соответствует одинаковое число отсчетов (равномерное распределение по полю). Для южной полярности можно записать аналогичную формулу.

Мы можем и другим способом определить вероятностную меру p_i . Рассмотрим в качестве «естественного» покрытия элементы CCD матрицы. Пусть площадка в центре солнечного диска, соответствующая одному пикселю, равна $\sigma \text{ см}^2$. Полагаем, что все наши измерения магнитного поля уже исправлены на эффекты проекции, то есть, «приведены к центру диска». Тогда магнитный поток F_{ij} в продольном поле в пикселе (i, j) равен $F_{ij} = \sigma B_{ij}$ мкс. Эту величину будем рассматривать в качестве «веса» данного элемента покрытия. Пусть F_{\max} — максимальное значение модуля потока в рассматриваемой выборке. Тогда $p_i = F_{ij}/F_{\max}$ относительный «вес» данного пикселя. Соответственно, энтропия H_n северной полярности равна:

$$H2_n = -\sum_{i=1}^{M_n} p_i \log(p_i) / \log(2). \quad (3.14)$$

Аналогично (с учетом знака!) может быть вычислена энтропия южной полярности. Несмотря на формальную идентичность формул (3.13) и (3.14) их физический смысл различен в силу различия в определении вероятностной меры. Энтропия, определяемая по формуле (3.13) соответствует энтропии Шеннона $H1$, тогда как в случае (3.14) — энтропии $H2$. Поскольку нас далее будут интересовать структура магнитного поля и распределение магнитных потоков по площади на уровне фотосферы, занимаемой рассматриваемой активной областью, то далее в качестве вероятностной меры рассматривается мера формулы (3.14).

Энтропию Реньи порядка q той и другой полярности можно вычислить по формуле (3.9). По этому алгоритму несложно написать компьютерную программу для расчета энтропий Реньи любого порядка. Результаты вычислений для двух рассматриваемых солнечных АО энтропий Реньи нулевого порядка (структурной энтропии), первого порядка (информационной энтропии) и второго порядка (корреляционной энтропии) представлены на рис. 3.8. Мы не будем здесь детально анализировать эти результаты. Это потребовало бы углубление в соответствующую предметную область, что выходит за рамки данной книги. Отметим лишь некоторые важные моменты, имеющие значение для анализа данных из любой области.

Энтропию Реньи нулевого порядка R_0 называют обычно структурной энтропией, так как эта энтропия несет информацию о том, как плотно исследуемый объект заполняет собой пространство вложения. В данном случае это просто логарифм числа пикселей, то есть площадь, занятая той или другой полярностью при выбранном уровне дискриминации (отсечки) по полю. При получении результатов, представленных на рис. 3.7, уровень отсечки был взят равным ± 100 гс, то есть рассматривались только умеренные и сильные поля. Для той и другой АО выборки представляют фрагменты рядов примерно из 70-ти последовательных магнитограмм с интервалом 96 минут, то есть временной интервал выборок составляет около пяти суток.

Две верхние панели на рис. 3.8 представляют временные вариации энтропии Реньи нулевого порядка N и S полярностей в той и другой АО, рассчитанных на покрытиях с минимально возможным размером элементов. Обращает на себя внимание синхронная депрессия структурных энтропий этих двух АО в районе 29.5 суток. Частично это сходство может быть обусловлено не с изменениями в структуре самих активных областей, а с воздействием на аппаратуру SOHO/MDI послевспышечных протонных потоков. Общим для этих АО является также тот факт, что в той и в другой области энтропия южной полярности была больше энтропии северной полярности.

Но в остальном АО ведут себя по-разному. Энтропии Реньи, как отмечалось выше, являются экстенсивными переменными. В рассматриваемом случае временные изменения R_{n0} и R_{s0} описывают изменение со временем логарифма площади магнитного потока. В растущей АО 10488 (правый верхний рисунок) рос магнитный поток и, соответственно, рос логарифм ее площади. В этой активной области тренды обеих полярностей однонаправленные, то есть с ростом магнитных потоков в обеих по-

лярностях практически синхронно росли и их энтропии. В АО 10486, как можно видеть из левой верхней панели на рис. 3.8, структурные энтропии N и S полярностей на рассматриваемой выборке демонстрируют разнонаправленные тренды. Иными словами, логарифм площади магнитного потока южной полярности увеличивался, в то время как поток в южной полярности стагнировал с небольшим понижением.

Две средние панели на рис. 3.8 показывают временные вариации энтропии Реньи первого порядка R_1 . Исторически эту энтропию называют информационной, так как она рассчитывается по формуле, похожей на формулу информационной энтропии Шеннона (см. предыдущий раздел). Ее величина дает количественную меру степени хаотичности распределения «массы» рассматриваемого объекта и позволяет проследить динамику переходов хаос–порядок и порядок–хаос. В рассматриваемом примере эта энтропия представляет меру хаотичности в распределении магнитного потока и в этом плане позволяет анализировать изменение со временем магнитных потоков в N и S полярностях в активных областях.

На левом среднем рисунке (АО 10486) можно видеть существенный рост со временем хаотичности распределения магнитного поля в S полярности и, напротив, упрощение поля в N полярности. В АО 10488 (правый средний рисунок) видим устойчивый рост этой энтропии в той и другой полярности. Следует отметить, что в силу определения этой энтропии шумы аппаратуры, как можно видеть из этих рисунков, меньше сказываются на общих тенденциях ее временных изменений.

Две нижние панели на рис. 3.8 демонстрируют временные вариации энтропии Реньи второго порядка. Как уже отмечалось в предыдущем разделе, эту энтропию называют корреляционной, так как она дает меру для оценки пространственной корреляции элементов структуры исследуемого объекта. В рассматриваемом примере это мера коррелированности по положению магнитных силовых трубок в той и другой полярности. Рисунок позволяет проследить временные вариации этой величины как в разных полярностях, так и в двух рассматриваемых АО. Уменьшение этой энтропии говорит о росте корреляций, а ее увеличение соответствует уменьшению коррелированности элементов и росту хаотичности в системе. Для случая АО 10486 (левая нижняя панель на рис. 3.8) мы видим рост этой энтропии в северной полярности и уменьшение ее в южной полярности. В АО 10488 (правая нижняя панель) можно наблюдать устойчивое падение этой энтропии, то есть рост коррелированности магнитных элементов в той и в другой полярности.

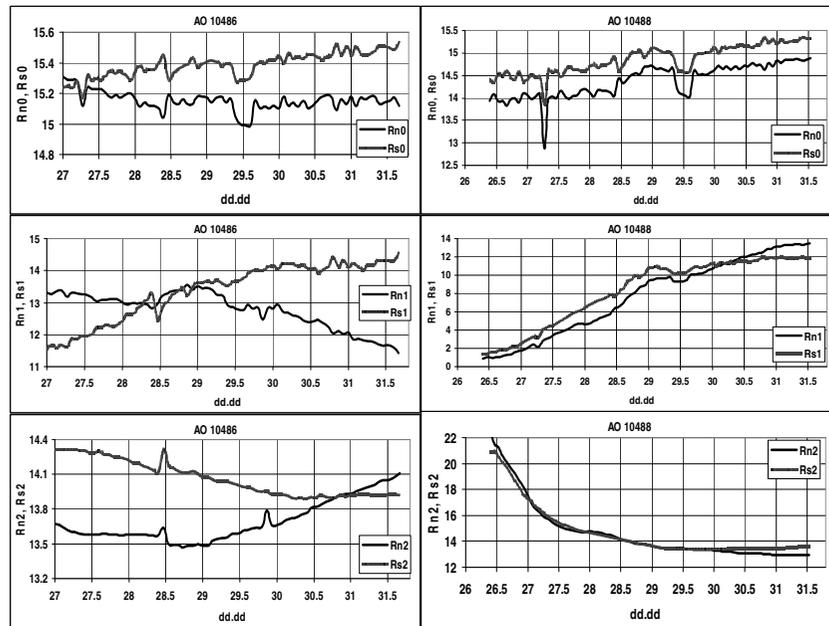


Рис. 3.8. Временные вариации энтропии Реньи порядка $q = 0, 1$ и 2 в двух солнечных активных областях. По оси абсцисс — время в сутках и долях суток октября 2003 г.; по оси ординат — энтропии полярностей: Rn_0, Rs_0 — энтропии Реньи нулевого порядка северной и южной полярностей, соответственно; Rn_1, Rs_1 — энтропии Реньи первого порядка; Rn_2, Rs_2 — энтропии Реньи второго порядка. Все энтропии рассчитывались для покрытий с минимально возможным размером ячеек

Итак, суммируя все вышесказанное по этим двум АО, можно отметить следующее. Характер роста площади и дисперсии магнитных потоков в S полярности АО 10486 при одновременном росте корреляций магнитопотоковых элементов говорит об особенностях развития во времени полного магнитного потока в этой полярности. Тенденции к уменьшению площади и дисперсии магнитных потоков в N полярности при уменьшении корреляции магнитопотоковых элементов описывают особенности стагнации и деградации поля в N полярности. Наличие таких двух противоположных тенденций во временном изменении магнитных потоков противоположной полярности в рамках одной активной области говорит об открытости магнитной конфигурации или замыкании избыточного потока в S полярности на другие АО.

В АО 10488 можно наблюдать другую эволюцию магнитного поля солнечной активной области. Здесь с синхронным ростом в обеих полярностях логарифма площади и дисперсии значений элементарных магнитных потоков наблюдается также синхронный рост коррелированности магнитопотоковых элементов в той и другой полярностях, переходящий в тенденцию насыщения при значениях $R_2 \sim 13$. Все это говорит о почти полной замкнутости структуры и быстром развитии системы магнитных потоков в данной АО.

Разумеется, анализ временных изменений трех энтропий Реньи не заменяет исследование других интегральных параметров, характеризующих общую динамику магнитного поля в солнечных активных областях. Однако наряду с анализом изменений других параметров, характеризующих магнитные поля солнечных АО, таких как полные потоки, параметры топологической сложности, имбаланс потоков в полярностях и другие (см., например, [12–18]), расчет энтропий Реньи дает в руки исследователей дополнительную важную информацию. Анализ этой информации позволяет получать более полную картину временной эволюции солнечных активных областей и более полные представления о генерации и последующей динамике сильных локальных магнитных полей в верхних слоях конвективной зоны Солнца.

Таким образом, как можно видеть из этого примера, расчет энтропий Реньи позволяет получить важную качественную и количественную информацию о структуре сложных физических систем, особенностях их динамической эволюции и о физических процессах, происходящих в таких системах.

Список литературы

1. Климонтович Ю. Л. Статистическая физика. М.: «Наука», 1982. 608 с.
2. Климонтович Ю. Л. Введение в физику открытых систем. М.: «Янус-К», 2002. 284 с.
3. Эбелинг В. Образование структур при необратимых процессах. М.: Мир, 1979. 279 с.
4. Коротеев С. М. Энтропия и информация — универсальные естественнонаучные понятия. 2002. 350 с. <http://www.chronos.msu.ru>.
5. Князева Е. Н., Курдюмов С. П. Синергетика и принципы коэволюции сложных систем // Категории. 1997. № 3. С. 37–49.
6. Вернадский В. И. Пространство и время в неживой и живой природе. М.: «Наука», 1975.

7. Eckmann J.-P. and Ruelle D. Rev. Mod. Phys. 57, 1985. P. 617.
8. Pesin Ya. B. Dimension theory in dynamical systems / university of Chicago Press, 1998.
9. Amigó J. M., Kennel M. B., Kocarev L. The permutation entropy rate equals the metric entropy rate for ergodic information sources and ergodic dynamical systems. Physica D: Nonlinear Phenomena V. 210, Iss. 1–2, 1 October 2005. P. 77–95.
10. Zunino L., Pérez D. G., Martín M. T., Garavaglia M., Plastino A., Rosso O. A. Permutation Entropy Of Fractional Brownian Motion And Fractional Gaussian Noise. Physics Letters A. V. 372. Iss. 27–28, 2008. P. 4768–4774.
11. <http://sohowww.nascom.nasa.gov/publications/CDROM1/papers/index.html>.
12. Чумак О. В., Кононович Э. В., Красоткин С. А. Интегральные характеристики солнечных активных областей. Известия РАН. Серия физическая. Т. 62. No. 9. 1998. С.1879–1882.
13. Chumak O., Zhang H. and Guo J. Integral Properties of the Magnetic Fields of Solar Active Regions Under Quiet and Flare Activity Conditions. Astron.&Astroph. Trans. 2004. Vol. 23. No. 6, P. 525–531.
14. Chumak O., Zhang H.-Q. Integrated Characteristics of the Radial Magnetic Field in Solar Active Regions during Quiet and Flare-Productive Phases of Their Evolution. Astronomical Reports. V. 49. No. 9, 2005. P. 755–766.
15. Guo J., Zhang H., Chumak O., Liu Y. A Quantitative Study on Magnetic Configuration for Active Regions. Solar Physics. V. 237. N. 1, 2006. P. 25–43.
16. Guo J., Zhang H., Chumak O. Magnetic properties of flare-CME productive active regions and CME speed. Astron&Astroph., 462, 2007. P. 1121–1126.
17. Chumak O. V., Zhang H., Guo J. Magnetic-Field Variations in the Active Region NOAA 10486 and Their Relationship to X-Ray Flares and Coronal Mass Ejections. Astronomy Reports, 2008. Vol. 52. No. 10. P. 852–865.
18. Guo J., Zhang H. Q., Chumak O. V., Lin J. B. The evolution of the quantified magnetic complexity of active regions in the 23rd solar cycle. Monthly Notices of the Royal Astronomical Society, 2010. V. 405, Issue 1. P. 111–114.

ГЛАВА 4 ФРАКТАЛЫ И ИХ СВОЙСТВА

4.1. Конструктивные фракталы

В этом разделе мы познакомимся с рядом абстрактно-математических фрактальных объектов, так называемых геометрических или конструктивных фракталов, и техникой расчета их фрактальных размерностей. Как уже отмечалось ранее, термин «фрактал» (лат. fractus — изломанный, дробный) был введен Бенуа Мандельбротом в 1975 г. [1]. Однако как предмет исследования объекты такого рода изучались гораздо раньше. Так, Карл Теодор Вильгельм Вейерштрасс (работы 1895 г.) привел пример непрерывной функции, не имеющей производной ни в одной точке. Еще один пример подобной функции, ставший классическим, привел Анри Лебег [2]. Рассмотрим пример Лебега более подробно.

4.1.1. Парадокс Лебега и производные Гельдера

Рассмотрим равносторонний треугольник ABC (рис. 4.1). Соединим середины трех сторон двумя отрезками, как это показано на рис. 4.1. Очевидно, что $AB + AC = BD + DE + EF + FC$. Продолжаем этот процесс и приближаем ломанную к стороне BC. Если BC есть предел ломанной, то $AB + AC = BC$. В чем причина парадокса?

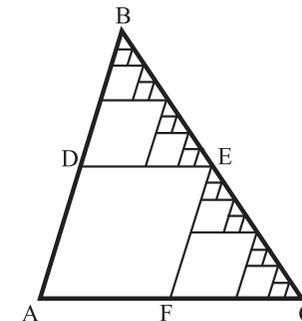


Рис. 4.1. Парадокс Лебега (рисунок из книги Лебега [2])

Здравый смысл подсказывает, что наша «пила» приближается к ВС по положению, но не по направлению. Фактически, при приближении к пределу элементы «пилы», оставаясь самоподобными, измельчаются, и производная становится неопределенной. Таким образом, все дело в гладкости. Для оценки степени гладкости функций О. Л. Гельдером (1889 г.) была предложена такая процедура. Зафиксируем некоторое произвольное значение аргумента x' функции $f(x)$ и рассмотрим отношение:

$$\frac{|f(x) - f(x')|}{|x - x'|^g} = H(x'); g \in [0, 1]. \quad (4.1)$$

Если на всей области определения функции $f(x)$ отношение (4.1) ограничено сверху, то говорят, что функция $f(x)$ удовлетворяет условию Гельдера в точке x' с показателем g и коэффициентом $H(x')$. Объекты с гельдеровскими показателями $g < 1$ называются фрактальными объектами.

В дальнейшем мы познакомимся с другими подходами, видами и определениями фрактальных объектов, но данное определение является ценным в плане рассмотрения различных природных объектов и стохастических временных рядов. Предел

$$h(x) = df/(dx)^g, \quad (4.2)$$

если он существует, называется гельдеровскими производными. Их две, так как, согласно (4.1), в каждой точке x' может существовать производная справа и слева, и эти производные могут быть, вообще говоря, неравны. В качестве примера можно привести броуновский процесс, в котором смещение $dx(t) \sim (dt)^{1/2}$. В этом случае показатель Гельдера $g = 1/2$. То есть функция $x(t)$ имеет только «половину» производной.

Подобные непрерывные, но нерегулярные функции, не имеющие обычных производных, были хорошо известны математикам 19 века. У многих они вызывали сложные чувства. «С омерзением и ужасом я отворачиваюсь от этой зловерной язвы — непрерывных функций, нигде не имеющих производных», — писал Эрмит Стильтессу в 1893 г. Однако те математики, которые творили этих «монстров», считали их важными «прецедентами», свидетельствующими о том, что в недрах чистой математики имеются неиссякаемые возможности для описания самых сложных и необычных структур с помощью простых моделей. Тем самым часто не до конца осознано они подтверждали известный феномен необычайной эффективности математики в приложении ее к естественным наукам. Ведь эти «монстры» говорили о том, что явления, процессы или

объекты, на первый взгляд, необычайно *сложные*, могут быть сведены к *простым* моделям, часто содержащим всего несколько элементов и итеративных правил. «Я намерен показать, — пишет в этой связи Б. Мандельброт (2002 г.) [1], — что за упомянутыми безумными творениями лежат необъятные миры, которых так и не увидели ни их создатели, ни несколько поколений последователей, — миры, которые будут небезынтересны тем, кто воспекает Природу, стремясь ей подражать».

Ниже для примера рассмотрим несколько таких объектов.

4.1.2. H-фрактал

Рассмотрим другое определение фрактального объекта. Пусть у нас имеется некоторый исходный геометрический элемент и итеративное правило, воспроизводящее этот элемент на разных масштабах. Таким путем мы можем получить целое семейство объектов, которое обычно называют геометрическими или конструктивными фракталами. Некоторые из них можно получить с помощью уже упомянутой программы Fractint, с другими можно познакомиться, прочитав монографии и учебные пособия по фрактальному анализу. Здесь мы частично следуем наиболее ясному, на наш взгляд, изложению этой темы, которое содержится в монографии Д. А. Морозова (2002 г.) [2]. Оттуда же взяты некоторые рисунки.

Одним из простейших объектов такого типа является так называемый H-фрактал (рис. 4.2).

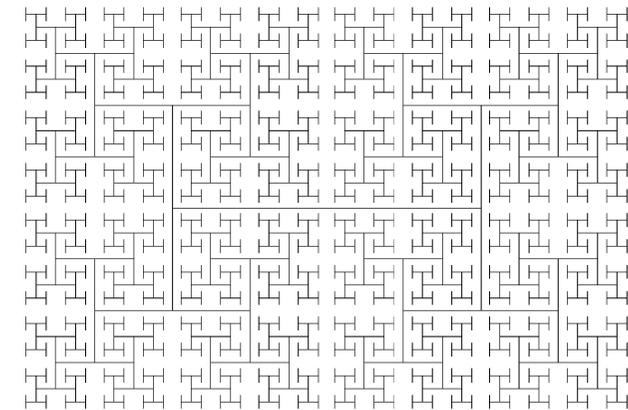


Рис. 4.2. H-фрактал

Исходным геометрическим элементом в этом случае является нечто, похожее на букву Н. Итеративное правило заключается в воспроизводстве этого элемента. Делается это так. На концах единичного горизонтального отрезка в центре рисунка помещаем два более коротких перпендикулярных отрезка. Далее эта операция повторяется с двумя более короткими отрезками и так далее. Теоретически эту процедуру можно продолжать до бесконечности. В итоге полученная фигура является фракталом, в котором один и тот же элемент повторяется многократно в разных масштабах. Этот фрактал демонстрирует одно из важнейших свойств геометрических (конструктивных) фракталов – их масштабную инвариантность.

4.1.3. Канторова пыль

Другим примером конструктивного фрактала является очень интересный фрактал Кантора. Способ его построения показан на рис. 4.3. Из единичного отрезка удаляется его средняя треть. С оставшимися двумя отрезками проделывается та же процедура и так далее. Как видно из рисунка, после трех шагов получаем $2^3 = 8$ отрезков. После N шагов, соответственно, $n = 2^N$ отрезков. После первого шага длина одного отрезка равна $1/3$, после второго — $1/9$, после третьего — $1/27$ и так далее. Соответственно, на N -том шаге длина одного отрезка равна $a = (1/3)^N$. Общая длина всех отрезков на N -том шаге равна $2^N \cdot (1/3)^N = (2/3)^N$. Устремляя число шагов N в бесконечность, видим, что общая длина фрактала Кантора стремится в ноль. Иными словами, и размерность Лебега или топологическая размерность фрактала Кантора равна нулю.

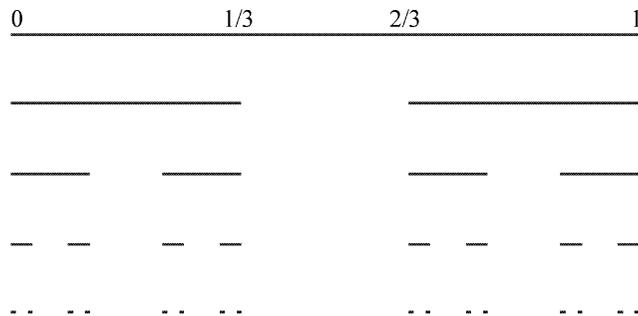


Рис. 4.3. Построение канторова множества «удаленной средней трети»: делим отрезок на три равных части и удаляем среднюю треть, затем то же проделываем с оставшимися двумя третями и так далее

Но можно ввести еще одну размерность, которая будет не нулевой. Рассуждаем следующим образом. Пусть у нас имеется отрезок единичной длины. Разобьем его на n равных частей длиной a . Тогда $n \cdot a = 1$, а $n = 1/a$. Пусть имеется квадрат со стороной, равной единице. Разобьем его на n равных квадратов со стороной a . Тогда $n \cdot a^2 = 1$, или $n = 1/a^2$. Для куба $n = 1/a^3$. В общем случае $n = 1/a^d$, где d — размерность объекта, для которой можно записать:

$$d = \frac{\log n}{\log \left(\frac{1}{a} \right)}. \quad (4.2)$$

Введенная таким образом размерность является частным случаем размерности Хаусдорфа. Ее также часто называют фрактальной размерностью.

В нашем случае число отрезков n на N -том шаге равно $n = 2^N$, а длина отрезка на N -том шаге, соответственно, равна $a = (1/3)^N$. Подставляя эти значения в формулу (4.2), получаем $d = \log(2)/\log(3) = 0.63\dots$ Как видим, размерность фигуры Кантора не является целым числом. Это дробь, значение которой заключено между двумя целыми размерностями: 0 — размерностью точки и 1 — размерностью линии. Образно говоря, фрактал Кантора еще не точка, но уже не линия. Фрактал Кантора обладает множеством интересных свойств, с которыми можно познакомиться в монографиях по фрактальному анализу.

4.1.4. Фигуры Коха

Интересным примером фрактального объекта является кривая Коха. Особенность этой кривой заключается в том, что она не имеет касательной. Строится она следующим образом. Как и в случае фрактала Кантора, из единичного отрезка вырезаем среднюю третью часть. Но в отличие от фрактала Кантора, заменяем ее на две стороны равностороннего треугольника, с длиной стороны, равной одной трети отрезка (см. рис. 4.4, шаг $N = 1$). Число отрезков в фигуре $n = 4$, а длина отрезка $a = 1/3$. На следующем шаге (см. рис. 4.4, шаг $N = 2$) $n = 4^2$, а длина отрезка равна $a = (1/3)^2$. На N -том шаге, соответственно, $n = 4^N$, а длина отрезка $a = (1/3)^N$. Подставляя эти значения в формулу (4.2), получаем значение фрактальной размерности кривой Коха $d = \log(4)/\log(3) = 1.26\dots$

Размерность кривой Коха, как видим, тоже дробная, но ее значение заключено между 1 — размерностью линии и 2 — размерностью плоскости. То есть она вышла за размерность линии и частично заполняет плоскость. Кривая Коха самоподобна, как мы видим, любой ее фрагмент представлен на разных масштабах.

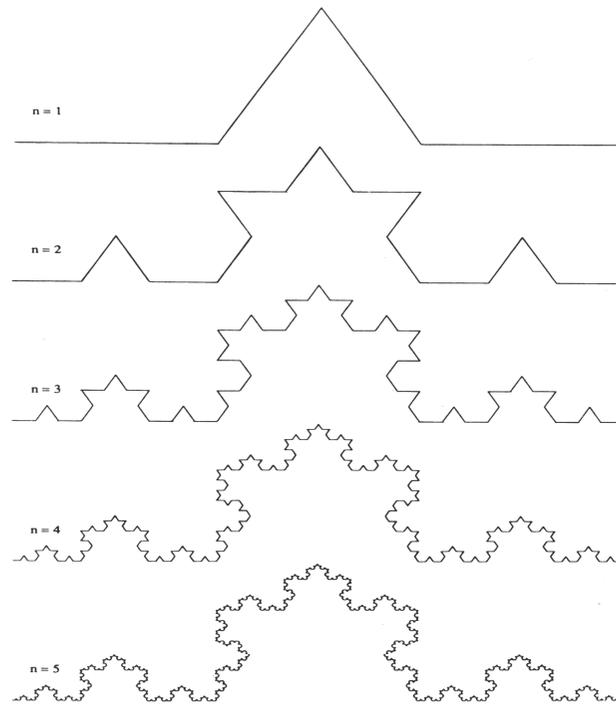


Рис. 4.4. Построение триадной кривой Коха: $n = 1$ — основной генерирующий элемент, $n = 2$ — заменяем все четыре отрезка основного элемента на генерирующий элемент в соответствующем масштабе, $n = 3$, $n = 4$ и $n = 5$ — следующие аналогичные итерации.

Есть много вариаций на тему кривой Коха. Мы можем взять равно-сторонний треугольник, квадрат или другой многоугольник и на каждой его стороне «вырастить» фигуру Коха. Получается множество разнообразных красивых объектов.

4.1.5. Фрактал Минковского

Подобным образом строится и фрактал Минковского. На рис. 4.5 представлен исходный элемент (основа) и фрагмент этого фрактала. Повторяя шаг за шагом процедуру, аналогичную рассмотренным выше для кривой Коха и пыли Кантора, мы получаем фрактал Минковского (рис. 4.6). На те-

му этого фрактала тоже можно найти множество разных вариаций. Снова воспользуемся формулой (4.2) для вычисления фрактальной размерности. Здесь $n = 8$, а длина отрезка $a = 1/4$. Соответственно, $d = \log(8)/\log(4) = 1.5$.

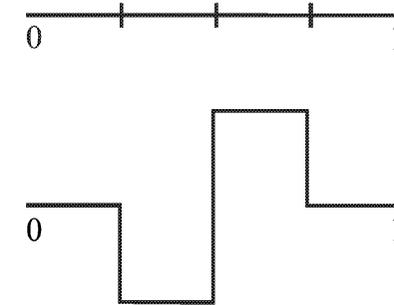


Рис. 4.5. Основа и фрагмент для построения фрактала Минковского

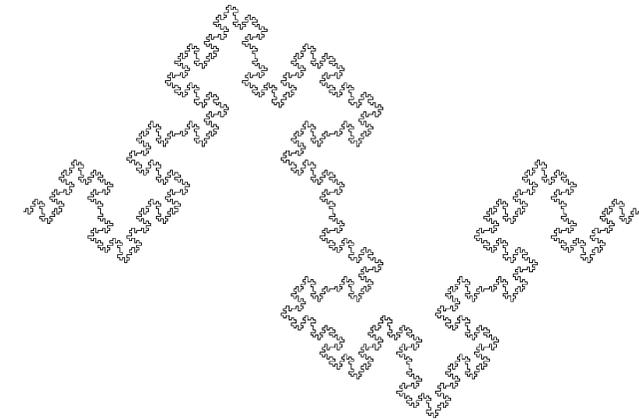


Рис. 4.6. Фрактал Минковского

Это значение размерности, как и значение размерности фрактала Коха, лежит между нулем и единицей, но по величине размерность фрактала Минковского больше размерности фрактала Коха. Иными словами, фрактал Минковского более плотно заполняет двумерье.

4.1.6. Фракталы Серпинского

Двумерным аналогом фрактала Кантора является показанный на рис. 4.7 прямоугольный ковер Серпинского. В единичном квадрате удаляется малый квадрат на пересечении прямоугольников, шириной которых являются центральные трети сторон (рис. 4.7, левая панель). На оставшихся восьми квадратах процедура повторяется (рис. 4.7, вторая панель), и так далее. Таким образом, здесь $a = 1/4$ и $n = 8$. Из формулы (4.2) получаем $d = 1.89\dots$ Такое значение размерности говорит о том, что прямоугольное решето Серпинского существенно плотнее заполняет плоскость, чем фракталы Коха и Минковского.

Этот вариант ковра Серпинского легко обобщается на единичный куб Серпинского – эдакий куб сыра с большой кубической дыркой в центре и множеством кубических дырок разных размеров во всем остальном объеме. Нетрудно сообразить, что в этом случае $n = 26$, $a = 1/3$ и $d = 2.96\dots$ Как видим, размерность куба Серпинского близка к размерности объема 3. Нетрудно также обобщить этот результат на k -мерный куб Серпинского: $n = (3^k - 1)$, $a = 1/3$, соответственно, $d = \log(3^k - 1)/\log(3)$. Это значение размерности всегда чуть меньше размерности пространства вложения k , так как k -мерный куб Серпинского — достаточно плотное тело с небольшими пустотами.

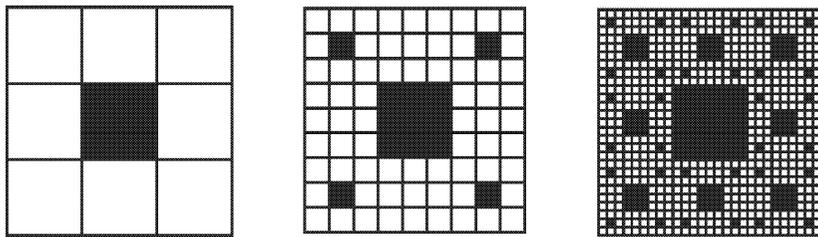


Рис. 4.7. Прямоугольный ковер Серпинского ($n = 8$, $a = 1/3$, $d = \log(8)/\log(3) = 1.89\dots$)

Другой вариант ковра Серпинского показан на рис. 4.8. Здесь исходным элементом является равносторонний единичный треугольник (рис. 4.8а). На первом шаге середины сторон соединяются отрезками, образованный ими центральный треугольник удаляется (рис. 4.8б).

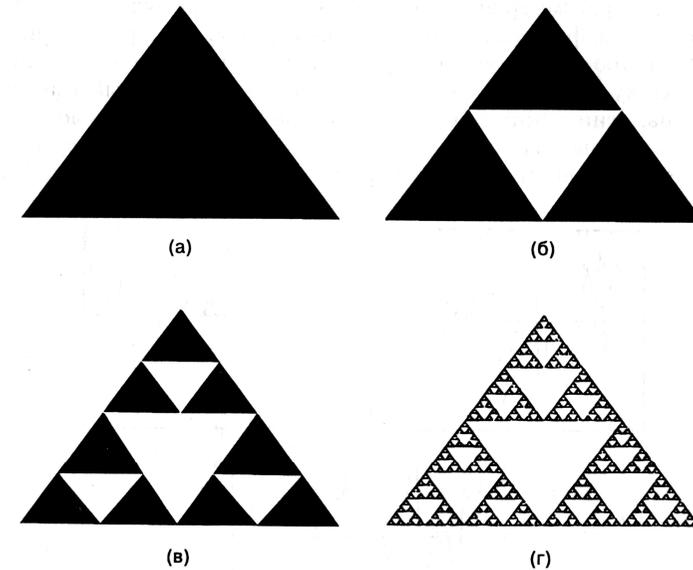


Рис. 4.8. Построение треугольного ковра Серпинского

То же самое проделываем с каждым из трех сплошных треугольников (рис. 4.8в) и, продолжая итерации, получаем множество треугольников в треугольнике (рис. 4.8г). В этом случае $n = 3$, $a = 1/2$, $d = \log(3)/\log(2) = 1.58\dots$

4.1.7. Круговые фракталы

В работах [14–19] рассматривались конструктивные круговые фракталы в качестве моделей «упаковки» магнитопотоковых трубок в солнечных пятнах и в солнечных активных областях. В отличие от ковров Серпинского они строятся не из прямоугольных и треугольных элементов, а из круговых. Варианты таких конструктивных кварталов приведены на рис. 4.9.

Для расчета размерностей этих объектов снова воспользуемся формулой (4.2). Как показано на рис. 4.9а, значение $n = 3$. Величина параметра a — отношение характерных длин соседних масштабов, то есть r_{i+1}/r_i — где r_i — радиус большей окружности, r_{i+1} — радиус окружности соседнего меньшего масштаба. Из простых геометрических соображений находим: $a = 0.4641$. Подставляя эти значения в формулу (4.2), получаем $d = 1,43\dots$

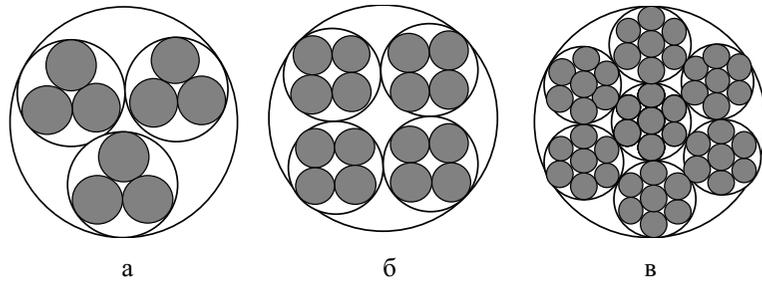


Рис. 4.9. Варианты круговых фракталов

Для варианта на рис. 4.9б, соответственно, $n = 4$, $a = 0.4142\dots$, $d = 1.57\dots$. Наконец, для варианта, изображенного на рис. 4.9в, имеем: $n = 7$, $a = 1/3$, размерность $d = 1.77\dots$. Эти значения размерностей круговых фракталов будут упоминаться ниже, при рассмотрении примеров применения фрактальной геометрии к анализу структурных особенностей сильных локальных магнитных полей в солнечных активных областях.

Резюмируем этот раздел. Математические фракталы отличаются от гладких аналитических функций и традиционных геометрических объектов такими принципиальными свойствами как отсутствие гладкости или наличием разрывов. Простые фракталы можно сконструировать путем выбора некоторого базового геометрического элемента и повторением его на бесконечной иерархии масштабов, вследствие чего такие фракталы обладают свойствами самоподобия и масштабной инвариантности. Фрактальная (Хаусдорфова) размерность таких объектов может быть рассчитана по простой формуле (4.2).

4.2. Природные фракталы

В случае конструктивных фракталов, как мы видели из предыдущего раздела, параметр, количественно характеризующий меру заполнения фрактальным множеством пространства вложения, — его фрактальная размерность — вычисляется просто. Сложнее обстоят дела с подобной параметризацией природных фракталов. Здесь, как правило, отсутствует точное самоподобие, вместо строго геометрического самоподобия элементов фрактала наблюдается лишь их топологическое или статистическое самоподобие, то есть *самоафинность* (Мандельброт, 1988 г.) [4]. Кроме того, глубина иерархии масштабов всегда конечна.

4.2.1. Размерность Хаусдорфа. Бокс алгоритм. Скейлинг

Чтобы разобраться, как параметризуют такие фракталы, рассуждаем следующим образом. Чтобы измерить длину, мы обычно берем некоторый масштаб a и прикладываем его к объекту измерений N раз. Величина $N(a) \sim a^{-1}$ и есть примерная (с точностью a) длина нашего объекта, выраженная в масштабе (шкале) a . Чтобы измерить площадь, мы покрываем ее квадратами со стороной a . Тогда число таких квадратов $N(a) \sim a^{-2}$ есть площадь нашего объекта. Для объема мы получаем, соответственно, $N(a) \sim a^{-3}$. Для d -мерного пространства получаем, соответственно, $N(a) \sim a^{-d}$. Коэффициент пропорциональности в этих соотношениях всегда есть некоторая конечная постоянная.

Точный размер объекта может быть получен при переходе к пределу:

$$\lim_{a \rightarrow 0} N(a)a^d. \quad (4.3)$$

Вопрос в том, всегда ли такой предел существует. И если да, то всегда ли при этом d является целым числом?

Как мы уже знаем, d является целым числом, если объект непрерывен, то есть всюду плотно заполняет собою пространство вложения. Если же это не так, а для природных объектов это всегда *не так*, то предел (4.3) при целых d может не существовать.

Хрестоматийный пример — определение длины побережья Англии, определенной Ричардсоном, о котором можно подробно прочитать практически во всех книгах по фрактальному анализу. Ричардсон, по-видимому, был первым, кто показал, что бывают случаи, когда при целых d предел (4.3) не существует, но существует при не целом d :

$$\lim_{a \rightarrow 0} Na^d = \text{const} \quad 0 < \text{const} < \infty. \quad (4.4)$$

d — это тот самый показатель, при котором произведение Na^d остается конечным и отличным от нуля. Такая размерность в силу своего дробного характера называется фрактальной.

Немецкий математик Феликс Хаусдорф (1868–1942) обобщил и обосновал идею Ричардсона о дробных размерностях. Размерность Хаусдорфа d обычно задают выражением

$$d_0 = \lim_{a \rightarrow 0} \frac{\log N(a)}{\log(1/a)}, \quad (4.5)$$

где N — минимальное число элементов размера a , необходимое для покрытия рассматриваемого множества.

На рисунке 4.10, взятом из книги Е. Федера (1991 г.) [4], показана техника получения фрактальной размерности побережья южной части

Норвегии. Покроем рассматриваемый объект квадратными клетками. Пусть эти клетки имеют размеры $a*a$, а $N(a)$ — число таких клеток, необходимых для того, чтобы покрыть рассматриваемый участок береговой линии на карте. Такой способ оценки размеров объекта обычно называют «*бокс алгоритмом*». В случае рассматриваемой береговой линии $N(a)$ примерно равно числу шагов, которое необходимо, чтобы пройти по карте вдоль этой линии циркулем с раствором a , как это делал Ричардсон, измеряя береговую линию Великобритании. Понятно, что, уменьшая размер a , мы увеличиваем число $N(a)$. Если бы длина береговой линии Великобритании или Норвегии имела определенную длину, то величина $L(a) = N(a)*a$ при уменьшении a стремилась бы к некоторой постоянной величине L_N . На самом деле это не так. С уменьшением a величина $L(a)$ неограниченно возрастает.

Ричардсон построил график $\log(L(a)) - \log(a)$ и показал, что в рассмотренном им диапазоне масштабов все точки ложатся на прямую линию с коэффициентом наклона $k = 1 - d$. То есть $L(a) = b*a^{1-d}$. Здесь значение размерности d является дробным числом и потому называется фрактальной размерностью. Процедуру многократного измерения длины, площади или объема с помощью разных по величине элементов покрытия обычно называют *скейлингом*, а график $\log(L(a)) - \log(a)$ — *графиком Ричардсона*. В настоящее время именно путем построения графиков Ричардсона обычно определяют фрактальные размерности различных природных структур в границах тех или иных рассматриваемых масштабов. Преимуществами метода Ричардсона являются его простота, наглядность, отсутствие предельных переходов к бесконечно малым элементам покрытия (что практически никогда невозможно реализовать в прикладных задачах), устойчивость при конечной иерархии масштабов. Определенную таким способом фрактальную размерность называют иногда «клеточной» размерностью.

Определенное методом Ричардсона значение фрактальной размерности для участка побережья Норвегии, представленного на рис. 4.10, оказалось равным $d_N = 1.52 \pm 0.01$; для побережья Великобритании — $d_G = 1.3...$

Для разных побережий фрактальная размерность d принимает различные значения, лежащие в пределах от 1 до 2. То есть изрезанные береговые линии, с геометрической точки зрения, это уже не одномерные объекты, но еще не двумерные. В данном случае параметр d , будучи числом геометрической величиной, характеризующей степень изрезанности береговых линий, в конечном счете отражает физику геологических процессов, которые привели к образованию того или иного вида побережья. Таким образом, расчет фрактальных размерностей оказался ценным ин-

струментом для практических приложений, так как дал в руки исследователей легко определяемый безразмерный параметр, позволяющий классифицировать сложные геологические процессы. Пример с береговой линией является типичным примером параметризации сложных объектов с помощью фрактальных размерностей.

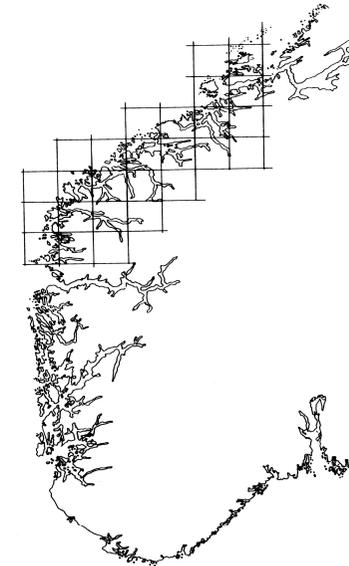


Рис. 4.10. Побережье южной части Норвегии из книги Е. Федера [5]. Изображенная сверху квадратная решетка имеет шаг $a \sim 50$ км

Таким образом, для фрактальной параметризации различных природных самоафинных объектов, таких как, например, береговые линии, обычно применяется так называемый «бокс алгоритм». То есть подсчитывается минимальное число клеток заданного размера $N(a)$, покрывающих участок побережья. Затем строится график Ричардсона: $\log N(a)$ от $\log(1/a)$. Если в этих координатах мы получали прямую, то это означает, что объект обладает масштабно инвариантными (*скейлинговыми*) свойствами, а тангенс угла наклона прямой дает нам размерность объекта, вообще говоря, фрактальную.

В случае природных объектов, в отличие от рассмотренных «идеальных» математических фракталов, базовый исходный элемент и итеративное правило в явном виде отсутствуют. Тем не менее, масштабно инвариантное (*скейлинговое*) поведение явно имеет место. Такой скейлинг

в литературе часто называют *стохастическим скейлингом*, а объекты, соответственно, стохастическими или самоафинными фракталами (Мандельброт, 1988) [4].

4.2.2. Соотношения «масса-радиус», «периметр-площадь», «площадь-объем»

Массовая размерность

Далеко не всегда при определении фрактальной размерности природных объектов удобно применять бокс алгоритм. Рассмотрим, например, сферический комочек пуха, фотографию шарового или рассеянного звездного скопления или другие подобные объекты. Для таких объектов удобнее рассматривать связь между радиусом и массой объекта, заключенного под этим радиусом:

$$M \sim R^d. \quad (4.6)$$

Для сплошной однородной плоской фигуры, как известно, $d = 2$. Например, масса диска с плотностью $\rho_0 = \text{const}$ и с площадью $S = \pi R^2$. Для однородного трехмерного тела, соответственно, $d = 3$. Однако для пуха, поролона или звездного скопления этот показатель может быть не целым. Манфред Шредер [6] в уже цитированной книге «Фракталы, хаос, степенные законы» пишет, что для пуха этот показатель равен ~ 1.6 , и чем он меньше, тем пух дороже. Размерность, определенную по такому правилу, называют «массовой» или «емкостной» размерностью, так как она определяется через подсчет количества элементов объекта, находящихся внутри сферы или круга заданного радиуса.

Возникает вопрос: равна ли фрактальная размерность, определенная по такому правилу, размерности Хаусдорфа?

Рассмотрим треугольный ковер Серпинского (рис. 4.8). Из точки пересечения медиан мысленно опишем окружность, проходящую через вершины треугольника. Сторона треугольника $a \sim R$. Переходя к следующему шагу, мы удваиваем масштаб, при этом количество треугольников, а следовательно, и их площадь увеличивается не в четыре, а в три раза. Таким образом, когда радиус круга удваивается, суммарная площадь треугольников S возрастает в три раза, а не в четыре, $S \sim R^{3/2}$. Если все треугольники изготовлены из материала с однородной плотностью, то их масса $M \sim S$ и, таким образом, $M \sim R^{3/2}$; $d_0 = \log 3 / \log 2 = 1.58\dots$ В данном случае массовая размерность совпадает с размерностью Хаусдорфа.

Рассмотрим квадратный ковер Серпинского. Здесь при увеличении описанной окружности в три раза площадь фигуры увеличивается в четыре раза. $d_0 = \log 8 / \log 3 = 1.89\dots$, что также совпадает с размерностью Хаусдорфа. Можно показать, что и для других самоподобных (идеальных) фракталов эти размерности совпадают. Однако в случае природных фракталов, как показал М. Шредер, это не всегда так. Мы не будем более подробно останавливаться на этом вопросе и заинтересованных читателей отсылаем к книге М. Шредера (2001) [6]. Отметим лишь, что на практике для таких природных объектов как различные дендроидные структуры, пористые среды, среды, в которых присутствуют агрегаты кластеров на разных масштабах, и для других подобных структур оценки массовой размерности оказываются очень полезными и информативными. Массовая размерность позволяет количественно оценивать физические и химические характеристики процессов, ведущих к образованию структур, дает информацию относительно получения новых материалов с заданными свойствами и так далее.

Соотношения «периметр – площадь» и «поверхность – объем»

Сейчас рассмотрим соотношения «периметр – площадь» и «поверхность – объем» и их связь с фрактальной размерностью. Этот способ определения размерностей объектов также очень полезен с практической точки зрения.

Как известно, у плоских евклидовых фигур — окружностей, квадратов, многоугольников — отношение периметра L к корню квадратному из площади S есть величина постоянная:

$$L/(S)^{1/2} = \text{const}. \quad (4.7)$$

Если периметр представляет собой фрактальный объект (это может быть береговая линия, стохастический временной ряд и так далее) и длина профиля объекта зависит (по степенному закону) от масштаба a , с помощью которого эта длина измеряется, то, как показал Мандельброт [1], соотношение (4.7) может быть записано в виде:

$$L^{1/d}/(S)^{1/2} = \text{const}, \quad (4.8)$$

где d — фрактальная (Хаусдорфова) размерность. Таким образом,

$$S \sim L^{2/d}. \quad (4.9)$$

Этим соотношением можно воспользоваться для определения фрактальных размерностей различных природных объектов.

Таким способом, например, автором в 1994 г. [7] была определена Хаусдорфова размерность теней солнечных пятен. Она оказалась равной

$d_0 = 1.35 \pm 0.03$, что соответствует размерности так называемых «эластичных скелетов», возникающих в перколяционных структурах плазменной турбулентности [8]. Оказалось также, что эта размерность равна фрактальной размерности границ облаков в земной атмосфере. Интересно отметить, что градовые облака имеют, как правило, большую размерность. Таким образом, было получено дополнительное подтверждение о магнитно-турбулентной природе солнечных пятен. Физический смысл этой размерности обсуждается в обзоре Л. М. Зеленого и В. А. Милованова [9], где показано, что «так называемая постоянная протекания $C \approx 1,327\dots$ — универсальная константа, описывающая топологию неравновесных (квази)стационарных состояний в сложных нелинейных динамических системах, допускающих самоорганизованное критическое поведение». Сравнивая это значение размерности с модельными структурами (рис. 4.9), заключаем, что в диапазоне масштабов, представленных на рис. 4.10, магнитопотоковые концентрации агрегированы более

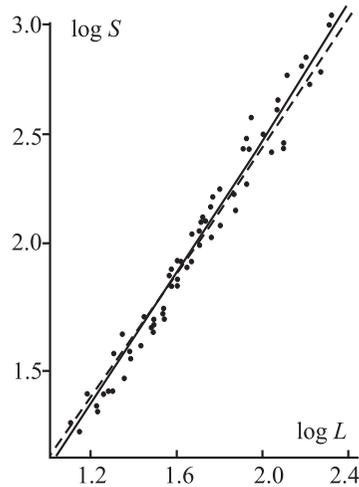


Рис. 4.10. Определение фрактальной размерности солнечных пятен по отношению «периметр – площадь». Здесь S — площадь, L — длина периметра

примеров применения этого соотношения для оценки фрактальной размерности поверхностей можно найти в книге Е. Федера [5].

разряженно, чем плотные упаковки рис. 4.9. Подобным же образом, с помощью соотношения «периметр – площадь», можно рассчитывать фрактальную размерность бассейнов рек и других природных объектов. Подробности можно посмотреть в книге Е. Федера [5].

Рассуждая аналогично, можно показать, что площадь поверхности S , если поверхность имеет фрактальную природу, связана с объемом $V \sim R^3$ соотношением:

$$S \sim R^{d-3}, \quad (4.10)$$

где R — радиус сферического элемента покрытия. Это соотношение часто используется для определения фрактальной размерности поверхностей различных геологических образцов по измерению молекулярной адсорбции. Множество

4.2.3. Информационная размерность

Как мы уже знаем, выборочная энтропия Шеннона определяется из соотношения

$$H(a) = -\sum_{i=1}^{n(a)} p_i \log_2 p_i, \quad (4.11)$$

где a — элемент покрытия. При уменьшении размеров покрытия $a \rightarrow 0$ и роста числа элементов этого покрытия $N(a) \rightarrow \infty$, энтропия Шеннона, как мы помним, неограниченно возрастает $H(a) \rightarrow \infty$. Однако предел

$$d_1 = \lim_{a \rightarrow 0} H(a) / \log_2(1/a) \quad (4.12)$$

может существовать. Если такой предел действительно существует, то им определяется величина d_1 , которая получила название информационной размерности. В том случае, если предела (4.12) не существует, вводят понятие верхней и нижней информационной размерности:

$$d_{\text{inf}} = \lim_{a \rightarrow 0} \inf(H(a) / \log_2(1/a)); \quad (4.13)$$

$$d_{\text{sup}} = \lim_{a \rightarrow 0} \sup(H(a) / \log_2(1/a)). \quad (4.14)$$

Мы знаем, что энтропия достигает своего максимального значения при равенстве вероятностей $p_i = 1/N(a)$. В этом случае из формулы (4.12) получаем

$$d_1 = \lim_{a \rightarrow 0} \log_2 N(a) / \log_2(1/a), \quad (4.15)$$

что с точностью до основания логарифмов (что не существенно) совпадает с формулой (4.5) для размерности Хаусдорфа, то есть в этом предельном случае $d_1 = d_0$.

В теории динамических систем предел (4.15) называют фрактальной размерностью или емкостью аттрактора. Если же предела (4.15) не существует, то говорят о нижней и верхней емкости аттрактора в смысле (4.13) и (4.14) соответственно.

4.2.4. Корреляционная размерность

Важную роль в различных приложениях играет так называемая корреляционная размерность, которую мы будем обозначать d_2 . Она определяется через корреляционный интеграл:

$$C(a) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i,j=1, i \neq j}^N \eta(a - |X_i - X_j|). \quad (4.16)$$

Здесь N — число выборок, X_i, X_j — векторы положения точек i и j , $\eta(Y)$ — функция Хевисайда (функция принадлежности к интервалу a). Значение $C(a)$ определяет относительное число пар точек, расстояние между которыми не больше a . Если предел

$$d_2 = \lim_{a \rightarrow 0} \frac{\ln C(a)}{\ln(1/a)} \quad (4.17)$$

существует, то его значение d_2 , называется корреляционной размерностью.

Подводя краткий итог, отметим следующее. Хаусдорфова размерность d_0 — характеризует «пористость» или плотность объекта, то есть то, как объект заполняет собою пространство вложения. Эта размерность может быть найдена разными способами в зависимости от конкретной задачи. Это важная характеристика любого объекта, и потому определение этой размерности чаще всего встречается в приложениях.

Информационная размерность d_1 играет важную роль в анализе нелинейных динамических систем, особенно при описании потери информации в ходе эволюции хаотических систем. В этом плане она связана с показателями Ляпунова и энтропией Колмогорова [20]. Однако она важна и в других приложениях, так как является важным количественным параметром системы, характеризующим меру его хаотичности. Если со временем система изменяется и эти процессы влекут за собой существенные вариации ее информационной размерности, то это говорит о том, что эти процессы могут быть связаны с ростом или диссипацией неравновесных структур, то есть переходам типа хаос–порядок или наоборот. Таким образом, информационная размерность позволяет количественно отслеживать направление и темп эволюционных процессов, таких как самоорганизация в среде структур разных уровней или, напротив, распад организованных структур и общая хаотизация.

Корреляционная размерность d_2 определяется корреляциями между элементами, составляющими среду, то есть вероятностью найти на расстоянии a от данного элемента множества один или несколько элементов того же множества. Эта размерность является также мерой иерархического скучивания, то есть характеризует упорядоченность внутренней структуры объекта. Ее теоретическая значимость обусловлена тесной связью с корреляционными функциями, которые определяют все физические особенности рассматриваемой среды.

Резюмируем этот раздел. Рассмотрены некоторые наиболее популярные способы нахождения значения Хаусдорфовой размерности природных фракталов: «бюкс алгоритм», соотношения «масса–радиус», «пери-

метр–площадь» и «площадь–объем». Бюкс алгоритм является наиболее универсальным. Он позволяет также рассчитывать информационную и корреляционную размерности природных фрактальных объектов. Эти фрактальные размерности являются полезными параметрами, позволяющие количественно характеризовать направление и темп процессов самоорганизации в среде и корреляционные связи между ее элементами.

4.3. Мультифракталы

В рамках конструктивных фракталов мультифрактал можно определить как «нагруженный» фрактал, в котором каждому элементу a_i ставится в соответствие некий «вес» или вероятность p_i . Для количественного описания мультифракталов рассчитывается спектр размерностей, соответствующих особенностям данного мультифрактала.

Понятие мультифрактала применимо и для природных фракталов. Признаком мультифрактальности для природных объектов является обнаружение различного скейленгового поведения на разных масштабах или наличие мультифрактального спектра размерностей. Таким образом, ключевым для обнаружения в природных объектах мультифрактальных структур и для количественной параметризации таких структур является умение рассчитывать мультифрактальный спектр размерностей. Этим мы и займемся в данном разделе.

4.3.1. Понятие мультифрактала

Рассмотрим сначала в качестве примера самоподобное множество, порожденное сегментами одинаковой длины с одним показателем скейлинга — размерностью Хаусдорфа. Далее каждому отрезку этого множества a_i припишем свой собственный вес p_i . Размерность Хаусдорфа задается, как мы помним, выражением

$$d_0 = \lim_{a \rightarrow 0} \frac{\log N(a)}{\log(1/a)}. \quad (4.18)$$

Оно вытекает из следующего соотношения:

$$\lim_{a \rightarrow 0} Na^d = c, \quad 0 < c < \infty, \quad (4.19)$$

где c — некоторая постоянная. Следует отметить, что предел (4.19) может существовать лишь в узких пределах значений d . Стоит хоть немного

изменить значение d , и произведение Na^d начнет стремиться либо в ноль, либо в бесконечность.

На n -том уровне самоподобного множества число элементов равно N_0^n , где N_0 — число элементов генератора (базовой структуры), соответственно, длина отрезка на n -том уровне равна a_0^n , где a_0 — длина отрезка генератора, если все отрезки имеют одинаковую длину. Тогда (4.19) можно записать так:

$$\lim_{n \rightarrow \infty} (N_0 a_0^d)^n = c. \quad (4.20)$$

Это предполагает, что $N_0 a_0^d = 1$, то есть

$$d = \frac{\log N_0}{\log(1/a_0)}. \quad (4.21)$$

То есть в случае строгого самоподобия, как мы уже знаем, необходимость в предельном переходе отпадает.

Если отрезки a имеют разную длину, то (4.20) записывается так:

$$\lim_{n \rightarrow \infty} \left(\sum_{i=1}^N a_i^d \right)^n = c. \quad (4.22)$$

Разумеется, при этом

$$\sum_{i=1}^N a_i^d = 1. \quad (4.23)$$

Для «нагруженного» фрактала вместо (4.22) рассматривают такой предел:

$$\lim_{n \rightarrow \infty} \left(\sum_{i=1}^N p_i^q a_i^\tau \right)^n = c \quad (4.24)$$

и находят такие q и τ , при которых предел (4.24) остается конечным, то есть

$$\sum_{i=1}^N p_i^q a_i^\tau = 1. \quad (4.25)$$

Из (4.25) видно, что может существовать множество $\tau = \tau(q)$ показателей τ и q , при которых это равенство выполняется, то есть имеем целый континуум фрактальных размерностей.

4.3.2. Спектр размерностей Реньи

В случае, если все a_i равны между собой (что из-за перехода к пределу не влияет на общность рассуждений), из (4.24) получаем:

$$\tau = \lim_{a \rightarrow 0} \frac{\log \left(\sum_{i=1}^N p_i^q \right)}{\log(1/a)}. \quad (4.26)$$

Заметим, что при $q = 0$, это выражение переходит в (4.18), то есть в этом случае $\tau = d_0$, то есть Хаусдорфовой размерности. Таким образом, τ можно рассматривать как некую новую обобщенную размерность, которая при $q = 0$ переходит в Хаусдорфову.

Рассмотрим предел

$$d_q = \lim_{a \rightarrow 0} R_q(a) / \log(1/a) = \frac{1}{q-1} \frac{\log \left(\sum_{i=1}^N p_i^q \right)}{\log(1/a)}, \quad (4.27)$$

где $R_q(a)$ — энтропия Реньи (3.9). Сравнивая (4.26) и (4.27), получаем $\tau = (1-q)d$. Таким образом, мультифракталы можно описать обобщенными фрактальными размерностями Реньи d_q или, что эквивалентно, показателями $\tau = (1-q)d_q$.

Из (4.27) можно получить как частные случаи первые три размерности Реньи. В самом деле, полагая (4.27) $q = 0$, получаем размерность Хаусдорфа (4.18). Предельный переход $q \rightarrow 1$ (3.13) дает нам информационную размерность (3.12–3.14). При $q = 2$ получаем корреляционную размерность (4.17), так как величина

$$C(a) = \sum_{i=1}^{N(a)} p_i^2,$$

стоящая в (4.27), при $q = 2$, и есть корреляционная сумма (4.16), если учесть, что p_i^2 описывает совместную вероятность пары соседних точек находиться в заданном элементе покрытия. Таким образом, при $q = 2$, формула (4.27) переходит в (4.17).

В формуле (4.27) порядок q может принимать значения в интервале $+\infty < q < -\infty$. Нетрудно показать (см., например, [6]), что d_q — монотонно невозрастающая функция q , то есть при $q < q'$, $d_q \geq d_{q'}$. Знак равенства возникает лишь в исключительных случаях идеальных самоподобных фракталов, например, все размерности d_q равны для канторова множества

[6]. Всю совокупность величин d_q рассматриваемого объекта называют спектром размерностей Реньи данного объекта.

Каждая из этих размерностей может быть определена множеством разных способов. Но вычисление обобщенных выборочных энтропий Реньи на достаточно широком диапазоне масштабов представляет нам общий и достаточно универсальный алгоритм для определения этих, важных в плане приложений, величин. В ряде работ речь идет о вычислении всего спектра размерностей Реньи. Но это в основном теоретические работы. На практике, как уже отмечалось, первых трех размерностей Реньи с их простотой вычислений и геометрической и физической наглядностью оказывается вполне достаточно для большинства практических приложений.

Отметим еще один важный момент. Знакомство с различными определениями и применениями энтропий в главах 1–3 позволяет при исследовании объектов или процессов осмысленно вводить вероятностную меру p_i и вычислять энтропии Реньи применительно к исследуемому объекту в рамках поставленной задачи. А знание энтропий Реньи в свою очередь, позволяет нам рассчитывать мультифрактальный спектр размерностей Реньи и физически адекватно интерпретировать получаемые результаты.

4.3.3. Расчет мультифрактальных спектров размерностей

Для практических расчетов главным является корректный способ определения иерархии вероятностных мер и расчет на их основе энтропий Реньи. В общих чертах мы уже касались этого вопроса в предыдущем разделе. Сейчас мы уделим больше внимания практической стороне вопроса.

Дело в том, что формула (4.27) мало пригодна для практических расчетов в силу ряда неопределенностей, связанных с предельным переходом $a \rightarrow 0$. Однако, как впервые было показано в работе [12] (*Meisel et al.*, 1992), можно наложить некоторые ограничения на процедуру выбора иерархии покрытий. Эти ограничения вполне приемлемы как с теоретической, так и с практической точек зрения. В этом случае задача создания компьютерного алгоритма для определения спектра размерностей Реньи существенно упрощается. Эти ограничения состоят в требовании дискретизации масштабов и равенства всех элементов покрытия на каждом масштабном уровне. Форма элементов покрытия особой роли не играет.

Точную формулировку этого прямого метода определения фрактальных размерностей — метода подсчета клеток — применительно к практическим расчетам спектров обобщенных размерностей Реньи можно найти в работе [12]. Однако в этой статье приводятся результаты тестирования метода только на модельных объектах типа кривой и снежинок Коха, то есть для получения размерностей конструктивных фракталов. Результаты работы [12] были обобщены в дальнейших исследованиях (см. [6], там же библиография), и в итоге был сформулирован подход для практического расчета фрактальных размерностей самых разнообразных наблюдаемых структур. В рамках этого подхода задаются меры p_i на последовательности покрытий структуры наборами конечного числа $R_q(a)$ равных элементов конечного размера a . Далее строится график Ричардсона и методом наименьших квадратов проводится прямая линия, если это имеет смысл, то есть если на заданной иерархии мер наблюдается скейлинговое поведение. Определяем далее тангенс угла наклона зависимости $\ln(R_q(a))$ от $\ln(1/a)$. Это и будет величина фрактальной размерности Реньи порядка q .

Итак, пусть $\{p_{i0}\}$ — есть мера, рассчитанная на упомянутом выше (a_0, N_0) покрытии ($i = 1, \dots, N_0$). Из этого покрытия формируем новое покрытие с равными ячейками размером $(a_k)^m$, где m — размерность пространства вложения, $a_k = 2, 3, 4, \dots, k = 1, 2, 3, \dots, M$. Число больших ячеек N_k равно целой части отношения $N_0 / (a_k)^m$, а мера большой ячейки равна сумме мер входящих в нее ячеек предыдущего масштабного уровня в силу аддитивности меры. Таким образом получаем ряд покрытий с размерами ячеек $\{a_k\}$ и мерами $\{p_{ik}, i = 1, 2, 3, \dots, N_k\}$.

Обычно для каждого разбиения рассчитывается обобщенная корреляционная функция

$$C_k(q) = \sum_{i=1}^{N_k} p_{ik}^q \quad (4.21)$$

для рассматриваемого набора q . Далее строим для каждой такой обобщенной корреляционной функции аналог графика Ричардсона в двух логарифмических осях ($\ln(a_k), \ln(C_k(q))$). Определив его наклон по методу наименьших квадратов, получаем оценку экспоненты обобщенной корреляционной функции $\tau(q)$, которая с размерностями Реньи (4.27) связана, как было показано выше, простым соотношением

$$\tau(q) = (q-1)d_q \quad (4.22)$$

Но можно рассчитывать энтропии Реньи порядка q и определять размерности Реньи d_q непосредственно путем построения графика ($\log(1/a_k), R_q(a_k)$).

Продемонстрируем технику расчета первых трех размерностей Реньи на примере ряда чисел Вольфа и рядов 5-минутных данных GOES по рентгеновскому потоку от Солнца. Здесь будет применен вышеописанный метод огрубленных покрытий в самом простом его варианте. Ниже этот метод будет развит и обобщен на случай обработки видео-изображений.

Алгоритм расчета в данном конкретном примере заключается в следующем. Находим максимальное и минимальное значения для рассматриваемой выборки. Делим этот интервал на некоторое достаточно большое число N_0 минимально возможных подинтервалов a_0 . Величина $a_0 = (\max(W) - \min(W))/N_0$ не должна быть меньше точности исходных данных.

На практике выбор N_0 и, соответственно, a_0 является достаточно тонким моментом, от которого могут сильно зависеть как величины, так и достоверности вычисляемых размерностей. Необходимо добиваться, чтобы достоверность аппроксимации (R^2 по Пирсону) прямыми линиями данных на графике $(\log(1/a), Rq)$ составляла порядка 0.9. Опыт показывает, что чаще всего это достигается при выборе минимального покрытия $a \geq 4m$.

Итак, мы вычисляем значения трех энтропий Реньи на покрытиях со все более и более грубыми ячейками $a = 4a_0, 6a_0 \dots$ и таким образом получаем данные для построения графиков, аналогичных графику Ричардсона $(\log(1/a), Rq)$, сразу для всех трех энтропий. Затем методом наименьших квадратов подбираем оптимальные аппроксимирующие прямые для всех трех энтропий и определяем тангенсы их углов наклонов, которые и дают нам значения соответствующих фрактальных размерностей. Достоверность аппроксимации R^2 прямыми, как уже отмечалось, должна быть около 0.9.

Здесь очень важен не только количественный, но и визуальный контроль результатов. Если разброс точек на графиках большой или сами графики не могут быть аппроксимированы прямыми, то скейлинговое поведение в данных отсутствует, и фрактальные размерности не могут быть определены. Фактически это означает отсутствие существования предела $\lim_{a \rightarrow 0} Rq/\log(1/a)$. Визуальный контроль можно осуществить путем вывода на дисплей графиков $(\log_2(1/a), Rq)$, что позволяет сразу оценить степень масштабной инвариантности (скейлинговость) рассматриваемых данных.

Результатами расчетов является таблица значений трех энтропий Реньи R_0, R_1, R_2 как функций размера ячейки покрытия, а также значения

трех фрактальных размерностей Реньи d_0, d_1, d_2 . Если мы имеем временной ряд магнитограмм, то можем отследить временные вариации всех трех размерностей. Средние значения, амплитуды вариаций, наличие или отсутствие периодичностей в этих рядах и другие их характеристики позволяют получить углубленное представление о динамике физических процессов в рассматриваемой среде.

4.3.4. Бокс алгоритм для ССД изображений

Мы продолжим рассмотрение метода огрубленных покрытий как базового инструмента для вычисления первых трех размерностей Реньи в приложении к экспериментальным и наблюдательным данным. В данном разделе мы обсудим приложение этого метода для вычисления размерностей при работе с цифровыми видео-изображениями.

Идея метода остается прежней: рассматриваем покрытие анализируемого изображения двумерной сеткой с минимально возможными ячейками, равными размеру одного пикселя (исходное или «базовое» покрытие). При этом, разумеется, предварительно должны быть проделаны все процедуры калибровки и коррекции, переводящие исходные ССД сигналы в размерные физические величины. Далее от размерных величин переходим к безразмерным вероятностям по обсуждавшейся ранее схеме и из полученных распределений вероятностей вычисляем три энтропии Реньи — R_0, R_1, R_2 . Затем огрубляем покрытие, объединяя несколько ячеек исходного покрытия, и снова проделываем ту же процедуру. Действуя таким образом, получаем последовательность значений трех энтропий Реньи, соответствующих набору огрубленных покрытий. Если изображение обладает скейлинговыми свойствами, то в координатах $(\log(1/a), Rq)$ получим три последовательности точек, соответствующих каждой из энтропий, которые должны быть хорошо (с достоверностью аппроксимации по Пирсону $R^2 \geq 0.9$) представлены линейными законами. Тангенсы углов наклона этих прямых дадут, соответственно, значения размерностей. Далее для удобства обозначаем через $H = d_0$ Хаусдорфову или структурную размерность, через $I = d_1$ — информационную размерность, через $C = d_2$ — корреляционную размерность.

На рис. 4.11 в качестве примера показаны три такие последовательности значений энтропий Реньи для S полярности в АО 10486, относящиеся к одной и той же дате. На данном рисунке видно четко выраженное скейлинговое поведение энтропий.

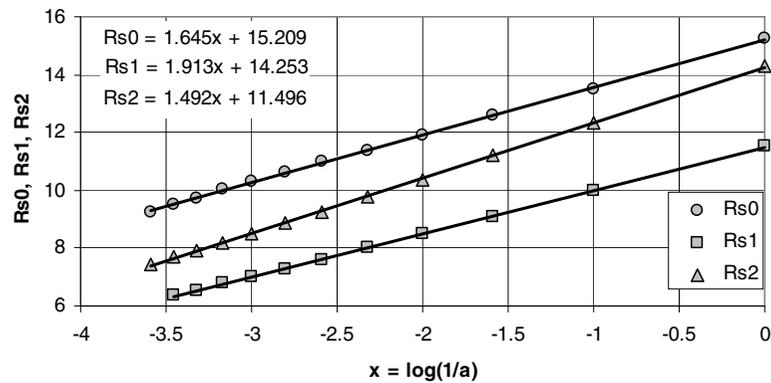


Рис. 4.11. Последовательности энтропий Реньи нулевого, первого и второго порядков (Rs_0 , Rs_1 и Rs_2 соответственно) для S полярности АО 10486 по 96-минутным данным SOHO/MDI на ноль часов 27 октября 2003 г. Во всех трех случаях достоверность линейной аппроксимации последовательностей значений энтропий близка к единице. Уравнения аппроксимационных прямых выписаны в левом верхнем углу рисунка. Коэффициенты наклона прямых (первый член в правой части уравнений) дают значения размерностей H , I и C на рассматриваемый момент времени

Сама идея вычислений и алгоритм ее реализации, в принципе, просты и понятны. Тем не менее, реализация их на практике сталкивается с рядом трудностей: значительные, даже для современных быстродействующих ЭВМ, объемы вычислений, проблема краевых эффектов при огрублении разбиений и др. Все это трудности чисто технического характера, и они могут быть успешно преодолены. Вместе с тем, нужно отметить, что трудно, да, пожалуй, и нецелесообразно разрабатывать программную реализацию данного алгоритма достаточно универсальную, чтобы она была применима к любым изображениям. Широкий класс решаемых задач, существенная специфика исходных данных при достаточной прозрачности алгоритма приводят к выводу, что целесообразнее строить специальные программы для реализации конкретных задач с учетом специфики конкретных данных. Такие специальные программы могут более просто реализовывать алгоритм и работать быстро, точно и надежно.

Тестируются такие программы как путем расчета размерностей конструктивных фракталов, например, ковров Серпинского, круговых фракталов, фракталов Кох, Минковского или любых других с известными размерностями, так и природных, например, упомянутого выше участка

побережья Норвегии или любых других природных фракталов, имеющих в литературе. Данные, разумеется, должны пройти тест на «скейлинговость» в рассматриваемом диапазоне масштабов.

Не станем поэтому подробно обсуждать саму схему вычислений для нижеприведенного конкретного примера — расчета размерностей магнитного поля солнечной активной области. Вместо этого уделим внимание анализу полученных в этом примере результатов, так как на них можно продемонстрировать некоторые типичные возможности и проблемы, возникающие при работе с размерностями, получаемыми из видеоизображений.

В отличие, скажем, от побережья Норвегии (см. рис. 4.9), диффузной туманности или галактики структура магнитного поля солнечной активной области находится в постоянной динамике. Она формируется конкурирующими физическими процессами: магнито-конвективно-турбулентной структуризацией плазмы и диссипативными процессами, разрушающими и упрощающими эти структуры. Поэтому естественно, ожидать, что величина структурной размерности может изменяться со временем. Интерес представляют характеристики этих изменений: амплитуды, тренды, стохастичность или периодичность вариаций и др. Все эти характеристики могут служить количественными критериями адекватности теоретических моделей структуризации солнечной плазмы в верхних подфотосферных слоях Солнца. Дополнительными критериями для уточнения и подгонки параметров таких моделей могут быть аналогичные характеристики временных рядов информационных и корреляционных размерностей, которые легко вычисляются параллельно с расчетом структурных размерностей.

Итак, чтобы получить общее представление о картине временных вариаций размерностей в сложных самоорганизующихся системах, продолжим рассмотрение нашего примера по анализу магнитных полей в солнечных активных областях NOAA10486 и NOAA10488. На рис. 4.12 представлены результаты расчета временных вариаций структурных — H , информационных — I и корреляционных — C энтропий N и S полярностей магнитного поля для этих двух АО. Как и в третьей главе, анализ основан на магнитограммах продольного поля, полученных командой SOHO/MDI и приведенных к центру солнечного диска. Иными словами, анализируется временное поведение фрактальных размерностей радиального поля этих активных областей. Данные SOHO/MDI по этим активным областям позволяют анализировать эквидистантные временные ряды магнитограмм с шагом 96 мин. и с шагом 1 мин. Для ниже приведенных примеров взяты 96-минутные данные.

На левой верхней панели (рис. 4.12) показаны временные вариации структурных размерностей солнечной активной области NOAA 10486 по 96-минутным магнитометрическим данным SOHO/MDI за период с 27 по 31 октября 2003 г. Из рисунка видно, что в среднем значения структурной размерности южной полярности были больше размерности северной полярности. То есть в целом магнитопотоковые трубки южной полярности были более плотно «упакованы» в кластеры и вся полярность в среднем была более компактной по сравнению с северной. В S полярности наблюдается растущий тренд, в N полярности — слабый растущий тренд наблюдается с 28 октября. В этом плане графики H_n и H_s похожи на соответствующие графики энтропий R_n и R_s на левой верхней панели (рис. 3.8).

Энтропии, как отмечалось выше, являются экстенсивными переменными, подобно таким переменным как магнитный поток, масса или количество тепла. Фрактальные размерности, в отличие от энтропий, являются интенсивными переменными, подобно магнитной индукции, плотности или температуры. Таким образом, энтропии и фрактальные размерности являются независимыми параметрами состояния рассматриваемой системы. Однако это вовсе не означает, что между интенсивными и экстенсивными переменными принципиально невозможно некоторое коррелированное поведение. Такая слабая корреляция между парами (H_n, R_n) , (H_s, R_s) наблюдается как в той, так и в другой АО. Однако оценки показывают, что достоверность по Пирсону этих корреляций не превышает 0.4.

В вариациях H_n и H_s , структурных размерностях N и S полярностей, в АО 10468 и АО 10488 можно наблюдать принципиально различное временное поведение: полную некоррелированность и даже антикорреляцию этих параметров в АО 10486 и положительную корреляцию (достоверность 0.85 по Пирсону) в АО 10488. И если коррелированное поведение структурных размерностей полярностей в АО 10488 можно считать вполне естественным для замкнутой системы магнитных потоков, то антикоррелированность этих размерностей в АО 10486 свидетельствует о принципиальной возможности более сложного поведения магнитных структур в солнечных АО.

Анализ и разработка физически содержательных моделей такого временного поведения структуры магнитных полей солнечных АО являются, разумеется, предметом специального исследования. То же относится и к анализу связей между заметными временными изменениями структурных размерностей по амплитуде на относительно короткой временной шкале и изменениями топологии магнитных потоков, и связанными с ними эруптивными процессами в этих активных областях. Эти временные вариации хорошо видны на обеих верхних панелях (рис. 4.12).

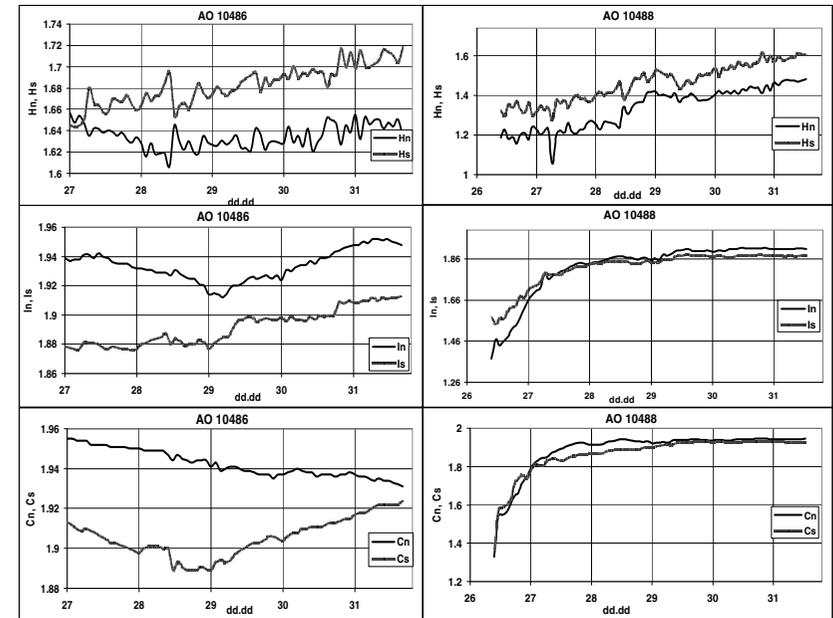


Рис. 4.12. Временные вариации структурных — H_n , H_s , информационных — I_n , I_s и корреляционных — C_n , C_s размерностей в солнечных активных областях АО 10486 и АО 10488 по 96-минутным магнитометрическим данным SOHO/MDI за 26–31 октября 2003 г. Время (ось X) в сутках и долях суток

Две средние панели (рис. 4.12) показывают изменение со временем информационной размерности радиальной компоненты магнитного поля в рассматриваемых АО. Информационную размерность, как уже отмечалось, можно рассматривать в качестве меры сложности структуры. Соответственно, из этих рисунков можно видеть, как на данном временном интервале изменялась сложность структуры магнитных полей северной и южной полярности в той и другой АО. Сравнивая изменения со временем H_s и I_s в АО 10486 (левый верхний и средний рисунки), можно видеть, что вместе с относительно более плотной «упаковкой» магнитопотоковых трубок в южной полярности ее структура была более однородной по сравнению со структурой северной полярности. В северной полярности, напротив, более рыхлая упаковка магнитных потоков демонстрирует большую сложность. Сложность в обеих полярностях АО 10486 достигла своих минимальных значений в начале 29 октября и затем, до 31 октября, шло усложнение структуры магнитных потоков в той и другой полярностях.

Обратимся теперь к АО 10488. На момент появления этой активной области на диске Солнца 24–25 октября она состояла из группы пор (небольших по площади магнитных концентраций), и все три ее размерности были меньше единицы. Затем структура начала усложняться и более плотно заполнять собою наблюдаемую область магнитограммы. Уже 26 октября, как видно из правой средней панели (рис. 4.12), все размерности превысили единицу и стали быстро увеличиваться, и к началу 27 октября значения структурной, информационной и корреляционной размерностей достигли 1.3, 1.66 и 1.8 соответственно, что характерно для вполне развитой АО. Таким образом, с 24 по 26 октября наблюдался «взрывообразный» процесс появления мощной и сложной локальной магнитной структуры на Солнце. Далее, после 28 октября, структурная размерность магнитного поля в обеих полярностях повышалась, а информационная и корреляционная размерности достигли своих предельных значений, сопоставимых с максимальными значениями этих размерностей АО 10486.

Важными отличиями в поведении магнитных полей АО 10488 в сравнении с АО 10486 являются почти равные значения и синхронные изменения информационной и корреляционной размерностей северной и южной полярности. Обращают на себя внимание также меньшие по амплитуде, в сравнении с АО 10486, короткопериодические вариации структурных размерностей и их хорошая корреляция. Поведение во времени корреляционных размерностей в этих АО, как это видно из нижних панелей (рис. 4.12), также существенно различалось.

Все эти особенности временного поведения структурных размерностей, безусловно, должны приниматься во внимание при исследовании процессов возникновения и эволюции мощных локальных магнитных полей в верхних слоях конвективной зоны Солнца. Применение фрактального анализа в исследовании структуры и динамики солнечных магнитных полей стало в последние годы важным и эффективным направлением в солнечной физике (см., например, [9, 13, 13] и другие публикации на эту тему). Стало понятно, что фрактальная размерность дает нам важную *количественную* информацию о динамической эволюции структуры сильных локальных магнитных полей и о связи между изменениями этой структуры и эруптивными процессами в солнечных активных областях.

Кратко резюмируем этот раздел. В начале раздела рассмотрено понятие мультифрактала как для абстрактно-математических, так и для природных объектов. Далее обсуждается концепция мультифрактального спектра размерностей и его связь с множеством фрактальных размерностей, получаемых из энтропий Реньи разных порядков. Рассмотрен бокс алгоритм расчета спектра размерностей Реньи для природных фракталов.

В качестве примера применения концепции мультифрактальных размерностей для анализа сложных неравновесных структур рассмотрен расчет временных рядов первых трех размерностей Реньи для последовательности CCD изображений сечения солнечной фотосферы сложных структур магнитных потоков в двух солнечных активных областях.

Показано, что расчет трех фрактальных размерностей как параметров состояния сложных неравновесных сред позволяет исследовать такие их свойства как появление и развитие в них неравновесных структур, а также качественно и количественно описать специфические особенности этих структур, их сложность и степень коррелированности элементов. Знание этих параметров и их динамики дает дополнительные возможности для решения ряда практических задач, таких как вопросы классификации структур, прогнозирование развития систем и др. Фрактальные размерности имеют также важное фундаментальное значение, так как дают количественные критерии адекватности физических моделей исследуемых систем. Таким образом, мультифрактальная параметризация любой сложной системы дает в руки исследователя ценный инструмент для анализа физических процессов в этой системе, в том числе тех, которые ответственны за возникновение, эволюцию и диссипацию в этой системе неравновесных динамических структур.

Список литературы

1. Мандельброт Б. Б. Фрактальная геометрия природы / под ред. В. А. Садовниченко и др. М.: ИКИ, 2002. 666 с.
2. Лебег А. Об измерении величин. М.: Учпедгиз, 1938. 210 с.
3. Морозов А. Д. Введение в теорию фракталов. М.–Ижевск: ИКИ, 2002, 160 с.
4. Мандельброт Б. Б. Самоаффинные фрактальные множества // «Фракталы в физике». М.: Мир, 1988. 672 с.
5. Федер Е. Фракталы. М.: Мир, 1991. 262 с.
6. Шредер М. Фракталы, хаос, степенные законы. М.–Ижевск: РХД, 2001. 528 с.
7. Chumak O. V., Chumak Z. N. Astron. and Astroph. Trans. 10, 1996. С. 329.
8. Фракталы в физике // Труды 6-го международного совещания по фракталам в физике. МСТФ, Триест, Италия, 9–12 июля, 1985. М., 1988. 670 с.

9. Зеленый Л. М., Милованов А. В. Фрактальная топология и странная кинетика: от теории преколяции к проблемам космической электродинамики // УФН. Т. 174. № 8. 2004. С. 809–852.
10. Кроновер Р. М. Фракталы и хаос в динамических системах. Основы теории. М.: Постмаркет, 2000.
11. Falconer K. Fractal Geometry. Mathematical Foundations and Applications. N.Y.: 1990. 155 p.
12. Meisel L. V., Johnson M. and Cote P. J. Box-counting multifractal analysis. Phys. Rev. A 45. 1992. P. 6989–6996.
13. Могилевский Э. И. Фракталы на Солнце. М.: Физматлит, 2001.
14. Abramenko V. Multifractal Analysis of Solar Magnetograms, Solar Physics. V. 228. Issue 1–2. P. 29–42.
15. Чумак О. В. Фрактальные размерности и соотношения «площадь – поток» для локальных магнитных полей на Солнце // *Астрономический Циркуляр*. № 1545. 1990.
16. Чумак О. В. Фрактальные размерности ассоциаций МФТ // *Астрономический Циркуляр*. № 1546. 1990.
17. Чумак О. В. Самоподобные фракталы с перекрывающимися элементами как модель фотосферных магнитных структур // *Астрономический Циркуляр*. № 1546. 1990.
18. Chumak O. V. Zhang H. Size-flux relation in active regions // *Chinese Journal Astron. and Astroph.* Vol. 3. No. 2. 2003. P. 175–182.
19. Chumak O. Self-similar and self-affine structures in observational data on solar activity // *Asrtion & Astroph. Trans.* V. 24. № 2. 2005. P. 93–99.
20. Bandt C. and Pompe B. Permutation entropy — a natural complexity measure for time series — *Phys. Rev. Lett.*, 2002, 29 апреля, 88 (17): 174102–174106.

ГЛАВА 5 СПЕЦИАЛЬНЫЕ МЕТОДЫ РАСЧЕТА ФРАКТАЛЬНЫХ РАЗМЕРНОСТЕЙ ВРЕМЕННЫХ РЯДОВ

Учитывая особую актуальность исследования временных рядов для анализа данных, в этой главе приводим некоторые проверенные практикой методы расчета фрактальных размерностей, которые могут оказаться в этом плане полезными. Здесь обсуждаются три таких метода. Это так называемый метод нормированного размаха, он же метод Херста, он же R/S анализ. Простой и быстрый алгоритм для оценки размерности Хаусдорфа, предложенный Хигучи, и алгоритм, предложенный Гассбергером и Прокачка (1983 г.) для оценки корреляционной размерности аттракторов динамических систем.

5.1. Элементы R/S-анализа

5.1.1. Эмпирический закон Херста

Одним из достаточно простых и эффективных методов анализа нестационарных временных рядов является так называемый R/S-анализ. Этот подход с успехом применяется для оценки текущих состояний и тенденций на товарных, фондовых и других рынках, на основе анализа соответствующих биржевых рядов. Этим методом анализируются ряды метеорологических и геофизических наблюдений – колебания температуры, давления, количества выпавших осадков, вариации магнитного поля Земли и др. В книге Енса Федера «Фракталы» (1991 г.) [1] R/S анализу посвящены 8 и 11 главы, где можно найти исторические и математические подробности этого метода. Ниже изложены суть метода, алгоритм и примеры его применения на практике.

Херст был гидрогеологом и проработал около сорока лет в районе Нила. Он занимался проблемой контроля уровня воды в водохранилище в районе старой нильской плотины (1907 г.). Задача состояла в том, чтобы найти такой оптимальный сток, при котором уровень воды в водо-

хранилище всегда находился бы между максимальной и минимальной фиксированной отметкой. Для этого нужно было изучить статистику вариаций притока воды, который, в свою очередь, зависел от частоты и интенсивности дождей в бассейне Нила. В итоге он открыл новый метод анализа стохастических временных рядов — так называемый метод нормированного размаха, или R/S-анализ. Метод устойчив, содержит минимальные предположения о системе, продуцирующей ряд, и его с успехом можно использовать для параметризации и классификации временных рядов.

Суть метода в следующем. Пусть $v(n)$ — приток, то есть объем воды, поступивший в водохранилище в n -том году. Пусть mv — сток, то есть объем воды, сброшенный из водохранилища за год. Этот объем должен равняться среднему притоку за год. Но дело в том, что величина этого среднего значения зависит от того, за сколько лет (N) мы берем это среднее. Средний приток за N лет равен:

$$mv_N = \frac{1}{N} \sum_{n=1}^N v(n). \quad (5.1)$$

Это среднее должно равняться объему, ежегодно спускаемому из водохранилища. В водохранилище при этом накапливается остаток: $dv(n)$ — накопившееся отклонение притока $v(n)$ от его среднего значения mv_N :

$$dv(n, N) = \sum_{i=1}^n (v(i) - mv_N). \quad (5.2)$$

Разность максимального и минимального значений этой величины на интервале N называется размахом R . При заданном стохастическом притоке $v(n)$, она должна быть меньше емкости водохранилища.

$$R(N) = \max_{1 \leq n \leq N} dv(n, N) - \min_{1 \leq n \leq N} dv(n, N). \quad (5.3)$$

Размах $R(N)$ зависит от периода N и с ростом этого периода не может убывать, более того, можно ожидать, что он будет расти. Если $R(N)$ разделить на стандартное отклонение $S(N)$ за тот же период N ,

$$S(N) = \sqrt{\frac{1}{N} \sum_{n=1}^N (v(n) - mv_N)^2}, \quad (5.4)$$

то получим безразмерную величину R/S , так называемый *нормированный размах*. Херст нашел, что для многих природных стохастических рядов

$$R/S = (aN)^H, \quad (5.5)$$

где показатель H был назван впоследствии показателем Херста, a — некоторая константа. Формула (5.5) называется *эмпирическим законом Херста*.

Для случайного процесса с некоррелированными значениями и конечной дисперсией

$$R/S = (\pi N/2)^{1/2} \quad (5.6)$$

То есть для таких процессов показатель Херста $H = 1/2$. Для большинства же случайных природных процессов $H > 1/2$. Это означает, что значения в природных рядах не являются независимыми. Текущее значение зависит от предыдущих значений, то есть от «истории» процесса. О таких рядах говорят, что они обладают «памятью». Причем это не кратковременная «марковская» память, когда только один-два предыдущих отсчета влияют на текущее значение. Показатель Херста реагирует на так называемую долговременную или «длинную» память, когда недавние события могут иметь влияние на текущее состояние (короткая память), но заметный вклад вносят также и события отдаленные (длинная память). Память о весьма отдаленных событиях, как показывает анализ, может сохраняться в рядах длительное время, иногда навсегда. Система, которая дает статистику Херста, есть результат длинного ряда взаимосвязанных событий. Таким образом, в таких рядах события не симметричны относительно текущего момента, то есть время имеет направление или, как говорят, имеется «стрела времени». Это, в частности, означает, что такие стохастические ряды в определенной мере прогнозируемы. Влияние настоящего на будущее может быть выражено корреляционным соотношением

$$C = 2^{2H-1} - 1, \quad (5.7)$$

где C — мера корреляции.

Для случайного процесса (процесса без памяти), как мы знаем, $H = 1/2$, и правая часть (5.7) обращается в нуль. Настоящее не влияет на будущее.

При $0.5 < H < 1$ имеется устойчивый тренд и ряд обладает памятью. Такие ряды называют устойчивыми или *персистентными*, поскольку если на данный момент есть какая тенденция (рост или убывание), то она сохранится и в будущем. Сила персистентности увеличивается при приближении H к 1. Напротив, чем ближе H к 0.5, тем больше зашумлен ряд и тем меньше выражен тренд. С этим все более или менее ясно.

Но вот интервал $0.0 < H < 0.5$ более загадочен. Получается, что существуют ряды, более случайные, чем просто случайный ряд. На самом деле такие ряды демонстрируют возврат к среднему, то есть если на дан-

ный момент существует в ряде некая тенденция, то существует вероятность ее переворота на противоположную. При этом эта вероятность тем выше, чем меньше H . Такие ряды называются эргодическими или *антиперсистентными*.

Персистентный временной ряд можно рассматривать как обобщенное броуновское движение, когда существует корреляция между событиями на временной шкале. Для классического броуновского движения фрактальная размерность равна 2. Как показал Мандельброт (2002 г.) [2], величина, обратная показателю Херста H , есть фрактальная размерность для объектов, вложенных в двумерье. То есть $d_0 = 1/H$. Классическое броуновское движение генерирует случайный ряд, для которого $H = 0.5$ и, таким образом, $d_0 = 2$. Для обобщенного броуновского движения, которое генерирует персистентные ряды, $H > 0.5$ и, соответственно, $d_0 < 2$.

5.1.2. Оценка показателя Херста

Полезность подхода Херста для анализа временных рядов не вызывает сомнений, поэтому рассмотрим сейчас простой метод оценки показателя Херста. Прологарифмировав (5.5), получаем:

$$\log(R/S) = H \cdot (\log(N) + \log(a)). \quad (5.8)$$

Если в двойных логарифмических координатах найдем тангенс угла наклона прямой, которая аппроксимирует наши данные, то это и будет показатель Херста. При этом всегда следует контролировать достоверность такой аппроксимации.

Программа, реализующая этот алгоритм, очень проста, и приводить ее здесь нет смысла.

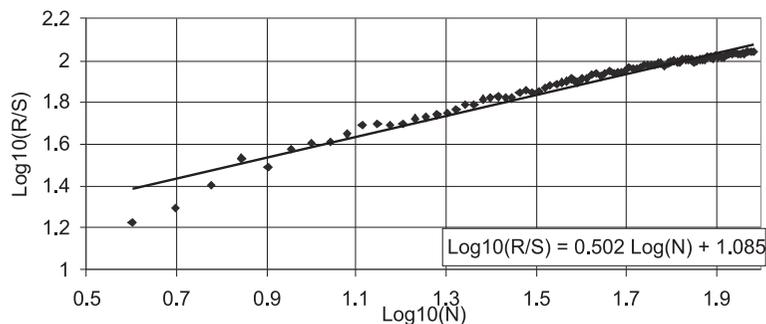


Рис. 5.1. Диаграмма Херста для случайного ряда

На рис. 5.1 показана диаграмма Херста для ряда, полученного с помощью компьютерного датчика случайных чисел. В правом углу рисунка представлено уравнение аппроксимирующей прямой. Из него видно, что показатель Херста $H = 0.502$, как и следует из теории. Достоверность аппроксимации данных прямой составляет по Пирсону $R^2 = 0.962$.

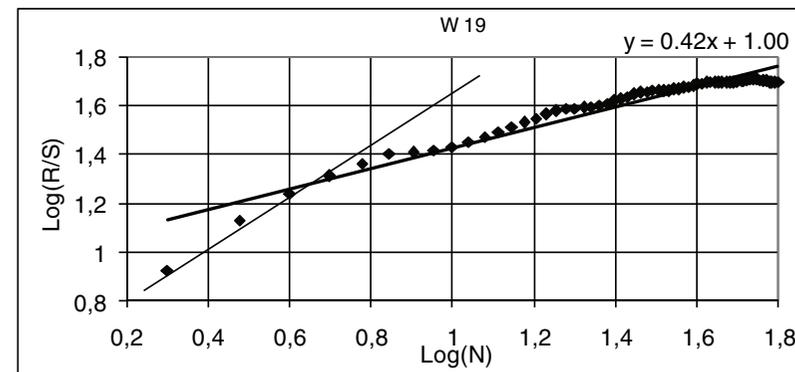


Рис. 5.2. Диаграмма Херста для 19 цикла ряда Вольфа

На рис. 5.2 представлена диаграмма Херста для 19 цикла ряда Вольфа. Волнообразное поведение данных свидетельствует о существовании на разных временных шкалах участков с разной степенью или, как говорят, силой персистентности. Начальный участок (левая прямая) дает $H = 0.87$. Весь ряд в целом — $H = 0.42$. То есть ряд Вольфа в зависимости от временного масштаба демонстрирует разные уровни персистентности.

5.2. Алгоритм Хигучи

Астрофизические и геофизические временные ряды генерируются сложными саморегулируемыми системами. Вследствие этого эти ряды, как правило, неоднородные и нестационарные и имеют фрактальную или чаще мультифрактальную временную структуру. Характерной особенностью астрофизических нелинейных процессов, например, в солнечно-земной физике, являются процессы взаимодействия (захвата, подстраивания, слипания) различных мод, что может нарушать случайное распределение фаз и приводить к их кластеризации. Это важное свойство временных рядов не может быть обнаружено и изучено в рамках традиционного спектрального анализа. Флуктуации амплитуды в таких рядах обнаруживают

свойство статистического самоподобия и могут рассматриваться как фрактальные объекты, подобно береговым линиям.

Для вычисления фрактальных размерностей таких временных рядов можно использовать простой и быстрый алгоритм, предложенный Хигучи [3]. Кривую разбиваем на n равных частей a , как мы это уже не раз проделывали. Длину кривой представляем в виде n^d , где d — Хаусдорфова размерность, величина которой находится, как и в случае с береговыми линиями, в пределах $1 < d < 2$.

Пусть анализируемый участок кривой (выборка) состоит из N отсчетов. Величина $a = \text{int}[N/n]$, где функция $\text{int}[x]$ является функцией выделения целой части. Из исходного ряда $f(t)$, в котором t принимает N дискретных значений, конструируем новые временные ряды $f(m, a)$:

$$f(m, a) = f(m), f(m + a), f(m + 2a), \dots \\ f(m + \text{int}[(N - m)/a]a); \quad m = 1, 2, \dots, a. \quad (5.9)$$

Длина кривой в методе Хигучи определяется так:

$$L_m(a) = \frac{(N - 1)}{a \text{int}[(N - m)/a]} \cdot \sum_{i=1}^{\text{int}[(N - m)/a]} |f(m + ia) - f(m + (i - 1)a)|. \quad (5.10)$$

Далее вычисляем среднее значение $L(a)$ по всем m и строим зависимость $\log(L(a))$, от $\log(a)$. Далее как обычно аппроксимируем эту зависимость прямой линией по способу наименьших квадратов. Тангенс угла наклона этой прямой дает нам значение фрактальной размерности по методу Хигучи.

Следует еще раз подчеркнуть, что рассчитанная таким образом размерность находится в интервале от 1 до 2 и не имеет отношения к размерностям аттрактора, который может быть определен из этого же временного ряда. Для вычисления этих размерностей необходимо восстановить фазовое пространство динамической системы, погрузить в это пространство траекторию системы и затем применить бокс-алгоритм или какой-либо иной метод для вычислений размерностей бассейнов притяжения фазовой траектории в этом пространстве. Такие размерности могут превышать 2, так как пространство погружения фазовой траектории может иметь существенно большее число размерностей.

Более подробно об алгоритме Хигучи и других о подобных алгоритмах, таких как алгоритм Катса и алгоритм Петерсона, а также о сравнении их достоинств и недостатков можно прочитать в [4].

5.3. Алгоритм Грассбергера–Прокаччия

Оригинальный метод расчета корреляционной размерности аттрактора динамической системы был предложен П. Грассбергером и И. Прокаччия [5] в 80-х годах прошлого века в связи с разработкой алгоритма по реконструкции фазового пространства динамической системы по одной наблюдаемой переменной [6], [7]. Фазовое пространство содержит в себе фазовую траекторию системы со всеми ее «бассейнами притяжения» — аттракторами и областями избегания — репеллерами. В этом смысле фазовое пространство можно рассматривать как пространство вложения (фазовой траектории) с минимальной целочисленной размерностью. Определение размерности фазового пространства основывается на методе запаздывания (delay-time reconstruction). Идея метода заключается в следующем.

Пусть x_k — некоторое значение из временного ряда, а n — может принимать значения 1, 2, 3 Тогда $\mathbf{x}_k = (x_k, x_{k-n}, x_{k-2n}, \dots, x_{k-(m-1)n})$ — точка в m -мерном пространстве. Переходя к последующим значениям ряда, получаем некоторую последовательность точек («наблюдаемую траекторию») в этом m -мерном пространстве. Показано (см. [6], [7]), что, так как текущие значения любой переменной в динамической системе являются детерминированными следствиями предыдущих значений остальных переменных, эта «наблюдаемая траектория» типологически эквивалентна «истинной» фазовой траектории. То есть наблюдаемую траекторию можно интерпретировать как некую проекцию «истинной» фазовой траектории системы из ее «истинного» фазового пространства в рассматриваемое m -мерное пространство. Меняя величину m начиная с минимальных значений 2 и 3, находим такое минимальное значение m , после которого вид фазовой траектории не меняется. Это значение m и будет размерностью минимального пространства вложения траектории, равной размерности «истинного» фазового пространства динамической системы.

Это последнее соображение связано с тем, что размерность аттрактора не изменится, если мы погружаем его в пространство более высокой размерности, чем его минимальное пространство вложения. Линия как в двумерном, так и в трехмерном пространстве остается одномерным объектом. Фрактальные размерности ковров Серпинского, фрактала Минковского или любого другого фрактала в любом пространстве вложения сохранят значения. Точно так же и аттрактор динамической системы (если наш временной ряд порожден именно такой системой) сохраняет свою размерность при увеличении размерности пространства вложения сверх минимальной.

Если же мы имеем дело со случайными временными рядами, то, в этом случае наблюдается иная картина: фазовые точки будут стремиться равномерно заполнить фазовое пространство любой размерности. Случайный ряд ведет себя в рамках этой процедуры как газ. Он равномерно распространяется по всему объему. Динамическая же система с ее далекодействующими силами, порождающими корреляции, дальний порядок и память в генерируемых ею временных рядах, будет порождать рассматриваемых пространствах последовательность точек, концентрирующихся в одних областях (аттракторы) и избегающих другие области, подобно звездным скоплениям или галактикам. Это облако точек можно погрузить в пространство любой размерности, но в силу своих внутренних связей его размерность от этого не изменится.

Проблема в том, как определить размерность фазового пространства m не на глаз (это, в принципе, возможно для простых систем с m не больше трех), а с помощью строгой количественной процедуры. Гассбергер и Прокачиа [5] предложили решить этот вопрос путем расчета корреляционной размерности аттрактора динамической системы. Если размерность аттрактора известна, то округление его величины до ближайшего целого в большую сторону и будет значением m .

Корреляционную размерность авторы предложили рассчитывать через корреляционный интеграл $C_m(a)$ (формула 4.16). Итак, из исходного ряда продуцируем ряды с набором времен запаздывания, начиная с минимальных, как было сказано выше. Исходный и запаздывающие ряды рассматриваем как разные координаты нашего пространства размерности m и оцениваем значение корреляционного интеграла для пар m -мерных точек X_i и X_j в этих пространствах:

$$C_m(a) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i,j=1, i \neq j}^N \eta(a - |X_i - X_j|), \quad (5.11)$$

где a — m -мерное расстояние между парой точек в анализируемом пространстве, η — функция Хэвисайда.

Мы ожидаем, что с ростом a $C_m(a)$ будет увеличиваться по закону

$$C_m(a) = a^d. \quad (5.12)$$

Поэтому далее как обычно, строим зависимость $(\log(C(a)), \log(a))$ и пытаемся определить корреляционную размерность d . То значение m , после которого значение d перестает изменяться и есть минимальное пространство вложения фазовой траектории, то есть фазовое пространство системы, а соответствующее ему d есть корреляционная размерность аттрактора.

В качестве примера рассмотрим аттрактор Энона. Отображение Энона определяется уравнениями

$$x(t+1) = 1 + y(t) - ax(t)^2, \quad y(t+1) = bx(t). \quad (5.13)$$

Не составляет большого труда написать простую компьютерную программу для построения аттрактора Энона. Результат работы такой программы для случая $a = 1.4$ и $b = 0.3$ показан на рис. 1 и рис. 2.

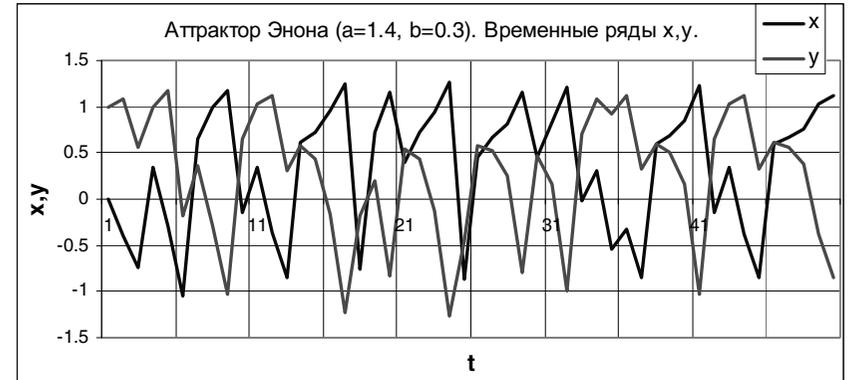


Рис. 5.3. Первые 50 значений временных рядов Энона

Рисунок 5.3 демонстрирует достаточно стохастическое поведение временных рядов $x(t)$ и $y(t)$ отображения Энона.



Рис. 5.4. Аттрактор Энона, вычисленный по 600 значениям временных рядов $x(t)$, $y(t)$.

Рисунок 5.4 показывает, что за хаотичным поведением временных рядов, скрывается вполне упорядоченные фазовые траектории.

Описанный алгоритм Грассбергера–Прокаччия при применении к ряду $x(t)$ или ряду $y(t)$ дает значение корреляционной размерности d для аттрактора Энона ~ 1.25 при значении $m = 2$. Дальнейшее увеличение m не изменяет величину d . Точность при вычислении d , равная ~ 0.02 , достигается уже при длине ряда ~ 600 значений, в то время как выше рассмотренный классический бокс алгоритм на m -мерном пространстве требует для этого не менее нескольких тысяч значений. Более подробно о применении алгоритма Грассбергера–Прокаччия для реконструкции фазового пространства динамической системы и определения размерности аттрактора можно прочитать, например, в книге [8].

Таким образом, в ряде случаев рассмотренные специальные алгоритмы расчета фрактальных размерностей могут оказаться предпочтительнее общего алгоритма огрубленных покрытий. Кроме того, эти методы могут служить хорошим контролем вычислений по общему алгоритму. Вместе с тем, необходимо отметить, что общий алгоритм вычислений фрактальных размерностей через бокс-алгоритм и энтропии Реньи более универсален и сразу позволяет рассчитывать весь спектр размерностей, что важно в случае мультифрактальных объектов.

5.4. Перестановочная размерность

По аналогии с вышеописанными размерностями Реньи можно ввести перестановочную размерность выборки. В соответствии с определением вероятностной меры перестановочной энтропии (см. раздел 3.3.3) значение $n!$ можно рассматривать в качестве числа ячеек покрытия в пространстве перестановок, где n — порядок перестановочной энтропии. Полагая объем этого пространства постоянным и равным единице, получаем размер ячейки покрытия a , равным $a = 1/n!$. Далее, как и в случае размерностей Реньи, строим зависимость $(\log_2 n! - H_n)$ и аппроксимируем ее прямой линией. Если такая аппроксимация имеет достоверность по Пирсону ~ 0.9 и выше, то коэффициент наклона этой прямой дает нам значение перестановочной размерности.

Как уже отмечалось, перестановочные энтропии, как и всякие другие энтропии, в случае однородного распределения стремятся к своим максимально возможным значениям, равным, в данном случае, $\log_2 N$, где $N = n!$ — число элементов покрытия. Соответственно, в случае случайного ряда, дающего однородное распределение по амплитудам, перестановочная размерность должна стремиться к единице при стремлении к бесконечности объема выборки.

Обратимся снова к Таблице 3.1. Из данных приведенных в этой таблице, легко найти, что перестановочная размерность Pd для случайного ряда, рассчитанная по $H_2 - H_5$, равна единице с достоверностью по Пирсону порядка единицы. Для случая реальных данных (столбцы 5, 6, 7) $Pd = 0.717, 0.655$ и 0.627 , соответственно, с достоверностью по Пирсону не ниже 0.996.

Перестановочные размерности дают удобный и легко вычисляемый параметр сложности выборки в смысле, вытекающем из определения перестановочной энтропии. Для скользящей последовательности выборок рассматриваемого временного ряда можно получить, соответственно, временную последовательность значений его сложности. Такой ряд сложности позволит качественно и количественно оценить эволюционные тенденции в системе, продуцирующей исходный временной ряд.

Список литературы

1. Федер Е. Фракталы. М.: «Мир», 1991.
2. Мандельброт Б. Б. Фрактальная геометрия природы / под. ред. В. А. Садовниченко и др. М.: ИКИ, 2002. 666 с.
3. Higuchi T. Approach to an irregular time series on the basis of the fractal theory // Physica D. Vol. 31. 1988. P. 277–283.
4. Esteller R., Vachtsevanos G., Echauz J. and Litt B. A Comparison of Waveform Fractal Dimension Algorithms // Ieee Transactions on Circuits and Systems — I: Fundamental Theory and Applications. Vol. 48. No. 2, 2001. P. 177–183.
5. Grassberger P. and Procaccia I. Characterization of strange attractors // Phys. Review Letters 50. № 5. 1983. P. 346–349.
6. Packard N. H., Crutchfield J. P., Farmer J. D. and Shaw R. S. Geometry from a time series // Phys. Rev. Lett. Vol. 45. 1980. P. 712–716.
7. Takens F. Detecting strange attractors in fluid turbulence in Dynamical Systems and Turbulence, D. a. y. L.-S. Rand, Ed. Berlin, Germany: Springer-Verlag, 1981. Vol. 898 (Lecture Notes in Mathematics). P. 366–381.
8. Кузнецов С. П. Динамический хаос. М.: Физматлит, 2001.

ЗАКЛЮЧЕНИЕ

Фракталы и детерминированный хаос. Фракталы в астрофизике. Фрактальная геометрия и фрактальная размерность. Самоподобие и математические фракталы. Природные фракталы и стохастическое самоподобие, мультифракталы и спектры размерностей. Фрактальные размерности, связанные с длинами, площадями, массами. Фрактальные размерности аттракторов динамических систем. Фрактальная параметризация сложных систем.

Фракталы и детерминированный хаос

В заключение взглянем на общую картину прикладного фрактального анализа и подведем некоторые итоги. «Фракталы вокруг нас повсюду: и в очертаниях гор, и в извилистой линии морского берега. Некоторые из фракталов непрерывно меняются, подобно движущимся облакам или мерцающему пламени, в то время как другие, подобно деревьям или нашим сосудистым системам, сохраняют структуру, приобретенную в процессе эволюции», — отмечают Х. О. Пайген и П. Х. Рихтер [1]. Подобные идеи стали движущей силой в развитии и применении методов фрактального анализа природных структур и физических процессов в сложных системах самой разной природы.

Как мы знаем, термин *фрактал* (от лат. fractus — изломанный, иррегулярный, фрагментарный) применяется к таким объектам в пространстве или процессам во времени, которые являются самоподобными или самоаффинными и которые на разных пространственно-временных интервалах демонстрируют скейлинговое поведение. Фракталы и мультифракталы являются рекуррентно иррегулярными объектами в пространстве и/или во времени. Фрагменты фрактального объекта или временного ряда являются геометрически подобными (в случае конструктивных фракталов), типологически подобными или статистически подобными копиями целого. В этом плане по фрагменту фрактального объекта, в принципе, можно воспроизвести целое путем серии сдвигов и растяжений для конструктивных и топологических фракталов или путем статистически подобных по-

вторений на разных масштабах — для статистических фракталов. В этом плане фракталы обладают некоторыми свойствами голограмм [2].

Фрактальные структуры наблюдаются во многих природных объектах и процессах. Фрактальная геометрия предлагает фундаментально новый взгляд как на неживые, так и на живые системы. Она рассматривает их как результат самоорганизации или самовыстраивания спонтанных самоподобных флуктуаций на протяжении длительного времени и на многих временных шкалах. Именно в результате таких процессов возникают системы, представляющие собой сложные рекурсивно построенные паттерны, организованные на множестве пространственных уровней и временных масштабов, в том числе в биологии и медицине [3], [13].

Вообще говоря, под системой обычно понимают все, что имеет больше чем одну часть. Такие разные объекты как атом водорода, Земля со своим спутником Луной или Галактика являются системами. Говорят, что система является *динамической*, если *состояние* системы, включая взаимосвязи и взаимодействия между ее частями, изменяется со временем по динамическим законам. Эти законы позволяют записать системы динамических уравнений, которые количественно описывают изменения системы со временем. Если взаимодействия являются нелинейными и, соответственно, уравнения, описывающие динамику системы, содержат нелинейные члены, то такая система может показывать нелинейное поведение, то есть поведение, при котором малые изменения в системе и их отдаленные последствия не пропорциональны. О таких системах говорят, что они демонстрируют *нелинейную динамику*. Например, в ряде систем, динамику которых можно описать законами классической механики, малые изменения начальных условий приводят к немалым расхождениям фазовых траекторий за конечные времена.

Даже простые системы с нелинейной динамикой часто демонстрируют очень сложное поведение, известное как *хаотическое* (см., например, [4]). Парадокс такого хаоса заключается в том, что он *детерминированный*, то есть он является результатом динамики, которая управляется детерминированными, а не вероятностными законами. Вследствие чрезвычайной чувствительности к начальным условиям и системным параметрам такие нелинейные системы могут проявлять полностью случайный характер. Однако принципиально важно то, что существует порядок, лежащий в основе такого поведения.

Например, компьютерный генератор случайных чисел представляет собой вполне упорядоченную детерминированную процедуру. В месте с тем, последовательность нулей и единиц, генерируемая таким датчиком, почти ничем не отличается от такой же последовательности, полученной из реального случайного процесса, имеющего два исхода, напри-

мер, подбрасывание монеты. Обе последовательности обладают одинаковыми вероятностно-статистическими свойствами. Таким образом, *случайность* — есть понятие относительное и контекстно зависимое. Определение случайности как отсутствия всякого порядка теряет смысл. Возникает идея, что за всяким случайным поведением может прятаться довольно простой детерминированный процесс, и нужно только найти способ, как его обнаружить.

Модельные системы детерминированного хаоса, в отличие от наблюдаемых в природе случайных процессов, могут быть легко контролируемы и перенастраиваемы. Таким образом, теория нелинейной динамики и динамического хаоса предоставляет нам методы анализа и моделирования природных хаотических явлений и снабжает нас эффективными способами количественного описания внутренней динамики и фрактальной структуры реальных систем.

Если мы проследим в течение длительного времени за поведением фазовой траектории динамической системы, то обнаружим, что системы большую часть времени проводят в области одного или более одного центров или бассейнов притяжения фазовых траекторий — аттракторов. Эти области можно рассматривать как равновесные или квазиравновесные состояния системы в ее фазовом пространстве.

Системы, порождающие динамический хаос, имеют хаотические или странные аттракторы. Здесь теория хаоса напрямую связана с фрактальной геометрией, поскольку все странные аттракторы являются фракталами. Построение фазового пространства для систем с неизвестной динамикой является непростой задачей. Наблюдаемые переменные, характеризующие систему, такие как амплитуды сигналов, продуцируемые системой в последовательные моменты времени, как бы погружены в многомерное фазовое пространство, определяемое самой системой (см., например, [5]). Кроме того, процессы в хаотических системах часто продуцируют структуры и/или сигналы, которые также являются фракталами в обычном пространстве и во времени.

Фракталы в астрофизике

Эволюционные процессы, происходящие в космосе, демонстрируют бесконечное разнообразие переходов от хаоса к порядку через самоорганизацию и возникновение диссипативных структур самых разных пространственно временных масштабов — от галактик и звезд, до планет-

ных систем и живых существ. Вот лишь несколько примеров широко об- суждаемых систем такого рода.

С бортов КА было зарегистрировано интересное явление, наблюдае- мое в кольцах Сатурна, — кратковременное появление радиальных структур, так называемых «спиц». Длина их обычно порядка 10000 км, ширина — ~1000 км. Типичное время жизни этих систем — порядка не- скольких часов, после чего они распадаются под действием дифференци- альных движений колец, чтобы через некоторое время появиться снова. Из-за не до конца понятой физики процессов надежная динамика этих систем до сих пор не построена, но возникновение дальнего порядка и стохастическое поведение структур явно указывает на спонтанную са- моорганизацию на нескольких пространственных масштабах.

Закон планетных расстояний и многочисленные резонансные соот- ношения в нашей планетной системе являются реликтами и отголосками процессов самоорганизации на широком диапазоне масштабов, прохо- дивших в эпоху формирования Солнечной системы. Таким образом, Солнечная система является впечатляющим примером возникновения порядка из первоначального хаоса протопланетного облака.

Солнце демонстрирует огромное разнообразие спонтанно возникаю- щих и затем диссипирующих магнитоплазменных структур [6]. Гранулы, супергранулы, сверхгранулы, факелы, протуберанцы — вот далеко не полный перечень спонтанно возникающих и исчезающих структур на Солнце. Для этих объектов обнаружено скейлинговое поведение, что го- ворит об их фрактальной природе. В этой связи можно отметить работы [7–9] и др., уже упоминавшиеся в четвертой главе в связи с обсуждением фрактальной природы солнечных пятен и групп солнечных пятен. В рабо- те [10] рассмотрено статистическое самоподобие флуктуаций солнечной активности. Найдено, что амплитуда флуктуаций суммарной площади солнечных пятен dS связана с их временем жизни t , скейлинговым соот- ношением $dS \sim t^{-2}$. Такой закон, возможно [6], обусловлен существованием «солитонного газа» в конвективной зоне Солнца, определяющем струк- турные особенности локальных солнечных магнитных полей разных пространственных масштабов.

Автоколебательная реакция Белоусова–Жаботинского стала класси- ческим примером демонстрации самоорганизации диссипативных струк- тур через так называемый *дальний порядок*, то есть корреляционные свя- зи самоподобных структур в широком диапазоне масштабов. Но автоколе- бания давно известны на физически переменных звездах — цефеидах. Пространственно-временная самоорганизация происходит в этих звездах благодаря согласованности колебаний ионизационного равновесия во внутренних зонах и механических колебаний внешних слоёв звезды.

Спирально-кольцевая структура галактик является другим ярким примером самоорганизации диссипативных структур через дальний по- рядок. Сами галактики и крупномасштабные структуры Вселенной, как теперь понято, также являются следствием самоорганизационных про- цессов, проходивших во Вселенной на разных этапах ее эволюции. Мож- но сказать, что современная астрофизика — это эволюционная теория, изучающая образование и устойчивость упорядоченных структур из пер- воначально неупорядоченного космического хаоса. Б. Мандельброт [11] посвятил специальную главу обсуждению этой темы. Основные доста- точно современные представления о крупномасштабной структуре Все- ленной, иерархическом сгущении (масштабном скейлинге) и эволюции малых возмущений (флуктуаций), приводящих к образованию иерархии космических систем, можно найти в монографии [12].

В первой главе упоминалась идея «тепловой смерти Вселенной», к которой пришел Клаузиус, основываясь на втором начале термодина- мики. Теория молекулярного хаоса, развитая Больцманом, находится в согласии с классической термодинамикой, подтверждает и объясняет ее со статистико-кинетических позиций. Классическая термодинамика и классическая кинетика не содержат идей спонтанного роста флуктуаций и появления систем в результате таких событий. Напротив, любые флук- туации в идеальном хаосе Больцмана неизбежно быстро рассасываются, а появление значительных флуктуаций и вовсе невероятное событие.

Теория динамического хаоса содержит иные идеи. В рамках этой тео- рии появление хаоса рассматривается как следствие динамики системы. При этом как переход «порядок – хаос», так и обратный переход «хаос – порядок» возможен в таких системах. Развитый хаос в нелинейных дина- мических системах может быть по своим статистическим характеристикам близок к идеальному хаосу Больцмана, подобно тому как последователь- ность нулей и единиц, порождаемая детерминированным датчиком слу- чайных чисел, близка по своим статистическим характеристикам после- довательности исходов при бросании монеты. И все же динамический хаос не тождественен молекулярному хаосу идеального газа. В динами- ческом хаосе всегда присутствуют взаимодействия и корреляции между элементами на разных масштабах. Эти взаимодействия могут при опре- деленных условиях способствовать спонтанному росту флуктуаций в широком диапазоне масштабов и, как следствие, порождать различные упорядоченные структуры. Пространственно-временные корреляции на- рушают однородность и отсутствие структуры в среде и путем самоорга- низации через дальний порядок порождают удивительные формы неживой и живой природы. Образно говоря, «идеальная» кинетика без памяти и корреляций и термодинамика «идеального» газа с ее непререкаемым

вторым началом — это теория *абсолютного хаоса* или наука о диссипации, распаде и смерти структур, что, как мы помним, понял еще Клаузиус с его тепловой смертью мира.

Современная кинетика открытых систем, с корреляциями и памятью и неравновесная термодинамика допускают не только рост, но и убыль энтропии в системе. Убыль энтропии в открытых системах приводит к появлению дальнего порядка путем корреляции самоафинных флуктуаций в широком наборе пространственно-временных масштабов. Таким образом, современная физика открытых систем, или теория *динамического хаоса*, является наукой о творении космоса (греч. *порядок, красота*) из хаоса. Во Вселенной есть и то и другое. Таким образом, классическая кинетика и теория динамического хаоса не находятся в противоречии друг с другом, но описывают разные стороны единого процесса — возникновения, эволюции и диссипации структур во Вселенной. Энтропии и мультифракталы — это те инструменты, которые позволяют количественно описывать разные фазы эволюции космических систем. Частично, библиографию работ по фракталам в астрофизике можно найти, например, на сайте <http://www.umich.edu/~lowbrows/biblio/fractals.html>

Фрактальная геометрия и фрактальная размерность

Как уже отмечалось, фрактальная размерность есть мера того, насколько сложной является самоподобная фигура. Грубо говоря, это мера того, сколько «физических» или «конечных» точек содержится в том или ином множестве. Так, интуитивно понятно, что плоскость должна содержать большее число таких «конечных точек», чем линия, в то время как, например, треугольник Серпинского находится где-то между этими двумя множествами. С другой стороны, все эти три множества имеют одно и то же число бесконечно малых «математических» точек, в том смысле, что все они имеют одну и ту же мощность, поскольку все являются несчетными множествами. Фрактальная размерность дает некое новое, интуитивно понятное и менее абстрактное, по сравнению с традиционными *кардинальными числами*, представление о том, «как велико множество». В этом плане фрактальный подход можно рассматривать как некий дополнительный раздел в классической теории множеств.

С другой стороны, *фрактальная геометрия* может рассматриваться как обобщение геометрии Евклида. Условно, мы можем рассматривать целые размерности как степени длины, то есть точка = длина⁰, линия =

= длина¹, поверхность = длина², объем = длина³. Показатель степени есть размерность. Фрактальная геометрия предполагает, что длина может иметь и любой дробный показатель степени. Такой дробный показатель степени называется по аналогии с целочисленной евклидовой размерностью *фрактальной размерностью*.

Фрактальные объекты, как мы уже знаем, обладают свойством самоподобия, которое означает, что конкретная морфологическая ситуация повторяется на разных масштабах. Фрактальная размерность представляет меру скорости добавления дополнительных деталей при переходе от одного масштаба к следующему. Таким образом, фрактальную размерность можно рассматривать и как некий *квантор сложности*.

Стоит обратить внимание еще на одну идею. Когда мы говорили о целых размерностях длины и полагали, что точка = длина⁰, линия = = длина¹ и так далее, мы имели в виду «математическую» или «идеальную» точку, линию, поверхность и так далее. Однако в реальном мире не существует бесконечно тонких линий и поверхностей, точек нулевых размеров. Нет и строго трехмерных объектов, в них на определенных масштабах всегда найдутся пустоты. Идея перехода к пределу хороша до определенной степени. В реальности всегда существует конкретный масштаб, за которым предельный переход теряет смысл. В такой ситуации попытка проникнуть в область меньших масштабов может привести нас неадекватным выводам.

Таким образом, бесконечно длинная и бесконечно тонкая линия, равно как и точка строго нулевых размеров или бесконечная плоскость нулевой толщины, есть лишь абстрактные модели реальных объектов, имеющие конечную область применения. Реальная «физическая» или «неидеальная» точка должна иметь размерность хотя бы «немного» больше 0, линия, соответственно, — больше 1, поверхность — больше 2. Соответственно, реальные тела, построенные из таких неидеальных объектов, будут иметь дробные размерности, значения которых будут тем больше, чем более сложным является составляющее их множество. Таким образом, идеальные конструкции Евклидовой геометрии часто лишь очень грубо описывают реальные тела [11], и фрактальная геометрия позволяет строить более адекватные модели реальных систем.

Самоподобие и математические фракталы

Центральными идеями фрактальной геометрии являются идеи *самоподобия* и *скейлинга* — повторение набора одних и тех же структур на

широком диапазоне масштабов. «Математические» или «идеальные» фракталы, типа Канторова множества, снежинок Коха, ковров Серпинского и другие, предельно четко демонстрируют такое поведение. В четвертой главе были приведены примеры таких объектов и проиллюстрированы эти в общем-то достаточно простые и интуитивно понятные идеи. Ценность математических фракталов состоит как раз в том, что они демонстрируют путь построения достаточно сложных и порой хаотических объектов с помощью набора всего нескольких элементов и простых рекуррентных правил, типа растяжений, поворотов, отражений и других топологических преобразований. Исследование уже известных математических фракталов и построение новых снабжает нас множеством идей в плане построения адекватных моделей реальных физических систем. Этими вопросами занимаются как серьезные математики, так и большая армия любителей, в основном программистов. Интересующимся этими вопросами можно порекомендовать сайт «Мультифрактальная Вселенная»: <http://worldmultimedia.biz/Science/> (там же можно найти ссылки на множество других сайтов).

«Природные» фракталы и стохастическое самоподобие, мультифракталы и спектры размерностей

В отличие от идеальных фракталов реальные объекты, как правило, более сложны. Эта сложность связана с конечным интервалом масштабов самоподобия, суперпозицией, взаимосвязью сразу нескольких скейлинговых соотношений (мультифрактальность) и др.

Одним из наиболее интересных и распространенных свойств природных объектов является *стохастическое самоподобие*. Такого рода стохастический скейлинг обнаруживают, например, ряд чисел Вольфа, ряды геофизических индексов, финансовые временные ряды, структура магнитного поля солнечных активных областей и др. Практически единственным надежным методом исследования структурных характеристик таких объектов является энтропийный подход, которому в этой книге уделено основное внимание. Вычисление последовательности энтропий Реньи и применение метода огрубленных покрытий позволяет найти не только *мультифрактальный спектр* таких объектов, но и определить интервалы масштабов, в которых обнаруживаются различные скейлинговые параметры. Анализ таких структурных свойств представляет особый интерес, так как свидетельствует о параллельном существовании в системе нескольких структурообразующих процессов.

Пример мультифрактального объекта был продемонстрирован в разделе 4.3.4. На этом примере продемонстрирован метод расчета первых трех фрактальных размерностей Реньи для неравновесной креативной природной среды, солнечной плазмы, в которой возникают и диссипируют сложные магнитоплазменные структуры — солнечные активные области. Пример показывает, что расчет фрактальных размерностей позволяет получить определенную дополнительную информацию о физических процессах в неравновесных средах. Рассмотрение фрактальных размерностей наряду с энтропиями и другими параметрами состояния неравновесных сред позволяет исследовать такие их свойства как появление, развитие и диссипация в них неравновесных структур, качественно и количественно описать специфические особенности этих структур, их сложность и степень коррелированности элементов.

Фрактальные размерности, связанные с длинами, площадями, массами. Фрактальные размерности аттракторов динамических систем

За недолгую, но достаточно бурную историю фрактальной геометрии накоплено множество разных способов определения фрактальных размерностей. Значительное число и разнообразие таких методов можно найти в уже цитированном сборнике [14]. По сути, любое скейлинговое соотношение, отражающее структурные особенности объекта исследования, определяет некую константу — показатель степенной зависимости, которую можно рассматривать как фрактальную размерность. Любой способ, позволяющий обнаружить скейлинговую зависимость и найти значение показателя степени, можно с полным правом называть методом или способом вычисления фрактальных размерностей.

Выше уже обсуждался целый ряд таких методов. Это метод «массарадиус», «периметр–площадь», «площадь поверхности – объем» и другие. Эти способы широко используются в изучении плазменной турбулентности, для анализа фрактальной структуры грозовых облаков с мощными конвективными движениями, в вычислительных алгоритмах по моделированию кластеров коррелированных спинов в модели Изинга и др. Особый вопрос в том, связаны ли эти размерности с размерностью Хауздорфа, и если да, то как? Для ряда размерностей, полученных с помощью таких методов, эта связь установлена. Для других – либо не установлена, либо ее вообще не существует. Некую разумную классификацию фрактальных размерностей дает нам энтропийный подход. Раз-

мерности Реньи разного порядка имеют достаточно четкий смысл, определяемый соответствующими энтропиями Реньи. Это дает возможность говорить о емкостных (Хаусдорфовых), информационных, корреляционных и других размерностях высших порядков.

Особо следует отметить определенное своеобразие методов фрактального анализа временных рядов, порождаемых сложными динамическими системами. Выше обсуждалось четыре таких специальных метода: R/S-анализ Херста, алгоритм Хигучи, метод Грассбергера–Прокаччия и расчет перестановочной размерности. Показатель Херста и размерность Хигучи связаны с размерностью Хаусдорфа. Метод Грассбергера–Прокаччия, в принципе, позволяет определить размерность аттрактора динамической системы (если это не простой полюс или вырожденный цикл, а так называемый «странный аттрактор»). Этот простой, мощный и информативный метод качественного анализа нелинейных динамических систем получил значительное развитие и имеет широкий круг приложений (см., например, [5]).

Фрактальная параметризация сложных систем

Параметризация — это процедура определения параметров, количественно характеризующих систему. Параметризация необходима, прежде всего, для количественной идентификации систем и их состояний. Если система параметризована, то можно количественно описать ее временное поведение, различать и сравнивать подобные друг другу системы, идентифицировать различные состояния одной и той же системы, оценивать влияние на систему внешних факторов и т. д.

Особую роль при этом играют бинарные или кратные отношения между параметрами. Они называются феноменологическими или эмпирическими законами. Известно много таких законов. Например, закон Гука связывает два параметра длину пружины и силу упругости. Закон Всемирного тяготения Ньютона — четыре параметра: массы двух тел, расстояние между ними и силу их взаимного притяжения. Уравнение Менделеева–Клапейрона — три параметра: давление, объем и температуру и т. д. Можно сказать, что физическая теория системы начинается с ее параметризации.

Феноменологическими параметрами, получаемыми в рамках подхода с использованием энтропийно-мультифрактального формализма, являются экстенсивные параметры — энтропии и интенсивные параметры, именуемые размерностями. В качестве информационной основы при вы-

боре таких параметров взяты идеи об информационной природе энтропии, скейлинговом поведении экспериментальных и наблюдательных данных, фрактальной симметрии и ее нарушениях в различных космических структурах или временных рядах. Анализ параметров, рассчитываемых при таком подходе, позволяет достаточно полно описать структурные особенности природных объектов, получать важную дополнительную количественную информацию о процессах в сложных неравновесных средах, в том числе и в динамических системах. Энтропийно-фрактальный анализ данных позволяет выявлять феноменологические соотношения, характерные масштабы и времена, прогнозировать на их основе возможное будущее поведение системы, что часто имеет важное практическое значение. Кроме того, феноменологическая информация, полученная в рамках энтропийно-фрактального подхода, может служить надежной количественной основой и четким критерием при разработке физических моделей различных космических систем.

Список литературы

1. Пайген Х.-О., Рихтер П. Х. Красота фракталов. Образы комплексных динамических систем. — М.: Мир, 1993.
2. Прибрам К. Нелокальность и локализация: голографическая гипотеза о функционировании мозга в процессах восприятия и памяти. // Синергетика и психология. М.: 1997.
3. Исаева В. В., Каретин Ю. А., Чернышев А. В., Шкуратов Д. Ю. Фракталы и хаос в биологическом морфогенезе / ИБМ ДВО РАН. Владивосток: 2004 г.
4. Лоскутов А. Ю., Михайлов А. С. Введение в синергетику. — М.: «Наука», 1990.
5. Кузницов С. П. Динамический хаос. — М.: Физматлит, 2001.
6. Могилевский Э. И. Фракталы на Солнце. — М.: Физматлит, 2001.
7. Зеленый Л. М., Милованов А. В. Фрактальная топология и странная кинетика: от теории преколяции к проблемам космической электродинамики. УФН: т. 174. № 8. 2004 г. С. 809–852.
8. Зеленый Л. М., Милованов А. В. Письма в Аастрон. Ж. 17, 1991 г. С. 425.
9. Чумак О. В., Чумак З. Н. Astron. and Astroph. Trans. 10, 1996 г. С. 329.
10. Кожевников Н. И. Труды ГАИШ, 53, 1983 г. С. 68.
11. Мандельброт Б. Б. Фрактальная геометрия природы / под. ред. В. А. Садовниченко и др. М.: Институт компьютерных исследований, 2002.

12. Пиблс Ф. Дж. Э. Структура Вселенной в больших масштабах. М.: Мир, 1983.
13. Ed. By Losa G. A., Merlini D., Nonnenmacher T. F., Welbel E. R. Fractals in Biology and Medicine. Birkhouser Verlag, Basel, 1997.
14. Фракталы в физике: Сборник. М.: Мир, 1988.

ПРИЛОЖЕНИЕ

Ральф Нельсон Эллиотт – автор оригинальной концепции изменчивости сложных систем

Введение. Лики времени

Размышляя об этом странном американце, я задумался об удивительной многоликости окружающего нас мира. Очень часто перед автором новой идеи мир вдруг раскрывается несколько по-иному. Сквозь призму своей идеи он начинает видеть такие его особенности и закономерности, которые никто ранее не замечал. Мир устроен удивительно. Он легко допускает, сразу и одновременно, несколько идей об одном и том же, в том числе и полностью противоположных, на первый взгляд, взаимоисключающих. Формальная логика и здравый смысл часто оказываются перед стеной непреодолимых противоречий при поиске ответа на вопрос: какая же все-таки из этих идей истинна.

Так, в III веке до новой эры «Евклид, сын Наукрата, известный под именем Геометр, ученый старого времени, по своему происхождению — грек, по месту жительства — сириец, родом из Тира...», один из величайших математиков понял, что мир состоит из точек, прямых линий, треугольников и параллелограммов. Этому видению мира нас до сих пор учат в школах... Вся природа «говорит» на языке гладких функций и дифференциальных уравнений, полагали величайшие математики нового времени. Этому видению мира мы учимся в вузах. Последние десятилетия, вслед за Бенуа Мандельбротом, мы постигаем, что в мире нет прямых линий, окружностей, треугольников, нет в нем и гладких функций. Весь мир и все его отдельные части — фракталы...

А вот, например, что пишет Секст Эмпирик о природе времени. «Одни говорят, что время — это то, в чем движется мир, другие же говорят: оно само движение мира. Аристотель и, как считают некоторые, Платон признают время числом (мерой или отношением) того, что находится в движении раньше или позже. Стратон ... признает его мерой движения и пребывания. Эпикур ... считает его свойством свойств, сопровождаю-

щим дни и ночи, и часы, и состояния, и отсутствие состояний, и движения, и пребывания. Последователи Энцидема говорят, что по сущности время – тело, другие говорят, что оно бестелесно.... Но все эти мнения не могут быть истинными (ибо большинство противоречат друг другу)... Кроме того, если признать ложью, что время является телом, а также ложью — что оно бестелесно, то будет признано само собою не существование времени, ибо кроме этого не может быть ничего другого. Но невозможно также понять, какие мнения о времени истинны, а какие ложны вследствие равносильного разногласия и затруднений по поводу критерия и доказательств. Стало быть, в силу этого мы ничего не можем утверждать о времени» (Секст Эмпирик. Три книги Пирроновых основоположений. Минск–Москва: 2000 г. С. 206–207).

Да, если рассуждать строго в рамках формальной логики, то легко можно прийти к агностицизму. Но все дело в том, что окружающий нас мир выходит далеко за пределы формальной логики и здравого смысла. Если мы пытаемся узнать о мире нечто новое, нам приходится выходить за пределы традиционных инструментов познания, принимать мир таким, какой он есть. Но (осторожно — математика!) тем самым мы становимся на тропу весьма рискованных приключений разума... «Снова я убеждаюсь, что надо вступить в область мрака, признать не вмещающее никаким рассудком совпадение противоположностей и искать истину там, где встает перед глазами невозможность», писал в этой связи Николай Кузанский.

В качестве хрестоматийного примера можно привести известный спор между Ньютоном и Гюйгенсом о природе света. Ньютон, как известно, полагал, что свет — это поток корпускул. Гюйгенс считал, что свет — это волновой процесс. Два формально взаимоисключающих друга представления! Но поскольку и та, и другая точка зрения подтверждалась наблюдениями, то невозможно было понять, какое из этих двух мнений о природе света истинно, а какое ложно *«в силу равносильного разногласия и затруднений по поводу критерия и доказательств»*. Далее, следуя Сексту Эмпирику, можно было сделать вывод: *«Стало быть, в силу этого мы ничего не можем утверждать о природе света»*.

Оказалось же, как мы теперь знаем, что и та, и другая точка зрения правильны. Свет — это и волновой процесс, и поток корпускул. Это два лика единой сущности. Основываясь на идее корпускул и множестве экспериментов, Ньютон создал изумительную по красоте теорию распространения света — геометрическую оптику и построил первый в мире зеркальный телескоп. Волновая идея Гюйгенса легла в основу всей современной спектроскопии, лазерной техники и других многочисленных приложений.

Время — достаточно сложный объект исследования, и мы должны быть готовы принять самые разные точки зрения на его природу. При этом не следует упускать из виду, что правильными могут оказаться сразу несколько идей, причем некоторые из них могут находиться, на первый взгляд, в явном и непримиримом противоречии друг с другом. Возможно, время является не двуликим или трехликим, а многоликим феноменом, и все эти лики времени могут оказаться истинными.

Например, что общего между монетой и песочными часами? Вроде бы, ничего. Но ведь и то и другое можно рассматривать как некие модели времени, причем концептуально противоположные! Как-то на археологических раскопках я разглядывал монету с изображением двуликого Януса — Бога входов и выходов, Бога времени и Творца мира. Посмотрев на нее с ребра, я подумал: «Вот еще одна “модель времени”, своеобразный “глобус” времени». Мне подумалось, что, возможно, я понял тех наших далеких предков, которые придумали и чеканили эту монету. В самом деле, один лик Бога смотрит влево, символизируя прошлое, другой смотрит вправо, символизируя будущее. Настоящее же есть нечто тонкое и неуловимое, как ребро монеты. Но только оно одно материально, как эта монета, только оно реально, только оно и существует. «Есть только миг между прошлым и будущим». Прошлое и будущее — это лишь зыбкие фигуры ума. Прошлого уже нет, а будущего еще нет. Прошлое — лишь в нашей памяти, будущее — лишь в нашем воображении. Монета как бы говорила: «Есть только миг, за него и держись!» Держись за настоящее! Все остальное – либо иллюзии, либо фантазии.

Иное дело песочные часы. Здесь время — это река, материальный поток, нечто вечное, существующее всегда. Будущее — наверху, прошлое — внизу. Настоящее же — это всего лишь быстрый, суетливый, неуловимый и всегда ускользающий ручеек, зажаты между этих двух неподвижных бесконечностей. Кончился определенный период, часы переворачиваются, время выходит на новый цикл, прошлое становится будущим, а будущее — прошлым, и мы снова должны повторять не усвоенный ранее урок. Песочные часы как бы говорят: «Помни о вечном! Настоящее — суетно и преходяще. Жизнь очень быстро протечет, как песок сквозь пальцы. Только вечное — реально. И усваивай, пожалуйста, прилежнее свои уроки! А то, не ровен час, время не сделает для тебя очередной свой оборот, и текущий цикл окажется для тебя последним. Иссякнет будущее, прекратится настоящее, кончится время...»

Кто же прав — Монета или Песочные часы? Всё дело в том, что правы оба! Время, возможно, всего лишь реляция, и в зависимости от субстанции эта реляция может принимать самые разные лики: как четвертая

координата в теории относительности, как мера изменчивости в сложных системах, как первичный атрибут мира или просто как упорядоченная последовательность событий. Это всего лишь лики одного и того же многоликого феномена — Его Величества Времени. И вот, нашему герою, финансовому аналитику из Калифорнии, удалось разглядеть еще один совершенно удивительный лик времени. Но начнем по порядку.

Краткая история знаменитого Ральфа Нельсона Эллиотта, великого аналитика, перевернувшего наши представления о принципах анализа рынка ценных бумаг и открывшего один из самых красивых законов природы¹⁰



Рис. 1. «Бычья» серия волн Эллиотта. Хорошо просматривается открытое Эллиоттом масштабное самоподобие изменения переменной. Оси координат умышленно отсутствуют, так как, по мнению Эллиота, такие серии волн носят универсальный характер. Их можно наблюдать во временных рядах любой природы и на любых временных масштабах (более подробное описание волн Эллиотта — в конце Приложения)

В начале прошлого века в Соединенных Штатах Америки важнейшими центрами деловой и социальной активности были Stocks markets —

¹⁰ В изложении биографии Р. Н. Эллиотта я следую в основном статьям Роберта Прехтера, появившихся в сентябрьских и октябрьских выпусках журнала *Futures* 1994 года. На русском языке биография публикуется впервые.

рынки товаров, ценных бумаг или фондовые биржи. Они и сейчас не утратили своего былого значения. В менталитете американцев, в отличие от нас, европейцев, все, что происходит на этих рынках, занимает значительно более важное место. Для американца рынки ценных бумаг это не только место, где можно купить или продать акции предприятий. Это, прежде всего, барометры экономического состояния компаний, деловой активности регионов и всей страны. Они же и источник заработка и одновременно увлекательная, азартная игра миллионов. Это место, где можно посоревноваться в уме и проницательности, предсказывая возможное развитие событий. Неудивительно, что дитя своей страны Эллиотт пришел к открытию своего «мирового закона» через рынок...

Свою теорию волн трудоспособный и упорный, как бык, Эллиотт подтвердил своей собственной биографией. Подъемы и спады на его жизненном пути сменяли друг друга, и в этих взлетах и падениях можно ясно различить все характерные признаки его знаменитой «бычьей» серии волн (рис. 1)...

Волна первая «Первичная экспансия»

Ральф Нельсон Эллиотт родился 28 июля 1871 года в Мерисвилле (Кан), в потомственной американской семье. В конце 1880 года юный Ральф Эллиотт вместе с семьей переезжает в Сан-Антонио (Техас). В возрасте 10 лет он уже бегло говорит, читает и пишет по-испански и с юношеским пылом влюбляется в Мексику, находящуюся всего в каких-то 150 милях к югу. В 1891 году, на 21 году жизни, Ральф навсегда покидает родительский дом и устраивается на работу в железнодорожную компанию в любезном его сердцу Мехико. Это случилось на самой вершине североамериканского железнодорожного бума. Здесь, в Мексике, он провел следующие 20 лет своей жизни. Он занимал самые разные должности. Был и путевым обходчиком, и диспетчером, и стенографом, и оператором телеграфа, и станционным смотрителем.

В 1896 году Эллиотт начинает работать в бухгалтерии (где он получил соответствующее образование – неизвестно). Его опыт, детальное практическое знание производства помогают ему разработать специальную систему экономического учета для железных дорог. В 1903 году Ральф женится на Мери Элизабет Фитцпатрик (1869–1941) из Нью-Йорка. И в течение 25 лет занимает руководящие посты в железнодорожных компаниях Мексики и Центральной Америки.

Эллиотт рассказывает нам о своих многочисленных приключениях в Латиноамериканском регионе. В этих рассказах он предстает человеком авантюрным и бесшабашным, чем это можно было предположить, принимая во внимание его профессию. Например, он вспоминает об одном инциденте, из которого можно сделать вывод, что он был превосходным наездником. Эллиотт также очень красочно описывает местные обычаи, рассказывая, как о невероятном богатстве и роскоши, так и о бедности, грязи и нищете. В его ярких бытовых зарисовках можно найти описания трапез местных жителей, когда «белки, ящерицы, попугаи, лисицы и даже змеи готовились и поедались с видимым удовольствием». Такие наблюдения не могли быть сделаны человеком, который проводил свое время в комфортабельных отелях или офисах корпораций. Неудивительно, что на своих 40-летних фотографиях он выглядит как человек крепкого телосложения (рис. 2), мужественный и уверенный в себе. Все это говорит о том, что в это время Эллиотт вел весьма активный образ жизни.

По мере занятий своей профессией его служебное положение все время повышалось. Его интересы были значительно шире интересов простого бухгалтера. Он прямо и окровенено пишет в одной из своих статей: «Я не люблю бухгалтерию, но я люблю работу, которую я делаю». Эта работа —



Рис. 2. Р. Н. Эллиотт. 1911 г.

организация бизнеса. Именно в целях более успешного ведения бизнеса он реорганизовывает финансовую структуру многочисленных корпораций, вводя новые системы ведения учета. Одна из идей — прогноз и учет возможных будущих затрат и их компенсация из специальных фондов предприятия, организуемых по принципу налоговых процентных отчислений. Все это в другой своей журнальной статье он назвал «только некоторым методом контроля успешности бизнеса». Похоже на то, что как независимый бизнес-консультант Эллиотт имеет в это время множество клиентов, но его неизменный принцип — оставаться в компании, пока ее реструктуризация не завершена. Со временем Эллиотт приобретает репутацию известного эксперта по организации бизнеса.

организация бизнеса. Именно в целях более успешного ведения бизнеса он реорганизовывает финансовую структуру многочисленных корпораций, вводя новые системы ведения учета. Одна из идей — прогноз и учет возможных будущих затрат и их компенсация из специальных фондов предприятия, организуемых по принципу налоговых процентных отчислений. Все это в другой своей журнальной статье он назвал «только некоторым методом контроля успешности бизнеса». Похоже на то, что как независимый бизнес-консультант Эллиотт имеет в это время множество клиентов, но его неизменный принцип — оставаться в компании, пока ее реструктуризация не завершена. Со временем Эллиотт приобретает репутацию известного эксперта по организации бизнеса.

Эллиотт имеет возможность оставаться в Мексике до конца своих дней, но обстоятельства вынуждают его возвратиться в Соединенные Штаты. Начиная с 1911 года, Мексику потрясает серия бешеных революций, продолжающаяся несколько лет. В конце концов, гражданская война в стране достигает критической точки, и в июне 1916 года, «когда Президент (Вильсон) приказал всем американцам покинуть» Мексику, Эллиотт собирается и со своей женой возвращается в Лос-Анжелес.

Волна вторая. «Ритренимент (откат) и консолидация»

Выброшенный из своих привычных условий работы и обитания, Эллиотт переживает период неудовлетворенности ситуацией, в которой оказался. В течение четырех лет он дважды меняет работу, дважды изучает возможность сменить работу и временами планирует снова возвратиться в Латинскую Америку.

Возможности Эллиотта как бизнес-консультанта были широко известны и высоко ценились. Одно из предложений, которое он рассматривал, было с Кубы. Он получил его по телеграфу. В нем, в частности, были такие слова: «Ваши услуги *крайне* необходимы». Позднее ему пришло предложение от Кубинской железнодорожной компании, несмотря на то, что эта должность была уже занята кем-то другим. Вице-президент и генеральный менеджер компании писали Эллиотту: «Если, как мы думаем, вы приедете, то это место мы освободим специально для вас». Подобные энергичные предложения не были для Эллиотта чем-то необычным. В то время он был одним из наиболее талантливых двуязычных бизнес-консультантов и реорганизаторов корпораций для многих Латиноамериканских компаний.

Несмотря на неблагоприятный политический климат, в 1918 году Эллиотт все же возвращается в Мексику. Он работает как аудитор в принадлежащей США компании Pierce Oil Corp. Его жена остается в Ашвилле (N.C.) вдали от потенциальной опасности. Наконец, в начале 1920 года, Эллиотт переезжает в Нью-Йорк. Возможно, этот переезд состоялся в результате служебного перевода, возможно — в связи с тем, что его жена была родом из Нью-Йорка. Как бы то ни было, но этот переезд обозначил крутой поворот в жизни Эллиотта. Вскоре он оставляет попытки найти работу и жилье в Латинской Америке. Неожиданно для себя здесь, в Нью-Йорке, Эллиотт находит занятие, вполне подходящее для его неумной и авантюристичной натуры.

Волна третья. «Длинный подъем»

В Нью-Йорке в 50 с небольшим лет в жизни Эллиота начинается весьма насыщенный, продуктивный период. Это подтверждается записями в его деловом расписании, которые он ведет протяжении семи лет. Позднее из его письма Чарльзу Дж. Коллинзу выясняется, что в этот период он также путешествовал по Канаде, посетил Германию, Англию и Францию, по каким причинам (личного или делового характера) — неизвестно. В этот же период он провел реструктуризацию крупной железнодорожной компании Amsinck & Co., экспортно-импортной организации со штатом в 500 человек. Одновременно он специализируется в новой для себя сфере — работает в качестве бизнес-консультанта в ресторанах, кафетериях и чайных. В целях поддержки и развития своей новой специальности летом 1924 года Эллиотт входит в редколлегию нью-йоркского ежемесячного бизнес-журнала «Чайная комната и Магазин подарков».

Термин «чайная комната» сегодня уже вышел из употребления, но в Америке 20-х годов чайные комнаты ассоциировались с очень крутым бизнесом. Популярность «чайных комнат» распространялась, как пожар, и все предприниматели, занятые в ресторанном и гостиничном бизнесе, пытались прибрать к рукам этот лакомый кусок. Журнал «Чайная комната и Магазин подарков» гордился в то время своими 3000 постоянными читателями, разбросанными по всему миру. Приход Эллиотта в журнал был отмечен торжественно: почти целая страница была посвящена его представлению читателям. Ежемесячные статьи Эллиотта появлялись в разделе «Коммерческий бюллетень» под рубрикой «Научный менеджмент». В них утверждалось, что правильно поставленный бухгалтерский учет является важнейшим инструментом организации бизнеса. Содержание этих статей выходило далеко за пределы традиционной бухгалтерии. «*Что я пытаюсь делать?*», — резюмировал автор, — «*я помогаю вам делать деньги*». Позднее эта цель была с блеском реализована в его фантастической карьере прогнозиста на фондовой бирже.

Оценкой статей Эллиотта в области ресторанной бухгалтерии и менеджмента может служить приглашение его в Колумбийский университет прочесть лекции о предмете. Эллиотт не принял приглашение, так как снова по делам покинул Нью-Йорк...

Предприимчивость, инициативность, настойчивость Эллиотта, его работа в корпорациях на протяжении многих лет сводили его с влиятельными людьми из академических и политических кругов. Одним из них был доктор И. У. Дженкс (Dr. Jeremiah Whipple Jenks) — знаменитый адвокат, академик, политолог, автор двух дюжины книг о политике, социологии, религии и бизнесе. Этот незаурядный человек также состоял в со-

вете директоров нескольких железнодорожных компаний, в том числе Тихоокеанской Железной дороги Никарагуа. Эллиотт, естественно, встречался с Дженксом и как член профессиональной ассоциации, и вследствие их общих интересов к железнодорожной индустрии, финансам и Центральной Америке. Их дружба, в конце концов, привела к неожиданному повороту в карьере Эллиотта.

В 1912 году старое либеральное правительство Никарагуа было свергнуто в результате революции. Американские морпехи тут же вошли в страну «для эффективного установления административного контроля за ситуацией, правительством соединенных Штатов и защиты американских интересов в Никарагуа». Однако спустя 12 лет Гос. Деп. США «устал» от этой своей роли и назначил Высшую Комиссию по Никарагуа для разработки предложений по стабилизации проамериканского правительства этой страны. Это нужно было сделать, прежде всего, для того, чтобы вывести из Никарагуа уже власть навоевавшихся и висящих бременем на бюджете США доблестных морпехов. Дженкс вошел в состав этой комиссии и был наделен полномочиями по установлению новой банковской системы, позволяющей контролировать все финансовые потоки в стране. Это давало возможность держать под контролем страну, не прибегая к военной силе. Впоследствии эта никарагуанская схема, с успехом применялась и в других странах.

По рекомендации Дженкса Эллиотт был избран Гос. Департаментом на должность главного финансиста Никарагуа. 18 декабря 1924 года Эллиотт встречается с Государственным Секретарем Эвансом Хагисом в Вашингтоне для получения официального назначения и инструкций. В Манагуа он выезжает в феврале 1925 г. Здесь Эллиотт получил возможность приложить свой опыт реорганизации корпораций для реструктуризации финансовой системы целой страны.

Как и было запланировано, Эллиотт исполнял свои официальные обязанности в течение двух лет, пока США вытаскивали себя из трясины Никарагуа. Наконец, к июню 1926 года США отзывает всех служащих Гос. Депа, выводит из страны всех морпехов, полагая, что спокойствие и *нужный* порядок наконец установлены. После краткого пребывания в Нью-Йорке в августе Эллиотт отправляется в Гватемала-Сити с другим важным деловым назначением — на должность главного аудитора Международной Железной дороги Центральной Америки. Акции этой компании котировались на Нью-Йоркской и Лондонской Фондовых Биржах. Это была последняя должность Эллиотта в его традиционной профессиональной сфере.

В августе 1926 года в издательстве Little, Brown & Co. выходит его книга «*Менеджмент чайных комнат и кафетериев*». Эта книга пред-

ставляет собой переработанную и дополненную версию его журнальных статей. Первые положительные отзывы на книгу появились в газете «Нью-Йорк Геральд Трибюн» и в Книжном Обозрении «Нью-Йорк Таймс» с комментариями в том духе, что «*м-р Эллиотт является авторитетом во всех этих областях, поскольку его широкий и разнообразный бизнес дал ему достаточно наблюдений и опыта*». Эллиотт здесь представлен как «эксперт-организатор» — эту роль он позже возьмет на себя и в следующей своей книге «*Принцип волн*» (“*The Wave Principle*”). В этой книге Эллиотт описывал циклы деловой активности, как «*приливы и отливы в потоке событий*». Позже, по аналогии с морскими приливами и отливами, он это свое понимание циклов деловой активности обозначит словом «*волны*». Конечно, этот термин в книгах Эллиотта следует воспринимать лишь как гидродинамическую метафору. О самой книге «*Принцип волн*» мы поговорим позже.

Возвращаясь к первой книге Эллиотта, следует отметить, что с ее публикацией он возобновляет свою карьеру специалиста по ресторанному менеджменту. Он оставляет должность (с зарплатой, которая по сегодняшнему курсу составляла \$170000 после всех вычетов) и возвращается в Нью-Йорк с совершенно определенной целью — поддержать эту книгу и утвердить себя в роли лучшего консультанта в этой области. Из своей временной резиденции в Уолкотт Отеле (WOLCOTT HOTEL), завязывает контакты с издателями для продвижения своей книги.

Однако из-за его бурной профессиональной активности ему не приходилось долго заниматься только одним делом. Гос. Деп. направляет его снова в Никарагуа, чтобы использовать его талант в другой области — в политике. Во время работы в Никарагуа Эллиотт предлагает вниманию издателей свой новый проект. Как можно заключить из письма, находящегося в папках спецхрана Гос.Деп., этот проект относится к его новой книге «*Будущее Латинской Америки*». Рукопись в сто страниц, которую Эллиотт направил в Государственный Департамент, составляла примерно половину от планируемой им книги. Эта рукопись была благополучно похоронена в Национальных Архивах США и открыта лишь 67 лет спустя в процессе исследований при написании его биографии Р. Прехтером.

Целью этой работы Эллиотта был анализ экономических и социальных проблем Латинской Америки, разработка плана мероприятий по созданию экономической стабильности и предпосылок для дальнейшего процветания региона. Эллиотт установил, что главным препятствием на пути экономического прогресса в Латинской Америке является непосильное бремя долгов и необдуманно высокомерная позиция работодателей в вопросах заработной платы.

Эллиотт обрисовал всеобъемлющий план, который следует реализовывать, если и когда «*какая-либо латиноамериканская страна обратится к Соединенным Штатам с просьбой о финансовой или политической поддержке*». План, прежде всего, предусматривал поддержку Соединенными Штатами выпуска страной облигаций национального займа (ГКО). Затем был предусмотрен выпуск под гарантии США специальных облигаций для реструктуризации и частичной оплаты чрезмерных долгов этой страны, а также для реформирования и развития инфраструктуры, реформирования сферы гражданских услуг и для финансирования компаний информационной поддержки всех этих мероприятий как в Латинской Америке, так и в США.

Многими путями эти его предложения позднее попадали в такие программы как «Добрые соседи» в политике администрации Франклина Д. Рузвельта, в современные программы Всемирного Банка, Международного Валютного Фонда¹¹ и др. Эллиотт получал жалование в Административной Коллегии, что следует из документа Государственного Департамента, датированного 2 февраля 1929 г. Разумеется, он был в постоянной обойме потенциальных кандидатов на правительственные должности в Никарагуа.

Однако какое бы влияние ни оказали политические идеи Эллиотта на Латинскую Америку, все равно они менее значимы, чем его более поздние достижения в открытии Принципа Волн. В этом плане его книгу «*Будущее Латинской Америки*» можно рассматривать лишь как предварительные и промежуточные догадки, которые затем были развиты, детализированы и нашли свое место в перспективной единой картине, которая позволила ему открыть и сформулировать Принцип Волн. Вот, например, всего один пассаж из «*Будущего Латинской Америки*», который, в частности, хорошо иллюстрирует взгляды Эллиотта на природные законы: «В предыдущих главах мы подвели читателя к выводу, что про-

¹¹ Эллиотт, предлагая свои финансовые схемы и инструменты, безусловно, преследовал благородные цели. Однако, как всякая научная разработка, они могут использоваться по-разному. В том числе они легко могут быть превращены в орудие подрыва национальных экономик, финансового и политического закабаления стран. Увеличивая процентные ставки по кредитам, молчаливо поощряя коррупцию и вывоз коррупционных доходов в «надежные» банки, можно добиться потрясающих результатов. Правящая верхушка и ее силовые структуры становятся *клиентами* этих банков и не способны проводить независимую, национально ориентированную политику. Страна постепенно попадает под внешнее управление, превращается в банкрота и скупается за бесценок, теряя всю свою национальную независимость. Никарагуа лишь одна из первых жертв этой политики, за ней были многие другие. Удивительно, что и Россия с ее неслабым интеллектуальным потенциалом также умудрилась наступить на эти примитивные грабли.

блемы Соединенных Штатов и Латинской Америки поддаются только совместному решению. Это может показаться случайностью, но скорее всего это связано с какими-то, еще не до конца понятыми природными законами, когда то, чего не достает в Соединенных Штатах, может быть в избытке найдено в Латинской Америке, и то, что требуется Латинской Америке проще всего получить от Соединенных Штатов».

Здесь Эллиотт предполагает, что *природа стремится к балансу: в каждой своей части недостатки в чем-то одном компенсируются избытком чего-то другого*. В этой идее уже содержится намек на ритмику, или динамический баланс, который он позже обнаружит на фондовой бирже.

Одна книга разошлась, новая в работе... Таким образом, Эллиотт имеет, по крайней мере две веские причины для возвращения в Соединенные Штаты. Но имеется и третья причина: Эллиотт страдает тяжелой формой местной болезни пищевого тракта, обусловленной амебным гистололизом. Стиль жизни Эллиотта и его характер не оставляли никаких шансов на серьезное лечение на месте. Соединенные Штаты предлагают ему пройти экспертное медицинское обследование.

В 1927 году он возвращается в Лос-Анжелес, чтобы вести более оседлый образ жизни после 36 лет интенсивной работы, путешествий и скитаний по гостиницам. Лишившись всех своих прежних деловых контактов, он сосредотачивается на перебазировании в Лос-Анжелес своего консалтингового бизнеса, пока предпринимаются попытки лечения его болезни. Отметим примечательный факт: у человека, который позднее прилагал последовательность чисел Фибоначчи к человеческой жизни, все эти изменения в карьере и стиле жизни случились, когда Эллиотту было 55 лет, — за 21 год до его смерти...

Волна четвертая. «Надлом. Коррекция и консолидация»

Известность Эллиотта, основанная на выдающейся карьере, его книге и длинном списке ссылок на него, однажды возросла еще больше. Отзывы на книгу оставались по-прежнему положительными. И вот Национальная Ассоциация Ресторанов приглашает его выступить в Буффало. *«Цена моих услуг, — сказано было в одном из его рекламных объявлений, публикуемых Ассоциацией, — выход на дополнительные доходы»*. С такой поддержкой он быстро приобретает много клиентов на новом месте жительства.

В то время, когда будущее Эллиотта казалось лучезарным, несчастье настигло его внезапно. Крайне сильное обострение его болезни в течение

1929 года привело к развитию разрушительной анемии, приковавшей его к постели. Безрассудно смелый, авантюристичный, чрезвычайно деятельный Эллиотт был вынужден уйти в отставку. В течение следующих пяти лет он несколько раз находился на грани жизни и смерти. Его фотографии в журнале «Financial World» десятилетием позже показывают, что безжалостный недуг сделал свое черное дело — Эллиотт на этих снимках выглядит далеко не так, как в более ранние годы, он изможден и худ...

Эллиотт крайне нуждался в чем-то, что могло бы полностью занять его ум в периоды ухудшения самочувствия, между жестокими накатами болезни... Как раз в эти годы рынок ценных бумаг в США переживал наиболее динамичный период своей истории: за пиком мощного «бычьего» рынка 20-х годов последовал оглушительный «медвежий» обвал. Эти события приковали его внимание. Он читает книгу Роберта Ри (Robert Rhea) «Теория Доу» (1932 г.) и становится одним из первых подписчиков периодического издания Р. Ри «Комментарии к теории Доу». Начиная с этого времени, Эллиотт полностью переключает свое внимание на изучение поведения рынка ценных бумаг.

Открытие. Подобно тому, что проделывал с Теорией Доу Роберт Ри, также в это время прикованный к постели из-за туберкулеза Эллиотт, проводивший долгие часы над изучением этой теории в периоды относительного затишья своей болезни, начал делать собственные наблюдения за движением цен на рынке акций. Вскоре он обнаруживает определенные повторяющиеся структуры в графиках движения цен, которые не укладываются в рамки Теории Доу. Он убежден, что открыл некую неизвестную никому ранее закономерность в поведении цен. Чтобы проверить это свое убеждение на рынке в реальных торгах, он исследует годовые, месячные, недельные, ежедневные, часовые и получасовые графики различных индексов поведения рынка на протяжении 75 лет. Он конструирует часовые графики из рядов индекса Доу–Джонса, которые начали регулярно публиковаться с 5 октября 1932 г., а также получасовые графики из чисел, которые он собирал в трейдинговом зале брокерской фирмы. Эллиотт целиком захвачен этой новой идеей, искренне полагая, что, как он писал в своей рукописи о Латинской Америке, *«Смысл есть во всем, и наша задача — попытаться его открыть»*.

В мае 1934 г., через два месяца после своей очередной встречи со смертью, все наблюдения за поведением рынков акций стали складываться у Эллиотта в некую стройную единую систему принципов, которая четко описывала все волновые движения падающих в этот период средних цен на акции. В наше время значительная часть наблюдательных фактов, подмеченных Эллиоттом и положенных им в основу его системы принципов временного поведения цен на рынках, обозначается емким

научным термином «фрактал». Этот термин сейчас ассоциируется с бурно развивающейся в последние десятилетия *хаотической динамикой*, хотя появился он намного раньше и употреблялся для описания компонентов различных самоорганизующихся структур и определенных связей между их компонентами. Таким образом, бывший «эксперт организатор бизнеса» в результате тщательных наблюдений и исследований открыл *самоорганизационный принцип*, которому следует движение рынков. Когда в течение нескольких последующих месяцев Эллиотт стал применять свои принципы к прогнозу поведения рынка, он испытал, как он позже писал, *«нечто подобное тому, что испытывает изобретатель, пытающийся научиться искусно управлять машиной, которую сам изобрел»*. По мере того как он все более и более профессионально учился пользоваться своими принципами и устранял мелкие ошибки в их формулировках, их точность изумляла его.

В это время финансовое состояние Эллиотта стало весьма непрочным из-за крупных трат, связанных с болезнью и с ограниченной возможностью вести его традиционный бизнес. Кроме того, часть активов его жены пострадала в результате обвала фондового рынка в 1929–1932 гг. Тяжелое финансовое положение, растущий интерес к процессам, происходящим на фондовом рынке, и сделанное им открытие в области динамики цен — все это вместе привело к тому, что Эллиотт принимает решение попробовать себя в новой профессии. Он решает *«начать все сначала»* и, как он пишет, *«работать в области, которую люблю и которая стала уже неотъемлемой частью ... жизни и борьбы»*. Вот так, в возрасте 64 лет, Эллиотт начинает свою новую карьеру, и в его биографии начинается период, который позже он будет называть как...

«Волна номер пять моей собственной жизни»

Ноябрь 1934 г. Р. Н. Эллиотт твердо уверен в истинности своей идеи и решает представить ее по крайней мере одному из членов финансового сообщества. В это время Эллиотт был подписчиком рыночного обозрения (биржевого бюллетеня), основанного и редактируемого Чарльзом Дж. Коллинзом и публиковавшимся Советом Корпорации Инвесторов Детройта. Эллиотт чувствовал, что он достаточно хорошо изучил Коллинза по его публикациям в бюллетене, чтобы доверить ему свое открытие. Это предположение, к счастью Эллиотта, оказалось верным.

19 февраля, после обмена письмами, Эллиотт посылает Коллинзу 17-страничный трактат, озаглавленный как *«Принцип Волн»*. Первая страница трактата содержит обоснование полезности Принципа Волн: *«Тщательное исследование явлений отката в структурах изменения цен*

позволяет выявить надежные признаки, не всегда ярко выраженные, но, тем не менее, позволяющие надежно выявить тренд и получить ясные предупреждения о развороте тенденции».

Говоря об искусстве адекватного пользования его теорией, он замечает: «Волны не ошибаются, неправильными могут быть те или иные мои представления о них. Чем ближе будут наши приближения к истинному закону, тем меньше ошибок мы будем делать». Говоря о пяти-волновой структуре фондовых рынков, Эллиотт заключает: «Я не понимаю, каковы возможные причины происхождения этих структур и, возможно, никогда об этом не узнаю, но структуры эти существуют и это закон природы. Законы природы, в том числе и в экономике, достаточно жестки, и именно они определяют то, что должно произойти».

Коллинз отбивался от своих многочисленных корреспондентов, которые присылали ему различные системы биржевой игры, следующим образом. Он предлагал им делать прогнозы будущего состояния рынка, справедливо полагая, что любая стоящая система должна показать, на что она способна, будучи опробована в реальном времени и на реальном рынке. Не удивительно, что после таких испытаний множество этих систем обращались в прах. С принципом Эллиотта, однако, все получилось по-другому...

Индекс Доу–Джонса в начале 1935 года понижался, и Эллиотт ежедневно посылал телеграммы с прогнозами, которые выполнялись с высокой точностью. Во второй неделе февраля железнодорожный Доу–Джонс, как и предсказывал Эллиотт, пробил свой самый низкий уровень 1934 года — 33.19. Консультанты были очень озабочены этой негативной тенденцией и, вспоминая обвал 1929–1932 гг., говорили об утверждении «медвежьей» тенденции в экономическом развитии страны. Далее происходят следующие события: промышленный Доу падает еще на 11 % и достигает отметки 96, а железнодорожный Доу (в то время более значимый) падает на 50% от своего пика в 1933 г. и выходит на уровень 27...

В среду, 13 марта 1935 г., сразу после закрытия торгов вблизи самого низкого значения за день индекса Доу–Джонса Эллиотт шлет Коллинзу знаменитую телеграмму: «Несмотря на медвежий (Доу), все индексы достигли *финального* минимума». Коллинз читает эту телеграмму утром следующего дня — в четверг, 14 марта 1935 г., — того самого дня, когда промышленный Доу достиг минимума этого года. Днем раньше упомянутой телеграммы, во вторник, 12 марта, рынок 1935 года прошел минимум железнодорожного Доу–Джонса. 13-месячная коррективная волна закончилась, и рынок сразу же развернулся на подъем.

Спустя два месяца, в то время, когда рынок продолжал карабкаться вверх, Коллинз, «впечатленный догматизмом и точностью» прогнозов Эллиотта, предлагает опубликовать его прогнозы в Investment Counsel

бюллетене и пишет в этой связи: «По нашему мнению, Принцип Волн является наилучшим методом прогноза, который когда-либо попадал в сферу нашего внимания».

Эллиотт принимает предложение Коллинза на публикацию прогнозов в течение двух лет. Если «Investment Counsel» подтвердит успешность прогнозов Эллиотта после этого периода, тогда Коллинз, в котором Эллиотт увидел прекрасного писателя, подготовит книгу по Принципу Волн для публичной продажи. На протяжении двух лет Коллинз внимательно отслеживает все сообщения Эллиотта на реальном рынке. Точность прогнозов подтверждается, и в конце второго года (в марте 1937 г.) Коллинз начинает работать над книгой, целиком основанной на оригинальных статьях и записках Эллиотта. Книга «*Принцип Волн*» (“The Wave Principle”) была опубликована 31 августа 1938 г.

Первая глава книги начинается словами: «Нет истины, которая бы встречала большее понимание, чем та, что мир управляется законами. Без законов, очевидно, был бы хаос, а где хаос, там ничего нет... Многочисленные исследования, относящиеся к ... социальной активности? указывают, что практически все явления, результатом которых являются наши социально-экономические процессы, следуют законам, которые обуславливают их самоподобные повторения, объединяющиеся в рекуррентные серии волн или импульсов определенного числа и вида... Фондовый рынок иллюстрирует волновые импульсы, соответствующие общей социально-экономической активности... Таков закон, и он справедлив также и для других процессов в природе».

Через несколько недель после публикации своей книги Эллиотт пакует вещи и переезжает со своей женой в Columbia Heights (Бруклин), находящийся на расстоянии короткой остановки метро от финансового района Манхэттена.

10 ноября Эллиотт публикует первое из длинной серии «Пояснительных Записок», в которых он описывает текущее состояние рынка и делает прогноз его дальнейшего развития в терминах Принципа Волн. Эллиотт публикует эти свои письма объемом от одной до четырех страниц не регулярно, а «как требуют обстоятельства» — по три–семь ежегодно, с 10 ноября 1938 г. по 6 августа 1945 г. За эти свои письма он получает по \$60 в год и продолжает продавать свою монографию, которую он называет «Трактат», по \$15. Ральф Эллиотт снова в седле. Он имеет свой независимый бизнес, как он и планировал 11 лет тому назад.

Постоянный автор журнала «Мир Финансов» (“*Financial World*”) Коллинз знакомит редакторов с Эллиоттом и его работой в начале 1939 г. Эллиотту поручают написать 12 статей о Принципе Волн. Эти статьи были опубликованы между апрелем и июлем 1939 г. Они сделали Элли-

отта известным среди сообщества инвесторов и создали ему имя ведущего аналитика.

После публикации статей в журнале «Мир финансов» Эллиотт начинает писать серию глубоких эссе, которые быстро вырастают в строгий Образовательный Курс (Educational Service), который публикуется с 1940 по 1944 г. В одном из ранних выпусков «Образовательного Бюллетеня» Эллиоттом была опубликована фундаментальная работа, которая подняла Принцип Волн от простого полного каталога фигур поведения рынка до продвинутой теории коллективного поведения толпы — предмета, который ранее находился вне поля зрения как экономики, так и социологии.

С 1935 г. Коллинз высылал Эллиотту книги, в которых обсуждалась природа последовательности Фибоначчи, — математической основы исследования растущих структур, на протяжении столетий восхищавшая исследователей своей гармоничностью и красотой. Коллинз, по видимому, был первым человеком, который заметил, что суммарные числа волн, которыми Эллиотт описывал последовательность цен на падающем тренде фондового рынка, порождают последовательности Фибоначчи. Собственные наблюдения Эллиотта после публикации его книги послужили для него еще одним дополнительным стимулом выйти за узкие рамки фондового рынка и посмотреть на свое открытие с более общих позиций.

В начале 40-х годов Эллиотт полностью развил свою новую концепцию, согласно которой приливы и отливы человеческих эмоций и активности следуют естественным процессам, управляемым природными законами. Кульминацией этого исследования явился трактат, по важности равный его оригинальной книге. 1 октября 1940 г. Эллиотт публикует свое первое обсуждение работы Леонардо Фибоначчи «Суммирование рядов динамической симметрии» в «Образовательном Бюллетене» под заголовком: «Основания Принципа Волн» и подзаголовком «Как работает Принцип Волн и как он связан с математическими законами». В этой работе он обсуждает формы коллективного поведения людей, опираясь на исследования Фибоначчи и на тесно связанный с ними принцип «золотого сечения», — математического феномена, известного на протяжении тысячелетий математикам, ученым, художникам, архитекторам и философам как одного из универсальных законов форм и креативной изменчивости.

Эта основополагающая работа повлекла за собой инцидент, который можно назвать не иначе как интеллектуальное пиратство. 19 мая 1941 г., меньше чем через восемь месяцев после трактата Эллиотта, небольшому числу его подписчиков была разослана издательством «Баррона» статья некоего «Эдсона Бирса». («Эдсон Бирс» — псевдоним Эдсона Голда, чье второе имя было Бирс). Статья называлась «Новая идея для трейдеров: Приложение принципов «динамической симметрии» к фондовому рын-

ку». В статье утверждалось, что предложена, якобы, «бесспорно новая идея» и при этом Эллиотт абсолютно не упоминался.

Голд позднее подтвердил, что редактор «Баррона» сообщил ему, что звонил Эллиотт, который был очень разгневан в связи с этой статьей, но от каких-либо дальнейших комментариев по этому вопросу отказался. Вся эта история, тем не менее, разрешилась благополучно, и в результате длительных переговоров приоритет Эллиотта был полностью подтвержден.

30 декабря 1941 г. супруга Эллиотта Мери Элизабет, которая была с ним во всех его скитаниях, взлетах и падениях его карьеры и длительных перипетиях со здоровьем, умерла в возрасте 72 лет. Они прожили вместе 38 лет.

В августе 1945 г. Эллиотт пишет свою последнюю Пояснительную Записку и в течение оставшихся месяцев 1945 г. и первой половине 1946 г. сосредотачивается на своей наиболее полной работе «Закон природы — главная тайна Вселенной». Такая в известной мере излишняя амбициозность названия имеет некоторые оправдания. Эта монография, которую Эллиотт публикует в возрасте 75 лет, включает почти все его мысли, касающиеся теории Принципа Волн. Книга была опубликована 10 июня 1946 г. тиражом 1000 экз. и быстро раскуплена членами Нью-Йоркского финансового сообщества.

Июль–декабрь 1946 г. Эллиотт руководит по крайней мере двумя полосами в периодических изданиях, но хроническая анемия опять настигает его. Проблемы со здоровьем снова сильно обостряются. В 1947 г. его друзья из Уол-Стрит уговаривают его лечь в Методистский Госпиталь Бруклина для обследования. Затем он переезжает в Государственный Госпиталь «Kings Park» — один из ведущих психиатрических госпиталей Нью-Йорка, который в то время выполнял также функции дома престарелых. Здесь Эллиотт имел все необходимое, чтобы прожить последние месяцы своей жизни. Эллиотт и Коллинз продолжали встречаться и проводить время со своими друзьями и корреспондентами. Согласно медицинским отчетам, Эллиотт сохранял ясное сознание и острый ум до своего последнего дня.

Эллиотт умер 15 января 1948 г. Смерть наступила от миокардита — хронического воспаления сердечной мышцы. Начиная с того момента, как он стал жертвой амебной инфекции, из-за которой он был вынужден прервать свою предыдущую карьеру, Эллиотту приходилось постоянно бороться со своей болезнью на протяжении всего времени его пребывания в Центральной Америке. В этом регионе хронический миокардит очень часто является следствием паразитных инфекций и называется болезнью Чагаса. Болезнь может привести к смерти спустя много лет после инфицирования.

Один из бывших трейдеров сообщил, что друзья Эллиотта пожертвовали деньги (что обычно делается, если у покойного нет прямых родственников) для его кремации, которая состоялась через два дня в Fresh Pond крематории Middle Village, N.Y.

Непосредственно после смерти Эллиотта сравнительно небольшое число инвесторов применяло его методы. Однако сегодня тысячи фирменных портфель-менеджеров, трейдеров и частных инвесторов используют Принцип Волн при принятии своих инвестиционных решений. Ральф Эллиотт, безусловно, был бы рад видеть все это.

На самом деле, наследие Эллиотта только начинает заявлять о себе. Его открытие не исчерпывается только динамикой рынка. Оно описывает динамику изменчивости социума в целом. Со временем будет понято, что вклад Эллиотта в социологию является прорывом, подобным тому, который имел место в физике в 1600–1700 гг. Вовремя принятый во внимание, Принцип Волн, безусловно, сможет вывести социологию из области чисто качественных рассуждений и обеспечить ей достойное место среди точных наук. И, возможно, как полагал сам Эллиотт, дело ограничится далеко не только социологией...

Принцип Волн Эллиотта и некоторые вопросы изменчивости сложных систем

Для нас основной интерес представляет последняя книга Эллиотта — «Закон природы — главная тайна Вселенной». Именно в процессе работы над этой книгой Эллиотт вплотную подошел к пониманию неких достаточно общих принципов, характеризующих изменчивость в сложных системах самой разной природы. Он понял, что эти принципы проявляются на самых разных пространственных масштабах (масштабно инвариантны), а серии последовательных изменений подобны друг другу (самоподобны). Самоподобие и масштабная инвариантность, как теперь известно, являются отличительными особенностями фрактальных структур. Таким образом, открытие Эллиоттом этих принципов позволяет считать его одним из основоположников фрактального анализа временных рядов. Но его открытие этим не ограничивается...

Особенностью является также и то, что Эллиотт обосновал эти свои принципы не на абстрактных геометрических объектах, а на наблюдательном материале, относящемся к *изменчивости сложных систем*. Осознав тот факт, что последовательности изменений не зависят от природы систем, Эллиотт приходит к выводу, что открытый им набор принципов или правил является достаточно универсальным, стоит как бы за или над пределами конкретных систем и, таким образом, относится

к числу фундаментальных законов природы. Он понимает, что эти принципы есть то, что управляет изменчивостью во Вселенной. Но изменчивость тесно связана с природой времени — величайшей тайной Вселенной. Таким образом, Эллиотт почувствовал, что, открыв *закон природы*, он прикоснулся к *главной тайне Вселенной*. Таково, по-видимому, происхождение названия книги.

Кратко остановимся на ее содержании. Во введении Эллиотт говорит о разнообразии природных ритмов и отмечает некие их общие закономерности. В первой главе, которая называется «Великая пирамида в Гизе», Эллиотт делает экскурс в историю золотого сечения и обсуждает роль чисел Фибоначчи в структурах растительного и животного мира. Он отмечает, что такие природные структуры как, например, раковины являются «записью» серий последовательных изменений организма моллюска, и, таким образом, пропорции их структур отражают пропорции ритмов изменчивости. Во второй главе Эллиотт формулирует свод правил, которым следует изменчивость сложных систем. В третьей главе «Человеческая активность» и четвертой «Отличительные особенности человеческой активности» Эллиотт показывает, что поведение «толпы», будь то римские рабы во время восстаний или участники биржевой игры, подчиняется характерным законам. По его мнению, психологическое поведение людей в процессе социально-массовых движений последовательно проходит стадии экспансии, энтузиазма, эйфории, успокоения, упадка, депрессии. Причем такая схема психологического поведения прослеживается на самых разных периодах — начиная от нескольких минут и заканчивая веками.

В пятой главе «Коррекции», шестой «Экспансии», седьмой «Неустойчивые вершины», восьмой «Чередования», девятой «Масштабы» подробно обсуждаются особенности всех фаз волнового цикла изменчивости и особенности переходов от одной фазы к другой. Несмотря на то, что обсуждения проводятся на хорошо знакомом Эллиотту материале — биржевых рядах — они, тем не менее, носят достаточно универсальный характер. Эллиотт подробно описывает особенности и характерные признаки изменчивости на разных фазах. Эти описания позволяют «узнать» волну и, таким образом, позволяют определить, на какой фазе волнового цикла находится система в ее текущем состоянии и что должно последовать за этим. В десятой и последующих главах приводятся многочисленные примеры и рассматривается применение общих принципов к самым разным конкретным ситуациям: динамика цен на золото, изменения железнодорожного Доу, Великая депрессия, эмоциональные циклы человека и др. В последних главах снова обсуждаются некоторые аспекты общих принципов, затем проводится обзор всех основных положений и делаются

выводы. Таково содержание этой очень разноплановой и довольно пестрой книги.

Рассмотрим теперь некоторые концептуальные положения Принципа Волн Эллиотта, которые могут представлять интерес в плане исследований явления изменчивости в сложных системах. Изменчивость сложных систем, утверждает Эллиотт, реализуется в форме последовательных импульсов или «волн», которые возникают не в произвольном порядке, а выстраиваются в определенную конфигурацию или волновую диаграмму. Таково важнейшее понятие Принципа Волн Эллиотта. Анализ соотношений между соседними волнами позволяет определить, в каком месте волновой диаграммы находится система и какие изменения можно ожидать в ней в дальнейшем.

В упрощенном варианте волновая диаграмма для фондовых рынков состоит из пяти волн роста и трех волн падения, что составляет полный цикл. Волны роста (Эллиотт присваивает им номера от 1 до 5) называются *импульсными* волнами, нисходящие (их Эллиотт обозначает буквами a, b, c) называются *корректирующими* волнами. Такая модель сочетания волн, утверждает Эллиотт, является устойчивой, и ее можно применять для исследования изменчивости самых разных систем на различных масштабах.

Другим важным понятием Принципа Волн является понятие *степени или масштаба* соответствующей *тенденции*. Для фондовых рынков Эллиотт определял девять различных масштабов развития тенденции, начиная с «Великого сверхцикла» протяженностью около 200 лет и заканчивая сверхкоротким масштабом, составляющим всего несколько часов. Важнейшее правило при этом гласит: независимо от масштаба, тенденция всегда будет развиваться по основному восьмиволновому циклу. При этом каждая такая волна разбивается на меньшие волны, которые в свою очередь разбиваются на еще меньшие и так далее, то есть имеет место *масштабная инвариантность*. Таким образом, каждая из волн является частью волны следующей иерархии. При этом соотношения между числом волн, входящих в разные уровни иерархии, как легко убедиться, являются числами из *числовой последовательности Фибоначчи*. Эта числовая последовательность играет важную роль в Принципе Волн, являясь по существу ее количественным оформлением. Поэтому прервем на какое-то время изложение интересующих нас аспектов Принципа Волн и напомним читателям некоторые моменты, относящиеся к этой числовой последовательности.

Леонард Пизанский (Фибоначчи, Fibonacci (ок. 1170–1250)) — один из величайших математиков Средневековья. Ему мы обязаны, в частности, тем, что пользуемся позиционной десятичной системой исчисления. Известна его «Книга абака» (арифметика), в которой Фибоначчи опи-

сывает арабско-индийскую систему исчисления и ее преимущества по сравнению с использовавшимися тогда римскими цифрами.

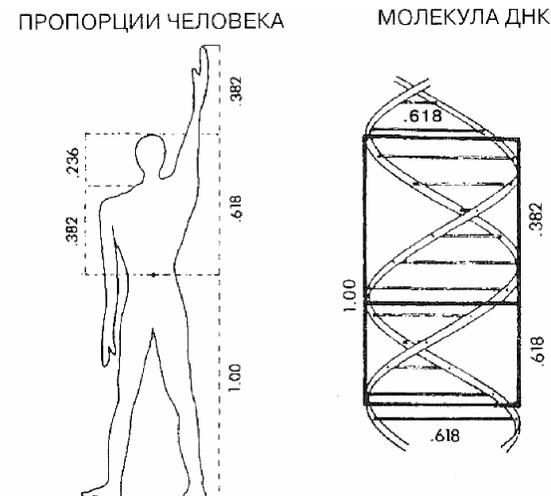
Свою знаменитую числовую последовательность Фибоначчи открыл, наблюдая рост потомства у семьи кроликов. Задача выглядела так: «Сколько кроликов можно получить от одной пары, если каждая пара, начиная с двухмесячного возраста, производит новую каждый месяц». Решением задачи является последовательность, которая задается начальными значениями $a_1 = a_2 = 1$ и рекуррентным соотношением $a_{n+1} = a_n + a_{n-1}$. Вот первые числа этой последовательности: 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987...

Эта последовательность чисел обладает многими замечательными свойствами. Например, отношение каждого числа к следующему за ним при увеличении номера числа стремится к 0.618, отношение последующего к предыдущему, соответственно, — к 1.618. Число 0.618 — это известный постоянный коэффициент в *золотом сечении*. Его называют еще золотым коэффициентом или золотой серединой. Если взять отношение чисел этой последовательности через одно, то в пределе получим два новых коэффициента — 0.382 и 2.618. Действуя в том же духе, получаем 0.236 и 4.235 и так далее...

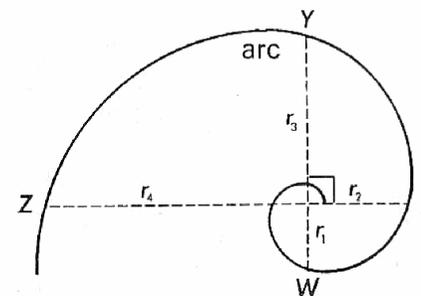
Самое интересное, что эти коэффициенты «используются» природой для построения ее частей на самых разных масштабах (рис. 3). Некоторые современные космологические модели полагают, что масштаб Вселенной изменяется по *золотой спирали*, в которой отношение радиусов через каждую четверть оборота равно 1.618. При этом все ее витки имеют одну и ту же форму независимо от масштаба. Это та самая спираль, которую мы видим на раковинах моллюсков, в водоворотах, в молодых ростках папоротника, в формах галактик и многом другом. В строении молекул ДНК, в пропорциях тела человека, в цветках и плодах растений и во множестве других природных структур мы находим все те же числа.

Числа из последовательности Фибоначчи буквально «пропитывают» весь Принцип Волн Эллиотта. Мы не можем здесь изложить Принцип Волн в полном объеме (для этого требуется специальная публикация) и отсылаем заинтересованных читателей как к упомянутым оригинальным трудам Эллиотта, так и к работам его последователей¹². Приведем в качестве примера лишь некоторые соотношения.

¹² См., например книгу, Frost A. and Prechter R. Elliott Wave Principle: Key to Stock Market Profits, 1990; Balan R. Elliott Wave Principle Applied to the Foreign Exchange Markets, 1989.



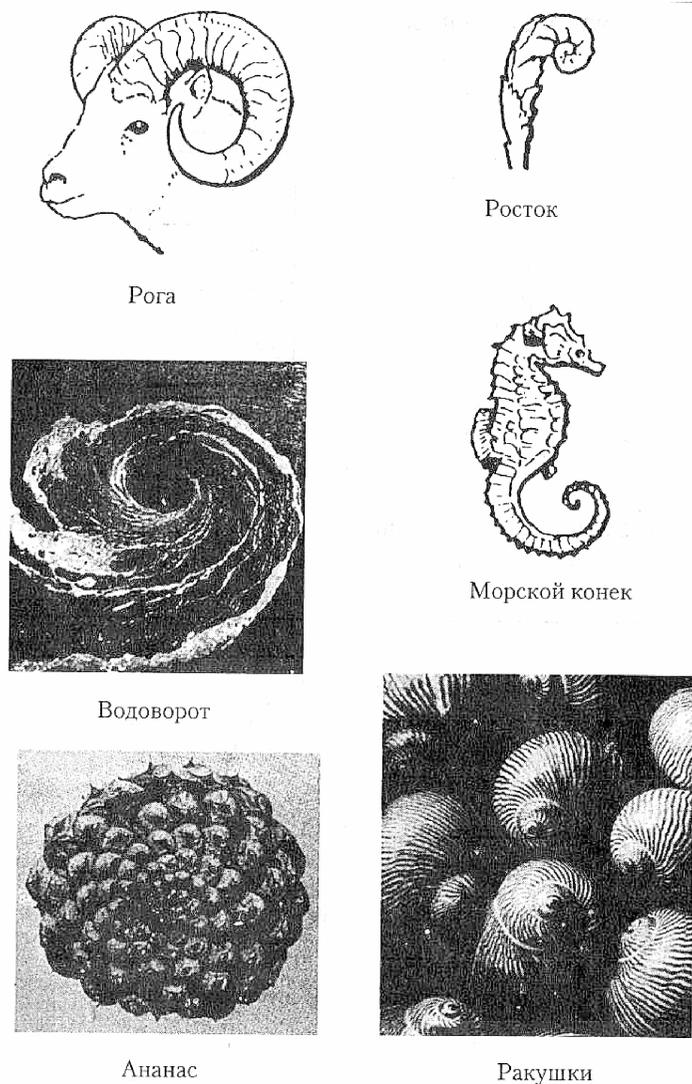
Коэффициенты Фибоначчи в строении человека и ДНК
Из кн.: А. Frost, R. Prechter "Elliott Wave Principle"



Золотая спираль
Некоторые из соблюдающихся соотношений:

$$\frac{r_2}{r_1} = \frac{r_3}{r_2} = \frac{r_4}{r_3} = \dots = \frac{r_n}{r_{n-1}} = 1.618$$

$$\frac{r_2 + r_4}{r_1 + r_3} = \dots = \frac{r_{n+1} + r_{n+3}}{r_n + r_{n+2}} = 1.618$$



Золотая Спираль в природе

Из кн.: А. Frost, R. Prechter "Elliott Wave Principle"

Рис. 3. Коэффициенты Фибоначчи и Золотая спираль в природе

Так, базовая волновая диаграмма состоит из 8 волн — 3 корректирующих и 5 импульсных. Все эти три числа — из последовательности Фибоначчи. Разбиение волн диаграммы на более мелкие дает 34 и 144 волны — числа из той же последовательности. Между волнами одной волновой диаграммы и между волнами, относящимися к разным масштабам, возникают пропорциональные отношения, соответствующие коэффициентам Фибоначчи. Так, координаты вершины волны 3 можно получить, умножив длину волны 1 на 1.618 и прибавив этот результат к основанию волны 2. Верхняя точка волны 5 может быть установлена путем умножения волны 1 на $2 \times 1.618 = 3.236$. Результат следует прибавить к вершине и основанию волны 1. Вершина волны 5 обязательно попадет в этот интервал. Таких соотношений открыто множество.

Безусловно, у Эллиотта были предшественники. Так, основоположник технического анализа Чарльз Доу ввел понятия и определения основных тенденций рынка и отмечал их «волнообразный» характер. Сам Эллиотт вначале писал, что его Принцип Волн является лишь полезным дополнением и уточнением теории Доу. Однако в процессе последующего развития Принцип Волн стал наиболее исчерпывающим техническим инструментом анализа фьючерсных и товарных рынков. После Эллиотта никто не внес в него каких-либо существенных дополнений. Усилия пользователей были в основном направлены на формирование более четкого понимания всех особенностей этой теории. Стало понятно, что разные положения Принципа Волн обладают не одинаковой универсальностью. Тем не менее, число последователей Эллиотта растет, а один из главных его современных последователей Роберт Прехтер стал «гуру» 80-х годов в области за прогнозирования движения индекса Доу–Джонса. Что касается более широкого применения Принципа Волн как универсального закона изменчивости сложных систем, о чем писал сам Эллиотт, то эта страница его наследия остается пока нераскрытой и непрочитанной...

На этом месте я вместе с Дж. Дж. Мерфи¹³ «хотел бы сделать еще одно отступление и привести цитату из книги Х. Бестона «Самый дальний дом», имеющую, на мой взгляд, отношение к морской тематике. Наткнулся я на нее летом во время отпуска, который проводил на пляжах Кейп-кода, и сразу понял, что она некоторым образом связана с теориями Доу и Эллиотта. Бестон описывает свое пребывание на Кейп-коде, где он провел целый год, в одиночестве бродя по пляжам. Он много наблюдал за океаном и заметил, что его движение подчиняется тройному ритму. Бестон пишет, что огромные океанские волны подходят к берегу группами по три: «Три огромные волны, затем короткая промежуточная серия

¹³ Мерфи Джон Дж. Технический анализ фьючерсных рынков. — М.: 1996 г. С. 407.

из более слабых, затем снова три большие». Словно в подтверждение того, что трех волновой ритм океана не плод его собственного воображения, он рассказывает об обычае служащих местной береговой охраны, которые во время сильного волнения стараются отойти от берега как раз в промежутке между сериями из трех больших волн. Как знать, может, на самом деле Доу с Эллиоттом знали о море больше, чем явствует из их книг, посвященных анализу рынка».

Заключение

Всякая теория имеет смысл и представляет ценность лишь тогда, когда она не только объясняет ту или иную группу фактов (часто это можно сделать десятками способов), но и предсказывает новые, ранее неизвестные. Если же идея претендует на открытие законов изменчивости сложных систем, то ее адекватность проще всего проверить на том, что она может сказать о будущем развитии тех или иных событий в этих системах. А это означает, что любые конструктивные соображения по анализу временных рядов, продуцируемых сложными системами, могут быть конвертированы, в частности, в некую стратегию биржевой игры и подвержены вышеописанному тесту Коллинза.

Благодаря Интернету временные ряды движения цены на большинстве рынков ценных бумаг, а также на товарных и валютных рынках сейчас стали доступны практически каждому. На этом уникальном материале можно проверить идеи, отработать свои методы и алгоритмы анализа временных рядов. Каждый человек может, внося относительно небольшую сумму, заключить фьючерсный контракт, опционный контракт или сформировать более сложный инвестиционный портфель и устроить практический экзамен своим алгоритмам. Таким образом, в наших руках уникальная лаборатория, в которой мы можем проверить практически любые наши соображения о поведении природных временных рядов. На этих рынках состязаются острейшие умы планеты. Именно от них каждый может получить самую квалифицированную «рецензию» на предмет адекватности своих представлений о развитии и течении процессов в сложных системах, причем сразу и в твердой валюте.

Ральф Нельсон Эллиотт создал свою теорию, благодаря которой он и его последователи неоднократно предсказывали будущее развитие событий. Теория Волн Эллиотта получила множество положительных «рецензий» в виде многих и многих миллионов долларов. Согласитесь, такие рецензии дают некоторое основание для того, чтобы относиться к идеям Эллиотта с определенным интересом...

Чумак О. В.

ЭНТРОПИИ И ФРАКТАЛЫ В АНАЛИЗЕ ДАННЫХ

*Дизайнер В. А. Толстолуцкая
Технический редактор А. В. Ширококов
Компьютерная верстка О. А. Печина
Корректор Е. В. Огородникова*

Подписано в печать 25.05.2011. Формат 60x84 ¹/₁₆.
Гарнитура «Times New Roman». Печать офсетная.
Усл. печ. л. 9,5. Уч.-изд. л. 10,2. Заказ № 11-20.

Научно-издательский центр «Регулярная и хаотическая динамика»
426034, г. Ижевск, ул. Университетская, 1
<http://shop.rcd.ru> E-mail: mail@rcd.ru

Уважаемые читатели!

Интересующие Вас книги нашего издательства можно заказать через наш Интернет-магазин <http://shop.rcd.ru> или по электронной почте subscribe@rcd.ru

Книги можно приобрести в наших представительствах:

МОСКВА

Институт машиноведения им. А. А. Благонравова РАН
ул. Бардина, д. 4, корп. 3, к. 414, тел.: (499) 135-54-37, (495) 641-69-38

ИЖЕВСК

Удмуртский государственный университет
ул. Университетская, д. 1, корп. 4, 2 эт., к. 211, тел./факс: (3412) 50-02-95

Также книги можно приобрести:

МОСКВА

Московский государственный университет им. М.В. Ломоносова
ГЗ (1 эт.), Физический ф-т (1 эт.), Гуманитарный ф-т (0 и 1 эт.),
Биологический ф-т (1 эт.).

Российский государственный университет нефти и газа
им. И. М. Губкина ГЗ (3-4 эт.), книжные киоски фирмы «Аргумент».

Магазины:

МОСКВА:	«Дом научно-технической книги» Ленинский пр., 40, тел.: 137-06-33 «Московский дом книги» ул. Новый Арбат, 8, тел.: 290-45-07
ДОЛГОПРУДНЫЙ:	Книжный магазин «Физматкнига», новый корп. МФТИ, 1 эт., тел.: 409-93-28
САНКТ-ПЕТЕРБУРГ:	«Санкт-Петербургский дом книги» Невский проспект, 28 Издательство СПбГУ, Магазин №1 Университетская набережная, 7/9