МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

МИНИСТЕРСТВО ПРОМЫШЛЕННОСТИ, НАУКИ И ТЕХНОЛОГИЙ
РОССИЙСКОЙ ФЕДЕРАЦИИ

РОССИЙСКАЯ АССОЦИАЦИЯ НЕЙРОИНФОРМАТИКИ

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ
(ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ)

НАУЧНАЯ СЕССИЯ МИФИ–2002

НЕЙРОИНФОРМАТИКА–2002

**IV ВСЕРОССИЙСКАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ
КОНФЕРЕНЦИЯ**

ЛЕКЦИИ ПО НЕЙРОИНФОРМАТИКЕ

Часть 2

По материалам Школы-семинара
«Современные проблемы нейроинформатики»

Москва 2002

УДК 004.032.26 (06)
ББК 32.818я5
М82

НАУЧНАЯ СЕССИЯ МИФИ–2002. IV ВСЕРОССИЙСКАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ «НЕЙРОИНФОРМАТИКА–2002»: ЛЕКЦИИ ПО НЕЙРОИНФОРМАТИКЕ. Часть 2. – М.: МИФИ, 2002. – 172 с.

В книге публикуются тексты лекций, прочитанных на Школе-семинаре «Современные проблемы нейроинформатики», проходившей 23–25 января 2002 года в МИФИ в рамках IV Всероссийской конференции «Нейроинформатика–2002».

Материалы лекций связаны с рядом проблем, актуальных для современного этапа развития нейроинформатики, включая ее взаимодействие с другими научно-техническими областями.

Ответственный редактор
Ю. В. Тюменцев, кандидат технических наук

ISBN 5–7262–0400–X

 ISBN 5–7262–0400–X

Содержание

Н. Г. МАКАРЕНКО
Институт математики, Алма-Ата, Казахстан
E-mail: makarenko@math.kz

ФРАКТАЛЫ, АТТРАКТОРЫ, НЕЙРОННЫЕ СЕТИ И ВСЕ ТАКОЕ

Аннотация

Известно, что *единое лучше, чем всё вместе, но врозь*. Лекции представляет собой попытку продемонстрировать этот тезис на примере интригующих связей между теорией фракталов, системами гиперболических итеративных функций, дискретными динамическими системами и нейронными сетями. Изложение рассчитано на широкий круг слушателей, которые не являются математиками.

N. MAKARENKO
Institute of Mathematics, Kazakhstan, Alma-Ata
E-mail: makarenko@math.kz

FRACTALS, ATTRACTORS, NEURAL NETWORKS AND ALL THAT

Abstract

It is known that *a single whole is better than all individual things together*. The lecture is to be an attempt to demonstrate this thesis on the example of intriguing connections between fractal theory, hyperbolic iterated function systems, discrete dynamical systems and neural networks. The exposition is counted on wide circle of listeners, who are not mathematicians.

Введение

Предлагаемый фрагмент из еще не завершенного романа публикуется досрочно в качестве раздумья об исподней сущности переживаемого момента, чем и как может он обернуться нам под воздействием нашей неосторожности

Леонид Леонов
«Мироздание по Дымкову»

«Все фигуры, которые я исследовал и назвал фракталами, в моем представлении обладали свойством быть нерегулярными, но самоподобными», — писал Бенуа Мандельброт, который в 1975 году ввел термин *фрактал* (от латинского *fractus* — дробный). Позднее оказалось, что фракталами являются и давно известные в анализе нерегулярные функции, вызывавшие отвращение аналитиков прошлого века [1] [2]. Процесс построения классических фракталов, таких как множество Кантора или ковер Серпинского, который можно принять за их определение, предельно прост. Сперва следует выбрать основную процедуру-генератор, а затем итеративно, до бесконечности применять ее к произвольному компакту. То, что получится в пределе — это и есть фрактал. Идея о том, что объектами геометрии могут быть предельные образы итеративной динамики, не вызывала особого восторга. Можно представить себе (и даже изобразить) на вещественной оси предел числовой последовательности. Однако, мускулы нашего тренированного воображения отказывают, когда пределом является компактный глобально несвязный объект, зачастую обладающий инфернальными свойствами: во многих случаях у него нет даже тени! Да и сами процедуры генерации фрактала напоминали скорее хирургию с элементами садизма, нежели приемы аналитики. Фракталы были чужды уютному евклидову миру с его регулярными структурами.

[1] Если ссылка на примечание представляет собой число, заключенное в круглые скобки, например, (3), то она обозначает номер примечания, помещенного в конце данного текущего раздела. Ссылка в виде числа, не заключенного в такие скобки — обычное подстраничное примечание. — Прим. ред.

[2] «С омерзением и ужасом я отворачиваюсь от этой зловредной язвы — непрерывных функций, нигде не имеющих производных», — писал Эрмит Стилтьесу в 1893 году.

Ситуация изменилась в 1981 году, когда Хатчинсон поместил пришельцев в их родную среду — пространство компактов. Именно здесь *Система Итеративных Функций* (СИФ) — сжимающих отображений, порождает фракталы простым и естественным путем, при котором ампутация отдельных фрагментов связных множеств трансформируется в корректную форму непрерывного сжатия в подходящей метрике. Реминисценции, навеянные предельной динамикой СИФ, ведут к теории *дискретных динамических систем*. Аналогами аффинных СИФ являются здесь нелинейные преобразования, порожденные, например, сечениями Пуанкаре многомерных фазовых потоков. Нелинейность приводит к существенной зависимости от начальных данных так, что малые погрешности экспоненциально растут в фазовом потоке и, начиная с некоторого момента времени, будущее состояние системы становится лишь ограниченно предсказуемым. Этот процесс чаще всего происходит в диссипативных системах, траектории которых заполняют низкоразмерное инвариантное притягивающее подмножество — *аттрактор* в фазовом пространстве. На аттракторе траектории разбегаются в неустойчивых направлениях и сжимаются в устойчивых. Вследствие диссипации сжатие преобладает и в устойчивых направлениях аттрактор копирует сам себя: сечение фазового потока приобретает самоподобную структуру канторова множества с дробной размерностью. Такой аттрактор называют *странным* или фрактальным.

Самоподобная структура фрактала позволяет восстановить СИФ по фрагменту, по крайней мере, в приближении конечного числа итераций. В 1991 году Бресслофф и Штарк показали, что процесс аппроксимации аттрактора СИФ эквивалентен работе бинарной нейронной сети. Таким образом, термины «фрактал» в геометрии и «странный аттрактор» в динамике оказываются синонимами, а СИФ можно рассматривать как рекуррентную асимметричную нейросеть. С другой стороны, Фернандо Ниньо в 2000 году установил, что случайная итеративная нейронная сеть (гипернейрон) топологически эквивалентна динамической системе с заданным аттрактором. Круг замкнулся, образовав Единый Контекст, объединяющий **фракталы**, **СИФ**, **аттракторы** и **нейронные сети**. Цель лекции — показать взаимную связь этих предметов, потому что *единое лучше, чем всё вместе, но по-отдельности*. Более серьезные мотивы указаны в эпиграфе ко введению.

Основой предлагаемой Лекции послужили *Заметки о фракталах*[3] и доклады автора на Фрактальном семинаре Института математики. В тексте опущены все доказательства, и это дает читателю восхитительную возможность *видеть, но не верить!* Поскольку предварительных знаний не всегда хватает как раз для понимания предварительных сведений — они помещены в конце текста, в Глоссарии. Он не обязателен для понимания основного текста: если вас начинает смущать терминология, следуйте простому правилу: *никогда не жуйте пилюлю, которую вас заставляют проглотить*. В конце каждого раздела помещен путеводитель по Литературе, список которой совсем не претендует на полноту.

Автор искренне благодарен своим молодым коллегам: Светлане Ким, Кате Данилкиной и Ерболу Куандыкову, которые взяли на себя неблагодарный труд по набору и редактированию Лекции.

Размерности, площади и объемы

> Необходимо число, различитель *инаковости*, без которого невозможно отличить одно от другого
>
> *Николай Кузанский*
> *«Игра в шар»*

Мы начнем с идей итальянского математика *Никколо Фонтана Тарталья* (1500–1557 гг.). Известно, что на три неколлинеарные точки плоскости можно натянуть треугольник, а четырем некомпланарным точкам в $\mathbb{R}^3$ соответствует тетраэдр. Это выпуклые геометрические тела, для которых существует мера — площади и объемы. Согласно Тарталья, *пространство является n-мерным, если для его $n+1$ точек, не принадлежащих одной гиперплоскости, существует полиэдр ненулевого объема.* Наиболее существенным здесь является то, что в основу определения размерности было положено понятие меры: именно этот момент служит основой современных конструкций. В этом определении неявно используется предположение, что необходимые $n+1$ точки *могут быть всегда выбраны так,* чтобы они не принадлежали одной гиперплоскости. Однако принцип общего положения, который столь широко используется в

[3] URL: http://www.keldysh.ru/dpt_17/works/mak/index.htm

современной математике, требует по меньшей мере понятий непрерывного отображения и многообразия. Поэтому последующие определения размерности рассматривались в именно в таком контексте.

До появления стандартной терминологии произвольное гладкое многообразие называли просто *непрерывностью*. Понятие размерности вводилось аналитически как минимальное число параметров, необходимых для идентификации точки в *непрерывности*. На привычном физическом языке — это n координат точки в $\mathbb{R}^n$. Однако почти сразу же оказалось, что эта **параметрическая размерность** неудовлетворительна по нескольким причинам. Прежде всего выяснилось, что такое определение не различает прямую и плоскость, поскольку между ними можно установить однозначное соответствие. Впервые это показал *Георг Кантор*, используя следующие соображения. Пусть точка $x = (x_1, x_2) \in I = [0,1] \times [0,1]$. Ее координаты $0 \leqslant x_1, x_2 \leqslant 1$ можно представить в виде бинарных дробей: $x_1 = 0, \alpha_1\alpha_2 \ldots; x_2 = 0, \beta_1\beta_2 \ldots$, где α_i, β_i — нули либо единицы. Для обеспечения единственности условимся записывать обрывающиеся разложения так, чтобы все двоичные цифры, начиная с некоторого места были тогда нулями. Сопоставим паре разложений x_1 и x_2 последовательность $z = 0, \alpha_1\beta_1\alpha_2\beta_2 \ldots$, наследующую *лексикографический* порядок исходных оригиналов. Мы получили отображение $z : I \to [0,1]$. Следовательно, существует способ параметризации точек плоскости только одной координатой — извилистой линией! Разумеется, это возможно лишь в том случае, если мы не связываем себя условием непрерывности так, чтобы двум бесконечно близким точкам прямой, соответствовали две бесконечно близкие точки плоскости. *Анри Пуанкаре* заметил, что отказ от непрерывности трудно примирить с интуитивными логическими принципами (1). То, что было предложено им взамен, основано на весьма остроумной концепции **физической непрерывности**. Здесь аналогом точки является элемент, неотличимый от соседних, но совокупность таких элементов образует непрерывную цепь, оба конца которой легко различимы. Размерность непрерывности вводится с помощью понятия *купюры*. Последняя образуется совокупностью элементов, изъятых из непрерывности. Если в результате такой операции непрерывность разделяется на две части, а сама купюра состоит лишь из конечного числа элементов, не образующих связного множества, то размерность непрерывности равна единице. Например, линия одномерна, потому что делится на две части нуль-мерной купюрой — точкой. Для того, чтобы разделить плоскость, необходима одномерная купюра — линия и т. д. Таким образом, *математическая непрерывность имеет n измерений, если ее можно разбить на части, произведя в ней одно или несколько сечений, которые сами являются непрерывностями $n-1$ измерения*. Это рекуррентное определение размерности, которое предполагает, что объемы — части пространства, поверхности — границы объемов, линии — границы поверхностей, а точки — границы линий.

Формализация этих идей привели *Брауэра, Урысона* и *Менгера* к индуктивному определению **топологической размерности**. Размерность любого конечного или счетного множества точек есть $d_t = 0$. Размерность любого связного множества[4] точек есть $d_t + 1$, если его можно разрезать на два несвязных куска, исключив как минимум d_t-мерное множество точек, т. е. сделав d_t-мерный разрез$^{(2)}$. Ясно, что топологическая размерность всегда есть целое число.

Следующее понятие размерности опирается на идею покрытия и практические способы вычисления длин и площадей, предложенные еще *Борхардом* и *Минковским*. Рассмотрим плоскую кривую $\mathcal{C}$, имеющую в каждой точке непрерывную касательную. Начнем перемещать отрезок длины 2ε так, чтобы его середина двигалась вдоль кривой, а сам он оставался нормалью: касательная определена, следовательно определена и нормаль. Отрезок заметает площадь $S(\varepsilon)$ (мы считаем, конечно, что площадь существует). То, что получается при этом, называется *шарфом Минковского*. При довольно широких допущениях можно показать, что существует $\lim S(\varepsilon)/2\varepsilon$, при $\varepsilon \to 0$, который мы и назовем длиной $\mathcal{C}$. Аналогичным образом можно определить площадь как предел: $\lim V(\varepsilon)/2\varepsilon$, $\varepsilon \to 0$, где $V(\varepsilon)$ — объем тела (*сосиски Минковского*), заметаемого отрезками 2ε, нормальными к поверхности.

Таким образом, меру геометрического объекта некоторой размерности можно определить как скорость изменения меры другого объекта, большей размерности, покрывающего исходный, при *линейном* убывании некоторого характерного масштаба. Заметим, что это определение не зависит от выбора координат.

[4]Топологическое пространство называют *связным*, если его нельзя представить в виде объединения двух непустых множеств.

Для определения размерности используется следующее обобщение. Регулярный многомерный объект можно представить в форме прямого произведения объектов низшей размерности. Тогда его мера (объем) выражается в виде степени характерного размера элемента покрытия — $\varepsilon^{\mathrm{dim}}$. Для сохранения аддитивности меры удобнее перейти от покрытия к разбиению. Степенная зависимость меры от ε называется **скейлингом**. Тогда размерность (dim) можно определить как скорость изменения меры при измельчении разбиения — т. е. уменьшения ε по *степенному* закону. Поскольку при этом нас интересует только функциональная связь, вместо суммарного объема элементов разбиения (покрытия) часто используют *число* таких элементов. Покрытия (разбиения) можно реализовать многими способами. Поэтому для получения корректных численных оценок берется *минимальное* число элементов покрытия или разбиения. В последнем случае подсчитываются только непустые кубы или шары.

Все эти соображения лежат в основе размерности по *Колмогорову*, которую чаще называют **емкостью**. Пусть (M, ρ) полное метрическое пространство с метрикой ρ и $\dim M = d$. Рассмотрим непустое компактное подмножество $\mathfrak{F} \subset M$. Для $\varepsilon > 0$ пусть $\mathcal{B}(x, \varepsilon)$ — замкнутый d-мерный шар с центром в точке $x \in M$.

Подсчитаем наименьшее число $N(\varepsilon)$ таких шаров, необходимых для покрытия $\mathfrak{F}$: $N(\varepsilon) =$ *наименьшему целому* m *такому, что* $\mathfrak{F} \subset \cup^m \mathcal{B}(x_n, \varepsilon)$ для $\{x_n \,|\, n = 1, 2, \ldots, m\} \subset M$. Такое число всегда существует, потому что $\mathfrak{F}$ — компакт, т. е. ограниченное и замкнутое множество. Следовательно, из каждого его бесконечного покрытия можно выбрать конечное подпокрытие. Тогда искомое число есть минимальная длина всех таких конечных последовательностей.

Если существует *скейлинг* $N(\varepsilon) \sim \varepsilon^{-d_c}$ при $\varepsilon \to 0$, то d_c — **колмогоровская емкость** множества. Заметим, что $d_c \leqslant d$ и не обязательно целое число! Это на первый взгляд кажется парадоксальным, потому что мы использовали d-мерные шары, где d — *целое* число. Однако, d_c определяется из скейлинга, который выполняется *асимптотически*, в пределе исчезающих ε. Для корректного определения d_c следует учесть еще нормировку, равную объему d-мерной единичной сферы $V(1)$. Тогда емкость можно определить как такое число d_c, при котором существует отличный от нуля предел(3): $\lim\{V(1) N(\varepsilon) \varepsilon^d\}$ при $\varepsilon \to 0$. При использовании кубов вместо шаров необходимость в нормировке исчезает.

Размерность Хаусдорфа d_H отличается от емкости тем, что следует брать разные шары, радиус которых $r_i \leqslant \varepsilon$. Тогда *крупнозернистая* α-мерная мера Хаусдорфа для любого $\alpha > 0$ определяется как

$$m_\alpha(\mathfrak{F}, \varepsilon) = \inf \sum_i (r_i)^\alpha,$$

где $\inf$ берется по всем покрытиям. Когда ε убывает, сумма возрастает или, во всяком случае, не убывает. Поэтому предел при $\varepsilon \to 0$, конечный или бесконечный, существует и называется *α-мерной мерой Хаусдорфа* для $\mathfrak{F}$: $m_\alpha(\mathfrak{F}) = \lim_{\varepsilon\to 0} m_\alpha(\mathfrak{F}, \varepsilon)$. *Хаусдорфова размерность* $\mathfrak{F}$ характеризует поведение $m_\alpha(\mathfrak{F})$ как функцию от α:

$$\dim_H \mathfrak{F} = \sup\{\alpha | m_\alpha(\mathfrak{F}) = \infty\} = \inf\{\alpha | m_\alpha(\mathfrak{F}) = 0\}.$$

Таким образом, $\dim_H$ отделяет значения α, дающие бесконечную меру, от значений α, приводящих к нулевой мере. В качестве примера рассмотрим фрагмент кривой длины L. Ясно, что для его покрытия (разбиения) требуется $N = L/\varepsilon$ шаров. Тогда мера $m = (L/\varepsilon)\varepsilon^\alpha = L\varepsilon^{\alpha-1}$. Если $\alpha < 1$, то $m \to \infty$ при $\varepsilon \to 0$. Если же $\alpha > 1$, то $m \to 0$ при $\varepsilon \to 0$. Единственным «правильным» значением будет $\alpha = 1$. Заметим, что d_H для некоторых относительно простых множеств может существенно отличаться от d_c. Тем не менее, мы будем использовать оба термина как синонимы, полагая $d_c = d_H$.

Приведенным выше определениям можно придать информационный смысл. Представим себе кассу из двух ящиков. В одном из них находится монета. Очевидно необходим всего один вопрос с возможными ответами «да» или «нет», чтобы идентифицировать ее нахождение. Четыре ящика требуют двух вопросов, а если касса состоит из $2^5 = 32$ ящиков, таких вопросов надо задать уже 5. На первый взгляд кажется, что каждый следующий вопрос следует выбирать только после получения ответа на предыдущий. Однако, венгерский математик *Реньи* показал, что можно задать всего один «сложный» вопрос, допускающий несколько «одновременных» ответов «да», «нет», позволяющих однозначно идентифицировать необходимый номер. Будем считать, что информация, которую содержит один ответ, равна одному биту. Тогда для идентификации системы с N возможными и равновероятными состояниями нам необходима информация в $\mathcal{I} = \log_2 N \equiv \lg N$ бит. Мы пришли к формуле, которую впервые получил *Хартли* в 1928 году.

Вопросы и ответы можно реализовать по-разному. Для нас удобно представить себе вопрос как элемент покрытия (нечто вроде фишки при игре в лото), идентифицирующий «точку» множества с точностью ε. Если $N(\varepsilon)$ — минимальное число таких элементов, $\mathfrak{I} = \log_2 N(\varepsilon)$ — количество информации, необходимое для идентификации всего множества с той же точностью. В этом контексте размерность:

$$d_c = -\lim_{\varepsilon\to 0} \mathfrak{I}/\log\varepsilon$$

является *скоростью изменения информации при бесконечном увеличении разрешения*.

Примечания

1. *А. Пуанкаре* писал по поводу концепций, не согласованных с интуицией: «Они заменяют определяемый предмет и интуитивное понятие этого предмета конструкцией, сделанной из более простых материалов. Мы видим, что из этих материалов действительно можно собрать такую конструкцию, но никогда непонятно, почему собрали эти материалы именно так, а не иначе. Я не хочу сказать, что арифметизация математики — плохая вещь, я утверждаю лишь, что она не составляет всего» [1]. Несмотря на указанные недостатки, параметрическая размерность осталась в физике, где она согласована с числом степеней свободы.

2. Такая размерность называется большой индуктивной размерностью — Ind. Кроме того, существует малая индуктивная размерность (ind) : считают, что $\mathrm{ind}X = n$, если $\forall x \in X$ существует сколь угодно малая окрестность $U : x \in U$, граница которой ∂U имеет $\mathrm{ind}(\partial U) = n-1$. Начало этой индуктивной цепочки — пустое множество $\varnothing$, для которого $\mathrm{ind}\varnothing = -1$.

3. Точнее, при $\varepsilon \to 0$ существует соотношение: $N(\varepsilon) \propto V(1)\varepsilon^d$, в котором знак $\propto$ понимается в следующем смысле. Для двух функций f и g, $f(\varepsilon) \propto g(\varepsilon)$, если $\lim\{\ln f(\varepsilon)/\ln g(\varepsilon)\} = 1$ при $\varepsilon \to 0$. Поэтому **емкость** определяется как предел, *если он вообще существует*, отношения
$$\{\ln N(\varepsilon) - \ln V(1)\}/\{\ln(1/\varepsilon)\}.$$

Очевидно, что $\ln V(1)/\ln(1/\varepsilon) \to 0$ при $\varepsilon \to 0$, поэтому этот член обычно опускают. Существуют две теоремы [2], которые позволяют перейти к дискретному варианту для ε. Первая из них утверждает, что емкость компактного подмножества $\mathfrak{F}$ совпадает с пределом:

$$\lim_{n\to\infty}\{\ln N(\varepsilon_n)\}/\{\ln(1/\varepsilon_n)\},$$

где $\varepsilon_n = cr^n$, $0 < r < 1$, $c > 0$, $n = 1, 2, \ldots$, так что радиусы шаров можно менять дискретно. Вторая теорема позволяет использовать боксы размером $1/2^n$. Если $N_n(\mathfrak{F})$ число таких боксов, пересекающих $\mathfrak{F}$, то емкость определяется как предел при $n \to \infty$:

$$\lim\{\ln N_n(\mathfrak{F})\}/\{\ln 2^n\}.$$

Путеводитель по литературе. Подход Пуанкаре к понятию размерности изложен им в книге [1]. О работах Минковского и Борхарда можно узнать из монографии Лебега [3]. Элементарное изложение современных понятий о математической непрерывности содержат учебные курсы [4, 5]. В монографии [6] можно найти много интересного по истории вопроса. Наконец, подробное описание всевозможных размерностей можно найти в обзорах [7–9]. С информацией Шеннона и ее применением можно познакомиться по книге [10] или замечательным беседам Реньи [11].

Дробные размерности

> Оставаясь на почве любой физической теории, мы не можем интерпретировать формулы, содержащие дробные показатели основных единиц.
>
> *В. Вильямс*

История началась с попыток английского физика *Ричардсона* измерить длину побережья Британии. Располагая подробной картой, он аппроксимировал линию побережья ломаной L_b, составленной из отдельных хорд длиной b, все вершины которой лежали на берегу. *Ричардсон* полагал, что в этом случае существует предел $L_b \to L$ при $b \to 0$, как бывает для гладких кривых. Однако, оказалось, что с уменьшением b суммарная длина ломаной растет до бесконечности по закону:

$$L_b = \lambda b^{1-D},\ D < 2.$$

Если построить на каждом звене b квадрат, то суммарная площадь квадратов: $b^2 N = b^2 L_b / b = \lambda b^{2-D} \to 0$ при $b \to 0$ и $D < 2$. Таким образом, береговая линия имеет бесконечную длину и порождает нулевую площадь! Первое, что приходит в голову, это сомнения в корректности такой аппроксимации. В математике уже давно известны случаи, когда метод хорд не работал.

Классический пример привел еще *Лебег*: возьмем равносторонний треугольник ABC и соединим середины трех сторон (см. рис. 1).

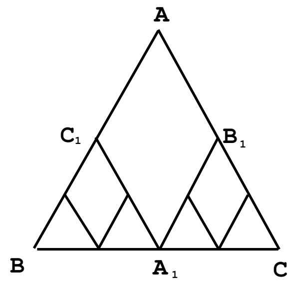

Рис. 1. Парадокс Лебега

Очевидно, $AB + AC = BC_1 + C_1A_1 + A_1B_1 + B_1C$. Продолжая процесс разбиения, мы приближаем ломаную к стороне BC. Если считать последнюю пределом ломаной, мы получим $AB + AC = BC$. Для того, чтобы разобраться в деталях парадокса, необходима нетривиальная математика: понятие тонкой сходимости, трансфинитные повторные пределы и т. п. Мы заменим все это здравым смыслом и простыми геометрическими соображениями. В действительности, сторона BC не является пределом зигзагообразной ломаной — эта ломаная действительно приближается к BC по положению, *но не по направлению*. Все дело здесь в степени гладкости. Вспомним, что длина кривой выражается интегралом, под знаком которого стоят производные от функции, задающей кривую в параметрическом виде. Таким образом, для вычисления длины необходима гладкость кривой по меньшей мере $\mathcal{C}^1$, т. е. ее уравнение имеет непрерывную первую производную. Длина спрямляемой кривой действительно ограничена верхней гранью суммарных длин вписанных ломаных. На такой кривой $x = x(s), s \in [0,1]$ конечной длины L в качестве параметра удобно выбрать длину дуги. В этом случае справедливо условие **Липшица**: $|x(s_1) - x(s_2)| \leqslant |s_1 - s_2|$ или при $s \to t, s = Lt : |x(t_2) - x(t_1)| \leqslant L|t_2 - t_1|$ и касательная вдоль кривой меняется непрерывно. Углы, которые образует произвольная хорда с касательными, не превосходят наибольшего из углов, образованного разными касательными в концах дуги. Поэтому, когда число вписанных в дугу звеньев ломаной возрастает, последняя приближается к дуге не только по положению, *но и по направлению*$^{(1)}$. В случае, если гладкость меньше чем $\mathcal{C}^1$, хорды при измельчении стремятся встать перпендикулярно к стороне треугольника, что и приводит к расходящемуся пределу.

Существует аналог парадокса *Лебега*. В конце прошлого века *Герман Шварц* показал, что триангуляция цилиндра единичного радиуса и единичной высоты дает для площади боковой грани произвольную величину. Причины обоих парадоксов — в функциях, гладкость которых ниже той, с которой мы привыкли иметь дело в обычном анализе.

В общем случае формула линейных приращений: $\triangle y = \mu \triangle x + O(\triangle x^2)$ заменяется на $\triangle y = \mu_H (\triangle x)^H$, где H — **показатель Гельдера**, а μ и μ_H — обычная и гельдеровская производные. Показатель $H = 1$ для гладких функций. В случае $H < 1$ вместо касательной имеется криволинейный конус $\triangle y \approx \triangle x^H$. Объекты, для которых $H < 1$, а показатель D в формуле *Ричардсона* строго больше единицы, называются *фракталами*. Рассмотрим несколько известных примеров.

Пример 1. **Множество Кантора.** Пусть $\mathfrak{F}_0 = [0,1]$. Выбросим из $\mathfrak{F}_0$ интервал $(1/3, 2/3)$, а то что останется, обозначим $\mathfrak{F}_1$. Затем выбросим из $\mathfrak{F}_1$ интервалы $(1/9, 2/9)$ и $(7/9, 8/9)$ и получим $\mathfrak{F}_2$ (рис. 2). Продолжая этот процесс, мы придем к убывающей последовательности замкнутых интервалов $\{\mathfrak{F}_n\}$. Множество $\mathfrak{F} = \cap \mathfrak{F}_n$ называют ***канторовым множеством***. Пусть $\mathfrak{S}$ — множество «выброшенных» кусков отрезка $[0,1]$ при построении $\mathfrak{F}$, т. е. $\mathfrak{S} = (1/3, 2/3) \cup (1/9, 2/9) \cup (7/9, 8/9) \cup \ldots$ Тогда

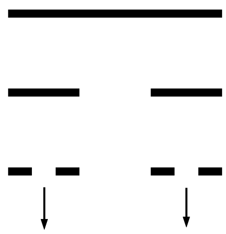

Рис. 2. Множество Кантора

лебегова мера $\mathfrak{S}$ равна сумме: $1/3 + 2 \times 1/9 + 4 \times 1/27 + \ldots = 1$. Таким образом, при построении $\mathfrak{F}$ мы выбросили всю длину отрезка! Но осталось бесконечное множество точек, *канторово множество*: $[0,1] \backslash \mathfrak{S}$, имеющее мощность континуума и *лебегову* меру нуль$^{(2)}$. Например, ему принадлежат точки $\{0, 1, 1/3, 2/3, 1/9, \ldots\}$ — то есть концы выбрасываемых интервалов, но не только они! Все точки $\mathfrak{F}_o \in \mathfrak{F}$ можно описать следующим образом. Запишем каждое из чисел $x \in [0,1]$ в троичной системе: $x = a_1/3 + a_2/3^2 + \ldots + a_n/3^n + \ldots$, где $a_n = 0, 1$ или 2. Некоторые числа в этом представлении допускают двойную запись, например: $1/3 = 1/3 + 0/3^2 + \ldots$ и $1/3 = 0/3 + 2/3^2 + 2/3^3 + \ldots$. Легко убедиться, что $\mathfrak{F}$ принадлежат только те x, которые могут быть записаны хотя бы одним способом так, что в последовательности числителей $a_1, a_2, \ldots, a_n, \ldots$ ни разу не встречается единица. Совокупность таких последовательностей имеет мощность континуума$^{(3)}$.

Вычислим размерность самоподобия канторова множества. При первой итерации имеем $\varepsilon = 1/3, N = 2$, при второй — $\varepsilon = 1/9, N = 2^2; \ldots$, при k-ой — $\varepsilon = 1/3^k, N = 2^k$. Поэтому $d_H = \ln 2 / \ln 3 \approx 0{,}63$ — меньше, чем размерность исходного прообраза.

Существует функциональный аналог множества Кантора — *функция Кантора*. Распределим равномерно на $\mathfrak{F}$ единичную массу (меру) с плотностью μ. Тогда функция $F(x) = \int_0^x d\mu(x)$ описывает распределение меры на канторовом носителе. Она является непрерывной возрастающей функцией, которая тем не менее почти всюду имеет нулевую производную (т. е. горизонтальна!). Ее называют «чертовой лестницей» (рис. 3).

Пример 2. **Кривая Коха.** Возьмем равносторонний треугольник и определим следующую элементарную операцию: каждая сторона делит-

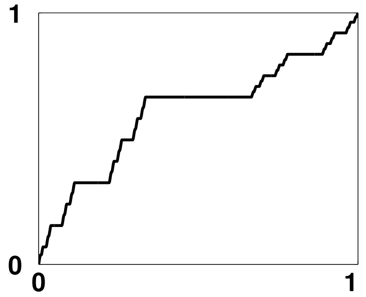

Рис. 3. Чертова лестница

ся на $r = 3$ части, после чего средний сегмент заменяется на равносторонний треугольник. Операция повторяется n раз. То, что получится при $n \to \infty$, и есть триада (снежинка) *Коха* (рис. 4)

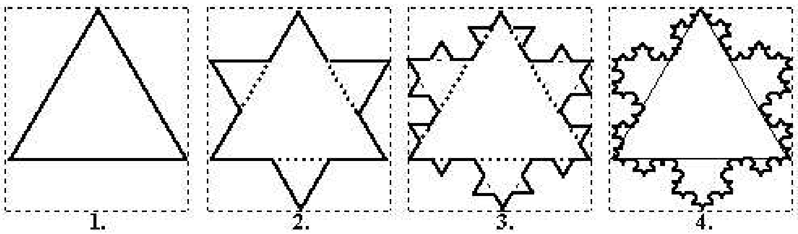

Рис. 4. Триада Коха. Четыре итерации

Ясно, что это множество самоподобно, точно так же, как и сам процесс его построения. На любом шаге его линейный элемент длины l заменяется на $N = 4$ элемента длиной $lr^{-1} = 1/3$ каждый. Поэтому размерность подобия или емкость $d_c = \ln N / \ln r = \ln 4 / \ln 3 \approx 1,2619$. Триада Коха является математической моделью кривой побережья, с которой работал *Ричардсон*. Любопытно, что в $\mathbb{R}^2$ можно построить множества с дробной размерностью, которые обладают инфернальным свойством *призраков* или *вурдалаков*: они *не отбрасывают тени*. Точнее, для таких объектов лебегова мера проекций на некоторые (или даже на все) направления равна нулю.

Примечания

1. Наиболее важным обстоятельством здесь является переход к пределу даже для гладкой кривой. Действительно, сегмент дуги рассматриваемый непосредственно, ничем не отличается от его половины удвоенной микроскопом. Целое однородно с частью и как заметил Пуанкаре [1]: «Здесь заключается противоречие, или скорее это было бы противоречием, если бы число членов предполагалось конечным; в самом деле, ясно, что часть, которая содержит менее членов сравнительно с целым, не может быть подобной целому».

2. Напомним, что множество $\mathfrak{X}$ на $\mathfrak{R}$ имеет меру нуль, если его можно покрыть интервалами, сумма длин которых произвольно мала. Например, $\mathfrak{X}$ состоит из рациональных чисел из $(0,1)$. Они могут быть записаны в виде последовательности: $a_1, a_2, a_3, \ldots$, например так:
$$1/2,\ 1/3,\ 2/3,\ 3/4,\ 1/5,\ 2/5, \ldots$$
Для любого $\varepsilon > 0$, a_1 можно заключить в интервал длины $\varepsilon/2$; a_2 — в интервал длины $\varepsilon/4; \ldots$ a_n — в интервал длины $\varepsilon/2^n$. Эти интервалы покрывают $\mathfrak{X}$ и сумма их длин равна ε. Следовательно, множество рациональных чисел имеет меру нуль.

3. Действительно, любому набору $\{a_1, \ldots, a_n, \ldots \,|\, a_i = 0 \text{ или } 2\}$ можно поставить в соответствие последовательность $b_1, \ldots, b_n, \ldots$, где $b_n = 0$, если $a_n = 0$ и $b_n = 1$, если $a_n = 2$. Но такая последовательность — просто двоичный код числа, принадлежащего $[0,1]$.

Путеводитель по литературе. Лучшими, по моему мнению, являются обзоры [7,9] написанные для физиков. Работы Ричардсона изложены в монографии [12], а в [13] им посвящена целая глава. Статьи Мандельброта в [14,15] весьма содержательны, но не подходят для первого чтения. Небольшая книга Лебега [3] содержит обсуждение упомянутых парадоксов. Описание функции Кантора можно найти в [16], а примеры фракталов-вурдалаков приведены в дополнении к книге [17]. Наконец, некоторые вопросы, связанные с тонкой сходимостью, строго рассмотрены в [18].

Фракталы, неполная автомодельность и контекстно-свободные грамматики

Plus ça change, plus c'est la même chose.
Чем больше оно изменяется, тем более остается тем же.

Французская пословица

Вернемся к построению *триады Коха*. Здесь все легко считается. Обозначим через L ребро треугольника на первом шаге. На n-ом шаге размер одной стороны $b = L/3^n$, а периметр $L_b = 3L(4/3)^n$ (рис. 4).

Когда $n \to \infty, b \to 0$, $L_b \to \infty$. Поскольку $n = \lg(L/b)/\lg 3$, имеем:

$$L_b = 3L10^{\alpha \lg(L/b)} = 3L(L/b)^\alpha,$$

где $\alpha = (\lg 4 - \lg 3)/\lg 3 \approx 0,2618$.

Мы легко получим формулу *Ричардсона*, если положим:

$$\lambda = L^{1-\alpha} = L^D;\ L_b = 3\lambda b^{1-D};\ D = 1 + \alpha = 1,2619.$$

Аналогия с береговой линией становится полной, поскольку число звеньев триады $N = L_b/b = \lambda b^{-D}$ и суммарная площадь квадратов, построенных на периметре, $Nb^2 = \lambda b^{2-D} \to 0$ при $b \to 0\,(D < 2)$. Мы можем получить конечную величину, если удастся найти такое D_*, что $Nb^{D_*} = \lambda$, где $1 < D_* < 2$; при этом $Nb = \infty$ и $Nb^2 = 0$, когда $b \to 0$.

Легко усмотреть основные особенности построения триады:

1. **Однородность**: каждая сторона треугольника n-го шага порождает одинаковое число звеньев $n + 1$.

2. **Самоподобие**: число звеньев $n + 1$ шага зависит только от отношения звеньев n-го и $n + 1$ шагов.

Оказывается, что этих двух условий, заданных локально, достаточно для получения *формулы Ричардсона*(1).

Рассмотрим произвольную кривую, которая аппроксимируется системой ломаных линий с уменьшающейся длиной звена. Пусть малая окрестность кривой содержит две соседних вершины ломаной с длиной звена a. Пусть N_{ab} — число вершин ломаной со звеном $b < a$, расположенных между теми же вершинами. Свойство *локального* самоподобия означает, что N_{ab} асимптотически при $a/b \to \infty$ можно представить разложением, главный член которого не зависит от a и b, а является функцией их отношения: $N_{ab} = f(a/b)$ при фиксированном $a/b \gg 1$. Возьмем новую ломаную с длиной звена $c \ll b$. В силу локальной однородности и самоподобия число ее звеньев, приходящихся на длину a, равно $f(a/c)$. С другой стороны, это же число равно произведению $(N_{ab}) \times (N_{bc})$. Здесь первый сомножитель — число b–звеньев внутри одного a–звена, а второй — число c–звеньев внутри одного b–звена. Таким образом, получаем функциональное уравнение:

$$f(a/c) = f(a/b)f(b/c).$$

Замена переменных $a/b = x$; $a/c = y$ дает:

$$f(y)/f(x) = f(y/x).$$

Дифференцируя обе части по y и полагая $y = x$, получим: $f'(x)/f(x) = (1/x)f'(1) = D/x$ или $f(x) = x^D$.

Следовательно, при упомянутых предположениях для длины ломаной справедлива асимптотика:

$$L_b = a^D b^{1-D} + \dots$$

Напомним, что процедура называется **автомодельной**, если ее характеристики на двух смежных уровнях связаны друг с другом преобразованием подобия. Асимптотика показывает, что фракталы обладают свойством неполной автомодельности. Поясним этот момент. В общем случае длина аппроксимирующей ломаной для непрерывной кривой между двумя точками, разделенными расстоянием a, зависит от a и длины звена b. Из соображений размерности $L_a = af(a/b)$. Для гладкой кривой при $a/b \to \infty$ $(b \to 0)$ функция f стремится к конечному пределу $f(\infty)$. Именно тогда величина $f(\infty)a$ является длиной отрезка гладкой кривой. Например, для окружности с диаметром a, $f(\infty) = \pi/2$. Таким образом, длина окружности автомодельна по параметру a/b при $a/b \to \infty$. Для фрактальных кривых конечного предела $f(a/b)$ не существует. Однако имеет место **неполная автомодельность**:

$$f(a/b) \propto (a/b)^{D-1}.$$

Разумеется, D зависит от геометрии кривой и не определяется из соображений размерности.

Существует любопытный подход к неполной автомодельности, связанный с *символической динамикой*. Рассмотрим **бинарное слово** (двоичную последовательность):

$$z = z_0 z_1 \dots z_{r-1},$$

где z_i — нули, либо единицы. Длина слова $\log_2 z = r$, а $s \leqslant r$ число мест, занятых единицей. Определим гомоморфизм $\mathbf{T} : [0,1] \to [0,1]$ уравнениями: $\mathbf{T}(0) = 0^r$, $\mathbf{T}(1) = z$. Пусть, например, $z = 101$. Возьмем точку $x_0 = 1$ и рассмотрим последовательность итераций: $x_1 = \mathbf{T}(x_0), \dots, x_k = \mathbf{T}(x_{k-1})$. В результате получим последовательность:

$$\begin{gathered} x_0 = 1 \\ x_1 = \mathbf{T}(1) = 101 \\ x_2 = \mathbf{T}(\mathbf{T}(1)) = \mathbf{T}(101) = \mathbf{T}(1)\mathbf{T}(0)\mathbf{T}(1) = 101000101 \end{gathered}$$

Можно убедиться, что неподвижной точкой x_f гомоморфизма $\mathbf{T}$ будет предел $\lim_{k\to\infty} x_k$, удовлетворяющий уравнению $\mathbf{T}(x) = x$:

$$x_f = 101000101000000000101000101\dots$$

Это символическая запись *множества Кантора*. Действительно, 101 представляет собой *код* первого шага построения этого фрактала, когда мы удаляем середину (0) единичного интервала. На втором шаге та же процедура применяется к получившимся подинтервалам: $[0, 1/3] \to 101$; $[2/3, 1] \to 101$ и т. д.

Примечания

1. Или на другом языке — получения **ренормгруппы**, относительно которой самоподобные системы являются инвариантными множествами. *Канторову пыль* $\mathfrak{K}$ можно рассматривать как подмножества интервала $[0, 1]$. Пусть преобразование S множества $\mathfrak{F}$ на вещественной оси дает образ $S(\mathfrak{F})$ и обладает следующим свойством:

образ объединения двух множеств равен объединению их образов. Рассмотрим два отображения:

$$\mathbf{S}_1(\mathfrak{F}) = [x/3 : x \in \mathfrak{F}]$$
$$\mathbf{S}_2(\mathfrak{F}) = [(x+2)/3 : x \in \mathfrak{F}]$$

$\mathbf{S}_1$ отображает *канторово множество* в его левую половину, а $\mathbf{S}_2$ — в правую. Очевидно, $\mathbf{T}(\mathfrak{F}) = S_1 \cup S_2$ — оставляет пыль неизменной. Преобразования, оставляющие множество неизменным, $\mathbf{T}(\mathfrak{K}) = \mathfrak{K}$ называются **симметриями** $\mathfrak{K}$. Если приложить $\mathbf{T}$ к интервалу $[0,1]$ несколько раз, то $\lim_{n\to\infty} \mathbf{T}^n([0,1]) = \mathfrak{K}$. Фактически $\mathfrak{K}$ можно генерировать из любого ограниченного множества вещественной оси, даже из любой единственной точки x_0 : $\mathbf{T}(x_0)$ тогда состоит из двух точек, $\mathbf{T}^2(x_0)$ — из четырех и т. д. На языке теории ренорм-групп $\mathfrak{K}$ — единственная притягивающая «точка» или инвариантное множество $\mathbf{T}$. В этом контексте $d_c = \ln 2/\ln 3$ результат того, что $\mathbf{T}$ объединение двух сжатий. Можно показать, что $\mathbf{T}, \mathbf{T}^2, \ldots$ — единственные симметрии канторова множества.

Путеводитель по литературе. Монография [13] — единственный известный мне источник, где фракталы изложены с позиций неполной автомодельности. Связь свойств самоподобия и однородности с теорией групп обсуждается в обзоре [19]. Связям фракталов с теорией чисел и символической динамикой посвящены работы [2, 20, 23].

Фракталы и системы итеративных функций

Но как же оно образуется, если не содержится в том же своем прообразе?

Николай Кузанский
«Игра в шар»

Отображение $f : X \to X$, действующее в метрическом пространстве (X, d), называется *сжимающим*, если существует $s \in [0,1)$ такое, что

$$d(f(x), f(y)) \leqslant sd(x,y)$$

для всех x и y из X; s называется коэффициентом сжатия отображения.

Системой итеративных функций (СИФ) $(X, \{f_i\}), i = 1, 2, \ldots, k$ называют набор сжимающих отображений[5] $\{f_i\}$, в компактном метрическом пространстве (X, d). Коэффициентом сжатия СИФ называют $s = \max\{s_i : i = 1, \ldots, k\}$.

Для пространства (X, d) можно определить другое метрическое пространство $(H(X), h)$, называемое «пространством фракталов».

Пусть $H(X)$ — множество непустых компактных подмножеств X. Определим на нем *метрику Хаусдорфа* h следующим образом. Пусть $B(x, r)$ — замкнутый шар, радиуса r с центром в точке x. Для произвольного множества $A \in X$ *дилатацией*[6] A_r радиуса r множества A называется $\bigcup_{x \in A} B(x, r)$. Таким образом, дилатация множества A — это добавление к A всех точек, лежащих на расстоянии $\leqslant r$ от его границы. Пусть A, B — непустые компактные подмножества из X. Тогда расстояние(метрика) Хаусдорфа определяется как (рис. 5)

$$h(A, B) = \min\{r > 0 | A \subset B_r; B \subset A_r\}.$$

Таким образом, это минимальное из двух чисел: первое из них получается расширением множества A, до тех пор, пока его образ не поглотит B, второе — дилатацией B, пока она не поглотит A[7].

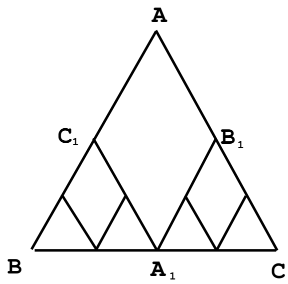

Рис. 5. Метрика Хаусдорфа

[5] Такую СИФ иногда называют гиперболической.

[6] Т.е. расширением.

[7] Другое определение дано в глоссарии.

Можно показать, что если (X, d) — полное метрическое пространство, то $(H(X), h)$ также является полным метрическим пространством. Определим преобразование $T : H(X) \to H(X)$ как

$$T(B) = \bigcup_{i=1}^{k} f_i(B), \forall B \in H(X),$$

где $T(B) = \{T(x)|x \in B\}$ — *оператор Хатчинсона*.

Пусть $B \in H$ и $T^{\circ n}$ — композиция [8] порядка n оператора T:

$$T^{\circ 0}(B) = B,\ T^{\circ(j+1)}(B) = T^{\circ j}(T(B)).$$

Последовательность множеств, полученная в результате итерирования $T(B)$, т. е.

$$\{B, T(B), T^{\circ 2}(b), T^{\circ 3}(B), ..., T^{\circ n}(B), ...\},$$

называется *орбитой* B для $(H(X), T)$. Пару $(H(X), T)$ можно рассматривать как *детерминированную дискретную динамическую систему* с пространством состояний $H(X)$ и преобразованием T.

Согласно *теореме Банаха о неподвижной точке*, действие сжимающего преобразования T на произвольную начальную точку $B_0 \in H$ в пространстве $(H(X), h)$ приводит к последовательности точек $B_0, B_1 = T(B_0), B_2 = T(B_1) \equiv T^{\circ 2}(B_0), \ldots$, которая сходится к некоторой точке $A \in H$. Она является единственным решением уравнения $T(A) = A$. Очевидно, что

$$\lim_{n\to\infty} T^{\circ n}(B) = A.$$

Неподвижная точка A называется *аттрактором* СИФ или *фракталом*.

При практическом использовании СИФ для построения фракталов оператор Хатчинсона применяют только к граничным точкам компакта.

Пример 1. Пусть $X = [0,1]$, $f_1 = 1/3x$, $f_2 = 1/3x + 2/3$. Пусть $T = f_1 \cup f_2$. Последовательность итераций $T(X)$ при $n \to \infty$

$$T([0,1]) = \{[0,1/3], [2/3,1]\},$$

$$T^{\circ 2} = \{[0,1/9], [2/9,1/3], [2/3,7/9], [8/9,1]\}, \ldots$$

[8] Не путать со степенью!

очевидно приводит к канторову множеству: $\mathfrak{F} = \bigcap T^{\circ n}[0,1]$.

Пример 2. Пусть $X \in \mathbb{R}^2$. Определим СИФ как

$$f_1\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$f_2\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1/2 \\ 0 \end{bmatrix},$$

$$f_3\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

Аттрактор этой СИФ (ковер Серпинского) приведен на рис. 6. Возьмем

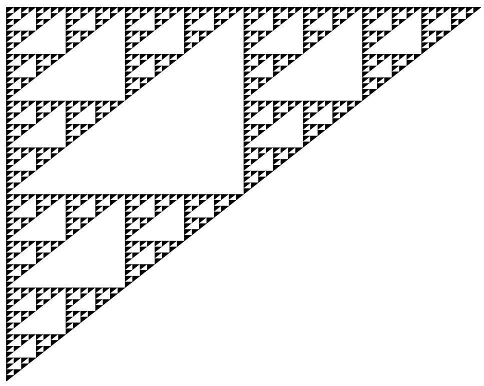

Рис. 6. Фрактал Серпинского

в качестве исходного множества при построении ковра прямоугольный треугольник B с координатами вершин $(0,0), (0,1), (1,1)$. Тогда все f_i — аффинные преобразования сжатия B с коэффициентом 0.5. Преобразование f_3 сдвигает образ B вправо на $1/2$, f_2 — вверх на эту же величину и f_1 оставляет его на месте. Таким образом, для каждого n $T^{\circ n}(B)$ репродуцирует сжатие и сдвиги. Заметим, что любое аффинное преобразование в $\mathbb{R}^2$ кодируется всего шестью коэффициентами, поэтому СИФ реализует фрактальное сжатие изображения ковра. Множество

$T(B) : f_1(B) \cup f_2(B) \cup f_3(B)$ представляет собой композицию (коллаж) трех уменьшенных копий B. Нахождение коэффициентов аффинных преобразований по фрагменту коллажа называют обратной задачей в теории фракталов. В ее решении важную роль играет следующая

Теорема о коллаже. Пусть $\{(X, d); f_1, f_2, \ldots, f_k\}$ — гиперболические СИФ с коэффициентом сжатия s, B — произвольное непустое компактное множество, принадлежащее $H(X)$ и A — аттрактор СИФ. Пусть

$$h(B, \bigcup_{i=1}^{k} f_i(B)) \equiv h(B, T(B)) < \varepsilon.$$

Тогда

$$h(A, B) \leqslant h(B, T(B)) \frac{\varepsilon}{1 - s}.$$

Таким образом, для решения обратной задачи следует выбирать B, «похожее» на фрагмент коллажа.

Системой *случайных итеративных функций* (ССИФ)

$$\{X; f_1, f_2, \ldots, f_k; p_1, p_2, \ldots, p_k.\}$$

называют СИФ, снабженную набором вероятностей $\{p_i | i = 1, 2, \ldots, k\}$ для каждого f_i, где $p_i > 0$ и $p_1 + p_2 + \ldots + p_k = 1$. Пусть $\sigma = (\sigma_1, \ldots, \sigma_N)$; $\sigma_n \in \{1, \ldots, k\}$. Тогда орбита ССИФ определяется как $x_{n+1} = f_{\sigma_n}(x_n)$, где f_{σ_n} выбирается с вероятностью[(1)] p_{σ_n}. Предположим, что преобразования f_i, $i = 1, 2, 3$ при построении ковра Серпинского выбираются с фиксированными вероятностями $p_1 > p_2 > p_3$. Тогда, образы $f_1(B)$ при итерациях T будут появляться «в среднем» чаще и, следовательно, соответствующие фрагменты в ковре окажутся «более черными» на каждом масштабе. При этом условие нормировки вероятностей при переходе к второй итерации индуцирует новые вероятности:

$$\begin{aligned}
(p_1) &\to (p_1 p_1, p_1 p_2, p_1 p_3) \\
(p_2) &\to (p_2 p_1, p_2 p_2, p_2 p_3) \\
(p_3) &\to (p_3 p_1, p_3 p_2, p_3 p_3)
\end{aligned}$$

Эта цепочка продолжается с ростом номера итерации и мы получим ковер Серпинского в черно-серо-белых тонах, причем интенсивность окраски (мера) приобретает скейлинговые свойства. На другом языке такую картину называют *мультифрактальной мерой*.

Заметим, что кроме ССИФ существуют их разновидности — рандомизированные СИФ. Это обычные итеративные функции, коэффициенты которых выбираются случайно на каждом шаге итерации из некоторого множества. Пример аттрактора такой СИФ приведен на рис. 7.

Примечания

1. Рассмотрим некоторое измеримое множество B, в которое может заходить орбита ССИФ. *Эргодическая теорема Элтона* утверждает, что для каждой последовательности символов $\{\sigma_n\}$ частота, с которой орбита $\{x_n\}$ посещает B, равна мере этого множества $\mu_F(B)$.

$$\lim_{N\to\infty} \frac{\sharp\{x_n \in B:\ 1 \leqslant n \leqslant N\}}{N} = \mu_F(B), \quad B \subset \mathbf{B}(X),$$

где $\sharp$ заменяет слово «количество». Другими словами, последовательность точек орбиты ССИФ $x_n = f_{\sigma_n} \circ f_{\sigma_{n-1}} \circ \ldots \circ f_{\sigma_1}(x_0)$ сходится к глобальному аттрактору в смысле расстояния Хаусдорфа для почти каждой точки $x_0 \in X$ и для любого набора вероятностей.

Путеводитель по литературе. Оригинальная работа Хатчинсона [21] трудна для первого чтения, однако существует ее упрощенный вариант [22]. Наиболее полное изложение СИФ и ССИФ дано в книге Барнсли [2] и в недавно переведенной книге Кроновера [23]. Хорошие введения в теорию детерминированных и случайных СИФ можно найти на *Web*-страничках, см., например, [24]. Основные идеи мультифрактального формализма можно найти в [20, 22].

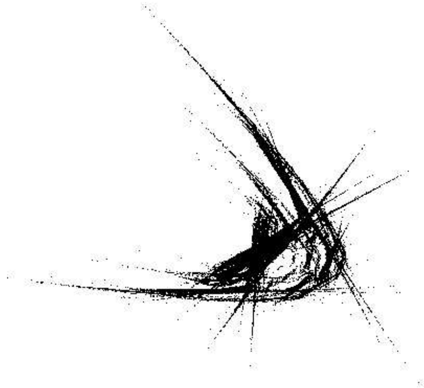

Рис. 7. Случайный Фрактал

Динамические системы и странные аттракторы

> Поэтому, если тебе сначала все покажется пустым бредом, знай, что причиной твоя слабость!
>
> *Николай Кузанский*
> *«Берилл»*

Дискретная *динамическая система* задается отображением[1]

$$x_{n+1} = f(x_n), x \in R^n.$$

Пространство, в котором работает отображение, обычно называют *фазовым*. Траектория такой системы получается последовательным применением оператора f:

$$x_1 = f(x_0);\ x_2 = f(x_1) = f(f(x_0)) = f^{\circ 2}(x_0)\ldots$$

Для некоторых систем траектории образуют множества, весьма экзотического вида. Одно из таких множеств приведено на рис. 8. Оно образовано итерациями системы:

$$\begin{cases} x_{n+1} = y_n - sign(x_n)|x_n|^{1/2} \\ y_{n+1} = 0.4 - x_n \end{cases}$$

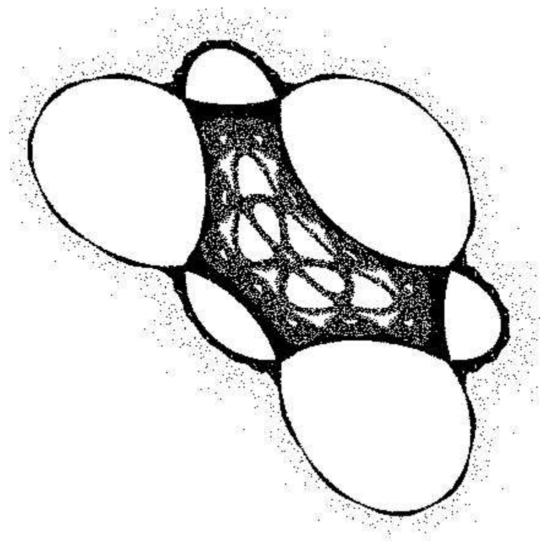

Рис. 8. Фазовый портрет дискретной системы

Как и в случае СИФ, *неподвижные точки* x_f отображения f определяются решением уравнения: $f(x_f) = x_f$. Неподвижные точки могут быть *устойчивыми*, если к ним сходиться последовательность точек $f^{\circ n}(x_0)$, $n \to \infty$, и неустойчивыми, в противном случае. Легко понять, в каких случаях возникает устойчивость. Рассмотрим «возмущенную» неподвижную точку $x = x_f + \delta x$, которую можно получить с помощью малого возмущения отображения f:

$$x_{n+1} = x_f + \delta x_{n+1} = f(x_f + \delta x_n) \simeq f(x_f) + \delta x_n f'(x_f).$$

Поскольку $\delta x_{n+1} = f'(x_f)\delta x_n$, понятно, что $|\delta x_n|$ будет увеличиваться, если $|f'(x_f)| > 1$. Таким образом, производная $|f'|$ играет роль локального коэффициента сжатия или расширения для нелинейного (в общем случае) отображения f. Очевидно, что устойчивой неподвижной точке соответствует сжатие: $|f'(x_f)| < 1$. Для многомерного случая надо вычислять производные по всем координатам, т. е. якобиан $J = det|\frac{\partial x^i_{n+1}}{\partial x^j_n}|$ отображения f в точке $x_f = (x^1_f, x^2_f, \ldots, x^n_f)$. В этом случае условие устойчивости $J < 1$ соответствует сжатию в J раз объема исходного «параллелепипеда», натянутого на координатные оси, которое происходит под действием отображения f.

Разновидностью неподвижных точек являются *периодические точки* отображения f: говорят, что точка x — *периодическая* с периодом p, если существует такое минимальное число p, что $f^{\circ p}(x) = x$. Все эти объекты *инвариантны* относительно действия отображения или его итераций. Это весьма напоминает фракталы — предельные образы СИФ, которые остаются инвариантными под действием оператора Хатчинсона. Для того, чтобы сделать эту аналогии более точной, введем некоторые дополнительные определения.

Точка y называется *ω-предельной точкой* для точки x, если

$$f^{\circ n}(x) \to y,\; n \to \infty.$$

Иными словами, предельная точка — это «конец» траектории, которая начинается в x. Все ω-предельные точки образуют *ω-предельное* множество. Следующим важным понятием является обобщение идеи инвариантности неподвижных и периодических точек относительно действия f на множества. А именно, множество $B \in R^n$ называют *инвариантным*, если для всех $n \geqslant 0,\; f^{\circ n}(B) = B$. *Окрестностью* множества B называют открытое множество U, которое содержит B и все его предельные и граничные точки. Можно представить себе это множество как шар без границ, внутри которого находится все B и все точки, которые служат пределами сходящихся последовательностей точек из B. Наконец, расстоянием между точкой x и множеством B называют наименьшее из всех расстояний, когда точка y пробегает по всему B, т. е. $l(x, B) = \inf_{y \in B} \|x - y\|$.

Замкнутое инвариантное множество $A \subseteq X$ называется *притягивающим множеством*, если для него существует окрестность U такая, что $\forall x \in U\; f^{\circ n}(x) \to A$ при $n \to \infty$. Наибольшее U, которое обладает таким свойством, называется *бассейном притяжения* для A. Иногда свойство притяжения формулируют на языке *асимптотической устойчивости.* Множество A называют *устойчивым по Ляпунову* если начальное расстояние $l(x, A) < \delta$ между любой точкой и множеством со временем становится меньше любого заданного числа ε, т. е. $l(f^{\circ n}(x), A) < \varepsilon$. Если, кроме того, $\varepsilon \to 0$ при $n \to \infty$ множество называют *асимптотически устойчивым.* Именно таким и является притягивающее множество.

Наиболее важное понятие — *аттрактор* — основано на выделении минимальной [9] структуры притягивающего множества. *Аттрактором* A называют притягивающее множество, содержащее точку x, для которой $\omega(x) = A$. Иными словами, аттрактор содержит целиком *всю траекторию* $f^{\circ n}(x)$. Иногда говорят, что A содержит всюду плотную траекторию x, понимая под этим, что в окрестности любой точки траектории на аттракторе, можно найти другую точку этой же траектории. Кроме того, если наугад выбранная точка аттрактора оказалась «пустой», не надо отчаиваться: в ее локальной окрестности обязательно найдется точка траектории. Часто на аттракторе можно ввести меру, т. е. некоторое число, аналогичное объему в евклидовой геометрии [10]. Это число в ряде интересных случаев можно оценить как относительную долю времени, которое проводит траектория в некотором компактном подмножестве аттрактора [11]. Тогда динамическая неразложимость аттрактора означает существование на нем инвариантной меры$^{(2)}$, которая не может быть представлена как взвешенное среднее нескольких ненулевых инвариантных мер.

Аттракторы возникают, как правило, для диссипативных динамических систем [12]. В том случае, если фазовым пространством является R^2, такие аттракторы имеют тривиальную структуру: это неподвижные точки и предельные циклы. Однако, в R^n, $n \geqslant 3$, диссипативные системы могут иметь экзотические аттракторы с фрактальной структурой.

Рассмотрим в качестве примера так называемый *аттрактор Хенона*. Исходная динамическая диссипативная система описывается нелинейными дифференциальными уравнениями в R^4. Для упрощения исследования эволюцию траекторий многомерных систем часто отображают на так называемое *сечение Пуанкаре* в фазовом пространстве. Этот прием заключается в следующем. Рассмотрим плоскость, ортогональную фазовым траекториям. Каждый момент прохождения орбиты через плоскость будем отмечать парой координат (x_n, y_n) той точки, в которой траектория «протыкает» плоскость. Тогда динамика исходной системы редуцируется

[9] Т.е. неразложимой на более мелкие элементы с теми же свойствами.

[10] Аналогичное лишь по свойствам: евклидов объем фрактального аттрактора равен нулю!

[11] Для *эргодической* меры, эта доля пропорциональна мере самого множества.

[12] Динамическая система называется *консервативной*, если отображение сохраняет ее фазовый объем, и *диссипативной* — в противном случае.

к дискретным преобразованиям $(x_n, y_n) \to (x_{n+1}, y_{n+1})$ точек плоскости. В частности, преобразование Хенона в R^2 задается оператором

$$T: \; x_{i+1} = y_i + 1 - ax_i^2, \; y_{i+1} = bx_i.$$

Этот оператор представляет собой композицию трех преобразований:

1. $T': \; x' = x, \; y' = y + 1 - ax^2$
2. $T'': \; x'' = bx', \; y'' = y'$
3. $T''': \; x''' = y'', \; y''' = x''$

Первое преобразование складывает фигуру и сохраняет площадь, второе — сжимает фигуру относительно оси x и уменьшает площадь, умножая ее на постоянный множитель $b < 1$, третье — поворачивает и сохраняет площадь, но меняет ее знак. Якобиан отображения равен

$$\frac{\partial(x_{i+1}, y_{i+1})}{\partial(x_i, y_i)} = -b.$$

Значение $|b| < 1$ соответствует сжатию. На рис. 9 приведен аттрактор преобразования Хенона для значений параметров $a = 1.4$; $b = 0.3$. Известно, что бокс-размерность этого аттрактора равна 1.26, так что это — фрактал.

Фрактальные аттракторы связаны с так называемыми *сценариями динамического хаоса*$^{(2)}$. Хаотические системы описываются полностью детерминированными уравнениями, однако прогноз их решений не может быть продолжен дальше некоторого ограниченного интервала времени. Это эффект так называемой *существенной зависимости* от начальных условий, которые всегда задаются с конечной точностью. Рассмотрим простой пример динамической системы [13], заданной на интервале $I = [0, 1]$ отображением: $x_{n+1} = 2x_n, \mod 1$. Динамика системы сводится просто к удвоению координаты точки и отбрасыванию единицы (это и обозначается как $\mod 1$), если полученное значение координаты становиться больше, чем размеры «фазового» пространства. Выберем начальную точку из интервала I и запишем ее координату (адрес) в виде двоичного кода. Это всегда можно сделать единственным образом, если

[13] Ее называют иногда *сдвигом Бернулли*.

заранее оговорить выбор способа записи для чисел, допускающих два варианта адреса. Итак, пусть например, $x_0 = 0,101101$, где адрес записан с точностью до шестого разряда. Последнее означает, что мы не знаем, какой из двух символов 0 или 1 стоит в следующем разряде. Заметим, что каждый разряд уточняет адрес:

$$0,1 \;\Rightarrow\; 1/2 < x_0 < 1$$
$$0,10 \;\Rightarrow\; 1/2 < x_0 < 3/4$$
$$0,101 \;\Rightarrow\; 5/8 < x_0 < 3/4 \;\cdots$$

Вспомним теперь, что умножение в двоичной арифметике — это просто сдвиг запятой вправо и отбрасывание единицы, с учетом модуля. Таким образом, эволюция нашей системы сводится просто к уточнению адреса начальной точки! Однако, через 6 итераций последний известный символ уйдет влево, исчерпав всю информацию, которая содержалась в начальных данных! Следовательно, на седьмой итерации мы не сможем сказать, в какую половину интервала I попадет очередная точка траектории. Понятно, что увеличение начальной точности не изменит ситуацию — увеличится лишь число «детерминированных» шагов. Легко понять причины потери точности: общее решение имеет вид: $x_n = 2^n x^0$, так что расстояние между сколь угодно близкими начальными точка-

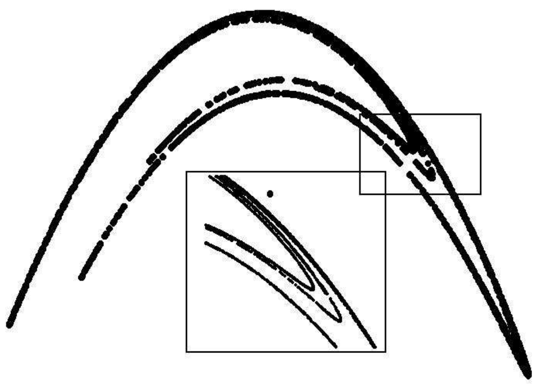

Рис. 9. Аттрактор Хенона

ми будет удваиваться при каждой итерации, пока не достигнет размеров всего интервала. Этот эффект называют *разбеганием близких траекторий*, существенной зависимостью от начальных данных или баттерфляй-эффектом [14]. Впервые этот эффект был обнаружен *Эдвардом Лоренцем* в 1961 году при численном решении простой системы из трех нелинейных уравнений, моделирующих конвекционные потоки в атмосфере. Уравнения были полностью детерминированными, т. е. они не содержали каких-либо случайных величин. Траектории системы заполняли фрактальное подмножество — аттрактор с хаусдорфовой размерностью 2.06. При этом, две численные траектории, стартующие из двух близких начальных точек очень быстро расходились, демонстрируя совершенно различное поведение. Поскольку начальные данные для метеосистем всегда получаются с конечной точностью, Лоренцу стало ясно что прогноз погоды будет всегда ограничен тем диапазоном времени, на котором расхождение близких траекторий еще не превышает некоторой заданной величины. Такой временной масштаб — его называют *горизонтом прогноза*, по-видимому, не превышает недели[(4)]. Существенно, что ограниченная предсказуемость является следствием нелинейности уравнений. В более общем случае оказалось, что динамический сценарий, которому следуют траектории хаотических систем определяется *типом* нелинейности, и почти не зависит от вида самого уравнения! Однако, обсуждение этих вопросов выходит за рамки данной лекции. В заключение еще раз заметим, что странные аттракторы хаотических динамических систем по своей сути являются синонимами предельных образов (аттракторов или фракталов) СИФ. Фракталы в геометрии — результат работы системы линейных сжимающих отображений (СИФ); в динамике, сжатие определяется нелинейными диссипативными уравнениями.

Примечания

1. В физике под динамической системой понимают обычно систему (автономных) обыкновенных дифференциальных уравнений $x' = f(x),\ x \in R^n$, где $x' \equiv dx/dt$, и $f : U \to R^n$ определено для некоторого открытого подмножества $U \in R^n$. Пространство R^n называют *фазовым пространством*, а $f(x)$ — *векторным полем*, так как ре-

[14] Из названия статьи *Э. Лоренца* «Предсказуемость: может ли взмах крылышек бабочки в Бразилии привести к образованию торнадо в Техасе?»

шение уравнения есть кривая с касательным вектором x'. В более общем случае под динамической системой в фазовом пространстве M понимают *однопараметрическую группу* преобразований $g^t : M \to M$, где параметром обычно служит время: непрерывное ($t \in R$) или дискретное ($t \in Z$). Точка $x_0 \in M$ под действием g^t переходит в точку $x_t = g^t(x_0)$, при этом групповое свойство означает, что $g^t g^s = g^{t+s}$. Последовательность точек $\{x_t\}$ образует *орбиту* группы. Если M — дифференцируемое многообразие, и g^t — взаимно однозначное дифференцируемое отображение, такое, что обратное к нему g^{-t} обладает аналогичными свойствами (такое отображение называют *диффеоморфизмом*), то пару $(M, \{g^t\})$ называют *фазовым потоком* или же *каскадом* — для случая, когда время дискретно. Легко показать, что «физическое» определение сводится к абстрактному. Рассмотрим динамическую систему заданную уравнением $x' = ax, \ x \in R$. Его решение можно записать в виде $x_t = g^t x_0 = e^{at} x_0$, так что поток индуцируется действием непрерывной группы $g^t = e^{at}$ однопараметрических преобразований в R.

2. Для того, чтобы ввести меру на траекториях динамической системы, заданной парой X, f, где X — топологическое пространство, выделяют некоторые интересные подмножества $B \subset X$, которые образованы обычно неподвижными точками или периодическими орбитами. Такие «элементарные» подмножества (их называют *борелевыми*) образуют «строительные блоки» для введения меры: она определяется на них, а затем обобщается на более сложные подмножества, которые можно построить из объединения борелевых «кирпичей». Например, для интервалов на прямой мера — это просто длина интервала. На плоскости, произведение интервалов дает прямоугольники, мера на которых — площадь. Объединение и пересечение прямоугольников порождает произвольные многоугольники. Их мера вычисляется по площадям «борелевых» прямоугольников, с учетом условия аддитивности. Говорят, что преобразование $f : X \to X$ *сохраняет меру* μ, если $\mu(f^{-1}(B)) = \mu(B)$. Такая запись позволяет рассматривать не только взаимно однозначные преобразования. Наиболее интересным утверждением о преобразованиях, сохраняющих меру на компактах, является *теорема Пуанкаре*

о возвращении. Пусть $B \in X$, μ — измеримо, X — компакт и обратимое преобразование f сохраняет меру на $X : \mu\{f(B)\} = \mu\{B\}$. Тогда существует некоторое n (возможно, зависящее от x_0), такое, что почти все точки $x_0 \in B$ возвращаются в B, т. е. $f^{\circ n}(x_0) \in B$. Иными словами, орбита любой точки $p \in B : \{p, f(p), f^{\circ 2}(p) \ldots\}$ содержит некоторые образы $f^{\circ k}(p)$, которые вновь попадают в B. Предположим, что это неверно и часть точек, образующих подмножество $A \subset B$, *никогда не возвращаются*. Тогда множества $\{A, f(A), f^{\circ 2}(A), \ldots\}$ *никогда не пересекаются!* Действительно, пусть точка $P \in f(A) \cap f^{\circ 3}(A)$. Очевидно, что $f^{-1}(P) \in A$, но она же принадлежит и $f^{\circ 2}(A)$, так как $P \in f^{\circ 3}(A)$. Но это невозможно, поскольку A не содержит возвращающихся точек. Следовательно, орбита $f^{circn}(A)$ разбивает X на бесконечную последовательность непересекающихся фрагментов, которые могут иметь только нулевую меру: X-компакт и его мера конечна.

3. Динамическая система (X, g) называется *хаотической*, если выполняется следующие условия:

   - пусть $x \in X$ и U — открытое подмножество, содержащее X. Если для некоторого $\delta > 0$ существует такое n и такая точка $y \in U$, что $d(g^n(x), g^n(y)) > \delta$, то g обладает *существенной зависимостью* от начальных условий;
   - g — *транзитивно*, т. е. для любой пары открытых множеств U, V существует такое n, что $g^n(U) \cap g^n(V) \neq \varnothing$;
   - периодические точки g плотны в X, т. е. в любой окрестности любой точки в X существует по крайней мере одна периодическая точка (и, следовательно, бесконечно много периодических точек).

4. Любопытно, что за 175 лет до Лоренца, ученик Эйлера, академик С. Я. Румовский писал в статье «Рассуждения о предсказании погод»: «Что же думать о предсказаниях погоды на целый год? Они не что иное суть, как тщетное и пустое умствование, которым легкомысленным людям во многих случаях вред нанести может». Статья была опубликована в № 1 Докладов Академии наук за 1786 год. Они назывались тогда «Новые ежемесячные сочинения к пользе и увеселению служащие».

Путеводитель по литературе. Связь фракталов и хаотических систем великолепно описана в книге Кроновера [23]. На сайтах [24] можно найти хорошие примеры дискретных преобразований на плоскости. Доступное изложение основ теории нелинейных динамических систем содержится в монографиях [25, 26] и журнальных обзорах [27, 28]. Строгому обсуждение концепции аттрактора посвящена статья Милнора [29]. Хорошие вводные курсы по топологической динамики можно найти на сайтах [30, 31], а современное изложение эргодической теории динамических систем на *Web*-страничке [31].

Нейронные сети, СИФ и гипернейрон

— Что это? Никак игрушка!
— Подберите фалды! ...
— Смотрите издали! ...

Козьма Прутков

Как известно, основной элемент модели нейронной сети — формальный нейрон — вычисляет скалярное произведение входного вектора

$$\mathbf{x} = (x_0, x_1, \ldots, x_n)^T$$

и вектора синаптических весов

$$\mathbf{w} = (\omega_0, \omega_1, \ldots, \omega_n)^T,$$

которое преобразуется затем функцией активации g, вычисляющей выход нейрона y:

$$y(\mathbf{x}, \mathbf{w}) = g(\mathbf{x} \cdot \mathbf{w})$$

В качестве g обычно используется либо функция Хевисайда (ступенчатая), либо сигмоида.

Пусть гиперболическая СИФ задана системой $\{(X, d); f_1, f_2, \ldots, f_q\}$. Пусть A_F — аттрактор СИФ. Существует два способа приближенного построения этого фрактала. Первый из них начинается с выбора одной или нескольких начальных точек x_0 и вычисления их дискретных орбит длиной N. Пусть, как и раньше, $\sigma = (\sigma_1, \ldots, \sigma_N)$; $\sigma_n \in \{1, \ldots, q\}$. Тогда при достаточно большом N множество:

$$A(x_0, N) = \bigcup_{\sigma} f_{\sigma_N} \dots f_{\sigma_1}(x_0)$$

является хорошим приближением A_F. Этот способ можно оптимизировать, если выбрать в качестве x_0 неподвижную точку одного из f_k так, что $x_0 \in A_F$. Заметим, что вычисление N орбит на один шаг вперед не эквивалентно вычислению одной орбиты на N шагов. Другой способ заключается в определении траектории начального компактного множества $A_0 \in H(X, d)$:

$$A_{n+1} = F(A_n) \equiv \bigcup_{i}^{N} f_i(A_n); \; n = 0, \dots, \infty.$$

Последовательность A_n сходится к A_F в метрике Хаусдорфа, так что при достаточно больших N, $F^{\circ N}(A_0) \approx A_F$. Рис. 10 иллюстрирует этот подход. Коллаж состоит из трех начальных множеств, которые эволюционируют под действием F и приводят к аппроксимации фрактала, показанной на рис. 11.

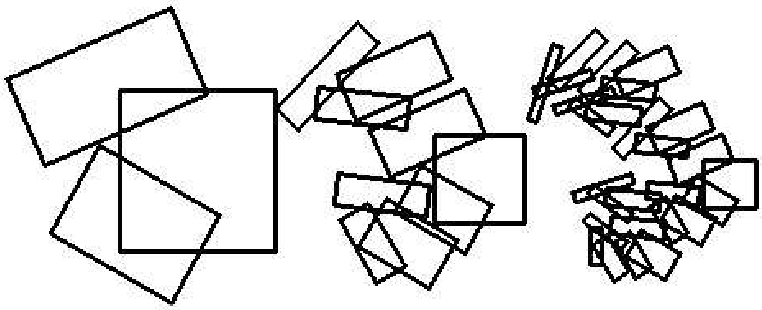

Рис. 10. Первые три итерации коллажа

Предположим, что мы собираемся построить на экране компьютера аттрактор СИФ. Будем считать, что экран состоит из квадратных пикселов и горизонтальное и вертикальное разрешения равны.

Пусть $X \equiv S^u \subset \mathbb{R}^2$, S^u — единичный квадрат с разбиением, соответствующим выбранному разрешению. Чтобы упростить обозначения,

Рис. 11. Аппроксимация фрактала, порожденного эволюцией коллажа

мы будем использовать значок s вместо координат пиксела x_{ij}. Снабдим каждое A_n индикаторной функцией (нейроном) $y : S^u \to \{0,1\}$ такой, что для каждого пиксела $s \in S^u$:

$$y_s(n) = \begin{cases} 1\,, & s \in A_n \\ 0\,, & s \notin A_n \end{cases}$$

Определим «веса»

$$\omega_{ss'} = \begin{cases} 1\,, & \text{если } f_i(s') = s \text{ для некоторого } i \\ 0\,, & \text{в противном случае} \end{cases}$$

и зададим динамику $y_s(n)$ как

$$y_s(n+1) = g\Big(\sum_{s' \in S} \omega_{ss'} y_{s'}(n)\Big),$$

где g — ступенчатая функция.

Очевидно, что по виду и смыслу полученное выражение определяет бинарную нейронную сеть с $|S^u|$ нейронами y_s и синаптическими весами $\omega_{ss'}$, которая реализует СИФ. Заметим, что каждый нейрон такой сети имеет только бинарные веса и максимальное количество нейронов, с которыми он связан, равно q. Несмотря на то, что число весов может быть очень большим, подавляющее большинство из них будут нулями. Так что результирующая сеть будет пространственно разреженной.

Детерминированная СИФ и эквивалентная ей нейронная сеть позволяют получать лишь черно-белые изображения аттракторов. Для генерации полутонов следует использовать ССИФ. Поскольку принципы построения соответствующей нейросети остаются прежними, мы опускаем здесь этот вариант.

Дальнейшее обобщение этих результатов приводит к рассмотрению случайной итеративной нейронной сети, которая в определенном смысле эквивалентна дискретной динамической системе. Представим себе формальный нейрон, который устроен следующим образом. На каждый его вход подается вектор $\mathbf{u}_i \in R^m$. Весами служат множество $\mathbf{W}$ векторов $\mathbf{w}_j \in R^k$, функцией активации $T_{\sigma,\mathbf{W}}$ является ССИФ, а выходом вектор $\mathbf{y} \in R^m$, так что $\mathbf{y} = T_{\sigma,\mathbf{W}}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$, где множество σ нумерует выбор отображений, входящих в T с соответствующими вероятностями, так же как и в разделе «Фракталы и системы итеративных функций» (см. с. 139–144). Такое устройство называют вероятностным *гипернейроном. Случайная итеративная нейронная сеть* (СИНС) состоит из трех гипернейронов и имеет архитектуру обычной рекуррентной нейронной сети, которая реализует аттрактор ССИФ. Динамику СИНС можно описать парой (X, A_f), где X — пространство состояний СИНС и A_f ее глобальный аттрактор[15]. Пусть (Y, A_g) — дискретная динамическая система с аттрактором A_g. В начале 2000 года Фернандо Ниньо (The University of Memphis) показал в своей диссертации, что для любой заданной пары (Y, A_g) существует (X, A_f) такая, что хаусдорфово расстояние $h(A_f, A_g)$ между аттракторами динамической системы и СИНС можно сделать сколь угодно малым. Иными словами, СИНС является аппроксиматором дискретной динамической системы с заданным аттрактором! Необходимый для аппроксимации набор ССИФ можно най-

[15]Такой аттрактор всегда существует в силу упомянутой в разделе «Фракталы и системы итеративных функций» теоремы Элтона.

ти генетическим алгоритмом. К сожалению, подробное обсуждение этой интересной проблемы выходит за рамки лекции.

Путеводитель по литературе. Автору известны всего три работы [37–39], специально посвященные связи СИФ и нейронных сетей. Описание гипернейрона можно найти в диссертации Ф. Ниньо [40].

Глоссарий

«Запятирующее тире» — отражает кумулятивное диффундирование смыслокачества в новизну подтекстовых проводок в ситуации резкого спада количества определяющих указателей.

Павел Таранов
«Маневры общения»

Этот раздел содержит предварительные сведения, т. е. определения и понятия, чуть более содержательные, чем просто определения. Они собраны из разных областей математики с целью сделать чтение ***Лекций*** по возможности независимым от сопутствующей литературы. Мы будем использовать ниже следующую символику:

- Запись $x \in \mathfrak{X}$ означает, что элемент x принадлежит множеству $\mathfrak{X}$.
- Запись $\forall\, x$ читается: для любых x (для всех x), запись $\exists\, y$ — существует y, а запись $\forall x \,\exists\, y\,|$ означает: для любого x существует y, такой что...
- Запись $\mathfrak{X} = \{x|A\}$ обозначает множество таких элементов x, для которых справедливо условие A.
- Запись $A \subset B$ ($A \subseteq B$) означает, что A содержится в B (содержится в B или совпадает с ним, соответственно).
- Наконец, $A \Rightarrow B$ читается: A влечет B.

Множество возникает благодаря объединению отдельных вещей в одно целое. Для того, чтобы избежать парадоксов, полагают, что элементы $a, b, c, \ldots$ особым *не подлежащим определению образом* порождают множество $\mathfrak{X}$. Иными словами $\mathfrak{X}$ задано, если известно, входит туда элемент или нет$^{(1)}$. Природа самих элементов безразлична: ими могут быть, например, Луна и прошлогодний снег.

Объединением множеств $\mathfrak{A}$ и $\mathfrak{B}$ называют множество $\mathfrak{C} = \mathfrak{A} \cup \mathfrak{B}$, состоящее из элементов $c \in \mathfrak{C}$, каждый из которых принадлежит $\mathfrak{A}$ или (и) $\mathfrak{B}$.

Пересечением $\mathfrak{A}$ и $\mathfrak{B}$ называют множество $\mathfrak{C} = \mathfrak{A} \cap \mathfrak{B}$, каждый элемент которого c входит и в $\mathfrak{A}$, и в $\mathfrak{B}$. Эти определения легко обобщаются. Для произвольного множества индексов $\mathfrak{I} = \{i\}$ каждому i поставим в соответствие множество $\mathfrak{A}_i$. Тогда объединение

$$\mathfrak{C} = \cup\mathfrak{A}_i = \{c \in \mathfrak{C} | \exists i \in \mathfrak{I} \Rightarrow c \in \mathfrak{A}_i\}.$$

Аналогично определяется пересечение:

$$\mathfrak{C} = \cap\mathfrak{A}_i = \{c \in \mathfrak{C} | \forall i \in \mathfrak{I} \Rightarrow c \in \mathfrak{A}_i\}.$$

Пустое множество обозначается $\varnothing$.

Семейство множеств $\mathfrak{A} \equiv \{A_i\}$ называют ***покрытием*** множества $\mathfrak{B}$, если $\mathfrak{B} \subseteq \cup\{A_i | A_i \in \mathfrak{A}\}$, т. е. каждая точка $\mathfrak{B}$ принадлежит некоторому элементу $\mathfrak{A}$. Частный случай покрытия, когда $A_i \cap A_j = \varnothing$, $\forall i, j$ называют ***разбиением*** $\mathfrak{B}$.

Два множества $\mathfrak{A}$ и $\mathfrak{B}$ можно поставить во взаимно однозначное соответствие, если $\forall a \in \mathfrak{A}\ \exists b \in \mathfrak{B}$ и наоборот. Так определяется ***отношение эквивалентности*** $\mathfrak{A} \equiv \mathfrak{B}$; говорят еще, что $\mathfrak{A}$ и $\mathfrak{B}$ имеют одинаковую мощность. Если $\mathfrak{A} \equiv \{1, 2, \ldots\}$ то $\mathfrak{A}$ называют ***счетно-бесконечным***.

Для описания мощности множества служат ***кардинальные числа***. Если множество конечно, это просто число элементов в нем. В случае бесконечного множества $\mathfrak{A}$, кардинальное число определяют присвоением имен кардинальным числам конкретных множеств, с которыми можно сравнить $\mathfrak{A}$. Кардинальное число счетно-бесконечного множества обозначают символом *алеф-нуль* — $\aleph_0$. Примером ***несчетного*** множества являются вещественные числа интервала $[0, 1]$. Его кардинальное число $|c|$ называют ***мощностью континуума***. Ясно, что мощность всей вещественной оси тоже $|c|$; в этом легко убедиться с помощью преобразования:

$x \to 1/2(1 + \operatorname{th} x)$. Кантор показал, что точки квадрата и точки отрезка эквивалентны, т. е. мощность $\mathfrak{R}^2$ равна $|c|$.

Соответствие между различными множествами удобно описывать функциями или отображениями: $f(*)$ или $f : \mathfrak{A} \to \mathfrak{Y}$. Если $A \in \mathfrak{X}$, то и $f[A] = \{f(x)|x \in A\}$. Аналогично, если $B \in \mathfrak{Y}$, то $f^{-1}[B] = \{x|f(x) \in B\}$ — подмножество в $\mathfrak{X}$. Подмножество $\{f(x)\} \in \mathfrak{Y}$ — это ***область значений*** функции, а $\{x\} \in \mathfrak{X}$ — ее ***область определения***. Функциональные отношения между множествами имеют определенную классификацию:

- Если образ всего множества $\mathfrak{A}$ совпадает с $\mathfrak{B}$ или, что тоже самое, каждый элемент из $\mathfrak{B}$ является образом по крайней мере одного элемента из $\mathfrak{A}$, то говорят, что f — ***сюръекция*** или отображение **на**.
- Функцию $f : \mathfrak{A} \to \mathfrak{B}$ называют ***инъективной*** или взаимно-однозначной если для $\forall y \in \mathfrak{B}$ существует не более одного элемента $x \in \mathfrak{A}$ такого, что $y = f(x)$. Очевидно, что в этом случае $f(x) = f(y) \Rightarrow x = y$.
- ***Биективное*** отображение или ***биекция*** — это функция, которая является одновременно сюръективной и инъективной.

Если $\mathfrak{A}$ и $\mathfrak{B}$ совпадают, то $f : \mathfrak{A} \to \mathfrak{A}$ и элемент x, удовлетворяющий условию $x = f(x)$, называется ***неподвижной точкой*** отображения f. Например, каждое непрерывное отображение $f : [a, b] \to [a, b]$ отрезка в себя имеет хотя бы одну неподвижную точку$^{(2)}$.

Абстрактное множество $\mathfrak{G}$ называют ***группой***, если:

1. Для любой пары его элементов g_1, g_2 определено произведение или бинарная операция $g_1 * g_2 \in \mathfrak{G}$, ассоциативная, но не обязательно коммутативная.
2. Существует единица группы $e : g * e = e * g = g, \ \forall g \in \mathfrak{G}$.
3. Для любого $g \in \mathfrak{G}$ существует $g^{-1} : g * g^{-1} = e$.

Рассмотрим группу преобразований плоскости $(x, y) \to (x^*, y^*)$, заданную в матричном виде: $\mathbf{x}^* = \mathbf{Ax}+\mathbf{B}$. Если $\mathbf{A}$ — ортогональная матрица (т.е. $\det \mathbf{A} = 1$), то обычную геометрию можно определить как совокупность свойств, инвариантных относительно действия нашей группы. Координаты в $\mathfrak{R}^n$, например, определены лишь с точностью до ортогональных преобразований. С другой стороны, в ***арифметическом*** пространстве X^n точка x определяется *абсолютно* набором n чисел $(x^1, \ldots, x^n)$ с точностью до *тождественных* преобразований.

Аффинная геометрия удовлетворяет более скромному требованию $\det \mathbf{A} \neq 0$. Можно задать группу в общем виде: $x^* = f(x, y); y^* = g(x, y)$, где f и g — нелинейные C^N–функции, т. е. функции, обладающие непрерывными производными до N-го порядка. Если такими же свойствами обладают обратные преобразования (а их существование гарантирует группа), то мы получим геометрию ***гладких многообразий***. Такое многообразие определяет объект с точностью до произвольных гладких деформаций. Вырожденный случай C^0 — это уже топология.

Метрическое пространство $(\mathbf{M}, d)$ — это множество $\mathbf{M}$ вместе с вещественнозначной функцией d, определенной на декартовом произведении $\mathbf{M} \times \mathbf{M}$, называемой метрикой на $\mathbf{M}$ и удовлетворяющей следующим требованиям:

- $d(x, y) \geqslant 0;\ d(x, y) = 0 \Rightarrow x = y$
- $d(x, y) = d(y, x)$
- $d(x, z) \leqslant d(x, y) + d(y, z)$

Для $\mathbf{M} \equiv \Re^n$ обычная евклидова метрика:

$$d(x, y) = \Big[\sum_i (x_i - y_i)^2\Big]^{1/2}$$

легко обобщается до $\boldsymbol{L}_\mathbf{p}$ метрики Минковского:

$$d_p(x, y) = \Big[\sum_i (x_i - y_i)^p\Big]^{1/p}$$

Располагая метрикой, нетрудно ввести понятие *сходимости* по аналогии с тем, что было в X^n. Говорят, что последовательность точек $\{x_n\}$, $n = 1, 2, \ldots$, метрического пространства $(\mathbf{M}, d)$ сходится к $x \in \mathbf{M}$, если $d(x, x_n) \to 0$, при $n \to \infty$. Различные метрики порождают разные сходимости.

Метрика d_1 эквивалентна d_2 в $\mathbf{M}$, если существуют две постоянные $0 < c_1 < c_2 < \infty$ такие, что:

$$c_1 d_1(x,y) \leqslant d_2(x,y) \leqslant c_2 d_1(x,y),\ \forall (x,y) \in \mathbf{M} \times \mathbf{M}$$

Последовательность $\{x_n\} \in (\mathbf{M}, d)$ называют ***последовательностью Коши***, если $\forall \varepsilon > 0\ \exists N \,|\, n, m > N \Rightarrow d(x_n, x_m) < \varepsilon$. Заметим, что эта последовательность не обязательно сходится к определенному пределу. Пусть $(\mathbf{M}, d)$ — метрическое пространство и $\{x_n\}$ — последовательность Коши, сходящаяся к точке $x \in \mathbf{M}$. Пусть также $f : \mathbf{M} \to \mathbf{M}$ — непрерывная функция. Тогда:

$$\lim_{n\to\infty} f(x_n) = f(x).$$

Пространство $(\mathbf{M}, d)$, в котором любая последовательность Коши сходится, называется ***полным***.

Пусть $(\mathbf{M}, d)$ — метрическое пространство и $r > 0$ — действительное число. Тогда ***открытым*** шаром в точке $a \in \mathbf{M}$ называется подмножество $B(a,r) \subset \mathbf{M} :\ B(a,r) = \{x \in \mathbf{M} \,|\, d(a,r) < r\}$. Шар называется ***замкнутым***, если $d(a,r) \leqslant r$.

Открытым множеством в $(\mathbf{M}, d)$ называется подмножество $A \subset \mathbf{M}$, для любой точки $x \in A$ которого существует такое $r > 0,\ B(x,r) \subset A$. Пустое множество $\varnothing$ и само пространство $\mathbf{M}$ — открыты.

Замкнутым множеством в $\mathbf{M}$ называют дополнение открытого множества; $\varnothing$ и $\mathbf{M}$ объявляются замкнутыми. Метрическое пространство $\mathbf{M}$ называют ***компактным***, если оно удовлетворяет *аксиоме Лебега—Бореля*: из любого покрытия $\mathbf{M}$ открытыми множествами можно выделить *конечное* подпокрытие.

$(\mathbf{M}, d)$ называют ***вполне ограниченным***, если $\forall \varepsilon > 0$ существует конечное покрытие пространства $\mathbf{M}$ шарами диаметром $< \varepsilon$.

Понятия компактности и ограниченности заменяют понятие *конечности* в чистой теории множеств. Компактным (вполне ограниченным) множеством в $(\mathbf{M}, d)$ называют такое подмножество $A \subset \mathbf{M}$, для которого подпространство A компактно (вполне ограничено). Наконец, для $(\mathbf{M}, d)$ следующие три условия эквивалентны: $\mathbf{M}$ — компактно, или любая бесконечная последовательность в $\mathbf{M}$ имеет по крайней мере одну предельную точку, или $\mathbf{M}$ — полное и вполне ограниченное.

Пусть $H(\mathbf{M})$ — пространство, «точками» которого являются компактные множества из $(\mathbf{M}, d)$. Пусть $A, B \in H$. Определим метрику h в H как

$$h(A,B) = \max\{\max_{x\in A}\min_{y\in B} d(x,y),\ \max_{y\in A}\min_{x\in B} d(x,y)\}$$

Величина $h(A,B)$ называется ***расстоянием Хаусдорфа*** и удовлетворяет всем аксиомам метрики. Можно показать, что $(H(\mathbf{M}), h)$ является полным метрическим пространством. В тексте Лекции дано эквивалентное определение этой метрики.

Пусть $(\mathbf{M}, d)$ — метрическое пространство и $s > 0$. Отображение $f : \mathbf{M} \to \mathbf{M}$ называют ***подобием*** с масштабным коэффициентом s, если $d(f(x), f(y)) = sd(x,y)$, $\forall x, y \in \mathbf{M}$ и ***сжимающим*** отображением с коэффициентом сжатия s, если существует такое число $0 \leqslant s < 1$, что $d(f(x), f(y)) < sd(x,y)$, $\forall x, y \in \mathbf{M}$. Иными словами, f — сжатие, если расстояние между образами двух произвольных точек в s раз меньше исходного. Определим для f множество итераций вперед $f^{\circ n} : \mathbf{M} \to \mathbf{M}$, $n = 1, 2, \ldots$ соотношениями:

$$f^{\circ 0}(x) = x;\ f^{\circ 1}(x) = f(x);\ f^{\circ n+1}(x) = f * f^{\circ n}(x) = f(f^{\circ n}(x)).$$

Если f — обратима, аналогично определяются итерации назад:

$$f^{-\circ n}(x) = (f^{\circ n})^{-1}(x).$$

Пусть $f : \mathbf{M} \to \mathbf{M}$ — сжатие в полном метрическом пространстве $(\mathbf{M}, d)$. Тогда справедлива ***Теорема о сжимающем отображении***: f имеет единственную[3] неподвижную точку $f(x) = x_f \in \mathbf{M}$ и $\forall y \in \mathbf{M}$ последовательность $f^{\circ n}(y)$ сходится к

$$x_f : \lim_{n\to\infty} f^{\circ n}(y) = x_f.$$

Множество $B \subset \mathbf{M}$ называют ***плотным*** в $\mathbf{M}$, если любое $x \in \mathbf{M}$ является пределом последовательности элементов из B.

Пусть X — множество точек x. Систему $\mathfrak{B}$ подмножеств X называют **σ-алгеброй**, если

1. $\varnothing, X \in \mathfrak{B}$
2. $A \in \mathfrak{B} \Rightarrow X - A \in \mathfrak{B}$

3. $A_n \in \mathfrak{B},\ n = 1, 2, \ldots \Rightarrow \bigcup A_n \in \mathfrak{B},\ \bigcap A_n \in \mathfrak{B}$

Наименьшая из σ-алгебр в метрическом пространстве $\mathbb{R}^n$ называется ***борелевой σ-алгеброй.***

Пара $(X, \mathfrak{B})$ называется измеримым пространством. Пусть $(X, \mathfrak{B})$ — измеримое пространство. Вещественная функция $\mu = \mu(A)$, $A \in \mathfrak{B}$, принимающая значения из интервала $[0, \infty]$, называется **мерой**, если

- $\mu(\varnothing) = 0$
- $\mu(A) \geqslant 0$, $\forall A \in \mathfrak{B}$
- если $\{A_n\}_{n=1}^{\infty}$ — непересекающиеся множества из $\mathfrak{B}$, то $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$

Тройка $(X, \mathfrak{B}, \mu)$ называется **пространством с мерой**. Например, *лебегова мера* на вещественной оси — это просто длина интервала вида (a, b), $[a, b)$, $(a, b]$, $[a, b]$. Мера Лебега не является универсальной: существуют множества, не являющиеся μ-измеримыми. Тем не менее, μ-мера является эталоном не только во многих разделах математики, но и в эргодической теории гладких динамических систем.

Множество $\mathfrak{S}$ ***имеет меру нуль***, если оно может быть покрыто системой интервалов, сумма длин которых произвольно мала. Иными словами, такое $\mathfrak{S}$ можно вложить в борелево множество $\mathfrak{B}$ мерой меньшей чем $\varepsilon, \forall \varepsilon > 0$. Если $\mathfrak{S}$ можно получить из борелева множества $\mathfrak{B}$ добавлением или выбрасыванием множества точек нулевой меры, то $\mu(\mathfrak{S}) = \mu(\mathfrak{B})$. Такие множества называют μ-измеримыми. Выражение «почти для всех x» понимают в смысле *для всех x, за исключением множества меры нуль*.

Если $\mu(X) = 1$, μ называют вероятностной мерой, а $(X, \mathfrak{B}, \mu)$ — **вероятностным пространством**.

Пусть $(X, \mathfrak{B}, \mu)$ — вероятностное пространство. Преобразование $T : X \to X$ называется **измеримым**, если для любого $A \in \mathfrak{B}$, $T^{-1}A = \{x : Tx \in A\} \in \mathfrak{B}$.

Говорят, что измеримое преобразование T **сохраняет меру** μ, если для каждого $A \in \mathfrak{B}$, $\mu(T^{-1}A) = \mu(A)$. Меру μ называют в этом случае T**-инвариантной**.

Четверка $(X, \mathfrak{B}, \mu, T)$, где T — преобразование, сохраняющее меру, называется **динамической системой**.

Примечания

1. При этом абсолютно неясен способ, посредством которого мы могли бы выяснить, входит элемент a в A или нет. Например, если A задано запахом, а B — вкусом, можем ли мы решить эквивалентны они или нет? Приведем пример одного парадокса Рассела. Пусть A множество всех натуральных чисел. Предположим, что каждое из этих чисел можно определить фразой, содержащей менее 20 русских слов. Содержит ли A весь натуральный ряд? Считая, что русский язык содержит не более чем n слов и, следовательно, не более чем n^{20} нужных нам фраз, можно полагать, что A конечно. Но тогда определим *наименьшее натуральное число, не входящее в множество* A. Это число по определению не входит в A, но должно входить в него, поскольку определяется фразой, не превышающей 20 слов. Очевидно, A содержит весь натуральный ряд, а парадокс возникает потому, что мы дополнили русский язык фразой, выделенной курсивом.

2. Определим новую функцию: $g(x) = f(x) - x$ для каждой $x \in [a, b]$. Она непрерывна и $g(a) > 0$, $g(b) < 0$, если концы интервала не являются неподвижными точками. Но если непрерывная функция меняет знак, найдется хотя бы одна точка x в которой $g(x) = 0 \Rightarrow f(x) = x$.

3. Действительно, пусть $f(x)$ имеет две неподвижных точки: $x^\star, y^\star$. Тогда, $d(f(x^\star), f(y^\star)) \leqslant sd(x^\star, y^\star)$. Но тогда $d(f(x^\star), f(y^\star)) = d(x^\star, y^\star) \leqslant sd(x^\star, y^\star)$. Следовательно, $d(x^\star, y^\star) = 0$.

Путеводитель по литературе. Основные понятия теории множеств можно найти в [33–35]. Небольшая книжка [41] содержит простое введение в теорию меры, история развития которой изложено в работе Лебега [3]. Отличное введение в теорию гладких многообразий содержит учебник [5]. Метрика Хаусдорфа в пространстве компактов и ее связь с фракталами описана в [2]. Сводку многих приведенных определений можно найти в справочнике [36].

Summa sine laude[16]

Разумеется, все что написано выше можно было бы изложить гораздо лучше, но ведь можно было и хуже. . .

Литература

1. *Пуанкаре А.* О науке. – М.: Наука, 1983. – 560 с.
2. *Barnsly M.* Fractals everywhere. – Academic Press, 1988. – 394 pp.
3. *Лебег А.* Об измерении величин. – М.: ГУПИМП, 1960. – 204 с.
4. *Стинрод Н., Чинн У.* Первые понятия топологии. – М.: Мир, 1967. – 224 с.
5. *Chillingworth D.* Differential topology with a view to applications. – Pitman Press, 1976. – 291 pp.
6. *Горелик Г. Е.* Размерность пространства. – М.: Изд-во МГУ, 1983. – 216 с.
7. *Зельдович Я. Б., Соколов Д. Д.* Фракталы, подобие, промежуточная асимптотика // УФН. – 1985. – т. 146. – с. 493–506.
8. *Perdang J.* Astrophysical fractals: An overview and prospects // Vistas in Astronomy. – 1990. – v. 33. – pp. 249–294.
9. *Соколов И. М.* Размерности и другие геометрические критические показатели в теории протекания // УФН. – 1986. – v. 150. – pp. 221–253.
10. *Голдман С.* Теория информации. – М.: Мир, 1967.
11. *Реньи А.* Трилогия о математике. – М.: Мир, 1980. – 376 с.
12. *Федер Е.* Фракталы. – М.: Мир, 1991. – 260 с.
13. *Баренблатт Г. И.* Подобие, автомодельность, промежуточная асимптотика. – Л.: Гидрометеоиздат, 1982. – 255 с.
14. *Мандельброт Б.* Общие свойства фракталов // В сб.: Фракталы в физике. – М.: Мир, 1988. – с. 9–47.
15. *Mandelbrot B.* Self-affine fractals and fractal dimension // Physica Scripta. – 1985. – v. 32. – pp. 257–260.
16. *Рихтмайер Р.* Принципы современной математической физики. т. 2. – М.: Мир, 1982. – 486 с.
17. *Иванов Л. Д.* Вариации множеств и функций. – М.: Наука, 1975.

[16] Итог без похвал (лат.)

18. *Янг Л.* Лекции по вариационному исчислению и теории оптимального управления. – М.: Мир, 1974. – 488 с.

19. *Pfeifer P., Obert M.* Fractals: Basic concepts and terminology // Chapter 1.2 in: Fractal Approach to Heterogeneous Chemistry: Surfaces, Colloids, Polymers / Ed.: David Avnir. – John Wiley & Sons Ltd., Chichester. – 1989. – p. 11.–40.

20. *Godreche C., Luck J. M.* Multifractal analysis in reciprocal space and the nature of the Fourier transform of self-similar structures // J. Physics. A.: Math. Gen. – 1990. – v. 23. – pp. 3769–3797.

21. *Hatchinson J.* Fractals and self-similarity // Indiana Univ. J. Math. – 1981. – v. 30. – pp. 713–747.

22. *Hatchinson J.* Fractals: A mathematical framework. – Department of Mathematics, School of Mathematical Sciences, Australian National University, Dec. 1996.
URL: http://www.csu.edu.au/ci/vol02/jeh2frac/jeh2frac.html

23. *Кроновер Р. М.* Фракталы и хаос в динамических системах. – М.: Постмаркет, 2000. – 350 с.

24. *Bourke P.* An introduction to fractals. – May 1991.
URL: http://astronomy.swin.edu.au/pbourke/fractals/fracintro
Paul Bourke's Home Page:
URL: http://astronomy.swin.edu.au/pbourke/

25. *Малинецкий Г. Г., Потапов А. Б.* Современные проблемы нелинейной динамики. – М.: Эдиториал, 2000. – 335 с.

26. *Лоскутов А. Ю., Михайлов А. С.* Введение в синергетику. – М.: Наука, 1990. – 272 с.

27. *Eckmann J. P., Ruelle D.* Ergodic theory of chaos and strange attractors // Rev. Mod. Phys. – 1985. – v. 57. – pp. 617–656.

28. *Schaw R.* Strange attractors, chaotic behavior, and information flow // Z. Naturforsch. – 1981. – v. 36a. – pp. 80–112.

29. *Milnor J.* On the concept of attractor // Commun. Math. Phys. – 1985. – v. 99. – pp. 177–195.

30. *Sprott J.* Strange attractors: Creating patterns in chaos. – New York: M & T Books, 1993. – 591 pp.
URL: http://sprott.physics.wisc.edu/sa.htm
http://sprott.physics.wisc.edu/fractals/booktext/sabook.pdf
Sprott's Home Page:
URL: http://sprott.physics.wisc.edu/sprott.htm

31. *Dynamical Systems* // In: Topics in Mathematics. Mathematical Archives, 1996–2001.
URL: http://archives.math.utk.edu/topics/dynamicalSystems.html
Surveys in dynamical systems available on-line:
URL: http://www.math.sunysb.edu/dynamics/surveys.html

32. *Young Lai-Sang*. Ergodic theory of chaotic dynamical systems. – Univ. of California, Department of Mathematics, Los Angeles, CA. – Oct. 1997. – 14 pp.
URL: http://www.math.ucla.edu/∼lsy/expository.html
Prof. Lai-Sang Young's Home Page:
URL: http://www.math.ucla.edu/∼lsy/

33. *Александров П. С.* Введение в теорию множеств и общую топологию. – М.: Наука, 1977. – 368 с.

34. *Келли Дж. Л.* Общая топология. – М.: Наука, 1968. – 432 с.

35. *Колмогоров А. Н., Фомин С. В* Элементы теории функций и функционального анализа. – М.: Наука, 1989. – 496 с.

36. *Фор Р., Кофман А., Дени-Папен М.* Современная математика. – М.: Мир, 1966. – 271 с.

37. *Severyanov V. M.* Automata network dynamical systems for construction of fractal objects // Electronic Proceedings of IMACS ACA'98 — the 4th International IMACS Conference on Applications of Computer Algebra. Czech Technical University, Prague, Czech Republic, August 9–11, 1998. – 8 pp. URL: http://www.math.unm.edu/ACA/1998/sessions/dynamical/sever

38. *Stark J.* Iterated function systems as neural networks // Neural Networks. – 1991. – v. 4. – pp. 679–690.

39. *Bressloff P. C., Stark J.* Neural networks, learning automata and iterated function systems // In: Fractals and Chaos / Eds.: A. J. Crilly et al. – 1991. – Springer-Verlag. – pp. 145–164.

40. *Niño F.* Random iterated neural networks: Properties, evolutionary design and applications. – Doctoral Dissertation, University of Memphis, May 2000. – 103 pp.
URL: http://www.msci.memphis.edu/∼ninol/diss_fn.ps
Fernando Niño's Home Page:
URL: http://www.msci.memphis.edu/∼ninol/index.html
URL: http://www.msci.memphis.edu/∼ninol/research.html

41. *Брудно А. Л.* Теория функций действительного переменного. – М.: Наука, 1971. – 119 с.

Николай Григорьевич Макаренко, ведущий научный сотрудник, кандидат физико-математических наук, руководитель группы в Лаборатории компьютерного моделирования (Институт математики, Алма-Ата, Казахстан). Область научных интересов: фрактальная геометрия, вычислительная топология, алгоритмическое моделирование, детерминированный хаос, нейронные сети, физика Солнца. Имеет более 50 научных публикаций.