

А. Ю. Лоскутов, А. С. Михайлов

ОСНОВЫ ТЕОРИИ СЛОЖНЫХ СИСТЕМ



Москва ♦ Ижевск

2007

УДК 530.1(075.8)
ББК 22.31
Л79

Интернет-магазин
MATHESIS

<http://shop.rcd.ru>

- физика
- математика
- биология
- нефтегазовые технологии

А. Ю. Лоскутов, А. С. Михайлов

Основы теории сложных систем. — М.-Ижевск: Институт компьютерных исследований, 2007. — 620 с.

Систематически изложена теория кооперативного поведения сильно неравновесных физических, химических, биологических и других нелинейных систем. Подробно исследованы свойства автоволн и диссипативных структур в активных средах, а также явление детерминированного хаоса. Большое внимание уделено описанию фрактальных множеств, отображений и пространственно-временной динамики. Рассмотрены методы аналоговой обработки информации с помощью распределенных активных сред и нейроноподобных сетей. Книга хорошо иллюстрирована, в ней содержится много поясняющих примеров.

В настоящее, второе, издание (1-е изд. — «Введение в синергетику», 1990 г.) вошли новые разделы, относящиеся к природе хаоса, фрактальной геометрии, управлению нелинейными хаотическими системами и подавлению хаоса, решеткам сцепленных отображений, анализу временных рядов и некоторым другим направлениям современной нелинейной динамики. Вследствие этого новое издание получило другое название, более полно отражающее содержание книги.

Для студентов, аспирантов и преподавателей, специалистов в области физики нелинейных систем, биологической и химической физики, физической информатики, а также всех, кто интересуется современными проблемами динамического хаоса.

Рекомендовано УМО по классическому университетскому образованию РФ в качестве учебного пособия для студентов высших учебных заведений, обучающихся по специальностям 010701.65 — Физика и 010710.65 — Физика открытых нелинейных систем.

ISBN 5-93972-???-?

**ББК 22.31
Л79**

© А. Ю. Лоскутов, А. С. Михайлов, 2007
© Институт компьютерных исследований, 2007

<http://rcd.ru>
<http://ics.org.ru>

Оглавление

Предисловие ко второму изданию	5
Введение	12
ГЛАВА 1. Волны и структуры в активных средах	17
1. Основные типы активных сред	17
2. Волны переключения и заселения в простых бистабильных средах	21
3. Общие свойства структур в простых бистабильных средах	32
4. Бегущие импульсы в возбудимых средах	39
5. Процессы в возбудимых средах, образованных клеточными автоматами	47
6. Спиральные волны в распределенных возбудимых средах	58
7. Кинематика автоволновых фронтов в возбудимых средах	66
8. Резонанс и дрейф спиральных волн	74
9. Автоволновые структуры в трехмерных возбудимых средах	81
10. Фазовая динамика в осциллирующих активных средах	90
11. Фазовые волны и пейсмекеры	96
12. Спиральные волны в автоколебательных активных средах	101
13. Стационарные диссипативные структуры	108
ГЛАВА 2. Динамический хаос	116
14. Гамильтоновы системы	116
15. Нелинейный резонанс	133
16. Элементы теории Колмогорова–Арнольда–Мозера (теории КАМ). Диффузия Арнольда	146
17. Природа хаоса	152
18. Основные свойства хаотических систем: эргодичность, перемешивание, расщепление корреляций	166
19. Бильярды. Газ Лоренца	177
20. Диссипативные динамические системы	208
21. Критерии динамического хаоса	234
22. Размерность странных аттракторов	258

23.	Фракталы	268
24.	Отображения и некоторые их свойства	280
25.	Хаос в одномерных отображениях	290
26.	Универсальность Фейгенбаума	297
27.	Отображения комплексной плоскости. Красота фракталов	308
28.	Бифуркации в динамических системах	329
29.	Типичные сценарии перехода к хаосу	346
30.	Подавление хаоса и управление динамическими системами	366
31.	Пространственно-временной хаос	383
32.	Динамика систем сцепленных отображений	393
33.	Временные ряды: анализ и прогноз	428
ГЛАВА 3. Обработка информации распределенными динамическими системами		
34.	Мозг и компьютер	462
35.	Клеточные автоматы	475
36.	Ассоциативная память	488
37.	Сложные задачи комбинаторной оптимизации	506
38.	Обучающиеся системы	517
39.	Эволюционные модели	533
Библиографический комментарий		543
Литература		555
Предметный указатель		613

Предисловие ко второму изданию

Со времени выхода первого издания этой книги минуло уже пятнадцать лет. За это время произошли большие перемены как во всем мире, так и в жизни каждого из нас. В последние годы один из авторов, А. С. Михайлов, живет в Берлине и работает в качестве заведующего лабораторией сложных систем в Институте им. Ф. Габера. Сегодня его научные интересы в первую очередь принадлежат химической физике и молекулярной биологии.

Синергетика, более известная теперь под названием *теории сложных систем*, прошла за пятнадцать лет бурное развитие от молодой перспективной отрасли к признанному разделу современной науки. В настоящее время выходят специальные научные журналы и организуются многочисленные научные конференции, посвященные этой тематике. Библиография работ по теории сложных систем уже насчитывает многие тысячи публикаций.

Для второго издания мы приняли решение не обновлять материал первой и третьей глав, которые были целиком написаны А. С. Михайловым. В основном это связано с огромным объемом материала, который следовало бы включить в книгу, чтобы вывести ее на современный уровень. На наш взгляд, даже без переработки она может служить хорошим пособием для первоначального знакомства с предметом. Цель настоящего предисловия — кратко осветить развитие теории сложных систем за время, прошедшее с момента выхода первого издания, и дать рекомендации для дальнейшего чтения.

Вместе с соавторами за истекшие годы нами было опубликовано несколько книг, продолжающих затронутую здесь тематику. В первую очередь, следует отметить два тома «Основ синергетики», которые вышли на английском языке двумя изданиями в немецком издательстве «Шпрингер» [1; 2]. Они пользуются популярностью и служат в качестве учебников во многих университетах по всему миру. Недавняя книга «От клеток к обществам: модели сложного когерентного действия» [3] продолжает этот курс и посвящена вопросам внутренней самоорганизации сложных биологических и социальных систем. Наконец, новая работа «Становление динамического порядка: явления синхронизации в сложных системах» [4] дает

подробный обзор эффектов самосинхронизации в больших ансамблях взаимодействующих периодических или хаотических осцилляторов и включает приложения к химическим и биологическим системам, а также к нейронным сетям и мозгу.

Перейдем теперь к более детальному обзору вопросов, затронутых в настоящей книге. Основной материал первой главы, посвященной теории волн и структур в активных средах, сохраняет свою актуальность. В первую очередь, дальнейшее развитие в этой области было связано с появлением мощных компьютеров, которые позволили проводить в широком масштабе численное интегрирование математических моделей. Кроме того, были значительно усовершенствованы экспериментальные методы и проведены эксперименты с новыми видами активных сред.

Результаты, относящиеся к волнам переключения и общим свойствам структур в простых бистабильных средах (§2, §3), являются классическими и не претерпели никаких изменений. Теория бегущих импульсов в одномерных возбудимых средах, описанная в §4, также не изменилась. Используя численные методы, удалось, однако, исследовать механизмы потери устойчивости таких решений. Клеточные автоматы (§5) полезны для понимания процессов, ответственных за образование структур в таких средах. В связи с наличием быстродействующих компьютеров они, однако, почти не применяются теперь для прямого их моделирования.

Кинематика (§7–§9) остается мощным методом описания и анализа автоволновых структур в возбудимых средах. Дальнейшее развитие этого метода было связано с учетом эффектов рефрактерности и процессов восстановления в среде после прохождения импульса. Как было показано В. Зыковым, это позволяет объяснить неустойчивость спиральных волн с ростом возбудимости среды и возникновение меандра, т. е. блуждания спиральных волн. Для детального знакомства с кинематической теорией автоволн можно рекомендовать два обзора [5; 6]. Впоследствии, в рамках кинематики, были рассмотрены новые типы автоволновых структур в системах со сложной анизотропией, которая меняется в зависимости от состояния среды [7; 8]. Кроме того, теория была обобщена на случай, когда ингибитор, отвечающий за процессы рефрактерности, обладает достаточно сильной диффузией [9].

Еще в шестидесятых годах прошлого века В. Кринским было показано, что автоволны могут рваться в среде с неоднородной рефрактерностью, приводя к образованию спиральных волн. В численных экспериментах А. Панфилова и других авторов [10–14] было установлено, что разрывы волн могут также возникать спонтанно во вполне однородных средах. В этом случае им предшествует неустойчивость спиральных волн, спо-

собствующая возникновению областей с затрудненным распространением. В результате происходит размножение спиральных волн и развитие режима автоволновой турбулентности. Эти эффекты были недавно обнаружены в реакции Белоусова–Жаботинского в экспериментах китайского ученого К. Уянга [15].

Вихри в трехмерных возбудимых средах были тщательно теоретически исследованы в работах ныне покойного американского физика А. Уинфри (см., например, его книгу [16]). Интересные явления, относящиеся к таким структурам, были рассмотрены А. Перцовым [17]. Описанная в §9 неустойчивость прямого вихря была действительно обнаружена в трехмерных численных расчетах В. Бикташева и др. [18]. Она приводит к возникновению особых режимов автоволновой турбулентности [19, 20], которые, по-видимому, играют существенную роль при возникновении фибрилляции и внезапной остановке сердца.

Фазовая динамика для осциллирующих активных сред и соответствующее описание пейсмекеров в таких средах (§10, §11), предложенные Й. Курамото, не претерпели никаких изменений. Было, однако, обнаружено, что подобные структуры могут также спонтанно возникать в отсутствие внешних неоднородностей в более сложных автоколебательных средах. В частности, самоорганизованные пейсмекеры были детально изучены в биритмических средах, где присутствуют два разных устойчивых предельных цикла с разными периодами колебаний [21; 22]. Взаимодействия между спиральными волнами и вихри в трехмерных автоколебательных средах были изучены И. Аронсоном (см. [23]).

В последние годы были подробно исследованы стационарные диссипативные структуры, которые лишь очень кратко описаны в §13. Дальнейшее развитие качественной теории локализованных структур изложено Б. Кернером и В. Осиповым в обзоре [24]. Подобные структуры были обнаружены Х. Суинни в экспериментах с химическими реакциями [25], а также детально исследованы для систем с газовыми разрядами в Германии в работах Х.-Г. Пурвинса с сотрудниками (см., например, [26]). Неустойчивости локализованных структур, приводящие к их статической деформации и развитию колебаний, рассмотрены японскими математиками Ф. Мимурой и Т. Ота [27]. А. С. Михайлов [28] изучил бифуркацию, приводящую к началу движения таких образований.

В настоящей книге полностью опущено обсуждение вопросов, относящихся к теории периодических стационарных структур Тьюринга. Эти структуры были обнаружены в экспериментах с химическими системами в лабораториях П. Де Кеппера во Франции [29] и Х. Суинни в США [30]. Следует заметить, что помимо стационарных структур в средах с как ми-

нимум тремя компонентами возможно также спонтанное возникновение бегущих волн. Детальное изложение соответствующих математических вопросов дано бельгийским ученым Д. Вальграфом [31].

Существует также еще один обширный класс активных систем, которые не рассмотрены в настоящей книге. Это системы, где наряду с реакциями и диффузией присутствуют также притяжение или отталкивание между составляющими их частицами. При термодинамическом равновесии в отсутствие реакций подобные конденсированные системы могут образовывать различные статические структуры и демонстрировать фазовые переходы между ними. Будучи выведенными с помощью химической реакции из состояния равновесия, они способны также к автоколебаниям и распространению автоволн [32]. Отличительной особенностью этих систем является то, что возникающие в них структуры могут иметь чрезвычайно малые размеры, гораздо меньше диффузионной длины. В результате явления самоорганизации становятся возможными на субмикронных и нанометровых масштабах, что особенно важно для молекулярной биологии. Примеры теоретического описания подобных структур в системах различного рода можно найти в работах [33–35].

Один из авторов, А. Ю. Лоскутов, подготовил для этого издания существенные обновления для главы 2. Здесь мы хотели бы добавить только несколько замечаний к материалу, изложенному в §30 и §31.

Исследования пространственно-временного хаоса стали по-настоящему возможны лишь в последние пятнадцать лет, когда впервые появились достаточно мощные компьютеры и численное моделирование турбулентности стало реальным. Для сред с реакциями и диффузией большой объем численных расчетов был проведен в рамках обобщенного уравнения Гинзбурга–Ландау. Сегодня мы знаем, что эта система обладает двумя возможными хаотическими режимами. Фазовая турбулентность характеризуется случайной эволюцией локальных фаз автоколебаний, в то время как их амплитуды остаются примерно постоянными. В отличие от этого, для режима амплитудной турбулентности характерны существенные случайные вариации как фазы, так и амплитуды колебаний. В двумерных системах турбулентность часто сопровождается размножением и аннигиляцией спиральных волн. В целом, в состоянии развитого пространственно-временного хаоса в системе присутствует большое число неустойчивых (седловых) локализованных структур и ее эволюция может быть описана как случайное блуждание между этими многочисленными отталкивающими структурами.

Ввиду присутствия огромного количества неустойчивостей управление развитым пространственно-временным хаосом не может быть осуществлено при помощи явного расчета корректирующих возмущений, как для хаоти-

ческих систем малой размерности. А. С. Михайловым с сотрудниками было показано, что турбулентность в обобщенном уравнении Гинзбурга–Ландау может быть подавлена путем введения глобальной запаздывающей обратной связи [36; 37]. В экспериментах соответствующая форма пространственно-временного хаоса была обнаружена для поверхностной химической реакции каталитического окисления угарного газа на платине. В недавних работах [38; 39] впервые была теоретически показана и экспериментально подтверждена возможность реализации контроля над развитой турбулентностью с помощью обратной связи в химических системах. Одновременно было найдено, что при помощи такой обратной связи система может удерживаться вблизи границы перехода к хаосу, и в этом состоянии появляется целый ряд новых пространственно-временных структур [39; 40]. Кроме того, недавно было показано, что стабилизация турбулентных режимов возбудимых сред может быть осуществлена посредством внешнего воздействия на некоторую достаточно малую (в некоторых случаях даже точечную) область среды [41–47]. При этом нетрудно найти эффективные частоты, на которых происходит такое подавление. После выхода среды на стабилизированный режим там остается только внешний источник. Управление трехмерной вихревой турбулентностью посредством периодических внешних воздействий также было недавно продемонстрировано [20].

Исследования по обработке информации распределенными динамическими системами (глава 3) наиболее бурно развивались за время, истекшее с публикации первого издания. Сегодня методы обработки информации искусственными нейронными сетями нашли широкое практическое применение в распознавании образов и речи с помощью компьютеров и стали неотъемлемой частью кибернетики. Кроме того, существенное развитие получили также эволюционные обучающиеся алгоритмы, кратко описанные в последнем параграфе. По всем этим вопросам уже имеется обширная специальная литература. Тем не менее, материал третьей главы может быть использован для первоначального ознакомления с предметом.

За рамками книги остались несколько важных направлений в современной теории сложных систем. В первую очередь, это проблемы синхронизации и образования динамических кластеров в больших системах связанных периодических или хаотических осцилляторов. Предварительную информацию о динамике таких систем читатель может получить из §32. К числу важных направлений также относится теория графов и динамических сетей, бурно развивающаяся в последнее время. Далее следует отметить теорию стохастических процессов и флуктуаций в неравновесных системах. Как мы теперь знаем, шум может не только вносить флуктуации, но и улучшать при определенных условиях поведение системы. Возможны

также качественные перестройки поведения системы и фазовые переходы, индуцированные шумом. Еще одним важным направлением исследований служит теория микроскопической самоорганизации, объясняющая спонтанное установление порядка в системах взаимодействующих макромолекул, ферментов и белковых машин. Обсуждение этих вопросов читатель может найти в других книгах, которые были цитированы выше.

Список литературы

1. A. S. Mikhailov, *Foundations of Synergetics I: Distributed Active Systems* (Springer, Berlin 1990, 2nd Ed. 1994).
2. A. S. Mikhailov, A. Yu. Loskutov, *Foundations of Synergetics II: Chaos and Noise* (Springer, Berlin 1991, 2nd Ed. 1996).
3. A. S. Mikhailov, V. Calenbuhr, *From Cells to Societies: Models of Complex Coherent Action* (Springer, Berlin 2002).
4. S. C. Manrubia, A. S. Mikhailov, D. H. Zanette, *Emergence of Dynamical Order: Synchronization Phenomena in Complex Systems* (World Scientific, Singapore 2004).
5. В. А. Давыдов, В. С. Зыков, А. С. Михайлов, УФН 161, 45 (1991).
6. A. S. Mikhailov, V. A. Davydov, V. S. Zykov, *Physica D* 70, 1 (1994).
7. A. S. Mikhailov, *Phys. Rev. E* 49, 5875 (1994).
8. F. Mertens, N. Gottschalk, M. Bar, et al., *Phys. Rev. E* 51, R5193 (1995).
9. V. S. Zykov, A. S. Mikhailov, S. C. Muller, *Phys. Rev. Lett.* 81, 2811 (1998).
10. A. V. Panfilov, A. Holden, *Phys. Lett. A* 151, 23 (1990).
11. A. V. Panfilov, P. Hogeweg, *Phys. Lett. A* 156, 295 (1993).
12. M. Courtemanche, A. T. Winfree, *Int. J. Bifurc. Chaos* 1, 219 (1991).
13. A. Karma, *Phys. Rev. Lett.* 71, 1103 (1993).
14. M. Bar, M. Eiswirth, *Phys. Rev. E* 48, R1635 (1993).
15. L. Q. Zhou, C. X. Zhang, Q. Ouyang, *Int. J. Mod. Phys. B* 17, 4072 (2003).
16. A. T. Winfree, *When Time Breaks Down* (Princeton Univ. Press, 1987).
17. M. Vinson, S. Mironov, S. Mulvey, A. Pertsov, *Nature* 386, 477 (1997).
18. V. N. Biktashev, A. V. Holden, H. Zhang, *Phil. Trans. R. Soc. Lond. A* 347, 611 (1994).
19. V. N. Biktashev, *Int. J. Bifurc. Chaos* 8, 677 (1998).
20. S. Alonso, F. Sagues, A. S. Mikhailov, *Science* 299, 1722 (2003).
21. M. Stich, M. Ipsen, A. S. Mikhailov, *Phys. Rev. Lett.* 86, 4406 (2001).
22. M. Stich, M. Ipsen, A. S. Mikhailov, *Physica D* 171, 19 (2002).
23. I. S. Aranson, L. Kramer, *Rev. Mod. Phys.* 74, 99 (2002).

24. Б. С. Кернер, В. В. Осипов, УФН 160, 1 (1990).
25. K. J. Lee, W. D. McCormick, J. E. Pearson, H. L. Swinney, *Nature* 369, 215 (1994).
26. I. Brauer, M. Bode, E. Anmel, H.-G. Purwins, *Phys. Rev. Lett.* 84, 4104 (2000).
27. T. Ohta, M. Mimura, R. Kobayashi, *Physica D* 34, 115 (1989).
28. K. Krischer, A. S. Mikhailov, *Phys. Rev. Lett.* 73, 3165 (1994).
29. P. De Kepper, V. Castets, E. Dulos, J. Boissonade, *Physica D* 49, 161 (1991).
30. Q. Ouyang, H. L. Swinney, *Nature* 352, 610 (1991).
31. D. Walgraef, *Spatio-Temporal Pattern Formation* (Springer, Berlin 1997).
32. A. S. Mikhailov, G. Ertl, *Science* 272, 1596 (1996).
33. M. Hildebrand, A. S. Mikhailov, G. Ertl, *Phys. Rev. Lett.* 81, 2602 (1998).
34. M. Hildebrand, M. Kuperman, H. Wio, et al., *Phys. Rev. Lett.* 83, 1475 (1999).
35. R. Reigada, F. Sagues, A. S. Mikhailov, *Phys. Rev. Lett.* 89, 038301 (2002).
36. D. Battogtokh, A. S. Mikhailov, *Physica D* 90, 84 (1996).
37. D. Battogtokh, A. Preusser, A. S. Mikhailov, *Physica D* 106, 327 (1997).
38. M. Kim, M. Bertram, M. Pollmann, et al., *Science* 292, 1357 (2001).
39. M. Bertram, C. Beta, M. Pollmann, et al., *Phys. Rev. E* 67, 036208 (2003).
40. M. Bertram, A. S. Mikhailov, *Phys. Rev. E* 67, 036207 (2003).
41. A. T. Stamp, G. V. Osipov, J. J. Collins, *Chaos* 12, 931 (2002).
42. G. V. Osipov, J. J. Collins, *Phys. Rev. E* 60, 54 (1999).
43. H. Zhang, B. Hu, G. Hu, *Phys. Rev. E* 68, 026134 (2003).
44. А. Ю. Лоскутов, Р. В. Черемин, С. А. Высоцкий, Доклады АН 404, 477 (2005).
45. S. A. Vysotsky, R. V. Cheremin, A. Loskutov, *J. Phys.: Conf. Ser.* 23, 202 (2005).
46. H. Zhang, Zh. Cao, N.-J. Wu, H.-P. Ying, G. Hu, *Phys. Rev. Lett.* 94, 188301 (2005).
47. А. Лоскутов, С. А. Высоцкий, Письма в ЖЭТФ 84, 616 (2006).

Введение

Долгое время жизнь рассматривалась как антипод неорганической природы. Сегодня, однако, происходит все более активное проникновение физических методов и подходов в биологию. Оказывается также, что основные формы кооперативного поведения, свойственные живым организмам, имеют свои аналоги среди неорганических систем.

Наиболее очевидная особенность биологических систем заключается в том, что они способны к самоорганизации, т.е. спонтанному образованию и развитию сложных упорядоченных структур. Как показал еще Э. Шредингер, это не противоречит законам термодинамики, поскольку все живые биологические системы не являются замкнутыми и обмениваются энергией (или веществом) с окружающей средой. Энтропия, служащая мерой беспорядка, может уменьшаться в открытых системах с течением времени. Необходимая предпосылка эффектов самоорганизации заключается, кроме того, в наличии потока энергии, поступающего в систему от внешнего источника и диссипируемого ею. Именно благодаря этому потоку система становится активной, т.е. приобретает способность к автономному образованию структур. Очевидно, что эффекты самоорганизации не могут быть исключительным свойством биологических объектов, и должны наблюдаться (пусть и в более простой форме) также в системах неорганического происхождения.

Большой интерес представляют распределенные среды, которые построены из дискретных элементов, локально взаимодействующих друг с другом и, таким образом, являющихся приближением естественных пространственно протяженных систем. Хотя разнообразие таких сред чрезвычайно велико, число математических моделей, которые используются для описания процессов образования и развития структур в таких системах, не столь значительно. По-видимому, даже когда отдельные элементы системы (например, живые клетки) обладают сложной внутренней структурой, вся их сложность не проявляется во взаимодействиях между ними и, с точки зрения макросистемы, они функционируют как достаточно простые объекты с малым числом эффективных степеней свободы. В противном случае никакие упорядоченные структуры в системе обычно не возникают.

Задача нелинейной динамики и теории сложных систем состоит в нахождении и подробном исследовании тех базовых математических моделей, которые исходят из наиболее типичных предположений о свойствах отдельных элементов, составляющих систему, и законах взаимодействия между ними. Поскольку главным отличительным свойством изучаемых сред являются протекающие в них процессы самоорганизации, теорию сложных систем можно также рассматривать и как общую теорию самоорганизации в средах различной природы.

Самоорганизация тесно связана с зарождением турбулентности. При макроскопическом течении жидкости к каждому ее малому элементу поступает энергия от крупномасштабных мод, которая превращается затем в теплоту за счет действия вязких сил. Наличие этого потока энергии превращает жидкость в своеобразную активную среду. Чем выше средняя скорость течения жидкости, тем интенсивнее поток энергии, проходящей через каждый ее элемент. Как известно, при больших средних скоростях течения оно, как правило, является турбулентным, т.е. характеризуется *хаотическими* пульсациями поля скоростей, давления, температуры, плотности и т.п. Переход от ламинарного течения к турбулентности может осуществляться постепенно, не скачком. В этом случае возникновению турбулентности предшествует особая стадия, характеризующаяся появлением все более сложного течения. Спонтанное образование таких упорядоченных структур можно рассматривать как пример самоорганизации. Для гидродинамических течений соответствующий процесс был впервые проанализирован в общем виде Л. Д. Ландау в 1944 г. В модели зарождения турбулентности, которая была предложена Л. Д. Ландау, спонтанное образование и усложнение структур связывалось с появлением новых эффективных степеней свободы, благодаря чему течение жидкости приобретало все большую автономию от внешних условий.

С ростом притока энергии регулярные структуры становятся все более сложными и причудливо меняющимися во времени. Где же лежит граница между регулярной, но сложно организованной структурой и хаосом? Критерием может служить устойчивость возникающих в процессе течения образований по отношению к малым возмущениям. Если такая устойчивость отсутствует, то детерминированное описание теряет смысл, и необходимо использовать статистические методы. Впервые на связь между статистическим подходом и неустойчивостью указывал еще Анри Пуанкаре.

Из сказанного ясно, что теоретический анализ процессов хаотизации (зарождения турбулентности) в различных средах также должен быть включен в круг проблем, изучаемых в рамках теории сложных систем. Естествен-

но отнести к ним и исследование общих свойств хаотических режимов, возникающих вслед за разрушением регулярных структур.

Как же возникает хаотическое движение? Казалось бы, путей его возникновения должно быть очень много. Однако выяснилось, что число сценариев процесса хаотизации совсем невелико. Более того, некоторые из них подчиняются универсальным закономерностям, и не зависят от природы рассматриваемого объекта. Одни и те же пути развития хаоса присущи самым разнообразным физическим, химическим, биологическим и др. системам. Универсальное поведение напоминает обычные фазовые переходы второго рода, а введение ренормгрупповых и скейлинговых методов, известных в статистической механике, открывает новые перспективы в изучении хаотической динамики.

Знание основных закономерностей в поведении хаотических сред позволяют перейти к целенаправленному конструированию искусственных систем, процессы самоорганизации в которых приводили бы к образованию нужных структур. Наиболее развитым приложением здесь является создание аналоговых устройств *обработки информации*. Действие таких устройств должно базироваться не на логических операциях над некоторыми символами, а представлять собой сложный процесс, связанный с эволюцией и взаимодействием пространственных и волновых структур в искусственно созданных средах. Это дает возможность при относительно малых энергетических затратах создать устройства принципиально нового типа, способные запоминать, шифровать и обрабатывать заданную информацию. Более того, экспериментальные данные свидетельствуют о том, что автоколебания (в том числе хаос) играют важную роль в процессе анализа информации нейроподобными системами. Следовательно, принцип организации памяти можно представить как динамический процесс. Такой подход привел к использованию теории динамических систем в проблеме обработки информации и при создании систем искусственного интеллекта.

Однако задачами информатики не исчерпываются потенциальные приложения теории самоорганизующихся систем. В своей автобиографической книге «Нарушая покой Вселенной» знаменитый физик-теоретик Фримен Дайсон, размышляя о будущем человечества, предвидит возникновение «зеленой» технологии, которая строилась бы на тех же принципах самоорганизации, что и живая природа. Ф.Дайсон противопоставляет ее традиционной «серой» технологии, в основе которой лежит жесткий механистический чертеж и пассивное следование командам.

Создание любой современной машины начинается с разработки ее детального плана, в соответствии с которым затем осуществляется ее сборка из предварительно изготовленных деталей. Процесс индивидуального раз-

вития живого организма не требует вмешательства внешних управляющих сил. Более того, нет никакого «подробного чертежа», который мог бы быть записан, например, в молекулах ДНК клетки. Легко видеть, что такой детальный план построения содержал бы колоссальное количество информации, которая просто не способна поместиться в клетке. По-видимому, развитие живого организма есть последовательность автономных актов самоорганизации. Управление этим процессом может осуществляться с помощью слабых воздействий, которые влияют на выбор того или иного конкретного пути развития в те моменты, когда развивающаяся структура оказывается в состояниях, характеризующихся наличием нескольких возможных равноправных продолжений. Именно эти управляющие воздействия закодированы, по всей видимости, в генетических последовательностях.

Любой живой организм представляет собой иерархию достаточно автономных подсистем, в которой исходящие от верхнего уровня сигналы управления не имеют характера жестких команд, подчиняющих себе активность всех индивидуальных элементов более низких уровней. Вместо этого от высших уровней иерархии поступают сигналы, которые определяют переходы подсистем из одного режима функционирования к другому. Иерархическое устройство сложных живых систем, которые представляют собой ансамбль слабо связанных подсистем более простого строения, позволяет избежать неустойчивостей и нежелательной динамики, которые неизбежно возникают в сложных системах с жестким централизованным управлением.

Использование этих принципов в технике позволило бы резко расширить ее возможности, построить ее «в формах самой жизни». Это постепенно привело бы к преодолению сегодняшних принципиальных различий между миром техники и миром живой природы. Одни и те же наиболее эффективные закономерности должны лежать в основе функционирования искусственно созданных технических устройств и живых организмов. Как пишет в упомянутой книге Ф.Дайсон, «по мере того, как будут углубляться наши знания в биологии, мы столкнемся с тем, что различия между биологией и электроникой будут все более стираться».

Знание закономерностей самоорганизации дает возможность в самом прямом смысле вмешиваться в деятельность существующих биологических систем и живых организмов и управлять ими. Можно также целенаправленно формировать живые (например, экологические) системы, чтобы они образовывали в своем развитии нужные нам пространственные структуры или обладали предписанным временным поведением.

Предупредим читателя, что в этой книге он не найдет широкого обсуждения перечисленных перспективных приложений. Наша задача является

более скромной — рассказать о теоретических результатах, которые были получены при исследовании математических моделей распределенных активных сред, нелинейных систем и сетей. Впрочем, некоторые их описанных результатов представлены в библиографических комментариях и списке литературы.

ГЛАВА 1

ВОЛНЫ И СТРУКТУРЫ В АКТИВНЫХ СРЕДАХ

Активные среды характеризуются непрерывным рассредоточенным притоком энергии от внешнего источника и ее диссипацией. Благодаря тому, что через каждый физически малый элемент среды протекает поток энергии от источника к термостату, этот элемент выводится из состояния теплового равновесия и приобретает способность совершать автоколебания, быть триггерным (бистабильным) либо возбудимым. Когда отдельные такие элементы локально связаны между собой и формируют распределенную активную среду, в подобной среде наблюдается образование различных стационарных или зависящих от времени пространственных структур. Эти процессы лежат в основе явлений самоорганизации в активных средах.

1. Основные типы активных сред

Как уже отмечалось во Введении, в изучении кооперативного поведения сложных систем важную роль играют типовые, или *базовые*, модели. Каждая такая модель «сама по себе» не соответствует какой-то конкретной системе; она строится исходя из чрезвычайно упрощенных представлений о свойствах отдельных элементов и процессов взаимодействия между ними.

Изучение кооперативного поведения — это важная составная часть физики конденсированных систем, занятой исследованием жидкостей, кристаллов, стекол и т. п. Однако рассматриваемые при этом системы обладают спецификой, не свойственной биологическим или сложным химическим системам. Обычно элементы, из которых состоит физическая система, являются пассивными — для каждого из них, как и для системы в целом,

существует выделенное равновесное состояние с минимумом энергии. Достигнув этого состояния, система уже не способна самостоятельно его покинуть.

Для биологии типична иная ситуация. Здесь отдельные элементы — живые клетки или макроорганизмы — являются активными. Их поведение не может быть представлено как эволюция в сторону какого-то выделенного устойчивого состояния абсолютного покоя. Разумеется, сохранение активности образующих систему элементов возможно лишь благодаря притоку к ним энергии от внешних источников. Если такой приток энергии имеет место, активными могут стать и довольно простые физические системы.

Можно выделить три простейших типа активных элементов: *бистабильные, возбудимые и автоколебательные*.

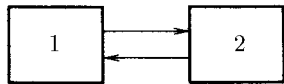


Рис. 1.1. Схема переходов в бистабильном элементе

Бистабильный (или триггерный) элемент обладает двумя стационарными состояниями (1 и 2), в каждом из которых он может находиться неограниченно долго (рис. 1.1). Внешние воздействия могут приводить к переходам из одного состояния в другое. Чтобы вызвать переход, интенсивность воздействия должна превысить некоторый пороговый уровень.

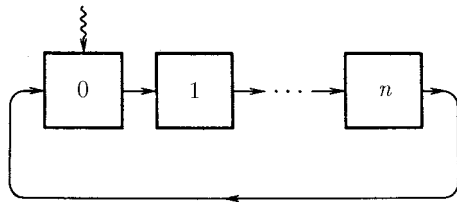


Рис. 1.2. Схема переходов в возбудимом элементе

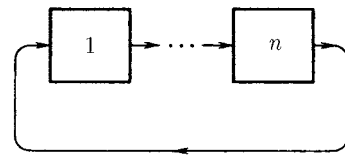


Рис. 1.3. Схема переходов в автоколебательном элементе

Возбудимый (или мультивибраторный) элемент имеет единственное выделенное состояние покоя, устойчивое по отношению к достаточно слабым внешним воздействиям. Однако такой элемент отличается от пассивного по своей реакции на воздействия, превышающие пороговый уровень. В ответ на такое достаточно интенсивное внешнее воздействие в элементе возникает вспышка активности: он совершает определенную последо-

вательность активных переходов и лишь затем возвращается к исходному состоянию покоя (рис. 1.2).

Автоколебательный элемент работает подобно «вечному двигателю». Он автономно совершает циклические переходы через некоторую группу состояний (рис. 1.3). Внешние воздействия способны лишь замедлить или ускорить это циклическое движение, но не приостановить его.

Рассмотрим вначале цепочку из бистабильных элементов. Естественно предположить, что в ней воздействие друг на друга оказывают лишь соседние элементы. В простейшем случае элементы, находящиеся в одинаковых состояниях, не влияют друг на друга. Если же состояния двух соседних элементов различны, они взаимодействуют. Элемент, находящийся в менее устойчивом (метастабильном) стационарном состоянии, переходит в более устойчивое состояние — такое же, что и у его соседа. Нетрудно убедиться, что в результате по цепочке бистабильных элементов распространяется *волна переключения* из метастабильного в абсолютно устойчивое состояние (рис. 1.4).

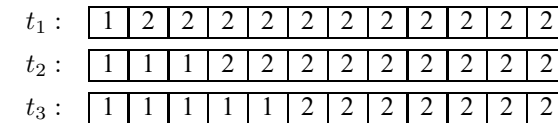


Рис. 1.4. Схема распространения волны переключения в цепочке из бистабильных элементов. Показаны состояния элементов цепочки в последовательные моменты времени t_1 , t_2 и t_3

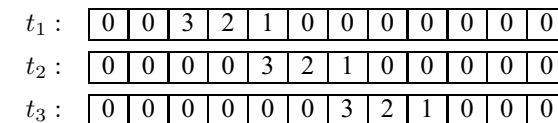


Рис. 1.5. Схема распространения волны возбуждения в цепочке из возбудимых элементов. Показаны состояния элементов цепочки в последовательные моменты времени t_1 , t_2 и t_3

Пусть цепочка состоит из возбудимых элементов. Каким может быть взаимодействие между этими элементами? Естественно считать, что элемент, перешедший из состояния покоя в активную форму, остается невосприимчивым к внешним воздействиям, пока он не совершит всю предписанную ему последовательность переходов. Поэтому достаточно рассмот-

t_1 :	3	2	1	3	2	1	3	2	1	3	2	1
t_2 :	1	3	2	1	3	2	1	3	2	1	3	2
t_3 :	2	1	3	2	1	3	2	1	3	2	1	3

Рис. 1.6. Схема распространения фазовой волны в цепочке из автоколебательных элементов. Показаны состояния элементов цепочки в последовательные моменты времени t_1 , t_2 и t_3

реть лишь случаи, когда по соседству оказались два элемента, из которых один находится в активной форме, а другой — в состоянии покоя. Можно представить себе два вида их взаимодействия. Во-первых, всякий элемент в активной форме мог бы выводить из состояния покоя оказавшийся рядом с ним элемент. Легко проверить, что в этом случае в ответ на однократное воздействие в цепочке возникнет незатухающая волновая активность. Более интересна ситуация, когда выводить из состояния покоя могут лишь элементы, находящиеся в первых фазах вспышки активности. Тогда в результате возбуждения крайнего элемента в цепочке по ней будет распространяться уединенный импульс активности (*волна возбуждения*), после прохождения которого элементы возвращаются в исходное состояние покоя (рис. 1.5).

Наконец, если составить цепочку из автоколебательных элементов, в ней могут наблюдаться *фазовые волны* (рис. 1.6). Для этого достаточно создать сдвиг по начальным фазам колебаний вдоль цепочки. Всем известный пример фазовых волн — это волны в гирляндах электрических лампочек: каждая из них загорается и гаснет через один и тот же промежуток времени, но моменты загорания сдвинуты для соседних лампочек. Взаимодействие между автоколебательными элементами может привести к появлению зависимости частоты фазовых волн от их пространственного периода либо даже обеспечивать синхронизацию, т. е. обеспечивать установление единой фазы колебаний.

Еще более сложные эффекты наблюдаются в двумерных или трехмерных сетях, образованных бистабильными, возбудимыми или автоколебательными элементами. Мы обсудим эти эффекты позднее.

До настоящего момента мы рассматривали активные элементы фактически в качестве определенных *автоматов*, т. е. объектов с дискретным набором состояний и некоторыми правилами переходов между ними. Более детальный уровень описания основывается на построении дифференциальных уравнений, характеризующих динамику отдельных элементов и их взаимодействия. Заметим, наконец, что активная среда может являться также распределенной, а не состоять из дискретных связанных между собой яче-

ек. В этом случае свойства бистабильности, возбудимости или способности к совершению незатухающих автоколебаний относятся к малым, условно выделенным элементам такой распределенной среды.

2. Волны переключения и заселения в простых бистабильных средах

Простейший бистабильный элемент описывается дифференциальным уравнением

$$\dot{u} = f(u), \quad (1.1)$$

в котором нелинейная функция $f(u)$ имеет характерный вид, изображенный на рис. 1.7. Очевидно, что всякое значение u^* , для которого $f(u^*) = 0$, отвечает стационарному состоянию такого элемента. Поведение малых отклонений $\delta u = u - u^*$ при этом подчиняется линеаризованному уравнению

$$\delta \dot{u} = f'(u^*) \delta u. \quad (1.2)$$

Следовательно, значение $u = u^*$ отвечает устойчивому (по отношению к малым возмущениям) стационарному состоянию элемента, лишь если производная $f'(u)$ в этой точке отрицательна. Поэтому для функции $f(u)$, изображенной на рис. 1.7, значения $u = u_1$ и $u = u_3$ соответствуют устойчивым состояниям, тогда как стационарное состояние со значением $u = u_2$ абсолютно неустойчиво.

Приведем несколько примеров различных систем, поведение которых действительно характеризуется уравнениями вида (1.1) с двумя устойчивыми состояниями.

1. Пусть элемент представляет собой ячейку, внутри которой может протекать процесс горения. При этом будем считать, что имеющийся запас горючей смеси в этой ячейке поддерживается постоянным за счет непрерывного притока от внешнего источника либо он настолько велик, что его расходом в ходе реакции можно пренебречь. Предположим, что при горении в единицу времени выделяется количество теплоты $q(\Theta)$, которое зависит, вообще говоря, от температуры Θ , при которой происходит горение. Ясно, что зависимость имеет ступенчатый характер (рис. 1.8): при температурах ниже некоторой критической Θ_c горение отсутствует и теплота не выделяется.

Если элемент изолирован от внешней среды, то вся выделяющаяся при горении теплота расходуется на его нагрев, т. е.

$$q(\Theta)\Delta t = C\Delta\Theta, \quad (1.3)$$

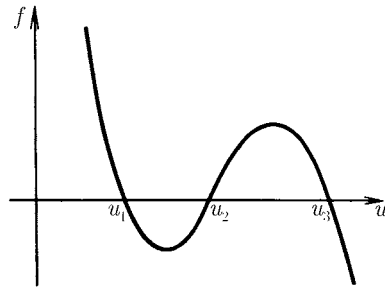


Рис. 1.7. Типичный вид функции $f(u)$

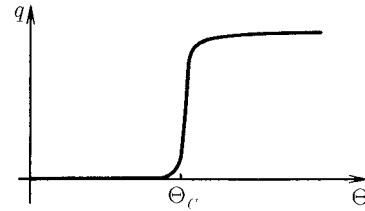


Рис. 1.8. Зависимость количества теплоты q , выделяющейся в единицу времени при горении, от температуры Θ

где $\Delta\Theta$ — приращения температуры за время Δt , а C — теплоемкость элемента. Поэтому изменение температуры со временем подчиняется уравнению

$$\dot{\Theta} = C^{-1}q(\Theta). \quad (1.4)$$

Поскольку функция $q(\Theta)$ неотрицательна, то согласно (1.4) температура элемента неограниченно возрастает с течением времени. В действительности, конечно, имеет место теплообмен между элементом и окружающей средой. Если температура окружающей среды есть Θ_1 , а коэффициент, характеризующий скорость теплообмена, равен γ , то модифицированное уравнение, учитывающее эффект теплообмена, имеет вид

$$\dot{\Theta} = C^{-1}q(\Theta) - \gamma(\Theta - \Theta_1). \quad (1.5)$$

Обозначив правую часть этого уравнения как

$$f(\Theta) = C^{-1}q(\Theta) - \gamma(\Theta - \Theta_1),$$

приходим к (1.1).

В данном случае устойчивые по отношению к малым возмущениям стационарные состояния имеют особенно простую интерпретацию (рис. 1.9а): состояние $\Theta = \Theta_1$ отвечает отсутствию горения, так что температура элемента совпадает с температурой окружающей среды, а состояние с $\Theta = \Theta_1$ соответствует установившемуся стационарному режиму, когда выделяющаяся при горении теплота полностью отводится в окружающую среду. Оба

состояния устойчивы по отношению к малым возмущениям, однако воздействия достаточно большой интенсивности могут приводить к переходам между двумя состояниями — элемент можно «зажечь» или «погасить».

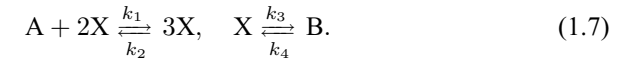
Кстати, тот же элемент при других соотношениях параметров может обладать всего одним устойчивым стационарным состоянием. Например, если тепловыделение слишком мало либо теплоотвод слишком велик, режим устойчивого самоподдерживающегося горения отсутствует (рис. 1.9б) и единственное стационарное состояние — это состояние с температурой $\Theta = \Theta_1$. Если температура окружающей среды Θ_1 превышает критическую температуру загорания Θ_c , элемент является самовоспламеняющимся (рис. 1.9в). Тогда единственное стационарное состояние отвечает самоподдерживающемуся режиму горения.

Заметим, наконец, что если взять цепочку из связанных между собой подобных элементов, то (в приближении непрерывной среды) она описывается уравнением

$$\frac{\partial\Theta}{\partial t} = f(\Theta) + \chi \frac{\partial^2\Theta}{\partial x^2}, \quad (1.6)$$

где χ — температуропроводность.

2. Рассмотрим гипотетическую схему химических реакций (модель Шлёгеля)



Пусть концентрации химических веществ A и B заданы и поддерживаются постоянными. Тогда кинетическое уравнение для изменения со временем концентрации X вещества X записывается как

$$\dot{X} = k_1AX^2 - k_2X^3 - k_3X + k_4B. \quad (1.8)$$

Таким образом, стоящая в правой части этого уравнения функция $f(X)$ является кубическим полиномом. В определенном интервале значений параметров k_1, k_2, k_3, k_4, A и B эта функция имеет вид, показанный на рис. 1.7, а следовательно, такая необратимая химическая реакция обладает двумя устойчивыми стационарными режимами, которые характеризуются разными значениями концентрации $X = X_1$ и $X = X_3$.

Кинетическое уравнение (1.8) записано в предположении, что внутри элемента происходит полное диффузионное перемешивание реагирующих веществ. Если взять цепочку из таких элементов, то между ними будет происходить обмен молекулами за счет диффузии. В континуальном приближении (т. е. для распределенной среды) распределение концентрации

вещества X вдоль цепочки меняется тогда с течением времени согласно уравнению

$$\frac{\partial X}{\partial t} = f(X) + D \frac{\partial^2 X}{\partial x^2}. \quad (1.9)$$

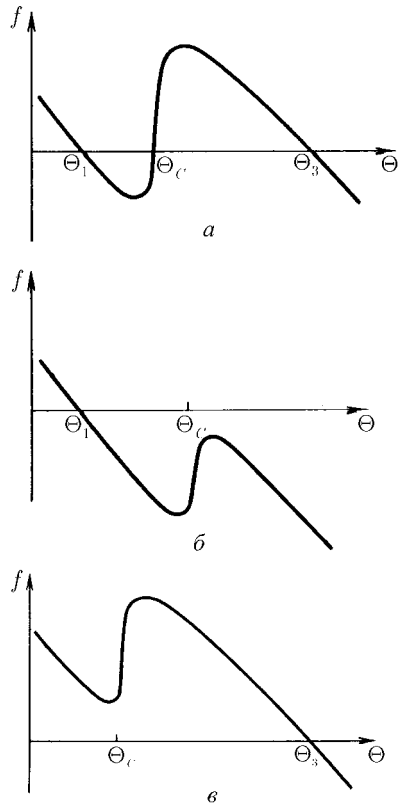


Рис. 1.9. Три возможные ситуации для ячейки, содержащей горючую смесь: a – бистабильный режим, b – воспламенение невозможно, v – самовоспламенение

3. Рассмотрим простую экологическую модель. Пусть в среде происходит размножение некоторых животных, численность которых в расчете на единицу объема мы обозначим как n . Предположим, что размножение имеет половой характер, т. е. вероятность рождения¹ нового индивида пропорциональна вероятности встречи мужской и женской особей, т. е. величине n^2 . Тогда с учетом смертности животных и их диффузионного блуждания по среде динамика такой системы описывается уравнением

$$\dot{n} = -\gamma n + \alpha m(n)n^2 + D\Delta n. \quad (1.10)$$

Скорость размножения пропорциональна массе пищи m , которая, конечно, зависит от числа поедающих ее животных. Изменение m со временем, вообще говоря, должно описываться отдельным дифференциальным уравнением, содержащим n .

Ограничимся рассмотрением ситуации, когда масса пищи быстро подстраивается к мгновенной численности животных, так что m представляет собой некоторую убывающую функцию от численности животных n . Чтобы

был невозможен неограниченный рост численности животных в популяции, убывание m должно быть достаточно быстрым, например, экспоненциальным:

$$m = m_0 \exp(-n/n_0), \quad (1.11)$$

где m_0 и n_0 – некоторые постоянные коэффициенты.

При таких условиях рассматриваемая экологическая система также представляет собой бистабильную среду, аналогичную рассмотренным выше, причем нелинейная функция $f(n)$ дается выражением

$$f = -\gamma n + \alpha m_0 \exp(-n/n_0) n^2. \quad (1.12)$$

Эта среда обладает двумя устойчивыми однородными состояниями. В первом из них $n \equiv 0$; такое состояние устойчиво, потому что при низких плотностях n вероятность встречи двух особей (а следовательно, вероятность размножения) очень мала. Во втором устойчивом стационарном состоянии ограничение роста достигается благодаря обеднению пищевых ресурсов.

Эти три простых примера показывают, что бистабильные активные среды, описываемые уравнением

$$\dot{u} = f(u) + D\Delta u, \quad (1.13)$$

могут иметь самую различную природу.

Основным типом структуры, характерным для простых бистабильных сред, является *волна переключения*; при ее распространении элементы среды переходят из одного устойчивого состояния в другое.

Волна переключения из состояния $u = u_3$ в состояние $u = u_1$, движущаяся со скоростью c , представляет собой частное решение уравнения (1.13) вида

$$u = u(\xi), \quad \xi = x - ct \quad (1.14)$$

с граничными условиями

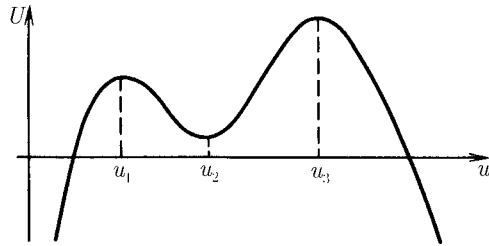
$$\begin{aligned} u &\rightarrow u_3, & \xi &\rightarrow -\infty; \\ u &\rightarrow u_1, & \xi &\rightarrow +\infty. \end{aligned} \quad (1.15)$$

После подстановки (1.14) в (1.13) приходим к уравнению

$$-cu' = f(u) + Du'', \quad (1.16)$$

в котором штрихами обозначены производные по переменной ξ .

¹Конечно, эта модель является крайне упрощенной. В частности, мы принимаем, что локальные численности мужских и женских особей всегда совпадают.

Рис. 1.10. Зависимость потенциала U от координаты u

Чтобы понять наиболее существенные особенности решения, удобно воспользоваться следующей аналогией. Введем функцию

$$U(u) = \int_0^u f(u) du \quad (1.17)$$

и перепишем уравнение (1.16) в форме

$$Du'' = -\frac{\partial U}{\partial u} - cu'. \quad (1.18)$$

Если рассматривать u как координату некоторой частицы, а ξ — как время, нетрудно заметить, что (1.18) совпадает с уравнением движения такой частицы массы D в потенциальном поле $U(u)$ при наличии вязкого трения, пропорционального скорости частицы u' , причем параметр c , т. е. скорость распространения фронта волны, играет в (1.18) роль динамической вязкости.

Легко убедиться, что в точках $u = u_1$ и $u = u_3$ функция $U(u)$ достигает максимума, а в промежуточной точке $u = u_2$ она имеет локальный минимум (рис. 1.10). Предположим вначале, что $U(u_3) > U(u_1)$; это осуществляется согласно (1.17) при выполнении условия $A > 0$, где

$$A = \int_{u_1}^{u_3} f(u) du. \quad (1.19)$$

Если бы вязкое трение отсутствовало ($c = 0$), частица в потенциальном поле, движение которой задано уравнением (1.18), сохраняла бы энергию

$$\mathcal{E} = U(u) + \frac{1}{2}D\left(\frac{du}{d\xi}\right)^2. \quad (1.20)$$

Тогда частице, находящейся в точке u_3 с более высоким значением потенциала U (рис. 1.10), потребовалось бы лишь конечное время ξ , чтобы попасть в точку u_1 , где потенциал U хотя и достигает локального максимума, но ниже, чем потенциал в точке u_3 . Затем она продолжала бы удаляться в сторону меньших значений координаты u , так что мы имели бы $u \rightarrow -\infty$ при $\xi \rightarrow +\infty$ и условие (1.15) было бы нарушено.

В то же время, если вязкость c слишком велика, частица быстро теряет свою энергию и не может вообще достигнуть перевальной точки u_3 . При $\xi \rightarrow +\infty$ она окажется в точке $u = u_2$, отвечающей локальному минимуму потенциала U .

При заданном потенциале $U(u)$, определяемом согласно (1.17) функцией $f(u)$, существует единственная вязкость $c = c_0$, при которой потеря энергии на трение в точности равна разности потенциалов в точках u_1 и u_3 , т. е. величине $\Delta\mathcal{E} = U(u_3) - U(u_1)$. Это значение и определяет скорость распространения фронта волны переключения (скорость волны).

Таким образом, скорость волны переключения и ее профиль однозначно определяются собственными характеристиками среды. Мы можем также заметить, что скорость волны c_0 убывает с уменьшением разности $\Delta\mathcal{E}$ и обращается в нуль при $U(u_3) = U(u_1)$. Если $U(u_3) < U(u_1)$, т. е. величина A , определяемая равенством (1.19), меньше нуля, волна переключения движется в обратном направлении: $u \rightarrow u_1$ при $\xi \rightarrow -\infty$ и $u \rightarrow u_3$ при $\xi \rightarrow +\infty$, т. е. при ее прохождении состояние среды u_3 сменяется состоянием u_1 .

Приведенные выше рассуждения можно было бы воспроизвести, не прибегая к аналогиям с механической системой. На рис. 1.11 изображены картины фазовых траекторий уравнения (1.18) на плоскости (u, u') в случае $A < 0$. При $c > c_0$ траектория, исходящая из стационарной точки $(u_3, 0)$, попадает в стационарную точку $(u_2, 0)$, тогда как при $c < c_0$ она уходит на бесконечность. Стационарные точки $(u_3, 0)$ и $(u_1, 0)$ являются по стандартной классификации седловыми, а траектория, исходящая из $(u_3, 0)$ или $(u_1, 0)$, есть сепаратриса седловой точки. Таким образом, значению $c = c_0$ (рис. 1.11б) соответствует случай, когда сепаратриса идет из седла $(u_1, 0)$ в седло $(u_3, 0)$.

Общих аналитических методов, позволяющих рассчитывать скорость волны переключения при произвольной функции $f(u)$, не существует.

Тем не менее, в важном частном случае, когда функция $f(u)$ задается полиномом третьего порядка

$$f = -\alpha(u - u_1)(u - u_2)(u - u_3), \quad u_1 < u_2 < u_3, \quad (1.21)$$

нетрудно получить точное решение задачи. Скорость волны переключения

тогда равна

$$c_0 = \frac{1}{2} \sqrt{\alpha D} (u_1 + u_3 - 2u_0). \quad (1.22)$$

Несколько других примеров, допускающих точное решение, приведено в [119].

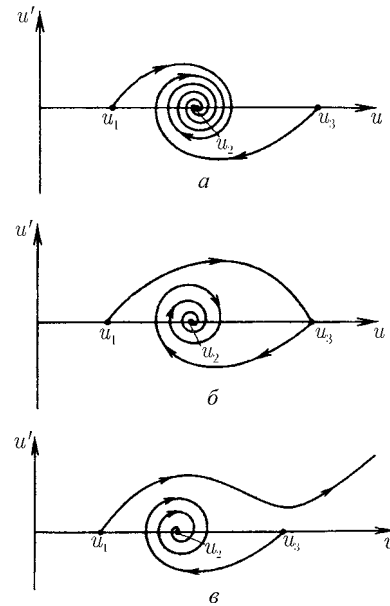


Рис. 1.11. Поведение фазовых траекторий, задаваемых уравнением (1.18), при $c > c_0$ (а), $c = c_0$ (б) и $c < c_0$ (в)

Обычно при расчете скорости волн переключения в бистабильных средах приходится использовать приближенные методы. Первый из них относится к расчету скорости медленных волн.

Как уже отмечалось, при выполнении условия $A = 0$ (см. (1.19)) скорость распространения фронта волны переключения обращается в нуль, т.е. граница раздела между областями с двумя различными стационарными состояниями покоится. Естественно ожидать, что при малых A скорость смещения границы раздела c также невелика.

Если домножить обе части уравнения (1.16) на производную $du/d\xi$ и проинтегрировать по ξ в пределах от $-\infty$ до ∞ , то с учетом условий (1.15) нетрудно получить

$$c \int_{-\infty}^{\infty} \left(\frac{du}{d\xi} \right)^2 d\xi = \int_{u_1}^{u_3} f(u) du. \quad (1.23)$$

Это равенство можно использовать для приближенного нахождения скорости c .

Именно если скорость волны c мала, то профиль этой волны $u(\xi)$ мало отличается от профиля покоящейся границы раздела $u^{(0)}(\xi)$ для $A = 0$. Подставляя в (1.23) вместо $u(\xi)$ функцию $u^{(0)}(\xi)$ и учитывая определение величины A , приближенно получаем

$$c \approx A \left[\int_{-\infty}^{\infty} \left(\frac{du^{(0)}}{d\xi} \right)^2 d\xi \right]^{-1}. \quad (1.24)$$

Заметим, что при $A = 0$ профиль стоячей волны переключения $u^{(0)}(\xi)$ дается решением уравнения

$$Du'' = -\frac{\partial U}{\partial u}, \quad (1.25)$$

которое имеет вид

$$\xi = \int_{u_1}^u \frac{1}{\sqrt{(2/D)(U(u_1) - U(u))}} du. \quad (1.26)$$

Формула (1.24) в сочетании с выражением (1.26) определяет скорость распространения *медленных волн переключения*; подчеркнем, что она оказывается пропорциональной величине A .

Другой приближенный способ расчета был предложен в 1938 г. Я. Б. Зельдовичем и Д. А. Франк-Каменецким [62] в теории теплового распространения пламени.

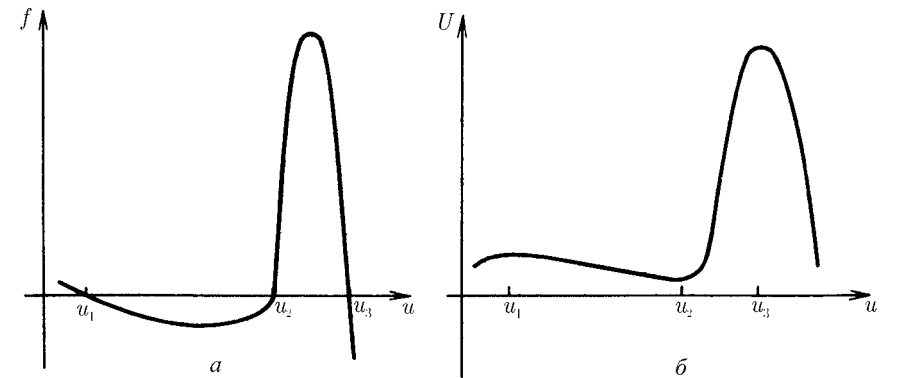


Рис. 1.12. Нелинейная функция $f(u)$ (а) и соответствующий ей потенциал $U(u)$ (б)

Допустим, что нелинейная функция $f(u)$ такова, как показано на рис. 1.12а: она достигает больших положительных значений в узком интервале от u_2 до u_3 и является отрицательной, но малой по модулю в интервале от u_1 до u_2 . Соответствующий потенциал $U(u)$, в котором согласно уравнению (1.18) движется фиктивная частица, изображен на рис. 1.12б.

Анализируя характер движения частицы в потенциале такой формы, можно заметить, что оно делится на две качественно различные стадии.

На участке $u_2 < u < u_3$ происходит быстрый разгон частицы; потенциальная сила $f(u)$ велика, и по сравнению с нею силой вязкого трения $-cu'$ можно пренебречь:

$$Du'' \approx -\frac{\partial U}{\partial u}. \quad (1.27)$$

Согласно этому уравнению, полная энергия частицы сохраняется:

$$\frac{1}{2}D(u')^2 + U(u) = U(u_3), \quad (1.28)$$

и при $u = u_2$ скорость частицы достигает значения

$$\left. \frac{du}{d\xi} \right|_{u=u_2} \approx -\sqrt{(2/D)(U(u_3) - U(u_2))}. \quad (1.29)$$

На участке $u_1 < u < u_2$, наоборот, основную роль в движении частицы играет сила вязкого трения, тогда как потенциальной силой $f(u)$ можно пренебречь:

$$Du'' \approx -cu'. \quad (1.30)$$

Уравнение (1.30) описывает торможение частицы, имевшей начальную скорость $du/d\xi$, даваемую выражением (1.29). Условие $u \rightarrow u_1$ при $\xi \rightarrow \infty$ означает, что длина пути торможения должна составить $u_2 - u_1$. Однако путь можно найти, непосредственно интегрируя уравнение (1.30):

$$u_2 - u_1 = \frac{D}{c} \left. \frac{du}{d\xi} \right|_{u_2}. \quad (1.31)$$

Сравнивая (1.31) и (1.29), находим

$$c = \frac{1}{(u_2 - u_1)} \sqrt{2D(U(u_3) - U(u_2))}. \quad (1.32)$$

Вспомним теперь, что, согласно сделанным выше предположениям величина $U(u_2)$ близка к $U(u_1)$, а значение u_2 мало отличается от u_3 . С учетом этого окончательно получим

$$c \approx \frac{\sqrt{2DA}}{u_3 - u_1}. \quad (1.33)$$

Это приближенное выражение определяет скорость распространения *быстрых волн переключения*.

В заключение кратко обсудим свойства волн переключения из состояния хотя и стационарного, но неустойчивого по отношению к слабым возмущениям. Может возникнуть вопрос: зачем вообще рассматривать эту ситуацию, поскольку, казалось бы, малые флуктуации должны всегда присутствовать и они приведут к разрушению неустойчивого состояния еще до прихода волны переключения. Существует, однако, важный класс задач о заселении среды, где такие флуктуации отсутствуют и неустойчивое стационарное состояние способно сохраняться неограниченно долго.

Пусть, например, в среде идет процесс размножения некоторых микроорганизмов, осуществляющийся путем их деления или репликации. Тогда число новых микроорганизмов, рождающихся в единицу времени в единице объема, пропорционально массе пищи m и числу уже имеющихся в этом объеме микроорганизмов n . С учетом диффузии микроорганизмов в среде распределение их численности подчиняется уравнению

$$\dot{n} = \alpha mn + D\Delta n. \quad (1.34)$$

Ясно, что масса пищи связана с числом имеющихся микроорганизмов. В простейшем случае эту связь можно считать линейной:

$$m = \begin{cases} m_0 - \beta n, & n < n_0, \\ 0, & n \geq n_0, \end{cases} \quad (1.35)$$

где $n_0 = m_0/\beta$.

Такая система имеет два стационарных состояния: $n = 0$ и $n = n_0$. Первое из них отвечает отсутствию в среде микроорганизмов, а второе представляет собой заселенное состояние среды, в котором численность микроорганизмов ограничена ресурсами пищи.

Формально, первое из таких двух стационарных состояний абсолютно неустойчиво по отношению к сколь угодно слабым возмущениям. Тем не менее, поскольку спонтанное рождение микроорганизмов запрещено, такие флуктуации в незаселенном состоянии отсутствуют и оно сохраняется, пока в него не будет внесена «затравка» в виде какого-то числа микроорганизмов.

Уравнение (1.34) было впервые исследовано А. Н. Колмогоровым, И. Г. Петровским и Н. С. Пискуновым в 1937 г. [84]. Оно описывает распространение *волн заселения* среды. В установившемся режиме такие волны являются решением уравнений вида

$$n = n(\xi), \quad \xi = x - ct \quad (1.36)$$

с граничными условиями

$$\begin{aligned} n &\rightarrow n_0 & \text{при } \xi &\rightarrow -\infty, \\ n &\rightarrow 0 & \text{при } \xi &\rightarrow +\infty. \end{aligned}$$

Скорость c такой волны заселения однозначно определяется свойствами самой среды; с точностью до численного коэффициента она равна $\sqrt{\alpha m_0 D}$ и не зависит от окончательной численности микроорганизмов n_0 в заселенном состоянии, устанавливаемом после прохождения волны.

3. Общие свойства структур в простых бистабильных средах

Проведенный в разделе 1.2 анализ оставил открытым ряд важных вопросов. В частности, устойчивы ли волны переключения в простых бистабильных средах? Являются ли они единственным возможным типом установившегося волнового режима? Что происходит при столкновении двух таких волн? Какова последующая эволюция произвольного начального распределения?

Обсуждение этих вопросов удобно проводить, переписав предварительно исходное уравнение (1.13) в эквивалентной форме

$$\frac{\partial u}{\partial t} = -\frac{\delta F[u]}{\delta u(\mathbf{r}, t)}, \quad (1.37)$$

где правая часть выражена как вариационная производная от функционала²

$$F = \int [-U(u(\mathbf{r})) + (1/2)D(\nabla u)^2] d\mathbf{r}, \quad (1.38)$$

а функция $U(u)$ дается выражением (1.17).

Из уравнения (1.37) следует, что величина F не возрастает с течением времени независимо от выбора исходного состояния (распределения) $u(\mathbf{r})$.

²Напомним определение вариационной производной:

$$\frac{\delta F[u]}{\delta u(\mathbf{r}_1)} \equiv \lim_{\int \delta u d\mathbf{r} \rightarrow 0} \frac{F[u + \delta u] - F[u]}{\int \delta u(\mathbf{r}) d\mathbf{r}},$$

где вариация $\delta u(\mathbf{r})$ локализована вблизи точки $\mathbf{r} = \mathbf{r}_1$.

Действительно, производную от F по времени можно выразить как

$$\frac{dF}{dt} = \int \frac{\delta F}{\delta u(\mathbf{r}, t)} \frac{\partial u}{\partial t} d\mathbf{r} \quad (1.39)$$

и с учетом (1.37) представить в следующем виде:

$$\frac{dF}{dt} = - \int \left(\frac{\delta F(u)}{\delta u(\mathbf{r}, t)} \right)^2 d\mathbf{r}. \quad (1.40)$$

Устойчивым стационарным состояниям соответствуют минимумы функционала F , а эволюция системы с течением времени заключается в приближении к одному из таких стационарных состояний.

Отметим прежде всего, что стационарные однородные состояния $u(\mathbf{r}) = u_1$ и $u(\mathbf{r}) = u_3$ всегда устойчивы по отношению к малым возмущениям. В самом деле, если слабо возмутить одно из таких решений, добавив к нему зависящую от координат \mathbf{r} поправку $\delta u(\mathbf{r})$, это приведет к изменению функционала F , складывающемуся из двух частей. Во-первых, из-за изменения u возрастает первое слагаемое под знаком интеграла в (1.38) (выбранное нами однородное решение отвечает максимуму U). Во-вторых, становится отличным от нуля второе слагаемое в (1.38), пропорциональное $(\nabla u)^2$ и поэтому также положительное. Мы видим, что обе появляющиеся добавки положительны и, следовательно, малое неоднородное возмущение δu приводит к возрастанию функционала F . Поэтому стационарные однородные состояния u_1 и u_3 являются устойчивыми по отношению к малым возмущениям.

Обсудим теперь эволюцию сильнонеоднородных распределений. Предположим вначале, что рассматриваемая нами система одномерна.

Пусть начальное распределение u таково, что оно нигде не превосходит значения u_2 , отвечающего минимуму функции $U(u)$ (см. рис. 1.10, 1.13 а). Такое распределение, очевидно, не может отвечать минимуму функционала F , поскольку, приближая в каждой точке x величину u к u_1 , можно монотонно уменьшать первое слагаемое $U(u)$ в этой точке, одновременно уменьшая неоднородность решения, что понижает вклад от второго слагаемого под знаком интеграла в (1.38). Подобное начальное распределение будет релаксировать к стационарному однородному состоянию $u = u_1$. Аналогично можно показать, что всякое распределение, для которого в каждой точке значение u оказывается выше u_2 , неустойчиво и релаксирует к однородному распределению $u = u_3$.

Начальное распределение может иметь также форму, показанную на рис. 1.13 б. Распределения такого типа неустойчивы, поскольку всегда

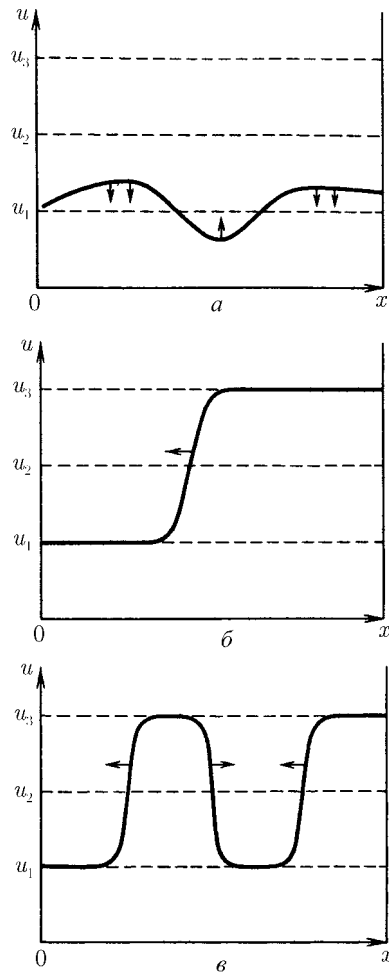
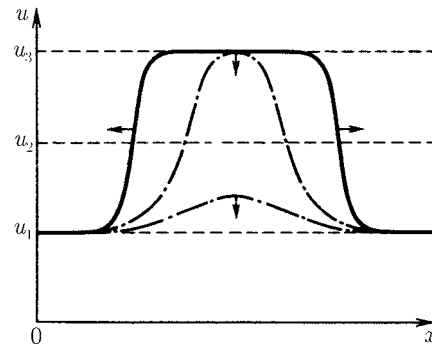


Рис. 1.13. Эволюция различных начальных распределений

можно понизить значение функционала F , сдвигая границу между областями (где u близко к u_1 или u_3) в левую или правую сторону, так что в конце концов установится стационарное однородное распределение. Так как вся-

Рис. 1.14. Пример эволюции со временем различных локальных возмущений



кое начальное распределение можно разбить на участки двух упомянутых типов (рис. 1.13в), оно всегда релаксирует к однородному распределению. Это, в частности, означает, что две столкнувшиеся волны переключения полностью гасят друг друга.

Выше мы отмечали, что оба однородных состояния $u = u_1$ и $u = u_3$ устойчивы относительно малых возмущений. Абсолютному минимуму функционала F соответствует, однако, лишь одно из них, для которого значение $U(u)$ является наибольшим (для ситуации, изображенной на рис. 1.10, это состояние $u = u_3$). Если система первоначально находилась в однородном состоянии, которое не отвечает абсолютному минимуму F , то, создав достаточно большое возмущение, ее можно перевести в наиболее устойчивое состояние, отвечающее наиболее глубокому (абсолютному) минимуму F . На рис. 1.14 показано, как эволюционируют со временем малые (штрихпунктирные линии) и большие (сплошная линия) возмущения однородного состояния $u = u_1$.

Существует значительная аналогия между рассматриваемыми явлениями и эффектами фазовых переходов 1-го рода в равновесных физических системах. Основываясь на указанной аналогии, однородные распределения, которые отвечают минимумам F , часто называют *фазами* данной системы, причем если данная фаза соответствует лишь локальному, а не абсолютному минимуму F , ее называют *метастабильной*. Метастабильная фаза неустойчива по отношению к достаточно большим возмущениям: если внутри ее возник довольно крупный *зародыш* устойчивой фазы, он начинает расти (рис. 1.14) и дает начало двум разбегающимся волнам переключения, после расхождения которых среда переходит в наиболее устойчивое стационарное однородное состояние. Критический размер зародыша определяется конкуренцией двух факторов. С одной стороны, образование зародыша, внутри которого u близко к u_3 , выгодно для системы, поскольку это уменьшает первое слагаемое в (1.38). С другой стороны, наличие зародыша означает возникновение в среде неоднородности, а поэтому — появление дополнительного положительного вклада в F , пропорционального $(\nabla u)^2$.

Особой является ситуация, когда оба минимума F при $u = u_1$ и $u = u_3$ имеют одинаковую глубину, т. е. $U(u_1) = U(u_3)$, или

$$\int_{u_1}^{u_3} f(u) du = 0. \quad (1.41)$$

В этом случае в бесконечной среде возможно стационарное сосуществование двух фаз, разделенных плоским переходным слоем. Такое состояние

не отвечает абсолютному минимуму функционала F из-за наличия неоднородности — переходного слоя между фазами. Тем не менее, если среда бесконечна, мы не можем понизить значение F , сдвигая переходный слой. Это означает, что подобное стационарное состояние обладает (нейтральной) устойчивостью.

Если размеры среды конечны и сравнимы с шириной переходного слоя, неоднородные состояния неустойчивы даже при выполнении условия (1.41). Область раздела между фазами выталкивается к границе среды, и, в конце концов, система становится однородной.

Отметим, что фактически среду следует считать бесконечной, когда ее размеры велики по сравнению с шириной переходного слоя. В этом случае при выполнении условия (1.41) среда разбивается на произвольное число областей (доменов, или страт), отвечающих различным фазам.

В конкретных приложениях часто возникает задача расчета *критического зародыша*, которого достаточно, чтобы инициировать переход всей распределенной системы из метастабильного в абсолютно устойчивое однородное состояние. В связи с этим заметим, что форме критического зародыша ставится в соответствие стационарное, но абсолютно неустойчивое неоднородное решение исходного уравнения (1.13).

В одномерном случае форма критического зародыша с центром в точке $x = 0$ определяется как решение дифференциального уравнения

$$D \frac{d^2 u}{dx^2} + f(u) = 0 \quad (1.42)$$

с граничными условиями

$$\left. \frac{du}{dx} \right|_{x=0} = 0, \quad u \rightarrow u_1, \quad \text{при } |x| \rightarrow \infty \quad (1.43)$$

(мы считаем состояние $u = u_1$ нестабильным). Если домножить уравнение (1.42) на du/dx , проинтегрировать его по x в пределах от 0 до $+\infty$ и учесть граничные условия (1.43), получим равенство

$$\int_{u_1}^{u^*} f(u) du = 0, \quad (1.44)$$

которое определяет значение $u = u^*$ в центре зародыша, т.е. при $x = 0$. Обратим внимание, что u^* всегда меньше u_3 .

В заключение приведем несколько количественных результатов, относящихся к двумерным и трехмерным бистабильным средам.

Ясно, что построенное в разделе 1.2 решение для волн переключения справедливо также для волн с плоским фронтом, распространяющихся в двух- или трехмерной бистабильной среде. Как, однако, будут распространяться волны с искривленным фронтом?

Рассмотрим вначале волны переключения в двумерных возбудимых средах. Очевидно, что всякую волну с искривленным фронтом можно разбить на совокупность достаточно малых участков, каждый из которых представляет собой сегмент окружности. Поэтому найдем закон расширения кругового фронта, радиус R которого велик по сравнению с шириной l переходного слоя.

В полярной системе координат с учетом аксиальной симметрии уравнение (1.13) имеет вид

$$\frac{\partial u}{\partial t} = f(u) + \frac{D}{r} \frac{\partial u}{\partial r} + D \frac{\partial^2 u}{\partial r^2}. \quad (1.45)$$

Заметим теперь, что фактически производная $\partial u / \partial r$ отлична от нуля лишь в пределах узкого переходного слоя порядка l вблизи значения $r = R$. Поскольку мы считаем, что $R \gg l$, во втором слагаемом справа в уравнении (1.45) можно приближенно положить $r = R$.

Пусть $c(R)$ есть мгновенная скорость распространения фронта с радиусом R . С учетом сделанных выше приближений такой фронт отвечает решению

$$u = u(\xi), \quad \xi = r - c(R)t, \quad (1.46)$$

дифференциального уравнения

$$-c(R)u' = f(u) + (D/R)u' + Du'' \quad (1.47)$$

с граничными условиями $u \rightarrow u_1$ при $\xi \rightarrow +\infty$ и $u \rightarrow u_3$ при $\xi \rightarrow -\infty$. (Заметим, что координата ξ , конечно, не может принимать сколь угодно большие отрицательные значения, но величина u оказывается экспоненциально мало отличающейся от u_3 уже на небольшом удалении от зоны перехода.) Перепишем теперь уравнение (1.47) в форме

$$-[c(R) + D/R]u' = f(u) + Du'' \quad (1.48)$$

и обратим внимание, что оно полностью совпадает с уравнением (1.16) для плоской волны переключения, если положить $c = c(R) + D/R$; идентичны также граничные условия, которым должны удовлетворять решения. Поэтому, если нам известна скорость c движения плоской волны переключения,

скорость распространения фронта с радиусом кривизны R можно выразить как

$$c(R) = c - \frac{D}{R}. \quad (1.49)$$

Следовательно, выпуклый фронт движется медленнее, чем плоский. Более того, если R достаточно мало, направление движения может смениться на обратное. Области достаточно малых размеров не растут, а сокращаются. Критический размер области, при котором она не растет и не сокращается, равен

$$R_{\text{кр}} = \frac{D}{c}. \quad (1.50)$$

Область с радиусом $R = R_{\text{кр}}$ представляет собой критический зародыш новой фазы: если радиус области чуть-чуть увеличить, она начинает расти, если же его слегка уменьшить — область сокращается и исчезает.

Следует, однако, напомнить, что все проведенное выше рассмотрение относится лишь к областям с радиусом гораздо больше ширины l переходного слоя. Поэтому выражение (1.50) применимо лишь в ситуациях, когда радиус критического зародыша достаточно велик, а именно, $R_{\text{кр}} \gg l$. Это условие выполнено, если скорость распространения плоского фронта мала.

Хотя выражение (1.49) было выведено нами только для фронта, выпуклого в направлении области, занимаемой метастабильной фазой, оно справедливо также для вогнутых участков фронта. Если ввести локальную кривизну фронта $K = 1/R$ и считать ее положительной, когда фронт выпуклый, и отрицательной для вогнутых фронтов, скорость движения такого искривленного участка фронта можно записать как

$$c(K) = c - DK. \quad (1.51)$$

Для волн распространения пламени зависимость (1.51) впервые получена в [204]. Заметим, что она обеспечивает устойчивость плоского фронта: если на нем появился выступ, т. е. выпуклый участок, он станет сдвигаться с меньшей скоростью и возмущение формы фронта затухнет.

Для волн переключения в трехмерных возбудимых средах рассмотрение проводится совершенно аналогично. Каждый участок фронтальной поверхности характеризуется своими двумя главными кривизнами K_1 и K_2 . Скорость его смещения равна

$$c(K_1, K_2) = c - D(K_1 + K_2). \quad (1.52)$$

Критический зародыш имеет радиус

$$R_{\text{кр}} = \frac{2D}{c}. \quad (1.53)$$

Выражение (1.53) справедливо лишь при $R_{\text{кр}} \gg l$, где l — ширина слоя, в котором непосредственно происходит переключение от одного стационарного состояния к другому.

4. Бегущие импульсы в возбудимых средах

Чтобы получить вместо бистабильной возбудимую активную среду, необходим какой-либо механизм, который обеспечивал бы возвращение элементов к исходному состоянию после прохождения волны переключения. Рассмотрим, например, следующую ситуацию. Пусть в среде осуществляется процесс горения, но скорость выделения теплоты q зависит не только от температуры Θ , но и от концентрации v некоторого вещества («ингибитора»): $q = q(\Theta, v)$. Ингибитор ухудшает условия горения: чем больше его концентрация, тем меньше выделение теплоты.

При заданном значении v распространение волны горения описывается уравнением

$$\frac{\partial \Theta}{\partial t} = f(\Theta, v) + \chi \frac{\partial^2 \Theta}{\partial x^2}, \quad (1.54)$$

где

$$f(\Theta, v) = -\gamma(\Theta - \Theta_1) + \frac{q(\Theta, v)}{C}. \quad (1.55)$$

Как отмечалось в разделе 1.2, скорость и направление перехода в распространяющейся волне зависят от величины

$$A(v) = \int_{\Theta_1}^{\Theta_3} f(\Theta, v) d\Theta. \quad (1.56)$$

Если $A > 0$, то при распространении волны происходит переключение из холодного состояния $\Theta = \Theta_1$ в состояние с высокой температурой $\Theta = \Theta_3$, а, значит, распространяющаяся по среде волна будет волной загорания. Если же $A < 0$, то после прохождения волны среда переходит из горячего в холодное состояние и мы имеем *волну гашения* (рис. 1.15).

Поскольку A зависит от концентрации v ингибитора, то, меняя последнюю, можно управлять волной переключения — изменять ее скорость и даже направление движения. При низких концентрациях v волна будет волной загорания, а при высоких — волной гашения.

Пусть теперь ингибитор выделяется в ходе горения как побочный продукт реакции. Одновременно необходимо предположить, что он способен

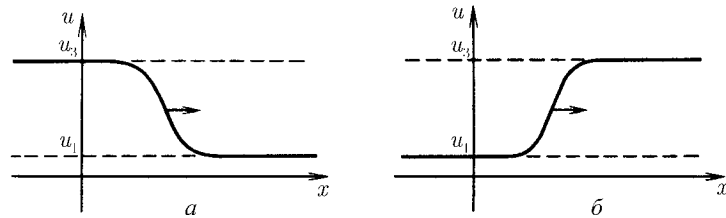


Рис. 1.15. Волны переключения при $A > 0$ (а) и $A < 0$ (б)

распадаться или уходить в окружающую среду — иначе ингибитор просто накапливался бы с течением времени. С учетом этих двух процессов изменение локальной концентрации ингибитора описывается уравнением

$$\frac{\partial v}{\partial t} = -\frac{1}{\tau}(v - \bar{v}(\Theta)). \quad (1.57)$$

Здесь $\bar{v}(\Theta)$ есть равновесная концентрация ингибитора, устанавливающаяся, если искусственно поддерживать заданную температуру Θ ; она является монотонно растущей функцией Θ . Мы будем считать, что характерное время установления концентрации ингибитора (τ) велико по сравнению с длительностью процесса перехода от холодного к горячему состоянию, т. е. по сравнению с длительностью вспышки загорания. Кроме того, предполагается, что диффузия ингибитора мала и ею можно пренебречь.

Уединенный бегущий импульс в среде, описываемой уравнениями (1.54) и (1.57), имеет вид, показанный на рис. 1.16. Его можно разбить на резкие фронт и спад, где происходит быстрое изменение температуры, а также на области более медленного изменения — вершину импульса и следующий за ним «хвост». Концентрация ингибитора не совершает быстрых изменений — она медленно возрастает при прохождении вершины импульса, а затем столь же плавно спадает до исходного нулевого значения.

Фронт и спад импульса — это фактически две волны (загорания и гашения), движущиеся друг за другом с одинаковой скоростью. Загорание происходит в исходном состоянии, т. е. в отсутствие ингибитора. После того как загорание произошло и температура резко повысилась, концентрация ингибитора возросла, и, когда она достигла некоторой критической величины, горение прекратилось. Температура резко снижается, выделение ингибитора приостанавливается, и его концентрация медленно спадает. В результате среда возвращается в исходное состояние. Теперь по ней можно пустить еще одну волну горения, и т. д.

Распространение уединенного бегущего импульса с последующим возвратом среды к исходному состоянию может быть обеспечено также другим механизмом, не связанным с выделением ингибитора. Рассмотрим в качестве гипотетического примера горение сухой травы в степи. При каждом прохождении волны пламени вся трава сгорает, но затем она очень медленно отрастает вновь, так что степь можно поджечь еще раз.

Два приведенных примера указывают на *типичные ситуации*, в которых можно ожидать появления уединенных бегущих импульсов. Разумеется, совершенно не обязательно, чтобы переменная Θ имела смысл именно температуры.

В химической активной среде с реакцией Белоусова–Жаботинского аналогом горения является процесс окисления одного из веществ. Реакция протекает в слабом растворе и не сопровождается существенными тепловыми эффектами. Если при горении активация («поджиг») элементов среды обусловлена теплопроводностью, то в данном случае главную роль в распространении волны играет диффузия молекул окислителя. Реакция является *самоускоряющейся* (или, как говорят, содержит автокаталитическую стадию).

Структуры, возникающие в среде с реакцией Белоусова–Жаботинского, очень удобны для экспериментального изучения. Реакция идет в обычных лабораторных условиях в тонком слое раствора. Движение волн возбуждения сопровождается локальным изменением цвета этого раствора. Соответствующие процессы достаточно медленные (импульс возбуждения движется со скоростью несколько миллиметров в минуту) и могут наблюдаться невооруженным глазом (размер самых мелких деталей около 1 мм).

Детальная схема реакции Белоусова–Жаботинского сложна и насчитывает более 20 промежуточных стадий. В упрощенном виде, учитывая различие в скоростях отдельных реакций, ее можно свести [251, 119] к системе уравнений для концентраций u и v двух веществ:

$$\frac{\partial u}{\partial t} = u(1 - u) - \frac{v(u - a)}{(u + a)} + D \frac{\partial^2 u}{\partial x^2}, \quad \frac{\partial v}{\partial t} = -\frac{1}{\tau}(v - bu). \quad (1.58)$$

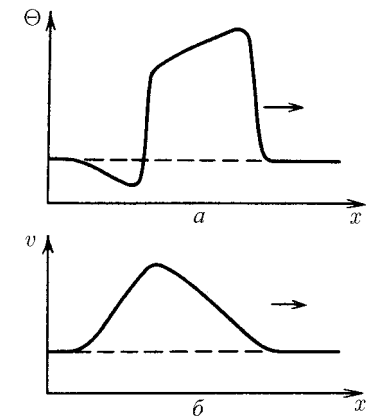


Рис. 1.16. Уединенный бегущий импульс в возбудимой среде. Показаны профили температуры (а) и концентрации ингибитора (б)

Вторая из переменных, v , медленно меняется во времени и играет роль концентрации ингибитора, подавляющего процесс окисления; его диффузия в уравнениях (1.58) не учитывается. Переменную u обычно называют концентрацией активатора.

Близкими уравнениями вида

$$\begin{aligned} C \frac{\partial U}{\partial t} &= I_{\text{Na}}(U) - I_{\text{K}}(U, \sigma) + \frac{1}{R} \frac{\partial^2 U}{\partial x^2}, \\ \frac{\partial \sigma}{\partial t} &= -\frac{1}{\tau} [\sigma - \bar{\sigma}(U)] \end{aligned} \quad (1.59)$$

приближенно описывается распространение импульса возбуждения вдоль нервного волокна. В первом из этих уравнений C — емкость мембраны на единицу длины волокна; R — удельное сопротивление аксоплазмы, заполняющей волокна; I_{Na} и I_{K} — значения локального натриевого и калиевого токов через мембрану; $\bar{\sigma}(U)$ — установившееся значение проводимости по отношению к калиевому току при заданной разности потенциалов. Обратим внимание, что процесс подстраивания калиевой проводимости σ к имеющейся разности потенциалов U на мембране является медленным и характеризуется большим временным масштабом τ .

Аналогичные уравнения справедливы для мышечных волокон, которые образуют поперечно-полосатые и гладкие мышцы, выстилающие полые органы, а также ткань сердца. Подробнее о биологических и других примерах возбуждимых сред можно прочитать в [36, 37, 89].

Итак, широкий класс возбуждимых сред самой различной природы описывается уравнениями типа

$$\frac{\partial u}{\partial t} = f(u, v) + D \Delta u, \quad \frac{1}{\epsilon} \frac{\partial v}{\partial t} = -v + \bar{v}(u), \quad (1.60)$$

где $\bar{v}(u)$ — монотонно возрастающая функция, а параметр ϵ мал ($\epsilon \ll 1$). Нас интересует случай, когда есть всего одно устойчивое однородное состояние такой среды. Это имеет место, если две нуль-изоклины $f(u, v) = 0$ и $v = \bar{v}(u)$ однородной системы (1.60) без диффузии пересекаются в одной единственной точке A , отвечающей состоянию покоя (рис. 1.17).

Согласно уравнениям (1.60) величина v является медленной — характерное время ее изменения имеет порядок $1/\epsilon$, тогда как величина u может существенно изменяться за гораздо меньшее время порядка единицы (в выбранных нами безразмерных переменных). Следовательно, на фронте и спаде бегущего импульса, где происходит резкое изменение u , ее в первом приближении можно считать постоянной. Это означает, что поведение u

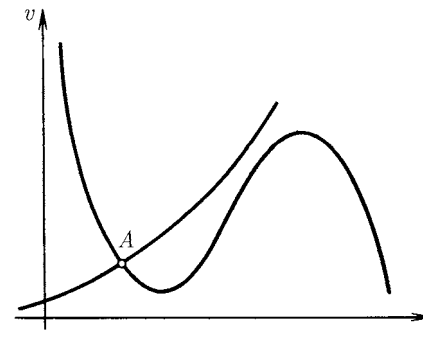


Рис. 1.17. Нуль-изоклины уравнений, описывающих возбудимую среду

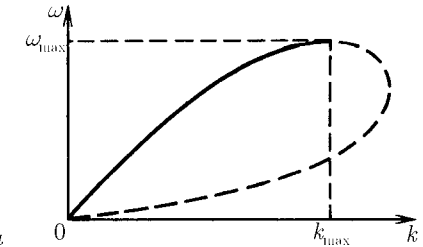


Рис. 1.18. Типичный вид закона дисперсии нелинейных периодических волн в возбудимой среде

на фронте и спаде описывается первым из уравнений (1.60), в котором v — заданный параметр. Но оно в точности совпадает с уравнением (1.13), рассматривавшемся при изучении волн переключения.

Следовательно, фронт импульса есть волна переключения, и ее скорость (а следовательно, скорость всего бегущего импульса) определяется решением первого из уравнений (1.60) при значении $v = v_0$, отвечающем состоянию покоя. Спад импульса — волна обратного переключения, которая должна бежать вслед за фронтом с той же самой скоростью. Это условие определяет значение переменной v , при котором происходит обратное переключение в установившемся бегущем импульсе.

Более строго задача о расчете уединенного бегущего импульса формулируется следующим образом. Перейдем в систему координат $\xi = x - V_0 t$, в которой импульс покоится. Тогда уравнения (1.60) перейдут в систему двух обыкновенных дифференциальных уравнений

$$-V_0 u' = f(u, v) + D u'', \quad -(V_0/\epsilon) v' = \bar{v}(u) - v, \quad (1.61)$$

которую мы должны решать, потребовав выполнения условий $u \rightarrow u_0$ и $v \rightarrow v_0$, $\xi \rightarrow \pm\infty$, где (u_0, v_0) — стационарное однородное состояние покоя. Наличие такого решения означает существование петли сепаратрисы стационарной седловой точки A , отвечающей состоянию покоя, а это возможно лишь при определенном значении параметра V_0 в уравнениях (1.61). Таким образом, скорость уединенного импульса находится путем решения нелинейной задачи на собственное значение. При малых значениях ϵ спра-

ведливо разложение [151]

$$V_0 \approx c_0(l - \zeta\epsilon), \quad (1.62)$$

где c_0 представляет собой скорость волны переключения, описываемой первым из уравнений (1.60) при $v = v_0$, а ζ есть некоторый коэффициент порядка единицы.

Разумеется, уравнения (1.60) допускают и периодическое распространение импульсов. Если период L следования импульсов достаточно велик, скорость их движения совпадает со скоростью одиночного импульса. Однако с уменьшением L скорость движения уменьшается, поскольку фронт следующего импульса начинает наступать на «хвост» предыдущего, и распространение фронта идет на фоне остаточной концентрации ингибитора, превышающей стационарный уровень v_0 .

Периодическая последовательность импульсов описывается решением вида $u = u(\eta)$, $v = v(\eta)$, где $\eta = kx - \omega t$; $u(\eta + 2\pi) = u(\eta)$, $v(\eta + 2\pi) = v(\eta)$. Оно удовлетворяет системе обыкновенных дифференциальных уравнений

$$-\omega u' = f(u, v) + Dk^2 u'', \quad -(\omega/\epsilon)v' = \bar{v}(u) - v \quad (1.63)$$

и соответствует предельному циклу этой системы. При произвольных значениях параметров ω и k период такого предельного цикла по фазовой координате η может быть различным. Потребовав, чтобы он был равен 2π , мы получим соотношение, связывающее ω и k , которое естественно назвать *законом дисперсии* нелинейных периодических волн: $\omega = \Omega(k, \epsilon)$.

Типичный вид закона дисперсии показан на рис. 1.18. Правая ветвь дисперсионной зависимости, обозначенная штриховой кривой, описывает медленные волны и соответствует неустойчивым режимам. Вообще говоря, область неустойчивости захватывает и часть области быстрых волн (сплошная кривая). Однако до настоящего времени отсутствуют аналитические расчеты положения границы устойчивости. Поэтому часто приближенно полагают, что устойчива вся левая ветвь дисперсионной зависимости.

Существует, таким образом, минимальный пространственный период устойчивого следования импульса $L_{\min} = 2\pi/k_{\max}$, по порядку величины он равен ширине отдельно бегущего импульса. Ему соответствует максимальная возможная частота следования импульсов $\omega_{\max} \sim \sqrt{\epsilon}$.

Как отмечалось в разделе 1.2, даже аналитический расчет скорости волны переключения в простых бистабильных средах обычно невозможен. В еще большей степени это относится к характеристикам бегущих импульсов в возбудимых средах. Обычно здесь приходится прибегать к численным

расчетам. Важным исключением является так называемая *модель Ринцеля–Келлера*, описываемая уравнениями

$$\frac{\partial u}{\partial t} = -u + \bar{u}\Theta(u - u_0) - v + \frac{\partial^2 u}{\partial x^2}, \quad \frac{1}{\epsilon} \frac{\partial v}{\partial t} = u, \quad (1.64)$$

где ступенчатая функция Хевисайда определена как $\Theta(z) = 0$ при $z < 0$ и $\Theta(z) = 1$ при $z \geq 0$.

Закон дисперсии нелинейных периодических волн в модели Ринцеля–Келлера был аналитически рассчитан в [207]. Максимальная частота одномерной последовательности волн дается в этой модели выражением

$$\omega_{\max} = \frac{\pi\epsilon}{\operatorname{arctanh}(\sqrt{3\epsilon/2\mu})}, \quad (1.65)$$

где $\mu = (1/2) - u_0/u$. Самоподдерживающееся распространение волн невозможно при $\mu < \sqrt{3\epsilon/2}$. Максимальное волновое число составляет

$$k_{\max} = \frac{\pi\sqrt{\epsilon/3}}{\operatorname{arctanh}(\sqrt{3\epsilon/2\mu})}. \quad (1.66)$$

При выполнении условия $\mu \ll 1$ закон дисперсии плоских волн вплоть до значения $k = k_{\max}$ хорошо аппроксимируется неявным выражением

$$\omega = 4k\mu\operatorname{th}\pi\nu - 3\nu k^2, \quad \text{где } \nu = \frac{\epsilon}{\omega}. \quad (1.67)$$

Заметим, что решение для периодической последовательности импульсов, имеющей пространственный период следования L , описывает также и движение одиночного импульса по окружности параметра L .

Для двумерной возбудимой среды одиночному бегущему импульсу соответствует уединенная волна возбуждения. Обсудим вкратце, какой может быть зависимость скорости распространения такой волны от кривизны ее фронта.

Чтобы найти зависимость скорости волны V от кривизны ее фронта K , введем на плоскости мгновенную полярную систему координат, в которой локально фронт описывается окружностью радиуса $R = 1/K$ с центром в начале координат. Если радиус кривизны гораздо больше ширины волны (т. е. ширины импульса возбуждения), приближенно можно записать

$$\frac{\partial u}{\partial t} = f(u, v) + \frac{D}{R} \frac{\partial u}{\partial r} + D \frac{\partial^2 u}{\partial r^2}, \quad \frac{1}{\epsilon} \frac{\partial v}{\partial t} = \bar{v}(u) - u. \quad (1.68)$$

Радиальному распространению со скоростью V отвечает решение $u = u(\xi)$, $v = v(\xi)$, где $\xi = r - Vt$. Оно удовлетворяет уравнениям

$$-(V + DK)u' = f(u, v) + Du'', \quad -(V/\epsilon)v' = \bar{v}(u) - u. \quad (1.69)$$

Нетрудно заметить, что они переходят в уравнения (1.61), описывающие распространение одиночного импульса, если положить $\tilde{V} = V + DK$ и заменить в (1.61) ϵ на $\epsilon^* = (V + DK)\epsilon/V$. Следовательно, \tilde{V} должно выражаться через ϵ^* по закону (1.62), полученному для одиночного импульса.

Подставляя выражения для \tilde{V} и ϵ^* в (1.62) и решая возникающее уравнение, находим

$$\frac{V}{c_0} = \frac{1}{2} \left(1 - \zeta\epsilon - \frac{DK}{c_0} \right) \pm \frac{1}{2} \left[\left(1 - \zeta\epsilon - \frac{DK}{c_0} \right)^2 - 4\zeta\epsilon \frac{DK}{c_0} \right]^{1/2}. \quad (1.70)$$

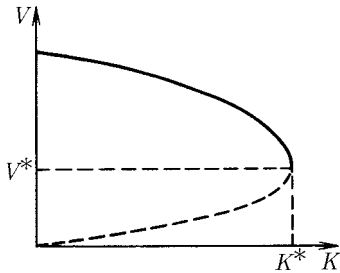


Рис. 1.19. Зависимость скорости распространения волны в двумерной возбудимой среде от кривизны ее фронта

Найденная зависимость показана на рис. 1.19. Никакие волны не могут распространяться с кривизной, превышающей

$$K^* = \frac{c_0}{D} (1 - 2\sqrt{\zeta\epsilon}). \quad (1.71)$$

При $K < K^*$ имеются две ветви, из которых нижняя (штриховая кривая) всегда неустойчива. Данные численного моделирования [65] свидетельствуют, что для верхней ветви (сплошная кривая) волна остается устойчивой, вплоть до значений кривизны ее фронта, довольно близких к предельному возможному значению K^* .

Почувительно сравнить особенности распространения искривленных волн переклочки в простых бистабильных средах с полученными выше результатами для движения искривленных волн возбуждения в двумерных возбудимых средах. Как отмечалось в разделе 1.3, скорость движения волны переключения связана с ее кривизной простым линейным законом $c = c_0 - DK$. В частности, при достаточно большой кривизне фронт может остановиться и даже начать движение в обратном направлении. Для волн возбуждения это невозможно. При достижении достаточно большой кривизны фронта, близкой к величине K^* , волна теряет устойчивость и разрушается. Она не может остановиться или поменять направление движения. Волны, кривизна фронта которых близка к пределу устойчивости, продвигаются вперед со скоростью, близкой к $V^* = \sqrt{\zeta\epsilon}$.

5. Процессы в возбудимых средах, образованных клеточными автоматами

Здесь рассматриваются волновые процессы в дискретных сетях, образованных связанными между собой возбудимыми элементами. В отличие от раздела 1.4 мы не предполагаем, что отдельный такой элемент описывается дифференциальными уравнениями с непрерывно меняющимися переменными. Вместо этого для каждого элемента теперь заданы определенные *правила перехода* между состояниями: состояние данного элемента в момент времени $n + 1$ однозначно определено состоянием этого элемента и его ближайших соседей по сети в предыдущий, дискретный момент времени n . Это означает, что мы моделируем возбудимую среду как сеть *клеточных автоматов*.

Прибегать к таким моделям приходится главным образом по двум обстоятельствам.

Даже если известны точные дифференциальные уравнения (например, вида (1.60)), которые описывают интересующую нас среду, как правило, не удается найти аналитические решения этих уравнений и необходимо проводить численные расчеты. Если задача заключается в исследовании волновых процессов, связанных с образованием и развитием структур в таких двумерных (а тем более трехмерных) средах, расчеты оказываются чрезвычайно трудоемкими. В этой ситуации (в особенности когда нас интересуют в первую очередь качественные результаты) можно отказаться от численного интегрирования соответствующих дифференциальных уравнений и обратиться к анализу гораздо более простых систем, представляющих собой сети клеточных автоматов. Как показывают расчеты, их динамика достаточно хорошо воспроизводит основные черты соответствующих явлений в распределенных возбудимых средах.

Кроме того, есть много примеров, когда именно модели дискретных сетей, а не приближение сплошной среды, являются наиболее адекватными. Важнейший из них — нейронные сети, а также живые ткани, образованные переплетенными мышечными волокнами (в частности, ткань сердечной мышцы).

В 1946 г. Н. Винер и А. Розенблют [38] предложили простую модель, позволяющую исследовать различные режимы распространения возбуждения по однородной нейронной сети (со связями лишь между соседними нейронами) и ткани сердечной мышцы. Согласно такой модели, среда представляет собой совокупность отдельных элементов, способных находиться в одном из трех состояний: покоя, возбуждения и рефрактерности. Будучи приведенным в возбужденное состояние, элемент пребывает в нем в те-

чение некоторого времени³, затем переходит в состояние рефрактерности и лишь после этого возвращается в состояние покоя. Элемент среды может перейти в возбужденное состояние благодаря внешнему воздействию либо если по соседству с ним оказался возбужденный элемент. Переход в возбужденное состояние возможен лишь из состояния покоя.

Ниже мы подробно рассмотрим обобщенную модель Винера–Розенблюта, в которой дополнительно учтен пороговый характер активации и временное суммирование приходящих сигналов [67].

Пусть двумерная сеть образована элементами, нумеруемыми парой целых индексов i и j . Состояние каждого элемента задается двумя величинами ϕ_{ij}^n и u_{ij}^n , верхний индекс n указывает дискретный момент времени. Переходы между состояниями осуществляются согласно следующему набору правил:

$$\phi_{ij}^{n+1} = \begin{cases} \phi_{ij}^n + 1, & \text{если } 0 < \phi_{ij}^n < \tau_e + \tau_r, \\ 0, & \text{если } \phi_{ij}^n = \tau_e + \tau_r, \\ 0, & \text{если } \phi_{ij}^n = 0, u_{ij}^{n+1} < h, \\ 1, & \text{если } \phi_{ij}^n = 0, u_{ij}^{n+1} \geq h \end{cases} \quad (1.72)$$

(τ_e — дискретная длительность возбужденного состояния, τ_r — дискретная длительность рефрактерного состояния). Переменная ϕ_{ij}^n имеет смысл фазы элемента, находящегося в узле (i, j) , в момент времени n . Значение фазы $\phi_{ij}^n = 0$ отвечает состоянию «покоя». Если $0 < \phi_{ij}^n \leq \tau_e$, элемент находится в состоянии «возбуждения». Состоянию рефрактерности соответствуют значения $\tau_e < \phi_{ij}^n \leq \tau_e + \tau_r$. Согласно (1.72), переход элемента из узла (i, j) от состояния покоя к возбужденному состоянию происходит при превышении величиной u_{ij}^{n+1} порога h . Затем элемент совершает фиксированную последовательность переходов, при каждом из которых значение фазы возрастает на единицу, а при достижении фазой значения $\tau_e + \tau_r$ возвращается к исходному состоянию покоя.

Величина u_{ij}^n есть концентрация активатора в узле (i, j) в момент времени n . Мы считаем, что в узлах сети идут процессы распада и производства активатора, причем активатор производится лишь элементами, находящимися в данный момент времени в возбужденном состоянии:

$$u_{ij}^{n+1} = gu_{ij}^n + \sum_{k,l} C(k,l)J_{i+k,j+l}^n, \quad (1.73)$$

³В [38] продолжительность возбужденного состояния предполагалась пренебрежимо малой (мгновенной); модель с конечным временем возбуждения была предложена позднее А. Розенблютом [239].

где

$$J_{ij}^n = \begin{cases} 1, & \text{если } 0 < \phi_{ij}^n \leq \tau_e, \\ 0, & \text{если } \tau_e < \phi_{ij}^n \leq \tau_e + \tau_r \text{ или } \phi_{ij}^n = 0. \end{cases} \quad (1.74)$$

Коэффициенты $C(k, l)$ быстро спадают с увеличением расстояния до данного узла. Обычно достаточно ограничиться учетом вклада от ближайших соседей, т. е. положить

$$C(k, l) = \begin{cases} 1, & \text{если } |k| \leq 1, |l| \leq 1, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (1.75)$$

Модель, задаваемая правилами перехода (1.72)–(1.74), переходит в первоначальную модель Винера–Розенблюта, если в ней положить $h = 1$, $g = 0$ и выбрать $C(k, l)$ в виде (1.75). Тогда возбужденный элемент возбуждает на следующем шаге по времени любой из покоящихся элементов, которые находятся по соседству с ним. Если же $h > 1$, $0 < g \leq 1$, то элемент выжидает перехода, пока не будет накоплено пороговое значение h концентрации активатора — либо за счет временной суммации, либо благодаря притоку активатора одновременно от нескольких соседних элементов. Время выжидания тем больше, чем меньше число соседних элементов возбуждено. Благодаря этому появляется зависимость скорости распространения волны возбуждения от формы ее фронта.

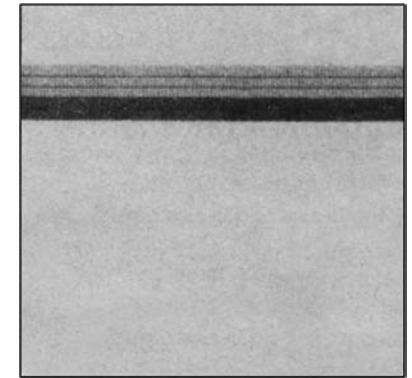


Рис. 1.20. Плоская волна возбуждения в среде из клеточных автоматов. Темные участки — область возбуждения, серые — область рефрактерности, белые — область покоя

Таким образом, параметр h имеет смысл порога возбуждения данной среды, а величина $1/(1-g)$ есть характерное время полураспада активатора.

При различных соотношениях параметров рассматриваемая модель демонстрирует значительное разнообразие режимов⁴.

Если $h = 3$, то по среде могут распространяться плоские волны возбуждения, движущиеся со скоростью $V = 1$ (рис. 1.20). Их фронт смещается

⁴Изолагаемые ниже результаты получены одним из авторов (А. С. М.) совместно с Б. Доржсурэнгийнгом.

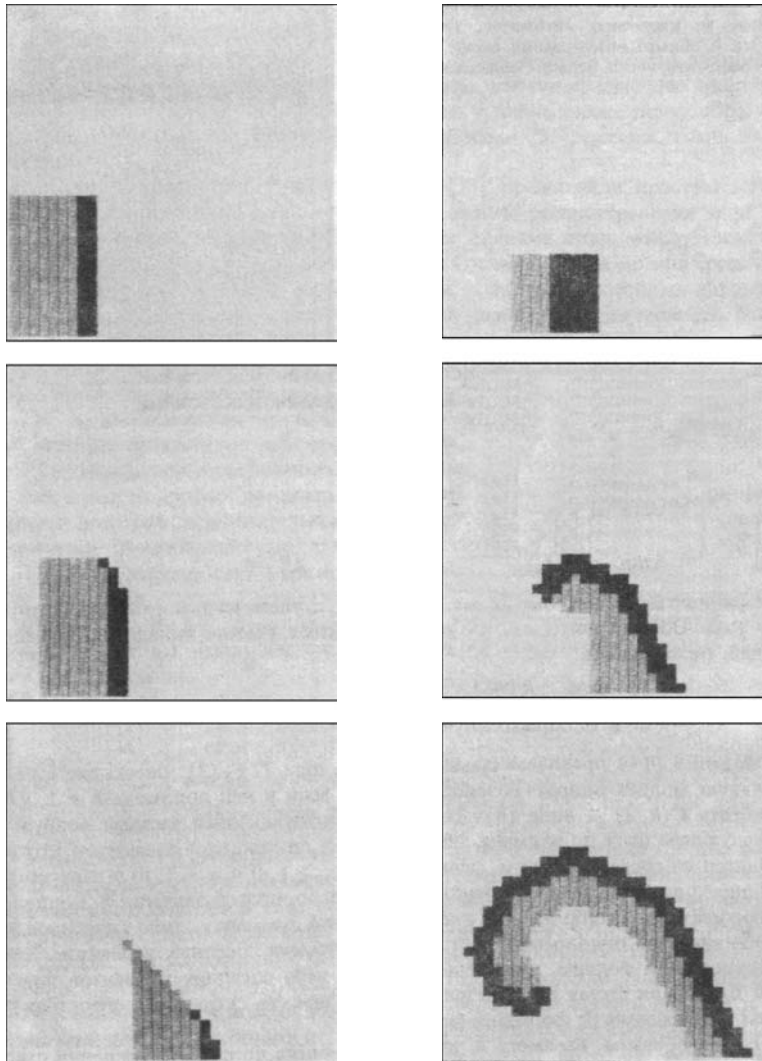


Рис. 1.21. Эволюция начального распределения в виде полуволны для среды с недостаточно высокой возбудимостью ($\tau_e = 2$, $\tau_r = 7$, $g = 0,3$, $h = 3$)

Рис. 1.22. Рождение спиральной волны из обрыва в среде с достаточно высокой возбудимостью ($\tau_e = 5$, $\tau_r = 7$, $g = 0,3$, $h = 3$)

на один элемент решетки за каждый шаг по времени; волна имеет суммарную ширину $\tau_e + \tau_r$. При $h > 3$ плоские волны движутся с меньшими скоростями ($V \leq 1/2$), т. е. фронт смещается на каждый второй либо на каждый третий и т. д. такт по времени. Чтобы распространение волны в этом случае было возможно, распад активатора не должен быть слишком быстрым, т. е. g не должно превышать определенное критическое значение, зависящее от h и τ_e .

Поскольку состояние среды до и после прохождения волны возбуждения одинаково, ничто не препятствует созданию начального распределения в виде оборванной полуволны (рис. 1.21). Какой будет последующая эволюция такого начального распределения? Оказывается, что ответ существенно зависит от эффективной *возбудимости* среды, определяемой значением порога возбуждения h , длительностью возбужденного состояния τ_e и значением коэффициента g .

Если возбудимость среды низка (рис. 1.21), полуволна сокращается с течением времени и начальное возмущение затухает. Подобрать значения параметров, можно добиться того, чтобы волна двигалась вперед, не сокращаясь и не «прорастая».

При более высокой возбудимости среды волна возбуждения «прорастает» на своем свободном конце, одновременно отставая. В результате из обрыва волны возбуждения рождается вращающаяся спиральная волна (рис. 1.22). При не очень высокой возбудимости эта волна является достаточно «рыхлой»: между витками спирали расположены участки, где элементы пребывают в состоянии покоя. Свободный конец совершает движение вокруг покоящейся области — ядра спиральной волны, которое в данном случае велико. Если возбудимость среды высока, спиральная волна «намотана» более плотно. В этом случае ядро спиральной волны малó.

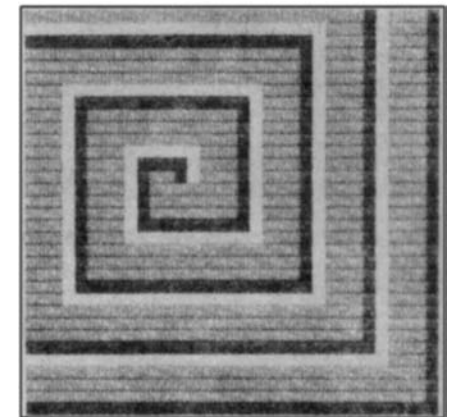


Рис. 1.23. Спиральная волна в модели Винера–Розенблюта

Спиральная волна для модели Винера–Розенблюта показана на рис. 1.23. Мы видим, что здесь ядро спиральной волны (покоящаяся область в ее центре) вообще отсутствует.

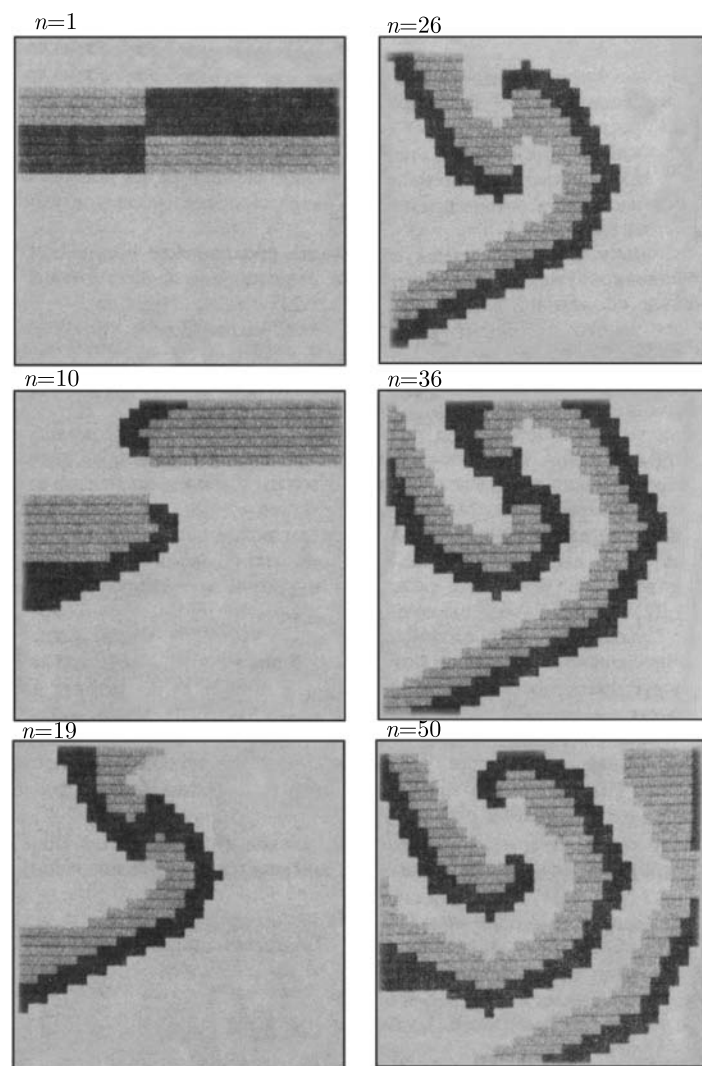


Рис. 1.24. Образование двухрукавной спиральной волны ($\tau_e = 5$, $\tau_r = 7$, $g = 0, 4$, $h = 3$)

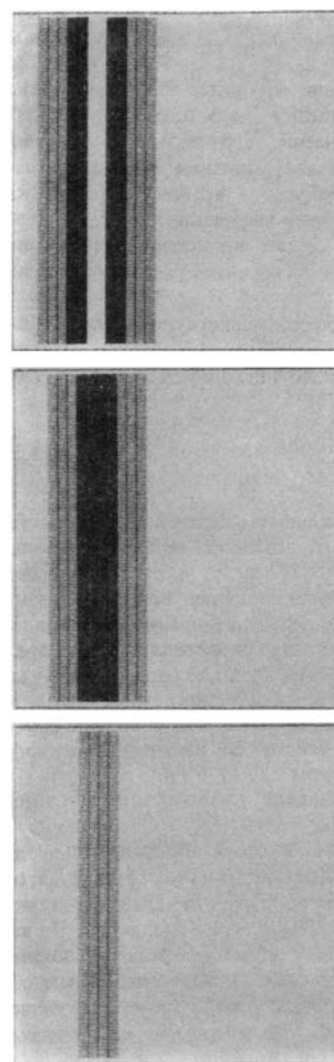


Рис. 1.25. Аннигиляция двух волн в результате их столкновения

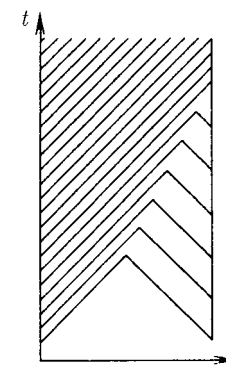


Рис. 1.26. Схематическое изображение временной развертки волн от двух источников с разными периодами генерации

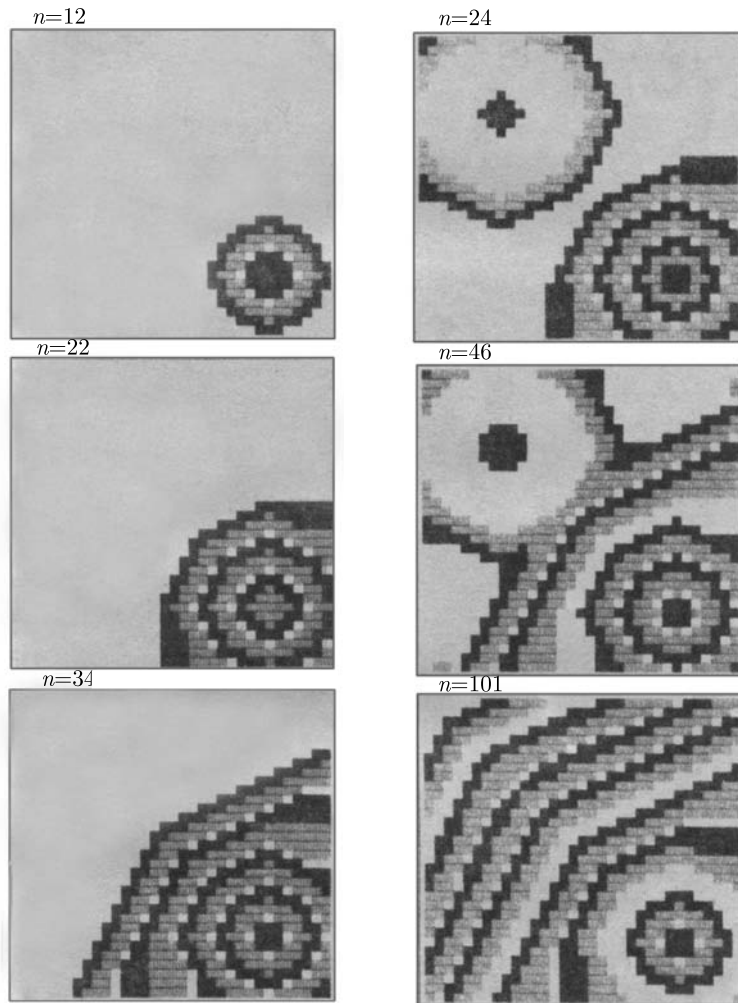


Рис. 1.27. Пейскекер, генерирующий периодическую последовательность волн ($\tau_e = 3$, $\tau_r = 4$, $g = 0, 5$, $h = 3$, $T = 9$)

Рис. 1.28. Подавление низкочастотного пейскекера высокочастотным

Если задать начальное распределение в виде двух полуволн, ориентированных навстречу друг другу, из него развивается двухрукавная спиральная волна (рис. 1.24); обратите внимание на нестационарность процесса в центральной области такой волны и последовательное «перезамыкание» волн.

Из приведенных иллюстраций видно, что сталкивающиеся волны возбуждения аннигилируют, т. е. гасят друг друга. Это весьма общее свойство волн в возбудимых средах. Действительно, при столкновении двух таких волн возбужденные элементы оказываются зажатыми с двух сторон элементами, находящимися в состоянии рефрактерности, и не способны передать возбуждение другим элементам среды (рис. 1.25).

Аннигиляция волн при их столкновении приводит к одному очень важному эффекту. Если в такой среде действует несколько периодических источников волн, то с течением времени источник, генерирующий волны с максимальной частотой, подавляет все остальные источники. Как видно из рис. 1.26, где схематически изображена временная развертка волн от двух источников с разными периодами генерации T_1 и T_2 , с каждым последующим столкновением точка, где оно происходит, смещается все ближе в направлении к более медленному источнику, и в конце концов деятельность его подавляется.

Чтобы получить локальный периодически действующий источник волн, или *пейскекер*, достаточно взять группу автоколебательных элементов, для которых фаза ϕ меняется не согласно (1.72), а по простому закону

$$\phi_{ij}^{n+1} = \begin{cases} \phi_{ij}^n + 1, & \text{если } \phi_{ij}^n < T - 1, \\ 0, & \text{если } \phi_{ij}^n \geq T - 1. \end{cases} \quad (1.76)$$

Ясно, что каждый такой элемент будет совершать, независимо от своего окружения, колебания с периодом T . Если поместить достаточно большую группу таких элементов в возбудимую среду, последняя играет роль источника концентрических распространяющихся волн возбуждения (рис. 1.27). Заметим, что фактический период генерации волн совпадает с T , лишь если $T \geq \tau_e + \tau_r + 1$. В противном случае среда «пропускает» некоторые такты активации, поскольку автоколебательные элементы, находящиеся в состоянии рефрактерности, и не способны породить волну возбуждения. В результате пейскекер начинает генерировать волны с более высоким кратным периодом $2T$, $3T$ и т. д.

Рис. 1.28 иллюстрирует эффект подавления низкочастотного пейскекера высокочастотным.

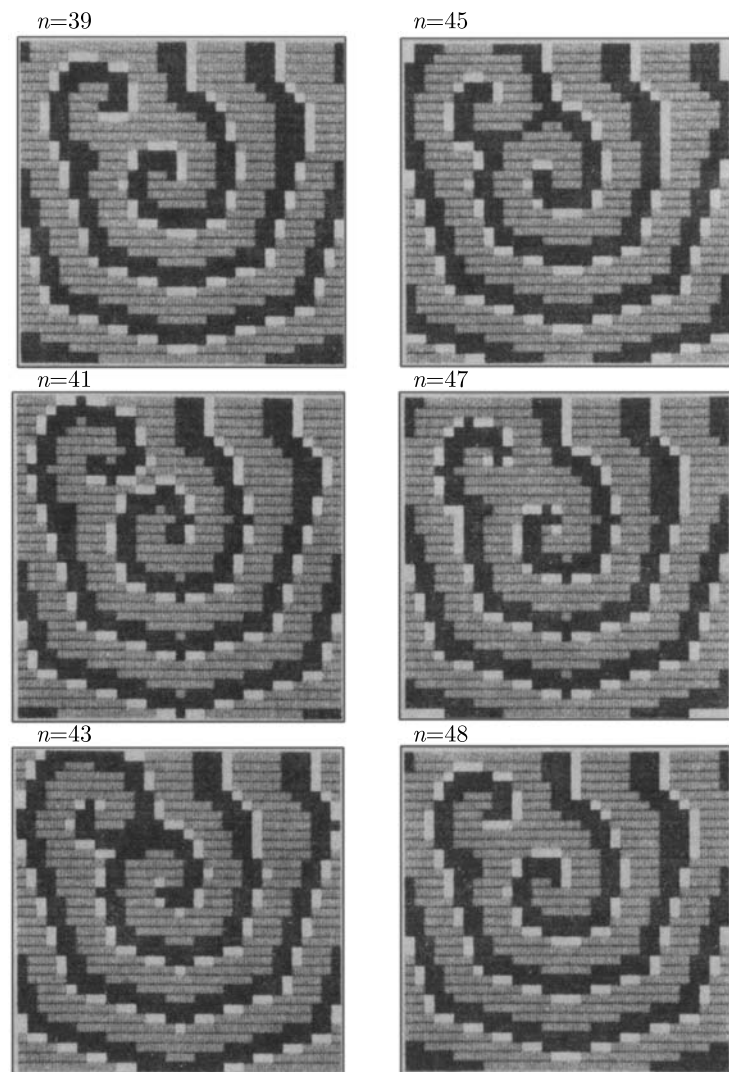


Рис. 1.29. Сосуществование двух спиральных волн в возбудимой среде ($\tau_e = 3$, $\tau_r = 5$, $g = 0, 55$, $h = 2, 5$)

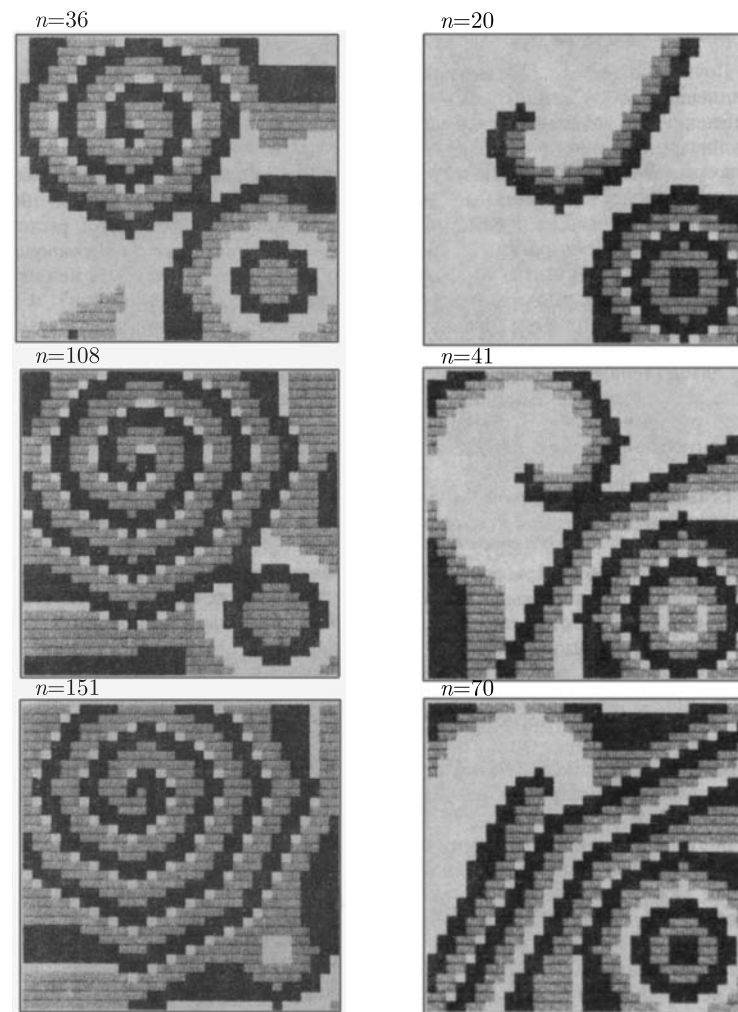


Рис. 1.30. Подавление низкочастотного пейсмекера спиральной волной

Рис. 1.31. Вытеснение спиральной волны высокочастотным пейсмекером ($\tau_e = 4$, $\tau_r = 4$, $g = 0, 4$, $h = 3$, $T = 6$)

В отличие от пейсмекеров, образование которых обусловлено неоднородностью активной среды, т. е. присутствием внутри нее группы автоколебательных элементов, спиральные волны являются локальными источниками автоволн, не привязанными к какой-либо неоднородности. Положение центра спиральной волны определяется лишь начальными условиями, сложившимися при ее зарождении. Все спиральные волны в данной среде имеют одинаковую частоту циркуляции. Поэтому две спиральные волны, будучи автоволновыми источниками, не подавляют с течением времени друг друга (рис. 1.29).

Своеобразные эффекты наблюдаются при взаимодействии спиральной волны с пейсмекером. Если частота генерации волн пейсмекером ниже частоты циркуляции спиральной волны, последняя подавляет пейсмекер (рис. 1.30). При обратном соотношении частот спиральная волна вырождается в дополнительную полуволну (своеобразную «дислокацию»), которая, испытывая последовательные перезамыкания, постепенно вытесняется все дальше на периферию пейсмекера (рис. 1.31).

6. Спиральные волны в распределенных возбудимых средах

Почти все эффекты, отмечавшиеся в разделе 1.5 для возбудимых сред, представляющих собой сети из дискретных автоматов, сохраняются для распределенных возбудимых сред. Об этом свидетельствуют как имеющиеся результаты численного моделирования автоволновых процессов в таких средах, так и данные экспериментальных исследований. Одним из наиболее удобных объектов для таких исследований служит химическая возбудимая среда, представляющая собой тонкий неперемешиваемый слой раствора, в котором протекает реакция Белоусова – Жаботинского⁵. Преимущества этой системы состоят в том, что автоволновые структуры в ней имеют макроскопические размеры (порядка нескольких миллиметров), а процессы эволюции структур являются достаточно медленными (с характерными временами порядка минуты). Это существенно упрощает регистрацию автоволновых эффектов. Недавно была предложена [188] модификация этой реакции с заменой ферроина в качестве катализатора на соединение рутения $\text{Ru}(\text{dipy})_3$. Модифицированная реакция оказалась светочувствительной: меняя освещенность, можно управлять возбудимостью среды.

⁵Подробное описание этой реакции, а также обзор работ по анализу ее химического механизма приведены в [119].

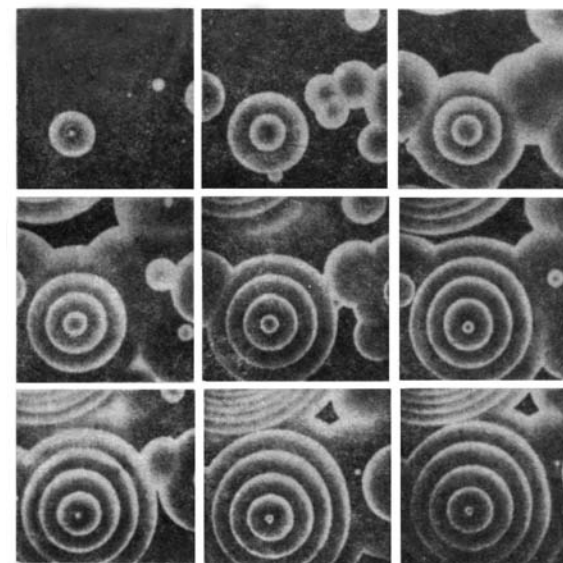


Рис. 1.32. Ведущие центры в среде с реакцией Белоусова – Жаботинского. Интервалы между кадрами 30 с

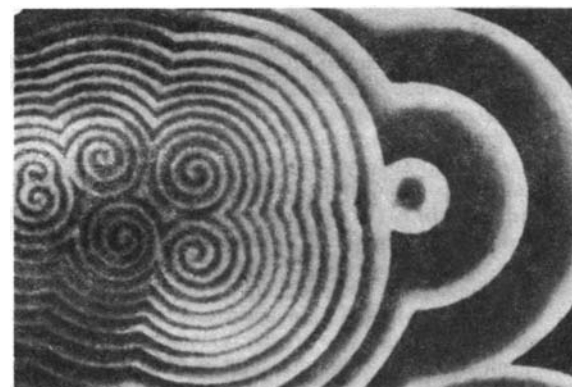


Рис. 1.33. Спиральные волны и ведущий центр в среде с реакцией Белоусова – Жаботинского

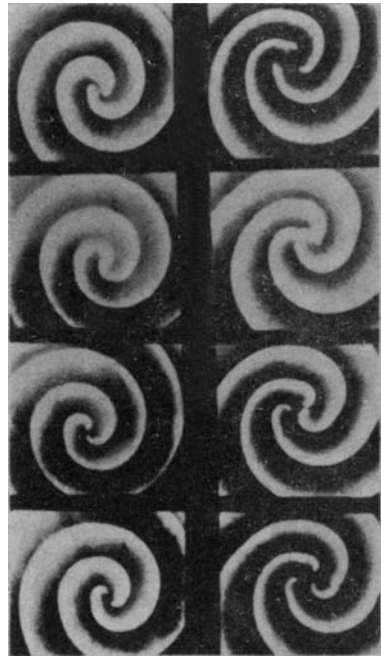


Рис. 1.34. Многорукавные спиральные волны в среде с реакцией Белоусова–Жаботинского. Интервалы между кадрами 15 с

Если приготовить такой раствор и предоставить его самому себе, в нем спонтанно возникают источники концентрически расходящихся волн — ведущие центры (рис. 1.32). Длительное время оставалось неясным, являются ли ведущие центры автономными источниками или же их наличие обусловлено присутствием в растворе посторонних включений (примесей, пузырьков газа), в окрестности которых элементы среды переходят в автоколебательный режим. Сегодня, по-видимому, можно уже считать твердо установленным, что все ведущие центры в реакции Белоусова–Жаботинского имеют внешнее происхождение. В эксперименте [134] на ведущий центр посылали периодическую последовательность волн с частотой выше его частоты генерации. В результате действие ведущего центра оказывалось подавленным. Однако стоило лишь прекратить подачу этой последовательности волн, как ведущий центр восстанавливался на прежнем месте.

В отличие от ведущих центров, спиральные волны в среде с реакцией Белоусова–Жаботинского не связаны с присутствием каких-либо посторонних включений, их обычно получают путем разрыва сложного фронта волны возбуждения. Все спиральные волны имеют одинаковую частоту. Поэтому, как видно из рис. 1.33, спиральные волны сосуществуют между собой, но гасят ведущий центр, являющийся более медленным автоволновым источником. В среде с реакцией Белоусова–Жаботинского [86] также наблюдались многорукавные спиральные волны (рис. 1.34; обратите внимание на нестационарность режима и «перезамыкание» волн в центральной области).

Приведенные примеры свидетельствуют, что спиральные волны представляют собой главный тип элементарных самоподдерживающихся структур в однородных возбудимых средах. Подобно вихрям в сверхпроводнике или в сверхтекучем гелии, они чрезвычайно устойчивы.

Расчет частоты вращения и формы спиральных волн — важная задача теории. Рассмотрим вначале распространение бегущего импульса по тонкому кольцу радиуса R . Если толщина кольца достаточно мала, задачу можно считать одномерной. Как отмечалось в разделе 1.4, ей отвечает то же решение, что и для периодической последовательности импульсов на бесконечной прямой с пространственным периодом $L = 2\pi R$. Скорость импульсов однозначно определялась периодом L и возрастала с увеличением L , стремясь к конечному пределу V_0 при $L \rightarrow \infty$.

Начнем увеличивать внешний радиус кольца и устремим его к бесконечности. В результате мы приходим к задаче о волне возбуждения, вращающейся вокруг отверстия радиуса R . Заметим прежде всего, что в установившемся режиме фронт волны не может представлять тогда прямую линию, вращающуюся с угловой скоростью ω . Действительно, это означало бы, что на расстоянии r от центра отверстия скорость фронта равняется $V = \omega r$ и неограниченно возрастает с увеличением r . Но скорость фронта не может превысить значение V_0 . Следовательно, удаленные участки фронта отстают, а сам фронт скручен в спираль.

Пусть в установившемся режиме спираль имеет форму $\varphi = \omega t - \chi(r)$ (где $\chi(r)$ — некоторая функция r) и вращается с угловой скоростью ω . Рассмотрим точку A , удаленную на расстояние r от центра спирали O (рис. 1.35а). Если провести через точку A и O окружность, видно, что точка пересечения линии фронта с такой окружностью бежит по ней со скоростью $V_r = \omega r$. Смещение этой точки связано с прохождением через точку A (рис. 1.35б) участка фронта спиральной волны, имеющего скорость нормального смещения V , и поэтому скорость такого смещения равна

$$V_r = \frac{V}{\cos \alpha}, \quad (1.77)$$

где α — угол между нормалью к фронту и касательной к окружности в точке A . Несложный расчет показывает, что если линия фронта задается как $\varphi = \omega t - \chi(r)$ и в данный момент фронт проходит через точку A , то

$$\cos \alpha = \frac{1}{\sqrt{1 + r^2 \chi_r^2}}, \quad (1.78)$$

где $\chi_r = d\chi/dr$. С учетом (1.77), (1.78) и равенства $V_r = \omega r$ находим

$$V = \frac{\omega}{\sqrt{\chi_r^2 + 1/r^2}}. \quad (1.79)$$

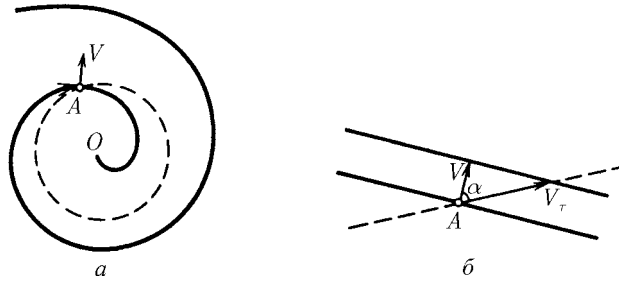


Рис. 1.35. К расчету формы спиральной волны

Предположим для простоты, что скорость нормального смещения всех участков фронта одинакова и равна скорости движения одиночной плоской волны возбуждения (одиночного импульса): $V = V_0$. Тогда (1.79) будет представлять собой уравнение, определяющее производную χ_r на различных удалениях r от центра волны при заданном значении ω .

Частоту вращения волны вокруг отверстия с радиусом R можно рассчитать исходя из следующих соображений. Поскольку отверстие непроницаемо, фронт волны подходит к нему под прямым углом, поэтому $\chi_r(R) = 0$. Полагая в (1.79) $r = R$, мы получаем с учетом этого граничного условия $V_0 = \omega R$. Следовательно, период $T = 2\pi/\omega$ спиральной волны равен времени одного оборота импульса, циркулирующего по окружности радиуса R : $T = 2\pi R/V_0$.

Учитывая этот результат, уравнение (1.79) можно переписать как

$$\chi_r^2 + \frac{1}{r^2} = \frac{1}{R^2}. \quad (1.80)$$

Функция $\chi(r)$, определяющая формулу спирали, вычисляется простым интегрированием:

$$\chi(r) = \int_R^r \sqrt{\frac{1}{R^2} - \frac{1}{r^2}} dr. \quad (1.81)$$

Таким образом, функция $\chi(r)$ монотонно возрастает с увеличением r , а следовательно, задаваемая ею в полярных координатах кривая $\varphi = \chi(r)$ является спиралью. Шаг этой спиральной волны $h_{св}$ можно найти из условия

$$\chi(r + h_{св}) - \chi(r) = 2\pi. \quad (1.82)$$

Вообще говоря, он зависит от расстояния r до ее центра. При больших удалениях

$$h_{св} \approx \frac{2\pi}{\chi_r(r)}. \quad (1.83)$$

Поскольку в этой области $\chi_r \approx 1/R$, мы видим, что вдали от центра шаг спирали постоянен и равен периметру отверстия: $h = 2\pi R$. Спираль с постоянным шагом называют *архимедовой*. Отметим также, что кривая, задаваемая в полярных координатах уравнением $\varphi = \chi(r)$ с функцией $\chi(r)$ вида (1.81), представляет собой *эвольвенту* отверстия радиуса R [38].

Итак, мы убедились, что вокруг отверстия в возбудимой среде может устойчиво циркулировать спиральная волна, и нашли в простейшем приближении ее форму и частоту циркуляции.

Проведенный расчет был неточным в двух отношениях. Во-первых, мы пренебрегли зависимостью скорости следования волн от интервала между ними. Этот недостаток нетрудно исправить. Поскольку угловая скорость вращения равна ω , в установившемся режиме вся картина полностью повторяется через промежутки времени $T = 2\pi/\omega$, а следовательно, интервал между прохождением двух последовательных волн через любую точку среды также равен T . Поэтому вместо скорости V_0 , отвечающей движению одиночного импульса, следует использовать скорость $V(T)$, которая характеризует движение последовательности импульсов с периодом T . Период T можно тогда найти, решив уравнение

$$V(T)T = 2\pi R. \quad (1.84)$$

Второе обстоятельство является более тонким, а его учет — не столь простым. Проведенное выше рассмотрение не принимало во внимание зависимость скорости движения фронта волны от его кривизны. Между тем, очевидно, что участки фронта, находящиеся на разном удалении от центра, имеют различную кривизну, а следовательно, различаются между собой и скорости их нормального смещения.

Кривизна линии, заданной в полярных координатах уравнением $\varphi = \chi(r)$, равна

$$K = \frac{r\chi_{rr} + 2\chi_r + r^2\chi_r^3}{\sqrt{(1 + r^2\chi_r^2)^3}}. \quad (1.85)$$

Из (1.80) следует, что для $\chi_{rr} = d\chi_r/dr$ справедливо выражение

$$\chi_{rr} = \frac{1}{r^2\chi_r}. \quad (1.86)$$

Поскольку волна подходит к отверстию под прямым углом, $\chi_r \rightarrow 0$ при $r \rightarrow R$ (мы уже использовали выше это условие). Но тогда, согласно (1.86), $\chi_{rr} \rightarrow \infty$ при $r \rightarrow R$, а поэтому кривизна найденной нами линии, описывающей форму фронта спиральной волны, обращается в бесконечность (ср. (1.85)) при подходе к границе отверстия. Ясно, однако, что пренебрегать зависимостью скорости движения фронта волны от его кривизны, как мы поступали до настоящего времени, можно лишь, если эта кривизна K достаточно мала. Мы видим, что по крайней мере вблизи от границы отверстия это приближение заведомо не справедливо.

Забега вперёд, заметим, что при учете такой зависимости найденное нами решение подправляется лишь в пределах узкого слоя шириной порядка $(D^2 R/V_0^2)$, где D — коэффициент, характеризующий зависимость скорости фронта от его кривизны; вне этого узкого пограничного слоя спираль имеет форму эвольвенты отверстия. Для исправленного решения при приближении к границе отверстия кривизна фронта не возрастает до бесконечности, а стремится к конечному значению порядка $\sqrt[3]{V_0/DR^2}$.

Циркуляция волны вокруг отверстия в предположении о постоянстве скорости движения фронта рассмотрена в классической работе Н. Винера и А. Розенблюта [38]. Именно они обратили внимание на то, что фронт спирали представляет собой эвольвенту отверстия.

Как свидетельствуют приведенные выше примеры, волны наблюдаются и в отсутствие отверстия.

В однородной возбудимой среде при установившемся вращении спиральной волны концевая точка ее фронта совершает вращательное движение, перемещаясь вдоль окружности — границы ядра. Внутри ядра элементы среды сохраняют состояние покоя, хотя эта область по своим свойствам ничем не отличается от других участков возбудимой среды. Все это выглядит так, как будто ядро представляет собой некоторое *эффективное* отверстие. Главный вопрос, который в связи с этим возникает, — почему возбуждение не проникает внутрь ядра?

Ранее мы отмечали, что для каждой среды существует свой минимальный возможный период устойчивого следования импульсов L_{\min} . Это означает, что на окружностях с радиусом меньше $R_{\min} = L_{\min}/2\pi$ устойчивая циркуляция импульсов невозможна.

Возьмем отверстие, вокруг которого циркулирует волна, и будем уменьшать его радиус R . Что будет происходить, если сделать радиус отверстия меньше R_{\min} ? Ясно, что волна не может циркулировать вдоль такого малого отверстия, будучи «приклеенной» к нему. В действительности волна возбуждения оторвется от отверстия и у нее образуется свободный оборванный

конец. Минимальный возможный радиус окружности (т. е. границы ядра), вдоль которой способен устойчиво циркулировать конец волны, — это радиус R_{\min} . Такая спиральная волна будет намотана максимально плотно — шаг ее спирали вдаль от центра равен $h_{\text{св}} = L_{\min}$, т. е. она испускает самую частую из возможных последовательностей волн и имеет частоту циркуляции $\omega = \omega_{\max}$.

Полученная простая оценка частоты спиральной волны и радиуса ее ядра оказывается довольно близкой к действительности в случае хорошо возбудимых сред (ср. раздел 1.5); она дает для них период спиральной волны, заниженный в среднем лишь примерно в два раза по отношению к истинному. Причина такого расхождения понятна. При выводе этой оценки мы пренебрегли кривизной фронта спиральной волны — точнее, ее влиянием на скорость движения фронта волны. Между тем, участки спиральной волны, лежащие ближе всего к ядру, имеют большую кривизну и, следовательно, должны двигаться несколько медленнее одиночного импульса. В результате частота циркуляции спиральной волны уменьшается по сравнению со значением ω_{\max} .

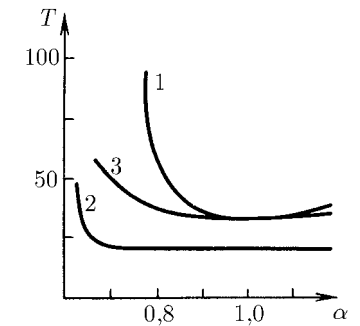


Рис. 1.36. Период одорукавной спиральной волны (кривые 1, 3) и минимальный возможный период следования импульсов (кривая 2) как функции параметра α , характеризующего возбудимость среды

Последовательный учет этого эффекта предпринят в [88, 207]; он позволяет получить достаточно хорошее согласие аналитических результатов с данными численного расчета. На рис. 1.36, взятом из [88], проведено сравнение величины $T_{\min} = 2\pi/\omega_{\max}$ (кривая 2) со значением периода циркуляции, полученным путем численного моделирования вращающейся спиральной волны в рамках полных, зависящих от времени уравнений возбудимой среды (кривая 1), а также с результатами приближенного аналитического расчета, учитывающего кривизну фронта распространяющихся волн (кривая 3). Мы видим, что согласие остается очень хорошим, пока период цир-

куляции спиральной волны не начинает резко возрастать по сравнению с T_{\min} .

Выше описан режим циркуляции спиральной волны, для которого характерны малый размер ядра (его радиус порядка R_{\min}) и малый период циркуляции порядка T_{\min} ; этот режим осуществляется в хорошо возбудимых средах. Если уменьшать возбудимость среды (что соответствует уменьшению параметра α на рис. 1.36), на смену ему постепенно приходит режим циркуляции с большим периодом ($T \gg T_{\min}$) и большим радиусом ядра ($R \gg R_{\min}$) спиральной волны. Описание таких спиральных волн и расчет их основных параметров требуют специального анализа, который проведен в разделе 1.7.

7. Кинематика автоволновых фронтов в возбудимых средах

Возбудимая среда нелинейна, и автоволны взаимодействуют друг с другом. При столкновении двух волн они аннигилируют; если же одна волна движется вслед другой, то она не ощущает идущую впереди волну, пока расстояние между ними значительно больше, чем ширина волны, т. е. величина порядка L_{\min} . Здесь мы ограничимся рассмотрением таких режимов, при которых каждая волна отделена от последующей расстоянием, гораздо большим L_{\min} . Тогда можно пренебречь шириной отдельной уединенной волны и считать, что она полностью задается ориентированной кривой своего фронта.

Поскольку состояния среды до и после прохождения импульса возбуждения совпадают, фронт автоволны может обрываться внутри среды, так что в ней волна имеет свободный конец (ср. раздел 1.5). Помимо смещения в нормальном направлении, на свободном конце волна может «прорасти» или сокращаться с течением времени. Скорость этого тангенциального смещения G зависит от кривизны K_0 фронта при подходе к точке обрыва. Пусть G_0 — скорость «прорастания» свободного конца плоской полуволны. Тогда при достаточно малых кривизнах K_0 скорость «прорастания» искривленной волны G можно рассчитать с помощью теории возмущений (см. приложение к [29]) и получить $G = G_0 - \gamma K_0$. Для интересующих нас возбудимых сред скорость «прорастания» уменьшается с ростом кривизны на свободном конце (т. е. $\gamma > 0$).

Как отмечалось в разделе 1.5, меняя возбудимость среды, можно управлять величиной G_0 и даже менять ее знак. При отрицательном значении G_0 свободный конец фронта сокращается и среда не допускает режимов в ви-

де спиральных волн. Поэтому мы предположим, что величина G_0 положительна, достаточно мала и линейная зависимость G от K сохраняется вплоть до значения кривизны на свободном конце $K_{\text{кр}} = G_0/\gamma$, при котором скорость «прорастания» обращается в нуль. Мы также считаем, что $DK_{\text{кр}} \ll V_0$. Эти два последних предположения существенно упрощают анализ благодаря появлению дополнительного малого параметра

$$\beta = \sqrt{\frac{DK_{\text{кр}}}{V_0}}. \quad (1.87)$$

Они выполняются, если возбудимость среды достаточно для существования спиральных волн, но все же не слишком высока.

Итак, в рамках кинематики математическое описание автоволн в двумерной возбудимой среде строится следующим образом. Будем характеризовать волну указанием только кривой ее фронта, для чего воспользуемся *натуральным уравнением* этой кривой $K = K(l)$. Оно устанавливает связь между длиной дуги l кривой (которую удобно отсчитывать от свободного конца фронта) и кривизной фронта K в соответствующей точке; как известно, натуральное уравнение задает кривую с точностью до ее расположения на плоскости. Если форма кривой меняется со временем, то $K = K(l, t)$.

Эволюция волны с течением времени определяется двумя условиями:

1. Каждый участок фронта смещается со временем по нормали к себе со скоростью

$$V = V_0 - DK. \quad (1.88)$$

2. Свободный конец волны сокращается ($G < 0$) или гладко «прорастает» в касательном направлении со скоростью

$$G = \gamma(K_{\text{кр}} - K_0), \quad (1.89)$$

где $K_0 = \lim_{l \rightarrow 0} K(l)$.

Сформулированных условий достаточно, чтобы получить уравнение, которому подчиняется функция $K(l, t)$. Его можно вывести, основываясь на методах дифференциальной геометрии [32]. Ниже мы изберем более простой способ, предложенный в [46].

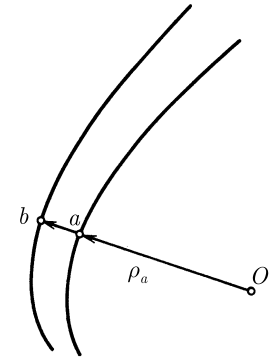


Рис. 1.37. К выводу уравнения (1.95)

Рассмотрим положение фронта волны в некоторый момент t . Пусть его кривизна в точке a была в этот момент K_a (рис. 1.37). Через промежуток времени dt малая окрестность точки a на фронте волны переместится в окрестность точки b , а кривизна в этой точке станет кривой K_b . Найдем связь между K_a и K_b .

Для этого удобно перейти в полярную систему координат с полюсом в центре кривизны малого участка фронта, лежащего в окрестности точки a . В этом случае справедливо соотношение

$$\rho_b = \rho_a + V dt, \quad (1.90)$$

где $\rho_a = 1/K_a$ — радиус кривизны фронта в точке a , а ρ_b — полярный радиус-вектор точки b . Кривизна кривой, описываемой в полярных координатах уравнением $\rho = \rho(\varphi)$, вычисляется по формуле

$$K = (\rho^2 - 2\rho' - \rho\rho'')(\rho^2 + \rho'^2)^{-3/2}. \quad (1.91)$$

Подставив (1.90) в выражение (1.91), получаем с точностью до членов первого порядка по dt :

$$K_b = K_a - V K_a^2 dt - K_a^2 \frac{d^2 V}{d\varphi^2} dt \quad (1.92)$$

(мы учли здесь, что $\partial\rho_a/\partial\varphi = 0$). Перейдя в (1.92) к дифференцированию по длине дуги l ($dl = \rho_a d\varphi$), находим

$$K_b - K_a = dK = - \left(K_a^2 V + \frac{d_2 V}{dl^2} \right) dt. \quad (1.93)$$

С другой стороны, $dK = (\partial K/\partial l) dl + (\partial K/\partial t) dt$. Для приращения длины дуги за время dt нетрудно получить выражение

$$dl = \left(\int_0^l KV d\xi \right) dt + G dt. \quad (1.94)$$

Первое слагаемое в (1.94) описывает увеличение длины дуги за счет изменения радиусов кривизны в каждой точке кривой; второму слагаемому отвечает изменение длины дуги за счет «прорастания» (или сокращения, если $G < 0$) фронта на своем свободном конце, который выбран за начало

отсчета l . Комбинируя два выражения для dK , приходим к уравнению

$$\frac{\partial K}{\partial t} + \left(\int_0^l KV d\xi + G \right) \frac{\partial K}{\partial l} = -K^2 V - \frac{\partial^2 V}{\partial l^2}, \quad (1.95)$$

в котором скорость «прорастания» G зависит от кривизны $K(0, t) = \lim_{l \rightarrow 0} K(l, t)$ на свободном конце, а скорость нормального смещения фронта V однозначно определяется локальной кривизной этого фронта в соответствующей точке, т. е. $V = V(K(l, t))$. Таким образом, уравнение (1.95) однозначно определяет эволюцию фронта со свободным концом (обрывом волнового фронта).

В частности, если G, V даются выражением (1.88) и (1.89), уравнение (1.95) приобретает вид

$$\frac{\partial K}{\partial t} + \left[\int_0^b KV_0 d\xi + \gamma(K_{\text{кр}} - K(0, t)) \right] \frac{\partial K}{\partial l} = -K^2 V_0 + D \frac{\partial^2 K}{\partial l^2} \quad (1.96)$$

— основного уравнения кинематики волновых фронтов в возбудимых средах. Мы учли здесь, что поправка к скорости нормального смещения $V = V_0 - DK$ при значениях кривизны $K \lesssim K_{\text{кр}}$ мала согласно сделанному нами выше предположению, а поэтому заменили V на V_0 всюду в (1.96), кроме последнего слагаемого, где это привело бы к отбрасыванию члена со старшей производной — вскоре мы увидим, что этот член играет весьма важную роль.

Уравнение (1.96) имеет тривиальное стационарное решение $K(l, t) \equiv 0$ при $l \geq 0$, которое отвечает плоскому фронту с обрывом, смещающемуся вперед со скоростью V_0 . Однако это решение абсолютно неустойчиво по отношению к малым возмущениям. Численные расчеты для уравнения (1.96) показывают (рис. 1.38), что независимо от характера начального возмущения из плоской полуволны со свободным концом всегда рождается вращающаяся вокруг некоторого центра с постоянной угловой скоростью спиральная волна. Основные характеристики установившейся стационарной циркуляции спиральных волн можно рассчитать аналитически, непосредственно исходя из (1.96).

При стационарной циркуляции форма волнового фронта постоянна, меняется со временем лишь положение кривой на плоскости. Поэтому зависимость K от t отсутствует. Кроме того, стационарная циркуляция возможна, лишь если свободный конец волны, совершающий равномерное круговое

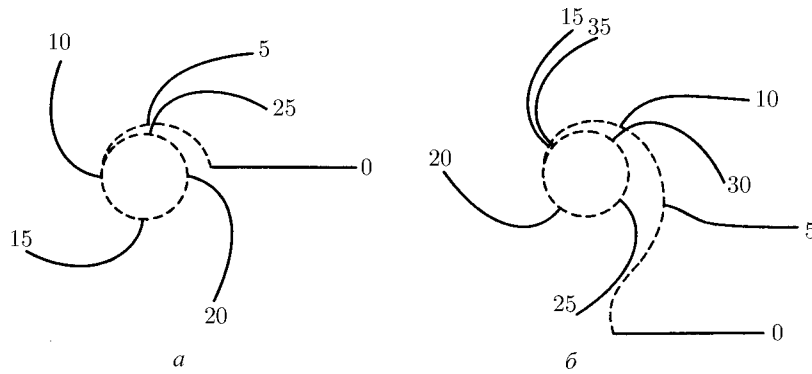


Рис. 1.38. Формирование спиральной волны (данные численного расчета) из плоской полуволны с двумя различными малыми начальными возмущениями (*a* и *b*)

движение вокруг центра, не «прорастает» и не сокращается со временем. Это означает, что кривизна его фронта должна быть равной критической, т. е. $K(0) = K_{кр}$. Круг, по границе которого движется свободный конец волны, представляет собой ее ядро; фронт спиральной волны подходит к ядру по нормали.

Заметим, что с помощью уравнения (1.96) можно описать также стационарную циркуляцию спиральной волны вокруг круглого отверстия в возбудимой среде. Если граница отверстия непроницаема и диффузионный поток через нее равен нулю, фронт ортогонален к границе. Форма фронта тогда определяется стационарным пределом уравнения (1.96) в отсутствие слагаемого, которое описывало бы «прорастание» свободного конца (поскольку такого свободного конца уже нет).

При стационарной циркуляции, когда $\partial K/\partial t = 0$, уравнение (1.96) можно один раз проинтегрировать, получив

$$K \int_0^l K V_0 d\xi - D \frac{dK}{dl} = \omega. \quad (1.97)$$

Постоянная интегрирования ω равна частоте вращения спиральной волны. Это легко понять, учтя, что фронт подходит к ядру (или отверстию) по нормали, и поэтому $dV/dl = -DdK/dl$ при $l = 0$ есть угловая скорость.

Рассмотрим, какой результат был бы получен в модели, где скорость V вообще не зависит от кривизны, т. е. $D = 0$, и второе слагаемое слева

в уравнении (1.97) отсутствует. Нетрудно заметить, что в этом случае (1.97) имеет решение

$$K = \sqrt{\frac{\omega}{2lV_0}}. \quad (1.98)$$

Пусть в среде вырезано отверстие с радиусом R . Тогда частота циркуляции спиральной волны вокруг отверстия равна $\omega = V_0/R$ и из решения (1.98) следует

$$K = \sqrt{\frac{1}{2Rl}}. \quad (1.99)$$

Можно проверить, что кривая, определяемая данным натуральным уравнением, есть эвольвента круга с радиусом R . При $l \gg R$ она представляет собой спираль с постоянным шагом $h = 2\pi R$, т. е. архимедову спираль. Таким образом, для модели, где $V \equiv V_0$, мы воспроизвели основные результаты, полученные в разделе 1.6.

Заметим, что при $l \rightarrow 0$ найденное выше решение даст $K \rightarrow \infty$. Это означает, что если коэффициент D в зависимости $V = V_0 - DK$ мал, но все же отличен от нуля, зависимостью скорости от кривизны нельзя пренебрегать при достаточно малых l , т. е. вблизи границы ядра (или отверстия).

Рассмотрим теперь задачу о расчете параметров спиральной волны в отсутствие отверстия с учетом зависимости скорости фронта от его кривизны. Частота вращения спиральной волны $\omega = \omega_0$, входящая в качестве неизвестного параметра в (1.97), определяется в этом случае из следующих соображений. Вдали от центра спирали кривизна ее фронта должна стремиться к нулю, т. е. $K \rightarrow 0$ при $l \rightarrow \infty$. Можно показать, что это условие уже однозначно определяет фазовую траекторию для дифференциального уравнения (1.97) при заданном значении ω . В том числе оказывается однозначно определенным и значение кривизны $K(0)$ при $l = 0$, т. е. на свободном конце волны, подходящем к ядру спирали. Потребовав теперь, чтобы $K(0) = K_{кр}$, получим уравнение для нахождения частоты ω_0 .

При достаточно больших значениях l можно пренебречь слагаемым с производной в левой части уравнения (1.97). В этой области решение дается формулой (1.98). Из уравнения (1.97) вытекает также, что при $l = \infty$ производная равна $dK/dl = -\omega/D$. Поэтому при малых l натуральное уравнение интересующей нас кривой есть

$$K = K_{кр} - \frac{\omega}{D}l. \quad (1.100)$$

Выражения (1.98) и (1.100) представляют собой внешнюю и внутреннюю аппроксимации для единой фазовой траектории, а поэтому они долж-

ны «сшиваться» между собой. Потребуем, чтобы в некоторой точке $l = l_0$ совпадали как значения двух функций, так и их первые производные. Эти условия дают два уравнения:

$$K_{\text{кр}} - \frac{\omega}{D} l_0 = \sqrt{\frac{\omega}{2V_0 l_0}}, \quad (1.101)$$

$$\frac{\omega}{D} = \frac{1}{2} \sqrt{\frac{\omega}{2V_0}} \sqrt{\frac{1}{l_0^3}}. \quad (1.102)$$

Их решения имеют вид

$$\omega_0 = \sqrt{\left(\frac{2}{3}\right)^3} \sqrt{DV_0} \sqrt{K_{\text{кр}}^3}, \quad (1.103)$$

$$l_0 = \frac{1}{2} \sqrt{\frac{3}{2}} \sqrt{\frac{D}{K_{\text{кр}} V_0}}. \quad (1.104)$$

Приведенный выше приближенный расчет, основанный на сшивке внутреннего и внешнего решений, является довольно очевидным подходом при решении уравнения (1.97) с малым параметром при старшей производной. Оказывается, однако, что для данного конкретного уравнения можно найти также и точное решение, а именно: переходя к новым переменным, уравнение (1.97) удается полностью обезразмерить, а затем проинтегрировать на ЭВМ. В результате мы получаем

$$\omega_0 = \zeta \sqrt{DV_0} \sqrt{K_{\text{кр}}^3}, \quad \zeta \approx 0,69 \dots, \quad (1.105)$$

что отличается от (1.103) лишь более точным значением числового коэффициента.

С точностью до членов порядка $DK_{\text{кр}}/V_0$ радиус ядра спиральной волны вычисляется из условия $\omega_0 R_0 = V_0$:

$$R_0 = \frac{1}{\zeta} \sqrt{\frac{V_0}{D}} \sqrt{\frac{1}{K_{\text{кр}}^3}}. \quad (1.106)$$

Обсудим полученные результаты. Как видно, при $K_{\text{кр}} \rightarrow 0$ частота вращения спиральной волны ω_0 убывает и стремится к нулю; при этом радиус R_0 ядра спиральной волны неограниченно возрастает. Всюду, кроме узкого пограничного слоя ширины порядка l_0 вблизи границы ядра, фронт

спирали имеет вид эвольвенты окружности радиуса R_0 и дается выражением (1.99) при $R = R_0$. Ширина погранслоя l_0 мала по сравнению с радиусом ядра: $l_0/R_0 \sim (DK_{\text{кр}}/V_0) \ll 1$.

Совершенно аналогично можно найти решение для спиральной волны, вращающейся вокруг отверстия. К отверстию с радиусом $R \gg D/V_0$ фронт волны подходит с кривизной

$$K(0) = \sqrt[3]{\frac{1}{\zeta^2} \sqrt[3]{\frac{V_0}{DR^2}}}. \quad (1.107)$$

Вне узкого погранслоя шириной порядка $D^2 R/V_0^2$ спираль имеет форму эвольвенты отверстия. С точностью до членов порядка $\sqrt[3]{(D/V_0 R)^2}$ частота вращения ω определяется из условия $\omega R = V_0$.

Заметим, что конец спиральной волны, движущейся вдоль непроницаемой для диффузии границы отверстия, нельзя считать свободным. Его кривизна может превышать значение критической кривизны $K_{\text{кр}}$ для свободного конца волны, но должна оставаться меньше значения критической кривизны K^* для распространяющегося сплошного фронта (см. выражение (1.71) и рис. 1.19). Это обуславливает эффект гистерезиса, наблюдавшийся в численном эксперименте [227].

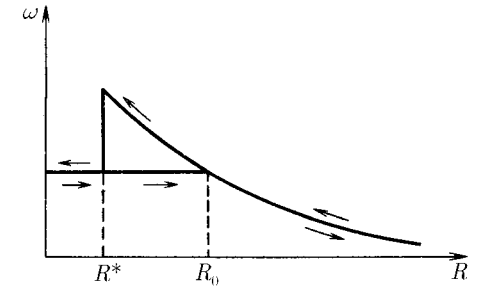


Рис. 1.39. Зависимость частоты вращения спиральной волны от радиуса отверстия

В некотором интервале значений радиуса отверстия ($R^* < R < R_0$) могут осуществляться два режима циркуляции спиральной волны, отличающихся частотой вращения (рис. 1.39). При первом режиме конец волны движется вдоль границы отверстия (волна как бы «приклеена» к отверстию), а при втором конце волны оказывается свободным и движется вокруг ядра с радиусом R_0 , так что спиральная волна не ощущает отверстия, расположенного в ее центре. Если вначале создать в среде достаточно большое

отверстие с радиусом $R > R_0$ и затем плавно уменьшать его, то спиральная волна будет оставаться «приклеенной» к отверстию вплоть до $R = R^*$, после чего конец волны отходит от отверстия, частота вращения скачком понижается и остается при $R \leq R^*$ равной частоте вращения спиральной волны в однородной среде (1.105). При движении в обратном направлении, т. е. при плавном увеличении R от нуля вплоть до $R = R_0$, сохраняется режим с «отклеившимся» концом волны, и лишь при $R > R_0$ конец спиральной волны подходит к границе отверстия.

В рамках использованной нами кинематической модели предполагалось, что скорость нормального смещения V зависит только от локальной кривизны данного участка волнового фронта, т. е. $V = V(K)$. Вообще говоря [32], для участков, близких к свободному концу, скорость V может также явно зависеть от расстояния до точки обрыва (т. е. от длины дуги l): $V = V(K, l)$. При такой дополнительной зависимости основные характеристики спиральных волн рассчитаны в [32]. Этот эффект, однако, не дает существенных различий, и мы его подробно не обсуждаем.

Подчеркнем, что все найденные здесь для спиральных волн результаты применимы лишь в том случае, если полученное значение радиуса R_0 ядра спиральной волны велико по сравнению с $R_{\min} = L_{\min}/2\pi$ (см. раздел 1.4), т. е. когда периметр ядра велик по сравнению с шириной отдельного импульса и фронт волны при своем движении не наталкивается на свой же «хвост».

8. Резонанс и дрейф спиральных волн

Основное уравнение кинематики волновых фронтов (1.96) можно использовать также для описания различных нестационарных режимов в определенных возбудимых средах.

Рассмотрим вначале процесс установления стационарной циркуляции спиральной волны в однородной среде. Допустим, что мы внесли малое локальное возмущение в форму фронта спиральной волны, локализованное на удалении l от свободного конца. Как следует из (1.96), оно будет с течением времени удаляться от центра волны на периферию, одновременно расплываясь и затухая по диффузионному закону с коэффициентом «диффузии» D . Согласно (1.96) скорость этого сноса вблизи от свободного конца, где $K \approx K_{\text{кр}}$, можно приближенно оценить как $K_{\text{кр}}V_0l$.

Итак, локализованное на удалении l возмущение смещается со скоростью порядка $K_{\text{кр}}V_0l$, а его ширина растет как \sqrt{Dt} (рис. 1.40). Можно поставить вопрос: на каком удалении должно быть первоначально локализовано возмущение, чтобы, несмотря на сносное движение, оно могло достиг-

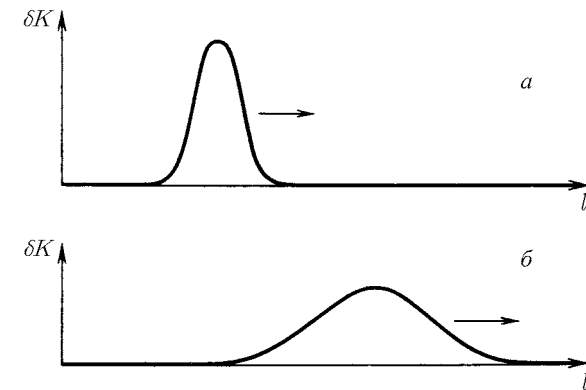


Рис. 1.40. Начальное возмущение формы спиральной волны (а) и его вид в последующий момент времени (б)

нуть свободного конца волны? Для этого необходимо, чтобы на протяжении некоторого интервала времени ширина расплывающегося возмущения (диффузионная длина) $L_{\text{диф}} \approx \sqrt{Dt}$ превышала удаление его центра от свободного конца волны $L_{\text{ц}} \approx l + K_{\text{кр}}V_0lt$. Сравнение двух этих выражений показывает, что $L_{\text{диф}}$ может превысить $L_{\text{ц}}$ на определенном интервале времени, лишь если начальное удаление l не превышает величины порядка $\sqrt{D/K_{\text{кр}}V_0}$. Но, согласно (1.104), именно такова ширина погранслоя l_0 вблизи границы ядра!

Следовательно, оказать влияние на движение свободного конца волны могут лишь те возмущения фронта, которые возникают в пределах узкого погранслоя, на удалении не более l_0 от свободного конца. Эти возмущения формы фронта затухают за время

$$\tau_D \sim \frac{l_0^2}{D} \sim \frac{1}{K_{\text{кр}}V_0}. \quad (1.108)$$

Заметим, что характерное время релаксации возмущения формы фронта τ_D спиральной волны всегда мало по сравнению с ее периодом циркуляции. Действительно, из (1.105) и (1.108) следует, что

$$\omega_0\tau_D \sim \sqrt{\frac{DK_{\text{кр}}}{V_0}} \ll 1. \quad (1.109)$$

Кроме τ_D , существует еще одно характерное время τ_G , связанное с эффектами «прорастания» свободного конца волны.

Если бы коэффициент γ в выражении (1.89) для скорости прорастания G равнялся нулю, величина $K_{кр}$ вообще не входила бы в основное уравнение кинематики (1.96). В этом случае имелось бы не одно, а много различных решений для спиральных волн, отличающихся значением кривизны фронта $K(0)$ при подходе к свободному концу. Каждое из таких решений описывало бы некоторую спиральную волну, стационарно циркулирующую с частотой ω_0 , которая определялась бы по-прежнему формулой (1.105) с заменой $K_{кр}$ на $K(0)$. Тогда внешнее воздействие, внесшее локальное возмущение в форму фронта в пределах погранслоя, могло бы вызвать переход к новому режиму циркуляции со значением кривизны фронта на свободном конце, отличающимся от прежнего на некоторую малую величину $\delta K(0)$.

Если коэффициент γ отличен от нуля и положителен, все отклонения кривизны $K(0)$ от значения $K_{кр}$ затухают со временем. Если γ достаточно мала, то соответствующее время релаксации велико ($\tau_G \gg \tau_D$) и осуществляется квазистационарный режим. В этом режиме форма фронта у ядра адиабатически подстраивается к мгновенному значению кривизны $K(0, t)$ на свободном конце волны, а эта кривизна, в свою очередь, медленно меняется из-за «прорастания» или сокращения согласно уравнению

$$\left(\frac{\partial K}{\partial t}\right)_{l=0} = -G\left(\frac{\partial K}{\partial l}\right)_{l=0}. \quad (1.110)$$

Производная $(\partial K/\partial l)_{l=0}$ определяется равенством

$$\left(\frac{\partial V}{\partial l}\right)_{l=0} = -D\left(\frac{\partial K}{\partial l}\right)_{l=0} = \omega, \quad (1.111)$$

в котором ω находится из (1.105) с заменой $K_{кр}$ на $K(0, t)$.

С учетом сделанных замечаний при квазистационарном режиме кривизна на свободном конце $K_0 \equiv K(0, t)$ меняется со временем, подчиняясь следующему обыкновенному дифференциальному уравнению:

$$\frac{dK_0}{dt} = -\zeta\gamma\sqrt{\frac{V_0}{D}}\sqrt{K_0^3(K_0 - K_{кр})}. \quad (1.112)$$

Для малых возмущений $\delta K_0 = K_0 - K_{кр}$ уравнение (1.112) линеаризуется и приобретает форму

$$\frac{d(\delta K_0)}{dt} = -\frac{\delta K_0}{\tau_G}, \quad (1.113)$$

где характерное время релаксации

$$\tau_G = \frac{D}{\gamma\omega_0}. \quad (1.114)$$

Заметим, что $\omega\tau_G = D/\gamma$, а поэтому, согласно (1.109), условие применимости квазистационарного приближения $\tau_D \ll \tau_G$ имеет вид

$$\frac{\gamma}{D} \ll \sqrt{\frac{V_0}{DK_{кр}}}. \quad (1.115)$$

По оценкам для известных возбудимых сред, коэффициент γ варьируется от нуля до значения порядка D . Поэтому, поскольку $\sqrt{DK_{кр}/V_0} \ll 1$, условие (1.115) почти всегда выполнено и квазистационарное приближение справедливо.

Если $\gamma \ll D$, то $\omega_0\tau_G \gg 1$ и подстройка кривизны фронта вблизи ядра к значению $K_{кр}$, а следовательно, окончательное установление стационарного режима циркуляции, занимают много оборотов волны. Можно сказать, что в таких возбудимых средах спиральная волна обладает *высокой* инерционностью. При выполнении обратного неравенства $\gamma \gg D$ установление стационарной циркуляции происходит за время, малое по сравнению с длительностью одного оборота, и спиральная волна имеет *низкую* инерционность.

При выводе основного уравнения кинематики (1.96) мы отмечали, что натуральное уравнение кривой $K = K(l, t)$ задает только ее форму, но не положение на плоскости. Чтобы полностью описать эволюцию волнового фронта, необходимо располагать дополнительными уравнениями, которые определяли бы изменение со временем ориентации этой кривой и движение ее начальной точки.

Введем на плоскости декартову систему координат (x, y) . Пусть $X_0(t)$ и $Y_0(t)$ — координаты свободного конца волны в момент времени t , а $\alpha_0(t)$ — угол между касательной к фронту в его концевой точке и осью x (рис. 1.41). Легко убедиться, что задание величин X_0 , Y_0 и α_0 полностью фиксирует положение фронта на плоскости.

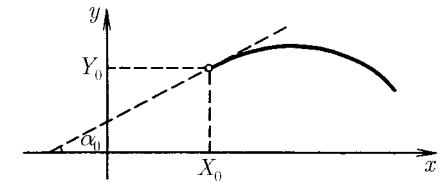


Рис. 1.41. К определению величин X_0 , Y_0 и α_0

Поскольку концевая точка смещается по нормали к фронту со скоростью $V(0) = V_0 - DK_0$ и смещается в тангенциальном направлении из-за «прорастания» или сокращения фронта со скоростью G , изменение ее декартовых координат со временем описывается уравнениями

$$\dot{X}_0 = -V(0) \sin \alpha_0 - G \cos \alpha_0, \quad \dot{Y}_0 = V(0) \cos \alpha_0 - G \sin \alpha_0. \quad (1.116)$$

При этом изменение со временем угла α_0 , характеризующего направление касательной к фронту в концевой точке, подчиняется уравнению

$$\dot{\alpha}_0 = \left(\frac{\partial V}{\partial l} \right)_{l=0} + GK_0, \quad (1.117)$$

или, с учетом $V = V_0 - DK$,

$$\dot{\alpha}_0 = -D \left(\frac{\partial K}{\partial l} \right)_{l=0} + GK_0. \quad (1.118)$$

В сочетании с (1.96) уравнения (1.116) и (1.118) полностью описывают движение по плоскости волнового фронта со свободным концом.

Особенно простой вид закон движения концевой точки приобретает в квазистационарном приближении, когда можно отделить эффекты «прорастания» от эффектов установления фронта спиральной волны вблизи ядра. В этом случае $(\partial K / \partial l)_{l=0} \approx -\omega / D$, где ω находится из (1.105) с заменой $K_{кр}$ на K_0 , и теперь вместо уравнения (1.118) имеем

$$\dot{\alpha} = \zeta \sqrt{V_0 D} \sqrt{K_0^3 + \gamma K_0 (K_{кр} - K_0)}. \quad (1.119)$$

Таким образом, при выполнении условия (1.115) для расчета движения концевой точки спиральной волны достаточно решить четыре обыкновенных дифференциальных уравнения (1.112), (1.116) и (1.119).

Варьируя параметры среды, можно менять ее возбудимость и, следовательно, величину $K_{кр}$. Пусть эта величина периодически меняется со временем как

$$K_{кр}(t) = K_{кр} + K_1 \cos(\omega_1 t + \varphi), \quad (1.120)$$

где K_1 — амплитуда модуляции, причем $K_1 \ll K_{кр}$, а частота модуляции ω_1 близка к собственной частоте ω_0 спиральной волны. Тогда, как было показано в [30], должен наблюдаться *резонансный эффект*, заключающийся в следующем. Оказывается, что при наличии модуляции $K_{кр}$ центр вращения спиральной волны уже не покоится, а движется с течением времени

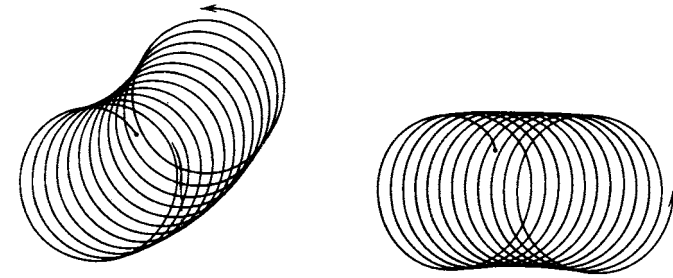


Рис. 1.42. Траектория движения кончика спиральной волны при двух различных отношениях между частотой модуляции и собственной частотой спиральной волны (численный расчет В. С. Зыкова)

по окружности (рис. 1.42), радиус $R_{рез}$ которой задается выражением

$$R_{рез} = \frac{3}{4} V_0 \frac{K_1}{K_{кр}} \frac{1}{|\omega - \omega_0|} \frac{1}{\sqrt{1 + (D/\gamma)^2}}; \quad (1.121)$$

он тем больше, чем ближе ω_1 к собственной частоте ω_0 . Скорость смещения центра

$$V_{рез} = \frac{3}{4} V_0 \frac{K_1}{K_{кр}} \frac{1}{\sqrt{1 + (D/\gamma)^2}} \quad (1.122)$$

пропорциональна амплитуде модуляции K_1 .

При точном совпадении частот, т. е. при *полном резонансе*, когда $\omega_1 = \omega_0$, центр фронта спиральной волны смещается с постоянной скоростью (1.122) вдоль прямой. Направление движения центра определяется начальной фазой модуляции φ и направлением вращения спиральной волны (по или против часовой стрелки).

Резонанс спиральных волн наблюдается в эксперименте [2] с использованием светочувствительной модификации реакции Белоусова – Жаботинского, катализатором в которой являлось соединение рутения. В среде первоначально создавали спиральную волну или пару таких волн. Затем, после измерения периода установившегося вращения спиральной волны, раствор начинали периодически однородно освещать с помощью интенсивного источника. Через равные интервалы времени производилось фотографирование волновой картины на фоне неподвижных реперных линий.

На рис. 1.43 показаны последовательные фотографии спиральной волны с интервалом времени 10 мин в условиях полного резонанса (частота

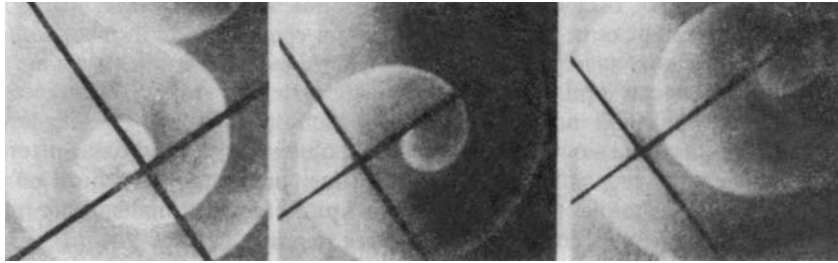


Рис. 1.43. Резонанс спиральной волны в химической возбудимой среде со светочувствительной реакцией Белоусова – Жаботинского

модуляции выбрана равной собственной частоте вращения волны). Видно, что центр вращения с постоянной скоростью смещается с течением времени вдоль одной из реперных линий; суммарное смещение составляет 2 см.

Если в среде имеются две спиральные волны с противоположными направлениями вращения, то, подбирая начальную фазу модуляции, можно добиться движения центров двух волн навстречу друг другу и их аннигиляции.

Перейдем теперь к рассмотрению явления *дрейфа* спиральной волны в неоднородной возбудимой среде [30]. Пусть критическая кривизна плавно меняется вдоль оси x , так что изменение $K_{кр}$ на расстоянии, равном размеру ядра R_0 , мало:

$$\left| \frac{\partial K_{кр}}{\partial x} \right| \ll \frac{K_{кр}}{R_0}. \quad (1.123)$$

При движении в неоднородной среде свободный конец спиральной волны последовательно проходит через области с различными значениями $K_{кр}$. Поэтому свободный конец движется так, как если бы критическая кривизна менялась со временем:

$$K_{кр}(t) = K_{кр}(X_0 + R_0 \cos \omega_0 t) \approx \bar{K}_{кр} + b_{кр} R_0 \cos \omega_0 t, \quad (1.124)$$

где X_0 — положение центра спиральной волны на оси,

$$\bar{K}_{кр} = K_{кр}(X_0), \quad b_{кр} = \left(\frac{\partial K_{кр}}{\partial x} \right)_{x=X_0}.$$

Следовательно, задача о дрейфе сводится к задаче о резонансе спиральной волны при точном совпадении частот. С течением времени центр

спиральной волны медленно смещается по прямой линии под некоторым направлением ϑ по отношению к оси x . Скорость дрейфа равна

$$V_d = \frac{(3/4)V_0}{\sqrt{1 + (D/\gamma)^2}} \frac{R_0}{K_{кр}} \left| \frac{\partial K_{кр}}{\partial x} \right|, \quad (1.125)$$

а тангенс угла дрейфа дается выражением

$$\operatorname{tg} \vartheta \approx -\frac{\gamma}{D}. \quad (1.126)$$

При высокой инерционности ($\gamma/D \ll 1$) спиральная волна смещается вдоль оси в направлении, обратном направлению увеличения $K_{кр}$. Если инерционность мала ($\gamma/D \gg 1$), дрейф осуществляется вдоль оси y , при этом его направление меняется на обратное при изменении направления вращения спиральной волны.

В заключение заметим, что дрейф спиральной волны не может продолжаться очень долго. В среде конечных размеров волна исчезает, когда ее ядро попадает на границу. Для неограниченной среды при движении волны в сторону меньших значений $K_{кр}$ радиус ее ядра растет: формально он равен бесконечности при $K_{кр} = 0$. При достижении этой области решение в виде спиральной волны теряется — свободный конец сокращается и уходит на бесконечность. Этот эффект, однако, уже не описывается простой линейной теорией, построенной выше, поскольку в такой области нарушается условие (1.123) и вариация критической кривизны на протяжении одного оборота спирали не является малой.

9. Автоволновые структуры в трехмерных возбудимых средах

Если взять на плоскости вращающуюся спиральную волну и продолжить ее прямо вверх, мы получим трехмерную автоволновую структуру — *простой вихрь*, имеющий вид вращающегося свитка (рис. 1.44). По своим свойствам такой вихрь аналогичен спиральной волне. В двумерном случае спиральная волна вращается вокруг некоторой точки; в трехмерном случае центры вращения образуют линию — ось вращения, называемую нитью вихря. Не всегда нить вихря — прямая линия. Она может быть искривлена или замкнута в окружность (рис. 1.45). В результате вихрь превращается в вихревое кольцо (рис. 1.46). Вихревое кольцо — это локальный автоволновой источник. На больших расстояниях от центра этот источник создает ту же картину расходящихся сферических волн, как и пейсмейкер.

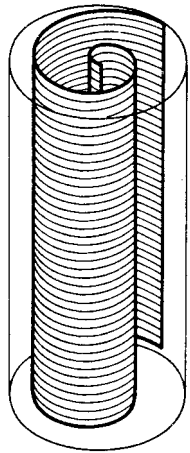


Рис. 1.44. Простой вихрь

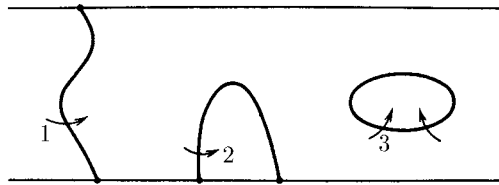


Рис. 1.45. Различные деформации нити вихря; стрелками указано направление вращения

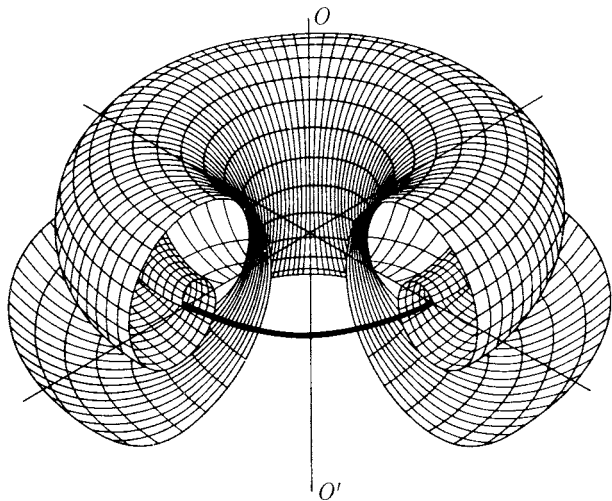


Рис. 1.46. Вихревое кольцо

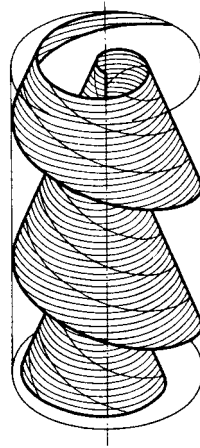


Рис. 1.47. Скрученный вихрь

Вихри и вихревые кольца представляют простейшие примеры автоволновых источников в трехмерных средах. Существуют и гораздо более сложные структуры. Так, если взять простую волну-свиток и скрутить ее вдоль вертикальной оси, мы получим объект, называемый *скрученным вихрем* (рис. 1.47). Пересекаясь с поверхностью воображаемого цилиндра, который охватывает нить такого вихря, волны возбуждения образуют винтовую линию. При замыкании нити скрученного вихря в окружность возникает объект, который естественно назвать *скрученным вихревым кольцом*. Продолжая такие построения, можно получить сложные структуры в виде зацепленных или завязанных узлом вихревых колец и т. п.

Существуют определенные топологические ограничения [267–270], которые значительно сокращают число допустимых трехмерных структур. Оказывается, что запрещено существование двух зацепленных простых вихревых колец или же одиночного скрученного вихревого кольца. Покажем, каким образом возникают такие топологические ограничения. Вспомним прежде всего, что если нить простого вихря окружить замкнутым контуром, то при обходе по такому замкнутому контуру фаза получит приращение, равное $2\pi N$, где N — число рукавов вихря. При обходе вдоль любого контура, не охватывающего нить, приращение фазы равно нулю.

Теперь нетрудно увидеть, что скрученное вихревое кольцо обязательно проткнуто нитью еще одного вихря. Окружим замкнутым контуром ось симметрии скрученного M раз вихревого кольца (величину M называют *полным кручением*). При обходе вдоль такого контура фаза получает приращение $2\pi M$. Следовательно, через такой контур проходит нить вихря с M рукавами или несколько вихревых нитей с суммарным топологическим зарядом ⁶ M .

Помимо описанных выше существуют многорукавные скрученные вихри, которые получаются путем скручивания вдоль оси многорукавных прямых вихрей. Если свернуть нить скрученного многорукавного вихря в окружность, мы приходим к многорукавному вихревому кольцу. Заметим, что полное кручение многорукавного вихревого кольца может принимать дробные значения: поскольку разрешено склеивать один рукав вихря с другим, кручение принимает значения $M = m/N$, где m — целое число.

Изменение фазы при обходе вдоль контура, охватывающего вертикальную ось кольца, равно NM . Повторяя предыдущие рассуждения, можно

⁶Понятие топологического заряда простого вихря предложено в [60]. Топологический заряд спиральной волны в двумерной среде равен деленному на 2π изменению фазы при обходе по направлению часовой стрелки вдоль замкнутого контура, окружающего ядро спиральной волны. Для спиральных волн, вращающихся в противоположных направлениях, знак топологического заряда будет обратным.

убедиться, что через центр такого кольца должен проходить вихрь с NM рукавами (или несколько вихрей с суммарным топологическим зарядом NM).

Нить вихря, проходящего через центр скрученного вихревого кольца, можно саму замкнуть в кольцо. В результате будет получена пара скрученных вихревых колец, зацепленных одно за другое.

Число рукавов и кручение двух вихревых колец не произвольны; они связаны определенным соотношением. Пусть N_1 — число рукавов, а M_1 — кручение первого кольца. Тогда, как следует из предыдущих рассуждений, число рукавов в втором кольце должно равняться $N_2 = N_1 M_1$. Аналогично для числа рукавов в первом кольце получаем $N_1 = N_2 M_2$, где M_2 — кручение второго кольца. Нетрудно видеть, что два равенства могут одновременно выполняться, лишь если

$$M_1 M_2 = 1, \quad \frac{N_1}{N_2} = \sqrt{\frac{M_1}{M_2}}. \quad (1.127)$$

Согласно этим соотношениям, если одно из колец перекручено целое число раз M больше единицы, то второе кольцо скручено лишь на $1/M$ оборотов. При этом число рукавов у второго кольца в M раз больше, чем у первого.

Соотношения (1.127) накладывают жесткие ограничения на допустимые кольцевые зацепления. Из них следует, например, что нельзя продеть сквозь друг друга два простых (нескрученных) вихревых кольца или два однорукавных вихревых кольца, каждое из которых перекручено дважды.

Вихревые нити можно не только замыкать в кольцо, но и завязывать их в узлы. Возникающие при этом объекты очень сложны. Их классификация и отбор разрешенных узлов требуют применения топологических методов [269].

К настоящему времени вихри в трехмерных активных средах наблюдались при реакции Белоусова–Жаботинского [257] и в численных экспериментах [113, 114].

Численные результаты свидетельствуют, что вихревые кольца обычно неустойчивы. Как правило, они медленно сжимаются со скоростью, обратно пропорциональной радиусу кольца, и в конце концов исчезают. В некоторых случаях, однако, наблюдается не сжатие, а раздувание кольца: его радиус растет и оно превращается в ограниченной среде в прямой вращающийся свиток.

Подобно спиральным волнам, в зависимости от значений параметров, характеризующих возбудимость среды, вихревые кольца могут быть рыхлыми или «плотными». В последнем случае испускаемые кольцом волны

следуют с периодом, близким к T_{\min} , так что фронт следующей волны на движется на хвост предыдущей.

Ниже мы кратко изложим методы теоретического описания рыхлых трехмерных автоволновых структур, основывающихся на кинематике фронтальных поверхностей [29].

Кинематическое описание, развитое для двумерных возбудимых сред, допускает естественное обобщение на трехмерный случай. Как и прежде, будем считать, что волна полностью описывается путем задания ее ориентированной поверхности фронта. Всякая поверхность в трехмерном пространстве обладает в каждой точке двумя главными радиусами кривизны: R_1 и R_2 . Можно сказать, что при малых кривизнах скорость нормального смещения участка фронтальной поверхности зависит лишь от суммы двух главных кривизн, т. е. от удвоенной средней кривизны $2H = 1/R_1 + 1/R_2$, согласно формуле

$$V = V_0 - 2DH. \quad (1.128)$$

На линии своего обрыва фронтальная поверхность может «прорасти» или сократиться. Скорость тангенциального смещения (скорость «прорастания» искривленной волны) G следующим образом зависит от средней кривизны H фронта при подходе к обрыву в данной точке и от тангенциальной (или геодезической) кривизны κ линии обрыва:

$$G = -\gamma_1(2H - K_{\text{кр}}) - \gamma_2\kappa. \quad (1.129)$$

Здесь γ_1 и γ_2 — некоторые положительные коэффициенты; кривизну линии обрыва мы считаем положительной, если она является выпуклой по отношению к фронтальной поверхности.

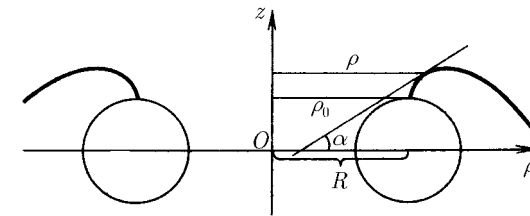


Рис. 1.48. К определению величин α и ρ

Рассмотрим в рамках кинематического описания эволюцию простого (нескрученного) вихревого кольца. Введем цилиндрическую систему координат z, ρ, φ с осью z , являющейся осью аксиальной симметрии кольца.

Вследствие цилиндрической симметрии достаточно рассмотреть эволюцию линии, образуемой сечением кольца плоскостью $\varphi = \text{const}$, т. е. меридиана вихревого кольца (рис. 1.48). Поскольку эта линия плоская, то, очевидно, кинематическое уравнение, описывающее ее эволюцию, совпадает с (1.95). Однако в отличие от случая спиральной волны на плоскости зависимость скоростей V и G от кривизны K меридиана вихревого кольца дается соотношениями (1.128) и (1.129). Средняя кривизна H для вихревого кольца легко вычисляется:

$$H = \frac{K - \rho^{-1} \sin \alpha}{2}, \quad (1.130)$$

а тангенциальная кривизна линии обрыва

$$\kappa = -\frac{1}{\rho_0} \cos \alpha_0, \quad (1.131)$$

где ρ — расстояние от оси z до соответствующего участка поверхности, α — угол между касательной к меридиану кольца на этом участке и осью ρ (см. рис. 1.48), α_0 — угол между касательной и осью ρ при подходе к свободному концу меридиана, ρ_0 — расстояние от оси z до свободного конца.

В условиях применимости квазистационарного приближения форма фронтальной поверхности вблизи линии обрыва быстро подстраивается к мгновенным значениям средней кривизны H_0 этой поверхности вблизи обрыва и кривизны самой линии обрыва κ . Поэтому достаточно следить за движением концевой точки линии, получаемой путем сечения вихревого кольца меридиональной плоскостью. Можно показать, что оно описывается уравнениями

$$\dot{\rho}_0 = -V(0) \sin \alpha_0 - G \cos \alpha_0, \quad (1.132)$$

$$\dot{z} = V(0) \cos \alpha_0 - G \sin \alpha_0, \quad (1.133)$$

$$\dot{\alpha}_0 = -D \left(\frac{\partial K}{\partial l} \right)_{l=0} - DK_0 \frac{\cos \alpha_0}{\rho_0} - \frac{D}{2\rho_0^2} \sin 2\alpha_0 + GK_0, \quad (1.134)$$

$$\dot{K}_0 = - \left(G - D \frac{\cos \alpha_0}{\rho_0} \right) - \frac{DK_0}{2\rho_0} (3 \cos 2\alpha_0 + 1), \quad (1.135)$$

где

$$\left(\frac{\partial K}{\partial l} \right)_{l=0} = \zeta \sqrt{\frac{V_0}{D}} \sqrt{K_0^3}.$$

В сочетании с (1.128)–(1.131) эти уравнения полностью определяют закон движения точки обрыва в меридиональной плоскости и, следовательно, временную эволюцию простого вихревого кольца.

Мы считаем, что радиус R вихревого кольца гораздо больше толщины его нити, которая близка по значению к диаметру ядра спиральной волны. Нетрудно убедиться, что при $R \rightarrow \infty$, т. е. для цилиндрического вихря, уравнения (1.132)–(1.135) имеют решение, для которого частота вращения совпадает с частотой спиральной волны ($\alpha_0 = \omega_0 t$), а кривизна K_0 равна $K_{\text{кр}}$. В линейном приближении по $1/R$ следует заменить в правых частях уравнений (1.134), (1.135) величину ρ_0 на R и положить $\alpha_0 \approx \omega_0 t$. В результате мы приходим к уравнениям, сходным с теми, которые описывают резонанс спиральных волн при полном совпадении двух частот. Можно поэтому ожидать, что в меридиональной плоскости центр нити вихря будет медленно смещаться со скоростью, пропорциональной амплитуде модуляции, т. е. величине $1/R$.

Подробный расчет [29] показывает, что в линейном по $1/R$ приближении радиус вихревого кольца изменяется с течением времени по закону

$$\dot{R} = -\frac{D}{R} \left(1 - \frac{3}{4\beta^2} \frac{q_1^2 + q_2}{q_1^2 + 1} \right), \quad (1.136)$$

где $\beta = \sqrt{DK_{\text{кр}}/V_0} \ll 1$ и использованы обозначения $q_1 = \gamma_1/D$ и $q_2 = (\gamma_2 - D)/D$. Одновременно кольцо смещается в вертикальном направлении (т. е. вдоль своей оси симметрии) со скоростью

$$\dot{z}_0 = -\frac{D}{R} \left(\frac{q_2}{2\zeta\beta} + \frac{3}{4\beta^2} \frac{q_1 q_2}{q_1^2 + 1} \right). \quad (1.137)$$

Если вихрь закручен в противоположную сторону, то направление его смещения вдоль оси симметрии меняется на обратное.

Таким образом, вихревое кольцо почти всегда неустойчиво. Оно либо коллапсирует ($\dot{R} < 0$) и исчезает, либо раздувается ($\dot{R} > 0$) с течением времени. Сжатие или раздувание сопровождаются дрейфом кольца вдоль оси своей симметрии.

Имеется, однако, узкий интервал значений параметров q_1 и q_2 , в котором скорость \dot{R} очень мала и проходит через нулевое значение (см. (1.136)). В пределах этого интервала наряду с линейным по $1/R$ слагаемым в выражении для \dot{R} необходимо учитывать также члены следующих, более высоких порядков. Детальный анализ показывает⁷, что слагаемые порядка $1/R^2$ отсутствуют, а член порядка $1/R^3$ входит в выражение для скорости \dot{R} с положительным коэффициентом. Поэтому, если в линейном по $1/R$ приближении вихревое кольцо сжимается, но скорость сжатия очень мала, это

⁷Эти расчеты были проведены одним из авторов (А. С. М.) совместно с Н. В. Хрустовой.

нелинейное положительное слагаемое порядка $1/R^3$ будет ограничивать сжатие и стабилизировать вихревое кольцо.

При выполнении неравенств $q_1^2 \gg q_2$ радиус стабильного вихревого кольца дается выражением

$$R = R_0 \sqrt{\frac{13}{32} \xi \beta q_1} \frac{1}{\sqrt{(q_1/\beta)^2 - 3/4}}, \quad (1.138)$$

где $R_0 = V_0/\omega_0$ есть радиус нити вихревого кольца, практически совпадающий с радиусом ядра спиральной волны в двумерной возбудимой среде.

Как следует из (1.138), устойчивые вихревые кольца существуют при $q_1^2 > (3/4)\beta^2$, причем непосредственно при $q_1^2 = (3/4)\beta^2$ радиус устойчивого вихревого кольца обращается в бесконечность. С увеличением q_1 радиус кольца быстро убывает. Выражение (1.138) справедливо, пока $R \gg R_0$. Решение в виде устойчивого вихревого кольца теряется, когда R оказывается порядка R_0 .

Заметим, что, согласно (1.137), стабильное вихревое кольцо медленно смещается со скоростью около $\beta(D/R)$ вдоль своей оси симметрии.

Стабильные вихревые кольца наблюдались в численном эксперименте [29] для одной конкретной модели возбудимой среды.

Мы уже обращали внимание, что на больших удалениях от центра вихревое кольцо неотличимо по картине испускаемых им волн от сферического пейсмейкера. Однако в отличие от истинного пейсмейкера, обязанного своим существованием неоднородности среды и присутствием в ней группы автоколебательных элементов, такой источник может быть создан в однородной среде; при этом он не покоится — его центр смещается со временем вдоль направления оси симметрии вихревого кольца.

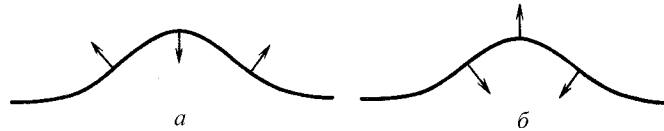


Рис. 1.49. Эволюция начального возмущения нити прямого вихря в среде с более высокой (а) и менее высокой (б) возбудимостью

Хотя приведенные выше результаты относились непосредственно лишь к эволюции простого вихревого кольца, они позволяют описать в простейшем приближении поведение со временем любого вихря, нить которого

произвольно (но не очень сильно) деформирована. В самом деле, если рассмотреть небольшой участок такого вихря, он будет выглядеть как часть вихревого кольца соответствующего радиуса.

Пусть исходно мы имели прямой вихрь (цилиндрический свиток) и немного деформировали локально его нить (рис. 1.49а). Если параметры среды таковы, что вихревое кольцо в ней сжимается, деформация с течением времени будет стремиться стать меньше. Иными словами, в этом случае нить вихря обладает упругостью и стремится к сокращению с течением времени. В обратном случае, когда вихревое кольцо расширяется, всякий «вырост» на нити раздувается со временем — нить вихря стремится к удлинению (рис. 1.49б). В результате прямой цилиндрический свиток оказывается неустойчивым по отношению к малым деформациям его нити. Это должно приводить к формированию очень сложных (и, возможно, даже хаотически организованных) структур в безграничных средах.

Недавно был рассмотрен [1] резонанс вихревых колец, т. е. их реакция на периодическую однородную модуляцию возбудимости среды. Оказывается, что при определенных условиях, подбирая амплитуду и фазу модуляции, можно стабилизировать раздувающиеся вихревые кольца и управлять скоростью их смещения вдоль центральной оси.

Обсудим в заключение свойства скрученных вихрей. Для цилиндрического скрученного свитка (рис. 1.47) линия обрыва волнового фронта представляет собой винтовую линию, намотанную на цилиндр — нить вихря. Скрученность такого вихря естественно характеризовать параметром $\mu = 2\pi/h$, где h — шаг винтовой линии. В зависимости от того, является ли винтовая линия обрыва правой или левой (положительное направление оси вихря определяется вектором угловой скорости вращения), скрученность μ будет, соответственно, положительной или отрицательной.

Кинематические расчеты показывают, что угловая скорость вращения скрученного свитка больше, чем нескрученного. При малых значениях скрученности μ угловая скорость вращения скрученного свитка равна [47]

$$\omega = \omega_0 \left[1 + \frac{3}{2} \frac{\gamma_2}{\gamma_1} \frac{|\mu|}{K_{\text{кр}}} \right]. \quad (1.139)$$

Результаты численного моделирования [208] свидетельствуют, что для каждой среды существует максимальное допустимое значение скрученности $\mu_{\text{кр}}$: если вихрь скрутить более сильно, он рвется. Максимально скрученный вихрь испускает волны с периодом порядка T_{min} , т. е. с периодом, близким к минимальному возможному интервалу следования волн возбуждения.

Если в начальный момент времени вихрь скручен неоднородно, т. е. $\mu = F(z)$, то по нему побегут «волны скрученности» [47]. Возмущения скрученности, имеющие разные знаки, распространяются вдоль вихря в противоположных направлениях, но с одинаковой скоростью

$$u = \frac{3}{2} \frac{\gamma_2}{\gamma_1} \frac{\omega_0}{K_{\text{кр}}}. \quad (1.140)$$

При столкновении двух волн скрученности образуются разрывы (аналог ударных волн).

Заметим, что приведенные результаты относятся, конечно, и к скрученным вихревым кольцам, для которых нить замкнута в окружность. Скрученные вихревые кольца сжимаются (или раздуваются) и дрейфуют вдоль аксиальной оси примерно так же, как и нескрученные кольца, поведение которых обсуждалось нами выше.

Если внутри конечного объема возбудимой среды создать цилиндрический скрученный свиток, так что его нить будет выходить на ограничивающие этот объем поверхности, то, как показывает численное моделирование, этот свиток нестационарен. Он раскручивается и с течением времени превращается в простой вихрь. Однако скрученные вихри естественно возникают в *неоднородных* возбудимых средах.

Пусть возбудимая среда состоит из двух областей. В верхней области спиральные волны (и нескрученные вихри) имеют период вращения T_1 , а в нижней период вращения спиральных волн равен T_2 , причем $T_2 > T_1$. Если теперь создать простой вихрь с нитью, перпендикулярной поверхности раздела двух областей возбудимой среды, то в нижней он превратится в скрученный вихрь (см. [208]). Степень скрученности вихря определяется условием совпадения периода его вращения T с периодом вращения T_1 нескрученного вихря в верхней области среды (напомним, что скручивание приводит к ускорению вращения вихря).

В неоднородных средах, где значения параметров плавно варьируются, вихри, как правило, должны быть скрученными. Будучи более быстрыми, такие вихри вытесняют простые нескрученные вихри из активной среды.

10. Фазовая динамика в осциллирующих активных средах

Пусть отдельно взятый элемент активной среды совершает устойчивые периодические автоколебания. Эти колебания характеризуются своей формой (не обязательно гармонической), амплитудой, частотой и начальной

фазой. При этом амплитуда, частота и форма колебаний однозначно определяются свойствами самого автоколебательного элемента, а начальная фаза колебаний φ может быть произвольной. Если произвести некоторое возмущение, форма осцилляций вернется к исходной, а их амплитуда возвратится к своему устойчивому значению по прошествии некоторого времени релаксации $\tau_{\text{рел}}$. Однако в результате возмущения фаза φ не вернется уже к исходному значению, а приобретет некоторую добавку $\delta\varphi$.

Допустим далее, что в нашем распоряжении имеются сеть из связанных между собой автоколебательных элементов либо распределенная активная среда, отдельный физически малый элемент которой обладает автоколебательными свойствами. Будем предполагать, что взаимодействие между элементами является локальным (т. е. взаимодействуют лишь соседние элементы), причем если два соседних элемента совершают синхронные колебания (с одинаковой формой, амплитудой и начальной фазой), то взаимодействие между ними отсутствует; если же рассогласование в колебаниях двух элементов мало, то и взаимодействие между ними будет слабым в меру такого рассогласования.

Очевидно, что при сделанных предположениях активная среда допускает режим синхронных автоколебаний, при котором все элементы среды совершают колебания с одной и той же фазой $\varphi(\mathbf{r}, t) = \text{const}$, а амплитуда и форма колебаний в каждой точке среды — те же, что и для отдельно взятого автоколебательного элемента.

Более того, если фаза колебаний меняется вдоль среды, но ее пространственное изменение является плавным и характеризуется достаточно большим пространственным масштабом L , следует ожидать, что тогда распределение фазы $\varphi(\mathbf{r}, t)$ будет медленно меняться со временем. Действительно, чем больше пространственный масштаб L изменения фазы, тем меньше рассогласование по фазам колебаний для двух соседних элементов среды и тем слабее взаимодействие.

С другой стороны, возмущения амплитуды и формы колебаний релаксируют быстро, независимо от того, каков их характерный пространственный масштаб. Их затухание определяется свойствами отдельно взятого элемента среды и происходит за время порядка $\tau_{\text{рел}}$.

Таким образом, для возмущений с достаточно большими пространственными масштабами характерные времена для изменения фаз $\varphi(\mathbf{r}, t)$ и амплитуд колебаний сильно различаются. Поскольку амплитуда колебаний в каждой точке стремится быстро (и независимо от других точек) возвратиться к своему устойчивому значению, при наличии слабого градиента фазы $\nabla\varphi$ амплитуды колебаний быстро адиабатически подстраиваются к мгновенному значению этого градиента. Поэтому их отклонения от рав-

новесного значения малы в меру малости градиента фазы $\nabla\varphi$ (т. е. в меру слабости взаимодействия между элементами среды).

Динамика плавных распределений фаз должна описываться отдельным уравнением, в которое уже не входят амплитуды и форма колебаний. Это уравнение можно вывести, воспользовавшись весьма общими соображениями.

В однородной активной среде скорость изменения фазы в данной точке среды $\partial\varphi/\partial t$ не может зависеть от направления градиента фазы. Поэтому при достаточно малых градиентах разложение скорости изменения фазы по степеням градиента фазы должно начинаться не с линейного, а с квадратичного слагаемого:

$$\frac{\partial\varphi}{\partial t} = a(\nabla\varphi)^2 + \dots \quad (1.141)$$

Согласно (1.141), для неоднородностей с пространственным масштабом L скорость изменения фазы пропорциональна $1/L^2$. Заметим теперь, что, наряду с градиентом, в правую часть уравнения (1.141) могут входить также слагаемые с более высокими пространственными производными. Нам необходимо сохранить только те из подобных слагаемых, которые приводят к вкладам порядка $1/L^2$ по обратному пространственному масштабу неоднородности фазы. Нетрудно убедиться, что есть всего одно такое слагаемое, пропорциональное лапласиану фазы $\Delta\varphi$.

Итак, уравнение для плавных изменений фазы φ должно иметь вид

$$\frac{\partial\varphi}{\partial t} = a(\nabla\varphi)^2 + b\Delta\varphi, \quad (1.142)$$

где a и b — некоторые коэффициенты с размерностью коэффициента диффузии (т. е. $[a] = [b] = \text{см}^2/\text{с}$). Чтобы определить значение этих коэффициентов, нужно использовать полные уравнения динамики для данной активной среды (см. ниже).

Заметим, что, подобно уравнениям кинематики для возбудимых сред, рассматривавшимся в разделах 1.8 и 1.9, выведенное уравнение фазовой динамики (1.142) универсально в том смысле, что всякая сложная осциллирующая среда описывается в нем всего двумя параметрами a и b , а все детали взаимодействия между элементами, формы и амплитуды колебаний оказываются несущественными.

Из уравнения (1.142) следует, что характерное время τ_L для изменения неоднородности фазы с пространственным масштабом L имеет порядок величины $\tau_L \sim L^2/b$. Поскольку при выводе этого уравнения предполагалось,

что характерное время изменения для распределения фазы велико по сравнению со временем релаксации амплитуды $\tau_{\text{рел}}$ в отдельном осцилляторе, полученное уравнение применимо лишь для описания таких плавных распределений фазы, для которых пространственный масштаб неоднородности удовлетворяет условию

$$L \gg \sqrt{b\tau_{\text{рел}}}. \quad (1.143)$$

Неравенство (1.143) ограничивает область применимости уравнения фазовой динамики.

Детальная схема вывода уравнения фазовой динамики, непосредственно исходя из уравнений в частных производных, описывающих автоколебательную распределенную среду, приведена в [190]. Ниже мы даем, в качестве иллюстрации, вывод этого уравнения для *квазигармонических* автоколебательных сред.

Допустим, что автоколебательная среда описывается двумя зависящими от времени и координат переменными $u = u(\mathbf{r}, t)$ и $v = v(\mathbf{r}, t)$, из которых удобно составить комплексную величину $\eta(\mathbf{r}, t) = u(\mathbf{r}, t) + iv(\mathbf{r}, t)$. Предположим, что изменение η со временем подчиняется следующему уравнению:

$$\frac{\partial\eta}{\partial t} = \lambda(\rho)\eta - i\omega(\rho)\eta + (D_1 + iD_2)\Delta\eta, \quad (1.144)$$

где λ и ω представляют собой некоторые функции от амплитуды $\rho = |\eta|$; функция $\lambda(\rho)$ обращается в нуль при $\rho = \rho_0$, отрицательна — при $\rho > \rho_0$ и положительна — при $\rho < \rho_0$.

Модель автоколебательной среды, характеризуемую уравнением вида (1.144), называют обычно « λ - ω -модель». Если комплексный коэффициент $D = D_1 + iD_2$, играющий роль коэффициента диффузии в уравнении (1.144), равен нулю, то среда распадается на совокупность не связанных между собой автоколебательных элементов. Нетрудно убедиться, что в установившемся режиме каждый элемент среды совершает гармонические колебания по закону

$$\eta(t) = \rho_0 \exp[-i(\omega_0 t + \varphi)] \quad (1.145)$$

с амплитудой ρ_0 , определяемой из условия $\lambda(\rho_0) = 0$, и частотой $\omega_0 = \omega(\rho_0)$. Начальная фаза φ этих колебаний остается произвольной.

Малые возмущения амплитуды $\delta\rho = \rho - \rho_0$ для отдельного элемента среды затухают со временем согласно уравнению (сп. (1.144))

$$\delta\dot{\rho} = \rho_0\lambda'(\rho_0)\delta\rho. \quad (1.146)$$

Таким образом, время релаксации амплитуды равно

$$\tau_{\text{рел}} = \frac{1}{|\rho_0(d\lambda/d\rho)_{\rho=\rho_0}|}. \quad (1.147)$$

Рассмотрим теперь общий случай, когда $D \neq 0$. Если представить η в виде

$$\eta(\mathbf{r}, t) \equiv \rho(\mathbf{r}, t) \exp[-i(\omega_0)t + \varphi(\mathbf{r}, t)], \quad (1.148)$$

то для неизвестных функций $\rho(\mathbf{r}, t)$ и $\varphi(\mathbf{r}, t)$ можно получить уравнения

$$\frac{\partial \rho}{\partial t} = \lambda(\rho)\rho + D_1 \Delta \rho - D_1 \rho (\nabla \varphi)^2 + D_2 \rho \Delta \varphi + 2D_2 \nabla \rho \nabla \varphi, \quad (1.149)$$

$$\frac{\partial \varphi}{\partial t} = [\omega(\rho) - \omega_0] + 2D_1 \frac{\nabla \rho \nabla \varphi}{\rho} - D_2 \frac{\Delta \rho}{\rho} + D_2 (\nabla \varphi)^2 + D_1 \Delta \varphi. \quad (1.150)$$

Ограничимся рассмотрением плавных распределений, характеризуемых большим пространственным масштабом L . Тогда следует ожидать, что амплитуда колебаний в каждой точке среды будет близка к ρ_0 , т. е. $\delta\rho/\rho_0 \ll 1$. Более того, отклонения амплитуд $\delta\rho(\mathbf{r}, t)$ будут адиабатически подстраиваться к значениям $\nabla\varphi$ и $\Delta\varphi$ в соответствующих точках. Поэтому с точностью до членов порядка $1/L^2$ мы имеем

$$\delta\rho = \rho_0 \tau_{\text{рел}} [D_2 \Delta \varphi - D_1 (\nabla \varphi)^2]. \quad (1.151)$$

Действительно, подставив (1.151) в (1.149), легко убедиться, что при выводе этого выражения отброшены как раз те слагаемые, которые отвечают более высокому порядку по параметру $1/L$.

Теперь остается только подставить $\rho = \rho_0 + \delta\rho$ в уравнение (1.150) для фазы и сохранить в нем лишь члены порядка $1/L^2$. В результате получаем

$$\frac{\partial \varphi}{\partial t} = D_1 \Delta \varphi + D_2 (\nabla \varphi)^2 + \omega'(\rho_0) \delta\rho, \quad (1.152)$$

где $\delta\rho$ дается выражением (1.151). С учетом формулы (1.147), справедливо

$$\frac{\partial \varphi}{\partial t} = \left(-D_1 \frac{\omega'(\rho_0)}{|\lambda'(\rho_0)|} + D_2 \right) (\nabla \varphi)^2 + \left(D_1 + \frac{\omega'(\rho_0)}{|\lambda'(\rho_0)|} D_2 \right) \Delta \varphi, \quad (1.153)$$

что совпадает с уравнением (1.142).

Частным случаем λ - ω -модели является *обобщенное уравнение Гинзбурга-Ландау*

$$\frac{\partial \eta}{\partial t} = (\alpha_1 - i\alpha_2)\eta - (\beta_1 + i\beta_2)|\eta|^2\eta + (D_1 + iD_2)\Delta\eta, \quad (1.154)$$

для которого $\lambda = \alpha_1 - \beta_1\rho^2$, $\omega = \alpha_2 + \beta_2\rho^2$. Как следует из (1.153), в этом случае параметры a и b в уравнении фазовой динамики (1.142) таковы:

$$a = -\frac{\beta_2}{\beta_1}D_1 + D_2, \quad b = \frac{\beta_2}{\beta_1}D_2 + D_1. \quad (1.155)$$

Перейдем теперь к рассмотрению процессов, описываемых уравнением фазовой динамики (1.142). Заметим прежде всего, что с помощью подстановки

$$\varphi(\mathbf{r}, t) = \frac{b}{a} \ln Q(\mathbf{r}, t) \quad (1.156)$$

оно сводится к линейному уравнению

$$\frac{\partial Q}{\partial t} = b \Delta Q. \quad (1.157)$$

Таким образом, решающую роль в эволюции системы играет знак коэффициента b . Если этот коэффициент положителен, уравнение (1.157) представляет собой обычное уравнение диффузии. Следовательно, любая локальная неоднородность Q (и, соответственно, φ) с характерным пространственным масштабом L полностью размывается и исчезает за время порядка L^2/b . В результате в среде устанавливается однородный режим, для которого Q и φ не зависят от пространственных координат. Это означает, что при $b > 0$ происходит *синхронизация* колебаний во всех точках среды.

Совершенно иная ситуация складывается, если коэффициент b отрицателен. В этом случае начальная неоднородность фазы не размывается, а, наоборот, сжимается и возрастает с характерным временем L^2/b . Вследствие этого однородный режим с синхронным автоколебаниями оказывается неустойчивым по отношению к малым возмущениям. Флуктуации усиливаются, причем быстрее всего идет рост наиболее коротковолновых флуктуаций. В результате в системе устанавливается хаотический, или *турбулентный*, режим [192].

Заметим, что, поскольку эволюция локальных неоднородностей фазы сопровождается их стягиванием, т. е. уменьшением пространственного масштаба, для каждой такой флуктуации по прошествии некоторого времени

условие (1.143) оказывается нарушенным, а уравнение (1.142) — неприменимым. Турбулентный режим характеризуется множественным рождением сингулярностей в распределении фазы. Для его описания необходимо использовать полные исходные уравнения распределенной автоколебательной среды. Свойства турбулентного режима оказываются поэтому не универсальными, а весьма чувствительными к конкретному виду уравнений, описывающих распределенную автоколебательную среду (см. раздел 2.31).

11. Фазовые волны и пейсмекеры

Как уже отмечалось в разделе 1.1, при распространении фазовой волны в автоколебательной среде в ней устанавливается постоянный градиент фазы колебаний

$$\nabla\varphi = -\mathbf{k} = \text{const.} \quad (1.158)$$

Благодаря этому вдоль направления вектора \mathbf{k} колебания соседних элементов происходят с постоянным сдвигом по фазе, и для наблюдателя этот режим выглядит так, как будто по среде движется волна. Эффектом того же происхождения являются «волны», бегущие по гирляндам из электрических лампочек, зажигаемых по очереди на короткое время.

Для возбуждения фазовых волн в среде необходимо создать исходный градиент фазы автоколебаний. Скорость волны при этом тем больше, чем меньше линейный градиент фазы.

Связь между элементами среды (осуществляемая, например, благодаря диффузии) не имеет принципиального значения для распространения быстрых фазовых волн, когда градиенты фазы достаточно малы⁸.

При уменьшении скорости фазовой волны, т. е. при возрастании градиента фазы автоколебаний, эффекты взаимодействия становятся все более существенными. Вообще говоря, при распространении по среде фазовой волны с волновым числом k элементы среды совершают колебания с частотой ω , несколько отличной от частоты однородных автоколебаний ω_0 .

Распространение фазовых волн можно исследовать в рамках уравнения фазовой динамики (1.142). Такая волна отвечает частному решению этого уравнения, имеющему вид

$$\varphi = -\mathbf{k}\mathbf{r} + ak^2t. \quad (1.159)$$

⁸В эксперименте [182] осциллирующую среду с реакцией Белоусова–Жаботинского, где были предварительно возбуждены быстрые волны, разделяли тонкой перегородкой из оргстекла и наблюдали, как волны «проходят» через перегородку, не ощущая ее.

Зависящее от времени слагаемое дает сдвиг частоты колебаний

$$\omega = \omega_0 + ak^2. \quad (1.160)$$

Таким образом, знак коэффициента a определяет, будет расти или же убывать частота колебаний с увеличением волнового числа. Скорость распространения фазовой волны равна

$$c = \frac{\omega_0}{k} + ak. \quad (1.161)$$

В рамках уравнения (1.142) все фазовые волны устойчивы, если $b > 0$. При $b < 0$ частное решение в виде фазовой волны по-прежнему существует, но оно уже не представляет интереса, поскольку в этом случае фазовая волна, как и другие регулярные волновые режимы, неустойчива и в среде формируется с течением времени хаотический пространственно-временной режим (раздел 2.31).

Волны с малым пространственным периодом, для которых $k \gg (b\tau_{\text{пер}})^{-1/2}$, не описываются уравнением фазовой динамики (1.142). Их исследование необходимо производить в рамках полных исходных уравнений распределенной автоколебательной среды. Обычно оказывается, что волны теряют устойчивость в области $k \sim (b\tau_{\text{пер}})^{-1/2}$.

Рассмотрим теперь, что происходит при столкновении двух волн с разными волновыми векторами. Для простоты будем считать, что волны движутся навстречу друг другу вдоль оси x , причем $\varphi(x) \rightarrow -k_1x$ при $x \rightarrow -\infty$ и $\varphi(x) \rightarrow k_2x$ при $x \rightarrow +\infty$. Такому режиму отвечает частное решение уравнения (1.157) вида

$$Q(x, t) = A_1 \exp\left[-\frac{a}{b}k_1x + \frac{a^2}{b}k_1^2t\right] + A_2 \exp\left[\frac{a}{b}k_2x + \frac{a^2}{b}k_2^2t\right], \quad (1.162)$$

где коэффициенты A_1 и A_2 находятся из начальных условий. Напомним, что

$$\varphi(x, t) = \frac{b}{a} \ln Q(x, t). \quad (1.163)$$

Поскольку Q представляет собой сумму двух экспоненциальных вкладов, независимо меняющихся по x и t , почти всюду на оси x будет доминировать какое-то одно из этих двух слагаемых. Простой анализ показывает, что вне узкого слоя шириной порядка

$$\delta x \sim \frac{b}{a(k_1 + k_2)} \quad (1.164)$$

с центром в точке

$$X(t) = a(k_1 - k_2)t \quad (1.165)$$

справа (т. е. при $x > X(t)$) преобладает второе, а слева (т. е. при $x < X(t)$) — первое слагаемое в выражении (1.162). Поскольку (см. (1.163)) первое слагаемое отвечает волне, движущейся в положительном, а второе — в отрицательном направлении оси x , можно сказать, что две движущиеся навстречу фазовые волны гасят друг друга в пределах узкого слоя шириной $\delta x \sim b/[a(k_1 + k_2)]$, а точка их встречи медленно смещается с постоянной скоростью $V_B = a(k_1 - k_2)$ вдоль оси x .

Заметим, что, учитывая зависимость $\omega = \omega_0 + ak^2$ для частоты фазовой волны от ее волнового числа, выражение для скорости смещения границы между областями, занятыми двумя волнами, можно представить также в виде

$$V_B = \frac{\omega_1 - \omega_2}{k_1 + k_2}. \quad (1.166)$$

Таким образом, с течением времени граница смещается вглубь области, занятой менее высокочастотной волной. Иными словами, область с более высокочастотными колебаниями постепенно растет и вытесняет ту область, в которой колебания имеют более низкую частоту.

До сих пор мы рассматривали только эффекты в однородных автоколебательных средах, где свойства всех элементов среды идентичны. Пусть в пределах некоторой малой области элементы среды имеют собственную частоту автоколебаний, превышающую частоту колебаний остальных элементов среды. Тогда эта область станет источником концентрически расходящихся фазовых волн, или *пейсмекером*.

Предположим, что собственная частота автоколебаний отдельно взятых элементов среды зависит от координат r по некоторому закону $\omega(r)$, причем $\omega(r) \rightarrow \omega_0$ при $r \rightarrow \infty$. При выполнении определенных условий (см. ниже) фазовая динамика в такой системе описывается уравнением

$$\frac{\partial \varphi}{\partial t} = \omega(r) - \omega_0 + a(\nabla \varphi)^2 + b \Delta \varphi. \quad (1.167)$$

После преобразования (1.156) оно сводится к линейному уравнению

$$\frac{\partial Q}{\partial t} = b \Delta Q + \frac{a}{b} [\omega(r) - \omega_0] Q, \quad (1.168)$$

которое оказывается эквивалентным квантовому уравнению Шредингера

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi + U(r) \psi, \quad (1.169)$$

если положить

$$\psi \rightarrow Q, \quad \frac{\hbar^2}{2m} \rightarrow b, \quad U(r) \rightarrow -\frac{a}{b} [\omega(r) - \omega_0] \quad (1.170)$$

и считать время мнимым: $(i/\hbar)t \rightarrow t$. Когда частота $\omega(r)$ превышает ω_0 внутри некоторой области, то в терминах уравнения Шредингера это означает, что в данной области имеется некоторая потенциальная яма. Пусть характерный размер такой области равен r_0 .

Уравнение (1.168) имеет общее решение

$$Q(r, t) = \sum_n C_n e^{\lambda_n t} \mu_n(r), \quad (1.171)$$

где λ_n и $\mu_n(r)$ являются решениями задачи на собственные значения

$$L\mu = \lambda\mu \quad (1.172)$$

для линейного оператора

$$L = b\Delta + \frac{a}{b} [\omega(r) - \omega_0]. \quad (1.173)$$

Положительные собственные значения λ_n соответствуют связанным состояниям частицы в потенциальной яме $U(r)$. Отрицательные собственные значения λ_n образуют сплошной спектр; поскольку, однако, вклад от таких членов в Q экспоненциально затухает со временем, его можно не учитывать. Наконец, поскольку $U(r) \rightarrow 0$ при $r \rightarrow \infty$, существуют также нулевое собственное значение $\lambda_0 = 0$ и соответствующая собственная функция $\mu_0(r)$. Как известно [94], для любого связанного состояния частицы в потенциальной яме волновая функция экспоненциально спадает на достаточно больших удалениях от центра ямы. Ввиду аналогии с уравнением Шредингера имеем

$$\mu_n(r) \approx \text{const} \cdot \exp\left(-\frac{r}{r_n}\right), \quad r \gg r_0, \quad (1.174)$$

где r_n — характерный радиус локализации,

$$r_n = \sqrt{\frac{b}{\lambda_n}}. \quad (1.175)$$

Для нулевого собственного значения собственная функция $\mu_0(r) \approx 1$ вдали от области, где сосредоточено возмущение $\omega(r)$.

Предположим вначале, что в потенциальной яме $U(r)$ имеется всего одно связанное состояние (т.е. есть только одно положительное значение λ_1 и соответствующая собственная функция $\mu_1(r)$). Тогда, пренебрегая затухающим вкладом от собственных функций сплошного спектра, с учетом (1.161) и (1.163), получаем

$$\varphi(r, t) = \frac{b}{a} \ln[C_0 \mu_0(r) + C_1 e^{\lambda_1 t} \mu_1(r)]. \quad (1.176)$$

При $r \gg r_0$, т.е. вдали от области, где локализовано возмущение $\omega(r)$, выражение (1.176) упрощается:

$$\varphi(r, t) = \frac{b}{a} \ln \left\{ C_0 + C_1 \exp \left[\lambda_1 t - \sqrt{\frac{\lambda_1}{b}} r \right] \right\}. \quad (1.177)$$

Это решение описывает рождающийся пейсмейкер. Внутри растущей сферической области радиуса $\mathcal{R}(t) = \sqrt{b\lambda_1}t$ мы имеем систему концентрических расходящихся фазовых волн

$$\varphi(r, t) \approx \frac{b}{a} \left[\lambda_1 t - \sqrt{\frac{\lambda_1}{b}} r \right] \quad (1.178)$$

с частотой $\omega = \omega_0 + (b/a)\lambda_1$ и волновым числом $k = \sqrt{\lambda_1 b/a^2}$. Вне этой области сохраняются однородные автоколебания с частотой ω_0 . С течением времени область, заполненная расходящимися волнами, расширяется и захватывает всю среду.

Если в потенциальной яме, отвечающей возмущению $\omega(r)$, содержится несколько связанных состояний (т.е. есть несколько различных положительных собственных значений λ_n), картина существенно не меняется. Вместо λ_1 в выражениях (1.177) и (1.178) для растущего пейсмейкера теперь следует брать максимальное из значений λ_n , отвечающее наиболее глубокому уровню в данной потенциальной яме.

Известно [94], что в одномерном и двумерном случаях любая сколь угодно мелкая потенциальная яма содержит хотя бы одно связанное состояние. Следовательно, для одномерной и двумерной автоколебательной среды достаточно сколь угодно слабого локального возмущения частоты $\omega(r)$, чтобы родился пейсмейкер, действие которого охватит со временем всю активную среду. В отличие от этого в трехмерном случае потенциальная яма содержит связанное состояние, лишь если она является достаточно

глубокой. Переформулируя соответствующее условие с учетом соотношений (1.170), можно получить условие, при выполнении которого имеется хотя бы одно положительное собственное значение λ . Это условие таково:

$$\delta\omega \gg \frac{b^2}{ar_0^2}, \quad (1.179)$$

где $\delta\omega$ — характерное возмущение частоты, а r_0 — характерный размер области, где такое возмущение локализовано. Когда условие (1.179) нарушено, рождение пейсмейкера в трехмерной неоднородной автоколебательной среде не происходит.

Заметим, наконец, что само описание с помощью уравнения фазовой динамики (1.173) справедливо, если для максимального из собственных значений λ_n радиус локализации соответствующей функции $\mu_n(r)$ превышает $\sqrt{b\tau_{\text{рел}}}$. Поскольку, согласно (1.175), этот радиус равен $\sqrt{b/\lambda_{\text{max}}}$, должно выполняться условие $\lambda_{\text{max}}\tau_{\text{рел}} \ll 1$. Поскольку всегда $\lambda_{\text{max}} \leq \delta\omega_{\text{max}}(a/b)$, условие применимости этого приближенного описания заведомо выполняется, если

$$\delta\omega_{\text{max}} \ll b/a\tau_{\text{рел}}. \quad (1.180)$$

Как было показано выше, при столкновении двух периодических фазовых волн с разными волновыми числами (и, следовательно, разными частотами) происходит их полное взаимное гашение. Узкий слой, в пределах которого осуществляется гашение, движется с течением времени в сторону области, занятой менее высокочастотной фазовой волной. Из этого результата следует, что если в распределенной активной среде есть два источника фазовых волн (какими являются пейсмейкеры), то наиболее высокочастотный из них полностью подавит с течением времени действие источника, имеющего меньшую частоту генерации. Таким образом, в среде происходит конкуренция между пейсмейкерами.

Если первоначально в среде имеется много различных пейсмейкеров со случайным набором частот, конкуренция между ними приведет к последовательному подавлению пейсмейкеров со все более высокими частотами, пока в среде не останется один-единственный пейсмейкер, для которого частота генерации максимальна [206].

12. Спиральные волны в автоколебательных активных средах

Спиральные волны наблюдаются не только в возбудимых, но и автоколебательных активных средах. Свойства этих волн во многом аналогичны

тем, что рассматривались нами в разделе 1.6. На больших удалениях от центра линии постоянной фазы имеют форму архимедовой спирали, частота вращения спиральной волны однозначно определяется характеристиками самой автоколебательной среды и не зависит от начальных условий, приведших к образованию такого автоволнового источника. С приближением к центру спиральной волны амплитуда колебаний монотонно уменьшается и обращается в нуль в центральной точке. Поэтому в автоколебательных средах также можно ввести понятие ядра спиральной волны как той области, где амплитуда колебаний существенно отличается от амплитуды однородных автоколебаний в данной среде. Однако в отличие от возбужденных сред размеры ядра в автоколебательной среде всегда малы и совпадают по порядку величины с диффузионной длиной $\sqrt{b\tau_{\text{рел}}}$.

Заметим, что, поскольку в пределах ядра амплитуда испытывает резкие изменения на длине порядка диффузионной, в этой области уравнение фазовой динамики (1.142) теряет применимость. Ввиду этого оказывается, что в рамках приближения фазовой динамики нельзя дать исчерпывающее описание спиральной волны и, в частности, рассчитать частоту ее вращения. Для этого необходимо обратиться к полным уравнениям в частных производных. Таким образом, свойства спиральной волны оказываются более тесно связанными с конкретным видом уравнений, характеризующих распределенную среду.

Ниже мы рассмотрим свойства спиральных волн в квазигармонических автоколебательных средах, описываемых λ - ω -моделью (1.144), частным случаем которой является обобщенное уравнение Гинзбурга–Ландау (1.154). Впервые такой расчет был произведен П. Хаганом в [169], ему мы и будем следовать в своем изложении.

Вначале ограничимся рассмотрением случая, когда $D_2 = 0$, т. е. мнимая добавка к эффективному коэффициенту диффузии отсутствует ($D_1 = D$). Напомним, что однородные автоколебания среды имеют амплитуду ρ_0 , определяемую условием $\lambda(\rho_0) = 0$, и осуществляются с частотой $\omega_0 = \omega(\rho_0)$. Диффузионная длина в данном случае равна

$$L_{\text{диф}} = \sqrt{b\tau_{\text{рел}}} = \sqrt{\frac{D}{|\rho_0 \lambda'(\rho_0)|}}. \quad (1.181)$$

Если ввести полярную систему координат (r, Θ) , спиральной волне, вращающейся с частотой ω , отвечает решение вида

$$\rho = \rho(r), \quad \varphi = \Theta - \chi(r) - (\omega_* - \omega_0)t. \quad (1.182)$$

Подстановка выражений (1.182) в уравнения (1.149) и (1.150) при $D_2 = 0$ и $D_1 = D$ дает

$$\rho_{rr} + \frac{1}{r}\rho_r + \left(\frac{1}{D}\lambda(\rho) + \frac{1}{r^2} - \chi_r^2\right)\rho = 0, \quad (1.183)$$

$$\chi_{rr} + \frac{1}{r}\chi_r + 2\left(\frac{\rho_r}{\rho}\right)\chi_r = \frac{\omega_* - \omega(\rho)}{D}. \quad (1.184)$$

Из требования, чтобы $\rho(r)$ и $\chi_r(r)$ оставались конечными при $r = 0$, следует

$$\rho(r) \sim r \quad \text{при} \quad r \rightarrow 0, \quad \chi_r(0) = 0. \quad (1.185)$$

Мы должны также потребовать, чтобы $\rho(r)$ асимптотически стремилось к какому-то отличному от нуля постоянному значению при $r \rightarrow \infty$. Действительно, амплитуда колебаний не может неограниченно возрастать с удалением от центра. С другой стороны, она не может и обращаться в нуль при $r \rightarrow \infty$, поскольку однородное состояние покоя в λ - ω -модели неустойчиво по отношению к малым возмущениям. Наконец, с помощью более тонких рассуждений можно показать, что неустойчивым будет также любое решение, для которого функция $\rho(r)$ является осциллирующей при больших r .

Пусть $\rho \rightarrow \rho_*$ при $r \rightarrow \infty$. Тогда из (1.183) и (1.184) следует, что

$$\omega_* = \omega(\rho_*), \quad \chi_r(r) \rightarrow \pm k_* \quad \text{при} \quad r \rightarrow \infty, \quad (1.186)$$

где $k_* = \sqrt{\lambda(\rho_*)}$.

Таким образом, на больших удалениях от центра спиральная волна (спираль) имеет постоянный шаг $h_{\text{св}} = 2\pi/|\chi_r| = 2\pi/k_*$, т. е. она является архимедовой. Шаг спирали равен пространственному периоду плоской фазовой волны с частотой ω_* . Два знака для $\chi_r(\infty)$ в (1.186) отвечают двум разным направлениям закручивания спирали.

Можно показать, что условия (1.185) и (1.186) являются совместными лишь при одном, вполне определенном выборе неизвестного параметра ω_* в уравнениях (1.183) и (1.184). Этим определяется частота вращения спиральной волны. Поскольку дифференциальные уравнения (1.183) и (1.184) не допускают полного аналитического решения, для нахождения величины ω_* , а также функций $\rho(r)$ и $\chi(r)$ приходится прибегать к приближенным методам.

Построим прежде всего решение в частном случае, когда $\omega(\rho) = \omega_0 = \text{const}$, т. е. отсутствует нелинейный сдвиг частоты. В этом случае взаимодействие между элементами среды не влияет на частоту совершаемых

ими колебаний. Очевидно, что частота спиральной волны также должна совпадать с ω_0 . Из (1.184) тогда следует, что при $\chi(r)(0) = 0$ должно выполняться равенство $\chi_r(r) = 0$ и, согласно (1.182), спираль вырождается в прямолинейный луч, вращающийся с угловой скоростью ω_0 .

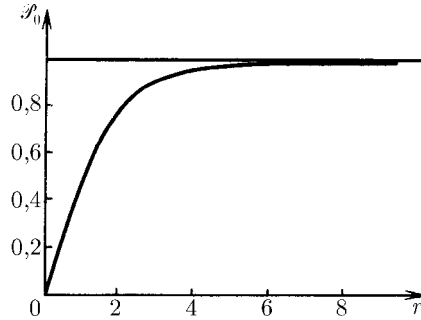


Рис. 1.50. Расчетная зависимость $\mathcal{P}_0(r)$ для среды, описываемой обобщенным уравнением Гинзбурга–Ландау (1.154)

В рассматриваемом случае амплитуда колебаний $\rho(r)$ подчиняется уравнению

$$\rho_{rr} + \frac{\rho_r}{r} - \frac{\rho}{r^2} + \frac{\rho\lambda(\rho)}{D} = 0, \quad (1.187)$$

которое должно быть дополнено граничными условиями $\rho(0) = 0$ и $\rho(r) \rightarrow \text{const}$ при $r \rightarrow \infty$. Заметим, что тогда $\rho(\infty) = \rho_0$, т.е. совпадает с амплитудой однородных колебаний. Если известен конкретный вид функции $\lambda(\rho)$, уравнение (1.187) с такими граничными условиями может быть численно проинтегрировано, в результате чего будет получена некоторая зависимость $\rho = \mathcal{P}_0(r)$. На рис. 1.50 показан полученный путем численного интегрирования вид функции $\mathcal{P}_0(r)$ для среды, описываемой обобщенным уравнением Гинзбурга–Ландау (1.154).

Рассмотрим теперь среду с нелинейным сдвигом частоты, в которой, однако, зависимость частоты от амплитуды колебаний является очень слабой:

$$\omega(\rho) = \bar{\omega} + \epsilon\delta\omega(\rho), \quad 0 < \epsilon \ll 1. \quad (1.188)$$

Введем малый параметр q , определив его как

$$q = \frac{\omega'(\rho_0)}{|\lambda'(\rho_0)|} = \frac{\epsilon\delta\omega'(\rho_0)}{|\lambda'(\rho_0)|}, \quad |q| \ll 1. \quad (1.189)$$

Решение для спиральной волны в такой среде может быть получено с помощью метода *сшиваемых асимптотических разложений*. Суть его в том,

чтобы независимо построить приближенные решения уравнений (1.183), (1.184) во *внутренней* ($r \ll L_{\text{диф}}e^{1/|q|}$) и *внешней* ($r \gg L_{\text{диф}}$) областях. Поскольку $|q| \ll 1$, эти две области перекрываются. Сшивка двух приближенных решений в интервале их перекрытия позволяет найти неизвестный параметр ω_* , входящий в уравнение (1.184), — частоту вращения спиральной волны.

Опуская детали расчета, которые можно найти в [169], приведем окончательные результаты.

Частота вращения спиральной волны ω_* и волновое число k_* равны

$$\omega_* = \omega_0 - qDk_*^2, \quad (1.190)$$

$$k_* = \frac{\zeta}{|q|L_{\text{диф}}} \exp\left[-\frac{\pi}{2|q|}\right]. \quad (1.191)$$

Числовой коэффициент ζ зависит от конкретного вида функций $\lambda(\rho)$ и $\omega(\rho)$; для обобщенного уравнения Гинзбурга–Ландау $\zeta = 0,509$.

Во внутренней области справедливо разложение по степеням q :

$$\begin{aligned} \rho(r) &= \mathcal{P}_0(r) + q^2\mathcal{P}_1(r) + \dots, \\ \chi_r(r) &= |q|v_0(r) + |q|^3v_1(r) + \dots \end{aligned} \quad (1.192)$$

В частности,

$$v_0(r) = -\frac{\lambda'(\rho_0)}{\delta\omega'(\rho_0)} \frac{1}{r\mathcal{P}_0^2(r)} \int_0^r y\mathcal{P}_0^2(y)[\delta\omega(\rho_0) - \delta\omega(\mathcal{P}_0(y))] dy, \quad (1.193)$$

причем при $L_{\text{диф}} \ll r \ll L_{\text{диф}}e^{1/|q|}$ справедливо соотношение

$$v_0(r) \left(\frac{L_{\text{диф}}}{r}\right) \left[\ln \frac{r}{L_{\text{диф}}} + C_1\right]. \quad (1.194)$$

Внешнюю область можно разделить на ближнюю ($L_{\text{диф}} \ll r \ll 1/k_*$) и дальнюю ($r \gg 1/k_*$). В ближней области имеем

$$\chi_r(r) \approx \left(\frac{L_{\text{диф}}}{r}\right) \text{tg}\left\{|q|\left[\ln \frac{r}{L_{\text{диф}}} + C_1\right]\right\}, \quad (1.195)$$

в дальней области —

$$\chi_r(r) \approx -\frac{k_*K'_0(|q|kr)}{K_0(|q|kr)}, \quad (1.196)$$

где $K_0(x)$ — модифицированная функция Бесселя с асимптотикой $K_0(x) \rightarrow 0$ при $x \rightarrow \infty$. В частности, при $qkr \ll 1$ функция χ_r дается разложением:

$$\chi_r(r) \approx k_* \left[1 + \frac{1}{2} \frac{1}{q|kr|} - \frac{1}{8} \frac{1}{(qkr)^2} + \dots \right]. \quad (1.197)$$

Во всей внешней области амплитуда колебаний $\rho(r)$ связана с $\chi_r(r)$ соотношением

$$\rho(r) = \rho_0 - \left(\frac{1}{r^2} + \chi_r^2 \right) L_{\text{диф}}^2. \quad (1.198)$$

При $r \rightarrow \infty$ амплитуда колебаний стремится к пределу

$$\rho_* = \rho_0 - (L_{\text{диф}} k_*)^2. \quad (1.199)$$

До сих пор здесь мы рассматривали только такие автоколебательные среды, для которых эффективный коэффициент диффузии D был действительным. Приведенные результаты нетрудно обобщить на случай, когда $D = D_1 + iD_2$. Все отличие заключается в том, что теперь в качестве малого параметра q необходимо выбрать величину

$$q = \frac{D_1 [\omega'(\rho_0)/|\lambda'(\rho_0)|] - D_2}{D_1 + [\omega'(\rho_0)/|\lambda'(\rho_0)|] D_2}, \quad (1.200)$$

а эффективная диффузионная длина равна

$$L_{\text{диф}} = \frac{1}{\sqrt{|\lambda'(\rho_0)\rho_0|}} \sqrt{D_1 + \frac{\omega'(\rho_0)}{|\lambda'(\rho_0)|} D_2}. \quad (1.201)$$

Заметим, что теперь уже для применимости метода сшиваемых асимптотических разложений недостаточно, чтобы нелинейный сдвиг частоты был мал. Как видно из (1.200), это еще не обеспечивает малости параметра q .

Интересно, что определенный формулой (1.200) параметр q легко выразить через коэффициенты a и b , входящие в уравнения фазовой динамики (1.142):

$$q = -\frac{a}{b}. \quad (1.202)$$

Учитывая также, что $L_{\text{диф}} = \sqrt{b\tau_{\text{рел}}}$, можно переписать (1.191) в форме

$$k_* = \zeta \sqrt{\left(\frac{b}{a^2 \tau_{\text{рел}}} \right) \exp \left[-\frac{\pi b}{2|a|} \right]}, \quad (1.203)$$

где ζ — некоторый безразмерный числовой коэффициент. Частота вращения спиральной волны при этом равна

$$\omega_* = \omega_0 + ak_*^2. \quad (1.204)$$

Следовательно, хотя частота вращения и зависит от конкретного вида уравнений автоколебательной распределенной среды, эта зависимость при $|a| \ll b$ является довольно слабой — от вида уравнений зависит лишь значение числового коэффициента ζ в (1.203). Благодаря этой универсальности можно ожидать, что выражения (1.203), (1.204), полученные выше лишь для случая квазигармонических автоколебаний, будут выполняться для произвольной автоколебательной среды в пределе $|a| \ll b$.

При увеличении параметра $|q|$ волновое число k_* генерируемых волн возрастает и становится порядка $L_{\text{диф}}^{-1}$. В этом режиме значения частоты вращения спиральной волны ω_* и волнового числа k_* существенно зависят от конкретного вида уравнений автоколебательной среды, а использованная выше процедура сшивки асимптотических разложений неприменима. Возможен только приближенный расчет [169] значений ω_* и k_* путем численного интегрирования уравнений (1.183), (1.184) с граничными условиями (1.185), (1.186).

Необходимо заметить, что при достаточно больших значениях q спиральные волны должны терять устойчивость. Действительно, на большом удалении от центра возникающая волновая картина имеет локально вид плоской фазовой волны с волновым числом k_* . С ростом параметра q волновое число k_* возрастает. Однако фазовые волны с достаточно большими волновыми числами неустойчивы; обычно это происходит при $k \sim 1/L_{\text{диф}}$. Следовательно, спиральные волны могут наблюдаться лишь в таких автоколебательных средах, где параметр $q = -a/b$ не слишком велик по модулю ($|q| \lesssim 1$). Например, для уравнения Гинзбурга–Ландау устойчивость фазовых волн теряется при выполнении неравенства

$$(kL_{\text{диф}}) > \frac{1}{\sqrt{3 + 2q^2}}. \quad (1.205)$$

Поэтому спиральные волны в среде, описываемой таким уравнением, существуют лишь при $q < 1,397 \dots$.

В [169] построены также решения для вращающихся спиральных волн, имеющих два или более рукава. Проведенный в [169] анализ показал, однако, что все такие режимы неустойчивы. Двухрукавная спиральная волна распадается на две спиральные волны с одним рукавом, и т. д.

Взаимодействие между спиральными волнами в автоколебательной распределенной среде рассматривали Я. Б. Зельдович и Б. А. Маломед [60]. Для автоколебательной среды, описываемой уравнением Гинзбурга–Ландау, при $D_2 = 0$ в пренебрежении нелинейным сдвигом частоты они обнаружили, что две спиральные волны отталкиваются друг от друга, если они вращаются в одинаковом направлении, и притягиваются, если их направления вращения противоположны.

Полученный результат позволяет доказать неустойчивость спиральной волны для ограниченной области, имеющей в сечении форму квадрата. В самом деле, эту квадратную область можно продолжить до бесконечной решетки; при этом решения в двух соседних областях получают одно из другого преобразованием зеркальной симметрии, т. е. направления вращения спиральных волн в соседних ячейках противоположны, вследствие чего вся эта система стремится аннигилировать. Аналогичное рассуждение доказывает неустойчивость отдельной спиральной волны в области, имеющей в сечении форму правильного треугольника. Однако для правильного шестиугольника оно, как легко понять из элементарного рассмотрения, уже не подходит. Это означает, что спиральные волны могут образовывать на плоскости устойчивую гексагональную решетку.

Хотя результаты Зельдовича и Маломеда, относящиеся к взаимодействию спиральных волн, были получены только для сред без нелинейного сдвига, где $q = 0$, можно ожидать, что они сохраняются также и для среды со слабым нелинейным сдвигом частоты ($q \ll 1$), если расстояние между центрами двух спиральных волн не слишком велико, а именно, если оно не превышает длины генерируемых фазовых волн $\lambda_* = 2\pi/k_*$. На больших удалениях взаимодействие между спиральными волнами должно быстро спадать.

13. Стационарные диссипативные структуры

Помимо рассмотренных автоволновых режимов, в распределенных активных средах наблюдаются также разнообразные стационарные диссипативные структуры, устойчивые по отношению к малым возмущениям.

Собственно говоря, с одним из примеров таких структур мы уже столкнулись в разделе 1.3 при изучении простых бистабильных сред, описываемых всего одной переменной. Как отмечалось в §3, при выполнении условия (1.41) межфазная граница, отделяющая области с двумя различными стационарными состояниями («фазами»), покоится. В результате среда разбивается на смесь доменов двух фаз.

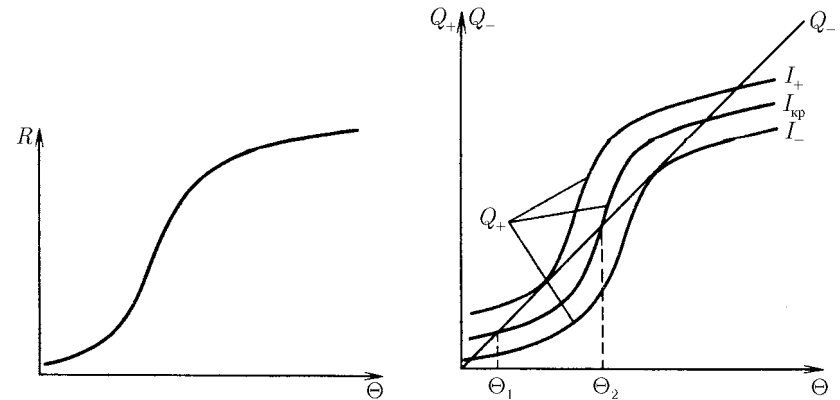


Рис. 1.51. Зависимость удельного сопротивления железной нити в атмосфере водорода от температуры нити

Рис. 1.52. Зависимость скоростей тепловыделения Q_+ и теплоотвода Q_- от температуры

Однако этот пример не вполне типичен. При выполнении условия равновесия фаз (1.41) размеры доменов произвольны, а следовательно формируемая ими пространственная структура не восстанавливается после возмущения, т. е. такая система находится в состоянии безразличного (или нейтрального) равновесия. Кроме того, само условие (1.41) выполняется лишь при определенном выборе параметров, характеризующих данную бистабильную среду. В общем случае граница между фазами не покоится, а смещается с течением времени. Направление и скорость ее движения зависят от значений параметров среды. Меняя эти параметры, можно управлять движением межфазных границ.

Устойчивую диссипативную структуру легче всего получить, предположив, что значения параметров не неизменны, а быстро подстраиваются к мгновенному положению межфазных границ. Иллюстрацией этого эффекта может служить теория барретирования, разработанная В. В. Барелко, Я. Б. Зельдовичем и др. [16].

Рассмотрим нагреваемую электрическим током железную нить, которая находится в атмосфере водорода. Экспериментально установлено, что удельное сопротивление R такой нити в расчете на единицу длины оказывается нелинейно зависящим от температуры. Как видно из рис. 1.51, функция $R = R(\Theta)$ довольно резко возрастает с увеличением температуры. Помимо выделения количества теплоты (тепловыделения) $Q_+ = I^2 R(\Theta)$,

где I — сила тока, идет также процесс теплообмена с окружающей средой, приводящий к потере количества теплоты (теплоотвода) $Q_- = \gamma(\Theta - \Theta_0)$. В этих условиях уравнение теплового баланса нити можно записать в виде

$$C \frac{\partial \Theta}{\partial t} = f(\Theta) + \chi \frac{\partial^2 \Theta}{\partial x^2}, \quad (1.206)$$

где $f(\Theta) = Q_+ - Q_-$, C — теплоемкость нити в расчете на единицу ее длины, χ — теплопроводность.

Положим вначале, что сила тока I поддерживается постоянной. Тогда уравнение (1.206) совпадает со стандартным уравнением простой бистабильной среды (1.13). В интервале $I_+ > I > I_-$ система обладает двумя устойчивыми однородными состояниями (рис. 1.52), одно из которых характеризуется высокой ($\Theta = \Theta_3$), а второй — низкой ($\Theta = \Theta_1$) температурой нити. Эти два состояния можно рассматривать в качестве двух «фаз». В распределенной системе, описываемой уравнением (1.206), переход между стационарными состояниями может осуществляться путем распространения температурной волны, вытесняющей (в зависимости от значений параметров) высокотемпературный либо низкотемпературный режим. Как было показано в разделе 1.2, направление смещения границы между двумя областями определяется знаком интеграла

$$A = \int_{\Theta_1}^{\Theta_3} f(\Theta) d\Theta. \quad (1.207)$$

Если $A > 0$, волна движется в сторону заполнения нити высокотемпературным режимом $\Theta = \Theta_3$, тогда как при $A < 0$ низкотемпературный режим вытесняет высокотемпературный. Равенство $A = 0$ является критическим условием смены направления фронта, при котором скорость распространения волны обращается в нуль. Это условие отвечает стационарному состоянию систем, при котором на нити могут устойчиво и стационарно существовать зоны высокотемпературного и низкотемпературного режимов.

До сих пор рассматривался случай, когда нить питалась от регулируемого источника тока, поддерживающего силу тока I постоянной. Если же питание производится от регулируемого источника напряжения, т. е. поддерживается постоянной не сила тока I , а разность электрических потенциалов U на концах нити, то процесс характеризуется новыми качественными особенностями. В этом случае сила тока I является уже не постоянной, а переменной величиной, поскольку она, согласно закону Ома, зависит от пол-

ного сопротивления нити:

$$I = U \left[\int_{-L_0}^{L_0} R(\Theta(x)) dx \right]^{-1}, \quad (1.208)$$

где $2L_0$ — длина нити.

Пусть, например, сила тока I , текущего через нить, находящуюся в состоянии с температурой Θ_1 , превысила критическое значение $I_{кр}$, для которого $A = 0$. В этом случае при внесении в систему локального возмущения вдоль нити начинает распространяться волна и нить переходит в состояние с температурой Θ_3 . По мере расширения высокотемпературной зоны полное сопротивление нити возрастает, что приводит к снижению силы тока I и, следовательно, обуславливает монотонное падение A . Когда значение I упадет до $I_{кр}$, фронт прекратит свое движение. Таким образом, нить окажется разбитой на участок с высокой температурой Θ_3 и большим сопротивлением $R(\Theta_3)$ и участок с низкой температурой и малым сопротивлением $R(\Theta_1)$. При увеличении напряжения U будут меняться длины участков и, следовательно, полное сопротивление нити, но не изменится сила тока $I_{кр}$, при которой горячий и холодный участки сосуществуют. Поэтому, пока высокотемпературная зона не заполнит всю нить, система работает как стабилизатор тока, поддерживая его значение, равное $I_{кр}$. При уменьшении U осуществляется аналогичный процесс. Этот эффект лежит в основе действия технического устройства, называемого *бареттором*.

Пусть ширина высокотемпературного участка $2l_3$ гораздо больше толщины переходного слоя l_0 . Тогда силу тока, протекающего через нить, можно приближенно выразить как

$$I = \frac{U}{2(R_3 l_3 + R_1 l_1)}, \quad l_1 = L_0 - l_3, \quad (1.209)$$

где $R_3 = R(\Theta_3)$ и $R_1 = R(\Theta_1)$. Учитывая, что в стационарном режиме устанавливается критическое значение силы тока $I_{кр}$, находим из (1.209)

$$l_3 = \frac{1}{R_3 - R_1} \left(\frac{U}{2I_{кр}} - R_1 L_0 \right). \quad (1.210)$$

Таким образом, стационарная диссипативная структура — разогретая зона — возникает при $U > U_{\min}$, где $U_{\min} = 2L_0 R_1 I_{кр}$. С увеличением U ширина этой зоны растет, а при $U = U_{\max}$, где $U_{\max} = 2L_0 R_3 I_{кр}$, эта зона заполняет собой всю нить. При $U > U_{\max}$ устойчиво только однородное высокотемпературное состояние.

До сих пор мы предполагали, что на нити располагается всего один высокотемпературный участок. Ясно, однако, что все приведенные выше рассуждения сохраняют силу и в том случае, если горячих участков несколько, — необходимо лишь, чтобы их суммарная длина совпадала со значением l_3 , задаваемым выражением (1.210). Сколько будет наблюдаться высокотемпературных участков и где они будут расположены — зависит от начального условия, т. е. от того, в каких именно областях был произведен разогрев нити.

Несмотря на это, казалось бы, частный характер, рассмотренный нами пример иллюстрирует основные свойства стационарных диссипативных структур. Прежде всего, каждая такая структура — это устойчивое образование, форма и размеры которого восстанавливаются после малых возмущений. Можно сказать, что каждая стационарная диссипативная структура представляет собой одно из устойчивых стационарных состояний распределенной системы. Каждое устойчивое стационарное состояние характеризуется своей *областью притяжения* (все начальные распределения, попадающие в область притяжения стационарного состояния, релаксируют к нему со временем). Одна и та же распределенная система может обладать огромным числом различных стационарных структур. Какое именно из этих стационарных состояний осуществляется в эксперименте, зависит от начальных условий.

Главная предпосылка для возникновения стационарных диссипативных структур заключается в том, чтобы, помимо локального взаимодействия, элементы активной среды были бы охвачены также действующей обратной связью. В рассмотренном нами примере обратная связь осуществляется через изменение силы тока, которое определяется состоянием (температурой) всех элементов распределенной системы. Возможны, конечно, и другие механизмы, приводящие к аналогичному эффекту.

Пусть, например, в среде происходит горение некоторого вещества. Предположим, что коэффициент диффузии этого вещества превышает теплопроводность, так что за счет диффузии концентрация горючего вещества a во всех элементах среды поддерживается практически постоянной. Будем также считать, что горючее вещество непрерывно поступает в среду с постоянной скоростью Q . Уравнения, описывающие процесс горения в такой среде, имеют вид

$$\frac{\partial \Theta}{\partial t} = aq(\Theta) - \gamma(\Theta - \Theta_0) + \chi \Delta \Theta, \quad \frac{da}{dt} = Q - a \int \tilde{q}(\Theta(r)) dr. \quad (1.211)$$

Здесь $q(\Theta)$ характеризует выделение теплоты в расчете на единицу массы

горючего вещества, а $\tilde{q}(\Theta)$ — скорость уменьшения этого вещества при горении. Поскольку процесс горения начинается лишь при достижении определенной температуры, обе эти функции имеют вид размытой ступеньки (рис. 1.8).

Уравнения (1.211) описывают образование стационарной диссипативной структуры. Действительно, предположим, что мы подожгли горючую смесь в некоторой малой области среды. Вначале зона горения будет расширяться. Но расширение зоны горения приведет к увеличению расхода горючего вещества, а следовательно, к понижению его концентрации a . В свою очередь, понижение a вызовет уменьшение скорости движения фронта горения. При достижении некоторого критического значения $a_{кр}$ фронт остановится. В результате в среде возникнет локализованный очаг горения — стационарная диссипативная структура.⁹ Заметим, что, как и в обсуждавшемся выше примере, в данном случае наблюдается *множественность* диссипативных структур. Вместо одного большого очага горения в среде может иметься несколько более мелких очагов с тем же суммарным расходом горючего. Расположение и число таких очагов определяются начальными условиями. Если очаги горения расположены не слишком далеко друг от друга, по сравнению с шириной фронта волны горения l_0 , между ними должно наблюдаться своеобразное взаимодействие: два расположенных поблизости очага притягиваются и стремятся слиться друг с другом.

Дальнейшая обратная связь может осуществляться также за счет хорошо диффундирующего ингибитора. В этом случае рост очага горения прекращается благодаря увеличению в среде концентрации ингибитора.

Приведенные простые примеры иллюстрируют главные механизмы, ведущие к образованию стационарных диссипативных структур. К настоящему времени имеется хорошо развитая математическая теория [72]–[83] стационарных диссипативных структур в реагирующих двухкомпонентных системах с диффузией, которые описываются уравнениями вида

$$\tau_u \frac{\partial u}{\partial t} = f(u, v) + l^2 \Delta u, \quad \tau_v \frac{\partial v}{\partial t} = g(u, v) + L^2 \Delta v. \quad (1.212)$$

Переменную u можно интерпретировать как концентрацию *активатора*, а переменную v — как концентрацию *ингибитора*. Нас сейчас интересуют

⁹Эти рассуждения предполагают, что горючее расходуется быстро по сравнению с характерными скоростями изменения температуры, а поэтому можно считать, что его концентрация a адиабатически подстраивается к мгновенному распределению температуры. В противном случае такая подстройка будет происходить с некоторым запаздыванием. Если запаздывание достаточно велико, оно может стать причиной незатухающих пульсаций зоны горения.

ситуации, когда коэффициент диффузии ингибитора значительно превышает коэффициент диффузии активатора ($L \gg l$), а изменения концентрации ингибитора происходят быстрее, чем изменения концентрации активатора ($\tau_v \ll \tau_u$). При этих условиях распределение ингибитора быстро подстраивается к мгновенному распределению активатора, оставаясь всегда достаточно плавным.

Свойства устанавливающихся в среде диссипативных структур существенно зависят от того, составлена ли данная активная среда из бистабильных, возбудимых или автоколебательных элементов.

1. Если среда составлена из бистабильных элементов, она разбивается на совокупность областей с размерами не меньше L , занятых двумя фазами, в которых состояние среды близко к одному из двух устойчивых состояний отдельно взятого элемента. Эти области отделены переходными слоями с шириной порядка l . Они могут иметь большие размеры: весь объем может, например, разбиться всего на две такие области с разными фазами. Кроме того, могут существовать также капли противоположной фазы с размером порядка l , вкрапленные в области постоянной фазы. Расстояние между каплями больше или порядка L .

2. Если среда составлена из возбудимых элементов, т. е. имеется всего одно однородное стационарное состояние покоя среды $u = u_0, v = v_0$ и оно устойчиво по отношению к малым возмущениям, то в одномерном случае стационарная диссипативная структура обычно представляет собой набор узких страт с шириной порядка l , расположенных на расстоянии порядка или более L друг от друга. В областях между стратами среда находится в состоянии, близком к состоянию покоя. Внутри страт резко меняется лишь концентрация активатора, тогда как распределение ингибитора является всюду плавным. Возможны локализованные диссипативные структуры, когда в длинном по сравнению с L отрезке расположены одна или несколько узких страт, разделенных промежутками порядка L , а в остальной области среда остается в состоянии покоя. В двумерных и трехмерных случаях диссипативная структура представляет собой совокупность малых капель с размером порядка l , разделенных расстояниями порядка или более L . В областях между каплями состояние среды близко к устойчивому одномерному состоянию. Исходя из однородного стационарного состояния, диссипативные структуры в таких средах можно создать лишь путем достаточно сильных начальных возмущений.

3. Если отдельно взятый элемент среды совершает периодические автоколебания, стационарное однородное состояние активной среды абсолютно неустойчиво. В ней возникают стационарные диссипативные структуры,

имеющие вид страт с длиной порядка L , отделенных переходными слоями с шириной порядка l . В двумерном и трехмерном случаях среда распадается на совокупность ячеек с размером порядка L , отделенных узкими переходными слоями. Локализованные стационарные структуры невозможны: ячейки заполняют весь объем среды.

Подчеркнем, что двухкомпонентные системы типа (1.212) допускают установление стационарных неоднородных структур при гораздо менее жестких ограничениях, чем однокомпонентные системы вида (1.13), где существование двух фаз возможно лишь при выполнении условия (1.41). Модели такого типа широко используются при изучении явления *морфогенеза*, т. е. позиционной дифференциации отдельных клеток в процессе индивидуального развития многоклеточного живого организма. Предполагается, что такая дифференциация обусловлена взаимодействием между клетками, осуществляемым посредством веществ двух типов — активаторов и ингибиторов, причем скорость диффузии ингибитора превышает скорость диффузии активатора. Устанавливающиеся стационарные неоднородные распределения этих двух веществ определяют структуру многоклеточного организма. Подробнее с этими вопросами можно ознакомиться в [20, 105, 127].

Заметим, наконец, что если условие $\tau_v \ll \tau_u$ нарушено и ингибитор обладает значительной инерционностью, картина явления может существенно измениться. Поскольку подстройка распределения ингибитора происходит теперь с некоторым запаздыванием, может возникнуть [77] режим осциллирующих диссипативных структур, параметры которых (например, ширина страт) периодически меняются со временем.

На этом мы заканчиваем рассмотрение явлений самоорганизации, связанных с образованием когерентных крупномасштабных структур в сильнонераспределенных средах. Разумеется, здесь у нас не было возможности подробнее остановиться на особенностях подобных структур в конкретных активных средах различного происхождения. Такой анализ должен быть предметом отдельного рассмотрения.

ГЛАВА 2

ДИНАМИЧЕСКИЙ ХАОС

Самоорганизация тесно связана с зарождением турбулентности. С изменением потока энергии или иных характеристик активной среды в ней происходит качественная перестройка имеющихся диссипативных или автоволновых структур. В ходе последовательного усложнения регулярных структур может быть достигнута граница, за которой их поведение становится неустойчивым по отношению к малым возмущениям. Такая неустойчивость служит отличительным признаком турбулентного режима. Хаотическая динамика структур при турбулентном режиме не является результатом действия каких-либо внешних случайных сил. Эффекты динамического хаоса представляют собой внутренне присущее свойство подобных систем. Они наблюдаются не только в распределенных активных средах, но и в более простых сосредоточенных системах, динамика которых описывается набором из нескольких обыкновенных дифференциальных уравнений. Аналогичные эффекты имеют место также и в гамильтоновых консервативных системах, где в отличие от активных сред нет ни диссипации, ни притока энергии.

14. Гамильтоновы системы

В основу гамильтонова подхода к описанию динамики физических систем положены уравнения Гамильтона, представляющие собой систему обыкновенных дифференциальных уравнений

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad (2.1)$$

где $i = 1, 2, \dots, n$ и при $t = t_0$ выполняются начальные условия $q_i(t_0) = q_i^0$, $p_i(t_0) = p_i^0$. Переменные q_i и p_i называются, соответственно,

обобщенными координатами и обобщенными импульсами, а их совокупность $(q_i, p_i) \equiv (\mathbf{q}, \mathbf{p})$ — каноническими переменными. Число n независимых обобщенных координат называется числом степеней свободы, а функция $H = H(\mathbf{q}, \mathbf{p})$, характеризующая физическое состояние системы, называется функцией Гамильтона, или гамильтонианом. Мы будем предполагать, что функция $H(\mathbf{q}, \mathbf{p})$ не зависит явно от времени и действующие силы потенциальны. В этом случае гамильтониан является полной энергией системы, выраженной через канонические переменные.

Решение уравнений (2.1) позволяет определить зависимость $\mathbf{q}(t)$ и $\mathbf{p}(t)$ по их начальным значениям \mathbf{q}_0 и \mathbf{p}_0 . Этот эволюционный процесс можно представить как движение точки в $2n$ -мерном пространстве с координатами \mathbf{q} и \mathbf{p} . Такое объединенное пространство (\mathbf{q}, \mathbf{p}) называется фазовым пространством системы, а точки этого пространства — фазовыми точками. Каждому мгновенному состоянию гамильтоновой системы (2.1) отвечает фазовая точка пространства (\mathbf{q}, \mathbf{p}) , и каждой точке этого пространства соответствует определенное и единственное состояние системы. Следовательно, динамику системы (2.1) можно представить как последовательное изменение положения фазовых точек, т.е. как движение этих точек, в фазовом пространстве. Траектории движения фазовых точек называются фазовыми траекториями, или фазовыми кривыми.

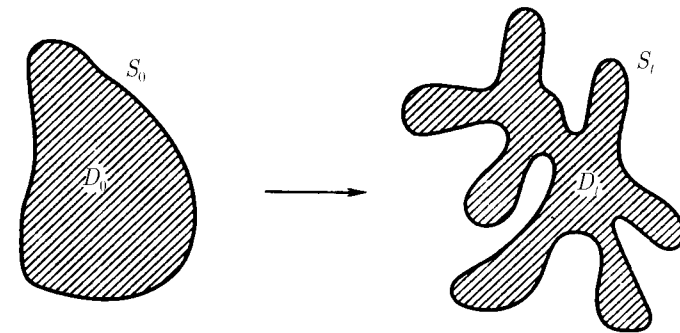


Рис. 2.1. Сохранение фазового объема при эволюции гамильтоновой системы

Фазовое пространство гамильтоновых систем обладает тремя важными свойствами. **1.** В любой момент времени траектории, определяемые уравнениями (2.1), в фазовом пространстве не пересекаются. Это свойство следует из теоремы существования и единственности решений обыкновенных дифференциальных уравнений. **2.** Сохраняется объем произвольной обла-

сти фазового пространства (*теорема Лиувилля*). Суть последнего утверждения состоит в том, что если в начальный момент времени t_0 фазовые точки непрерывно заполняли некоторую конечную область D_0 фазового пространства (\mathbf{q}, \mathbf{p}) и с течением времени перешли в другую область D_t этого пространства (рис. 2.1), то соответствующие фазовые объемы равны между собой:

$$\int_{D_0} d\mathbf{q}_0 d\mathbf{p}_0 = \int_{D_t} d\mathbf{q} d\mathbf{p}.$$

Следовательно, по аналогии с гидродинамикой, движение точек, изображающих состояние системы в пространстве (\mathbf{q}, \mathbf{p}) , можно представить как движение несжимаемого потока жидкости. Поэтому часто совокупность фазовых кривых в пространстве (\mathbf{q}, \mathbf{p}) называют *потоком*. **3.** Непрерывная граница S_0 , охватывающая в момент времени t_0 некоторую область D_0 фазового пространства, с течением времени трансформируется (не разрываясь) в непрерывную границу S_t , охватывающую ту область D_t , в которую D_0 эволюционировала к моменту t (рис. 2.1). Системы, удовлетворяющие свойству **2**, часто называют *консервативными*.

Описание динамики физических систем составляет основную задачу механики. Существует ряд достаточно общих способов, позволяющих в некоторых случаях построить решение уравнений Гамильтона. Однако это возможно далеко не всегда. В большинстве случаев уравнения (2.1) оказываются неинтегрируемыми в элементарных функциях. Тем не менее, часто удается получить достаточно сведений относительно поведения системы, не интегрируя полностью ее уравнений, а отыскивая *интегралы движения*, т. е. функции, которые остаются постоянными при движении: $f(\mathbf{q}, \mathbf{p}) = \text{const}$. Для любой системы, в которой функция $H(\mathbf{q}, \mathbf{p})$ не зависит явно от времени, гамильтониан является интегралом движения: для данных начальных условий постоянное значение гамильтониана есть просто полная энергия системы $E = H(\mathbf{q}, \mathbf{p})$. Отсюда находим $q_1 = g(q_2, \dots, q_n, \mathbf{p}, E)$, где g — некоторая функция.

Следовательно, задавая E , мы определяем в фазовом пространстве (\mathbf{q}, \mathbf{p}) некоторую гиперповерхность (или пространство уровня энергии) размерности $2n - 1$, и все траектории системы, отвечающие данному значению энергии, должны располагаться на такой энергетической гиперповерхности. Если в дополнение к гамильтониану имеются и другие интегралы движения, то все фазовые кривые лежат на поверхности, размерность которой меньше $2n - 1$.

Если фазовые траектории системы (2.1) при данном значении энергии E не уходят на бесконечность, то говорят, что движение является *финитным*. Ниже рассматриваются только финитные движения.

Уравнения Гамильтона обладают тем важным свойством, что допускают широкий класс преобразований канонических переменных $(\mathbf{q}, \mathbf{p}) \rightarrow (\mathbf{q}', \mathbf{p}')$, называемый каноническими преобразованиями, при которых не изменяется общая форма уравнений для любой гамильтоновой системы, т. е. $\dot{q}'_j = \partial H' / \partial p'_j$, $\dot{p}'_j = -\partial H' / \partial q'_j$, $j = 1, 2, \dots, n$. Здесь $H' = H'(\mathbf{q}', \mathbf{p}')$ — новая функция Гамильтона, а $\mathbf{q}' = \mathbf{q}'(\mathbf{q}, \mathbf{p})$, $\mathbf{p}' = \mathbf{p}'(\mathbf{q}, \mathbf{p})$ — новые канонические переменные. Сами по себе эти преобразования не изменяют свойств и динамики происходящих в системе процессов, однако могут оказаться полезными при построении решений уравнений Гамильтона и анализе физической картины движения.

14.1 Переменные действие–угол

Одним из важных и часто используемых канонических преобразований (которые не всегда существуют!) является преобразование $(\mathbf{q}, \mathbf{p}) \rightarrow (\boldsymbol{\alpha}, \mathbf{J})$, при котором в новых переменных $(\boldsymbol{\alpha}, \mathbf{J})$ функция Гамильтона H' не зависит от координат $\boldsymbol{\alpha}$: $H' = H'(\mathbf{J})$. В этом случае переменные \mathbf{J} называют переменными *действия*, а соответствующие сопряженные переменные $\boldsymbol{\alpha}$ — переменными типа *угол*. Тогда система уравнений (2.1), которая запишется в виде

$$\dot{\alpha}_i = \frac{\partial H'}{\partial J_i}, \quad \dot{J}_i = 0, \quad j = 1, 2, \dots, n, \quad (2.2)$$

легко интегрируется: $\mathbf{q} = \mathbf{q}(\boldsymbol{\alpha}, \mathbf{J})$, $\mathbf{p} = \mathbf{p}(\boldsymbol{\alpha}, \mathbf{J})$, $\mathbf{J} = \mathbf{J}_0 = \text{const}$, $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{\omega}t$, $\omega_i = \partial H' / \partial J_i$. Следовательно, задача нахождения решений канонических уравнений (2.1) сводится к отысканию канонического преобразования, при котором возможно перейти к переменным действие–угол.

Гамильтонова система называется *полностью интегрируемой* (а соответствующий гамильтониан *интегрируемым*), если существует такое каноническое преобразование, с помощью которого можно перейти к переменным действие–угол: $(\mathbf{q}, \mathbf{p}) \rightarrow (\boldsymbol{\alpha}, \mathbf{J})$.

Рассмотрим системы с одной степенью свободы, $n = 1$. Для таких систем фазовое пространство двумерно, а преобразование вида $(q, p) \rightarrow (\alpha, J)$ осуществимо всегда. Поскольку гамильтониан $H(q, p) = E$ является интегралом движения, то, разрешив это уравнение относительно импульса, найдем $p = p(q, E)$. Из первого уравнения Гамильтона системы (2.1) имеем: $t = \int_{q_0}^q (\partial H / \partial p)^{-1} dq$. Используя выражение для импульса, таким образом приходим к выводу, что решение гамильтоновой системы с одной

степенью свободы сводится к квадратуре — вычислению интеграла от известной функции.

Обратимся к простому примеру консервативной системы с одной степенью свободы — математической модели маятника. Характерная особенность этого примера заключается в том, что подобный гамильтониан появляется во многих системах, и поэтому ряд понятий, возникающих здесь, будет использован в дальнейшем.

Гамильтониан маятника длины l и массы m равен сумме потенциальной $\Pi = -mgl \cos \varphi$ и кинетической $K = p^2/2ml^2$ энергий:

$$H = \frac{p^2}{2ml^2} - mgl \cos \varphi, \quad (2.3)$$

где φ — угол отклонения от вертикали, а g — ускорение свободного падения. Уравнения движения маятника имеют вид $\dot{p} = -mgl \sin \varphi$, $\dot{\varphi} = p/ml^2$, или, в более привычной форме,

$$\ddot{\varphi} + \omega_0^2 \sin \varphi = 0, \quad (2.4)$$

где $\omega_0 = \sqrt{g/l}$ — частота колебаний.

Когда полная энергия $H = E$ маятника превышает наибольшее значение потенциальной энергии, $E = E_{rot} > mgl$, импульс p всегда будет отличен от нуля, что приводит к неограниченному росту угла φ . Это означает, что маятник будет вращаться. На фазовой плоскости (рис. 2.2) такое поведение изображается траекториями E_{rot} , отвечающими движению фазовой точки слева направо для $p > 0$ и справа налево для $p < 0$. Колебания маятника соответствует энергия $E = E_{osc} < mgl$. Если же $E \simeq E_s \equiv mgl$, то период колебаний стремится к бесконечности и движение происходит по *сепаратрисе* — линии, разделяющей два качественно различных типа движения: колебания ($E = E_{osc}$) и вращение ($E = E_{rot}$).

Легко видеть, что в окрестности точек с координатами $(p, \varphi) = (0, 2\pi k)$, $k = 0, \pm 1, \pm 2, \dots$, семейство фазовых кривых имеет вид эллипсов. Поэтому такие точки называются *эллиптическими* точками системы. Семейство траекторий вблизи точек $(p, \varphi) = (0, \pi + 2\pi k)$, $k = 0, \pm 1, \pm 2, \dots$, имеет вид гипербол, и такие точки называются *гиперболическими*. Эллиптические точки являются устойчивыми и соответствуют нижнему положению равновесия маятника, а гиперболические точки, соответствующие верхнему положению равновесия маятника, являются неустойчивыми. Фазовая кривая, начавшаяся в окрестности гиперболической точки, удаляется от нее, в то время как траектория вблизи эллиптической точки всегда остается в ее окрестности.

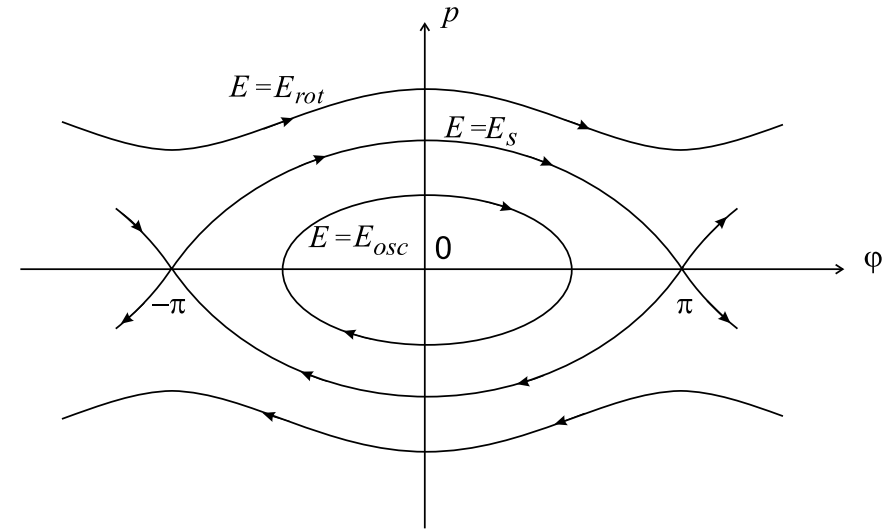


Рис. 2.2. Фазовый портрет маятника с гамильтонианом (2.3)

Нетрудно найти уравнение сепаратрисы маятника. Из условия $E = mgl$ и (2.3) имеем $p_s = \pm 2\omega_0 ml^2 \cos \varphi_s / 2$, где знак \pm соответствует различным ветвям сепаратрисы (см. рис. 2.2). Используя уравнение Гамильтона, связывающее φ и p , после его интегрирования и несложных преобразований находим: $\varphi_s = 4 \arctg e^{\omega_0 t} - \pi$.

Теперь, применяя описанную выше процедуру интегрирования систем с одной степенью свободы, можно найти период колебаний: $T = l\sqrt{m/2} \int d\varphi (E + mgl \cos \varphi)^{-1/2}$. В частности, на сепаратрисе, при $\varphi \rightarrow \pi + 2\pi k$, $k = 0, \pm 1, \pm 2, \dots$, период $T \rightarrow \infty$.

К введенным понятиям и аналогии с маятником и гамильтонианом (2.3) мы обратимся, когда будем изучать явление нелинейного резонанса.

Вернемся к рассмотрению систем с одной степенью свободы общего вида. Когда такая система исследуется в переменных действие–угол, то для описания динамики в пространстве (α, J) очень удобной оказывается полярная система координат, в которой величине J ставится в соответствие расстояние от начала координат до некоторой точки фазового пространства, а величине α — угол (рис. 2.3). Тогда, как легко видеть, движение при заданном J будет происходить по кривой, представляющей собой окружность с радиусом J . Меняя радиус, получим множество вложенных центрирован-

ных окружностей, которые в совокупности заполняют фазовое пространство такой консервативной системы. В конечном счете двумерное фазовое пространство оказывается разбитым окружностями на совокупность колец. Скорость движения по индивидуальной окружности, как следует из уравнений (2.2), в общем случае зависит от значения переменной действия J , т. е. $\omega = \omega(J)$.

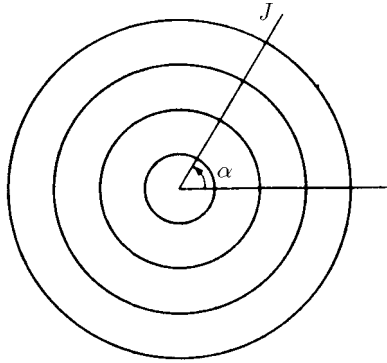


Рис. 2.3. Фазовое пространство гамильтоновой системы с одной степенью свободы в переменных действие – угол

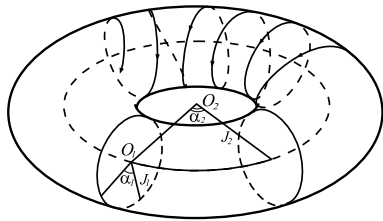


Рис. 2.4. Тороидальная поверхность, представляющая фазовое пространство системы двух гармонических осцилляторов

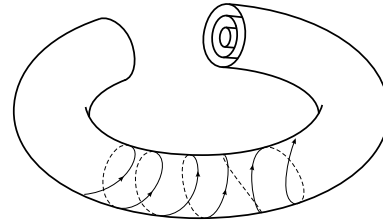


Рис. 2.5. Фазовый портрет интегрируемой системы с двумя степенями свободы

Для систем с двумя степенями свободы фазовое пространство четырехмерно. При переходе к переменным действие – угол (если каноническое преобразование $(\mathbf{q}, \mathbf{p}) \rightarrow (\boldsymbol{\alpha}, \mathbf{J})$ возможно) динамика таких систем наглядно может быть представлена следующим образом: движение по окружности с центром O_1 , образованной переменными α_1, J_1 , и одновременно

по окружности с центром O_2 (лежащей в плоскости, перпендикулярной плоскости окружности O_1), образованной переменными α_2, J_2 (рис. 2.4). Суперпозиция таких вращений задает движение по тору (т. е. поверхности «бублика»), размерность которого равна двум. Траектория, располагающаяся на торе, называется *обмоткой* тора.

Чтобы нагляднее представить себе сказанное, рассмотрим в качестве примера систему двух гармонических осцилляторов единичной массы. Ее динамика описывается в нормальных координатах гамильтонианом вида

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(\omega_1^2 q_1^2 + \omega_2^2 q_2^2). \quad (2.5)$$

Фазовое пространство такой системы четырехмерно (q_1, q_2, p_1, p_2) , и уравнения Гамильтона запишутся как

$$\dot{q}_i = p_i, \quad \dot{p}_i = -\omega_i^2 q_i, \quad i = 1, 2. \quad (2.6)$$

Каноническое преобразование, позволяющее перейти к переменным действие – угол, имеет вид

$$p_i = \sqrt{2\omega_i J_i} \cos \alpha_i, \quad q_i = -\sqrt{\frac{2J_i}{\omega_i}} \sin \alpha_i, \quad i = 1, 2. \quad (2.7)$$

Тогда гамильтониан и решение системы запишутся, соответственно, как

$$H = \omega_1 J_1 + \omega_2 J_2, \quad J_1 = J_1^0, \quad J_2 = J_2^0, \quad (2.8)$$

$$\alpha_1 = \omega_1 t + \alpha_1^0, \quad \alpha_2 = \omega_2 t + \alpha_2^0.$$

Поскольку угол всегда определяется по модулю 2π , эволюцию системы удобно представить на поверхности тора. Условимся, что переменные действия J_1 и J_2 обозначают, соответственно, внутренний радиус и радиус поперечного сечения тора, а угловые переменные α_1 и α_2 меняются в соответствии с поворотами радиусов J_1 и J_2 (рис. 2.4). Тогда при фиксированных J_1 и J_2 значения α_1 и α_2 определяют точку на торе. С течением времени координаты α_1 и α_2 меняются согласно (2.8) и, таким образом, задают движение по поверхности тора. В этом случае, в силу произвольности J_1 и J_2 , все фазовое пространство оказывается расслоенным на двумерные торы, которые в обычном трехмерном пространстве можно представить как совокупность вложенных друг в друга торов, большой и малый радиусы которых определяются значениями J_1 и J_2 (рис. 2.5).

Аналогичная картина будет иметь место для любой интегрируемой гамильтоновой системы с двумя степенями свободы, т. е. ее фазовые траектории будут располагаться на концентрических двумерных торах. При этом для нелинейных систем частоты обращения ω_i , $i = 1, 2$, в силу связей $\omega_1 = \omega_1(J_1, J_2)$, $\omega_2 = \omega_2(J_1, J_2)$ — когда они есть, — могут меняться от тора к тору. Фиксируя величины J_k , $k = 1, 2$, можно определить, по поверхности какого тора происходит движение. В общем случае при переходе от одного тора к другому могут меняться не только частоты ω_i , но и их отношение:

$$\frac{\omega_1}{\omega_2} = \frac{\omega_1(J_1, J_2)}{\omega_2(J_1, J_2)}. \quad (2.9)$$

Поведение фазовых траекторий существенным образом зависит от того, является ли отношение (2.9) рациональным или иррациональным. Действительно, если оно рационально, т. е. $\omega_1/\omega_2 = k/m$, где k и m — целые числа, то, задав интервал времени $T = 2\pi(k/\omega_1 + m/\omega_2)$, из равенства (2.8) найдем

$$\begin{aligned} \alpha_1(t+T) &= \omega_1 t + 4\pi k + \alpha_1^0, \\ \alpha_2(t+T) &= \omega_2 t + 4\pi m + \alpha_2^0. \end{aligned} \quad (2.10)$$

Отсюда следует, что выражения $\alpha_1(t+T)$, $\alpha_2(t+T)$ и $\alpha_1(t)$, $\alpha_2(t)$ при фиксированных J_1 , J_2 задают одну и ту же точку на торе. Другими словами, через время T траектория возвращается в точку, из которой она вышла в момент времени t . Таким образом, если отношение (2.9) является рациональным (в этом случае частоты ω_1 и ω_2 часто называют *соизмеримыми*), движение системы является периодическим с частотами ω_1 и ω_2 , а фазовая траектория представляет собой замкнутую непересекающуюся линию на торе. Если же отношение частот (2.9) иррационально (ω_1 и ω_2 *несоизмеримы*), т. е. выражение (2.9) не может быть представлено как $\omega_1/\omega_2 = k/m$, k, m — целые числа, то фазовая траектория никогда не замкнется сама на себя и с течением времени будет подходить сколь угодно близко к любой заданной точке поверхности тора. При этом говорят, что фазовая кривая образует *всюду плотную обмотку тора*, а движение системы называют *квазипериодическим* или *условно периодическим*.

В исходных переменных q, p , если частоты ω_1 и ω_2 соизмеримы, то траектории движения системы в пространстве (q, p) представляют собой определенные замкнутые кривые. Проекция этих кривых на координатные плоскости называются *фигурами Лиссажу* (рис. 2.6а, б). Если же частоты ω_1 и ω_2 несоизмеримы, то фазовые кривые никогда не замыкаются и всюду плотно заполняют определенные области фазового пространства

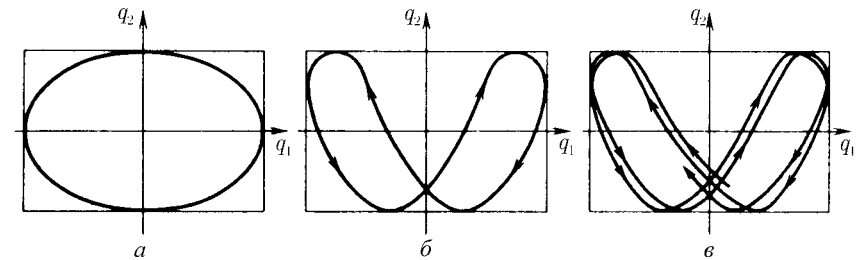


Рис. 2.6. Проекция траекторий гамильтоновой системы (2.6) на координатную плоскость (q_1, q_2) : а — $\omega_2 = \omega_1$, б — $\omega_2 = 2\omega_1$, в — $\omega_1/\omega_2 \neq k/m$

(рис. 2.6в). Из приведенного примера легко видеть, что даже простейшая система двух гармонических осцилляторов может демонстрировать весьма сложное поведение.

Вернемся к полностью интегрируемым системам с n степенями свободы. В этом случае фазовое пространство $2n$ -мерно и в переменных действие–угол имеет структуру множества определенных n -мерных торов. Любая возможная траектория располагается на одном из них. При этом некоторые траектории могут оказаться замкнутыми, другие же будут всюду плотно покрывать поверхность соответствующего тора.

n -мерный тор ($n \geq 2$) с заданными значениями J_1, \dots, J_n называется *резонансным*, если для набора частот $\{\omega_i(J_1, \dots, J_n)\}$ выполняется соотношение $\sum_{i=1}^n k_i \omega_i(J_1, \dots, J_n) = 0$, где величины k_i , $i = 1, 2, \dots, n$, — некоторые отличные от нуля целые числа.

Для нерезонансного тора имеет место квазипериодическое движение системы, и траектория плотно заполняет всю его поверхность. Поскольку частоты непрерывно зависят от значений J_1, \dots, J_n , $\omega_i = \omega_i(J_1, \dots, J_n)$, то, меняя J_i , мы переходим от резонансных торов к нерезонансным и обратно.

14.2 Отображение Пуанкаре

Введем важное понятие, часто используемое при анализе гамильтоновой динамики, — *отображение Пуанкаре*. Для простоты рассмотрим сначала систему с двумя степенями свободы. В этом случае, если гамильтониан явно не зависит от времени, траектория будет находиться на некоторой трехмерной энергетической гиперповерхности $H(p, q) = E$ в четырехмер-

ном фазовом пространстве (q_1, q_2, p_1, p_2) . Последнее равенство позволяет локально выразить любую из этих четырех переменных (например, p_2) как функцию трех остальных: $p_2 = p_2(q_1, q_2, p_1, E)$. Таким образом, усеченное фазовое пространство исходной системы при условии, что дополнительные интегралы движения отсутствуют, будет трехмерным (q_1, q_2, p_1) .

Выберем в этом пространстве некоторую двумерную поверхность S и рассмотрим ее последовательные пересечения фазовой траекторией только в одном направлении (рис. 2.7). Такую поверхность S в предположении, что она пересекается фазовыми кривыми без касания, называют *секущей поверхностью* потока динамической системы. В результате движения некоторая точка пересечения A выбранной секущей поверхности S с фазовой траекторией переходит в другую точку $B = \Phi(A)$ на этой же поверхности. Функция Φ , последовательно переводящая точки пересечения одна в другую, $A \rightarrow B = \Phi(A) \rightarrow C = \Phi(B) \rightarrow A \dots$ (рис. 2.7а), называется *функцией последования*, или *отображением Пуанкаре*.

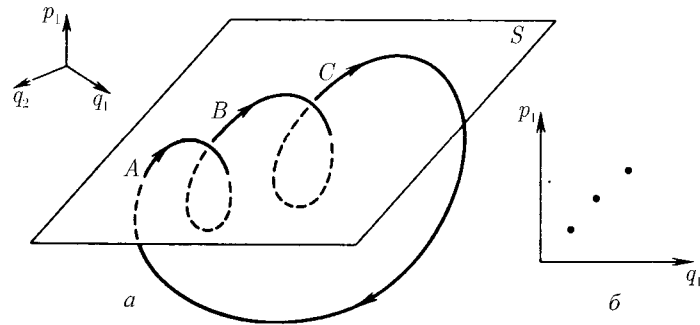


Рис. 2.7. Построение отображения Пуанкаре в фазовом пространстве автономной гамильтоновой системы с двумя степенями свободы

Часто в качестве секущей поверхности S выбирают некоторую плоскость. Тогда задача изучения поведения траекторий в трехмерном фазовом пространстве сводится к анализу отображения Φ двумерного куска плоскости S в себя. Удобным способом наглядного представления поведения системы являются проекции полученного сечения на фазовые плоскости (q_i, p_j) (рис. 2.7б).

Понятие отображения Пуанкаре можно обобщить и на системы с числом степеней свободы $n > 2$. Для автономных систем размерность энергетической гиперповерхности, на которой расположены фазовые кривые,

в этом случае равна $2n - 1$. Исключим, как и раньше, одну из переменных и рассмотрим последовательные точки пересечения траектории динамической системы с $2n - 2$ -мерной гиперповерхностью S . При условии, что поток нигде не касается S , можно определить отображение Φ , которое любой точке M гиперповерхности S ставит в соответствие ближайшую, следующую за M точку M' пересечения фазовой кривой с этой же гиперповерхностью S . Если, помимо интеграла энергии, существуют k дополнительных интегралов движения, то все пересечения необходимо брать с гиперповерхностью, размерность которой равна $[2n - (k + 1)] - 1$. Затем, для наглядности, можно спроецировать полученные точки пересечения на одну из фазовых плоскостей (q_i, p_j) . Тогда анализ динамики исходной системы сведется к исследованию проекций найденных пересечений.

Использование отображения Пуанкаре существенно упрощает описание поведения соответствующей системы дифференциальных уравнений. Однако сведение динамической системы (2.1) к отображению в большинстве случаев удается сделать только численно. Тем не менее, если функция последования Φ построена или если известна структура следов фазовых кривых на секущей S , это дает возможность наглядно представить динамику происходящих в системе процессов.

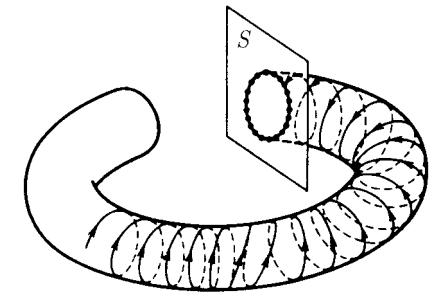


Рис. 2.8. Отображение Пуанкаре для квазипериодического движения

Действительно, когда отображение Пуанкаре представляет собой конечное число точек, последовательно переходящих друг в друга, т. е. $A \rightarrow B \rightarrow C \rightarrow A \rightarrow \dots$ (рис. 2.7а), в фазовом пространстве этому будет соответствовать движение по замкнутой траектории и, таким образом, поведение системы будет периодическим. Если множество точек отображения Пуанкаре плотно заполняет определенную замкнутую кривую, то этому будет отвечать квазипериодическое движение (рис. 2.8).

Наконец, существуют системы, для которых при некоторых условиях траектория на секущей поверхности S представлена «случайным» множеством точек в том смысле, что точки последовательных пересечений нерегулярным образом распределены на S . Следовательно, в этом случае фазовая траектория хаотически блуждает по фазовому пространству. Это указывает на то, что режим эволюции таких систем не является ни периоди-

ческим, ни квазипериодическим. Рассмотрим достаточно простой пример, иллюстрирующий такое нерегулярное поведение, — модель Эно–Эйлеса (1964) [341, 347, 349]. Авторы моделировали движение звезды в среднем поле галактики. После ряда упрощений и учета интегралов движения задача свелась к анализу поведения в четырехмерном фазовом пространстве частицы единичной массы с гамильтонианом

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3. \quad (2.11)$$

Для такой редуцированной задачи имеется интеграл энергии $H = E$, с помощью которого можно исключить одну из переменных: $p_1 = p_1(q_1, q_2, p_2, E)$.

Если бы система с гамильтонианом (2.11) была полностью интегрируемой, то имелись бы переменные действие–угол и с помощью канонического преобразования можно было бы привести функцию Гамильтона к виду $H = H(J_1, J_2)$. Тогда, используя интеграл энергии, можно найти $J_1 = J_1(J_2, E)$. Вторая переменная давала бы дополнительный интеграл движения. Таким образом, при заданных E и J_2 динамическая система была бы представлена движением по поверхности, топологически эквивалентной двумерному тору. В общем случае такая поверхность была бы всюду плотно покрыта фазовыми траекториями, и множество точек отображения Пуанкаре тогда образовало бы некоторую замкнутую кривую.

Эно и Эйлес провели численные исследования системы с гамильтонианом (2.11), используя отображение Пуанкаре. Отмечались последовательные точки пересечения фазовой траектории с координатной плоскостью (q_2, p_2) (т. е. при $q_1 = 0$) для различных значений энергии E . Результаты оказались весьма неожиданными. При достаточно малых энергиях точки отображения Пуанкаре выглядели как совокупность замкнутых и вложенных друг в друга линий, что в фазовом пространстве отвечало движению по поверхности тора с плотной обмоткой (рис. 2.9а). Каждая из замкнутых кривых соответствовала определенному значению дополнительного интеграла движения $J_2 = \text{const}$. С ростом энергии E картина существенно менялась. Некоторые инвариантные кривые начинали разрушаться, образуя густое множество точек, расположенных случайным на вид образом на плоскости (q_2, p_2) . Так, при $E = 0,125$ в сечении $q_1 = 0$ наблюдались как инвариантные замкнутые кривые 1, аналогичные инвариантным кривым при более низких энергиях E , так и множество точек 2, хаотически разбросанных между этими кривыми (рис. 2.9б). Другими словами, в фазовом пространстве при данных значениях энергии E имеются островки регулярного (1) и нерегулярного (2) поведения системы. Для «случайной»

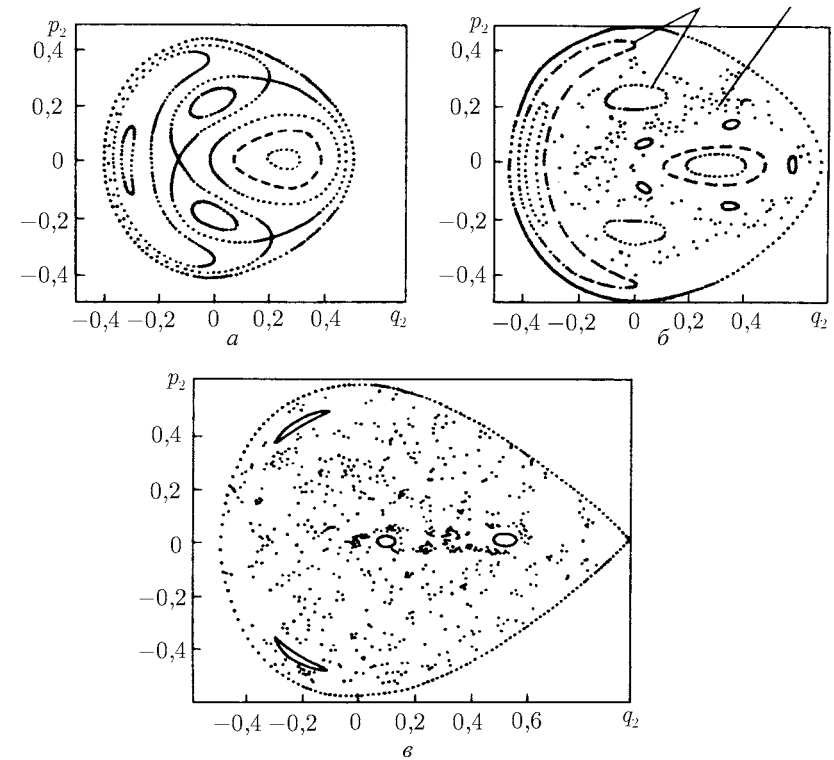


Рис. 2.9. Множество точек отображения Пуанкаре в плоскости (q_2, p_2) для модели Эно–Эйлеса при значениях E , равных: 0,04 (а), 0,125 (б) и 0,167 (в)

траектории 2 переменные действия J_i уже не являются интегралами движения. С дальнейшим увеличением энергии E область, занятая торами, сокращается до ничтожных размеров (рис. 2.9в). Любая траектория вне этих малых участков случайным образом проходит по всей энергетически доступной гиперповерхности $H = E$. Если энергия E достаточно велика, то движение в модели Эно–Эйлеса становится инфинитным.

Дальнейшие исследования системы с гамильтонианом (2.11) показали [116], что и при малых энергиях E (рис. 2.9а) наряду с регулярными траекториями имеются очень тонкие (экспоненциально малые по E^{-1})

и неразличимые при численных расчетах слои, в которых фазовая кривая ведет себя «случайным», хаотическим образом. Это демонстрирует невозможность установить численными методами факт полной интегрируемости системы.

Кроме того, пример Эно–Эйлеса показывает, что даже простые динамические системы с двумя степенями свободы при некоторых условиях обладают поведением, которое выглядит как случайное: их фазовая траектория хаотически блуждает по всей энергетически доступной гиперповерхности.

14.3 Системы, близкие к интегрируемым

В общем случае уравнения Гамильтона (для системы с n степенями свободы) не являются полностью интегрируемыми; для них не существует канонического преобразования $(q, p) \rightarrow (\alpha, J)$, при котором новый гамильтониан не зависел бы от α , т.е. $H = H(J)$. Тем не менее, в ряде случаев функцию Гамильтона $H(\alpha, J)$ можно разбить на интегрируемую часть $H_0 = H_0(J)$ и неинтегрируемую, но представимую в виде возмущения функции H_0 :

$$H(\alpha, J) = H_0 + \epsilon H_1, \quad (2.12)$$

где $H_0 = H_0(J)$, $H_1 = H_1(\alpha, J)$ и ϵ — некоторый малый параметр. Такие системы, гамильтониан которых представим в виде (2.12), называются *системами, близкими к интегрируемым*.

Приведем пример, наглядно демонстрирующий изменение динамики системы, близкой к интегрируемой, в зависимости от параметра ϵ (не предполагая его малым). Это задача о быстром вращении по инерции тяжелого несимметричного твердого тела [30, 94, 315]. Для нее функция Гамильтона может быть записана в виде (2.12). Поскольку нас интересует лишь качественное поведение системы при разных значениях параметра ϵ , мы, не вдаваясь в подробности решения задачи, приведем только конечный результат [30, 315] (рис. 2.10). На этом рисунке представлена проекция на одну из координатных плоскостей множества точек отображения Пуанкаре, возникающих при пересечении определенной гиперповерхности фазовыми кривыми системы. Если $\epsilon = 0$, то система является полностью интегрируемой и точки отображения Пуанкаре образуют семейство регулярных кривых (интегрируемый случай Эйлера) (рис. 2.10 а). При появлении возмущения ($\epsilon > 0$) первоначальная конфигурация следов фазовых кривых нарушается и с увеличением ϵ становится все более хаотичной (рис. 2.10 б, в). Хорошо видно, что при малом ϵ динамика системы по сравнению с невозмущенной меняется мало (рис. 2.10 а, б).

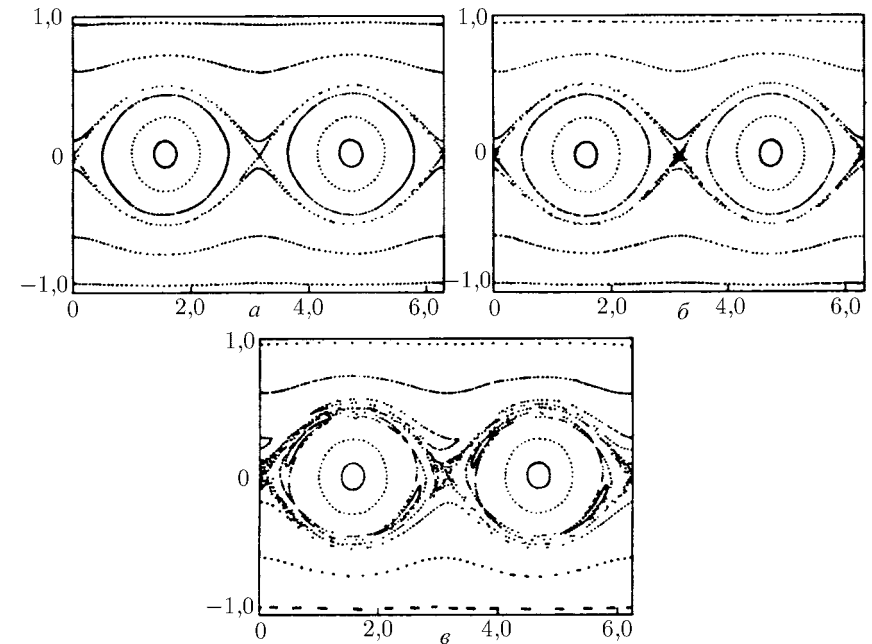


Рис. 2.10. Множество точек отображения Пуанкаре в задаче о вращении твердого тела при значениях ϵ , равных: 0 (интегрируемый случай Эйлера) (а), 0,1 (б), 1,0 (в) [315]

Вернемся к общему анализу системы с гамильтонианом типа (2.12). В выбранных переменных α, J их канонические переменные примут вид

$$\begin{aligned} \dot{J}_j &= -\epsilon \frac{\partial H_1}{\partial \alpha_j}, \\ \dot{\alpha}_j &= \frac{\partial H_0}{\partial J_j} + \epsilon \frac{\partial H_1}{\partial J_j}, \\ j &= 1, 2, \dots, n, \end{aligned} \quad (2.13)$$

где $\epsilon \ll 1$. Если $\epsilon = 0$, то система (2.13) является полностью интегрируемой, и ее решения будут покрывать n -мерные торы. Пусть теперь $\epsilon \neq 0$.

Насколько сильно в этом случае изменится характер интегрируемой системы? Можно было бы ожидать, что при наличии возмущения первоначальная структура фазового пространства сильно нарушится даже при очень малом ϵ . Кроме того, встает вопрос об инвариантных торах: сохранится ли в этом случае квазипериодическое движение?

Ответы на эти вопросы существенным образом зависят от соотношения между частотой возмущающей силы и собственной частотой системы. Так, если собственная частота системы близка к частоте внешней силы, это приведет к нарастанию амплитуды — *резонансу*. Но поскольку для нелинейных систем с возмущением (2.13) частота оказывается зависимой от амплитуды, через некоторое время она выйдет из резонанса. Это приведет к уменьшению амплитуды, что, в свою очередь, повлечет изменение частоты. Поэтому вскоре система вновь возвратится в окрестность резонанса. Появятся так называемые фазовые колебания.

Резонансы могут возникать не только между системой и внешним воздействием, но также и между разными степенями свободы самой системы, что соответствует автономности гамильтониана H_1 в соотношении (2.12). В этом случае говорят о *внутренних* резонансах.

Если резонанс не является изолированным, перекрытие резонансов приводит к появлению очень сложного движения в системе. Кроме этого, резонансы препятствуют нахождению решений уравнений методом канонической теории возмущений. При использовании теории возмущений исходная система аппроксимируется близкой к ней интегрируемой системой и налагаемым на нее малым возмущением, а решение задачи ищется в виде разложения по степеням ϵ . Наличие резонансов нарушает сходимость таких рядов. Это связано с тем, что при таком подходе неявно предполагается, что исходные уравнений являются интегрируемыми. Однако в большинстве случаев это не так. Даже достаточно простые системы не интегрируемы и при некоторых начальных условиях могут обладать очень сложной динамикой (см. выше). Теория возмущений не может описать такое сложное поведение, что формально и выражается в расходимости рядов. Если же начальные условия системы отвечают регулярным траекториям (квазипериодическому движению), то в окрестности резонансов под действием возмущения такие траектории претерпевают качественное изменение. Например, плотная траектория на торе превращается в цикл, замыкающийся через конечное число оборотов. Обычная теория возмущений не отражает такого изменения, что влечет появление малых знаменателей и снова расходимость рядов.

Таким образом, для систем, близких к интегрируемым, резонансы и их взаимодействие играют важную роль. Поэтому остановимся на этом явлении подробнее.

15. Нелинейный резонанс

Теория нелинейного резонанса замечательна тем, что позволяет относительно легко получить аналитический критерий возникновения нерегулярного движения в гамильтоновых системах. Этот критерий впервые был введен Б. Чириковым [194, 263]. Он основан на переходе к так называемым резонансным переменным и оценке степени перекрытия резонансов. В результате такого преобразования мы переходим во вращающуюся систему координат. Рассмотрим это с более формальной и общей точки зрения, следуя [81, 83] (см. также [84, 116]).

15.1 Малые знаменатели

Остановимся сначала на внутреннем резонансе. Разложим функцию $H_1(\alpha, \mathbf{J})$ (см. (2.12)) в ряд Фурье, $H_1(\alpha, \mathbf{J}) = \sum_k H_1^k(\mathbf{J}) e^{ik\alpha}$, и подставим разложение в уравнения Гамильтона (2.13). Это дает:

$$\begin{aligned} \dot{J}_j &= -i\epsilon \sum_k k_j H_1^k e^{ik\alpha}, \\ \dot{\alpha}_j &= \omega_j(\mathbf{J}) + \epsilon \sum_k \frac{\partial H_1^k}{\partial J_j} e^{ik\alpha}, \quad j = 1, 2, \dots, n. \end{aligned} \quad (2.14)$$

Здесь \mathbf{k} — вектор с целыми вещественными компонентами, H_1^k — коэффициенты Фурье и $\omega_j(\mathbf{J}) = \partial H_0(\mathbf{J}) / \partial J_j$.

Будем искать решение возмущенной системы (2.14) в виде рядов по степеням малого параметра ϵ :

$$\begin{aligned} J_j &= J_j^{(0)} + \sum_{s=1}^{\infty} \epsilon^s J_j^{(s)}, \\ \alpha_j &= \alpha_j^{(0)} + \sum_{s=1}^{\infty} \epsilon^s \alpha_j^{(s)}, \quad j = 1, 2, \dots, n. \end{aligned} \quad (2.15)$$

Теперь, используя это разложение и уравнения (2.14), отберем члены, стоящие при одинаковых степенях параметра ϵ . Тогда в нулевом приближении находим: $\dot{J}_j^{(0)} = 0$, $\dot{\alpha}_j^{(0)} = \omega_j(\mathbf{J}^{(0)})$, $j = 1, 2, \dots, n$. Отсюда

$$J_j^{(0)} = \text{const}, \quad \alpha_j^{(0)} = \omega_j(\mathbf{J}^{(0)}) t + \text{const}. \quad (2.16)$$

Учитывая, что $\omega_j(\mathbf{J}^{(0)} + \epsilon \mathbf{J}^{(1)}) \approx \omega_j(\mathbf{J}^{(0)}) + \epsilon \sum_m \frac{\partial \omega_j(\mathbf{J}^{(0)})}{\partial J_m^{(0)}} J_m^{(1)}$, нетрудно найти уравнения первого приближения:

$$\begin{aligned} \dot{J}_j^{(1)} &= -i \sum_k k_j \tilde{H}_1^k(\mathbf{J}^{(0)}) e^{ik\omega(\mathbf{J}^{(0)})t}, \\ \dot{\alpha}_j^{(1)} &= \sum_m \frac{\partial \omega_j(\mathbf{J}^{(0)})}{\partial J_m^{(0)}} J_m^{(1)} + \sum_k \frac{\partial \tilde{H}_1^k(\mathbf{J}^{(0)})}{\partial J_j} e^{ik\omega(\mathbf{J}^{(0)})t}, \end{aligned} \quad (2.17)$$

где \tilde{H}_1^k означает учет константы в коэффициентах H_1^k . Эти уравнения сразу интегрируются:

$$\begin{aligned} J_j^{(1)} &= - \sum_k \frac{k_j \tilde{H}_1^k(\mathbf{J}^{(0)})}{k\omega} e^{ik\omega t} + \text{const}, \\ \alpha_j^{(1)} &= i \sum_m \frac{\partial \omega_j(\mathbf{J}^{(0)})}{\partial J_m^{(0)}} \sum_k \frac{k_m \tilde{H}_1^k(\mathbf{J}^{(0)})}{(k\omega)^2} e^{ik\omega t} - \\ &\quad - i \sum_k \frac{\partial \tilde{H}_1^k(\mathbf{J}^{(0)})}{\partial J_j} \frac{e^{ik\omega t}}{k\omega}. \end{aligned} \quad (2.18)$$

Легко видеть, что при выполнении условия

$$k\omega = k_1\omega_1 + k_2\omega_2 + \dots + k_n\omega_n \approx 0, \quad (2.19)$$

которое называют *соотношением резонанса*, в (2.18) появляются члены с равными или близкими к нулю знаменателями. Это приводит к значительному росту поправок $\alpha_j^{(s)}$, $J_j^{(s)}$, что явно нарушает сходимость рядов (2.15).

Если соотношение резонанса не выполняется в уравнениях первого приближения (2.17), то оно может иметь место для уравнений более высокого приближения $s > 1$. Резонанс, проявляющийся в s -ом порядке теории возмущений, называется *резонансом порядка s* .

Пусть теперь возмущение H_1 периодически с периодом $T = 2\pi/\Omega$ зависит от времени, т. е. $H_1(\alpha, \mathbf{J}, t) = H_1(\alpha, \mathbf{J}, t+T)$. Поступая аналогично предыдущему случаю, имеем:

$$H_1(\alpha, \mathbf{J}, t) = \sum_{k,m} H_1^{km}(\mathbf{J}) \exp[i(k\alpha - m\Omega)t].$$

Уравнения Гамильтона при этом переходят в следующие:

$$\begin{aligned} \dot{J}_j &= -i\epsilon \sum_{k,m} k_j H_1^{km} \exp[i(k\alpha - m\Omega)t], \\ \dot{\alpha}_j &= \omega_j(\mathbf{J}) + \epsilon \sum_{k,m} \frac{\partial H_1^{km}}{\partial J_j} \exp[i(k\alpha - m\Omega)t], \\ j &= 1, 2, \dots, n. \end{aligned}$$

Как и ранее, в нулевом приближении решение будет иметь вид (2.16). Нетрудно получить уравнения первого приближения, которые легко интегрируются и дают:

$$\begin{aligned} J_j^{(1)} &= - \sum_{k,m} \frac{k_j \tilde{H}_1^{km}(\mathbf{J}^{(0)})}{k\omega - m\Omega} \exp[i(k\omega - m\Omega)t] + \text{const}, \\ \alpha_j^{(1)} &= i \sum_l \frac{\partial \omega_j(\mathbf{J}^{(0)})}{\partial J_l^{(0)}} \sum_{k,m} \frac{k_l \tilde{H}_1^{km}(\mathbf{J}^{(0)})}{(k\omega - m\Omega)^2} \exp[i(k\omega - m\Omega)t] - \\ &\quad - i \sum_{k,m} \frac{\partial \tilde{H}_1^{km}(\mathbf{J}^{(0)})}{\partial J_j} \frac{\exp[i(k\omega - m\Omega)t]}{k\omega - m\Omega}. \end{aligned}$$

Здесь \tilde{H}_1^{km} — это значения H_1^{km} с учетом константы, которая появляется в нулевом порядке. Таким образом, если выполняется *соотношение резонанса*

$$k\omega - m\Omega \approx 0, \quad (2.20)$$

то ряды теории возмущений (2.15) расходятся. В этом и заключается проблема малых знаменателей. Интегрирование систем по теории возмущений в этом случае если и возможно, то лишь в отсутствие резонансов и при условии сходимости рядов (2.15).

Обратимся к выражению (2.20). Для систем с одной степенью свободы оно означает, что частоты ω и Ω соизмеримы, а динамика является периодической. В общем случае резонансные соотношения определяют резонансные торы, на которых происходит движение. Следовательно, резонанс соответствует случаю, когда траектория системы замыкается через конечное число оборотов на торе.

Если же соотношения (2.19) или (2.20) выполняются не для всех частот ω_j , то фазовая траектория будет плотно заполнять тор, число измерений которого равно числу рационально независимых частот. Стало быть, явление резонанса является препятствием также и в определении инвариантных торов.

Для преодоления этих трудностей было предложено использовать каноническое преобразование к специальным (резонансным) переменным. Такое преобразование физически соответствует переходу во вращающуюся систему координат. При этом на фазовой плоскости центру резонанса отвечает эллиптическая точка, а новые переменные описывают медленные колебания на нем.

Остановимся на этом подробнее и с общих позиций, следуя [81, 83, 116]. Изучим сначала системы с одной степенью свободы ($n = 1$).

15.2 Резонансные переменные

Пусть возмущение H_1 является периодической функцией времени с периодом $T = 2\pi/\nu$, и движение описывается гамильтонианом

$$H = H_0(J) + \epsilon H_1(\alpha, J, t), \quad (2.21)$$

где $\epsilon \ll 1$. Тогда функцию H_1 можно разложить в ряд Фурье:

$$H_1(\alpha, J, t) = \sum_{k,m} H_1^{km}(J) e^{i(k\alpha - m\nu t)}. \quad (2.22)$$

С учетом этого разложения уравнения Гамильтона (2.13) примут вид

$$\begin{aligned} \dot{J} &= -i\epsilon \sum_{k,m} k H_1^{km}(J) e^{i(k\alpha - m\nu t)}, \\ \dot{\alpha} &= \omega(J) + \epsilon \sum_{k,m} \frac{dH_1^{km}(J)}{dJ} e^{i(k\alpha - m\nu t)}, \end{aligned} \quad (2.23)$$

где $H_1^{-k,-m} = H_1^{*k,m}$. Легко видеть, что если в этих уравнениях выполняется условие

$$k\omega(J) - m\nu \approx 0, \quad (2.24)$$

то исходную систему нельзя исследовать методом теории возмущений: возникает резонанс.

Чтобы изучить динамику в окрестности резонанса, выделим из разложения (2.22) резонансный член и рассмотрим поведение системы, обусловленное только этим членом. Для этого зафиксируем тройку чисел k_0, m_0, J_0 , так чтобы условие резонанса выполнялось точно,

$$k_0\omega(J_0) = m_0\nu, \quad (2.25)$$

а в уравнениях (2.23) сохраним *только* резонансную гармонику. Имеем:

$$\begin{aligned} \dot{J} &= -i\epsilon k_0 H_1^{k_0 m_0}(J) e^{i(k_0\alpha - m_0\nu t)} + i\epsilon k_0 H_1^{*k_0 m_0}(J) e^{-i(k_0\alpha - m_0\nu t)} = \\ &= 2\epsilon k_0 \frac{1}{2i} \left[\left| H_1^{k_0 m_0}(J) \right| e^{i(\psi + k_0\alpha - m_0\nu t)} - \left| H_1^{k_0 m_0}(J) \right| e^{-i(\psi + k_0\alpha - m_0\nu t)} \right] = \\ &= 2\epsilon k_0 \left| H_1^{k_0 m_0}(J) \right| \sin(k_0\alpha - m_0\nu t + \psi). \end{aligned}$$

Аналогично

$$\begin{aligned} \dot{\alpha} &= \omega(J) + \epsilon \frac{d}{dJ} H_1^{k_0 m_0}(J) e^{i(k_0\alpha - m_0\nu t)} + \epsilon \frac{d}{dJ} H_1^{*k_0 m_0}(J) e^{-i(k_0\alpha - m_0\nu t)} = \\ &= \omega(J) + 2\epsilon \frac{d}{dJ} \left| H_1^{k_0 m_0}(J) \right| \cos(k_0\alpha - m_0\nu t + \psi). \end{aligned}$$

Введем новые обозначения: $\varphi \equiv k_0\alpha - m_0\nu t + \psi$, $H_1^0 \equiv 2 \left| H_1^{k_0 m_0} \right|$. Тогда уравнения (2.23) переписутся как

$$\begin{aligned} \dot{J} &= \epsilon k_0 H_1^0(J) \sin \varphi, \\ \dot{\varphi} &= k_0\omega(J) - m_0\nu + \epsilon k_0 \frac{dH_1^0(J)}{dJ} \cos \varphi. \end{aligned} \quad (2.26)$$

Сохранив в разложении (2.22) только резонансный член, находим выражение для гамильтониана (2.21):

$$H = H_0(J) + \epsilon H_1^0 \cos \varphi. \quad (2.27)$$

Допустим теперь, что величина J достаточно близка к J_0 , так что отклонение $\Delta J \equiv J - J_0$ является малой величиной. В этом случае $H_0(J)$ и $\omega(J)$ можно разложить в ряд по ΔJ :

$$H_0(J) = H_0(J_0) + \frac{\partial H_0}{\partial J} \Delta J + \frac{1}{2} \frac{\partial^2 H_0}{\partial J^2} (\Delta J)^2 + \dots, \quad (2.28)$$

$$\omega(J) = \omega(J_0) + \frac{\partial \omega}{\partial J} \Delta J + \frac{1}{2} \frac{\partial^2 \omega}{\partial J^2} (\Delta J)^2 + \dots \quad (2.29)$$

Отбросим в выражении (2.28) члены выше второго порядка малости, а в выражении (2.29) — выше первого порядка. Кроме того, величину $H_1^0(J)$ возьмем в точке J_0 , учтем равенство (2.25) и во втором уравнении (2.26) пренебрежем членом $\sim \epsilon$. Тогда система (2.26) и гамильтониан (2.27) преобразуются, соответственно, к виду:

$$\frac{d}{dt} \Delta J = \epsilon k_0 H_1^0 \sin \varphi, \quad \dot{\varphi} = k_0 \frac{d\omega(J_0)}{dJ} \Delta J, \quad (2.30)$$

$$H = H_0(J_0) + \omega(J_0) \Delta J + \frac{1}{2} \frac{\partial \omega(J_0)}{\partial J} (\Delta J)^2 + \epsilon H_1^0 \cos \varphi. \quad (2.31)$$

Здесь для удобства переобозначено $H_1^0 \equiv H_1^0(J_0)$. Канонически сопряженными переменными для системы (2.30) являются переменные $(\Delta J, \varphi)$, а сама система (2.30), как нетрудно видеть, порождается гамильтонианом

$$\bar{H} = \frac{1}{2} k_0 \frac{d\omega(J_0)}{dJ} (\Delta J)^2 + \epsilon k_0 H_1^0 \cos \varphi, \quad (2.32)$$

который называется *универсальным гамильтонианом нелинейного резонанса* [263].

Каноническое преобразование, позволяющее совершить переход от системы с гамильтонианом (2.31) к системе к универсальным гамильтонианом нелинейного резонанса, имеет вид:

$$(\alpha, J) \longrightarrow (k_0 \alpha - m_0 \nu t, \Delta J), \quad \nu = \omega(J_0), \quad (2.33)$$

$$\bar{H} = k_0 H - m_0 \nu \Delta J - k_0 H_0(J_0).$$

Физически оно соответствует переходу во вращающуюся систему координат с частотой $m_0 \nu$.

Проанализируем выражение (2.32) и уравнения (2.30). Сразу видно сходство гамильтониана \bar{H} с гамильтонианом нелинейного маятника (2.3), если в последнем положить $l = 1$. Действительно, роль импульса p играет переменная ΔJ , а эффективной массы — величина $m = \left(k_0 \frac{d\omega(J_0)}{dJ} \right)^{-1}$. Более того, систему (2.30) с помощью замены $\Omega_0^2 = \epsilon k_0^2 H_1^0 |d\omega(J_0)/dJ|$ можно записать как

$$\ddot{\varphi} - \Omega_0^2 \sin \varphi = 0. \quad (2.34)$$

Это уравнение совпадает с точностью до сдвига по фазе на π с уравнением (2.4). Подобно тому, как мы поступили при анализе движения маятника,

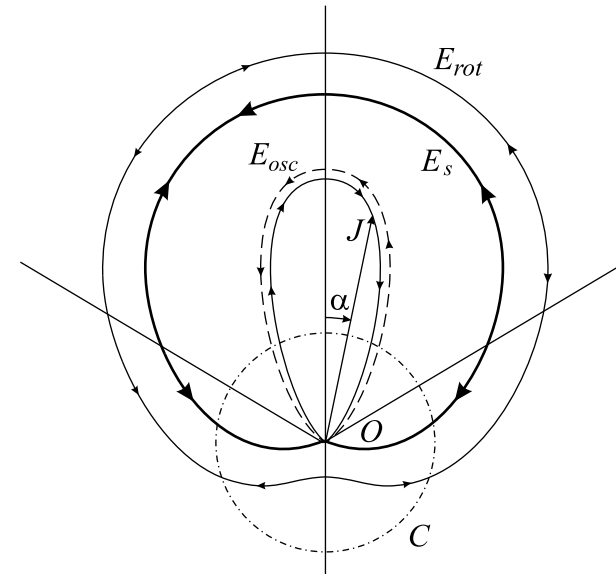


Рис. 2.11. Нелинейный резонанс первого порядка в переменных (α, J)

представив динамику на фазовой плоскости (α, J) (см. §14), рассмотрим поведение системы в окрестности резонанса.

Пусть $k_0 = 1$, $m_0 = 1$ (резонанс первого порядка). В исходных переменных (α, J) фазовый портрет будет иметь вид, представленный на рис. 2.11. Локальным осцилляциям системы отвечает двойная каплеобразная кривая, соответствующая малой энергии, $E = E_{osc}$. Если обратиться к аналогии с маятником, то это будет отвечать его колебаниям. Наружняя кривая соответствует энергии $E = E_{rot}$, что для маятника означает вращение. Сепаратриса ($E = E_s$) разделяет эти два качественно различных типа движения.

Перейдем теперь во вращающуюся систему координат согласно (2.33). Тогда в каждый момент времени к мгновенному значению переменной J необходимо добавить (вычесть) постоянную величину, равную радиусу окружности C , по которой происходит вращение, т.е. величину J_0 (рис. 2.11). В результате фазовый портрет системы будет выглядеть, как показано на рис. 2.12.

Поясним этот рисунок, пользуясь аналогией с маятником. Если маятник вращается в сторону, противоположную вращению системы координат

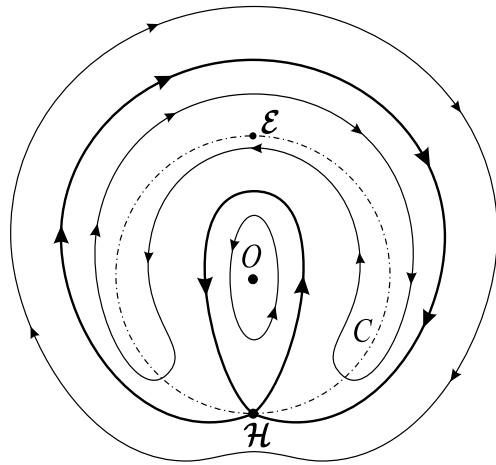


Рис. 2.12. Нелинейный резонанс первого порядка во вращающейся системе координат. Тонкие линии — фазовые колебания, штрихпунктирная линия — невозмущенная траектория ($J = J_0$), жирная кривая — сепаратриса [81]

нат, то такому движению отвечают линии, расположенные вокруг точки O внутри малой петли сепаратрисы. Если же направление вращения маятника совпадает с направлением вращения системы координат, то этому соответствуют линии, находящиеся снаружи большой петли сепаратрисы. Подковообразные замкнутые кривые, расположенные между двумя петлями сепаратрисы, отвечают колебаниям маятника. Они происходят вокруг эллиптической точки, обозначенной на рисунке как \mathcal{E} . Через гиперболическую точку \mathcal{H} проходят обе (большая и малая) петли сепаратрисы. Окружности C отвечает нижнее состояние равновесия маятника.

Для исходной системы окружности C соответствует невозмущенная траектория при $J = J_0$. Направления движения на рис. 2.12 указаны стрелками. Таким образом, во вращающейся системе координат вся картина фазовых траекторий сохраняется.

Максимальное расстояние между двумя (большой и малой) петлями сепаратрисы называют *шириной нелинейного резонанса* (а также *шириной сепаратрисы*). Нетрудно оценить эту величину. Из (2.32) имеем для ширины резонанса по действию:

$$\max(\Delta J) = \sqrt{2\epsilon H_1^0 |d\omega(J_0)/dJ|^{-1}}, \quad (2.35)$$

и по частоте

$$\max(\Delta\omega) = \left| \frac{d\omega(J_0)}{dJ} \right| \max(\Delta J) = \sqrt{2\epsilon H_1^0 \left| \frac{d\omega(J_0)}{dJ} \right|}. \quad (2.36)$$

Определим теперь, насколько оправданы приближения, сделанные нами выше, а также найдем порядки величин (2.35), (2.36).

Из (2.27) следует, что $H_0 \sim H_1^0$. Но поскольку $H_0 \sim J_0\omega(J_0)$, то

$$H_0 \sim H_1^0 \sim J_0\omega(J_0). \quad (2.37)$$

Поэтому, учитывая (2.35), (2.36), получаем:

$$\frac{\max(\Delta J)}{J_0} \sim \sqrt{\frac{\epsilon}{\gamma}}, \quad \frac{\max(\Delta\omega)}{\omega(J_0)} \sim \sqrt{\epsilon\gamma}, \quad (2.38)$$

где $\gamma \equiv \left(\frac{d\omega(J_0)}{dJ} \right) \left(\frac{J_0}{\omega(J_0)} \right)$ — так называемый *параметр нелинейности*. При выводе системы (2.30) мы отбросили член $\sim \epsilon$ в уравнении (2.26) для $\dot{\varphi}$. Это возможно (см. (2.26), (2.30)), если $(dH_1^0(J_0)/dJ_0) \epsilon \ll \left| \frac{d\omega(J_0)}{dJ} \right| \Delta J$ или, оценивая порядок величин, $\epsilon H_1^0/J_0 \ll \left| \frac{d\omega(J_0)}{dJ} \right| \Delta J$. Отсюда, согласно (2.37), (2.38), находим

$$\gamma \gg \epsilon. \quad (2.39)$$

Кроме того, мы заменили величину $H_1^0(J)$ на $H_1^0(J_0)$. Эта замена эквивалентна условию $J - J_0 \ll 1$ или $\Delta J \ll J_0$. Последнее соотношение, как легко видеть из (2.38), приводит к неравенству (2.39).

Далее, сохранение только резонансного слагаемого и отбрасывание нерезонансных членов допустимо, когда другие частоты в разложении (2.29) качественно не изменяют динамику, т. е. при $\omega(J_0) \gg \left| \frac{d\omega(J_0)}{dJ} \right| \Delta J$. Используя оценку (2.36), получаем $\gamma\epsilon \ll 1$, которое имеет место при малых γ и ϵ . Теперь, объединяя его с (2.39), находим

$$\epsilon \ll \gamma \ll \frac{1}{\epsilon}. \quad (2.40)$$

Это неравенство называется *условием умеренной нелинейности* [81, 263].

Таким образом, изложенная выше приближенная теория нелинейного резонанса справедлива при условии выполнения неравенства (2.40). Следовательно, предельного перехода к линейной задаче ($\gamma \rightarrow 0$) не существует.

Условие резонанса (2.25) может выполняться, вообще говоря, при любом значении k_0 . В этом случае все допущения и формулы остаются прежними. Однако, поскольку фаза определяется выражением $\varphi = k_0\alpha - \nu t + \psi$, это приводит к появлению гирлянды из петель сепаратрисы. Число таких петель равно k_0 . Следовательно, количество пар эллиптических и гиперболических точек также будет равным k_0 (рис. 2.13).

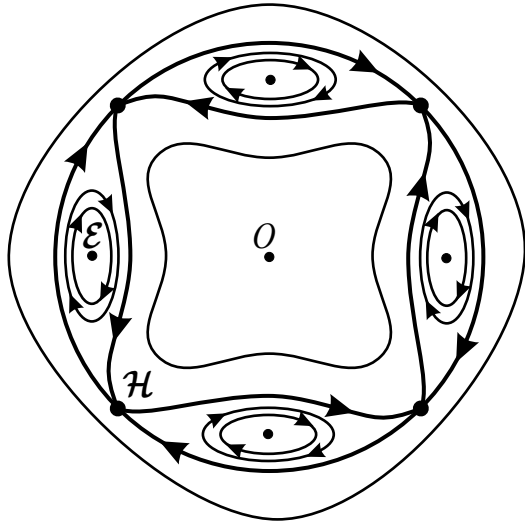


Рис. 2.13. Нелинейный резонанс при $k_0 = 4$, $m_0 = 1$

Приведенное описание нелинейного резонанса легко обобщается на системы со многими степенями свободы. Однако при $n \geq 2$ возможно появление внутренних резонансов, т. е. резонансов между степенями свободы самой системы. Применяя допущения, аналогичные изложенным выше, исследуем такие *внутренние* резонансы [81, 84].

Рассмотрим, для наглядности, систему с двумя степенями свободы ($n = 2$). Гамильтониан такой системы можно записать как

$$H = H_{01}(J_1) + H_{02}(J_2) + \epsilon H_1(J_1, J_2, \alpha_1, \alpha_2),$$

а уравнения движения представятся в виде

$$\dot{J}_j = -\epsilon \frac{\partial H_1}{\partial \alpha_j}, \quad \dot{\alpha}_j = \frac{\partial}{\partial J_j} (H_{01} + H_{02}) + \epsilon \frac{\partial H_1}{\partial J_j}, \quad j = 1, 2.$$

Согласно соотношению (2.19) внутренний резонанс возникает, если условие

$$k\omega_1(J_{01}) - m\omega_2(J_{02}) = 0$$

выполняется для некоторых целых чисел (k, m) и значений переменных действия (J_{01}, J_{02}) . Сохраняя в разложении

$$H_1 = \sum_{k,m} H_1^{km}(J_1, J_2) e^{i(k\alpha_1 - m\alpha_2)}$$

только резонансную гармонику, раскладывая функции H_{0j} , $\omega_j = dH_{0j}/dJ_j$, $j = 1, 2$, в окрестности резонанса (J_{01}, J_{02}) и используя приближения, описанные выше, находим уравнения движения

$$\begin{aligned} \dot{J}_1 &= \epsilon k H_1^0 \sin \varphi, \\ \dot{J}_2 &= -\epsilon m H_1^0 \sin \varphi, \end{aligned} \quad (2.41)$$

$$\dot{\varphi} = k \frac{d\omega_1(J_{01})}{dJ_1} \Delta J_1 - m \frac{d\omega_2(J_{02})}{dJ_2} \Delta J_2$$

и универсальный гамильтониан

$$\bar{H} = \frac{1}{2} \frac{d\omega_1(J_{01})}{dJ_1} (\Delta J_1)^2 + \frac{1}{2} \frac{d\omega_2(J_{02})}{dJ_2} (\Delta J_2)^2 + \epsilon H_1^0 \cos \varphi.$$

Здесь введены обозначения $\Delta J_j = J_j - J_{0j}$, $j = 1, 2$, $H_1^0 e^{i\psi} \equiv 2H_1^{km}(J_{01}, J_{02}) = 2 |H_1^{km}(J_{01}, J_{02})| e^{i\psi}$ и $\varphi = k\alpha_1 - m\alpha_2 + \psi$.

Легко видеть, что, умножив первое уравнение системы (2.41) на m , а второе — на k и сложив их, получим дополнительный интеграл движения $mJ_1 + kJ_2 = \text{const}$. Следовательно, уравнения (2.41) являются интегрируемыми. Можно, однако, пойти иным путем. Действительно, продифференцировав уравнение для $\dot{\varphi}$ по времени, приходим к выражению (2.34), где $\Omega_0^2 = \epsilon H_1^0 |k^2 d\omega_1(J_{01})/dJ_1 + m^2 d\omega_2(J_{02})/dJ_2|$. Если величины $d\omega_1(J_{01})/dJ_1$ и $d\omega_2(J_{02})/dJ_2$ имеют разные знаки, то выражение под знаком модуля может обратиться в нуль, т. е. в системе допустимо дополнительное вырождение.

Для систем с более чем двумя степенями свободы ($n \geq 3$) число дополнительных интегралов движения увеличивается, но вся динамика по-прежнему будет определяться уравнением типа (2.34).

Таким образом, при исследовании систем, близких к интегрируемым, существенную роль играет отношение частот: их несоизмеримость в типичном случае определяет квазипериодическую траекторию, плотно покрывающую тор. Но рациональное отношение приводит к появлению резонансов и изменяет структуру инвариантных поверхностей.

15.3 Перекрытие резонансов

Весь анализ нелинейного резонанса был проведен при допущении, что величина J (см. (2.25)) оставалась фиксированной. Иными словами, мы предполагали, что существует *единственный* изолированный *первичный* нелинейный резонанс. На самом деле, согласно (2.24), резонансное соотношение может выполняться и для других значений J_m , $m = 1, 2, \dots$. Иначе говоря, первичных резонансов в системе может быть много, и соотношение (2.25) при $k_0 = 1$ нужно переписать как

$$\omega(J_m) = m\nu.$$

Следовательно, можно ввести расстояние между резонансами по действию

$$\delta J_m = J_{m+1} - J_m$$

и по частоте

$$\delta\omega_m = \omega(J_{m+1}) - \omega(J_m).$$

Если все значения J_m расположены достаточно далеко друг от друга, т. е.

$$\delta J_m \gg \Delta J_m, \quad (2.42)$$

где ΔJ_m определяется соотношением (2.35), то вблизи резонанса фазовые траектории на плоскости (φ, J) качественно можно представить в виде, изображенном на рис. 2.14.

Таким образом, условие (2.42) означает, что сепаратрисы резонансов не пересекаются, а сами резонансы не взаимодействуют друг с другом. При попадании начальных условий системы в область одного из них описание динамики может быть выполнено в приближении влияния только этого резонанса. Если же начальная точка попала в область между резонансами, то система может быть изучена в рамках нерезонансного приближения [49, 116].

Аналогично можно определить условие отсутствия перекрытия (взаимодействия) резонансов по частотам:

$$\delta\omega_m \gg \Delta\omega_m. \quad (2.43)$$

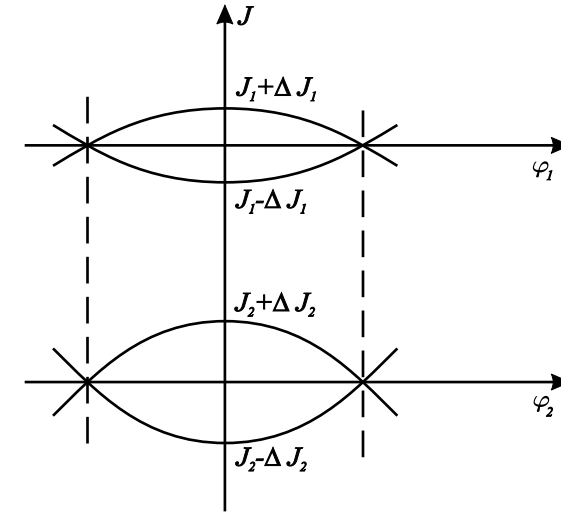


Рис. 2.14. Невзаимодействующие резонансы

Допустим теперь, что соотношение (2.42) или (2.43) не выполняется, а резонансы близки друг к другу, так что соответствующие им сепаратрисы могут перекрываться. Что произойдет в этом случае, который называют сильным взаимодействием резонансов?

Введем параметр K , определяющий степень перекрытия резонансов:

$$K = \frac{\Delta\omega}{\delta\omega} \sim \frac{\Delta J}{\delta J}.$$

Величина K называется *параметром перекрытия резонансов*. Он был введен Чириковым для характеристики динамики гамильтоновых систем [194] (см. также [195, 263, 296]). При условии малости взаимодействия, т. е. при $K \ll 1$ движение в системе должно быть регулярным, и в общем случае фазовые траектории будут покрывать всюду плотно $(n - m)$ -мерные торы. Однако при сильном перекрытии, $K > 1$, динамика системы оказывается очень сложной и отличной от периодического и квазипериодического режимов.

Степень перекрытия резонансов как критерий возникновения нерегулярного движения подтверждается большим количеством численных исследований (см. [84, 116, 263]). В частности, в работе [522] было обнаружено рождение хаоса при перекрытии всего лишь двух резонансов.

Особенность этого критерия состоит в том, что при изучении конкретных систем его относительно просто использовать. Действительно, достаточно применить описанную выше технику в окрестности только одного резонанса в приближении отсутствия всех остальных. Это делает критерий перекрытия резонансов очень удобным в практическом отношении. Вместе с тем, он работает не всегда, и его в ряде случаев необходимо уточнить [116, 263, 296].

Какова же структура резонансов вблизи эллиптических и гиперболических точек? Как мы увидим далее, этот вопрос тесно связан с происхождением и природой хаотичности.

16. Элементы теории Колмогорова – Арнольда – Мозера (теории КАМ). Диффузия Арнольда

Итак, нелинейный резонанс показывает, что даже небольшие возмущения могут весьма существенным образом повлиять на динамику интегрируемой гамильтоновой системы. Резонансы изменяют топологию фазовых траекторий и приводят к образованию цепи островов в фазовом пространстве. Теория возмущений не описывает такие резонансы, поскольку регулярные решения вблизи них сильно возмущены, а это влечет появление малых знаменателей и расходимость рядов. На эту проблему обратил внимание еще А. Пуанкаре, назвав ее фундаментальной проблемой классической механики. Она была решена только в начале 60-х годов с появлением знаменитой теории Колмогорова – Арнольда – Мозера (теории КАМ) [20, 21, 96, 148, 149].

Чтобы подчеркнуть физическую значимость этой теории, совершим небольшой экскурс в историю. Иоганн Кеплер, анализируя свои многолетние астрономические наблюдения, пришел к выводу, что планеты движутся вокруг Солнца по эллиптическим орбитам. Однако, строго говоря, он нашел только первое приближение траекторий планет. Если бы в Солнечной системе существовала только одна планета, то ее движение действительно было бы эллиптическим. Иными словами, в отсутствие влияния планет друг на друга их орбиты будут эллиптическими. Именно такой вывод и сделал Иоганн Кеплер, опираясь на свои расчеты. Но эти расчеты были основаны на недостаточно точных наблюдениях, не позволяющих в то время сделать заключение об истинных траекториях движения планет. Полное решение задачи Кеплера в 1710 г. дал И. Бернулли.

На самом деле в Солнечной системе планеты хотя и слабо, но все же влияют друг на друга, возмущая эллиптическое движение. Такое возмущение приводит к тому, что фактически орбиты представляют собой де-

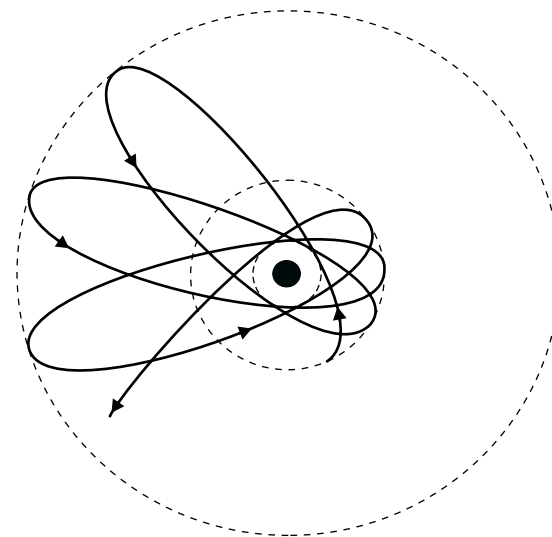


Рис. 2.15. Прецессия перигелия планеты, вращающейся вокруг Солнца (смещение сильно преувеличено)

формированные эллипсы с медленной прецессией, т. е. при каждом обороте планеты вокруг Солнца их перигелий смещается на некоторую величину (рис. 2.15). Максимальное значение прецессии демонстрирует Меркурий. Смещение его перигелия, обусловленное влиянием остальных планет, оценивается величиной $\sim 532''$ за 100 лет [58]¹.

Эта планетарная задача, сводящаяся в общем случае к известной проблеме N тел, относится к одному из самых важных вопросов, появившихся в результате развития как математики, так и физики. В свое время в 1885 году королем Швеции и Норвегии Оскаром II была даже учреждена премия за ее решение: «Дана система из произвольно большого числа материальных точек, взаимодействующих друг с другом по законам Ньютона. Попытайтесь представить (при условии, что никакие две точки никогда не столкнутся) координаты каждой точки в виде рядов от переменной, которая является некоторой известной функцией времени и для всех значений которой ряд сходится равномерно»². Таким образом, основной задачей было не только найти формальное разложение решений задачи N тел, но и доказать

¹Подчеркнем, что мы рассматриваем только ньютоновское взаимодействие.

²Недавно вышла замечательная книга [75], где описана история развития небесной механики и нелинейной динамики от Исаака Ньютона до наших дней.

сходимость этих рядов. Последняя часть как раз и представляла основную трудность.

Не останавливаясь на истории вопроса, отметим только, что премию получил Анри Пуанкаре за неоценимый вклад в развитие задачи N тел и связанные с ней фундаментальные задачи динамики.

Понадобилось очень много времени, чтобы понять, что проблема N тел не проста, и для ее решения требуется использовать мощные методы, которые как раз и были развиты создателями КАМ-теории — А. Н. Колмогоровым, В. И. Арнольдом и Ю. Мозером. КАМ-теория позволяет построить не только сходящуюся процедуру разложения, но и является важным пунктом в понимании природы возникновения хаоса. Кроме того, эта теория и ее следствия оказались очень важными для многих разделов современной науки (чистой и прикладной математики, механики, физики и даже численного анализа!).

За последние годы в теории КАМ достигнут значительный прогресс, и она продолжает быть весьма активной областью исследований (см. [30, 184, 205] и данные там ссылки). Ниже мы представим основные положения этой теории и опишем некоторые важные ее следствия.

16.1 Теорема Колмогорова

Рассмотрим невозмущенную полностью интегрируемую систему с гамильтонианом $H = H_0(\mathbf{J})$. В переменных действие–угол ее динамика описывается уравнениями

$$\begin{aligned} \dot{\alpha}_i &= \frac{\partial H}{\partial J_i}, \\ \dot{J}_i &= 0, \quad i = 1, 2, \dots, n. \end{aligned} \quad (2.44)$$

Фазовое пространство такой системы $2n$ -мерно и представляет собой совокупность n -мерных инвариантных торов. Что произойдет, если немного возмутить систему (2.44), так что $H = H_0(\mathbf{J}) + \epsilon H_1(\boldsymbol{\alpha}, \mathbf{J})$, $\epsilon \ll 1$? Сохранятся или разрушатся в этом случае торы? Казалось бы, возмущение изменит гамильтониан, что позволит фазовым кривым покинуть тор и свободно блуждать по пространству уровня энергии. Например, для задачи о двух гармонических осцилляторах (см. §14) возмущение нелинейным образом свяжет их и, как следствие, разрушит торы.

Однако эти рассуждения не являются верными. Теория КАМ утверждает, что при достаточно малых ϵ большинство нерезонансных торов сохраняется и лишь немного деформируется. «Большинство» в данном случае

означает, что все резонансные торы (отвечающие периодическому движению) и часть нерезонансных торов разрушатся, но это множество будет мало по сравнению с множеством сохранившихся в результате возмущения нерезонансных торов.

Условия применимости КАМ-теории следующие.

- Невозмущенный гамильтониан должен удовлетворять условию невырожденности:

$$\det \left| \frac{\partial \omega_i}{\partial J_k} \right| = \det \left| \frac{\partial^2 H_0}{\partial J_i \partial J_k} \right| \neq 0, \quad i = 1, 2, \dots, n.$$

Это означает, что частоты невозмущенной системы функционально независимы.

- Возмущение должно быть гладким, т. е. необходимо, чтобы гамильтониан H_1 имел достаточное число непрерывных производных.
- Система должна находиться вне окрестности резонанса, т. е.

$$\left| \sum_j k_j \omega_j \right| > c |\mathbf{k}|^{-r}, \quad \mathbf{k} = (k_1, k_2, \dots, k_n), \quad (2.45)$$

где r зависит от числа степеней свободы n , а постоянная c определяется величиной возмущения ϵH_1 и параметром нелинейности γ (см. §15). Заметим, что из этого неравенства можно получить условие умеренной нелинейности (2.40).

При выполнении приведенных требований смысл теории КАМ состоит в следующем. При достаточно малых $\epsilon < \epsilon_0$ энергетическую гиперповерхность $H = E$ системы можно разбить на две области ненулевого объема. Большая из них содержит деформированные нерезонансные торы невозмущенной задачи, а меньшей области (объем которой стремится к нулю при $\epsilon \rightarrow 0$) движение оказывается очень сложным; оно отлично как от периодического, так и от квазипериодического типов движений.

Иными словами, для большинства начальных условий в системе, близкой к интегрируемой, квазипериодическая динамика сохраняется. Имеются, однако, начальные условия, для которых существовавшие при $\epsilon = 0$ (главным образом резонансные) торы разрушатся и движение станет нерегулярным. Именно эти разрушающиеся под действием возмущения торы и приводят систему к хаосу. Вот почему теория нелинейного резонанса, кратко

представленная в §15, играет столь важную роль. Траектории, начавшиеся в области разрушенных торов, могут свободно блуждать по энергетической поверхности. Это проявляется на поверхности сечения как множество случайно разбросанных точек.

Заметим, что в большинстве случаев при общем анализе гамильтоновых систем получаются заниженные значения параметра ϵ_0 по сравнению с реальной величиной, точное определение которой при исследовании конкретного гамильтониана вида (2.12) является достаточно сложной задачей и обычно проводится численно.

Интересно, что при малых ϵ далекие от резонанса торы сохраняются для произвольных гладких возмущений. Но с ростом ϵ картина качественно меняется: торы начинают разрушаться, а область хаотичности — расти. В конце концов это приводит к перекрытию первичных резонансов (см. §15) и возникновению явления, которое в гамильтоновой динамике называется *глобальным хаосом* [116]. При этом в системе не остается ни одного из тех торов, что были близки к торам невозмущенной задачи, хотя могут появиться и другие торы. Фазовая траектория в такой системе способна двигаться поперек хаотических слоев (см. §17).

Однако при $\epsilon \ll 1$ резонансы не перекрываются и решения лежат на слегка деформированных инвариантных торах. Здесь имеется качественное различие в динамике систем с двумя степенями свободы $n = 2$ и систем с числом степеней свободы $n > 2$.

16.2 Диффузия Арнольда

В случае двух степеней свободы фазовое пространство является четырехмерным, энергетическая гиперповерхность (или пространство уровня энергии) $H = E$ — трехмерной, а инвариантные торы — двумерными. Это позволяет представить такие торы «погруженными» в трехмерное пространство уровня энергии $H = E$ (рис. 2.16). Следовательно, они будут разделять его на непересекающиеся области. Тем самым, разрушенные торы оказываются зажатыми между сохранившимися после возмущения торами (рис. 2.17). Фазовая траектория, начавшаяся в области такого разрушенного тора (т. е. в щели между двумя инвариантными торами), поскольку траектории не могут пересекаться, останется запертой здесь навсегда. Это означает, что соответствующие переменные действия практически не меняются и при движении остаются вблизи своих начальных значений. Поэтому для систем с двумя степенями свободы, удовлетворяющих условию КАМ-теории, при всех начальных условиях эволюция отсутствует и *сохраняется глобальная устойчивость* [23, 30].

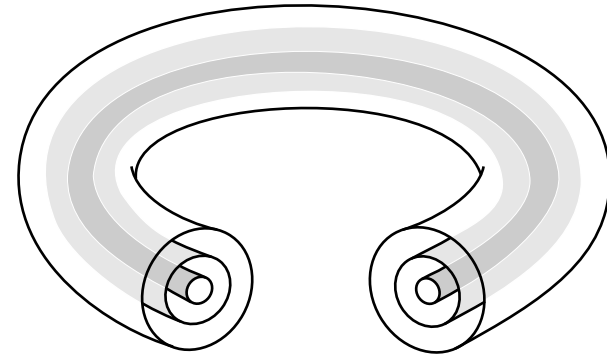


Рис. 2.16. Семейство торов полностью интегрируемой системы с двумя степенями свободы

При $n > 2$ инвариантные торы уже не делят $(2n - 1)$ -мерную энергетическую гиперповерхность на непересекающиеся части, и области разрушенных торов сливаются, образуя сложную единую сеть — т. н. хаотическую паутину Арнольда. Двигаясь по нитям этой паутины, фазовая точка на гиперповерхности заданной энергии может сколь угодно близко подойти к любой его точке или удалиться на значительное расстояние от своего первоначального положения. Это случайное блуждание по резонансам вокруг инвариантных торов называется *диффузией Арнольда* [22, 23, 30, 116].

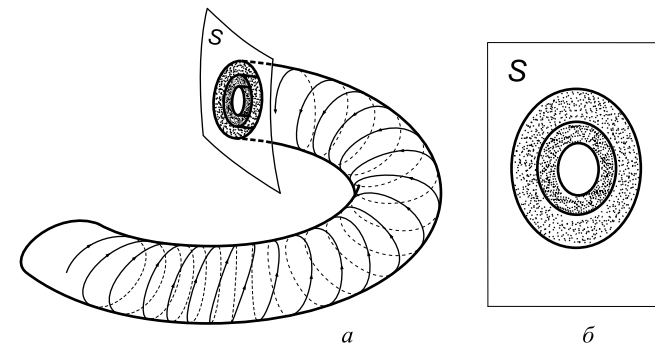


Рис. 2.17. Сохранение глобальной устойчивости системы с двумя степенями свободы, близкой к интегрируемой: *a* — разрушенные торы зажаты в щелях между сохранившимися и слегка деформированными торами; *б* — соответствующая картина в отображении Пуанкаре

Таким образом, малые возмущения в системах с более чем двумя степенями свободы могут не только качественно преобразовать динамику системы, но и изменить топологию фазовых траекторий, преобразовав их в связанную сеть.

Характерной особенностью диффузии Арнольда является ее универсальность, т. е. не существует критического значения возмущения ϵ , необходимого для ее возникновения. Иными словами, диффузия имеет место всегда, даже при сколь угодно малых ϵ . При этом, очевидно, скорость диффузии стремится к нулю при уменьшении возмущения. Следовательно, для гамильтоновых систем число степеней свободы которых $n > 2$, переменные действия могут медленно эволюционировать, что соответствует отсутствию глобальной устойчивости. Эта эволюция в общем случае происходит достаточно медленно и может быть различной в разных частях фазового пространства.

Еще одно свойство систем, которым свойственна диффузия Арнольда, состоит в том, что в их динамике не может наблюдаться резкого перехода к глобальному хаосу, возникающему вследствие перекрытия резонансов. Это объясняется тем, что в таких системах области с хаотическим поведением уже связаны между собой в единую сеть. При этом диффузия Арнольда является очень медленной по сравнению с движением траекторий в области глобальной хаотичности.

Поясним формально, почему инвариантные торы делят пространство уровня энергии только при $n = 2$ [523]. Рассмотрим сначала некоторую k -мерную область. Эта область разобьется на непересекающиеся части только в том случае, если размерность разделяющей ее гиперповерхности равна $k - 1$. Для автономных гамильтоновых систем размерность пространства D уровня энергии равна $2n - 1$. Следовательно, размерность границ, которые бы делили D на непересекающиеся области, должна быть равна $2n - 2$. Для того, чтобы инвариантные торы разделяли пространство D , их размерность n должна удовлетворять условию $n \geq 2n - 2$, или $n \leq 2$.

Рассмотрим подробнее, как же происходит «разрушение» резонансных торов и рождается хаос.

17. Природа хаоса

Вопрос о разрушении торов с резонансными частотами и рождении хаоса восходит к знаменитой теореме о неподвижной точке Пуанкаре–Биркгофа [48, 458]. Несмотря на некоторую абстракцию, этот результат имеет непосредственное отношение к движению гамильтоновых систем,

т. к. помогает установить существование периодических решений. Это следует из того, что неподвижные точки отображения Пуанкаре соответствуют периодическим траекториям систем дифференциальных уравнений.

17.1 Отображение поворота

Будем рассматривать систему с двумя степенями свободы $n = 2$ и гамильтонианом $H(q_1, p_1, q_2, p_2)$, хотя изложенные ниже результаты можно обобщить и на многомерный случай. Данное упрощение позволит получить наглядное представление о всей сложности явлений, происходящих в гамильтоновых системах. В такой системе величина H является полной энергией и потому постоянна. Следовательно, всегда при данном значении $H = E$ поток (т. е. совокупность решений) будет *трехмерным*. Это дает возможность рассмотреть вместо непрерывной эволюции системы соответствующее отображение Пуанкаре. Таким образом, утверждение о существовании инвариантных торов будет эквивалентно утверждению о том, что в отображении Пуанкаре множество точек образует инвариантную замкнутую кривую. В резонансном случае, очевидно, вместо замкнутой кривой получится конечный набор точек, последовательно переходящих друг в друга.

Рассмотрим, как аналитически можно построить это отображение. Для интегрируемой системы в переменных действие–угол фазовое пространство представляет собой множество вложенных торов. Запишем поток на одном из них:

$$\alpha_1(t) = \omega_1 t + \alpha_1(0),$$

$$\alpha_2(t) = \omega_2 t + \alpha_2(0),$$

где $\omega_1 = \omega_1(J_1, J_2) = \partial H / \partial J_1$ и $\omega_2 = \omega_2(J_1, J_2) = \partial H / \partial J_2$. Очевидно, полный оборот вдоль координаты α_2 происходит за время $t_2 = 2\pi / \omega_2$. В течение этого промежутка изменение переменной α_1 составит

$$\alpha_1(t + t_2) = \alpha_1(t) + \omega_1 t_2 = \alpha_1(t) + \frac{2\pi\omega_1}{\omega_2} = \alpha_1(t) + 2\pi\rho(J_1).$$

Введенная здесь величина $\rho = \omega_1 / \omega_2$ называется *числом вращения*. Поскольку движение происходит в пространстве уровня энергии, то $J_2 = J_2(J_1, E)$. Поэтому при заданном E число ρ является функцией только переменной J_1 .

Допустим теперь, что в качестве секущей поверхности выступает плоскость (α_1, J_1) , т. е. $\alpha_2 = \text{const}$. Тогда точки x_k пересечений фазовой траектории с этой плоскостью можно записать как $x_k = (\alpha_1(t + kt_2), J_1)$

(рис. 2.18а). Стало быть, решения исходной системы на данном торе можно представить как преобразование P_0 , которое переводит одну точку пересечения x_k в следующую за ней точку x_{k+1} по инвариантной окружности «радиуса» J_1 . Вводя обозначение $J_k \equiv J_1(t+kt_2)$, такое перемещение точек можно записать как

$$P_0 : \begin{cases} \alpha_{k+1} = \alpha_k + 2\pi\rho(J_k), \\ J_{k+1} = J_k. \end{cases} \quad (2.46)$$

Как видно из этого соотношения, число вращения ρ , вообще говоря, зависит от радиуса окружности.

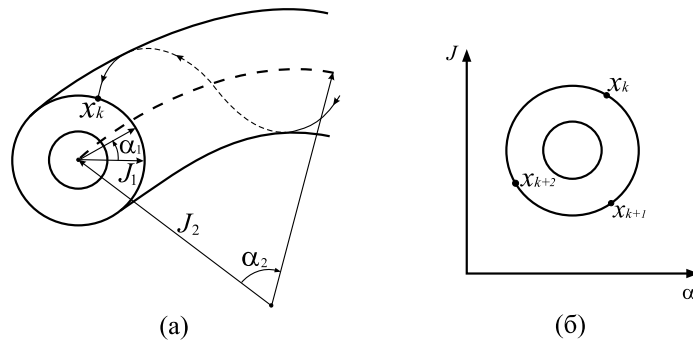


Рис. 2.18. Система с двумя степенями свободы (а) и ее отображение Пуанкаре (б)

Когда ρ иррационально, то при $k \rightarrow \infty$ точки x_k заполняют всю окружность. При рациональном $\rho = l/m$ они будут через каждые m шагов последовательно переходить друг в друга (рис. 2.18б). Поэтому в таком представлении можно говорить о нерезонансных и резонансных *окружностях*. Обозначим такие окружности как \mathcal{C} .

При переходе от окружности к окружности число вращения меняется. Предположим для определенности, что $\rho(J)$ возрастает с увеличением J . Тогда, по мере удаления от начала координат, угол, на который поворачивается окружность, в среднем будет возрастать. Это приводит к закручиванию радиальной линии точек под воздействием P_0 (рис. 2.19). Поэтому преобразование (2.46) называется *закручивающим отображением* (или иногда *отображением поворота*). Поскольку это отображение переводит точки окружности \mathcal{C} в некоторые точки *той же окружности* \mathcal{C} , то говорят, что оно преобразует (отображает) *окружность в себя*. Поэтому, обобщая соот-

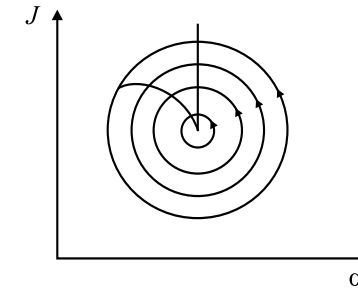


Рис. 2.19. Действие невозмущенного закручивающего отображения

ношение (2.46), можно записать:

$$P_0(\mathcal{C}) = \mathcal{C}.$$

Иными словами, окружность \mathcal{C} является инвариантной относительно преобразования P_0 . Важным свойством отображения P_0 , которое будет использовано ниже, является его консервативность, т. е. фазовый объем под действием P_0 сохраняется.

Рассмотрим теперь возмущенную систему (2.13) с гамильтонианом (2.12). На поверхности сечения учет возмущения соответствует введению в закручивающее отображение дополнительных слагаемых:

$$P_\epsilon : \begin{cases} \alpha_{k+1} = \alpha_k + 2\pi\rho(J_k) + \epsilon f(J_k, \alpha_k), \\ J_{k+1} = J_k + \epsilon g(J_k, \alpha_k), \end{cases} \quad (2.47)$$

где f и g — периодические функции по α . При этом преобразование P_ϵ должно по-прежнему сохранять площадь, иначе инвариантные окружности, как правило, не существуют. Согласно теории КАМ (см. §16), под действием малого возмущения, $\epsilon \ll 1$, большинство окружностей с иррациональным ρ сохраняются и лишь немного деформируются. Рассмотрим, что происходит с окружностями в области *рациональных* значений $\rho = l/m$, т. е. резонанса, где соотношение (2.45) не выполняется. Именно в этой области, как уже отмечалось в §16, рождается хаотическое движение.

17.2 Теорема о неподвижной точке

Вернемся пока к интегрируемому случаю. При $\rho(J) = l/m$ любая точка на окружности через m шагов перейдет в свое первоначальное положение, т. е. будет *периодической* с периодом m . Обозначим эту резонансную

окружность как C и рассмотрим две нерезонансные инвариантные окружности C^+ и C^- , лежащие по обе стороны от нее (рис. 2.20а). Поскольку $\rho(J)$ возрастает с J (см. рис. 2.19), то для таких окружностей их иррациональные числа вращения будут удовлетворять соотношениям $\rho > l/m$ и $\rho < l/m$, соответственно. После действия отображения P_0 подряд m раз (обозначим такое действие как P_0^m) точки окружности C^+ повернутся на угол, больший, чем 2π , а точки окружности C^- — на угол, меньший, чем 2π . Поэтому относительно C кажется, что отображение закручивает C^+ против, а C^- — по часовой стрелке (см. рис. 2.20а).

Применение m раз отображения P_0 означает, что через m шагов начальная точка вернется в свое исходное положение. Поэтому такая точка называется периодической с периодом m . При $m = 1$ отображение P_0 оставляет точку на месте. В этом случае говорят, что точка является неподвижной. Соответственно, для P_0^m периодическая точка периода m будет неподвижной.

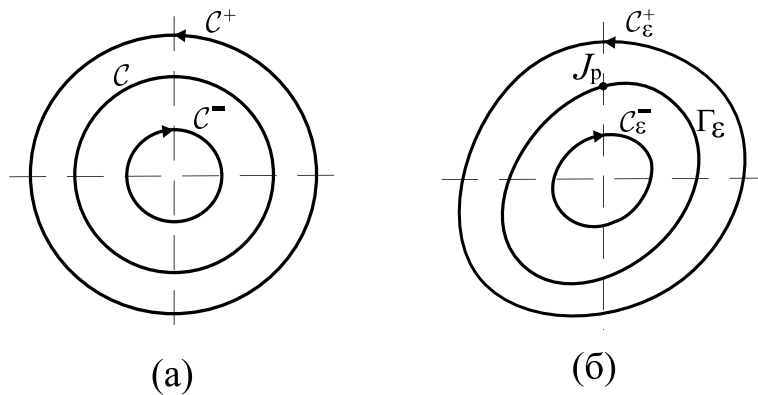


Рис. 2.20. Инвариантные окружности C , C^+ и C^- невозмущенного закручивающего отображения P_0 при $\rho = l/m$, $\rho > l/m$ и $\rho < l/m$, соответственно (а). Результат действия возмущения P_ϵ (б)

Рассмотрим теперь возмущенное отображение (2.47). Согласно теории КАМ, окружности C^+ и C^- при действии возмущения сохраняются и только немного деформируются. Обозначим такие замкнутые кривые как C_ϵ^+ и C_ϵ^- , соответственно. Они инвариантны относительно преобразования P_ϵ :

$$P_\epsilon(C_\epsilon^+) = C_\epsilon^+, \quad P_\epsilon(C_\epsilon^-) = C_\epsilon^-.$$

Допустим, что параметр ϵ достаточно мал, так чтобы относительный пово-

рот C_ϵ^- и C_ϵ^+ сохранялся при действии P_ϵ^m . Тогда на каждом радиусе $\alpha = \text{const}$ между кривыми C_ϵ^- и C_ϵ^+ найдется некоторая точка $J_p(\alpha, \epsilon)$, угловые координаты которой поворачиваются отображением P_ϵ^m ровно на 2π . Следовательно, для этой точки угловые координаты при действии P_ϵ^m сохраняются. На каждом из радиусов, проведенных из центра, лежит одна такая точка. А так как возмущение P_ϵ гладко, то эти точки $J_p(\alpha, \epsilon)$ образуют некоторую замкнутую кривую Γ_ϵ (рис. 2.20б), которая стягивается к C при $\epsilon \rightarrow 0$. Эта кривая Γ_ϵ не будет инвариантной относительно преобразования P_ϵ . Действие P_ϵ^m состоит в смещении каждой точки кривой Γ_ϵ в радиальном направлении. Значит, вместо Γ_ϵ образуется новая кривая $P_\epsilon^m(\Gamma_\epsilon)$ (рис. 2.21а).

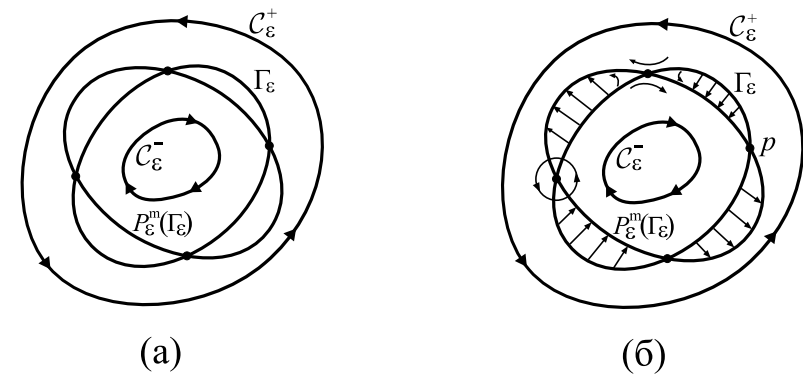


Рис. 2.21. Преобразование кривой Γ_ϵ в кривую $P_\epsilon(\Gamma_\epsilon)$ при действии отображения P_ϵ (а) и возникающие при этом эллиптические и гиперболические точки (б)

Далее вспомним, что отображение P_ϵ консервативно. Стало быть, обе кривые, Γ_ϵ и $P_\epsilon^m(\Gamma_\epsilon)$, должны ограничивать одинаковые площади. Поэтому кривая $P_\epsilon^m(\Gamma_\epsilon)$ не может находиться ни внутри, ни снаружи кривой Γ_ϵ . Следовательно, в общем случае они должны пересекаться³ в четном числе точек (рис. 2.21а). В свою очередь, каждая из таких точек пересечения будет неподвижной для возмущенного преобразования P_ϵ^m .

В этом состоит основной смысл теоремы Пуанкаре–Биркгофа о неподвижной точке [48, 458], согласно которой возмущенное закручивающее отображение (2.47) с числом вращения $\rho = l/m$ имеет $2im$, $i = 1, 2, \dots$, неподвижных точек. Таким образом, в результате возмущения резонансной

³Мы не рассматриваем возможность касания, т. к. это исключительная ситуация.

окружности (для нее, напомним, каждая точка является неподвижной для преобразования P_0^m) сохраняется лишь четное число $2im$ неподвижных точек.

Рассмотрим одну из точек p пересечения кривых Γ_ϵ и $P_\epsilon^m(\Gamma_\epsilon)$ (см. рис. 2.21б). Для отображения P_ϵ^m она будет неподвижной. Преобразование же P_ϵ , действуя на p , порождает последовательность точек $p, P_\epsilon p, P_\epsilon^2 p, \dots, P_\epsilon^{m-1} p$, т. е. после m итераций точка p возвратится в исходное положение. Но каждая из них будет неподвижной точкой для P_ϵ^m . Следовательно, имеется m неподвижных точек, которые связаны с исходной точкой p . В силу того что пересечение кривых Γ_ϵ и $P_\epsilon^m(\Gamma_\epsilon)$ происходит в четном числе точек, всего получается $2im$ неподвижных точек.

17.3 Эллиптические и гиперболические точки

Если детально рассмотреть действие отображения P_ϵ в окрестности различных точек p , то можно заметить качественную разницу. Вблизи одних соседние точки остаются рядом с p , как бы вращаясь вокруг них, вблизи других — стремятся покинуть окрестность p . При этом такие точки чередуются (рис. 2.21б). Движение подобного рода наблюдается в фазовом пространстве нелинейного маятника (см. рис. 2.2). Поэтому такие точки называются эллиптическими и гиперболическими, соответственно. Эллиптические точки окружены семейством замкнутых траекторий, инвариантных относительно P_ϵ^m , которые образуют некие «островки», тогда как гиперболические соединяются посредством сепаратрис. Эта картина является типичной для нелинейных систем с малым возмущением и всегда возникает вблизи резонанса.

Каждый островок удовлетворяет КАМ-теории. Следовательно, большинство нерезонансных окружностей сохранится. Однако здесь существуют и резонансы. В окрестности каждого из них в соответствии с теоремой Пуанкаре–Биркгофа о неподвижной точке появится череда эллиптических и гиперболических точек, но в более мелком масштабе (рис. 2.22). В свою очередь, островок вблизи каждой из этих эллиптических точек в миниатюре будет воспроизводить всю структуру в целом. Замкнутые кривые в окрестности эллиптических точек соответствуют меньшим тором. Некоторые из этих мелких торов будут, согласно КАМ-теории, сохраняться, другие вследствие резонансов более высоких порядков — разрушатся на меньшие. И так до бесконечности. Таким образом, резонанс приводит к очень сложной картине, которая при любом изменении масштаба повторяется, т. е. является самоподобной.

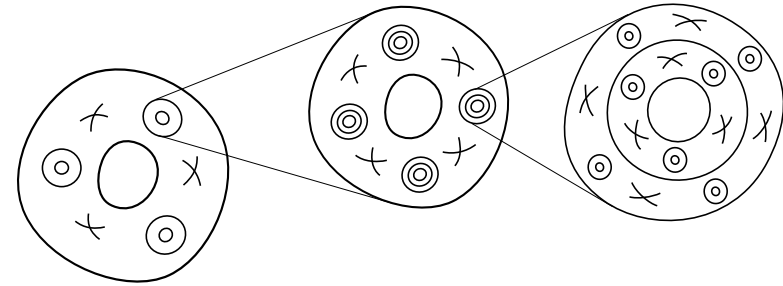


Рис. 2.22. Распад торов с рациональным отношением частот и возникновение самоподобной структуры

Рассмотрим теперь, что происходит в окрестности гиперболических точек, следуя изложению [26, 234]. Каждая из них характеризуется четырьмя инвариантными направлениями (или ветвями сепаратрисы): двумя устойчивыми (w^s), которые входят в гиперболическую точку \mathcal{H} , и двумя неустойчивыми (w^u), которые выходят из \mathcal{H} (рис. 2.23). Поскольку мы рассматриваем отображение Пуанкаре, то в исходном фазовом пространстве эти кривые будут соответствовать поверхностям или инвариантным устойчивым W^s и неустойчивым W^u многообразиям гиперболической точки.⁴

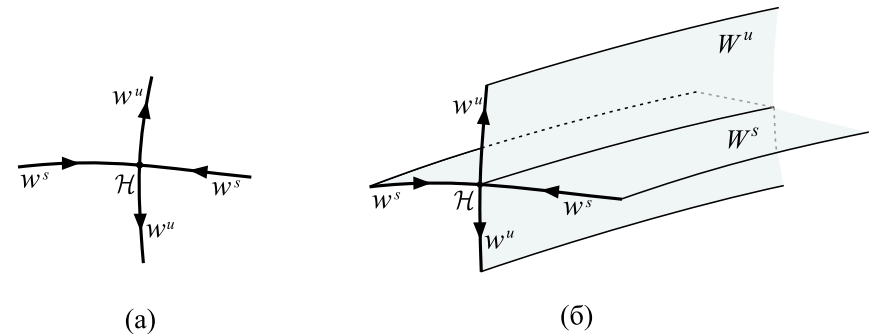


Рис. 2.23. Устойчивые (w^s) и неустойчивое (w^u) направления гиперболической точки \mathcal{H} в отображении Пуанкаре (а) и соответствующие многообразия W^s и W^u в фазовом пространстве (б)

⁴Индексы s и u происходят от английских слов *stable* — устойчивый и *unstable* — неустойчивый.

На сепаратрисе период колебаний бесконечен (см. §14). Поэтому, если некоторая точка q принадлежит устойчивой ветви w^s , то под действием отображения P она будет экспоненциально медленно приближаться к \mathcal{H} , т. е.

$$\lim_{k \rightarrow \infty} P^k q \rightarrow \mathcal{H}.$$

Если же $q \in w^u$, то точка q будет экспоненциально медленно стремиться покинуть окрестность \mathcal{H} , т. е.

$$\lim_{k \rightarrow \infty} P^{-k} q \rightarrow \mathcal{H}.$$

17.4 Расщепление сепаратрис

В случае интегрируемой системы устойчивые и неустойчивые многообразия гиперболических точек могут соединяться друг с другом, образуя гладкие структуры. В отображении Пуанкаре такое строение выглядит как плавный переход от неустойчивой ветви сепаратрисы в устойчивую. При этом они могут замыкаться на одну гиперболическую точку (рис. 2.24а) или, как мы видели выше (см. §15), соединять несколько таких точек, образуя «гирлянды» (рис. 2.24б). В первом случае устойчивая w^s и неустойчивая w^u ветви сепаратрисы образуют петлю, которая называется *гомоклинической кривой*. Она представляет собой двоякоасимптотическую траекторию, которая характеризуется тем, что любая точка q , находящаяся на такой петле, всегда стремится к \mathcal{H} , т. е. $\lim_{k \rightarrow \pm\infty} P^k q \rightarrow \mathcal{H}$. Внутри гомоклинической кривой существует эллиптическая точка.

Во втором случае кривые, состоящие из устойчивых и неустойчивых ветвей сепаратрисы, называются *гетероклиническими траекториями*. Двигаясь по таким траекториям, мы экспоненциально медленно удаляемся от одной гиперболической точки и приближаемся к другой.

В случае возмущения ветви сепаратрисы уже не образуют такие гладкие гомоклинические и гетероклинические соединения, а могут пересекаться. Если точка пересечения образована устойчивой и неустойчивой ветвями сепаратрисы одного и того же резонанса, то она называется *гомоклинической точкой* (рис. 2.25а). Если же пересекаются устойчивая и неустойчивая ветви от разных гиперболических точек (резонансов), то образуется *гетероклиническая точка* (рис. 2.25б). Рассмотрим, как эволюционируют такие точки при действии отображения P_ϵ .

Пусть q — гомоклиническая точка, $q' \in w^s$ и $q'' \in w^u$ — соседние с ней точки (см. рис. 2.26а). После действия P_ϵ эти две точки должны отображаться в точки $P_\epsilon q'$ и $P_\epsilon q''$. Где должна находиться точка $P_\epsilon q$, в которую

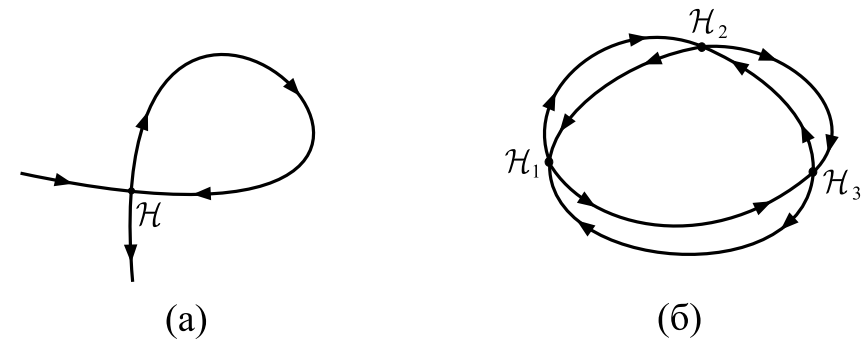


Рис. 2.24. Гомоклиническая (а) и гетероклиническая (б) траектории, образованные ветвями сепаратрисы

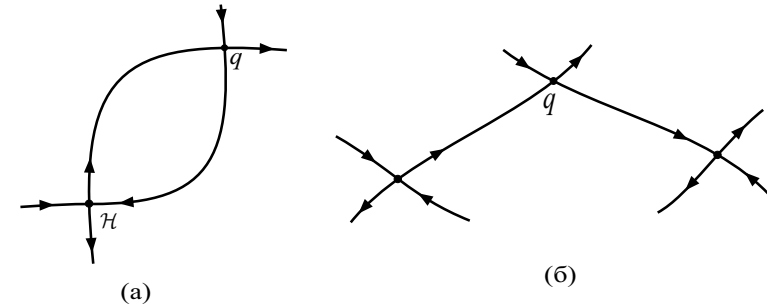


Рис. 2.25. Гомоклиническая (а) и гетероклиническая (б) точки

перейдет q после действия P_ϵ ? Как видно на рис. 2.26а, следуя направлениям движения по устойчивой и неустойчивой ветвям, q расположена перед точками q' и q'' . Поэтому, в силу непрерывности отображения P_ϵ , она должна перейти в точку, которая также будет расположена одновременно перед точками $P_\epsilon q'$ и $P_\epsilon q''$. Значит, должно образоваться новое пересечение, т. е. новая гомоклиническая точка $P_\epsilon q$ (рис. 2.26б).

Иными словами, поскольку ветви w^s и w^u являются инвариантными, то после действия отображения мы должны остаться как на одной так и на другой ветвях сепаратрисы. Следовательно, появляется новая точка пересечения $P_\epsilon q$. При этом между точками q и $P_\epsilon q$ образуется петля.

Рассуждая аналогичным образом, легко видеть, что точка $P_\epsilon q$ отображится в точку $P_\epsilon^2 q$, формируя другую петлю (рис. 2.27). Поскольку новая точка $P_\epsilon^2 q$ оказывается расположенной ближе к гиперболической точке \mathcal{H} ,

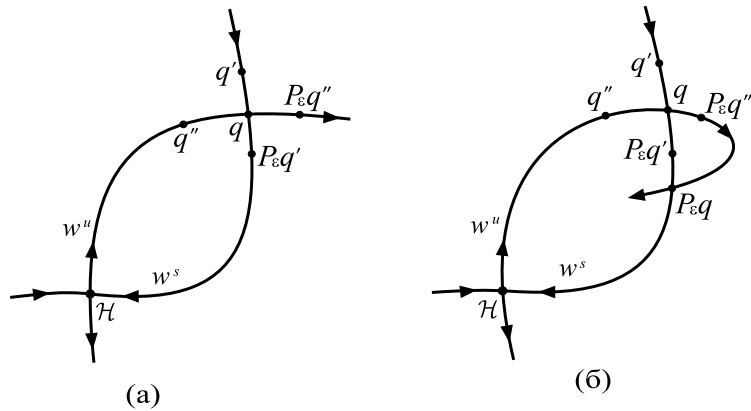


Рис. 2.26. Отображение соседних точек q' и q''

то расстояние между $P_\epsilon^2 q$ и $P_\epsilon q$ будет меньше, чем расстояние между $P_\epsilon q$ и q . Вспомним, что одним из свойств отображения P_ϵ является сохранение фазового объема. Это означает, что площади, ограниченные петлями между q , $P_\epsilon q$ и $P_\epsilon^2 q$ должны быть равны. Следовательно, вторая петля будет более вытянутой и изогнутой, чем первая.

Продолжая рассуждения, мы приходим к выводу, что в результате появляется бесконечное множество пересечений ветвей сепаратрисы, которые все более и более сближаются, а появляющиеся при этом петли становятся все более длинными и тонкими (рис. 2.28а). Подобным же образом ведет себя и устойчивая ветвь сепаратрисы w^s (рис. 2.28б). Это нетрудно понять, если двигаться по w^s в противоположном направлении. Таким образом,

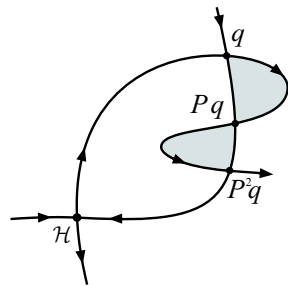


Рис. 2.27. Образование петель из сепаратрис

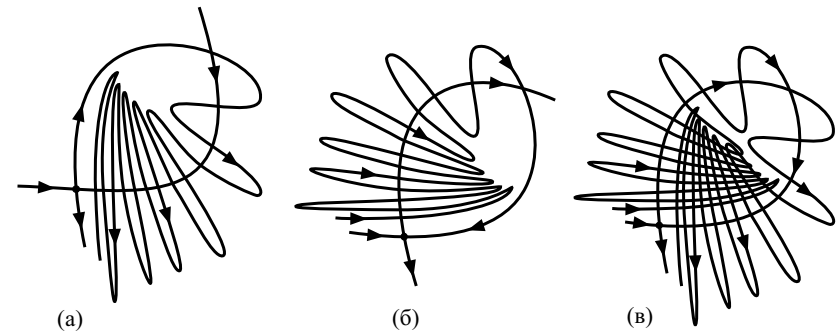


Рис. 2.28. Гомоклинические сплетения в окрестности гиперболической точки

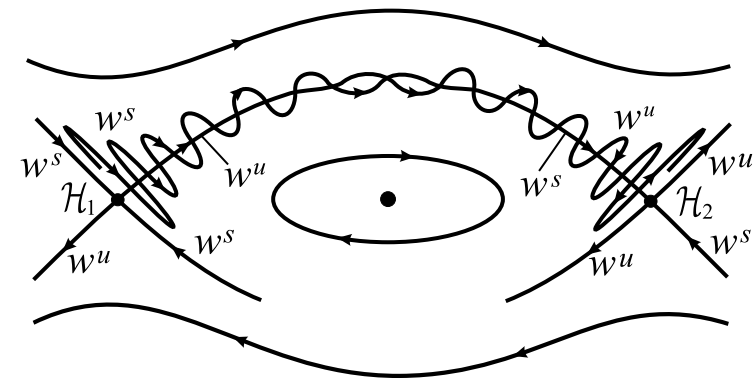


Рис. 2.29. Образование гетероклинических сплетений

в целом картина в окрестности гиперболических точек оказывается чрезвычайно сложной (рис. 2.28в).

Для гетероклинических траекторий в результате пересечения сепаратрис образуются несколько иные структуры. Неустойчивая ветвь w^u , выходящая из гиперболической точки \mathcal{H}_1 , осциллирует, приближаясь к гиперболической точке \mathcal{H}_2 , расположенной справа. Устойчивая же ветвь w^s осциллирует, когда удаляется от \mathcal{H}_1 (рис. 2.29). То же самое происходит с устойчивой и неустойчивой ветвями, расположенными в нижней части рисунка (для удобства они не показаны).

Таким образом, возмущение, вносимое в интегрируемую систему, приводит к тому, что сепаратрисы уже не гладкие, как на рис. 2.24, а сложным образом расщепляются. Поэтому такое явление называется *расщепле-*

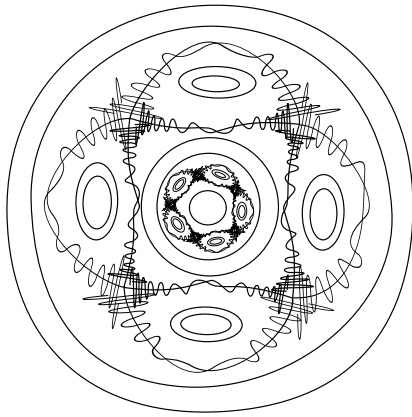


Рис. 2.30. Сечение фазового пространства гамильтоновых систем с малым возмущением

нием *сепаратрис*. При этом структуры, образованные ветвями сепаратрисы в области гиперболических точек, называются *гомоклиническими* и *гетероклиническими сплетениями*. Именно такое сложное поведение сепаратрис и является причиной возникновения хаоса в детерминированных системах. На рис. 2.10 видно, что хаотические траектории образуются как раз вблизи гиперболических точек.

В области гомоклинических сплетений не могут существовать инвариантные торы. Это означает, что здесь системы оказываются неинтегрируемыми и, как следствие, обладающими хаотическими свойствами. Однако при других начальных условиях из окрестности сохранившихся торов динамика систем будет регулярной.

Теперь можно представить всю картину в целом, возникающую в фазовом пространстве и на поверхности сечения множества инвариантных торов. Разрушение резонансных окружностей вследствие возмущения сопровождается появлением *2im* гиперболических и эллиптических точек. В окрестности каждой эллиптической точки имеется семейство замкнутых инвариантных кривых, часть из которых тоже разрушается из-за возмущения. Это приводит к появлению более мелкой цепи из эллиптических и гиперболических точек. В области каждой из гиперболических точек сепаратрисы расщепляются и возникают их сплетения (рис. 2.30).

Однако вспомним, что все это происходит не на поверхности сечения, а в фазовом пространстве, образованном совокупностью торов. Таким образом, общая картина движения фазовых траекторий получается очень

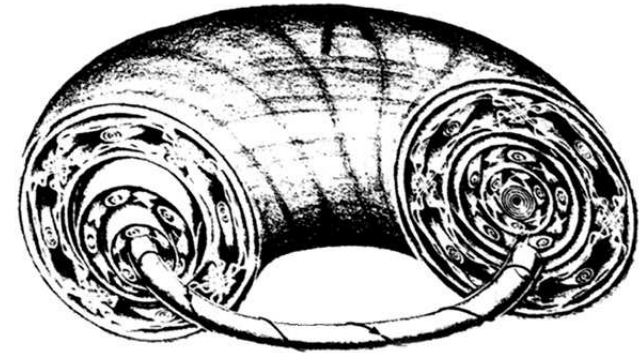


Рис. 2.31. Структура фазового пространства возмущенных гамильтоновых систем с двумя степенями свободы [346]

сложной (рис. 2.31). Такая структура повторяется на все меньшем и меньшем масштабе и характерна для систем, близких к интегрируемым. В этом многообразии движений и состоит причина их хаотического поведения.

С увеличением возмущения разрушаются также и некоторые нерезонансные окружности. Это особенно четко видно на примере задачи Эно-Эйлеса (рис. 2.9). Но если возмущение H_1 мало, то хаотические траектории будут существовать только в ограниченной инвариантными кривыми области фазового пространства.

Впервые доказательство существования гомоклинических точек и сепаратрис, образующих сплетения, было получено А. Пуанкаре при изучении задачи трех тел [167]. Он достаточно образно описал эти структуры в заключительном томе «Новых методов...»: «Если попытаться представить себе фигуру, образованную этими двумя кривыми и их бесчисленными пересечениями, каждое из которых соответствует двоякоасимптотическому решению, то эти пересечения образуют нечто вроде решетки, ткани, сети с бесконечно тесными петлями; ни одна из двух кривых никогда не должна пересечь саму себя, но она должна навиваться на саму себя очень сложным образом, чтобы пересечь бесконечно много раз все петли сети.»

Таким образом, КАМ-теория и теорема о неподвижной точке выявляют всю сложность и многообразие динамики нелинейных гамильтоновых систем с внешними возмущениями. С количественным критерием, позволяющим оценить расстояние между сепаратрисами и получить условия возникновения гомо- и гетероклинического хаоса, мы познакомимся в §21.5.

18. Основные свойства хаотических систем: эргодичность, перемешивание, расщепление корреляций

Для дальнейшего изучения эволюции систем типа (2.1) в фазовом пространстве необходимо ввести некоторые новые понятия и определения.

Рассмотрим динамическую систему, задаваемую обыкновенными дифференциальными уравнениями

$$\dot{x} = f(x), \quad (2.48)$$

где $f = \{f_1, \dots, f_n\}$ — векторная функция, обычно предполагаемая гладкой, а $x = \{x_1, \dots, x_n\}$ — n -мерный вектор с компонентами x_1, \dots, x_n , характеризующий состояние динамической системы. Функция $F(t, c_i)$, $c_i = \text{const}$, $i = 1, 2, \dots, n$, которая при подстановке в уравнения (2.48) обращает их в тождества, называется решением системы. Задание начальных условий $x(0) = x_0$ однозначно определяет решение в любой момент времени t :

$$x(t) = F(t, x_0). \quad (2.49)$$

Если уравнения (2.48) описывают динамику гамильтоновой системы, то в этом случае компоненты x_i вектора $x = \{x_1, \dots, x_n\}$ обозначают совокупность канонических переменных $x_j \equiv q_j$, $x_k \equiv p_k$, $k, j = 1, 2, \dots, n/2$.

Всякое решение $x(t)$ (2.49) системы (2.48) геометрически можно представить как линию в n -мерном пространстве, образованном переменными x_1, \dots, x_n . Это n -мерное пространство называется *фазовым пространством* системы; обозначим его M . Каждому состоянию динамической системы соответствует точка в этом пространстве, и каждой точке из пространства M соответствует единственное состояние системы. Изменение состояния системы можно интерпретировать как движение некоторой точки (называемой изображающей) в фазовом пространстве. Траектория такой изображающей точки, т. е. ее последовательные положения в фазовом пространстве M , называется *фазовой траекторией*.

Предположим, что в момент времени t_0 система (2.48) с фазовым пространством M находилась в состоянии x_0 . Тогда в другой момент времени $t \neq t_0$ ее состояние, вообще говоря, будет иным. Обозначим это новое состояние как $F^t x_0$. Таким образом, для каждого t мы определили оператор эволюции, или отображение сдвига, $F^t: M \rightarrow M$ фазового пространства M в себя. Это отображение F^t переводит систему из состояния, в котором она находилась в момент времени t_0 , в состояние в момент времени t . Другими словами, решение (2.49) уравнений (2.48) ставит в соответствие точке

$x_i(t_0)$, $i = 1, 2, \dots, n$, фазового пространства M в момент времени t_0 определенную точку $x_i(t)$ фазового пространства в момент t :

$$F^t x_0 = x(t). \quad (2.50)$$

Следовательно, любая область Ω_0 фазового пространства под действием отображения F^t перейдет через время t (если, конечно, F^t — не тождественное отображение) в некоторую другую область $\Omega_t = F^t \Omega_0$. Схематически это изображено на рис. 2.32. Отображение $F^t: M \rightarrow M$ называют также *фазовым потоком*, а функцию $f(x)$ — *векторным полем фазовой скорости* данной динамической системы с фазовым пространством M .

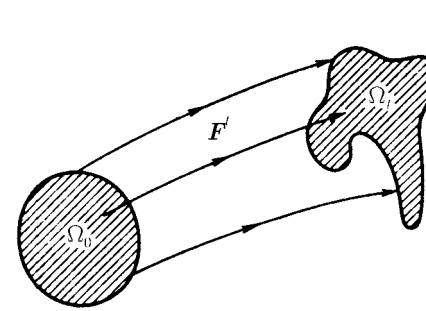


Рис. 2.32. Действие отображения $F^t: M \rightarrow M$ фазового пространства в себя для гамильтоновой системы

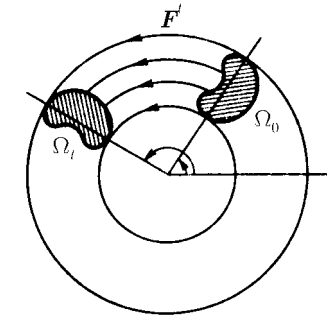


Рис. 2.33. Действие фазового потока F^t для гармонического осциллятора

Рассмотрим два характерных примера, поясняющих действие фазового потока.

1. Гармонический осциллятор единичной массы.

Гамильтониан для такой задачи имеет вид $H = (1/2)p^2 + (1/2)\omega_0^2 q^2$, а уравнения движения запишутся как $\dot{q} = p$, $\dot{p} = -\omega_0^2 q$. Переход к переменным действие-угол осуществляется следующими каноническими преобразованиями: $p = \sqrt{2\omega_0 J} \cos \alpha$, $q = -\sqrt{2J/\omega_0} \sin \alpha$. Тогда функция Гамильтона и система канонических уравнений, соответственно, примут вид $H = \omega_0 J$, $\dot{J} = 0$, $\dot{\alpha} = \omega_0$. Отсюда сразу находим решение: $J(t) = J_0$, $\alpha(t) = \alpha_0 + \omega_0 t$. Таким образом, действие отображения F^t в переменных действие-угол для гармонического осциллятора сводится к повороту (рис. 2.33).

2. Рассмотрим теперь систему, фазовый объем которой не сохраняется со временем. Пусть

$$\dot{x} = \alpha x, \quad \dot{y} = \alpha y, \quad \alpha > 0. \quad (2.51)$$

В этом случае начальный объем Ω_0 с течением времени неограниченно возрастает: $x(t) = x_0 \exp(\alpha t)$, $y(t) = y_0 \exp(\alpha t)$. Действие фазового потока для такой системы показано на рис. 2.34.

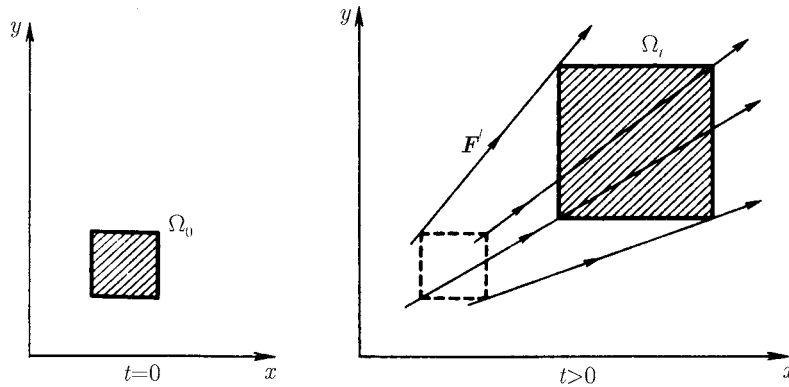


Рис. 2.34. Действие фазового потока F^t для системы, фазовый объем которой со временем не сохраняется

Ниже рассматриваются только консервативные системы, для которых имеет место сохранение с течением времени фазового объема.

18.1 Эргодичность

Напомним определения *среднего по времени* значения функции (или *среднего вдоль траектории* динамической системы) и *фазового среднего*.

Рассмотрим некоторую интегрируемую функцию n переменных $h(x_1, \dots, x_n) \equiv h(\mathbf{x})$. Если величины x_i — динамические переменные, то, вообще говоря, h будет (через x_i) зависеть от времени: $h = h(\mathbf{x}(t))$. Средним по времени значением такой функции называется величина

$$\bar{h}(\mathbf{x}_0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T h(\mathbf{x}(t)) dt. \quad (2.52)$$

Строгое рассмотрение вопроса об условиях применимости формулы (2.52) можно найти в [115, 179]. Вспомнив, что $F^t \mathbf{x}_0 = \mathbf{x}(t)$ (см. (2.50)), из (2.52) найдем

$$\bar{h}(\mathbf{x}_0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T h(F^t \mathbf{x}_0) dt. \quad (2.53)$$

Можно показать, что для *почти всех* (т. е. для всех, кроме множества меры нуль) начальных условий для \mathbf{x} функция $\bar{h}(\mathbf{x}_0)$ является инвариантной, т. е. не изменяется вдоль траектории [44].

Предположим, что фазовые кривые динамической системы не уходят на бесконечность и движение происходит в некоторой ограниченной области D с объемом V_D фазового пространства. Тогда фазовым средним функции $h(\mathbf{x})$ называется величина

$$\langle h \rangle = \frac{1}{V_D} \int_{(D)} h(\mathbf{x}) dV, \quad dV \equiv dx_1 \dots dx_n. \quad (2.54)$$

Движение динамической системы (2.48) называется *эргодическим* (или динамическая система (2.48) называется *эргодической*), если для произвольной интегрируемой функции $h(\mathbf{x})$ и почти всех начальных условий \mathbf{x}_0 имеет место равенство временных и фазовых средних функции $h(\mathbf{x})$:

$$\bar{h}(\mathbf{x}_0) = \langle h \rangle. \quad (2.55)$$

Таким образом, для эргодической системы среднее по времени не зависит от \mathbf{x}_0 , т. е. от начальной точки траектории.

Рассмотрим произвольную область Ω внутри D , а в качестве $h(\mathbf{x})$ возьмем такую функцию $h_\Omega(\mathbf{x})$, что

$$h_\Omega(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega, \\ 0, & \mathbf{x} \notin \Omega, \end{cases} \quad (2.56)$$

т. е. равную единице, когда \mathbf{x} принадлежит области Ω , и нулю — в противном случае. Эта функция, очевидно, интегрируема:

$$\frac{1}{V_D} \int_{(D)} h_\Omega(\mathbf{x}) dV = \frac{V_\Omega}{V_D}. \quad (2.57)$$

Следовательно, фазовое среднее функции (2.56) есть «относительный объем» области Ω . Если движение динамической системы эргодично, то, как

следует из (2.55), относительное время, проведенное фазовой траекторией внутри любой области Ω , равно относительному объему этой области и не зависит от выбора начальных условий. Иными словами, фазовая кривая эргодической системы, в силу (2.55), (2.56), будет равномерно и плотно заполнять всю область Ω .

Заметим, что для автономных гамильтоновых систем поток в некотором ограниченном объеме фазового пространства, вообще говоря, не может быть эргодическим согласно приведенному выше определению. Действительно, поскольку всегда имеется интеграл энергии, фазовая кривая обязательно должна располагаться на гиперповерхности $H = E$ и поток может быть эргодическим только на этой гиперповерхности. Если же существуют и другие k интегралов движения, то при заданных значениях этих интегралов эволюция гамильтоновой системы будет происходить на гиперповерхности размерности $n - k - 1$, определяемой всеми $k + 1$ интегралами движения, в n -мерном фазовом пространстве. Именно эта гиперповерхность и является энергетически доступной областью фазового пространства. С учетом этого замечания условие эргодичности (2.55) следует переписать в виде

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T h(\mathbf{F}^t \mathbf{x}_0) dt = \frac{1}{S_G} \int_{(G)} h(\mathbf{x}) dS, \quad (2.58)$$

где интегрирование ведется по гиперповерхности G , а S_G — «площадь» этой гиперповерхности.

Таким образом, если движение гамильтоновой системы эргодично, то с течением времени фазовая траектория равномерно и плотно будет покрывать определяемую всеми интегралами движения гиперповерхность в n -мерном фазовом пространстве.

Относительный объем V_Ω/V_D или «относительная площадь» S_Ω/S_D некоторой области Ω представляет собой меру $\mu(\Omega)$ этой области.⁵ Используя понятие меры, определение эргодичности (2.55) можно записать как

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T h(\mathbf{F}^t \mathbf{x}_0) dt = \int_{(D)} h(\mathbf{x}) d\mu, \quad (2.59)$$

⁵Вообще говоря, меру можно вводить по-разному. Приведенный пример — это лишь частный случай меры. Подробное рассмотрение вопросов, связанных с введением понятия меры, можно найти в [88, 100]. Для наших целей, однако, достаточно понимать под мерой только относительный объем области.

где интегрирование ведется по всей доступной области (гиперповерхности) фазового пространства; при этом

$$\mu(D) = \int_{(D)} d\mu = 1. \quad (2.60)$$

Примером эргодического поведения может служить квазипериодическое движение гамильтоновой системы, имеющее место, например, в модели двух гармонических осцилляторов с иррациональным отношением частот (см. § 14). В этом случае фазовая кривая плотно покрывает поверхность тора и движение на нем является эргодическим [116, 166].

18.2 Перемешивание. Непредсказуемость и необратимость

Свойство эргодичности — необходимое, но не достаточное условие хаотичности динамических систем [6, 44, 166]. В самом деле, рассмотрим действие фазового потока \mathbf{F}^t на малую область Ω_0 , принадлежащую поверхности двумерного инвариантного тора (рис. 2.35). Площадь этой области задается как $J_1 J_2 \Delta\alpha_1 \Delta\alpha_2$. В силу теоремы о сохранении фазового объема, имеем (при $\Delta J \equiv \Delta J_1 \Delta J_2$ и $\Delta\alpha \equiv \Delta\alpha_1 \Delta\alpha_2$)

$$\frac{d(\Delta J \Delta\alpha)}{dt} = \Delta\alpha \frac{d(\Delta J)}{dt} + \Delta J \frac{d(\Delta\alpha)}{dt} = 0. \quad (2.61)$$

Поскольку $d(\Delta J)/dt = 0$, то из (2.61) получим $d(\Delta\alpha)/dt = 0$. Следовательно, характерной особенностью эргодического движения гамильтоновой системы на двумерном торе является неизменность не только объема, но и формы области Ω_0 . Действие же фазового потока \mathbf{F}^t сводится просто к перемещению Ω_0 по поверхности тора (рис. 2.35). Вместе с тем, существуют гамильтоновы системы (см. § 14) с более сложными режимами движения. Для таких систем начальный элементарный объем, перемещаясь по всей энергетически доступной гиперповерхности (как для эргодических потоков), сильно деформируется, выпуская нерегулярного вида отростки, но при этом не разрывается (рис. 2.36). Это приводит к тому, что спустя некоторое время начальная область Ω_0 таким образом распределится по всей гиперповерхности, что ее кусочки можно будет обнаружить в любой части гиперповерхности независимо от размеров, формы и расположения исходной области Ω_0 . Такие системы, как говорят, обладают *перемешиванием*. Как будет видно из дальнейшего, свойство перемешивания, в отличие от эргодичности, как раз и может служить критерием хаотичности движения динамических систем.

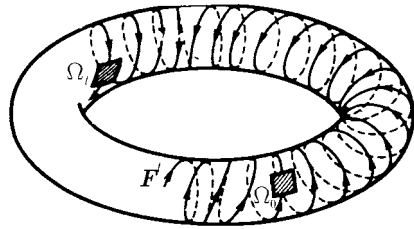


Рис. 2.35. Действие фазового потока F^t на малую область Ω_0 при квазипериодическом движении

30% чернил и 70% воды. Формализация подобного процесса при условии, что и чернила, и вода не разрываются на отдельные капли, приводит к понятию перемешивания.

Как правило, перемешивающий поток имеет место тогда, когда в фазовом пространстве близкие в начальный момент времени точки будут двигаться по сильно расходящимся траекториям.

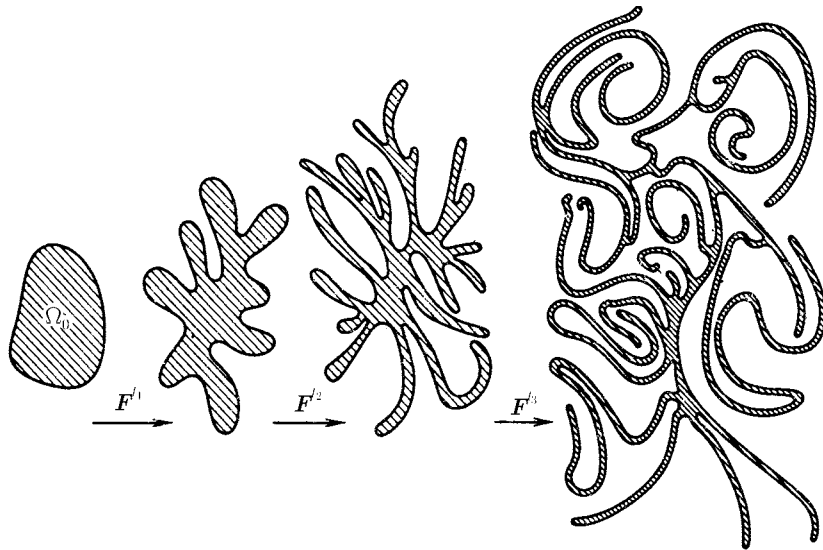


Рис. 2.36. Действие фазового потока F^t на малую область Ω_0 при перемешивании

Для введения понятия перемешивания рассмотрим две произвольные малые области A и B с мерами $\mu(A)$ и $\mu(B)$. Будем считать, что область B остается неподвижной, а область A эволюционирует с течением времени согласно непрерывному отображению F^t , порождаемому системой дифференциальных уравнений (2.48), т. е. $A_t = F^t A$. Пусть $A_t \cap B$ представляет собой совокупность всех частей A_t , оказавшихся в момент времени t внутри неподвижной области B (рис. 2.37). Динамическая система (2.48) называется *перемешивающей* (и, соответственно, поток — *перемешивающим*), если существует предел при $t \rightarrow \infty$, равный

$$\lim_{t \rightarrow \infty} \frac{\mu(A_t \cap B)}{\mu(B)} = \mu(A). \quad (2.62)$$

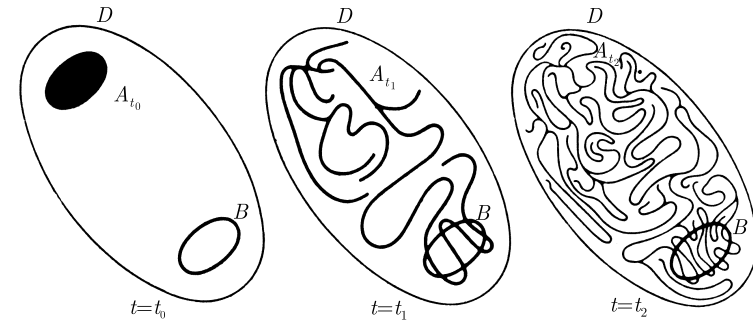


Рис. 2.37. Эволюция области A в случае перемешивания

Поясним смысл введенного определения. Прежде всего напомним, что, поскольку рассматриваемое движение консервативно, мера области A при движении сохраняется, т. е. $\mu(A_t) = \mu(A)$. Более того, $\mu(A) = \mu(A_t)/\mu(D)$ есть относительная доля объема, занимаемая областью A в D . Однако отношение $\mu(A_t \cap B)/\mu(B)$ есть относительная доля объема, занимаемая в области B попавшими внутрь нее в момент t кусочками области A_t . Определение (2.62) утверждает, что в пределе, когда $t \rightarrow \infty$, эти два отношения совпадают независимо от того, какими были размеры, форма и взаимное расположение областей A и B .

Из определения перемешивания (2.62) следует несколько выводов.

Неподвижную область B можно взять сколь угодно малой и расположить где угодно. Несмотря на это, по прошествии достаточно длительного

отрезка времени t внутри B можно будет обнаружить кусочки области $A_t = F^t A$, получаемой в результате эволюции исходной области A . Это означает, что при $t \rightarrow \infty$ кусочки A_t оказываются в любой сколь угодно малой окрестности произвольно выбранной точки энергетически доступной части фазового пространства. Иными словами, исходная область A с течением времени превращается в очень тонкую всепронизывающую паутинку, нити которой можно обнаружить внутри любого сколь угодно малого элемента объема. При этом подчеркнем, что суммарный объем, занимаемый этой паутинкой, равен начальному объему области A . Более того, получающаяся при $t \rightarrow \infty$ паутинка равномерно и однородно пронизывает фазовое пространство: внутри любой, наугад взятой области нити паутинки занимают одну и ту же относительную долю объема.

Заметим также, что исходную область A можно выбрать сколь угодно малой и расположить в любой доступной области фазового пространства. Тем не менее, как следует из определения (2.62), при $t \rightarrow \infty$ эта область превратится в пронизывающую все паутинку, о которой говорилось выше. Это означает, что фазовые траектории динамической системы с перемешиванием всегда абсолютно неустойчивы по отношению к малым возмущениям и разбегаются с течением времени.

Разбегание траекторий означает *непредсказуемость* поведения системы. Если в начальный момент времени положение фазовой точки было известно с конечной точностью, т. е. при $t = 0$ мы знаем лишь, что эта точка принадлежит некоторой области Ω_ϵ с характерным размером ϵ , то сказать, где именно она окажется через достаточно длительный промежуток времени, невозможно.

Перемешивание приводит к *необратимости*. Действительно, внутри любой доступной части фазового пространства с течением времени оказываются кусочки из самых различных начальных областей. Следовательно, зная лишь, что в конечный момент времени частица находилась в пределах малой области с размерами порядка ϵ , мы не сможем сказать, где она находилась в начальный момент времени. Поскольку в любом физическом эксперименте измеряется значение характерных величин лишь с какой-то конечной точностью, то детерминированное описание систем с перемешиванием не имеет смысла, и здесь, естественно, возникает потребность в построении статистических теорий.

Таким образом, понятие перемешивания соответствует нашему интуитивному представлению о том, какими должны быть системы со сложными и нерегулярными режимами движения.

Можно показать [26, 44], что перемешивание влечет за собой эргодичность. Однако обратное утверждение неверно: из эргодичности не следует

перемешивание. Действительно, вернемся к примеру двух гармонических осцилляторов с иррациональным отношением частот. Для такой системы, как было показано выше, действие фазового потока F^t на некоторую малую область Ω_0 сводится просто к перемещению этой области по тору (рис. 2.35) без искажения формы. При этом мера пересечения Ω_t с малой фиксированной областью B поверхности тора в различные моменты времени либо равна нулю, либо не равна нулю. Поэтому $\mu(\Omega_t \cap B)$ при $t \rightarrow \infty$ не стремится к определенному значению. Следовательно, существуют системы, которые обладают только эргодическим поведением, но не являются системами с перемешиванием.

Таким образом, перемешивание — это гораздо более сильное свойство, чем эргодичность. Если система обладает перемешиванием, то ее поведение естественно считать хаотическим. Однако не всякое хаотическое движение обязано быть перемешивающим. Определение (2.62) накладывает ряд жестких ограничений, не всегда выполняющихся в действительности.

18.3 Расщепление временных корреляций

Следствием перемешивания в системе является расщепление временных корреляций [81, 101].

Рассмотрим произвольные интегрируемые функции f и g в замкнутой области D фазового пространства динамической системы. Их фазовые средние определяются как (см. (2.54))

$$\langle f \rangle = \frac{1}{V_D} \int_{(D)} f(\mathbf{x}) dx_1 \dots dx_n, \quad \langle g \rangle = \frac{1}{V_D} \int_{(D)} g(\mathbf{x}) dx_1 \dots dx_n. \quad (2.63)$$

Пусть в начальный момент времени состояние системы определяется начальными условиями $\mathbf{x}(0) = \mathbf{x}_0$. Тогда через некоторое время система перейдет в другое состояние, характеризуемое фазовой точкой $\mathbf{x}(t)$, которая в силу единственности решения зависит от \mathbf{x}_0 : $\mathbf{x} = \mathbf{x}(\mathbf{x}_0, t)$. Определим коррелятор функций f и g как

$$\langle f(\mathbf{x}(\mathbf{x}_0, t))g(\mathbf{x}_0) \rangle = \frac{1}{V_D} \int_{(D)} f(\mathbf{x}(\mathbf{x}_0, t))g(\mathbf{x}_0) d\mathbf{x}_0. \quad (2.64)$$

Это соотношение можно переписать в виде

$$\langle f(\mathbf{x})g(\mathbf{x}_0(\mathbf{x}, t)) \rangle = \frac{1}{V_D} \int_{(D)} f(\mathbf{x})g(\mathbf{x}_0(\mathbf{x}, t)) d\mathbf{x}. \quad (2.65)$$

Как мы сейчас докажем, если система обладает перемешиванием, то при $t \rightarrow \infty$

$$\langle f(\mathbf{x})g(\mathbf{x}_0) \rangle \rightarrow \langle f \rangle \langle g \rangle \quad (2.66)$$

для любых функций f и g . Свойство (2.66) называется *расцеплением временных корреляций*.

Для того чтобы показать, что из перемешивания (2.62) следует расцепление корреляций, рассмотрим сначала в качестве f и g функции следующего вида:

$$f_B(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in B, \\ 0, & \mathbf{x} \notin B, \end{cases} \quad g_A(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in A, \\ 0, & \mathbf{x} \notin A, \end{cases} \quad (2.67)$$

где A и B — некоторые подобласти области D . Допустим, что область B остается неизменной, а область A эволюционирует под действием фазового потока: $A_t = \mathbf{F}^t A$. Тогда при больших t из условия перемешивания (2.62) найдем

$$\mu(A_t \cap B) \approx \mu(A)\mu(B). \quad (2.68)$$

Но мера $\mu(A_t \cap B)$ — это относительный объем пересечения образа A_t области A и области B :

$$\mu(A_t \cap B) = \frac{1}{V_D} \int_{(D)} f_B(\mathbf{x})g_{A_t}(\mathbf{x}) d\mathbf{x} = \frac{1}{V_D} \int_{(D)} f_B(\mathbf{x})g_A(\mathbf{x}_0(\mathbf{x}, t)) d\mathbf{x}, \quad (2.69)$$

а произведение $\mu(A)\mu(B)$ есть

$$\mu(A)\mu(B) = \frac{1}{V_D} \int_{(D)} f_B(\mathbf{x}) d\mathbf{x} \frac{1}{V_D} \int_{(D)} g_A(\mathbf{x}) d\mathbf{x}. \quad (2.70)$$

Поэтому из (2.68) при $t \rightarrow \infty$ находим

$$\frac{1}{V_D} \int_{(D)} f_B(\mathbf{x})g_A(\mathbf{x}_0(\mathbf{x}, t)) d\mathbf{x} \approx \frac{1}{V_D} \int_{(D)} f_B(\mathbf{x}) d\mathbf{x} \frac{1}{V_D} \int_{(D)} g_A(\mathbf{x}) d\mathbf{x}, \quad (2.71)$$

или

$$\langle f_B(\mathbf{x})g_A(\mathbf{x}_0) \rangle \approx \langle f_B \rangle \langle g_A \rangle. \quad (2.72)$$

Разобьем теперь область D фазового пространства на непересекающиеся области A_j . Тогда при достаточно мелком разбиении любую гладкую

функцию в области D можно приблизительно записать как

$$f(\mathbf{x}) = \sum_j \gamma_j f_{A_j}(\mathbf{x}), \quad f_{A_j}(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in A_j, \\ 0, & \mathbf{x} \notin A_j. \end{cases} \quad (2.73)$$

Отсюда получим, что для перемешивающих потоков

$$\begin{aligned} \langle f(\mathbf{x}(\mathbf{x}_0, t))g(\mathbf{x}_0) \rangle &= \sum_{j,k} \gamma_j \gamma'_k \langle f_{A_j}(\mathbf{x})g_{A_k}(\mathbf{x}_0) \rangle = \\ &= \sum_{j,k} \gamma_j \gamma'_k \langle f_{A_j} \rangle \langle g_{A_k} \rangle = \langle f \rangle \langle g \rangle, \end{aligned} \quad (2.74)$$

т. е. имеет место расцепление временных корреляций. Система с перемешиванием с течением времени как бы «забывает» о своих начальных условиях.

19. Бильярды. Газ Лоренца

Долгое время считалось, что перемешивание и хаотическое движение возможны только в системах с чрезвычайно большим числом степеней свободы, например, в системах статистической механики. Однако в 1962 г. Я. Г. Синай доказал [177] (см. также [178]), что очень простая модель, состоящая из двух плоских упругих дисков в ящике, один из которых закреплен, может обладать свойством перемешивания (рис. 2.38а). Если движущийся диск заменить материальной точкой, то мы приходим к так называемой бильярдной задаче (рис. 2.38б) — динамической системе, отвечающей движению по инерции материальной точки («шара») внутри ограниченной области с условием упругого отражения от ее границы по закону «угол падения равен углу отражения» [101]. Очевидно, такая динамическая система строится как последовательное преобразование от соударения к соударению бильярдного шара о границу.

Впервые понятие бильярда в теоретической и математической физике возникло после того, как Д. Биркгоф (см. [48]) рассмотрел задачу о движении шара на поверхности, ограниченной некоторой замкнутой кривой. Позже глубокие работы Н. С. Крылова [103], посвященные проблеме перемешивания в системе из упругих шаров, привели исследователей к необходимости рассмотрения задач бильярдного типа (см. [53, 81, 502] и приведенные там ссылки), поскольку они являются достаточно удобными моделями целого ряда физических систем. Например, многим динамическим задачам

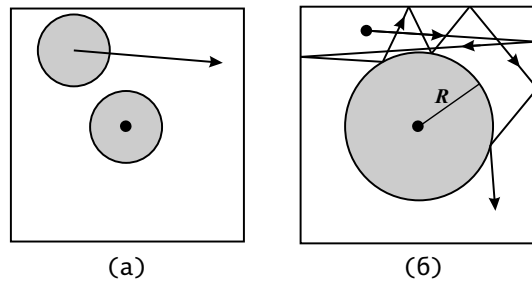


Рис. 2.38. Простейший вид рассеивающего бильярда (бильярда Синая)

могут быть поставлены в соответствие уравнения, описывающие траектории частиц в бильярдах заданной формы. Бильярдные модели возникают также в физике твердого тела и в некоторых геометрических приближениях отражения волн, особенно в акустике [322].

Сейчас теория бильярдных систем — это самостоятельно развивающаяся область математической физики. В данном параграфе мы подробно рассмотрим некоторые результаты этой теории, поскольку они неожиданны, чрезвычайно интересны и оказываются важными в приложениях.

Математически плоский бильярд представляет собой обычный бильярд, только с произвольной конфигурацией стола и без луз. Хотя, строго говоря, бильярды относятся к сильно упрощенным моделям классической статистической механики, имеется естественная аналогия между некоторыми физическими задачами и системами бильярдного типа [81, 314]. Одной из таких моделей является так называемый газ Лоренца. На ее основе удается получить ряд строгих и очень глубоких результатов.

Рассмотрим неограниченную область D в двумерном евклидовом пространстве. Допустим, что в этой области находятся определенным образом расположенные непересекающиеся шары B_i (рассеиватели) радиуса R (рис. 2.39). При условии, что шары неподвижны, такая система, очевидно, будет бильярдной. Она называется *газом Лоренца*: частица движется между рассеивателями, упруго отражаясь от них. Впервые газ Лоренца был введен в связи с проблемой описания движения электронов в металлах (см. [101]).

При периодическом расположении рассеивателей с периодом a существует три качественно различных модели газа Лоренца: газ Лоренца с ограниченным горизонтом, $R \geq a/2$, открытым горизонтом, $a\sqrt{3}/4 \leq R < a/2$ и бесконечным горизонтом $R < a\sqrt{3}/4$. В первом случае движение частицы ограничено одной ячейкой. Поэтому для ограниченного горизонта динамика газа Лоренца аналогична динамике частицы в рассеивающем бильярде

(рис. 2.38б). Во втором и третьем случаях бильярдной частице будет доступна вся область D (рис. 2.39).

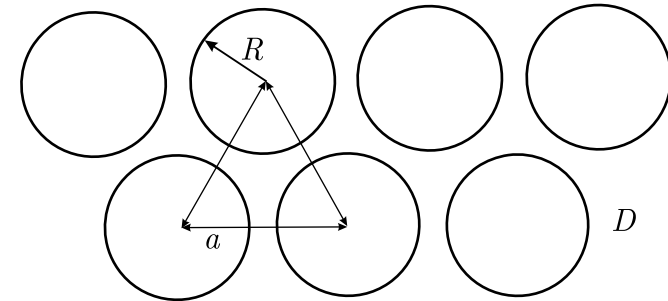


Рис. 2.39. Пример газа Лоренца с ограниченным горизонтом

Для бесконечного горизонта из-за увеличения вероятности длинных перелетов изменяются статистические свойства бильярда. В частности, не сходится среднее значение длины свободного пробега [53, 225, 252, 261, 317]. Эта величина является одним из основных показателей газа Лоренца. Она определяется как $l = \pi A/P$, где A — доступная частице область бильярда, а P — периметр рассеивателей. Для системы с открытым горизонтом $l = (a^2 - \pi R^2)/2R$, а для бильярда с бесконечным горизонтом значение l не ограничено сверху.

Для газа Лоренца с ограниченным и открытым горизонтами было доказано, что в двумерном (плоском) случае движение частицы является эргодическим, обладает свойством перемешивания и, более того, сводится к броуновскому. Также в такой системе существует коэффициент диффузии [252]. Это явилось первым строгим подтверждением рождения хаоса в динамических системах.

Появление перемешивания в бильярде, показанном на рис. 2.38, и газе Лоренца можно легко понять, если заметить, что два параллельных луча поле отражения от границы становятся расходящимися. Это означает наличие в бильярдной системе локальной неустойчивости траекторий, которая фактически и лежит в основе явления перемешивания (см. рис. 2.36).

Рассмотрим бильярды на плоскости с более формальной точки зрения и проведем анализ динамики частиц. Бильярдным столом Q является область M с кусочно-гладкой границей ∂Q . Бильярдная динамическая система в M порождается свободным движением материальной точки (бильярдного шара) с условием упругого отражения от ∂Q . Это означает, что

точка всегда движется по прямой в M с постоянной скоростью до соударения с границей. Если граница в точке соударения гладкая, то шар отражается от нее так, что тангенциальная компонента скорости не изменяется, в то время как нормальная составляющая изменяет знак. Если шар попадает в угол, то его дальнейшее движение определено неоднозначно или же не определено совсем.

Как правило, граница каждого бильярда состоит из m компонент ∂Q_i , $i = 1, 2, \dots, m$. Точки сшивки компонент ∂Q_i (т. е. «углы» стола Q) называются особыми. Остальные точки q — регулярными. Если для каждой компоненты ∂Q_i рассмотреть единичные нормали $\mathbf{n}(q)$ в каждой точке $q \in \partial Q_i$, направленные внутрь области Q , то они будут определять кривизну $k(q)$ кривой ∂Q_i во всех регулярных точках q . Компонента ∂Q_i является *рассеивающей* и обозначается Γ^d , если $k(q) > 0$. Для $k(q) = 0$ и $k(q) < 0$ получим, соответственно, *нейтральную* Γ^0 и *фокусирующую* Γ^f компоненты границы бильярда. Таким образом, каждая из компонент границы может быть рассеивающей, фокусирующей или нейтральной. Объединение всех рассеивающих, нейтральных и фокусирующих компонент обозначим, соответственно, как ∂Q^+ , ∂Q^0 и ∂Q^- .

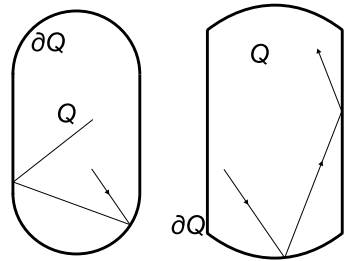


Рис. 2.40. Бильярд с границей в форме стадиона

Когда граница состоит только из рассеивающих и нейтральных компонент, бильярд называется *рассеивающим* (или бильярдом Синая), рис. 2.38б.

Бильярд, граница которого состоит только из фокусирующих или нейтральных и фокусирующих компонент, называется *фокусирующим*. Самым популярным примером такого бильярда является «стадион», граница которого состоит из двух дуг окружностей одинакового радиуса и двух параллельных отрезков, соединяющих концы этих дуг (рис. 2.40). Несмотря на наличие фокусирующих компонент, для такого

бильярда тоже характерно свойство перемешивания. Это связано с тем, что параллельные лучи после прохождения точки фокусировки начинают расходиться.

Возможны также бильярды, граница которых составлена только из нейтральных компонент. Простейший из них — бильярд, образованный границей в форме прямоугольника. Для исследования такого бильярда представим каждый участок траектории шара между двумя последовательными отражениями от вертикальных стенок на отдельном листе. Если затем скле-

ить полученные листы, то получится бесконечная развертка поверхности бильярдного стола (рис. 2.41).

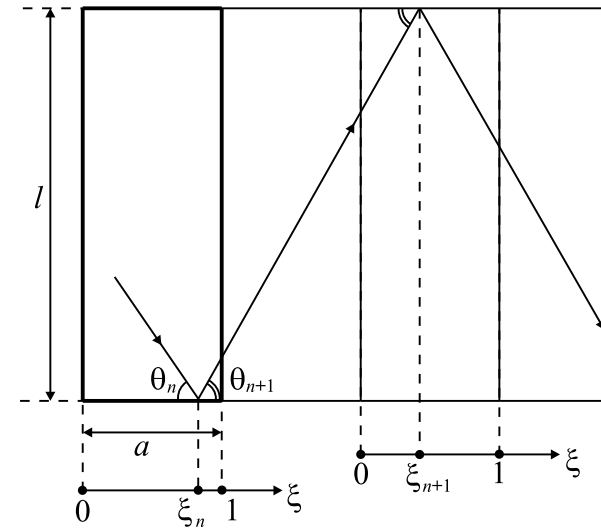


Рис. 2.41. Развертка бильярдного стола в форме прямоугольника

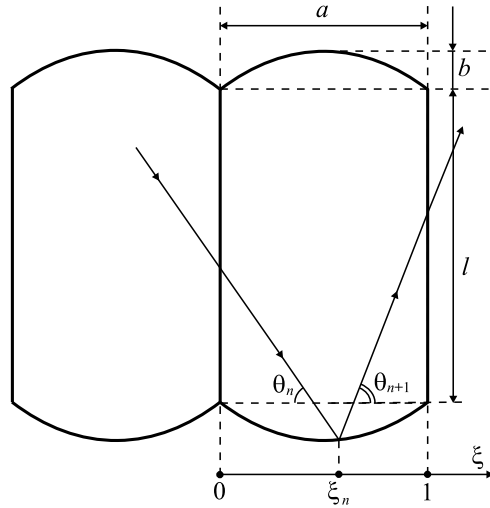
Введем дискретные динамические переменные θ_n и ξ_n , как показано на рисунке. Тогда преобразование, переводящее значение (θ, ξ) в момент n -го столкновения с ∂Q в их значения в момент $(n + 1)$ -го столкновения можно записать как

$$\begin{aligned} \theta_{n+1} &= \theta_n \equiv \theta_0 = \text{const}, \\ \xi_{n+1} &= \left\{ \xi_n + \frac{l}{a} \text{ctg} \theta_0 \right\}, \end{aligned} \quad (2.75)$$

где фигурные скобки означают взятие дробной части. Легко понять, что эта система является полностью интегрируемой, и переменные θ, ξ играют роль действия и угла соответственно. Поэтому, в зависимости от начальных условий, она может проявлять два качественно различных типа движения:

(1) если $l \text{ctg} \theta_0 / a = p/q$, где p и q — целые, то имеет место периодическая динамика;

(2) если $l \text{ctg} \theta_0 / a \neq p/q$, то в системе будет наблюдаться квазипериодическое поведение.

Рис. 2.42. Координаты (θ, ξ) , бильярд типа «стадион»

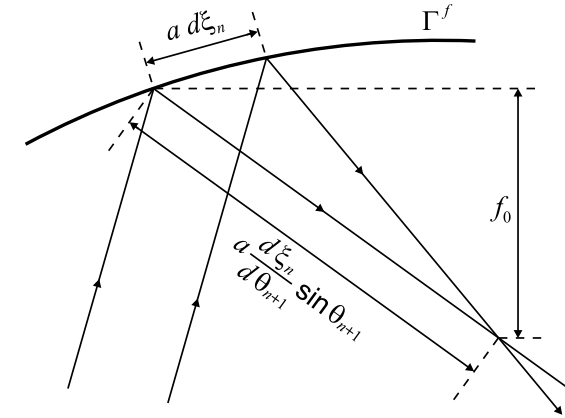
Рассмотрим теперь бильярд с границей ∂Q в форме, напоминающей стадион, состоящей из двух параллельных прямых Γ^0 длины l и двух фокусирующих дуг Γ^f (рис. 2.42). Пусть

$$l \gg a \gg b. \quad (2.76)$$

Это условие упрощает задачу и дает возможность аналитического рассмотрения. Предположим, что $\chi(\xi) = 4b\xi(1 - \xi)$ — функция, описывающая форму дуг Γ^f границы ∂Q , $\max(\chi) = b$. Тогда, вводя аналогично предыдущему случаю динамические переменные (см. рис. 2.42), легко получить следующее преобразование, описывающее изменение динамики бильярдной частицы от столкновения к столкновению:

$$\begin{aligned} \theta_{n+1} &= \theta_n - 2 \operatorname{arctg} \left\{ \frac{4b}{a} (1 - 2\xi_n) \right\}, \\ \xi_{n+1} &= \left\{ \xi_n + \frac{l}{a} \operatorname{ctg} \theta_{n+1} \right\}, \end{aligned} \quad (2.77)$$

где учтены условия (2.76) и фигурные скобки также означают взятие дробной части. Сравнивая (2.75) и (2.77), легко видеть, что они отличаются только малым членом $\sim b/a$ в соотношении для θ .

Рис. 2.43. Отражение узкого пучка траекторий от фокусирующей компоненты Γ^f бильярда типа стадион

Из матрицы преобразования $\partial(\theta_{n+1}, \xi_{n+1})/\partial(\theta_n, \xi_n)$ соотношения (2.77) найдем корни характеристического уравнения: $\lambda_{1,2} = \left(2 - K_0 \pm \sqrt{K_0^2 - 4K_0} \right) / 2$, $K_0 = |d\xi_{n+1}/d\xi_n - 1| = K/\sin^2 \theta_{n+1}$, $K = 16bl/a^2$. Коэффициент K_0 называется *коэффициентом растяжения фаз* [81].

Для выяснения физического смысла этого коэффициента проведем геометрический анализ бильярда. Рассмотрим отражение узкого параллельного пучка траекторий от дуги Γ^f границы бильярда (рис. 2.43). Ширина пучка после n -го отражения равна $a d\xi_n \sin \theta_{n+1}$, а путь, проходимый пучком до точки фокусировки, определяется как $a \sin \theta_{n+1} d\xi_n / d\theta_{n+1}$. Его проекция f_0 на вертикальную ось равна

$$f_0 = a \frac{d\xi_n}{d\theta_{n+1}} \sin^2 \theta_{n+1}. \quad (2.78)$$

Далее, из (2.77) получаем, что $d\theta_{n+1}/d\xi_n = 16b/a$. Следовательно, $f_0 = f \sin^2 \theta_{n+1}$, где $f = a^2/16b$. Поэтому $K_0 = l/f_0$, $K = l/f$. Значит, если граница ∂Q удовлетворяет условию, что до очередного отражения от Γ^f успевают произойти фокусировка первоначально параллельного пучка, то K , $K_0 \sim 1$. При $K_0 \gg 1$ расстояние, проходимое пучком от компоненты Γ^f до точки фокусировки, оказывается много меньше, чем расстояние от точки фокусировки до следующего отражения. Иными словами, если $K_0 \gg 1$, то пучок траекторий от столкновения к столкновению сильно расходится.

Следовательно, соотношение $K_0 \gg 1$ задает условие возникновения хаоса. Его мы используем для анализа динамики частиц в бильярде типа «стадион» (см. ниже).

В общем случае, т. е. при произвольных соотношениях между параметрами бильярда, если его граница имеет как рассеивающие, так и фокусирующие компоненты, динамика частицы будет определяться соотношением времен, которые она проводит в окрестности ∂Q^+ и ∂Q^- . Точный смысл этого утверждения описан в работах [52, 249], где доказано, что если на фокусирующую компоненту $\Gamma^f \subset \partial Q$, имеющую постоянную кривизну, падает параллельный пучок траекторий и испытывает серию из j отражений от Γ^f , то компонента Γ^f эффективно не влияет на уменьшение ширины фронта, а его кривизна остается ограниченной сверху.

Основные результаты, касающиеся бильярдов с фокусирующими и нейтральными компонентами, сводятся к следующему [52].

а) если кривизна каждой компоненты ∂Q^- постоянна;

б) если никакие две фокусирующие компоненты не являются дугами одной и той же окружности;

в) если дуга, дополняющая любую фокусирующую компоненту до полной окружности, лежит внутри области Q , то бильярд в области, граница которого имеет рассеивающие и фокусирующие компоненты, обладает свойством перемешивания. При этом не накладывается никаких условий на соотношение длин рассеивающих и фокусирующих компонент.

Условие а) можно немного ослабить, чтобы граница фокусирующей компоненты слабо менялась. Кроме того, если Γ^f пересекает только нейтральные компоненты, то вместо условия в) достаточно, чтобы центр окружности лежал внутри Q . Примеры таких бильярдов показаны на рис. 2.40 и рис. 2.44.

Более общие условия существования хаотичности в двумерных бильярдах на плоскости описаны в работах [249, 250] (см. также данную там литературу).

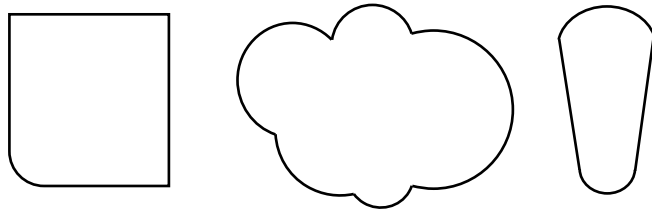


Рис. 2.44. Примеры перемешивающих бильярдов с фокусирующими компонентами

Хотя строгие результаты, полученные для бильярдных систем, относятся только к нескольким взаимодействующим частицам, можно ожидать, что аналогичными свойствами обладают ансамбли из большого числа упруго сталкивающихся частиц. Фактически, свойство перемешивания всегда предполагается при построении кинетической теории газов.

19.1 Бильярды с возмущаемыми границами⁶

Естественным обобщением бильярдных систем являются бильярды, границы которых могут изменяться с течением времени. К моделям такого типа сводятся некоторые важные проблемы математической физики. Например, анализ движения частиц внутри возмущенных ядер привел к необходимости изучения квадрупольных, или «динамических», резонансов. В таких задачах рассматривалось столкновение частиц с движущимися границами (полями) и, как следствие, обмен энергиями. В физике плазмы ускорение в магнитных ловушках моделируется бильярдами с изменяющейся во времени границей. Наконец, если рассматривать газ Лоренца как систему, описывающую динамику электронов в металлах, то вследствие ненулевой температуры рассеиватели должны слегка «осциллировать» с небольшой амплитудой вблизи своего равновесного состояния. Более того, как сейчас ясно, бильярды с зависящими от времени границами — это область физики, открывающая большие перспективы в исследовании многих давно известных, но малоизученных проблем (см. [132, 407, 409] и приводимые там ссылки). Например, возмущение границы может приводить к замедлению частиц (см. ниже), что может оказаться очень важным в приложениях.

В бильярдах с возмущаемыми границами скорость шара уже не является постоянной величиной, а от столкновения к столкновению меняется. При единичных актах столкновения шар приобретает или отдает энергию в зависимости от того, движется ли шар и граница навстречу друг другу (т. е. происходит их *встречное* столкновение) или же шар нагоняет границу (т. е. имеет место *сопутствующее* столкновение). Вопрос состоит в том, чтобы определить, каких в среднем столкновений больше. При сопутствующих столкновениях частица теряет энергию, так что ее скорость будет уменьшаться. Если же преобладают встречные столкновения, то бильярдная частица должна ускоряться.

Кажется вполне понятным, что если бильярд обладает хаотической динамикой, то возмущение границ должно приводить к ускорению бильярдной частицы. В самом деле, направление скорости частицы в хаотическом

⁶Излагаемые ниже результаты получены одним из авторов (А. Ю. Л.) совместно с А. Б. Рябовым, который также принимал участие в написании пп. 19.1–19.3

бильярде можно считать случайной величиной. При введении возмущения границы число встречных столкновений будет преобладать над числом сопутствующих. Поэтому в среднем скорость частицы должна увеличиваться. Таким образом, для бильярда с зависящими от времени границами существенными оказываются его динамические свойства: если бильярд проявляет хаотическую динамику, то возмущение границы может привести к неограниченному росту скорости частицы. С другой стороны, бильярд с достаточно гладкой границей, например, в форме эллипса или окружности, не хаотичен, и ее колебания не приводят к разгону бильярдной частицы.

Первоначально идея стохастического ускорения частиц была предложена Энрико Ферми [308] для объяснения происхождения космических лучей высоких энергий. Она состояла в том, что при столкновениях заряженных частиц с беспорядочно движущимися магнитными облаками в межзвездном пространстве частицы должны в среднем ускоряться. Рассматривая облако как тело большой массы, нетрудно понять причину ускорения. Если скорости облаков, с которыми сталкивается частица, распределены случайно, то можно сказать, что число облаков, движущихся в одном направлении, равно числу облаков, движущихся в обратном направлении. Поэтому частица будет чаще сталкиваться с теми из них, которые движутся ей навстречу. Отсюда следует, что частица чаще будет приобретать энергию, чем отдавать ее. Так возникает эффективное ускорение, называемое *ускорением Ферми*.

Для бильярдов ситуация оказалась значительно сложнее: не всякому хаотическому бильярду при возмущении границ присуще ускорение Ферми. Более того, в некоторых бильярдах, например, в форме стадиона, когда кривизна фокусирующих компонент мала (справа на рис. 2.40), наблюдается *замедление* частиц [407].

19.2 Газ Лоренца с возмущаемыми границами

В данной части мы рассмотрим возникновение ускорения Ферми на примере обобщенного бильярда — газа Лоренца с открытым горизонтом с возмущаемой границей [132]. При этом изучаются два разных случая, соответствующие стохастическим и регулярным (гармоническим) колебаниям границы. Представленный анализ легко переносится на любой другой бильярд, для которого известно вероятностное распределение угла между нормалью к поверхности в точке соударения и вектором скорости частицы.

19.2.1 Неподвижная граница

В качестве канонических переменных можно выбрать азимутальный угол ϕ и угол падения α между нормалью к поверхности и вектором ско-

рости частицы перед соударением. Введем угол отражения α^* между нормалью и вектором скорости после соударения (см. рис. 2.45). Очевидно, что $\phi \in [0, 2\pi]$, а углы α и α^* изменяются в интервале $[-\pi/2, \pi/2]$.

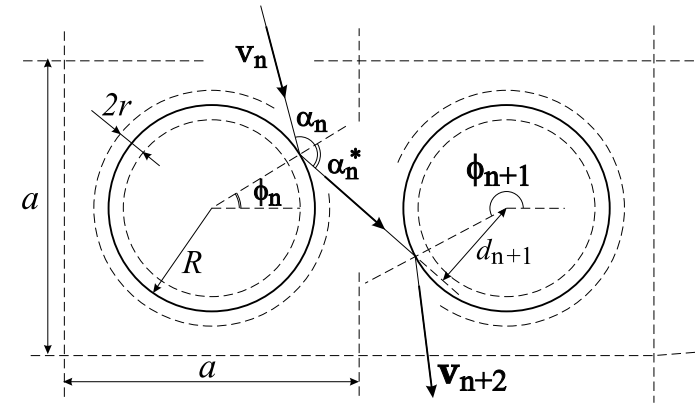


Рис. 2.45. Координаты в модели газа Лоренца

Для описания динамики невозмущенного бильярда необходимо получить преобразование $(\alpha_n, \phi_n) \rightarrow (\alpha_{n+1}, \phi_{n+1})$, переводящее значения переменных (α, ϕ) в момент n -го столкновения с ∂Q в их значения в момент $(n+1)$ -го столкновения. Из геометрических соображений понятно, что

$$\phi_n + \alpha_n^* + \pi = \phi_{n+1} + \alpha_{n+1}. \quad (2.79)$$

Кроме того, $\alpha_n^* = -\alpha_n$, так как углы отсчитываются в противоположные стороны.

Введем систему координат с центром в круге, на котором произошло очередное рассеяние, и найдем расстояние, на котором частица пройдет мимо другого центра, отстоящего на p ячеек по горизонтали и q ячеек по вертикали:

$$d_{n+1} = a [p \sin(\phi_n + \alpha_n^*) - q \cos(\phi_n + \alpha_n^*)] - R \sin \alpha_n^*. \quad (2.80)$$

Значение p считается положительным, если частица движется вправо, и отрицательным в противном случае; соответственно, q положительно, если частица движется вверх, и отрицательно, если вниз. Величины p и q определяются из условия рассеяния, т.е. как минимальные по модулю целые

числа, при которых выполняется $|d_{n+1}| \leq R$. Вычислив прицельный параметр d_{n+1} , легко найти угол, под которым произойдет столкновение со следующим рассеивателем:

$$\alpha_{n+1} = \arcsin\left(\frac{d_{n+1}}{R}\right). \quad (2.81)$$

Якобиан преобразования (2.79)–(2.81) равен $\partial(\phi_{n+1}, \alpha_{n+1})/\partial(\phi_n, \alpha_n) = \cos \alpha_n / \cos \alpha_{n+1}$. Таким образом, система сохраняет фазовый объем $\cos \alpha d\alpha d\phi$. Отсюда, в частности, следует что если рассматриваемый бильярд является эргодическим, то величина α_n будет распределена по закону

$$\rho_\alpha(\alpha) = \frac{1}{2} \cos \alpha, \quad (2.82)$$

где $1/2$ — нормировочный множитель.

19.2.2 Осциллирующая граница рассеивателей

Допустим, что рассеивающая компонента ∂Q^+ границы ∂Q сжимается и расширяется (см. рис. 2.45), так что ее радиус меняется по закону $R = R(t) = R + r(t)$, где $\max |r(t)| \ll R$. Тогда скорость границы будет зависеть от времени по закону $u(t) = \dot{r}(t)$. Далее для определенности будем считать, что $u(t) = u_0 \cos(\omega t)$, где $u_0 = \omega r_0$. В этом случае, помимо величин α и ϕ , для описания поведения системы необходимо ввести еще две динамические переменные: скорость частицы v и время столкновения t . Учитывая, что при отражении от рассеивателя меняется только нормальная (радиальная) компонента скорости, а тангенциальная остается неизменной, получим следующее соотношение для модуля скорости частицы после столкновения:

$$v_{n+1} = \sqrt{v_n^2 - 4u_n v_n \cos \alpha_n + 4u_n^2}, \quad (2.83)$$

где $u_n \equiv u_0 \cos \omega t_n$ — скорость границы рассеивателя в момент n -го столкновения. В свою очередь, связь между углом падения и углом отражения можно записать в виде $\alpha_n^* = -\arcsin(v_n \sin \alpha_n / v_{n+1})$. Теперь, вычисляя расстояние между последовательными соударениями, нетрудно получить преобразование для момента столкновения t_n :

$$t_{n+1} = t_n + \frac{l_{n+1}}{v_{n+1}},$$

$$l_{n+1} = \sqrt{\left(R(\cos \phi_{n+1} - \cos \phi_n) - pa\right)^2 + \left(R(\sin \phi_{n+1} - \sin \phi_n) - qa\right)^2}, \quad (2.84)$$

где l_n — длина свободного пробега. В предположении $r \ll R$ соотношения для переменной ϕ и прицельного параметра d останутся теми же, что и для невозмущенного бильярда (см. (2.79) и (2.80)).

Во время соударений с возмущаемой границей величина скорости бильярдного шара постоянно меняется. Как показывают исследования [132, 409], эти изменения скорости носят случайный характер. Поэтому далее рассмотрим ансамбль частиц и найдем распределение частиц по скоростям и среднюю скорость в зависимости от времени t и числа столкновений n . Дело в том, что число столкновений и время не пропорциональны друг другу, так как за одно и тоже время «быстрая» частица испытывает больше соударений чем «медленная».

• Стохастически возмущаемая граница рассеивателей

Пусть скорость границы рассеивателя, с которым происходит n -е столкновение, равна

$$u_n(t) = u_0 \cos \varphi_n, \quad (2.85)$$

где u_0 — это амплитудное значение скорости колебания границы, а совокупность $\{\varphi_n\}$ — множество некоррелированных между собой случайных величин, равномерно распределенных в интервале $[0, 2\pi)$. Найдем функцию распределения частиц по скоростям и зависимость средней скорости ансамбля частиц от числа соударений n и времени t . В случае малых скоростей частицы ($v \ll u_0$) в выражении для скорости (2.83) основную роль играет последний член и, следовательно, $v_{n+1} \approx 2|u(t_n)|$.

Если колебания границы заданы в соответствии с выражением (2.85), то $\langle v_{n+1} \rangle \approx 2 \langle |u(t_n)| \rangle_t = 4u_0/\pi$. Таким образом, уже после первого столкновения средняя скорость становится больше величины u_0 .

Теперь найдем изменение скорости в случае $v \gg u_0$. Раскладывая (2.83) в ряд Тейлора по параметру u/v , получим выражение для изменения скорости:

$$\Delta v_n = v_{n+1} - v_n = -2u_n \cos \alpha_n + 2\frac{u_n^2}{v_n} \sin^2 \alpha_n + v_n O\left(\left(\frac{u_n}{v_n}\right)^3\right), \quad (2.86)$$

где u_n скорость границы рассеивателя при n -ом столкновении.

Используя соотношение (2.82) и условие равномерного распределения по фазе колебаний в момент соударения, найдем $\langle \Delta v_n \rangle$ и $\langle (\Delta v_n)^2 \rangle$:

$$\mu_s \equiv \langle \Delta v_n \rangle = \frac{M_s}{v},$$

$$\sigma_s^2 \equiv \langle (\Delta v_n)^2 \rangle = \frac{4}{3} u_0^2. \quad (2.87)$$

Здесь для простоты дальнейших выкладок введена величина $M_s \equiv u_0^2/3$; индекс s обозначает стохастический случай. При подсчете среднего ненулевой вклад дает только второе слагаемое, а при вычислении дисперсии достаточно учесть вклад, связанный с первым слагаемым.

При достаточно больших n от первого уравнения системы (2.87) можно перейти к дифференциальному уравнению

$$\mu_s = \frac{\langle \Delta v_n \rangle}{1} = \frac{\partial v(n)}{\partial n} = \frac{M_s}{v(n)}.$$

Его решение с начальным условием $v(0) = v_0$ дает зависимость наиболее вероятной скорости от числа соударений: $v(n) = \sqrt{2M_s n + v_0^2}$. Поскольку скорость частиц представляется в виде суммы случайных независимых величин Δv_n с известным средним и дисперсией, то на основании центральной предельной теоремы Ляпунова можно утверждать, что плотность распределения случайной величины $v_n = v_0 + \sum_{i=1}^n \Delta v_i$ будет стремиться к нормальной со средним $v(n)$ и дисперсией $n\sigma_s^2$. Таким образом, распределение частиц по скоростям имеет вид расплывающегося гауссова пакета. Максимум распределения дается наиболее вероятной скоростью $v(n)$, которая растет как корень из n .

Приведенные рассуждения верны только в случае достаточно больших скоростей частицы ($v \gg u_0$). Чтобы описать распределение при малых скоростях, введем еще одно условие — отсутствие потока частиц в область отрицательных скоростей: $(v \partial \rho / \partial v)|_{v=0} = 0$. Хорошо известно, что удовлетворяющее этому условию гауссово распределение имеет вид:

$$\rho(v, n) = \frac{1}{\sigma_s \sqrt{2\pi n}} \left[\exp\left(-\frac{(v - v(n))^2}{2\sigma_s^2 n}\right) + \exp\left(-\frac{(v + v(n))^2}{2\sigma_s^2 n}\right) \right]. \quad (2.88)$$

Это позволяет найти среднюю скорость ансамбля частиц в зависимости от числа столкновений:

$$V(n) = \sigma_s \sqrt{\frac{2n}{\pi}} \exp\left(-\frac{v(n)^2}{2\sigma_s^2 n}\right) + v(n) \Phi\left(\frac{v(n)}{\sigma_s \sqrt{2n}}\right),$$

где $\Phi(x) = 2/\sqrt{\pi} \int_0^x \exp(-x^2) dx$ — известный интеграл ошибок. Здесь и далее большой буквой V обозначена средняя скорость ансамбля частиц. Подставляя все значения и раскладывая в ряд выражение для скорости, можно найти, что

$$V(n) = C\sqrt{n} + O\left(\frac{1}{\sqrt{n}}\right), \quad (2.89)$$

где постоянная $C = \sqrt{2} \left(\sigma_s e^{-M_s/\sigma_s^2} / \sqrt{\pi} + \Phi(\sqrt{M_s}/\sigma_s) \sqrt{M_s} \right) \approx 1.143u_0$.

Таким образом, соотношения (2.88) и (2.89) дают *распределение по скоростям и среднюю скорость ансамбля частиц в зависимости от числа столкновений*.

Для нахождения *средней скорости частиц в зависимости от времени* воспользуемся уравнением Фоккера–Планка:

$$\frac{\partial \rho(v, t)}{\partial t} = -\frac{\partial}{\partial v} (A \rho(v, t)) + \frac{1}{2} \frac{\partial^2}{\partial v^2} (B \rho(v, t)),$$

где коэффициенты A и B имеют следующий вид: $A \equiv \langle \Delta v / \tau \rangle = M_s / l$, $B \equiv \langle \Delta v^2 / \tau \rangle = \sigma_s^2 v / l$. Здесь среднее время между столкновениями $\tau = l / v$, l — средняя длина свободного пробега, а Δv и Δv^2 взяты из (2.87). Теперь, подставляя полученные коэффициенты в уравнение, найдем:

$$\frac{\partial \rho(v, t)}{\partial t} = -\frac{M_s}{l} \frac{\partial}{\partial v} \rho(v, t) + \frac{1}{2} \frac{\sigma_s^2}{l} \frac{\partial^2}{\partial v^2} (v \rho(v, t)).$$

Если параметры M_s и σ_s заданы в соответствии с выражениями (2.87), то решение этого уравнения при скоростях частиц, много больших начальной скорости, т. е. через достаточно большой интервал времени, будет стремиться к $\rho(v, t) = \exp\left(-\frac{v}{2tA}\right) / \sqrt{2tA\pi v}$, где $A = M_s / l$. Используя его, получаем среднюю скорость частиц:

$$V(t) = \frac{M_s}{l} t + v_0 = \frac{1}{3} \frac{u_0^2}{l} t + v_0. \quad (2.90)$$

Таким образом, в исследуемой системе наблюдается ускорение Ферми, причем скорость растет линейно со временем.

• Периодически возмущаемая граница рассеивателей

Пусть все границы рассеивателей сжимаются и расширяются по некоторому периодическому закону и в одной и той же фазе. Тогда в течение одной половины периода скорость частицы при столкновениях будет увеличиваться, а в течение другой — уменьшаться. При достаточно больших скоростях время свободного пробега τ_s становится меньше периода колебаний радиусов рассеивателей. Это приводит к корреляциям изменений скорости частицы, так что ее последовательные приращения (2.86) уже нельзя считать независимыми.

Для проведения анализа изменений скорости в хаотических бильярдах с периодически осциллирующими границами можно предложить следующий подход. Рассмотрим приближенное преобразование для скорости (2.86). Поскольку корреляции величин α_n по-прежнему экспоненциально спадают (так как это определяется геометрией бильярда), то соотношения можно усреднить по α , пользуясь (2.82). Тогда

$$\langle \Delta v \rangle_\alpha = -\frac{\pi}{2} u_0 \cos \omega t_n + \frac{u_0^2 \cos \omega t_n}{v_n}. \quad (2.91)$$

В течение периода колебаний наибольший вклад в изменение скорости дает первое слагаемое. Следовательно, для учета корреляций в первом приближении достаточно учитывать изменения скорости, связанные только с первым слагаемым, а вторым слагаемым можно пренебречь. С другой стороны, учет корреляционных поправок во втором слагаемом приводит к членам более высокого порядка малости, чем среднее от него. Таким образом, можно не учитывать корреляционные эффекты во втором слагаемом. Поэтому допустимо независимо вычислить две величины, связанные с первым и вторым слагаемым, т. е. $\langle \Delta v \rangle = \langle \Delta v_I \rangle + \langle \Delta v_{II} \rangle$, причем $\langle \Delta v_{II} \rangle = u_0^2 / (3v)$, что совпадает с величиной μ_s для стохастического случая (см. (2.87)), а $\langle \Delta v_I \rangle$ — поправка, обусловленная наличием корреляций. Отбрасывая в выражении (2.91) второе слагаемое, для вычисления величины $\langle \Delta v_I \rangle$ получим:

$$\begin{aligned} v_{n+1} &= v_n + \gamma \cos \theta_n, \\ \theta_{n+1} &= \theta_n + \frac{l_{n+1} \omega}{v_{n+1}}. \end{aligned}$$

Здесь $\gamma = -\pi u_0 / 2$, и вместо времени была введена фаза столкновения $\theta_n \equiv \omega t_n$. Приведенное соотношение в точности соответствует известному преобразованию Улама (см. [80, 116, 168, 169, 240, 389, 394, 516]) с той лишь разницей, что в данном случае длина свободного пробега l_n является случайной величиной, которая может принимать значения из определенного интервала.

Оценим дисперсию и скорость спада корреляций изменений скорости. Предположим, что скорость частицы столь велика, что ее изменение за n соударений намного меньше ее значения. Понятно, что для выполнения этого условия при любом n достаточно соответствующим образом выбрать v и u_0 . Найдем корреляции приращений скорости Δv_m и Δv_{m+n} (см. (2.86)) при $n \rightarrow \infty$. Учитывая в первом приближении только первые слагаемые,

из (2.86) получим:

$$R(n) \equiv \langle \Delta v_m \Delta v_{m+n} \rangle = u_0^2 \frac{\pi^2}{4} \langle \cos \omega t_m \cos \omega t_{m+n} \rangle,$$

где в соответствии с (2.82) учтено, что $\langle \cos \alpha_n \rangle = \pi/4$. Полагая частоту колебаний равной единице и вводя обозначения $S_n \equiv \sum_{i=1}^n \tau_{m+i}$, где $\tau_i = t_i - t_{i-1}$, найдем, что $\langle \cos t_m \cos t_{m+n} \rangle = \langle \cos t_m \cos (t_m + S_n) \rangle$. Величину S_n можно представить в виде $S_n = \sum_{i=1}^n (l + \Delta l_i) / v$, где Δl_i — отклонение от средней длины свободного пробега при i -м перелете. Поскольку S_n является суммой независимых случайных величин, то ее распределение при больших n будет стремиться к нормальному $N(nl, n\sigma_l^2)$, где σ_l^2 — дисперсия длины свободного пробега. Раскладывая косинус суммы и производя усреднение по S_n , получим следующее выражение для корреляционной функции приращений скорости:

$$R(n) \simeq \frac{\pi^2}{8} u_0^2 \cos(\omega n \tau) \exp\left(-\frac{n}{N}\right),$$

где ω — частота колебаний рассеивателя и $N = v^2 / (\omega^2 \sigma_l^2)$. Таким образом, последовательные изменения скорости скоррелированы тем сильнее, чем выше скорость. При этом число соударений N , после которых корреляции спадают e раз, увеличивается как v^2 . Отметим, что число столкновений за период растет пропорционально v . Таким образом, для правильного вычисления дисперсии необходимо учитывать тем большее число периодов колебаний, чем выше скорость частицы. Однако вопрос о том, каким образом это можно сделать, пока остается открытым.

Для того, чтобы в первом приближении оценить дисперсию, рассмотрим изменение скорости только после двух последовательных соударений с границей. При этом будем считать, что корреляции между тремя и более изменениями пренебрежимо малы. В пределе больших скоростей бильярдной частицы коррелятор последовательных приращений скорости можно оценить как

$$\langle \Delta v_n \Delta v_{n+1} \rangle = u_0^2 \frac{\pi^2}{4} \langle \cos^2 \omega t_n (1 - O(\tau^2)) \rangle = u_0^2 \frac{\pi^2}{8} + O\left(\frac{u_0^2}{v^2}\right).$$

Учитывая это выражение и соотношение (2.87), получим:

$$\sigma_r^2 = \frac{\langle (\Delta v_n + \Delta v_{n+1})^2 \rangle}{2} \approx \left(\frac{4}{3} + \frac{\pi^2}{8}\right) u_0^2.$$

Таким образом, приведенные оценки свидетельствуют о том, что ускорение частиц должно наблюдаться в хаотических бильярдах с периодически осциллирующими границами. По-видимому, можно утверждать, что *детерминированная случайность — это достаточное условие возникновения ускорения Ферми*.

• Численный анализ

Приведем численно полученные зависимости скорости частиц от числа столкновений и времени и сопоставим их с аналитическими, полученными выше. Вычисления производились на основе модели газа Лоренца со следующими параметрами: амплитудное значение скорости колебаний границы рассеивателя $u_0 = 0,01$; радиус рассеивателей $R = 0,4$; расстояние между центрами $a = 1$; частота осцилляций границы $\omega = 1$; начальная скорость $v_0 = 1$. Таким образом, соответствующее аналитически полученное значение длины свободного пробега составляет $l = 0,6216815$. Численно найденные длина свободного пробега (см. (2.84)) и ее дисперсия при данной геометрии бильярда составляют, соответственно, $l = 0,62163 \pm 0,00003$ и $\sigma_l^2 = 0,657 \pm 0,001$.

В каждом случае было построено 500 реализаций динамики бильярдной частицы, отличающиеся друг от друга начальными значениями α и ϕ , которые выбирались случайным образом. Исследовались два различных случая: стохастические колебания границ рассеивателей при равномерном распределении фазы колебаний и регулярные колебания границ. Динамика бильярдного шара (частицы) определялась из соотношений, полученных выше. При этом скорость колебаний границ рассеивателей в момент столкновения в первом случае определялась как $u_n = u_0 \cos \varphi_n$ (где φ_n — равномерно распределенная на интервале $[0, 2\pi)$ случайная величина), а во втором — как $u_n = u_0 \cos \omega t_n$ (где t_n — момент соударения частицы с границей).

Усредненные зависимости скорости от числа столкновений и времени представлены на рис. 2.46а, б. На обоих графиках сплошные линии отвечают регулярному случаю возмущения границы, а пунктир — стохастическому. При этом в стохастическом случае величины M_s и σ_s были найдены из уравнений (2.87), а в регулярном — взяты их предельные значения M_r и σ_r , полученные численно.

Как следует из рисунков, рост средней скорости близок к линейному, и аналитические аппроксимации средней скорости на основании соотношений (2.89), (2.90), дают неплохое согласие с численным экспериментом.

Таким образом, можно выделить два основных механизма ускорения. Во-первых, это «сносовый» механизм, возникающий вследствие выполнения соотношения $\langle \Delta v \rangle > 0$ (см. (2.87)) и приводящий к сносу всех ча-

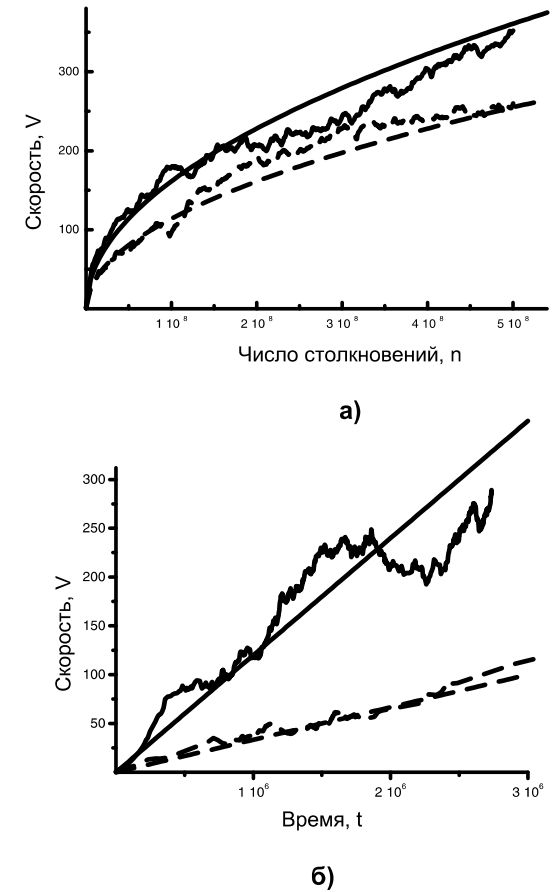


Рис. 2.46. Зависимости средней скорости частиц в газе Лоренца: от числа столкновений (а), от времени (б) и их аппроксимации на основе уравнений (2.89) и (2.90), соответственно

стиц в направлении положительных скоростей. Во-вторых, это дисперсионный (или флуктуационный) механизм, проявляющийся по двум причинам: (а) $\langle \Delta v^2 \rangle > 0$, и поэтому вероятностный пакет «расплывается» со временем; (б) модуль скорости не может быть отрицательным и, таким образом, расплывание пакета не симметрично, а направлено в сторону больших ско-

ростей, что выражается в замене нормального распределения распределением (2.88). При этом флуктуации и среднее увеличение скорости частицы выше в случае *регулярных* колебаний границ рассеивателей, что приводит к большему росту скорости. Следовательно, можно также говорить о механизме, обусловленном корреляциями между последовательными изменениями скорости.

19.3 Бильярд в форме «стадион». Замедление частиц

Как уже отмечалось, механизмы возникновения хаоса в рассеивающих и фокусирующих бильярдах различны. В первом случае два близких пучка частиц, попав на рассеивающую компоненту, начинают сразу расходиться, тогда как во втором случае после отражения от фокусирующей компоненты два близких пучка сходятся до точки пересечения. Хаос в таком бильярде возникает, если время, в течение которого пучки сходятся, меньше времени расхождения.

19.3.1 Неподвижная граница

Рассмотрим фокусирующий бильярд с невозмущенными границами и сравним результаты *точного* и *приближенного* анализа в предположении, что кривизна фокусирующих компонент мала, т. е. выполняется соотношение (2.76). Это позволит более полно понять динамику бильярдной частицы.

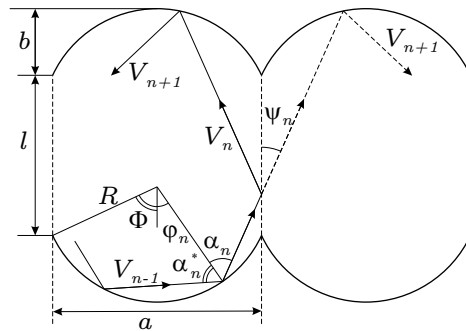


Рис. 2.47. Бильярд типа стадион с границей в виде дуги окружности

Пусть фокусирующие компоненты представлены дугами окружности радиуса R и угловой меры 2Φ , симметричными относительно вертикальной оси бильярда (рис. 2.47). Из геометрических соображений вытекают

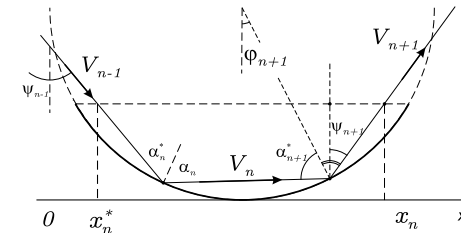


Рис. 2.48. Парные столкновения с границей

следующие соотношения:

$$R = \frac{a^2 + 4b^2}{8b}; \quad \Phi = \arcsin \frac{a}{2R}.$$

Условие стохастичности для такого бильярда выполнено, если дуга, дополняющая фокусирующую компоненту до полной окружности, принадлежит области Q (см. выше). При $b \ll a$ это дает:

$$\frac{l}{2R} \approx \frac{4bl}{a^2} > 1.$$

Введем динамические переменные, как показано на рис. 2.47. Положительным направлением для углов φ_n и α_n^* будем считать поворот против часовой стрелки, а для угла α_n — по часовой. Если граница бильярда неподвижна, то угол падения α_n^* равен углу отражения α_n . Пусть V_n — скорость частицы, а t_n — время n -го столкновения с границей.

Построим *точное преобразование*, описывающее динамику частицы в таком бильярде. Для этого необходимо рассмотреть два случая: после очередного столкновения с фокусирующей компонентой частица сталкивается с ней же (парные столкновения), либо следующее столкновение происходит с другой фокусирующей компонентой.

В случае парных столкновений геометрические рассуждения [407] приводят к соотношению вида (см. рис. 2.48):

$$\begin{aligned} \alpha_{n+1}^* &= \alpha_n, & \alpha_{n+1} &= \alpha_{n+1}^*, \\ \varphi_{n+1} &= \varphi_n + \pi - 2\alpha_n \pmod{2\pi}, & t_{n+1} &= t_n + \frac{2R \cos \alpha_n}{V_n}. \end{aligned} \quad (2.92)$$

Если $|\varphi_{n+1}| < \Phi$, то частица продолжает каскад столкновений с одной границей. В противном случае $(n + 1)$ -е столкновение произойдет с другой фокусирующей компонентой.

При переходе от одной фокусирующей компоненты к другой эти преобразования переписываются как

$$\begin{aligned} \alpha_{n+1}^* &= \arcsin \left[\sin(\psi_n + \Phi) - \frac{x_{n+1}^*}{R} \cos \psi_n \right], \\ \alpha_{n+1} &= \alpha_{n+1}^*, \\ \varphi_{n+1} &= \psi_n - \alpha_{n+1}^*, \\ t_{n+1} &= t_n + \frac{R(\cos \varphi_n + \cos \varphi_{n+1} - 2 \cos \Phi) + l}{V_n \cos \psi_n}, \end{aligned} \quad (2.93)$$

где введено $\psi_n = \alpha_n - \varphi_n$, $x_n = R / \cos \psi_n [\sin \alpha_n + \sin(\Phi - \psi_n)]$, $x_{n+1}^* = x_n + \text{tg} \psi_n \pmod{a}$.

На рис. 2.51a представлен фазовый портрет бильярда типа стадион, задаваемый соотношениями (2.92)–(2.93). В качестве координаты ξ_n выбиралась величина $\xi_n = 1/2 + (R \sin \varphi_n)/a$, равная нормированной на ширину бильярда проекции точки столкновения частицы с фокусирующей компонентой на ось Ox . Поскольку бильярд обладает осевой симметрией, то и исследуемые преобразования должны обладать симметрией относительно замены $\xi \rightarrow 1 - \xi$, $\psi \rightarrow -\psi$. Следовательно, достаточно получить результаты в области неотрицательных ψ . При численном анализе интервал изменения величин ξ и ψ разбивался на 300×300 ячеек. Интенсивность черного цвета на рисунке пропорциональна числу точек, попавших в данную ячейку. Для построения графика использовалось четыре траектории. Одна из них начиналась в стохастической области и включала $5 \cdot 10^8$ итераций. Три другие начинались в регулярной области и содержали по 10^7 итераций. Начальные условия отмечены на графике крестиком.

Из приведенного рисунка следует, что в зависимости от начальных условий фазовая точка либо совершает нерегулярное блуждание, формируя на фазовой плоскости так называемый *стохастический слой*⁷ (серый цвет), либо образует замкнутые инвариантные кривые, соответствующие регулярным колебаниям. Бильярдная частица, начав свое движение в стохастическом слое, посещает все доступные ей области случайным образом.

Ширина стохастического слоя определяется величиной нелинейности в системе. Если бильярд близок к прямоугольнику (подобно изображенному справа на рис. 2.40), то стохастический слой оказывается разделенным инвариантными кривыми. Эти инвариантные кривые окружают устойчивые неподвижные точки, т. е. точки x^* , в которых для преобразования (2.92)–(2.93) выполняется $x^* = f(x^*)$, $x = (\psi, \xi)$ (см. §24).

⁷Для обозначения хаоса в гамильтоновых системах часто также используется термин «стохастика». Отсюда и название — стохастический слой.

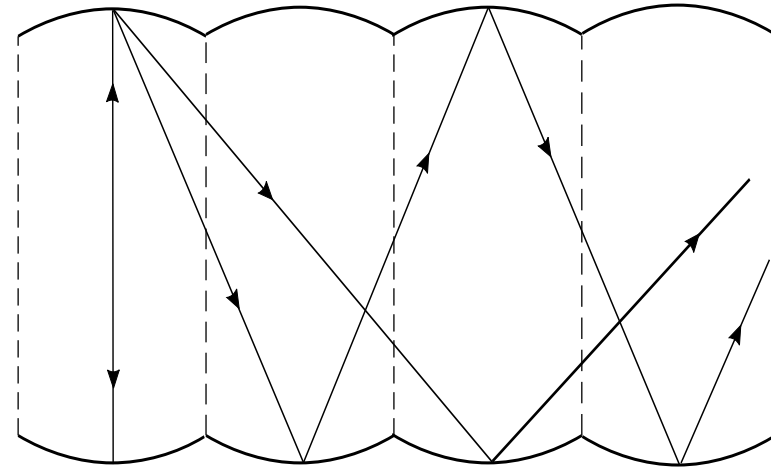


Рис. 2.49. Траектории в конфигурационном пространстве, отвечающие неподвижным точкам

В конфигурационном пространстве неподвижные точки отвечают траекториям, показанным на рис. 2.49, т. е. периодическому движению бильярдной частицы. Если бильярдную частицу, движущуюся по таким траекториям, немного возмутить, то после возмущения она останется в их малой окрестности, совершая в общем случае квазипериодическое движение. Нетрудно видеть, что на фазовой плоскости таким движениям как раз и соответствует множество инвариантных кривых, окружающих неподвижные точки.

С ростом нелинейности (т. е. увеличением параметра b , рис. 2.47) неподвижные точки могут потерять свою устойчивость. В результате образуется область глобальной стохастичности, в которой частице доступно уже все фазовое пространство, и она может диффундировать в любом направлении.

Для нахождения неподвижных точек и анализа их устойчивости используем определенные упрощения в (2.92)–(2.93). Аппроксимируем дуги окружности частью параболы, $\chi(x) = 4bx(x-a)/a^2$, а качестве координат выберем угол ψ между вектором скорости и вертикалью, $\psi \in [-\pi/2, \pi/2]$, а также величину x — проекцию на ось Ox точки столкновения частицы с границей, $x \in [0, a)$ (рис. 2.50). Тогда, принимая во внимание приближение (2.76), найдем:

$$x_{n+1} = x_n + \text{tg} \psi_{n+1} \pmod{a}, \quad \psi_{n+1} = \psi_n - 2\beta(x_{n+1}),$$

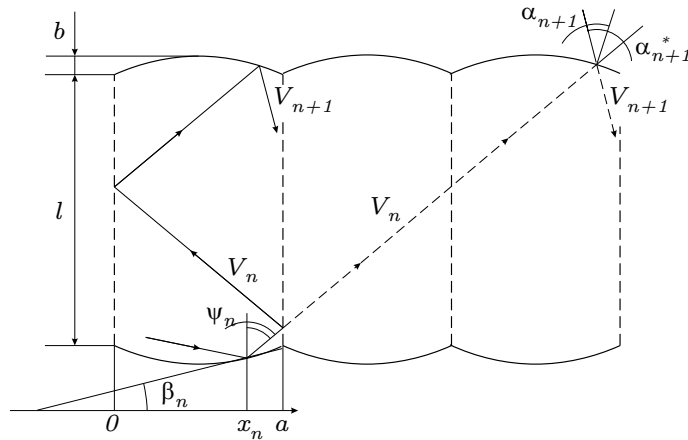


Рис. 2.50. Развертка бильярда типа «стадион»

где $\beta(x) = \arctg(\chi'(x))$ — угол наклона касательной в точке столкновения. В этом случае очевидно, что $\beta \approx 4b(2x - a)/a^2$.

Для дальнейших упрощений сделаем замену: $\xi = x/a$, $\xi \in [0, 1)$. Таким образом,

$$\xi_{n+1} = \xi_n + \frac{l}{a} \operatorname{tg} \psi_n \pmod{1}, \quad \psi_{n+1} = \psi_n - \frac{8b}{a} (2\xi_{n+1} - 1). \quad (2.94)$$

Одно из семейств неподвижных точек бильярда, очевидно, можно записать в виде $\{\xi = 1/2, \psi_s = \arctg(ma/l)\}$, где m — целое. Эти точки соответствуют столкновениям частицы с центром дуги. При этом, если $m = 0$, то частица движется строго вертикально, если $m = 1$, то при каждом соударении частица смещается на одну ячейку и т. д. (см. рис. 2.49). Исследуем устойчивость таких точек. Для этого проведем линеаризацию, сделав замену $\xi_n = \Delta\xi_n + 1/2$, $\psi_n = \Delta\psi_n + \arctg(ma/l)$. Тогда, раскладывая в ряд по $\Delta\psi$, найдем:

$$\Delta\xi_{n+1} = \Delta\xi_n + \frac{l}{a \cos^2 \psi_s} \Delta\psi_n + O(\Delta\psi_n^2), \quad \Delta\psi_{n+1} = \Delta\psi_n - \frac{16b}{a} \Delta\xi_{n+1},$$

где $\psi_s = \arctg(ma/l)$. При этом матрица преобразования имеет вид:

$$A = \begin{pmatrix} 1 & \frac{l}{a \cos^2 \psi_s} \\ -\frac{16b}{a} & 1 - \frac{16bl}{a^2 \cos^2 \psi_s} \end{pmatrix}.$$

Легко убедиться, что $\det A = 1$. Следовательно, полученное преобразование сохраняет площадь. Критерий устойчивости неподвижных точек можно записать как $|\operatorname{Tr} A| \leq 2$. Отсюда $\cos^2 \psi_s \geq 4bl/a^2$ или $m^2 \leq l/(4b) - l^2/a^2$.

Разрушение всех резонансов произойдет, если $4bl/a^2 > 1$. Нетрудно понять [407], что движение частицы вокруг устойчивой точки описывается отображением поворота с числом вращения (см. §28)

$$\rho = \arccos \left(1 - \frac{8bl}{a^2 \cos^2 \psi_s} \right).$$

Движение частицы внутри бильярда с определенной скоростью V порождает поток, в котором можно ввести время t . Время между последовательными столкновениями приблизительно равно $\tau \approx \frac{l}{\cos \psi_s} \frac{1}{V}$. Поэтому период вращения

$$T_1 = \frac{2\pi}{\rho} \tau = \frac{2\pi l \cos \psi_s}{(\cos^2 \psi_s - 8bl/a^2) V}.$$

На рис. 2.51б представлен фазовый портрет бильярда с фокусирующими компонентами в форме параболы, задаваемого *приближенным преобразованием* (2.94). Из его сравнения с рис. 2.51а, построенного на основе точного соотношения (2.92)–(2.93), видно, что во втором случае траектория равномерно распределена по всей стохастической области. Однако в первом случае большая плотность соответствует области $\psi \geq 0$, а область вблизи $\psi \approx \pi/2$ практически пуста. Разницу результатов объясняет рис. 2.52. Приближение малой глубины фокусирующей компоненты, сделанное при выводе (2.94), буквально означает следующее: при пересечении отрезка, соединяющего концы параболы в точке B , мы считаем, что соударение происходит в точке A , являющейся проекцией точки B на дугу. На самом деле соударение происходит в точке C . В результате для точного преобразования при большом ψ соударение будет происходить с большей вероятностью с правой (на рисунке) частью дуги, так что угол ψ будет уменьшаться. Таким образом, частицы как бы выталкиваются в область малых ψ и большая часть их движения проходит в окрестности регулярных областей.

19.3.2 Периодически возмущаемая граница

В случае невозмущенного бильярда движение частицы вокруг устойчивой точки происходит по эллипсу. Если граница бильярда возмущается, то частица может переходить из стохастической области в регулярную и обратно. Когда скорость частиц много больше скорости движения границ, можно

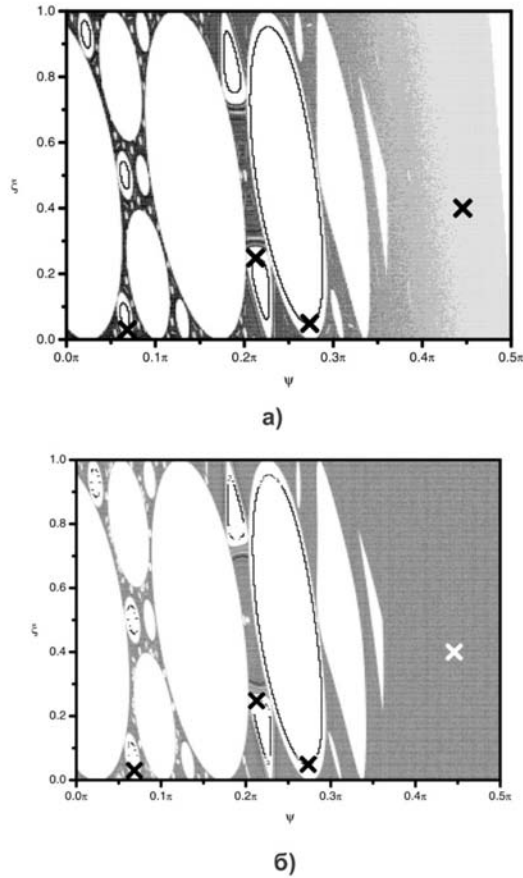


Рис. 2.51. Фазовый портрет бильярда типа стадион с границей в виде дуги окружности (а) и параболической границей (б) при $a = 0,5$; $b = 0,01$ и $l = 1$

показать [407], что частицы, попадая в окрестность устойчивых точек, будут двигаться по спирали, сначала приближаясь к центру, а потом удаляясь от него. При этом частота вращения вокруг устойчивой точки остается такой же, как и в невозмущенной системе. Кроме того, можно показать, что при определенной скорости частиц наступает резонанс между вращением частицы вокруг устойчивой точки и возмущением границ. В результате вся область в ее окрестности становится доступной. Резонансную скорость

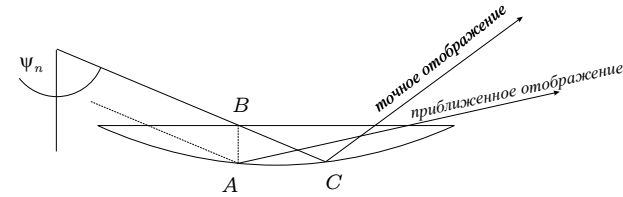


Рис. 2.52. Разница траекторий частицы в точном и приближенном преобразованиях

можно выразить следующим образом:

$$V_r = \frac{l}{2U_0 \cos \psi_s \arccos(1 - 8bl/(a \cos \psi_s)^2)}. \quad (2.95)$$

Пусть фокусирующие компоненты возмущаются таким образом, что скорость их движения в каждой точке одинакова по модулю и направлена по нормали к поверхности. Предположим, что модуль скорости границы зависит от времени по некоторому периодическому закону $U(t) = U_0 f(\omega(t + t_0))$, где ω — частота колебаний. Будем считать, что $U_0/\omega \ll l$, т. е. смещение границы мало. В этом случае соотношение, описывающее движение бильярдного шара, можно записать как

$$\begin{aligned} V_n &= \sqrt{V_{n-1}^2 + 4V_{n-1} \cos \alpha_n^* U_n + 4U_n^2}, \\ \alpha_n &= \arcsin \left(\frac{V_{n-1}}{V_n} \sin \alpha_n^* \right), \end{aligned} \quad (2.96)$$

$$\left. \begin{aligned} \alpha_{n+1}^* &= \alpha_n, \\ \varphi_{n+1} &= \varphi_n + \pi - 2\alpha_n \pmod{2\pi}, \\ t_{n+1} &= t_n + \frac{2R \cos \alpha_n}{V_n}, \end{aligned} \right\} \text{если } |\varphi_{n+1}| \leq \Phi \quad (2.97)$$

$$\left. \begin{aligned} \psi_n &= \alpha_n - \varphi_n, \\ x_n &= \frac{R}{\cos \psi_n} [\sin \alpha_n + \sin (\Phi - \psi_n)], \\ x_{n+1}^* &= x_n + l \operatorname{tg} \psi_n \pmod{a}, \\ \alpha_{n+1}^* &= \arcsin \left[\sin (\psi_n + \Phi) - \frac{x_{n+1}^*}{R} \cos \psi_n \right], \\ \varphi_{n+1} &= \psi_n - \alpha_{n+1}^*, \\ t_{n+1} &= t_n + \frac{R(\cos \varphi_n + \cos \varphi_{n+1} - 2 \cos \Phi) + l}{V_n \cos \psi_n}. \end{aligned} \right\} \text{если } |\varphi_n + \pi - 2\alpha_n| > \Phi \quad (2.98)$$

Отметим, что при его выводе *не использовалось* соотношение (2.76). Единственное приближение, которое было сделано, — учет малости смещения границ бильярда. Выражения (2.97) соответствуют серии последовательных столкновений частицы с одной фокусирующей компонентой, а (2.98) — переходу от одной границы к другой.

• Численный анализ

Численное исследование бильярда типа «стадион» с возмущаемой границей, описываемого преобразованиями (2.96)–(2.98), проводилось при тех же значениях геометрических параметров, что и в предыдущем пункте. При этом амплитуда и частота осцилляций границы составляли $U_0 = 0,01$ и $\omega = 1$, соответственно.

Как указывалось ранее, для разных скоростей частиц фазовые портреты такой системы должны отличаться. На рис. 2.53 представлена зависимость резонансной скорости от угла ψ_s , задаваемой выражением (2.95). Из него следует, что в области от 0 до $\psi_{s \max}$ ее значение меняется в узких пределах.

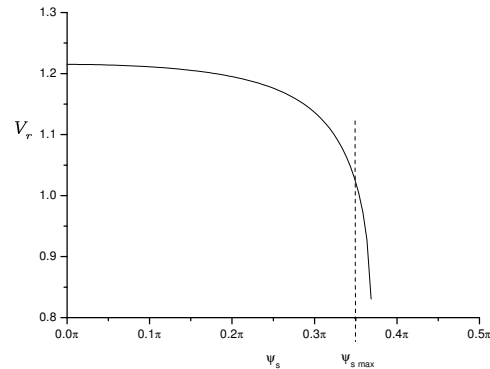


Рис. 2.53. Резонансная скорость в зависимости от ψ_s

Фазовые портреты для возмущенного бильярда представлены на рис. 2.54. Для исследования рассматривались ансамбли частиц с начальными скоростями V_0 разных направлений, выбираемыми случайным образом в стохастической области невозмущенного бильярда. Для построения фазового портрета плоскость была поделена на 300×300 ячеек. При попадании координат частицы (ξ_n, ψ_n) в определенную ячейку значение в ячейке увеличивалось на величину $\Delta V_n = V_n - V_{n-1}$. Таким образом, фазовый портрет дает представление о том, как в среднем происходит изменение скорости.

Вертикальная штриховка соответствует увеличению скорости, горизонтальная — уменьшению. Черным цветом показаны области, недоступные частице.

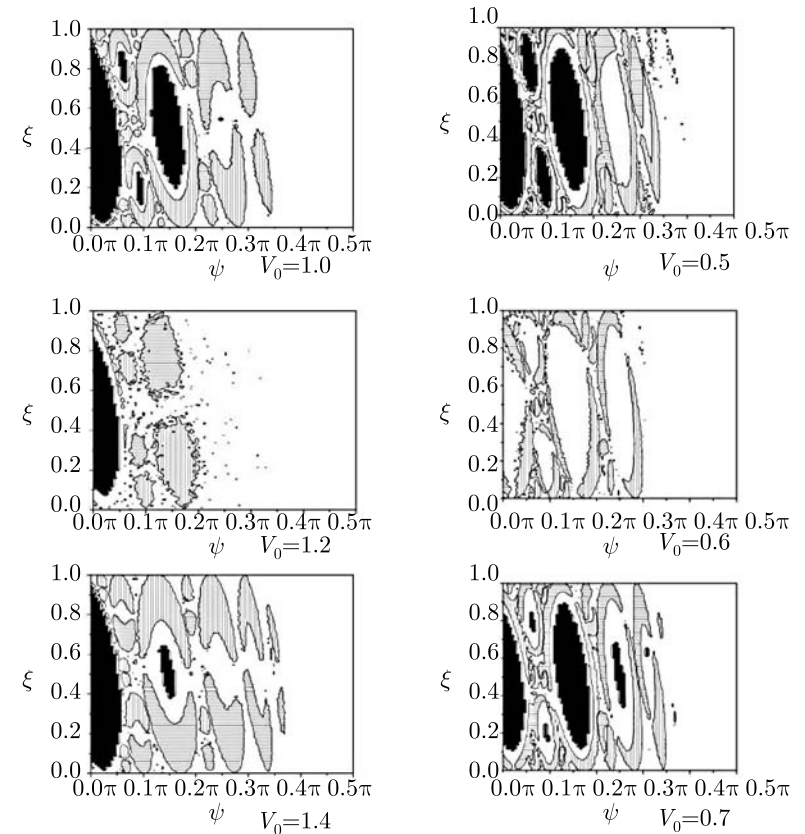


Рис. 2.54. Фазовые портреты изменения скорости в бильярде с возмущаемыми границами (2.96)–(2.98) при $b = 0.01$, $a = 0.5$, $l = 1$, $U_0 = 0.01$ и $\omega = 1$, $V_r = 1.2$

Нетрудно видеть, что если $V_0 \neq V_r$, то вокруг устойчивых точек существуют области, окруженные инвариантными кривыми. В окрестности устойчивых точек, ставшей доступной частице в результате возмущения системы, хорошо выделяются области преимущественного увеличения и уменьшения скорости. Они меняются местами в зависимости от того,

больше или меньше резонансной начальная скорость частиц. Если $V_0 = V_r$, то все окрестности устойчивых точек, кроме центральной ($\psi_0 = 0$, $\xi_0 = 1/2$), становятся доступными частице. При этом нет ярко выраженных областей ускорения частиц.

Можно показать [407], что резонансная скорость в окрестности центральной устойчивой точки в два раза меньше, чем получается из уравнения (2.95), и будет равна $V_r^0 \approx 0,6$. Из рис. 2.54 следует, что при $V_0 = V_r^0$ все области фазового пространства становятся доступными частице. Кроме того, области ускорения и замедления частицы в окрестности центральной устойчивой точки меняются местами при переходе через V_r^0 .

Анализ соотношений (2.96)–(2.98) был проведен в двух вариантах: для полностью хаотического и близкого к интегрируемому бильярду (рис. 2.55). В первом случае бильярд — классический стадион, т. е. $\Phi = \pi/2$, и граница бильярда состоит из двух полуокружностей и двух параллельных прямых, касательных к ним (левый бильярд на рис. 2.40). Близкий к интегрируемому случай означает, что фокусирующие компоненты — почти прямые (правый бильярд на рис. 2.40), так что выполняется соотношение (2.76). Для хаотического бильярда были выбраны следующие параметры: $a = 0,5$; $b = 0,25$; $l = 1$; $U_0 = 0,01$; $\omega = 1$ и $V_0 = 0,1$. Скорость частицы была вычислена как среднее значение ансамбля из 5000 траекторий с различными начальными условиями (сплошная кривая 1 на рис. 2.55). Начальные условия выбирались случайным образом на фокусирующей компоненте так, чтобы вектор скорости частицы был направлен внутрь бильярдной области. Как следует из численного анализа, полученная зависимость имеет приблизительно корневой характер ($V(n) \sim \sqrt{n}$).

В близком к интегрируемому бильярду параметр b достаточно мал и кривизна фокусирующей компоненты дает лишь слабую нелинейность в системе. В такой конфигурации невозмущенный бильярд имеет устойчивые точки, окруженные инвариантными кривыми. В их окрестности движение частицы квазипериодично. В то же время вне этих регионов динамика хаотична. Таким образом, если граница бильярда не возмущена, то, в зависимости от начальных условий, движение бильярдной частицы может быть либо хаотическим, либо регулярным.

Теперь обратимся к возмущаемому бильярду. В такой системе частица может двигаться из стохастической области в регулярную и обратно. В достаточно малой окрестности неподвижной точки поведение частицы имеет определенный период вращения, $T_1 = 2\pi l/(\rho V)$, где ρ — число вращения, l — средняя длина свободного пробега и V — скорость частицы. В то же время, период осцилляций границы $T_2 = 2\pi/\omega$. Таким образом, в системе при некоторой скорости V_r будет наблюдаться резонанс. Как

следует из численного анализа, зависимости скорости частицы от числа столкновений по разные стороны резонанса отличаются. Если начальная скорость $V_0 < V_r$, тогда скорость частицы уменьшается до конечной величины $V_{fin} < V_r$, и распределение частиц по скоростям стремится к стационарному в интервале $(0, V_{fin})$. Если же $V_0 > V_r$, то частицы могут достигать высоких скоростей. В этом случае распределение частиц не стационарно и неограниченно возрастает.

На рис. 2.55 показана зависимость скорости от числа столкновений (кривые 2–5, параметры бильярда те же, что и для бильярда «классический стадион» (кривая 1), за исключением $b = 0.01$).

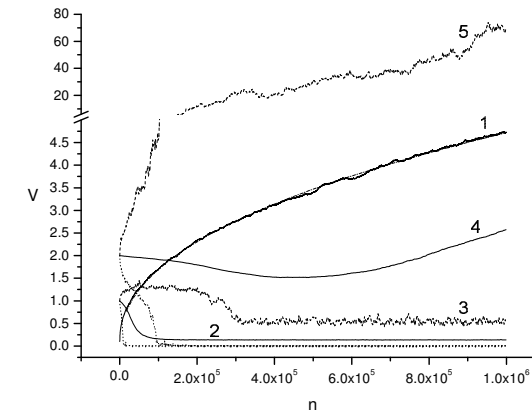


Рис. 2.55. Зависимость средней скорости частицы от числа итераций

На основании 5000 реализаций для каждой начальной скорости было построено три кривых: средняя, минимальная и максимальная скорости, достигнутые к n -ой итерации. Таким образом, можно определить интервал изменения скорости. Как следует из рисунка, если $V_0 < V_r$, тогда средняя скорость частиц (сплошная кривая 2) постепенно уменьшается и стремится к константе. Максимальная скорость частиц (пунктирная кривая 3) также уменьшается до величины V_{fin} и флуктуирует около этого значения. В конечном счете скорости частиц оказываются в интервале $0 < V < V_{fin}$. В случае $V > V_r$ минимальная скорость частиц также уменьшается. Это означает, что в ансамбле существуют частицы, попавшие в область малых скоростей. В численном анализе доля таких частиц составляла пример-

но 3/4. В то же время существуют частицы с большой скоростью (штриховая кривая 5, которая соответствует максимальной скорости в ансамбле). В результате средняя скорость (сплошная кривая 4) частиц будет расти.

Таким образом, в бильярде типа «стадион» динамика роста скорости зависит от его формы. Если форма такова, что поведение частицы в нем полностью хаотично, то, как и в газе Лоренца, средняя скорость ансамбля частиц растет как корень из числа столкновений. В то же время в близком к интегрируемому случаю наблюдается новое явление. В зависимости от начальной скорости ансамбль частиц может как *ускоряться*, так и *замедляться*, т. е. происходит *разделение частиц по скоростям*.

20. Диссипативные динамические системы

Таким образом, некоторые простые гамильтоновы системы могут обладать нерегулярными, или хаотическими, режимами движения. Появление таких режимов — это свойство самих систем, не связанное с воздействием на их динамику каких-либо случайных сил. В этом состоит отличительная особенность хаотичности. Но гамильтоновы системы, вообще говоря, — это достаточно узкий класс систем. Большинство же изучаемых систем — биологические, химические, экологические и т. д. — диссипативные, т. е. системы, для которых не справедлива теорема Лиувилля и фазовый объем которых со временем не остается постоянным, а сжимается. Сокращение фазового объема приводит к тому, что при $t \rightarrow \infty$ все решения диссипативной системы будут стягиваться к некоторому подмножеству фазового пространства, называемому *аттрактором*.

Изучим здесь некоторые свойства диссипативных динамических систем. Как мы увидим в дальнейшем, эти системы тоже могут иметь очень сложные режимы движения: в зависимости от того, каким является аттрактор — простым или сложным, — их динамика может быть, соответственно, либо регулярной, либо хаотической.

20.1 Свойства диссипативных систем

Поведение диссипативных динамических систем в большинстве случаев описывается n дифференциальными уравнениями 1-го порядка

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad (2.99)$$

где $\mathbf{x} = \{x_1, \dots, x_n\}$ — n -мерный вектор с компонентами x_1, \dots, x_n , а $\mathbf{v} = \{v_1, \dots, v_n\}$ — векторная функция, компонентами которой служат

величины v_1, \dots, v_n . Будем предполагать, что функция \mathbf{v} не зависит явно от времени, т. е. система (2.99) является автономной. Заметим, что любое дифференциальное уравнение, разрешенное относительно старшей производной, с помощью соответствующей замены переменных можно свести к виду (2.99). Если же система не автономна, то, выбрав в качестве нового времени независимую переменную τ и добавив уравнение $d\tau/dt = 1$, мы превратим исходную систему в автономную.

Фазовое пространство системы (2.99) (будем, как прежде, обозначать его M) образовано компонентами $\{x_1, \dots, x_n\}$ вектора \mathbf{x} и имеет размерность n . Решение уравнений (2.99) $x_i(t) = F_i(t, \mathbf{x}_0)$, $i = 1, 2, \dots, n$, задает определенную линию, или фазовую траекторию, в пространстве M . Совокупность фазовых траекторий $\mathbf{x}(t)$ для системы (2.99) называется n -мерным потоком. Если \mathbf{v} — гладкая функция, то решение $\mathbf{x}(t) = \mathbf{F}(t, \mathbf{x}_0)$ существует при всех t и каждой точке фазового пространства ставится в соответствие единственное состояние системы.

В силу теоремы Лиувилля, характерной особенностью движения гамильтоновых систем является сохранение их фазового объема. В противоположность этому в диссипативных динамических системах фазовый объем в среднем сжимается.

Найдем изменение малого элемента объема

$$\Delta\Omega = \prod_i \Delta x_i \quad (2.100)$$

в момент времени t , близкий к начальному ($t \rightarrow 0$) [116]:

$$\begin{aligned} \frac{d(\Delta\Omega)}{dt} &= \frac{d(\Delta x_1)}{dt} \prod_{i \neq 1} \Delta x_i + \frac{d(\Delta x_2)}{dt} \prod_{i \neq 2} \Delta x_i + \dots = \\ &= \prod_i \Delta x_i \frac{d(\Delta x_1)}{dt} \frac{1}{\Delta x_1} + \prod_i \Delta x_i \frac{d(\Delta x_2)}{dt} \frac{1}{\Delta x_2} + \dots \end{aligned} \quad (2.101)$$

Отсюда, учитывая (2.100), получаем

$$\frac{d(\Delta\Omega)}{dt} = \Delta\Omega \sum_i \frac{1}{\Delta x_i} \frac{d(\Delta x_i)}{dt}. \quad (2.102)$$

Обозначим

$$\Xi \equiv \frac{1}{\Delta\Omega} \frac{d(\Delta\Omega)}{dt} = \sum_i \frac{1}{\Delta x_i} \frac{d(\Delta x_i)}{dt}. \quad (2.103)$$

Принимая во внимание, что

$$\Delta x_i = \frac{\partial x_i}{\partial x_i(0)} \Delta x_i(0) \quad (2.104)$$

и, значит,

$$\frac{d(\Delta x_i)}{dt} = \Delta x_i(0) \frac{d}{dt} \frac{\partial x_i}{\partial x_i(0)}, \quad (2.105)$$

из (2.103) и (2.99) при $t \rightarrow 0$ находим

$$\Xi \Big|_{t \rightarrow 0} = \sum_i \frac{\Delta x_i(0)}{\Delta x_i} \frac{d}{dt} \frac{\partial x_i}{\partial x_i(0)} \Big|_{t \rightarrow 0} = \sum_{i=1}^n \frac{\partial v_i}{\partial x_i} = \operatorname{div} \mathbf{v}. \quad (2.106)$$

Таким образом, если

$$\operatorname{div} \mathbf{v} = 0, \quad (2.107)$$

то система является консервативной: при движении ее фазовый объем сохраняется. Если же

$$\operatorname{div} \mathbf{v} < 0, \quad (2.108)$$

то система будет диссипативной и ее фазовый объем со временем сокращается.

В конечном счете (т. е. при $t \rightarrow \infty$) все решения диссипативной системы сосредоточатся на некотором подмножестве \mathcal{A} фазового пространства M . Это подмножество \mathcal{A} называется аттрактором.

20.2 Простые аттракторы

В литературе *аттрактором* обычно называют такое подмножество \mathcal{A} фазового пространства M , которое удовлетворяет следующим условиям [115, 202]:

- а) \mathcal{A} инвариантно относительно потока динамической системы;
- б) существует окрестность U , которая сжимается к \mathcal{A} под действием потока;
- в) \mathcal{A} нельзя разложить на два и более непересекающихся инвариантных подмножества.

Поясним смысл приведенного определения. Решения системы $\mathbf{x}(t) = \mathbf{F}(t, \mathbf{x}_0)$ можно рассматривать как закон, согласно которому начальной точке $\mathbf{x}(0)$ ставится в соответствие определенная точка $\mathbf{x}(t)$ в момент времени t . Поэтому, как и ранее (см. §18), можно ввести оператор эволюции,

действующий в фазовом пространстве, $\mathbf{F}^t : M \rightarrow M$. Инвариантность аттрактора \mathcal{A} относительно потока означает, что $\mathbf{F}^t \mathcal{A} = \mathcal{A}$, т. е. попав на аттрактор, мы остаемся там при любом $t \rightarrow \infty$.

Второе утверждение означает, что \mathcal{A} является подмножеством множества U , $\mathcal{A} \subset U \subset M$, и при $t \rightarrow \infty$ имеем $\mathbf{F}^t U \rightarrow \mathcal{A}$. Фактически здесь речь идет об *области притяжения* аттрактора \mathcal{A} , т. е. совокупности таких начальных точек $\mathbf{x}_0 \in U$, что при $t \rightarrow \infty$ фазовые траектории, начавшиеся в этих точках, стремятся к аттрактору \mathcal{A} .

Третье утверждение необходимо для того, чтобы в фазовом пространстве разделять аттракторы между собой.

Отметим, что в настоящее время еще нет единого мнения по поводу наилучшего определения аттрактора и, помимо приведенного, принадлежащего Лэнфорду, существуют и другие определения (см., например, [72, 429, 438]).

Обычно динамическая система имеет конечное число аттракторов $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_k$, однако известны простые системы, которые могут иметь бесконечно много аттракторов [202, 439]. Для любой динамической системы, имеющей аттракторы, все начальные точки в фазовом пространстве, кроме множества меры нуль, лежат в области притяжения одного из них.

Вернемся к системе (2.99). Точки \mathbf{x}^0 , в которых правая часть $\mathbf{v}(\mathbf{x})$ уравнений (2.99) обращается в нуль, т. е. $\mathbf{v}(\mathbf{x}^0) = 0$, называются *положениями равновесия*, или *стационарными* (или иногда *особыми*), точками системы (2.99). Положение равновесия может быть как устойчивым, так и неустойчивым. Понятие устойчивости динамических систем имеет большое прикладное значение и является одним из наиболее важных понятий. В частности, характер эволюции некоторой системы из состояния равновесия существенно зависит от того, каким это равновесие является: устойчивым или неустойчивым. В случае неустойчивого равновесия в результате даже очень малых начальных отклонений система может быть отброшена от стационарного состояния и движение станет либо очень сложным, либо система перейдет в другое стационарное состояние, весьма далекое от исходного. Поэтому важное значение имеет изучение характера устойчивости динамических систем.

Говорят, что система находится в положении *устойчивого равновесия* \mathbf{x}^0 , если при малых отклонениях от него система останется вблизи \mathbf{x}^0 при любых t . На математическом языке это утверждение формулируется следующим образом. Стационарная точка \mathbf{x}^0 системы (2.99) называется устойчивой или, точнее, *устойчивой по Ляпунову*, если для любого $\epsilon > 0$ существует $\delta(\epsilon) > 0$, такое, что для всякого решения $\mathbf{x}(t)$ той же системы,

начальные значения которого удовлетворяют неравенству

$$\rho(\mathbf{x}(t_0), \mathbf{x}^0) < \delta(\epsilon), \quad (2.109)$$

для всех $t \geq t_0$ справедливо

$$\rho(\mathbf{x}(t), \mathbf{x}^0) < \epsilon. \quad (2.110)$$

Здесь $\rho(\dots)$ — расстояние между фазовыми точками с координатами $\mathbf{x}(\cdot)$, \mathbf{x}^0 .

Если же положение равновесия \mathbf{x}^0 не только устойчиво, но, кроме того, удовлетворяет условию

$$\lim_{t \rightarrow \infty} \rho(\mathbf{x}(t), \mathbf{x}^0) = 0, \quad (2.111)$$

то стационарная точка \mathbf{x}^0 в этом случае называется *асимптотически устойчивой*.

Предположим, что система (2.99) в некоторый момент времени t_0 находилась в состоянии $\mathbf{x}(t_0)$, близком к состоянию равновесия:

$$\mathbf{x}(t_0) = \mathbf{x}^0 + \boldsymbol{\xi}(t), \quad (2.112)$$

где $\boldsymbol{\xi}(t)$ — малая по сравнению с \mathbf{x}^0 величина. При каких условиях такая система и для $t > t_0$ останется вблизи \mathbf{x}^0 , т. е. будет устойчивой? Ответ на этот вопрос дает *теорема Ляпунова об устойчивости по первому приближению*. Суть ее заключается в следующем. Рассмотрим систему (2.99) при условии (2.112). Тогда, переходя в (2.99) к переменной $\boldsymbol{\xi}(t)$ и разлагая правую часть в ряд, получаем

$$\dot{\xi}_i = v_i(\mathbf{x}^0 + \boldsymbol{\xi}) = v_i(\mathbf{x}^0) + \sum_k \left. \frac{\partial v_i}{\partial x_k} \right|_{\mathbf{x}^0} \xi_k + R_i, \quad (2.113)$$

$$R_i = \frac{1}{2} \sum_{j,m} \left. \frac{\partial^2 v_i}{\partial x_j \partial x_m} \right|_{\mathbf{x}^0} \xi_j \xi_m + \dots$$

Поскольку $v(\mathbf{x}^0) = 0$, то уравнение (2.113) примет вид

$$\dot{\xi}_i = \sum_{k=1}^n \left. \frac{\partial v_i}{\partial x_k} \right|_{\mathbf{x}^0} \xi_k + R_i. \quad (2.114)$$

Если теперь отбросить нелинейные члены, то вместо (2.99) получим линеаризованную систему

$$\dot{\xi}_i = \sum_{k=1}^n a_{ik} \xi_k, \quad (2.115)$$

где $a_{ik} = \left. \frac{\partial v_i}{\partial x_k} \right|_{\mathbf{x}^0}$ — элементы *матрицы линеаризации* \mathbf{A} системы (2.99). Теорема Ляпунова утверждает, что если все собственные значения λ_i матрицы \mathbf{A} удовлетворяют неравенству $\text{Re } \lambda_i < 0$, $i = 1, 2, \dots, n$, то положение равновесия \mathbf{x}^0 исходной системы (2.99) асимптотически устойчиво. Если же среди всех собственных значений λ_i существует хотя бы одно такое, что $\text{Re } \lambda_k > 0$, то особая точка \mathbf{x}^0 системы (2.99) является неустойчивой. Наконец, если среди λ_i встречаются такие, для которых $\text{Re } \lambda_k = 0$, то устойчивость положения равновесия \mathbf{x}^0 нелинейной системы (2.99), вообще говоря, не может быть определена из анализа линеаризованной системы.

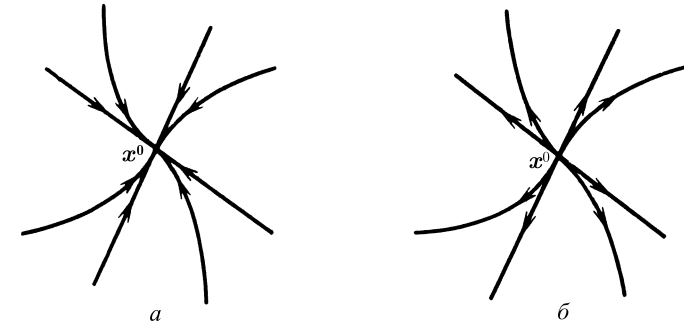


Рис. 2.56. Устойчивый (а) и неустойчивый (б) узлы

Собственные значения λ_i матрицы линеаризации \mathbf{A} , как известно, определяются из соотношения $\det[a_{ik} - \lambda \delta_{ik}] = 0$, которое называется *характеристическим уравнением* матрицы \mathbf{A} . Следовательно, изучение устойчивости стационарных точек системы сводится к алгебраической задаче на собственные значения и, в конечном счете, к исследованию корней характеристического уравнения.

Рассмотрим в качестве иллюстрации систему двух дифференциальных уравнений

$$\dot{x} = v_1(x, y), \quad \dot{y} = v_2(x, y). \quad (2.116)$$

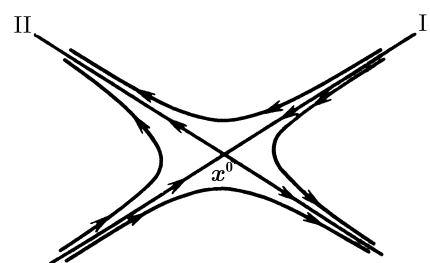


Рис. 2.57. Седловая точка

В зависимости от значения ее характеристических чисел λ_i особые точки такой системы могут быть только четырех различных типов.

1. Корни λ_1 и λ_2 характеристического уравнения действительные и одного знака. Особая точка системы называется *узлом* (рис. 2.56).

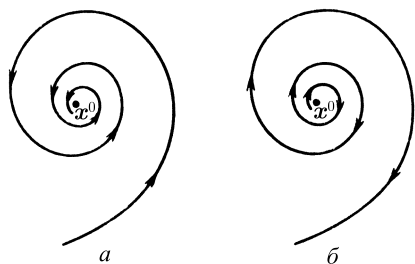


Рис. 2.58. Устойчивый (а) и неустойчивый (б) фокусы

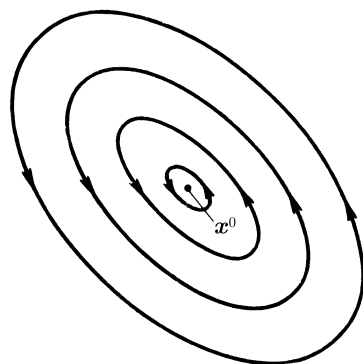
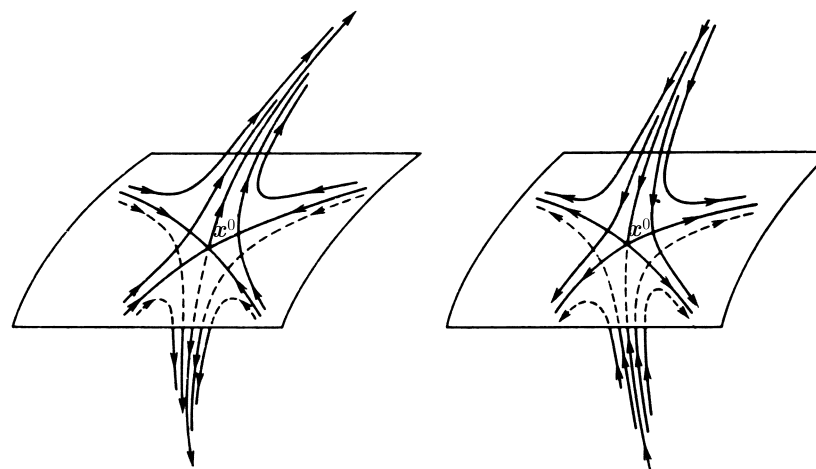
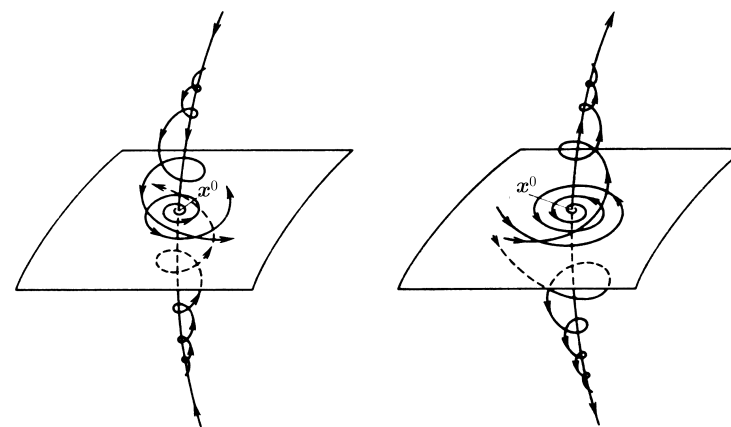


Рис. 2.59. Центр

2. Корни λ_1 и λ_2 действительные и разных знаков. Такая неустойчивая стационарная точка системы называется *седлом* (рис. 2.57). Траектории I и II, проходящие через седло, называются *сепаратрисами*.

3. Корни λ_1 и λ_2 характеристического уравнения — комплексно-сопряженные (но не чисто мнимые). В этом случае особая точка называется *фокусом* (рис. 2.58).

4. Корни λ_1 и λ_2 чисто мнимые. Этот случай встречается только в консервативных (но не диссипативных) системах. Особая точка называется *центром* (рис. 2.59).

Рис. 2.60. Стационарная точка x^0 типа седло-узелРис. 2.61. Стационарная точка x^0 типа седло-фокус

Кроме перечисленных, в многомерных системах встречаются и более сложные типы стационарных точек. Например, возможны комбинации 1-го

и 2-го, а также 2-го и 3-го типов. Особые точки в этих случаях называются, соответственно, *седло-узел* (рис. 2.60) и *седло-фокус* (рис. 2.61).

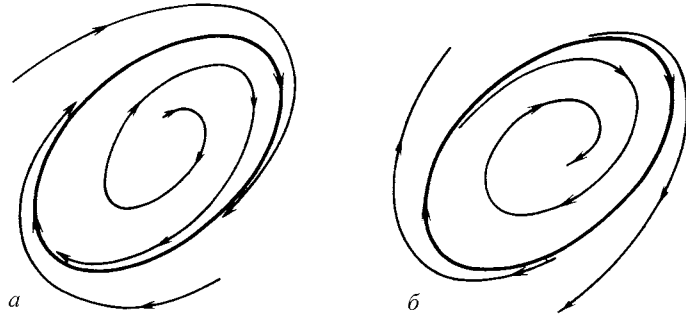


Рис. 2.62. Устойчивый (а) и неустойчивый (б) предельный циклы

Нетрудно заметить, что притягивающие особые точки, такие как устойчивый узел и устойчивый фокус, являются аттракторами. Но аттракторами в диссипативных системах могут быть не только устойчивые стационарные точки, но и замкнутые фазовые кривые, соответствующие периодическому движению. Такие изолированные замкнутые траектории называются *предельными циклами*. Устойчивые предельные циклы являются аттракторами. Они обладают тем свойством, что в их достаточно малой окрестности нет других замкнутых траекторий, а все остальные фазовые кривые из этой окрестности наматываются на эту единственную замкнутую траекторию (рис. 2.62а). Если все траектории сматываются с предельного цикла, то он является *абсолютно неустойчивым* (рис. 2.62б). В этом случае предельный цикл, очевидно, не будет аттрактором.

20.2.1 Элементы теории Флоке

Вопросы устойчивости предельных циклов рассматриваются в теории Флоке, объектом исследования которой служат линейные дифференциальные уравнения с периодическими коэффициентами. Эта теория связывает, как будет понятно из дальнейшего, отображение Пуанкаре в окрестности предельного цикла и уравнение в вариациях (т. е. линеаризованную систему).

Пусть $x(t)$ — периодическое решение системы с периодом T : $x(t) = x(t + T)$. Рассмотрим малое возмущение такого решения в некоторый

момент времени t_0 ,

$$\tilde{x}(t_0) = x(t_0) + \xi(t_0),$$

и проследим, при каких условиях из малости $\xi(t_0)$ вытекает малость $\xi(t)$ при $t > t_0$. Подставив $\tilde{x}(t_0)$ в исходную систему (2.99), разложив функцию v вблизи значений $x(t)$ в ряд по степеням ξ и оставив в разложении члены, содержащие ξ в степени не выше первой, получим линейную систему дифференциальных уравнений первого приближения с периодическими коэффициентами:

$$\dot{\xi}_i = \sum_{k=1}^n a_{ik}(t) \xi_k(t), \quad a_{jk}(t) = \left. \frac{\partial v_j}{\partial x_k} \right|_{x(t)}.$$

Эта система часто называется системой уравнений в вариациях вдоль фазовой кривой. Запишем его в матричной форме:

$$\dot{\xi} = A\xi, \quad (2.117)$$

где $A(t) = A(t + T)$ — T -периодическая $n \times n$ матрица. Фундаментальное матричное решение системы (2.117) — это зависящая от времени матрица $\hat{\Xi}$, такая, что $\xi(t) = \hat{\Xi}(t)\mathbf{l}$, где \mathbf{l} — некоторый постоянный вектор. Таким образом, $\hat{\Xi}(t)$ представляет собой матрицу, столбцы которой являются компонентами линейно независимых решений системы (2.117).

Согласно теории Флоке каждое фундаментальное матричное решение можно представить в виде

$$\hat{\Xi}(t) = \hat{Y}(t)e^{\hat{B}t}, \quad (2.118)$$

где $\hat{Y}(t + T) = \hat{Y}(t)$ — некоторая дифференцируемая по t матрица $n \times n$, \hat{B} — постоянная $n \times n$ матрица и

$$e^{\hat{B}} = \hat{I} + \hat{B} + \frac{1}{2!} \hat{B}^2 + \dots + \frac{1}{k!} \hat{B}^k + \dots,$$

$$e^{\hat{B}t} = e^{t\hat{B}} = \hat{I} + t\hat{B} + \frac{t}{2!} \hat{B}^2 + \dots + \frac{t^k}{k!} \hat{B}^k + \dots$$

Далее, из T -периодичности A вытекает, что $\hat{\Xi}(t + T)$ тоже будет фундаментальным решением. Поэтому существует постоянная невырожденная матрица \hat{L} , такая, что

$$\hat{\Xi}(t + T) = \hat{\Xi}(t)\hat{L}. \quad (2.119)$$

Тогда из общей теории (см., например, [190]) и условия $\det \hat{L} \neq 0$ следует существование (но не единственность) матрицы \hat{B} , обладающей свойством $\hat{L} = e^{\hat{B}T}$ или $\ln \hat{L} = \hat{B}T$. Значит, $\hat{\Xi}(t+T) = \hat{\Xi}(t)e^{\hat{B}T}$. Теперь из соотношения $\hat{Y}(t) = \hat{\Xi}(t)e^{-\hat{B}t}$ (см. (2.118)) легко найти, что

$$Y(t+T) = \hat{\Xi}(t+T)e^{-\hat{B}(t+T)} = \hat{\Xi}(t+T)e^{-\hat{B}T}e^{-\hat{B}t} = \hat{\Xi}(t)e^{-\hat{B}t} = Y(t).$$

Заметим, что матрица \hat{B} и ее собственные значения определяются неоднозначно системой (2.117). Например, в соотношении (2.118) матрицу $\hat{Y}(t)$ можно заменить матрицей $\hat{Y}(t)e^{-2\pi it}$, а матрицу \hat{B} — матрицей $\hat{B} + 2\pi i\hat{I}$. Однако собственные значения матрицы $e^{\hat{B}T}$ единственным образом определяются матрицей \mathbf{A} . Сама матрица $e^{\hat{B}T}$ определяется фундаментальной матрицей $\hat{\Xi}(t)$ так, что при замене $\hat{\Xi}(t)$ на $\hat{\Xi}(t)\hat{L}_0$, где \hat{L}_0 — постоянная невырожденная матрица, $e^{\hat{B}T}$ заменяется на $\hat{L}_0 e^{\hat{B}T} \hat{L}_0^{-1}$.

Матрица \hat{L} называется *матрицей монодромии*⁸, а ее собственные значения $\alpha_j, j = 1, 2, \dots, n$, — *мультипликаторами* системы (2.117). Очевидно, они могут быть как вещественными, так и комплексными. Когда фазовая траектория вблизи предельного цикла совершает один оборот, возмущение, отвечающее каждому собственному вектору, умножается на соответствующий мультипликатор.

Собственные числа β_j матрицы \hat{B} называются *показателями Флоке*. Они связаны с мультипликаторами соотношением $\alpha_j = e^{\beta_j T}$. Таким образом, вещественные компоненты показателей Флоке определены однозначно, однако мнимые части задаются только с точностью до $2\pi ik/T, k = \pm 1, \pm 2, \dots$.

Пусть $\xi(t)$ — некоторое T -периодическое решение системы (2.117). Тогда, поскольку $\xi(t) = \hat{Y}(t)e^{\hat{B}t}\mathbf{1}$, где $\mathbf{1}$ — постоянный вектор, найдем: $e^{\hat{B}T}\mathbf{1} = \mathbf{1}$. Стало быть, $\mathbf{1}$ будет собственным вектором матрицы $\hat{L} = e^{\hat{B}T}$, которому соответствует мультипликатор $\alpha = 1$. Поэтому матрица \hat{B} будет иметь собственное значение $\beta = 0 \pmod{2\pi ik/T}$.

Качественно данный результат соответствует ситуации, когда начальное возмущение ξ выбирается по касательной вдоль траектории. Тогда с течением времени это возмущение будет изменяться периодически с тем же периодом, что и невозмущенная траектория $x(t)$. Следовательно, один

⁸Слово *монодромия* происходит от двух составляющих: «моно» — один и греч. «дром» — круг. Таким образом, матрица монодромии — это матрица, полученная через один оборот (период) цикла.

из показателей Флоке периодического решения всегда оказывается равным нулю, а соответствующий ему мультипликатор — единице.

Теперь нетрудно понять основной результат, касающийся устойчивости периодических решений: если один из мультипликаторов предельного цикла γ равен единице, а все остальные по модулю меньше единицы, то такой цикл асимптотически устойчив. Если хотя бы один из мультипликаторов по модулю больше единицы, то периодическое решение неустойчиво, т. е. существуют фазовые траектории, удаляющиеся от γ . Таким образом, цикл γ устойчив, если $n-1$ мультипликатор на комплексной плоскости ($\text{Im}\alpha, \text{Re}\alpha$) лежит внутри единичной окружности.

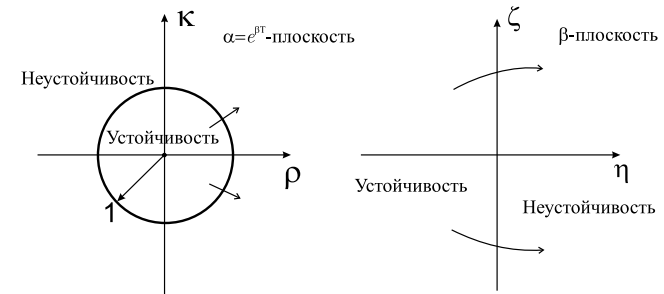


Рис. 2.63. Мультипликаторы и показатели Флоке [90]

Все изложенное можно обобщить следующим образом. Пусть $\alpha = \rho \pm i\kappa, \beta = \eta \pm i\zeta$. Представим собственные значения на плоскостях α и β (рис. 2.63). Тогда повторяющиеся точки $i\zeta \pm 2\pi ik/T$ на мнимой оси β -плоскости отобразятся в одни и те же точки α -плоскости. Если пара комплексно-сопряженных показателей Флоке пересекает мнимую ось β -плоскости, то при этом пара комплексно-сопряженных мультипликаторов выходит из единичного круга (на рис. 2.63 это обозначено стрелками). При пересечении показателем Флоке мнимой оси в начале координат ($\beta = 0$) мультипликатор выходит из единичного круга в точке $\alpha = \rho = 1$. Когда показатель Флоке пересекает мнимую ось в точках $\beta = \pm i\pi/T$, это соответствует мультипликатору $\alpha = \rho = -1$.

Описанные способы потери устойчивости предельным циклом являются типичными, т. е. они встречаются в динамических системах общего положения (см. §28).

Рассмотрим теперь, как геометрически связаны мультипликаторы α и показатели Флоке β . Очевидно, эта связь реализуется через отображение Пуанкаре.

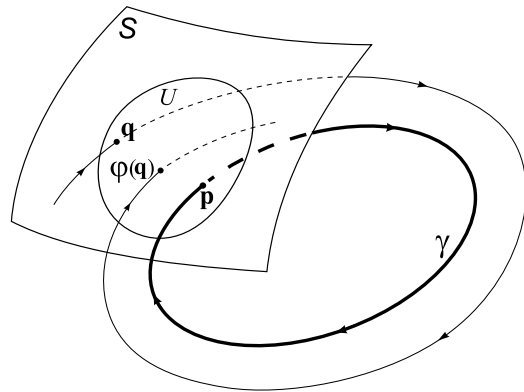


Рис. 2.64. Отображение Пуанкаре в окрестности предельного цикла

Предположим для наглядности, что исходная система (2.99) трехмерна, $n = 3$. Пусть $\gamma = \mathbf{x}(t)$ — предельный цикл и S — секущая поверхность, которую этот цикл пересекает в точке \mathbf{p} . Выберем некоторую окрестность $U \subset S$ этой точки (рис. 2.64) и рассмотрим первое возвращение траектории, вышедшей из близкой к \mathbf{p} точки $\mathbf{q} \in U$. Тогда можно определить отображение Пуанкаре $\Phi : U \rightarrow S$, переводящее точку \mathbf{q} в точку $\mathbf{q}' = \varphi(\mathbf{q})$ на поверхности S .

Очевидно, в координатном представлении в окрестности точки \mathbf{p} отображение Пуанкаре можно записать в виде

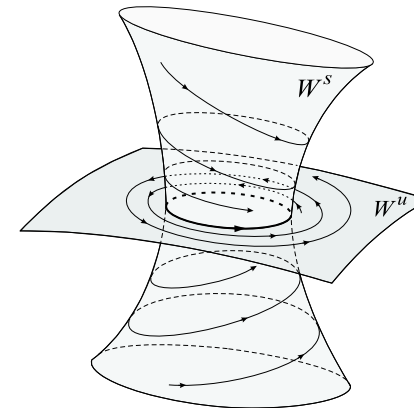
$$\mathbf{y}' = \hat{L}\mathbf{y} + g(\mathbf{y}),$$

где \hat{L} — матрица 2×2 с элементами $\left. \frac{\partial \varphi_i}{\partial y_k} \right|_{\mathbf{y}=\mathbf{p}}$. Таким образом, линеаризованное отображение есть

$$\mathbf{y}' = \hat{L}\mathbf{y}.$$

Матрица \hat{L} и есть матрица монодромии. Каждому ее собственному значению отвечает собственный вектор, задающий инвариантное направление. Один из таких векторов всегда будет направлен по касательной вдоль траектории. Поэтому соответствующий ему мультипликатор оказывается равным единице.

При каждом обходе предельного цикла и возвращении траектории на S возмущение умножается на соответствующий мультипликатор. Следовательно, цикл будет устойчив, когда все мультипликаторы, за исключением

Рис. 2.65. Устойчивое W^s и неустойчивое W^u многообразия седлового предельного цикла

одного, отвечающего возмущению вдоль траектории, удовлетворяют условию $|\alpha| < 1$.

Предельные циклы γ , для которых один из мультипликаторов лежит внутри единичного круга, а другой — снаружи, называются *седловыми*. Седловой предельный цикл характеризуется тем, что при $t \rightarrow \infty$ одни фазовые кривые динамической системы будут наматываться на него, образуя *устойчивое многообразие* W^s седлового цикла γ , а другие — сматываться с него, образуя *неустойчивое многообразие* W^u (рис. 2.65).

В общем случае часть мультипликаторов будет лежать внутри единичного круга, а часть — снаружи. Количество мультипликаторов, лежащих внутри единичного круга, определяет размерность пересечения устойчивого многообразия W^s с секущей S . Аналогично, размерность пересечения неустойчивого многообразия W^u и S равна числу мультипликаторов, лежащих вне единичного круга.

В двумерных диссипативных динамических системах, поскольку фазовые траектории не могут пересекаться, возможны аттракторы только двух типов: устойчивые стационарные точки и устойчивые предельные циклы. Однако для систем с размерностью фазового пространства $n = 3$ динамика уже не исчерпывается этими двумя простыми случаями. Кроме стационарных точек и предельных циклов, в таких системах могут существовать и более сложные аттракторы, в частности *двумерные инвариантные торы*, отвечающие квазипериодическому движению с двумя рационально независимыми частотами.

В многомерных системах $n > 3$ возможно возникновение еще более сложного квазипериодического движения, когда в фазовом пространстве рождаются торы более высокой размерности. Но в общем случае многомерные притягивающие торы являются неустойчивыми и под действием всегда присутствующих в динамической системе возмущений разрушаются. Кроме того, возникновению режима движения с большим числом несоизмеримых частот препятствует явление, называемое *синхронизацией* колебаний (см. в [113] § 30). Синхронизация может происходить несколькими способами и заключается в том, что в многомерных системах колебания с независимыми частотами сложным образом влияют друг на друга, что приводит к исчезновению квазипериодического и установлению периодического режима движения с соизмеримыми частотами, т. е. предельному циклу на торе. По этим причинам появление многомерных торов в фазовом пространстве диссипативных динамических систем не является случаем общего положения, т. е. это скорее исключение, чем правило.

Исследовать квазипериодическое движение на устойчивость значительно сложнее, чем периодическое, и для произвольных систем аналитическое рассмотрение, аналогичное анализу устойчивости периодического движения, возможно далеко не всегда. Однако имеется подход, основанный на применении отображения Пуанкаре. Рассечем тор некоторой поверхностью S без касания так, чтобы в сечении получилась инвариантная замкнутая кривая γ . Тогда при некоторых условиях [156, 158] отображение последования в окрестности γ может быть рассмотрено как отображение сдвига (см. раздел 2.18), т. е. на секущей S представлено дифференциальными уравнениями.

Периодическому решению этих уравнений будет соответствовать замкнутая кривая γ , и исследование устойчивости исходного тора сводится к уже рассмотренной ранее задаче — задаче об устойчивости периодического движения.

Наглядный пример использования подобной процедуры для системы, обладающей высокой степенью симметрии, приведен в [192]. Но если изучаемая система достаточно сложна, то применение указанных методов становится чрезвычайно громоздким, если вообще возможным, и, как правило, наталкивается на непреодолимые трудности. Поэтому в настоящее время широко используются численные методы при исследовании устойчивости квазипериодического движения диссипативных динамических систем.

20.3 Странные аттракторы

Все перечисленные аттракторы — устойчивые стационарные точки, устойчивые предельные циклы и инвариантные торы — называются *про-*

стыми аттракторами, поскольку динамика систем с такими аттракторами не является хаотической и носит, самое сложное, эргодический характер. Простые аттракторы являются подмножествами фазового пространства динамических систем. Подмножеством пространства M называют всякое подмножество W в пространстве M' ($M' \subseteq M$), такое, что локально оно выглядит как «кусочек» пространства M' и имеет в каждой точке единственную касательную гиперплоскость, т. е. W вложено в M гладко. Например, предельный цикл и двумерный инвариантный тор — это, соответственно, одномерное и двумерное подмножества. Но в диссипативных динамических системах, размерность фазового пространства которых $n \geq 3$, в дополнение к простым аттракторам, могут существовать ограниченные притягивающие множества, которые являются аттракторами и одновременно не являются подмножествами. Такие аттракторы, различные, в том числе, и от конечного объединения подмножеств фазового пространства, называются *странными аттракторами* [29]. Термин «странный аттрактор» был введен Д. Рюэлем и Ф. Такенсом в [173] и означал аттрактор, отличный от стационарной точки и предельного цикла.

Сжатие фазового объема диссипативной динамической системы приводит к тому, что фазовые кривые с течением времени стягиваются к предельному множеству — странному аттрактору и, попав в область, занятую им, остаются там навсегда. На самом же аттракторе движение является неустойчивым: любые две траектории системы расходятся экспоненциально быстро, оставаясь, разумеется, на странном аттракторе. Иначе говоря, поведение системы со странным аттрактором характеризуется сочетанием глобального сжатия фазового объема с локальной неустойчивостью фазовых траекторий.

Чтобы лучше представить себе множество, являющееся странным аттрактором, рассмотрим простой пример [482].

Возьмем бесконечно тонкий прямоугольный лист, растянем его вдоль и сложим вдвое только с одной стороны (см. рис. 2.66а). Затем сложенную сторону растянем вдвое и склеим с противоположной стороной (рис. 2.66б). Для гладкости склейки и вследствие того, что траектории на такой поверхности могут пересекаться, необходимо, чтобы исходный лист состоял из бесконечного числа плотно прижатых непересекающихся поверхностей. Тогда фазовые кривые после совершения одного оборота каждый раз будут попадать на новый лист. В результате получим множество, объем которого равен нулю, а траектории расходятся.

Поскольку фазовые кривые на странном аттракторе расходятся, то динамика системы с таким аттрактором во многом аналогична динамике в ограниченном объеме гамильтоновой системы с перемешиванием, т. е.

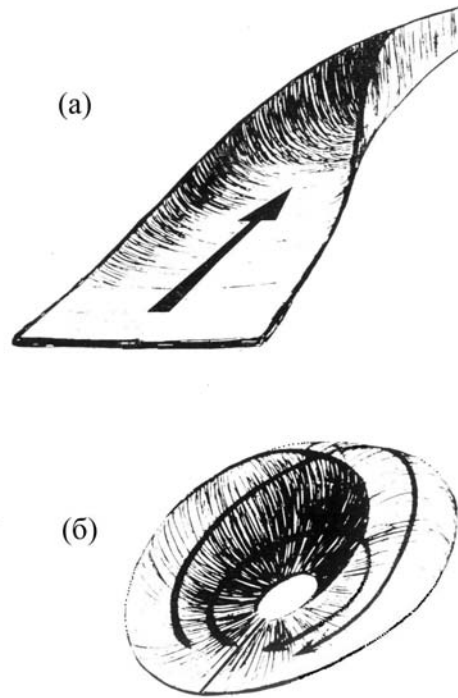


Рис. 2.66. Модель странного аттрактора [482]

является хаотической. Математическим образом такого хаотического движения в диссипативных динамических системах и служит странный аттрактор. При этом малые возмущения системы хотя и могут изменить структуру странного аттрактора, но не разрушают его: для всех близких динамических систем движение будет хаотическим.

20.3.1 Аттрактор Лоренца

Приведем достаточно простую систему, обладающую странным аттрактором. Это — известная *система Лоренца* [117]. Она была предложена в работе Б. Зальцмана [478] и представляет собой систему обыкновенных дифференциальных уравнений, которая получается из уравнений гидродинамики в задаче о термоконвекции в подогреваемом снизу горизонтальном

слое жидкости:

$$\dot{x} = \sigma y - \sigma x, \quad \dot{y} = rx - y - xz, \quad \dot{z} = xy - bz. \quad (2.120)$$

Здесь σ — число Прандтля, r — приведенное число Рэлея, b — постоянная, характеризующая размеры физической системы. Исторически уравнения Лоренца явились первой динамической системой, в которой было доказано наличие странного аттрактора. Нас эта система в основном будет интересовать как модельный пример, изучая который, можно понять, что такое странный аттрактор и как он может появиться.

Перечислим сначала ряд общих свойств модели Лоренца.
Дивергенция системы (2.120)

$$\sum_{i=1}^3 \frac{\partial v_i}{\partial x_i} = -(b + \sigma + 1) < 0. \quad (2.121)$$

Следовательно, фазовый объем со временем постоянно сжимается, т. е. система Лоренца является диссипативной. Можно показать [51, 117], что система (2.120) не допускает интегральных кривых, уходящих на бесконечность, т. е. любое решение (2.120) в конце концов попадает в шар:

$$x^2 + y^2 + z^2 < R^2, \quad R = R(\sigma, r, b), \quad (2.122)$$

и остается там при любых $t > 0$.

Система Лоренца симметрична относительно замены переменных $x \rightarrow -x$, $y \rightarrow -y$, $z \rightarrow z$, следовательно, ее фазовый портрет симметричен относительно оси z . Данное свойство уравнений Лоренца облегчает численное исследование системы (2.120).

Опишем теперь изменение режимов движения в системе (2.120) в зависимости от значений ее параметров r , σ и b .

1. Положения равновесия — стационарные точки системы Лоренца — находятся из уравнений

$$\sigma y - \sigma x = 0, \quad rx - y - xz = 0, \quad xy - bz = 0. \quad (2.123)$$

Начало координат, т. е. точка $O(x = y = z = 0)$, является стационарной точкой при любых r , σ и b . Ее характеристическое уравнение имеет вид

$$[\lambda^2 + (\sigma + 1)\lambda + \sigma(1 - r)](\lambda + b) = 0. \quad (2.124)$$

Следовательно, точка O устойчива и является устойчивым узлом, если $r < 1$.

2. Когда $r > 1$, точка O теряет устойчивость, превращается в седло-узел и в системе возникают еще две стационарные точки:

$$O_1(x, y, z) = (\sqrt{b(r-1)}, \sqrt{b(r-1)}, (r-1)),$$

$$O_2(x, y, z) = (-\sqrt{b(r-1)}, -\sqrt{b(r-1)}, (r-1)).$$

При всех $r > 1$ система (2.120) имеет только эти три стационарные точки: O , O_1 и O_2 . Тип точек O_1 и O_2 определяется из характеристического уравнения

$$\lambda^3 + (\sigma + b + 1)\lambda^2 + (r + \sigma)b\lambda + 2\sigma b(r - 1) = 0. \quad (2.125)$$

Отсюда находим, что O_1 и O_2 устойчивы, если $\sigma > b + 1$ и

$$1 < r < r^* = \sigma \frac{\sigma + b + 3}{\sigma - b - 1}. \quad (2.126)$$

При $r > r^*$ точки O_1 и O_2 становятся неустойчивыми. В этом случае характеристическое уравнение (2.125) имеет один действительный отрицательный корень и два комплексно-сопряженных с положительной действительной частью, т.е. точки O_1 и O_2 будут стационарными точками типа седло-фокус.

Для дальнейшего изучения поведения траекторий требуется численное интегрирование уравнений (2.120), поскольку локальный анализ окрестностей неустойчивых стационарных точек O , O_1 и O_2 не дает сведений о характере движения в системе Лоренца.

Численное исследование системы Лоренца было выполнено многими авторами (сводка результатов имеется в [55, 158, 198, 497]). Следуя [198], рассмотрим эволюцию режимов движения в уравнениях (2.120) при $\sigma = 10$, $b = 8/3$ и $10 \leq r \leq 28$.

1. При $10 \leq r < r_1 \approx 13,926$ система имеет три состояния равновесия: O , O_1 и O_2 . Точка O неустойчива и представляет собой точку типа седло-узел, имеющую двумерное устойчивое многообразие W^s и две выходящие кривые (сепаратрисы) Γ_1 и Γ_2 . Две другие стационарные точки, O_1 и O_2 , являются устойчивыми. Характер поведения фазовых траекторий для этого случая представлен на рис. 2.67.

2. Когда $r = r_1 \approx 13,926$, каждая из кривых Γ_1 , и Γ_2 превращается в замкнутую гомоклиническую петлю; при этом точки O_1 и O_2 остаются устойчивыми (рис. 2.68).

3. $r_1 < r < r_2 \approx 24,06$. В этой ситуации точки O_1 и O_2 по-прежнему являются устойчивыми, а от каждой из гомоклинических петель Γ_1 и Γ_2

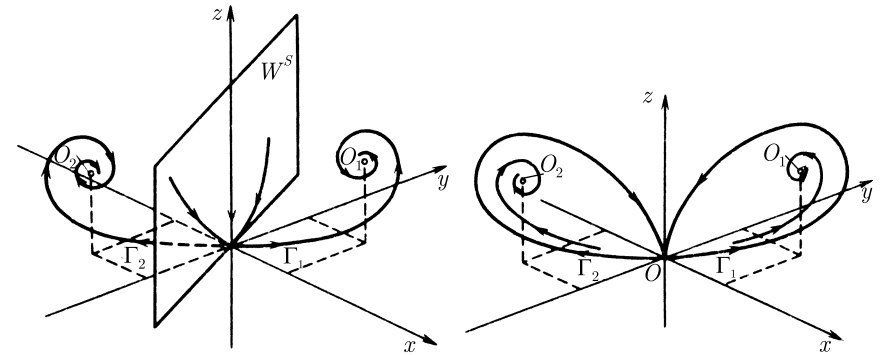


Рис. 2.67. Фазовый портрет системы Лоренца при $\sigma = 10$; $b = 8/3$, $10 < r < 13,926$

рождается по седловой периодической траектории L_1 и L_2 . Сепаратрисы Γ_1 и Γ_2 стремятся теперь, соответственно, к точкам O_2 и O_1 . Кроме того, появляются кривые, идущие от седлового предельного цикла L_1 к L_2 и от L_2 к L_1 . Эти кривые образуют одномерное множество B_1 , но это множество не является притягивающим и, следовательно, аттрактором. Устойчивые многообразия седловых периодических траекторий L_1 и L_2 являются границами областей притяжения стационарных точек O_1 и O_2 . Фазовая кривая, начавшаяся вне этих областей, может совершать колебания из окрестности L_1 в окрестность L_2 и обратно, пока не попадет в область притяжения одного из аттракторов O_1 или O_2 , причем по мере приближения параметра r к значению r_2 число колебаний существенно возрастает [67, 89]. Такое поведение системы называют *метастабильным хаосом* [89]. Общий вид основных элементов фазового портрета системы Лоренца для этого случая показан на рис. 2.69а, а его проекция на координатную плоскость (x, z) — на рис. 2.69б.

4. При $r = r_2 \approx 24,06$ стационарные точки O_2 и O_1 , согласно (2.126), еще являются устойчивыми, но сепаратрисы Γ_1 и Γ_2 уже не стремятся к ним, а наматываются на седловые траектории L_2 и L_1 , соответственно (рис. 2.70). Теперь на месте B_1 возникает множество B_2 , которое при $r > r_2$ становится притягивающим.

5. $r_2 < r < r_3 \approx 24,74$. В этом интервале точки O_1 и O_2 (см. (2.126)) по-прежнему устойчивы. Однако кроме них в фазовом пространстве системы имеется предельное множество B_2 , называемое *аттрактором Лоренца*

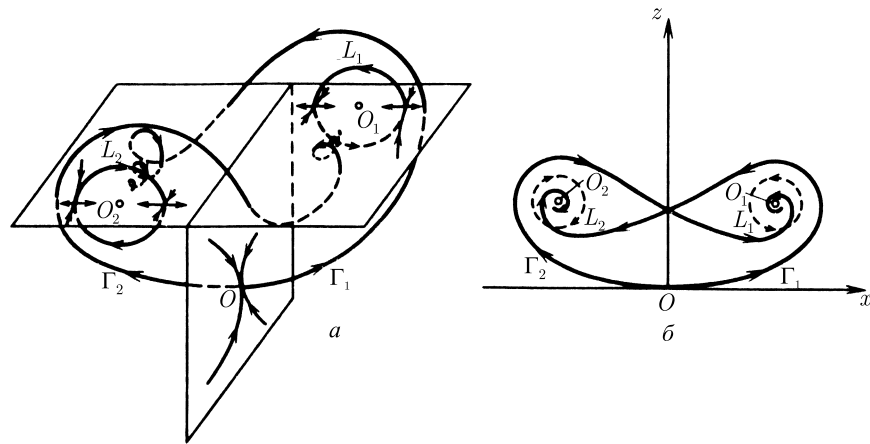


Рис. 2.69. Основные элементы фазового портрета системы Лоренца при $\sigma = 10$, $b = 8/3$ и $13,926 < r < 24,06$ (а) [158] и его проекция на координатную плоскость (x, z) (б) [198]

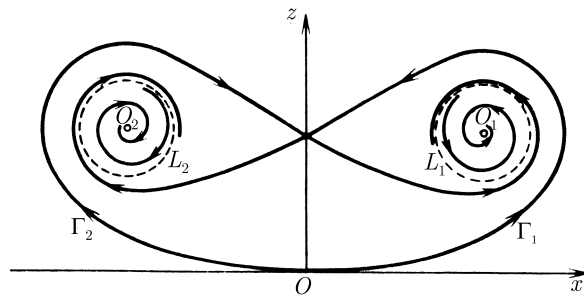


Рис. 2.70. Проекция фазового портрета системы Лоренца на координатную плоскость (x, z) при $\sigma = 10$, $b = 8/3$, $r = 24,06$

ца. Множество B_2 представляет собой совокупность интегральных кривых, идущих от L_1 к L_2 и обратно; кроме них, B_2 содержит седловую точку O и ее сепаратрисы Γ_1 и Γ_2 .

Таким образом, при $r \in (r_2, r_3)$ в системе Лоренца имеются три аттрактора: стационарные точки O_1 и O_2 и аттрактор Лоренца B_2 . Область притя-

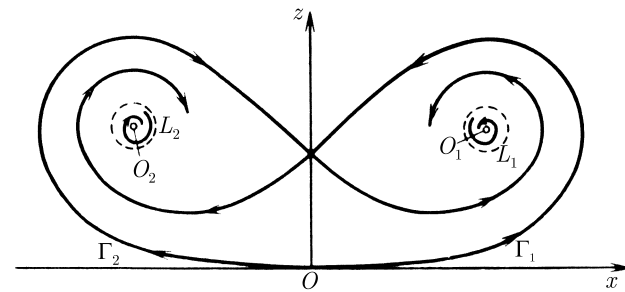


Рис. 2.71. Проекция фазового портрета системы Лоренца на координатную плоскость (x, z) при $\sigma = 10$, $b = 8/3$ и $24,06 < r < 24,74$

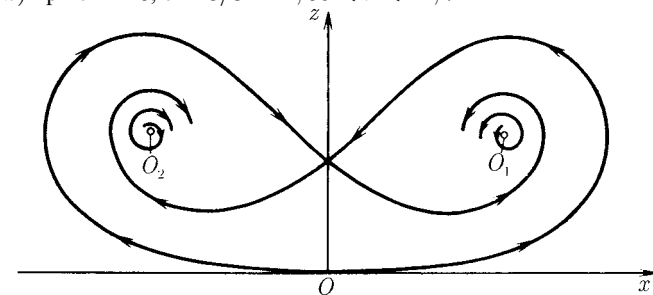


Рис. 2.72. Проекция фазового портрета системы Лоренца на координатную плоскость (x, z) при $\sigma = 10$, и $b = 8/3$, $24,74 < r < 28$

жения B_2 ограничена устойчивыми многообразиями седловых предельных циклов L_1 и L_2 (рис. 2.71). Фазовые траектории системы в зависимости от начальных условий с течением времени стремятся либо к точке O_1 , либо к точке O_2 , либо совершают колебания, случайным образом переходя от вращения вокруг точки O_1 к вращению вокруг точки O_2 и обратно. Следовательно, в зависимости от начальных условий в этом случае в системе (2.120) могут реализоваться существенно различные режимы движения: стационарный или хаотический.

6. Когда $r \rightarrow r_3 \approx 24,74$, седловые предельные циклы L_1 и L_2 стягиваются, соответственно, к стационарным точкам O_1 и O_2 и при $r = r_3$ исчезают, сливаясь с ними. В этот момент точки O_1 и O_2 , согласно (2.126), становятся неустойчивыми.

7. В интервале $r_3 < r \leq r_4 = 28$ все стационарные точки O , O_1 и O_2 являются неустойчивыми. Единственным устойчивым предельным множеством — аттрактором — будет B_2 , т. е. аттрактор Лоренца (рис. 2.72). Сле-

довательно, в динамической системе при любых (кроме множества меры нуль) начальных условиях реализуется хаотический режим движения.

Чтобы лучше представить себе, что происходит при таком движении, обратимся к модели, предложенной Лоренцом [67, 117]. Эта модель представляет собой ветвящуюся поверхность S , расположенную близко к вертикальной плоскости $x = y$, содержащую неподвижные точки O_1 и O_2 и симметричную относительно оси z (рис. 2.73). Фазовая траектория, начавшаяся слева от линии Oz , разделяющей поверхность S , будет оборачиваться вокруг точки O_1 по часовой стрелке, раскручиваясь по спирали до тех пор, пока не пересечет Oz , после чего начнет оборачиваться вокруг точки O_2 , но в противоположном направлении (против часовой стрелки). Фазовая траектория будет продолжать раскручиваться от точки O_2 до того момента, пока снова не пересечет линию Oz . Теперь вращение будет происходить вокруг точки O_1 и т.д. При этом последовательные числа оборотов относительно точек O_1 и O_2 изменяются нерегулярным образом, и движение выглядит хаотическим.

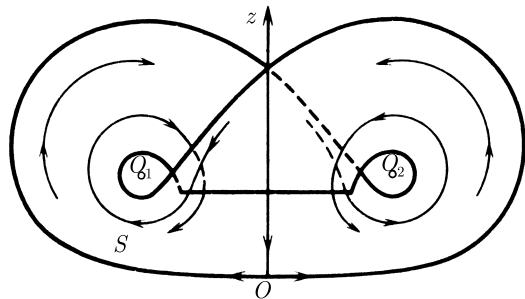


Рис. 2.73. Модель аттрактора Лоренца

Заметим, что описанная картина не может быть точной, поскольку траектории, расположенные на поверхности S , могут пересекаться. Поэтому, вообще говоря, аттрактор должен состоять из бесконечного числа листов, связанных вместе и тесно прижатых друг к другу, причем отдельные листы не сливаются, а отстоят один от другого на очень малое расстояние (ср. рис. 2.66). Фазовая траектория будет расположена на такой бесконечно ветвящейся поверхности, которая и называется аттрактором Лоренца.

Результат численного интегрирования уравнений Лоренца (2.120) показан на рис. 2.74 [114]. Хаотическая траектория аттрактора просчитывалась при $\sigma = 10$, $b = 8/3$, $r = 28$ и начальных условиях $x(0) = y(0) = z(0) = 0$.

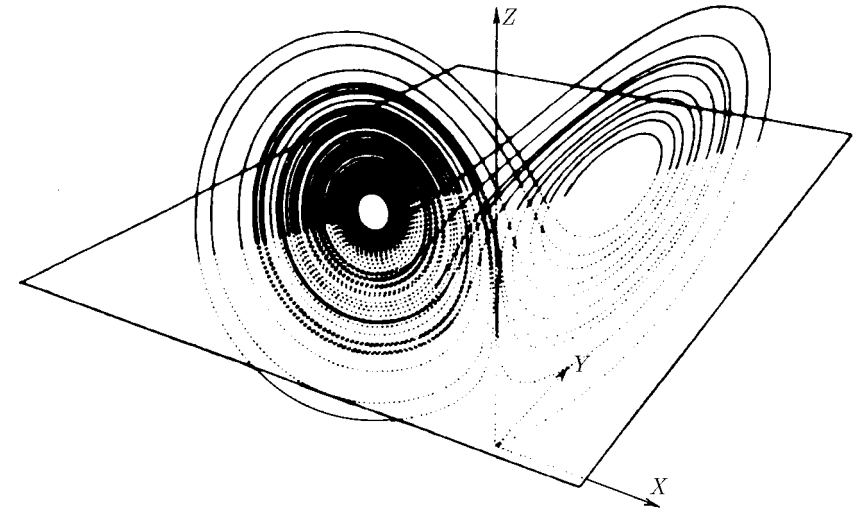


Рис. 2.74. Хаотическая траектория на аттракторе Лоренца при $\sigma = 10$, $b = 8/3$ и $r = 28$. Плоскость (x, y) соответствует $z = 27$ [114]

Описанный аттрактор не является ни двумерной поверхностью, ни склеенным двумерным многообразием. Более того, поверхность, проходящая через такой аттрактор, пересекается траекториями в точках, образующих множество, известное в математике как канторово множество (см. §22).

В заключение этого примера отметим, что хаотический режим движения в системе Лоренца имеет место и при $r > r_4 = 28$ ($\sigma = 10$, $b = 8/3$) вплоть до значения $r \approx 148,4$, когда в фазовом пространстве странный аттрактор сменяется предельным циклом. Движение в этом случае становится периодическим.

20.3.2 Аттрактор Смейла–Вильямса

Другой хорошо известный пример притягивающего предельного множества, сочетающего в себе свойство локальной неустойчивости с общим сжатием фазового объема, — *соленоид Смейла–Вильямса* [181, 524]. Этот пример носит модельный характер: аттракторы такого типа, по-видимому, не встречаются в физических системах, хотя во многих отношениях являются похожими на них.

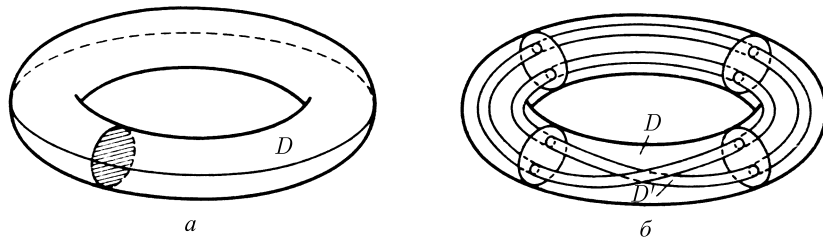


Рис. 2.75. Построение соленоида Смейла–Вильямса

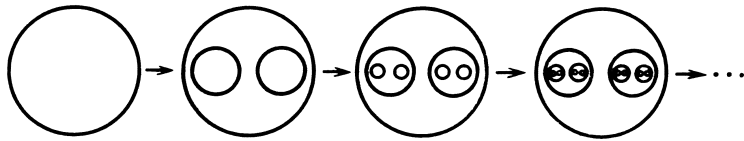


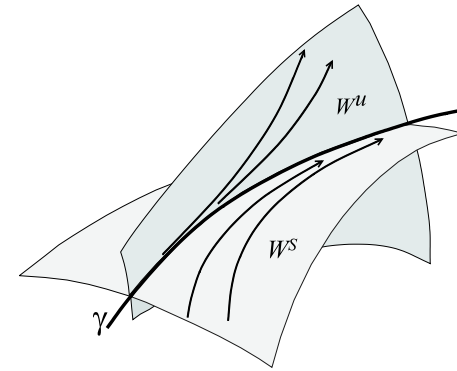
Рис. 2.76. Иерархия структур, возникающая при построении соленоида Смейла–Вильямса (в сечении)

Соленоид Смейла–Вильямса строится в пространстве размерности три или больше. Рассмотрим тороидальную область D , т. е. внутренность двумерного тора (рис. 2.75а). Растянем ее, сожмем вдоль меридиана, затем перекрутим и изогнем так, чтобы она перешла в область D' , лежащую в D (рис. 2.75б) [181]. Следующее действие состоит в применении аналогичного преобразования к области D' . Тогда получим область D'' , расположенную в D' , и т. д. В результате бесконечной последовательности таких преобразований в сечении тороидальной области вертикальной плоскостью получим иерархию структур, типа изображенных на рис. 2.76. При этом первоначально соседствующие точки будут расходиться экспоненциально быстро [158], а объем исходной области D — стремиться к нулю. Однако возникающее предельное множество — бесконечно тонкая линия, «протыкающая» секущую плоскость неограниченное число раз, не замкнется и не уйдет на бесконечность.

В заключение этого параграфа отметим следующее. В литературе часто подчеркивается, что динамика системы является хаотической благодаря наличию в ее фазовом пространстве странного аттрактора. В этих случаях понятие «странный аттрактор» имеет собирательный смысл, и его ино-

гда заменяют словосочетанием «хаотический аттрактор». Под хаотическим аттрактором может подразумеваться несколько типов аттракторов, однако гиперболические аттракторы, стохастические (или квазигиперболические) аттракторы и квазистохастические аттракторы (или квазиаттракторы) являются наиболее распространенными.

Гиперболический аттрактор — это притягивающее множество, каждая траектория которого устроена так, что она имеет устойчивое W^s и неустойчивое W^u многообразия (рис. 2.77). Устойчивое многообразие характеризуется тем свойством, что в малой окрестности γ соседние траектории, принадлежащие W^s , при $t \rightarrow \infty$ будут экспоненциально быстро стремиться к γ . На неустойчивом многообразии W^u , наоборот, такие траектории при $t \rightarrow \infty$ с экспоненциальной скоростью покидают окрестность γ .

Рис. 2.77. Устойчивое W^s и неустойчивое W^u многообразия траектории γ , принадлежащей гиперболическому аттрактору.

Системы с гиперболическим аттрактором имеют наиболее выраженные хаотические свойства. Малые возмущения таких систем не могут привести к качественным перестройкам как самого аттрактора, так и поведения систем в целом. Динамические системы с гиперболическим типом аттрактора являются моделями систем со строго хаотическими свойствами [34, 179]. Однако в настоящее время гиперболических аттракторов построено немного. Один из них — это соленоид Смейла–Вильямса.

Стохастический аттрактор является математическим образом наблюдаемого развитого хаотического поведения физических систем. Системы со стохастическими аттракторами обладают свойством перемешивания. Самый известный пример — аттрактор Лоренца при $b = 8/3$, $\sigma = 10$, $r = 28$.

Малые возмущения могут приводить к модификациям такого аттрактора, но в то же время динамика системы будет оставаться хаотической.

Гиперболический аттрактор является стохастическим. В то же время стохастические аттракторы не обязательно являются странными (см. [27, 34, 53, 175, 429]).

Однако подавляющее большинство аттракторов хаотических динамических систем принадлежат к квазистохастическому типу, т.е. являются *квазиаттракторами* [212, 254]. Квазистохастические аттракторы содержат в себе помимо седловых предельных циклов еще и устойчивые предельные циклы, период которых достаточно велик, а область притяжения столь мала, что они не выявляются численно. Например, аналитические результаты показывают, что в системе Лоренца с параметрами, близкими к значениям $\sigma = 10, 2$; $r = 30, 2$; $b = 8/3$, существуют устойчивые предельные циклы [38, 57]. За счет флуктуаций фазовая кривая системы может перескакивать с одной устойчивой траектории на другую. Вследствие этого при численных расчетах или в эксперименте движение выглядит хаотическим.

Таким образом, суммируя изложенное, можно сделать вывод, что не только простые консервативные, но и совсем несложные диссипативные динамические системы (например, система Лоренца), размерность фазового пространства которых больше или равна трем, могут иметь наряду с регулярными и очень сложные, хаотические режимы движения. Математическим образом такого хаотического поведения диссипативных систем является притягивающее множество сложной структуры — странный аттрактор. Естественно поэтому теперь перейти к определению некоторых критериев отличия регулярных движений от хаотических.

21. Критерии динамического хаоса

Итак, аттракторы диссипативных динамических систем могут быть нескольких типов: устойчивая стационарная точка, устойчивый предельный цикл, инвариантный тор (все они относятся к простым аттракторам) и странный аттрактор. Простому аттрактору соответствует регулярное движение системы (например, инвариантному тору — квазипериодическое движение); напротив, динамика систем со странным аттрактором является хаотической. Как видно, существует принципиальная разница в поведении систем со странными и простыми аттракторами. Поэтому необходимо иметь критерии, позволяющие отличать аттракторы одного типа от аттракторов другого типа. Не менее важно, чтобы такие критерии давали возможность различать регулярное и хаотическое движение и в случае гамильтоновых

систем, т.е. систем, не имеющих аттракторов (см. раздел 2.14). Поскольку хаотичность является следствием неустойчивости фазовых траекторий, так что близкие в фазовом пространстве интегральные кривые с течением времени расходятся, то представляется вполне естественным в качестве одного из таких критериев выбрать именно меру разбегания фазовых кривых динамической системы.

21.1 Энтропия Колмогорова–Синяя

Пусть, как и прежде, некоторая динамическая система задается обыкновенными дифференциальными уравнениями

$$\dot{x} = v(x), \quad x(0) = x_0. \quad (2.127)$$

Рассмотрим в фазовом пространстве в начальный момент времени $t = 0$ две близкие фазовые точки $x_1(0)$ и $x_2(0)$, «выпустим» из них траектории и проследим, как при эволюции системы ($t > 0$) будет изменяться расстояние $d(t) = |\xi(t)| = |x_2(t) - x_1(t)|$ между этими точками. Тогда, если режим движения является хаотическим, $d(t)$ с течением времени будет экспоненциально возрастать, так что на больших временах

$$d(t) \approx d(0)e^{kt}. \quad (2.128)$$

Отсюда найдем среднюю скорость экспоненциального разбегания траекторий

$$k \approx \frac{1}{t} \ln \frac{d(t)}{d(0)}. \quad (2.129)$$

Но это определение, вообще говоря, не является приемлемым. В самом деле, при финитности движения (а только такое движение мы и рассматриваем) $d(t)$ не может увеличиваться всегда. Поэтому при больших t величина (2.129) в любом случае независимо от режима — хаотического или регулярного — будет близка к нулю. Однако чем меньше мы выберем начальное расстояние $d(0) = |\xi(0)|$, тем дольше можно следить за возрастанием $d(t)$, т.е. в течение большего промежутка времени величина $d(t)$, согласно (2.129), не достигнет максимального значения. Следовательно, в формуле (2.129) необходимо положить $d(0) \rightarrow 0$ и $t \rightarrow \infty$:

$$h = \lim_{\substack{d(0) \rightarrow 0 \\ t \rightarrow \infty}} \frac{1}{t} \ln \frac{d(t)}{d(0)}. \quad (2.130)$$

Величину h в физической литературе часто называют *энтропией Колмогорова–Синая*, или *КС-энтропией* (см., например, [6]). Используя КС-энтропию, можно определить, каким является исследуемый режим движения — хаотическим или регулярным. В частности, если динамика системы является периодической или квазипериодической, то с течением времени расстояние $d(t)$ не возрастает, так что значение энтропии равно нулю, $h = 0$. Если движению системы отвечает устойчивая стационарная точка, то $d(t) \rightarrow 0$ и $h < 0$. Однако в случае хаотического поведения КС-энтропия всегда положительна, $h > 0$.

Энтропия h (2.130) — величина размерная ($[h] = c^{-1}$) и по существу является не только качественной, но и количественной характеристикой режима движения: величина, обратная энтропии (при условии $h > 0$), определяет характерное время перемешивания $t_{\text{mix}} = h^{-1}$ в системе; по прошествии промежутка времени $t \gg t_{\text{mix}}$ начальная область Ω_0 расплывается по всей энергетически доступной гиперповерхности (в отсутствие диссипации, см. раздел 2.18) или по предельному подмножеству фазового пространства — странному аттрактору (для диссипативных динамических систем); при $t \gg t_{\text{mix}}$ описание системы может быть только вероятностным. Однако на малых временах $t \ll t_{\text{mix}}$ поведение системы можно предсказать с достаточной точностью (не превышающей, естественно, точность ϵ задания начального положения фазовой точки).

Таким образом, КС-энтропия, определяемая (2.130), является важной характеристикой динамической системы, и в настоящее время она широко используется в качестве критерия хаотичности движения. Подробно с методами численных расчетов величины h можно ознакомиться в [227, 229, 312].

На первый взгляд, в определении (2.130) есть некоторая неясность. В самом деле, не может ли зависеть значение предела от того, как мы выбрали направление начального смещения $\xi(0)$? Как будет показано ниже, хотя такая зависимость в принципе и есть, она несущественна: если мы выбираем вектор начального смещения *случайно*, то получаемое предельное значение h (2.130) будет практически всегда одним и тем же.

Вообще говоря, значение h может зависеть и от выбора начальной точки траектории $x_1(0)$. Дело в том, что динамическая система может одновременно обладать несколькими аттракторами со своими областями притяжения, и в зависимости от того, в какой области притяжения окажется $x_1(0)$, мы будем получать различные значения величины h . Например, система Лоренца (2.120) при $\sigma = 10$, $b = 8/3$ и r в интервале от 24,06 до 24,74 имеет один странный и два простых аттрактора. Не исключены и случаи, когда в фазовом пространстве некоторой динамической системы могут присутствовать сразу несколько странных аттракторов.

21.2 Показатели Ляпунова

В сущности, величина h , определяемая по формуле (2.130), не является истинной энтропией Колмогорова–Синая, введенной в [97, 98, 176] для характеристики свойств динамических систем. В строгом смысле используемая в физической литературе величина h — это максимальный из характеристических показателей Ляпунова. Поскольку характеристические показатели Ляпунова играют важную роль в теории динамических систем и тесно связаны с другими критериями хаотического движения (см. раздел 2.22), мы приведем их определение и поясним физический смысл этих характеристик.

Возьмем фазовую траекторию $x(t)$ динамической системы (2.127), выходящую из точки $x(0)$, а также какую-то траекторию, близкую к ней: $x_1(t) = x(t) + \xi(t)$. Рассмотрим функцию

$$\Lambda(\xi(0)) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{|\xi(t)|}{|\xi(0)|}, \quad (2.131)$$

определенную на векторах начального смещения $\xi(0)$, таких, что $|\xi(0)| = \epsilon$, где $\epsilon \rightarrow 0$.

Мы утверждаем, что при всевозможных поворотах вектора $\xi(0)$ функция Λ будет меняться скачками и принимать конечный ряд значений $\{\lambda_j\}$, $j = 1, 2, \dots, n$. Эти значения называют (глобальными) *показателями Ляпунова*.

Поясним теперь смысл предложенного определения. Если с течением времени исходная $x(t)$ и возмущенная $x_1(t)$ траектории остаются близкими, то изменение $\xi(t)$ подчиняется линеаризованному уравнению

$$\dot{\xi} = A\xi, \quad (2.132)$$

где A — матрица с элементами

$$a_{ik} = \left. \frac{\partial v_i}{\partial x_k} \right|_{x(t)}. \quad (2.133)$$

Вообще говоря, матрица A зависит от времени, $A = A(t)$. Однако вначале для простоты мы рассмотрим случай, когда она является постоянной. Предположим также, что все собственные значения матрицы A действительные. Тогда общее решение линейного уравнения (2.132) имеет вид

$$\xi = \sum_j C_j e_j \exp(\lambda_j t), \quad (2.134)$$

где C_j — постоянные интегрирования, определяемые из начальных условий, а e_j — собственные векторы, отвечающие различным собственным значениям λ_j матрицы A . Каждый собственный вектор e_j задает инвариантное направление в пространстве всех векторов ξ . Если вектор начального смещения $\xi(0)$ направлен вдоль одного из векторов e_j , то тогда $\xi(t)$ остается параллельным ему, и с течением времени расстояние между траекториями возрастает или убывает по экспоненциальному закону: $|\xi(t)| \sim \exp(\lambda_j t)$. При таком выборе вектора начального смещения $\xi(0)$ имеем

$$\Lambda(\xi(0)) = \lambda_j. \quad (2.135)$$

В общем случае вектор начального смещения $\xi(0)$ имеет составляющие вдоль нескольких или всех векторов e_j (рис. 2.78). Поэтому на больших временах изменение расстояния между двумя траекториями $|\xi(t)|$ определяется, как видно из (2.134), слагаемым с наибольшей экспонентой в разложении вектора $\xi(t)$. В зависимости от того, какие именно собственные векторы e_j присутствуют в разложении вектора начального смещения $\xi(0)$, доминирующая экспонента будет различна. Однако всегда на достаточно больших временах расстояние $|\xi(t)|$ пропорционально какой-либо из экспонент $\exp(\lambda_j)$, а следовательно, функция $\Lambda(\xi(0))$ может принимать значения, равные только λ_j , и при повороте вектора $\xi(0)$ могут происходить только скачкообразные изменения функции $\Lambda(\xi(0))$.

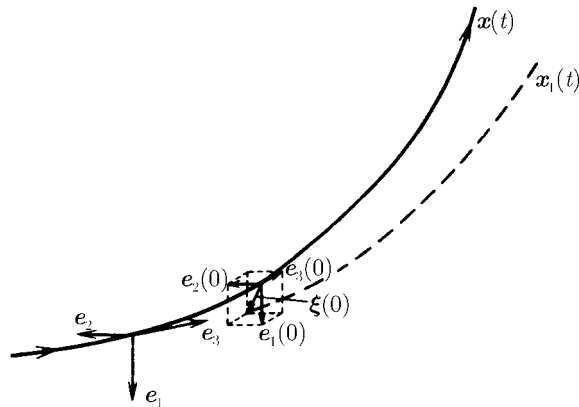


Рис. 2.78. Близкие траектории $x(t)$ и $x_1(t)$ в фазовом пространстве

Кроме действительных, матрица A (см. (2.132)) может иметь и комплексные собственные значения, которые мы обозначим α . Поскольку матрица A действительная, то, наряду с собственными значениями $\alpha = \lambda + i\omega$, обязательно есть и комплексно-сопряженные собственные значения $\alpha^* = \lambda - i\omega$. Этой паре комплексно-сопряженных собственных значений отвечает некоторая инвариантная плоскость, задаваемая двумя векторами e и e' . На этой плоскости решение системы (2.132) имеет вид

$$\xi(t) = (Ce \cos \omega t + Ce' \sin \omega t) \exp(\lambda t). \quad (2.136)$$

Следовательно, если выбрать начальное смещение $\xi(0)$ принадлежащим этой плоскости, то в среднем по времени расстояние между траекториями меняется по закону $|\xi(t)| \sim \exp(\lambda t)$, и, согласно формуле (2.131), мы получаем $\Lambda(\xi(0)) = \lambda$. Таким образом, паре комплексно-сопряженных собственных значений α, α^* отвечает характеристический показатель Ляпунова $\lambda = \text{Re } \alpha$, который следует считать дважды вырожденным, поскольку ему соответствует движение на инвариантной плоскости, а не вдоль инвариантного направления.

Итак, для постоянной матрицы $A \neq A(t)$ все характеристические показатели Ляпунова совпадают с действительными частями ее собственных значений.

Разумеется, в общем случае элементы матрицы A зависят от времени. Из теории линейных дифференциальных уравнений с зависящими от времени коэффициентами известно, что тогда общее решение этой системы может быть записано в виде суперпозиции n фундаментальных частных решений $\{\xi_j(t)\}$:

$$\xi(t) = \sum_j C_j \xi_j(t), \quad (2.137)$$

где постоянные C_j определяются из разложения начального смещения $\xi(0)$ по векторам $\xi_j(0)$. Вообще говоря, зависимость от времени для фундаментальных частных решений $\{\xi_j(t)\}$ является более сложной, чем экспоненциальная. Однако, учитывая, что матрица $A(t)$ не возрастает неограниченно с течением времени, на больших временах можно записать $|\xi_j(t)| = \Phi_j(t) \exp(\lambda_j t)$, где функция $\Phi_j(t)$ имеет более медленную, чем экспоненциальная, степень роста. Поэтому если мы выберем вектор начального смещения $\xi(0)$ параллельно одному из векторов $\xi_j(0)$, тогда $\Lambda(\xi(0)) = \lambda_j$, что совпадает с формулой (2.135), хотя теперь уже λ_j не являются действительными частями собственных значений матрицы A . Легко убедиться, что никаких других значений величина $\Lambda(\xi(0))$ не принимает.

Покажем, что при произвольном выборе начального смещения $\xi(0)$ предел (2.131) дает максимальный показатель Ляпунова. Действительно, наугад взятый вектор, скорее всего, будет иметь проекции на все векторы $\xi(0)$, в том числе и на тот из них, который отвечает максимальному значению λ_j . На больших временах экспонента с этим значением λ_j в разложении (2.137) будет доминировать, а поэтому предел (2.131) будет давать значение λ_{\max} . Таким образом, введенная вначале по формуле (2.130) величина h , которую мы называли ранее «энтропия Колмогорова–Синая», является в действительности *максимальным* характеристическим показателем Ляпунова.

Перечислим основные свойства характеристических показателей Ляпунова для диссипативных и консервативных систем.

Число характеристических показателей равно размерности n фазового пространства исходной динамической системы, и их можно упорядочить по убывающей:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n. \quad (2.138)$$

Один из показателей Ляпунова, отвечающий смещению вдоль траектории, не оканчивающейся в особой точке, всегда равен нулю.

Для гамильтоновых систем показатели обладают следующей симметрией [116]:

$$\lambda_j = -\lambda_{2k-j+1}, \quad (2.139)$$

где k — число степеней свободы ($n = 2k$). Следовательно, для гамильтоновых систем по крайней мере два показателя равны нулю. В случае полностью интегрируемых систем все λ равны нулю [116, 255].

Сумма всех ляпуновских показателей равна среднему вдоль траектории значению дивергенции $\operatorname{div} \mathbf{v} = \operatorname{div} \dot{\xi} = \sum_k a_{kk}(t)$:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \operatorname{div} \dot{\xi} dt = \sum_{j=1}^n \lambda_j. \quad (2.140)$$

Относительное изменение фазового объема, как известно, определяется знаком при дивергенции: если $\operatorname{div} \mathbf{v} < 0$, то система диссипативна; если $\operatorname{div} \mathbf{v} = 0$, то фазовый объем при движении сохраняется. Значит, для гамильтоновых систем $\sum_j \lambda_j = 0$, в то время как для диссипативных систем

$$\sum_j \lambda_j < 0.$$

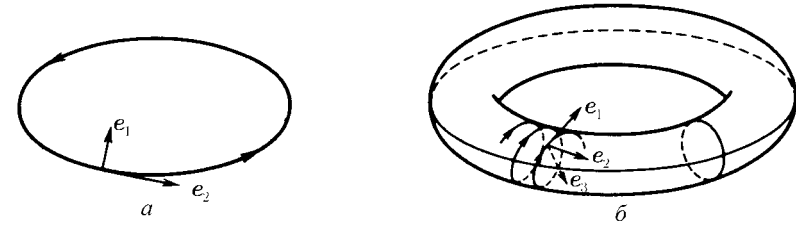


Рис. 2.79. В случае предельного цикла (а) и инвариантного тора (б) равны нулю показатели Ляпунова, отвечающие соответственно направлениям e_2 , и e_1, e_2

Динамическую систему называют *гиперболической*, если она имеет как положительные, так и отрицательные показатели Ляпунова.

Сформулируем теперь критерии, позволяющие, основываясь на характеристических показателях Ляпунова, различать типы аттракторов динамических систем [189].

Для одномерной системы, в которой аттракторами могут быть только устойчивые особые точки, существует один показатель Ляпунова, который является отрицательным, $\lambda_1 = (-)$.

В двумерных системах аттракторы встречаются только двух типов: устойчивые стационарные точки и предельные циклы. Если оба показателя отрицательны, $(\lambda_1, \lambda_2) = (-, -)$, то аттрактором является устойчивая стационарная точка. Если же $(\lambda_1, \lambda_2) = (-, 0)$, то аттрактор — предельный цикл: один из показателей, соответствующий направлению e_2 (параллельно потоку), равен нулю (рис. 2.79а).

В трехмерном случае, помимо устойчивых особых точек и предельных циклов, аттракторами являются инвариантные торы и странные аттракторы. Когда два характеристических показателя Ляпунова равны нулю, а третий отрицателен, $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, -)$, это означает, что в фазовом пространстве системы имеется инвариантный устойчивый тор (рис. 2.79б). Если же один из показателей является отрицательным, другой равен нулю, а третий положителен, $(\lambda_1, \lambda_2, \lambda_3) = (-, 0, +)$, то аттрактор считается странным. Итак, для трехмерных динамических систем встречается только четыре типа аттракторов:

- $(\lambda_1, \lambda_2, \lambda_3) = (-, -, -)$ — устойчивый фокус или узел;
- $(\lambda_1, \lambda_2, \lambda_3) = (-, -, 0)$ — устойчивый предельный цикл;
- $(\lambda_1, \lambda_2, \lambda_3) = (-, 0, 0)$ — устойчивый тор;
- $(\lambda_1, \lambda_2, \lambda_3) = (-, 0, +)$ — странный аттрактор.

В частности, для системы Лоренца (2.120) при $\sigma = 16$, $b = 4$, $r = 40$ характеристические показатели Ляпунова равны $\lambda_1 = 1, 37$, $\lambda_2 = 0$, $\lambda_3 = -22, 37$ [485]. Следовательно, при указанных значениях параметров в фазовом пространстве системы имеется странный аттрактор, и движение будет хаотическим.

Аналитическое определение характеристических показателей Ляпунова и «энтропии» $h = \lambda_{\max}$ для большинства физически интересных задач не представляется возможным, поскольку для этого необходимо знать решения системы дифференциальных уравнений. Однако существуют достаточно надежные алгоритмы, позволяющие найти все показатели Ляпунова, используя численные методы. Интересующихся читателей мы отсылаем к [116, 158, 228, 312, 485].

21.3 Автокорреляционная функция и спектральная плотность

Помимо характеристических показателей Ляпунова при изучении сложных режимов движения иногда используется *автокорреляционная функция*. Автокорреляционная функция, если ее удастся вычислить, является достаточно эффективной характеристикой поведения рассматриваемой системы. Так, при периодической или квазипериодической динамике автокорреляционная функция тоже будет, соответственно, периодической или квазипериодической. Однако если с течением времени автокорреляционная функция стремится к нулю и система не имеет устойчивых стационарных точек, то следует ожидать, что будет наблюдаться хаотический режим движения. Стремление к нулю автокорреляционной функции и используется в качестве критерия динамического хаоса.

Наряду с автокорреляционной функцией (особенно при экспериментальных исследованиях) для установления хаотичности движения часто применяется спектральная плотность (или просто спектр), которая, как мы увидим ниже, тесно связана с автокорреляционной функцией. Характер спектральной плотности — один из самых простых и в то же время вполне надежных критериев, используемых при анализе режимов движения. Рассчитывая ее численно или измеряя экспериментально, сравнительно легко определить, каким является поведение системы — хаотическим, периодическим или квазипериодическим. Например, если система демонстрирует периодическую динамику с периодом T_0 (что математически отвечает движению фазовой точки по предельному циклу), то спектр такого движения будет дискретным и состоящим из узких линий, соответствующих частоте движения $\omega_0 = 2\pi/T_0$ и кратным ей гармоникам $2\omega_0, 3\omega_0, \dots$

Если поведение системы является квазипериодическим с несоизмеримыми частотами $\omega_1, \dots, \omega_k$ (чему, как известно, отвечает движение по k -мерному нерезонансному тору), то спектр будет состоять из k линий, соответствующих этим частотам и кратным им гармоникам. Иными словами, спектр будет оставаться *дискретным*, хотя, если число k частот движения достаточно велико, линии в спектре могут быть расположены весьма густо. Однако для случая хаотических режимов, когда динамика системы не является регулярной, соответствующий спектр будет *сплошным*. Это и используется при спектральном анализе движения динамических систем.

Остановимся подробнее на понятиях автокорреляционной функции и спектральной плотности. Прежде всего дадим их определения.

Пусть $x_i(t) = \{x_1(t), \dots, x_n(t)\}$ — решение системы (2.127). Рассмотрим только одну компоненту решения $x_m \equiv x(t)$. *Автокорреляционной функцией* называется среднее по некоторому временному интервалу T (при $T \rightarrow \infty$) произведений x , взятых в два различных момента времени t и $t + \tau$:

$$g(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t + \tau) dt. \quad (2.141)$$

Спектральная плотность определяется формулой

$$S(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} |X(\omega)|^2, \quad (2.142)$$

где $X(\omega)$ — коэффициенты Фурье функции $x(t)$:

$$X(\omega) = \int_0^T x(t)e^{-i\omega t} dt. \quad (2.143)$$

Автокорреляционная функция и спектральная плотность тесно связаны. В самом деле, из (2.142) после несложного преобразования найдем

$$S(\omega) = \lim_{T \rightarrow \infty} \left[\frac{1}{\pi} \int_0^T \cos(\omega\tau) d\tau \frac{1}{T} \int_0^{T-\tau} x(t)x(t + \tau) dt \right]. \quad (2.144)$$

Отсюда, допустив, что имеется возможность перестановки порядка выполнения операций, получим

$$S(\omega) = \frac{1}{\pi} \int_0^\infty \cos(\omega\tau) g(\tau) d\tau, \quad (2.145)$$

или

$$S(\omega) = \frac{1}{\pi} \int_0^{\infty} e^{-i\omega\tau} g(\tau) d\tau. \quad (2.146)$$

Предположим, что $x(t)$ — периодическая функция с периодом T_1 , $x(t) = x(t + T_1)$. Тогда ее можно разложить в ряд Фурье:

$$x(t) = \sum_{n=-\infty}^{\infty} C_n e^{-in\omega_1 t}, \quad (2.147)$$

где $\omega_1 = 2\pi/T_1$, а C_n — коэффициенты Фурье. В этом случае автокорреляционная функция равна

$$\begin{aligned} g(\tau) &= \frac{1}{T} \int_0^T x(t)x(t+\tau) dt = \\ &= \frac{1}{T} \sum_{n=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} C_n C_{n'} e^{-in'\omega_1\tau} \int_0^T e^{-i(n+n')\omega_1 t} dt = \\ &= \sum_{n=-\infty}^{\infty} C_n C_{-n} e^{-in\omega_1\tau} = \sum_{n=-\infty}^{\infty} |C_n|^2 e^{-in\omega_1\tau}, \end{aligned} \quad (2.148)$$

поскольку из-за того, что функция $x(t)$ действительная, имеет место равенство $C_n = C_{-n}^*$. Следовательно, автокорреляционная функция $g(\tau)$ тоже является периодической с периодом T_1 .

Используя (2.148), из (2.146) найдем спектральную плотность периодической функции

$$S(\omega) = \sum_{n=-\infty}^{\infty} |C_n|^2 \frac{1}{\pi} \int_0^{\infty} e^{-i(n\omega_1 - \omega)\tau} d\tau = \sum_{n=-\infty}^{\infty} |C_n|^2 \delta(\omega - n\omega_1). \quad (2.149)$$

Таким образом, спектр периодической функции $x(t)$, отвечающей периодическому движению динамической системы (2.127), будет дискретным и содержащим, помимо основной частоты движения ω_1 , гармоники $n\omega_1$, $n = 2, 3, \dots$.

В качестве иллюстрации приведем примеры спектров в экспериментах по исследованию конвекции Рэлея–Бенара [323] и химической реакции Белоусова–Жаботинского [515].

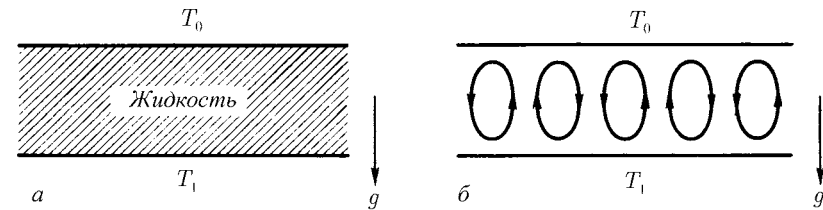


Рис. 2.80. Конвекция Рэлея–Бенара: *a* — схема эксперимента, *б* — конвективные потоки

На рис. 2.80*a* показана схема опыта при изучении конвекции Рэлея–Бенара. Слой жидкости подогревается снизу при постоянной разности температур $\Delta T = T_1 - T_0 > 0$. При малых ΔT в системе имеется устойчивое состояние равновесия, однако выше некоторого критического значения $\Delta T > \Delta T_c$ это равновесие становится неустойчивым, стационарный режим переходит в периодический, и в системе возникают конвективные потоки (рис. 2.80*б*). На рис. 2.81 показан спектр скорости жидкости в эксперименте Рэлея–Бенара [323] при стационарной конвекции для случая $R_a/R_c = 21$, где R_a и R_c — безразмерное и критическое числа Рэлея. Хорошо видны основная частота движения и кратные ей гармоники.

Рис. 2.81. Спектральная плотность, полученная в эксперименте Рэлея–Бенара при $R_a/R_c = 21$ (периодический режим) [323]

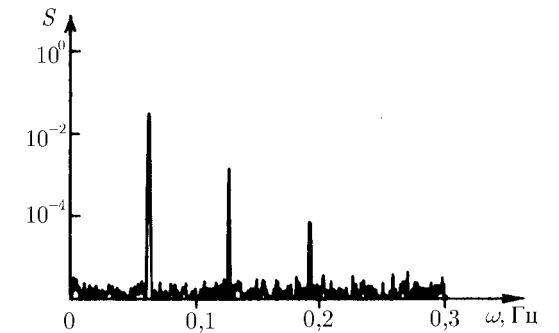


Рис. 2.82 иллюстрирует динамику химической реакции Белоусова–Жаботинского (подробнее см. [189], с. 31 и [61, 515]): при смешивании реагентов в определенных пропорциях в реакторе непрерывного перемешивания возникают колебания концентрации промежуточных соединений. В замкнутой системе такие колебания с течением времени затухают, однако при непрерывной подаче реагентов в реактор и отводе конечных продуктов

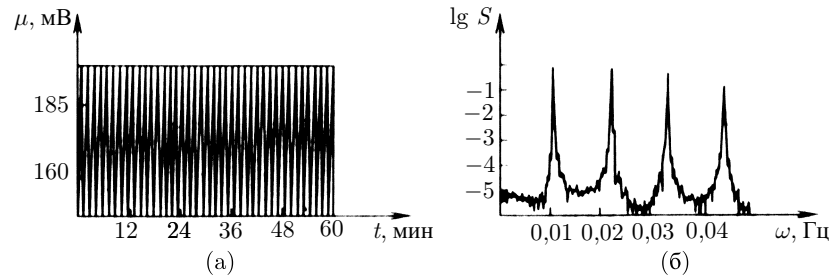


Рис. 2.82. Периодическая динамика химической реакции Белоусова–Жаботинского: *a* — колебания концентраций промежуточных соединений, *б* — соответствующая спектральная плотность [515]

колебания будут продолжаться неограниченно долго. В зависимости от концентрации одного из реагентов можно получить периодические или хаотические колебания. На рис. 2.82*a* показаны концентрационные колебания, возникающие в реакции Белоусова–Жаботинского в проточном реакторе. Соответствующие спектральные плотности приведены на рис. 2.82*б*. Последний рисунок наглядно демонстрирует периодическую динамику исследуемой системы.

Пусть теперь $x(t)$ — квазипериодическая функция,

$$x(t) = x(\varphi_1(t), \dots, \varphi_k(t)), \quad (2.150)$$

где $\varphi_i(t) = \omega_i t$, $i = 1, 2, \dots, k$, причем $x(t)$ по каждой переменной φ_i имеет период 2π :

$$x(\varphi_1 + 2\pi, \dots, \varphi_k + 2\pi) = x(\varphi_1, \dots, \varphi_k). \quad (2.151)$$

Для квазипериодичности необходимо, чтобы между частотами ω_i не выполнялось рациональных соотношений, т. е. не должно существовать такой частоты ω_0 , чтобы все ω_i были целыми кратными от нее.

Разложим $x(t)$ в многократный ряд Фурье по $\{\varphi_i\}$:

$$x(t) = \sum_{n_1, \dots, n_k} C_{n_1} \dots C_{n_k} e^{-i(n_1\varphi_1 + \dots + n_k\varphi_k)}. \quad (2.152)$$

Подставив сюда значение $\varphi_i = \omega_i t$, окончательно получим

$$x(t) = \sum_{n_1, \dots, n_k} C_{n_1} \dots C_{n_k} e^{-i(n_1\omega_1 + \dots + n_k\omega_k)t}. \quad (2.153)$$

Используя последнее соотношение, для автокорреляционной функции и спектральной плотности найдем

$$g(\tau) = \sum_{n_1, \dots, n_k} |C_{n_1} \dots C_{n_k}|^2 e^{-i(n_1\omega_1 + \dots + n_k\omega_k)\tau}, \quad (2.154)$$

$$S(\omega) = \sum_{n_1, \dots, n_k} |C_{n_1} \dots C_{n_k}|^2 \delta(\omega - n\omega_1 - \dots - n_k\omega_k). \quad (2.155)$$

Следовательно, спектр квазипериодической функции содержит всевозможные линейные комбинации $n_1\omega_1 + n_2\omega_2 + \dots + n_k\omega_k$. Если число частот движения ω_k велико, то линии в спектре будут расположены достаточно тесно, однако спектр будет оставаться дискретным.

Наконец, предположим, что динамика системы (2.127) является хаотической. В этом случае функция $x(t)$ есть аперриодическая функция времени. Следовательно, спектр $x(t)$ будет *сплошным* (или непрерывным).

Проиллюстрируем сказанное на упоминавшихся выше двух примерах: конвекции Рэлея–Бенара [323] и химической реакции Белоусова–Жаботинского [61, 515].

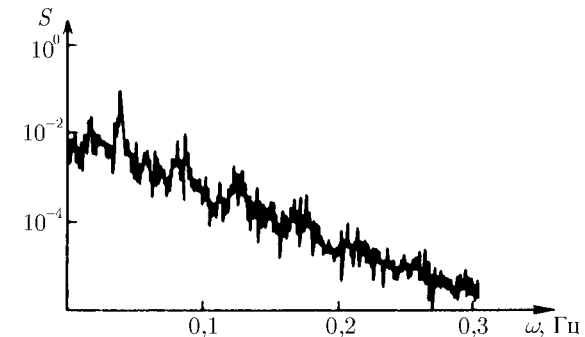


Рис. 2.83. Спектральная плотность, полученная в эксперименте Рэлея–Бенара при $R_g/R_c = 36,9$ (хаотический режим) [323]

При исследовании конвекции Рэлея–Бенара с увеличением приведенного числа Рэлея R_a/R_c периодическое движение конвективных потоков (рис. 2.80*б*) теряет устойчивость, и при $R_a/R_c = 36,9$ поведение жидкости становится хаотическим. При этом спектр из дискретного переходит в непрерывный (рис. 2.83).

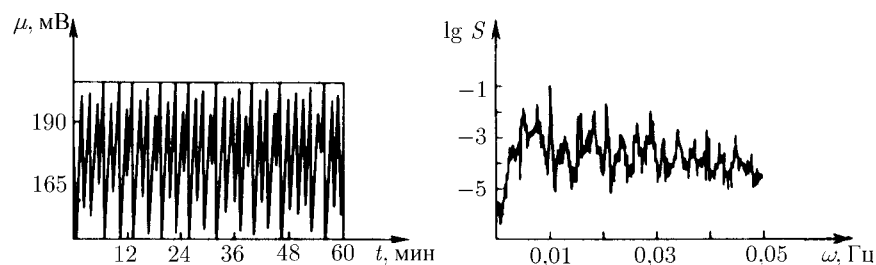


Рис. 2.84. Хаотическая динамика реакции Белоусова – Жаботинского [515]

Для реакции Белоусова – Жаботинского, помимо регулярных, возможно и хаотическое изменение концентраций промежуточных соединений. При определенных значениях скорости протока реагентов через реактор наблюдается переход от периодических колебаний концентраций (рис. 2.82) к хаотическим со сплошным спектром (рис. 2.84).

21.4 Отображение Пуанкаре

Кроме описанных критериев динамического хаоса — характеристических показателей Ляпунова, автокорреляционной функции и спектральной плотности — очень часто (в особенности при исследовании гамильтоновых систем) применяется еще один весьма простой критерий, основанный на использовании отображения Пуанкаре. Построение отображения Пуанкаре для гамильтоновых систем было подробно описано в разделе 2.14. Здесь же мы кратко остановимся на основных случаях, которые наиболее часто встречаются при анализе динамики диссипативных систем.

Предположим, что фазовый поток, порождаемый динамической системой (2.127), имеет секущую поверхность S . Отметим на этой поверхности точки 1, 2, 3, 4, в которых фазовый поток пересекает S в определенном направлении (рис. 2.85а). Пусть для простоты фазовое пространство системы является трехмерным; обобщение на n -мерный случай дано в разделе 14.2. Тогда, если в сечении S мы имеем конечный набор точек, при эволюции системы последовательно переходящих друг в друга (рис. 2.85а), в фазовом пространстве этому соответствует движение фазовой точки по предельному циклу, т. е. периодическая динамика системы. Если точки пересечения потока с поверхностью S образуют инвариантную замкнутую кривую γ , в фазовом пространстве этому соответствует движение по тору (рис. 2.85б). Следовательно, в этом случае поведение системы будет квазипериодическим.

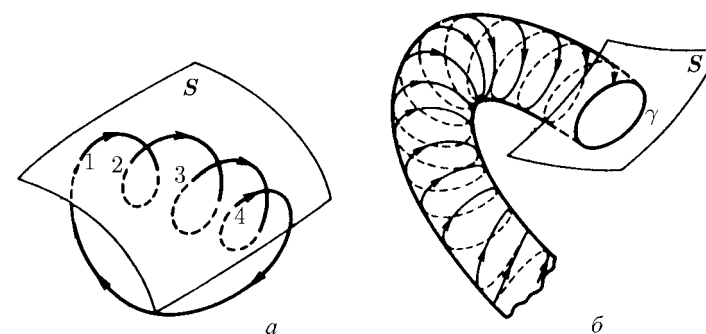


Рис. 2.85. Периодический (а) и квазипериодический (б) режимы движения

Когда же траектория динамической системы на поверхности S представлена множеством точек, нерегулярным образом заполняющих некоторые области (рис. 2.86), это указывает на присутствие в фазовом пространстве странного аттрактора, т. е. на хаотичность режима движения.

Таким образом, построив отображение Пуанкаре, достаточно легко определить характер поведения динамической системы.

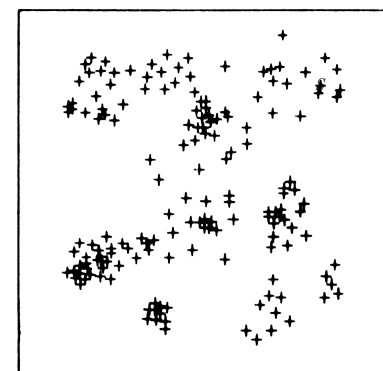


Рис. 2.86. Пример отображения Пуанкаре при хаотическом режиме движения

В заключение отметим, что из-за конечного времени наблюдения (в эксперименте) или интегрировании (при численном анализе), используя перечисленные критерии — спектральную плотность и отображение Пуанкаре, —

строго говоря, невозможно определить, является ли динамика системы периодической с очень большим периодом или же аperiodической. В самом деле, на временном интервале τ , меньшем периода движения T (т. е. при $\tau < T$), спектр будет непрерывным. Кроме того, из-за ограниченности времени наблюдения ширина линий в спектре будет конечной, но не бесконечно тонкой. По этой причине дискретный спектр может быть практически неотличим от сплошного. Сказанное относится и к отображению Пуанкаре. Действительно, при его исследовании заранее неизвестно, сколько точек необходимо отследить, чтобы заметить периодичность движения. Однако во всех практических отношениях нет особой разницы, является ли динамика системы периодической с очень большим периодом или аperiodической.

21.5 Метод Мельникова

Как мы видели в §17, в возмущенных консервативных системах в окрестности сепаратрисы динамика будет хаотической. Для диссипативных систем это не всегда так. В данном разделе, следуя [116] (см. также [66, 72, 355]), мы опишем эффективный метод, позволяющий в приближении малой диссипации и малой амплитуды возмущения аналитически получить достаточно общие условия существования хаоса. Этот метод впервые был предложен В. Мельниковым [146] и заключается в оценке расстояния между сепаратрисами. Однако он применим только в малой окрестности гомоклинической кривой.

Для упрощения изложения рассмотрим двумерную автономную систему с внешним периодическим возмущением

$$\dot{\mathbf{x}} = \mathbf{f}_0(\mathbf{x}) + \varepsilon \mathbf{f}_1(\mathbf{x}, t), \quad (2.156)$$

где $\mathbf{x} = (x_1, x_2)$, $\mathbf{f}_0 = (f_{01}, f_{02})$, $\mathbf{f}_1 = (f_{11}, f_{12})$ и $\mathbf{f}_1(\mathbf{x}, t) = \mathbf{f}_1(\mathbf{x}, t + T)$, которая имеет единственную гиперболическую точку \mathcal{H}_0 . Допустим, что невозмущенная система ($\varepsilon = 0$) имеет гомоклиническую петлю $w_0(t)$ (рис. 2.87а): $\lim_{t \rightarrow \pm\infty} w_0(t) = \mathcal{H}_0$. Возмущение приводит к ее расщеплению, так что входящая w_0^s и выходящая w_0^u ветви уже не образуют единую кривую. Они либо вообще не пересекаются (при этом любая из них будет охватывать другую, рис. 2.87б, в), либо пересекаются в бесконечном числе гомоклинических точек (рис. 2.87г). Хаос возникает только в последнем случае (см. §17).

Следовательно, чтобы найти условие существования хаотичности, необходимо методом теории возмущений вычислить расстояние $D(t, t_0)$ между ветвями сепаратрисы для некоторого момента времени t_0 . Тогда,

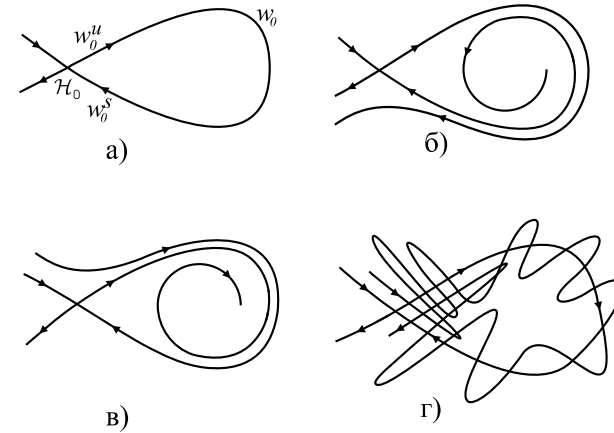


Рис. 2.87. Расщепление петли сепаратрисы

если эта величина не меняет знак, то пересечения ветвей w_0^s и w_0^u нет. Однако, если при каком-либо t_0 функция $D(t, t_0)$ изменит свой знак, возникают гомоклинические сплетения (рис. 2.87г) и хаотическое движение.

Предложенный в работе [146] метод основан на сравнении членов первого порядка в разложении сепаратрисных решений в ряды по параметру возмущения ε . Таким образом, для вычисления $D(t, t_0)$ достаточно найти в первом приближении устойчивую w^s и неустойчивую w^u ветви. Пусть

$$w^{s,u}(t, t_0) = w_0(t - t_0) + \varepsilon w_1^{s,u}(t, t_0).$$

Тогда, используя (2.156), в первом порядке по ε получим:

$$\frac{dw_1^{s,u}}{dt} = \hat{M}(w_0) w_1^{s,u} + \varepsilon \mathbf{f}_1(w_0(t - t_0), t), \quad (2.157)$$

где

$$\hat{M}(w_0) = \begin{pmatrix} \frac{\partial f_{01}}{\partial x_1} & \frac{\partial f_{01}}{\partial x_2} \\ \frac{\partial f_{02}}{\partial x_1} & \frac{\partial f_{02}}{\partial x_2} \end{pmatrix}$$

— матрица-столбец Якоби, взятая на невозмущенной гомоклинической траектории $w_0(t - t_0)$.

При внесении возмущения в интегрируемую систему исходная гиперболическая точка \mathcal{H}_0 сместится в точку \mathcal{H}_p . При этом, очевидно, устойчивая и неустойчивые ветви будут удовлетворять соотношению $\lim_{t \rightarrow \infty} w^s = \lim_{t \rightarrow -\infty} w^u = \mathcal{H}_p$. Введем вектор

$$\mathbf{d}(t, t_0) = w^s(t, t_0) - w^u(t, t_0).$$

В первом приближении, очевидно,

$$\mathbf{d}(t, t_0) = w_1^s(t, t_0) - w_1^u(t, t_0).$$

Следуя работе [146], расстояние $D(t, t_0)$ между устойчивой и неустойчивой ветвями сепаратрисы определяется как проекция \mathbf{d} на нормаль \mathbf{N} к невозмущенной петле w_0 в момент времени t (рис. 2.88):

$$D(t, t_0) = \mathbf{N} \cdot \mathbf{d}. \quad (2.158)$$

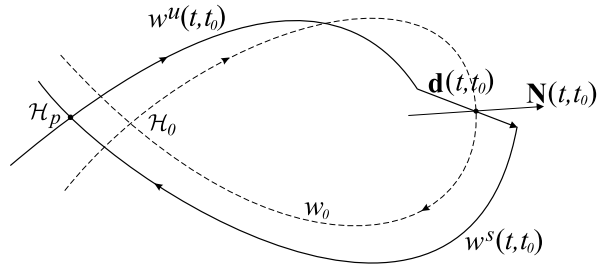


Рис. 2.88. Определение расстояния $D(t, t_0) = \mathbf{N} \cdot \mathbf{d}$ между устойчивой и неустойчивой ветвями сепаратрисы [116]

При $\epsilon = 0$ из (2.156) можно определить вектор \mathbf{N} как

$$\mathbf{N}(t, t_0) = \begin{pmatrix} -f_{02}(w_0) \\ f_{01}(w_0) \end{pmatrix}.$$

Если ввести оператор $\mathbf{x} \wedge \mathbf{y} = x_1 y_2 - x_2 y_1$, то соотношение (2.158) можно представить в более простом виде:

$$D(t, t_0) = \mathbf{f}_0 \wedge \mathbf{d}.$$

Перепишем это выражение как

$$D = D^s - D^u, \quad (2.159)$$

где $D^{s,u}(t, t_0) = \mathbf{f}_0 \wedge w_1^{s,u}$. Теперь, дифференцируя по времени, для D^s найдем:

$$\dot{D}^s = \dot{\mathbf{f}}_0 \wedge w_1^s + \mathbf{f}_0 \wedge \dot{w}_1^s = (\hat{M}(w_0) \cdot \dot{w}_0) \wedge w_1^s + \mathbf{f}_0 \wedge \dot{w}_1^s.$$

Поэтому из (2.157), учитывая, что $\dot{w}_0 = \mathbf{f}_0$, получим:

$$\dot{D}^s = \dot{\mathbf{f}}_0 \wedge w_1^s + \mathbf{f}_0 \wedge \dot{w}_1^s = (\hat{M}(w_0) \cdot \mathbf{f}_0) \wedge w_1^s + \mathbf{f}_0 \wedge (\hat{M}(w_0) \cdot w_1^s) + \mathbf{f}_0 \wedge \mathbf{f}_1$$

либо

$$\dot{D}^s = \text{Sp} \hat{M}(w_0) \mathbf{f}_0 \wedge w_1^s + \mathbf{f}_0 \wedge \mathbf{f}_1 = \text{Sp} \hat{M}(w_0) D^s + \mathbf{f}_0 \wedge \mathbf{f}_1. \quad (2.160)$$

Асимптотически расстояние $D^s(t, t_0)|_{t \rightarrow \infty} = \mathbf{f}_0(\mathcal{H}_0) \wedge w_1^s = 0$. Если невозмущенная система является слабодиссипативной, то $\text{Sp} \hat{M} \simeq 0$. Стало быть, интегрируя уравнение (2.160) на полуинтервале $[t_0, \infty)$, имеем:

$$D^s(t_0, t_0) = - \int_{t_0}^{+\infty} \mathbf{f}_0 \wedge \mathbf{f}_1 dt.$$

Аналогично для D^u находим, что

$$D^u(t_0, t_0) = \int_{-\infty}^{t_0} \mathbf{f}_0 \wedge \mathbf{f}_1 dt. \quad (2.161)$$

Таким образом, вспоминая, что искомая функция D определяется выражением (2.159), приходим к окончательному соотношению:

$$D = - \int_{-\infty}^{\infty} \mathbf{f}_0 \wedge \mathbf{f}_1 dt. \quad (2.162)$$

Полученная функция (которая иногда называется *функцией Мельникова*) описывает расщепление сепаратрисы гиперболической точки при внесении в интегрируемую систему малого возмущения. Если она знакопеременна,

то устойчивая и неустойчивая ветви пересекаются, образуя гомоклинические сплетения, что влечет появление хаотической динамики.

Нетрудно понять, что этот критерий работает также для траекторий гетероклинического типа.

В качестве примера использования представленного подхода рассмотрим достаточно известную модель, на которой исследуются всевозможные приложения методов аналитического анализа нелинейных систем, — осциллятор Дуффинга [292] и его обобщение, систему Дуффинга–Холмса.

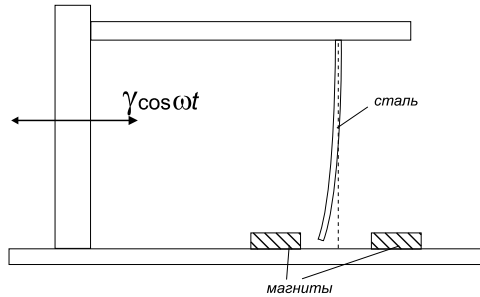


Рис. 2.89. Схема осциллятора Дуффинга [72]

Осциллятор Дуффинга можно представить как систему с двумя равновесными состояниями (или с двумя потенциальными ямами): например, как закрепленный за один конец стальной стержень, находящийся в поле двух магнитов (рис. 2.89). Эта модель имеет одно неустойчивое состояние равновесия (центральное положение) и два устойчивых, соответствующих положению вблизи каждого магнита. В отсутствие внешнего возбуждения свободный конец стержня будет находиться около одного из магнитов.

Уравнение, описывающее такую систему, имеет вид $\ddot{x} + \delta \dot{x} - x + x^3 = 0$. Нетрудно понять, что оно решается аналитически. В отсутствие трения ($\delta = 0$) фазовый портрет данной системы представляется седловой точкой, расположенной в начале координат, и двумя центрами, которые окружены сепаратрисой, похожей на символ ∞ (рис. 2.90а). Сепаратриса разделяет два качественно различных типа движения: колебания вблизи одного из магнитов и колебания от магнита к магниту. При включении диссипации ($\delta > 0$) центры переходят в устойчивые фокусы (рис. 2.90б).

Если теперь начать периодически возмущать стержень (см. рис. 2.89), т. е. добавить в правую часть гармоническую функцию, то динамика системы сильно усложнится и при некоторых значениях параметров может стать хаотической. Этот процесс не является переходным: при фиксированных

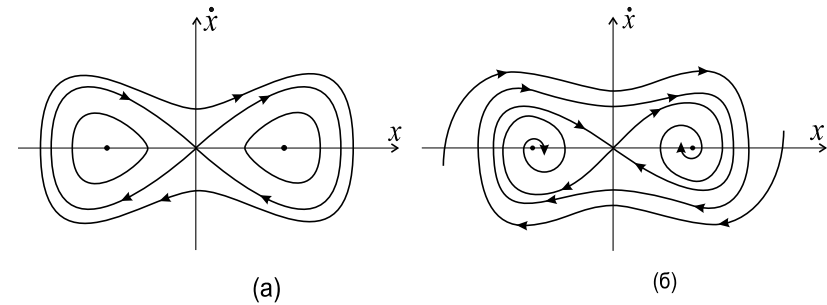


Рис. 2.90. Фазовое пространство системы Дуффинга в отсутствие трения (а) и с диссипацией (б)

значениях амплитуды γ и частоты возбуждения ω хаотическое движение может наблюдаться неограниченно долго.

Уравнение, описывающее движение стержня, называется *уравнением Дуффинга*:

$$\ddot{x} + \delta \dot{x} - x + x^3 = \gamma \cos \omega t. \quad (2.163)$$

В настоящее время это уравнение весьма обстоятельно изучено как численно и аналитически, так и экспериментально (см., например, [72, 104, 155, 158] и приведенные там ссылки). В частности, при определенных значениях параметров оно имеет хаотические решения (рис. 2.91).

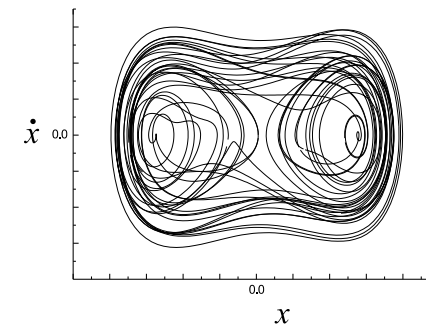


Рис. 2.91. Хаотическое поведение системы Дуффинга

Иногда рассматривается так называемое уравнение Дуффинга–Холмса — осциллятор с параметрическим возмущением кубического члена:

$$\ddot{x} + \delta \dot{x} - x + \beta(1 + \eta \cos \Omega t)x^3 = \gamma \cos \omega t, \quad (2.164)$$

где $\eta \ll 1$ — амплитуда и Ω — частота параметрического возмущения. На примере этого уравнения вычислим функцию Мельникова. Перепишем его в более удобной форме:

$$\dot{x} = y, \quad \dot{y} = x - \beta(1 + \eta \cos \Omega t)x^3 - \delta y + \gamma \cos \omega t. \quad (2.165)$$

Соответствующий невозмущенный гамильтониан можно записать как $H_0 = y^2/2 - x^2/2 + \beta x^4/4$.

При $H_0 = 0$ система (2.165) имеет единственную гиперболическую точку $x = y = 0$ и две гомоклинические петли (рис. 2.90a). Нетрудно найти закон движения по такой сепаратрисе:

$$x_0(t) = \kappa / \operatorname{ch} t, \quad y_0(t) = \dot{x}_0(t) = -\kappa \frac{\operatorname{sh} t}{\operatorname{ch}^2 t}, \quad (2.166)$$

где $\kappa = \sqrt{2/\beta}$. Теперь, используя общий вид уравнения (2.156), из (2.165) получим:

$$f_{01} = y; \quad f_{11} = 0;$$

$$f_{02} = x - \beta x^3; \quad f_{12} = \gamma \cos \omega t - \eta \beta x^3 \cos \Omega t - \delta y.$$

Поэтому $\mathbf{f}_0 \wedge \mathbf{f}_1 = y_0 [\gamma \cos \omega t - \eta \beta x_0^3 \cos \Omega t - \delta y_0]$. Следовательно, из (2.162) найдем:

$$D = - \int_{-\infty}^{+\infty} dt [\gamma y_0(t-t_0) \cos \omega t - \eta \beta x_0^3(t-t_0) \cdot y_0(t-t_0) \cos \Omega t - \delta y_0^2(t-t_0)]. \quad (2.167)$$

Таким образом, окончательно выражение для функции Мельникова системы Дуффинга–Холмса запишется как (см. [279]):

$$D(t_0) = -A \sin \omega t_0 + B \sin \Omega t_0 + C, \quad (2.168)$$

где постоянные

$$A = \pi \kappa \gamma \omega \operatorname{sch} \frac{\pi \omega}{2}, \quad B = \frac{1}{24} \pi \beta \eta \kappa^4 \Omega^2 (4 + \Omega^2) \operatorname{csch} \frac{\pi \Omega}{2}, \quad C = \frac{2}{3} \kappa^2 \delta.$$

Введем подмножество Δ_c множества параметрических значений δ , такое, что если $\delta \in \Delta_c$, то система (2.165) проявляет хаотические свойства.

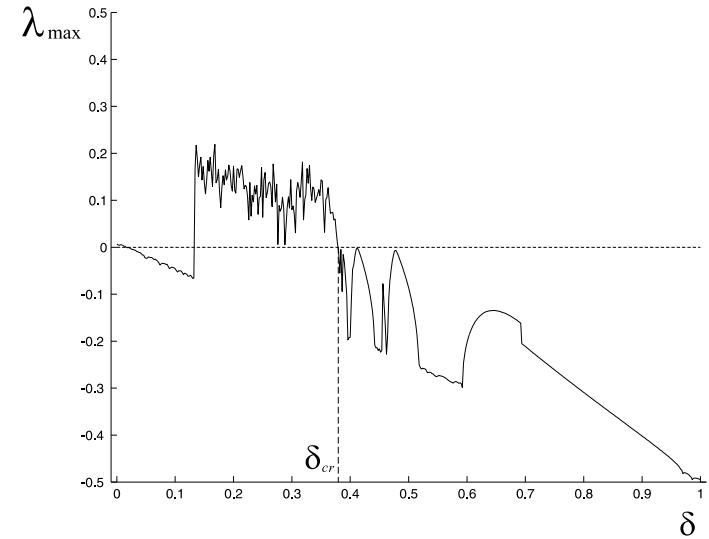


Рис. 2.92. Зависимость старшего показателя Ляпунова от параметра δ для системы (2.164) при $\omega = \Omega = 1, 1$; $\gamma = 0, 114$; $\eta = 0, 03$; $\beta = 8$

Чтобы найти это хаотическое подмножество, проверим, при каких значениях δ функция D может менять знак. Если $\omega = \Omega$, то эти значения выписываются явно:

$$\delta < \delta_{cr} \equiv \frac{3}{2\kappa^2} |B - A|. \quad (2.169)$$

Если же $\omega \neq \Omega$, тогда необходимо дополнительно рассмотреть, когда частоты ω и Ω будут соизмеримы, а когда несоизмеримы. Если ω/Ω — число иррациональное, то нетрудно найти, что

$$\delta < \delta_{cr} \equiv \frac{3}{2\kappa^2} (A + B). \quad (2.170)$$

Когда $\omega/\Omega = m/k$, где m и k — целые числа, существует общее выражение для δ_{cr} (см. [410]). Следовательно, для каждого конкретного набора параметров область хаотичности дается выражением $\Delta_c = [0, \delta_{cr}]$.

Проверим, насколько точна полученная величина δ_{cr} и, тем самым, насколько корректен данный критерий возникновения хаоса. Для этого положим в уравнении (2.164) $\beta = 8$; $\gamma = 0, 114$; $\eta = 0, 03$; $\omega = \Omega = 1, 1$, а параметр δ оставим свободным. Тогда из соотношения (2.169) получим,

что $\delta_{cr} = 0,3798$. Теперь проанализируем систему численно, выбрав в качестве критерия возникновения хаоса зависимость старшего показателя Ляпунова λ_{max} от параметра δ . Результат такого анализа показан на рис. 2.92. Из рисунка видно, что $\delta_{cr} = 0,378$. Это находится в прекрасном согласии с аналитически найденным значением.

Таким образом, метод Мельникова позволяет в ряде случаев аналитически определять границу возникновения гомо- и гетероклинического хаоса. И хотя мы описали использование этого метода на примере неавтономных двумерных систем, его удастся обобщить и на многомерные системы [352]. Подчеркнем, однако, что это все же локальный критерий, который можно использовать в окрестности невозмущенных сепаратрисных гомоклинических и гетероклинических петель, и он ничего не говорит о появлении странного аттрактора.

22. Размерность странных аттракторов

Диссипативные динамические системы, как отмечалось выше, обладают тем свойством, что их решения при $t \rightarrow \infty$ притягиваются к некоторому подмножеству меры нуль в фазовом пространстве. Это подмножество для случая регулярной динамики может быть либо устойчивой стационарной точкой, либо устойчивым предельным циклом, либо инвариантным тором. Все эти подмножества являются подмногообразиями фазового пространства. Математическим образом хаотических колебаний диссипативных систем служит странный аттрактор, который уже не обладает гладкой структурой и достаточной непрерывностью, предполагаемой в понятии подмногообразия. Геометрическое строение странных аттракторов более сложное. Они обладают *геометрической (масштабной) инвариантностью*, или, как иногда говорят, *скейлинговой структурой*.

22.1 Геометрическая инвариантность

Чтобы лучше представить себе, о чем идет речь, рассмотрим характерный пример — аттрактор Эно [203], возникающий в простой модели, описываемой отображением Эно. Мы уже знаем, что странные аттракторы могут появляться в системах дифференциальных уравнений, размерность фазового пространства которых больше или равна трем, $n \geq 3$. Однако сложные геометрические притягивающие множества могут возникать и в так называемых отображениях — динамических системах с дискретным временем (подробнее см. разделы 2.24 и 2.25). Используя отображения, можно описывать системы самой различной природы — от физических до биологиче-

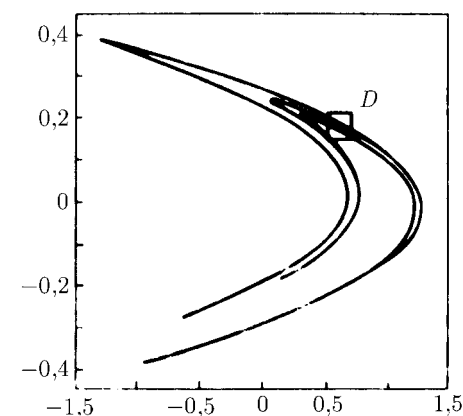


Рис. 2.93. Аттрактор Эно

ских. Отображение Эно — это обратимое двумерное отображение, которое в принципе можно рассматривать как отображение Пуанкаре для некоторой двумерной секущей поверхности и трехмерного потока (см. раздел 2.24).

Если аттрактор в динамической системе является странным, то отображение на секущей будет представлено сложным множеством точек, притягивающим соседние. Такое множество точек, возникающее в отображении Эно, и образует аттрактор Эно, показанный на рис. 2.93. Внимательно анализируя рис. 2.93, можно заметить, что аттрактор Эно состоит из кривых, которые имеют конечную толщину, т. е. фактически внутреннюю структуру. Чтобы убедиться в этом, увеличим определенный участок D аттрактора Эно (рис. 2.93). Тогда мы увидим, что часть линий, образующих аттрактор, действительно состоит из нескольких «кривых» (рис. 2.94а). Продолжая процесс последовательного увеличения, можно убедиться, что каждая из этих «кривых» образована бесконечным числом компонент (рис. 2.94б, в), т. е. структура аттрактора Эно при увеличении масштаба практически повторяется, причем сохраняется и масштабный множитель [203]. О множествах, имеющих такую структуру, говорят, что они обладают геометрической (масштабной) инвариантностью.

22.2 Канторовы множества

Чтобы научиться количественно описывать такие сложные множества, в частности странные аттракторы, обратимся сначала к хорошо известным

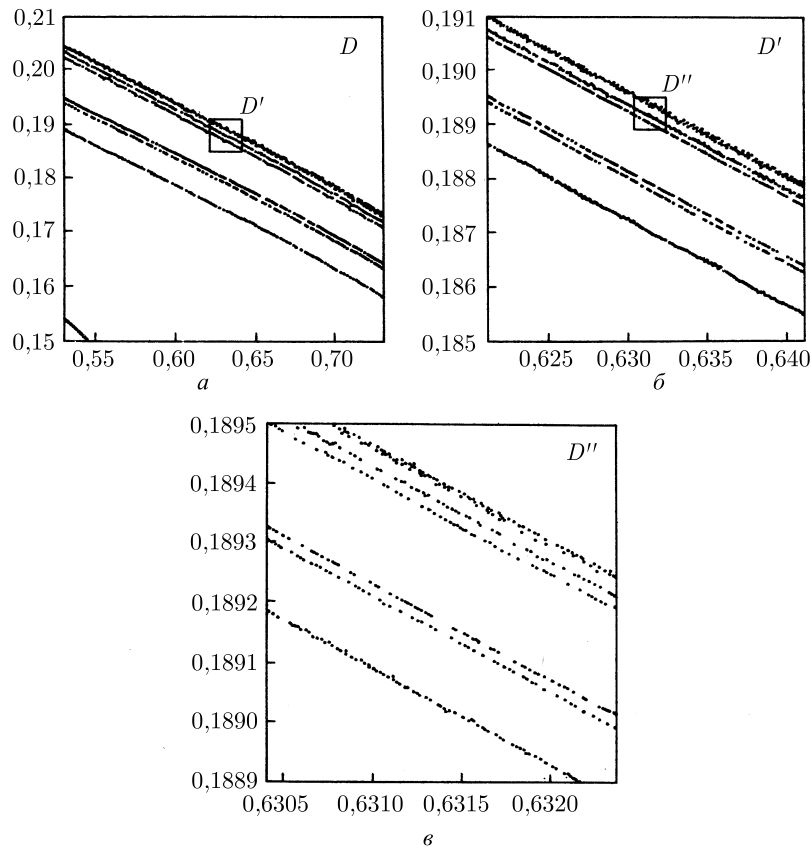


Рис. 2.94. Последовательное увеличение малого участка аттрактора Эно. Отчетливо видна геометрическая инвариантность аттрактора

в математике *канторовым множествами*, обладающим той же геометрической инвариантностью. Самым известным примером здесь является так называемое «*множество средних третей*».

Возьмем отрезок единичной длины $[0, 1]$, разделим его на три равные части и вырежем среднюю из них — интервал $(1/3, 2/3)$. С каждым из оставшихся отрезков поступим точно так же. Тогда мы получим последовательность отрезков все убывающей длины (рис. 2.95). На первом этапе

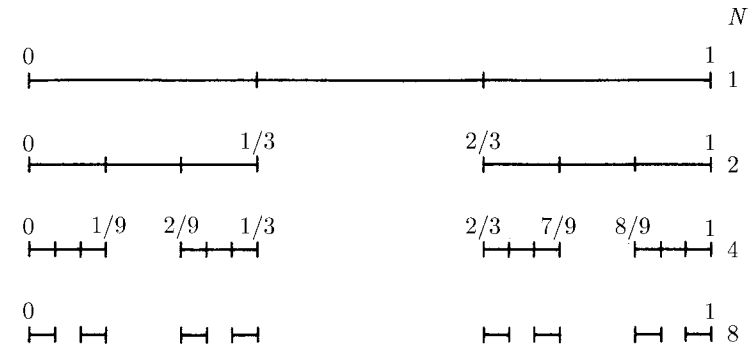


Рис. 2.95. Построение множества Кантора

построения имеем один отрезок, на втором — два, на третьем — четыре. На k -м этапе будем иметь 2^k отрезков, не связанных друг с другом, длиной 3^{-k} каждый. При $k \rightarrow \infty$ получим некоторое множество точек, которое и называется *множеством Кантора*. Суммарная длина всех вырезанных отрезков при этом равна единице:

$$l = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots = \frac{1}{2} \sum_{k=1}^{\infty} \left(\frac{2}{3}\right)^k = \frac{1/3}{1 - 2/3} = 1. \quad (2.171)$$

Заметим, что при построении канторова множества вовсе не обязательно разбивать отрезок на три и притом равные части: канторово множество можно получить, разделяя отрезок каждый раз на произвольное число ($i \geq 3$) частей.

Рассмотрим теперь другой пример масштабно-инвариантного множества. Это непосредственное обобщение канторова множества средних третей на случай плоских фигур — *ковер Серпиньского*. Возьмем квадрат со стороной, равной единице, и разделим его на девять равных квадратов прямыми, параллельными его сторонам. Затем при первой итерации ($k = 1$) удалим все внутренние точки центрального квадрата. Аналогично поступим с каждым из оставшихся восьми квадратов (вторая итерация, $k = 2$) и т. д. (рис. 2.96). Пересечение полученных при $k \rightarrow \infty$ множеств — это ковер Серпиньского.

Таким образом, очень грубо можно сказать, что канторово множество является как бы «всюду дырявым». Оказывается, что странные аттракторы для динамических систем и отображений обычно имеют структуру канто-

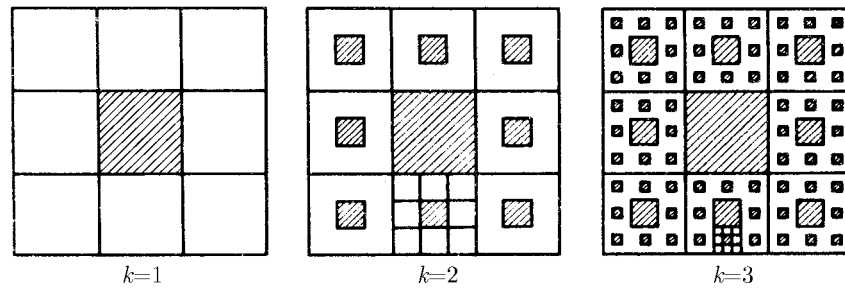


Рис. 2.96. Построение ковра Серпиньского. Заштрихованные области удаляются в процессе итераций

рова множества [116, 203, 264]. Этим они с геометрической точки зрения отличаются от простых аттракторов, являющихся подмножествами.

Существует важная количественная характеристика канторова множества — дробная размерность, указывающая на близость этого множества к соответствующему гладкому подмножеству. Есть несколько различных определений дробной размерности. Обычно они основаны на метрических свойствах множеств. Но для странных аттракторов предложены также вероятностные размерности, при вычислении которых главную роль играет частота, с которой типичная фазовая траектория посещает различные области аттрактора. Ниже мы дадим определение наиболее распространенной метрической размерности — емкости множества. Обзор основных вопросов, касающихся различных размерностей странных аттракторов, взаимосвязи между ними и способов их вычисления, можно найти в [54, 298, 476].

22.3 Фрактальная размерность аттракторов

Понятие дробной размерности, по существу, опирается на понятие целой размерности. Хорошо известно, например, что отрезок и предельный цикл являются одномерными, плоскость, сфера и тор — двумерными, и т. д. Это можно интерпретировать так, что для однозначного задания точки на отрезке достаточно знать только одну координату, а на торе — две, и т. д. Рассмотрим теперь более сложный объект — множество с масштабнo-инвариантной структурой. Трудность состоит в том, что в этом случае мы заранее не знаем, какое минимальное число координат нужно для однозначного описания его точек. В принципе, здесь можно поступить следующим образом. Допустим, что координатная линия все более плотно заполняет

круг, заключающий в себе геометрически инвариантное множество. Тогда положение любой точки этого множества можно определить с помощью одной координаты. Близким способом, в частности, задается адрес человека в городе. Однако такой подход мало эффективен. Поэтому поступают по-другому.

Разобьем поверхность круга, заключающего в себе некоторое множество, на ячейки с характерным размером ϵ и подсчитаем число ячеек, в которые попали его части. Обозначим их число через N . При уменьшении размера ячейки ($\epsilon \rightarrow 0$) число N будет возрастать ($N \rightarrow \infty$). Но если рассмотреть отношение числа N к величине $1/\epsilon$, то, поскольку порядок роста N и $1/\epsilon$ одинаков, можно ожидать, что полученное значение будет конечной и вполне определенной величиной для данного множества. Ясно, что это отношение может и не быть целым числом. Тогда, округлив его до большего целого, получим минимальное число координат, необходимое для однозначного задания точек на этом множестве.

Формализация подобного подхода и приводит к понятию дробной размерности.

Рассмотрим в n -мерном фазовом пространстве динамической системы некоторое множество A . Покроем это множество n -мерными кубиками со стороной ϵ так, чтобы эти кубики содержали все точки множества A . Пусть N — минимальное число кубиков, необходимых для покрытия A . Рассмотрим предел

$$d(A) = \lim_{\epsilon \rightarrow 0} \frac{\ln N(\epsilon)}{\ln(1/\epsilon)}. \quad (2.172)$$

Величина $d(A) \equiv d_F$ является метрической размерностью и называется емкостью [99, 298] или *фрактальной*⁹ *размерностью* [142, 416]. Заметим, что в литературе емкость иногда называют также *хаусдорфовой* или *энтропийной размерностью* [193, 298, 476]. Это оправдано тем, что для типичных странных аттракторов (которые встречаются в численных расчетах и физических экспериментах) значения емкости, хаусдорфовой и энтропийной размерности, по-видимому, очень близки [298, 476].

Для регулярных множеств (например, «куска» трехмерного евклидова пространства, поверхности или линии) фрактальная размерность d_F равна целому числу (соответственно 3, 2, 1) и совпадает с обычной размерностью. Действительно, при малых ϵ из (2.172) получаем

$$N \sim K\epsilon^{-d}. \quad (2.173)$$

⁹От латинского *fractus* — фрагментированный

Число кубиков со стороной ϵ , необходимое для покрытия некоторого объема в трехмерном пространстве, пропорционально $1/\epsilon^3$, для покрытия поверхности — $N \sim 1/\epsilon^2$ и линии — $N \sim 1/\epsilon$. Следовательно, если мы имеем простой аттрактор, например, предельный цикл, то его фрактальная размерность равна единице. Для инвариантного тора значение d_F равно двум и т. д.

Однако для нерегулярных множеств, обладающих масштабно-инвариантной структурой, фрактальная размерность имеет дробное значение. Остановимся на этом более детально.

Определим вначале фрактальную размерность множеств, рассмотренных выше, — множества средних третей и ковра Серпиньского.

Из построения множества средних третей следует, что оно состоит из $N = 2^m$ разделенных интервалов длиной $\epsilon = 1/3^m$ каждый. В самом деле, при $k = 0$ имеем $N = 1$, $\epsilon = 1$. Если $k = 1$, то $N = 2$, $\epsilon = 1/3$; для $k = 2$ $N = 4$, $\epsilon = 1/9$, и при $k = m$ $N = 2^m$, $\epsilon = 1/3^m$ (рис. 2.95). Здесь k обозначает число итераций построения множества. Следовательно, используя определение (2.172), получим

$$d_F = \lim_{m \rightarrow \infty} \frac{\ln 2^m}{\ln 3^m} = \frac{\ln 2}{\ln 3} \approx 0,631. \quad (2.174)$$

Определим теперь фрактальную размерность ковра Серпиньского. Имеем

$$\begin{array}{lll} k = 1, & N = 8 = 8^1, & \epsilon = 1/3^1, \\ k = 2, & N = 8 \cdot 8 = 8^2, & \epsilon = 1/3^2, \\ k = 3, & N = 8 \cdot 8 \cdot 8 = 8^3, & \epsilon = 1/3^3, \\ \vdots & \vdots & \vdots \\ k = m & N = 8^m, & \epsilon = 1/3^m, \end{array}$$

отсюда

$$d_F = \lim_{m \rightarrow \infty} \frac{\ln 8^m}{\ln 3^m} = \frac{\ln 8}{\ln 3} \approx 1,893. \quad (2.175)$$

Таким образом, ковер Серпиньского — это уже не линия, размерность которой равна единице, но еще и не поверхность, поскольку размерность поверхности равна двум. Это что-то «между» линией и двумерной поверхностью. Неожиданным является то, что реальные фрактальные множества, рождающиеся в физических экспериментах, представляют собой аналог ковра Серпиньского в том смысле, что их размерность d_F тоже имеет дробное значение (см. §23).

Как мы уже отмечали, странные аттракторы обычно близки по своей структуре к канторовым множествам, поэтому следует ожидать, что размерность странного аттрактора будет дробной. Таким образом, значение размерности можно использовать в качестве критерия отличия простых аттракторов от странных. В основополагающей работе Рюэля и Такенса [173] термин «странный аттрактор» был введен авторами именно для того, чтобы подчеркнуть, что такие аттракторы не являются гладкими многообразиями.

Ввиду чрезвычайной важности фрактальной размерности возникает вопрос о явном ее вычислении для тех или иных аттракторов динамических систем.

22.4 Гипотеза Каплана–Йорка

Существует гипотеза, выдвинутая Капланом и Йорком [383], согласно которой фрактальная размерность связана с характеристическими показателями Ляпунова λ_i . Эта гипотеза предполагает, что фрактальная размерность d_F совпадает с *ляпуновской размерностью* d_L , определяемой как

$$d_L = j + \sum_{i=1}^j \frac{\lambda_i}{|\lambda_{j+1}|}, \quad (2.176)$$

где все показатели λ_i упорядочены: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ (n — размерность фазового пространства), а число j определяется из условий

$$\lambda_1 + \lambda_2 + \dots + \lambda_j \geq 0, \quad \lambda_1 + \lambda_2 + \dots + \lambda_{j+1} < 0. \quad (2.177)$$

Для трехмерных динамических систем (т. е. для систем с размерностью фазового пространства $n = 3$) в предположении, что движение происходит на странном аттракторе, формула (2.176) сводится к виду

$$d_L = 2 + \frac{\lambda_1}{|\lambda_3|}, \quad (2.178)$$

поскольку для таких систем $\lambda_1 \geq \lambda_2 \geq \lambda_3$ и (см. раздел 2.21) $(\lambda_1, \lambda_2, \lambda_3) = (+, 0, -)$.

Физически кажется ясным, что для того, чтобы установить геометрическую структуру странного аттрактора, необходимо взять какую-либо малую область фазового пространства и проследить, как с течением времени она эволюционирует. Информацию об изменении малого элемента фазового объема динамической системы дают характеристические показатели Ляпунова.

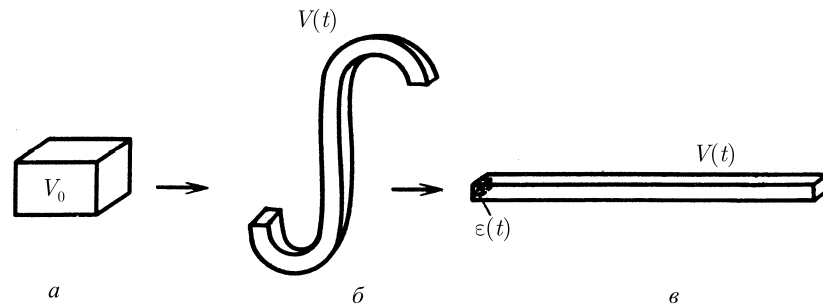


Рис. 2.97. К пояснению гипотезы Каплана–Йорка (2.178)

Покажем на примере трехмерной ($n = 3$) динамической системы со странным аттрактором, как можно объяснить смысл гипотезы Каплана–Йорка. Рассмотрим в начальный момент времени $t = 0$ малый элемент V_0 фазового объема и проследим за его эволюцией при $t > 0$. Поскольку система является диссипативной, а движение происходит на странном аттракторе, то с течением времени элемент объема V_0 будет сжиматься и сильно деформироваться (рис. 2.97а, б). Будем постоянно (мысленно) распрямлять этот деформированный объем $V(t)$, покрывая его в каждый момент времени трехмерными кубиками со стороной $\epsilon(t)$ (рис. 2.97в). Как известно (см. раздел 2.21), характеристические показатели Ляпунова определяют степень сжатия фазового объема в одних направлениях и растяжения — в других. Используя показатели Ляпунова, изменение со временем элемента объема V можно записать следующим образом:

$$V(t) \sim V_0 e^{(\lambda_1 - |\lambda_3|)t}, \quad (2.179)$$

где нумерация λ_i выбрана так, что $\lambda_1 \geq \lambda_2 \geq \lambda_3$ и $\lambda_1 \geq 0$, $\lambda_2 = 0$, $\lambda_3 \leq 0$. Изменение $\epsilon(t)$ записывается в виде

$$\epsilon(t) \sim e^{-|\lambda_3|t}, \quad (2.180)$$

$$\epsilon^3(t) \sim e^{-3|\lambda_3|t}, \quad (2.181)$$

где $\epsilon^3(t)$ — объем кубиков со стороной $\epsilon(t)$, покрывающих элемент $V(t)$ фазового объема в момент времени t . Тогда число $N(t)$ таких кубиков

$$N(t) \sim \frac{V(t)}{\epsilon^3(t)} = e^{\lambda_1 t + 2|\lambda_3|t}, \quad (2.182)$$

отсюда

$$d_F = \lim_{t \rightarrow \infty} \frac{\ln N(t)}{\ln(1/\epsilon(t))} = 2 + \frac{\lambda_1}{|\lambda_3|}, \quad (2.183)$$

что совпадает с (2.178).

Из соотношения (2.176) легко получить, что $d_F < n$, как и следовало ожидать. В самом деле, из условия диссипативности динамической системы вытекает, что ее фазовые траектории с течением времени стягиваются к некоторому подмножеству фазового пространства. Размерность этого подмножества всегда меньше n .

Гипотеза (2.176) выполняется в ряде частных случаев [297, 476, 533], а вообще говоря, она несправедлива [54]. Можно показать лишь, что ляпуновская размерность дает оценку сверху для фрактальной (хаусдорфовой) размерности [273, 391]. Поэтому при расчетах размерностных характеристик странных аттракторов динамических систем часто приходится непосредственно использовать определение (2.172) величины d_F и опираться исключительно на численные расчеты.

Чтобы определить d_F , численно исходя из (2.172), необходимо фазовое пространство динамической системы разбить на ячейки со стороной ϵ , а затем произвести численное интегрирование до того момента, пока не закончится переходный процесс, после которого движение является стационарным (установившимся) и происходит на странном аттракторе. При последующем интегрировании необходимо помечать те ячейки, в которые попадает траектория. После достаточно большого числа итераций, т. е. по прошествии длительного промежутка времени численного интегрирования динамической системы, число таких ячеек будет стремиться к $N(\epsilon)$. Далее, из формулы (2.172) нетрудно увидеть, что зависимость $y \equiv \ln N$ от $x \equiv \ln(1/\epsilon)$ при малых ϵ является линейной. Следовательно, если графически изобразить зависимость y от x , то, как следует из (2.172), фрактальная размерность d_F равна тангенсу угла наклона прямой $y(x)$. Численные методы определения фрактальной размерности обсуждаются в [135, 476].

Приведем примеры значений фрактальной размерности странных аттракторов. Для системы Лоренца (2.120) при $\sigma = 10$, $b = 8/3$ и $r = 28$ она равна $d_F = 2,05 \pm 0,01$. Тем самым, оказывается, что аттрактор Лоренца при указанных значениях параметров близок к некоторой двумерной поверхности.

В [170] исследовалась следующая система:

$$\dot{x} = y(z - 1 + x^2) + \gamma x, \quad \dot{y} = x(3z + 1 - x^2) + \gamma y, \quad \dot{z} = -2z(\nu + xy), \quad (2.184)$$

при $\nu = 1,1$ и $\gamma = 0,87$. Д. А. Рассел и др. [476] численно определили размерность d_F странного аттрактора в такой системе и нашли, что он занимает промежуточное положение между поверхностью и трехмерной фигурой, $d_F = 2,318 \pm 0,002$.

Заметим, что когда размерность d_F оказывается близкой к целому числу и трудно оценить погрешность численных расчетов, следует всегда дополнительно использовать другие критерии динамического хаоса (см. раздел 2.21).

23. Фракталы

Масштабная инвариантность, характеризующая меру «странности» хаотических аттракторов, по существу, является универсальным свойством, присущим большому числу явлений природы. Обобщение этого понятия, использование его для обработки экспериментальных данных постепенно привели к формированию новой дисциплины — *фрактальной геометрии*, изучающей так называемые фрактальные объекты.

Термин фрактал впервые был введен Бенуа Мандельбротом (см. [143, 415, 416]) для обозначения объектов, емкость которых всегда больше, чем их топологическая размерность d_T . Например, фрактальная размерность множества средних третей (см. рис. 2.95) составляет $d_F = \ln 2 / \ln 3$, в то время как его топологическая размерность (т.е. фактически размерность множества точек) $d_T = 0$. При этом, однако, вовсе необязательно, чтобы величина d_F была нецелой (см. ниже).

Фрактальная геометрия долгое время не находила широкого применения, пока не было обнаружено большое число задач, где геометрическая инвариантность и размерность служат основными характеристиками системы [186, 200, 336]. Так, если рассмотреть скорость турбулентного потока как функцию пространственных переменных и времени, то она будет представлять собой фрактал того же типа, что и броуновская кривая [143, 200, 520].

У фрактальной теории много точек соприкосновения с методом ренормгруппы и теорией фазовых переходов [452, 499]. Кроме того, в настоящее время очень популярна область, находящаяся на стыке финансовой математики и теории хаоса. Здесь можно многое описать, используя фрактальный подход [144, 454, 455]. Неожиданно важные приложения теории фрактальных множеств были выявлены в достаточно новых областях современной науки — теоретической биологии и математической медицине [200, 248, 443]. Наконец, фрактальные множества интересны с точки зрения создания моста между математикой и искусством [163] (см. §27).

Удивительные свойства фрактальных множеств, их повсеместное распространение и многочисленные парадоксы, связанные с ними, описаны в книгах [143, 200].

23.1 Примеры искусственных и естественных фракталов

Простейшее фрактальное множество можно представить, например, как узор из неких идеализированных «снежинок», показанный на рис. 2.98 и называемый *фракталом Вичека* [200, 520]. Каждый отдельно выбранный элемент фрактала Вичека состоит из пяти идентичных элементов меньшего размера. В свою очередь, соединив пять отдельных частей, подобных фракталу Вичека, получим объект абсолютно той же структуры, что и исходный, и т.д. Очевидно, такое множество является масштабно-инвариантным: на каждом «уровне» часть фрактала выглядит в точности так же, как и все множество.

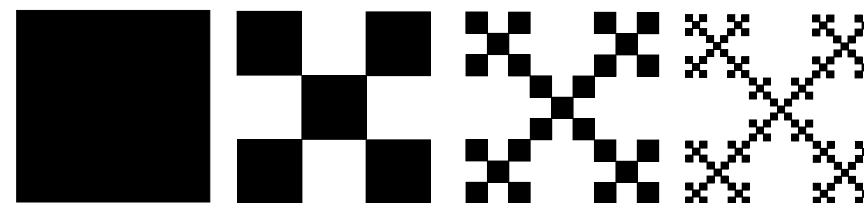


Рис. 2.98. Построение фрактала Вичека

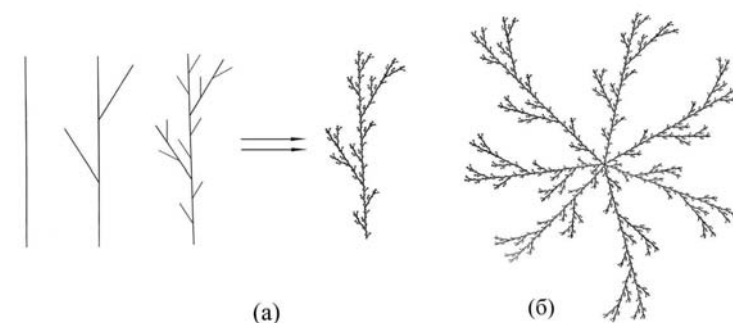


Рис. 2.99. Один из способов построения фрактального «дерева»

Другой простой пример фрактала строится подобно росту «дерева» [186]. От первоначального ствола по обе стороны отрастают две ветви, длина которых в три раза меньше длины ствола. Затем от каждой ветви и ствола отрастает в разные стороны по еще меньшей ветви, и т. д. (рис. 2.99а). Теперь, если обобщить полученное множество и допустить, что «дерево» может иметь несколько стволов, растущих в разные стороны, мы приходим к картине, показанной на рис. 2.99б.

Можно рассмотреть несколько иное построение фрактального дерева, когда на каждой последующей итерации ветвь раздваивается под определенным углом. В этом случае необычна зависимость формы полученного фрактала от угла, под которым «прорастают» ветви (рис. 2.100).

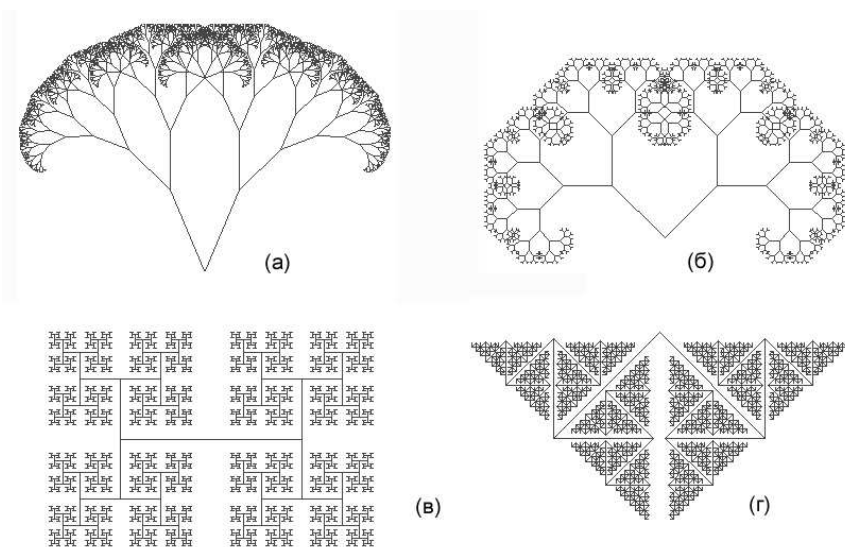


Рис. 2.100. Зависимость формы фрактального «дерева» от угла расхождения «ветвей». а — 45° , б — 90° , в — 180° , г — 270°

Фрактальные структуры, аналогичные показанным на рис. 2.98–2.100, встречаются в природе и рождаются в физических экспериментах. Чтобы убедиться в этом, рассмотрим в качестве примера возникновение так называемых «вязких пальцев» — сложных рыхлых образований, отдаленно напоминающих по форме пальцы перчатки (рис. 2.101). Вязкие пальцы появляются в ячейке Хеле–Шоу, состоящей из двух близко расположенных

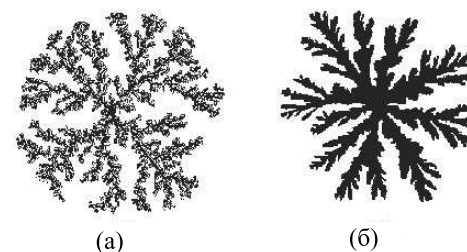


Рис. 2.101. «Вязкие пальцы», полученные в ячейке Хеле–Шоу в случае смешивающихся (а) и несмешивающихся (б) жидкостей

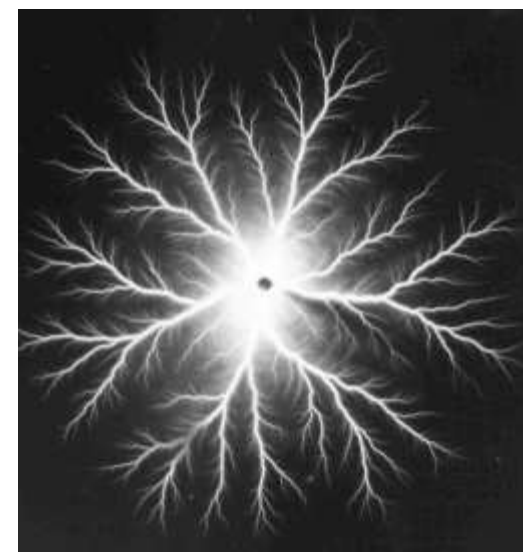


Рис. 2.102. Фигура Лихтенберга

друг к другу параллельных пластин, между которыми заключена жидкость. Если впрыснуть через отверстие в одной из них другую жидкость, то в случае их смешиваемости возникает структура, показанная на рис. 2.101а. При условии *несмешиваемости* двух жидкостей «пальцы» гораздо толще, поскольку в этом случае поверхностное натяжение препятствует образованию тонких отростков (рис. 2.101б). Поразительно сходство вязких пальцев с

фрактальным деревом (рис. 2.99б). Как мы увидим в дальнейшем, это сходство является не только качественным, но и количественным.

Другой пример природного объекта, по форме сильно напоминающий фрактальное дерево, — это *фигура Лихтенберга*. Фигура Лихтенберга образуется вследствие пробоя диэлектрика между точечными положительным и центрально-симметрично окружающим его отрицательным потенциалами (рис. 2.102).

Еще один пример фрактала, также получаемого в эксперименте, — результат осаждения на электроде ионов металла, диффундирующих в тонком слое электролитического раствора. Итог такого эксперимента показан на рис. 2.103.

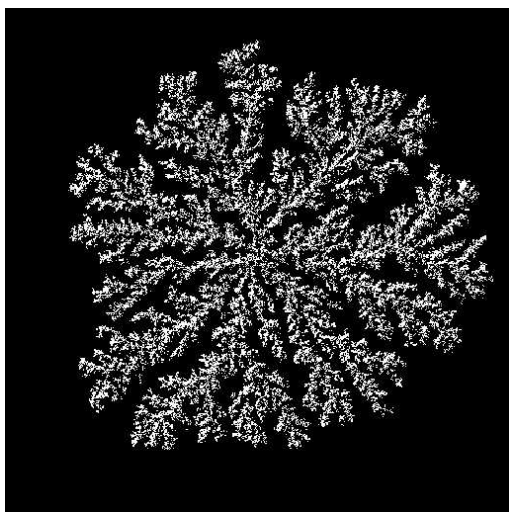


Рис. 2.103. Результат осаждения ионов золота в тонком слое электролита

23.2 Агрегация, ограниченная диффузией

Представляет интерес, как же могут рождаться такие сложные объекты, как фракталы. Один из теоретических механизмов, предложенных для объяснения их образования, — это *агрегация, ограниченная диффузией* (diffusion-limited aggregation, DLA) [423–425, 525, 526]. Согласно этому механизму определенная разновидность фракталов может быть получена в процессе неупорядоченного необратимого роста.

Представим себе объект (кластер), растущий следующим образом: с течением времени к нему присоединяется молекула, причем когда эта молекула приходит в соприкосновение с кластером, она сразу же прилипает к нему. Такой процесс называется *агрегацией*. Допустим теперь, что частицы (молекулы) диффундируют к растущему кластеру абсолютно случайным образом (т. е. по закону движения броуновской частицы). Агрегация частиц, протекающая в условиях случайного блуждания, — это и есть DLA. Формирование кластеров посредством DLA является крайне неравновесным. Тем не менее, в качественном отношении с помощью него можно легко объяснить некоторые свойства роста фрактальных структур.

Предположим, что процесс идет в очень тонком слое раствора, расположенном между двумя пластинами, и рост начинается с гладкой поверхности — зародыша. В результате случайного блуждания молекул в какой-то части поверхности зародыша их налипнет больше, чем в другой. Тогда рост в этом месте будет происходить быстрее, так как вероятность наткнуться на образовавшийся выступ выше, чем вероятность попасть в другое, менее доступное место. Это приводит к тому, что когда следующая частица прилипнет к выступу, он станет еще больше. В итоге исходная форма зародыша, сначала лишь немного искаженная, деформируется еще сильнее. Этот процесс называется *неустойчивостью роста*. В результате такого роста образуется неупорядоченная, сильно разреженная структура — фрактальный кластер.

Значение предложенного подхода состоит в том, что он показывает связь между фрактальными объектами и процессом роста. Кроме того, процесс DLA легко смоделировать на компьютере следующим образом.

Экран компьютера разбивается на квадратную, шестиугольную или треугольную решетку с достаточно мелким шагом. Кластер начинает свой рост с одной частицы, расположенной в центре экрана (т. е. в начале координат). Затем в случайно выбранной точке на окружности R , ограничивающей кластер извне, появляется другая частица, двигающаяся по решетке по броуновскому закону. В тот момент, когда расстояние между центром и второй частицей не превышает шага решетки, она замирает на месте. Затем в случайно выбранном месте на расстоянии R от центра появляется следующая частица и случайным образом начинает движение по решетке. Это блуждание продолжается до тех пор, пока она не окажется на расстоянии, не большем шага решетки, от какой-либо из двух уже имеющихся частиц. И так далее. В результате такого моделирования образуются структуры (рис. 2.104), удивительно похожие на естественные фракталы (см. рис. 2.101–2.103).

Модель DLA нетрудно усложнить во многих направлениях. Например, можно допустить, что частица, прежде чем остановиться, отскакивает

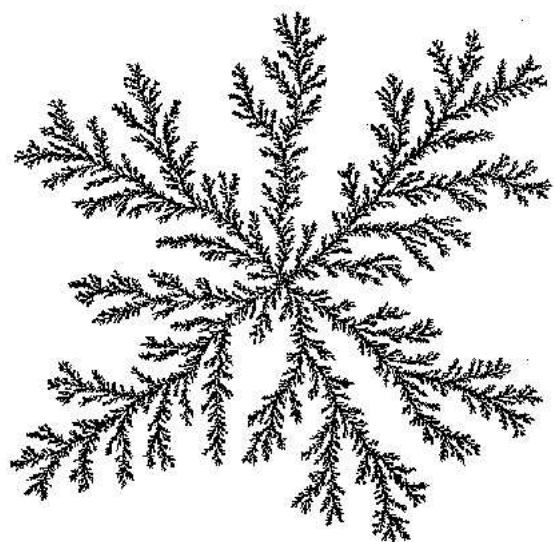


Рис. 2.104. Классический DLA-кластер, состоящий из 10^6 частиц

от поверхности растущего кластера. При этом, однако, основная структура кластера сохраняется; в зависимости от числа допустимых отскоков утолщаются только ветви образующегося кластера.

В случае трехмерных процессов образование некоторых фракталов также удается объяснить с помощью модели агрегации, ограниченной диффузией.

Сходство фракталов, полученных в процессе DLA и в экспериментах не случайно. Образование фрактальных кластеров описывается уравнением Лапласа, которому удовлетворяет потенциал в задачах электростатики, диффузии, потока несжимаемой жидкости и др. Для фигуры Лихтенберга пробой происходит в направлении наибольшего градиента потенциала, а труднодоступные области кластера являются хорошо экранированными. Это и приводит к появлению структуры типа представленной на рис. 2.102.

Для случая диффузии кластер отвечает эквипотенциальной поверхности, а градиент потенциала соответствует полю диффузии. Частицы больше адсорбируются там, где выше градиент, т. е. вблизи выступов, что и является причиной образования фрактала (рис. 2.103, 2.104). При рождении вязких пальцев решающую роль играет гидродинамическая неустойчивость между жидкостями. Градиент давления будет наибольшим у выступов, так что лю-

бая выпуклость на границе раздела имеет тенденцию к росту, в результате чего образуется структура, подобная показанной на рис. 2.101.

23.3 Вероятностная модель роста фрактала¹⁰

Нетрудно построить вероятностную модель роста фрактала, образующегося в процессе DLA [401].

Пусть частицы, образующие кластер, имеют диаметр d . Разобьем плоскость, на которой растет кластер, на n_{\max} концентрических колец шириной d , так чтобы центр колец совпадал с центром кластера. Тогда средний радиус кольца можно определить как $r_n = nd$, $n = 1, \dots, n_{\max}$ (рис. 2.105). Движение частицы в такой модели определим на основе следующих правил:

- Броуновское движение частицы состоит в перемещении в пределах кольца или скачкообразных переходах на соседние кольца. Направление движения определяется вероятностным образом.
- При переходе частица может с определенной вероятностью столкнуться с одной из частиц кластера, что соответствует агрегации частицы на его поверхности, после чего в систему вводится новая частица.

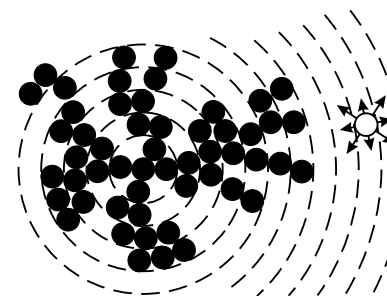


Рис. 2.105. Кольцевая модель DLA-кластера

Значения вероятностей перехода и агрегации частиц зависят от рассматриваемой модели распределения частиц внутри кольца. Наиболее простой, но все еще дающей хорошие результаты, моделью будет ячеечное распределение частиц внутри кольца. В этом случае предполагается, что кольцо разбито на $M_n = [2\pi r_n/d]$ ячеек, внутри которых могут находиться

¹⁰Результаты этой части получены одним из авторов (А. Ю. Л.) совместно с А. Б. Рябовым.

частицы. Здесь $[\dots]$ обозначает выделение целой части. В этом случае вероятность попадания в одну из N_n частиц кластера n -го кольца, очевидно, составит

$$\tilde{p}_n = \frac{N_n}{M_n}. \quad (2.185)$$

Таким образом, вероятность перехода частицы из n -го слоя в $(n+1)$ -й слой равна $p_n^1 = (1 + d/r_n)/3$, а в $(n-1)$ -й слой — $p_n^{-1} = (1 - d/r_n)/3$. При этом вероятность остаться в том же слое $p_n^0 = 1/3$. Следовательно, $p_n^k = (1 + kd/r_n)/3$, $k = -1, 0, 1$.

Наличие слагаемого $\pm d/r_n$ следует из предположения, что вероятность перехода из одного кольца в другое должна быть пропорциональна соотношению длин этих колец. Поэтому вероятность адсорбции частицы, находящейся на n -м слое, определяется как

$$P_n = \sum_{k=-1}^1 \tilde{p}_{n+k} p_n^k. \quad (2.186)$$

Если адсорбции не произошло, то с вероятностью $Q_n = 1 - P_n$ происходит переход. Соответствующие (условные) вероятности перехода могут быть легко найдены на основе формулы Байеса полной вероятности:

$$q_n^k = \frac{p_n^k (1 - \tilde{p}_{n+k})}{Q_n}. \quad (2.187)$$

Построенный таким образом кластер представляет собой набор слоев, каждый из которых характеризуется парой чисел N_n и M_n . Для простоты положим массу одной частицы равной единице. Тогда масса всего кластера будет равна $m = \sum_{n=1}^{n_{\max}} N_n$.

Фрактальную размерность такой системы можно определить исследуя зависимость массы кластера от его радиуса. В случае реального DLA $m(n_{\max}) \propto n_{\max}^{d_F}$, где d_F — размерность кластера. В логарифмических координатах график такой зависимости будет представлять собой прямую с коэффициентом наклона d_F . Продифференцировав его, можно получить локальную размерность как функцию n_{\max} :

$$d_F(n_{\max}) = \frac{d}{d \log n_{\max}} \log m(n_{\max}).$$

Данная модель легко обобщается на трехмерный случай. Тогда пространство разбивается на слои концентрическими сферами. При этом максимальное число частиц внутри слоя можно оценить как $M_n = [\pi r_n^2/d^2]$,

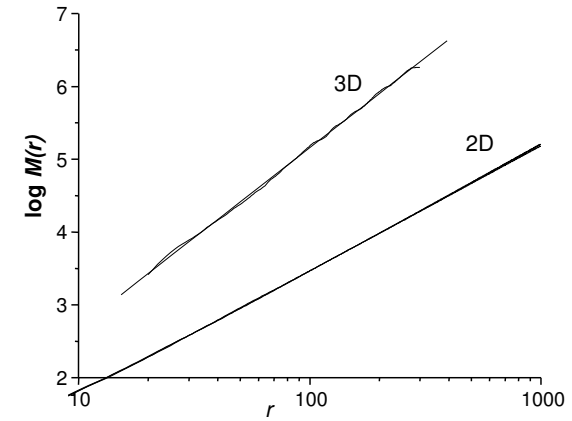


Рис. 2.106. Зависимости массы кластера от его радиуса в двумерном (2D) и трехмерном (3D) случаях

а вероятности перехода

$$p_n^k = \frac{1}{3} \left(1 + k \frac{d}{r^2} \right).$$

Остальные соотношения (2.186)–(2.187) остаются без изменений.

Численный анализ построенной модели для двумерного и трехмерного случаев показан на рис. 2.106. Фрактальные размерности, полученные путем линейной аппроксимации графиков, равны $d_F^{(2)} = 1.727 \pm 0.002$ для двумерного случая и $d_F^{(3)} = 2.471 \pm 0.005$ — для трехмерного.

Таким образом, простейший вероятностный подход дает очень хорошее согласие с известными результатами $d_F^{(2)} = 1.71$, $d_F^{(3)} = 2.49$.

23.4 Жирные фракталы

Среди фрактальных множеств встречаются также такие, которые устроены сложнее, чем самоподобные структуры. Более того, сложный фрактал может иметь целочисленную размерность и при этом иметь структуру, подобную канторову множеству. Множества, фрактальная размерность которых равна натуральному числу и совпадает с обычной размерностью

объемлющего пространства, а сами они сложно геометрически устроены, называются *толстыми* или *жирными* фракталами.

Для характеристики жирных фракталов было предложено использовать вместо фрактальной размерности так называемую метаразмерность [518]. Эта величина вводится следующим образом. Допустим, что существует предел

$$\lim_{\epsilon \rightarrow 0} \frac{\ln N(\epsilon)}{-\ln \epsilon} = d_F. \quad (2.188)$$

Предположим также, что $\lim_{\epsilon \rightarrow 0} N(\epsilon)\epsilon^{d_F} = \nu_0$. *Метаразмерностью* называется величина

$$d_\delta = \lim_{\epsilon \rightarrow 0} \frac{\ln |N(\epsilon)\epsilon^{d_F} - \nu_0|}{\ln \epsilon}.$$

Признак жирного фрактала — целая емкость и конечная метаразмерность. Из этого определения нетрудно получить, что

$$N(\epsilon) = \epsilon^{d_F} (\nu_0 \pm a\epsilon^\delta),$$

где a — некоторое положительное число. При $0 < \delta < 1$ отношение (2.188) под знаком предела, т. е. $\ln N(\epsilon)/\ln(1/\epsilon)$, может существенно отличаться от d_F на большом интервале значений ϵ .

23.5 Обобщенная размерность

Мы уже отмечали, что среди фрактальных множеств встречаются такие, которые не являются однородными. Можно построить фрактал, размерность которого, например, меняется при переходе от одной его части к другой. Более того, фракталы со строгим самоподобием вроде ковра Серпиньского или фрактала Вичека скорее являются исключением, чем правилом. Легко понять, что неоднородность не может быть описана только емкостью. Поэтому было предложено использовать целую совокупность размерностей D_q , $q = 0, 1, 2, \dots$, характеризующих *статистическую структуру* (т. е. степень неоднородности или некий аналог функции плотности) фрактального множества [350].

Величина D_q называется *обобщенной размерностью* и определяется следующим образом:

$$D_q = \lim_{\epsilon \rightarrow 0} \frac{1}{q-1} \frac{\ln \sum_{i=1}^{N(\epsilon)} p_i^q}{\ln \epsilon}, \quad q = 0, 1, 2, \dots, \quad (2.189)$$

где p_i — вероятность того, что точка, принадлежащая фракталу, содержится в i -й ячейке размером ϵ , $i = 1, 2, \dots, N(\epsilon)$.

Если фрактальное множество полностью однородно, то вероятность p_i одинакова для каждой ячейки, т. е. $p_i = 1/N(\epsilon)$. Поэтому

$$\sum_{i=1}^{N(\epsilon)} p_i^q = N(\epsilon) \left[\frac{1}{N(\epsilon)} \right]^q = [N(\epsilon)]^{1-q}$$

и, таким образом, обобщенная размерность (2.189) совпадает с емкостью d_F (2.172). Однако когда фрактал неоднороден, значение D_q отличается от емкости. Как показано в работе [350], $D_q \geq D_{q'}$ для $q < q'$, и равенство имеет место только тогда, когда фрактальное множество является однородным.

Поясним смысл обобщенной размерности D_q для различных q . Если $q = 0$, то $p_i^q = 1$ и сумма в (2.189) есть просто $N(\epsilon)$. Следовательно, соотношение (2.189) сводится к (2.172). Поэтому $D_0 = d_F$.

Если $q = 1$, то числитель и знаменатель в соотношении (2.189) обращаются в нуль, поскольку $\ln \sum_{i=1}^{N(\epsilon)} p_i = 1$. Для раскрытия неопределенности положим $\kappa = q - 1$ и преобразуем (2.189) следующим образом:

$$D_1 = \lim_{\substack{\epsilon \rightarrow 0 \\ \kappa \rightarrow 0}} \frac{\ln \left[\sum_{i=1}^{N(\epsilon)} p_i e^{\kappa \ln p_i} \right]}{\kappa \ln \epsilon}.$$

Тогда, используя разложение экспоненты и логарифма, найдем:

$$D_1 = \lim_{\substack{\epsilon \rightarrow 0 \\ \kappa \rightarrow 0}} \frac{\ln \left(1 + \kappa \sum_{i=1}^{N(\epsilon)} p_i \ln p_i \right)}{\kappa \ln \epsilon} = \lim_{\epsilon \rightarrow 0} \frac{\sum_{i=1}^{N(\epsilon)} p_i \ln p_i}{\ln \epsilon}.$$

Поскольку соотношение $S(\epsilon) = -\sum_i p_i \ln p_i$ определяет энтропию, то величина D_1 показывает, как при уменьшении размеров ячеек (т. е. при $\epsilon \rightarrow 0$) возрастает количество информации о рассматриваемом фрактальном множестве. Поэтому D_1 называется *информационной размерностью*.

Если $q = 2$, то соотношение (2.189) дает:

$$D_2 = \lim_{\epsilon \rightarrow 0} \frac{\ln \sum_{i=1}^{N(\epsilon)} p_i^2}{\ln \epsilon}. \quad (2.190)$$

Выражение, стоящее здесь под знаком суммы, p_i^2 , — это вероятность того, что по крайней мере две точки, принадлежащие фракталу, находятся в i -й ячейке. Следовательно, значение D_2 определяется через бинарные корреляции в распределении вероятностей. Поэтому величину D_2 называют *корреляционной размерностью*. Она часто используется при анализе распределенных систем (см. §31).

Для $q = 3, 4, \dots$ размерности D_3, D_4, \dots связаны с корреляциями высших порядков и дают дополнительную информацию о геометрической структуре фрактальных множеств. Кроме того, существует и определена обобщенная размерность $D|_{q \rightarrow \infty}$ [350].

Таким образом, используя приведенные характеристики, можно подробно изучить фрактальное множество и определить его статистическую структуру.

24. Отображения и некоторые их свойства

Теория отображений — это самостоятельный раздел теории динамических систем, где изучаются объекты не с непрерывным, а с дискретным временем.

24.1 Простейшие свойства отображений

Пусть, как и прежде, динамическая система задается дифференциальными уравнениями вида

$$\dot{\mathbf{y}} = \mathbf{v}(\mathbf{y}), \quad \mathbf{y}(0) = \mathbf{y}_0. \quad (2.191)$$

Выберем в фазовом пространстве этой динамической системы секущую поверхность, которую фазовые кривые пересекают не касаясь. Будем отмечать точкой на этой секущей поверхности каждое пересечение фазовой траекторией, происходящее в определенном направлении (например, снизу вверх). Тогда на этой поверхности мы получим некоторый набор точек A, B, C, D, \dots , последовательно переходящих друг в друга. Поскольку

начальные условия \mathbf{y}_0 динамической системы (2.191) полностью и однозначно задают всю фазовую траекторию, положение каждой точки в этой последовательности однозначно определяется положением предыдущей. Иными словами, существует некоторая функция φ , связывающая между собой положение двух следующих одна за другой точек. Обозначим через \mathbf{x}_n координаты n -й точки пересечения фазовой траектории с секущей поверхностью. Тогда координаты последующих точек пересечения связаны между собой некоторым соотношением

$$\mathbf{x}_{n+1} = \varphi(\mathbf{x}_n). \quad (2.192)$$

Говорят, что это соотношение определяет *отображение*. Поочередно применяя отображение (2.192), получаем бесконечную последовательность точек

$$\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{x}_{n+1}, \dots, \quad (2.193)$$

однозначно определяемую координатой начальной точки \mathbf{x}_0 .

Использование отображения вместо дифференциальных уравнений при исследовании динамики конкретных систем оказывается весьма полезным как в силу их наглядности, так и в вычислительном отношении, поскольку при переходе к отображению размерность изучаемой системы уменьшается на единицу. При этом оказывается, что свойства рассматриваемой динамической системы во многом определяются свойствами порождаемого ею отображения. Например, периодическим решениям дифференциальных уравнений (или, что то же самое, предельным циклам) ставятся в соответствие неподвижные точки отображения.

Разумеется, отображения вида (2.192) могут быть определены и вне всякой связи с какими-либо конкретными системами дифференциальных уравнений. Рассмотрим простой пример из области экологии. Допустим, у нас есть популяция бабочек, размножающихся в определенное время года. Далее, примем, что x_n — численность этой популяции в n -м году. Тогда ее численность на следующий год однозначно определяется тем, сколько бабочек было в предыдущем году, т. е. величиной x_n . Следовательно, изменение численности популяции подчиняется отображению (2.192) с определенной функцией φ . В данном конкретном случае отображение является одномерным. Можно, однако, представить себе и другие ситуации, когда x_n является набором из нескольких независимых переменных.

Точка \mathbf{x}^* называется неподвижной точкой отображения φ , если $\mathbf{x}^* = \varphi(\mathbf{x}^*)$. В теории отображений неподвижные точки играют ту же роль, что и стационарные точки в дифференциальных уравнениях. Найдем условия

устойчивости неподвижной точки \mathbf{x}^* отображения (2.192). Пусть $\mathbf{x} = \mathbf{x}^* + \delta\mathbf{x}$, причем $\delta\mathbf{x}$ — малая по сравнению с \mathbf{x}^* величина. Тогда отображение (2.192) примет вид

$$\mathbf{x}^* + \delta\mathbf{x}_{n+1} = \varphi(\mathbf{x}^* + \delta\mathbf{x}_n). \quad (2.194)$$

Разлагая правую часть (2.194) в ряд в окрестности точки \mathbf{x}^* и отбрасывая члены порядка малости выше первого, находим

$$\mathbf{x}^* + \delta\mathbf{x}_{n+1} = \varphi(\mathbf{x}^*) + \left. \left(\frac{\partial\varphi}{\partial\mathbf{x}_n} \right) \right|_{\mathbf{x}^*} \delta\mathbf{x}_n. \quad (2.195)$$

Отсюда, поскольку точка \mathbf{x}^* является неподвижной, получим

$$\delta\mathbf{x}_{n+1} = \left. \left(\frac{\partial\varphi}{\partial\mathbf{x}_n} \right) \right|_{\mathbf{x}^*} \delta\mathbf{x}_n, \quad (2.196)$$

или, вводя обозначение $a_{ik} = \left. (\partial\varphi_i / \partial x_n^k) \right|_{\mathbf{x}^*}$, будем иметь покомпонентно:

$$\delta x_{n+1}^i = \sum_k a_{ik} \delta x_n^k. \quad (2.197)$$

Это есть линейное отображение. Собственные значения α_k матрицы A с элементами a_{ik} называются *мультипликаторами*. Неподвижная точка \mathbf{x}^* исходного отображения (2.192) является *устойчивой* (см. раздел 20.2.1), если все мультипликаторы α_k удовлетворяют условию $|\alpha_k| < 1$. Если среди мультипликаторов α_k имеются такие, для которых $|\alpha_j| > 1$, то неподвижная точка отображения (2.192) будет *неустойчивой*.

Для двумерных отображений ($k = 2$) имеются только три возможных случая: $|\alpha_1| > 1, |\alpha_2| > 1$ — полностью неустойчивая неподвижная точка; $|\alpha_1| < 1, |\alpha_2| < 1$ — устойчивая неподвижная точка; $|\alpha_1| > 1, |\alpha_2| < 1$ (или $|\alpha_1| < 1, |\alpha_2| > 1$) — седловая неподвижная точка. В последнем случае соответствующая фазовая траектория дифференциального уравнения — *седловой предельный цикл* (рис. 2.107).

В качестве примера двумерного отображения рассмотрим отображение Эно [203]:

$$\begin{aligned} x_{n+1} &= \varphi_1(x_n, y_n) = y_n + 1 - ax_n^2, \\ y_{n+1} &= \varphi_2(x_n) = bx_n. \end{aligned} \quad (2.198)$$

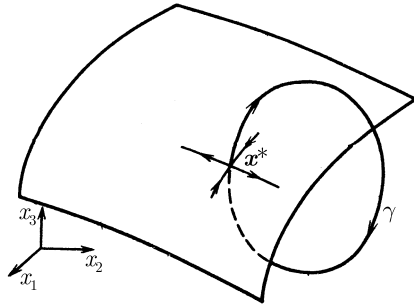


Рис. 2.107. Седловая точка \mathbf{x}^* отображения и соответствующий этой точке седловой предельный цикл γ

Неподвижные точки такого отображения получаем из уравнений

$$x = y + 1 - ax^2, \quad y = bx. \quad (2.199)$$

Отсюда находим, что при $a > a_0 = -(1 - b)^2/4$ ($a \neq 0$) координаты неподвижных точек

$$\begin{aligned} x_{\pm} &= \frac{1}{2a} (-(1 - b) \pm \sqrt{(1 - b)^2 + 4a}), \\ y_{\pm} &= bx_{\pm}. \end{aligned} \quad (2.200)$$

Точка (x_+, y_+) устойчива при $a < a_1 = 3(1 - b)^2/4$ и неустойчива при $a > a_1$. Точка (x_-, y_-) всегда неустойчива.

Можно показать, что отображение Эно диссипативно, так что всегда существует некоторое притягивающее множество меры нуль (аттрактор), к которому при итерациях отображения сходятся все последовательности точек. Это множество может быть чрезвычайно сложным и иметь фрактальную структуру (см. рис. 2.93, 2.94).

24.2 Одномерные отображения

Ниже мы рассмотрим только одномерные отображения вида

$$x_{n+1} = \varphi(x_n). \quad (2.201)$$

Динамика одномерного отображения может быть наглядно представлена графически. Построим на плоскости (x_{n+1}, x_n) график функции $x_{n+1} = \varphi(x_n)$ и проведем прямую $x_{n+1} = x_n$. Тогда последовательные итерации отображения (2.201) можно представить с помощью геометрического построения, называемого *диаграммой Ламерея* (рис. 2.108). Используя диаграмму Ламерея, легко найти неподвижные точки отображения (2.201). Ими являются точки пересечения графика $x_{n+1} = \varphi(x_n)$ и биссектрисы $x_{n+1} = x_n$.

Условие устойчивости неподвижных точек x^* отображения (2.201) сводится к выполнению неравенства $|\varphi'(x^*)| < 1$. Если $|\varphi'(x^*)| > 1$, неподвижная точка x^* неустойчива.

На рис. 2.108 последовательность x_1, x_2, x_3, \dots сходится к устойчивой неподвижной точке x_1^* . Неподвижная точка x_2^* является неустойчивой и при малом возмущении дает начало расходящейся последовательности x'_1, x'_2, x'_3, \dots

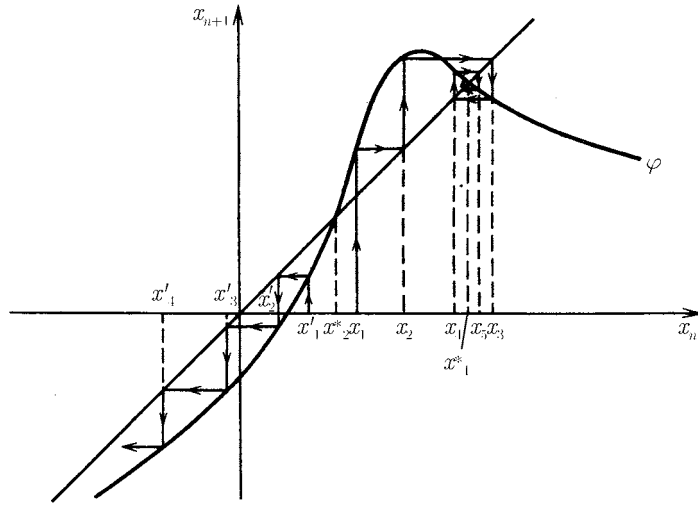


Рис. 2.108. Диаграмма Ламерея некоторого одномерного отображения φ . Неподвижная точка x_1^* устойчива, точка x_2^* неустойчива

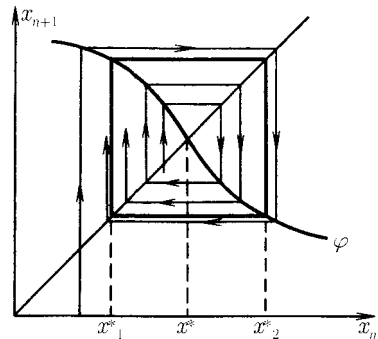


Рис. 2.109. Двукратный устойчивый цикл одномерного отображения

Кроме неподвижных точек одномерные отображения могут иметь циклы. Циклом периода t (или t -кратным циклом) отображения называется последовательность точек x_1, x_2, x_3, \dots , удовлетворяющих условиям

$$\begin{aligned} x_2 &= \varphi(x_1), \\ x_3 &= \varphi(x_2), \\ &\dots, \\ x_1 &= \varphi(x_m), \end{aligned} \tag{2.202}$$

причем никакие два элемента в наборе x_1, x_2, \dots, x_m не совпадают. Заметим также, что точки цикла x_1, x_2, \dots, x_m называются иногда t -кратными неподвижными точками.

На рис. 2.109 изображен двукратный устойчивый цикл одномерного отображения, соответствующий двукратным неподвижным точкам x_1^* и x_2^* . При этом точка x^* является неустойчивой.

Поскольку для t -кратных неподвижных точек

$$x_1 = \varphi^t(x_1) = \underbrace{\varphi(\varphi(\dots\varphi(x_1)\dots))}_t, \tag{2.203}$$

то каждая из t точек, соответствующих t -кратному циклу, одновременно является однократной неподвижной точкой отображения φ^t . Условие устойчивости и неустойчивости t -кратной неподвижной точки одномерного отображения сводится, соответственно, к выполнению неравенств

$$\begin{aligned} |\varphi'(x_1^*) \dots \varphi'(x_t^*)| &< 1, \\ |\varphi'(x_1^*) \dots \varphi'(x_t^*)| &> 1. \end{aligned} \tag{2.204}$$

Одномерные отображения бывают взаимно-однозначными и не взаимно-однозначными. Остановимся сначала на первых.

24.2.1 Взаимно-однозначные отображения

Точечное отображение (2.201) называется *взаимно-однозначным*, если ни при каких различных x и y не может иметь место равенство $\varphi(x) = \varphi(y)$. Следовательно, при монотонном изменении x функция $\varphi(x)$ тоже меняется монотонно, и при всех x имеем $\varphi'(x) \geq 0$ или $\varphi'(x) \leq 0$.

Рассмотрим отображение, для которого $\varphi'(x) \geq 0$ для любого x [56]. В качестве примера возьмем функцию, представленную на рис. 2.110. Отображение с такой функцией $\varphi(x)$ имеет неподвижные точки x_1^*, x_2^*, x_3^* . Легко видеть, что x_2^* — неустойчивая, а x_1^* и x_3^* — устойчивые точки, причем всякая начальная точка x_0 при последовательном применении отображения стремится либо к x_1^* , если $x_0 \in (-\infty, x_2^*)$, либо к x_3^* , если $x_0 \in (x_2^*, \infty)$. Следовательно, вся прямая x_n разбивается неустойчивой неподвижной точкой x_2^* на области притяжения устойчивых точек x_1^* и x_3^* (рис. 2.110).

Для взаимно-однозначного отображения общего вида, для которого $\varphi'(x) \geq 0$, также имеет место разбиение всей прямой x_n на области притяжения устойчивых неподвижных точек. Обозначим как

$$\dots < x_{k+1}^* < x_k^* < x_{k-1}^* < \dots < x_1^* < \dots \tag{2.205}$$

совокупность неподвижных точек отображения $\varphi(x)$. Легко видеть, что в цепочке (2.205) устойчивые и неустойчивые точки чередуются. Поэтому для

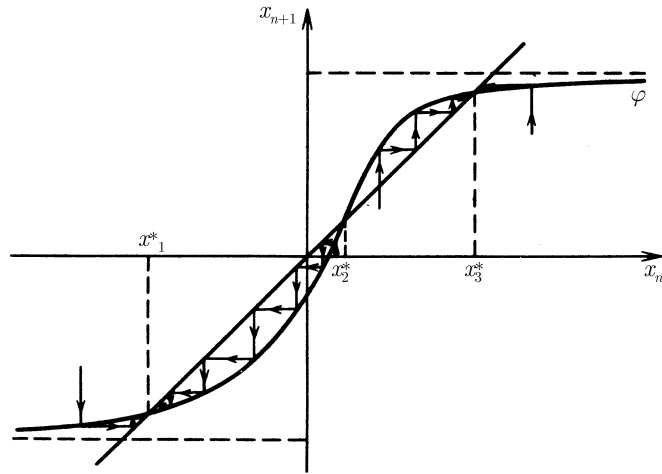


Рис. 2.110. Пример взаимно-однозначного отображения $\varphi(x)$, для которого $\varphi'(x) \geq 0$

каждой устойчивой точки x_i^* интервал (x_{i-1}^*, x_{i+1}^*) будет ее областью притяжения.

Отметим особый случай, когда график функции $\varphi(x)$ касается биссектрисы $x_{n+1} = x_n$. Такая ситуация соответствует *бифуркации* (см. раздел 28) отображения. В результате бифуркации происходит перестройка качественной картины движения. Так, случай, изображенный на рис. 2.111, отвечает слиянию и исчезновению устойчивой и неустойчивой неподвижных точек отображения, если переход осуществляется от кривой 1 к кривой 3. При переходе от 3 к 1, наоборот, происходит возникновение (рождение) двух новых неподвижных точек. Если рассматривать отображение (2.201) как некоторое отображение Пуанкаре для системы дифференциальных уравнений, то рассмотренному бифуркационному случаю соответствует рождение или исчезновение в фазовом пространстве системы пары предельных циклов — устойчивого и неустойчивого.

Теперь рассмотрим взаимно-однозначное отображение, для которого $\varphi'(x) \leq 0$ [56]. Здесь имеется единственная неподвижная точка x^* , которая в зависимости от вида функции φ может быть устойчивой или неустойчивой. Чтобы исследовать такое отображение, заметим, что функ-

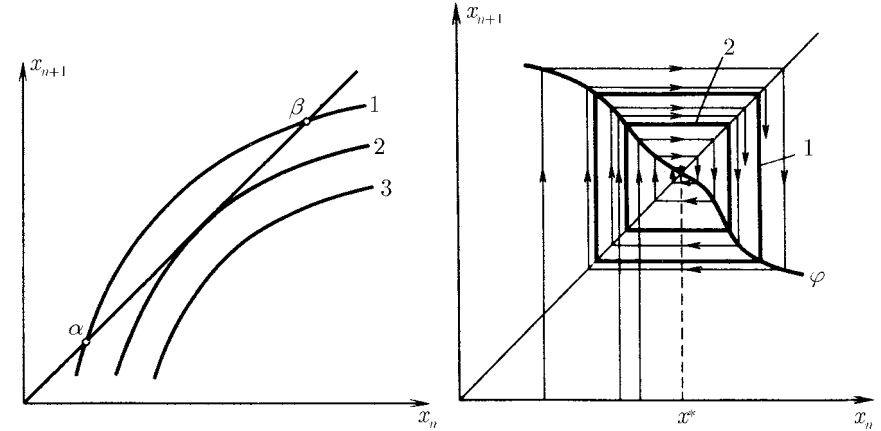


Рис. 2.111. Рождение (исчезновение) пар неподвижных точек: устойчивой β и неустойчивой α

ция $\varphi(\varphi(x)) = \varphi^2$ является отображением уже рассмотренного выше типа с $(\varphi^2)' \geq 0$. Действительно, отображение φ^2 можно записать в виде [56]

$$x_{n+2} = \varphi(\varphi(x_n)) = F(x_n), \tag{2.206}$$

отсюда

$$F'(x) = \left(\frac{d\varphi}{dx} \right) \Big|_{\varphi(x)} \left(\frac{d\varphi}{dx} \right) \Big|_x \geq 0. \tag{2.207}$$

Но, как известно, всякой неподвижной точке x_i^* , отличной от x^* отображения $F(x_n) = \varphi^2(x_n)$, соответствует двукратный цикл отображения φ , т. е. две двукратные неподвижные точки x_i^* и x_i^{**} , $x_i^{**} = \varphi(x_i^*)$. Следовательно, вся прямая x_n разбивается на область притяжения неподвижной точки x^* (если x^* является устойчивой) и область притяжения двукратных циклов — устойчивого и неустойчивого (рис. 2.112).

Для взаимно-однозначного отображения с $\varphi'(x) \leq 0$ возможны два типа бифуркаций: смена устойчивости неподвижной точки, при которой происходит рождение или исчезновение двукратного цикла (рис. 2.113a, б), и рождение (исчезновение) двукратных циклов (рис. 2.114a, б).

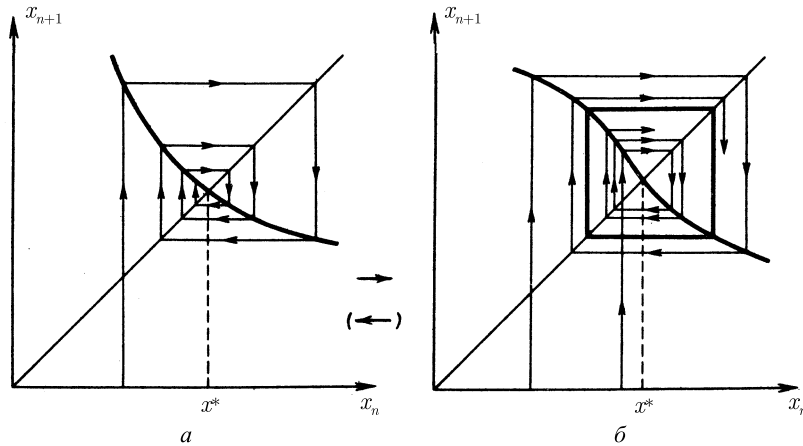


Рис. 2.113. Бифуркация потери устойчивости неподвижной точки x^* отображения (а) и рождение двукратного цикла (б) [56]

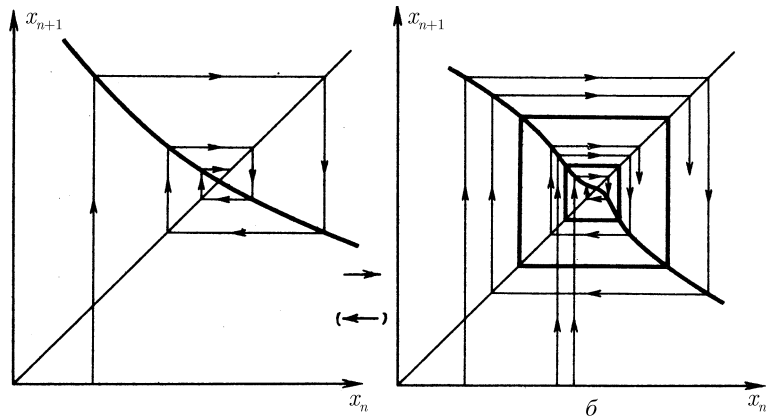


Рис. 2.114. Рождение (исчезновение) двукратных циклов отображения [56]

24.2.2 Не взаимно-однозначные отображения

Остановимся теперь на рассмотрении более сложных однозначных, но не взаимно-однозначных отображений (рис. 2.115). Характерной осо-

бенностью таких отображений является возможность появления очень сложной динамики — аperiodических последовательностей x_0, x_1, x_2, \dots (рис. 2.115), когда при некоторых начальных условиях отображение не обладает никакими устойчивыми неподвижными точками любой кратности. Это следует из того факта, что если мы попытаемся восстановить последовательность точек в обратном направлении, то, поскольку обратное отображение φ^{-1} неоднозначно (например, двузначно, как на рис. 2.115), всякий раз возникает вопрос, какую из ветвей отображения φ выбрать.

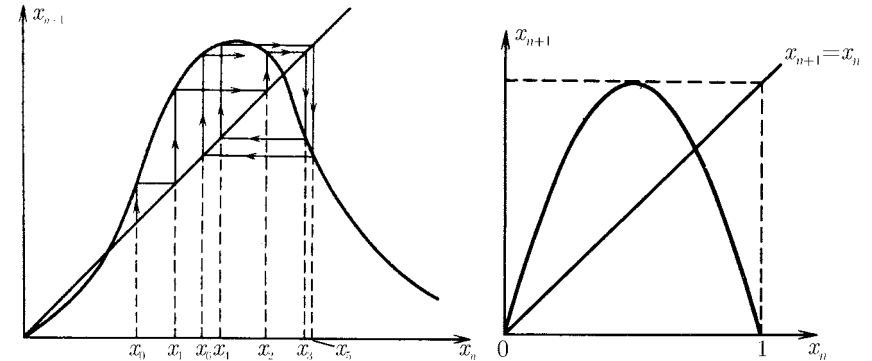


Рис. 2.115. Пример однозначного, но не взаимно-однозначного отображения

Рис. 2.116. Квадратичное отображение (2.209)

Приведем формальные рассуждения, показывающие, что немонотонные отображения могут иметь бесчисленное множество различных кратных неподвижных точек [56].

Предположим, что обратное отображение φ^{-1} имеет p ветвей, т.е. состоит из совокупности p взаимно-однозначных отображений φ_i^{-1} , $i = 1, 2, \dots, p$. Тогда φ_i^{-1} можно записать в виде $x_n = g_i(x_{n+1})$, $i = 1, 2, \dots, p$. Рассмотрим произведение m взаимно-однозначных отображений

$$\Phi = \varphi_{i_m}^{-1} \varphi_{i_{m-1}}^{-1} \dots \varphi_{i_2}^{-1} \varphi_{i_1}^{-1}, \tag{2.208}$$

где i_m — любые целые числа от 1 до p . Это отображение монотонно и имеет хотя бы одну неподвижную точку $x_{i_1 i_2 \dots i_m}^*$. Но эта неподвижная точка является также неподвижной точкой отображения φ^{-m} , а следовательно, и отображения φ^m . Неподвижной точке φ^m отвечает либо неподвижная точка отображения φ , либо цикл определенного периода. Первый случай

может иметь место только тогда, когда $i_1 = i_2 = \dots = i_m$. Поскольку можно взять любое m и произвольно выбрать числа i_1, i_2, \dots, i_m от 1 до p (неподвижные точки различных отображений (2.208) обязательно разные), значит, немонотонное отображение будет иметь бесчисленное множество различных кратных неподвижных точек.

Рассмотрим простой пример [56, 197]. Пусть

$$x_{n+1} = \varphi(x_n, \mu) = \mu x_n(1 - x_n), \quad (2.209)$$

где $\mu = 4$. Это отображение представляет собой частный случай так называемого *квадратичного отображения* (см. ниже) и называется *логистическим отображением*. Его график представлен на рис. 2.116. Логистическое отображение преобразует отрезок $[0, 1]$ в себя и, как легко убедиться, имеет бесконечное множество различных циклов, т. е. различных кратных неподвижных точек, причем все эти точки неустойчивы. В самом деле, из (2.209) имеем $dx_{n+1}/dx_n = -8(x-1/2)$, откуда $dx_{n+1}/dx_n \leq 1$ при $|x-1/2| \leq 1/8$ и $|dx_{n+1}/dx_n| > 1$ при $|x-1/2| > 1/8$; m -кратный цикл отображения (2.209) устойчив или неустойчив, если (см. (2.204)) произведение

$$|\varphi'(x_1^*)| |\varphi'(x_2^*)| \dots |\varphi'(x_m^*)| = F'_m(x_1^*), \quad (2.210)$$

соответственно, меньше или больше единицы. Нетрудно показать [56], что соотношение (2.210) всегда больше единицы. В силу этого, все неподвижные точки отображения (2.209) будут неустойчивы. Следовательно, любая последовательность точек

$$x_1, x_2 = \varphi(x_1), \quad x_3 = \varphi(x_2), \quad \dots \quad (2.211)$$

является неустойчивой. Таким образом, поведение последовательностей, порождаемых квадратичным отображением (2.209), будет чрезвычайно сложным.

25. Хаос в одномерных отображениях

Таким образом, из предыдущего материала можно сделать вывод, что чрезвычайно сложное, нерегулярное поведение одномерных отображений, которые для удобства мы запишем еще раз как

$$x_{n+1} = \varphi(x_n, \mu), \quad (2.212)$$

состоит в свойстве приобретать при определенных значениях параметра μ неустойчивость траекторий. В этом случае последовательности, порождаемые преобразованием (2.212) не стремятся с ростом n к неподвижным точкам или циклам.

25.1 Растягивающие отображения

Чтобы лучше представить такую нерегулярную динамику, рассмотрим так называемое преобразование типа палатки (tent map), или *треугольное отображение* (рис. 2.117):

$$x_{n+1} = \Delta(x_n) = \begin{cases} 2x_n, & 0 \leq x \leq \frac{1}{2}, \\ 2(1-x_n), & \frac{1}{2} < x \leq 1. \end{cases} \quad (2.213)$$

Это отображение топологически эквивалентно логистическому отображению (2.209) при $\mu = 4$. Иными словами, каждой последовательности, порождаемой отображением (2.213), отвечает последовательность, порождаемая отображением (2.209).

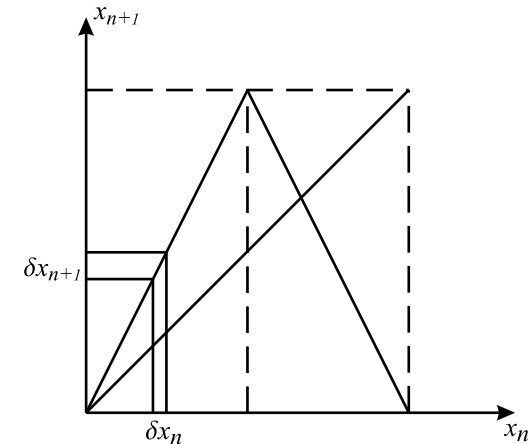


Рис. 2.117. Симметричное треугольное отображение (2.213), $\delta x_n > \delta x_{n+1}$

Два отображения f и g называются *топологически эквивалентными (сопряженными)*, если существует такое обратимое преобразование $h(x)$, что $f = h^{-1} \cdot g \cdot h$. Для логистического отображения (2.209) функция $h(x) = (2/\pi) \arcsin \sqrt{x}$ и $h^{-1}(x) = \sin \pi x/2$. Поскольку $0 \leq h^{-1}(x) \leq 1/2$ на интервале $0 \leq x \leq 1/2$, а $1/2 < h^{-1}(x) \leq 1$, если $1/2 < x \leq 1$, то в первом случае $h^{-1} \cdot \Delta \cdot h = \sin^2(2 \arcsin \sqrt{x}) = (2\sqrt{x}\sqrt{1-x})^2 = \varphi(x)$. Во втором случае $h^{-1} \cdot \Delta \cdot h = \sin^2[2(\pi/2 - \arcsin \sqrt{x})] = \sin^2(2 \arcsin \sqrt{x}) = \varphi(x)$.

Таким образом, для исследования свойств отображения (2.209) доста-

точно рассмотреть более простое преобразование (2.213). Легко видеть, что оно удовлетворяет условию

$$\left| \frac{d\Delta}{dx} \right| > 1, \quad (2.214)$$

т. е. растягивает интервалы (рис. 2.117). Следовательно, близкие точки под действием отображения расходятся. Отображения, удовлетворяющие условию (2.214) во всем диапазоне изменения переменной x , называются *растягивающими*.

Чтобы убедиться в том, что динамика растягивающего отображения хаотична, вычислим его показатель Ляпунова. Мы уже знаем, что показатель Ляпунова определяет среднюю скорость разбегания близких траекторий (см. §21). Поэтому для некоторого N найдем: $|\varphi^N(x_0 + \epsilon) - \varphi^N(x_0)| \simeq \epsilon e^{N\lambda(x_0)}$. Таким образом,

$$\lambda(x_0) = \lim_{N \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \frac{1}{N} \ln \left| \frac{\varphi^N(x_0 + \epsilon) - \varphi^N(x_0)}{\epsilon} \right| = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \left| \frac{d\varphi^N(x_0)}{dx_0} \right|.$$

Поскольку

$$\left. \frac{d\varphi^2(x)}{dx} \right|_{x_0} = \left. \frac{d}{dx} \varphi(\varphi(x)) \right|_{x_0} = \varphi'(\varphi(x_0))\varphi'(x_0) = \varphi'(x_1)\varphi'(x_0),$$

$$x_1 = \varphi(x_0).$$

Поэтому

$$\lambda(x_0) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \left| \frac{d\varphi^N(x_0)}{dx_0} \right| = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \left| \prod_{i=0}^{N-1} \varphi'(x_i) \right|.$$

Следовательно, показатель Ляпунова можно определить как

$$\lambda = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i \ln |\varphi'(x_i)|. \quad (2.215)$$

Для простоты предположим, что растягивающее отображение задается в виде

$$\delta x_{n+1} = \{\gamma \delta x_n\}, \quad 0 < x < 1, \quad (2.216)$$

где $\delta x = x - x'$, $\gamma > 1$ — параметр, а выражение в скобках $\{\cdot\}$ означает дробную часть аргумента. Тогда можно записать, что $\delta x_{n+1} = \gamma \delta x_n = \gamma^2 \delta x_{n-1} = \gamma^3 \delta x_{n-2} = \dots = \gamma^n \delta x_1 = e^{n \ln \gamma} \delta x_1$. Отсюда, используя

(2.215) и (2.216), находим: $\lambda = \ln \gamma$, что указывает на хаотичность преобразования (2.216).

В общем случае, если значения переменной x ограничены некоторым конечным интервалом, то динамика растягивающих отображений оказывается во многом аналогичной поведению систем с перемешиванием.

Отметим один интересный факт: для логистического отображения (2.209) можно найти точное решение итерационного процесса:

$$x_n = \sin^2(2^n y_0), \quad (2.217)$$

где $y_0 \in [0, 1]$. Действительно, из (2.209) получаем:

$$\begin{aligned} x_{n+1} &= 4 \sin^2(2^n y_0) [1 - \sin^2(2^n y_0)] = \\ &= 4 \sin^2(2^n y_0) \cos^2(2^n y_0) = \sin^2(2^{n+1} y_0). \end{aligned}$$

Значит, простое аналитическое выражение (2.217) может порождать хаотические последовательности чисел!

25.2 Порядок Шарковского

Существуют строгие результаты, используя которые можно определить, в каких случаях одномерные отображения будут иметь сложную нерегулярную динамику. Один широко известных таких результатов — теорема Шарковского о сосуществовании циклов [196, 197, 426], смысл которой заключается в следующем.

Установим сначала определенные соотношения между различными циклами одного и того же отображения. Будем говорить, что между целыми числами m и k существует *отношение порядка* $m \triangleright k$, если из существования цикла p_m периода m следует, что у этого же отображения есть и цикл p_k периода k . Как установлено в работе [196], циклы отображения можно упорядочить следующим образом:

$$\begin{aligned} 1 \triangleleft 2^1 \triangleleft 2^2 \triangleleft 2^3 \triangleleft \dots \triangleleft 2^n \triangleleft \dots \triangleleft 2^m \cdot 7 \triangleleft 2^m \cdot 5 \triangleleft 2^m \cdot 3 \triangleleft \dots \\ \dots \triangleleft 2^3 \cdot 7 \triangleleft 2^3 \cdot 5 \triangleleft 2^3 \cdot 3 \triangleleft \dots \triangleleft 2^2 \cdot 7 \triangleleft 2^2 \cdot 5 \triangleleft 2^2 \cdot 3 \triangleleft \dots \\ \dots \triangleleft 2 \cdot 7 \triangleleft 2 \cdot 5 \triangleleft 2 \cdot 3 \triangleleft \dots \triangleleft 7 \triangleleft 5 \triangleleft 3. \end{aligned} \quad (2.218)$$

Это означает, что если непрерывное отображение имеет цикл периода m , то оно имеет также и циклы периода $k \triangleleft m$. Отметим, что обратное утверждение неверно: для всякого m существует отображение, имеющее цикл периода m и не имеющее циклов периода k при $m \triangleleft k$.

Первое отношение в ряду (2.218) означает, что если существует цикл p_2 периода два, то значит есть и цикл p_1 (неподвижная точка). Это очень легко проверить. В самом деле, из существования цикла p_2 вытекает, что найдутся такие значения x_1 и x_2 , что $x_1 = \varphi(x_2)$, $x_2 = \varphi(x_1)$, $x_2 > x_1$. Рассмотрим функцию $g(x) = x - \varphi(x)$ в точках x_1 и x_2 : $g(x_1) = x_1 - \varphi(x_1) = x_1 - x_2 < 0$, $g(x_2) = x_2 - \varphi(x_2) = x_2 - x_1 > 0$. Но так как функции $g(x)$ и $\varphi(x)$ непрерывны, то существует точка x^* , в которой $g(x^*) = 0$, а значит, $\varphi(x^*) = x^*$.

Из отношения порядка (2.218) вытекает, что цикл периода три является самым сложным. В работе [392] независимо было показано, что если отображение имеет цикл периода три, то это отображение имеет циклы всех (в том числе и сколь угодно больших) периодов и несчетное множество непериодических траекторий. По этой причине данная работа была названа «Период три рождает хаос».

Таким образом, если известно, что непрерывное отображение имеет p_m -цикл, то оно имеет также и такие циклы, периоды которых соответствуют всем числам, стоящим в соотношении (2.218) слева от m . Теорема А. Н. Шарковского о существовании циклов была одним из первых результатов, указывающих на существование сложного, непериодического поведения одномерных отображений.

25.3 Унимодальные отображения

Другой результат позволяет в явном виде найти условия, при выполнении которых в одномерных отображениях возникает хаос. Он известен как теорема Огнева – Мисюревича [161, 180, 426, 430].

Введем несколько дополнительных определений. Допустим, что функция $\varphi(x)$ имеет единственный максимум или минимум x_c . Тогда, очевидно, $\varphi'(x_c) = 0$. Такие точки называются *критическими* точками отображения (2.212). Если в дополнение этому функция φ : (а) непрерывна; (б) монотонна слева и справа от критической точки, т. е. монотонно возрастающая (убывающая) слева от x_c и монотонно убывающая (возрастающая) справа от x_c , когда x_c — максимум (минимум), то отображение называется *унимодальным*.

Кроме унимодальности, важным понятием является так называемая *производная Шварца*, или *шварциан*:

$$S\varphi \equiv \frac{\varphi'''}{\varphi'} - \frac{3}{2} \left(\frac{\varphi''}{\varphi'} \right)^2. \quad (2.219)$$

Основной результат, касающийся хаотического поведения унимодальных отображений с отрицательным шварцианом (2.219) сводится к следующе-

му [180, 197]. Если критическая точка за конечное число шагов попадает в неустойчивый цикл или неустойчивую неподвижную точку, то отображение хаотично, т. е. порождает непериодические, чрезвычайно сложные последовательности. Кроме того, если унимодальное отображение с хаотическим поведением имеет отрицательную производную Шварца, то оно не может иметь устойчивых циклов. Можно наложить и более слабые условия на траекторию критической точки x_c . Например, в некоторых случаях достаточно потребовать, чтобы она проходила в некоторой окрестности x_c [430]. Отображения, удовлетворяющие этому свойству, называют *отображениями Мисюревича*.

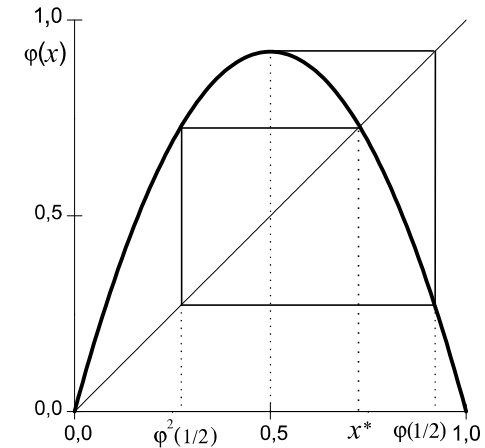


Рис. 2.118. Логистическое отображение при $\mu = 3, 678 \dots$

Рассмотрим в качестве примера логистическое отображение (2.209). Оно унимодально и на всем интервале $[0, 1]$ имеет отрицательную производную Шварца, поскольку $\varphi''' = 0$. При $\mu = 3, 678 \dots$ критическая точка $x_c = 1/2$ этого отображения за три шага попадает в неустойчивую неподвижную точку $x^* = 1 - 1/\mu$ (см. рис. 2.118). Значит последовательности, порождаемые преобразованием (2.209) при данном значении параметра, будут иметь хаотические свойства. В этом случае интервал $I = [\varphi^2(1/2), \varphi(1/2)] \simeq [0, 27; 0, 92]$ является аттрактором отображения. Иными словами, для любого $I' \subset (0; 1)$ найдется такое m , что $\varphi^m I' \subset I$.

На данном интервале I логистическое отображение с $\mu = 3, 678 \dots$ обладает всеми теми свойствами, которые оно имеет при $\mu = 4$ на отрезке $[0; 1]$, т. е. его поведение во многом аналогично динамике системы с пе-

решиванием. Это нетрудно понять, если заметить, что на I отображение (2.209) при $\mu = 3, 678 \dots$ сопряжено с кусочно-линейным преобразованием (рис. 2.119):

$$x_{n+1} = g(x) = \begin{cases} \frac{2}{3}(1+x), & x \leq \frac{1}{2}, \\ 2(1-x), & x > \frac{1}{2}. \end{cases} \quad (2.220)$$

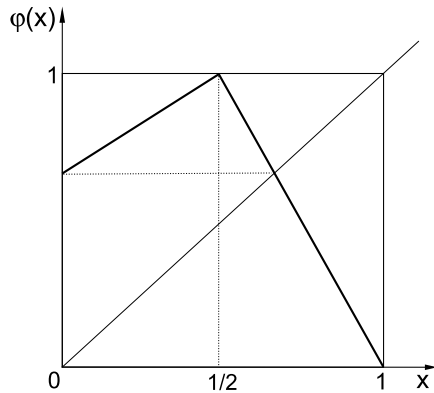


Рис. 2.119. Отображение (2.220)

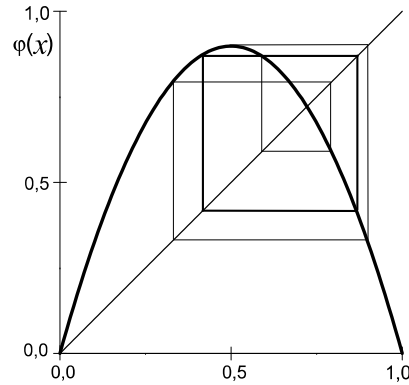


Рис. 2.120. Логистическое отображение при $\mu = 3, 593 \dots$

Легко найти другие значения параметра μ , при которых орбита критической точки логистического отображения переходит в неподвижную точку или неустойчивый цикл. Например, при $\mu = 3, 593 \dots$ траектория $x_c = 1/2$ за четыре итерации попадает на неустойчивый цикл периода два (рис. 2.120). Этот цикл образуют точки $x_2 = \varphi(x_1)$, $x_1 = \varphi(x_2)$, $x_{1,2} = (\mu + 1 \pm \sqrt{\mu^2 - 2\mu - 3}) / 2\mu$, которые определяются из соотношения $\varphi^2(x) = x$.

Описанное свойство унимодальных отображений позволяет легко строить преобразования, обладающие хаотической динамикой и не имеющие устойчивых точек и циклов. Для этого достаточно проследить, чтобы одна из итераций критической точки попала в неустойчивую неподвижную точку или цикл.

Вернемся к общей ситуации. Если вблизи единственного экстремума x_c производная $\varphi' = 0$, но $\varphi'' \neq 0$, то отображение будет локально квадратич-

ным. Его разложение в ряд Тейлора около x_c приводит к *квадратичному отображению* общего вида:

$$x_{n+1} = ax_n^2 + bx_n + c. \quad (2.221)$$

Если совершить линейное преобразование, то это отображение можно записать как

$$x_{n+1} = 2x_n^2 + 2d \cdot x_n.$$

Теперь, производя замену $x = -(\mu/2)y$, $d = \mu/2$, получим логистическое отображение $y_{n+1} = \mu y_n(1 - y_n)$, частные случаи которого мы уже рассмотрели.

Глубокий анализ логистического отображения показал, что при изменении параметра $\mu \in (0, 4]$ оно демонстрирует ряд интересных свойств [187, 302, 303, 337, 426]. Более того, эти свойства не связаны только со своеобразием логистического отображения. Оказывается, при некоторых ограничениях переход к хаосу, присущий логистическому отображению, встречается во всех одномерных отображениях, где после соответствующего масштабирования функция $\varphi(x)$ в интервале $[0, 1]$ имеет единственный максимум. Это свойство получило название *явление универсальности Фейгенбаума*, которое мы подробно рассмотрим в следующем параграфе.

26. Универсальность Фейгенбаума

В этом разделе мы познакомимся с одним из самых замечательных явлений, встречающихся в нелинейной динамике, — бесконечной последовательности удвоений периода при переходе к хаосу. Это явление присуще достаточно широкому классу динамических систем и объясняется при помощи так называемой *ренормализации*. В основе идеи ренормализации лежит естественное представление, что ряд свойств динамических систем повторяется на разных масштабах (т. е. идея самоподобия). При этом поразительным аспектом ренормализации служат количественные предсказания момента рождения хаоса.

26.1 Последовательность удвоений периода

Рассмотрим логистическое отображение

$$x_{n+1} = \varphi(x_n, \mu) = \mu x_n(1 - x_n), \quad (2.222)$$

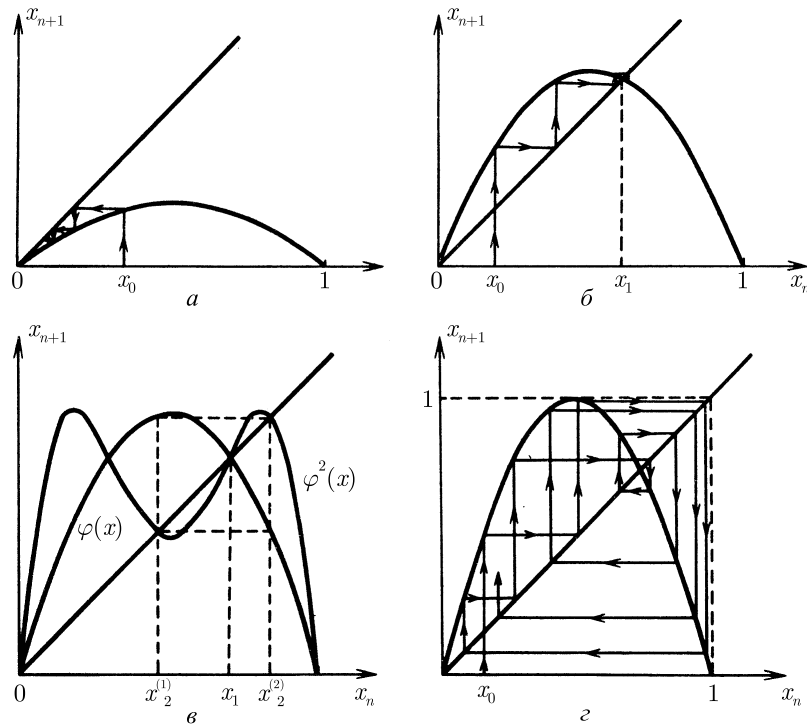


Рис. 2.121. Логистическое отображение (2.222) при значениях μ , равных: $0 < \mu < 1$ (а), $1 < \mu < 3$ (б), $3 < \mu \leq 1 + \sqrt{6}$ (в), $\mu = 4$ (г)

Будем постепенно увеличивать параметр μ в интервале от 0 до 4 и следить за изменениями динамики этого отображения. Тогда на отрезке $[0, 1]$, который отображение (2.222) преобразует в себя, получим следующую картину.

а) $0 < \mu \leq 1$. В этом случае квадратичное отображение имеет единственную неподвижную точку $x = 0$, которая является устойчивой (рис. 2.121а).

б) $1 < \mu \leq 3$. При $\mu > 1$ неподвижная точка $x = 0$ теряет устойчивость, поскольку теперь $\varphi'(0) > 1$, и на отрезке $[0, 1]$ появляется еще одна неподвижная точка $x_1 = 1 - 1/\mu$. Ее мультипликатор равен $\alpha(x_1) = \varphi'(x_1) = 2 - \mu$, и, следовательно, неподвижная точка x_1 будет устойчивой

(притягивающей) (рис. 2.121б). При $\mu = 3$ точка x_1 все еще остается притягивающей, хотя $|\alpha(x_1)| = 1$.

в) $3 < \mu \leq 1 + \sqrt{6}$. Когда $\mu > 3$, отображение (2.222) претерпевает новую бифуркацию: неподвижная точка x_1 становится неустойчивой, и вместо нее появляется устойчивый двукратный цикл (рис. 2.121в), который образуют две двукратные неподвижные точки

$$x_2^{(1),(2)} = \frac{\mu + 1 \pm \sqrt{\mu^2 - 2\mu - 3}}{2\mu}. \quad (2.223)$$

Их значения определяются из соотношения $\varphi^2(x_2) = x_2$, поскольку для такого отображения точки $x_2^{(1),(2)}$ являются неподвижными (рис. 2.121в).

г) $1 + \sqrt{6} < \mu < \mu_\infty = 3,5699456\dots$. При переходе параметра μ через значение $1 + \sqrt{6} \approx 3,45$ происходит следующая бифуркация: двукратный цикл $\{x_2^{(1)}, x_2^{(2)}\}$ теряет устойчивость, но при этом появляется притягивающий четырехкратный цикл. При $\mu > 3,54\dots$ этот цикл становится неустойчивым, и его сменяет устойчивый цикл периода 8, и т. д. Последовательные бифуркации удвоения периодов притягивающего цикла отображения (2.222) происходят до значения $\mu = \mu_\infty$, при котором притягивающий цикл достигает бесконечно большого периода, а циклы периодов 2^m , $m = 1, 2, \dots$, будут отталкивающими. Циклов других периодов квадратичное отображение (2.222) в этом случае не имеет.

д) При $\mu_\infty < \mu \leq 4$ отображение имеет циклы с любым периодом, в том числе и аperiodические траектории. Такие траектории при последовательных итерациях будут нерегулярным, хаотическим образом блуждать внутри единичного квадрата (рис. 2.121з).

Таким образом, при $\mu < \mu_\infty$ отображение (2.222) имеет единственный устойчивый цикл периода 2^m , $m = m(\mu)$, который, кроме множества меры нуль, притягивает все точки из отрезка $[0, 1]$. Когда $\mu > \mu_\infty$, динамика отображения (2.222) становится более сложной. В этом случае существуют аperiodические траектории, не притягивающиеся к циклам.

Описанную качественную перестройку динамики отображения (2.222) при изменении параметра μ удобно представить с помощью так называемой бифуркационной диаграммы (рис. 2.122): по вертикальной оси откладываются точки последовательностей, образующие устойчивый цикл периода 2^m , который появляется при $\mu = \mu_m$ из цикла периода 2^{m-1} , а по горизонтальной — значения μ .

Последовательность значений μ_m , при которых наблюдаются удвоения

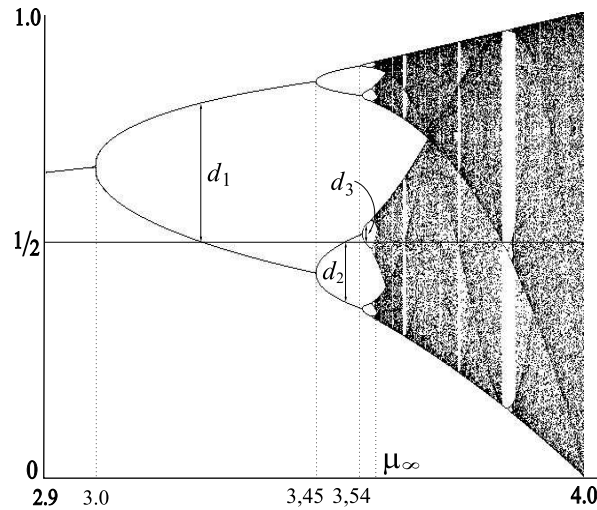


Рис. 2.122. Бифуркационная диаграмма логистического отображения (2.222)

периода, удовлетворяет простому закону

$$\lim_{m \rightarrow \infty} \frac{\mu_m - \mu_{m-1}}{\mu_{m+1} - \mu_m} = \delta = 4,66920160910299\dots \quad (2.224)$$

Введем расстояния $d_1, d_2, \dots, d_m, \dots$ между линией $x = 1/2$ и ближайшим к ней элементом цикла периода 2^m (рис. 2.122). Тогда оказывается, что отношение d_m/d_{m+1} при $m \rightarrow \infty$ имеет предел

$$\lim_{m \rightarrow \infty} \frac{d_m}{d_{m+1}} = -\alpha = 2,50290787509589\dots, \quad (2.225)$$

где знак минус означает изменение ориентации (отсчета).

Числа δ и α называются *универсальными постоянными Фейгенбаума*. Постоянная δ характеризует скорость сходимости последовательности μ_m , $m = 1, 2, 3, \dots$ к предельному значению $\mu_\infty = 3,5699456\dots$, а число α — изменение масштаба при последовательных бифуркациях. Иначе говоря, при увеличении в α раз очередная бифуркация вблизи $x_c = 1/2$ будет выглядеть точно так же, как и предыдущая.

Отметим, что последовательности удвоения периода наблюдаются только в таких дифференцируемых отображениях с единственным макси-

мумом, аппроксимируемым квадратичной функцией, которые имеют отрицательную производную Шварца (2.219). А поскольку для появления фейгенбаумовского перехода к хаосу достаточно, чтобы в сечении Пуанкаре система была близка к такому отображению, последовательности удвоения оказываются присущи самым разнообразным нелинейным системам. Более того, от вида конкретной системы не зависят и значения пределов (2.224), (2.225).

26.2 Элементы теории ренормализации

Объяснение явления универсальности Фейгенбаума дает теория ренормализации [187, 302, 303]. Остановимся на простейших деталях этой теории, следуя работам [187] и [357].

Отметим сначала важные свойства отображения (2.222). k -й элемент последовательности, порождаемый этим отображением, равен $x_k = \underbrace{\varphi(\varphi(\dots\varphi(x_0)))}_{k \text{ раз}} \equiv \varphi^k(x_0)$. Поэтому $\varphi^k(\varphi^m(x)) = \varphi^m(\varphi^k(x)) =$

$= \varphi^{k+m}(x)$. Построим графики логистической функции $\varphi(\mu, x)$ (2.222) и $\varphi^2(\mu, x)$ при таком $\mu = \mu_1$ (рис. 2.123а, б), чтобы критическая точка $x_c = 1/2$ была неподвижной точкой отображения φ^2 (рис. 2.123б) и одной из точек, формирующих устойчивый цикл периода 2 исходного отображения φ (рис. 2.123а). Такой цикл называется *суперустойчивым* (или *сверхустойчивым*); для него первое из условий (2.204) обращается в нуль.

Далее, повторим конструкцию (см. рис. 2.123 в, г) и представим графически функции $\varphi^2(\mu, x)$ и $\varphi^4(\mu, x)$ при таких значениях параметра $\mu = \mu_2$, чтобы неподвижная точка $x_c = 1/2$ для φ^4 была элементом цикла периода 2 отображения φ^2 . Этот цикл также будет суперустойчивым.

Теперь сравним график функции φ^2 , заключенный в квадрат, обозначенный штрихом на рис. 2.123в, с графиком всей функции φ (рис. 2.123а). Хорошо видно, что при соответствующем изменении масштаба и отражения от горизонтальной оси они практически совпадают (рис. 2.124). Если сравнивать графики функций при больших значениях параметра μ , но также соответствующих суперустойчивым циклам отображений, задаваемых функциями φ^{2^k} , то можно заметить, что после масштабного преобразования описанные кривые для φ^{2^m} и $\varphi^{2^{m+1}}$ сходятся к одному пределу. Это означает, что величина изменения масштаба от преобразования к преобразованию также сойдется к одному предельному значению. Следовательно, должна существовать определенная *универсальная функция* $g(x)$, которая

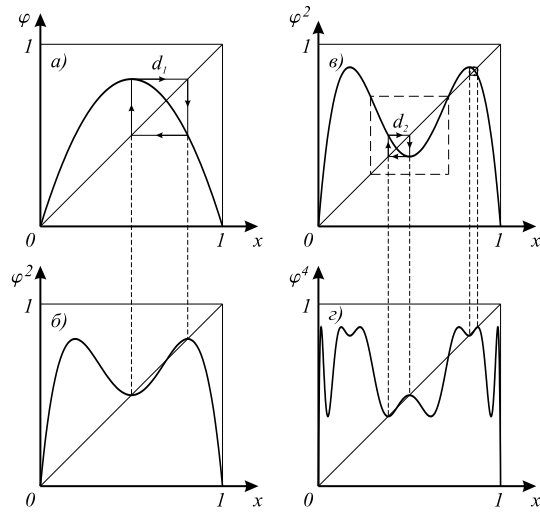


Рис. 2.123. Суперустойчивые циклы отображений $\varphi(\mu, x)|_{\mu=\mu_1}$ и $\varphi^2(\mu, x)|_{\mu=\mu_2}$ (а, в) и формирующие их неподвижные точки (б, г)

после одной итерации и масштабирования с коэффициентом α не изменяется. Найдем эту функцию.

Из условия существования сверхустойчивого цикла периода 2^m следует, что значения d_m (см. рис. 2.122) в точности равны расстоянию между его элементами:

$$d_m = \varphi^{2^{m-1}}\left(\mu_m, \frac{1}{2}\right) - \frac{1}{2}.$$

Для удобства сделаем преобразование координат $x = 1/2 \rightarrow x = 0$. Тогда

$$d_m = \varphi^{2^{m-1}}(\mu_m, 0).$$

Теперь, используя соотношение (2.225), найдем:

$$d_m = (-\alpha)^m d_{m+1}.$$

Следовательно, существует предел $\lim_{m \rightarrow \infty} (-\alpha)^m \varphi^{2^m}(\mu_{m+1}, 0)$. Далее, из рис. 2.124 следует, что этот предел — есть значение преобразованной в $(-\alpha)^m$ раз функции φ^{2^m} в точке $x = 0$. Значит можно записать, что

$$g_1(x) = \lim_{m \rightarrow \infty} (-\alpha)^m \varphi^{2^m}\left(\mu_{m+1}, \frac{x}{(-\alpha)^m}\right). \quad (2.226)$$

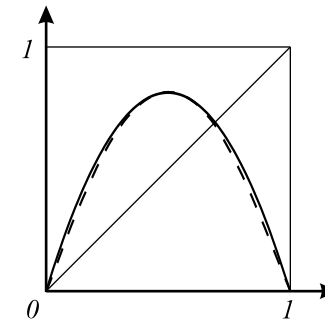


Рис. 2.124. Исходная функция $\varphi(\mu_1, x)$ (сплошная линия) и часть графика $\varphi^2(\mu_1, x)$, заключенная в квадрат на рис. 2.123в, после масштабирования (штриховая линия)

Так как с ростом итераций в данном соотношении фигурирует все меньшая и меньшая окрестность максимума x_c , то функция $g_1(x)$ будет одной и той же для любых φ , имеющих квадратичный максимум. Поэтому g_1 является универсальным пределом. На рис. 2.123а показана аппроксимация g_1 при $m = 0$. На рис. 2.123в (в квадрате, обозначенном штриховой линией) построено более точное приближение g_1 .

Если рассмотреть вторую итерацию, g_1^2 , функции (2.226), то, очевидно, она тоже будет универсальной. Аппроксимацию этой функции можно видеть на рис. 2.123г. После ее масштабирования получим:

$$g_0(x) \equiv (-\alpha)g_1^2\left(-\frac{x}{\alpha}\right).$$

Следовательно, из (2.226) имеем:

$$g_0(x) = \lim_{m \rightarrow \infty} (-\alpha)^m \varphi^{2^m}\left(\mu_m, \frac{x}{(-\alpha)^m}\right),$$

что определяет положение элементов цикла вблизи экстремума.

Таким образом, можно ввести семейство g_k универсальных функций:

$$g_k(x) = \lim_{m \rightarrow \infty} (-\alpha)^m \varphi^{2^m}\left(\mu_{m+k}, \frac{x}{(-\alpha)^m}\right). \quad (2.227)$$

Смысл этого соотношения состоит в том, что по каждой из этих функций g_k , $k = 0, 1, \dots$, можно определить положение 2^k элементов цикла на каждый

экстремум. Как следует из рис. 2.123, с увеличением итераций рост этих экстремумов замедляется. Поэтому должен существовать предел

$$g(x) = \lim_{k \rightarrow \infty} g_k(x). \quad (2.228)$$

Тогда, используя соотношение (2.227), найдем:

$$g(x) = \lim_{m \rightarrow \infty} (-\alpha)^m \varphi^{2^m} \left(\mu_\infty, \frac{x}{(-\alpha)^m} \right).$$

Далее, поскольку

$$\begin{aligned} g_{k-1}(x) &= \lim_{m \rightarrow \infty} (-\alpha)^m \varphi^{2^m} \left(\mu_{m+k-1}, \frac{x}{(-\alpha)^m} \right) = \\ &= \lim_{m \rightarrow \infty} (-\alpha)(-\alpha)^{m-1} \varphi^{2^{m-1+1}} \left(\mu_{m-1+k}, -\frac{1}{\alpha} \frac{x}{(-\alpha)^{m-1}} \right) = \\ &= \lim_{j \rightarrow \infty} (-\alpha)(-\alpha)^j \varphi^{2^j} \left(\mu_{j+k}, \frac{1}{(-\alpha)^j} (-\alpha)^j \varphi^{2^j} \left(\mu_{j+k}, -\frac{1}{\alpha} \frac{x}{(-\alpha)^j} \right) \right) = \\ &= -\alpha g_k \left(g_k \left(-\frac{x}{\alpha} \right) \right), \end{aligned}$$

то все g_k оказываются связаны через оператор \mathcal{T} :

$$g_{k-1}(x) = \mathcal{T}g_k(x) = -\alpha g_k \left(g_k \left(-\frac{x}{\alpha} \right) \right), \quad (2.229)$$

который называется *преобразованием удвоения*. Следовательно, из условия существования предела (2.228) вытекает, что функция g оказывается неподвижной точкой преобразования \mathcal{T} : $g(x) = \mathcal{T}g(x)$, т. е.

$$g(x) = -\alpha g \left(g \left(-\frac{x}{\alpha} \right) \right) \quad \text{или} \quad -\frac{1}{\alpha} g(-\alpha x) = g(g(x)). \quad (2.230)$$

Это соотношение иногда называют *уравнением Фейгенбаума–Цветановича*.

Нетрудно видеть, что при любом ν величина $\nu g(x/\nu)$ является решением уравнения (2.230) с тем же значением α . Следовательно, нужно зафиксировать масштаб. Это можно сделать, положив без потери общности $g(0) = 1$. Тогда из (2.230) получим $\alpha = -1/g(1)$.

Чтобы решение уравнения (2.230) было единственным, необходимо дополнительно потребовать гладкость универсальной функции g с вполне определенным максимумом (например допустить, что он квадратичный). При таких условиях функциональное уравнение (2.230) будет определять масштабный коэффициент α вместе с функцией g однозначно.

Для решения уравнения (2.230) воспользуемся полиномиальной аппроксимацией функции $g(x)$ и ограничимся только простейшим степенным разложением:

$$g(x) = 1 - ax^2.$$

Тогда, сохраняя только квадратичные члены, из (2.230) находим:

$$(1 - a) + 2a^2x^2 = -\frac{1}{\alpha} + a\alpha x^2.$$

Устанавливая теперь равенства при одинаковых степенях, получим следующую систему:

$$1 - a = -\frac{1}{\alpha}, \quad 2a^2 = a\alpha.$$

Отсюда $a \simeq 1,366$ и $\alpha \simeq 2,732$. Полученные значения лишь немного отличаются от численных результатов Фейгенбаума:

$$\begin{aligned} g(x) &= 1 - 1,52763x^2 + 0,104815x^4 - 0,0267057x^6 + \dots, \\ \alpha &= 2,50290787\dots \end{aligned}$$

Функциональное уравнение (2.230) можно использовать и для нахождения постоянной δ . Рассмотрим малое возмущение функции $g(x)$:

$$g^*(x) = g(x) + \varepsilon h(x).$$

Тогда

$$\begin{aligned} g^*(g^*(x)) &= g^*(g(x) + \varepsilon h(x)) = g(g(x) + \varepsilon h(x)) + \varepsilon h(g(x) + \varepsilon h(x)) \approx \\ &\approx g(g(x)) + \varepsilon [g'(g(x))h(x) + h(g(x))]. \end{aligned}$$

Следовательно, уравнение на собственные значения функции $h(x)$ будет следующим:

$$g'(g(x))h(x) + h(g(x)) = -\frac{\rho_k}{\alpha} h(\alpha x).$$

Как было показано в работах [269, 302, 303], для случая удвоения периода существует единственное собственное значение $\rho_k = \delta = 4,6692016\dots$

После бесконечной последовательности бифуркаций удвоения, т.е. при $\mu = \mu_\infty$ период цикла становится бесконечным. В этом случае в логистическом отображении появляется так называемый *аттрактор Фейгенбаума*, имеющий канторову структуру и размерность $d_F = 0,548\dots$. Логистическое отображение с аттрактором Фейгенбаума еще не является хаотическим. Оно имеет циклы периодов 2^m (и все они неустойчивые) и не имеет циклов других периодов.

При $\mu > \mu_\infty$ в логистическом отображении будут рождаться все новые и новые циклы, в том числе с периодами, отличными от 2^m . Более того, в этом случае оно при определенных значениях параметра μ проявляет хаотические свойства. Однако в интервале $\mu_\infty < \mu < 4$ хаотическое поведение перемежается с окнами периодичности, где динамика регулярна. Например, при $\mu \simeq 3,74$ в логистическом отображении наблюдается устойчивый цикл периода пять, а при $\mu = 1 + \sqrt{8} \simeq 3,83$ — цикл периода три (см. рис. 2.122). Следовательно, согласно теореме Шарковского, данное отображение имеет циклы любых периодов.

26.3 Универсальный спектр

При экспериментальных исследованиях и численном анализе очень часто используется спектральная плотность (см. §21). Это очень удобный критерий, позволяющий легко установить характер динамики изучаемой системы.

С рождением цикла основного периода в спектре присутствует только основная частота и ее гармоники. После удвоения периода появляется новая спектральная компонента и ее субгармоники. Затем, при дальнейших удвоениях периода, точно посередине интервала появившимися ранее компонентами возникают все новые и новые составляющие. С достижениями предельного значения μ_∞ в спектре будут наблюдаться гармоники всех уровней.

В работах [304, 305] были получены универсальные закономерности появления гармоник, отвечающих очередной бифуркации удвоения. Это позволило количественно сравнивать экспериментально получаемые спектральные плотности с предсказанными теорией. Опишем здесь основные элементы этого анализа, следуя [104, 304].

Для упрощения перепишем отображение (2.222) как

$$x_{n+1} = 1 - \mu x_n^2. \quad (2.231)$$

Нетрудно видеть, что обе формы записи сводятся друг к другу заменой пе-

ременной $x \rightarrow \frac{x-1/2}{\mu/4-1/2}$ и параметра $\mu \rightarrow \mu \left(\frac{\mu}{4} - \frac{1}{2} \right)$. Пусть x_n и x'_n — значения переменной, отвечающей циклам периодов 2^m и 2^{m+1} , соответственно.

Спектр цикла периода $2^m \equiv N$ можно выразить через Фурье-компоненты X_l :

$$x_n = \sum_{l=0}^{N-1} X_l \exp\left(\frac{2\pi i n l}{N}\right), \quad X_l = \frac{1}{N} \sum_{n=0}^{N-1} x_n \exp\left(-\frac{2\pi i n l}{N}\right).$$

При этом частоты m -й компоненты задаются как $f = l/N$, а амплитуды — соотношением $S(f) = |X_l|^2$ (см. §21).

В результате бифуркации появляется цикл удвоенного периода $2^{m+1} = 2N$. Представим его в виде суперпозиции четных и нечетных номеров динамической переменной. Тогда

$$\begin{aligned} X'_l &= \frac{1}{2N} \sum_{l=0}^{2N-1} x'_n \exp\left(-\frac{2\pi i n l}{2N}\right) = \\ &= \frac{1}{2N} \sum_{n=0}^{N-1} x'_{2n} \exp\left(-\frac{2\pi i n l}{N}\right) + \frac{1}{2N} \sum_{n=0}^{N-1} x'_{2n+1} \exp\left(-\frac{2\pi i n l}{N} - \frac{\pi i l}{N}\right). \end{aligned}$$

В силу масштабных свойств бифуркации удвоения можно записать $x'_{2n} \simeq x_n/\alpha$ и

$$x'_{2n+1} = 1 - \mu_\infty x'_{2n}{}^2 \simeq 1 - \frac{\mu_\infty x_n^2}{\alpha^2} = 1 - \frac{1 - x_{n+1}}{\alpha^2},$$

где предельное значение параметра μ для отображения (2.231) равно $\mu_\infty = 1,401155\dots$. Поэтому

$$\begin{aligned} X'_l &\simeq \frac{1}{2N} \sum_{n=0}^{N-1} \frac{x_n}{\alpha} \exp\left(-\frac{2\pi i n l}{N}\right) + \frac{1}{2N} \sum_{n=0}^{N-1} \left(1 - \frac{1}{\alpha^2}\right) \exp\left(-\frac{2\pi i n l}{N} - \frac{\pi i l}{N}\right) + \\ &+ \frac{1}{2N} \sum_{n=0}^{N-1} \frac{x_{n+1}}{\alpha^2} \exp\left(-\frac{2\pi i n l}{N} - \frac{\pi i l}{N}\right). \end{aligned}$$

Можно показать, что первое слагаемое — это просто $\frac{X_l}{2\alpha}$. Вторая сумма дает вклад только при $l = 0$ и $l = N$. Последнее, третье, слагаемое при $0 < l < N$ равно $\frac{X_l}{2\alpha^2} \exp\left(\frac{\pi il}{N}\right)$. Следовательно,

$$X'_l = X_l \frac{\alpha^{-1} + \alpha^{-2} \exp(\pi il/N)}{2}, \quad X'_{l+N} = X_l \frac{\alpha^{-1} - \alpha^{-2} \exp(\pi il/N)}{2}.$$

Пользуясь этими соотношениями, можно найти $S'(f/2)$ как функцию $S(f)$. В пределе бесконечно большого периода штрих можно опустить, что окончательно дает [104]:

$$S\left(\frac{f}{2}\right) = \left(\frac{1+\alpha^2}{4\alpha^4} + \frac{1}{2\alpha^3} \cos \pi f\right) S(f),$$

$$S\left(\frac{1}{2} + \frac{f}{2}\right) = \left(\frac{1+\alpha^2}{4\alpha^4} - \frac{1}{2\alpha^3} \cos \pi f\right) S(f).$$

Таким образом, интенсивность новых спектральных субгармоник, появляющихся в результате бифуркации удвоения, всегда в среднем уменьшается на фактор $(1+\alpha^2)/4\alpha^4$ и не зависит от номера бифуркации.

Если в системе имеется внешний шум, то для того, чтобы следующая бифуркация удвоения была *наблюдаема*, его дисперсия должна быть в 6,619... раз меньше, чем амплитуда субгармоник [278, 294, 306].

В заключение отметим несколько свойств бифуркаций удвоения.

Допустим, что в динамической системе произошла m -я бифуркация удвоения периода предельного цикла. Тогда вероятность появления $(m+1)$ -й бифуркации после возникновения m -й больше, чем вероятность появления m -й бифуркации после $(m-1)$ -й для любого $m > 1$.

Последовательности бифуркаций удвоения периода наблюдаются не только в диссипативных системах. Аналогичный сценарий существует и для гамильтоновых систем и отображений [267, 334], для которых при движении характерно сохранение фазового объема. В этом случае универсальные постоянные равны $\alpha = 4, 018077\dots$ и $\delta = 8, 72108\dots$

27. Отображения комплексной плоскости. Красота фракталов

Таким образом, одномерное квадратичное преобразование (2.221) и его частный случай — логистическое отображение, — играют важную роль в ре-

шении проблемы выявления причин рождения хаоса. И хотя такие системы демонстрируют весьма обширный спектр поведения, достаточно большое множество процессов ими не описывается. Более широкое представление о многообразии происходящих в нелинейных системах и отображениях явлений можно обнаружить, если от действительных переменных перейти в комплексную область и рассмотреть *отображения комплексной плоскости в себя*.

Такие отображения вследствие эстетической привлекательности рождающихся множеств в последнее время приобрели большую популярность. Более того, их анализ выявляет новые свойства фракталов и хаоса и дает наглядное представление о бесконечной геометрической сложности нелинейных систем (см., например, рис. 2.130–2.132 и цветные иллюстрации). При этом естественным образом проявляется еще одна особенность: границы областей притяжения аттракторов оказываются фрактальными.

Как мы видели в §22, в случае обыкновенных дифференциальных уравнений (или отображений) странные аттракторы обладают свойством самоподобия. Но такое свойство могут иметь также области притяжения *простых аттракторов* в пространстве начальных условий. Если аттрактор — стационарная (неподвижная) точка, то его областью притяжения будет область, имеющая гладкую границу. Однако для некоторых систем форма границы фрактальна. Наглядный пример такого явления как раз и демонстрируют отображения, заданные на комплексной плоскости.

Первые исследования в этом направлении были проведены Гастоном Жюлиа и Пьером Фату [301, 372]. Однако эти классические работы оставались незамеченными, поскольку в отсутствие компьютеров невозможно было передать всю красоту получаемых результатов. Положение коренным образом изменилось, когда в 70-х годах XX века была установлена связь отображений с фрактальными множествами (см. [147, 163, 287]).

В настоящем параграфе мы представим некоторые результаты, полученные в этой области, и дадим наглядное описание рождающихся множеств.

27.1 Множество Жюлиа

Рассмотрим множество всех комплексных чисел $z = x + iy$. Говорят, что последовательность $\{z_n\}_{n=1}^{\infty}$ таких чисел стремится к бесконечности, $\lim_{n \rightarrow \infty} z_n = \infty$, если для любого заданного R существует такое $N > 0$, что для всех $n > N$ выполняется $|z_n| > R$. Это означает, что, начиная с достаточно большого n , все точки z_n лежат вне круга радиуса R .

Введем на множестве комплексных чисел некоторое отображение. Тогда можно записать, что

$$z_{n+1} = f(z_n, c), \quad (2.232)$$

где f — определенная функция, а $c = a + ib$ символизирует совокупность параметров. Легко видеть, что если в этом преобразовании разделить действительную и мнимую части, то получится двумерное вещественное отображение. Однако благодаря особым свойствам комплексных чисел соотношение (2.232) приводит к чрезвычайно большому разнообразию рождающихся при итерациях множеств.

Если начать итерировать отображение (2.232), то, вне зависимости от конкретного вида функции f , значения z_n будут либо оставаться в ограниченной области, стремясь к аттрактору, либо уходить на бесконечность. Эти области будут разделены границей, которая называется *множеством Жюлиа*. Следовательно, множество Жюлиа можно определить как границу области притяжения бесконечно удаленной точки:

$$J_c(f) = \text{граница} \left\{ z : \lim_{n \rightarrow \infty} f^n(z) = \infty \right\}.$$

Рассматривая множество начальных условий z_0 , при которых для заданных значений параметра c итерации z_n остаются ограниченными, получим *заполненное множество Жюлиа* $J(f)$ для функции f . Визуально такие множества намного более привлекательны. Очевидно, границей заполненных множеств $J(f)$ и будет множество Жюлиа $J_c(f)$. Дополнение к множеству Жюлиа называется *множеством Фату*.

Выбирая функцию f и различные значения параметров, каждый раз будем получать новое множество Жюлиа. Как показал компьютерный анализ, множество Жюлиа отличается неожиданной эстетической привлекательностью и поразительным разнообразием, причем даже малейшие изменения параметров могут привести к существенным метаморфозам в его структуре.

Обычно множество Жюлиа фрактально, т. е. обладает геометрической инвариантностью, а траектория изображающей точки на этом множестве хаотична. Именно эти два свойства и создают приятные для восприятия образы.

Ограничимся далее рассмотрением таких функций f , которые представляют собой полином степени $k \geq 2$ с комплексными коэффициентами:

$$f(z) = c_0 + c_1 z + c_2 z^2 + \dots + c_k z^k. \quad (2.233)$$

Простейшее множество Жюлиа соответствует $f(z) = z^2$. Тогда, поскольку при $n \rightarrow \infty$ последовательность $f^n(z) \rightarrow \infty$ только когда $|z| > 1$,

то граничной областью, т. е. множеством Жюлиа, будет единичная окружность $\{z : |z| = 1\}$. Оно, конечно, не является фракталом, однако отображение, порождаемое функцией $f(z) = z^2$, хаотично на J_c [102].

Наиболее исследованным примером отображения (2.232) с функцией (2.233) является квадратичное преобразование

$$z_{n+1} = f_c(z_n) = z_n^2 + c, \quad (2.234)$$

которое, очевидно, сводится к логистическому (2.222), определенному на комплексной плоскости. Разделив вещественную и мнимую части, это отображение можно записать как

$$\begin{aligned} x_{n+1} &= x_n^2 - y_n^2 + a, \\ y_{n+1} &= 2x_n y_n + b. \end{aligned} \quad (2.235)$$

Для нашего анализа наиболее важными свойствами множества Жюлиа являются следующие [163, 154].

- 1) Траектория произвольной точки $p \in J_c$ либо периодическая (т. е. представляет собой цикл), либо хаотическая.
- 2) Множество J_c инвариантно относительно преобразования f .
- 3) Все неустойчивые циклы расположены в J_c .
- 4) Если $f(z)$ — рациональная функция, т. е. $f(z) = P(z)/Q(z)$, то при $\bar{z} \in J_c$ и $J^* = \{z \in J_c : |z - \bar{z}| < \varepsilon, \varepsilon > 0\}$ найдется целое n , такое, что $f^n(J^*) = J_c$.

Если в отображении (2.234) $c \neq 0$, то множество Жюлиа уже не гладкая кривая, как при $c = 0$, но оно будет фракталом. В этом случае изменения параметра c приводят к большому разнообразию и богатству его форм, а также к рождению необычных и неожиданно привлекательных картин (см. ниже). На рис. 2.125 показано множество Жюлиа, которое образовано границей области притяжения единственной устойчивой неподвижной точки отображения (2.234) при $a = -0,12375$ и $b = 0,56508$. Нетрудно видеть, что основная его особенность — это самоподобие: одна и та же структура встречается на различных масштабах. Как следует из свойства 4, такую границу можно целиком восстановить, если использовать ее сколь угодно малую часть и конечное число итераций функции $f_c(z)$.

Опишем изменения J_c в зависимости от значений параметра c в отображении (2.234). Сначала рассмотрим перестройки множества Жюлиа, если

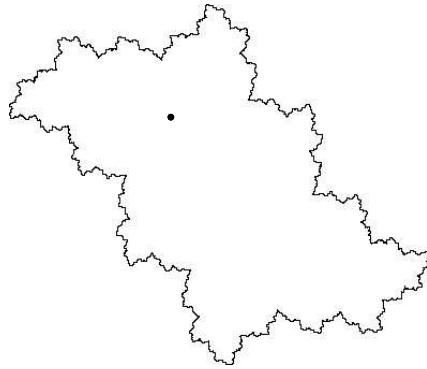


Рис. 2.125. Множество Жюлиа отображения (2.234) при $a = -0,12375$ и $b = 0,56508$

изменять c как действительный параметр, т. е. при $b = 0$. В этом случае на оси $y = 0$ соотношение (2.235) перейдет в обычное квадратичное преобразование

$$x_{n+1} = x_n^2 + a, \quad (2.236)$$

которое при $a = \mu/2 - \mu^2/4$ и $\bar{x}_n = 1/2 - x_n/\mu$ есть логистическое отображение (2.222). Как показано в разделе 26.1, для логистического отображения параметр $\mu = 3$ соответствует потере устойчивости неподвижной точки $\bar{x} = 1 - 1/\mu = 2/3$. Для преобразования (2.236) это отвечает значениям $a = -3/4$, $x = -1/2$. Следовательно, неподвижная точка $(-1/2, 0)$ исходного преобразования (2.235) или (2.234) находится на границе области устойчивости, т. е. принадлежит множеству Жюлиа. Проводя дальнейший анализ, легко прийти к выводу, что точки на оси $y = 0$ будут подчиняться правилу Фейгенбаума.

Это элементарное исследование показывает, что при $b = 0$ множество Жюлиа должно состоять из бесконечного количества островов, которые попарно касаются друг друга на оси x (рис. 2.126). При этом понятно, что множество Жюлиа будет включать неустойчивые неподвижные точки и циклы всех периодов. Кроме того, оно будет содержать и хаотические последовательности, не стремящиеся к какому-либо конечному периоду.

Другая предельная ситуация соответствует значению $c = i$. В этом случае множество Жюлиа не будет иметь «внутренности», поскольку оно не содержит аттракторов, отличных от бесконечно удаленной точки (рис. 2.127). Таким образом, здесь J_c представляет собой границу области притяжения

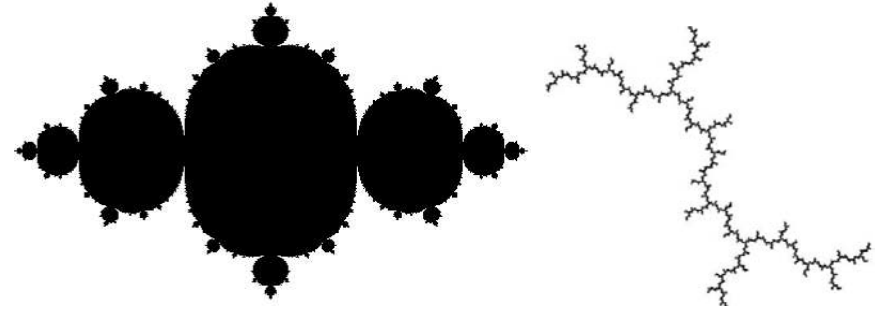


Рис. 2.126. Заполненное множество Жюлиа при $a = -3/4$ и $b = 0$ Рис. 2.127. Дендрит ($a = 0$, $b = 1$)

единственного аттрактора на бесконечности. Множество такой формы называется *дендрит*.

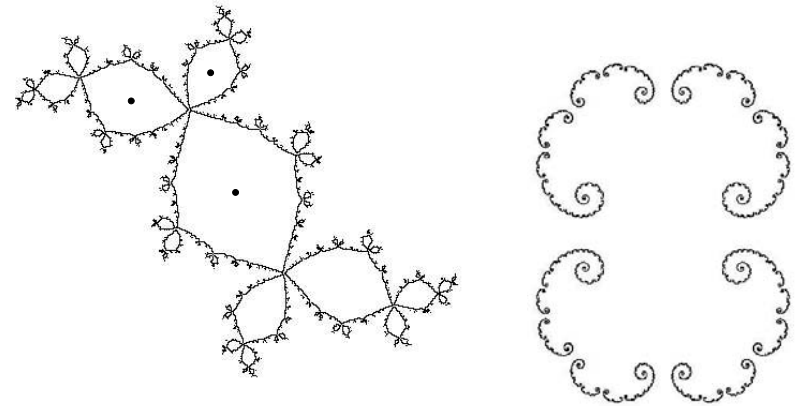


Рис. 2.128. Множество Жюлиа — совокупность замкнутых деформированных кривых ($a = -0,12$; $b = 0,74$) Рис. 2.129. Пыль Фату ($a = 0,3$; $b = 0$)

С множеством Жюлиа могут происходить и более сложные метаморфозы. Например, из фрактально деформированной окружности (рис. 2.125)

оно может превратиться либо в бесконечное число замкнутых деформированных кривых (рис. 2.128), либо, подобно множеству Кантора, распасться на совокупность различных компонент (рис. 2.129).

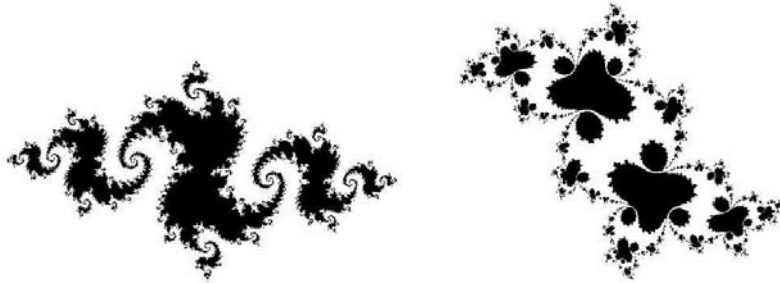


Рис. 2.130. Заполненное множество Жюлиа при $a = -0,8$; $b = 0,15$ (слева) и $a = -0,11$; $b = 0,65$ (справа)

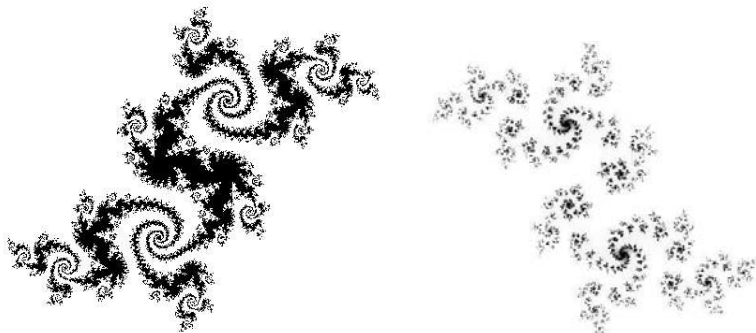


Рис. 2.131. Заполненное множество Жюлиа при $a = -0,068$; $b = -0,667$ и пыль Фату при $a = 0,019$; $b = 0,693$

Говорят, что множество является *связным*, если любые две его точки могут быть соединены дугой, принадлежащей этому множеству. Множество *вполне несвязно*, если все его компоненты — одиночные точки. Таким образом, в первом случае (рис. 2.128) множество Жюлиа остается связным. Во втором случае (рис. 2.129) оно словно рассыпается в пыль, и, следовательно, будет вполне несвязным. Такие множества, очевидно, являются

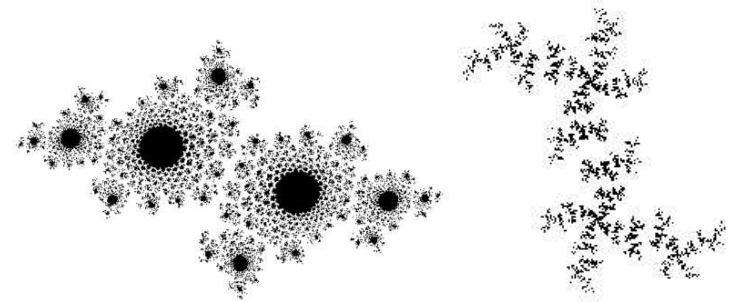


Рис. 2.132. Пыль Фату: $a = -0,68$; $b = 0,312$ (слева) и $a = -0,03$; $b = 0,875$ (справа)

множествами Кантора. Их обычно называют *пылью Фату*. Пыль Фату всегда фрактальна, а итерации на этом множестве хаотичны.

Некоторые наиболее яркие примеры заполненных множеств Жюлиа показаны на рис. 2.130–2.132.

Нетрудно построить цветные изображения множеств Жюлиа, используя понятие так называемого *времени убегания*. Пусть R — некоторое достаточно большое число. Временем убегания точки $z_0 \notin J_c$ называется первое значение k , для которого z_k , определяемое соотношением (2.234), по модулю превышает R .

Как показано в работе [290], полученные таким образом линии можно интерпретировать как эквипотенциальные, если соответствующее множество рассматривать как заряженный проводник. Когда изображающая точка находится во внутренней области множества Жюлиа, то цвет выбирается в соответствии с тем, сколько итераций потребуется, чтобы та или иная точка попала в область, содержащую аттрактор. На цветных иллюстрациях 2.141–2.144 показаны завораживающие результаты такой раскраски.

27.2 Множество Мандельброта

В общем случае, как уже отмечалось, для каждого c на множестве Жюлиа имеет место чрезвычайная чувствительность к изменениям начальных данных. Поэтому близкие параметры могут порождать множества Жюлиа абсолютно разных форм. Однако, несмотря на такое большое разнообразие, существует общее правило, которое позволяет установить форму множества Жюлиа при каждом значении параметра c и понять, когда оно является связным. Это правило приводит к множеству Мандельброта M [417]

(рис. 2.133). Множество \mathcal{M} удивительным образом появляется во многих задачах, в том числе не связанных с итерациями функции $z^2 + c$. Именно это обстоятельство делает его столь важным объектом исследования.

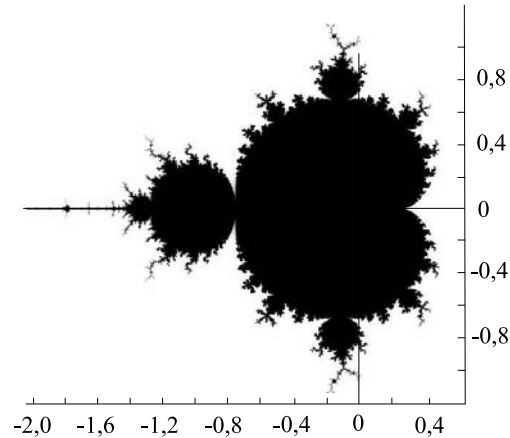


Рис. 2.133. Множество Мандельброта для преобразования $z^2 + c$

Множество Мандельброта \mathcal{M} — это совокупность таких значений параметра c , для которых соответствующее множество Жюлиа J_c связно. Из результатов Г. Жюлиа и П. Фату [301, 372] следует, что множество J_c связно только тогда, когда $f_c^n(0)$ при $n \rightarrow \infty$ остается конечной величиной. Таким образом, можно ввести эквивалентное определение множества Мандельброта:

$$\mathcal{M} = \left\{ c : \lim_{n \rightarrow \infty} f_c^n(0) \neq \infty \right\}. \quad (2.237)$$

Поясним подробнее смысл введения множества Мандельброта. Прежде всего, относительно \mathcal{M} , как и для последовательности Фейгенбаума, доказано свойство *структурной устойчивости*, или *универсальности* (см. [77]). Это означает, что множество Мандельброта встречается в достаточно общих семействах преобразований комплексной плоскости в себя. Более точно, если семейство алгебраических функций в некоторой области близко к преобразованию (2.234), то множество, порожаемое этим семейством, будет иметь ту же форму, что и \mathcal{M} . Следовательно, имея представление о структуре множества Мандельброта, можно предсказать тип поведения изучаемой системы и получить представление о соответствующих множествах Жюлиа.

Рассмотрим, как форма множества Жюлиа, отвечающая заданному параметру c , связана с множеством Мандельброта. На плоскости (a, b) значения c могут быть только двух типов: $c \in \mathcal{M}$ и $c \notin \mathcal{M}$. При этом соответствующие множества Жюлиа будут качественно различаться. В первом случае J_c будет связным (см., например, рис. 2.130), а во втором — распадется на несчетное множество точек (рис. 2.132). Более того, различным частям множества Мандельброта отвечают определенные закономерности в структуре множества J_c . Опишем основные из них.

Как следует из рис. 2.133, структура множества Мандельброта достаточно сложна. Прежде всего, бросается в глаза основная область, ограниченная определенной замкнутой кривой с острием в точке $(0, 25; 0)$ и вершиной в точке $(-0, 75; 0)$. Внутренняя часть этой области отвечает параметру c , при котором множество Жюлиа имеет одну притягивающую неподвижную точку. Ее устойчивость определяется соотношением $|df_c(z)/dz| < 1$. Следовательно, граница области устойчивости дается выражением $|2z| = 1$ или $z = e^{i\theta}/2$, $0 \leq \theta \leq 2\pi$. Теперь, используя уравнение существования неподвижной точки $z^2 + c = z$, находим:

$$c = \frac{1}{2}e^{i\theta} - \frac{1}{4}e^{2i\theta}.$$

Это соотношение описывает кривую, которая называется кардиоидой. Таким образом, основная область множества Мандельброта представляет собой заполненную кардиоиду. Если значение c выбрать там, то множество Жюлиа всегда будет фрактальной замкнутой кривой (за исключением случая $a = b = 0$), охватывающей область притяжения неподвижной точки (рис. 2.125).

Кардиоиды касается круг с центром в точке $(-1, 0)$ и радиусом $0, 25$. К этим двум основным элементам множества \mathcal{M} примыкает бесчисленное количество меньших круговых областей. Все они соответствуют устойчивым циклам определенного периода. Если c лежит в одной из таких областей, то множество Жюлиа будет образовано неограниченным числом фрактально деформированных окружностей, заключающих точки устойчивого цикла (рис. 2.128).

Области множества Мандельброта, отвечающие существованию устойчивых циклов различных периодов, показаны на рис. 2.134.

Если параметр c выбрать в месте касания одного из кругов и кардиоиды (т. е. на границе \mathcal{M}), то множество Жюлиа будет вплотную подходить к аттрактору (рис. 2.135). Такой случай называется параболическим. Он характерен тем, что при малом изменении параметра c множество J_c пре-

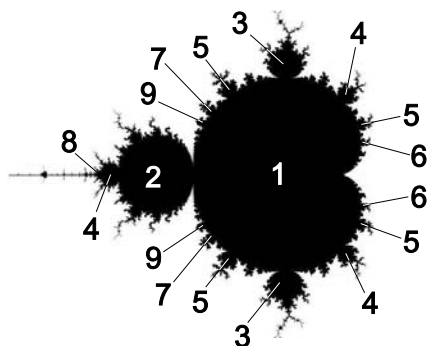


Рис. 2.134. Участки множества Мандельброта, отвечающие существованию устойчивых циклов

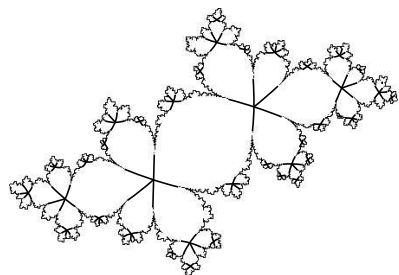


Рис. 2.135. Множество Жюлиа, соответствующее точке касания кардиоиды и одного из кругов множества Мандельброта ($a = -0,481762$; $b = -0,531657$)

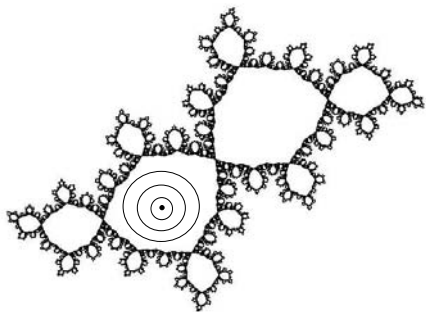


Рис. 2.136. Множество Жюлиа, содержащее диск Зигеля

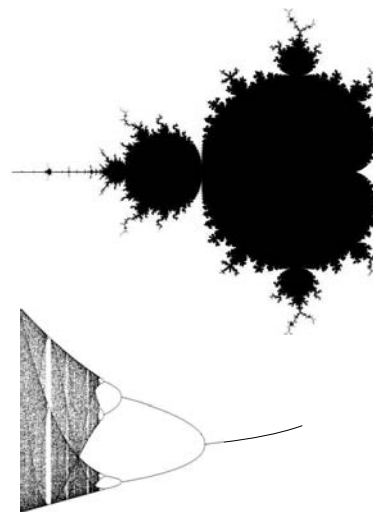


Рис. 2.137. Сравнение множества Мандельброта для преобразования $z^2 + c$ и бифуркационной диаграммы отображения $x^2 + a$

вратится либо во фрактальную замкнутую кривую или в неограниченную совокупность таких кривых, либо рассыплется в пыль Фату.

В других точках границы \mathcal{M} (при некоторых дополнительных условиях) множество Жюлиа включает в себя так называемый диск Зигеля [147, 163]. Это множество определяет поведение изображающей точки внутри множества Жюлиа: сначала она перескакивает из одной области в другую до тех пор, пока не попадет внутрь диска, содержащего неподвижную точку. После этого изображающая точка начинает вращаться вокруг нее (рис. 2.136).

Описанные четыре примера дают представление о всех типичных формах множества Жюлиа в системе (2.234), охватывающей область, содержащую неподвижные точки или цикл.

На оси a множество Мандельброта описывается отображением (2.236). Как мы знаем, это отображение заменой переменных сводится к логистическому (2.222). Для него, в соответствии с универсальностью Фейгенбаума, бифуркации удвоения периода происходят при $a = -0,75$; $-1,25$; $-1,3681\dots$; $-1,3940\dots$; $-1,3996\dots$; \dots ; $-1,4012\dots$. Здесь последнее значение соответствует точке накопления a_∞ . Именно эта последовательность соответствует точкам касания круговых областей (рис. 2.133).

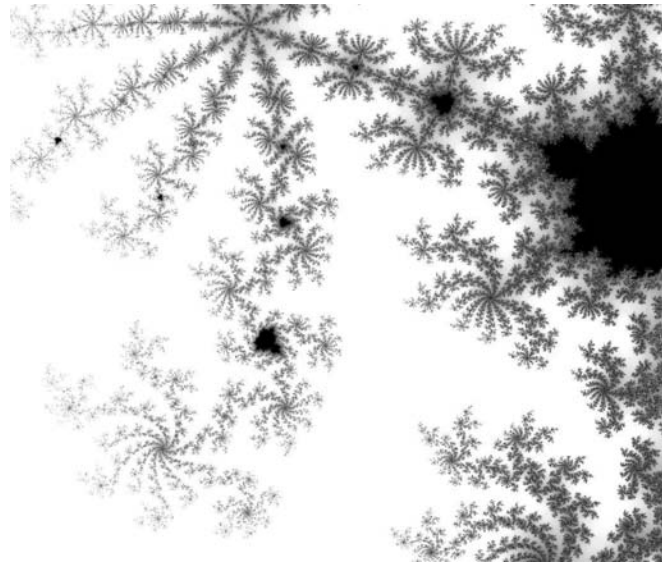


Рис. 2.138. Увеличенный фрагмент множества Мандельброта. Хорошо видны вторичные множества

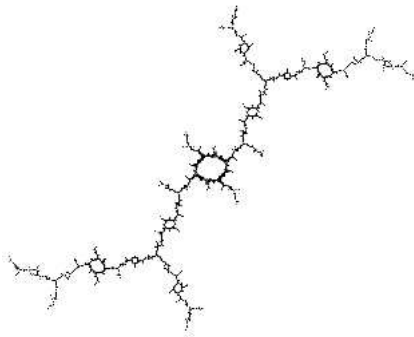


Рис. 2.139. Множество Жюлиа для значения параметра c из вторичного множества Мандельброта: $a = 0,158888889$; $b = 1,031818182$

Далее, в логистическом отображении есть характерное значение параметра $\mu \simeq 3,83$ (см. рис. 2.122), при котором оно обладает устойчивым циклом периода 3. Соответствующая величина для преобразования (2.236)

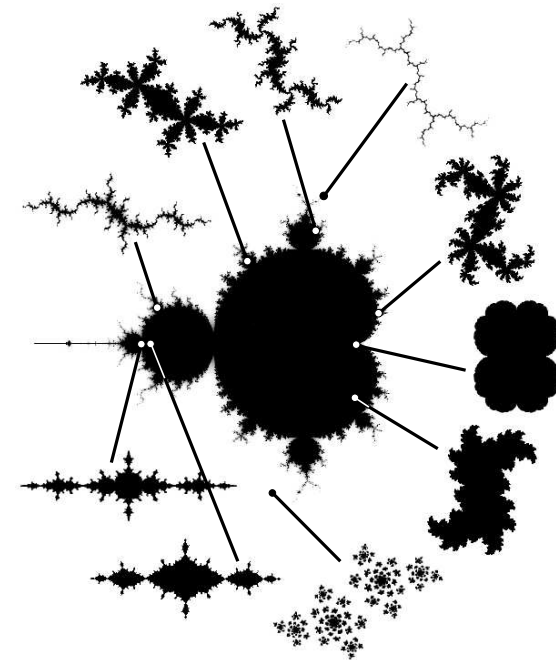


Рис. 2.140. Множество Мандельброта и множества Жюлиа, отвечающие заданному параметру $c \in \mathcal{M}$

— это $a = -1,75$. Для множества Мандельброта параметры $a = -1,75$; $b = 0$ отвечают существованию маленькой копии большой кардиоиды, центр которой находится в точке $-1,754877666\dots$. К этой маленькой кардиоиде примыкает бесконечное семейство кругов, так что вместе они составляют уменьшенную копию всего множества \mathcal{M} . При этом окно периодичности при $a = -1,75$ в отображении (2.236) отвечает сечению такой копии. На рис. 2.137 показана связь множества Мандельброта и бифуркаций удвоения периода отображения (2.236). Видно, насколько более полная картина наблюдается при выходе в комплексную плоскость.

Как впервые обнаружил Б.Мандельброт, число компонент множества \mathcal{M} , соответствующих малым копиям исходной кардиоиды, бесконечно велико. В окрестности каждой из них всегда находится множество примыкающих круговых областей. При этом все кардиоиды и сопровождающие их компании кругов связаны с большой кардиоидой нитями, на которые

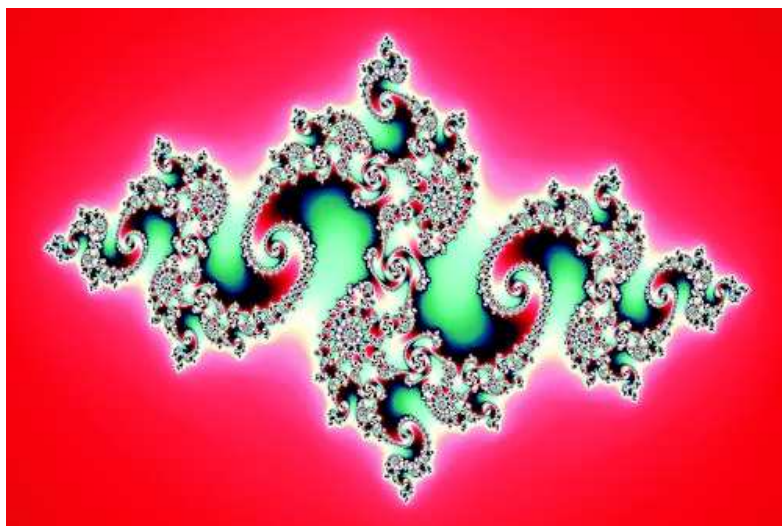


Рис. 2.141. Множество Жюлиа

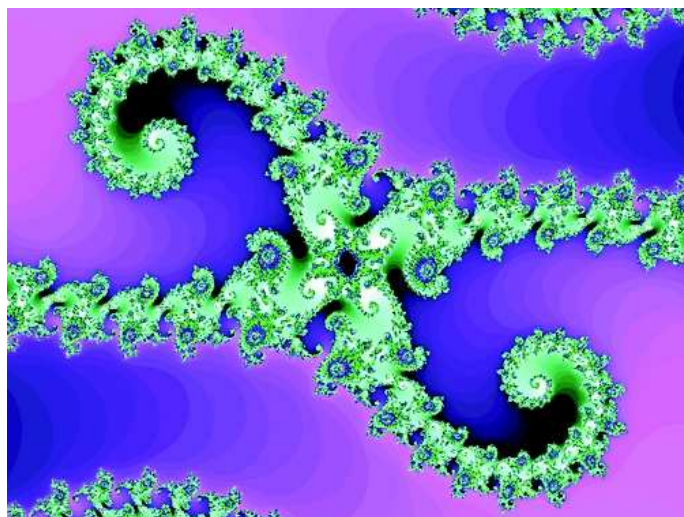


Рис. 2.142. Множество Жюлиа

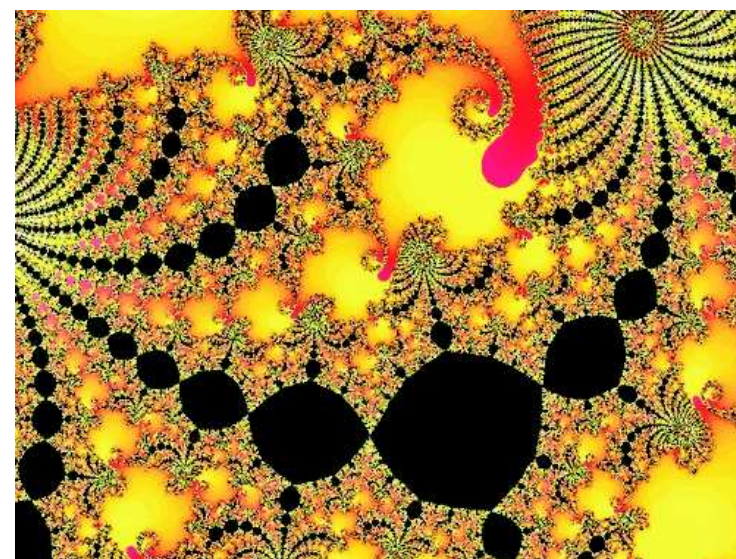


Рис. 2.143. Множество Жюлиа

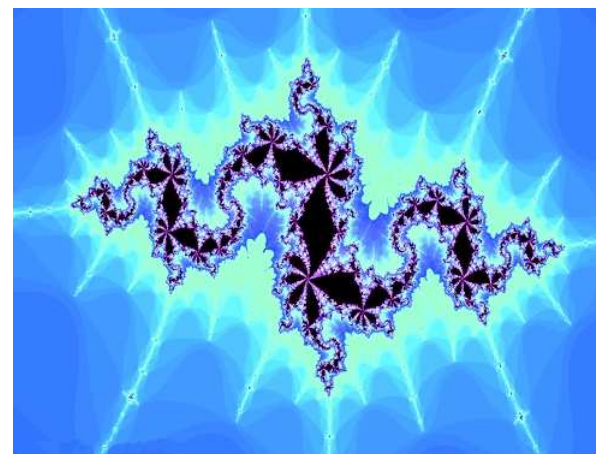


Рис. 2.144. Множество Жюлиа

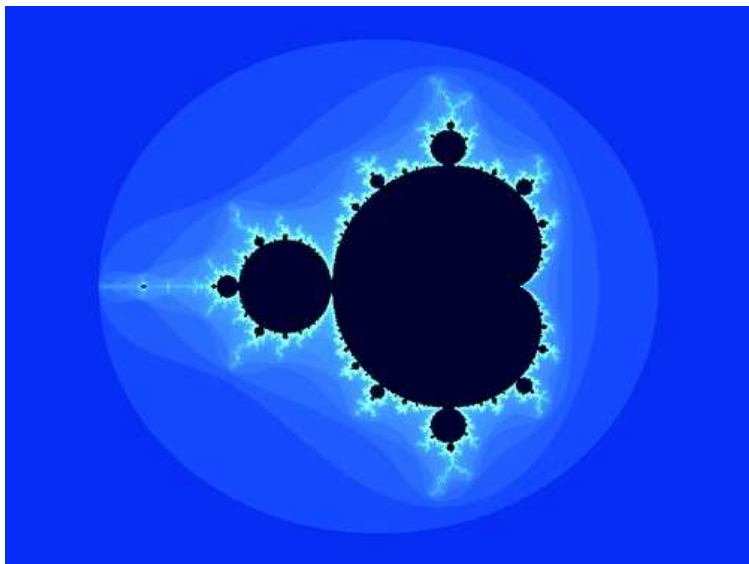


Рис. 2.145. Исходное множество Мандельброта

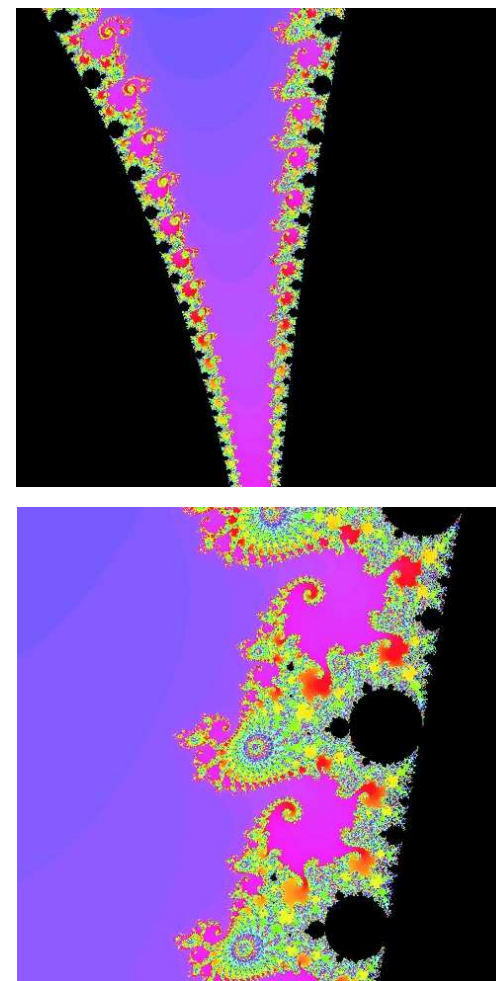


Рис. 2.146. Последовательное увеличение фрагмента множества Мандельброта

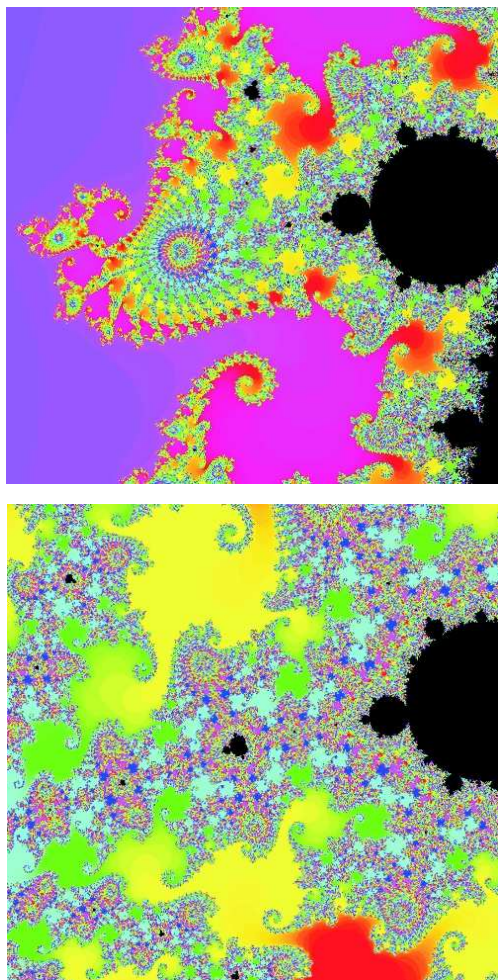


Рис. 2.146. Окончание

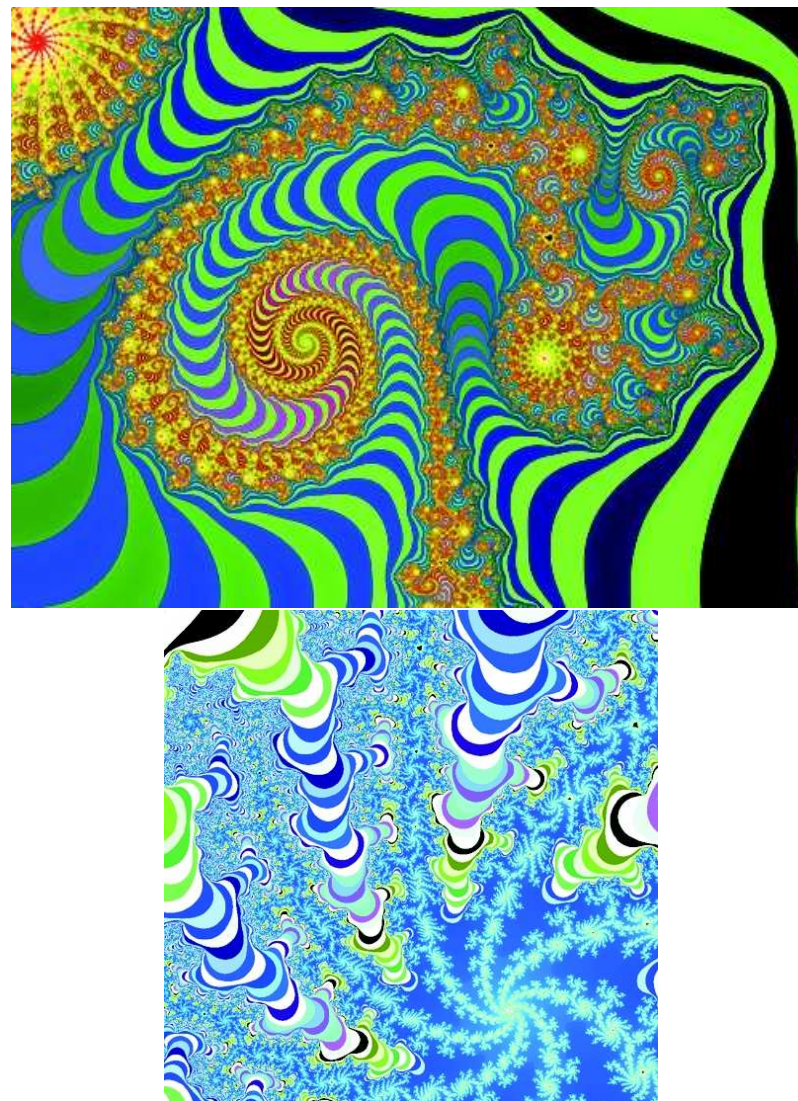


Рис. 2.147. Фрагменты множества Мандельброта

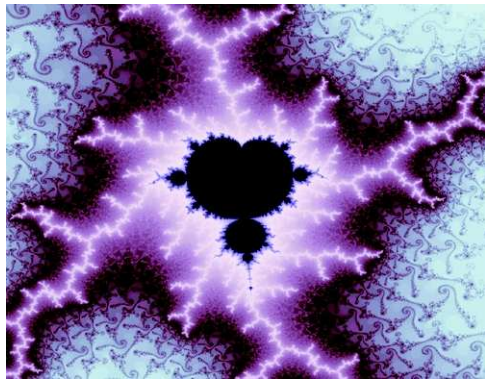


Рис. 2.148. Окрестность одного из вторичных множеств Мандельброта

словно нанизаны другие (еще меньшие) кардиоиды. Такие нити образуют великолепные узоры (рис. 2.138). Благодаря им множество Мандельброта связно.

Если взять параметр c в одной из таких копий, то получится множество, которое является комбинацией дендрита и множества Жюлиа, полученного выбором c из основной части M (рис. 2.139).

Общая картина соответствия множества Мандельброта множеству Жюлиа показана на рис. 2.140.

В отличие от множеств Жюлиа, множество Мандельброта в классическом смысле не обладает свойством самоподобия. Хотя при последовательном увеличении действительно наблюдаются малые копии исходного множества, это подобие сильно зависит от значений параметров a и b . Более того, выбрав малые копии множества M , нетрудно заметить, что отношение их размеров к расстоянию между ними определяется не только местом их расположения на комплексной плоскости, но и увеличением, с которым мы их рассматриваем.

Используя понятие времени убегания, как это было сделано при построении цветных иллюстраций J_c , множество Мандельброта M также можно представить в цвете. Для этого достаточно рассмотреть ту же последовательность z_k , но для начальной точки $z_0 = 0$. Таким образом, для M время убегания зависит только от параметра c . Некоторые эффектные иллюстрации множества Мандельброта представлены на рис. 2.145–2.148. Глядя на них, вспоминаются слова Д. Рюэля: «Это область исследования, в которой будут открыты новые гармонии» [472].

28. Бифуркации в динамических системах

Как было показано в предыдущих параграфах, динамика диссипативных систем может быть самой разнообразной. В зависимости от вида аттрактора, присутствующего в фазовом пространстве, в системе может реализоваться либо регулярный режим движения — стационарный, периодический или квазипериодический, что соответствует простым аттракторам (стационарной точке, предельному циклу и инвариантному тору), либо хаотический, которому отвечает странный аттрактор. Если динамическая система зависит от параметра (в качестве параметра может выступать, например, температура, число Рейнольдса и т. п.), то при его изменении в общем случае аттрактор и, соответственно, поведение системы меняются плавно. Например, при небольшом изменении параметра может немного сдвинуться устойчивая точка, изменить свою форму и период предельный цикл или деформироваться инвариантный тор. При переходе параметра через некоторое критическое значение аттрактор может претерпеть качественную перестройку, а динамика системы резко измениться. В частности, из предельного цикла может возникнуть инвариантный тор и периодическое движение сменится квазипериодическим.

Значения параметров, при которых происходит топологическая (или качественная) перестройка установившихся режимов движения в системе, называются *бифуркационными значениями*, а сама перестройка — *бифуркацией*. При непрерывном изменении параметров могут возникать последовательности бифуркаций.

28.1 Бифуркации и кризисы

При бифуркациях аттракторы динамических систем могут не только усложняться, но и упрощаться. Так, инвариантный тор при соответствующем изменении параметра может «схлопнуться» в предельный цикл. Однако при переходе параметра через бифуркационное значение может произойти не только усложнение или упрощение, но и вообще исчезновение аттрактора. В связи с этим бифуркации аттракторов делят на *внутренние бифуркации и кризисы* [27, 33].

Предположим, что динамическая система, задаваемая дифференциальными уравнениями, зависит от некоторого управляющего параметра μ :

$$\dot{x} = v(x, \mu). \quad (2.238)$$

Предположим также, что система (2.238) имеет стационарное решение x^0 . Естественно, что это решение, вообще говоря, зависит от управляющего па-

раметра, $x^0 = x^0(\mu)$. Допустим далее, что стационарная точка $x^0(\mu)$ системы (2.238) устойчива при $\mu < \mu_0$ и неустойчива при $\mu > \mu_0$. Следовательно, при $\mu = \mu_0$ реальная часть некоторых собственных значений матрицы линеаризации становится положительной, т. е. пересекает мнимую ось слева направо (рис. 2.149). Если при этом изменится топологическая структура разбиения фазового пространства системы на траектории, то точка $\mu = \mu_0$ будет точкой бифуркации потока $x(t)$ динамической системы (2.238).

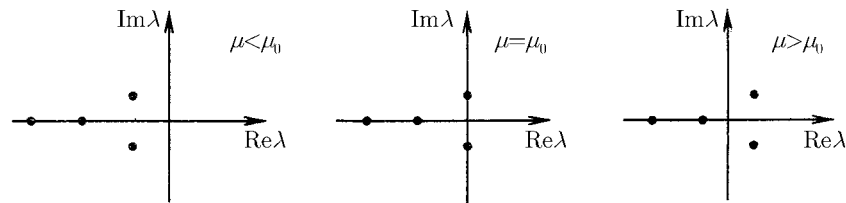


Рис. 2.149. Поведение корней λ при бифуркации Андронова – Хопфа: пара комплексно-сопряженных собственных значений переходит в правую полуплоскость

При значениях $\mu < \mu_0$ поток $x(t)$ в окрестности x^0 легко описать, поскольку при $t \rightarrow \infty$ все траектории системы из этой окрестности стремятся к x^0 . Однако при $\mu > \mu_0$ это уже не так, и поэтому при переходе через значение μ_0 характер потока может внезапно измениться. В силу того, что при $\mu > \mu_0$ стационарная точка x^0 является неустойчивой, нас будет интересно состояние системы при $\mu > \mu_0$. Мы рассмотрим только наиболее часто встречающиеся смены характера движений в системе при переходе через бифуркационное значение $\mu = \mu_0$.

Наиболее известным примером бифуркации стационарного состояния является *бифуркация Андронова – Хопфа*, когда стационарное решение x^0 динамической системы теряет устойчивость в результате того, что пара комплексно-сопряженных собственных значений $\lambda_{1,2} = \xi \pm i\omega$ матрицы линеаризации переходит в правую полуплоскость, так что $\xi > 0$ (рис. 2.149). При этом в системе возбуждаются периодические колебания с периодом $T \approx 2\pi/\omega$ и происходит бифуркация рождения из первоначально устойчивой стационарной точки предельного цикла с радиусом, растущим как $\sqrt{|\mu - \mu_0|}$. Если в этом случае родившийся цикл устойчив, то говорят о *мягкой потере устойчивости*, или мягком возникновении колебаний в системе (рис. 2.150), поскольку при малом отклонении параметра μ от би-

фуркационного значения устанавливающийся колебательный режим мало отличается от состояния равновесия.

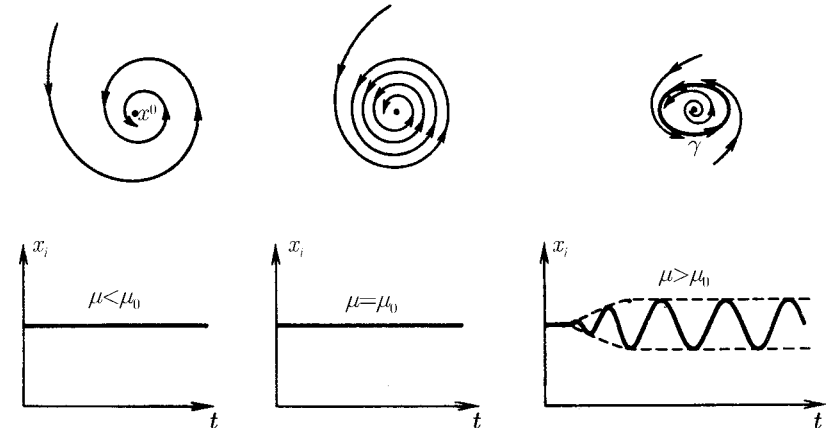


Рис. 2.150. Перестройка фазового портрета системы при мягкой потере устойчивости

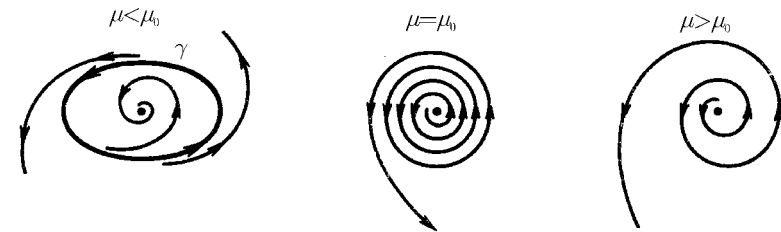


Рис. 2.151. Перестройка фазового портрета системы при жесткой потере устойчивости

Мягкое возникновение колебаний — это внутренняя бифуркация. В результате такой бифуркации из аттрактора одного типа (устойчивого фокуса) появляется аттрактор качественно другого типа (предельный цикл).

Другой часто встречающийся механизм бифуркации потери устойчивости стационарного состояния x^0 динамической системы, — это так назы-



Рис. 2.152. Рождение двумерного тора из теряющего устойчивость предельного цикла

ваемое жесткое возбуждение, когда x^0 сливается со стягивающимся к ней неустойчивым предельным циклом. В этом случае с приближением управляющего параметра μ к бифуркационному значению μ_0 область притяжения U стационарного состояния x^0 системы одновременно с размерами предельного цикла уменьшается до нуля и при $\mu = \mu_0$ цикл исчезает, сливаясь с x^0 и передавая ему свою неустойчивость. При $\mu \geq \mu_0$ все фазовые кривые покидают окрестность точки x^0 (рис. 2.151). Отметим, что перед тем, как аттрактор x^0 потеряет устойчивость, его область притяжения U станет очень малой, так что всегда присутствующие в системе возмущения выбивают фазовую кривую из этой области немного раньше того момента, когда U совсем исчезнет. Иными словами, реализуется механизм *жесткой потери устойчивости*: при прохождении через μ система скачком переходит на другой режим движения.

При жесткой потере устойчивости имеет место кризис: в результате него аттрактор — устойчивая стационарная точка — исчезает.

Еще одна бифуркация, достаточно часто встречающаяся в динамических системах, — это слияние узла и седла и обоюдное их исчезновение. В этом случае также имеет место кризис.

Бифуркации стационарного состояния мы описали в предположении, что динамическая система зависит от параметра, который *не* меняется со временем. Если же параметр *медленно* эволюционирует с течением времени, то при переходе пары собственных значений через мнимую ось возникает интересный эффект — *затягивание потери устойчивости* [27, 159, 199].

Затягивание потери устойчивости состоит в том, что система определенное время t_0 еще остается в потерявшем устойчивость состоянии равновесия. За это время родившийся цикл успевает вырасти до конечных (не малых) размеров. По прошествии времени t_0 система скачком переходит на предельный цикл, так что в целом весь процесс выглядит как жесткая потеря устойчивости. При этом существенную роль играет аналитичность

системы. Если функция v имеет только конечную гладкость, то время затягивания оказывается малым.

Механизм затягивания перехода на цикл объясняется тем, что вследствие быстрого затухания колебаний фазовая кривая динамической системы оказывается настолько близка к устойчивому фокусу x^0 , что даже после потери устойчивости требуется время, чтобы отойти от x^0 на конечное расстояние.

28.2 Типичные бифуркации предельного цикла

Предположим, что в результате бифуркации Андронова – Хопфа родился устойчивый предельный цикл. Каковы дальнейшие возможные бифуркации в динамической системе при изменении управляющего параметра? Ответ на это вопрос не является однозначным: здесь, как и при бифуркации потери устойчивости стационарного состояния, может реализоваться несколько случаев. Мы остановимся только на самых характерных и важных из них, при реализации которых дальнейшие бифуркации в системе могут привести к появлению странного аттрактора:

- 1) рождение цикла примерно удвоенного периода;
- 2) столкновение устойчивого и неустойчивого (седлового) циклов и их взаимное исчезновение;
- 3) потеря устойчивости исходного цикла и рождение инвариантного двумерного тора («с цикла слезает шкур», по выражению А. Андронова рис. 2.152).

Допустим, что поток, порождаемый динамической системой, имеет секущую S . Тогда наглядно ее поведение можно представить на этой секущей: устойчивому циклу на S будет отвечать устойчивая точка, а устойчивому двумерному тору — замкнутая кривая. Стало быть, последующая эволюция динамической системы определяется характером собственных значений (мультипликаторов) α преобразования, задающего отображение Пуанкаре.

Из результатов раздела 20.2.1 следует, что если мультипликаторы α_i находятся внутри единичной окружности, т. е. они удовлетворяют соотношению $|\alpha_i| < 1$, то цикл устойчив. Он теряет устойчивость, когда мультипликаторы выходят из единичной окружности. Рассмотрим подробнее, как это может происходить, следуя [104]. Наш анализ также прояснит, почему при переходе к хаосу типичными оказываются перечисленные выше три бифуркации.

Пусть для простоты исходная динамическая система, обладающая предельным циклом γ , трехмерна, $n = 3$. Введем в рассмотрение вблизи этого цикла секущую и определим отображение Пуанкаре. В общем виде его

можно записать как двумерное преобразование

$$\begin{aligned} x_{n+1} &= \varphi_1(x_n, y_n), \\ y_{n+1} &= \varphi_2(x_n, y_n). \end{aligned} \quad (2.239)$$

На секущей поверхности предельный цикл γ будет представлен неподвижной точкой отображения, $x^* = \varphi_1(x^*, y^*)$, $y^* = \varphi_2(x^*, y^*)$, устойчивость которой определяется характером собственных значений (мультипликаторов) α_1, α_2 матрицы линеаризации (см. раздел 20.2.1):

$$\hat{A} = \begin{pmatrix} \frac{\partial \varphi_1(x^*, y^*)}{\partial x} & \frac{\partial \varphi_1(x^*, y^*)}{\partial y} \\ \frac{\partial \varphi_2(x^*, y^*)}{\partial x} & \frac{\partial \varphi_2(x^*, y^*)}{\partial y} \end{pmatrix}.$$

Каждый собственный вектор задает инвариантное направление, которое локально определяет многообразие предельного цикла. При каждом обходе предельного цикла возмущение, отвечающее собственному вектору, умножается на соответствующий мультипликатор. Следовательно, цикл получается устойчивым, если все мультипликаторы по модулю меньше единицы, и неустойчив, когда хотя бы один из них больше единицы.

Введем след $S = \text{Sp} \hat{A}$ и детерминант $D = \det \hat{A}$ матрицы \hat{A} . Тогда собственные значения можно найти как корни уравнения

$$\alpha^2 - S\alpha + D = 0. \quad (2.240)$$

Таким образом,

$$\alpha_{1,2} = \frac{S}{2} \pm \sqrt{\frac{S^2}{4} - D}. \quad (2.241)$$

Определим область устойчивости цикла на плоскости (S, D) . Это можно сделать, если задать значения мультипликаторов на границе. Полагая в (2.240) $\alpha = +1$ и $\alpha = -1$, получим условия $1 - S + D = 0$ и $1 + S + D = 0$, определяющие две граничные линии области устойчивости. Третье условие определяется из отрицательности подкоренного выражения в соотношении (2.241): $\alpha_{1,2} = S/2 \pm i\sqrt{D - S^2/4}$. Поэтому $|\alpha| = +1$ при $D = 1$. Таким образом, на плоскости (S, D) получается треугольная область (рис. 2.153).

Очевидно, в типичной ситуации выход из этой области может происходить через одну из его сторон. Рассмотрим, какие при этом будут происходить перестройки в динамике системы.

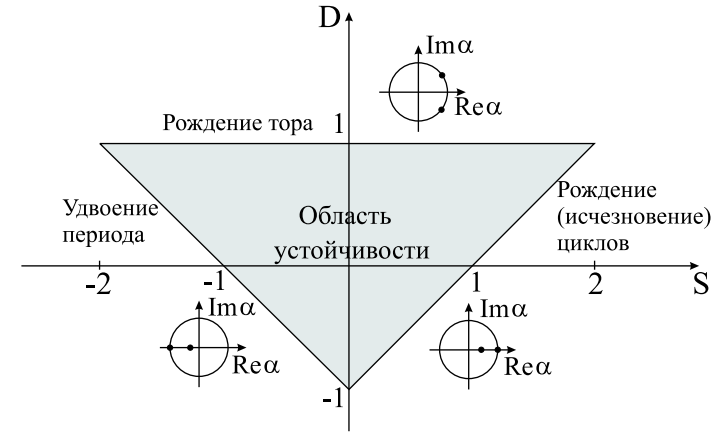


Рис. 2.153. Область устойчивости неподвижной точки системы (2.239) на плоскости (S, D) [104]

При пересечении левой границы мультипликатор $\alpha = -1$. Это означает, что с обходом цикла вектор возмущения меняет направление (рис. 2.154). Тогда, чтобы траектория замкнулась, необходимо совершить еще один оборот около теряющего устойчивость цикла. При этом, если рождающийся цикл устойчив, то в результате появится цикл с периодом вдвое большим исходного. Это — *бифуркация удвоения*.

Если первоначально в границу области притяжения исходного цикла входил седловой цикл приблизительно удвоенного периода, то с прохождением мультипликатора через -1 эти циклы сольются. При этом в рассматриваемой области не остается устойчивых траекторий, а исходный цикл станет седловым. Это соответствует *жесткому возбуждению* (или обратной бифуркации).

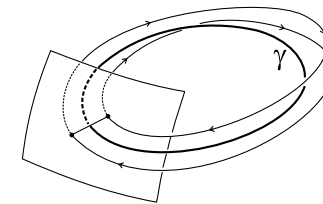


Рис. 2.154. Бифуркация цикла γ при пересечении левой границы области устойчивости на рис. 2.153

Однако возможна ситуация, когда траектория, покидая окрестность неустойчивого цикла, через некоторые промежутки времени вновь и вновь возвращается к циклу. Такое поведение траектории называется *перемежаемостью*: поведение системы почти периодическое, прерывающееся хаотическими всплесками. С увеличением параметра число хаотических пульсаций увеличивается и постепенно наступает развитый хаос. В данном случае это перемежаемость III рода (см. раздел 29.3).

Рассмотрим теперь изменения в динамике системы при пересечении правой границы области устойчивости (рис. 2.153). В этом случае мультипликатор $\alpha = +1$. С приближением к границе к устойчивому циклу приближается неустойчивый предельный цикл и сливается с ним. При этом происходит их обоюдное исчезновение, т.е. имеет место кризис. В отображении Пуанкаре такая бифуркация выглядит как слияние устойчивой и неустойчивой неподвижных точек (рис. 2.111). Эта бифуркация называется *касательной* и связана с переходом к хаосу через перемежаемость I рода (см. раздел 29.3).

Наконец, рассмотрим выход из границы области устойчивости через верхнюю границу $D = 1$ (рис. 2.153). В этом случае пара комплексно-сопряженных собственных значений $\alpha = \rho + ik$, $\alpha^* = \rho - ik$ переходит через единичную окружность. Наличие мнимой части в мультипликаторе заставляет возмущенную траекторию при обходе предельного цикла навиваться на него (рис. 2.152). При этом угол поворота вектора возмущения определяется аргументом комплексного числа α . Если с потерей устойчивости циклом динамика системы стабилизируется, то при иррациональном значении аргумента, $(\arg \alpha)/2\pi \neq p/q$, происходит *бифуркация рождения тора* и мягкое возникновение квазипериодических колебаний. При $(\arg \alpha)/2\pi = p/q$ имеет место резонанс, и траектория замыкается, совершив q оборотов на торе. Если же стабилизации не происходит, то возникает режим с перемежаемостью II рода (см. раздел 29.2).

Типичные бифуркации потери устойчивости предельным циклом показаны на рис. 2.155.

28.3 Бифуркации тора

При бифуркации рождения тора из предельного цикла пара комплексно-сопряженных мультипликаторов пересекает единичную окружность. С последующим изменением параметра в системе в принципе допустима последовательность бифуркаций, приводящая к появлению в фазовом пространстве торов все возрастающей размерности. В результате динамика системы будет характеризоваться большим числом несоизмеримых частот, см. §29.

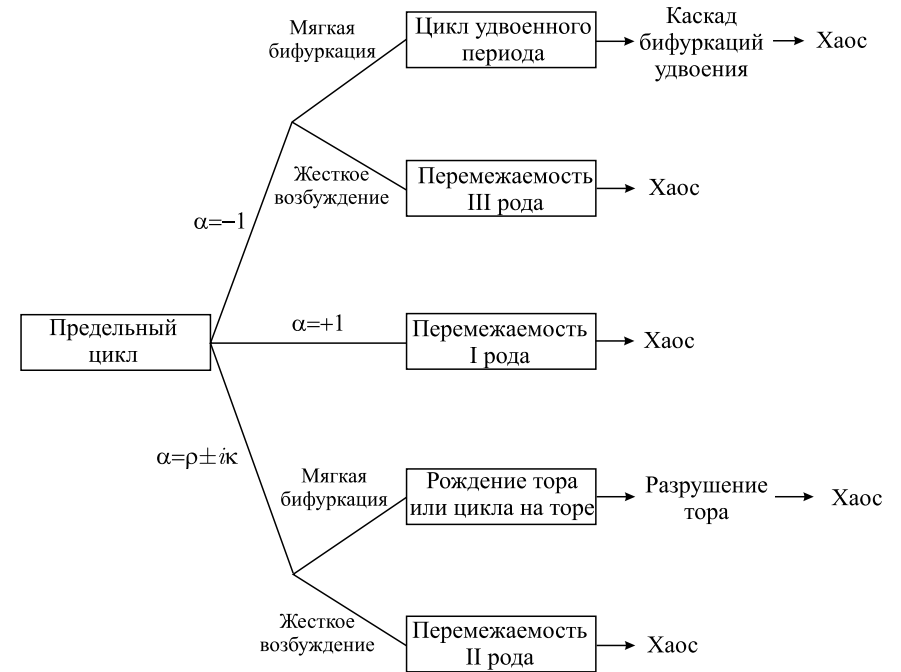


Рис. 2.155. Типичные бифуркации предельного цикла [104]

Однако существует и иная возможность, когда двумерный тор претерпевает ряд резонансов. Это приводит к появлению и исчезновению предельных циклов, расположенных на торе. С дальнейшим ростом параметра тор может потерять гладкость и превратиться в странный аттрактор, отвечающий хаотическому поведению динамической системы (см. §29).

Нарушение гладкости тора удобно рассмотреть на примере отображения, порождаемого фазовым потоком в его окрестности. Мы уже встречались с подобным построением в §17.

Предположим, что динамическая система имеет устойчивый инвариантный двумерный тор, на котором определена глобальная секущая Σ . Это означает, что любая фазовая кривая, расположенная на торе, пересекает Σ без касания. Следовательно, через эту секущую можно провести некоторую поверхность S , которая будет пересекать траектории, лежащие в окрестности тора. Легко понять, что на S такие траектории образуют некоторое кольцо, а сам тор представляется замкнутой инвариантной кривой (рис. 2.156а).

Таким образом, в общем случае задача сводится к изучению диссипативного отображения кольца в себя. При этом конкретный вид отображения не играет роли [27, 33]. Перестройки фазовых портретов в таком кольце имеют место и в общей ситуации, что позволяет рассмотреть модельное отображение.

Однако представления об основных свойствах отображения кольца можно получить, рассматривая отображение окружности в себя, естественным образом возникающим на секущей Σ .

28.3.1 Отображение окружности и его свойства

Введем координаты на торе, как показано на (рис. 2.156а). При соизмеримости частот движения, $\omega_1/\omega_2 = p/q$ (резонанс), фазовая траектория, расположенная на торе, замыкается, совершив p оборотов по координате x и q оборотов по координате ϑ (рис. 2.156б). Таким образом, в отображении Пуанкаре будет наблюдаться m точек, переходящих друг в друга, что соответствует периодической динамике системы.

Когда частоты ω_1 и ω_2 несоизмеримы, поведение системы будет квазипериодическим, а точки в отображении Пуанкаре при $t \rightarrow \infty$ плотно заполняют инвариантную замкнутую кривую (см.рис. 2.8), которую подходящим преобразованием можно превратить в окружность.

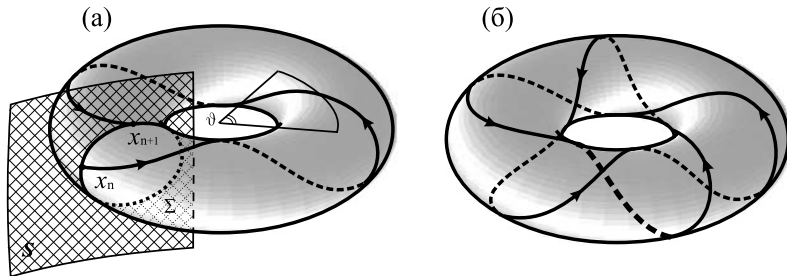


Рис. 2.156. Отображение Пуанкаре как функция x_{n+1} от x_n на секущей Σ (а) и резонансная фазовая траектория на торе при $\omega_1/\omega_2 = 5/2$ (б)

Выберем в качестве координаты на торе переменную x (рис. 2.156а). Тогда в общем случае в отображении Пуанкаре будем иметь:

$$x_{n+1} = \varphi(x_n) = x_n + a + g(x_n), \quad (2.242)$$

где $g(x)$ — некоторая 2π -периодическая функция, и параметр a задает сдвиг по x за один оборот по координате ϑ (т.е. за одну итерацию). Очевидно, значение x определяется с точностью до 2π .

Преобразование координат (2.242) можно интерпретировать как последовательное перемещение точек на окружности. Поэтому соотношение (2.242) называется *отображением окружности*.

Функция φ может быть как монотонно возрастающей на $[0, 2\pi]$ (и потому взаимно-однозначной), так и имеющей максимум и минимум (т.е. не взаимно-однозначной). Для монотонной функции φ характер динамики отображения (2.242) определяется *числом вращения*, которое можно записать как

$$\rho = \lim_{n \rightarrow \infty} \frac{1}{2\pi} \frac{\varphi^n(x_0) - x_0}{n},$$

где $\varphi^n = \underbrace{\varphi(\varphi(\dots(\varphi)))}_{n \text{ раз}}$ — n -я итерация отображения. По-существу число

вращения представляет собой количество оборотов по координате x в течение одного оборота по координате ϑ . Для преобразования (2.242) по смыслу это просто параметр a .

Действительно, рассмотрим простейшее преобразование поворота окружности на постоянный угол a : $x \rightarrow \{x + a\}$, где фигурные скобки означают взятие аргумента с точностью до 2π . Тогда $x^n = x + na$. Следовательно, $a = \lim_{n \rightarrow \infty} \frac{x^n - x}{n}$.

Таким образом, если ρ является рациональным числом, $\rho = p/q$, то начальное значение x_0 будет повторяться через определенное число периодов $x_{n+q} - x_n = 2\pi p$, так что динамика отображения будет периодической. Когда число вращения иррационально, то траектория почти любой точки оказывается всюду плотной на окружности.

Для не взаимно однозначной функции φ отображение окружности может обладать очень сложными режимами поведения. Более того, для такой функции φ в различных областях параметрического пространства возможно сосуществование двух различных хаотических режимов или цикла и хаоса.

Для иллюстрации сказанного рассмотрим известное *стандартное* отображение окружности:

$$x_{n+1} = \varphi_{a,b}(x) = x_n + a + b \sin x_n, \quad (2.243)$$

определенное на интервале $[0, 2\pi]$. Когда $b < 1$, функция $\varphi_{a,b}$ является монотонной, а отображение — взаимно-однозначным (рис. 2.157а). При $b = 1$ оно еще остается монотонным и взаимно-однозначным, но теперь обратная

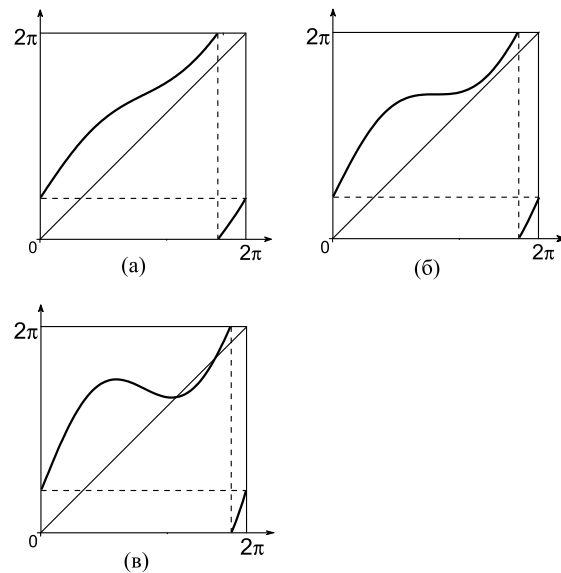


Рис. 2.157. Функция $\varphi_{a,b}(x)$ (2.243) при $a = 1, 4$ и различных значениях параметра b : $b < 1$ (а), $b = 1$ (б), $b > 1$ (в)

функция уже не гладкая, поскольку в точке перегиба производные претерпевают разрыв (рис. 2.157б). Если же $b > 1$, то $\varphi_{a,b}$ не будет взаимнооднозначной функцией (рис. 2.157в), так что при некоторых значениях параметров в отображении (2.243) может возникнуть хаос (см. раздел 24.2.2). Следовательно, интервалы $b < 1$ и $b > 1$ соответствуют двум качественно различным режимам поведения. По этой причине значение $b = 1$ называется *критическим*.

В параметрическом пространстве отображения (2.243) каждой паре (a, b) отвечает либо рациональное, либо иррациональное значение числа вращения. Следовательно, все пространство (a, b) разбивается на резонансные зоны (т. е. области синхронизации) — так называемые *языки Арнольда*, отвечающие числам $\rho = p/q$ (см. рис. 2.158). Когда $b \rightarrow 0$, то число вращения $\rho = a/2\pi$, так что динамика будет определяться значением параметра a . В этом случае языки Арнольда своими клювами подходят к рациональным точкам оси $a/2\pi$. В малой окрестности оси их суммарная площадь мала по сравнению с площадью самой окрестности. С ростом b площадь, занимаемая языками, увеличивается, они растут и при $b > 1$ могут перекрываться.

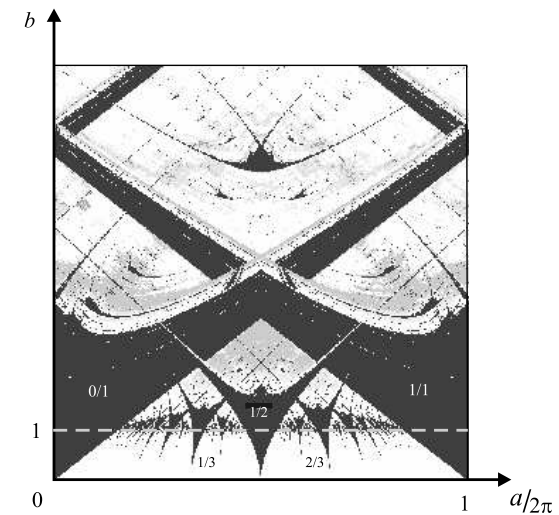


Рис. 2.158. Бифуркационная диаграмма отображения (2.243). Области, отмеченные серым цветом, соответствуют периодической динамике с рациональными значениями числа вращения $\rho = p/q$.

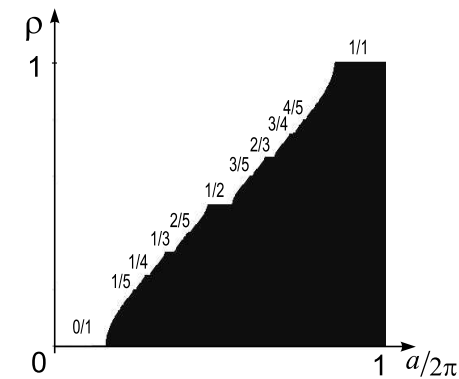


Рис. 2.159. «Чертова лестница» отображения окружности (2.243) при $b = 1$. Хорошо видна ее самоподобная структура

Языки Арнольда существуют для всех рациональных отношений p/q , где p и q — взаимно-простые числа. Это означает, что имеется бесконечное число языков, отвечающих всем возможным отношениям частот движения.

Между любыми двумя языками с числами вращения p/q и p'/q' существует другая область синхронизации с отношением $(p+p')/(q+q')$. Несмотря на это, в области $0 \leq b < 1$ всегда найдутся значения параметров a , соответствующие квазипериодическому поведению отображения окружности.

Таким образом, в обратимой области $0 < b < 1$ для каждого фиксированного рационального числа вращения $\rho = p/q$ существует некоторый интервал значений параметра a , который определяется шириной соответствующего языка Арнольда. Отношения p/q можно упорядочить в последовательность, которая называется *деревом*, или *рядом Фарей* (Фэри). Ряд Фарей — это возрастающая последовательность неотрицательных несократимых дробей. При этом, если $p/q, p''/q'', p'/q'$ — любые три последовательные дроби дерева Фарей, то $p''/q'' = (p+p')/(q+q')$ и $p''q - q''p = 1$. Теперь, если для отображения окружности представить значения ρ в зависимости от параметра $a/2\pi$, то мы получим непрерывную монотонную функцию, содержащую бесконечное число горизонтальных отрезков. Она будет локально постоянной для каждого $a/2\pi$, отвечающего рациональному числу ρ . Функция, обладающая такими свойствами, называется «*чертовой лестницей*» (см. рис. 2.159).

Когда $b \rightarrow 0$, общая длина l всех ступеней стремится к нулю, поскольку мера рациональных чисел на отрезке $[0, 1]$ нулевая. При увеличении параметра b величина l монотонно возрастает и при $b = 1$ становится равной единице. Напротив, мера значений параметра $a/2\pi$, соответствующих иррациональным числам вращения ρ (т.е. не принадлежащих ступенькам «чертовой лестницы»), при $b = 0$ равна единице. С ростом b мера таких значений убывает и при $b = 1$ становится равной нулю (рис. 2.160).

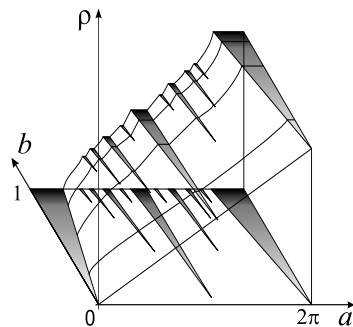


Рис. 2.160. «Чертова лестница» отображения окружности (2.243) в пространстве a, b, ρ

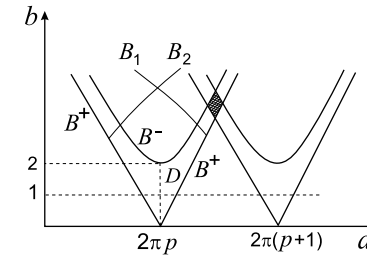


Рис. 2.161. Структура языков Арнольда отображения окружности (2.243) [33]

«Чертова лестница» отображения окружности с критическим значением параметра ($b = 1$) называется *полной*. Она обладает свойством самоподобия и почти всюду горизонтальна. В этом случае, хотя множество точек, отвечающих квазипериодической динамике, имеет нулевую меру, его фрактальная размерность положительна и равна примерно 0,87 [223, 369].

При $b > 1$ языки Арнольда вытягиваются, расщепляясь на ветви. Следовательно, в этой области они могут перекрываться, образуя очень сложную пространственную структуру, также обнаруживающую свойства самоподобия (см. рис. 2.158). В области пересечения двух языков имеет место *бистабильность*, когда для одних и тех же значений параметров a, b существуют два устойчивых цикла отображения с различными числами вращения. Каждый из этих циклов можно наблюдать при попадании начальных условий в его область притяжения. Кроме того, внутри языков Арнольда существуют зоны, где имеет место переход к хаосу через бифуркации удвоения, начиная со основного для данного языка значения периода.

Описанные самоподобные свойства отображения окружности наводят на мысль изучить это отображение методом ренормализации аналогично логистическому отображению. Исследования такого рода привели к открытию универсальных свойств при переходе от квазипериодического движения к хаосу (см. раздел 29.1.1).

Общая структура языков Арнольда отображения окружности оказывается следующей (рис. 2.161) [27, 33]. Неподвижные точки этого отображения с числом вращения $2\pi\rho$ существуют внутри области, ограниченной прямыми B^+ , которые находятся из соотношения $b = \pm(2\pi\rho - a)$. Эти точки соответствуют предельным циклам на торе, замыкающимся через p оборотов по координате x . На прямых B^+ неподвижные точки сливаются, образуя точку с мультипликатором $+1$.

Линия, отвечающая бифуркации удвоения периода, находится из условия равенства мультипликатора значению -1 в неподвижной точке. Это

условие, следовательно, можно записать как

$$\begin{aligned} b \cos x &= -2, \\ b \sin x &= 2\pi p - a. \end{aligned}$$

Легко видеть, что это гипербола $b^2 = (2\pi p - a)^2 + 4$, которая обозначена на рис. 2.161 как B^- .

Из общей теории одномерных отображений следует [426], что если критическая точка за конечное число итераций попадает в неподвижную точку, то возникают траектории с числом вращения, отличным от $2\pi p$. Это условие определяет бифуркационные кривые B_1 и B_2 (рис. 2.161). Очевидно, критические точки x_c для отображения (2.243) появляются, когда $b > 1$ (см. рис. 2.157).

Таким образом, поведение отображения окружности (2.243) внутри языка Арнольда определяется перечисленными бифуркационными кривыми (рис. 2.161), которые на плоскости (a, b) образуют область D . Если $(a, b) \in D$, то в отображении имеется одна устойчивая неподвижная точка. При переходе через B^+ , т. е. выходе из резонансной зоны, все зависит от значения параметра b . Когда $b < 1$ и число вращения рационально, отображение будет иметь четное число устойчивых и неустойчивых циклов. Для иррационального ρ динамика является квазипериодической. При $b > 1$ может возникнуть хаотическая динамика. С переходом через гиперболу B^- в отображении происходит бифуркация удвоения периода, а через кривые $B_{1,2}$ — рождение циклов, имеющих числа вращения $\rho \neq 2\pi p$.

Для чисел вращения $2\pi(p+1)$ структура резонансной зоны оказывается абсолютно аналогичной. Следовательно, на плоскости параметров (a, b) будет существовать область, отвечающая наличию неподвижных точек с числами вращения $2\pi(p+1)$ и $2\pi p$ (на рис. 2.161 она заштрихована). В этой области режим движения, как уже отмечалось, является бистабильным. В зависимости от выбора начальных условий будет наблюдаться то или иное поведение системы.

28.3.2 Свойства отображения кольца

Вернемся к рис. 2.156 *a* и рассмотрим отображение, порождаемое фазовым потоком на поверхности сечения S в окрестности тора. Пусть

$$\Phi_{a,b} : \begin{cases} x_{n+1} = x_n + a + y_n + b \sin x_n, \\ y_{n+1} = e^{-a}(y_n + b \sin x_n), \end{cases} \quad (2.244)$$

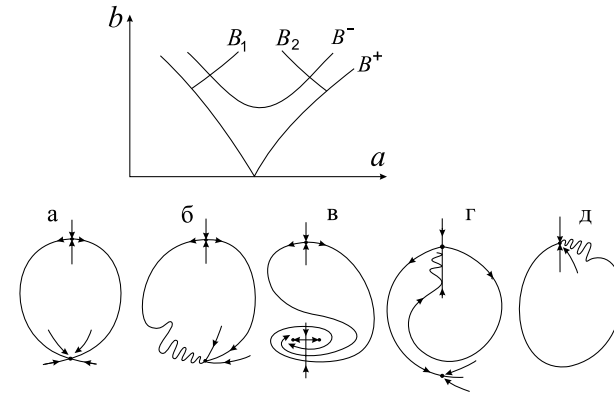


Рис. 2.162. Бифуркационная диаграмма отображения кольца (2.244) и соответствующие фазовые портреты [33]

где $a > 0$, $b > 0$ и $0 \leq x \leq 2\pi$. Хотя это отображение является модельным, все присущие ему бифуркационные явления имеют место и в общей ситуации [33].

Отображение (2.244) взаимно-однозначно и переводит кольцо $0 \leq x \leq 2\pi$, $|y| \leq y_0$, $y_0 > be^{-a}/(1 - e^{-a})$ в себя. Его свойства во многом аналогичны свойствам только что описанного отображения окружности (рис. 2.162). В частности, это отображение имеет кривые B^+ и B^- , отвечающие наличию неподвижной точки с мультипликаторами $+1$ и -1 , соответственно, и кривые B_1 , B_2 . Уравнения для B^+ , B^- даются выражениями $b = \pm(2\pi p - a)(1 - e^{-a})$ и $b^2 = (2\pi p - a)^2(1 - e^{-a}) + 4(1 + e^{-a})^2$. Легко видеть, что при $a \gg 1$ B^+ и B^- аппроксимируются соответствующими кривыми отображения окружности. Соотношения, которыми описываются кривые $B_{1,2}$, при $a \gg 1$ также близки к уравнениям для соответствующих кривых отображения окружности.

Рассмотрим бифуркации в системе, возникающие при пересечении этих кривых, следуя [33, 40].

При $b + e^{-a} < 1$ отображение кольца имеет инвариантную замкнутую кривую, которая включает неподвижные точки, одна из которых устойчива, а другая седловая, причем ее неустойчивые сепаратрисы замыкаются на устойчивую точку, образуя тем самым замкнутую кривую (см. рис. 2.162 *a*). В пространстве параметров (a, b) такая ситуация соответствует области, ограниченной бифуркационными кривыми B^+ , B^- и $B_{1,2}$. Эта замкнутая кривая становится негладкой (при $b > 1$, рис. 2.162 *б*), но существует при всех значениях параметров a и b .

С переходом через B^- происходит бифуркации удвоения периода (рис. 2.162 *е*).

Кривые $B_{1,2}$ отвечают появлению гомоклинической траектории на ветвях неустойчивой сепаратрисы (рис. 2.162 *з*). При пересечении $B_{1,2}$ возникает бесконечное множество траекторий со счетным числом неустойчивых седловых циклов.

При переходе через кривые B^+ седло и узел сливаются, появляется седло-узел, причем инвариантную кривую в этот момент образует его неустойчивая сепаратриса (рис. 2.162 *д*). Если эта кривая гладкая, то после исчезновения седло-узла рождается гладкая замкнутая кривая, к которой притягиваются все точки в кольце. Если же сепаратриса негладкая, то возникает хаотическая динамика (она реализуется в зависимости от значения параметра b). При этом в системе будет наблюдаться режим с перемежаемостью.

Таким образом, классификация смены режима движения на торе определяется тремя основными видами бифуркаций [27]:

- сменой устойчивости цикла на торе;
- рождением гомоклинических структур;
- слиянием и исчезновением циклов на негладком торе.

Этот факт известен как теорема о разрушении тора [40] (см. также [375]).

29. Типичные сценарии перехода к хаосу

Установление в динамической системе хаотического режима движения в результате той или иной последовательности бифуркаций принято называть *сценарием*, или *картиной*, развития хаоса. Как мы видели в предыдущем параграфе, в типичном случае это может происходить путем последовательности бифуркаций удвоения периода предельного цикла, через перемежаемость или разрушение тора. Здесь мы обсудим в целом эти сценарии и опишем универсальные свойства при переходе к хаосу через квазипериодичность и перемежаемость.

29.1 Разрушение тора

Допустим, что в результате потери устойчивости предельным циклом в фазовом пространстве динамической системы родился двумерный тор (рис. 2.152)

С последующим изменением параметра μ в фазовом пространстве многомерной динамической системы может произойти потеря его устойчивости и рождение трехмерного тороидального многообразия. При этом поведение системы будет характеризоваться тремя независимыми частотами. Дальнейшее изменение управляющего параметра может привести к последовательности бифуркаций, в результате которых в фазовом пространстве диссипативных динамических систем возникают инвариантные торы все возрастающей размерности. В конечном счете мы приходим к сложному квазипериодическому движению с k несоизмеримыми частотами, которое при очень большом k будет выглядеть как хаотическое. Считая, что такой путь развития хаоса действительно возможен, Ландау [112, 113] и независимо Хопф [356] выдвинули гипотезу, согласно которой хаотическая динамика диссипативных систем есть не что иное, как движение по инвариантному тору большой размерности. Такой тор будет занимать в фазовом пространстве область, соответствующую всевозможным наборам начальных фаз, и наматывающаяся на него фазовая траектория будет с течением времени проходить практически через любую сколь угодно малую часть этой области.

Однако, как было показано в разделе 18.2, квазипериодическое движение, пусть даже с очень большим числом несоизмеримых частот, не может быть названо хаотическим, поскольку для такого движения отсутствует разбегание фазовых кривых, ответственное за появление хаотической динамики. Кроме того, следует сказать, что многомерное квазипериодическое притягивающее движение с большим числом несоизмеримых частот не является типичным и встречается чрезвычайно редко. Под действием всегда присутствующих возмущений такое движение, скорее всего, с течением времени вырождается в периодическое, отвечающее появлению в фазовом пространстве предельного цикла, или же разрушается и дает начало хаотическому режиму. Отметим также, что картина Ландау–Хопфа не подтверждается экспериментально: после небольшого числа бифуркаций обычно наблюдается резкий переход к хаотическому движению.

Впервые на возможность разрушения притягивающего тороидального многообразия, в результате которого происходит рождение странного аттрактора, обратили внимание Д. Рюэль и Ф. Такенс [173] (см. также [440]). Авторы исследовали поведение решения динамической системы вида (2.238) при весьма общих предположениях относительно характера векторного поля v . Ими было показано, что если при изменении управляющего параметра μ после трех бифуркаций (начиная со стационарного состояния) и возникает трехчастотное квазипериодическое движение, как и в модели развития хаоса Ландау–Хопфа, то оно, как правило, неустойчиво,

легко разрушается, и на месте разрушенного трехмерного тора появляется странный аттрактор.

Остановимся на этом сценарии развития хаоса более детально.

Напомним, что каждая динамическая система, как предполагалось выше, задается соответствующим векторным полем v , т. е. правыми частями дифференциальных уравнений (2.238). Совокупность всех возможных векторных полей v образует некоторое функциональное пространство Φ . Каждая точка этого функционального пространства отвечает одной из возможных динамических систем, и наоборот, каждая динамическая система соответствует единственной точке функционального пространства Φ . Если слегка возмутить правые части уравнений (2.238), т. е., как еще говорят, произвести малые шевеления соответствующего векторного поля v , мы получим новую динамическую систему, для которой векторное поле будет близко к исходному. Все динамические системы, векторные поля которых могут быть получены малым шевелением векторного поля данной динамической системы, образуют ее окрестность в функциональном пространстве Φ .

Говорят, что векторное поле v динамической системы является *грубым*, или *структурно-устойчивым*, если имеется окрестность U поля v , такая, что для каждого векторного поля v' из этой окрестности фазовый портрет качественно не отличается от фазового портрета динамической системы, задаваемой полем v . Соответственно, свойство динамической системы структурно-устойчиво, если при малых возмущениях системы это свойство сохраняется. Понятие структурной устойчивости (грубости) впервые было введено для дифференциальных уравнений А. А. Андрономым и Л. С. Понтрягиным [14].

При построении математической модели любого реального процесса всегда приходится прибегать к определенным упрощениям, пренебрегать влиянием малосущественных факторов и т. п. Следовательно, векторное поле, входящее в правую часть динамических уравнений, всегда будет известно лишь с какой-то степенью точности, т. е. в пределах некоторой малой окрестности U функционального пространства Φ . Поэтому очевидно, что если рассматриваемое свойство динамической системы не является структурно-устойчивым, то для реальных систем оно в эксперименте наблюдаться не будет.

Теорема Рюэля и Такенса [173, 440] утверждает, что если существует векторное поле v на трехмерном торе, отвечающее трехчастотному квазипериодическому движению, то в любой окрестности U соответствующей точки функционального пространства Φ найдутся векторные поля v' , обладающие странными аттракторами. Аналогичное утверждение справедливо для квазипериодических движений и на торах большей размерности. Ина-

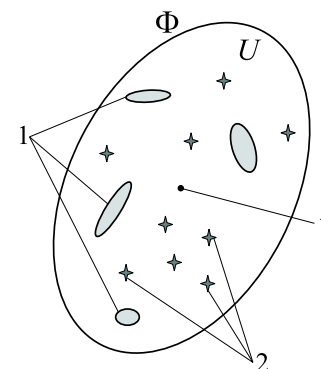


Рис. 2.163. Окрестность U векторного поля v в функциональном пространстве Φ : 1 — множества векторных полей, обладающих странными аттракторами; 2 — множества полей с регулярной динамикой [202]

че говоря, в принципе достаточно слабо возмутить правые части системы (2.238), чтобы движение из квазипериодического с тремя несоизмеримыми частотами перешло в хаотическое. Однако это выполняется не для всех векторных полей v' , имеющих в окрестности U . Если бы векторное поле v было структурно-неустойчиво, то, действительно, какие угодно малые возмущения привели бы к разрушению трехмерного тора и появлению режима движения, качественно отличного от квазипериодического. Теорема Рюэля — Такенса утверждает, что при малых возмущениях поля v может возникнуть странный аттрактор: в окрестности U существуют поля v' , среди которых есть такие, которые имеют странные аттракторы, и такие, которые не обладают ими. Доля областей, занятых векторными полями со странными аттракторами, конечна, однако эти области не заполняют всю окрестность U (см. рис. 2.163).

Подтверждением того, что не для всякого векторного поля с трехчастотным квазипериодическим движением малые возмущения приводят к появлению странного аттрактора, являются, например, исследования [331–333, 507, 508, 521] (см. также [201]). В частности, в работе [331] было показано, что добавление гладких нелинейных возмущений необязательно разрушает трехчастотную динамику, а в работе [521] экспериментально была обнаружена возможность движения с четырьмя и даже пятью несоизмеримыми частотами.

Более того, анализ некоторых систем показал, что переход от квазипериодического движения к хаосу может происходить через появление

странных нехаотических аттракторов [333] (см. также [33]). При этом слово «странный» используется как характеристика геометрии аттрактора в том смысле, что он является фрактальным множеством с нецелой размерностью. В системе с таким аттрактором хаос отсутствует, т. е. она не имеет положительных показателей Ляпунова. Следовательно, тор, прежде чем разрушиться, теряет гладкость (см. §28). Хотя потеря гладкости тором не обязательно приводит к рождению хаоса, переходу к хаосу в системе всегда предшествует стадия возникновения недифференцируемого тора. Такой сценарий развития хаотичности описан в §28.

29.1.1 Универсальные свойства перехода от квазипериодической динамики к хаосу

Количественные аспекты сценария перехода к хаосу через квазипериодическое движение можно изучить в рамках ренормализационного анализа, переходя от системы дифференциальных уравнений к отображению [307, 463, 464, 481]. Мы кратко представим элементы такого подхода, следуя, в основном, монографиям [72, 201] и учебнику [104].

В общем случае в окрестности двумерного тора фазовый поток порождает двумерное отображение кольца в себя (см. раздел 28.3.2). В свою очередь, основные свойства отображения кольца можно получить, используя модельное отображение окружности, например,

$$x_{n+1} = \varphi_{a,b}(x) \equiv x_n + a + b \sin x_n, \quad 0 \leq x \leq 2\pi, \quad (2.245)$$

которое уже было подробно рассмотрено в разделе 28.3.1. Такой анализ позволяет выявить ряд универсальных закономерностей, получивших экспериментальное подтверждение.

При переходе от квазипериодичности к хаосу в отображении окружности необходимо изменять сразу два параметра, чтобы сохранять число вращения ϱ равным заданному иррациональному числу. Ренормализационный анализ, в основном, проводился для числа вращения, равного золотому среднему.

Как известно, любое число $0 < \chi \leq 1$ единственным образом можно разложить в непрерывную дробь:

$$\chi = \frac{1}{s_0 + \frac{1}{s_1 + \frac{1}{s_2 + \dots}}} = \langle s_0, s_1, s_2, \dots \rangle, \quad (2.246)$$

где s_0, s_1, s_2, \dots — натуральные числа. Для рационального числа такое разложение будет всегда конечным, $\chi = p/q = \langle s_0, s_1, s_2, \dots, s_N \rangle$. Иррациональные же числа представляются бесконечными дробями. Если оборвать дробь, например, на m -ом элементе, то получится наилучшее приближение иррационального числа рациональным: $\chi_m = p_m/q_m = \langle s_0, s_1, s_2, \dots, s_m \rangle$. При этом

$$p_{m+1} = p_m s_m + p_{m-1}, \quad p_0 = 0, \quad p_{-1} = 1, \quad (2.247)$$

$$q_{m+1} = q_m s_m + q_{m-1}, \quad q_0 = 1, \quad q_{-1} = 0.$$

Золотое среднее — это простейшая бесконечная непрерывная дробь, у которой $s_0, s_1, s_2, \dots = 1$:

$$w \equiv \langle 1, 1, 1, \dots \rangle = \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}} = \frac{1}{1+w}.$$

Отсюда находим, что $w = (\sqrt{5} - 1)/2 = 0,61803398874989\dots$. Легко видеть, что геометрически это число определяется золотым сечением отрезка, т. е. таким его делением, когда отношение большей части ко всей длине равно отношению меньшей части к большей.

Золотое среднее наилучшим образом приближается так называемыми числами Фибоначчи F_n , которые определяются как $F_{n+1} = F_n + F_{n-1}$, $F_0 = 0, F_1 = 1$, где $n = 1, 2, \dots$. Для таких чисел

$$\chi_n = \frac{F_n}{F_{n+1}} = \frac{1}{1 + \frac{F_{n-1}}{F_n}} = \frac{1}{1 + \underbrace{\frac{1}{1 + \dots}}_{n \text{ раз}}},$$

Таким образом,

$$w = \lim_{n \rightarrow \infty} \chi_n.$$

Нетрудно видеть, что число $-w$ есть собственное значение матрицы

$$\hat{T} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

При этом, очевидно, \hat{T}^n — это матрица

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix},$$

где F_i — i -е число Фибоначчи. Поэтому

$$F_{n-1} - F_n w = (-w)^n. \quad (2.248)$$

Следовательно, можно записать $R_w^{F_n}(x) - x = F_n w$, где R_w обозначает поворот окружности на угол w . Поэтому $|R_w^{F_n}(x) - x| = w^n$. Значит, последовательность $R_w^{F_n}$ сходится со скоростью геометрической прогрессии к тождественному преобразованию. Кроме того, с помощью рекуррентного соотношения для чисел Фибоначчи найдем, что $R_w^{F_{n+1}}$ порождается композицией $R_w^{F_n} \left(R_w^{F_{n-1}} \right)$.

Используя данные соображения, можно найти универсальные масштабные свойства при переходе к хаосу через квазипериодичность.

Введем новую переменную $\theta = x/2\pi + 1/2$. Тогда отображение окружности запишется в виде

$$\theta_{n+1} = f(\theta) = \theta_n + r - \frac{b}{2\pi} \sin 2\pi\theta_n, \quad (2.249)$$

где $r = a/2\pi$. При этом точка перегиба при $b = 1$ будет расположена в начале координат, а значения θ заданы на интервале $[0, 1]$.

Универсальные свойства перехода к хаосу описываются ренормализационным уравнением для так называемой критической точки, которая соответствует иррациональному числу вращения ρ для $b = 1$ и определенного $r = r_c$ (см. ниже). При этом ρ приближается рациональными числами $p_m/q_m = \langle s_0, s_1, \dots, s_m \rangle$ и рассматривается последовательность операторов эволюции для интервалов времени, заданных величиной q_m .

Обозначим $f_m(\theta) \equiv \underbrace{f(f(\dots f(\theta) \dots))}_{q_m \text{ раз}} - p_m \equiv f^{q_m}(\theta) - p_m$, где функция

f определяется соотношением (2.249). Тогда, поскольку $f(\theta + p) = f(\theta) + p$ для любого целого p , из (2.247) найдем

$$f_{m+1}(\theta) = f_{m-1}(f_m^{s_m}(\theta)).$$

Теперь, аналогично тому, как это было сделано для случая удвоения периода, (см. раздел 26.2), на каждом шаге преобразования необходимо ввести соответствующую нормировку на некоторую величину d_m . Тогда оператор эволюции за число шагов q_m представится как

$$g_m(\theta) = d_m f_m \left(\frac{\theta}{d_m} \right).$$

Положим $g_m(0) = 1$, т. е. $d_m = 1/f(0)$.

Используя обозначение $\alpha_m = d_m/d_{m-1}$, уравнения, задающие последовательность функций g_m , можно записать как

$$\alpha_{m+1} = \frac{1}{\alpha_m g_{m-1} \left(\alpha_m^{-1} g_m^{s_m}(0) \right)}, \quad (2.250)$$

$$g_{m+1}(\theta) = \alpha_{m+1} \alpha_m g_{m-1} \left(\alpha_m^{-1} g_m^{s_m} \left(\frac{\theta}{\alpha_{m+1}} \right) \right).$$

Таким образом, мы получили преобразование, позволяющее совершить переход от функций g_{m-1} , g_m к функции g_{m+1} :

$$g_{m+1} = \mathcal{T}_m(g_m, g_{m-1}). \quad (2.251)$$

Это ренормализационное уравнение играет ту же роль в теории универсальности при переходе от квазипериодичности к хаосу, как и уравнение (2.229) в теории Фейгенбаума. Однако, в отличие от него, соотношение (2.251) является двумерным отображением. Кроме того, оно неавтономно. Зависимость от итераций здесь реализуется через последовательность значений s_m разложения соответствующего числа вращения в рациональную дробь. Это, в частности, означает, что в общем случае невозможно найти решение, которое не зависело бы от m .

Тем не менее, для числа вращения, заданного золотым средним w , все элементы s_m непрерывной дроби (2.246) равны единице. Поэтому индекс s_m в уравнениях (2.250) при таком выборе ρ можно опустить. В этом случае функция g_m представляет собой перенормировку за число шагов, равное числу Фибоначчи F_m .

При $b = 1$ значению числа вращения $\rho = w$ на плоскости параметров (b, r) соответствует определенная точка, которая называется *критической*. Основное ее отличие от остальных точек параметрического пространства состоит в том, что здесь окружность теряет гладкость, так что соответствующий ей тор становится фрактальным множеством (см. раздел 28.3). Для нашего отображения окружности (2.249) координаты критической точки $(b, r) = (b, r_c) = (1; 0,6066610634702\dots)$.

Если $m \rightarrow \infty$, то из (2.250) получаем для функции неподвижной точки оператора \mathcal{T} :

$$g(\theta) = \alpha^2 g \left(\alpha^{-1} g \left(\frac{\theta}{\alpha} \right) \right).$$

Это уравнение называется *уравнением Фейгенбаума–Каданова–Шенкера*.

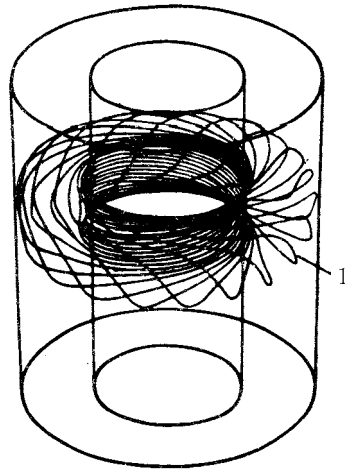


Рис. 2.164. Схема эксперимента по исследованию вихрей Тейлора: 1 — численно найденная траектория частицы внутри вихрей Тейлора [189]

Его решение методом полиномиальной аппроксимации дает:

$$g(\theta) = 1 + 0,765184 \theta^3 - 0,215464 \theta^6 - 0,053469 \theta^9 + \\ + 0,032921 \theta^{12} + 0,001231 \theta^{15} - 0,004304 \theta^{18} \dots,$$

$\alpha = -1,288574553954368866 \dots$ (см. [307, 481]). Кроме того, анализ поведения малого возмущения, которое отвечает сдвигу по параметрам из критической точки, приводит к двум основным собственным числам:

$$\delta_1 = -2,833610655891167799 \dots,$$

$$\delta_2 = \alpha^2 = 1,6604243810987068 \dots$$

Эти постоянные α , δ_1 и δ_2 характеризуют универсальность скейлинга при переходе к хаосу через квазипериодическое движение с отношением частот, близким к золотому среднему.

Некоторые экспериментальные данные свидетельствуют о том, что сценарий перехода к хаосу Рюэля–Такенса действительно выполняется для ряда систем. Так, в [324] исследовались вихри Тейлора, возникающие в

жидкости между двумя concentрическими цилиндрами, внутренний из которых может вращаться с постоянной скоростью (рис. 2.164). Динамика вихрей изучалась методом светового рассеяния. Свет, рассеянный небольшим объемом жидкости, несет информацию о частотах фурье-компонент радиальных составляющих скорости жидкости. Исследования проводились для различных значений числа Рейнольдса. Переход к хаотическому режиму в такой системе наблюдался вслед за тем, как последняя бифуркация приводила к квазипериодическому движению с тремя независимыми частотами [67, 324], которое являлось структурно-неустойчивым и под действием естественных возмущений разрушалось, приобретая хаотические черты.

Были проведены эксперименты по исследованию конвекции Рэля–Бенара в подогреваемом снизу горизонтальном слое жидкости [500]. С увеличением градиента температуры перед переходом к хаотическому движению в спектре скорости жидкости наблюдалась сначала одна, а затем две независимые частоты (рис. 2.165а, б). Хаотический режим, характеризующийся сплошным спектром, появлялся сразу вслед за квазипериодическим двухчастотным течением (рис. 2.165в).

Хаотическая динамика при исследовании конвекции Рэля–Бенара в ртути [188] была обнаружена после того, как в спектре последовательно возникали три рационально независимые частоты. После появления третьей частоты для системы характерно возникновение низкочастотного шума со сплошным и широким спектром.

Эти и ряд других экспериментальных данных [7, 134, 218, 280, 285, 470] подтверждают сценарий Рюэля–Такенса: в спектрах движения появляется сначала первая, затем вторая и (иногда) третья независимые частоты. На пороге появления третьей частоты внезапно возникает широкополосный

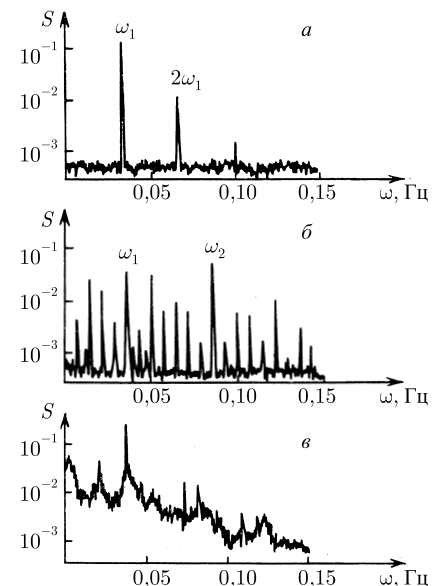


Рис. 2.165. Частотная зависимость спектральной плотности скорости жидкости в эксперименте Рэля–Бенара: а — периодический, б — квазипериодический двухчастотный и в — хаотический режимы [500]

шум, свидетельствующий о хаотичности происходящих в системе процессов.

Хотя сценарий Рюэля–Такенса является достаточно типичным, кроме него есть еще как минимум два столь же часто осуществляющихся сценария: каскад бифуркаций удвоения и перемежаемость.

29.2 Сценарий Фейгенбаума

Этот сценарий и соответствующие ему бифуркации удвоения уже достаточно подробно были рассмотрены в §26 и §28. Поэтому здесь мы остановимся только на общей картине развития хаоса путем последовательности таких бифуркаций.

Сценарий Фейгенбаума отвечает ситуации, когда в результате потери устойчивости исходного цикла в фазовом пространстве рождается цикл удвоенного периода. При дальнейшем увеличении управляющего параметра может снова произойти бифуркация потери устойчивости, в результате которой появится цикл учетверенного (по сравнению с исходным) периода, и т. д. В частности, возможна бесконечная последовательность удвоений периода исходного предельного цикла. Эта последовательность бифуркаций происходит на конечном интервале изменения управляющего параметра и переводит систему от устойчивого периодического движения к хаотической динамике.

Рассмотрим это явление более подробно, иллюстрируя сказанное на примере системы Реслера [158, 277, 469]

$$\dot{x} = -(y + z), \quad \dot{y} = x + \frac{y}{5}, \quad \dot{z} = \frac{1}{5} + z(x - \mu). \quad (2.252)$$

Уравнения (2.252) описывают динамику абстрактной химической реакции [65, 340]. При определенном значении параметра $\mu = \mu_1$ из устойчивой точки посредством бифуркации Андронова–Хопфа рождается предельный цикл периода τ_1 (рис. 2.166а). При увеличении μ , $\mu > \mu_1$, этот цикл остается устойчивым, пока не будет достигнуто следующее бифуркационное значение $\mu = \mu_2$. В этот момент цикл периода τ_1 превращается в сложный «восьмеркой» устойчивый предельный цикл вдвое большего периода $\tau_2 = 2\tau_1$ (рис. 2.166б). Он замыкается после двух оборотов теряющего устойчивость цикла, а в спектре движения появляются кратные гармоники.

С дальнейшим увеличением параметра μ , $\mu > \mu_2$, при $\mu = \mu_m$, $m = 3, 4, \dots$ в системе будут происходить последовательные бифуркации удвоения, приводящие к возникновению устойчивого периодического движения, соответственно, с периодами $2^m \tau_1$, $m = 3, 4, \dots$ (рис. 2.166в, г).

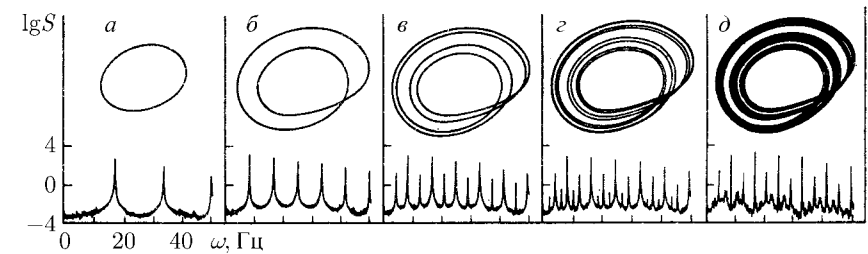


Рис. 2.166. Переход к странному аттрактору через последовательность бифуркаций удвоения периода предельного цикла в системе Реслера (2.252) при значениях μ , равных: 2,6 (а), 3,5 (б), 4,1 (в), 4,18 (г), 4,23 (д) [277]

Значения управляющего параметра $\mu = \mu_m$, при которых происходят очередные бифуркации удвоения, образуют сходящуюся последовательность:

$$\lim_{m \rightarrow \infty} \mu_m = \mu_\infty. \quad (2.253)$$

Когда $\mu = \mu_\infty$, предельный цикл достигает бесконечно большого периода, т. е. превращается в незамыкающуюся притягивающую фазовую траекторию, из которой при $\mu > \mu_\infty$ формируется странный аттрактор (рис. 2.166д). Динамика системы в этом случае характеризуется сплошным спектром и разбеганием близких фазовых кривых. Скорость сходимости бесконечной последовательности μ_m определяется универсальной постоянной — числом Фейгенбаума:

$$\lim_{m \rightarrow \infty} \frac{\mu_m - \mu_{m-1}}{\mu_{m+1} - \mu_m} = \delta = 4,6692 \dots \quad (2.254)$$

Описанному сценарию присуща универсальность: константа Фейгенбаума δ не зависит от конкретного вида динамической системы.

Переход к хаотическому режиму через последовательность бифуркаций удвоения периода наблюдается также для неавтономных систем. Так, в [363] рассматривалось уравнение движения маятника с вязким трением под действием периодической вынуждающей силы:

$$\ddot{\theta} + \beta \dot{\theta} + \omega_0^2 \sin \theta = \nu \cos \omega_1 t. \quad (2.255)$$

Оказывается, при определенном выборе интенсивности вынуждающей силы — так, чтобы маятник доходил до вертикального перевернутого положения ($\theta = \pi$), — с изменением частоты внешнего воздействия ω_1 наблюдается последовательность бифуркаций удвоения.

Аналогичное явление для движения частицы в потенциальном поле под действием периодической вынуждающей силы при наличии вязкого трения обнаружено в [362].

Сценарий развития хаоса Фейгенбаума хорошо подтверждается численными исследованиями. Последовательности бифуркаций удвоения периода найдены во многих системах и отображениях (см., например, [116, 118, 158, 450]), в том числе и в системе Лоренца при больших значениях параметра r [467]. Ряд физических, химических и многие другие эксперименты также обнаруживают эти бифуркации и некоторые признаки универсальности [92, 236, 321, 323, 393, 399, 470, 511].

Заметим, что в реальных экспериментах и численных расчетах, где всегда имеются физические шумы или ошибки округления, *бесконечную* последовательность бифуркаций удвоения наблюдать не удается. Вместо этого после нескольких бифуркаций удвоения движение сразу становится хаотическим. Влияние шумов на рассматриваемый сценарий перехода к хаосу обсуждается в [202, 278].

29.3 Перемежаемость

Кроме описанных, широко распространен механизм возникновения хаотических колебаний, лежащий в основе сценария И. Помо и П. Манневилля [351, 421, 460] и связанный с переходом к хаотическому движению через *перемежаемость*¹¹. При таком движении всплески хаотического поведения чередуются (перемежаются) с участками, на которых происходят почти периодические колебания (рис. 2.167).

Чтобы пояснить механизм, лежащий в основе перемежаемости, рассмотрим одномерное отображение вида

$$x_{n+1} = \varphi(\mu, x_n). \quad (2.256)$$

Будем считать, что график отображения (2.256) при определенном критическом значении параметра $\mu = \mu_c$ касается биссектрисы $x_{n+1} = x_n$. Для удобства предположим, что точкой касания является начало координат $x = 0$ (рис. 2.168). Разложим вблизи ее функцию φ в ряд по степеням x_n :

$$\varphi(\mu, x_n) = 0 + (\mu - \mu_c)\varphi'|_{x=0} + x_n\varphi'|_{x=0} + \frac{1}{2}x_n^2\varphi''|_{x=0} + \dots \quad (2.257)$$

Теперь, определив соответствующим образом параметр μ и величину x_n , значение $(1/2)\varphi''|_{x=0}$ можно обратить в единицу. Кроме того, легко видеть,

¹¹Отметим, что впервые об этом эффекте говорилось в [37, 136].

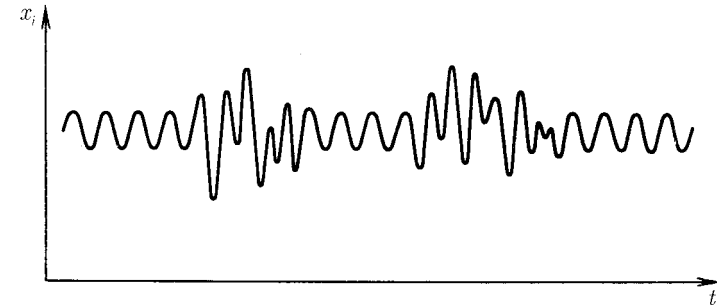


Рис. 2.167. Типичная временная зависимость динамических переменных x_i при переходе к хаосу через перемежаемость

что $\varphi'|_{x=0} = 1$. Следовательно, оставляя в разложении (2.257) члены не выше второго порядка малости, из (2.256) найдем

$$x_{n+1} = (\mu - \mu_c) + x_n + x_n^2. \quad (2.258)$$

Для такого отображения в случае $\mu < \mu_c$ существуют две неподвижные точки

$$x_{1,2}^* = \pm\sqrt{\mu_c - \mu}, \quad (2.259)$$

одна из которых (x_1^*) устойчива, а другая (x_2^*) неустойчива (рис. 2.168). Если $\mu = \mu_c$, точки x_1^* и x_2^* сливаются и при $\mu > \mu_c$ исчезают. Это соответствует случаю, когда действительное собственное значение отображения пересекает единичную окружность в точке $+1$ (см. раздел 28.2). При этом перестройка режимов движения в системе представляет собой кризис.

Когда значение параметра μ лишь немного превышает критическое, т. е. при $\mu - \mu_c \gtrsim 0$, расстояние между линией (2.258) и биссектрисой $x_{n+1} = x_n$ будет малым вблизи начала координат $x = 0$ (рис. 2.168). В этом случае нетрудно убедиться, что при x_n , близких к нулю, потребуется очень большое число итераций отображения (т. е. много шагов по дискретному времени n), чтобы покинуть окрестности точки $x = 0$.

Представим себе теперь, что в отображении φ есть также участок, порождающий сложную динамику. Пусть начальная точка x_0 такого отображения находится в интервале $[a, b]$, обозначенном на рис. 2.169. Тогда этому положению точки x_0 будет отвечать определенная последовательность x_0, x_1, x_2, \dots , порождаемая преобразованием φ . На некотором шаге точка x_k этой последовательности обязательно попадает в область вблизи максимума

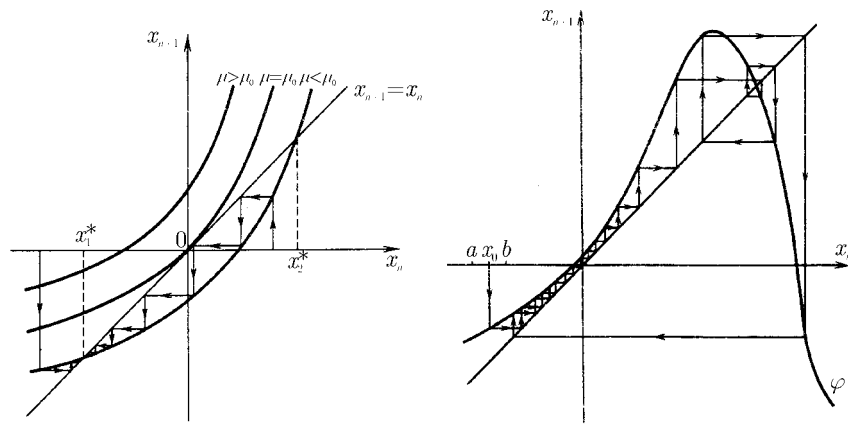


Рис. 2.168. Отображение (2.258) в окрестности начала координат

Рис. 2.169. Пример одномерного отображения, демонстрирующего режим с перемежаемостью

функции φ и, совершив здесь несколько нерегулярных колебаний, опять будет отброшена в область малых значений x , после чего процесс повторится (рис. 2.169). Получаемая в результате такого процесса последовательность значений x_0, x_1, x_2, \dots не будет содержать циклов никакого конечного периода.

Обратимся теперь к системам дифференциальных уравнений. Для того чтобы система могла обладать хаотическим режимом, обусловленным странным аттрактором, число описывающих ее динамических переменных должно быть больше или равно трем. Если в такой системе перейти к отображению Пуанкаре, то в этом отображении фазовые траектории будут представлены некоторым множеством точек, нерегулярным образом заполняющих области секущей поверхности. При сильном сжатии отображение Пуанкаре для многих систем оказывается близким к одномерному, и поведение фазовых траекторий может быть приближенно описано с помощью одномерного отображения. Когда изображающая точка такого одномерного отображения долгое время проводит в окрестности начала координат $x = 0$, для динамической системы это означает, что ее поведение будет почти периодическим.

Покинув область с почти периодическим поведением, фазовая точка системы попадает в другую область фазового пространства, характеризу-

ющуюся сложной динамикой. (Эта область соответствует области максимума отображения, как на рис. 2.169.) Проведя некоторое время там, фазовая точка выбрасывается из нее и вновь попадает в ту область фазового пространства, где происходят почти периодические колебания. При этом фазовая траектория динамической системы никогда не замыкается и не уходит на бесконечность.

Оценим длительность регулярных колебаний в системе. Для этого заметим, что выражение (2.258) при малых x_n (т. е. вблизи начала координат) можно просто заменить дифференциальным уравнением

$$\frac{dx}{dl} = (\mu - \mu_c) + x^2,$$

где l характеризует число итераций в регулярной области. Интегрирование этого уравнения дает

$$l = \frac{1}{\sqrt{\mu - \mu_c}} \operatorname{arctg} \left[\frac{x}{\sqrt{\mu - \mu_c}} \right] \Big|_{x_1}^{x_2}.$$

Теперь, чтобы найти среднюю длину $\langle l \rangle$ регулярной области допустим, что расстояние между точками x_1 и x_2 , расположенных по обе стороны от начала координат $x = 0$, много больше чем $\sqrt{\mu - \mu_c}$, но не настолько, чтобы нарушалось разложение (2.258). Тогда на этом интервале $\operatorname{arctg}(\cdot)$ практически не изменяется и

$$\langle l \rangle \sim \frac{1}{\sqrt{\mu - \mu_c}}. \tag{2.260}$$

Таким образом, при подходе к значению $\mu = \mu_c$ со стороны $\mu > \mu_c$ (рис. 2.168) длительность пребывания системы в области с нерегулярной динамикой существенно не меняется, тогда как время, проводимое в области с почти периодическими колебаниями, возрастает в соответствии с соотношением (2.260).

Описанный режим с перемежаемостью появляется путем слияния устойчивой и неустойчивой неподвижных точек, т. е. через касательную бифуркацию (рис. 2.168). Однако подобное поведение может возникать и еще двумя основными путями. Поэтому перемежаемость, которая образуется в результате касательной бифуркации, называют *перемежаемостью I рода*.

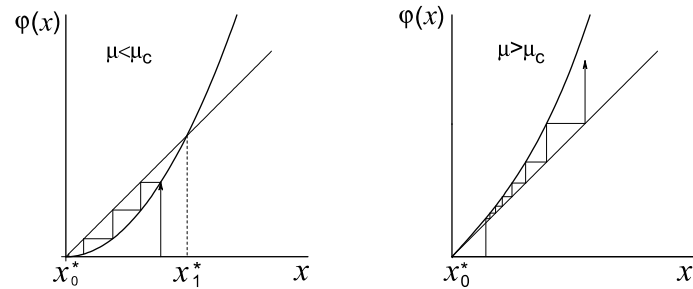


Рис. 2.170. Возникновение перемежаемости II рода

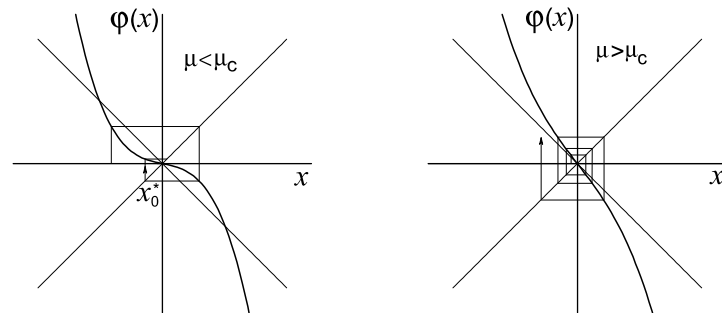


Рис. 2.171. Возникновение перемежаемости III рода

Если при увеличении параметра μ функция $\varphi(\mu, x)$ изменяется согласно рис. 2.170, то, как нетрудно видеть, также возникает режим с перемежаемостью. С переходом μ через бифуркационное значение μ_c устойчивая неподвижная точка $x_0^* = 0$ теряет свою устойчивость, сталкиваясь с неустойчивой неподвижной точкой x_1^* , но при этом не исчезает. Это соответствует случаю, когда два комплексно-сопряженных собственных значения отображения одновременно пересекают единичную окружность (см. раздел 28.2). Возникающий в результате этой перестройки режим движения называется *перемежаемостью II рода*.

Перемежаемость III рода возникает при изменении функции $\varphi(x)$ как показано на рис. 2.171. При этом действительное собственное значение отображения пересекает единичную окружность в точке -1 , так что первоначально устойчивая неподвижная точка x_0^* становится неустойчивой.

На рис. 2.170, 2.171 динамика систем, демонстрирующих перемежаемость II и III родов, показана вблизи точки $x = 0$ при помощи диаграмм

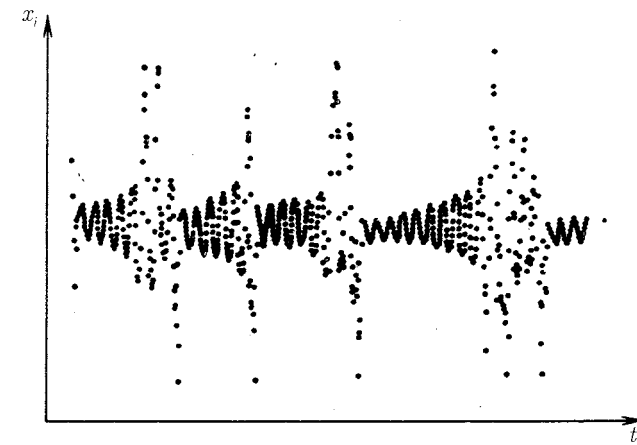


Рис. 2.172. Режим с перемежаемостью, обнаруженный в простой радиотехнической системе [361]

мы Ламерея. Нетрудно видеть, что в обоих случаях должны наблюдаться длительные регулярные колебания, пока изображающая точка находится в области начала координат. При выходе из этой области система может демонстрировать хаотические всплески.

Таким образом, при увеличении управляющего параметра, в зависимости от вида динамической системы, допустимы три вида перемежаемости: I, II и III родов.

Развитие хаоса по сценарию Помо–Манневилля наблюдалось и численно, и экспериментально. Так, исследования квадратичного отображения подтверждают этот сценарий [202, 427]. Система Лоренца в некотором узком интервале параметров также демонстрирует переход к хаосу через перемежаемость [420].

В другой работе [361] авторы изучали явление перемежаемости на примере простой радиотехнической системы. Результаты наблюдений выводились на осциллограф (рис. 2.172). Хорошо видно, что регулярные участки перемежаются с хаотическими.

Аналогичный переход к хаосу наблюдался во многих экспериментах, в частности, в исследованиях конвекции Рэля–Бенара [232], химической реакции Белоусова–Жаботинского [61, 470] и др. [217, 320, 422, 532].

29.3.1 Ренормализационный анализ перемежаемости

Перемежаемость также поддается описанию в рамках ренормализационного подхода. В отличие от ренормгруппового анализа сценария Фейгенбаума и Рюэля–Такенса, режим с перемежаемостью допускает *точное* решение функционального уравнения.

Коротко остановимся на теории ренормализации перемежаемости, следуя [201].

Обобщим функцию φ , определяемую как (2.258) таким образом, чтобы при $x \rightarrow 0$ она имела вид

$$\varphi = x + b|x|^z \quad (2.261)$$

(рис. 2.173 а). Нетрудно убедиться в том, что в интервале $x \in [0, 1]$ функция $\varphi(\varphi(x)) = \varphi^2(x)$ после соответствующего изменения масштаба демонстрирует то же поведение, что и $\varphi(x)$. Действительно, вторая итерация, $\varphi^2(x)$, будет выглядеть как рис. 2.173 б. В свою очередь, из рис. 2.173 следует, что в квадрате, обозначенном пунктиром, $\varphi^2(x)$ оказывается подобной исходной функции $\varphi(x)$. Поэтому снова, как это было сделано при изучении универсальности Фейгенбаума, можно записать

$$\varphi_{k+1}(x) = \mathcal{T}\varphi_k(x) = \alpha\varphi_k\left(\frac{x}{\alpha}\right).$$

При $k \rightarrow \infty$ последовательность функций φ_k сходится к неподвижной точке $g(x)$, т. е.

$$g(x) = \alpha g\left(g\left(\frac{x}{\alpha}\right)\right), \quad (2.262)$$

или

$$g(\alpha x) = \alpha g(g(x)).$$

Граничные условия, однако, для этого случая будут иными:

$$\begin{aligned} g(0) &= 0, \\ \left. \frac{dg}{dx} \right|_{x=0} &= 1. \end{aligned} \quad (2.263)$$

Метод решения функционального уравнения (2.262), предложенный в работе [359], заключается в том, что исходное отображение $x' = \varphi(x)$ представляется в неявном виде:

$$F(x') = F(x) - a,$$

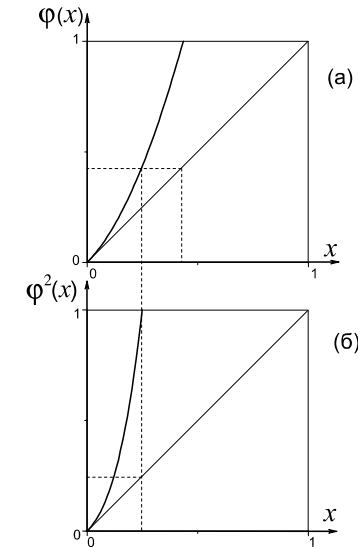


Рис. 2.173. Графики $\varphi(x)$ и $\varphi^2(x)$ обобщенной функции (2.261)

$x'(x) = F^{-1}[F(x) - a] = \varphi(x)$, где a — некоторый параметр. Тогда уравнение (2.262) преобразуется к виду $\alpha x''(x) = x'(\alpha x)$, где x'' означает второй шаг итерации, $x'' = \varphi(\varphi(x))$. Следовательно, после действия оператора F получим:

$$F(\alpha x'') = F[x'(\alpha x)] = F(\alpha x) - a. \quad (2.264)$$

Теперь, поскольку $F(x'') = F(x') - a = F(x) - 2a$, приходим к выражению:

$$\frac{1}{2}F(x'') = \frac{1}{2}F(x) - a. \quad (2.265)$$

Рассмотрим соотношения (2.264) и (2.265). Они показывают, что для того, чтобы F соответствовала уравнению (2.262), оператор должен подчиняться равенству $1/2F^*(x) = F^*(\alpha x)$. Учтеть его позволит следующий выбор F и α : $F^*(x) = |x|^{-(z-1)}$, $\alpha = 2^{1/(z-1)}$. Отсюда находим выражение для $g(x)$:

$$g(x) = (F^*)^{-1}[F^*(x) - a] = \left[|x|^{-(z-1)} - a\right]^{-\frac{1}{z-1}}.$$

Легко видеть, что при $a = b(z-1)$ (см. (2.261)) эта функция будет удовлетворять граничным условиям (2.263).

Таким образом, в зависимости от показателя степени z , универсальная постоянная α будет иметь различное значение.

Нельзя утверждать, что мы привели исчерпывающий перечень всех возможных сценариев развития хаоса. Известны другие, менее распространенные сценарии, связанные с жестким возникновением странного аттрактора из двоякоасимптотических траекторий [28, 35, 158] (для системы Лоренца такой сценарий реализуется при $\sigma = 10$, $b = 8/3$, $10 < r < 24,74$), с последовательностью удвоения и дальнейшим разрушением двумерных торов [10, 19, 215, 374], с утроением, учетверением и т. д. предельных циклов [28, 70, 360], сценарий перехода к хаосу с перемежаемостью типа «тор–хаос» [283, 479] и др. Мы рассказали лишь о тех путях развития хаоса, которые достаточно просты и часто наблюдаются в экспериментах и при численном моделировании.

В заключение отметим, что уравнение (2.262) по виду совпадает с функциональным уравнением Фейгенбаума–Цветановича (2.230). Это говорит о том, что два таких непохожих сценария перехода к хаосу описываются в рамках одного и того же формализма. Такое совпадение имеет глубокие причины. В частности, в логистическом отображении (2.222) появлению цикла периода 3 при $\mu = 1 + \sqrt{8}$ предшествует режим с перемежаемостью. С превышением критического значения этот цикл претерпевает обычные бифуркации удвоения периода (рис. 2.122).

30. Подавление хаоса и управление динамическими системами

Из всего изложенного материала данной главы следует, что динамический хаос оказывается типичным явлением. Хаотические свойства могут проявлять самые разнообразные системы, и если хаос не обнаруживается, то, возможно, лишь потому, что либо он возникает в очень малых областях параметрического пространства, либо при нефизических значениях параметров.

В связи с этим не так давно стало интенсивно развиваться новое направление в нелинейной динамике, посвященное проблемам предсказуемости поведения хаотических систем, управления их динамикой и возможности подавления хаоса. Теоретические и экспериментальные работы в этой области выявили одно неожиданное и вместе с тем замечательное свойство хаотических динамических систем: они являются весьма податливыми к внешним воздействиям. Таким образом, была обнаружена возможность *управлять* динамикой хаотических систем, т. е. посредством *достаточно*

слабых воздействий переводить первоначально хаотические системы на *требуемый* динамический режим и тем самым стабилизировать их поведение. Более того, было показано, что для распределенных сред внешнее воздействие при некоторых условиях приводит к рождению сложных пространственно протяженных структур с заданными свойствами.

Под стабилизацией неустойчивого или хаотического поведения обычно подразумевается искусственное создание в изучаемой системе *устойчивых* (как правило, периодических) колебаний посредством внешних мультипликативных или аддитивных воздействий. Иными словами, для стабилизации необходимо найти такие возмущения, которые вывели бы систему из хаотического режима на регулярный. При внешней простоте формулировки этой проблемы ее решение для ряда динамических систем оказывается достаточно сложной задачей. Более того, хотя в настоящее время имеется большое число работ, посвященных этому вопросу (см. [120, 238, 257, 260, 398, 400, 445, 486, 487, 492]), развить последовательную теорию и более или менее *строго* обосновать возможность стабилизации хаотического поведения удалось пока только для достаточно общих *семейств* динамических систем.

Стабилизация хаотического поведения может быть осуществлена двумя различными способами. Первый из них обеспечивает выведение системы из хаотического на регулярный режим посредством внешних возмущений, реализованных *без* обратной связи. Другими словами, этот метод не учитывает текущее состояние динамических переменных системы. Качественно отличный от данного метод реализуется посредством корректирующего воздействия в соответствии с требуемым значением динамических переменных и, таким образом, *вовлекает* обратную связь как необходимую компоненту динамической системы. По установившемуся соглашению, первый способ стабилизации хаотической динамики называется *подавлением хаоса*, или контролем (иногда управлением или регулированием) хаотической динамики без обратной связи. Второй способ носит название *контролирование хаоса* с обратной связью (controlling chaos). В свою очередь, реализация каждого из этих методов может быть проведена параметрическим или силовым способами.

Введение обратной связи является определенным преимуществом, поскольку в большинстве случаев такой способ управления приводит к требуемому результату: выбранный заранее седловой предельный цикл стабилизируется и, таким образом, исследуемая система выводится на предписанный режим движения. Однако этот метод эффективен, если только изображающая точка находится вблизи выбранного цикла. В противном случае необходимо использовать дополнительные способы воздействия [388, 488–490].

В то же время, методы без обратной связи не требуют введения постоянного компьютерного слежения за состоянием системы и менее подвержены воздействиям шумов, что существенно упрощает их использование в приложениях.

Гипотеза о стабилизации хаотических колебаний *параметрическим* образом в области значений параметров A_c , отвечающей существованию только хаотического поведения (так чтобы можно было говорить именно о *подавлении хаоса*) впервые была выдвинута и численно получила подтверждение в публикациях [4, 5] (см. также обзор [400]), где был рассмотрен класс дифференциальных уравнений с неполиномиальной правой частью. Впоследствии метод подавления хаоса (без обратной связи) был аналитически обоснован в работах [123, 408]. Однако данное направление получило широкое распространение после публикаций работ группы из Мэриленда [446, 468], где было показано, что при помощи достаточно слабых параметрических возмущений возможно стабилизировать практически любой седловой предельный цикл, вложенный в хаотический аттрактор. Эти результаты стимулировали изучение вопросов стабилизации хаотического поведения и вызвали большой интерес к вопросам управления неустойчивыми системами.

Поскольку подавление хаоса и управление динамическими системами подразумевает введение внешнего возмущения, рассмотрим сначала общие свойства динамических систем, подверженных внешним воздействиям.

30.1 Системы с внешними возмущениями и их свойства

Предположим, что рассматривается динамическая система с некоторым возмущением:

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, a), \quad (2.266)$$

$\mathbf{x} = \{x_1, x_2, \dots, x_n\}$, $\mathbf{v} = \{v_1, v_2, \dots, v_n\}$, $\mathbf{x}(0) \equiv \mathbf{x}_0$, где a — параметр. Если такое возмущение реализуется посредством мультипликативного воздействия по отношению к динамическим переменным x_i , то говорят, что имеет место *параметрическое* (или *мультипликативное*) управление, поскольку, как правило, параметры мультипликативно включаются в динамическую систему. В этом случае регулирование состоит в такой модификации функции \mathbf{v} в соотношении (2.266), чтобы новая система $\dot{\mathbf{x}} = \mathbf{v}'(\mathbf{x}, a', t)$ имела бы требуемое (выбранное заранее) поведение. Здесь $\mathbf{v}'(\mathbf{x}, a', t) = \mathbf{v}(\mathbf{x}, a_0 + a_1(t))$ и параметр $a_1(t)$ обычно является T — периодической функцией времени. Для параметрических возмущений *с обратной связью* учитывается текущее состояние системы, $\mathbf{v}'(\mathbf{x}, a', t) = \mathbf{v}(\mathbf{x}, a(\mathbf{x}(t)))$, так

что параметр a изменяется специальным путем, и необязательно периодически.

Достаточно часто в приложениях встречается ситуация, когда мультипликативное введение внешних возмущений в систему невозможно. Тогда фазовый поток $\mathbf{F}^t(\mathbf{x}, G)$ (см. §18) разлагается на две составляющие: часть, соответствующую невозмущенному фазовому потоку, $\mathbf{F}^t(\mathbf{x})$, и компоненту $\mathbf{F}^t(G)$, которая инициируется исключительно возмущениями, $\mathbf{F}^t(\mathbf{x}, G) = \mathbf{F}^t(\mathbf{x}) + \mathbf{F}^t(G)$. В этом случае имеет место аддитивное возмущение, т. е. $\mathbf{v}'(\mathbf{x}, a', t) = \mathbf{v}(\mathbf{x}, a) + \mathbf{g}(t)$, где $\mathbf{g}(t)$ обозначает внешнее воздействие. Таким образом, управление динамикой системы подразумевает приложение силовой компоненты к векторной функции. Поэтому данный тип управления поведением динамической системы называется *силовым*. В свою очередь, если в силовом контроле учитывается обратная связь, то функция \mathbf{v} модифицируется как $v'_i = v_i(\mathbf{x}, a) + g_k(x_i(t))$, $i = 1, 2, \dots, n$, $1 \leq k \leq n$.

По ряду причин параметрический метод имеет определенные преимущества перед силовым. Во-первых, в приложениях к физическим, химическим, биологическим и другим важным системам часто рассматриваются величины, которые являются пропорциональными динамическим переменным x_i . Для таких систем $\mathbf{v}(0, a_1, \dots, a_m) = 0$, а гиперповерхности x_i являются инвариантными. Это означает, что система (2.266) отражает реальные процессы только на симплексе $\mathbf{X} = \left\{ \mathbf{x} \mid x_i > 0, \sum_{i=1}^n x_i < \text{const} \right\}$. В этом случае внешнее аддитивное воздействие может привести к тому, что фазовые траектории покинут множество \mathbf{X} , пересекая гиперповерхности $x_j = 0$. Поэтому силовое воздействие часто является причиной вырождения системы или выхода ее на нежелательный режим эволюции. Например, для биологических систем это означает вымирание части особей. В то же время, параметрическое воздействие означает изменение ресурсов системы и, таким образом, является более мягким по сравнению с силовым.

Во-вторых, силовое возмущение гораздо труднее реализовать. Так, для химических систем силовой контроль подразумевает введение (и соответственно удаление) дополнительных веществ; для биологических систем такой метод может быть реализован через стерилизацию части особей или введением в сообщество дополнительных видов.

С другой стороны, в противоположность параметрическому воздействию, силовой метод, как правило, приводит к требуемому результату для почти всех систем, поскольку во многих случаях ее естественное поведение может быть буквально «задавлено» внешней силой.

Формально проблема управления поведением системы (2.266) заключается в том, чтобы найти такое внешнее возмущение G , при котором фа-

зовый поток $F^t(x, G)$, порождаемый возмущенной динамической системой $\dot{x} = v'(x, a, G)$, стремился бы к выбранному подмножеству $X(G)$ ее фазового пространства. Вследствие воздействия G подмножество $X(G)$ может как преобразоваться в аттрактор так и остаться неустойчивым множеством. В последнем случае возмущения G модифицируют систему (2.266) таким образом, что фазовые траектории подходят к подмножеству $X(G)$ и остаются в достаточно малой его окрестности $U \supset X(G)$ под действием G . Как правило, в приложениях в качестве подмножества $X(G)$ выбирается неустойчивый цикл определенного периода.

Для того, чтобы выявить некоторые общие свойства системы (2.266) с внешними возмущениями, рассмотрим отображение, порождаемое этой системой:

$$x_{n+1} = f(x_n, a), \quad (2.267)$$

где $a \in A$, $f = \{f_1, \dots, f_s\}$, $x = \{x_1, \dots, x_s\} \in M$. Определим возмущение G , действующее на множестве допустимых значений параметров A как

$$a_{n+1} = g(a_n), \quad a_n \in A. \quad (2.268)$$

Тогда результирующее возмущенное отображение можно записать в виде

$$\begin{aligned} x_{n+1} &= f(x_n, a_n), & x &\in M, \\ a_{n+1} &= g(a_n), & a_n &\in A. \end{aligned} \quad (2.269)$$

Далее ограничимся только *периодическими* возмущениями периода τ , т. е. $a_{i+1} = g(a_i)$, $i = 1, 2, \dots, \tau - 1$, $a_1 = g(a_\tau)$, $a_i \neq a_j$, $i \neq j$.

Теперь, анализируя отображение (2.269), легко понять, что период t любого цикла периодически возмущаемого отображения (2.267) определяется из соотношения: $t = \tau k$, где τ — период возмущения и k — положительное целое. Действительно, если возмущенное отображение имеет t -периодический цикл, то последовательность координат точек, которые его формируют, также является периодической с периодом t . Но последовательность параметров a по определению периодична с периодом τ . Поэтому всегда $t = \tau k$, где k — целое.

Здесь, однако, необходимо сделать одно существенное замечание. Если спроектировать полученный t -периодический цикл на пространство M (т. е. рассмотреть систему (2.269) как неавтономную), то возможно получить цикл, который не может быть назван циклом в обычном понимании. Причина состоит в том, что точки цикла, которые отличаются друг от друга только в значении координаты a (если они существуют), спроектируются в

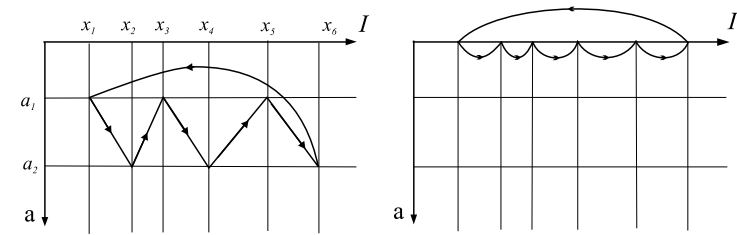


Рис. 2.174. Цикл возмущенного одномерного отображения в пространстве (x, a) и его проекция на исходное пространство I

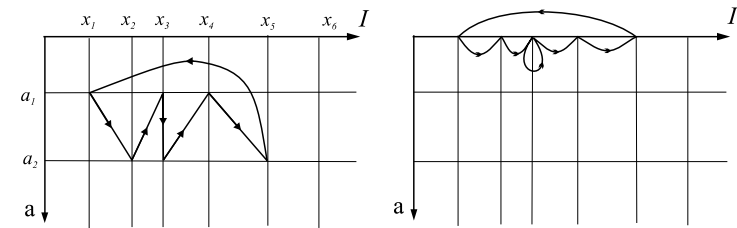


Рис. 2.175. Цикл возмущенного одномерного отображения с совпадающими x -координатами в пространстве (x, a) и его проекция

одну и ту же точку пространства M . Поэтому изображающая точка возмущенного отображения будет по несколько раз попадать в некоторые точки, формирующие цикл.

Например, для одномерных отображений, $s = 1$, в общем случае в проекции на исходное пространство $M = I$ получится цикл периода $k\tau$ (рис. 2.174).

Однако в I возможно получить цикл с совпадающими x -координатами, когда $x_i = x_m$, $a_i \neq a_m$, $i \neq m$, где (x_i, a_i) и (x_m, a_m) — точки цикла отображения (2.269). В этом случае на координатной оси получится $(P-l)$ точек (см. рис. 2.175), где l — число совпадений. В частности, при $P = 2$ ($\tau = 2$) вполне возможно в проекции наблюдать только одну фиксированную точку. Для $P > 2$ вероятно появление более экзотических циклов. Описанная ситуация, однако, не является случаем общего положения, и, как правило, встречается только при специально подобранных возмущениях.

Введение τ -циклического преобразования (2.268) в отображение (2.267)

означает, что систему (2.269) можно записать как

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{f}(\mathbf{x}_n, a_1) \equiv \mathbf{f}_1, \\ \mathbf{x}_{n+2} &= \mathbf{f}(\mathbf{x}_{n+1}, a_2) \equiv \mathbf{f}_2, \\ &\dots\dots\dots, \\ \mathbf{x}_{n+\tau} &= \mathbf{f}(\mathbf{x}_{n+\tau-1}, a_\tau) \equiv \mathbf{f}_\tau. \end{aligned} \quad (2.270)$$

Введем в рассмотрение τ функций следующего вида:

$$\begin{aligned} \mathbf{h}_1 &= \mathbf{f}_\tau(\mathbf{f}_{\tau-1}(\dots \mathbf{f}_2(\mathbf{f}_1(\mathbf{x})) \dots)), \\ \mathbf{h}_2 &= \mathbf{f}_1(\mathbf{f}_\tau(\mathbf{f}_{\tau-1}(\dots \mathbf{f}_3(\mathbf{f}_2(\mathbf{x})) \dots))), \\ &\dots\dots\dots, \\ \mathbf{h}_\tau &= \mathbf{f}_{\tau-1}(\mathbf{f}_{\tau-2}(\dots \mathbf{f}_1(\mathbf{f}_\tau(\mathbf{x})) \dots)), \end{aligned}$$

где $\mathbf{x} = \{x_1, \dots, x_s\}$ и $\mathbf{f}_i = \{f_i^{(1)}, \dots, f_i^{(s)}\}$, $\mathbf{h}_i = \{h_i^{(1)}, \dots, h_i^{(s)}\}$, $i = 1, 2, \dots, \tau$. Тогда возмущенное отображение (2.269) представится как преобразование

$$\mathbf{x}_{\tau n+(j-1)} = \mathbf{h}_j(\mathbf{x}_{\tau(n-1)+(j-1)}, a_1, \dots, a_\tau), \quad (2.271)$$

$j = 1, \dots, \tau$, для которого начальные условия определяются следующим образом:

$$\mathbf{x}_1 = \mathbf{f}_1(\mathbf{x}_0), \quad \mathbf{x}_2 = \mathbf{f}_2(\mathbf{x}_1), \quad \dots, \quad \mathbf{x}_{\tau-1} = \mathbf{f}_{\tau-1}(\mathbf{x}_{\tau-2}). \quad (2.272)$$

Теперь легко показать [123], что если отображение, порождаемое функцией \mathbf{h}_k , $1 \leq k \leq \tau$, имеет цикл p периода t и функции $\mathbf{f}_k(\mathbf{x})$ являются непрерывными, тогда отображение, порождаемое функцией \mathbf{h}_m , $m = k + 1 \pmod{\tau}$, также имеет цикл p' того же периода t . Более того, если цикл p устойчив, то и цикл p' также будет устойчивым.

Основной смысл описанного построения (2.270), (2.271) заключается в том, что исследование отображения с периодическим возмущением можно существенно упростить. Вместо исходной неавтономной системы (2.269) достаточно рассмотреть одно из автономных отображений, определяемое выражением (2.271). Таким образом, вся динамика исходного отображения (2.269) будет задаваться совокупностью отображений (2.271), которые действуют независимо друг от друга и связаны лишь начальными условиями (2.272). Это справедливо для любого множества допустимых значений A параметра a динамической системы (2.267) с периодическим возмущением (2.268).

30.2 Силовое возбуждение

В работах [364, 365, 411] для управления поведением динамических систем было предложено использовать так называемое *силовое резонансное* воздействие. Хотя такое воздействие можно реализовать далеко не во всех случаях, его достоинство состоит в том, что оно применимо не только к хаотическим системам.

Для достижения возможности контроля посредством резонансных возбуждений в динамическую систему, находящуюся в хаотическом режиме, необходимо аддитивно включить внешнее возмущение $\mathbf{F}(t)$:

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, a) + \mathbf{F}(t). \quad (2.273)$$

Далее, пусть требуемая динамика задается функцией $\mathbf{y}(t)$, которая удовлетворяет так называемому уравнению предписанного движения:

$$\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y}). \quad (2.274)$$

Теперь, выбирая возмущение в виде $\mathbf{F} = \mathbf{g}(\mathbf{y}(t)) - \mathbf{v}(\mathbf{y}(t), a)$ и подставляя его в (2.273), получим уравнение контролирования:

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, a) + \mathbf{g}(\mathbf{y}) - \mathbf{v}(\mathbf{y}, a). \quad (2.275)$$

Таким образом, если устремить $\mathbf{x} \rightarrow \mathbf{y}$ при $t \rightarrow \infty$, то в конечном счете динамика будет представлена уравнением (2.274).

Отличительной чертой метода управления при помощи резонансных возбуждений является тот факт, что его практическое применение не ограничивается только хаотическими системами. Однако это возможно далеко не всегда и существуют начальные условия, для которых поведение не будет задаваться функцией $\mathbf{y}(t)$. Кроме того, этот метод контроля сильно зависит от знания динамики системы, и малые ошибки в модели (2.273) могут расти вследствие возмущения $\mathbf{F}(t)$. Тем не менее, можно усовершенствовать контролирование (2.273)–(2.275). Это приведет к большей эффективности метода резонансных возбуждений [411].

30.3 Метод Отга–Гребоджи–Йорка

Известный параметрический метод с обратной связью, предложенный в работах [446, 468] и получивший широкое продолжение во многих других публикациях (ссылки см. в [120, 238, 260, 398, 445, 486, 487, 492]), основывается на предположении, что параметры a_i системы могут быть преобразованы в неявные (зависящие от динамических переменных \mathbf{x}) функции времени.

Этот метод позволяет достаточно легко стабилизировать выбранный цикл, вложенный в хаотический аттрактор. В общем случае такой цикл является седловым. В отображении Пуанкаре¹² это соответствует седловой точке \mathbf{x}^* (рис. 2.176). В ее достаточно малой окрестности поведение системы представляется неустойчивыми траекториями (штриховые линии на рис. 2.176): фазовая точка p_1 сначала стремится к \mathbf{x}^* , повиная устойчивому направлению (движение к точке p_2), но затем на нее начнет сильно влиять неустойчивое направление (далее по штриховой линии), и система покидает окрестность седла \mathbf{x}^* . Однако, зная динамику системы, можно подобрать такие возмущения параметра, чтобы точка p_1 двигалась по направлению к устойчивому многообразию. Очевидно, после введения возмущения точка \mathbf{x}^* сместится. Это состояние обозначено на рис. 2.176 как \mathbf{x}_1^* . Следовательно, шаг за шагом выбирая возмущения так, чтобы p_i оставалась вблизи устойчивой сепаратрисы седла, можно добиться практически устойчивого поведения.

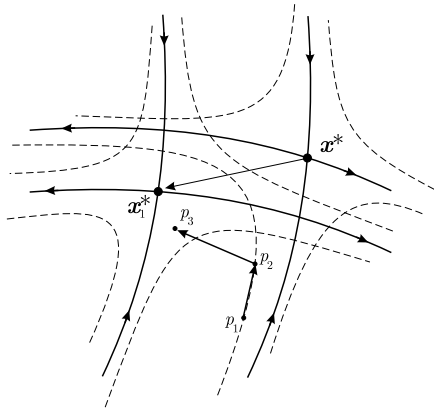


Рис. 2.176. Геометрическая иллюстрация метода Отта–Гребоджи–Йорка.

Формально представленный метод можно описать следующим образом. Предположим, что вблизи седлового предельного цикла, который нужно стабилизировать, система задается отображением Пуанкаре $\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n, a)$. Допустим, что при $a = a_0$ это отображение имеет неподвижную точку \mathbf{x}^* , т. е. $\mathbf{f}(\mathbf{x}^*, a_0) = \mathbf{x}^*$, отвечающую данному циклу. В окрестности \mathbf{x}^* для значений параметра a , близких к a_0 , поведение отображения дается

¹²Для сложного цикла, имеющего несколько оборотов, рассматривается соответствующая итерация отображения.

линейным преобразованием

$$\mathbf{x}_{n+1} - \mathbf{x}^* = \hat{A}(\mathbf{x}_n - \mathbf{x}^*) + \hat{B}(a - a_0), \quad (2.276)$$

где \hat{A} — k -мерная матрица Якоби, \hat{B} — k -мерный вектор-столбец, $\hat{A} = \partial \mathbf{f} / \partial \mathbf{x}$, $\hat{B} = \partial \mathbf{f} / \partial a$ при $\mathbf{x} = \mathbf{x}^*$ и $a = a_0$. Если от итерации к итерации параметр a изменяется, то, определяя \mathbf{x}_n через линейное отображение (2.276), можно задать подходящее малое отклонение в значении a от номинального a_0 . В линейном приближении это изменение параметра можно записать в виде

$$a_n - a_0 = -\hat{L}^T(\mathbf{x}_n - \mathbf{x}^*),$$

где \hat{L} — k -мерный вектор-столбец и T означает операцию транспонирования. Следовательно, из (2.276) находим, что

$$\delta \mathbf{x}_{n+1} = (\hat{A} - \hat{B}\hat{L}^T) \delta \mathbf{x}_n,$$

где $\delta \mathbf{x}_n = \mathbf{x}_n - \mathbf{x}^*$. Таким образом, неподвижная точка \mathbf{x}^* будет стабилизирована, если определить \hat{L} так, чтобы матрица $(\hat{A} - \hat{B}\hat{L}^T)$ имела собственные значения по модулю меньше единицы.

Очевидно, возмущение параметра a вблизи его номинального значения не должно быть слишком большим. Максимально допустимое отклонение дается выражением $|\hat{L}^T(\mathbf{x}_n - \mathbf{x}^*)|$. Кроме того, процедура стабилизации неустойчивых циклов является эффективной, когда фазовая траектория близка к нужному циклу. Но если она проходит вдали от требуемого положения, то может пройти достаточно долгое время, прежде чем контролируемое окажется возможным. Если аттрактор эргодический, то практически любая окрестность оказывается достижимой. Однако когда аттрактор системы не эргодический и, например, включает устойчивые предельные циклы (т. е. является квазиаттрактором), то этот метод может быть применен только для стабилизации *некоторых* траекторий. Для преодоления этих трудностей было предложено использовать различные процедуры [388, 488–490], позволившие по-новому подойти к проблеме стабилизации неустойчивых циклов, а также разработать другие способы контроля хаотических динамических систем (см. [238]).

Хотя описанные методы могут быть использованы достаточно широко (от стабилизации поведения механических систем до управления сокращениями сердечной мышцы), основной их недостаток сводится к тому, что, применяя их на практике, необходимо не только каждый раз *задавать положение* изображающей точки (что не всегда возможно), но и учитывать

уровень шума, поскольку они оказываются весьма податливы к влиянию шумовых факторов. Чтобы избежать этих трудностей, нужно *исключить* обратную связь. Используя силовое воздействие, в достаточно общем виде такую задачу можно решить [121], опираясь на теорию Мельникова (см. раздел 21.5).

30.4 Подавление хаоса в окрестности сепаратрисы¹³

Исследуем проблему стабилизации поведения систем с сепаратрисными контурами. Допустим, что исходные уравнения имеют вид

$$\dot{\mathbf{x}} = \mathbf{f}_0(\mathbf{x}) + \varepsilon \mathbf{f}_1(\mathbf{x}, t), \quad (2.277)$$

где $\mathbf{f}_0(\mathbf{x}) = (f_{01}, f_{02})$, $\mathbf{f}_1 = (f_{11}, f_{12})$. Для этой системы мельниковскую функцию $D(t, t_0)$ можно записать следующим образом:

$$D(t, t_0) = - \int_{-\infty}^{\infty} \mathbf{f}_0 \wedge \mathbf{f}_1 dt \equiv I[g(\mathbf{x}, t)]. \quad (2.278)$$

Предположим, что величина $D(t, t_0)$ знакопеременна, т. е. сепаратрисы пересекаются. Найдем такое внешнее возмущение $\mathbf{f}^*(\omega, t)$, приложение которого к системе приводит к ситуации, когда сепаратрисы не могут пересекаться, т. е. динамика стабилизируется¹⁴:

$$\dot{\mathbf{x}} = \mathbf{f}_0(\mathbf{x}) + \varepsilon [\mathbf{f}_1(\mathbf{x}, t) + \mathbf{f}^*(\omega, t)], \quad (2.279)$$

где $\mathbf{f}^* = (f_1^*, f_2^*)$. Обозначим отрезок, на котором функция $D(t, t_0)$ меняет знак, как $[s_1, s_2]$. После приложения внешнего возмущения $\mathbf{f}^*(\omega, t)$ возможны два варианта:

$$D^*(t, t_0) > s_2 \quad (2.280)$$

или

$$D^*(t, t_0) < s_1, \quad (2.281)$$

где $D^*(t, t_0)$ — мельниковское расстояние для системы (2.279). Остановимся на случае (2.280). При выполнении неравенства (2.281) анализ проводится аналогично. Тогда

$$I[g(\mathbf{x}, t)] + I[g^*(\omega, \mathbf{x}, t)] > s_2. \quad (2.282)$$

¹³Излагаемые ниже результаты получены одним из авторов (А.Ю.Л.) совместно с А.Р. Джаноевым.

¹⁴Можно условно назвать такое возмущение \mathbf{f}^* «регуляризирующим» или «стабилизирующим».

Здесь $I[g^*(\omega, \mathbf{x}, t)] = - \int_{-\infty}^{+\infty} \mathbf{f}_0 \wedge \mathbf{f}^* dt$. Выражение (2.282) верно для всех значений левой части неравенства, больших s_2 . Поэтому всегда найдётся такое χ , что $I[g(\mathbf{x}, t)] + I[g^*(\omega, \mathbf{x}, t)] = s_2 + \chi = \text{const}$, где χ, s_2 — положительные. Отсюда

$$I[g^*(\omega, \mathbf{x}, t)] = \text{const} - I[g(\mathbf{x}, t)]. \quad (2.283)$$

С другой стороны,

$$I[g^*(\omega, \mathbf{x}, t)] = - \int_{-\infty}^{\infty} \mathbf{f}_0 \wedge \mathbf{f}^* dt. \quad (2.284)$$

Выберем $\mathbf{f}^*(\omega, t)$ из класса функций, абсолютно интегрируемых на бесконечном участке и разложимых в интеграл Фурье. Примем $\mathbf{f}^*(\omega, t) = \text{Re}\{\hat{A}(t)e^{-i\omega t}\}$, где $\hat{A}(t) = (A(t), A(t))$. Поэтому $-\int_{-\infty}^{\infty} \mathbf{f}_0 \wedge \{\hat{A}(t)e^{-i\omega t}\} dt = \text{const} - I[g(\mathbf{x}, t)]$. Обратное Фурье-преобразование дает:

$$\mathbf{f}_0 \wedge \hat{A}(t) = \int_{-\infty}^{\infty} (I[g(\mathbf{x}, t)] - \text{const}) e^{i\omega t} d\omega.$$

Следовательно,

$$A(t) = \frac{1}{f_{01} - f_{02}} \int_{-\infty}^{\infty} (I[g(\mathbf{x}, t)] - \text{const}) e^{i\omega t} d\omega.$$

Величина $A(t)$ может быть интерпретирована как амплитуда «регуляризирующего» возмущения.

Таким образом, для систем, представимых в виде (2.279), внешнее стабилизирующее возмущение имеет вид:

$$f_{1,2}^*(\omega, t) = \text{Re} \left[\frac{e^{-i\omega t}}{f_{01} - f_{02}} \int_{-\infty}^{\infty} (I[g(\mathbf{x}, t)] - \text{const}) e^{i\omega t} d\omega \right].$$

Если в качестве примера рассмотреть нелинейный маятник с диссипацией, то для него стабилизирующая функция будет иметь смысл серии ударов, между которыми движение является свободным. Такое внешнее воздействие может быть найдено в явном виде [121].

30.5 Резонансная и высокочастотная стабилизация

Другой метод подавления хаоса без обратной связи, в котором также используется теория Мельникова, подразумевает резонансную и высокочастотную стабилизации [257, 313, 386, 397, 457].

В качестве примера использования этой процедуры рассмотрим классический нелинейный осциллятор Дюффинга (2.163), который подробно описан в разделе 21.5. Перепишем это уравнение в более удобной форме:

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= x - x^3 - \delta y + \gamma \cos \omega t.\end{aligned}\tag{2.285}$$

Как показано в п. 21.5, при $H_0 = 0$ в такой системе имеется единственная гиперболическая точка $x = y = 0$ с сепаратрисной петлей (рис. 2.90а). Возмущения разрушают эту петлю, что приводит к возникновению хаоса.

Нетрудно найти условие пересечения сепаратрис [116]. В самом деле, из представления (2.277) и (2.285) получим:

$$\begin{aligned}f_{01} &= y, & f_{11} &= 0, \\ f_{02} &= x - x^3, & f_{12} &= \gamma \cos \omega t - \delta y.\end{aligned}$$

Поэтому (см. (2.162))

$$D = - \int_{-\infty}^{+\infty} [\gamma y_0(t - t_0) \cos \omega t - \delta y_0^2(t - t_0)] dt.$$

Таким образом, заменяя переменную интегрирования $\tau = t - t_0$ и учитывая закон движения по сепаратрисе (2.166), приходим к соотношению:

$$D = \sqrt{2} \gamma \sin \omega t_0 \int_{-\infty}^{+\infty} d\tau \frac{\operatorname{sh} \tau}{\operatorname{ch}^2 \tau} \sin \omega \tau + 2\delta \int_{-\infty}^{+\infty} \frac{\operatorname{th}^2 \tau}{\operatorname{ch}^2 \tau} d\tau$$

или, окончательно,

$$D = \sqrt{2} \pi \gamma \omega \operatorname{sch} \frac{\pi \omega}{2} \sin \omega t_0 + \frac{4}{3} \delta.$$

Величина D меняет знак (т. е. сепаратрисы пересекаются и возникает хаотическое движение), если

$$\delta < \frac{3\sqrt{2}}{4} \pi \gamma \omega \operatorname{sch} \frac{\pi \omega}{2}.$$

Для выявления возможности подавления хаоса в системе (2.285) в работах [313, 397, 457] было предложено использовать мультипликативное возмущение кубической нелинейности:

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= x - \beta(1 + \eta \cos \Omega t)x^3 - \delta y + \gamma \cos \omega t,\end{aligned}\tag{2.286}$$

Очевидно, эта система — есть уравнение Дуффинга–Холмса (2.165). Для него, как показано в п. 21.5, расстояние Мельникова D дается соотношением (2.168).

Для того, чтобы величина D оставалась положительной для всех t_0 (что, собственно, и означает стабилизацию хаотического поведения), необходимо выполнение неравенства

$$\eta > \left| \frac{6\beta(A - C)}{\pi(\Omega^4 - 6\Omega^2 + 1)\operatorname{csch}(\pi\Omega/2)} \right|.$$

Однако это условие не является достаточным. Оно будет таковым, если частоты Ω и ω являются соизмеримыми. Более того, если отношение Ω/ω иррационально, то существует значение t_0 , когда $D(t_0)$ меняет знак. При этом период времени τ , в течение которого происходит двойная смена знака, можно определить из соотношения $A - B - C \simeq 0$. Величина τ , в зависимости от Ω , претерпевает скачки в точках, где частоты Ω и ω являются соизмеримыми. Используя численный анализ, можно убедиться, что хаос подавляется на частотах $\Omega \sim \Omega_R^{(k)} \equiv k\Omega_R^{(1)}$, где $\Omega_R^{(k)}$ — гармоники частоты возбуждения ω уравнения (2.286).

Таким образом, стабилизация хаотической динамики в уравнении Дюффинга наблюдается при резонансном соотношении частоты внешнего параметрического возмущения и частоты силовой составляющей.

Если параметрическое возмущение уравнения Дюффинга ввести иначе [386],

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= a(t)x - \beta x^3 - \delta y + \gamma \cos \omega t,\end{aligned}\tag{2.287}$$

где $a(t) = a(1 + \eta \cos \Omega t)$, то для стабилизации хаотической динамики достаточно, чтобы частота Ω была много больше частоты ω .

30.6 Подавление хаоса и стабилизация заданных циклов

Неоспоримым преимуществом описанных методов является то, что они позволяют развить аналитический подход и, в то же время, не требуют введения обратной связи. Однако ни один из них не дает возможность *управлять* системами с неустойчивым или хаотическим поведением. Тем не менее, если усовершенствовать внешние возмущения, то нетрудно добиться полного контроля над динамикой системы. Это возможно, если проблему управления представить как обратную задачу, когда неизвестными являются параметры, которые можно найти как решения уравнений на заданный цикл [120, 400].

Для одномерных отображений с критическими точками x_c (т. е. такими, что $\partial f(x, a)/\partial x|_{x=x_c} = 0$) эта задача решается точно¹⁵.

Допустим, что отображение интервала I , $x_{n+1} = f(x_n, a)$, $x \in I$, $a \in A$, удовлетворяет следующим свойствам: (i) существует такое подмножество $\sigma \subset I$, что для любых $x_1, x_2 \in \sigma$ найдется значение $a^* \in A$, для которого $f(x_1, a^*) = x_2$; (ii) при любом $a \in A$ существует критическая точка $x_c \in \sigma$. Тогда для любых $x_2, x_3, \dots, x_\tau \in \sigma$ найдутся такие x_1 и периодические возмущения параметра, a_1, a_2, \dots, a_τ , что цикл $(x_1, x_2, \dots, x_\tau)$ будет устойчивым циклом возмущенного отображения T_a .

Действительно, выберем произвольные величины x_1, x_2, \dots, x_τ . В силу условия (i) решение системы уравнений $f(x_1, a_1) = x_2$, $f(x_2, a_2) = x_3, \dots, f(x_\tau, a_\tau) = x_1$ относительно параметрических значений можно записать как $\hat{a} = (a_1, a_2, \dots, a_\tau)$. Это означает, что при периодическом возмущении $\hat{a} = (a_1, a_2, \dots, a_\tau)$ последовательность $(x_1, x_2, \dots, x_\tau) = p$ является циклом периода τ отображения T_a . Чтобы этот цикл p сделать устойчивым, достаточно выбрать элемент x_1 близким к критическому значению x_c , поскольку мультипликатор $\alpha(p) = \prod_{i=1}^{\tau} \partial f(x_i, a_i)/\partial x$ и $\partial f(x, a)/\partial x|_{x=x_c} = 0$ при любом a . Это гарантирует выполнение условия устойчивости $|\alpha(p)| < 1$.

Очевидно, условиям (i), (ii) удовлетворяют семейства отображений с минимумами или максимумами (см. §24). Поскольку любой цикл вида $(x_c, x_2, x_3, \dots, x_\tau)$ при произвольных $x_i \in \sigma$ является устойчивым, то приведенный анализ позволяет практически использовать данный метод управления динамикой систем, которые эффективно описываются такими семействами, например, для кодирования информации и пересылке зашифрованных сообщений [130, 131].

¹⁵Эти результаты получены одним из авторов (А. Ю. Л.) совместно с С. Д. Рыбалко.

В прикладных задачах очень важно знать максимально допустимый уровень шума, не разрушающий стабилизированные циклы. Нетрудно получить такие оценки.

Пусть стабилизированному циклу $(x_c, x_2, x_3, \dots, x_\tau)$ соответствует возмущение $(a_1, a_2, \dots, a_\tau)$. Предположим, что значения a_i слегка изменились: $(a'_1, a'_2, \dots, a'_\tau) = (a_1 + \Delta a_1, a_2 + \Delta a_2, \dots, a_\tau + \Delta a_\tau)$, $|\Delta a_i| \leq \varepsilon_a$. Найдем максимально допустимое значение ε_a , при котором возмущенный цикл сохраняет устойчивость, и исследуем, как в этом случае он исказится, т. е. определим Δx_i для $(x'_1, x'_2, \dots, x'_\tau) = (x_c + \Delta x_1, x_2 + \Delta x_2, \dots, x_\tau + \Delta x_\tau)$. Результаты таких вычислений даются следующей точной оценкой [400].

Если

$$|\Delta a_i| \leq \varepsilon_a = \frac{1}{\tau S_a L S_x^{\tau-1} \sum_{i=1}^{\tau} S_x^i},$$

где $S_a = \max_{x,a} |\partial f(x, a)/\partial a|$, $L = \max_{x,a} |\partial^2 f(x, a)/\partial x^2|$, $S_x = \max_{x,a} |\partial f(x, a)/\partial x|$, то возмущенное отображение имеет устойчивый цикл $p' = (x_c + \Delta x_1, x_2 + \Delta x_2, \dots, x_\tau + \Delta x_\tau)$ периода τ при $\hat{a}' = (a_1 + \Delta a_1, a_2 + \Delta a_2, \dots, a_\tau + \Delta a_\tau)$ и

$$|\Delta x_i| \leq \varepsilon_x = \frac{1}{L S_x^{\tau-1}}.$$

Полученные оценки позволяют в каждом конкретном случае эффективно определять предельно допустимые ошибки в задании необходимых управляющих параметров.

Рассмотрим в качестве примера возмущенное логистическое отображение:

$$\begin{aligned} x_{n+1} &= a_n x_n (1 - x_n), \\ a_{n+1} &= g(a_n). \end{aligned} \quad (2.288)$$

Для него множество σ — это интервал $[x_b, x_e]$, где x_b и x_e — решение уравнения $x_{int} = f(x, 4)$, x_{int} — точка пересечения дуги $y = 4x(1-x)$ и $y = x$. Таким образом, $[x_b, x_e] = [1/4, 3/4]$. Найдем возмущения $\hat{a} = (a_1, a_2, \dots, a_\tau)$, при которых в отображении (2.288) существует устойчивый цикл периода $t = m\tau$. Очевидно, что при $t = \tau$ точки, формирующие цикл $p = (x_1, x_2, \dots, x_t)$, будут подчиняться следующей системе уравнений:

$$x_2 = a_1 x_1 (1 - x_1), \quad x_3 = a_2 x_2 (1 - x_2), \quad \dots, \quad x_1 = a_t x_t (1 - x_t).$$

Чтобы решить обратную задачу, т. е. найти значения параметров, при которых отображение (2.288) имеет данный цикл p , необходимо отсюда выразить значения a_i :

$$\begin{aligned} a_1 &= \frac{x_2}{x_1(1-x_1)}, \\ a_2 &= \frac{x_3}{x_2(1-x_2)}, \\ &\dots\dots\dots, \\ a_t &= \frac{x_1}{x_t(1-x_t)}. \end{aligned} \quad (2.289)$$

Ясно, что не для всех возможных $x_i \in (0; 1)$ полученные значения будут удовлетворять условию $a_i \in [0; 4]$. Однако если это верно, то для любого цикла $p = (x_1, x_2, \dots, x_t)$ найдутся (a_1, a_2, \dots, a_t) . Теперь, если мультипликатор $|\alpha(p)| = \left| \prod_{i=1}^t a_i(1-2x_i) \right| < 1$, то данный цикл будет устойчивым. С учетом уравнений (2.289) это приводит к условию

$$|\alpha(p)| = \left| \prod_{i=1}^t \frac{x_{i+1}}{x_i(1-x_i)}(1-2x_i) \right| = \left| \prod_{i=1}^t \frac{1-2x_i}{1-x_i} \right| < 1. \quad (2.290)$$

Когда среди точек цикла существует критическая, $x_c = 1/2$, то $(1-2x_c)/(1-x_c) = 0$ и данное неравенство выполняется, т. е. цикл устойчив.

Множество значений $p = (x_1, x_2, \dots, x_t)$, для которых $a_i \in [0; 4]$ и удовлетворяется неравенство (2.290), образует определенную область в координатном пространстве. Каждой точке этой области соответствует устойчивый цикл возмущенного отображения. Используя систему уравнений (2.289), можно получить соответствующую область в *параметрическом* пространстве [400].

В простейшем случае двухпериодического возмущения ($\tau = 2$) в пространстве (x_1, x_2) область существования устойчивого цикла, как следует из соотношений (2.289) и (2.290), определяется следующей системой неравенств:

$$0 < \frac{x_2}{x_1(1-x_1)} \leq 4, \quad 0 < \frac{x_1}{x_2(1-x_2)} \leq 4, \quad \left| \frac{1-2x_1}{1-x_1} \frac{1-2x_2}{1-x_2} \right| < 1.$$

Выполнение первого и второго неравенств соответствует множеству всех допустимых значений (x_1, x_2) . Третье неравенство выделяет из этого множества область существования *устойчивых* циклов. Решая эту систему, получим их область существования (рис. 2.177 а). Нетрудно построить соответствующую область в пространстве параметров (a_1, a_2) . Для этого

достаточно использовать соотношение (2.289). Эта область показана на рис. 2.177 б.

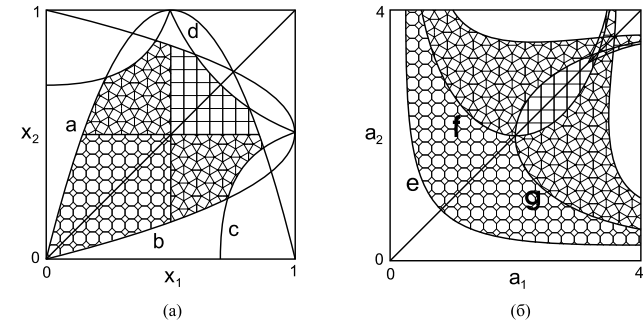


Рис. 2.177. Область существования устойчивых циклов периода 2 возмущенного ($\tau = 2$) отображения (2.288), задаваемая кривыми $x_2 = 4x_1(1-x_1)$ (а), $x_1 = 4x_2(1-x_2)$ (б), $5x_2 = (3x_1 - 2)/(5x_1 - 3)$ (с), $x_2 = x_1/(3x_1 - 1)$ (д), в пространстве (x_1, x_2) и кривыми $a_2 = 1/a_1$ (е), $a_2 = 8/[a_1(4 - a_1)]$ (ф), $a_1 = 8/[a_2(4 - a_2)]$ (г) в параметрическом пространстве (a_1, a_2)

Для выявления более детального действия отображения (2.288) область на рис. 2.177 а разбита прямыми $x_1 = 1/2$, $x_2 = 1/2$ на четыре части, окрашенные мозаичным способом. Соответствующие области в пространстве (a_1, a_2) имеют тот же мозаичный рисунок. Поскольку на рис. 2.177 б существуют перекрывающиеся области, то отображение (2.288) является однозначным, но не взаимно-однозначным. Кроме того, наличие перекрывающихся областей указывает на то, что при определенных возмущениях отображение бистабильно и обладает двумя устойчивыми циклами одновременно.

31. Пространственно-временной хаос

До сих пор мы рассматривали только сосредоточенные системы с конечным числом степеней свободы, которые могут обладать как свойствами регулярного движения, так и свойствами динамического хаоса. Однако кроме них существуют и распределенные системы. Типичным примером распределенных систем могут служить активные среды, о которых шла речь в гл. 1. В распределенных системах кроме регулярного поведения возможно возникновение хаотических пространственно-неоднородных автоколебаний, или турбулентности. Такие режимы эволюции распределенной

системы называют *пространственно-временным хаосом*. К их числу относятся гидродинамическая турбулентность, фибрилляции сердечной мышцы, сложные химические колебания и т. п. Здесь мы расскажем о некоторых результатах, относящихся к теории пространственно-временного хаоса для распределенных диссипативных систем.

Фазовое пространство распределенной системы является бесконечномерным. Каждая точка этого функционального пространства отвечает определенному распределению величин, характеризующих систему (например, полю температуры, давления, плотности и т. п.). Динамика распределенной системы описывается дифференциальными уравнениями в частных производных и состоит в последовательном изменении различных пространственных распределений. Поскольку каждому такому распределению ставится в соответствие определенная точка в фазовом пространстве, решение уравнений задает фазовую траекторию в этом пространстве. На случай распределенных систем легко обобщаются и многие другие понятия, введенные выше.

Нас будут интересовать прежде всего *диссипативные* распределенные системы. Их эволюция, как известно, сопровождается сжатием фазового объема. Это означает, что спустя достаточно большой промежуток времени все фазовые траектории окажутся сосредоточенными на каком-то подмножестве исходного бесконечномерного пространства. В частности, в фазовом пространстве распределенной системы могут существовать различные (конечномерные) аттракторы: стационарные точки, замкнутые притягивающие траектории и т. п. Могут иметься и более сложные конечномерные аттракторы, отвечающие хаотической динамике распределенной системы¹⁶.

Существование конечномерных аттракторов в распределенных системах было показано во многих работах (см., например, [43, 71, 109, 275, 366, 441]). Мы приведем один характерный пример — конвекцию в ячейке Хеле–Шоу [139]. Такая ячейка представляет собой плоский прямоугольный сосуд, наполненный жидкостью, который равномерно подогревается снизу. При слабом подогреве жидкость остается в покое. С ростом притока тепла жидкость приходит в движение, совершая сначала периодические, а затем, при дальнейшем увеличении подогрева, аperiodические колебания. Ясно, что жидкость в ячейке Хеле–Шоу является распределенной системой, и полностью ее движение должно описываться уравнениями в частных производных. Однако оказывается, что установившееся поведение такой системы может быть очень хорошо аппроксимировано достаточно простой динами-

¹⁶В принципе возможны также бесконечномерные аттракторы, но этот вопрос до сих пор недостаточно исследован.

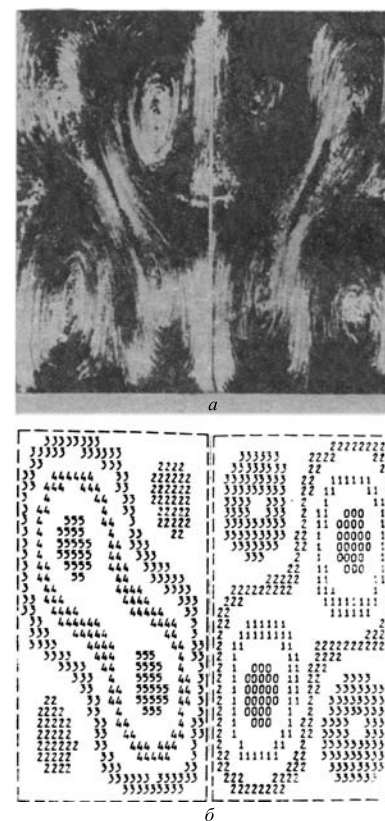


Рис. 2.178. Исследование конвекции в ячейке Хеле–Шоу: *a* — физический эксперимент; *b* — результат численного интегрирования восьми обыкновенных дифференциальных уравнений, полученных из исходной распределенной системы в рамках галеркинской аппроксимации [139]

ческой системой из восьми обыкновенных дифференциальных уравнений (рис. 2.178*a, б*).

31.1 Размерность вложения

Наименьшее число независимых переменных, однозначно определяющих установившееся движение диссипативной распределенной систе-

мы, называют (минимальной) *размерностью вложения* и обозначают как d_e [135, 447, 504]. В приведенном только что примере размерность вложения равнялась восьми. Размерность вложения *аттрактора* — это минимальная размерность фазового подпространства, в которое без самопересечений может быть помещено подмножество, целиком содержащее этот аттрактор. Ясно, что размерность подпространства, куда вкладывается аттрактор, больше размерности аттрактора.

Приведем несколько примеров. Тороидальную поверхность можно вложить в трехмерное пространство. Плоскую кривую — разместить без самопересечений на плоскости. Однако пространственная кривая (скажем, достаточно сложный цикл на торе) может быть помещена без самопересечений только в трехмерное пространство. Существует теорема, согласно которой любое гладкое многообразие размерности m всегда может быть вложено (помещено без самопересечений) в пространство как минимум размерности $2m + 1$ [165, 297]. Это означает, в частности, что любые кривые ($m = 1$) могут быть вложены в трехмерное пространство.

Странный аттрактор, характеризующийся фрактальной размерностью d_F , всегда можно вложить в пространство с целочисленной размерностью k , которая удовлетворяет неравенству $k \geq 2d_F + 1$ [418, 504]. Разумеется, в особых случаях для некоторых конкретных аттракторов размерность вложения может быть и меньше, вплоть до значения $[d_F] + 1$, где $[d_F]$ — целая часть d_F .

Возникает вопрос, как по экспериментальным данным установить размерность вложения распределенной системы? Ответ дает процедура, предложенная в [447, 504] и основанная на теореме Такенса ([504], см. также [418]). Согласно этой процедуре, для определения размерности вложения достаточно знать, как меняется со временем лишь *единственная* переменная в какой-то точке (например, температура или скорость жидкости в эксперименте по исследованию конвекции Рэлея–Бенара, см. рис. 2.80). Поясним идею этого метода.

31.2 Функциональный метод определения размерности вложения

Допустим, что мы изучаем распределенную систему, находящуюся в установившемся режиме. Тогда, регистрируя изменение какой-либо одной величины, например, x , через *равные* промежутки времени τ , мы получаем последовательность значений

$$x(t) \equiv y_1, x(t + \tau) \equiv y_2, \dots, x(t + (n - 1)\tau) \equiv y_n, \dots \quad (2.291)$$

Вначале для простоты будем считать, что динамика этой системы эффективно описывается всего одним дифференциальным уравнением 1-го порядка. Тогда, как легко видеть, последовательность (2.291) будет однозначно определяться начальным значением y_1 , а остальные значения y_n будут функционально зависимыми от него, т.е. $y_2 = f(y_1)$, $y_3 = f(f(y_1))$ и т.д. Это означает, что все точки последовательности (2.291) на плоскости (y_{n+1}, y_n) будут ложиться на определенную линию $y_{n+1} = f(y_n)$.

Предположим теперь, что для описания системы требуется два обыкновенных дифференциальных уравнения 1-го порядка. В этом случае точки y_n уже не ложатся на линию, а будут сложным образом разбросаны по плоскости (y_{n+1}, y_n) . Теперь для однозначного задания последовательности (2.291) уже недостаточно только y_1 ; требуется дополнительно знать значение y_2 . Повторяя эти рассуждения для двух произвольных последовательных значений y_{n-1}, y_n , мы видим, что они должны однозначно определять величину y_{n+1} , т.е. должна существовать функция f , такая, что $y_{n+1} = f(y_{n-1}, y_n)$.

В общем случае, если размерность динамической системы равна k , зная k последовательных значений $y_n, y_{n-1}, \dots, y_{n-k+1}$, можно однозначно восстановить по ним следующее значение y_{n+1} , т.е. должна существовать некоторая функция f , такая, что $y_{n+1} = f(y_n, y_{n-1}, \dots, y_{n-k+1})$ для любого n . Иными словами, величина y_{n+1} оказывается функционально зависимой от $y_n, y_{n-1}, \dots, y_{n-k+1}$.

Метод, предложенный в [447, 504], формулируется следующим образом. Вначале через равные промежутки времени τ строится последовательность $y_1, y_2, \dots, y_n, y_{n+1}, \dots$ значений какой-либо наблюдаемой величины. Начиная с $k = 1$, осуществляется проверка того, будет ли при любом n значение y_{n+1} в этой последовательности функционально зависимым от предшествующих k значений $y_n, y_{n-1}, \dots, y_{n-k+1}$. Если такая функциональная зависимость отсутствует, число k увеличивается на единицу и проверка повторяется. То значение k , для которого функциональная зависимость будет найдена, дает минимальную размерность вложения $d_e = k$.

Рассмотрим теперь практические методы проверки на функциональную зависимость. Описанный выше способ, заключающийся в геометрическом построении функции $y_{n+1} = f(y_n, y_{n-1}, \dots)$, уже для размерности вложения $d_e = 3$ является невыполнимым. Вместо него можно использовать простой геометрический критерий, предложенный в [135] (см. также [245]).

Выберем в последовательности измерений (2.291), начиная с n -го элемента, произвольную последовательность длины k и рассмотрим всевозможные векторы $w^{(n)} = \{y_{n-k+1}, y_{n-k+2}, \dots, y_{n-1}, y_n\}$. Очевидно, что все такие векторы принадлежат некоторому k -мерному пространству. За-

фиксируем одну из подпоследовательностей, начинающуюся с какого-то номера n_0 , и обозначим соответствующий вектор как $\mathbf{w}^{(0)}$. Рассмотрим расстояние $\rho_k(n, n_0)$ между произвольным вектором $\mathbf{w}^{(n)}$ и этим вектором $\mathbf{w}^{(0)}$:

$$\rho_k(n, n_0) = |\mathbf{w}^{(n)} - \mathbf{w}^{(0)}| = \sqrt{\sum_{i=1}^k (y_{n-k+i} - y_{n_0-k+i})^2}, \quad (2.292)$$

а также определим величину

$$r(n, n_0) = |y_{n+1} - y_{n_0+1}|. \quad (2.293)$$

Построим далее для различных n зависимость $r(\rho)$, откладывая по оси абсцисс значения $\rho_k(n, n_0)$, а по оси ординат — соответствующие им величины $r(n, n_0)$.

Наличие функциональной зависимости между y_{n+1} и предыдущими k членами $y_n, y_{n-1}, \dots, y_{n-k+1}$ означает существование некоторой функции f , такой, что $y_{n+1} = f(\mathbf{w}^{(n)})$. Поэтому

$$r(n, n_0) = |f(\mathbf{w}^{(n)}) - f(\mathbf{w}^{(0)})|. \quad (2.294)$$

Расстояние $\rho_k(n, n_0)$ стремится к нулю, когда вектор $\mathbf{w}^{(n)}$ стремится к $\mathbf{w}^{(0)}$. При этом, как видно из (2.294), $r(n, n_0)$ также должно стремиться к нулю. Если же функциональная зависимость отсутствует, то вектор $\mathbf{w}^{(n)}$ не определяет однозначно значение y_{n+1} и, следовательно, $r(n, n_0)$ не будет стремиться к нулю при уменьшении $\rho_k(n, n_0)$.

Таким образом, процедура нахождения размерности вложения заключается в следующем. Для различных n необходимо найти значения $\rho_k(n, n_0)$ и $r(n, n_0)$, согласно (2.292), (2.293), и отложить соответствующие точки на плоскости (r, ρ) , последовательно соединив их друг с другом, так что получится некоторая ломаная линия. Если функциональная зависимость при выбранном k достигнута, то в области малых значений ρ все участки этой ломаной располагаются вблизи начала координат. Размерность вложения — это минимальное значение k , начиная с которого ломаная обладает этим свойством.

В качестве иллюстрации описанного метода приведем результаты экспериментальной работы [183], в которой изучалась геометрия аттракторов для различных режимов хаоса в распределенной системе. Реализация алгоритма вычисления d_e , согласно формулам (2.292)–(2.294), показана на рис. 2.179. Видно, что размерность вложения $d_e = 5$.

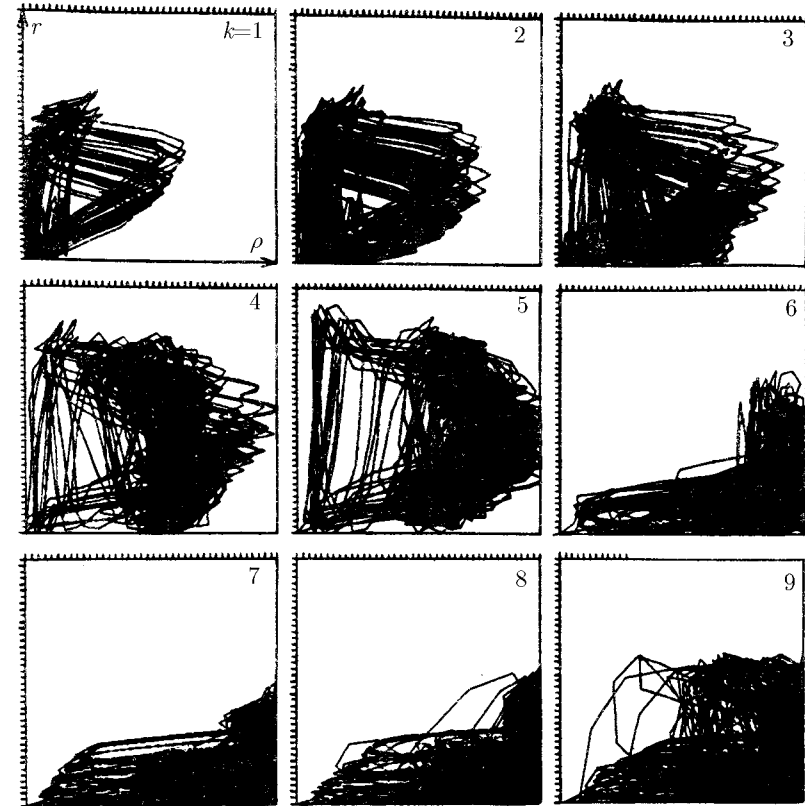


Рис. 2.179. Проверка функциональной зависимости $(k+1)$ -й координаты от предыдущих в распределенной системе [183]

31.3 Корреляционная размерность

Более точный метод оценки размерности вложения для распределенных систем предложен П. Грассберггером и И. Прокаччиа [326, 327]. Этот метод позволяет также рассчитать так называемую *корреляционную размерность* странного аттрактора.

Предположим, как и прежде, что из эксперимента известна последовательность (2.291), построенная по единственной наблюдаемой с фиксиро-

ванным временем запаздывания τ . Аналогично предыдущему выделим из этой последовательности различные произвольные подпоследовательности длины k и построим всевозможные векторы $w^{(n)}$. Определим по формуле (2.292) расстояние $\rho_k(n, n')$ для любой пары $w^{(n)}, w^{(n')}$ таких векторов. Возьмем некоторое значение l и вычислим корреляционный интеграл $C_k(l)$ как

$$C_k(l) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{n, n'=1}^N \Theta[l - \rho_k(n, n')]. \quad (2.295)$$

Здесь $\Theta(z)$ — ступенчатая функция Хевисайда, такая, что $\Theta(z) = 0$ при $z < 0$ и $\Theta(z) = 1$ при $z \geq 0$, а N — полное число элементов в последовательности (2.291)¹⁷.

Как следует из (2.295), в корреляционный интеграл $C_k(l)$ дают вклад только те пары векторов $w^{(n)}$ и $w^{(n')}$, расстояние между которыми меньше или равно заданной величине l . При малых l корреляционный интеграл должен стремиться к нулю, т. е.

$$C_k(l) \sim l^{\alpha_k}. \quad (2.296)$$

Величину α_k можно найти, построив график зависимости $\ln C_k(l)$ от $\ln l$. Таким образом, можно вычислить значение α_k для различных k , начиная с $k = 1$. Оказывается [327], что начиная с некоторого номера величина α_k , т. е. характерный наклон графика $\ln C_k(l)$ от $\ln l$ перестает возрастать с увеличением k . Это значение k дает размерность вложения аттрактора d_e , а предельный тангенс угла наклона — корреляционную размерность ν данного аттрактора. В общем случае значение корреляционной размерности не превышает фрактальной размерности d_F аттрактора, введенной в разделе 22.3 [326, 327]. Эти две размерности совпадают лишь при равномерном распределении точек на аттракторе [327].

Приведем пример использования процедуры Грассбергера–Прокаччия.

В работе [414] методом поглощения светового потока изучалась конвекция Рэлея–Бенара. Узкий пучок света, рассеянный небольшим объемом жидкости, несет информацию о частотах фурье-компонент скорости жидкости. По экспериментальной последовательности измерений вычислялась зависимость $\ln C_k(l)$ от $\ln l$ для различных значений k (рис. 2.180). Из рисунка видно, что при $k > 4$ наклон графика не меняется, т. е. $d_e = 5$. Корреляционная размерность при этом равна $\nu \approx 2,8$.

¹⁷В реальных условиях вместо предельной величины $C_k(l)$ вычисляется соответствующее значение для максимальной зарегистрированной последовательности измерений.

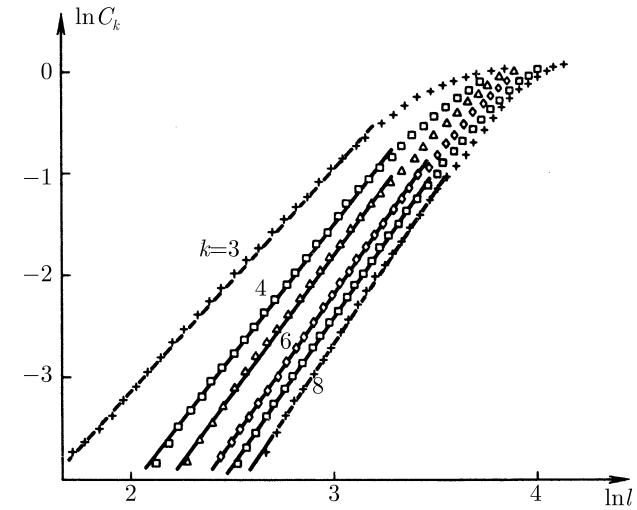


Рис. 2.180. Экспериментальная зависимость корреляционного интеграла C_k от l для ряда значений k (в логарифмическом масштабе) [414]

В [222, 293] процедура Грассбергера–Прокаччия использовалась для обработки электроэнцефалограмм, регистрирующих активность мозга. Метод Грассбергера–Прокаччия применялся для определения размерности вложения d_e и корреляционной размерности ν также во многих других экспериментальных исследованиях [71, 286, 295, 431, 531].

Заметим, что для систем с большой размерностью вложения использовать метод Грассбергера–Прокаччия становится трудно, поскольку при этом быстро возрастает требуемое число экспериментальных точек N [494].

Процедура Грассбергера–Прокаччия позволяет не только определять размерность вложения d_e и корреляционную размерность ν , но и отличать динамический хаос от случайного шума, всегда присутствующего в эксперименте [230, 435]. Обобщение понятия корреляционного интеграла дает возможность оценить снизу КС-энтропию [328]. Для наиболее эффективного использования процедуры Грассбергера–Прокаччия имеются рецепты оптимального выбора времени запаздывания τ [310], а также некоторые способы увеличения скорости сходимости метода [209, 512].

31.4 Аналитические оценки размерности аттрактора

Ранее обсуждался главным образом вопрос о том, как определить некоторые количественные характеристики аттрактора, если известна лишь по-

следовательность измерений какой-то физической величины, полученной, например, из экспериментальных данных. Существует, однако, и другая задача. Допустим, что нам известны динамические уравнения, описывающие некоторую распределенную активную среду. Как, зная эти уравнения, рассчитать размерность вложения и другие количественные характеристики пространственно-временного хаоса в такой системе?

В общем случае эта задача не решена. Поэтому ниже мы ограничимся рассмотрением одного конкретного примера, который играет важную роль в общей теории активных сред. Рассмотрим осциллирующую активную среду, которая описывается обобщенным уравнением Гинзбурга–Ландау (см. раздел 10 главы 1):

$$\dot{\eta} = (\alpha_1 - i\alpha_2)\eta - (\beta_1 + i\beta_2)|\eta|^2\eta + (D_1 + iD_2)\frac{\partial^2\eta}{\partial x^2}. \quad (2.297)$$

В [289] получена аналитическая оценка ляпуновской размерности d_L для аттрактора этого уравнения в одномерном случае при периодических граничных условиях. Пусть x в (2.297) меняется в интервале от 0 до L_0 . Удобно ввести параметр $Re = \alpha_1 L_0^2 / D_1$, который играет роль эффективного числа Рейнольдса для рассматриваемой системы. Если $|\beta_2/\beta_1| \leq \sqrt{3}$, то ляпуновские характеристические показатели имеют вид

$$\lambda_n = \left(\frac{D_1}{L_0^2}\right) \left(Re - (2\pi)^2 \left[\frac{n}{2}\right]\right), \quad (2.298)$$

где $[n/2]$ означает целую часть числа $n/2$, $n = 1, 2, \dots$. Тогда ляпуновская размерность d_L легко рассчитывается по формуле Каплана–Йорке (2.176). Полученная зависимость d_L от Re показана на рис. 2.181.

Нетрудно также получить очень простую верхнюю оценку ляпуновской размерности d_L :

$$d_L \leq 2\sqrt{\frac{3Re}{4\pi^2} + \frac{1}{4}}. \quad (2.299)$$

Она показана штриховой линией на рис. 2.181. Эта оценка совпадает с точным значением d_L в точках, где d_L равна любому нечетному целому числу. Напомним, что ляпуновская размерность d_L дает верхнюю оценку для фрактальной размерности (см. раздел 22.4).

Кроме того, удается найти нижнюю оценку фрактальной размерности

$$d_F \geq 1 + 2\left[\frac{\sqrt{Re}}{2\pi}\right], \quad (2.300)$$

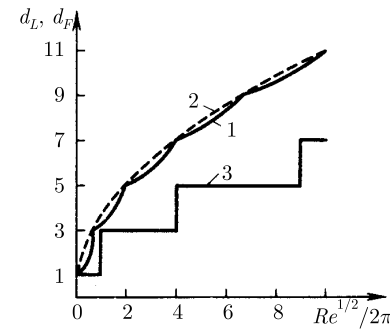


Рис. 2.181. Зависимость размерностей d_L и d_F от числа Рейнольдса Re для уравнения Гинзбурга–Ландау: точное значение d_L (1), верхняя оценка d_L (2), нижняя оценка фрактальной размерности d_F (3) [289]

где квадратные скобки вновь обозначают взятие целой части. На рис. 2.181 она изображена ступенчатой линией.

При выполнении неравенства $|\beta_2/\beta_1| > \sqrt{3}$ можно получить лишь верхнюю оценку для ляпуновской размерности [289].

Из поведения верхней и нижней оценок d_F (2.299), (2.300) при больших Re следует, что

$$d_F \sim \frac{\sqrt{Re}}{2\pi}, \quad Re \gg 1. \quad (2.301)$$

Этот результат имеет простую физическую интерпретацию: размерность аттрактора примерно равна «числу независимых степеней свободы» для рассматриваемой системы при заданном Re . Число степеней свободы можно оценить как L_0/L_g , где $L_g = 2\pi\sqrt{D_1/\alpha_1}$ есть самая короткая длина волны для самовозбуждающихся мод. (Все моды, имеющие более короткую длину волны, являются затухающими.) Легко убедиться, что, согласно (2.301), $d_F \sim L_0/L_g$ при $Re \gg 1$. Понятие эффективного числа степеней свободы при турбулентном движении впервые ввел в рассмотрение Л. Д. Ландау [112, 113].

32. Динамика систем сцепленных отображений¹⁸

Описание распределенных систем может быть выполнено посредством иного подхода, когда среда представляется совокупностью взаимодействующих элементов различной природы. Такое дискретное представление сейчас весьма популярно, поскольку является достаточно эффективным методом

¹⁸Результаты этого параграфа получены одним из авторов (А.Ю.Л.) совместно с С. Д. Рыбалко.

исследования. При этом дискретизация может быть как пространственная, так и пространственно-временная. При пространственной дискретизации исходная система аппроксимируется конечным или счетным набором элементов, между которыми введен определенный вид связи. Каждый такой элемент представляет собой некоторую динамическую систему с малым числом переменных. Если в дополнение к этому динамическая система задается отображением, то говорят о пространственно-временной дискретизации.

Пространственно-временные дискретные модели называют *решеточными* и *сеточными*. Отдельные элементы при этом представляют собой точки физического пространства, а характер связи между ними моделирует взаимодействие исходя из основных физических принципов.

В зависимости от размерности исходной моделируемой системы решетки сцепленных отображений могут быть одномерными, двумерными или трехмерными. К исследованию таких моделей сводятся многие задачи теории нелинейных распределенных сред. Например, некоторые проблемы статистической физики могут быть эффективно изучены посредством их представления через пространственно-временной аналог, распределенные возбудимые системы часто описывают дискретными аналогами и т. п. (см. [251, 258, 376–378, 434, 513] и приведенные там ссылки).

Решеточные системы возникают не только при рассмотрении распределенных сред, но также и при изучении процессов, происходящих в системах, имеющих существенно дискретную структуру как по пространству, так и во времени. К таким моделям непосредственно приводит ряд задач теории синхронизации радиогенераторов, биологии, медицины, а также изучение поведения клеточных автоматов и нейронных сетей (см. [251, 258]).

При исследовании решеток сцепленных отображений естественным образом возникают следующие вопросы. Если поведение отдельных элементов является хаотическим, будет ли решетка также проявлять хаотические свойства? Хаос подавит окружающий порядок или же порядок в отдельных элементах распространится на всю решетку? Наконец, может ли в первоначально пространственно однородной решетке образоваться определенная структура, т. е. произойти кластеризация и самоорганизация в решетке? На эти вопросы нетрудно найти ответ, если иметь критерий локального поведения элементов и решетки в целом (см. ниже).

Замена распределенных сред их дискретным аналогом подразумевает адекватность непрерывного физического пространства дискретным с правильно выбранной связью между элементами. Однако даже при удачном выборе полученная решеточная модель далеко не всегда описывает богатство всех явлений, наблюдаемых в распределенных системах. Поэтому за-

частую прибегают к рассмотрению таких решеток, чтобы при изменении управляющих параметров они демонстрировали широкий спектр явлений, присущих распределенной среде [376–378].

Таким образом, при построении решеток, аппроксимирующих исходную среду, прежде всего необходимо подобрать вид отображения, которое задает временную эволюцию в каждом элементе. Такой выбор определяет локальное фазовое пространство X , на котором действует преобразование $f : X \rightarrow X$,

$$x_{n+1} = f(x_n, a), \quad x \in X, \quad (2.302)$$

где a — параметр. При этом фазовым пространством всей решетки является прямое произведение всех локальных пространств отдельных элементов. Как правило, в качестве преобразования (2.302) выбирается одномерное отображение.

Второй основной задачей является определение связей между элементами. Для двумерной решетки, состоящей из $L \times K$ элементов, в общем случае такую связь можно определить преобразованием

$$x(l, k) = g\left(x(1, 1), x(1, 2), \dots, x(1, K), x(2, 1), x(2, 2), \dots, x(2, K), \dots, x(L, 1), x(L, 2), \dots, x(L, K)\right). \quad (2.303)$$

Следовательно, динамика решетки связанных отображений может быть представлена в виде композиции двух отображений (2.302) и (2.303). Тогда за один шаг по времени значение элемента под номером (l, k) , $l = 1, \dots, L$, $k = 1, \dots, K$, будет изменяться по закону

$$x_{n+1}(l, k) = g\left(f(x_n(1, 1), a), f(x_n(1, 2), a), \dots, f(x_n(1, K), a), f(x_n(2, 1), a), f(x_n(2, 2), a), \dots, f(x_n(2, K), a), \dots, f(x_n(L, 1), a), f(x_n(L, 2), a), \dots, f(x_n(L, K), a)\right), \quad (2.304)$$

так что поведение всей решетки распадается на временное, задаваемое функцией f , и пространственное, определяемое преобразованием g .

Хотя большинство результатов, относящихся к связанным отображениям, получены именно для такого разложения, было бы совершенно неправильно утверждать, что это представление является общим. Например, при дискретизации некоторых уравнений в частных производных, описывающих реальную распределенную среду, возникают преобразования, когда оба отображения f и g действуют одновременно.

Наиболее часто исследуются решетки, когда каждый элемент тем или иным образом взаимодействует только со своими ближайшими соседями (см., например, [251, 258, 434, 513] и приведенные там ссылки). При таком взаимодействии, как правило, изучается диффузионный вид связи элементов. Он используется, в основном для моделирования таких явлений, как пространственно-временной хаос, структурообразование, самоорганизация и др. [226, 253, 258, 376–378, 433, 513, 536] (см. также цитированную там литературу).

В данном параграфе мы подробно исследуем только *одномерные* системы (т. е. цепочки) сцепленных отображений. Для них *диффузионная* связь означает, что состояние элементов (2.304) задается преобразованием в следующей форме:

$$x_{n+1}(k) = f(x_n(k), a) + \frac{\varepsilon}{2} \left(f(x_n(k-1), a) - 2f(x_n(k), a) + f(x_n(k+1), a) \right), \quad (2.305)$$

где ε — коэффициент диффузии, a — параметр. Такая связь характерна тем, что поведение каждого элемента как бы сглаживается действием окружающих его соседей, и это сглаживание тем сильнее, чем больше величина ε (рис. 2.182).

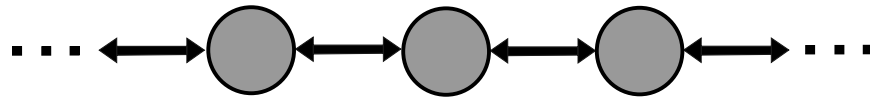


Рис. 2.182. Диффузионное сцепление элементов (2.305) в цепочке связанных отображений

Наряду с диффузионным типом связи вида (2.305) часто исследуется поведение отображений, задаваемых соотношением

$$x_{n+1}(k) = f(x_n(k), a) + \frac{\varepsilon}{2} \left(x_n(k-1) - 2x_n(k) + x_n(k+1) \right), \quad (2.306)$$

где $x \in [0, 1]$, a — параметр нелинейности и ε — коэффициент взаимодействия между элементами. Такой выбор связи означает, что в системе (2.306) пространственное и временное преобразования происходят одновременно. Данное приближение можно использовать при анализе однокомпонентных активных сред (см. ниже).

Помимо представленных видов взаимодействия иногда рассматривается так называемая *поточковая* модель, аппроксимирующая свободное течение жидкости. В этом случае цепочка представляет собой систему односто-

ронне связанных отображений [380, 381] (рис. 2.183):

$$x_{n+1}(k) = f(x_n(k), a) + \varepsilon \left(f(x_n(k-1), a) - f(x_n(k), a) \right), \quad (2.307)$$

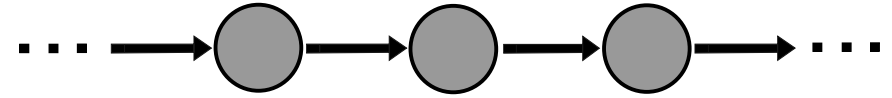


Рис. 2.183. Поточковая модель (2.307) цепочки связанных отображений

Еще один вид связи предполагает глобальное сцепление, когда каждый элемент сети связан с каждым [373, 379, 453].

Очевидно, что в дополнение к виду связи необходимо указать поведение системы на границе. Наиболее часто связанные отображения задаются со свободными или периодическими граничными условиями. Свободные граничные условия определяются соотношениями $x(0) = x(K+1) \equiv 0$ (или $x(0) \equiv 0$ для систем вида (2.307)). При периодических граничных условиях необходимо положить $x(k) = x(K+k)$, $k = 1, \dots, K$.

Если все элементы, составляющие решетку сцепленных отображений, одинаковы, то такая решетка называется *однородной*. Меняя управляющий параметр и величину диффузии (или силу связи), можно исследовать фазовую диаграмму, возможные фазовые переходы и возникающие при этом пространственно-временные картины.

Когда параметры в элементах различны или же некоторые из них представлены другими отображениями, то такая система уже не будет однородной и ее исследование является намного более сложной задачей [128, 129]. При этом неоднородности (которые часто называют «дефектами») могут быть самые разные: от отдельных дефектов до периодической неоднородности по всему пространству.

Практически во всех работах, посвященных изучению диффузионно взаимодействующих отображений, исследуются (и, как правило, численно) однородные решетки. Однако очевидно, что с физической точки зрения однородность пространства (в данном случае идентичность всех элементов) — это идеализация, принятая для упрощения анализа.

Математически решетки связанных отображений с конечным числом элементов представляют собой динамическую систему с конечным числом степеней свободы. Вычисление основных характеристик такой динамической системы сопряжено с большими трудностями либо в принципе невозможно. Более того, многие величины, на основе которых можно судить

об эволюции отдельных элементов и всей решетки (показатели Ляпунова, пространственный и временной спектры, скорость убывания временных корреляций и др.), отражают характер только *асимптотического* поведения при $t \rightarrow \infty$ и в целом по пространству [319, 370]. Очевидно, для решеток связанных отображений знание этих величин не представляет значительного интереса. Важнее найти характеристики, описывающие их *локальные* свойства и эволюцию во времени и пространстве.

Настоящий параграф посвящен исследованию неоднородных систем одномерных сцепленных отображений. Мы покажем, как описать некоторые из таких моделей аналитически и представим достаточно эффективный метод численного анализа, позволяющий определять как поведение отдельных элементов и динамику всей решетки в целом, так и визуализировать эволюцию системы [128]. Он выгодно отличается от других имеющихся методов (см. [376–378, 433, 434]) своей простотой и наглядностью.

32.1 Модель однокомпонентной активной среды

Остановимся сначала на модели (2.306). Она допускает аналитическое решение и поэтому мы опишем ее подробно.

32.1.1 Однородная среда

Перепишем систему K диффузионно сцепленных отображений (2.306) в виде

$$x_{n+1}(k) = G(x_n(k), \alpha, \gamma) + \varepsilon(x_n(k+1) - 2x_n(k) + x_n(k-1)) \quad (2.308)$$

с периодическими граничными условиями $x_n(k+K) \equiv x_n(k)$, где $k = 1, 2, \dots, K$ — дискретная пространственная координата, $n = 0, 1, 2, \dots$ — дискретная временная координата, $\varepsilon > 0$ — коэффициент диффузии, который считается постоянным, и $\alpha > 0$, γ — параметры. В качестве функции G будем рассматривать следующую:

$$G(x, \alpha, \gamma) = \begin{cases} (1 - 2\alpha)x - \gamma, & x < 1/2; \\ (1 - 2\alpha)x - \gamma + 2\alpha, & x > 1/2. \end{cases} \quad (2.309)$$

Функция $G(x, \alpha, \gamma)$ выбрана в таком виде не случайно. Система отображений (2.308) с элементами вида (2.309) является дискретным вариантом

базовой модели однокомпонентной одномерной активной среды, описываемой уравнением типа Колмогорова–Петровского–Пискунова:

$$\frac{\partial U(x, t)}{\partial t} = D \frac{\partial^2 U(x, t)}{\partial x^2} + \Phi(U).$$

Как известно, это уравнение часто встречается в задачах биофизики, теории горения, химической кинетике (например, в реакции Белоусова–Жаботинского), твердотельной электронике и т. д. (см., например, [1, 59, 86, 221]).

Легко видеть, что $G(x, \alpha, \gamma)$ — кусочно-линейная функция, график которой имеет постоянный наклон, равный $(1 - 2\alpha)$. В зависимости от значений параметра α отображение, порождаемое функцией G , может обладать качественно различными типами поведения. При $0 < \alpha < 1$ (см. рис. 2.184а) в зависимости от величины γ оно имеет одну или две устойчивые точки A и B . При $\alpha > 1$ наклон графика отображения превышает по модулю единицу (см. рис. 2.184б), что соответствует экспоненциальному разбеганию близких траекторий. При этом, если движение фазовой точки является финитным, динамика будет хаотической (см. раздел 25.1)).

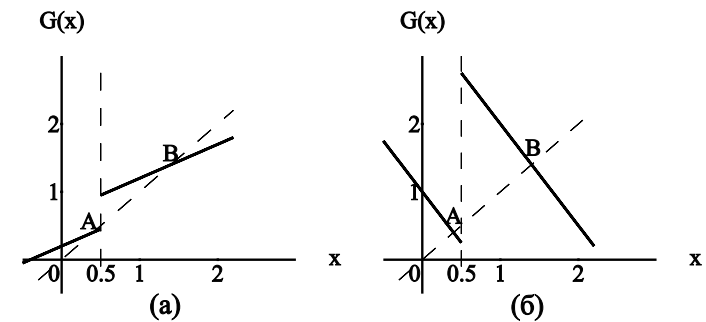


Рис. 2.184. Отображение, порождаемое функцией $G(x, \alpha, \gamma)$ при $\alpha = 0, 25$, $\gamma = -0, 2$ (а) и $\alpha = 1, 25$, $\gamma = -1$ (б)

Однородная модель (2.308), составленная из сцепленных функций вида (2.309), была подробно описана в монографии [39]. Приведем основные результаты этого исследования. Систему (2.308), (2.309) можно рассматривать как K -мерное кусочно-линейное отображение. Поскольку $G(x, \alpha, \gamma)$ — кусочно-линейная функция с одинаковой производной, то дифференциал D_f отображения f — постоянная матрица. Поэтому легко определить ля-

пуновские показатели λ_i такого отображения. Если ρ_s — собственные числа D_f , то $\lambda_s = \ln |\rho_s|$. Когда среди всех ρ_s имеется собственное число, лежащее на комплексной плоскости вне единичного круга, то любая траектория отображения f неустойчива. В противном случае динамика модели будет регулярной.

Таким образом, вопрос о том, каким является поведение системы, сводится к изучению расположения корней ρ_s . Характеристический многочлен дифференциала K -мерного отображения выражается через определители трехдиагональных матриц разной размерности, для которых можно получить рекуррентные соотношения. Рассматривая эти соотношения как разностные уравнения с заданными начальными условиями, нетрудно найти их решения, которые будут пропорциональны полиномам Чебышёва от линейной функции собственных чисел D_f . Теперь, используя свойства полиномов Чебышёва, можно показать, что характеристический многочлен сводится к квадратному трехчлену от известной функции ρ_s . Вычисление значений ρ_s и анализ их расположения на комплексной плоскости приводят к разбиению пространства параметров системы отображений (2.308), (2.309) на две области (см. рис. 2.185):

- область D_1 , задаваемую такими параметрами α и ε , что все собственные числа дифференциала отображения f по модулю меньше единицы, так что модель обладает регулярной динамикой;
- область D_2 , определяемую условием, что среди собственных значений ρ_s есть по крайней мере один корень, удовлетворяющий неравенству $|\rho_s| > 1$. При этом в общем случае динамика может быть инфинитной. Однако при финитности движения можно говорить о хаотическом поведении системы (2.308), (2.309) [258].

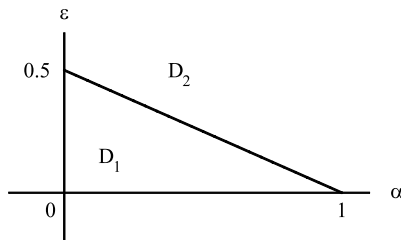


Рис. 2.185. Пространство параметров однородной кольцевой цепочки диффузионно связанных отображений (2.308), (2.309)

Система (2.308), (2.309) является пространственно однородной в том смысле, что все составляющие ее элементы одинаковы. Однако наша цель — рассмотреть пространственно неоднородную модель. Неоднородность может быть связана либо с тем, что в разных точках пространства действуют различные отображения, либо с тем, что отображения одинаковы во всех элементах, но их параметры различны.

Для возможности проведения аналитического исследования остановимся на последнем случае, т.е. допустим, что каждый элемент по-прежнему является кусочно-линейным отображением (см. (2.309)), но, в отличие от однородного случая, будем считать, что имеются два типа элементов, отличающиеся друг от друга значением параметров α_i функции $G(x_i, \alpha_i, \gamma)$ ¹⁹. Предположим, что одному типу элементов соответствует параметр α , а другому — параметр β . Чтобы избежать краевых эффектов, как и для однородного случая (2.308), будем считать, что цепочка замкнута в кольцо, т.е. $x_n(k+K) \equiv x_n(k)$.

Мы подробно исследуем два качественно различных случая: модель с периодически расположенными неоднородностями и модель с единственным дефектом.

32.1.2 Кольцевая модель с периодической неоднородностью по пространству

Периодическая пространственная неоднородность означает, что в цепочке чередуются элементы с различными значениями параметра функции G . При чередовании «через один элемент» такую цепочку можно записать как

$$x_{n+1}(k) = \begin{cases} G(x_n(k), \alpha, \gamma) + \varepsilon(x_n(k+1) - 2x_n(k) + x_n(k-1))), & \text{для нечетных } k, \\ G(x_n(k), \beta, \gamma) + \varepsilon(x_n(k+1) - 2x_n(k) + x_n(k-1))), & \text{для четных } k, \end{cases} \quad (2.310)$$

где $x_n(k+K) \equiv x_n(k)$, функция G по-прежнему задается соотношением (2.309) и число элементов K четное (в противном случае цепочка (2.310) не замкнется в кольцо). Исследуем, при каких значениях параметров α, β

¹⁹Вообще говоря, элементы могут различаться и величиной параметра γ . Но, как будет видно из дальнейшего изложения, характер динамики как неоднородной, так и однородной системы не зависит от γ .

и ε периодически неоднородная кольцевая модель (2.310), (2.309) будет обладать качественно различными режимами поведения. Для этого вычислим показатели Ляпунова, используя описанную выше технику.

Дифференциал K -мерного отображения, порождаемого системой вида (2.310), с учетом граничных условий имеет вид:

$$D_f \equiv Q_K = \begin{pmatrix} 1 - 2\varepsilon - 2\alpha & \varepsilon & 0 & \dots & \varepsilon \\ \varepsilon & 1 - 2\varepsilon - 2\beta & \varepsilon & \dots & 0 \\ 0 & \varepsilon & 1 - 2\varepsilon - 2\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon & 0 & 0 & \dots & 1 - 2\varepsilon - 2\beta \end{pmatrix}.$$

Чтобы найти собственные значения ρ_s , $s = 1, 2, \dots, K$, матрицы Q_K , вычислим детерминант

$$\det(Q_K - \rho I_K) = \begin{vmatrix} 2z_1\varepsilon & \varepsilon & 0 & 0 & \dots & \varepsilon \\ \varepsilon & 2z_2\varepsilon & \varepsilon & 0 & \dots & 0 \\ 0 & \varepsilon & 2z_1\varepsilon & \varepsilon & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varepsilon \\ \varepsilon & 0 & 0 & 0 & \dots & 2z_2\varepsilon \end{vmatrix},$$

где z_1 и z_2 определяются из соотношений $1 - 2\varepsilon - 2\alpha - \rho = 2z_1\varepsilon$, $1 - 2\varepsilon - 2\beta - \rho = 2z_2\varepsilon$. Разлагая его по первой строке, получим: $\det(Q_K - \rho I_K) = 2z_1\varepsilon B_{K-1} - \varepsilon^2 B_{K-2} - 2\varepsilon^K - \varepsilon^2 \hat{B}_{K-2}$, где

$$B_K = \begin{vmatrix} 2z_1\varepsilon & \varepsilon & 0 & \dots & 0 \\ \varepsilon & 2z_2\varepsilon & \varepsilon & \dots & 0 \\ 0 & \varepsilon & 2z_1\varepsilon & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2z_2\varepsilon \end{vmatrix}, \quad \hat{B}_K = \begin{vmatrix} 2z_2\varepsilon & \varepsilon & 0 & \dots & 0 \\ \varepsilon & 2z_1\varepsilon & \varepsilon & \dots & 0 \\ 0 & \varepsilon & 2z_2\varepsilon & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2z_1\varepsilon \end{vmatrix}.$$

Легко видеть, что $\hat{B}_K(z_1, z_2)$ получается из $B_K(z_1, z_2)$ заменой $z_1 \leftrightarrow z_2$. Поэтому нетрудно найти $B_K(z_1, z_2)$ с помощью рекуррентного соотношения. Разлагая B_K по первой строке, найдем:

$$B_K = \begin{cases} 2z_1\varepsilon B_{K-1} - \varepsilon^2 B_{K-2}, & \text{для четных } K; \\ 2z_2\varepsilon B_{K-1} - \varepsilon^2 B_{K-2}, & \text{для нечетных } K. \end{cases}$$

Отсюда легко выразить детерминанты нечетного порядка через детерминанты четного порядка,

$$B_{2K+1} = \frac{B_{2K+2} + \varepsilon^2 B_{2K}}{2z_1\varepsilon}, \quad (2.311)$$

и получить рекуррентное соотношение для детерминантов четного порядка:

$$B_{2K} = \varepsilon^2(4z_1z_2 - 2)B_{2K-2} - \varepsilon^4 B_{2K-4}. \quad (2.312)$$

Это соотношение можно интерпретировать как разностное уравнение с начальными условиями

$$B_0 = 1, \quad B_2 = \varepsilon^2(4z_1z_2 - 1). \quad (2.313)$$

Тогда, решая систему (2.312), (2.313) и учитывая (2.311), получаем:

$$B_K(z_1, z_2) = \begin{cases} \varepsilon^K U_K(z), & \text{для четных } K, \\ \varepsilon^K \sqrt{\frac{z_2}{z_1}} U_K(z), & \text{для нечетных } K, \end{cases}$$

где $U_K(z)$ — полином Чебышёва второго рода. В свою очередь, соотношения для $\hat{B}_K(z_1, z_2)$ имеют вид:

$$\hat{B}_K(z_1, z_2) = \begin{cases} \varepsilon^K U_K(z), & \text{для четных } K, \\ \varepsilon^K \sqrt{\frac{z_1}{z_2}} U_K(z), & \text{для нечетных } K. \end{cases}$$

Используя соотношения для B_{K-1} , B_{K-2} и \hat{B}_{K-2} , находим:

$$\det(Q_K - \rho I_K) = 2\varepsilon^K (zU_{K-1}(z) - U_{K-2}(z) - 1). \quad (2.314)$$

Теперь нетрудно определить собственные значения дифференциала отображения (2.310), (2.309). Из (2.314) следует, что

$$zU_{K-1}(z) - U_{K-2}(z) - 1 = 0, \quad (2.315)$$

или $T_K(z) - 1 = 0$. Здесь мы воспользовались соотношением между полиномами Чебышёва второго рода $U_K(z)$ и первого рода $T_K(z)$. Поскольку

$$T_K(z) = \frac{(z - \sqrt{z^2 - 1})^K + (z + \sqrt{z^2 - 1})^{-K}}{2},$$

где $z^2 = z_1 z_2$, то, полагая $t = z - \sqrt{z^2 - 1}$, из (2.315) получим:

$$\frac{1}{2} \left(t^K + \frac{1}{t^K} \right) - 1 = 0,$$

или $(t^K - 1)^2 = 0$. Следовательно, $t_s = e^{i2\pi s/K}$ и $z_s = \cos(2\pi s/K)$, $s = 1, 2, \dots, K$. Вспоминая, что $2z_1\varepsilon \cdot 2z_2\varepsilon = 4z^2\varepsilon^2$, приходим к несложному уравнению:

$$(1 - 2\varepsilon - 2\alpha - \rho)(1 - 2\varepsilon - 2\beta - \rho) = 4\varepsilon^2 \cos^2 \left(s \frac{2\pi}{K} \right).$$

Решая его относительно ρ , найдем набор собственных значений дифференциала отображения (2.310), (2.309):

$$\rho_{1,2}^s = 1 - 2\varepsilon - \alpha - \beta \pm \sqrt{(\alpha - \beta)^2 + 4\varepsilon^2 \cos^2 \left(s \frac{2\pi}{K} \right)}, \quad (2.316)$$

где $s = 1, 2, \dots, K/2$. Динамика системы (2.310), (2.309) будет полностью регулярной, если все значения ρ_s будут находиться внутри единичного круга на комплексной плоскости. Поскольку все $\rho_{1,2}^s$ действительны, то условие существования регулярной динамики выполнено, когда

$$2\varepsilon + \alpha + \beta > \sqrt{4\varepsilon^2 + (\alpha - \beta)^2}, \quad \varepsilon + \frac{\alpha + \beta}{2} + \sqrt{\varepsilon^2 + \left(\frac{\alpha - \beta}{2} \right)^2} < 1.$$

Так как $\varepsilon, \alpha, \beta > 0$, то первое неравенство этой системы выполняется всегда. В свою очередь, область в пространстве параметров, где $\lambda_s > 0$, удовлетворяет условиям $\varepsilon + (\alpha + \beta)/2 + \sqrt{\varepsilon^2 + (\alpha - \beta)^2/4} > 1$.

Граница раздела областей с качественно различной (регулярной и хаотической) динамикой — это поверхность в трехмерном пространстве параметров $(\alpha, \beta, \varepsilon)$, которая, очевидно, задается выражением

$$\varepsilon + \frac{\alpha + \beta}{2} + \sqrt{\varepsilon^2 + \frac{(\alpha - \beta)^2}{4}} = 1. \quad (2.317)$$

Для наглядности приведем сечения фазового пространства системы (2.310), (2.309) плоскостями $\delta = \text{const}$, где $\delta = \beta - \alpha$ — параметр неоднородности (рис. 2.186). Пусть D_1 обозначает область регулярной динамики, а D_2 —

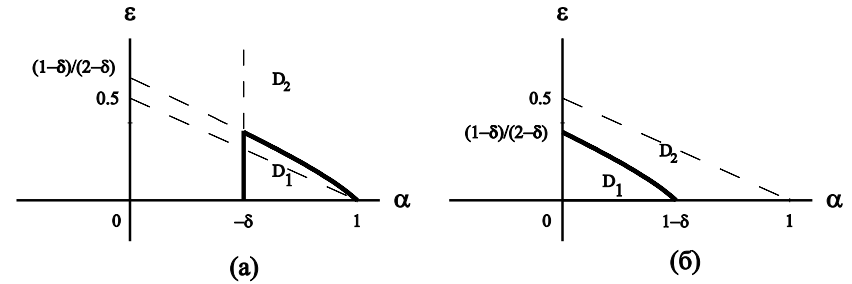


Рис. 2.186. Сечения плоскостями $\delta = \text{const}$ пространства параметров модели (2.309), (2.310) при $-1 < \delta < 0$ (а) и $0 < \delta < 1$ (б)

область, где $\lambda_s > 0$. Граница раздела этих областей задается уравнением, вытекающим из (2.317). Для сравнения на этом же рисунке пунктирной линией представлена граница раздела областей D_1 и D_2 для *однородной* ($\delta = 0$) системы сцепленных отображений (см. рис. 2.185).

Легко видеть, что при $\delta > 1$ при любых значениях ε, α модель проявляет хаотическую динамику. При $\delta < 1$ появляется область D_1 , где динамика регулярна. Эта область в определенном диапазоне значений параметров $0 < \delta < 1$ будет подобластью области регулярного поведения для однородной модели. При $-1 < \delta < 0$ среди значений параметров ε, α существуют такие, при которых неоднородная модель имеет регулярную динамику, тогда как однородная — хаотическую. Ограничение $\alpha > -\delta$ соответствует условию $\beta > 0$. При $\delta < -1$ область D_1 исчезает, и при всех возможных значениях ε, α модель (2.310), (2.309) проявляет только хаотические свойства.

Непосредственно из рис. 2.186 следует возможность двух качественно различных вариантов: параметр α принадлежит области порядка, β — области хаоса с $\lambda_s > 0$, а общая динамика является регулярной, и наоборот, общая динамика периодически неоднородной модели может быть хаотической. Реализация того или иного случая зависит от величины параметра диффузии ε .

32.1.3 Кольцевая модель с единственным дефектом

Рассмотрим теперь другой интересный случай пространственно неоднородной модели диффузионно сцепленных отображений: систему с единственным «дефектом». Этот случай означает, что один из составляющих ее K элементов имеет параметр β (см. (2.309)), а остальные $(K - 1)$ элементов — параметр α . Без потери общности будем считать, что выделенный

элемент находится в точке $k = 1$:

$$x_{n+1}(k) = \begin{cases} G(x_n(k), \alpha, \gamma) + \varepsilon(x_n(k-1) - 2x_n(k) + x_n(k+1)), & k = 2, 3, \dots, K, \\ G(x_n(1), \beta, \gamma) + \varepsilon(x_n(K) - 2x_n(1) + x_n(2)), & k = 1, \end{cases} \quad (2.318)$$

где $x_n(k+K) \equiv x_n(k)$, $k = 1, 2, \dots, K$. Пусть по-прежнему функция $G(x, \alpha, \gamma)$ задается выражением (2.309). Определим дифференциал соответствующего отображения для системы (2.318), (2.309):

$$D_f \equiv Q_K = \begin{pmatrix} 1 - 2\varepsilon - 2\beta & \varepsilon & 0 & \dots & \varepsilon \\ \varepsilon & 1 - 2\varepsilon - 2\alpha & \varepsilon & \dots & 0 \\ 0 & \varepsilon & 1 - 2\varepsilon - 2\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon & 0 & 0 & \dots & 1 - 2\varepsilon - 2\alpha \end{pmatrix}. \quad (2.319)$$

Если использовать для нахождения ляпуновских показателей модели (2.318), (2.309) ту же технику, что и ранее, то решение характеристического уравнения для дифференциала отображения (2.319) сведется к определению корней многочлена порядка $2N + 2$. Это можно сделать только численно, и поэтому для исследования этой модели применим несколько иной подход.

Пользуясь оценкой собственных значений матрицы (2.319) по теореме Гершгорина [191], построим оценку области регулярной динамики системы (2.318), (2.309). Согласно утверждению этой теоремы, все собственные числа матрицы $A = \{a_{ij}\}_{k \times k}$ принадлежат объединению кругов на комплексной плоскости:

$$\rho_s \in \bigcup_{i=1}^k \left\{ z \in \mathbf{C} : |z - a_{ij}| \leq R'_i \right\}, \quad s = 1, 2, \dots, K,$$

где R'_i — либо строчная почти-норма матрицы A , $R'_i = \sum_{j=1, j \neq i}^k |a_{ij}|$, либо столбцевая почти-норма матрицы A , $R'_j = \sum_{i=1, i \neq j}^k |a_{ij}|$. Так как дифференциал (2.319) представляет собой действительную симметричную матрицу, то, во-первых, строчная и столбцевая почти-нормы совпадают, а, во-вторых, собственные значения D_f являются действительными. Следовательно, круги Гершгорина преобразуются в интервалы на действительной

оси:

$$\rho_s \in \bigcup_{i=1}^k \left\{ z \in \mathbf{R} : |z - a_{ij}| \leq R'_i \right\}, \quad s = 1, 2, \dots, K.$$

Так как $\varepsilon > 0$, из (2.319) легко определить R'_i :

$$R'_i = \sum_{\substack{j=1 \\ j \neq i}}^k |a_{ij}| = |\varepsilon| + |\varepsilon| = 2\varepsilon.$$

Поскольку R'_i не зависит от i и в дифференциале D_f (2.319) присутствует только два типа различных диагональных элементов, то собственные значения D_f принадлежат объединению двух интервалов:

$$\left\{ z \in \mathbf{R} : |1 - 2\varepsilon - 2\alpha - z| \leq 2\varepsilon \right\} \cup \left\{ z \in \mathbf{R} : |1 - 2\varepsilon - 2\beta - z| \leq 2\varepsilon \right\}.$$

Тогда, учитывая, что $\alpha, \beta, \varepsilon > 0$, получим *оценку сверху* для модуля собственных значений D_f :

$$|\rho_s| \leq \begin{cases} \beta + 2\varepsilon, & \alpha < \beta, \\ \alpha + 2\varepsilon, & \alpha > \beta. \end{cases}$$

Таким образом, область \tilde{D}_1 в пространстве параметров $(\alpha, \beta, \varepsilon)$, удовлетворяющая условиям

$$\tilde{D}_1 : \begin{cases} \beta + 2\varepsilon < 1, & \text{при } \alpha < \beta, \\ \alpha + 2\varepsilon < 1, & \text{при } \alpha > \beta, \end{cases} \quad (2.320)$$

соответствует регулярной динамике неоднородной модели (2.318), (2.309), а область

$$\tilde{D}_2 : \begin{cases} \beta + 2\varepsilon > 1, & \text{при } \alpha < \beta, \\ \alpha + 2\varepsilon > 1, & \text{при } \alpha > \beta, \end{cases} \quad (2.321)$$

соответствует положительным λ_s (рис. 2.187).

Область \tilde{D}_1 является *нижней оценкой* области регулярной динамики, так как она получена с помощью оценки сверху собственных значений

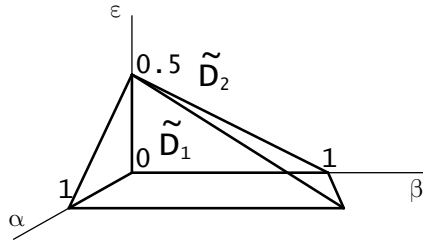


Рис. 2.187. Оценка области регулярной динамики неоднородных цепочек с двумя типами элементов в пространстве параметров $(\alpha, \beta, \varepsilon)$

дифференциала отображения. Любая точка пространства параметров из \tilde{D}_1 будет соответствовать регулярной динамике модели, а любая точка из \tilde{D}_2 — либо регулярной, либо, если движение финитно, хаотической динамики.

Полученная оценка верна не только для рассмотренной модели с одним «дефектом», но и для *целого класса* кольцевых систем отображений с диффузионной связью вида (2.308), характеризующихся наличием двух типов элементов (с разными параметрами α_i функции $G(x_i, \alpha_i, \gamma)$). Дело в том, что в данном случае количество элементов каждого типа и их взаимное расположение *не имеют значения*. В самом деле, почти-нормы дифференциалов отображения всех таких систем совпадают и центры двух возможных интервалов, в которых, по теореме Гершгорина, содержатся собственные значения D_f , также общие.

Согласно следствию из теоремы Гершгорина, если два интервала, в которых проводится оценка ρ_s , не пересекаются, происходит «кластеризация» собственных значений, т. е. q собственных значений ρ_s , $s = 1, 2, \dots, q$, будут лежать внутри одного интервала, а $(K - q)$ собственных значений ρ_s , $s = q + 1, q + 2, \dots, K$, — внутри другого интервала. Здесь q — число элементов с параметром α и, соответственно, $(K - q)$ — число элементов с параметром β . Однако учет этого свойства не позволяет улучшить оценку области регулярной динамики.

Сравним оценку (2.320) с точным результатом, полученным для кольцевой модели с пространственно периодической неоднородностью (см. выше). Очевидно, эта модель относится к классу, для которого приведенная оценка остается справедливой. Для наглядности построим сечения $\delta = \text{const}$ области D_1 , соответствующей регулярной динамике системы (2.310), (2.309), и ее оценки \tilde{D}_1 (рис. 2.188). Из этого рисунка следует, что оценка (2.320) области регулярной динамики достаточно хорошо приближает истинное поведение модели, но не отражает того факта, что при $-1 < \delta < 0$ область

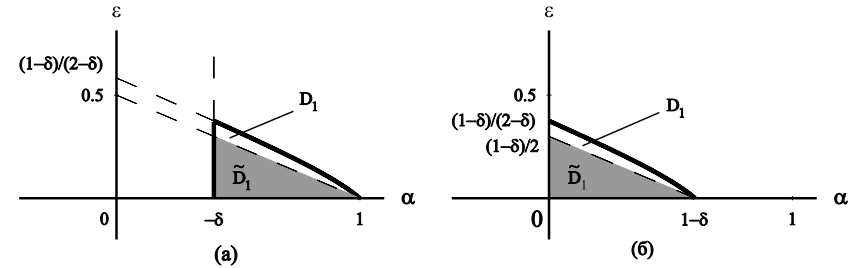


Рис. 2.188. Оценка и точный результат для области регулярной динамики модели с периодической неоднородностью при $-1 < \delta < 0$ (а) и $0 < \delta < 1$ (б)

регулярной динамики для системы (2.310), (2.309) шире области с тем же типом поведения для пространственно однородного случая.

Вернемся теперь к системе отображений (2.318), (2.309), пространственная неоднородность которой связана с наличием только одного «дефекта», и *численно* найдем собственные значения дифференциала отображения (2.319). Для этого случая характеристический многочлен можно записать следующим образом:

$$\det(Q_K - \rho I_K) = \begin{vmatrix} 2z\varepsilon + 2(\alpha - \beta) & \varepsilon & 0 & \dots & \varepsilon \\ \varepsilon & 2z\varepsilon & \varepsilon & \dots & 0 \\ 0 & \varepsilon & 2z\varepsilon & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon & 0 & 0 & \dots & 2z\varepsilon \end{vmatrix},$$

где $2z\varepsilon = 1 - 2\varepsilon - 2\alpha - \rho$. Разлагая $\det(Q_K - \rho I_K)$ по первой строке, получим:

$$\det(Q_K - \rho I_K) = [2z\varepsilon + 2(\alpha - \beta)]B_{K-1} - 2\varepsilon^2 B_{K-2} - 2(-1)^K \varepsilon^K, \quad (2.322)$$

где

$$B_K = \begin{vmatrix} 2z\varepsilon & \varepsilon & 0 & \dots & 0 \\ \varepsilon & 2z\varepsilon & \varepsilon & \dots & 0 \\ 0 & \varepsilon & 2z\varepsilon & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2z\varepsilon \end{vmatrix}.$$

Как и ранее, используя разложение B_K по первой строке, найдем рекуррентное соотношение: $B_K = 2z\varepsilon B_{K-1} - \varepsilon^2 B_{K-2}$. Его можно интерпретировать как разностное уравнение с начальными условиями $B_1 = 2z\varepsilon$,

$B_2 = \varepsilon^2(4z^2 - 1)$, решение которого дается соотношением

$$B_K = \varepsilon^K U_K(z), \quad (2.323)$$

где $U_K(z)$ — полином Чебышёва второго рода. Учитывая (2.323) и вид $U_K(z)$,

$$U_K(z) = \frac{(z + \sqrt{z^2 - 1})^{K+1} - (z - \sqrt{z^2 - 1})^{K+1}}{2\sqrt{z^2 - 1}} = \frac{(1/t)^{K+1} - t^{K+1}}{1/t - t},$$

где $t = z - \sqrt{z^2 - 1}$, из (2.322) получим следующее соотношение:

$$\det(Q_K - \rho I_K) = 2\varepsilon^K \left[\frac{t^{2K+1} + 1}{2t^K} + \left(\frac{\alpha - \beta}{\varepsilon} \right) \frac{1 - t^{2K}}{(1 - t^2)t^{K-1}} - (-1)^K \right].$$

Таким образом, характеристическое уравнение для дифференциала отображения (2.319) можно записать как

$$t^{2K+2} + 2 \left(\frac{\alpha - \beta}{\varepsilon} \right) t^{2K+1} - t^{2K} - 2(-1)^K t^{K+2} + 2(-1)^K t^K + t^2 - 2 \left(\frac{\alpha - \beta}{\varepsilon} \right) t - 1 = 0. \quad (2.324)$$

Теперь, численно определив корни уравнения (2.324), легко вычислить ρ_s и получить для всех возможных значений параметров α , β , ε области D_1 и D_2 с качественно различной динамикой.

На рис. 2.189 приведены сечения этих областей плоскостями $\delta = \text{const}$, где $\delta = \beta - \alpha$. Пунктиром указана оценка (2.320) области регулярной динамики. Так же, как для модели с периодической неоднородностью, в данном случае при положительных δ эта область шире, чем предсказывает оценка. Однако, в отличие от периодической неоднородности, при $\delta < 0$ оценка полностью совпадает с численным результатом и одновременно с результатом для однородной цепочки. Таким образом, если один из элементов системы имеет параметр меньший, чем остальные, то на характер динамики всего ансамбля такой дефект не окажет никакого влияния. Если же параметр отличается в большую сторону, это сужает область регулярной динамики модели по сравнению с однородным случаем.

Все три рассмотренных случая, т. е. оценка области регулярной динамики (2.320) для класса моделей с двумя типами элементов, точные результаты для периодически неоднородной системы (рис. 2.186) и модели

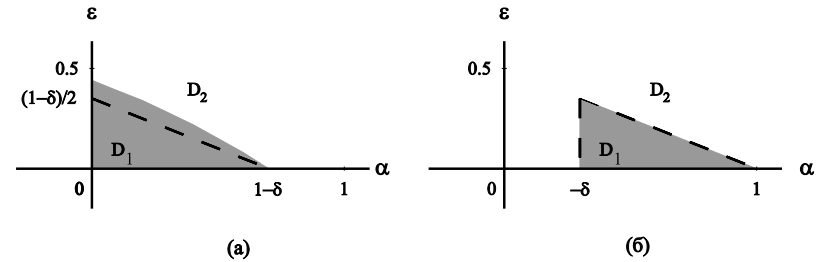


Рис. 2.189. Характерный вид сечений плоскостями $\delta = \text{const}$ пространства параметров цепочки с $N = 20$ элементами и одним дефектом при $\delta = -0, 3$ (а) и $\delta = 0, 3$ (б)

с одним «дефектом» (рис. 2.189) отражают возможность двух вариантов общей динамики системы. В первом случае, когда параметр одного типа элементов принадлежит области регулярной динамики, а параметр другого типа элементов — области неустойчивости, то поведение модели является регулярным. Во втором случае наоборот, динамика всей модели может быть неустойчивой, или, в случае финитности, хаотической. Реализация того или иного типа поведения зависит от параметра диффузии ε и возможна только при его малых значениях. Очевидно, что при достаточно больших значениях ε общая динамика всех цепочек из рассматриваемого класса не будет регулярной.

32.2 Локальный критерий динамики связанных отображений

В данной части мы опишем достаточно эффективный метод численного анализа решеток связанных отображений, позволяющей определять поведение отдельных элементов и динамику всей системы в целом [128]. Он основан на разбиении всего временного интервала на короткие отрезки времени, на которых проводится анализ степени неперIODичности траекторий всех элементов решетки. Это дает возможность определять образование областей синхронизации, их временной период, трансформацию со временем и разрушение путем пространственной хаотизации. Хотя в данном параграфе мы изучаем только одномерные цепочки связанных отображений, метод может быть использован для исследования и многомерного случая.

В процессе временной эволюции значения каждого элемента решетки в последовательные моменты времени образуют временной ряд:

$$x_1(k), x_2(k), x_3(k), \dots, x_n(k), \dots, \quad (2.325)$$

где k — номер элемента, а n — дискретное время. Для того чтобы выявить

периодичность этого ряда и определить, как изменяется степень периодичности со временем, необходимо, очевидно, взять только некоторый отрезок этого ряда и попытаться выяснить, насколько данный отрезок ряда близок к периодическому. Разбивая весь ряд на небольшие отрезки одной длины, мы производим определенное огрубление по времени. Анализируя периодичность ряда на каждом таком отрезке и выражая ее некоторым числом, можно проследить эволюцию данной характеристики при переходе от одного отрезка к другому.

Опишем теперь данную процедуру более формально. Пусть ряд состоит из T элементов, т. е. он имеет вид x_1, x_2, \dots, x_T (для удобства мы опускаем номер элемента k). Разобьем его на одинаковые временные интервалы τ . Выберем T и τ так, чтобы $T = m\tau$, где m — некоторое целое. Теперь исследуем каждый полученный кусок на периодичность. Для этого с каждым отрезком ряда x_1, \dots, x_τ сделаем следующую процедуру.

- 1) Сравним его со специально построенным рядом $x_1, x_1, x_1 \dots, x_1$, т. е. таким, все элементы которого равны x_1 . Результат сравнения выразим в виде числа λ_1 . Способ вычисления λ_1 будет представлен ниже. Сейчас пока для нас важно, чтобы число λ_1 обладало следующим свойством: чем λ_1 меньше, тем отрезок ряда ближе к построенному $x_1, x_1, x_1 \dots, x_1$.
- 2) Сравним искомый ряд с рядом $x_1, x_2, x_1, x_2, \dots, x_1, x_2, \dots$, т. е. с рядом, состоящим из периодических подпоследовательностей периода 2. Результат выразим числом λ_2 .
- 3) Прделаем аналогичную операцию для рядов, составленных из трех элементов x_1, x_2, x_3 , четырех элементов x_1, x_2, x_3, x_4 и т. д. вплоть до $x_1, x_2, \dots, x_{\tau/2}$. Результаты сравнения также выразим числами $\lambda_3, \lambda_4, \dots, \lambda_{\tau/2}$, которые вновь будут отражать некоторую степень отклонения исследуемого ряда (2.325) от модельного, составленного из отдельных последовательных его элементов.
- 4) Найдем величину

$$\lambda \equiv \lambda_{\min} = \min \{ \lambda_1, \lambda_2, \dots, \lambda_{\tau/2} \}. \quad (2.326)$$

В общем виде сравнительный показатель λ_p для каждого $p = 1, \dots, \tau/2$ вычисляется по формуле

$$\lambda_p = \sqrt{\frac{\sum_{i=1}^{\tau} (x_i - x_{(i \bmod (p+1))})^2}{\tau - p}}. \quad (2.327)$$

Легко видеть, что λ_p — это среднеквадратичное отклонение элементов остатка исходного ряда x_{p+1}, \dots, x_τ от модельного периодического ряда $x_1, x_2, \dots, x_p, x_1, x_2, \dots, x_p, \dots$. Согласно (2.327), при каждом p выполняется $\lambda_p \geq 0$. Ясно, что если для некоторого p имеет место равенство $\lambda_{\min} = \lambda_p = 0$, то ряд x_1, \dots, x_τ — периодический с периодом, кратным p . Чтобы точно определить период исходного ряда, необходимо остановить процедуру вычисления показателей λ_p на первом значении p , при котором $\lambda_p \simeq 0$. Если, однако, в процессе вычисления получим $\lambda = \min \{ \lambda_1, \lambda_2, \dots, \lambda_{\tau/2} \} > 0$, то можно с определенностью сказать, что ряд x_1, \dots, x_τ не будет периодическим.

Значение p , на котором достигается данный минимум, можно назвать периодом, который приближает данный аperiodический ряд. Определив величину λ для отдельных отрезков длиной τ , можно проследить их динамику на всем промежутке длины T . Вычисляя λ и периоды, на которых достигается минимум, для всех точек цепочки, можно эффективно выявлять области синхронизации в пространстве рассматриваемой системы сцепленных отображений и изучать динамику отдельных элементов.

Легко понять, что предлагаемый метод анализа решеток связанных отображений, строго говоря, не является точным. Прежде всего, необходимо удачно выбрать значение τ , т. е. длину анализируемого отрезка. Это можно сделать, вычисляя наборы λ_p для различных τ . Другим недостатком является тот факт, что показатель λ_{\min} будет положительным для элементов, динамика которых квазипериодична. Тем не менее, он будет достаточно близок к нулю, так как квазиопериодическое движение можно сколь угодно точно приблизить периодическим. Однако поскольку нас интересует *качественная* картина распределения λ_{\min} , а не конкретные значения, то данный недостаток не является столь существенным.

32.3 Примеры однородных систем

Продемонстрируем эффективность предложенного метода на примерах, описанных выше.

32.3.1 Потокковая модель

Потоковая модель (2.307) обладает широким спектром различных режимов поведения в зависимости от степени нелинейности, задаваемой параметром a , и значения коэффициента диффузии ε . Более того, можно аналитически вычислить наиболее простые виды ее предельного поведения. Остановимся на этом подробнее.

В качестве единичного элемента, составляющего потоковую цепочку, возьмем квадратичную функцию $f(x, a) = 1 - ax^2$. Этот выбор обусловлен тем фактом, что в диапазоне $0 \leq a \leq 2$ одиночное отображение с такой функцией проявляет весь спектр сложного движения, поскольку заменой координат оно сводится к логистическому (см. раздел 26.1).

Формально система (2.307) в момент времени n задается K -мерным вектором $\xi_n = (x_n(1), x_n(2), \dots, x_n(K))$. Следовательно, динамика всей цепочки выражается некоторым преобразованием фазового пространства \mathcal{X} в себя: $F : \mathcal{X} \mapsto \mathcal{X}$, $\xi_{n+1} = F(\xi_n)$. При этом, очевидно, функция F задается как композиция отображений во времени (2.302) и пространстве (2.303).

Таким образом, изучение устойчивых стационарных и периодических состояний сводится к анализу спектра собственных значений матрицы Якоби DF

$$DF(\xi) = \begin{pmatrix} -(1-\varepsilon)2ax(1) & 0 & 0 & \dots & 0 \\ \varepsilon & -(1-\varepsilon)2ax(2) & 0 & \dots & 0 \\ 0 & \varepsilon & -(1-\varepsilon)2ax(3) & \dots & 0 \\ 0 & 0 & \varepsilon & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -(1-\varepsilon)2ax(K) \end{pmatrix} \quad (2.328)$$

и ее степеней $DF^{(p)}$, где p — период. Матрица DF является треугольной. Поэтому все ее степени $DF^{(p)}$ также являются треугольными матрицами.

Вычислим однородные стационарные состояния модели (2.307). Однородность означает, что при любом $k = 1, \dots, K$ выполняется $x_n(k) = x_n^*$, а стационарность выражает независимость однородного состояния от времени, т.е. $x_n(k) = x_n^* \equiv x^*$. Следовательно, значения x^* должны подчиняться условию

$$x^* = f(x^*, a) + \varepsilon(f(x^*, a) - f(x^*, a)) = f(x^*, a). \quad (2.329)$$

Иными словами, x^* — это неподвижные точки отображения, задаваемого функцией $f(x, a)$:

$$x_{1,2}^* = \frac{-1 \pm \sqrt{1+4a}}{2a}. \quad (2.330)$$

Таким образом, система (2.307) имеет два однородных состояния: $\xi_1 = (x_1^*, x_1^*, \dots, x_1^*)$ и $\xi_2 = (x_2^*, x_2^*, \dots, x_2^*)$. Их устойчивость определяется величинами собственных значений матриц $DF(\xi_1)$ и $DF(\xi_2)$. Поскольку $DF(\xi)$ — треугольная матрица, то из (2.328) найдем, что все собственные

значения $\lambda_i \equiv \lambda_1 = -(1-\varepsilon)2ax_1^*$ для ξ_1 и $\lambda_i \equiv \lambda_2 = -(1-\varepsilon)2ax_2^*$ для ξ_2 . Состояния ξ_1 и ξ_2 устойчивы, если $|\lambda_{1,2}| \leq 1$. Или

$$-1 < -(1-\varepsilon)2ax_1^* < 1, \quad -1 < -(1-\varepsilon)2ax_2^* < 1. \quad (2.331)$$

Подстановка x_1^*, x_2^* приводит к соотношениям между параметрами a и ε :

$$1 - \frac{1}{1 + \sqrt{1+4a}} < \varepsilon < 1 + \frac{1}{1 + \sqrt{1+4a}} \quad \text{для } x_1^*, \quad (2.332)$$

$$1 - \frac{1}{\sqrt{1+4a} - 1} < \varepsilon < 1 + \frac{1}{\sqrt{1+4a} - 1} \quad \text{для } x_2^*.$$

На рис. 2.190 представлен результат вычисления областей (2.332) для $\varepsilon \in [0; 1]$ и $a \in [0; 2]$. Области, обозначенные $S1$ и $S2$, соответствуют устойчивым точкам x_1^* и x_2^* .

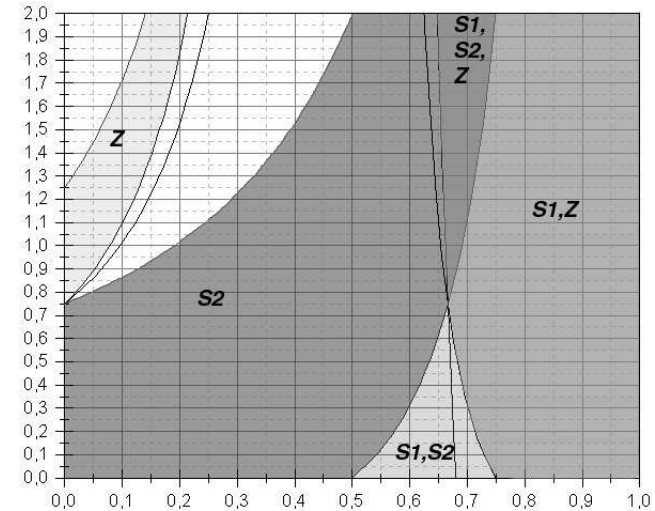


Рис. 2.190. Области существования и устойчивости стационарных состояний x_1^* и x_2^* и зигзагообразного поведения цепочки (2.307). $S1$ и $S2$ — области существования устойчивых точек x_1^* и x_2^* , соответственно. Z — область существования устойчивых зигзагообразных решений (2.334)

Цепочка связанных отображений (2.307) в широком диапазоне параметров (a, ε) обладает одним замечательным типом периодического пове-

дения. При $a \in [1, 6; 2, 0]$ она проявляет так называемую зигзагообразную динамику, когда всем элементам (либо большинству из них) присуще периодическое поведение периода 2 как по пространству так, и по времени, а соседние элементы осциллируют в противофазе.

Проанализируем подобное состояние аналитически. Период 2 в каждом элементе выражается двумя значениями x_1 и x_2 . Цикличность по пространству и по времени отражают соотношения

$$\begin{aligned} x_2 &= (1 - \varepsilon)f(x_1) + \varepsilon f(x_2), \\ x_1 &= (1 - \varepsilon)f(x_2) + \varepsilon f(x_1), \end{aligned} \quad (2.333)$$

где для удобства иначе переписана правая часть (2.307). Решая эти соотношения и исключая при этом однородные состояния $x_1 = x_2$, получим:

$$x_{1,2} = \frac{1 \pm \sqrt{4(1 - 2\varepsilon)^2 a + 4\varepsilon - 3}}{2(1 - 2\varepsilon)a}. \quad (2.334)$$

Устойчивость периодического состояния нетрудно найти из анализа собственных значений произведения матриц $DF(\hat{\xi}_1)DF(\hat{\xi}_2)$, где DF определяется соотношением (2.328), а $\hat{\xi}_1 = (x_1, x_2, \dots, x_1, x_2, \dots)$, $\hat{\xi}_2 = (x_2, x_1, \dots, x_2, x_1, \dots)$. Очевидно, произведение треугольных матриц $DF(\hat{\xi}_1)$ и $DF(\hat{\xi}_2)$ — треугольная матрица DF^2 . На диагонали DF^2 расположены произведения диагональных элементов матриц $DF(\hat{\xi}_1)$ и $DF(\hat{\xi}_2)$. Поэтому собственные значения матрицы DF^2 можно записать в виде

$$\lambda_i \equiv \lambda = (1 - \varepsilon)^2 f'(x_1)f'(x_2) = (1 - \varepsilon)^2 4a^2 x_1 x_2.$$

Отсюда, с учетом (2.334), приходим к выражению:

$$\lambda = \frac{4(1 - \varepsilon)^2}{(1 - 2\varepsilon)^2} - 4(1 - \varepsilon)^2 a.$$

Окончательно для построения области, где наблюдается зигзагообразное поведение (x_1, x_2) , необходимо учесть систему неравенств

$$\begin{cases} 4(1 - 2\varepsilon)^2 a + 4\varepsilon - 3 \geq 0, \\ -1 < \lambda < 1, \end{cases} \quad (2.335)$$

отражающую условия существования и устойчивости такого поведения. Условия $a \in [0; 2]$, $\varepsilon \in [0; 1]$ приводят к окончательному соотношению:

$$\begin{cases} \frac{4(1 - \varepsilon)^3 - (1 - 2\varepsilon)^2}{4(1 - \varepsilon)^2(1 - 2\varepsilon)^2} < a < \frac{4(1 - \varepsilon)^3 + (1 - 2\varepsilon)^2}{4(1 - \varepsilon)^2(1 - 2\varepsilon)^2}, \\ a \geq \frac{3 - 4\varepsilon}{4(1 - 2\varepsilon)^2}. \end{cases} \quad (2.336)$$

Область, определяемая соотношением (2.336), показана на рис. 2.190, где она обозначена через Z . Как следует из этого рисунка, область Z может перекрываться с областями $S1$ и $S2$ существования устойчивых точек. Это означает, что при некоторых фиксированных значениях параметров a и ε из областей пересечения предельное состояние цепочки (2.307) зависит от начального распределения $\{x_0(k)\}_{k=1}^K$. Динамику таких систем называют мультистабильной.

Во всех вычислениях стационарных и периодических состояний и их устойчивости (см. (2.329), (2.331), (2.333) и (2.335)) не учитывалось значение граничного условия $x(0) \equiv 0$. Иначе говоря, все результаты были получены в отсутствие границы, т. е. при $K \rightarrow \infty$. При конечном числе элементов граничное условие можно рассматривать как постоянное внешнее возмущение связанных отображений, которое корректирует их динамику.

Очевидно, богатство возможных режимов поведения системы (2.307) с $f(x, a) = 1 - ax^2$ отнюдь не ограничивается решениями (2.330) и (2.334). Для иллюстрации разнообразия динамики приведем результаты численного анализа данной модели, состоящей из $K = 100$ элементов на основе критерия, описанного выше.

1. $a = 1, 7$; $\varepsilon = 0, 45$. В цепочке наблюдается временная периодичность и удвоение периода по пространству при передвижении от элемента к элементу слева направо. Это иллюстрирует рис. 2.191, где показаны значения показателя λ и p в зависимости от номера элемента (горизонтальная ось) и времени в масштабе τ (вертикальная ось). Чем меньше λ и p , тем темнее соответствующая точка на графике. Самым темным точкам диаграммы для λ соответствует $\lambda = 0$. Легко видеть, что значения λ для всех элементов и во все моменты времени нулевые, а величина периодов удваиваются слева направо вплоть до $p = 32$.

Для анализа пространственной картины на рис. 2.191 также показаны величины $x_n(k)$ в некоторые последовательные моменты времени n . Интуитивно ясно, что в данной цепочке наблюдается пространственный хаос.

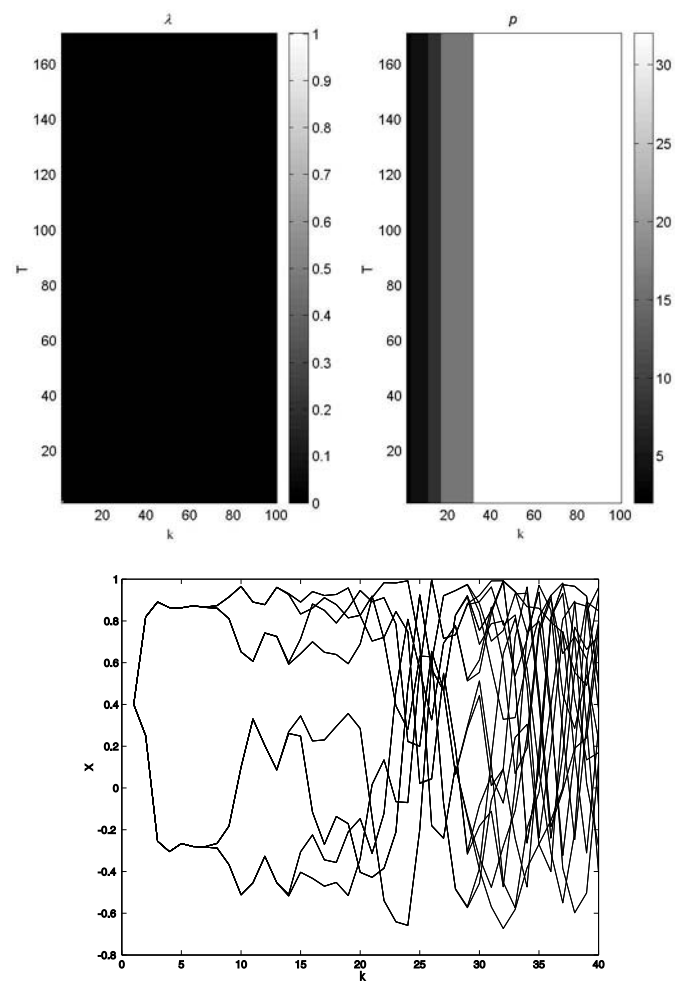


Рис. 2.191. Динамика системы (2.307) с $f(x, a) = 1 - ax^2$, состоящей из $K = 100$ элементов при $a = 1,7$ и $\varepsilon = 0,45$. Слева показаны значения показателя λ в зависимости от номера элемента (горизонтальная ось) и периода p в масштабе τ (вертикальная ось). Внизу представлено распределение значений $x_n(k)$ в некоторые последовательные моменты времени n

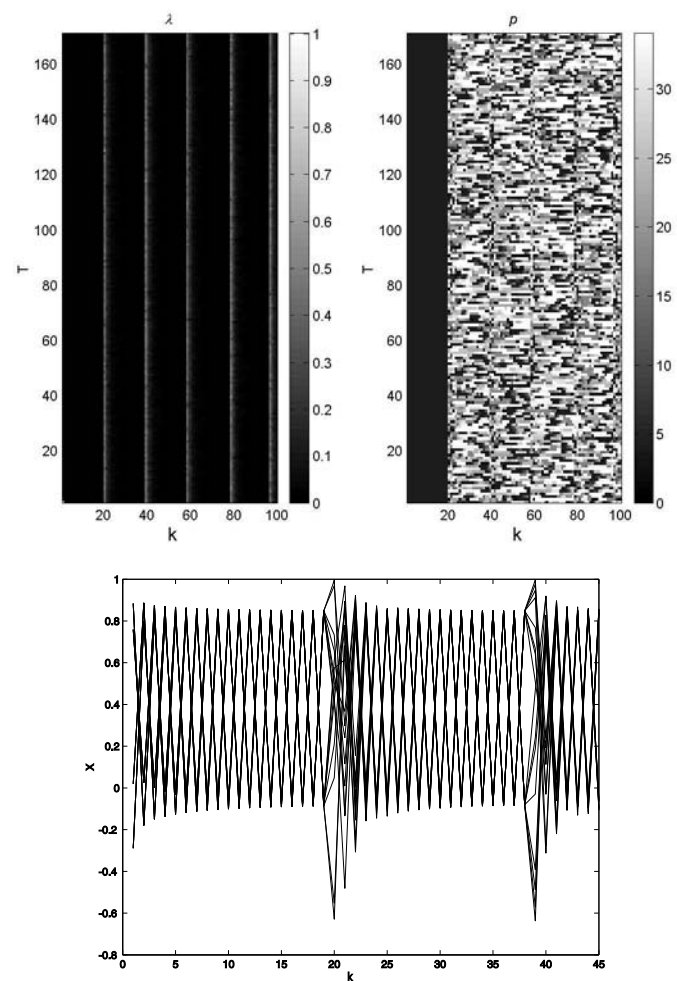


Рис. 2.192. То же, что на рис. 2.191, но при $a = 1,7$ и $\varepsilon = 0,11595$

Более строго это можно подтвердить на основе изучения стационарных состояний при помощи специально построенного отображения (см. [380]). Детальное исследование такого состояния показало, что значение макси-

мального периода, до которого происходит удвоение, зависит лишь от a и ε , а номер элемента, после которого прекращается дальнейшее удвоение, увеличивается с ростом K .

2. $a = 1,7$; $\varepsilon = 0,11595$. При данных значениях степени нелинейности и силы связи авторами работы [381] было обнаружено поведение, при котором большинство элементов проявляют зигзагообразную динамику, а отдельные единичные точки, равноотстоящие друг от друга, — хаотическую. При этом расстояние между хаотическими элементами логарифмически зависит от величины $\varepsilon - \varepsilon_c$, где $\varepsilon_c \approx 0,11525$ — некоторое критическое значение. Описанное явление отражено на рис. 2.192. Хаотические элементы выражены четкими «мигающими» линиями, расстояние между которыми может быть легко найдено.

32.3.2 Диффузионно связанные отображения

Исследуем теперь одномерную однородную систему (2.305) диффузионно связанных отображений вида $f(x, a) = 1 - ax^2$. Такая модель интенсивно исследовалась в ряде работ (см., например, [376, 377, 513]), где было показано, что она обладает большим разнообразием возможных режимов поведения: от синхронизации пространственных структур до полностью развитой турбулентности.

Аналогично цепочке (2.307), динамика системы (2.305) существенно зависит от значений параметров a и ε . Рассмотрим некоторые из видов ее поведения, анализируя траектории точек путем вычисления показателей λ и значений периодов p в различные промежутки времени.

1. $a = 1,44$; $\varepsilon = 0,1$. В этом случае, в зависимости от начальных условий, вдоль цепочки образуются устойчивые во времени области с регулярной и нерегулярной динамикой. Несмотря на диффузионную связь между элементами, пространственная картина также является устойчивой. Такое поведение можно назвать состоянием «замороженных» структур. Типичная картина представлена на рис. 2.193.

2. $a = 1,88$; $\varepsilon = 0,3$. Увеличение степени нелинейности, задаваемой параметром a в отдельном отображении $f(x, a)$, ведет к разрушению всех пространственных областей синхронизации. При этом динамика всей цепочки, равно как и ее отдельных элементов, быстро меняется как во времени так и в пространстве, что совпадает с результатами работы [376]. Это хорошо отражает величина λ (рис. 2.194). Аналогичная картина наблюдается и для распределения периодов p отдельных элементов. В качестве подтверждения на этом же рисунке показаны значения $x_n(k)$ в несколько последовательных моментов времени n . Обычно такое поведение цепочек

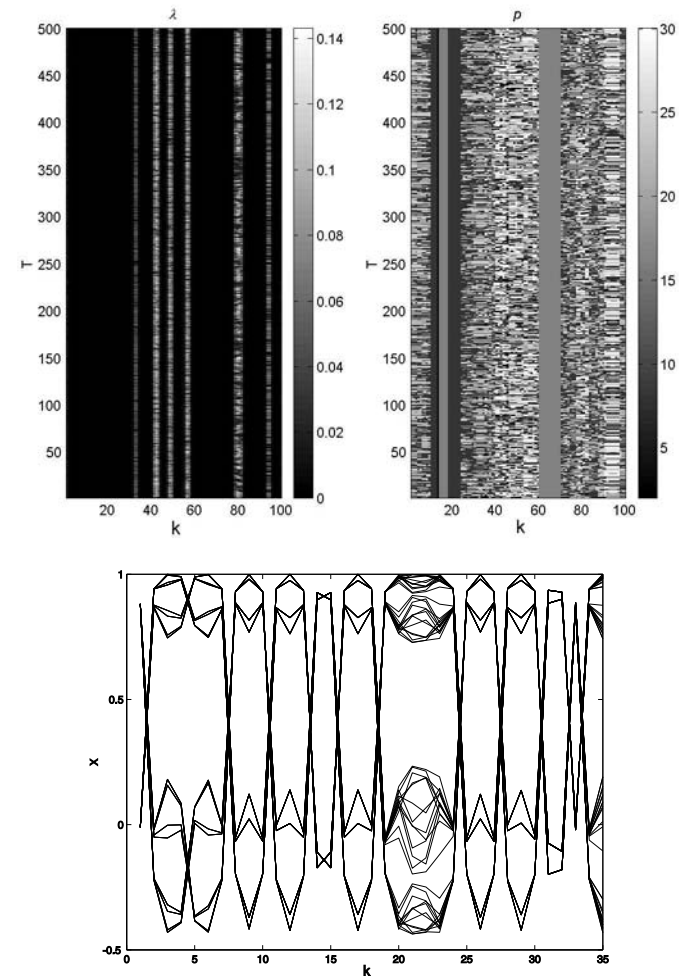


Рис. 2.193. Динамика кольцевой цепочки (2.305) с $f(x, a) = 1 - ax^2$ при $a = 1,44$ и $\varepsilon = 0,1$

связанных отображений называют пространственно-временным хаосом или полностью развитой турбулентностью [376].

3. $a = 1,8$; $\varepsilon = 0,3$. В этом случае при случайно выбранных начальных условиях области регулярной и хаотической динамики перемежаются

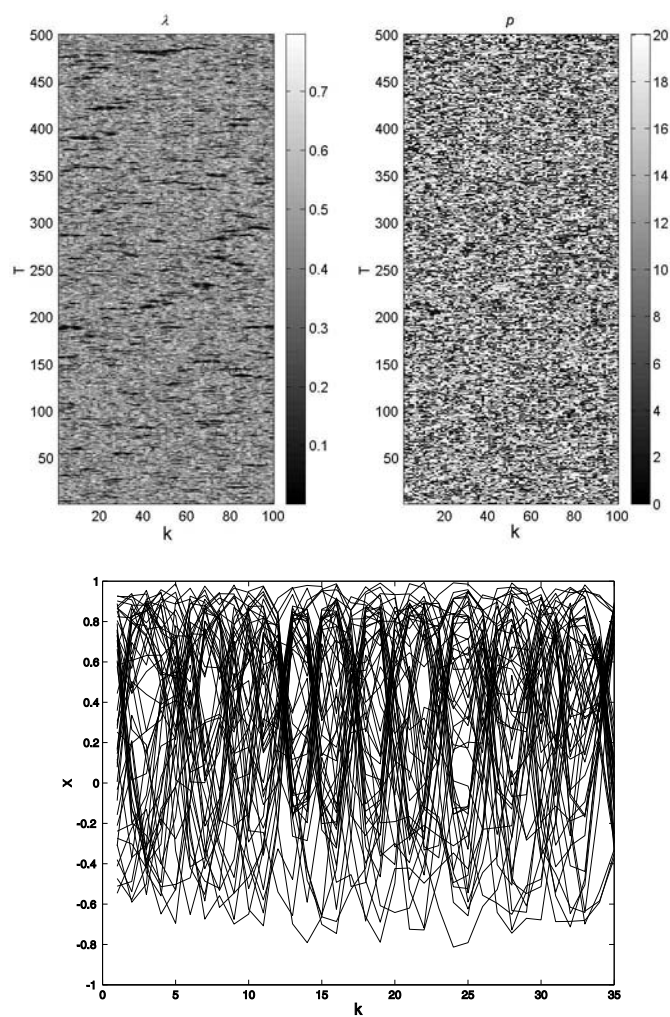


Рис. 2.194. То же, что на рис. 2.193, но при $a = 1,88$ и $\varepsilon = 0,3$

между собой. Подобное поведение наблюдается довольно длительное время. Однако в конечном счете почти все элементы цепочки начинают вести себя регулярно. Анализ показывает, что цепочка стремится разделиться на

ячейки размером по два элемента, в каждой из которых динамика имеет период 2.

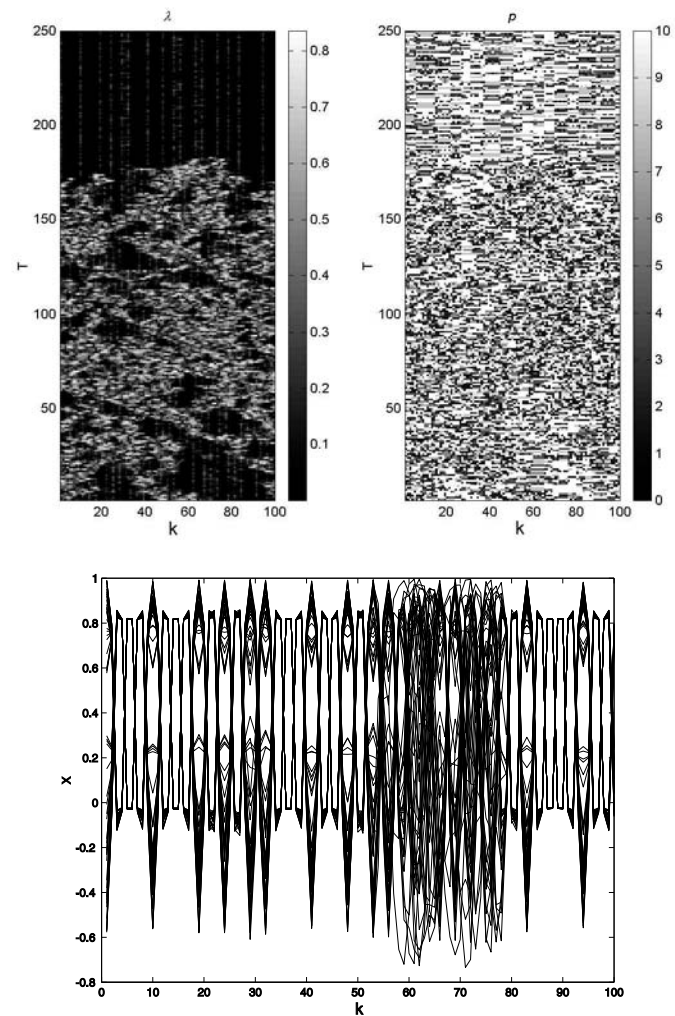


Рис. 2.195. То же, что и на рис. 2.193, при $a = 1,8$ и $\varepsilon = 0,3$

Это демонстрирует рис. 2.195, где показано поведение системы (2.305) после предварительных $2 \cdot 10^6$ итераций.

32.4 Неоднородные системы связанных отображений

Как отмечалось выше, однородные цепочки, составленные из одинаковых отображений, — это определенное приближение, введенное для упрощения исследования. Для описания реальных сред необходимо использовать *неоднородные* модели, включающие элементы с другим значением параметров или вовсе иной природы. Очевидно, с физической точки зрения система с различными значениями параметров вдоль цепочки является более адекватным представлением неоднородных сред.

Формально наличие хотя бы одного элемента с таким дефектом увеличивает количество управляющих параметров. Это приводит к тому, что возрастает размерность параметрического пространства решетки как динамической системы, что, в свою очередь, ведет к расширению множества возможных режимов движения.

В данной части мы покажем при помощи разработанного метода (см. раздел 32.2), к чему приводит наличие определенных дефектов в одномерных системах диффузионно связанных отображений.

1. Рассмотрим цепочку (2.305) с $f(x, a) = 1 - ax^2$ длиной $K = 100$, у которой все элементы (кроме центральных) имеют одно значение параметра, $a = a_1 = 1,44$, а центральные (с номерами $k = 47-53$) — иное, $a = a_2 = 1,97$. При $a = a_1 = 1,44$ и широком диапазоне значений ε в *однородной* цепочке (2.305) наблюдается регулярное поведение в виде пространственных синхронизованных структур малого размера (см. рис. 2.193). В случае $a = a_2 = 1,97$ она проявляет все свойства пространственно-временного хаоса. Таким образом, условно говоря, дефект состоит в том, что в цепочку с регулярным поведением вкраплены несколько «хаотических» элементов.

Динамика такой цепочки после предварительных 10^5 итераций представлена на рис. 2.196. Видно, что центральные элементы синхронизованы с периодом 2, а хаос перемещается на другие элементы. Однако, если провести $5 \cdot 10^6$ предварительных итераций, то хаос будет локализован в элементах с $a_2 = 1,97$ и некоторых соседних элементах справа и слева. Это говорит о том, что хаотическое поведение благодаря диффузии распространяется на элементы с регулярным поведением.

2. Рассмотрим теперь другой случай неоднородности. Напомним, что система (2.305) при $a = 1,8$ и небольших значениях диффузии ε проявляет перемежаемость синхронизации и нерегулярного поведения как в пространстве так и во времени (см. рис. 2.195). Построим цепочку из $K = 100$

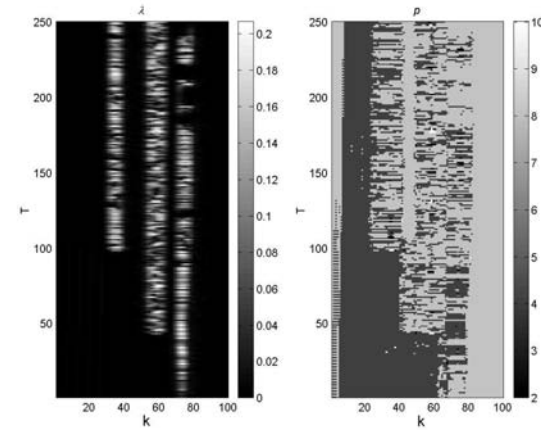


Рис. 2.196. Динамика неоднородной цепочки (2.305) после предварительных 10^5 итераций. Центральные элементы с номерами $k = 47-53$ имеют значение параметров $a = a_2 = 1,97$. Остальные — $a = a_1 = 1,44$. Величина диффузии $\varepsilon = 0,7$

элементов так, чтобы 50 из них имели одно значение нелинейности ($a = a_1 = 1,8$), а остальные 50 — другое ($a = a_2 = 1,97$), причем элементы с различными a последовательно чередовались в пространстве:

$$x_{n+1}(k) = \begin{cases} (1-\varepsilon)f(x_n(k), a_1) + \frac{\varepsilon}{2}(f(x_n(k-1), a_1) + f(x_n(k+1), a_1)), & k - \text{четн.}, \\ (1-\varepsilon)f(x_n(k), a_2) + \frac{\varepsilon}{2}(f(x_n(k-1), a_2) + f(x_n(k+1), a_2)), & k - \text{нечетн.} \end{cases} \quad (2.337)$$

На относительно небольших временах $T \approx 10^4$ такая неоднородная система проявляет пространственно-временной хаос, аналогичный однородной цепочке при $a = 1,97$. Однако если асимптотически проанализировать динамику (скажем, провести $5 \cdot 10^7$ итераций), то можно обнаружить неожиданное явление: вся цепочка синхронизируется. Это выражается в том, что в пространстве значений параметра λ наблюдается регулярная картина (рис. 2.197): всплески нерегулярности возникают строго через равные промежутки времени и на одинаковом расстоянии друг от друга, т.е. со временем хаос словно перемещается по цепочке с постоянной скоростью.

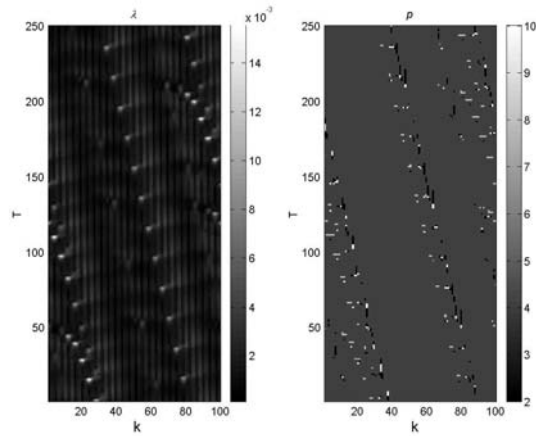


Рис. 2.197. Динамика неоднородной цепочки (2.337) с чередованием параметра нелинейности a после предварительных 10^7 итераций. $a_1 = 1,8$; $a_2 = 1,97$ и $\varepsilon = 0,7$

На рис. 2.197 видны также небольшие «загнутые» ответвления от прямых линий. Это означает, что хаотические всплески движутся не равномерно, а с небольшим ускорением. Однако через некоторое время они исчезают. Если подробно исследовать это явление, то можно обнаружить, что скорость перемещения хаотических дефектов существенно зависит от значений a_1 и a_2 , а также от диффузии ε , но не зависит от распределения начальных условий. От начальных условий зависит только их направление движения.

3. В заключение опишем динамику неоднородной кольцевой цепочки (2.305) с *единственным* дефектом и сравним ее с однородной. Пусть все элементы, кроме выделенного, имеют общую степень нелинейности $a = a_1 = 1,8$, а выделенный элемент $a = a_2 = 1,99$. Выберем небольшой коэффициент диффузии $\varepsilon = 0,1$.

Сначала рассмотрим систему без дефекта (рис. 2.198). В течение всего времени наблюдения большинство элементов синхронизованы, но при этом существуют мелкие «хаотические» структуры, которые блуждают в пространстве. Как хорошо видно на рис. 2.198, встречаясь, они взаимно уничтожаются. Соответственно и рождаются они парами. Подобное поведение в работе [376, 377] сравнивалось с аналогом броуновского движения.

Теперь введем в систему дефект: увеличим нелинейность до значения $a = a_2 = 1,97$ в элементе с номером $k = 75$. Исследования показыва-

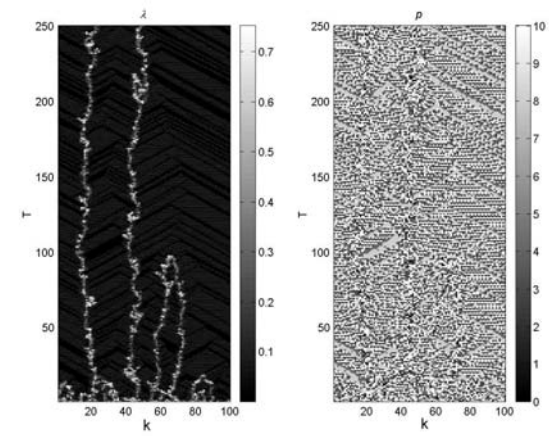


Рис. 2.198. Динамика однородной цепочки (2.305) при $a = 1,97$ и $\varepsilon = 0,1$

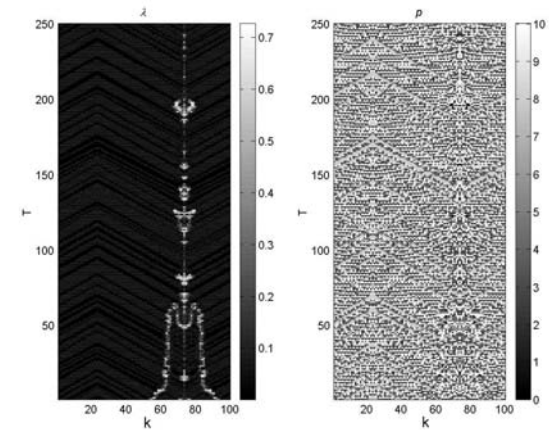


Рис. 2.199. То же, что на рис. 2.198, но при наличии дефекта с $a = a_2 = 1,97$ в элементе с номером $k = 75$

ют, что его наличие приводит к похожей картине, за исключением того, что в выделенном элементе всегда существует расстройка синхронизации, а количество блуждающих структур увеличивается (рис. 2.199). Их дина-

мика полностью повторяет динамику структур в однородной системе с тем лишь отличием, что они могут также исчезать, встречаясь с дефектом ($k = 75$), и рождаться на нем. Условно говоря, элемент с дефектом является зародышем-поглотителем «броуновских» частиц. В связи с этим интересно выяснить состояние такой системы на больших временах. Оказывается, что независимо от начальных условий, предельное состояние является *симметричным*.

На рис. 2.199 представлена подобная картина через $5 \cdot 10^7$ итераций. «Хаотические частицы» рождаются парами на дефекте и движутся по симметричным относительно него траекториям. Они также исчезают путем столкновения друг с другом и с дефектом. Отметим, что хаотические элементы являются причиной небольшой расстройки синхронизации во всех элементах. Эта расстройка распространяется по пространству в разные стороны и сходится прямо противоположно (на кольце) относительно дефекта.

Приведенные примеры показывают, что неоднородные цепочки отображений обладают непредсказуемым поведением и большим разнообразием возможных режимов движения. Таким образом, разработанный метод визуализации динамики распределенной среды, аппроксимируемой решетками сцепленных отображений, является эффективным способом исследования пространственно-временной эволюции. Он выгодно отличается от других прежде всего тем, что позволяет наблюдать в системе развитие хаоса как во времени, так и по пространству.

33. Временные ряды: анализ и прогноз²⁰

Из изложенного выше материала следует один важный вывод: несмотря на то, что нелинейные системы могут значительно отличаться в конкретных проявлениях и деталях, существуют глубокие аналогии в их организации и функционировании. Это предопределило интерес к методам, которые развиваются в рамках теории сложных систем, как универсальному инструменту исследования объектов самой различной природы. Наиболее явно такой синергетический подход проявляется при анализе временных рядов.

Как известно, большинство систем (природных, таких, например, как атмосфера, или искусственных, таких, например, как биржа) в силу их сложности не могут быть смоделированы с достаточной точностью. Однако их описание может быть выполнено посредством иного подхода, основанного на наблюдении за их поведением (см. §31). Наблюдаемая (сигнал, реализа-

²⁰Результаты этого параграфа получены одним из авторов (А. Ю. Л.) совместно О. Л. Котляровым и И. А. Истоминым.

ция) — это функция от времени, по которой судят о процессе в исследуемой системе. Иными словами, наблюдаемая — это временной ряд. Например, для атмосферы в качестве наблюдаемой может выступать, скажем, изменение температуры, для биржи — ежедневный курс ценных бумаг и т. п. Если такую наблюдаемую определенным образом обработать, то при некоторых условиях возможно с большой точностью произвести оценку ее будущего значения.

В данном параграфе рассмотрена возможность использования методов синергетики и нелинейной динамики для прогноза временных рядов различной природы. Эти методы, в основном, можно разделить на локальные и глобальные. Такое деление проводится по области определения параметров так называемой аппроксимирующей функции [141]. Исторически первыми были разработаны глобальные методы, поэтому рассмотрение начнем именно с них.

33.1 Глобальные методы

В глобальных методах параметры аппроксимирующей функции идентифицируются посредством использования всех известных значений ряда. Основное приложение таких методов — получение глобальных характеристик системы. При этом прогнозирование если и необходимо, то главным образом как дополнительная информация. Одним из подходов здесь является использование авторегрессионных моделей.

33.1.1 Авторегрессия

Методы, основанные на авторегрессии, уже достаточно давно применяются в статистике, экономике и метеорологии. Их основная идея — выражение следующих значений ряда через предыдущие. Вероятно, это практически единственный способ, который можно использовать в ситуации, когда никакой другой информации о системе, кроме заключенной в предыдущих значениях ряда, нет.

Линейная модель авторегрессии порядка p имеет вид:

$$x_{t+1} = a_1 x_t + a_2 x_{t-1} + \dots + a_p x_{t-p+1} + \xi_{t+1}, \quad (2.338)$$

где a_1, \dots, a_p — некоторые постоянные коэффициенты, т. е. параметры модели, $\{\xi_t\}$ — последовательность случайных величин, образующая белый шум, и x_t — значение ряда на шаге t . Модель авторегрессии порядка p относится к классу линейных параметрических моделей.

Идентификация модели включает в себя статистическую оценку ее параметров, основанную на решении системы линейных уравнений Юла–Уокера [371], связывающих между собой неизвестные параметры модели и автокорреляции анализируемого ряда. При этом оценки параметров получаются глобальными, т. е. едиными для всего ряда. Однако прежде чем оценить сами параметры в рамках идентификации модели, нужно определить количество этих параметров, т. е. выбрать порядок авторегрессии. Обычно выбор осуществляется по виду частной автокорреляционной функции (см. [309]). Прогнозирование с использованием авторегрессионных моделей, как правило, сводится к рекуррентному расчету значений ряда, следующих за последним известным значением по формуле (2.338).

В случаях, требующих быстрого обновления прогноза на основе вновь поступивших данных, используются адаптивные методы прогноза. К ним относится, например, метод экспоненциального сглаживания (метод Брауна) [247]. Следуя ему, каждому значению ряда в процессе идентификации модели присваивается весовой коэффициент, экспоненциально убывающий со временем, отделяющим это значение от последнего известного значения ряда. Таким образом, самые «старые» значения ряда практически не влияют на результаты прогноза, тогда как последние известные величины имеют наибольший вес. Тем самым этот метод приближается к локальным методам, так как основной вклад при прогнозе дает лишь небольшая часть самых последних по времени значений ряда. Однако заметим, что в локальных методах, которые рассматриваются в следующем разделе, для построения прогноза также выбираются наиболее близкие значения к последнему известному значению ряда, но эта близость определяется не по времени.

Методы авторегрессии разработаны наиболее тщательно и применяются, как правило, в прикладных задачах (см., например, [171, 428, 493] и приведенные там ссылки). Они реализованы практически во всех программных пакетах статистической обработки данных. По этим причинам мы их не рассматриваем, а опишем один из достаточно новых и эффективных глобальных методов — так называемый сингулярный спектральный анализ.

33.1.2 Сингулярный спектральный анализ

Метод сингулярного спектрального анализа (SSA) имеет строгое обоснование в рамках теории динамических систем и используется главным образом для определения основных составляющих временного ряда, а так-

же для подавления шума [125, 318, 519]. Однако существуют основанные на этом методе оригинальные алгоритмы прогноза [68], а также прогнозы по методу авторегрессии.

В основе SSA, как, впрочем, и большинства подходов, связанных с обработкой временных рядов $\{x_1, \dots, x_N\}$, явно или неявно лежит построение множества векторов задержек [245]:

$$\mathbf{x}_t = (x_t \ x_{t-1} \ \dots \ x_{t-p+1})^T, \quad t = p, p+1, \dots, N. \quad (2.339)$$

Метод задержек, посредством которого строятся указанные векторы, устанавливает переход от исходного одномерного (скалярного) временного ряда к многомерному (векторному) представлению, сходному с применяемым в авторегрессии. При этом каждый многомерный вектор образуется из некоторого числа p следующих друг за другом значений исходного временного ряда. Результат можно представить в виде набора «фотографий» ряда, сделанных через скользящее вдоль ряда окно, в которое одновременно попадает лишь p последовательных значений ряда:

$$\mathbf{X}_{p \times (N-p+1)} = \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ \left[\begin{array}{c} x_{p+1} \\ x_p \\ \vdots \\ x_2 \\ x_1 \end{array} \right] & \left[\begin{array}{c} x_{p+2} \\ x_{p+1} \\ \vdots \\ x_3 \\ x_2 \\ x_1 \end{array} \right] & \dots & \left[\begin{array}{c} x_N \\ \vdots \\ x_{N-p+2} \\ x_{N-p+1} \\ x_{N-p} \\ \vdots \end{array} \right] \end{pmatrix}$$

Здесь каждая квадратная скобка — вектор в p -мерном пространстве задержек; последовательность таких векторов задает матрицу наблюдений $\mathbf{X}_{p \times (N-p+1)}$, где N — число элементов исходного ряда. Эта матрица, в каждом столбце которой стоят части одного и того же ряда, сдвинутые друг относительно друга, и будет многомерным представлением исходного скалярного ряда в пространстве задержек. Она называется *матрицей задержек*.

Особенностью SSA является обработка матрицы \mathbf{X} по алгоритму, близкому к методу главных компонент [2]. Использование этого алгоритма является основным положением SSA, отличающим его от других методов нелинейной динамики, применяемых для анализа и прогноза временных рядов²¹.

²¹В методах локальной аппроксимации, рассматриваемых далее, также возможно приме-

Суть метода главных компонент состоит в снижении размерности исходного пространства факторов (у нас это пространство задержек) посредством перехода к более информативным переменным (координатам). Полученные таким образом новые переменные и называют главными компонентами. Этот переход осуществляется с помощью ортогонального линейного преобразования. Главные компоненты обладают многими важными свойствами. В SSA получающееся разложение используется для выделения наиболее значимых составляющих ряда и отсева случайных возмущений. Применение этого метода позволяет сгладить исходный ряд, снизить уровень случайных возмущений, повысить отношение сигнал/шум.

Реализация возможности прогнозирования на базе SSA осуществляется двумя основными способами. Первый имеет собственное название «Гусеница» [68] и использует специфические особенности SSA-разложения. Второй [519] предусматривает использование для прогнозирования метод авторегрессии, который применяется по отдельности к каждой из выбранных компонент разложения.

В качестве иллюстрации рассмотрим два примера, в которых метод SSA-«Гусеница» применялся для прогнозирования временных рядов.

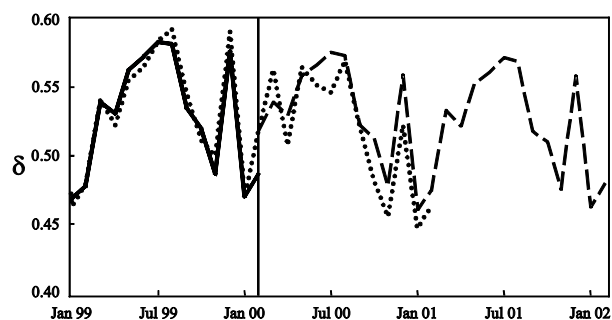


Рис. 2.200. Фактическая динамика, реконструкция и прогноз отношения товарных запасов к продажам товаров длительного пользования в США [325]

Первый пример взят из работы [325] и демонстрирует возможности метода для прогноза временного ряда отношения товарных запасов к продажам товаров длительного пользования в США. Для построения использовались месячные данные за период с января 1981 года по февраль 2000 года

нение метода главных компонент, но лишь для обработки отдельных столбцов матрицы X , выбранных в соответствии с определенным критерием.

включительно. Прогноз строился на период с марта 2000 года до февраля 2002 года и затем сравнивался с фактическими данными (но только до февраля 2001 года). Результат прогноза иллюстрирует рис. 2.200. Точками показаны фактические значения отношения, сплошной линией — результат реконструкции и штриховой — прогноз. Видно, что все двенадцать прогнозируемых значений оказались достаточно точными, остальные на момент написания работы [325] сопоставить с фактическими данными было невозможно ввиду отсутствия последних.

Второй пример связан с прогнозированием солнечной активности, характеризуемой значениями чисел Вольфа [125]. Прогноз на 18 лет, т.е. примерно на полтора 11-летних солнечных цикла оказался весьма точным (см. рис. 2.201). Он был выполнен по тому же принципу, что и в предыдущем примере, т.е. последние 216 значений ряда (12 месяцев \times 18 лет) — на рисунке они отделены вертикальной линией — не учитывались при реконструкции ряда и построении прогноза, а использовались только для оценки точности прогноза.

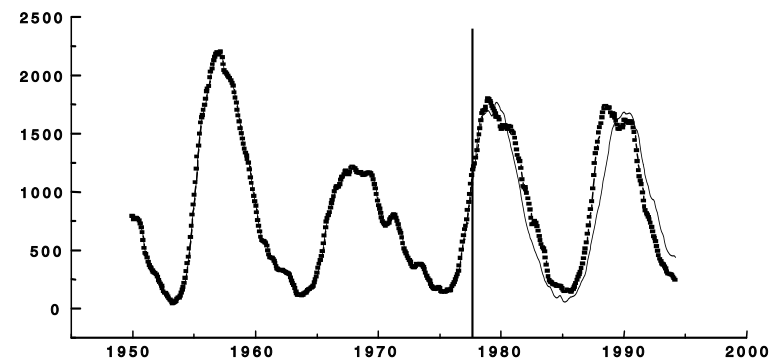


Рис. 2.201. Прогноз по методу SSA-«Гусеница» и реальные значения чисел Вольфа, характеризующие солнечную активность. Вертикальная линия отделяет начало прогноза [125]

33.2 Локальные методы

Локальные методы прогноза, основанные на принципе локальной аппроксимации (LA), имеют преимущество перед остальными методами исследования и прогноза временных рядов в задачах, связанных с прогнозированием *нерегулярных* (хаотических, квазипериодических) рядов. Принцип

LA предложен впервые в работе [299]. Его применение не требует априорной информации о системе, породившей временной ряд, и нет необходимости в построении специфической модели динамики исследуемого ряда. При этом возможность использования LA для прогнозирования нерегулярных временных рядов определяется базовым принципом, лежащим в его основе, т. е. локальностью строящейся аппроксимации. Этот принцип подразумевает отказ от явного использования для прогнозирования всех (уже известных) значений ряда и ограничивает их число лишь наиболее близкими в некотором смысле к стартовой точке, после которой начинается прогноз. Под стартовой точкой (стартовым вектором) понимается последнее или несколько последних известных значений ряда.

Применение локальной аппроксимации позволяет избежать главной проблемы, возникающей при прогнозировании нерегулярных временных рядов существующими глобальными методами (см., например, [31, 50, 243]). Эта проблема связана с трудностями адекватного описания динамики квазипериодических и хаотических временных рядов на основе глобальной линейной модели (2.338). Введение в указанную модель нелинейных членов требует резкого увеличения количества исходных данных для того, чтобы получить устойчивые оценки коэффициентов. В свою очередь увеличение количества данных означает увеличение длины интервала, на котором применяется модель. А чем длиннее интервал, тем больше параметров должно содержаться в модели, чтобы обеспечить ее адекватность на всем интервале (последнее характерно прежде всего для работы с экономическими данными).

Модель LA допускает использование вместо глобально-линейной аппроксимации кусочно-линейного приближения. При этом вид уравнения модели остается, как правило, линейным типа (2.338), но его параметры полагаются неизменными лишь в ближайшей окрестности стартового значения, причем близость определяется не во времени, а в пространстве задержек [245].

Наиболее важный результат перехода от глобальной аппроксимации к локальной при прогнозировании достаточно длинных хаотических временных рядов состоит в том, что длина прогноза ограничивается в первую очередь не возможностями метода (как при глобальной аппроксимации), а лишь особенностями динамики ряда.

Принцип локальной аппроксимации лег в основу целого класса методов прогноза временных рядов (локальная аппроксимация применяется исключительно в целях прогнозирования динамики ряда, а анализ основных закономерностей осуществляется, как правило, другими методами, например SSA), различающихся способами построения прогноза и приме-

няемыми моделями аппроксимации, но сохранившими общее название по объединяющему их базовому принципу. Классификация этих методов была предложена самими авторами принципа локальной аппроксимации [300], но вопрос о выборе оптимального метода (варианта метода) решается, как правило, эмпирически [256].

Ниже мы опишем универсальный подход, объединяющий все методы и варианты локальной аппроксимации и позволяющий выбрать наиболее оптимальный вариант с точки зрения использования данных, точности и устойчивости прогноза на основе анализа каждого из методов. Разработка единого алгоритма расчета по любому из вариантов LA может помочь в поиске способов повышения точности получаемых прогнозов (когда это возможно) и расширения области применения методов типа LA.

33.3 Общий алгоритм локальной аппроксимации

В данной части рассматривается первая и главная часть алгоритма LA, определяющая способ построения прогноза на один шаг по времени вперед. Возможность прогноза на большую длину описана ниже.

33.3.1 Построение прогноза на один шаг вперед

При анализе временных рядов главной задачей является реконструкция породившей этот ряд динамической системы [208, 284]. В соответствии с теорией Такенса–Мане [418, 504] приемлемое описание фазового пространства динамической системы можно получить, если взять вместо реальных переменных системы p -мерные векторы задержек (2.339) из значений ряда в последовательные моменты времени [245, 246]. При выполнении условия $p \geq 2d + 1$, где d — размерность вложения, возможно реконструировать фазовое пространство (пространство состояний) системы [418]. При условии стационарности временного ряда на базе этой реконструкции строится прогноз его дальнейшей динамики.

Главная проблема, возникающая при использовании многомерного представления временного ряда в виде матрицы задержек, — это определение величины p [141, 505]. На сегодняшний день наиболее часто используемым алгоритмом для оценки величины p является алгоритм Грассбергера–Прокаччия [326, 327] (см. раздел 31.3). Однако и он оказывается неэффективным при работе с короткими (до 10^4 точек) временными рядами. Другие методы тоже имеют свои недостатки: сложность реализации, большая длительность расчетов, неоднозначность либо сомнительность результатов [73]. В связи с этим величина p , за исключением модельных примеров, в которых она достоверно известна, как правило, определяется эмпирически. Главный

критерий в этом случае — выбор такого p , начиная с которого прекращается качественное изменение прогноза [206].

Ниже мы опишем основные этапы построения алгоритма LA.

• 1. Построение матрицы задержек и выбор локального представления

Первым шагом как алгоритма LA, так и других методов анализа временных рядов, основанных на принципах нелинейной динамики, является метод задержек [245, 318]. В результате его применения скалярный временной ряд, содержащий N значений наблюдаемой, превращается в матрицу задержек:

$$\{x_1, x_2, \dots, x_N\} \rightarrow \mathbf{X}_{p \times (N-p+1)} = \begin{pmatrix} x_p & x_{p+1} & \cdots & x_N \\ \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & \cdots & x_{N-p+2} \\ x_1 & x_2 & \cdots & x_{N-p+1} \end{pmatrix}.$$

Размерность матрицы \mathbf{X} определяется количеством задержек p .

После построения матрицы задержек выбирается вид локального представления, т. е. вид функции, связывающей следующие значения ряда с предыдущими:

$$x_{t+1} = f(x_t, \mathbf{a}), \quad (2.340)$$

где \mathbf{a} — вектор параметров представления. Наиболее распространенный вариант — это линейная аппроксимация типа (2.338):

$$x_{t+1} = a_0 + \mathbf{x}_t^T \mathbf{a}. \quad (2.341)$$

Кроме него, используются еще два варианта:

$$x_{t+1} = a_0 \quad (2.342)$$

и

$$x_{t+1} = a_0 + \mathbf{x}_t^T \mathbf{a} + \mathbf{x}_t^T \mathbf{B} \mathbf{x}_t, \quad (2.343)$$

где $\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{12} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1p} & b_{2p} & \cdots & b_{pp} \end{pmatrix}$ — матрица дополнительных параметров

представления. Следуя авторам работы [300], модели (2.341)–(2.343) будем называть аппроксимацией первого (LA1), нулевого (LA0) и второго (LA2) порядков, соответственно. Подробнее особенности каждого вида аппроксимации будут рассмотрены в п. 33.3.2.

• 2. Выбор ближайших соседей — выделение локальной подобласти фазового пространства

Как уже отмечалось, основная идея LA состоит в прогнозировании не по всем исходным данным, а по «ближайшим» к последнему известному значению вектора. После введения матрицы задержек легко определить понятие «ближайшие соседи» или просто «соседи».

Будем считать соседними к стартовому вектору \mathbf{x}_{N-p+1} векторы, удовлетворяющие условию:

$$\{\mathbf{x}_s\} : \sum_{s \in \omega_{\Xi}} \|\mathbf{x}_{N-p+1} - \mathbf{x}_s\| \rightarrow \min_{\omega_{N-p+1}}, \quad (2.344)$$

где $s \in \omega_{\Xi} \subset \omega_{N-p+1}$, $\omega_{N-p+1} \equiv \{1, \dots, N-p+1\}$, Ξ — количество соседей, ω_{Ξ} — набор номеров векторов-соседей. Норма обычно берется евклидова, хотя возможно использование и других норм. Выбор соседей автоматически определяет локальную подобласть, в которой параметры представления полагаются неизменными.

Допустим также и обратный вариант: выбрав некоторую окрестность, т. е. задавшись некоторым критерием вида

$$\|\mathbf{x}_{N-p+1} - \mathbf{x}_s\| \leq \varepsilon, \quad (2.345)$$

можно также определить соседей. Однако этот критерий практически не применяется, так как выбор величины ε является в достаточной мере произвольным, тогда как из-за необходимости оценки параметров на количество соседей налагается существенное ограничение снизу: $\Xi \geq p+1$. В реальных расчетах стараются придерживаться соотношения $\Xi \cong 3(p+1)$, чтобы получить более устойчивые оценки параметров. Таким образом, если выбрать условие (2.345) и подобрать величину ε , обеспечивающую нужное число соседей, все сведется к условию (2.344).

Данный этап алгоритма — главный отличительный признак метода и самый важный его этап. Выбор «хороших» соседей определяет качество прогноза.

• 3. Оценка параметров модели и построение прогноза на один шаг вперед

Выбором вида представления (2.341), (2.342) или (2.343) и размерности реконструкции p устанавливается количество неизвестных параметров, которые требуется оценить. При этом круг используемых данных ограничивается набором соседей стартового вектора.

Параметры модели (вектор \mathbf{a}) оцениваются методом наименьших квадратов (МНК). Это самый распространенный и наиболее эффективный метод

решения подобных задач, хотя и не единственный [78]. Оценка по МНК для вектора \mathbf{a} , обозначим ее как $\hat{\mathbf{a}}$, находится из условия:

$$\hat{\mathbf{a}} : \sum_{\omega_s} \left(x_{s+1} - f(\mathbf{x}_s, \hat{\mathbf{a}}) \right)^2 \rightarrow \min_{\hat{\mathbf{a}}} \quad (2.346)$$

Однако в большинстве случаев применение МНК в чистом виде невозможно из-за вырожденности матрицы факторов (подробнее об этом см. раздел 33.4), что является следствием взаимной близости соседей. Поэтому обычно применяется так называемое сингулярное разложение (SVD — Singular Value Decomposition) [91]. При этом не всегда учитывается, что оценки, даваемые этим методом, в общем случае смещенные [390], сильно зависящие от машинной точности и выбора минимального значащего сингулярного числа (последний факт упоминается редко). Следовательно, при использовании SVD важно всегда контролировать устойчивость получаемых результатов.

Оценив значения параметров аппроксимации, нетрудно построить прогноз следующего значения ряда (стартовый вектор, который в большинстве случаев совпадает с последним известным вектором, обозначим индексом L):

$$\hat{x}_{L+1} = f(\mathbf{x}_L, \hat{\mathbf{a}}). \quad (2.347)$$

Таким образом, пройдя три описанных шага алгоритма LA, можно построить прогноз одного нового значения ряда. Вопрос о способах построения более длинного прогноза подробно рассматривается в разделе 33.3.3.

Краткой иллюстрацией изложенного алгоритма может служить пример из работы [300]. Один из возможных способов построения прогноза температуры на следующий день состоит в том, чтобы, найдя в какой из предшествующих дней температура была максимально близкой к сегодняшней (случай $p = 1$), взять в качестве прогноза температуры на завтра ее величину в следующий за найденным день. Аппроксимация более высокого порядка позволяет учитывать влияние на прогноз отклонений «сегодняшней» температуры от температуры в день, наиболее похожий на сегодняшний.

33.3.2 Порядок аппроксимации

В этом разделе представлены результаты сравнения основных вариантов и способов прогноза, использующих принцип локальной аппроксимации. Сравнение проводится на временных рядах, полученных из системы Лоренца (см. раздел 20.3.1) с при $\sigma = 5$, $r = 15$, $b = 1$.

При ограниченном количестве данных важным фактором становится минимально необходимое количество точек для оценки параметров. Оно

зависит от порядка аппроксимации: для нулевого порядка достаточно одной точки, так как есть всего один неизвестный параметр a_0 (см. (2.342)). При линейной аппроксимации LA1 неизвестных параметров $p + 1$ (см. (2.341)). Роль порядка аппроксимации возрастает с уменьшением длины исходного ряда. Ниже будет показано, что для очень длинных рядов роль порядка аппроксимации значительно снижается.

Для получения надежных оценок количество соседей должно быть достаточно большим. Существует эмпирическое правило, в соответствии с которым для успешного применения МНК соседей должно быть как минимум втрое больше, чем число оцениваемых параметров. Таким образом, для LA1 в каждой локальной подобласти желательно иметь не менее $3(p + 1)$ известных векторов.

Порядок аппроксимации можно повышать и дальше, однако при этом резко возрастает количество необходимых соседей и увеличивается риск выхода из области справедливости гипотезы о постоянстве коэффициентов, т. е. происходит глобализация метода с потерей преимуществ локальной аппроксимации. Поэтому локальную аппроксимацию выше второго порядка (LA2) обычно не используют.

LA2 отличается от LA1 добавлением в модель парных произведений координат векторов-соседей (2.343):

$$x_{t+1} = a_0 + \sum_{i=1}^p a_i x_{t-p+i} + \sum_{j=1}^p \sum_{k \geq j}^p b_{jk} x_{t-p+j} x_{t-p+k}.$$

Минимальное количество соседей здесь равно $p + 1 + p(p + 1)/2$, и их число растет как квадрат размерности реконструкции. Поэтому в число соседей могут попасть векторы, которые очень непохожи на стартовый. Таким образом, главная причина неэффективности локальной аппроксимации старших порядков — одновременное использование большого числа векторов задержек. Результат использования большого числа соседей (часть которых с трудом можно так назвать) показан на рис. 2.202.

Сгенерированный ряд длиной 3600 точек разбивался на две части, первая из которых использовалась для оценки параметров аппроксимации, а по второй (последние 164 точки) проверялось качество прогноза. На рисунке показана только вторая часть исходного ряда и результаты прогнозов, полученные при разных порядках аппроксимации. Параметр p принят равным пяти. Число соседей взято втрое больше минимально необходимого, т. е. для LA0 — 3 соседа, LA1 — $3(5 + 1) = 18$ соседей, LA2 — 63. Видно, что на коротких интервалах времени LA0 уступает остальным в точности, а LA1 и LA2 дают близкие результаты. Естественно, по одному примеру

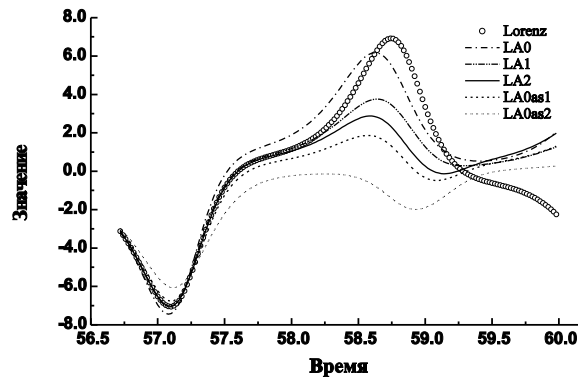


Рис. 2.202. Фактическая динамика и результат прогнозирования x -компоненты системы Лоренца методом LA нулевого, первого и второго порядков

нельзя судить о том, какой вариант (LA1 или LA2) лучше, но важно другое — роль числа соседей.

Для сравнения на этом же рисунке приведены прогнозы методом LA0 с тем же числом соседей, как для LA1 (линия LA0as1) и LA2 (линия LA0as2). Легко видеть, что с ростом количества соседей происходит «сглаживание» прогноза LA0, вызванное увеличением различий между векторами, попавшими в число соседей стартового вектора, а следовательно, и их прогнозами. Поэтому использование LA старших порядков (начиная со 2-го) при экстраполяции достаточно коротких рядов вряд ли может быть оправдано.

Таким образом, локальная аппроксимация первого порядка имеет по сравнению с аппроксимацией второго порядка два существенных преимущества. Во-первых, она требует меньше данных, что важно для коротких рядов. Во-вторых, получить более точное приближение легче в малой области, чем в большой, так как в первом случае различия между величинами менее выражены.

33.3.3 Прогноз на несколько шагов вперед

Рассмотренный в разделе 33.3.1 общий алгоритм метода LA завершился построением прогноза на один шаг. Чаще всего необходим более долгосрочный прогноз. Метод LA позволяет его построить. Однако к его результатам надо относиться осторожно, постоянно контролируя устойчивость прогноза по отношению к небольшим изменениям исходного ряда. Поэтому лучше заранее оценить максимально возможную длину прогноза.

В работе [300] предложено два способа, основанных на принципе локальной аппроксимации и позволяющих строить прогнозы динамики ряда на несколько шагов: итеративный и прямой. Рассмотрим сначала их алгоритмы, а сравнение отложим до построения общей модели (см. п. 33.4). К этим двум методам также можно добавить еще один способ — итеративный вариант с пересчетом, который также будет рассмотрен в этом разделе.

• Итеративный способ прогноза

Пусть требуется построить прогноз на S шагов вперед. Идея итеративного способа состоит в последовательном построении прогноза на один шаг с добавлением его результата к исходным данным. Иными словами, сначала делается прогноз на один шаг, а затем на основе полученного значения строится новый вектор в пространстве задержек:

$$\hat{x}_{L+1} = \begin{pmatrix} \hat{x}_{L+1} \\ x_L \\ \vdots \\ x_{L-p+2} \end{pmatrix}. \quad (2.348)$$

Этот вектор подставляется в правую часть соотношения (2.347) и делается прогноз еще на один шаг. И так далее. При этом параметры \hat{a} в (2.347) выбираются те же, что и на первом шаге. Таким образом, каждое следующее прогнозируемое значение выражается через предыдущее (предыдущие). Отсюда название способа — итеративный.

Основное преимущество данного способа — простота и легкость прогнозирования следующих значений, так как оценка параметров модели проводится всего однажды — на первом шаге. В работе [300] обоснована предпочтительность именно итеративного способа при работе с *хаотическими* системами. Однако на практике он не всегда оказывается лучшим. Основная причина невысокой точности прогноза, получаемого этим способом, связана, по-видимому, с неявно используемой предположкой о том, что соседями всех строящихся по принципу (2.348) новых векторов задержек являются те же векторы, которые были соседями стартового. Этот факт следует из неизменности параметров в уравнении (2.347), что, очевидно, не всегда верно.

• Итеративный способ с пересчетом параметров

Можно попытаться улучшить итеративный способ, отказавшись от предположки идентичности соседей всех новых прогнозируемых векторов. При этом требуется пересчитать заново на каждом шаге параметры в выражении (2.347). Следовательно, чтобы построить прогноз на два шага, после получения прогноза на один шаг и определения нового вектора (2.348)

в пространстве задержек, который будет теперь считаться стартовым, необходимо полностью повторить процедуру прогнозирования на один шаг и так далее. В этом случае время на построение длительного прогноза может сильно возрасти по сравнению с простым итеративным способом.

Вариант с пересчетом параметров свободен от основного недостатка простого итеративного способа, но так же имеет свои минусы. Главный — это высокая вероятность ошибочного выбора соседей. Поскольку выбор соседей в этом варианте проводится на каждом шаге, то также на каждом шаге возрастает ошибка прогноза, связанная с вероятностью неправильного выбора. Вследствие этого, особенно при работе со стохастическими рядами, высока вероятность перехода на смежную траекторию. Контроль правильности выбора «нужных» соседей затруднителен, и поэтому применение рассматриваемого способа также может не обеспечивать желаемой надежности прогноза.

• Прямой прогноз

При прямом способе прогноза на S шагов вперед стартовый вектор и все его соседи остаются неизменными, а прогноз строится независимо для всех $t \in [1, \dots, S]$:

$$\hat{x}_{L+t} = f(x_L, \hat{a}^{(t)}).$$

Таким образом, для выбранных векторов-соседей строится зависимость значения ряда через t шагов от координат исходного вектора, оцениваются параметры этой зависимости, т. е. коэффициенты $\hat{a}^{(t)}$ и с помощью них затем строится прогноз на t шагов для стартового вектора. Коэффициенты аппроксимации для всех t оцениваются независимо.

При прямом прогнозе отсутствуют недостатки двух первых способов — гипотеза о неизменности соседей здесь не вызывает сомнения, а накопления ошибки не происходит. Однако этот вариант более требователен к длине ряда, так как для прогноза на t шагов необходимы соседи, после которых имеется еще как минимум t известных векторов. Как будет показано в разделе 33.4, статистически этот способ должен обеспечивать наиболее надежные результаты.

В качестве примера рассмотрим уравнение Маккея–Гласса (MG), которое используется как модель процесса регенерации кровяных телец [412]:

$$\frac{dx}{dt} = \frac{ax(t - \Delta t)}{1 + x^c(t - \Delta t)} - bx(t),$$

где $a = 0, 2$; $b = 0, 1$; $c = 10$, $\Delta t = 17$. Как известно, при данных значениях параметров и времени задержки это уравнение проявляет хаотические свойства.

Прогноз для уравнения MG на $S = 164$ шага тремя рассмотренными способами иллюстрирует рис. 2.203. Видно, что при итеративном способе прогнозируемые значения, начиная с четвертого шага экспоненциально быстро отклоняются от истинных (для $S = 164$ это отклонение уже порядка 10^{15}).

Прогноз с пересчетом давал приемлемые результаты дольше, но приблизительно с сорокового шага он начал существенно уступать прямому прогнозу.

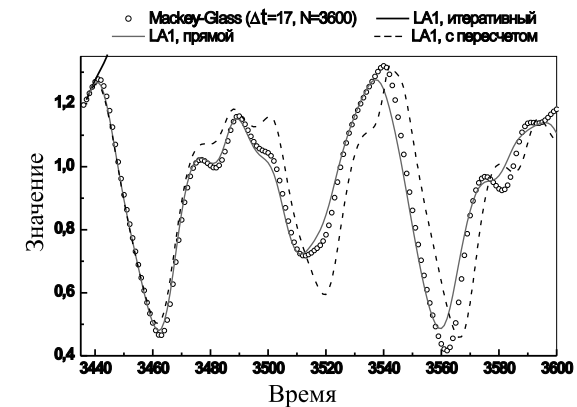


Рис. 2.203. Итеративный, прямой и прямой с пересчетом прогнозы методом LA1

На основе этого примера можно предположить, что итеративный вариант будет лучшим при очень малых величинах S , когда справедливо предположение о том, что предсказанный вектор остается соседом стартового. В остальных случаях, вероятно, прямой прогноз даст более достоверный результат.

33.4 Общее решение задачи прогноза

Рассмотренные выше примеры использования LA для экстраполяции временных рядов позволяют сделать ряд выводов об эффективности его различных вариантов. Однако для достоверности этих выводов, а также с целью получения инструмента для выбора оптимального варианта без

проведения численных расчетов необходимо найти аналитическое решение задачи прогноза методом LA.

33.4.1 Построение системы уравнений LA

Вновь обратимся к задаче построения прогноза на один шаг вперед, а точнее — к выбору локального представления. В общем случае функцию $f(\mathbf{x}_t, \mathbf{a})$ из (2.340) можно представить как

$$f(\mathbf{x}_t, \mathbf{a}) = a_0 + a_1 x_t + a_2 x_{t-1} + \dots + a_M x_{t-M+1} + \\ + a_{11} x_t^2 + a_{12} x_t x_{t-1} + \dots + a_{111} x_t^3 + a_{112} x_t^2 x_{t-1} + \dots \quad (2.349)$$

Если положить все коэффициенты кроме a_0 равными нулю, то получим LA0. Оставив коэффициенты только с одним индексом, приходим к LA1, с одним и двумя индексами — к LA2, и т.д. Следовательно, соотношение (2.340) в обобщенном виде можно записать следующим образом:

$$x_{t+1} = a_0 + \mathcal{X}_t^T \mathcal{A}, \quad (2.350)$$

где

$$\mathcal{X}_t^T = (x_t x_{t-1} \dots x_{t-M+1} x_t^2 x_t x_{t-1} \dots x_t^3 x_t^2 x_{t-1} \dots)$$

и

$$\mathcal{A}^T = (a_1 a_2 \dots a_M a_{11} a_{12} \dots a_{111} a_{112} \dots)$$

— обобщенный вектор в пространстве задержек и обобщенный вектор коэффициентов, соответственно. Для упрощения записи будем использовать единую нумерацию координат этих векторов, а сами векторы обозначать маленькой буквой. Какой именно вектор имеется в виду — обычный или обобщенный — зависит от рассматриваемого порядка LA. Если анализируется любой вариант LA, кроме LA1, значит используются обобщенные векторы коэффициентов и задержек.

Таким образом, далее используется единая запись для уравнения LA любого порядка:

$$x_{t+1} = a_0 + \mathbf{x}_t^T \mathbf{a}, \quad (2.351)$$

где

$$\mathbf{a}^T = (a_1 a_2 \dots), \quad \mathbf{x}_t^T = (x_t x_{t-1} \dots).$$

При этом постоянная \mathbf{a}_0 , как и ранее, выделяется не случайно. Вопрос о ее учете оказывается не простым. С одной стороны, наличие постоянной аналогично добавлению строки из единиц в матрицу задержек. С другой

стороны, без ее введения велика вероятность систематических ошибок при численных расчетах. Рассмотрим этот вопрос подробнее.

Запишем уравнение (2.351) для LA1 в терминах матрицы задержек:

$$x_{t+1} = \left(1 \quad \underbrace{x_t \cdots x_{t-p+1}} \right) \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix}.$$

Выделенная фигурной скобкой часть — это транспонированный вектор в пространстве задержек, т.е. один из столбцов матрицы \mathbf{X} . Таким образом, при прогнозировании матрица задержек неявно расширяется за счет добавления строки из единиц. А это несколько противоречит принципу ее построения.

Казалось бы, такого необоснованного расширения в большинстве случаев можно избежать уже в рамках линейного приближения, увеличив размерность реконструкции на единицу:

$$x_{t+2} = a_0^{\text{II}} + \sum_{i=1}^p a_i^{\text{II}} x_{t-i+2} \Rightarrow x_{t+2} = \sum_{i=1}^{p+1} b_i x_{t-i+2}, \\ x_{t+1} = a_0^{\text{I}} + \sum_{i=1}^p a_i^{\text{I}} x_{t-i+1}$$

$$b_1 = a_1^{\text{II}} + \frac{a_0^{\text{II}}}{a_0^{\text{I}}}, \quad b_{p+1} = -a_p^{\text{I}} \frac{a_0^{\text{II}}}{a_0^{\text{I}}},$$

$$b_i = a_i^{\text{II}} - a_{i-1}^{\text{I}} \frac{a_0^{\text{II}}}{a_0^{\text{I}}}, \quad a_0^{\text{I}} \neq 0, \quad i = 2, \dots, p.$$

В последнем выражении постоянная уже не присутствует. Однако в реальной ситуации при построении регрессионных зависимостей для оценки коэффициентов ее отсутствие может привести к систематическим ошибкам (см. рис. 2.204). Поэтому постоянную приходится сохранять, хотя это может быть недостаточно обосновано с точки зрения метода задержек.

Чтобы оценить неизвестные коэффициенты (a_0, \mathbf{a}) в уравнении (2.351), воспользуемся гипотезой их постоянства для всех соседей стартового вектора, т.е. основной гипотезой LA. В результате получится следующая система

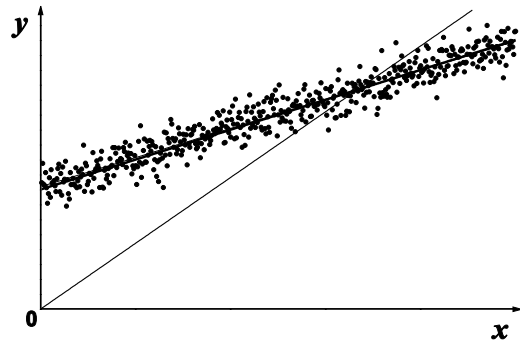


Рис. 2.204. Учет постоянной при построении линейной регрессии. Жирная линия — постоянная учитывалась, тонкая — нет. Точки — исходные данные

уравнений:

$$\begin{aligned} x_{s_1+S} &= a_0 + a_1x_{s_1} + a_2x_{s_1-1} + \dots + a_px_{s_1-p+1}, \\ x_{s_2+S} &= a_0 + a_1x_{s_2} + a_2x_{s_2-1} + \dots + a_px_{s_2-p+1}, \\ &\dots\dots\dots \\ x_{s_{\Xi}+S} &= a_0 + a_1x_{s_{\Xi}} + a_2x_{s_{\Xi}-1} + \dots + a_px_{s_{\Xi}-p+1}. \end{aligned} \tag{2.352}$$

Она построена для общего случая прогноза значения ряда, удаленного на S шагов от последнего известного вектора. Для прогноза на один шаг можно положить $S = 1$. Индексы s_1, \dots, s_{Ξ} указывают номера значений ряда.

Перепишем (2.352) в следующем виде:

$$\begin{pmatrix} x_{s_1+S} \\ x_{s_2+S} \\ \vdots \\ x_{s_{\Xi}+S} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} a_0 + \begin{pmatrix} x_{s_1} & x_{s_1-1} & \dots & x_{s_1-p+1} \\ x_{s_2} & x_{s_2-1} & \dots & x_{s_2-p+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{s_{\Xi}} & x_{s_{\Xi}-1} & \dots & x_{s_{\Xi}-p+1} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}. \tag{2.353}$$

Здесь большая матрица (матрица соседей) — это транспонированный фрагмент матрицы задержек, состоящий лишь из одних соседей стартового вектора (в случае LA1). Левая часть представляет собой результаты эволюции каждого из соседей за S шагов и $\{a_0, \dots, a_p\}$ — параметры аппроксимации, которые надо определить. Для нулевого порядка система (2.353) трансфор-

мируется в соотношение

$$\begin{pmatrix} x_{s_1+S} \\ x_{s_2+S} \\ \vdots \\ x_{s_{\Xi}+S} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} a_0.$$

Для упрощения записи обозначим x_{s_k+S} как y_k , где k — номер соседа, а матрицу соседей (возможно, обобщенную) — как X . Для вектора коэффициентов оставим обозначение (a_0, \mathbf{a}) . Тогда система уравнений (2.352) примет вид

$$\mathbf{Y} = \mathbf{I}a_0 + \mathbf{X}\mathbf{a}. \tag{2.354}$$

Требуется оценить параметры (a_0, \mathbf{a}) .

33.4.2 Аналитическое решение задачи прогноза методом LA

Решение системы (2.354) будем искать с помощью МНК. Для нее критерий МНК (ср. (2.346)) имеет вид:

$$\Phi = (\mathbf{Y} - \mathbf{I}a_0 - \mathbf{X}\mathbf{a})^T (\mathbf{Y} - \mathbf{I}a_0 - \mathbf{X}\mathbf{a}) \rightarrow \min_{\{a_0, \mathbf{a}\}}.$$

Условия первого порядка этой задачи можно записать как

$$\frac{\partial \Phi}{\partial a_0} : -2\mathbf{I}^T (\mathbf{Y} - \mathbf{I}a_0 - \mathbf{X}\mathbf{a}) = 0, \tag{2.355}$$

$$\frac{\partial \Phi}{\partial \mathbf{a}} : -2\mathbf{X}^T (\mathbf{Y} - \mathbf{I}a_0 - \mathbf{X}\mathbf{a}) = 0.$$

Выразим теперь из первого соотношения a_0 :

$$a_0 = \frac{\mathbf{I}^T (\mathbf{Y} - \mathbf{X}\mathbf{a})}{\mathbf{I}^T \mathbf{I}}. \tag{2.356}$$

Здесь $\mathbf{I}^T \mathbf{I}$ просто равно Ξ . Подставляя a_0 во второе соотношение (2.355), получим оценку вектора коэффициентов \mathbf{a} :

$$\hat{\mathbf{a}} = \left[\mathbf{X}^T \left(\mathbf{E} - \frac{\mathbf{I}\mathbf{I}^T}{\mathbf{I}^T \mathbf{I}} \right) \mathbf{X} \right]^{-1} \mathbf{X}^T \left(\mathbf{E} - \frac{\mathbf{I}\mathbf{I}^T}{\mathbf{I}^T \mathbf{I}} \right) \mathbf{Y}, \tag{2.357}$$

где $\mathbf{E}_{\Xi \times \Xi}$ — единичная матрица. Это выражение определяет все необходимые для построения прогноза коэффициенты, если, конечно, матрица в квадратных скобках обратима²². Оценку \hat{a}_0 можно получить из соотношения (2.356), подставив вместо коэффициентов \mathbf{a} их оценки, полученные из (2.357).

Полученные оценки коэффициентов позволяют получить аналитическое выражение для прогнозируемого значения ряда \hat{x}_{L+S} в момент $L+S$. Сделаем это поэтапно, одновременно анализируя получаемые результаты.

Сначала подставим в уравнение (2.351) стартовый вектор \mathbf{x}_L (возможно, обобщенный по принципу (2.350))

$$\hat{x}_{L+S} = \hat{a}_0 + \mathbf{x}_L^T \hat{\mathbf{a}}$$

и только что полученную оценку \hat{a}_0

$$\hat{x}_{L+S} = \frac{\mathbf{I}^T}{\mathbf{I}^T \mathbf{I}} (\mathbf{Y} - \mathbf{X} \hat{\mathbf{a}}) + \mathbf{x}_L^T \hat{\mathbf{a}}.$$

Преобразуем полученное выражение следующим образом:

$$\hat{x}_{L+S} = \bar{Y} + (\mathbf{x}_L^T - \bar{\mathbf{X}}) \hat{\mathbf{a}}. \quad (2.358)$$

Здесь \bar{Y} — среднее значение вектора \mathbf{Y} , т.е. прогноз по LA0, $\bar{\mathbf{X}}$ — вектор средних значений координат в матрице соседей \mathbf{X} , т.е. «усредненный» сосед.

Соотношение (2.358) имеет особый смысл. Оно показывает, что прогноз, полученный методом локальной аппроксимации любого порядка, есть линейная комбинация прогноза нулевого порядка и отклонений стартового вектора от «усредненного» соседа. Отсюда вытекают два важных следствия.

1. Поправка, обусловленная отклонением стартового вектора от «усредненного» соседа, корректирует неравномерность распределения соседей вокруг стартового вектора. Это следует из того, что в случаях, когда стартовый вектор одновременно является и центральным (с учетом выбранного порядка аппроксимации), т.е. совпадает с «усредненным» соседом, поправка к прогнозу по LA0 в соответствии с (2.358) оказывается равной нулю. Частным случаем такого распределения является полное совпадение соседей и стартового вектора. Такая ситуация может возникнуть при анализе периодического ряда.

²²Способы расчета коэффициентов, если матрица необратима, требуют отдельного рассмотрения и специальных методов, что существенно усложняет процедуру [243].

2. Поправка прогноза зависит именно от отклонения стартового вектора от центрального вектора для соседей, а не от абсолютных значений координат векторов. Следовательно, при численных расчетах методом LA правильнее рассматривать не сами векторы-соседи, а их отклонения от гипотетического центрального вектора.

Возвращаясь к вопросу о включении константы в уравнение (2.351), заметим, что метод LA0 при ее отсутствии не реализуется и, следовательно, рассмотренные свойства решения задачи прогноза LA без постоянной не проявляются.

Запишем в явном виде решение задачи прогноза по методу LA. На основе выражений (2.357), (2.358) имеем:

$$\hat{x}_{L+S} = \bar{Y} + (\mathbf{x}_L - \bar{\mathbf{X}}) \left[(\mathbf{X} - \mathbf{I}\bar{\mathbf{X}})^T (\mathbf{X} - \mathbf{I}\bar{\mathbf{X}}) \right]^{-1} (\mathbf{X} - \mathbf{I}\bar{\mathbf{X}})^T (\mathbf{Y} - \mathbf{I}\bar{Y}). \quad (2.359)$$

Здесь использовано свойство

$$\left(\mathbf{E} - \frac{\mathbf{I}\mathbf{I}^T}{\mathbf{I}^T \mathbf{I}} \right) = \left(\mathbf{E} - \frac{\mathbf{I}\mathbf{I}^T}{\mathbf{I}^T \mathbf{I}} \right)^T \left(\mathbf{E} - \frac{\mathbf{I}\mathbf{I}^T}{\mathbf{I}^T \mathbf{I}} \right).$$

Из соотношения (2.359) видно, что в выражении для прогноза все исходные данные, кроме \bar{Y} («усредненного» прогноза, совпадающего с прогнозом LA0) входят в виде отклонений: $(\mathbf{x}_L - \bar{\mathbf{X}})$ — отклонение стартового вектора от «усредненного» соседа; $(\mathbf{X} - \mathbf{I}\bar{\mathbf{X}})^T$ — матрица отклонений соседей от «усредненного» соседа; $(\mathbf{Y} - \mathbf{I}\bar{Y})$ — вектор отклонений индивидуальных «результатов эволюции» соседей за S шагов от среднего варианта эволюции. Таким образом, произошел переход к центрированным величинам.

Уравнение (2.359) может быть непосредственно использовано для получения прогноза методом LA при численных расчетах. Единственным условием его применения является обратимость матрицы, стоящей в квадратных скобках. К сожалению, при достаточно длинном ряде и соседях, близких к стартовому вектору, эта матрица плохо обусловлена и для решения уравнения требуется применение специальных методов.

33.5 Выбор оптимального варианта LA при прогнозировании на несколько шагов вперед

Полученное в предыдущем разделе в явном виде решение задачи прогноза методом LA может быть использовано для анализа различных свойств прогноза, таких как точность, устойчивость и др. без проведения численных

расчетов. Сравнение же свойств прогнозов, получаемых в разных вариантах LA, позволяет выбрать оптимальный вариант метода.

33.5.1 Способы решения задачи прогноза на несколько шагов вперед

Основные варианты построения прогноза на несколько шагов вперед методом LA качественно были описаны выше. Остановимся на их реализации в рамках общей модели LA.

Первый шаг во всех вариантах одинаков — построение методом LA выбранного порядка прогноза для момента времени $L + 1$. При этом оцениваются параметры аппроксимации (a_0, \mathbf{a}) . На основе полученного прогноза достраивается матрица \mathbf{X} посредством добавления к ней нового вектора, содержащего в качестве первой координаты найденное значение (2.348). Дальше способы прогноза расходятся.

● *Итеративный способ прогноза.* В этом варианте предполагается, что получившийся вектор — сосед стартового. Как было отмечено в разделе 33.3.3, такое предположение весьма сомнительно, особенно если рассматривается случай больших временных интервалов между значениями ряда. Применительно к модели LA оно означает, что коэффициенты \mathbf{a} для вектора (2.348) при прогнозе $(L + 2)$ -го значения ряда совпадают с коэффициентами, рассчитанными для стартового вектора, и, следовательно, прогнозируемое выражение для $(L + S)$ -го значения ряда (ср. (2.358)) имеет вид:

$$\hat{x}_{L+S} = \bar{Y}^{(L+1)} + \left(\left(\hat{\mathbf{x}}_{L+S-1}^{(L+S)} \right)^T - \bar{\mathbf{X}}^{(L+1)} \right) \hat{\mathbf{a}}^{(L+1)}. \quad (2.360)$$

Индексы в скобках показывают, на каком шаге прогноза рассчитывается соответствующая величина: $(L + 1)$ означает, что она вычисляется только на первом шаге и далее не пересчитывается; значение с индексом $(L + S)$ вычисляется заново на каждом шаге.

В выражении (2.360) все величины за исключением вектора $\hat{\mathbf{x}}_{L+S-1}^{(L+S)}$, рассчитанного на предыдущем шаге прогноза, остаются неизменными для всех прогнозируемых моментов времени. Этот факт отражен в их индексах $(L + 1)$. Вектор координат «усредненного» соседа $\bar{\mathbf{X}}^{(L+1)}$ рассчитан по исходному набору соседей и также не зависит от номера шага прогноза.

Основное преимущество итеративного способа — быстрота и легкость расчета, поскольку не надо искать заново соседей и пересчитывать коэффициенты. Главный недостаток — необходимость принятия не бесспорного предположения, что любой вектор $\hat{\mathbf{x}}_{L+S-1}^{(L+S)}$ является соседом стартового вектора.

В работе [300] показано, что теоретически этот вариант наиболее обоснован. Мы не будем рассматривать это доказательство, отметим только, что для базового варианта LA0 прогноз этим способом даст универсальный результат:

$$\bar{x}_{L+S} = \bar{Y}^{(L+1)} = \text{const}(S). \quad (2.361)$$

Таким образом, прогноз по LA0 итеративным способом для любого ряда и на любой период — постоянная, не зависящая от времени. Этот результат при всей его очевидной ошибочности является следствием предположения, на котором основан итеративный способ.

● *Прямой прогноз для каждого следующего момента времени.* Прогноз этим способом проводится независимо для всех искомым значений ряда, т. е. прогноз значения ряда в момент $L + 1$ никак не влияет на построение величины в момент $L + 2$. Прямой способ прогноза реализуется подстановкой в левую часть основного уравнения LA (2.353) результатов эволюции соседей за нужное число шагов. Следовательно, для прогноза ряда к моменту $L + S$ в левую часть (2.353) нужно подставить те значения векторов-соседей, в которые они перешли за S шагов. Общий вид прогноза $(L + S)$ -го значения ряда прямым способом в тех же обозначениях, которые использовались для описания итеративного способа, дается выражением:

$$\hat{x}_{L+S} = \bar{Y}^{(L+S)} + \left(\left(\hat{\mathbf{x}}_L^{(L+1)} \right)^T - \bar{\mathbf{X}}^{(L+1)} \right) \hat{\mathbf{a}}^{(L+S)}.$$

При использовании этого способа на каждом шаге пересчитываются как прогноз LA0 (значение $\bar{Y}^{(L+S)}$), так и коэффициенты линейной комбинации (вектор $\hat{\mathbf{a}}^{(L+S)}$). При этом стартовый вектор и соседи остаются неизменными.

В отличие от итеративных вариантов, прямой способ прогнозирования не может дать неограниченно возрастающие значения, так как здесь не происходит накопления ошибки прогноза, а ошибка за один шаг не влияет на прогноз следующего. В случае итеративного прогноза ошибка на первом шаге, в соответствии с (2.360), повлияет на прогноз для второго шага и так далее. Следовательно, единственная значительная ошибка на одном из шагов при итеративном способе прогнозирования сделает бессмысленным весь дальнейший прогноз.

● *Итеративный способ с пересчетом параметров.* Этот вариант, оставаясь по существу итеративным, не использует предположения, что каждый новый предсказанный вектор является соседом стартового. При прогнозировании этим способом для каждого нового вектора прогноза определяются

свои соседи. Таким образом, алгоритм предсказания данным вариантом LA следующий: определить соседей стартового вектора (в отличие от простого итеративного варианта здесь эта процедура должна осуществляться на каждом шаге), построить прогноз на один шаг вперед, добавить полученный вектор как последний в матрицу \mathbf{X} и выбрать его в качестве стартового. Повторением этой процедуры S раз можно получить прогноз для значений ряда в моменты $L+1, L+2, \dots, L+S$. Если в частном случае окажется, что на каждом шаге прогноза соседи стартового вектора были одни и те же, то результат прогноза совпадет с результатом простого итеративного способа.

Общий вид решения задачи прогноза для $(L+S)$ -го значения ряда при использовании итеративного способа с пересчетом параметров можно записать как

$$\hat{x}_{L+S} = \bar{Y}^{(L+S)} + \left(\left(\hat{x}_{L+S-1}^{(L+S)} \right)^T - \hat{\mathbf{X}}^{(L+S)} \right) \hat{\mathbf{a}}^{(L+S)}.$$

Для расчета вектора $\hat{\mathbf{X}}^{(L+S)}$ («усредненный» сосед) могут использоваться результаты прогнозов предшествующих значений ряда, поэтому к его обозначению добавлена «крышечка». Заметим, что эта особенность не всегда ведет к повышению устойчивости получаемого результата.

• *Качественное сравнение рассмотренных способов.* Вернемся к примеру из работы [300] о прогнозировании погоды и рассмотрим, как работает каждый из рассмотренных способов. Ограничимся вариантами нулевого порядка (LA0) и выбором только одного ближайшего соседа. Тогда алгоритм действий при итеративном способе можно представить следующим образом: находим в истории день с наиболее похожей на сегодняшний день погодой, смотрим, какая погода была в следующий за ним день и утверждаем, что такая же погода будет завтра, послезавтра и т. п.

При использовании прямого способа прогнозом на завтра будет та погода, что была в день, следующий за найденным в истории; на послезавтра — та, что была через день после найденного и так далее. Здесь может возникнуть некоторая сложность, связанная с тем, что наиболее похожим на сегодняшний день может оказаться вчерашний. Тогда для прогноза на послезавтра придется искать другого соседа, так как мы не знаем достоверно, какая «была» (точнее, «будет») погода через день после вчерашнего дня, т. е. завтра. Но эту проблему можно легко снять, ограничив заранее выбор соседей только теми из них, которые удалены от сегодняшнего дня на необходимую длительность прогноза.

Наконец, при использовании итеративного способа с пересчетом получится следующая последовательность действий: предсказываем погоду на

завтра; считая, что угадали, для прогноза на послезавтра ищем день с погодой, наиболее похожей на только что сделанный прогноз и утверждаем, что послезавтра погода будет такой же как в день, последовавший за ним.

Основным результатом проведенного сравнения является предположение, что наиболее устойчивым методом, дающим самый точный прогноз среди рассмотренных способов, является прямой способ прогноза. Он более полно, чем простой итеративный способ, использует имеющуюся информацию и более устойчив к ошибкам в данных и в вычислениях, чем любой из итеративных вариантов.

33.5.2 Величина ошибки прогноза: предпосылки выбора порядка аппроксимации и способа прогнозирования

Имея общий вид выражения для прогнозируемого значения ряда в момент $L+S$, можно оценить насколько в среднем оно отличается от истинного значения ряда в этот момент. Для этого предположим, во-первых, что истинное значение ряда в момент $L+S$ существует и, во-вторых, что оно задается некоторой гладкой функцией вида

$$x_{L+S} = f_S(x_L). \quad (2.362)$$

Введем ошибку прогноза на шаге S как

$$E_S = |x_{L+S} - \hat{x}_{L+S}|. \quad (2.363)$$

Цель настоящего раздела — получить явное выражение для ошибки прогноза в случае итеративного и прямого способов прогнозирования. В случае итеративного способа с пересчетом параметров оценить ошибку значительно сложнее, поскольку возможно изменение набора соседей стартового вектора.

При нулевом порядке аппроксимации, следуя (2.361), прогнозируемое значения ряда при итеративном способе будет представляться как

$$\hat{x}_{L+S}^{(iter)} = \frac{1}{\Xi} \sum_{s=1}^{\Xi} f_1(x_L + \hat{\delta}_s),$$

где суммирование ведется по всем соседям стартового вектора. Вектор $\hat{\delta}_s$ задает отклонение соседа с номером s от стартового вектора. Выражение для того же значения ряда, полученное прямым способом, имеет вид:

$$\hat{x}_{L+S}^{(direct)} = \frac{1}{\Xi} \sum_{s=1}^{\Xi} f_T(x_L + \hat{\delta}_s).$$

Подставляя эти соотношения в (2.363), получим значения ошибок прогноза итеративным и прямыми способами, соответственно:

$$E_S^{(iter)} = \left| x_{L+S} - \left\langle f_1 \left(\mathbf{x}_L + \hat{\delta} \right) \right\rangle \right|, \quad (2.364)$$

$$E_S^{(direct)} = \left| x_{L+S} - \left\langle f_S \left(\mathbf{x}_L + \hat{\delta} \right) \right\rangle \right|. \quad (2.365)$$

Здесь усреднение, обозначенное угловыми скобками, ведется по всем соседям.

Используя разложение в ряд функции $f_T \left(\mathbf{x}_L + \hat{\delta}_s \right)$, соотношения (2.364), (2.365) в первом порядке по $\hat{\delta}$ можно переписать как

$$E_S^{(iter)} = |x_{L+S} - x_{L+1}| + O \left(\left| \left\langle \hat{\delta} \right\rangle \right| \right), \\ E_S^{(direct)} = O \left(\left| \left\langle \hat{\delta} \right\rangle \right| \right).$$

В предельном случае, когда соседи распределены равномерно вокруг стартового вектора и, следовательно, $\left| \left\langle \hat{\delta} \right\rangle \right| \rightarrow 0$, прогноз по прямому способу стремится к истинному значению, тогда как прогноз по итеративному способу имеет систематическую неубывающую ошибку. Таким образом, более предпочтительным оказывается прямой способ прогноза.

При аппроксимации ненулевого порядка в соответствии с (2.360) решение задачи прогноза итеративным способом можно представить в виде:

$$\hat{x}_{L+S}^{(iter)} = \left(\bar{Y}^{(L+1)} - \bar{X}^{(L+1)} \hat{\mathbf{a}}^{(L+1)} \right) \hat{b}_0 + \mathbf{x}_L^T \hat{\mathbf{b}}.$$

Здесь $\left(\hat{b}_0 \hat{\mathbf{b}}^T \right)$ — некоторые коэффициенты. Их значения зависят только от координат найденного на первом шаге вектора параметров $\left(\hat{a}_0 \hat{\mathbf{a}}^T \right)$ и длительности прогноза. Таким образом, итеративный способ в лучшем случае дает линейную по координатам стартового вектора аппроксимацию функции (2.362). Последнее означает, что если выбран порядок аппроксимации p , то имеет место нелинейная аппроксимация p -го порядка с постоянными коэффициентами.

Независимо от выбранного порядка прямой вариант LA в общем случае дает аппроксимацию первого порядка. Тем не менее и в этом варианте увеличение порядка LA может повышать точность аппроксимации.

Рассмотрим прямой вариант LA1 в одномерном случае. Ошибку прогноза можно представить как

$$E_S^{(direct)} = \left| \left(f'_S(x) - \hat{f}'_S \right) \frac{\sum_{s=1}^{\Xi} \delta_s}{\Xi} - f''_S(x) \frac{\sum_{s=1}^{\Xi} \delta_s^2}{2\Xi} + O(\delta^3) \right|. \quad (2.366)$$

Следовательно, переход от нулевого к первому порядку позволяет снизить вклад в ошибку линейного члена по δ за счет оценки величины производной. В случае равномерного распределения соседей вокруг стартового вектора, так же как и при нулевом порядке, получается аппроксимация с точностью до второго порядка по отклонению. Иногда к хорошим результатам приводит увеличение порядка метода, т. е. переход к LA2.

Главное достоинство прямого варианта LA — это уменьшение ошибки с приближением соседей к стартовому вектору: при ограниченности $f''_S(x)$ и $\delta \rightarrow 0$, $E_S^{(direct)} \rightarrow 0$. Такого не происходит в итеративном варианте: ошибка $E_S^{(iter)}$ определяется поведением функции f_S и независимо от величины отклонений может быть сравнима с самим значением прогноза, начиная уже со второго шага.

Общий вывод из проведенного анализа следующий: итеративный вариант нулевого порядка полностью соответствует определению и дает аппроксимацию нулевого порядка, т. е. прогноз на любое количество шагов аппроксимируется значением прогноза на один шаг вперед. В то же время при прямом варианте LA0 сразу получается аппроксимация первого порядка, а при равномерном распределении соседей вокруг стартового вектора — второго порядка.

Таким образом, при использовании прямого варианта LA0 прогноз тем точнее, чем ближе и равномернее относительно стартового вектора распределены соседи, тогда как итеративный вариант подобен стоящим часам, которые иногда показывают точное время. Использование в итеративном варианте LA высших порядков позволяет качественно улучшить прогноз, но с увеличением порядка для оценки параметров требуется больше соседей. Это приводит к снижению качества прогноза. Поэтому при использовании итеративного варианта важен оптимальный выбор порядка аппроксимации. В прямом варианте LA выбрать порядок легче. Если ряд короткий и соседи распределены неравномерно вокруг стартового вектора, то оптимален LA1, так как в этом случае минимизируется вклад в величину ошибки прогноза линейного по отклонениям слагаемого в (2.366). Дальнейшее увеличение порядка аппроксимации к качественному улучшению не приведет. В случае длинного ряда с равномерным распределением соседей линейное слагаемое

в (2.366) в силу своей малости практически не влияет на величину ошибки, и здесь может оказаться полезным применение LA2.

Оценим для прямого и итеративного вариантов ошибку прогноза, связанную с конечной точностью вычислений. Введем несколько дополнительных предположений.

1. Исходный ряд бесконечно длинный, т. е. всегда можно найти необходимое количество истинных соседей для стартового вектора.

2. Аппроксимация функции $f_S(\mathbf{x})$, $x_{L+S} = f_S(x_L)$ любым из рассматриваемых вариантов прогноза является абсолютно точной, т. е. дает те же значения, которые были бы получены с помощью истинной функции.

3. Единственная причина ошибки при прогнозировании — конечность представления числовых значений ряда (ограниченность машинной точности представления чисел). Поэтому в процессе расчетов вместо истинного значения ряда подставляется ближайшее округленное значение, отличающееся от истинного на величину, не превосходящую δ (параметр, характеризующий машинную точность представления чисел). Тогда ошибка, возникающая из-за этого ограничения, распределена равномерно на интервале $(-\delta, \delta)$. Дисперсия такой величины есть $\delta^2/3$. Для обозначения ошибки введем величину $\varepsilon(0, \delta^2/3)$, где первая цифра в скобках — математическое ожидание ошибки, вторая — дисперсия.

4. Будем рассматривать одномерный случай и предположим, что $|df_S(\mathbf{x})/d\mathbf{x}| \approx k^S$, где k — величина, которая определяется параметрами системы и не зависит от S и \mathbf{x} . Несмотря на то, что \mathbf{x} теперь скаляр, оставим для нее прежнее полужирное написание. Используя показатель Ляпунова λ , для достаточно больших t найдем: $|df_t(\mathbf{x})/d\mathbf{x}| \approx e^{\lambda t}$. Учитывая, что $t = \tau S$, где τ — промежуток времени между значениями ряда, и считая, что τ велико, получим:

$$\left| \frac{df_t(\mathbf{x})}{d\mathbf{x}} \right| \approx e^{\lambda \tau S} = k^S.$$

Оценим в рамках указанных предположений ошибку прогноза на шаге S . При условии бесконечности ряда для любого стартового значения x_L всегда найдется ближайший сосед \mathbf{x} , совпадающий с ним в пределах ошибки округления: $\mathbf{x} = x_L + \varepsilon(0, \delta^2/3)$. В таком случае значение прогноза на шаге S будет с точностью до ошибки округления совпадать со значением, в которое за S шагов перейдет ближайший сосед стартового вектора:

$$\hat{x}_{L+S} = f_S(\mathbf{x}) + \varepsilon\left(0, \frac{\delta^2}{3}\right) = f_S\left(x_L + \varepsilon\left(0, \frac{\delta^2}{3}\right)\right) + \varepsilon\left(0, \frac{\delta^2}{3}\right). \quad (2.367)$$

Разложим функцию эволюции в (2.367) в ряд по степеням отклонения соседа от стартового вектора (до первого порядка включительно). Тогда

$$\hat{x}_{L+S} \approx x_{L+S} + \varepsilon\left(0, \frac{\delta^2}{3}\right) k^S + \varepsilon\left(0, \frac{\delta^2}{3}\right).$$

Считая ошибки некоррелированными, отсюда находим:

$$\hat{x}_{L+S} = x_{L+S} + \varepsilon\left(0, \frac{\delta^2}{3}(k^{2S} + 1)\right). \quad (2.368)$$

Следовательно, в прямом способе математическое ожидание прогноза должно совпадать с истинным значением, а дисперсия прогноза равна

$$D_S^{direct} = \frac{\delta^2}{3}(k^{2S} + 1). \quad (2.369)$$

В случае итеративного прогноза на первом шаге в соответствии с (2.368) получим:

$$\hat{x}_{L+1} = x_{L+1} + \varepsilon\left(0, \frac{\delta^2}{3}(k^2 + 1)\right).$$

Аналогично, значение прогноза на втором шаге в соответствии с (2.367) примет вид

$$\hat{x}_{L+2}^{iter} = x_{L+2} + \varepsilon\left(0, \frac{\delta^2}{3}[(k^2 + 1)k^2 + 1]\right).$$

Значит, происходит накопление ошибки округления на каждом шаге. Продолжая построение прогноза для следующих шагов, можно получить общее выражение для дисперсии ошибки прогноза в итеративном варианте: $D_S^{iter} = \delta^2 d_S/3$, $d_S = d_{S-1}k^2 + 1$, $d_0 = 1$ или

$$D_S^{iter} = \frac{\delta^2}{3} \sum_{j=0}^S k^{2j} = \frac{\delta^2}{3} \left(k^{2S} + \sum_{j=1}^{S-1} k^{2j} + 1 \right). \quad (2.370)$$

Математическое ожидание в итеративном варианте также должно (при точной аппроксимации функции преобразования) совпадать с истинным значением ряда.

Сравнивая выражения (2.370) и (2.369), можно заметить, что $D_S^{iter} \geq D_S^{direct}$, причем равенство возможно только либо при $k = 0$, либо при $S = 1$. В остальных случаях $D_S^{iter} > D_S^{direct}$, т. е. ошибка прогноза в итеративном варианте должна быть не меньше, чем при прямом варианте прогноза. Поэтому прямой вариант прогнозирования снова оказывается предпочтительнее итеративного. В случае нехаотического поведения системы $k = 1$ и $D_S^{direct} = 2\delta^2/3$, $D_S^{iter} = (S + 1)\delta^2/3$. Рост D_S^{iter} с увеличением S — это следствие увеличения ошибки за счет последовательных округлений.

То же самое происходит, если вместо того, чтобы измерять длину отрезка целиком, разбить его на части и измерять каждую часть отдельно, округляя получаемые результаты. В предельном случае, если длина любой части отрезка окажется меньше половины деления линейки (при стандартных правилах округления в сторону ближайшего целого), то общая длина отрезка, измеренная таким образом, получится нулевой, вне зависимости от его реальной длины.

Основным результатом проведенного сравнения вариантов LA является тот факт, что наиболее универсальным методом при прогнозировании достаточно длинных стационарных нерегулярных временных рядов будет прямой вариант LA0. С одной стороны, он может обеспечивать прогнозирование с величиной ошибки, убывающей пропорционально среднему квадрату отклонений соседей от стартового вектора. С другой стороны, расчет по этому методу является самым простым из всех рассмотренных. Кроме того, при его использовании не вносятся дополнительные ошибки, связанные с методом поиска коэффициентов. Это полностью соответствует интуитивно понятному методу прогноза по историческим аналогиям. В случае ограниченного набора данных или их неравномерного распределения вокруг стартового вектора можно воспользоваться прямым вариантом LA1.

33.5.3 Результаты численного сравнения

Осуществить проверку соотношений (2.369) и (2.370) для стохастических рядов численно не представляется возможным, так как первые два предположения, на основе которых они были выведены, на практике нереализуемы. Однако можно сравнить точности прогнозов, получаемых различными вариантами LA. В качестве объектов сравнения точности прогнозов выберем ряды, полученные из тех же уравнений MG и Лоренца, а сравнивать получаемые прогнозы будем, следуя работе [300], по величине

нормированной среднеквадратичной ошибки или ошибки прогноза:

$$E = \sqrt{\frac{\langle (f_{t+S} - \hat{f}_{t+S})^2 \rangle_t}{\langle (f_t - \langle f_t \rangle_t)^2 \rangle_t}}.$$

Здесь усреднение проводится по моментам начала прогноза. Естественно, чем меньше ошибка прогноза, тем он лучше. Если величина E превышает единицу, то прогноз неудачный и вместо него лучше взять среднее значение ряда. Для большей надежности получаемой оценки используется не арифметическое среднее по начальным моментам прогноза, а медиана.

Рассмотрим сначала ряд значений, полученный из уравнения MG при тех же условиях, что и в работе [300]: $N = 10\,000$, $\tau = 6$, $p = 4$, количество соседей — вдвое больше минимального, т. е. $\Xi = 2$ для LA0, $\Xi = 10$ для LA1 и $\Xi = 30$ для LA2. Результаты сравнения представлены на рис. 2.205. Из рисунка следует, что в данном примере лучшим оказался итеративный вариант с пересчетом; причем если для нулевого порядка его преимущество незначительно, то при применении старших порядков аппроксимации этот вариант имеет абсолютное преимущество перед прямым вариантом LA. Прогноз по итеративному варианту LA оказался совершенно неудачным и просто несравнимым с другими вариантами (см. врезку на рис. 2.205). Этот результат принципиально отличается от полученного при тех же условиях в работе [300]. Там итеративный вариант оказался лучше прямого.

Во всех вариантах лучшие результаты получились при использовании первого порядка аппроксимации. Второй порядок дал худшие результаты. Прогноз при нулевом порядке аппроксимации оказался существенно хуже, правда, в прямом варианте только при небольшой длине прогноза.

Таким образом, здесь лучшим оказался итеративный вариант с пересчетом первого порядка. В прямом варианте лучшие результаты также были получены при аппроксимации первого порядка, но на достаточном удалении от стартовой точки результат практически не зависит от порядка аппроксимации. Однако если немного изменить параметры, в первую очередь значение τ , т. е. рассмотреть ряд MG при $\tau = 1$ и выбрать количество соседей втрое больше минимального ($\Xi = 3$ для LA0, $\Xi = 15$ для LA1 и $\Xi = 45$ для LA2), результаты окажутся другими. Они представлены на рис. 2.206. В этом примере результат итеративного варианта LA снова оказался худшим, но лучшим стал уже прямой вариант LA.

Обрыв линии, соответствующей прогнозу по итеративному варианту с пересчетом второго порядка, связан с особенностями численной реали-

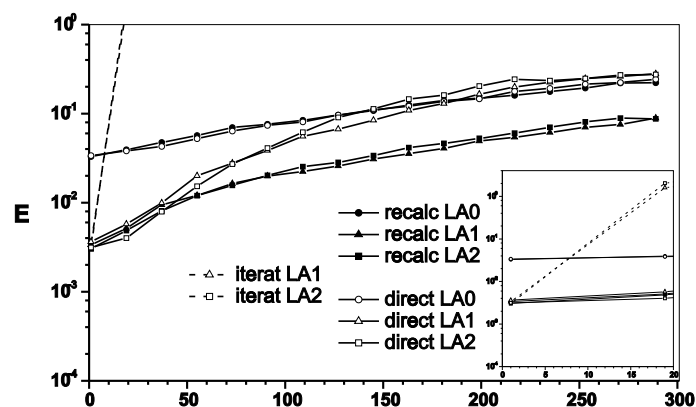


Рис. 2.205. Зависимость ошибки от длины прогноза для ряда MG ($\tau = 6$). Итеративный вариант обозначен как *iterat*, прямой — *direct*, итеративный с пересчетом — *recalc*. Медианное усреднение по 500-м точкам старта

зации алгоритмов расчета. В данном случае невозможность дальнейшего построения прогноза была вызвана превышением максимально допустимого числового значения, что обусловлено конечной точностью проводимых расчетов. Это, а также накопление ошибки, характерное для итеративного варианта LA, может приводить к некорректным результатам, как это было в примере на рис. 2.203.

Из рассмотренных примеров следует, что для рядов, ограниченных по количеству исходных данных, в каждом случае лучшим оказался первый порядок аппроксимации. Это связано с небольшой длиной применяемых рядов, поэтому результаты при аппроксимации нулевого порядка оказались хуже, чем в случае линейной аппроксимации. Проявившееся в большинстве случаев превосходство линейной аппроксимации над аппроксимацией второго порядка связано с ограниченной точностью расчетов. С ростом порядка аппроксимации отрицательное влияние точности на качество прогноза возрастает. Все сказанное относится в первую очередь к прямому и итеративному с пересчетом вариантам, так как итеративный способ ведет к неприемлемой точности прогноза.

Итеративный вариант с пересчетом, который дает в среднем результаты, сопоставимые по точности с прямым вариантом, в конкретных задачах прогноза может применяться только как LA0. Это связано с тем, что при использовании старших порядков высока вероятность появления систематических ошибок за счет характерного для этого варианта быстрого нарастания

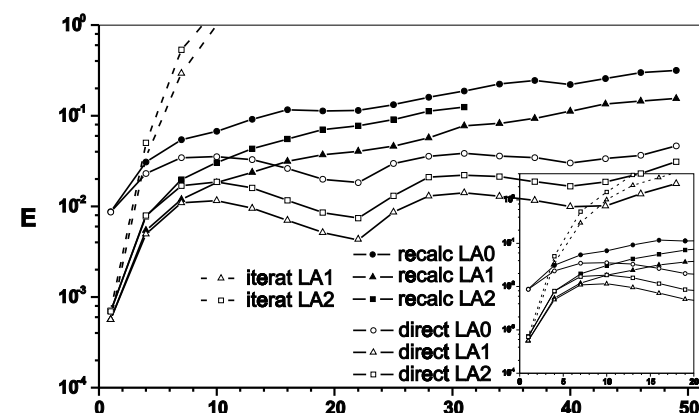


Рис. 2.206. То же, что на рис. 2.205, но при $\tau = 1$

тания погрешности прогноза, обусловленного погрешностями численных расчетов.

Таким образом, как следует из всего изложенного материала, в каждом конкретном случае необходимо систематизировать существующие разновидности LA и выбрать из них оптимальную. Справиться с этим позволяет предложенная общая модель методов LA. Она также дает возможность выяснить некоторые существенные особенности получаемого решения задачи прогноза. Например, оптимальным по соотношению точности и устойчивости прогноза является прямой вариант LA нулевого порядка. В случае коротких незашумленных рядов более эффективным может оказаться прямой вариант LA1. Но при этом важно учитывать связанные с ним вычислительные сложности.

В первую очередь представленные методы применимы к системам, временные масштабы существования которых значительно превосходят периоды наблюдений и прогнозов. Кроме того, процесс измерений, как и его результаты, не должны влиять на динамику системы. Поэтому в качестве примеров мы не рассматривали увлекательные задачи прогноза курсов акций или валют.

В заключение отметим, что для выявления всех достоинств и недостатков хорошо было бы провести сравнение алгоритма LA и результатов прогноза посредством LA с другими известными методами, которые близки по идеологии, но не являются локальными [68, 519]. Кроме того, желательна проверка на *реальных* рядах, рассмотренных, например, в работах [124, 125, 318, 402].

ГЛАВА 3

ОБРАБОТКА ИНФОРМАЦИИ РАСПРЕДЕЛЕННЫМИ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

Знание основных закономерностей образования структур в активных средах, а также в сетях, состоящих из большого числа активных элементов, позволяет перейти к целенаправленному созданию распределенных динамических систем, которые формируют те или иные пространственные структуры. Одним из основных приложений при этом являются задачи аналоговой обработки информации. Использование в качестве элементарной единицы обработки информации не отдельных сигналов, а протяженных пространственных структур дает возможность резко повысить эффективность компьютера в проблемах искусственного интеллекта. Имеется ряд свидетельств, что именно аналоговые механизмы лежат в основе работы человеческого мозга.

34. Мозг и компьютер

Человеческий мозг — это гигантская сеть из десятков миллиардов нервных клеток (нейронов), связанных между собой отростками (дендритами и аксонами). Число связей одного нейрона может достигать десятков тысяч. Благодаря работам нейрофизиологов достаточно хорошо известен механизм действия отдельного нейрона. Отвлекаясь от быстрых переходных процессов, можно сказать, что нервная клетка способна находиться в одном из трех дискретных состояний — покое, возбуждении и рефрактерности (состоянии невозбудимости). Переходы между состояниями управляются как процессами внутри самой клетки, так и электрическими сигналами, поступающими к ней по отросткам от других нейронов. Переход от состояния покоя к возбуждению происходит пороговым образом при почти одновременном поступлении достаточно большого числа импульсных сигналов возбуждения.

Оказавшись в возбужденном состоянии, нейрон находится в нем в течение определенного времени, а затем самостоятельно переходит в состояние рефрактерности. Это состояние характеризуется очень высоким порогом возбуждения: нейрон практически не способен реагировать на приходящие к нему сигналы возбуждения. Через некоторое время способность к возбуждению восстанавливается и нейрон возвращается в состояние покоя.

При переходе в возбужденное состояние в выходном отростке (аксоне) генерируется импульс возбуждения, распространяющийся по нему со скоростью от 1 до 100 м/с; в основе процесса распространения лежит изменение локальной проводимости мембраны аксона по отношению к ионам натрия и калия. Между нейронами нет прямых электрических контактов. Перенос сигнала с аксона на входной отросток (дендрит) другого нейрона происходит химическим путем в специальной области — синапсе, где окончания двух нервных клеток подходят близко друг к другу. Некоторые из синапсов являются особыми. Под действием сигналов возбуждения они формируют в дендрите импульс торможения, имеющий обратную электрическую полярность. При одновременном поступлении в нейрон такие импульсы способны гасить сигналы возбуждения.

Помимо устройства отдельной нервной клетки относительно хорошо изучены глобальные аспекты деятельности мозга — специализация его больших областей, функциональные связи между ними и т. п. В то же время, мало известно, как осуществляется обработка информации на промежуточном уровне, в участках нейронной сети, содержащих всего десятки тысяч нервных клеток.

Иногда мозг уподобляют колоссальной вычислительной машине, отличающейся от привычных компьютеров лишь значительно большим числом составляющих его элементов. Считается, что каждый импульс возбуждения переносит единицу информации, а нейроны играют роль логических переключателей в полной аналогии с устройством ЭВМ. Такая точка зрения полностью ошибочна¹. Работа мозга должна основываться на совершенно иных принципах. В мозге нет жесткой структуры связей между нейронами, которая была бы подобна электрической схеме ЭВМ. Надежность его отдельных элементов (нейронов) гораздо ниже, чем элементов, используемых для создания современных компьютеров. Разрушение даже таких участков, которые содержат довольно большое число нейронов, зачастую почти не влияет на эффективность обработки информации в этой области мозга. Часть нейронов отмирает при старении организма. Никакая вычислитель-

¹ Поверхностную аналогию между мозгом и компьютером подвергал критике один из основателей молекулярной биологии Ф. Крик [64].

ная машина, построенная на традиционных принципах, не сможет работать при таких обширных повреждениях.

Современные вычислительные машины выполняют операции последовательно, по одной операции на такт (подобно человеку с арифмометром или логарифмической линейкой). Число извлекается из *памяти*, помещается в *процессор*, где над ним производится некоторое действие в соответствии с диктуемой *программой* инструкцией, и результат вновь направляется в память (рис. 3.1). Вообще говоря, при выполнении отдельной операции электрический сигнал должен пробежать по соединительным проводам расстояние порядка размера центрального процессора. Отсюда возникает ограничение на быстродействие такой вычислительной машины.

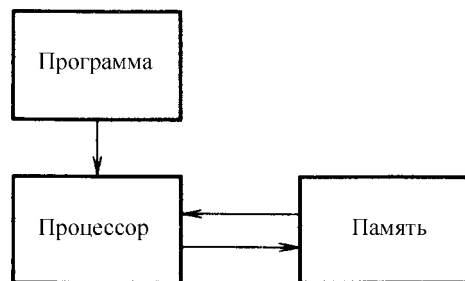


Рис. 3.1. Упрощенная блок-схема традиционной ЭВМ

Пусть размер процессора равен 30 см. Электрический сигнал пробежит это расстояние по металлическим проводам со скоростью света за время 10^{-9} с. Поэтому, если все операции выполняются последовательно, теоретический предел быстродействия этой вычислительной машины составляет миллиард операций в секунду. В действительности быстродействие, кроме того, ограничивается скоростью срабатывания отдельных элементов компьютера. В целом быстродействие современных машин уже довольно близко подошло к своему теоретическому пределу. Этого быстродействия, однако, совершенно недостаточно, чтобы организовать управление сложными системами, непрерывный анализ больших потоков информации, решение типичных задач «искусственного интеллекта».

Если распространить приведенные выше рассуждения на человеческий мозг, результаты будут абсурдными. Ведь скорость распространения сигналов по нервному волокну не превышает 100 м/с, что в сотни миллионов

раз меньше скорости света. Если бы мозг работал, как традиционная ЭВМ, теоретический предел его быстродействия равнялся бы всего тысяче логических операций в секунду. Даже по самым грубым оценкам этого совершенно недостаточно, чтобы объяснить высокую эффективность работы мозга.

Очевидно, что деятельность мозга связана с параллельной обработкой информации (известно, например, что мозг решает за 100 шагов задачи, требующие от традиционного компьютера выполнения нескольких миллионов операций). К настоящему времени организация параллельных вычислений уже используется в наиболее совершенных ЭВМ, чтобы преодолеть ограничения, связанные с конечной скоростью распространения сигналов. Принцип проведения вычислений в несколько параллельных потоков реализован в мощных суперкомпьютерах, но в наиболее полной мере параллелизм выполнения операций используется в компьютерах с *матричными* процессорами.

Матричный процессор можно представить себе как сеть из более простых (примитивных) процессоров. Каждый примитивный процессор обладает небольшой собственной памятью и может обмениваться информацией лишь со своими ближайшими соседями по сети (рис. 3.2).

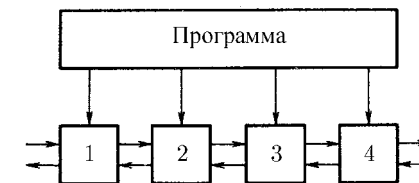


Рис. 3.2. Упрощенная блок-схема ЭВМ с матричным процессором. Каждый примитивный процессор 1, 2, 3, ..., обладает своей собственной памятью и обменивается информацией с ближайшими соседями

Техника параллельного вычисления наиболее эффективна для решения задач, удовлетворяющих условию локальности, когда изменение состояния данного малого элемента среды за малый промежуток времени определяется лишь мгновенным состоянием этого элемента и его ближайших соседей. Примитивный процессор «знает» лишь о состоянии своего малого элемента среды. Основываясь на этой информации и на полученных от соседних процессоров сведениях о состоянии ближайших элементов, каждый процессор вычисляет, согласно инструкции, поступающей от центрального программного блока, состояние своего элемента в следующий момент времени. Ограничение на быстродействие, обусловленное конечной скоростью распространения сигналов, здесь отсутствует, ибо обмен сигналами идет только между ближайшими соседями, а расстояние между ними можно сделать очень малым.

Допустим, что наша задача заключается в решении уравнения теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}. \quad (3.1)$$

Соответствующая явная разностная схема имеет вид

$$u_j^{n+1} = u_j^n + \frac{\tau}{h^2}(u_{j+1}^n + u_{j-1}^n - 2u_j^n), \quad (3.2)$$

где τ — шаг по времени, h — шаг по координате.

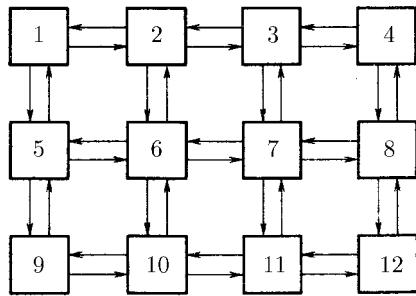


Рис. 3.3. Упрощенная блок-схема вычислительной сети. Каждый примитивный компьютер 1, 2, 3, ..., обменивается информацией с соседями по сети

Моделирование процесса теплопроводности на матричном процессоре осуществляется крайне просто. Каждый примитивный процессор в цепочке хранит в памяти только одно число u_j^n — свое состояние в текущий момент времени — и получает от соседей сведения об их состояниях u_{j-1}^n и u_{j+1}^n в тот же момент времени. На основе этой информации по алгоритму (3.2) примитивный процессор вычисляет свое состояние в следующий, $(n + 1)$ -й, момент времени, после чего цикл повторяется.

Очевидно, что аналогичным образом на матричном процессоре могут моделироваться и другие явления, описываемые уравнениями электродинамики, диффузии, гидро- и аэродинамики и т. п. Этим способом можно моделировать также эволюцию квантовых систем, для которых изменение со временем волновой функции подчиняется уравнению Шредингера. Отдельный класс задач, для решения которых с успехом применяются матричные процессоры, — обработка изображений. Требуемые для такой обработки основные операции (усиление контрастности, выделение контуров и т. п.) являются локальными и могут выполняться одновременно для разных областей обрабатываемой картины.

Работа матричного процессора устойчива по отношению к локальным повреждениям. Если какой-то из примитивных процессоров вышел из строя, это скажется лишь на правильности вычислений, производимых в его малой окрестности, и не приведет к срыву работы всего компьютера.

Следующим этапом в развитии идеи параллельных вычислений явилось создание *вычислительных сетей*. В отличие от описанного выше матричного процессора, где все отдельные примитивные процессоры проводят вычисление одновременно, по единой команде из центрального программного блока, вычислительная сеть представляет собой систему связанных между собой относительно автономных примитивных компьютеров. Каждый из них содержит все основные компоненты самостоятельного компьютера, т. е. память, процессор и программный блок (рис. 3.3). Такое своеобразное «сообщество» примитивных компьютеров проводит коллективно параллельные вычисления, обмениваясь информацией друг с другом.

Вычислительная сеть напоминает многоклеточный организм. После того как указана программа (т. е. репертуар реакций) каждого примитивного компьютера и начальные состояния их всех, сеть как бы начинает «жить своей жизнью» и прямо моделирует в ходе своей эволюции интересующий нас процесс.

Довольно очевидно, что функционирование вычислительной сети как сообщества компьютеров не зависит существенно от того, как именно устроен каждый отдельный компьютер, какими процессами внутри его обеспечена обработка информации.

В настоящее время уже выпускаются ЭВМ, представляющие собой вычислительные сети. Они состоят из относительно малого числа примитивных компьютеров, каждый из которых обладает сложной внутренней структурой и является, в сущности, традиционным малым компьютером с последовательным выполнением операций. Можно представить себе, однако, и иную ситуацию, когда число примитивных компьютеров очень велико, но каждый из них — это чрезвычайно простое устройство, способное совершать всего несколько операций и хранить в своей памяти мгновенные значения нескольких величин.

С математической точки зрения подобные сети, состоящие из элементов с простым репертуаром реакций, принято рассматривать как клеточные автоматы (см. раздел 3.35).

Мозг гораздо ближе к матричному процессору, чем к традиционной ЭВМ с последовательным выполнением операций. Однако существует фундаментальное различие между мозгом и любым параллельным компьютером. Дело в том, что нейронные сети мозга вообще не заняты никакими вычислениями.

Это утверждение может показаться странным. Чем же мы занимались в школе на уроках арифметики? Тут нужно прежде всего напомнить, что сейчас мы обсуждаем не интегральные аспекты деятельности мозга, а характер функционирования его небольших участков. Способность к абстрактно-

му мышлению и, в частности, к обращению с числами и математическими символами свойственна только человеку, т. е. приобретена на очень поздней стадии эволюционного развития. Это не сопровождалось какими-либо существенными изменениями в устройстве нервных клеток и организации нейронной сети. Абстрактное мышление вторично, надстроечно по отношению к фундаментальным механизмам работы мозга.

Бессмысленно допускать, что когда кошка настигает в прыжке птичку, ее мозг решает в считанные доли секунды сложные дифференциальные уравнения, описывающие траекторию прыжка. Глубинные механизмы, на которых базируется человеческое мышление, должны быть общими с теми, которые были у наших эволюционных предшественников. Язык и способность манипулировать с абстрактными символами должны строиться из чего-то, что уже было у «неразумных» животных.

В связи с этим хотелось бы привести высказывание А. Эйнштейна: «Слова и язык, по-видимому, не играют никакой роли в моем механизме мышления. Физические сущности, которые в действительности, видимо, служат элементами мышления, — это определенные знаки и более или менее ясные образы, которые могут «произвольно» воспроизводиться и комбинироваться (...). Такая комбинационная игра является, возможно, существенным признаком продуктивного мышления — когда еще нет никакой связи с логической конструкцией из слов или каких-то признаков, предназначенных для сообщения другим людям. В моем случае указанные элементы принадлежат к зрительному и отчасти даже мускульному типу. Обычные слова приходится подбирать лишь на второй стадии, когда упомянутая игра уже достаточно упорядочена и может воспроизводиться по желанию» [1].

Разумеется, те образцы, о которых говорит Эйнштейн, относятся уже к довольно высокому уровню функционирования мозга. И все же его слова проливают, как кажется, свет на характер процессов, лежащих в основе мышления. Можно предположить, что мозг хранит и перерабатывает информацию в образах. Он работает как колоссальная «аналоговая» машина, где окружающий мир находит отражение в пространственно-временных структурах активности нейронов. Образное моделирование позволяет строить предсказания будущих событий и планы действий.

Подобный механизм работы мозга мог, естественно, возникнуть в ходе биологической эволюции. Для простейшего животного основная функция нервной системы состоит в том, чтобы преобразовать ощущения, вызываемые внешним миром, в определенную двигательную активность. Дождевой червь стремится покинуть освещенное место, лягушка мгновенно атакует любой передвигающийся малый предмет, оказавшийся в ее поле зрения. Существенные для поведения животного аспекты внешнего мира формиру-

ют определенный образ — картину активности некоторой группы нейронов. Этот образ, в свою очередь, должен вызывать соответствующий акт движения, т. е. трансформироваться в определенную активность другой группы нейронов, управляющих сокращениями мышечных волокон. На ранних стадиях эволюции связь между образом-ощущением и образом-движением является прямой, однозначной и наследственно закрепленной в исходной структуре соединений между нейронами. На более поздних стадиях эта связь усложняется, появляется способность к индивидуальному обучению. Теперь уже образ-ощущение не связан жестко с планом действий. Вначале осуществляется его промежуточная обработка и сравнение с хранящимися в памяти картинками. Промежуточная обработка образов становится все более сложной по мере движения вверх по эволюционной лестнице. В конечном счете после длительного развития формируется процесс, называемый нами *мышлением*.

Приведенные выше общие рассуждения нуждаются в дальнейшей конкретизации. Прежде всего, необходимо уточнить, что именно мы вкладываем в понятие «образ».

Ясно, что мозг не в состоянии полностью моделировать все явления в окружающем человека (или животное) мире. Вместо этого он использует в своей работе построенную в нем картину внешнего мира, которая включает только те свойства, которые существенны для жизнедеятельности человека (или животного). Для этого необходима первичная обработка поступающего из внешнего мира информационного потока, с тем чтобы выделить из него существенные характеристики. Подобная первичная обработка сопровождается, по-видимому, *сжатием информации*.

Сжатие информации означает проведение классификации: различные состояния окружения необходимо отнести к одному и тому же классу, если они обладают одними и теми же свойствами. Ясно, что классификация может быть успешной, лишь если она основывается на определенных инвариантах внешнего мира. Распознавание этих инвариантных характеристик (или *закономерностей*) является одной из наиболее важных функций мозга.

Например, все буквы «а», независимо от своего написания, должны распознаваться как единый объект; все треугольники с разными ориентациями и отношениями сторон должны рассматриваться как объекты, отличные от квадратов или окружностей. Мы умеем узнавать людей, т. е. способны определить, что перед нами один и тот же человек, независимо от его возраста, одежды, выражения лица и т. п. Мозг способен распознавать временные последовательности. В частности, нам нетрудно установить идентичность мелодии независимо от того, исполняется ли она какими-то музыкальными инструментами, человеческим голосом, хором голосов или

же целым симфоническим оркестром. Наконец, решение задач человеком часто представляет собой их распознавание по типу требуемой для решения последовательности действий.

Образ — это типичная картина (мелодия, последовательность действий), отвечающая целому классу различных допустимых картин. *Распознавание* образов предполагает, что мозг хранит в своей памяти наборы разных типичных картин и может сравнивать с ними любую поступающую на вход картину, основываясь на определенных критериях.

Типичные картины, хранящиеся в мозге, являются либо генетически унаследованными, либо индивидуально приобретенными. Различие между этими двумя ситуациями не очень глубокое. В обоих случаях имеет место процесс *обучения*, в ходе которого типичные картины каким-то путем формируются на основе опыта, — разница лишь в характерных временах обучения.

В простейшей форме процесс обучения можно описать следующим образом. Имеется набор картин, о которых известно, что они принадлежат к одному и тому же классу. Требуется найти такую картину-прототип, которая будет достаточно близкой ко всем картинам из этого класса и далекой от картин из других классов согласно определенному критерию. Чтобы решить эту задачу, достаточно случайно генерировать большое число различных картин и произвести сравнение каждой из них со всеми имеющимися в наличии картинами из данного набора. Тогда в качестве картины-прототипа можно выбрать ту картину, для которой совпадение будет наиболее полным.

Приведенная простая схема нуждается в нескольких существенных уточнениях. Во-первых, было бы крайне нерационально генерировать потенциальные прототипы чисто случайным образом. Намного эффективнее другой способ. Допустим, что нами было случайно порождено некоторое число картин и путем сравнения мы отобрали среди них одну наиболее близкую. Теперь мы могли бы взять эту картину за основу и подвергнуть ее различным случайным модификациям (мутациям), получив вновь некоторый набор картин. После этого весь процесс может быть повторен снова — и так до тех пор, пока не будет достигнут приемлемый результат. Это рассуждение обнаруживает сходство между обучением и процессом *эволюционного* развития: генерируемые картины конкурируют между собой, причем лишь наиболее эффективная («приспособленная») дает потомство.

Во-вторых, обучение далеко не всегда сводится к построению картин-прототипов, которые буквально напоминали бы предъявляемые изображения. Если цель состоит, например, в том, чтобы научиться отличать все изображения, обладающие зеркальной симметрией, то очевидно, что никакого единого «прототипа» для всех зеркально-симметричных картин

не существует. Вообще говоря, обучение есть процесс, в ходе которого система вырабатывает деление на категории, или классы. После обучения система ставит в соответствие всем «входным» изображениям из данного класса одну определенную «выходную» картину, но эта выходная картина может вовсе не быть того же типа, что и входное изображение.

Наконец, результатом обучения может являться распознавание не только изображений, но и мелодий или последовательностей событий. Путем обучения часто осуществляется *решение задач*. Действительно, решить некоторую задачу — значит найти (или распознать!) оптимальную картину, т. е. траекторию на карте, последовательность событий, маршрут «по дереву решений» некоторой игры.

Как уже отмечалось, деятельность мозга принципиально отличается от работы универсальной цифровой вычислительной машины. Это означает, что все операции по распознаванию, генерации, сравнению образов должны осуществляться мозгом без перевода их на язык числовых последовательностей.

Известно, что альтернативой вычислениям служит *аналоговая* обработка информации. В простейшем случае она представляет собой прямую имитацию. Техника электронной аналоговой имитации широко использовалась в середине нашего столетия до изобретения цифровых ЭВМ. Чтобы решить, например, задачу об определении траектории снаряда, можно попытаться построить специальную электрическую схему с лампами, сопротивлениями, конденсаторами и т. п., так что динамика токов в ней будет подчиняться тем же самым дифференциальным уравнениям, как и движение снаряда. Тогда, подобрав подходящие начальные условия и включив эту электрическую схему, по ее динамике можно предсказать траекторию движения снаряда. Из приведенного примера видно, что аналоговая машина не является программируемой — точнее, программа жестко встроена в ее физическую организацию. Это, наряду с невысокой точностью имитации, послужило причиной отказа в технике от электронных аналоговых машин.

Сегодня интерес к аналоговым машинам возрождается, но уже на качественно ином уровне. Главное внимание привлекают *распределенные* аналоговые машины, представляющие собой пространственные сети из взаимодействующих между собой элементов. «Единицами» обработки информации в подобных машинах являются целые пространственные картины.

Узкую специализацию аналоговых машин можно преодолеть, наделив их способностью к обучению. Рассмотрим этот вопрос подробнее. Предположим, что в законы взаимодействия между элементами, формирующими сеть аналоговой машины, первоначально жестко («физически») встроена только одна программа — программа обучения. Повинуясь этой программе,

в процессе предварительной тренировки система перестраивает свою структуру: в ней устанавливаются новые или разрываются старые связи между элементами сети, модифицируются параметры, характеризующие отдельные элементы, и т. п. Пройдя обучение, эта аналоговая машина приобретает способность к решению некоторой задачи — распознаванию образов, принадлежащих определенному набору, либо отысканию оптимального пути (траектории, последовательности действий), и т. п.

В настоящее время получает все более широкое признание точка зрения, согласно которой мозг человека (и животных) представляет собой именно аналоговое обучающееся устройство. Из экспериментов, например, известно, что процесс обучения у человека (и животных) сопровождается установлением новых синаптических контактов между нейронами и модификацией уже имеющихся синаптических связей. Установлено, что память не локализована в отдельных нейронах или небольших группах нервных клеток. Хранящиеся в памяти образцы не теряются, а лишь как будто тускнеют при повреждениях отдельных участков головного мозга.

Представление о мозге как о гигантской обучающейся аналоговой машине заимствует, в сущности, многое из общей теории условных рефлексов Павлова, отказываясь от упрощенных механических интерпретаций этой теории.

Работа мозга по распознаванию образов носит многоступенчатый иерархический характер. Очевидно, что анализ зрительной картины или сложного звукового сигнала нельзя осуществить с применением какой-то одной простейшей процедуры. Рассмотрим, например, как могла бы выглядеть последовательность шагов, если задача заключается в том, чтобы отличить треугольник от всех других многоугольников.

Первый шаг состоит в выделении линий и проверке того, что их достаточно большие участки близки к отрезкам прямых. Затем необходимо выделить вершины, т. е. точки, где сходятся две или несколько таких линий. Далее необходимо построить как бы промежуточное, сжатое изображение картины — ее *граф*. На этом графе фиксируется лишь наличие связей между вершинами, независимо от длины или ориентации линий, осуществляющих такие связи. Теперь, когда каждый граф отвечает целому классу (например, всем различным треугольникам), следует осуществить распознавание графов, что дает окончательный ответ на поставленную задачу.

Мы видим, что даже решение такой простой задачи требует ее перевода на некоторый внутренний «язык» и построения, фактически, определенных *семантических структур* — графов. Элементы такой структуры — это уже, в сущности, абстрактные объекты или понятия (в описанном выше случае — вершины и связи, осуществляемые отрезками прямых линий).

Каждый граф, в свою очередь, также может рассматриваться в качестве некоторого понятия (например, граф треугольника) и служить элементом, входящим в состав семантической структуры следующего иерархического уровня.

Основной вид деятельности у человека и высших животных — это операции с семантическими структурами: их распознавание, генерация, передача, преобразование и сравнение. С этой точки зрения, появление у человека языка не было чем-то совершенно новым. По словам М. Минского, «задолго до того как наши предки научились говорить, у них уже возникли специальные механизмы мозга для представления объектов, различий и причин; эти механизмы позднее легли в основу нашего языка (и грамматики в том числе)» [22].

Все необходимые операции с семантическими структурами должны осуществляться в мозге аналоговым образом, как некоторые динамические процессы в сложной распределенной нелинейной системе. Мозг как бы является средой, где семантические структуры «живут своей жизнью»: эволюционируют, взаимодействуют и конкурируют между собой.

Хотя выше довольно много говорилось о том, как устроен (или может быть устроен) мозг, необходимо подчеркнуть, что все эти вопросы привлекают внимание прежде всего из-за потенциальных практических приложений. Не ожидая полного выяснения всех вопросов, относящихся к работе мозга, можно ставить задачу построения таких устройств для обработки информации, в которых были бы воплощены некоторые из известных или кажущихся правдоподобными принципов работы мозга.

Главная тенденция в развитии современной вычислительной техники состоит в переходе к использованию распределенных систем, которые образованы из логических элементов с довольно простой внутренней структурой. Большие надежды здесь связывают с *молекулярной микроэлектроникой*. Современный уровень развития технологии позволяет создавать схемы с размерами отдельного логического элемента порядка размеров полимерной молекулы [24, 52, 62]. Разрабатываются методы массового химического синтеза таких молекулярных элементов и способы их соединения в сети на основе механизмов самосборки.

Однако даже если будут окончательно решены проблемы технологического характера, на пути к созданию молекулярного компьютера остается несколько принципиальных теоретических проблем.

Очевидно, что устройство молекулярных размеров не может работать как традиционная ЭВМ с последовательным выполнением операций. На молекулярном уровне невозможно избавиться от локальных дефектов структуры, обусловленных примесями и «дислокациями», а также от влияния теп-

ловых флуктуаций. Все это требует использовать схемотехнику с большой устойчивостью по отношению к локальным повреждениям. Кроме того, как отмечалось выше, сам по себе последовательный характер выполнения операций уже накладывает жесткие ограничения на быстрдействие ЭВМ.

Поэтому молекулярный компьютер должен быть основан на принципе параллельных вычислений. Из нескольких сотен или тысяч молекулярных элементов можно формировать блоки, исполняющие роль отдельного примитивного процессора информации, или клеточного автомата. Сеть из таких блоков, связанных между собой, формирует распределенную вычислительную среду.

Основная трудность, возникающая на этом пути, — это «кризис программируемости». Чем меньше размеры отдельных блоков и чем плотнее их упаковка, тем труднее программировать работу такого компьютера. Лишь в простейших случаях, при прямой аналоговой имитации процессов с локальными взаимодействиями, протекающих в однородных условиях, или при первичной простейшей обработке изображений, все элементы сети должны выполнять идентичные инструкции. Решение более сложных задач обработки информации предполагает определенное разнообразие инструкций, исполняемых различными элементами (блоками) одной и той же сети. Невозможно создать такой единый программный блок, который выдавал бы индивидуальные инструкции для каждого примитивного процессора с молекулярными размерами.

Эту трудность необходимо преодолеть, если наша цель заключается в том, чтобы создать достаточно универсальные молекулярные компьютеры, а не устройство, узкоспециализированное на решении небольшого круга задач.

Выход здесь видят в создании обучающихся систем. В таких распределительных системах на элементном уровне жестко запрограммирована лишь способность к обучению. Приспособление подобной вычислительной сети к решению конкретных задач достигается затем в процессе индивидуального ее обучения или «тренировки». Все разработанные к настоящему времени модели обучающихся сетей в той или иной степени основаны на попытках имитации процессов в нейронных сетях мозга. В связи с этим по отношению к разрабатываемому новому поколению вычислительной техники, широко используемому принципу обучения, часто применяют термины «нейрокомпьютер» или «биокомпьютер».

Хотя наиболее адекватной базой для будущих нейрокомпьютеров является молекулярная электроника, это не исключает создания обучающихся вычислительных сетей на основе относительно традиционной полупроводниковой пленочной технологии или оптоэлектроники.

После этого необходимого введения мы переходим к рассмотрению различных теоретических моделей и подходов, используемых при аналоговой обработке информации в распределенных системах.

35. Клеточные автоматы

Клеточными автоматами принято называть сети из элементов, меняющих свое состояние в дискретные моменты времени по определенному закону в зависимости от того, каким было состояние самого элемента и его ближайших соседей по сети в предыдущий дискретный момент времени.

Что значит задать сеть из элементов? Пусть мы имеем некоторый набор из N элементов, перенумерованных последовательно числами $j = 1, 2, \dots, N$. Нам нужно указать связи между элементами. Для этого необходимо для каждого j -го элемента задать ту группу элементов, которые являются его ближайшими соседями. Будем обозначать множество ближайших соседей j -го элемента как $O(j)$: если $j' \in O(j)$, то элемент с номером j' является ближайшим соседом элемента j .

Состояние отдельно взятого j -го элемента в момент времени n характеризуется некоторой переменной $a_j^{(n)}$. Эта переменная может быть целым числом, действительным или комплексным, либо представлять собой набор из нескольких чисел (т. е. быть «векторной» переменной).

Теперь необходимо определить правила перехода между состояниями. Здесь мы ограничимся рассмотрением однородных клеточных автоматов, для которых все элементы в сети и связи между ними одинаковы. Тогда правила перехода должны быть одними и теми же для любого из элементов. Идентичность элементов означает также, что все ближайшие соседи данного элемента находятся в равном положении и в равной мере влияют на то, каким будет его следующее состояние. Например, закон перехода может иметь вид

$$a_j^{(n+1)} = F\left(a_j^{(n)}, \sum_{j' \in O(j)} a_{j'}^{(n)}\right). \quad (3.3)$$

В этом случае состояние данного элемента в момент времени $n+1$ является однозначной функцией F от двух переменных — состояния этого элемента и суммы состояний его ближайших соседей в предшествующий момент времени n .

При таком определении клеточный автомат не обладает памятью. Клеточные автоматы с памятью можно получить, предположив, что функция F

зависит, например, также от состояния элемента $a_j^{(n-1)}$ в еще более ранний момент времени.

Кроме того, поскольку переходы между состояниями однозначно определены, клеточный автомат, отвечающий правилу перехода (3.3), является *детерминированным*. Можно, однако, указать и такие ситуации, когда переходы имеют случайный характер. Тогда вместо функции F необходимо задать набор вероятностей переходов, например,

$$w = w\left(a_j^{(n+1)} \mid a_j^{(n)}, \sum_{j' \in O(j)} a_{j'}^{(n)}\right), \quad (3.4)$$

которые показывают, какой будет вероятность перехода j -го элемента из состояния $a_j^{(n)}$ в n -й момент времени в состояние $a_j^{(n+1)}$ в последующий $(n+1)$ -й момент времени при условии, что состояния его ближайших элементов в n -й момент принимали определенные значения. Подобные клеточные автоматы называют *вероятностными*.

Приведенные выше общие определения не предполагают регулярность сети. В частности, случайно варьироваться может число связей отдельного элемента, т.е. число его ближайших соседей. Ниже, однако, мы ограничимся рассмотрением лишь регулярных сетей, элементы которых занимают узлы правильной решетки. Ею может быть простая квадратная или гексагональная плоская решетка, кубическая пространственная решетка и т.п.

Хотелось бы подчеркнуть, что тип решетки не связан прямо с размерностью пространства, в которое она погружена. Например, на плоскости вполне можно разметить сеть с топологией кубической решетки. Чтобы представить себе это, возьмем кубическую решетку, где связи между узлами сделаны из гибких изолированных «проводников». Сбросим теперь эту трехмерную сетку на подстилающую плоскость. Очевидно, что тогда мы получим сеть на плоскости, по топологии своих связей полностью эквивалентную правильной кубической решетке.

Кроме задания типа решетки следует указать, какие узлы из окрестности данного элемента мы будем считать его ближайшими соседями. Так, для простой квадратной решетки соседями данного элемента можно считать либо только элементы, расположенные вверх–вниз и налево–направо

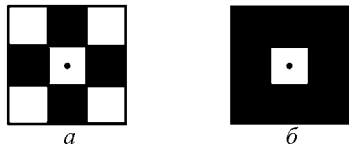


Рис. 3.4. Два возможных способа (а, б) выбрать соседей элемента на квадратной решетке

от него (рис. 3.4а), либо добавить к ним еще и диагональные элементы (рис. 3.4б).

Рассмотрим вначале общие свойства детерминированных клеточных автоматов с конечным числом состояний.

Если отдельный элемент сети может находиться лишь в одном из K состояний, все эти состояния можно перенумеровать, так что переменная $a_j^{(n)}$ будет принимать ряд целых значений от 0 до $K-1$. Предположим, что сеть является правильной и каждый элемент имеет ровно r соседей. Каково общее число возможных правил перехода (3.3) для такой сети?

Согласно (3.3), выбор будущего состояния элемента определяется комбинацией из двух целых чисел: состояния $a_j^{(n)}$ самого этого элемента и суммы состояний $a_{j'}^{(n)}$ его ближайших соседей r в данный дискретный момент времени n . Первое из них принимает K , а второе — rK различных значений. Поэтому число различных комбинаций равно rK^2 . Каждой такой комбинации необходимо сопоставить одно из K возможных значений величины $a_j^{(n+1)}$. Следовательно, полное число различных клеточных автоматов с заданными K и r составляет

$$N_{\text{полн}} = K^{rK^2}. \quad (3.5)$$

Оно неожиданно велико. Например, для квадратной решетки с четырьмя ближайшими соседями ($r=4$), в узлах которой расположены элементы, способные находиться лишь в одном из двух состояний ($K=2$), мы получаем $N_{\text{полн}} = 256$. Если же к ближайшим соседям отнеси также четыре диагональных элемента, то $r=8$ и $N_{\text{полн}} = 2^{32} = 65\,536$.

Динамика рассматриваемых детерминированных клеточных автоматов с конечным числом состояний является необратимой. Действительно, при задании алгоритма переходов каждой из rK^2 возможных комбинаций двух чисел $a_j^{(n)}$ и $\sum a_{j'}^{(n)}$ мы должны сопоставить одно из K возможных значений величины $a_j^{(n+1)}$. Поскольку число исходных комбинаций rK^2 гораздо больше числа конечных состояний K , различным исходным комбинациям соответствуют одни и те же конечные состояния. Но это означает, что по конечному состоянию всей сети невозможно однозначно восстановить ее исходное состояние.

Лишь сравнительно недавно в работах С. Уолфрама [230, 232] было предпринято систематическое моделирование на ЭВМ простейших клеточных автоматов с малым числом состояний. Как показал С. Уолфрам, по своему поведению эти автоматы делятся на четыре класса.

Клеточные автоматы *класса I* достигают за конечное число шагов однородного состояния, в котором величины $a_j^{(n)}$ для всех элементов сети одинаковы и не зависят от времени. Это однородное состояние устанавливается независимо от того, каким было исходное состояние (картина активности) сети. В процессе эволюции для таких клеточных автоматов полностью теряется информация о начальных условиях.

Приведем простой пример клеточного автомата класса I. Пусть элементы занимают узлы одномерной цепочки и каждый элемент имеет два ближайших соседа — слева и справа от него. Пусть также каждый элемент может находиться в двух состояниях: $a_j = 0$ (покой) и $a_j = 1$ (активность). Рассмотрим следующее правило перехода:

$$a_j^{(n+1)} = \begin{cases} 1, & \text{если } a_j^{(n)} = a_{j-1}^{(n)} = a_{j+1}^{(n)} = 1, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (3.6)$$

Нетрудно убедиться, что, согласно этому правилу, все элементы сети приходят со временем к состоянию покоя. Действительно, после каждого следующего шага исчезают все одиночные активные элементы, а в группах, состоящих из двух и более активных элементов, в состояние покоя переходят два крайних элемента.

Клеточные автоматы *класса II* генерируют локализованные простые структуры. Эти простые структуры могут быть стационарными или периодическими по времени.

Рассмотрим два примера клеточных автоматов этого класса, представляющих собой одномерную цепочку из элементов с двумя возможными состояниями — покоя и активности.

1. Пусть элемент цепочки сохраняет состояние активности, если в предшествующий момент времени среди его соседей слева и справа было не больше одного активного элемента, и переходит из состояния покоя в состояние активности, если в этот момент времени были активными оба его соседа:

$$a_j^{(n+1)} = \begin{cases} 1, & \text{если } a_j^{(n)} = 1, \quad a_{j-1}^{(n)} + a_{j+1}^{(n)} = 0, \\ 1, & \text{если } a_j^{(n)} = 1, \quad a_{j-1}^{(n)} + a_{j+1}^{(n)} = 1, \\ 1, & \text{если } a_j^{(n)} = 0, \quad a_{j-1}^{(n)} + a_{j+1}^{(n)} = 2, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (3.7)$$

Такой клеточный автомат генерирует простые периодические по времени структуры.

На рис. 3.5 показаны состояния активности цепочки в последовательные дискретные моменты времени $n = 1, 2, 3, \dots$. Использовано периодическое граничное условие: крайний правый элемент считается соседом крайнего левого элемента. Кроме периодических по времени структур (мигающие тройки) присутствуют также уединенные элементы, сохраняющие состояние активности с течением времени.

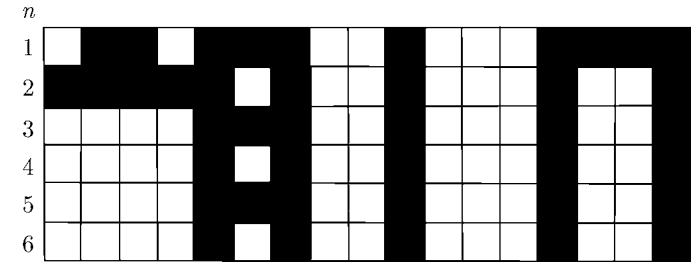


Рис. 3.5. Пример действия клеточного автомата, задаваемого уравнением (3.7). Белые клетки соответствуют элементам в состоянии покоя, черные клетки — активным элементам. Каждая строка характеризует картину активности цепочки в определенный дискретный момент времени $n = 1, 2, 3, \dots$

2. Пусть каждый элемент имеет по два соседа слева и справа в цепочке, а переходы между состояниями осуществляются по правилу

$$a_j^{(n+1)} = \begin{cases} 1, & \text{если } q_j^{(n)} = 3, \\ 0 & \text{в остальных случаях,} \end{cases} \quad (3.8)$$

где

$$q_j^{(n)} = \sum_{l=-2}^2 a_{j+l}^{(n)}. \quad (3.9)$$

Этот клеточный автомат генерирует стационарные локальные структуры — тройки из активных элементов (рис. 3.6).

Клеточные автоматы *классов III и IV* обладают более сложной динамикой. Вообще говоря, в пределе $N \rightarrow \infty$ различные картины активности сети для таких автоматов сменяют друг друга, никогда не повторяясь. Локализованное локальное возмущение порождает процесс изменения активности, который захватывает с течением времени все большую часть сети. Отличительной особенностью клеточных автоматов класса III является то,

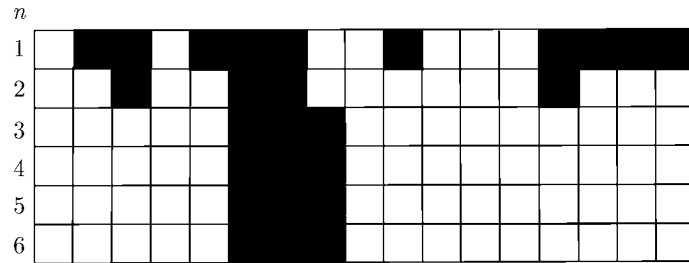


Рис. 3.6. Пример действия клеточного автомата, задаваемого уравнениями (3.8), (3.9) (использованы периодические граничные условия)

что по прошествии достаточно большого числа шагов статистические свойства разворачивающегося процесса теряют зависимость от начальных условий [230]. Тем самым автоматы класса III обладают «эргодическим», или «турбулентным», поведением. Примеры таких клеточных автоматов можно найти в [230].

Динамика клеточных автоматов класса IV существенно зависит от начальной картины активности сети. Подбирая определенные начальные условия, можно генерировать самые различные последовательности сменяющих друг друга картин. Клеточные автоматы класса IV могут, подобно «машине Тьюринга», осуществлять универсальные вычисления (подробнее см. [230–232]).

Хотя имеется несколько кандидатов на роль клеточного автомата класса IV, в настоящее время твердо установлено только, что таким автоматом является игра «Жизнь», предложенная в 1970 г. в качестве математического развлечения Джоном Конвеем. Правила этой игры очень просты: каждый элемент может находиться в состоянии покоя или активности; элемент переходит из состояния покоя в активное состояние, если по соседству с ним оказались три активных элемента². Состояние активности сохраняется при наличии среди ближайших соседей двух или трех активных элементов.

Наблюдать за эволюцией автомата «Жизнь» чрезвычайно интересно. Уже зрительно заметно, что его динамика не является хаотической, а скорее «регулярна». Какие именно из регулярных последовательностей картин будут осуществляться — зависит от начальных условий. Стараниями сотен любителей компьютерных игр к сегодняшнему дню зафиксировано уже

²В число соседей включены только четыре ближайших элемента на квадратной решетке.

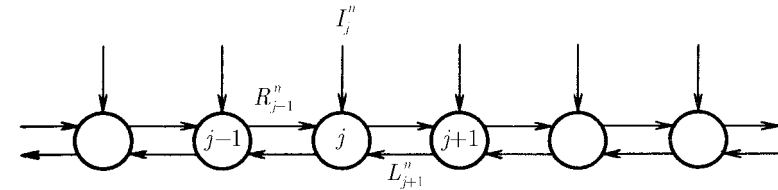


Рис. 3.7. Схема клеточного автомата — детектора движущихся объектов

огромное число различных начальных условий, приводящих к качественно различным последовательным картинам развития [45, 99].

Используя клеточные автоматы, можно решать различные задачи обработки информации. В качестве конкретного примера рассмотрим сеть, способную выделять движущиеся объекты на фоне шума [152].

Пусть в каждый момент времени n на вход системы поступает линейное изображение, представляющее собой некоторую последовательность I_j^n из единиц и нулей (1 — если элемент j освещен, и 0 — в отсутствие света). Известно, что это изображение является суперпозицией шума и изображения нескольких регулярно смещающихся объектов (для простоты мы будем считать, что максимальная скорость их перемещения V не превосходит единицы, т. е. за один шаг по времени изображение объекта не может передвинуться более чем на один элемент цепочки). Задача системы состоит в том, чтобы подавить шум и выделить изображения перемещающихся объектов.

Рассмотрим цепочку, состоящую из элементов, каждый из которых связан с двумя своими соседями (рис. 3.7). Пусть мгновенное состояние отдельного элемента характеризуется набором из трех целых чисел S_j^n , R_j^n и L_j^n , причем S_j^n принимает только значения 0 и 1, а значения R_j^n и L_j^n лежат в интервале от 0 до M . На эту цепочку непрерывно проецируется изображение, т. е. на вход j -го элемента в момент n поступает двоичный сигнал I_j^n .

Предположим, что правила переходов между различными состояниями имеют вид

$$\begin{aligned}
 L_j^{n+1} &= \begin{cases} L_{j+1}^n + (1 - S_j^n), & \text{если } I_j^{n+1} = 1, \\ L_{j+1}^n - 1, & \text{если } I_j^{n+1} = 0, \end{cases} \\
 R_j^{n+1} &= \begin{cases} R_{j-1}^n + (1 - S_j^n), & \text{если } I_j^{n+1} = 1, \\ R_{j-1}^n - 1, & \text{если } I_j^{n+1} = 0, \end{cases} \\
 S_j^{n+1} &= I_j^{n+1}.
 \end{aligned} \tag{3.10}$$

Если вычисленное по правилам (3.10) значение R_j^{n+1} оказывается отрицательным, следует положить $R_j^{n+1} = 0$, если же оно превосходит максимально допустимое значение M — мы полагаем $R_j^{n+1} = M$. Аналогично поступим по отношению к значениям L_j^{n+1} .

Таким образом, по цепочке слева направо передаются сигналы R_j^n , а в обратном направлении — сигналы L_j^n . Передавая сигнал своему соседу, элемент: а) увеличивает его значение на единицу, если в настоящий момент времени этот элемент освещен, а в предыдущий момент он был темным; б) сохраняет значение сигнала неизменным, если он был освещен в предыдущий момент времени и остался освещенным; в) уменьшает значение сигнала на единицу, если он не оказался освещенным в наступивший момент времени. При достижении предельных значений 0 или M значение сигнала соответственно более не уменьшается или не увеличивается.

Если построить клеточный автомат с правилами перехода (3.10) и подождать некоторое время, по состояниям его элементов можно определить координаты движущихся объектов и направление их движения. Именно, те элементы в цепочке, где достигнуто максимальное значение $R_j^n = M$, отмечают положение объектов, совершающих регулярное перемещение слева направо. Если же объект смещается справа налево, то в соответствующем его мгновенному положению элементе цепочки L_j^n достигает максимального значения. Случайные сигналы, обусловленные шумом, усредняются и не приводят к достижению предельно больших значений этих величин.

По данным [152], даже при уровнях шума в 25% такой алгоритм обеспечивает правильное выделение движущихся объектов более чем в половине случаев. Заметим также, что он допускает простое обобщение на случай двумерных изображений.

Клеточные автоматы используются также для моделирования гидродинамических течений. Хорошо известно, что уравнения гидродинамики описывают макроскопические усредненные движения в системе, состоящей из огромного числа взаимодействующих друг с другом молекул. Оказывается, однако, что на макроуровне теми же самыми уравнениями описывается и несравненно более простая система — решеточный газ, являющийся одним из примеров клеточного автомата.

Первая модель жидкости как решеточного газа была предложена в работах Д. Харди, О. де Пацциса и И. Помо [117, 118], которые рассматривали регулярную квадратную решетку со связями единичной длины. Считалось, что в каждом узле решетки могут находиться не более четырех молекул равной массы с единичной скоростью, направленной вдоль одного из четырех направлений (налево, направо, вверх или вниз). Одновременное пребывание

в узле двух молекул с одинаковым направлением скорости запрещено. Время также является дискретным. Изменение состояния на следующем шаге по времени происходит согласно следующим правилам. Молекула смещается на одну длину решетки — в ближайшую ячейку, на которую направлен вектор ее скорости. При этом любая конфигурация, состоящая в точности из двух молекул с направленными навстречу друг другу скоростями в одном узле (лобовое столкновение), заменяется на другую конфигурацию, где скорости повернуты на угол 90 градусов по сравнению с первоначальными. Все остальные конфигурации остаются без изменений.

Эта модель обладает рядом важных свойств. Главным из них является существование термодинамического равновесия. Хотя эргодическая теорема не доказана, численные расчеты свидетельствуют о релаксации к равновесию. Возникающие равновесные распределения обладают свободными непрерывными параметрами — средней плотностью и импульсом. Равновесные функции распределения не зависят от положения узла, но зависят от направления скорости (если только средний импульс частиц не обращается в нуль). Когда плотность и импульс медленно меняются в пространстве и по времени, справедливо «макродинамическое» описание, уравнения которого отличаются от нелинейных уравнений Навье–Стокса в двух отношениях — отсутствием галилеевой инвариантности и изотропии.

Галилеевская инвариантность, т. е. эквивалентность всех инерциальных систем отсчета, нарушается из-за наличия решетки. Это нарушение проявляется в том, что в выражение для тензора плотности потока импульса входят не только квадратичные по средней скорости u члены, но также и слагаемые, имеющие более высокий порядок. Эти нелинейные поправки оказываются, однако, малыми в пределе низких чисел Маха, когда жидкость можно рассматривать как несжимаемую. Поэтому, если ограничиться изучением лишь ситуаций, когда сжимаемость жидкости несущественна, нарушение галилеевской инвариантности не будет заметно.

Более серьезная трудность связана с отсутствием изотропии, т. е. инвариантности относительно произвольных поворотов. Решеточный газ на квадратной решетке инвариантен лишь при поворотах на угол, кратный $\pi/2$. В результате тензор плотности потока импульса для такой модели имеет вид

$$P_{\alpha\beta} = p\delta_{\alpha\beta} + T_{\alpha\beta\gamma\epsilon}u_\gamma u_\epsilon + O(u^4). \quad (3.11)$$

Здесь $p = \rho/2$ есть давление, ρ — плотность числа частиц; члены, нечетные по u , исчезают из-за симметрии по отношению к пространственным отражениям. Тензор T симметричен по индексам α, β и γ, ϵ . У. Фриш, Б. Хасслахер и И. Помо [90] обратили внимание, что если бы решеточный газ был инвариантен по отношению к гексагональной группе поворотов (т. е. по отноше-

нию к вращениям на угол $\pi/3$), тензор T был бы изотропным и выражение (3.11) приняло бы следующий вид:

$$P_{\alpha\beta} = (p + \mu u^2)\delta_{\alpha\beta} + \lambda u_{\alpha}u_{\beta} + O(u^4). \quad (3.12)$$

При низких числах Маха (малые средние скорости u) выражение (3.12) для тензора плотности потока импульса совпадает с входящим в уравнение Навье–Стокса. С учетом этого в [90] предложена модель газа на гексагональной плоской решетке — клеточный автомат, правила которого следующие.

Частицы могут находиться в узлах правильной треугольной решетки. Скорости всех частиц одинаковы по модулю (обычно их выбирают равными 1), но различны по направлению. Скорость частицы может быть направлена лишь вдоль одной из шести связей, которые соединяют тот узел, где она находится, с его ближайшими соседями. В каждом узле может находиться не более одной частицы с данным направлением скорости. Удобно перенумеровать все шесть допустимых направлений скорости с помощью индекса i , положив $\varphi = i\pi/3$, где $i = 0, \pm 1, \pm 2, \dots$, а φ — угол, определяющий направление скорости. Очевидно, что значения i , отличающиеся на любое целое число, кратное 6, отвечают одному и тому же направлению. Обозначим скорость частицы, имеющую направление i , как e_i .

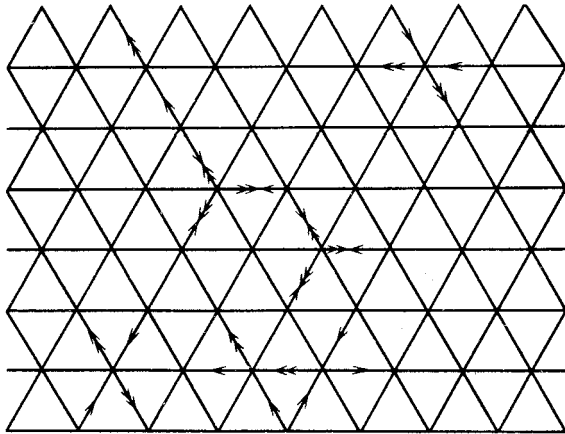


Рис. 3.8. Клеточный автомат для уравнения Навье–Стокса

Положения частиц и их скорости меняются в дискретные моменты времени синхронно для всех узлов решетки. За один шаг по времени каждая

частица смещается в тот смежный узел решетки, на который был направлен вектор ее скорости. Если в результате такого смещения в одном узле оказалось две или три частицы, в тот же момент времени происходит акт рассеяния, приводящий к изменению скоростей столкнувшихся частиц. Если сталкивается пара частиц с противоположными направлениями скорости e_i, e_{i+3} , то скорости частиц после столкновения равны e_{i+1}, e_{i-2} или e_{i-1}, e_{i+2} , причем оба варианта рассеяния равновероятны. При столкновении трех частиц со скоростями e_i, e_{i+2}, e_{i-2} новые направления скорости равны $e_{i+3}, e_{i+1}, e_{i-1}$. Во всех остальных случаях направления скорости в результате рассеяния не изменяются.

Так, например, на рис. 3.8 показаны положения и направления скорости частиц в два последовательных момента времени n и $n+1$. Простая стрелка, выходящая из узла решетки, означает, что в момент n в этом узле находилась частица с соответствующим направлением скорости. Двойными стрелками указаны положения и скорости частиц на следующем, $(n+1)$ -м шаге по времени.

Приведем вкратце схему вывода уравнений Навье–Стокса для решеточного газа, описываемого таким клеточным автоматом (подробное рассмотрение этого вопроса можно найти, например, в [234]).

Пусть N_i — среднее число частиц с направлением скорости i , приходящееся на один узел решетки. Усреднение производится по макроскопической пространственно-временной области, охватывающей много узлов решетки и много шагов по времени, так что N_i является плавной функцией от пространственных координат и от времени. Определим медленно меняющуюся плотность числа частиц ρ и плотность импульса $\rho\mathbf{u}$ как

$$\rho = \sum_i N_i, \quad \rho\mathbf{u} = \sum_i N_i \mathbf{e}_i. \quad (3.13)$$

При равновесии с заданными значениями ρ и \mathbf{u} средние числа заполнения N_i даются распределением Ферми–Дирака:

$$N_i = \frac{1}{\exp[f(\rho, \mathbf{u}) + g(\rho, \mathbf{u})\mathbf{e}_i\mathbf{u}] + 1}, \quad (3.14)$$

где f и g есть некоторые функции от ρ и \mathbf{u} . Их можно разложить в ряд по степеням средней скорости \mathbf{u} . С помощью этих разложений удастся вычислить плотность потока массы и плотность импульса с точностью до вторых степеней по макроскопическим градиентам, что приводит к следующим

гидродинамическим уравнениям:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0, \quad (3.15)$$

$$\begin{aligned} \frac{\partial}{\partial t}(\rho u_\alpha) + \sum_{\beta} \frac{\partial}{\partial x_\beta} [g(\rho) \rho u_\alpha u_\beta + O(u^4)] = \\ = -\frac{\partial p}{\partial x_\alpha} + \eta_1(\rho) \Delta u_\alpha + \eta_2(\rho) \operatorname{grad} \operatorname{div} \mathbf{u}, \end{aligned} \quad (3.16)$$

где $g(\rho) = (\rho - 3)/(\rho - 6)$; $\eta_1(\rho)$ и $\eta_2(\rho)$ представляют собой коэффициенты вязкости (ср. [20]).

Линеаризация этих уравнений и отбрасывание членов, учитывающих вязкость, дают волновое уравнение, описывающее распространение звуковых волн со скоростью $\sqrt{2}$. Система уравнений (3.15), (3.16) переходит в уравнение Навье–Стокса для несжимаемой вязкой жидкости, если рассмотреть предел, когда число Маха $M = u\sqrt{2}$ стремится к нулю, а гидродинамическая длина L неограниченно возрастает так, чтобы их произведение ML оставалось конечным. В этом пределе вариации плотности следует учитывать лишь в давлении $p = \rho/2$, а уравнение непрерывности (3.15) сводится к $\operatorname{div} \mathbf{u} = 0$. Тогда множитель $g(\rho)$ в уравнении (3.16) постоянен, и при $0 < \rho < 3$ его можно исключить, перейдя к новым единицам времени. В результате число Рейнольдса окажется равным

$$\operatorname{Re} = \frac{ML\rho g(\rho)}{\sqrt{2}\eta_1(\rho)}. \quad (3.17)$$

Так, на макроскопическом уровне при не слишком больших числах Маха M (от 0,3 до 0,5) рассматриваемый клеточный автомат дает в точности те же гидродинамические течения, что и реальная жидкость³. Это открытие чрезвычайно важно для компьютерных расчетов — на современных ЭВМ и в особенности на будущих ЭВМ с высокой степенью параллелизма. Поскольку все вычисления осуществляются в целых числах, здесь нет ошибок округления, дающих нефизический источник шума. Кроме того, сам алгоритм расчета идеально подходит для реализации на распределенных вычислительных сетях.

³Правильное граничное условие $\mathbf{u}_\tau = 0$ можно получить, полагая, что при столкновении с неподвижной границей происходит случайное рассеяние частиц обратно в полуплоскость падения.

Приведенная модель клеточного автомата, имитирующего двумерные гидродинамические течения несжимаемой вязкой жидкости, не является единственно возможной. Тем же самым макроскопическим поведением обладает, например, несколько другая модель (рис. 3.9). В этой модели частицы располагаются на связях гексагональной решетки (связи, занятые частицами, отмечены на рис. 3.9 стрелками). К настоящему времени имеются обобщения этой модели на случай магнитогидродинамики [54]. Первые результаты моделирования турбулентных течений с помощью клеточных автоматов сообщены в [77].

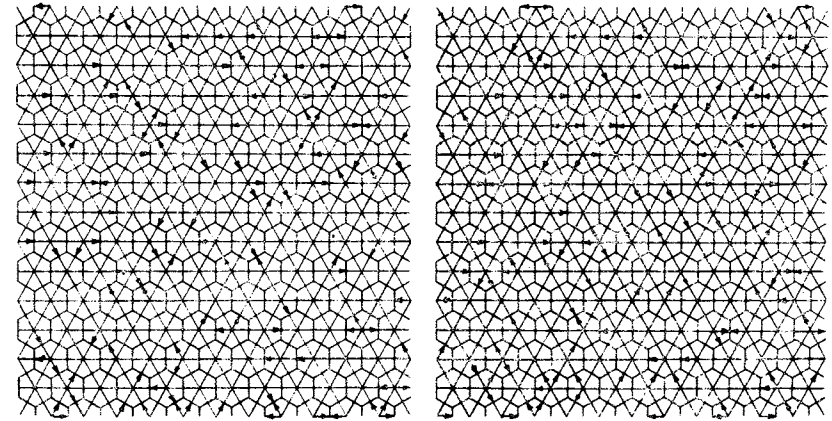


Рис. 3.9. Две последовательные микроскопические конфигурации для клеточного автомата, имитирующего гидродинамические течения

К сожалению, до сих пор отсутствуют модели клеточных автоматов для имитации трехмерных течений жидкости. Главная трудность здесь заключается в том, что для макроскопической изотропии в трехмерном случае необходима икосаэдрическая инвариантность, но не существует ни одной регулярной кристаллографической решетки со столь высокой симметрией. Поэтому приходится искать более сложные обходные пути.

Мы столь подробно рассмотрели этот частный, хотя и важный в практическом отношении пример, поскольку он иллюстрирует одну все более отчетливую тенденцию: переход от цифровых расчетов на универсальных ЭВМ к прямой имитации интересующих нас процессов с использовани-

ем параллельно действующих вычислительных сетей. Можно ожидать, что в ближайшем будущем появится много других клеточных автоматов, имитирующих самые различные виды процессов в физических, химических и биологических средах.

Однако, как уже отмечалось в разделе 3.34, возможности аналоговых распределенных систем не исчерпываются прямой имитацией процессов в распределенных средах. Далее обсуждаются более сложные функции по обработке информации, реализуемые с помощью таких систем.

36. Ассоциативная память

Мы говорим, что система обладает ассоциативной памятью, если при подаче на ее вход некоторой картины она автоматически отбирает и подает на выход наиболее близкую к ней хранящуюся в памяти картину. Иными словами, по достаточно большому фрагменту или искаженному изображению такая система может восстановить полное изображение. Очевидно, что ассоциативная память является частным случаем распознавания образов.

Один из возможных путей реализации аналоговой ассоциативной памяти состоит в том, чтобы построить распределенную динамическую систему (или сеть из дискретных элементов), аттракторами которой в ее конфигурационном пространстве являлись бы типичные картины-образы. Каждая такая картина будет обладать своей областью притяжения, и всякое начальное условие, представляющее собой какую-то допустимую картину, обязано попасть в одну из ее областей притяжения. С течением времени в ходе эволюции эта начальная структура трансформируется в наиболее близкую из хранящихся в памяти структур-аттракторов, а именно в ту, области притяжения которой она принадлежала. Следовательно, подавая на вход в качестве начального условия для такой распределенной системы некоторую структуру, мы будем осуществлять ее автоматическое (т.е. аналоговое) распознавание, которое будет параллельным, поскольку оно выполняется в ходе однократной эволюции данной динамической системы. Итак, мы приходим к задаче целенаправленного создания распределенных динамических систем с заданными аттракторами.

Рассмотрим здесь конкретную реализацию динамической системы с ассоциативной памятью, предложенную в работах Дж. Хопфилда [136, 137]. Поскольку стимулом к созданию этой модели послужили теоретические исследования особых физических систем — спиновых стекол, вначале приведем некоторые сведения о таких системах.

Предположим, что рассматриваемая система состоит из N элементов («изинговских спинов»), каждый из которых может находиться всего в одном из двух состояний: $S_i = +1$ или $S_i = -1$. Энергия E взаимодействующих спинов дается выражением

$$E = -\frac{1}{2} \sum_{i,j} J_{ij} S_i S_j, \quad (3.18)$$

где J_{ij} — элементы некоторой матрицы взаимодействия (весовые коэффициенты связей). Не конкретизируя пока временную динамику взаимодействующих спинов, будем считать лишь, что в процессе эволюции эта система стремится понизить свою полную энергию, а в пределе $t \rightarrow \infty$ приходит в состояние с минимумом полной энергии.

Если все весовые коэффициенты J_{ij} положительны, в состоянии с минимумом полной энергии E все спины имеют одинаковое направление $S_i \equiv +1$ или $S_i \equiv -1$. При этом, как легко убедиться, энергия взаимодействия любой пары спинов $\delta E_{ij} = -J_{ij} S_i S_j$ достигает своего минимального значения $\delta E_{ij} = -J_{ij}$. Такое состояние системы взаимодействующих спинов принято называть «ферромагнитным».

В спиновых стеклах матрица $\mathbf{J} = [J_{ij}]$ состоит из случайных элементов, принимающих как положительные, так и отрицательные значения. В состоянии с минимумом полной энергии E для спинового стекла энергии взаимодействий *всех* пар спинов, как правило, не могут быть одновременно минимальными. Действительно, среди N спинов обычно найдутся такие тройки S_i, S_j, S_k , что произведение $J_{ij} J_{jk} J_{ki}$ будет отрицательным. Но тогда, как бы мы ни выбирали направления этих трех спинов, нам не удастся сделать минимальными все три энергии парных взаимодействий $\delta E_{ij}, \delta E_{jk}$ и δE_{ki} . Условия минимальности энергии взаимодействия для различных пар спинов несовместимы (конфликтуют) между собой. В результате, как говорят, система оказывается *фрустрированной* («frustration» (англ.) — расстройство планов, крушение надежд).

Вследствие фрустрации спиновое стекло имеет много состояний с минимумами полной энергии, отвечающих различным спиновым конфигурациям — пространственным картинам ориентации спинов S_i . Для больших N все эти состояния характеризуются примерно одной энергией. Иными словами, при заданных случайных весовых коэффициентах J_{ij} спиновое стекло как бы хранит в памяти большое число определенных пространственных картин. Всякая начальная картина взаимной ориентации спинов с течением времени превращается в одну из устойчивых картин, а именно в ту,

наиболее близкой к которой она является. В этом смысле спиновое стекло обладает не только памятью, но и способностью к распознаванию образов.

К сожалению, однако, сам набор хранящихся в спиновом стекле картин-образов является случайным. Неплохо было бы уметь так выбирать весовые коэффициенты J_{ij} для системы спинов, чтобы устойчивыми оказались не случайные, а некоторые вполне определенные картины, которые нам хотелось бы сохранить в памяти в качестве образов-эталонов. Оказывается, это возможно.

Пусть нам необходимо записать образ, характеризующийся своей определенной ориентацией $S_i = \xi_i$ для каждого спина i . Выберем весовые коэффициенты J_{ij} так, чтобы $J_{ij} = \xi_i \xi_j$. Нетрудно убедиться, что записанная спиновая конфигурация автоматически отвечает минимуму энергии, т. е. является устойчивой притягивающей пространственной картиной. Действительно, при этом для всякой наугад выбранной пары спинов энергия взаимодействия $\delta E_{ij} = -J_{ij} \xi_i \xi_j = -\xi_i^2 \xi_j^2 = -1$, т. е. достигает наименьшего возможного значения.

Полученная таким путем система, однако, не способна распознавать образы. Она хранит в своей памяти всего одну картину, к которой эволюционирует с течением времени любая начальная спиновая конфигурация⁴. Как показано Дж. Хопфилдом, этот недостаток можно устранить. Допустим, что нам хотелось бы записать M различных образов, каждый из которых характеризуется своим набором (последовательностью) ориентации спинов: $\{\xi_j^{(m)}\}$, где $m = 1, 2, \dots, M$, причем различные наборы ортогональны,

$$\frac{1}{N} \sum_{j=1}^N \xi_j^{(m)} \xi_j^{(m')} = \delta_{mm'}. \quad (3.19)$$

Выберем весовые коэффициенты J_{ij} в виде (*правило Хебба* [119])

$$J_{ij} = \sum_{m=1}^M \xi_i^{(m)} \xi_j^{(m)}. \quad (3.20)$$

Тогда для не слишком большого числа записанных картин можно гарантировать, что все они окажутся устойчивыми конфигурациями этого искусственно сконструированного спинового стекла, т. е. будут отвечать минимуму энергии.

⁴Точнее, таких картин две, поскольку кроме конфигурации $S_i = \xi_i$ минимуму энергии отвечает также зеркально-симметричная конфигурация $S_i = -\xi_i$.

Чтобы доказать это, заметим прежде всего, что полную энергию системы можно представить как

$$E = -\frac{1}{2} \sum_{i=1}^N h_i S_i, \quad (3.21)$$

введя поле

$$h_i = \sum_j J_{ij} S_j, \quad (3.22)$$

действующее на спин S_i . Ясно, что в состоянии с минимумом полной энергии все спины должны быть направлены по полю, т. е. знак S_i должен совпадать со знаком h_i . Пусть весовые коэффициенты J_{ij} даются формулой (3.20), а направления спинов отвечают одной из записанных картин $\{\xi_i^{(m')}\}$. Проверим тогда выполнение указанного выше условия

$$h_i S_i = \sum_j \sum_m \xi_i^{(m)} \xi_j^{(m)} \xi_i^{(m')} \xi_j^{(m')} = N \xi_i^{(m')} \xi_i^{(m')} = N > 0. \quad (3.23)$$

Итак, каждый из M записанных образов действительно отвечает некоторой притягивающей спиновой конфигурации, т. е. состоянию с минимумом полной энергии.

Помимо записанных в память образов, характеризующихся набором ориентации спинов $\{\xi_j^{(m)}\}$, система всегда имеет некоторое число ложных образов, также отвечающих состояниям с минимумом энергии. В частности, наряду с каждым записанным образом всегда хранится и его зеркальное отражение, т. е. образ с набором ориентации спинов $\{-\xi_j^{(m)}\}$. Легко проверить, что он отвечает тому же значению полной энергии.

Для ортогональных образов матрица $[J_{ij}]$, построенная по правилу (3.20), является проективной. Действуя на произвольный вектор с N компонентами ξ_j , эта матрица проецирует его на M -мерное подпространство, которое содержит векторы $\{\xi_j^{(m)}\}$. В пределах такого подпространства действие матрицы на любой вектор сводится к умножению на N . Это означает, что для любого вектора с компонентами $\xi_j = \sum_m C_m \xi_j^{(m)}$, представляющего собой линейную комбинацию M записанных образов, справедливо $h_j \xi_j = (\mathbf{J}\xi)_j \xi_j = N \xi_j^2 > 0$. Далеко не всякая линейная комбинация, однако, является допустимой. Для того чтобы отвечать какому-то мыслимому образу, все компоненты ξ_j в этой линейной комбинации должны принимать лишь значения ± 1 . Если записанные в память образы таковы, что они позволяют построить линейную комбинацию, являющуюся некоторым новым

образом, этот дополнительный ложный образ также отвечает минимуму полной энергии и является притягивающей спиновой конфигурацией.

Модель Хопфилда можно переформулировать на языке нейронных сетей. Когда суммарный потенциал, создаваемый на нейроне импульсами возбуждения и торможения, приходящими по дендритам от синапсов, превышает пороговый уровень, нейрон испускает импульс возбуждения, распространяющийся от него по аксону к другим нервным клеткам. Если поддерживать этот потенциал примерно постоянным, нейрон будет генерировать частую периодическую последовательность импульсов возбуждения, разделенных короткими интервалами рефрактерности, — он будет находиться в «активном» состоянии. При потенциале, не превышающем пороговый уровень или имеющем обратную полярность, нервная клетка сохраняет состояние покоя. Будем характеризовать состояние отдельного нейрона с номером j дискретной переменной S_j , присвоив ей значение $+1$, если нейрон активен в рассматриваемый малый промежуток времени, и значение -1 , если он находится в состоянии покоя.

Как уже отмечалось в разделе 3.34, нейроны связаны друг с другом посредством синапсов, которые делятся на тормозящие и возбуждающие. Пусть есть всего два нейрона, связанных через тормозящий синапс. Тогда, если первый нейрон активен ($S_1 = +1$), генерируемый им сигнал преобразуется в сигнал торможения для второго нейрона и последний находится в состоянии покоя ($S_2 = -1$). Ситуация меняется, когда нейроны связаны через возбуждающий синапс. Тогда активность первого нейрона вызывает (при достаточно низком пороге) активность второго и реализуется состояние $S_2 = S_1 = +1$. Удобно характеризовать синаптическую связь некоторым коэффициентом J_{12} , считая его положительным, если синапс возбуждающий, и отрицательным — для тормозящего синапса; модуль коэффициента J_{12} характеризует «силу» связи.

Синаптическая связь является однонаправленной. Чтобы сигналы распространялись от каждого из двух нейронов, необходимо иметь два разных синапса, коэффициенты J_{12} и J_{21} которых, вообще говоря, различны. Для простоты будем считать, что связи симметричны: $J_{21} = J_{12}$. Тогда можно формально построить энергию $E_{12} = -J_{12}S_1S_2$ и убедиться, что возможные устойчивые картины активности системы из двух нейронов действительно отвечают минимуму этой величины. Нетрудно также проверить, что для системы, состоящей из большого числа нейронов, устойчивые картины активности отвечают минимуму величины, получаемой путем суммирования значений E_{ij} , для всех пар нейронов (предполагается, что каждый нейрон связан в сети со всеми другими).

Итак, аналогия со спиновым стеклом вполне очевидна. Чтобы записать образ, характеризующийся активностями нейронов ξ_i , необходимо каким-то

способом сформировать картину синаптических связей, характеризующуюся коэффициентами $J_{ij} = \xi_i \xi_j$. Если мы хотим записать M различных образов, система синаптических связей должна характеризоваться коэффициентами, получаемыми путем суммирования соответствующих коэффициентов для всех этих образов.

Из сказанного, разумеется, не следует, что в нейронных сетях мозга реализуется именно такой механизм длительной памяти. Прямые экспериментальные свидетельства этого до настоящего времени отсутствуют, хотя и известно, что процесс обучения сопровождается установлением новых синаптических контактов между нейронами, а также модификацией уже имевшихся синаптических связей. Кроме того, установлено, что память об отдельных событиях не локализована в отдельных нейронах или небольших группах нервных клеток. Хранящиеся в памяти образы не теряются, а лишь как будто тускнеют при повреждениях отдельных участков головного мозга.

Требование строгой ортогональности записываемых в память образов (3.19) является слишком жестким. Оказывается, что при большом полном числе спинов N случайно взятые образы обычно будут почти ортогональными и этой приближенной ортогональности уже достаточно для работы устройства ассоциативной памяти.

Предположим, что мы взяли наугад две случайные последовательности $\{\xi_i^{(1)}\}$ и $\{\xi_i^{(2)}\}$, состоящие из $+1$ и -1 , причем каждый элемент в них выбирается независимо от остальных, а значения $+1$ и -1 равновероятны. Для двух таких образов каждая из величин $r_i = \xi_i^{(1)} \xi_i^{(2)}$ представляет собой случайное число с двумя равновероятными значениями $r_i = \pm 1$, а входящее в условие ортогональности (3.19) скалярное произведение

$$R_{12} = \sum_{i=1}^N \xi_i^{(1)} \xi_i^{(2)} \quad (3.24)$$

является суммой N независимых случайных чисел r_i . Поэтому при больших N дисперсия случайной величины R_{12} растет как

$$\langle R_{12}^2 \rangle = N. \quad (3.25)$$

Следовательно, при $N \rightarrow \infty$ справедлива оценка

$$\frac{1}{N} \sqrt{\langle R_{12}^2 \rangle} \sim \frac{1}{\sqrt{N}}. \quad (3.26)$$

Итак, в пределе большого полного числа спинов N два случайно выбранных образа обычно оказываются почти ортогональными.

Отклонения от ортогональности приводят к появлению дополнительных случайных вариаций в поле h_i (3.22), действующем на спины, причем каждый из M хранящихся образов вносит свой независимый вклад в интенсивность таких случайных вариаций. Влияние этого шума на правильность распознавания образов было детально изучено в [32–34], где получены следующие результаты.

1. При любом конечном значении отношения M/N записанные в память образы воспроизводятся системой с некоторыми искажениями; кроме того, всегда имеются также ложные устойчивые образы, представляющие собой суперпозицию из M записанных в память картин.

2. Примерно при $M/N = 0,05$ в модели наблюдается своего рода «фазовый переход»: минимумы энергии, отвечающие истинным образам, становятся менее глубокими, чем те, которые соответствуют ложным образам. В результате даже содержащий незначительные случайные искажения предъявленный образ может эволюционировать с течением времени не к ближайшему из записанных образов, а к одному из ложных. Следовательно, при отношениях M/N , превышающих это критическое значение, опасность неверного срабатывания устройства ассоциативной памяти становится существенной.

3. Примерно при $M/N = 0,4$ в системе происходит второй «фазовый переход», после которого записанные картины перестают отвечать минимумам энергии, а все притягивающие конфигурации соответствуют только ложным образам. Это означает полную потерю способности к ассоциативной памяти.

В существенной степени приведенные результаты зависят от предположения о полной случайности записываемых в память образов. Если записываемые образы коррелированы, то интенсивность «шума» будет более значительной, а ошибки в распознавании возрастут. Это ограничение является довольно серьезным, поскольку на практике все записываемые образы часто принадлежат одному и тому же определенному классу и, следовательно, не полностью случайны. Чтобы приспособить систему к работе с коррелированными образами, можно отказаться от правила Хебба (3.20) и использовать другие схемы записи.

Заметим, что при использовании правила Хебба запись дополнительного $(M + 1)$ -го образа заключается в следующей простой модификации матричных элементов:

$$J_{ij}^{(M+1)} = J_{ij}^{(M)} + \xi_i^{(M+1)} \xi_j^{(n+1)}. \quad (3.27)$$

Если все образы ортогональны, то выполняются следующие равенства:

$$\sum_j J_{ij}^{(M)} \xi_j^{(m)} = \sum_j \xi_i^{(m)} \xi_j^{(m)} \xi_j^{(m)} = N \xi_i^{(m)}. \quad (3.28)$$

Это означает, что каждый из векторов $\{\xi_i^{(m)}\}$ является тогда собственным вектором матрицы J .

Предположим, что новый записываемый образ не ортогонален образам, уже хранящимся в памяти. Тогда в качестве простой характеристики корреляции можно рассмотреть, насколько «старая» матрица узнает «новый» вектор:

$$\xi_i^{\parallel} = \frac{1}{N} \sum_j J_{ij}^{(M)} \xi_j^{(M+1)}. \quad (3.29)$$

Если бы новый вектор $\xi^{(M+1)}$ был ортогонален ко всем старым векторам, мы получили бы $\xi^{\parallel} = 0$. Если же новый вектор уже хранится в памяти, то $\xi^{\parallel} = \xi^{(M+1)}$. В данном случае было бы неверно по-прежнему использовать правило Хебба, поскольку тогда мы бы эффективно удвоили глубину энергетического минимума, отвечающего такому вектору. Более глубокий минимум обычно имеет более широкую область притяжения, а это уменьшает области притяжения соседних минимумов и искажает соответствующие притягивающие векторы. В предельном случае это может привести к тому, что в системе останется всего один очень глубокий минимум энергии, область притяжения которого захватывает все возможные спиновые конфигурации. У системы формируется как бы «навязчивая идея», что, конечно же, крайне нежелательно. Чтобы избежать этой опасности, есть простой выход: оставить матрицу элементов без изменений, если окажется, что $\xi^{\parallel} = \xi^{(M+1)}$. Этому требованию отвечает следующее правило:

$$J_{ij}^{(M+1)} = J_{ij}^{(M)} + \xi_{ij}^{\perp} \xi_j^{(M+1)}, \quad (3.30)$$

где $\xi_{ij}^{\perp} = \xi_j^{(M+1)} - \xi_j^{\parallel}$.

Правило (3.30) имеет два недостатка. Во-первых, получаемая матрица $[J_{ij}]$ оказывается несимметричной, а следовательно, все предыдущие рассуждения, основывающиеся на рассмотрении минимумов энергии, несправедливы⁵. Во-вторых, такое правило некоммутативно: точность, с которой хранится записанный образ, зависит от того, в какой последовательности

⁵Как определить динамику системы с ассоциативной памятью в случае несимметричной матрицы $[J_{ij}]$ — см. ниже.

предъявляется системе список образов в процессе обучения. Последний из записанных образов стремится подавить все предыдущие. От этого недостатка свободно правило обучения

$$J_{ij}^{(M+1)} = J_{ij}^{(M)} + \epsilon \xi_i^\perp \xi_j^{(M+1)}, \quad (3.31)$$

где $0 < \epsilon < 1$. Если набор записываемых образов предъявлять вновь и вновь (в случайной последовательности), можно доказать, что такая процедура сходится и приводит к формированию хорошей матрицы взаимодействий $[J_{ij}]$. Использование алгоритма (3.31) в системах ассоциативной памяти обсуждалось в [38, 70].

Наконец, можно указать также более сложное правило обучения, сохраняющее свойство симметрии матрицы $[J_{ij}]$, а именно

$$J_{ij}^{(M+1)} = J_{ij}^{(M)} + \frac{\xi_i^\perp \xi_j^\perp}{\xi^\perp \xi^\perp}, \quad (3.32)$$

где

$$\xi^\perp \xi^\perp = \frac{1}{N} \sum_i (\xi_i^\perp)^2. \quad (3.33)$$

При $\xi^\perp = 0$ последнее слагаемое в (3.32) полагается равным нулю. Можно убедиться, что такая схема устраняет все трудности, связанные с возможной корреляцией между записываемыми образами. Однако она является более сложной и нелокальной, что может оказаться серьезным препятствием, если процесс обучения системы также должен реализоваться в аналоговой форме (см. раздел 3.38). Существует также несколько других схем записи коррелированных образов, описание которых можно найти в [70].

Во всех упомянутых нами схемах получающиеся элементы матрицы J_{ij} могут принимать непрерывный ряд отрицательных и положительных значений. Это не очень удобно при создании конкретных устройств ассоциативной памяти в виде электронных сетей и т. п. Поэтому интересно обсудить, как меняются свойства системы, если взять матрицу взаимодействий в виде

$$J'_{ij} = \text{sign}(J_{ij}), \quad (3.34)$$

так что ее матричные элементы равны лишь ± 1 . Удивительно, но оказывается [70], что во многих случаях правило Хебба (3.20) лишь выигрывает от такого огрубления. По всей видимости, это происходит от того, что появление самых больших значений элементов матрицы взаимодействия связано

с наличием корреляций. Из этого, однако, не следует, что на практике нужно использовать именно правило Хебба. Другие правила приводят к лучшим результатам, хотя их действенность и ухудшается при огрублении.

Заметим также, что свойство распознавания образов в модели Хопфилда сохраняется при случайном разрыве даже достаточно большого числа связей, т. е. если положить $J_{ij} = 0$ для случайно выбранных пар спинов i и j . Разрыв связей, как и неортогональность образов при использовании правила Хебба, приводит к появлению шума в поле, действующем на спины. Если связи разорваны случайно, шум состоит из большого числа независимых случайных вкладов, которые в значительной мере уничтожают друг друга.

До сих пор мы лишь предполагали, что в ходе своей эволюции система стремится к минимуму энергии (3.18), но не указывали конкретных динамических законов, приводящих к такой эволюции. Ниже мы рассмотрим конкретные реализации модели Хопфилда.

Предположим, что в нашем распоряжении имеется сеть из N связанных попарно между собой формальных «нейронов». Каждый такой нейрон представляет собой очень простой автомат с двумя возможными состояниями, которые мы будем характеризовать спиновой переменной $S_i = \pm 1$. Связь между автоматами i и j характеризуется своим весовым коэффициентом J_{ij} , принимающим как положительные, так и отрицательные значения.

Зададим закон, по которому осуществляется эволюция этой системы взаимодействующих автоматов — «нейронов». Будем предполагать, что в ней имеют место только переходы, связанные с изменением состояния отдельно взятых элементов. Пусть мы знаем состояние всех «нейронов» S_i^n в n -й момент времени. Чтобы найти их состояние в следующий $(n+1)$ -й момент, необходимо наугад выбрать один из N элементов, вычислить действующее на него поле $h_i^n = \sum_j J_{ij} S_j^n$ и определить новое состояние этого

элемента как $S_i^{n+1} = \text{sign} h_i^n$. Вкратце этот алгоритм динамики можно записать как

$$S_i^{n+1} = \text{sign} \left(\sum_{j=1}^N J_{ij} S_j^n \right), \quad (3.35)$$

где $i = i(n)$ — случайно выбираемый для n -го момента времени номер «нейрона», в котором происходит переход.

Как следует из (3.35), новое значение S_i^{n+1} для i -го элемента совпадает со старым значением S_i^n , если $S_i^n = \text{sign} h_i^n$, т. е. если соответствующий спин был направлен в n -й момент времени по полю; в противном случае, если $S_i^n = -\text{sign} h_i^n$, произойдет переворот спина: $S_i^{n+1} = -S_i^n$. Нетрудно

но убедиться, что в ходе эволюции, согласно алгоритму (3.35), «энергия» такой системы автоматов, определяемая выражением (3.18), ни на одном шаге по времени не возрастает. При достижении конфигурации с минимумом энергии, когда для *всех* спинов $S_i = \text{sign } h_i$, переходы прекращаются и устанавливается устойчивое по отношению к переворотам отдельных спинов стационарное состояние. Всякое начальное распределение релаксирует с течением времени к одной из устойчивых конфигураций, отвечающих минимуму энергии.

Необходимо заметить, что, вообще говоря, конечное состояние такой системы не вполне однозначно определяется ее начальным состоянием: результат эволюции может зависеть от конкретного выбора случайной последовательности элементов, в которых на каждом шаге по времени осуществляются переходы. Это означает, что иногда начальная конфигурация, характеризующаяся некоторым значением энергии E , может в зависимости от выбора этой случайной последовательности перейти в различные конечные конфигурации, отвечающие минимуму энергии. Если, однако, начальная конфигурация достаточно близка к одному из записанных образов, т. е. устойчивых конечных конфигураций, эта возможность маловероятна и распознавание образов не нарушается. Для задачи распознавания обычно неинтересны ситуации, когда предъявленный образ лишь чуть-чуть ближе к одному из записанных, чем к остальным.

Помимо рассмотренного выше, существует иной вариант реализации модели Хопфилда — для динамических систем с непрерывным временем, описываемых обыкновенными дифференциальными уравнениями.

Рассмотрим динамическую систему, которая представляет собой совокупность N связанных между собой элементов, описываемых непрерывно меняющимися действительными переменными σ_i . Предположим, что эволюция такой системы со временем подчиняется уравнениям

$$\dot{\sigma}_i = -\frac{\partial W}{\partial \sigma_i}, \quad i = 1, 2, \dots, N, \quad (3.36)$$

где функция W имеет вид

$$W = -\sum_{i,j} J_{ij} \sigma_i \sigma_j + \alpha \sum_i (\sigma_i^4 - \sigma_i^2). \quad (3.37)$$

Если параметр α достаточно велик ($\alpha \rightarrow +\infty$), то в устойчивых стационарных состояниях переменные σ_i могут принимать лишь значения, равные ± 1 . Следовательно, в пределе $\alpha \rightarrow +\infty$ всякий аттрактор динамической системы (3.36) представляет собой некоторый набор $\{S_i\}$ со значениями S_i ,

равными ± 1 . Но на классе таких точек, где все переменные σ_i принимают лишь значения ± 1 , функция W совпадает с рассматривавшимся ранее выражением (3.18) для энергии E системы из N бинарных спинов. Следовательно, для устойчивых стационарных состояний динамической системы (3.36) справедливы все выводы, полученные выше в случае дискретных элементов.

Никаких других аттракторов, кроме устойчивых стационарных точек, система (3.36) не имеет из-за ее особого потенциального вида — W играет роль функции Ляпунова.

В N -мерном фазовом пространстве каждому состоянию $\{\sigma_i\}$ отвечает некоторая точка, а все состояния, в которых σ_i принимают лишь значения ± 1 , соответствуют вершинам N -мерного куба. В зависимости от конкретного выбора элементов матрицы J_{ij} некоторые из этих вершин будут отвечать устойчивым стационарным состояниям динамической системы. Каждая такая точка — аттрактор — обладает своей областью притяжения, причем, поскольку других аттракторов система не имеет, (почти) любая точка фазового пространства принадлежит области притяжения одного из стационарных состояний.

Если каждой притягивающей точке соответствует какой-то записанный в память образ, такая система способна к автоматическому распознаванию образов. Для этого достаточно задать предъявляемый для распознавания образ в качестве начального условия и проследить за последующей временной эволюцией системы.

Рассмотренный нами простой пример иллюстрирует принцип аналогового распознавания образов динамическими системами, но не слишком удобен для практического воплощения. Поэтому ниже мы расскажем о другом примере, допускающем прямую реализацию в виде электронной схемы [140].

Допустим, что в нашем распоряжении имеются нелинейные усилители, для которых напряжения на входе (u_i) и на выходе (v_i) связаны зависимостью $v_i = g(u_i)$, такой, что функция g является монотонно возрастающей и ограниченной сверху и снизу (рис. 3.10); удобно так выбрать начало отсчета и единицы измерения выходного напряжения v_i , чтобы его минимальное

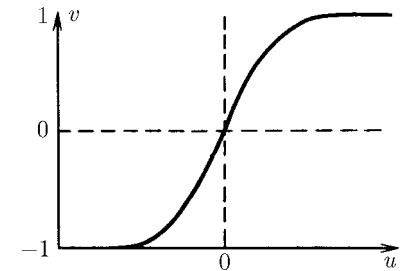


Рис. 3.10. Вид передаточной функции $v = g(u)$ для отдельного нелинейного усилителя

возможное значение равнялось -1 , а максимальное значение было равно $+1$. Предположим, что входное сопротивление усилителя i равно ρ_i , а его входная емкость равна C_i . На вход каждого усилителя поступают электрические сигналы со всех остальных усилителей, причем связь между усилителями осуществляется через переходные сопротивления ρ_{ij} . Помимо нормального каждый из усилителей имеет еще и инвертированный выход с сигналом обратной электрической полярности. Чтобы получить положительный коэффициент связи $T_{ij} = 1/\rho_{ij}$ между усилителями i и j , сигнал на вход i -го усилителя следует подать с нормального выхода j -го усилителя. Отрицательный коэффициент связи $T_{ij} = -1/\rho_{ij}$ получается, если снимать подаваемый на вход сигнал с инвертированного выхода j -го усилителя.

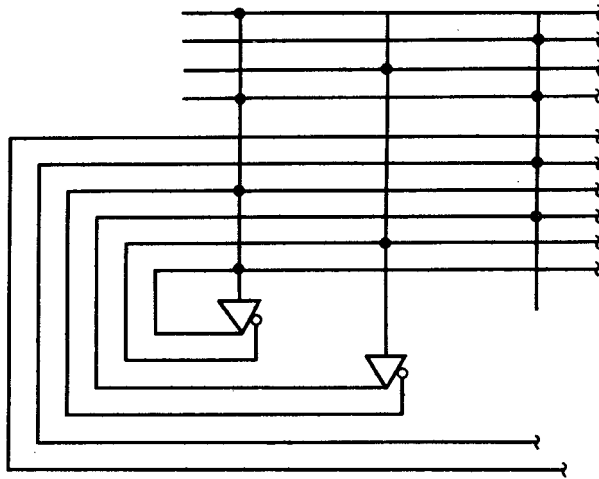


Рис. 3.11. Электрическая схема, реализующая модель Хопфилда. Треугольниками обозначены нелинейные усилители, белыми кружками отмечены инвертированные выходы этих усилителей. Черные кружки — сопротивления, осуществляющие связь между входами и выходами разных усилителей

Соответствующая электрическая схема изображена на рис. 3.11. Изменение со временем входных напряжений в такой системе описывается уравнениями

$$C_i \dot{u}_i = \sum_{j=1}^N T_{ij} v_j - \frac{u_i}{R_i}, \quad v_j = g(u_j), \quad (3.38)$$

где сопротивление R_i получается путем параллельного соединения сопротивлений ρ_i и ρ_{ij} ;

$$\frac{1}{R_i} = \frac{1}{\rho_i} + \sum_j \frac{1}{\rho_{ij}}. \quad (3.39)$$

Рассмотрим функцию W , определяемую как

$$W = -\frac{1}{2} \sum_{ij} T_{ij} v_i v_j + \sum_j \frac{1}{R_j} \int_0^{v_j} \frac{1}{g(v)} dv, \quad (3.40)$$

где функция $1/g(v)$ является обратной к $v = g(u)$. Вычислим производную от функции W по времени:

$$\begin{aligned} \frac{dW}{dt} &= -\sum_{ij} T_{ij} v_i \dot{v}_j + \sum_j \frac{\dot{v}_j}{g(v_j) R_j} = -\sum_j \left[\sum_i T_{ij} v_i - \frac{1}{R_j} u_j \right] \dot{v}_j = \\ &= -\sum_j C_j \dot{u}_j \dot{v}_j = -\sum_j C_j g'(v_j) (\dot{u}_j)^2 \leq 0. \end{aligned} \quad (3.41)$$

Следовательно, W является функцией Ляпунова для системы обыкновенных дифференциальных уравнений (3.38).

Заметим, что функция W ограничена снизу, поскольку второе слагаемое в выражении (3.40) для W стремится к $+\infty$, когда одно или несколько напряжений v_j достигают своих минимальных (-1) или максимальных ($+1$) значений. Все локальные минимумы функции отвечают устойчивым стационарным состояниям рассматриваемой системы. Действительно, из $dW/dt = 0$ следует $\dot{v}_j = \dot{u}_j = 0$. Однако

$$\frac{dW}{dt} = \sum_j \frac{\partial W}{\partial v_j} \dot{v}_j, \quad (3.42)$$

а поэтому всякий минимум ($\partial W/\partial v_j = 0$) отвечает некоторому стационарному состоянию. Устойчивость этого стационарного состояния обеспечена тем, что функция W не может возрастать в процессе эволюции.

Поскольку минимальное и максимальное значения переменных v_i равны, соответственно, -1 и $+1$, областью изменения этих переменных является N -мерный куб $\{-1 < v_j < +1\}$, где N — полное число усилителей. Если предположить, что передаточная функция усилителя $v = g(u)$ имеет почти ступенчатую форму (рис. 3.12), то тогда последнее слагаемое в (3.40)

отлично от нуля лишь в непосредственной близости от границ куба, а внутри него функция W дается фактически лишь первым слагаемым в этом выражении. Оно может иметь минимум внутри этой области, лишь если матрица $[T_{ij}]$ является отрицательно-определенной. В обратном случае — а он как раз и представляет для нас основной интерес — все минимумы лежат на границах N -мерного куба. Если к тому же $T_{ii} = 0$, все минимумы будут лежать в вершинах этого куба, т. е. отвечать наборам параметров $\{v_j\}$, состоящих из ± 1 . Итак, в пределе, когда передаточная функция $g(u)$ близка к ступеньке, устойчивые стационарные состояния системы будут теми же, что и для системы спинов S_i , характеризуемой энергией взаимодействия (3.18) с элементами матрицы $J_{ij} = T_{ij}/2$. Аналогично системе (3.36), в ходе своей временной эволюции эта электрическая схема будет автоматически решать задачу распознавания образов.

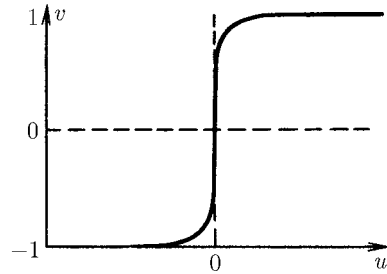


Рис. 3.12. Передаточная функция усилителя $v = g(u)$, имеющая почти ступенчатую форму

элементов со значениями ± 1 .

В табл. 3.2 показано, как выглядела временная эволюция системы вида (3.38) при подаче в качестве начального условия последовательности, отвечающей записи «john s» (время и энергия измеряются в условных единицах). Мы видим, что даже по такому небольшому кусочку система правильно восстанавливает всю запись. Достаточно было изменить всего один символ и задать в качестве начального условия запись «john h», как результат эволюции кардинально изменился (табл. 3.3) и система правильно воспроизвела имя, фамилию и номер телефона другого сотрудника лаборатории.

Возможности системы не ограничиваются восстановлением пропущенных символов. Она способна также восстанавливать запись по *искаженно-*

В заключение продемонстрируем несколько примеров использования модели Хопфилда в качестве устройства ассоциативной памяти.

В [70] в качестве записываемых в память «картин» были взяты имена и номера служебных телефонов шести сотрудников лаборатории фирмы «Белл» (табл. 3.1). Каждая запись состояла из 25 символов, которые кодировались в двоичную форму с затратой 5 бит на один символ. Следовательно, каждую запись можно было представить в виде некоторой двоичной последовательности $\{\xi_i\}$, состоящей из 125

му фрагменту. Табл. 3.4 показывает, как шло восстановление правильной записи по искаженному имени «larry» одного из сотрудников.

Шесть записей, использовавшихся в эксперименте [70], не были ортогональными. Поэтому, кроме истинных, система обладала также ложными притягивающими образами. Эволюция к одному из таких фиктивных образов показана в табл. 3.5.

В [55] ставилась задача распознавания букв русского алфавита. Каждая буква задавалась в виде совокупности черных и белых точек на решетке размером 20×20 . Использовался набор из 15 или менее букв. Поскольку в изображениях букв имелись сильные корреляции, простая схема, основанная на правиле обучения Хебба (3.20), оказалась неэффективной: правильность распознавания терялась уже при трех различных буквах и незначительном уровне наложенного шума. Поэтому перед записью изображения подвергались предварительной ортогонализации и матрица связей $[J_{ij}]$ строилась по правилу

$$J_{ij} = \sum_{m_1, m_2=1}^M \left(\frac{1}{R}\right)^{m_1 m_2} \xi_i^{(m_1)} \xi_j^{(m_2)}, \quad (3.43)$$

где $1/R$ — матрица, обратная к матрице

$$R^{m_1 m_2} = \frac{1}{N} \sum_{i=1}^N \xi_i^{(m_1)} \xi_i^{(m_2)}. \quad (3.44)$$

Здесь $\xi_i^{(m)}$ есть изображение буквы с номером m , перекодированное в линейную последовательность двоичных символов ($\xi_i = \pm 1$). В численном эксперименте [55] $N = 20 \times 20 = 400$, $M \leq 15$.

Для распознавания предъявлялись буквы с наложенным на них случайным шумом, уровень которого достигал 90%. На рис. 3.13 показаны некоторые результаты распознавания букв по их искаженным изображениям.

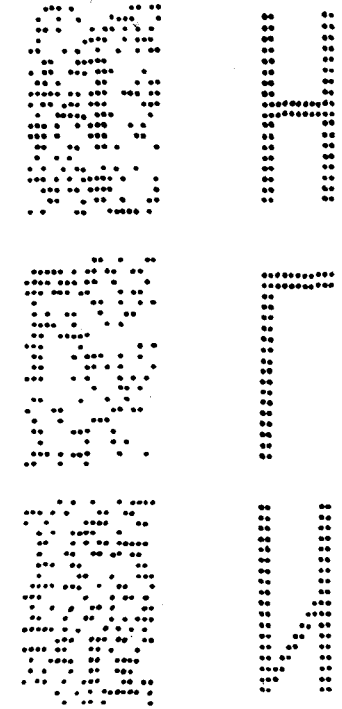


Рис. 3.13. Примеры распознавания букв русского алфавита по их искаженным изображениям

Таблица 3.1

John Stewart Denker	8128
Lawrence David Jacke	17773
Richard Edwin Howard	5952
Wayne P. Hubbard	7077
Brian V. Straughn	3126
John Henry Scofield	8109

Таблица 3.2

Время	Энергия		
0	0	john s	
0,20	-0,0784	john sdewirubneoimv	8109
0,40	-0,8426	john sdewirtbnenimw	8129
0,60	-0,8451	john sdewirtbnenimv	8129
0,80	-0,8581	john sdewirt nenkmv	8128
1,00	-0,9099	john sdewart denker	8128
1,20	-0,9824	john stewart denker	8128

Таблица 3.3

Время	Энергия		
0	0	john h	
0,20	-0,0665	john hdnwybqbofmalt	8109
0,40	-0,8579	john henwybqcofield	8109
0,60	-0,9094	john henry scofield	8109

В табл. 3.6 приведены данные о качестве распознавания (процент правильно идентифицированных изображений) в зависимости от числа записанных в память образов и уровня шума (в процентах).

В модели Хопфилда и ее различных обобщениях распознаваемый образ подается в качестве начального условия, а каждому из запомненных образов

Таблица 3.4

Время	Энергия		
0	0	larry	
0,20	-0,0213	larrynce david jacke	17773
0,40	-0,7977	larrynce david jacke	17773
0,60	-0,8713	larrynce david jacke	17773
0,80	-0,8813	lasrence david jacke	17773
1,00	-1,001	lasrence david jacke	17773
1,20	-1,001	lawrence david jacke	17773

Таблица 3.5

Время	Энергия		
0	0	garbage	
0,20	-0,0244	garbagee lafj naabd	5173
0,40	-0,6280	garbaged derjd naabd	7173
0,60	-0,6904	garbaged derjd naabd	7173
0,80	-0,6904	gasbafed derjd naabd	7173
1,00	-0,7595	gasbated derjd naabd	7173
1,20	-0,7709	fasjebad derjd naabd	7173
1,40	-0,8267	fasjebad derjd naabd	7173
1,60	-0,8282	fasjeb d derjd naabd	7173

соответствует аттрактору (притягивающая стационарная точка) некоторой динамической системы. Возможен, однако, и альтернативный подход.

Рассмотрим такую динамическую систему, которая описывается набором параметров $\{q_k\}$, причем при заданных значениях всех этих параметров она обладает только одним аттрактором. Пусть, однако, эта динамическая система довольно чувствительна к изменению своих параметров, так что при изменении величин $\{q_k\}$ один аттрактор сменяется другим. Ясно, что тогда каждому из возможных качественно различных типов аттрактора отвечает своя область в пространстве параметров, образованном всевозможными наборами $\{q_k\}$. Если теперь выбрать в качестве набора $\{q_k\}$ рассматриваемый образ и проследить, притяжением к какому из аттракторов за-

вершается эволюция системы, это также позволит осуществить аналоговое распознавание. Конкретный пример распознавания образов, использующий согласно этой схеме, свойства динамической системы с конкуренцией мод, рассмотрен в [6].

Таблица 3.6. Процент правильно опознанных символов

Уровень шума, %	Число записанных в память образов		
	2	5	7
10	100	100	100
30	100	100	43
50	100	100	28
70	100	40	14
90	50	20	8

37. Сложные задачи комбинаторной оптимизации

В самых различных приложениях приходится сталкиваться с проблемой выбора оптимального варианта из огромного числа возможностей. Типичным примером этого служит так называемая «задача о коммивояжере». Предположим, что на плоскости случайно разбросаны N точек (городов) A, B, C, D, \dots и известны все расстояния между ними $l_{AC}, \dots, l_{BC}, \dots$. Задача заключается в том, чтобы выбрать минимальный по длине замкнутый путь (маршрут обхода), соединяющий все эти точки.

Подсчитаем полное число имеющихся вариантов. Выбрав некоторую последовательность точек B, F, E, G, D, \dots, W , мы получим определенный маршрут обхода длиной

$$l = l_{BF} + l_{FE} + l_{EG} + l_{GD} + \dots + l_{WB}. \quad (3.45)$$

Полное число подобных последовательностей составляет $N!$, но не все они отвечают разным маршрутам обхода. Начальную точку и направление обхода можно выбрать произвольно, а поэтому каждому маршруту соответствует $2N$ различных последовательностей. Итак, полное число вариантов равно $N!/2N$.

При решении задачи о коммивояжере на ЭВМ путем последовательного перебора вариантов необходимо выполнить не меньше отдельных операций, чем полное число имеющихся вариантов, а с увеличением N число вариантов растет быстрее любой степени N . Поэтому при увеличении числа пунктов обхода затраты машинного времени на поиск оптимального маршрута будут возрастать столь быстро, что точное решение такой задачи на ЭВМ путем последовательного перебора вариантов практически невозможно. Для задачи о коммивояжере удается доказать, что никакого более короткого алгоритма поиска точного решения, имеющего степенную по N сложность, не существует.

Рассмотренный пример является частным случаем задачи *неполиномиальной сложности*, или *NP-проблемы* [100]. Для таких задач отсутствуют строгие алгоритмы поиска решения с длиной, пропорциональной некоторой степени от полного числа N имеющихся элементов. В результате длительность процесса решения экспоненциально быстро растет с увеличением N и никакая сколь угодно мощная ЭВМ не в состоянии его осуществить.

NP-проблемы не являются редким исключением, они типичны для задач комбинаторной оптимизации. Их наличие накладывает жесткие ограничения на сферу применимости современных вычислительных машин. Однако мозг человека (и животных) не испытывает особых трудностей в решении сложных задач комбинаторной оптимизации, возникающих в регулярно повторяющихся условиях. Это, конечно, не противоречит строгим математическим теоремам, так как мозг не дает *точного* решения, обеспечивающего выбор оптимального варианта. Вместо этого довольно быстро отыскивается один из *достаточно* оптимальных вариантов.

Специалистами по прикладной математике предложено много различных приближенных способов решения сложных задач комбинаторной оптимизации. Обычно, однако, такие способы узкоспецифичны. Они пригодны только для решения задачи какого-то одного типа и основываются на результатах, полученных путем сложного математического анализа именно этой задачи. Ясно, что способности к решению задач оптимизации у человека (и животных) должны базироваться на иных принципах.

В последнее время предложено несколько аналоговых методов решения задач комбинаторной оптимизации, для которых приближенный поиск решения осуществляется в ходе прямой эволюции определенной нелинейной системы. Наиболее популярен *метод имитации отжига*, выдвинутый в работе С. Киркпатрика, С. Гелатта и М. Веччи [158]. Мы проиллюстрируем вначале сущность этого метода на примере задачи о разбиении графа на две равные части. Эту задачу приходится часто решать при проектировании сверхбольших полупроводниковых интегральных схем. Известно, что точ-

ное решение такой задачи требует выполнения экспоненциально большого числа операций.

Пусть граф состоит из четного числа элементов и совокупности связей между ними. Задача оптимизации заключается в поиске такого разбиения элементов на две равные по числу элементов группы, чтобы полное число связей между элементами, принадлежащими разным группам, было минимальным (рис. 3.14).

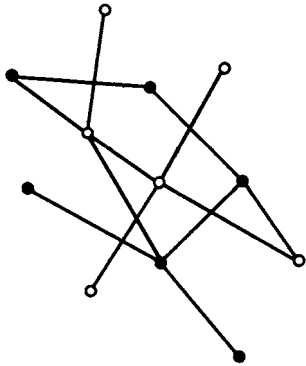


Рис. 3.14. Пример разбиения графа на две равные по числу элементов группы (белые и черные кружки)

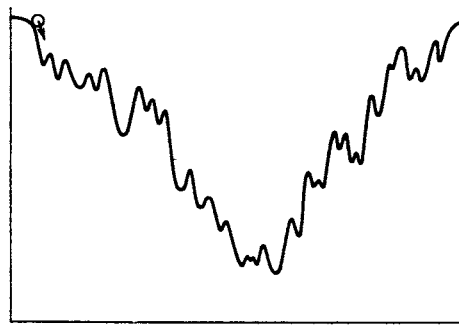


Рис. 3.15. Движение частицы с большим вязким трением по сильно изрезанному потенциальному рельефу

Чтобы дать аналоговый способ решения, нужно сопоставить этой задаче некоторую искусственно сконструированную физическую систему. Именно, поставим в соответствие каждому элементу графа некоторый бистабильный физический элемент (автомат), состояния которого характеризуется спиновой переменной S_i , принимающей значения ± 1 . Введем матрицу связей $[T_{ij}]$, такую, что $T_{ij} = 1$, если элементы с номерами i и j связаны в графе между собой, и $T_{ij} = 0$, если связь отсутствует; кроме того, положим $T_{ii} = 0$. Будем считать, что все элементы в состоянии $S_i = +1$ относятся к первой группе, а элементы в состоянии $S_i = -1$ принадлежат второй группе.

Тогда полное число связей $N_{\text{св}}$ между элементами двух групп можно представить в виде

$$N_{\text{св}} = \frac{1}{8} \sum_{i,j} T_{ij} (S_i - S_j)^2. \quad (3.46)$$

Действительно, если между элементами i и j есть связь (т. е. $T_{ij} = 1$) и они принадлежат к разным группам (т. е. $S_i = -S_j$), это дает вклад в общую сумму, равный четырем, а каждая пара элементов учитывается в сумме дважды.

Простые преобразования выражения (3.46) дают

$$N_{\text{св}} = \frac{1}{2} N_{\text{полн}} - \frac{1}{4} \sum_{i,j} T_{ij} S_i S_j, \quad (3.47)$$

где

$$N_{\text{полн}} = \sum_{i < j} T_{ij} \quad (3.48)$$

есть полное число связей в рассматриваемом графе. Заметим далее, что

$$K = \sum_i S_i \quad (3.49)$$

представляет собой разность числа элементов в двух группах, которая по условию задачи для оптимального разбиения должна равняться нулю.

Оптимальное разбиение графа на две равные по числу элементов группы будет поэтому отвечать абсолютному минимуму функции

$$E = N_{\text{св}} + \lambda K^2 \quad (3.50)$$

при больших положительных значениях множителя λ . Подставляя в (3.50) приведенные выше выражения для $N_{\text{св}}$ и K , получаем

$$E = \frac{1}{2} N_{\text{полн}} + \sum_{i,j} \left(\lambda - \frac{1}{4} T_{ij} \right) S_i S_j. \quad (3.51)$$

Теперь видно, что функцию E можно интерпретировать как энергию системы связанных изинговских спинов (3.18). Каждая спиновая конфигурация, т. е. распределение проекций спинов S_i , отвечает некоторому разбиению графа. Основное состояние такой спиновой системы, для которого энергия E минимальна, при $\lambda \rightarrow \infty$ отвечает оптимальному разбиению графа на две равные по числу элементов группы⁶.

Следовательно, аналоговое решение этой задачи оптимизации возможно, если подобрать такую динамику, т. е. алгоритм переходов между состояниями в системе, которая обеспечивала бы ее переход с течением времени в состояние с абсолютным минимумом энергии.

⁶Если коэффициент λ велик, но конечен, числа элементов в двух группах могут оказаться слегка различными, что обычно не столь существенно в конкретных приложениях.

Казалось бы, для этого можно воспользоваться детерминированным алгоритмом перехода (3.35), введенным при рассмотрении модели Хопфилда. Согласно этому алгоритму, каждый переход, т. е. переворот одного из спинов, осуществляется только в том случае, если он приводит к понижению полной энергии спинов. В результате переходы продолжаются лишь до тех пор, пока не достигается состояние с минимумом энергии.

Однако использование этого простого детерминированного алгоритма, эквивалентного методу наибо́льшего спуска в поиске минимума функции большого числа переменных, не дает правильного ответа. В действительности рассматриваемая нами система имеет огромное число различных стационарных состояний, каждое из которых отвечает какому-то локальному минимуму энергии, лежащему выше абсолютного минимума. Если разрешены только переходы, приводящие к уменьшению энергии, очень скоро система «застревает» в одном из локальных минимумов, из которого она уже не способна выйти.

Это можно пояснить на примере частицы, движущейся с вязким трением по сильно изрезанному потенциальному рельефу (рис. 3.15). Перемещаясь только вниз по градиенту потенциала, частица быстро застрянет в одном из локальных минимумов (промежуточных «долин») и никогда не достигнет дна потенциальной ямы.

Аналогичный эффект давно известен в физике конденсированного состояния вещества. При быстром охлаждении расплава последний обычно затвердевает, приобретая аморфную, т. е. неупорядоченную, структуру расположения отдельных атомов. Эта неупорядоченная структура представляет собой одно из метастабильных состояний с локальным минимумом энергии, тогда как абсолютному минимуму энергии отвечает кристаллическая структура, характеризующаяся регулярным периодическим расположением атомов. Однако перевести систему из расплава в кристаллическое состояние все же можно, если воспользоваться техникой *отжига*, заключающейся в очень медленном охлаждении системы с течением времени.

Оказывается, что метод имитации отжига может быть использован и для того, чтобы перевести в состояние с абсолютным минимумом энергии построенную выше искусственную спиновую систему.

Рассмотрим следующий вероятностный алгоритм переходов [174]. Пусть каждый такой переход заключается в перевороте случайно взятого спина: $S_i \rightarrow S'_i$. Если в результате переворота полная энергия спиновой системы понизилась ($\Delta E \leq 0$), он принимается безоговорочно. Если же переворот приводит к возрастанию полной энергии ($\Delta E > 0$), то он принимается лишь с некоторой вероятностью

$$p(S_i \rightarrow S'_i) = \exp\left(-\frac{\Delta E}{\Theta}\right), \quad (3.52)$$

которая тем меньше, чем больше прирост энергии. Можно показать, что применение этого вероятностного алгоритма переходов приводит с течением времени к установлению равновесного теплового распределения Больцмана с температурой Θ , для которого вероятность обнаружить систему в состоянии с энергией E равна

$$p(E) = \frac{1}{Z} \exp\left(-\frac{E}{\Theta}\right). \quad (3.53)$$

Удобно отсчитывать энергию E от ее значения, отвечающего абсолютному минимуму, т. е. положить, что в этом минимуме $E = 0$. Допустим, что исходная конфигурация характеризуется достаточно большим значением энергии $E \gg \Theta$. Тогда на начальном этапе эволюции энергия системы в среднем со временем понижается: тепловые флуктуации (3.52) не позволяют ей застрять в метастабильных высокоэнергетических состояниях, отвечающих локальным минимумам энергии. Когда система оказывается уже в окрестности самого глубокого энергетического минимума, тепловые флуктуации, приводящие к переходам с увеличением энергии, нежелательны, поскольку они препятствуют нахождению той единственной конфигурации, которая соответствует абсолютному минимуму энергии. Действительно, как видно из (3.53), в установившемся распределении Больцмана все состояния, для которых $0 < E \lesssim \Theta$, примерно равновероятны.

Метод имитации отжига [158] заключается в том, чтобы, придерживаясь больцмановского правила переходов, постепенно понижать температуру Θ , в конце концов доводя ее до нулевого значения. Тогда в окончательном состоянии влияние тепловых флуктуаций исключается и состояние соответствует конфигурации с абсолютным минимумом энергии E , которая, согласно проведенному нами построению, дает оптимальный вариант разбиения графа.

Необходимо подчеркнуть, что, вообще говоря, метод имитации отжига гарантирует нахождение точного оптимума лишь в пределах бесконечно медленного охлаждения. Если полное число элементов достаточно велико, то в окрестности абсолютного минимума энергии может располагаться много локальных минимумов с почти совпадающими энергиями. Проведя охлаждение недостаточно медленно, после выключения тепловых флуктуаций можно оставить систему в одном из таких метастабильных состояний. Впрочем, совершенная при этом ошибка невелика, поскольку все подобные состояния имеют энергию, близкую к абсолютному минимуму, и отвечают решениям, лишь незначительно худшим, чем оптимальное.

Далеко не всегда удается построить такую физическую систему, для которой оптимизируемая функция играла бы роль энергии, как это было

проделано выше, в задаче о разбиении графов. Тем не менее, и в этих случаях можно воспользоваться методом имитации отжига. Рассмотрим, например, как получить с помощью этого метода приближенное решение задачи о коммивояжере⁷.

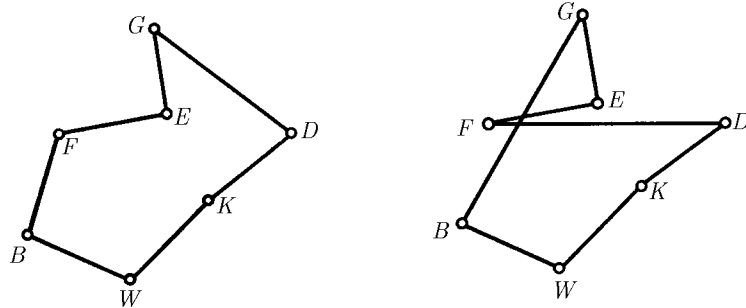


Рис. 3.16. Два маршрута обхода, отличающиеся последовательностью посещения точек F и G

Пусть изменения маршрута обхода осуществляются посредством локализованных «мутаций», затрагивающих только два произвольно взятых города. Допустим, что на данном шаге в качестве таких городов были случайно выбраны точки F и G . Тогда мутация заключается в том, что мы поменяем F и G в последовательности обхода (рис. 3.16). Это приведет к новому маршруту обхода с другой суммарной длиной

$$l' = l_{BG} + l_{GE} + l_{EF} + l_{FD} + \dots + l_{WB}, \quad (3.54)$$

отличающейся от длины исходного маршрута (3.45) на величину

$$\Delta l = l_{BG} + l_{GE} + l_{EF} + l_{FD} - l_{BF} - l_{FE} - l_{EG} - l_{GD}. \quad (3.55)$$

Условимся, что результат мутации безоговорочно принимается, если она привела к сокращению маршрута обхода ($\Delta l < 0$), и принимается с вероятностью

$$p(\Gamma \rightarrow \Gamma') = \exp\left(-\frac{\Delta l}{\Theta}\right), \quad (3.56)$$

⁷Путем введения m -компонентных спинов и последующего рассмотрения формального предела $m \rightarrow 0$ эту задачу можно привести к некоторой задаче со спинами [176]. Это позволяет изучить статистические свойства оптимальных траекторий и т. п., но не дает способа прямой аналоговой имитации.

если в результате мутации маршрут обхода удлинился ($\Delta l > 0$). Когда мутация принята, новая последовательность B, G, E, F, D, \dots, W принимается в качестве исходной, в ней опять наугад выбираются два города, и весь процесс повторяется заново. Если же от мутации отказались, в качестве исходной берется старая последовательность B, F, E, G, D, \dots, W . Имитация отжига заключается в том, чтобы понемногу уменьшать от мутации к мутации значение параметра Θ в (3.56), играющего роль температуры, и обратить его в конечном счете в нуль. Подбирая подходящую скорость уменьшения «температуры» Θ , можно добиться того, что получаемая в итоге траектория обхода будет минимальной.

Обсудим теперь сходства и различия между задачами комбинаторной оптимизации и задачами ассоциативной памяти (или распознавания образов). При решении задач комбинаторной оптимизации типичной является ситуация, когда есть много локальных минимумов, а мы разыскиваем тот единственный (абсолютный) минимум, который хотя бы чуть-чуть глубже всех остальных. При этом, чтобы наш поиск давал верные результаты, его исход не должен зависеть от начальных условий. Стремясь предотвратить остановку поиска на каких-то локальных минимумах, мы вводим эквивалент тепловых флуктуаций в алгоритм последовательных мутаций.

В задачах ассоциативной памяти, наоборот, результат в решающей степени *должен* зависеть от начальных условий. Необходимо, чтобы процесс поиска приостановился при достижении первой же устойчивой стационарной точки, а вовсе не при достижении состояния с абсолютным минимумом энергии. Следовательно, ассоциативная память и оптимизация представляют собой как бы две взаимно дополняющие проблемы.

Введение тепловых флуктуаций и имитация отжига не являются единственным подходом к аналоговому решению сложных задач комбинаторной оптимизации. Дж. Хопфилд и Д. Танк [138] указали на возможность альтернативного подхода, основанного на использовании многокомпонентных нелинейных динамических систем с детерминированной динамикой.

Идея Хопфилда и Танка заключалась в том, чтобы построить непрерывную динамическую систему, которая двигалась бы по направлению быстрой спуска к состоянию с минимумом «энергии», или, точнее, с минимумом функции Ляпунова. Это состояние задает последовательность обхода городов для оптимального маршрута. Для решения задачи с N городами Хопфилду и Танку пришлось использовать нелинейную динамическую систему с N^2 переменными V_{xi} , каждая из которых менялась в пределах от 0 до 1, так что допустимая область фазового пространства представляла собой внутренность N^2 -мерного куба. Возможные последовательности обхода городов отвечали некоторым вершинам этого многомерного куба. Динамиче-

ские уравнения были составлены таким образом, чтобы устойчивые стационарные состояния могли достигаться только на вершинах куба и, более того, лишь на тех из вершин, которые отвечали разрешенным маршрутам обхода (посещение каждого города ровно один раз). По построению функция Ляпунова достигала абсолютного минимума на тех вершинах N^2 -мерного куба, которые соответствовали кратчайшему маршруту обхода⁸.

Моделирование такой динамической системы на ЭВМ показало [138], что ее эволюция позволяет неожиданно хорошо находить оптимальные или близкие к оптимальным маршруты обхода, несмотря на то, что в системе отсутствует механизм, разрешающий выходить из состояний с локальными минимумами функции Ляпунова. Причины этого, по-видимому, следующие. Во-первых, число устойчивых притягивающих точек было невелико по сравнению с размерностью фазового пространства системы, а поэтому области притяжения разных точек, отвечающих метастабильным состояниям, не могли «экранировать» область притяжения состояния с абсолютным минимумом функции Ляпунова. Во-вторых, путем должного выбора параметров различия в глубинах локальных и абсолютного минимумов были сделаны большими, так что их области притяжения имели резко различающиеся размеры. В-третьих, использовались особые, почти симметричные начальные условия, не создававшие каких-либо специальных преимуществ для определенных маршрутов обхода.

Для задачи с 10 точками Хопфилд и Танк примерно в 50% случаев наблюдали сходимость к одному из двух самых коротких маршрутов. Поскольку общее число различных маршрутов для такой задачи составляет $10!/20 = 181\,440$, этот результат представляется довольно хорошим. В задаче с 30 точками наилучший найденный авторами [138] маршрут обхода был лишь на 19% длиннее кратчайшего.

Еще более эффективным является аналоговый метод решения задачи о коммивояжере, предложенный Р. Дурбином и Д. Уиллшоу [78]. Его суть состоит в том, что на плоскость, где разбросаны N точек, как бы набрасывается упругая петля. Эта петля стремится сжаться, уменьшая свою длину, однако ее коллапсу препятствуют силы, заставляющие петлю проходить хотя бы раз через каждую из N точек. Точная формулировка метода Дурбина и Уиллшоу такова.

На плоскости, где разбросаны N точек, положения которых заданы радиус-векторами R_i , находятся M связанных упруго в петлю частиц, положения которых мы будем задавать векторами r_j , $j = 1, 2, \dots, M$. Полная

⁸Каждому маршруту обхода соответствовало $2N$ различных вершин — по числу эквивалентных последовательностей посещения городов.

упругая энергия системы равна

$$E = -\alpha\epsilon \sum_{i=1}^N \ln \left[\sum_{j=1}^M \exp \left(-\frac{|R_i - r_j|}{2\epsilon^2} \right) \right] + \beta \sum_{j=1}^M |r_{j+1} - r_j|^2, \quad (3.57)$$

где α , β и ϵ — некоторые параметры и $r_{M+1} \equiv r_1$. В пределе, когда $\epsilon \rightarrow 0$, первое слагаемое в выражении для E не будет неограниченно возрастать, только если для каждой точки i найдется хотя бы одна такая частица j , что $|R_i - r_j| \rightarrow 0$; иными словами, в этом пределе в состоянии с минимумом энергии упругая петля должна проходить через все точки. Если, кроме того, полное число частиц M очень велико, второе слагаемое в (3.57) будет достигать минимума, когда все M частиц расположены на равном расстоянии \mathcal{L}/M друг от друга, где \mathcal{L} — длина петли. Тогда оно принимает значение \mathcal{L}^2/M и тем меньше, чем более короткой оказывается петля.

В результате упругая энергия E минимальна в пределе $\epsilon \rightarrow 0$ и $M \rightarrow \infty$, если петля представляет собой кратчайший контур, соединяющий все N точек. Очевидно, что это условие останется примерно справедливым и для конечных значений ϵ и M . Чтобы система частиц при своем движении стремилась к состоянию с минимумом энергии, достаточно задать ее динамику уравнениями

$$\dot{r}_j = -\gamma \frac{\partial E}{\partial r_j}. \quad (3.58)$$

При расчетах на ЭВМ вместо дифференциальных уравнений (3.58) приходится использовать разностную схему:

$$r_j \rightarrow r'_j, \quad r'_j = r_j - \gamma\tau \frac{\partial E}{\partial r_j}. \quad (3.59)$$

Удобно выбрать $\gamma\tau = \epsilon$. С учетом (3.57) получим

$$r'_j = r_j + \alpha \sum_{i=1}^N \kappa_{ij} (R_i - r_j) + \beta\epsilon (r_{j+1} - 2r_j + r_{j-1}). \quad (3.60)$$

Коэффициент κ_{ij} играет роль жесткости связи, соединяющей частицу j упругого контура с точкой i ; он равен

$$\kappa_{ij} = \frac{\phi(R_i - r_j)}{\sum_k \phi(R_i - r_k)}, \quad (3.61)$$

где

$$\phi(d) = \exp\left(-\frac{d^2}{2\epsilon^2}\right). \quad (3.62)$$

Таким образом, жесткость κ_{ij} нелинейно зависит от расстояний между точкой i и всеми частицами контура. Эту зависимость приходится вводить, чтобы сделать полное воздействие каждой точки на все частицы контура одинаковыми:

$$\sum_{j=1}^M \kappa_{ij} = 1. \quad (3.63)$$

При использовании метода упругой петли для нахождения оптимального маршрута обхода вначале задают петлю в виде небольшой окружности в центре квадрата и значение $\epsilon \sim 1$. После этого осуществляют последовательность итераций по формуле (3.60), которая приводит к деформации контура, и одновременно уменьшают ϵ , доводя его до значений 0,01–0,02.

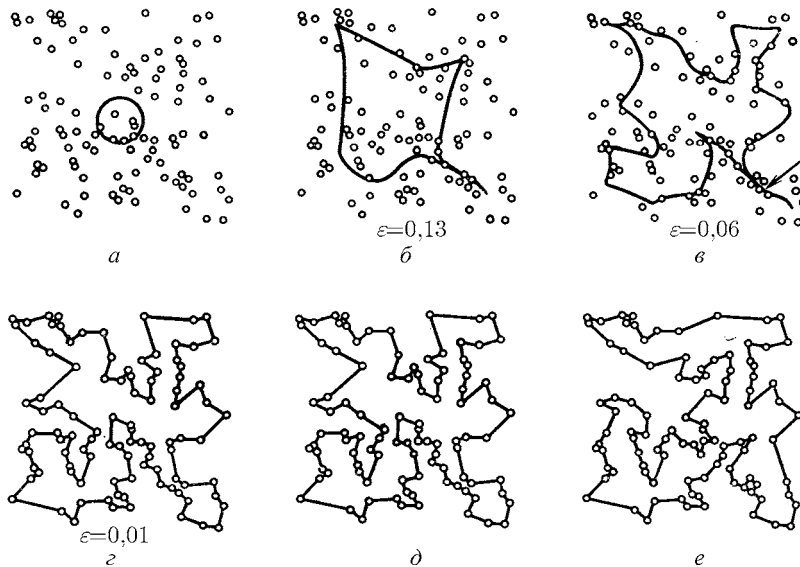


Рис. 3.17. Пример использования метода упругой петли для нахождения кратчайшего маршрута обхода

На рис. 3.17 показан пример использования метода упругой петли для нахождения кратчайшего маршрута обхода для 100 пунктов, которые случайно разбросаны внутри квадрата. Начальный контур изображен на рис. 3.17а; последующие контуры, получаемые при постепенном уменьшении параметра ϵ , изображены на рис. 3.17б–г. Окончательный маршрут (длина 7,78), найденный с помощью метода упругой петли, приведен на рис. 3.17д. Для сравнения на рис. 3.17е изображен кратчайший известный маршрут для данного расположения пунктов (его длина равна 7,70). Расчет проводился при значениях параметров $M = 2,5N$, $\alpha = 0,2$, $\beta = 2,0$.

Согласно данным [78], для задачи с 50 городами в зависимости от (случайного) выбора их расположения на карте метод упругой петли давал маршруты обхода с длинами, лишь на $3,0 \pm 1,4\%$ большими, чем у самых коротких известных маршрутов. В той же ситуации метод имитации отжига при затрате примерно того же машинного времени давал превышения на $1,5 \pm 0,7\%$, т. е. был лишь немного более эффективным. Высокая эффективность чисто детерминированного алгоритма, на котором базируется метод упругой петли, свидетельствует, что введение тепловых флуктуаций при решении сложных задач комбинаторной оптимизации не является обязательным⁹.

Заметим в заключение, что все описанные в настоящем разделе методы решения сложных задач оптимизации допускают естественную реализацию в параллельных расчетах, и поэтому они будут особенно удобными при использовании матричных процессоров или вычислительных сетей [88].

38. Обучающиеся системы

Простым примером обучающейся системы служит модель Хопфилда, в которой использован алгоритм обучения Хебба. Такая система представляет собой сеть, состоящую из N бистабильных элементов и $N(N-1)/2$ перестраиваемых связей между ними. Каждая связь характеризуется своим весовым коэффициентом взаимодействия J_{ij} . Процесс обучения заключается в следующем. К каждому из элементов i подводится значение $\xi_i^{(m)}$, которое должна принимать его спиновая переменная S_i , для m -го записываемого образа. Автомат, осуществляющий связь между элементами i и j , получает от них информацию о соответствующих значениях $\xi_i^{(m)}$ и $\xi_j^{(m)}$

⁹Необходимо, однако, указать на одно отличие. Метод имитации отжига позволяет, в принципе, находить решение с любой степенью точности, но для этого охлаждение необходимо производить все медленнее и медленнее, что требует растущих затрат машинного времени. Точность детерминированных алгоритмов фиксирована.

и изменяет хранящееся у него в памяти значение J_{ij} по правилу

$$J_{ij} \rightarrow J_{ij} + \xi_i^{(m)} \xi_j^{(m)}. \quad (3.64)$$

После этого к элементам сети подводятся новые значения $\xi_i^{(m+1)}$, отвечающие следующему записываемому образу, и цикл обучения повторяется¹⁰. Эта процедура продолжается до тех пор, пока в сети не будет записано нужное число образов.

После того как обучение закончено, сеть можно использовать для распознавания образов. В этом режиме весовые коэффициенты J_{ij} фиксированы. К каждому из элементов сети подводится значение ξ_i для распознаваемого образа, которое формирует начальное состояние $S_i(t=0) = \xi_i$ данного элемента. Затем внешнее воздействие отключается и бистабильные элементы начинают менять свои состояния по правилу переходов

$$S_i \rightarrow \text{sign} \left(\sum_j J_{ij} S_j \right). \quad (3.65)$$

По прошествии некоторого времени переходы прекращаются и сеть переходит к устойчивому состоянию, отвечающему ближайшему из хранящихся в памяти сети прообразов $\{\xi_i^{(m)}\}$ (если их не слишком много!). Считывая установившиеся значения $S_i = \xi_i^{(m)}$, можно непосредственно регистрировать этот ближайший прообраз. Все поступающие на вход образы система разделяет по их близости к определенному числу классов, задаваемых хранящимися у нее в памяти прообразами, и, следовательно, такая обучающаяся распределенная система действительно способна решать некоторые задачи классификации.

Существенное достоинство модели Хопфилда заключается в очень простом алгоритме обучения. Однако она обладает и рядом недостатков. Главный из них состоит в том, что при ее использовании мы лишены возможности управлять критерием распознавания, т. е. выбором областей притяжения для каждого из прообразов. Эксперименты показывают, что в результате обучения система обычно указывает в качестве ближайшего к распознаваемому образу $\{\xi_i\}$ на тот из прообразов $\{\xi_i^{(m)}\}$, для которого максимальна степень перекрытия, т. е. величина $\sum_i \xi_i \xi_i^{(m)}$. Часто при распознавании изображений или структур следует использовать более изощренные критерии, отделяющие существенные признаки от незначительных деталей. Задача более сложных обучающихся систем заключается в выработке существенных

¹⁰В начале обучения для всех связей следует положить $J_{ij} = 0$.

признаков, что эквивалентно *поиску закономерностей* во входных данных. Первым обучающимся устройством этого типа явился *перцептрон*, схема действия которого была предложена около тридцати лет назад [180].

Рассмотрим систему, образованную бистабильными элементами, которые подразделяются на два слоя — входных и выходных элементов (рис. 3.18). Состояния входных элементов будем характеризовать двоичными переменными $\{I_j\}$, а выходных — двоичными переменными $\{O_j\}$; переменные I_j и O_j , принимают только значения 0 и 1. Входной элемент j соединяется с выходным элементом i посредством весового коэффициента J_{ij} . При заданных состояниях входных элементов и известных коэффициентах весовых связей состояния выходных элементов определяются по правилу

$$O_i = \sigma(a_i), \quad (3.66)$$

где

$$a_i = \sum_j J_{ij} I_j,$$

а $\sigma(z)$ представляет собой ступенчатую функцию: $\sigma(z) = 0$ при $z \leq 0$ и $\sigma(z) = 1$ при $z > 0$. Следовательно, каждой картине активности входных элементов будет отвечать некоторая картина активности на выходе системы. Иными словами, такая система *реагирует* определенным образом на каждое внешнее воздействие.

Выработка реакций системы осуществляется в процессе обучения, заключающегося в определении нужных значений весовых коэффициентов J_{ij} . При обучении системе последовательно предъявляются различные входные картины $\{I_j\}$ и для каждой из них указывается желательная картина¹¹ на выходе $\{D_j\}$. Сравнение картины на выходе при заданном входном

¹¹Число желательных картин не может быть большим; цель обучения — построить классифицирующую систему, а поэтому многим разным входным картинам должны соответствовать идентичные реакции.

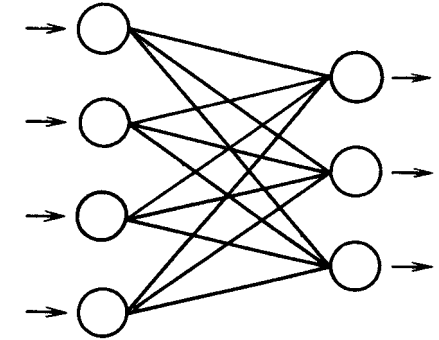


Рис. 3.18. Схема перцептрона. Показаны слой входных и слой выходных элементов, а также связи между ними

изображении с желаемым выходным изображением позволяет сформулировать сигналы ошибки $\{E_j\}$:

$$E_j = D_j - O_j.$$

Зная сигналы ошибки, можно попытаться исправить значения весовых коэффициентов J_{ij} , чтобы уменьшить ошибку. Для этого используется алгоритм

$$\Delta J_{ij} = \epsilon E_i I_j, \quad (3.67)$$

где ϵ — малый параметр, управляющий скоростью обучения.

Для двоичных переменных $\{I_j\}$, $\{O_j\}$ и $\{D_j\}$ алгоритм (3.67) сводится к изменению весовых коэффициентов связей для данного выходного элемента i , только если его состояние O_i ошибочно (т. е. не совпадает с желаемым D_i), и к изменению весовых коэффициентов связей, приходящих только от активных входных элементов j (для которых $I_j = 1$).

Обучение заключается в том, что системе многократно предъявляются разные входные изображения с известными желаемыми картинками на выходе и на каждом шаге, после предъявления очередного изображения, производится подстройка весов (3.67).

Вместо детерминированного правила (3.66) иногда используют вероятностный алгоритм — считают, что при заданной величине a_i выходной элемент i находится в активном состоянии ($O_i = 1$) с вероятностью

$$p_i = \frac{1}{1 + \exp(-a_i/\Theta)}, \quad (3.68)$$

где Θ — аналог температуры. При $\Theta = 0$ вероятностный алгоритм (3.68) совпадает с детерминированным (3.66).

Описанная выше схема перцептрона с вероятностным поведением ($\Theta \neq 0$) была использована в одной лингвистической задаче [204]. На вход системы подавались (закодированные в двоичные последовательности) английские слова, отвечающие различным глаголам в настоящем времени. «Реакция» системы должна была состоять в том, чтобы указать форму этих глаголов в прошедшем времени. Как известно, в английском языке, наряду с регулярными правилами построения прошедшей формы глаголов, имеется большое число различных исключений, не подчиняющихся четким правилам, но обладающих все же определенными закономерностями. Хотя системе не сообщалось никаких правил построения, она самостоятельно научилась на опыте достаточно верно угадывать прошедшие формы глаголов.

Тем не менее, возможности обучения, свойственные перцептронам, остаются очень ограниченными. Например, такая машина не может воспроизвести функцию «исключенное ИЛИ». Рассмотрим этот пример подробнее.

Функция «исключенное ИЛИ» заключается в том, что только при подаче сигнала на один из двух входных элементов (A и B) выходной элемент (C) должен переходить в активное состояние:

$$D_C = \begin{cases} 0, & \text{если } I_A = I_B = 0, \\ 1, & \text{если } I_A = 1, I_B = 0, \\ 1, & \text{если } I_A = 0, I_B = 1, \\ 0, & \text{если } I_A = I_B = 1. \end{cases} \quad (3.69)$$

Однако по правилу (3.66) состояние выходного элемента есть

$$O_C = \sigma(J_A I_A + J_B I_B). \quad (3.70)$$

В частности,

$$O_C = \begin{cases} 0, & \text{если } I_A = I_B = 0, \\ \sigma(J_A), & \text{если } I_A = 1, I_B = 0, \\ \sigma(J_B), & \text{если } I_A = 0, I_B = 1, \\ \sigma(J_A + J_B), & \text{если } I_A = 1, I_B = 1. \end{cases} \quad (3.71)$$

Чтобы $\sigma(J_A) = \sigma(J_B) = 1$, необходимо выполнение условий $J_A > 0$ и $J_B > 0$. Но тогда $\sigma(J_A + J_B) = 1$ и при подаче сигнала на оба входных элемента выходной элемент оказывается в активном состоянии, что противоречит условию.

Преодолеть серьезные ограничения, свойственные перцептронам, можно, переходя к использованию более сложных систем, где между входом и выходом расположено какое-то число слоев *скрытых* элементов. Алгоритм обучения для такого обобщенного перцептрона, называемый *методом обратного распространения ошибок*, был предложен Д. Румелхартом, Дж. Хинтоном и Р. Уильямсом [203].

Пусть обучающаяся система состоит из N слоев (рис. 3.19). Элементы первого слоя являются входными — на них подается подлежащее анализу изображение $\{I_j\}$. Последний слой элементов — выходной, с него снимается изображение $\{O_j\}$. Между входным и выходным расположены $N - 2$ слоя скрытых элементов. В системе имеются связи только между элементами,

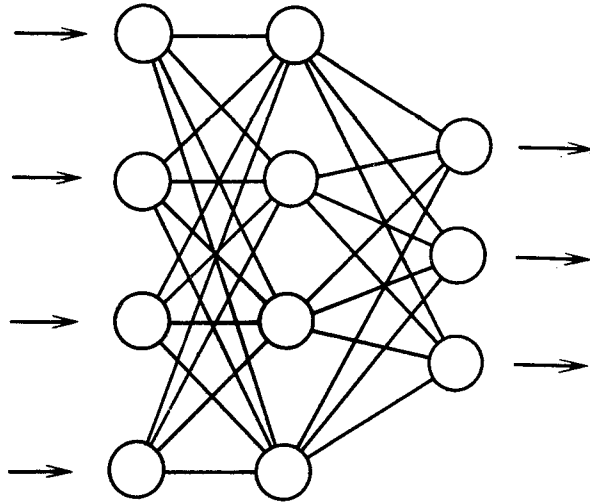


Рис. 3.19. Схема обобщенного перцептрона, используемого в методе обратного распространения ошибок

принадлежащими соседним слоям; число элементов в каждом слое может быть различным.

При поступлении некоторого сигнала $x_i^{(n)}$ на вход i -го элемента в слое с номером n этот элемент вырабатывает выходной сигнал $y_i^{(n)}$ по закону¹²

$$y_i^{(n)} = \frac{1}{1 + \exp(-x_i^{(n)})}. \quad (3.72)$$

В свою очередь, входной сигнал, поступающий на элемент j в следующем, $(n + 1)$ -м слое, формируется как взвешенная сумма выходных сигналов от элементов предыдущего n -го слоя:

$$x_j^{(n+1)} = \sum_i J_{ji}^{(n)} y_i^{(n)}. \quad (3.73)$$

¹²Конкретный вид зависимости $y = f(x)$ не очень существен. Необходимо лишь, чтобы при изменении сигнала x от $-\infty$ до $+\infty$ значение сигнала y монотонно возрастало от 0 до 1. Тем не менее, использование функции (3.72) несколько упрощает расчеты.

Весовые коэффициенты связей $J_{jn}^{(n)}$ могут принимать положительные и отрицательные значения.

Согласно сказанному выше, входным сигналом для элементов первого слоя является предъявляемое изображение, т. е. $x_j^{(1)} = I_j$. Выходной сигнал элементов последнего слоя дает результат обработки, т. е. $y_j^{(N)} = O_j$.

Если весовые коэффициенты связей случайны, случайной будет и реакция такой системы — формируемое на ее выходе изображение. Задача обучения состоит в том, чтобы путем подбора весовых коэффициентов связей добиться требуемой реакции системы на предъявляемые изображения.

Конкретная формулировка такова. Допустим, что в нашем распоряжении имеются M различных эталонных изображений $\{I_{j,m}\}$, для каждого из которых известна желаемая реакция — выходное изображение $\{D_{j,m}\}$. Необходимо так подобрать значения весовых коэффициентов $\{J_{ji}^{(n)}\}$ и, вообще говоря, число промежуточных слоев, чтобы при подаче на вход системы любого из эталонных изображений на ее выходе появлялась соответствующая заданная нами картина.

Определим ошибку E как

$$E = \frac{1}{2} \sum_{m=1}^M \sum_j (D_{j,m} - O_{j,m})^2, \quad (3.74)$$

где $D_{j,m} = y_j^{(N)}$ представляют собой выходные сигналы элементов последнего слоя при подаче на первый слой m -го эталонного изображения $\{I_{j,m}\}$. Очевидно, что эта функция достигает своего минимального значения, лишь если для всех M эталонных изображений результат обработки совпадает с желаемым. Цель обучения сводится к нахождению значений коэффициентов $J_{ji}^{(n)}$, минимизирующих функцию (3.74). Мы видим, что обучение оказывается сложной задачей оптимизации.

Чтобы минимизировать E по способу быстрейшего спуска, необходимо вычислить частные производные от E по отношению к весовому коэффициенту каждой связи. Любая такая производная будет даваться суммой величин, относящихся к различным эталонным изображениям. По отдельности для каждого эталонного изображения все частные производные от функции ошибки можно вычислить, воспользовавшись методом обратного распространения ошибок [203, 207, 225].

Заметим прежде всего, что справедливо тождество

$$\frac{\partial E}{\partial y_i^{(n-1)}} = \sum_j \frac{\partial E}{\partial y_j^{(n)}} \frac{\partial y_j^{(n)}}{\partial x_j^{(n)}} \frac{\partial x_j^{(n)}}{\partial y_i^{(n-1)}}, \quad (3.75)$$

в котором значения входных и выходных сигналов взяты для одного из эталонных входных изображений. Поскольку последующий расчет ведется при фиксированном эталонном изображении, мы временно опускаем индекс m (например, $y_{j,m}^{(n)}$), указывающий на номер этого изображения.

С учетом зависимостей (3.72) и (3.73) из (3.75) следует рекуррентное соотношение

$$\frac{\partial E}{\partial y_i^{(n-1)}} = \sum_j \frac{\partial E}{\partial y_j^{(n)}} y_j^{(n)} (1 - y_j^{(n)}) J_{ji}^{(n-1)}. \quad (3.76)$$

С его помощью можно вычислить производные $\partial E/\partial y$ для некоторого слоя, если нам известны значения таких производных для следующего слоя. Заметим, что для выходного слоя эти производные легко найти, непосредственно продифференцировав равенство (3.74):

$$\frac{\partial E}{\partial y_i^{(N)}} = y_i^{(N)} - D_i. \quad (3.77)$$

Итак, начиная с нижнего N -го слоя и двигаясь вверх, можно последовательно вычислять по формуле (3.76) частные производные $\partial E/\partial y$ для каждого из элементов.

Далее можно воспользоваться равенством

$$\frac{\partial E}{\partial J_{ji}^{(n-1)}} = \frac{\partial E}{\partial y_j^{(n)}} \frac{dy_j^{(n)}}{dx_j^{(n)}} \frac{\partial x_j^{(n)}}{\partial J_{ji}^{(n-1)}}, \quad (3.78)$$

которое с учетом (3.72) и (3.73) переписывается в виде

$$\frac{\partial E}{\partial J_{ji}^{(n-1)}} = \frac{\partial E}{\partial y_j^{(n)}} y_j^{(n)} (1 - y_j^{(n)}) y_i^{(n-1)}. \quad (3.79)$$

Если предварительно были вычислены производные $\partial E/\partial y$, то по формуле (3.79) мы находим интересующие нас значения производных по весовым коэффициентам $\partial E/\partial J_{ji}$ для каждой из связей при заданном входном эталонном изображении. Найденными значениями производных $\partial E/\partial J_{ji}$ можно воспользоваться для внесения поправок в весовые коэффициенты связей $J_{ji}^{(n)}$ после предъявления каждого из эталонных изображений. Альтернативный подход заключается в том, чтобы для каждого нового цикла

обучения предъявить системе все M эталонных изображений, просуммировать значения $\partial E/\partial J_{ji}$, относящиеся к различным эталонным изображениям, и лишь после этого внести необходимые поправки в значения J_{ji} . Так или иначе, каждый цикл обучения заключается в предъявлении эталонов, вычислении производных от функции ошибки E по весовым коэффициентам всех связей и небольшом изменении коэффициента каждой связи на величину

$$\Delta J_{ji}^{(n)} = -\epsilon \frac{\partial E}{\partial J_{ji}^{(n)}}, \quad (3.80)$$

после чего цикл обучения повторяется с новыми значениями весовых коэффициентов.

Процесс обучения следует продолжать, пока не будет достигнута приемлемая достаточно малая ошибка (3.74). Если в процессе обучения существенно понизить ошибку не удастся, его следует повторить заново, введя дополнительные слои скрытых элементов. Вместо правила (3.80) при изменении весовых коэффициентов лучше использовать формулу

$$\Delta J_{ji}^{(n)}(t) = -\epsilon \frac{\partial E}{\partial J_{ji}^{(n)}}(t) + \alpha \Delta J_{ji}^{(n)}(t-1), \quad (3.81)$$

где время t изменяется на 1 после окончания каждого цикла обучения, а коэффициент α , заключенный в интервале от 0 до 1, характеризует инерционность обучающейся системы. В начальный момент времени весовым коэффициентам следует приписать малые случайные значения.

В некоторых случаях удобно воспользоваться системой, в которой элементы имеют постоянные смещения, так что вместо соотношения (3.73) справедливо соотношение

$$x_j^{(n+1)} = \sum_j J_{ji}^{(n)} y_i^{(n)} + b_j^{(n)}. \quad (3.82)$$

Смещения $b_j^{(n)}$ для различных элементов могут подбираться в процессе обучения. Заметим, что введение смещений не меняет сколь-либо существенно приведенные выше рассуждения. Действительно, наличие постоянного смещения эквивалентно тому, что на вход элемента поступает дополнительно единичный входной сигнал $Y^{(n)} = 1$ через связь с весом $b_j^{(n)}$. Поэтому смещения можно рассматривать совершенно аналогично весовым коэффициентам связей $J_{ji}^{(n)}$.

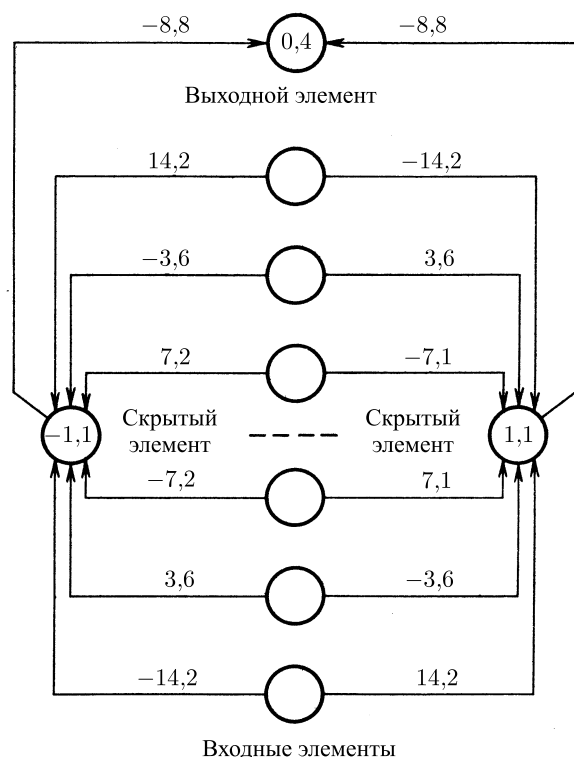


Рис. 3.20. Схема сети, обучившейся распознавать зеркальную симметрию. Числа, стоящие над соединительными линиями, представляют собой весовые коэффициенты этих связей, а числа внутри кружков, обозначающих элементы, указывают постоянное смещение для соответствующего элемента

Простой задачей, которую нельзя решить, соединяя входные элементы непосредственно с выходными (т. е. с помощью обычного перцептона), является обнаружение симметрии. Чтобы определить, обладает ли зеркальной симметрией одномерная картина активности, задаваемая состояниями цепочки входных элементов, необходимо использовать слой промежуточных элементов. В численном эксперименте путем обучения удалось найти красивое решение этой задачи [210], использующее всего два промежуточных элемента (рис. 3.20). Для обучения были выбраны 64 эталонные картины. Каждый цикл обучения заключался в том, что для всех таких

картин по методу обратного распространения ошибок вычислялись значения $\Delta E/\partial J$ и $\partial E/\partial b$ для различных элементов, затем они суммировались по всем эталонным картинкам и согласно правилу (3.81) проводилось изменение весовых коэффициентов и смещений, после чего цикл обучения повторялся заново. Начальные весовые коэффициенты были случайно распределены в интервале от $-0,3$ до $0,3$. Параметры ϵ и α в (3.81) имели значения $\epsilon = 0,1$, $\alpha = 0,9$. Всего потребовалось провести 1425 циклов обучения, после чего система приобрела способность распознавать зеркальную симметрию в *любой* одномерной картине активности (это можно проверить, воспользовавшись значениями весовых коэффициентов и смещений, которые приведены на рис. 3.20).

К числу более сложных задач, решение которых получено путем обучения в численном эксперименте [203], относится запоминание родственных связей у группы людей («у Кристофера и Пенелопы есть сын Артур и дочь Виктория, а у Эндрю и Кристины — сын Джеймс и дочь Дженифер; Виктория и Джеймс женаты и имеют детей Колина и Шарлотту»). После обучения (разумеется, с предварительным кодированием) эта сеть родственных связей, запомненная машиной, может быть использована, чтобы получить, например, ответ на вопрос: «Кто является тетей Колина?». Сообщается также [208, 225], что с использованием метода обратного распространения ошибок машину удалось обучить читать вслух английский текст.

Подчеркнем, что, обучившись распознавать некоторую закономерность, такая машина «не знает», как она это делает, и «не умеет объяснить», в чем эта закономерность состоит. В отличие от традиционного подхода, основанного на использовании программируемых ЭВМ, при работе с обучающимися машинами нет необходимости, чтобы природа имеющихся закономерностей и их конкретный вид были заранее известны человеку.

Метод обратного распространения ошибок использует детерминированный алгоритм оптимизации при поиске требуемых весовых коэффициентов связей. Ввиду этого имеется опасность, что при движении строго вниз по градиенту функции ошибки E будет достигнут какой-то из локальных минимумов этой функции и процесс обучения приостановится, не найдя истинного решения. Если ошибка остается значительной, часто достаточно добавить еще несколько элементов. В результате размерность пространства весовых коэффициентов возрастает и возникают пути обхода барьеров, отделяющих плохие локальные минимумы в подпространстве более низкой размерности.

Существуют, однако, и модели обучения, использующие вероятностные алгоритмы. Наиболее известной моделью является так называемая

машина Больцмана, предложенная в 1984 г. Дж. Хинтоном, Т. Сейновским и Д. Экли [129].

В отличие от рассмотренного выше обобщенного перцептрона машина Больцмана состоит из бистабильных элементов, состояния которых описываются двоичными переменными s_i , со значениями 0 и 1. Элементы не группируются в слои — каждый из них может быть связан со всеми остальными, причем связи являются симметричными. Для каждой картины активности сети определена «энергия» \mathcal{E} , даваемая выражением

$$\mathcal{E} = -\frac{1}{2} \sum_{i,j} J_{ij} s_i s_j + \sum_i b_i s_i, \quad (3.83)$$

где J_{ij} — весовые коэффициенты связей ($J_{ii} = 0$), а b_i — смещения.

Элементы сети делятся на *видимые* (входные и выходные) и *скрытые*. К входным элементам подводится анализируемая картина, а с выходных снимается реакция сети. Как и для перцептрона, обучение заключается в постепенном подборе весовых коэффициентов, с тем чтобы минимизировать рассогласование между желаемым и имеющимся результатами обработки входных картин.

Подача определенной картины заключается в том, что мы фиксируем в нужных состояниях входные элементы, позволяя остальным элементам сети совершать переходы между состояниями согласно описанному ниже алгоритму. Заметим, что постоянные смещения b_i в (3.83) можно считать результатом связи элемента i с дополнительным входным элементом, который все время находится во включенном (активном) состоянии; тогда b_i есть весовой коэффициент связи, соединяющей элемент i с этим дополнительным элементом. Следовательно, достаточно рассмотреть более простую сеть, не имеющую смещений, для которой

$$\mathcal{E} = -\frac{1}{2} \sum_{i,j} J_{ij} s_i s_j. \quad (3.84)$$

Если элемент i переходит из состояния покоя в активное состояние, это сопровождается (при заданных состояниях остальных элементов) уменьшением энергии на величину

$$\Delta \mathcal{E}_i = \sum_j J_{ij} s_j. \quad (3.85)$$

Пусть вероятностный алгоритм переходов в системе таков, что при заданных состояниях остальных элементов на очередном шаге по времени эле-

мент i принимает активное состояние ($s_i = 1$) с вероятностью

$$p_i = \frac{1}{1 + \exp(-\Delta \mathcal{E}_i / \Theta)}, \quad (3.86)$$

где параметр Θ играет роль температуры.

В результате таких переходов с течением времени в системе устанавливается «тепловое равновесие», при котором вероятность обнаружить различные картины активности сети $\{s_i\}$ дается распределением Больцмана:

$$P(\{s_j\}) = \frac{1}{Z} \exp \left[-\frac{\mathcal{E}(\{s_j\})}{\Theta} \right], \quad (3.87)$$

где Z — некоторый нормировочный коэффициент.

Перейдем теперь к описанию процедуры, посредством которой производится обучение машины Больцмана. Прежде всего сделаем одно общее замечание. Цель обучения будем понимать как построение машиной Больцмана своего рода «внутренней модели», воспроизводящей с достаточно высокой степенью точности закономерности связей между различными структурами в окружающем машину мире.

Предположим, что в окружающем мире встречаются структуры двух классов A и B ; структуры, относящиеся к классу A , будем нумеровать индексом $\alpha = 1, 2, \dots, K$, а структуры, относящиеся к классу B , — индексом $\beta = 1, 2, \dots, M$. Допустим, что связь между структурами двух классов носит статистический характер и определяется совместными вероятностями $p(\beta, \alpha)$ наблюдения различных структур β и α . Удобно ввести также вероятности

$$\pi(\beta | \alpha) = \frac{p(\beta, \alpha)}{p(\alpha)} \quad (3.88)$$

наблюдения структуры β , если в среде присутствует структура α (здесь $p(\alpha)$ — вероятность наблюдения структуры α).

Пусть цель обучения машины Больцмана состоит в построении «внутренней модели», воспроизводящей вероятностную взаимосвязь между структурами двух классов. Это означает, что различные структуры α и β мы должны каким-то образом кодировать в соответствующих картинах активности входных $I^{(\alpha)} = (I_1^{(\alpha)}, \dots, I_n^{(\alpha)})$ и выходных $O_1^{(\beta)} = (O_1^{(\beta)}, \dots, O_{n'}^{(\beta)})$ групп элементов. Разным «внутренним моделям» отвечают свои весовые коэффициенты связей J_{ij} между элементами сети. Если все эти коэффициенты заданы, то при фиксации входных элементов в состоянии с некоторой

картиной активности $\mathbf{I}^{(\alpha)}$ сетью будут генерироваться с какими-то вероятностями разные картины активности \mathbf{O} выходных элементов. Наблюдая за поведением системы, можно найти условные вероятности $\tilde{\pi}(\beta | \alpha)$ того, что при подаче на вход картины α на выходе будет наблюдаться картина β . Процесс обучения сводится к изменению весовых коэффициентов связей J_{ij} , приближающему условные вероятности $\tilde{\pi}(\beta | \alpha)$, которые порождаются «внутренней моделью», к условным вероятностям $\pi(\beta | \alpha)$, характерным для окружающего мира.

Критерием близости двух распределений служит информационно-теоретическая мера

$$G = \sum_{\alpha, \beta} p(\beta, \alpha) \ln \frac{\pi(\beta | \alpha)}{\tilde{\pi}(\beta | \alpha)}. \quad (3.89)$$

Если для всех пар (β, α) вероятности π и $\tilde{\pi}$ совпадают, G достигает своего абсолютного минимума, равного нулю.

Вычислим производные $\partial G / \partial J_{ij}$. Для этого прежде всего заметим, что от весовых коэффициентов связей в выражении (3.89) зависят лишь величины $\tilde{\pi}$, а следовательно,

$$\frac{\partial G}{\partial J_{ij}} = - \sum_{\alpha, \beta} p(\beta, \alpha) \frac{\partial}{\partial J_{ij}} \ln \tilde{\pi}(\beta | \alpha). \quad (3.90)$$

Найдем логарифмическую производную от $\tilde{\pi}$ по J_{ij} для произвольно выбранных картин активности входных \mathbf{I} и выходных \mathbf{O} групп элементов. Удобно обозначить в качестве $\mathbf{S} = \{s_j\}$ картину активности скрытых элементов. Тогда при подаче на вход заданной постоянной картины \mathbf{I} вероятность обнаружить скрытые и выходные элементы в состояниях \mathbf{S} и \mathbf{O} дается выражением

$$P_{\mathbf{I}}(\mathbf{S}, \mathbf{O}) = \frac{1}{Z_{\mathbf{I}}} \exp \left[- \frac{\mathcal{E}_{\mathbf{I}}(\mathbf{S}, \mathbf{O})}{\Theta} \right], \quad (3.91)$$

где статистическая сумма $Z_{\mathbf{I}}$ определена как

$$Z_{\mathbf{I}} = \sum_{\mathbf{S}, \mathbf{O}} \exp \left[- \frac{\mathcal{E}_{\mathbf{I}}(\mathbf{S}, \mathbf{O})}{\Theta} \right]. \quad (3.92)$$

Поскольку условная вероятность $\tilde{\pi}(\mathbf{O} | \mathbf{I})$ связана с $P_{\mathbf{I}}(\mathbf{S}, \mathbf{O})$ соотношением

$$\tilde{\pi}(\mathbf{O} | \mathbf{I}) = \sum_{\mathbf{S}} P_{\mathbf{I}}(\mathbf{S}, \mathbf{O}), \quad (3.93)$$

имеем

$$\tilde{\pi}(\mathbf{O} | \mathbf{I}) = \frac{1}{Z_{\mathbf{I}}} \sum_{\mathbf{S}} \exp \left[- \frac{\mathcal{E}_{\mathbf{I}}(\mathbf{S}, \mathbf{O})}{\Theta} \right]. \quad (3.94)$$

Введя статистическую сумму

$$Z_{\mathbf{I}, \mathbf{O}} = \sum_{\mathbf{S}} \exp \left[- \frac{\mathcal{E}_{\mathbf{I}}(\mathbf{S}, \mathbf{O})}{\Theta} \right], \quad (3.95)$$

которая вычисляется при заданных состояниях входных и выходных элементов, перепишем равенство (3.94) в виде

$$\tilde{\pi}(\mathbf{O} | \mathbf{I}) = \frac{Z_{\mathbf{I}, \mathbf{O}}}{Z_{\mathbf{I}}}. \quad (3.96)$$

Следовательно,

$$\frac{\partial}{\partial J_{ij}} \ln \tilde{\pi}(\mathbf{O} | \mathbf{I}) = \frac{1}{Z_{\mathbf{I}, \mathbf{O}}} \frac{\partial Z_{\mathbf{I}, \mathbf{O}}}{\partial J_{ij}} - \frac{1}{Z_{\mathbf{I}}} \frac{\partial Z_{\mathbf{I}}}{\partial J_{ij}}. \quad (3.97)$$

Учитывая определения $Z_{\mathbf{I}}$ и $Z_{\mathbf{I}, \mathbf{O}}$, а также зависимость энергии от весовых коэффициентов связей (см. (3.85)), получаем из (3.97)

$$\frac{\partial}{\partial J_{ij}} \ln \tilde{\pi}(\mathbf{O} | \mathbf{I}) = \frac{\langle s_i s_j \rangle_{\mathbf{I}, \mathbf{O}} - \langle s_i s_j \rangle_{\mathbf{I}}}{\Theta}, \quad (3.98)$$

где $\langle s_i s_j \rangle_{\mathbf{I}}$ и $\langle s_i s_j \rangle_{\mathbf{I}, \mathbf{O}}$ представляют собой средние по времени от произведения активностей двух элементов i и j , взятые при фиксированном состоянии только входных (\mathbf{I}) или же как входных, так и выходных (\mathbf{O}) элементов, соответственно.

Подстановка (3.98) в (3.90) дает

$$\frac{\partial G}{\partial J_{ij}} = \frac{1}{\Theta} \sum_{\alpha, \beta} p(\beta, \alpha) (\langle s_i s_j \rangle_{\alpha} - \langle s_i s_j \rangle_{\alpha, \beta}). \quad (3.99)$$

После того как производные $\frac{\partial G}{\partial J_{ij}}$ известны, их можно использовать для модификации весовых коэффициентов связей J_{ij} , чтобы уменьшить величину G . Это можно сделать, например, по формуле¹³

$$\Delta J_{ij} = -\epsilon \frac{\partial G}{\partial J_{ij}}. \quad (3.100)$$

¹³ Можно использовать и другие варианты метода быстрого спуска. Например, $\Delta J_{ij} = -\epsilon \text{sign} [\partial G / \partial J_{ij}]$.

В важном частном случае, когда различные картины α равновероятны и каждой из них в результате обучения сеть должна сопоставлять какую-то однозначно определенную картину $\beta(\alpha)$, выражение (3.99) приобретает вид

$$\frac{\partial G}{\partial J_{ij}} = \frac{1}{K\Theta} \sum_{\alpha} (\langle s_i s_j \rangle_{\alpha} - \langle s_i s_j \rangle_{\alpha, \beta(\alpha)}), \quad (3.101)$$

где K — полное число различных входных картин. Тогда процедура обучения машины Больцмана выглядит особенно просто.

Обучение заключается в последовательном повторении тренировочных циклов. Каждый цикл состоит из следующих стадий.

1. Подать на вход машины одну из картин α и зафиксировать выходные элементы в состояниях, которые они должны принимать в соответствующей требуемой выходной картине активности $\beta(\alpha)$. Путем усреднения по большим интервалам времени вычислить средние значения $\langle s_i s_j \rangle_{\alpha, \beta(\alpha)}$ для всех пар скрытых элементов i и j .

2. Повторить стадию **1** с тем, однако, отличаем, что теперь не нужно фиксировать состояния выходных элементов; это позволяет найти средние значения $\langle s_i s_j \rangle_{\alpha}$.

3. Выполнить процедуры, описанные в **1** и **2** для всех имеющихся входных картин α , и вычислить для каждой связи производную $\partial G / \partial J_{ij}$ по формуле (3.101).

4. Зная производные, найти, согласно (3.100) или другим аналогичным формулам, новые модифицированные значения весовых коэффициентов J_{ij} .

Последовательные циклы обучения следует повторять, пока не будет достигнута приемлемая частота правильного распознавания¹⁴. Если снижения числа ошибок не происходит, обучение можно произвести заново, увеличив число скрытых элементов.

Машины Больцмана обладают хорошими способностями к обучению. Так, в [200] описано использование такой машины для распознавания гласных звуков английского языка, произносимых разными людьми¹⁵. Машина научилась также делить предложения на отдельные слова и даже «понимать» простые предложения.

Недостаток подобных устройств заключается в том, что процесс их обучения занимает довольно длительное время, поскольку необходимо мно-

¹⁴При практической реализации этого алгоритма обучения на каждой из стадий **1** и **2**, описанных выше, применяют процедуру имитации отжига [200]. Начальная температура Θ выбирается довольно большой, а затем ее медленно понижают. Средние значения $\langle s_i s_j \rangle$ вычисляют для конечных состояний с низкой температурой.

¹⁵Звуковые сигналы предварительно обрабатывались для получения спектров, а спектры кодировались в форме двоичных входных последовательностей.

гократное накопление статистики. Это относится, конечно, главным образом к случаям, когда машина Больцмана имитируется на ЭВМ с последовательным выполнением операций. При реализации с помощью сетей, производящих параллельные вычисления, машина Больцмана должна быть более эффективной.

39. Эволюционные модели

В предыдущих параграфах были приведены примеры динамических систем, которые могут использоваться для решения задач аналоговой обработки информации. Все они относились к числу систем потенциального типа — с течением времени эти системы стремились прийти в состояние, отвечающее минимуму некоторого потенциала или «энергии». Флуктуации в рассмотренных нами системах описывались больцмановским распределением вероятности с некоторой эффективной температурой. Все эти свойства типичны для систем, близких к состоянию теплового равновесия.

Как уже отмечалось, вся живая природа основана на системах, находящихся вдали от состояния теплового равновесия. Проходящий через живые системы поток энергии делает их активными, т. е. способными к самоорганизации. Главную роль в развитии биологических систем играют процессы эволюционного типа. У них также имеются аналоги среди физических систем. Наиболее близкий пример — процесс конкуренции мод в лазере непрерывного действия. В упрощенном виде его можно описать следующим образом.

Лазер представляет собой нелинейную сильнонеравновесную распределенную систему с внешней энергетической накачкой. Когда мощность накачки превышает определенный порог, в системе начинается экспоненциальный рост различных мод. С ростом интенсивности различных мод между ними устанавливается конкуренция. Действительно, на поддержание каждой такой моды расходуется определенная часть потока энергии, поступающего от источника накачки. Чем больше интенсивность данной моды, тем больше расход энергии на ее поддержание. Поскольку полная мощность источника накачки ограничена, обычно в результате конкуренции выживает всего одна наиболее эффективная мода.

Простейшая модель, описывающая процесс конкуренции мод, такова:

$$\dot{n}_j = \left(s_j - \beta \sum_{i=1}^N n_i \right) n_j. \quad (3.102)$$

Нетрудно проверить, что с течением времени в такой системе выживает лишь одна из мод j , а именно та, для которой эффективность s_j максимальна. Ее установившаяся интенсивность равна

$$\bar{n} = \max_j \frac{s_j}{\beta}. \quad (3.103)$$

Близким поведением обладает также несколько иная модель, а именно

$$\dot{n}_j = (s_j - \langle s \rangle) n_j, \quad (3.104)$$

где

$$\langle s \rangle \equiv \sum_{i=1}^N s_i n_i \left(\sum_{i=1}^N n_i \right)^{-1}. \quad (3.105)$$

Как и система (3.102), она имеет единственное устойчивое стационарное состояние; в этом состоянии отлична от нуля лишь интенсивность той моды, для которой коэффициент s_i максимален. Главное отличие системы (3.104) заключается в том, что она сохраняет суммарную интенсивность всех мод. Действительно, как следует из (3.104) и (3.105),

$$\frac{d}{dt} \sum_i n_i = \sum_i s_i n_i - \langle s \rangle \sum_i n_i = 0. \quad (3.106)$$

Таким образом, в модели (3.104) рост одной из мод может осуществляться лишь за счет уменьшения интенсивности каких-то других мод.

Помимо конкуренции за общий источник питания, различные моды могут также непосредственно влиять на скорости роста друг друга, оказывая перекрестное усиливающее или подавляющее воздействие. Учет этого эффекта можно произвести¹⁶, вводя дополнительные члены в уравнение (3.102):

$$\dot{n}_j = \left(s_j - \beta \sum_i n_i \right) n_j + \sum_{i \neq j} \alpha_{ij} n_i n_j. \quad (3.107)$$

Коэффициент α_{ij} описывает влияние i -й моды на скорость роста j -й моды: если $\alpha_{ij} > 0$, то имеет место усиление, а если $\alpha_{ij} < 0$, то происходит подавление.

Все приведенные выше модели используются для описания биологических сообществ. Каждой моде j при этом соответствует один из биологических видов, а интенсивность n_j представляет собой его численность.

¹⁶Аналогичные слагаемые можно ввести также и в правые части уравнений (3.104).

Первое слагаемое справа в (3.107) отвечает конкуренции различных видов за общий источник питания («возобновляемый ресурс»), а второе слагаемое характеризует межвидовое взаимодействие (например, оно возникает в системах типа «хищник – жертва», «паразит – хозяин» и т. п.).

Заметим, что при наличии межвидового взаимодействия (т. е. когда $\alpha_{ij} \neq 0$) поведение системы (3.107) может быть очень сложным. Помимо устойчивых стационарных точек, она может обладать различными предельными циклами и даже странными аттракторами (заметим, что формально к тому же типу относится и модель Лоренца, см. раздел 20.3 главы 2). Это означает, что в установившемся режиме биологическое сообщество способно обладать весьма сложной структурой соподчинения различных видов живых организмов, демонстрировать регулярные или хаотические изменения со временем своей внутренней структуры. Разумеется, приведенные нами модели являются лишь простейшими: взаимодействия членов сообщества с окружающей средой и между собой могут описываться и более сложными уравнениями [206].

Мутации приводят к рождению новых видов и, вообще говоря, к межвидовым превращениям. Последний эффект (т. е. тот факт, что потомство одного вида из-за мутации теоретически могло бы оказаться принадлежащим другому виду) нетрудно учесть, добавив соответствующие члены в уравнения естественного отбора:

$$\dot{n}_j = \left(s_j - \beta \sum_i n_i \right) n_j + \sum_{i \neq j} \alpha_{ij} n_i n_j + \sum_i D_{ij} n_i - \sum_i D_{ji} n_i. \quad (3.108)$$

Здесь D_{ij} — частота мутаций, превращающих i в j . Гораздо интереснее и важнее первый эффект — рождение в результате мутаций новых видов.

Рождение нового вида $j = N + 1$ означает, что в динамических уравнениях (3.108) появятся дополнительные слагаемые, описывающие его рост, взаимодействие с источником питания и с другими видами. Иными словами, результатом такого случайного события оказывается изменение не значений параметров, а изменение самой структуры динамических уравнений эволюционной модели. С математической точки зрения, это чрезвычайно существенное обстоятельство. Последовательный учет мутаций требует от нас перехода к изучению нового математического объекта — динамических уравнений, сама *форма* которых случайно меняется со временем.

Заметим, что виды не только рождаются, но и погибают. Формально, в рамках уравнений (3.108) численность n_j вымирающего вида может лишь стремиться со временем к нулю, оставаясь всегда положительной. Ясно, однако, что при очень малых численностях n_j эти уравнения неприменимы

и необходимо принимать во внимание дискретный состав популяции. Фактически, если в популяции осталось порядка сотни особей, она обычно обречена на быстрое вымирание за конечный промежуток времени. Гибель видов можно учесть в уравнениях (3.108), дополнительно предположив, что, когда численность какого-то вида опускается ниже порога $n_j \sim 100$, его необходимо исключить из динамических уравнений модели.

Итак, полная эволюционная модель описывается динамическими уравнениями, сама структура и набор переменных которых непрерывно меняются: случайно (за счет мутаций) и детерминированно (из-за вымирания), обусловленного естественным отбором. Эта структурная подвижность позволяет автоматически строить динамические системы с заданными (в определенных пределах) свойствами и поведением, а это как раз необходимо для задач обучения. Некоторые примеры использования эволюционных моделей в задачах обучения рассматривались в [83, 135].

При решении задач оптимизации нам обычно известна некоторая функция $U = U(x)$, ставящая в соответствие каждому возможному варианту x его эффективность U . Пусть, например, x — координата частицы, $U(x)$ — значение потенциала в точке с координатой x . Задача состоит в том, чтобы найти точку с абсолютным минимумом потенциала.

Аналоговое решение такой задачи оптимизации с использованием метода имитации отжига выглядело бы следующим образом. Предположим, что частица совершает вязкое движение в потенциальном силовом поле $U(x)$, так что ее координата меняется со временем согласно уравнению

$$\dot{x} = -\gamma \frac{\partial U}{\partial x}. \quad (3.109)$$

Тогда эта частица будет стремиться попасть в точку минимума потенциала U и оставаться в ней неограниченно долго. Если функция $U(x)$ обладает всего одним минимумом, решение получено. Если минимумов U много, то частица может попасть в любой из них в зависимости от своего начального положения и далее уже не перемещаться. Абсолютный минимум U при этом найден не будет.

При наличии тепловых флуктуаций частица совершает случайные блуждания в потенциальном силовом поле $U(x)$, так что вероятность $p(x, t)$ обнаружить частицу в момент t в точке с координатой x подчиняется диффузионному уравнению

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x} \left[\gamma \frac{\partial U}{\partial x} p \right] + D \frac{\partial^2 p}{\partial x^2}. \quad (3.110)$$

Легко проверить прямой подстановкой, что стационарное распределение вероятностей для (3.110) дается выражением

$$p_s(x) = \frac{1}{Z} \exp \left[-\frac{\gamma}{D} U(x) \right], \quad (3.111)$$

где $1/Z$ является *нормировочным множителем*. При выполнении соотношения Эйнштейна

$$D = \gamma \Theta, \quad (3.112)$$

устанавливающего связь между подвижностью γ и коэффициентом диффузии D , мы получаем равновесное распределение Больцмана с температурой Θ .

Тепловые флуктуации ведут к перебросам частицы между минимумами потенциала $U(x)$. Если два минимума отделены потенциальным барьером ΔU , то среднее время перехода между ними можно оценить как

$$\tau \sim \exp \left(\frac{\Delta U}{\Theta} \right). \quad (3.113)$$

Взяв в качестве ΔU максимальный потенциальный барьер ΔU_{\max} между двумя минимумами, согласно (3.113) получим характерное время установления равновесного распределения вероятностей при температуре Θ ; оно тем больше, чем выше потенциальные барьеры и чем ниже температура Θ . Ввиду экспоненциального закона убывания вероятностей (3.111) равновесное распределение вероятностей оказывается сосредоточенным в области x , где значение потенциала $U(x)$ отличается от его абсолютного минимума на величину, меньшую порядка Θ . Если в эту область попадают другие локальные минимумы потенциала U , то при наличии флуктуации наша динамическая система не может отличить их от абсолютного минимума.

Чтобы устранить неоднозначность в окончательном результате, необходимо постепенно понижать температуру Θ , устремив ее в конце к нулю, т. е. осуществлять медленный *отжиг* системы. Чтобы длительность отжига, гарантирующего правильное отыскание абсолютного минимума U , не была экспоненциально велика, его нужно начинать с температуры $\Theta_0 \sim \Delta U_{\max}$. Следовательно, при прочих равных условиях начальная температура равна Θ_0 и, значит, полная длительность процесса поиска оптимума по методу имитации отжига будет тем больше, чем выше потенциальные барьеры, отделяющие локальные минимумы функции U .

Рассмотрим теперь, как можно осуществить аналоговый поиск оптимума, используя модели с естественным отбором [46, 79, 80].

Основная идея эволюционного метода чрезвычайно проста. Мы знаем, что в ходе естественного отбора выживает тот вид, для которого эффективность s максимальна. Следовательно, если построить такую искусственную эволюционирующую систему, для которой эффективности различных видов x тем больше, чем меньше значение функции $U(x)$, то в процессе эволюции выживет тот вид, для которого $U(x)$ обладает абсолютным минимумом, что дает прямое решение поставленной задачи. В частности, выбрав $s(x) = U_{\max} - U(x)$, мы получим из уравнений (3.104)

$$\frac{\partial}{\partial t} n(x, t) = -[U(x) - \langle U \rangle] n(x, t), \quad (3.114)$$

где

$$\langle U \rangle = \frac{\int n(x, t) U(x) dx}{\int n(x, t) dx}. \quad (3.115)$$

Согласно (3.114), с течением времени распределение $n(x)$, сохраняя свою суммарную интенсивность

$$N = \int n(x, t) dx, \quad (3.116)$$

локализуется в точке абсолютного минимума функции $U(x)$, т. е. при $t \rightarrow \infty$ мы имеем

$$n(x, t) \rightarrow N \delta(x - x_0), \quad (3.117)$$

где x_0 — точка абсолютного минимума U .

Действительно, общее решение уравнения (3.114) есть

$$n(x, t) = N \frac{n(x, 0) \exp[-U(x)t]}{\int n(x, 0) \exp[-U(x)t] dx}. \quad (3.118)$$

При $t \rightarrow \infty$ интеграл в знаменателе можно приближенно вычислить по методу перевала:

$$\int n(x, 0) \exp[-U(x)t] dx \approx n(x_0, 0) \sqrt{\frac{2\pi}{U''(x_0)}} \exp[-U(x_0)t]. \quad (3.119)$$

Разлагая числитель по степеням отклонения $x - x_0$, находим

$$n(x, t) \approx N \frac{n(x, 0)}{n(x_0, 0)} \sqrt{\frac{\nu}{\pi}} \exp[-\nu(x - x_0)^2], \quad (3.120)$$

где $\nu = (1/2)U''(x_0)t$. Но произведение последних двух сомножителей в пределе $\nu \rightarrow \infty$ дает $\delta(x - x_0)$.

Заметим, что асимптотика (3.117) справедлива, лишь если $n(x_0, 0) \neq 0$, т. е. если в начальный момент времени точка x_0 была заселена. Поскольку нам предварительно неизвестно положение точки абсолютного минимума функции $U(x)$, при использовании описанного метода следует исходить из начального состояния, в котором численности *всех* видов x отличны от нуля¹⁷. Например, можно выбрать $n(x, 0) = \text{const}$.

Сравнивая эволюционный метод решения задач оптимизации, основанный на использовании динамических уравнений (3.114) с конкурентной мод (т. е. естественным отбором), и метод имитации отжига, использующий уравнение (3.110) с зависящим от времени коэффициентом диффузии $D(t) = \gamma\Theta(t)$, можно заметить, что в ряде случаев эволюционный метод позволяет более быстро найти решение.

Как видно из (3.118), время выхода на асимптотику (3.117) *не зависит* от размера барьеров, разделяющих минимумы функции $U(x)$, а следовательно, высота подобных барьеров не будет отражаться и на длительности процесса поиска оптимума. Напомним, что в методе имитации отжига длительность поиска растет с увеличением высоты потенциальных барьеров. Пример использования эволюционного метода для решения конкретной задачи оптимизации приведен в [46].

Рассмотрим теперь возможность использования систем с естественным отбором (т. е. конкуренцией мод) в задачах распознавания образов.

При построении динамических систем с ассоциативной памятью мы исходили из того, что динамическая система должна иметь столько аттракторов, каково число запомненных образов. Тогда распознаваемый образ подается в качестве начального условия, и результат распознавания определяется тем, области притяжения какого из аттракторов он принадлежит. Возможен, однако, и альтернативный подход, а именно, рассмотрим такую динамическую систему, которая описывается некоторым набором параметров $\{s_j\}$, причем при любом заданном наборе этих параметров она обладает *всего одним* аттрактором. Пусть, однако, эта динамическая система довольно чувствительна к изменению своих параметров, так что при изменении величин s_j один аттрактор сменяется другим. Ясно тогда, что каждому из возможных аттракторов будет отвечать своя область *в пространстве параметров*, образованном всевозможными наборами $\{s_j\}$. Если теперь выбрать в качестве набора $\{s_j\}$ распознаваемый образ и проследить, притяжением

¹⁷Альтернатива состоит в том, чтобы допустить мутации, приводящие к порождению всех возможных видов в процессе эволюции.

к какому из аттракторов будет завершаться эволюция системы, это также позволит осуществить аналоговое распознавание.

В качестве примера рассмотрим способ построения гипотетической биохимической системы, реализующей такой механизм распознавания изображений [6].

Пусть мы имеем плоский слой клеток, погруженных в жидкость. Все клетки идентичны. Каждая из них отделена от среды мембраной, пропускающей одни молекулы и непрозрачной для остальных. Внутри клеток могут происходить определенные фотохимические реакции, скорость которых зависит от интенсивности падающего сверху света.

Опишем вначале процессы, протекающие в системе на стадии *записи образов*.

При наличии молекул A_j сорта j в отдельной клетке идет синтез соответствующих молекул-ферментов χ_j , причем скорость синтеза пропорциональна освещенности I на этой клетке:



Чтобы записать изображение с номером j , следует спроецировать его сверху на слой клеток и создать картину освещенности $I_j(x)$. Одновременно с этим ко всем клеткам слоя через межклеточную среду нужно подвести в равных концентрациях молекулы сорта A_j , необходимые для синтеза. Тогда после записи j -го изображения в клетках слоя будут накоплены молекулы χ_j в концентрациях, пропорциональных локальной освещенности $I_j(x)$ в точках x для j -го изображения.

Перед записью следующего изображения необходимо удалить из клеток молекулы A_j и подать в них новые молекулы A_{j+1} с той же начальной концентрацией. Проецируя затем на слой клеток новое изображение, мы накапливаем в них молекулы χ_{j+1} с концентрацией $\chi_{j+1}(x) \sim I_{j+1}(x)$.

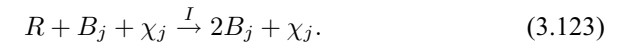
Допустим, что в слое клеток записано N изображений $I_j(x)$, $j = 1, 2, \dots, N$, которые мы будем рассматривать в качестве образов. Задача распознавания образов состоит в том, чтобы для всякого дальнейшего изображения $I(x)$ указать наиболее близкий к нему образ. Простейший критерий близости таков: образ j следует считать самым близким (среди прочих) к анализируемому изображению $I(x)$, если для него максимальна величина¹⁸

$$s_j = \int I_j(x)I(x) dx. \quad (3.122)$$

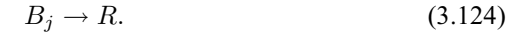
¹⁸Для простоты мы предположим, что интегральная освещенность $\int I_j(x) dx$ для всех образов одинакова.

На стадии автоматического *распознавания изображений* в системе протекают следующие процессы.

При наличии молекул субстрата R в клетке идут катализируемые молекулами-ферментами χ_j реакции репликации молекул B_j ; скорость этих фотохимических реакций пропорциональна освещенности I :



Происходит также реакция распада молекул B_j с выделением молекул субстрата:



Мы считаем, что клеточные мембраны прозрачны для молекул субстрата R и молекул B_j , так что клетка осуществляет свободный обмен ими с межклеточной средой. Кроме того, мы полагаем, что скорости реакций (3.123) и (3.124) малы по сравнению со временем диффузионного перемешивания для молекул B_j и R по всему слою. Это означает, что концентрации этих молекул поддерживаются одинаковыми для всех клеток, в отличие от концентрации молекул-ферментов χ_j , для которых мембраны непрозрачны.

С учетом сказанного кинетические уравнения для изменения со временем концентраций R и B_j имеют вид

$$\begin{aligned} \dot{B}_j &= k_+ R B_j \int I(x) \chi_j(x) dx - k_- B_j, \\ \dot{R} &= -k_+ R \sum_{j=1}^N B_j \int I(x) \chi_j(x) dx + k_- \sum_{j=1}^N B_j, \end{aligned} \quad (3.125)$$

где $k_+ I$ и k_- — скорости реакций (3.123) и (3.124). Считая, что размеры отдельных клеток достаточно малы, мы перешли в (3.125) к континуальному описанию.

Как следует из уравнений (3.125), в процессе эволюции сохраняется величина

$$C = R + \sum_{j=1}^N B_j. \quad (3.126)$$

Она имеет смысл начальной концентрации R молекул субстрата (предполагается, что в начальный момент времени концентрации B_j всех молекул-индикаторов были очень малы, хотя и отличны от нуля). Выразив R из (3.126)

и подставив его в (3.125), получим

$$\dot{B}_j = k_+ w_j \left(C - \sum_j B_j \right) B_j - k_- B_j, \quad (3.127)$$

где

$$w_j = \int I(x) \chi_j(x) dx. \quad (3.128)$$

Согласно (3.127), при выполнении неравенства

$$\max_j w_j < \frac{k_-}{k_+ C} \quad (3.129)$$

с течением времени все моды j вымирают, так что концентрации B_j стремятся к нулю. В обратном случае конкуренция приводит к выживанию той моды, для которой коэффициент w_j максимален, и вымиранию всех остальных мод.

Выше отмечалось, что концентрация $\chi_j(x)$ катализаторов, формируемая при записи образов, пропорциональна локальной освещенности $I_j(x)$ для соответствующего образа j , т.е. $\chi_j(x) = \zeta I_j(x)$. Поэтому $w_j = \zeta s_j$, а следовательно, конечная стадия эволюции будет прямо указывать, к какому из записанных образов наиболее близко анализируемое изображение. Действительно, каждому из образов соответствуют свои молекулы-индикаторы B_j . В конечном состоянии будут присутствовать лишь молекулы того индикатора, для которого s_j максимален. В частности, если каждый индикатор окрашивает клетки в свой цвет, система будет реагировать на подачу изображений изменением цвета слоя клеток, автоматически выбирая тот цвет, который отвечает наиболее близкому из хранящихся в ней образов.

БИБЛИОГРАФИЧЕСКИЙ КОММЕНТАРИЙ

К главе 1

Широкий круг вопросов, связанных с изучением автоволн и диссипативных структур в различных активных средах, рассмотрен в монографиях [37, 65, 69, 101, 105, 108, 119, 122, 126–128, 153, 192, 262, 266]. Имеется также ряд общих обзоров: [11, 20, 36, 39, 45, 46, 55, 59, 82, 98, 106, 132, 154]. В популярной форме эта проблемы обсуждаются в брошюрах [9, 87, 89, 93].

Конкретные модели различных активных сред приведены, например, в [10, 11, 15, 18, 42–44]. Большое число работ посвящено рассмотрению химических активных сред и в особенности среды с химической реакцией Белоусова–Жаботинского [27, 28, 54, 174, 198, 251, 252].

Существует обширная литература о волнах переключения в простых бистабильных системах. В первую очередь здесь необходимо отметить монографию [58] и подробный обзор [106]. Впервые волны переключения были рассмотрены в [62, 84, 155]. Различные задачи, возникающие в теории таких волн, изучаются в [19, 21, 45, 57, 125, 145, 146, 209]. В цикле работ [123, 124] и др. рассматриваются близкие вопросы образования пространственных нестационарных структур в средах с сильной нелинейной зависимостью теплопроводности от температуры или коэффициента диффузии от концентрации реагирующего вещества.

Подробный теоретический анализ механизмов, лежащих в основе распространения уединенных бегущих импульсов в возбудимых средах, проведен в [110, 140, 221, 222, 235] и [237]. Устойчивость и другие свойства бегущих импульсов изучались в [92, 131, 149, 151, 152, 199]. Бегущие импульсы в конкретных моделях возбудимых сред рассматривались в [76, 80, 150, 157, 229–231], стохастические последовательности бегущих импульсов — в [91].

Современные работы по использованию клеточных автоматов для моделирования волновых процессов в возбудимых средах восходят к статье [38], в которой была предложена так называемая аксиоматическая модель возбудимой среды и с ее помощью проанализирована циркуляция спи-

ральной волны вокруг отверстия. В ее рамках позднее было показано, что наличие отверстия не является необходимым условием [8, 46] и что спиральная волна может возникнуть из разрыва фронта в неоднородной по рефрактерности среде [85]. Различные модификации аксиоматической модели обсуждались в [41, 171, 232, 233, 271].

Спиральные волны в среде с химической реакцией Белоусова–Жаботинского были открыты в [56, 259]. Различные подходы к теоретическому описанию спиральных волн в возбудимых средах развивались в [46, 88, 111, 116, 118, 207, 226, 241]. Процессы взаимодействия спиральных волн между собой, а также с другими волновыми структурами в возбудимых средах изучались в [51, 52, 187, 240, 261]. Многообразные спиральные волны были обнаружены в экспериментах [86, 133]. Особой тщательностью и высокой техникой эксперимента выделяются исследования [211–214]. Результаты численного моделирования спиральных волн в возбудимых средах приведены в [65], а также в [66, 116, 227].

Кинематический подход был использован еще в 1951 г. при описании спирального роста кристалла вокруг винтовой дислокации [139]. Применительно к возбудимым средам этот подход был предложен в [63–65, 68]. При этом, однако, не учитывалась зависимость скорости прорастания от кривизны волнового фронта. Уравнения кинематики с учетом такой зависимости были получены в [31, 32, 46].

Вывод уравнений квазистационарного приближения дан в [30]. Обобщение кинематического подхода на случай трехмерных возбудимых сред проведено в [29]. Кинематический подход развивался также в цикле работ [178, 179, 255]. В указанных работах, однако, не рассматривалась динамика обрывов автоволнового фронта. Механизм возникновения спиральных волн из обрывов волнового фронта изучался в [112, 117]. Кинематический подход к описанию взаимодействий между бегущими импульсами в возбудимых средах развит в [210, 238].

Дрейф спиральных волн в неоднородных возбудимых средах изучался экспериментально в [242]; резонанс спиральных волн в светочувствительной возбудимой среде с реакцией Белоусова–Жаботинского был открыт в [2]. Простое качественное объяснение дрейфа спиральных волн предложено в [173]. Последовательная теория резонанса и дрейфа спиральных волн в возбудимых средах дана в [30].

Трехмерные вихри в химической среде с реакцией Белоусова–Жаботинского изучались в экспериментах [257, 260]. Численное моделирование сложных трехмерных автоволновых структур, а также рассмотрение отдельных теоретических вопросов, связанных с их описанием, предпринято в [113–115, 205, 208, 215, 224, 225, 274]. В [264, 267–270] дана подробная

классификация всех возможных трехмерных структур и сформулированы правила отбора для таких структур, основанные на некоторых топологических ограничениях (см. также [168]). Кинематическая теория вихревых колец в трехмерных возбудимых средах построена в [29]. В рамках кинематического подхода свойства скрученных вихрей и распространение волн скрученности изучались в [47].

Уравнения фазовой динамики для осциллирующих активных сред были выведены и изучались в [189, 191, 192]. Численные и аналитические исследования процессов, связанных с образованием регулярных или хаотических структур в осциллирующих активных средах, проводились в [4–7, 13, 14, 40, 59, 60, 96, 98, 107, 121, 122, 156, 160, 172, 176, 183, 200–203, 216–218, 220, 248, 256, 258, 260, 273]. Частота вращения и форма спиральной волны в осциллирующей активной среде были рассчитаны в работе [169] (см. также [97]). В выполненных ранее работах [59, 60, 107, 143, 147, 160, 161, 164, 165, 180, 181, 186] не удавалось правильно рассчитать частоту вращения спиральных волн (либо рассматривались лишь некоторые частные случаи осциллирующих сред, где спираль имеет особую вырожденную форму, а ее частота вращения совпадает с частотой однородных автоколебаний в такой среде).

Ведущие центры в среде с химической реакцией Белоусова–Жаботинского открыты в экспериментах [275, 276]. Длительное время считалось, что такие автоволновые структуры являются автономными и не связаны с наличием каких-либо неоднородностей или посторонних включений. Впоследствии было установлено [134, 254], что это не так: ведущие центры в среде с реакцией Белоусова–Жаботинского обусловлены присутствием примесей, локально переводящих среду в автоколебательный режим. Тем не менее имеется ряд теоретических моделей [90, 130, 182, 183, 219], объясняющих возникновение таких структур в однородных средах (при этом обычно остается открытым вопрос об устойчивости подобных структур). В [34] рассмотрена модель ведущих центров в трехкомпонентной активной среде. Теория пейсмекеров, основанная на уравнениях фазовой динамики, приведена в [192] (см. также [206]).

Стационарные диссипативные структуры в распределенных активных средах впервые рассматривались в [250]. Большой вклад в их изучение был внесен в работах И. Пригожина и его сотрудников (подробную библиографию см. в [108]). Различные вопросы, связанные с образованием и устойчивостью подобных структур, анализировались также в [12, 20, 22–24, 33, 35, 37, 49, 50, 71–76, 78, 79, 95, 103, 104, 120, 136–138, 141, 142, 152, 167].

Подробная классификация различных типов стационарных пульсирующих и бегущих структур в двухкомпонентных реагирующих системах

с диффузией, зависящая от соотношений между характерными временами и длинами, а также от типа нелинейности, проведена в [74–83].

К главе 2

В данных комментариях затронуты не только вопросы, освещенные в гл. 2, но и другие важные аспекты динамического хаоса.

Литература по теории динамического хаоса и близким разделам чрезвычайно обширна. Библиография по динамическим системам [235] содержит около четырех с половиной тысяч наименований. Для первоначального ознакомления с предметом можно рекомендовать [44, 51, 56, 74, 81, 83, 85, 104, 110, 115, 116, 157, 158, 189, 201, 202, 432], где вводятся все основные понятия теории динамического хаоса. В качестве математического введения в проблему обратим внимание на работы [26, 29, 33, 54, 179, 197, 213, 270, 271, 294, 439, 473–475]. Различные аспекты динамического хаоса и проблемы описания зарождения турбулентности обсуждаются в [137, 151, 152, 231, 233, 274, 291, 339, 368, 375, 419, 437, 473, 514].

Основные принципы механики достаточно полно изложены в [23, 30, 44]. Методы построения отображения Пуанкаре имеются в [6, 116, 348]. Модель Эно–Эйлеса рассматривалась во многих работах [116, 341, 347, 349, 523]. В [353] приведено доказательство ее неинтегрируемости. Вопросы теории интегрируемых систем обсуждаются в [30, 95, 150]. Начало теории КАМ положено в [96]; дальнейшее развитие эта теория получила в [20, 21, 148]. Обсуждение возникающих здесь вопросов имеется в [30]. Обзор методов доказательства неинтегрируемости гамильтоновых систем содержится в [95]. Математические аспекты многочастотных колебаний нелинейных систем со многими степенями свободы обсуждаются в монографии [174].

Впервые наличие диффузионных траекторий было строго доказано для гамильтоновой нелинейной системы специального вида [22]. В общем случае доказательства слияния хаотических траекторий в единую сеть не существует, но известно достаточно много примеров, где наблюдается это явление [85, 116, 263, 354, 382, 395, 396, 413, 510]. Строгая оценка сверху скорости диффузии Арнольда получена в работе [160] (см. также [30] и приведенные там ссылки). Различные случаи диффузии, возникающие при этом физические явления и оценки ее скорости для различных систем приведены в монографиях [30, 116]. В частности, в [529, 530] показано, что в задаче трех тел присутствует диффузия Арнольда. Диффузия в задаче N тел обсуждается в статье [442]. В работе [262] содержится доказательство наличия

диффузии Арнольда в задаче Ж. Л. Д'Аламбера о вращении сплюсненной у полюсов планеты по орбите вокруг неподвижной звезды. Астрономические следствия хаоса и диффузии Арнольда в Солнечной системе описаны в [224, 456]. Интересные приложения диффузии Арнольда и структура диффузионных траекторий, которая сейчас называется *стохастической наутиной*, описаны в книгах [82, 84]. Качественное рассмотрение диффузии Арнольда можно найти в [44, 81, 166].

Вопрос о природе хаотичности восходит к знаменитой проблеме пересечения сепаратрис в динамических системах. Это явление было открыто А. Пуанкаре при изучении задачи трех тел [167]. Элементарное изложение основ теории, вопросов, касающихся физических аспектов пересечения сепаратрис, а также история предмета, представлены в монографии [75]. Строгое изложение близких вопросов дается в работах [26, 30, 72, 234]. Теорему о неподвижной точке, ее качественное изложение и приложения можно найти в книге [48]. С математическими проблемами пересечения сепаратрис можно ознакомиться по работам [66, 72, 184, 484].

Эргодичность, перемешивание и некоторые их следствия обсуждаются в монографиях [44, 81, 83, 201]. Физические аспекты этих явлений можно найти в [6, 81, 83, 116]. Математические вопросы эргодической теории рассматриваются в [54, 60, 101, 220]. Бильярды обсуждаются в [53, 54, 101, 177, 178, 261, 502] (математические аспекты) и в [81, 83, 477] (физическое рассмотрение). Динамические системы статистической механики и кинетические уравнения с математической точки зрения рассматриваются в [76].

Доказательство движения частиц в ограниченной области плоскости или пространства тесно связано со знаменитой эргодической гипотезой Л. Больцмана. Рассмотрение бильярдных задач восходит к работам Ж. Адамара [342] (см. также [172]), где рассматривалось движение на закрученной поверхности отрицательной кривизны. Позже бильярды как динамические системы изучались Д. Биркгофом [48]. Более полное рассмотрение вопросов, касающихся динамики материальных точек в ограниченной области, было выполнено Н. С. Крыловым [103]. Затем класс бильярдных систем, которые способны проявлять хаотические свойства, был значительно расширен (см. [53, 502]). Свойства бильярда типа «стадион» описаны в [52, 249]. Наиболее общие условия возникновения хаотического поведения в системах бильярдного типа представлены в работах [53, 250]. На основе изучения двумерного газа Лоренца было доказано, что движение в чисто детерминированных системах может быть подобно броуновскому [252]. Этот результат явился первым строгим подтверждением проявления хаотичности динамическими (т. е. без какого-либо случайного механизма) си-

стемами. Описание бильярдов с возмущаемыми границами можно найти в [132, 407, 409]

Аттракторам диссипативных динамических систем посвящено большое число работ. В качестве введения в предмет можно рекомендовать обзоры [29, 115, 202] и монографии [189, 201, 233, 432]. С методами исследования устойчивости систем можно познакомиться по книгам [45, 90, 158, 204]. Теория показателей Флоке рассматривается, например, в [51, 90, 189, 190, 241]. Различные понятия и определения аттрактора имеются в [29, 33, 115, 173, 201, 202, 294, 338, 438]. Математическое определение физически наблюдаемого стохастического аттрактора предложено в [54, 179]. Понятие квазиаттрактора можно найти в статьях [32, 211, 212]. Система Лоренца впервые введена в [117, 478] и исследовалась во многих работах [35, 54, 55, 67, 89, 114, 116, 145, 158, 198, 467, 497]. С различными математическими аспектами странных аттракторов можно познакомиться по [28, 29, 54, 173, 179, 202, 294, 429, 440, 475]. Физические концепции хаотических колебаний изложены в [6, 51, 110, 158, 233, 368, 432, 471, 514].

Критерии динамического хаоса рассматриваются в [116, 158, 179, 201, 432]. Теория характеристических показателей Ляпунова и методы их расчета излагаются в [6, 115, 116, 158, 189, 228, 294, 312, 485, 528]. Энтропия динамических систем впервые была введена в [97, 98, 176]. С дальнейшим развитием математической теории энтропии можно познакомиться по [60]. Физические аспекты и способы численных расчетов энтропии обсуждаются в [116, 201, 227, 229, 312]. Спектральная плотность и ее характеристики изучаются в [116, 158, 164, 201, 294, 305, 432, 436]. Численные методы расчета спектральной плотности и автокорреляционной функции можно найти, например, в [162].

Геометрические характеристики фрактальных объектов рассматриваются в [87, 99, 142, 143, 343, 416, 448]. Различные аспекты метрической и вероятностной размерностей обсуждаются в [298]. Размерностные характеристики инвариантных множеств динамических систем исследуются в [54, 193, 294]. Методы расчета фрактальной размерности (емкости) аттракторов можно найти в [135, 476]. Гипотеза Каплана – Йорка предложена в [383] и далее обсуждалась в [54, 273, 311, 391, 448, 476, 533]. Соотношения различных размерностей, а также способы их вычисления рассмотрены в [135, 158, 201, 298]. Недостатки некоторых методов расчета обсуждаются в [327, 335].

Теории отображений (динамическим системам с дискретным временем) посвящено большое число работ. В качестве учебного пособия можно рекомендовать [56], а для более углубленного ознакомления с теорией — обзоры и монографии [156, 157, 197, 201, 426]. Э르고дическая теория одно-

мерных отображений рассматривается в [54]. Теорема Шарковского о сосуществовании циклов впервые была опубликована в [196]. Ее частный случай исследовался в широко известной работе [392]. Универсальному поведению квадратичных отображений посвящены статьи [54, 187, 201, 302, 303, 305]. Рассмотрение квадратичного отображения восходит к работе [517]. Одномерные отображения сыграли существенную роль в развитии метода ренорм-группы применительно к динамическим системам [116, 201, 271]. Отображение Эно впервые введено и изучалось в [203]. Различные аспекты этого отображения позднее исследовались в [281, 282, 449]. Отображения с более чем одним максимумом анализировались в [276, 358]. Для гамильтоновых отображений универсальные постоянные определены в статье [239] (см. также [116, 201]). Рассмотрение некоторых отображений, возникающих в физических задачах и приложениях, имеется в монографиях [81, 83, 116].

Исследование отображений комплексной плоскости восходит к работам Г. Жюлиа и П. Фату [301, 372]. Однако только относительно недавно вследствие большой эстетической привлекательности получаемых результатов и в связи с повсеместным использованием компьютеров это направление приобрело популярность. Хорошее введение в теорию дается в монографиях [102, 287]. С математическими аспектами отображений комплексной плоскости в себя можно ознакомиться по работам [147, 154]. Подробное описание множеств Жюлиа, Фату и Мандельброта, их бесконечное разнообразие и методы их раскраски представлены в книгах [163, 200].

Для первоначального ознакомления с теорией бифуркаций рекомендуем [13, 46, 90, 204]. С математической стороны вопроса можно ознакомиться по работам [27, 28, 33, 145]. Обратим внимание на книгу [265], где содержится обширная библиография. Некоторые методы и приемы исследования бифуркации рождения цикла имеются в [45, 56, 145, 192]. В частности, в [192] приведена программа для ЭВМ, позволяющая проанализировать такую бифуркацию. В монографиях [45, 145] и учебном пособии [56] (см. также [185]) описаны алгоритмы вычисления коэффициентов Ляпунова, характеризующих устойчивость родившегося цикла. Отметим работу [36], где обнаружена бифуркация рождения двумерного тора при потере устойчивости положения равновесия.

Сценарии развития хаоса (турбулентности) в диссипативных системах рассматриваются во многих работах. Сценарий Ландау – Хопфа был впервые предложен в [112, 356] и обсуждается в [113, 115, 116, 202]. Сценарий Рюэля – Такенса впервые рассматривался в статье [173] и позднее в [440]. Некоторые аспекты развития турбулентности по Рюэлю – Такенсу в рамках ренорм-группового подхода обсуждаются в [201]. Подтверждение этого сценария получено в ряде экспериментов (см., например, [7, 218, 280, 285, 466]).

Отметим статьи [331, 521] (см. также [201]), где показана возможность существования систем с трехчастотным [331] и более [521] режимами движения. Механизмы перемежаемости (I, II и III рода) подробно рассматриваются в [37, 136, 201, 421, 459, 460] (численные и аналитические результаты) и [217, 232, 237, 361, 532] (эксперименты). Сценарий развития хаоса Фейгенбаума обсуждается, например, в [92, 116, 201, 236, 450, 470, 501]. С другими сценариями можно познакомиться по следующим публикациям: [19, 33, 70, 158, 215, 270, 283, 360, 374, 375]. Понятие кризиса аттрактора впервые предложено в статье [329] и обсуждается в [28, 33, 201, 330]. Рассмотрение имеющихся здесь экспериментальных аспектов см. в [465].

В настоящее время существуют достаточно мощные методы, позволяющие стабилизировать динамику систем и/или управлять их поведением. Такая задача впервые была поставлена и численно решена в 1983 г. на примере модели, описывающей взаимодействие популяций [4, 5]. Позже возможность стабилизации хаотической динамики была строго обоснована в работах [123, 408]. Большое распространение проблема управления получила после публикации [446]. Аналитические методы подавления хаоса развиты в работах [121, 257, 386, 397]. Исследования показали, что проблема управления хаотическими динамическими системами и подавления хаоса имеет непосредственное отношение ко многим областям естественных наук, поскольку на этом пути удастся найти подходы к таким важным и насущным приложениям как обработка (запись, кодирование и расшифровка) информации [403, 404, 461], скрытая пересылка зашифрованных сообщений [127, 130, 131, 207, 216, 288, 344, 345, 387], проблема самоорганизации и искусственное создание когерентных структур в распределенных системах, обладающих пространственно-временным хаосом [122, 242, 406, 491], стабилизация сильно неупорядоченных сокращений сердечной мышцы и дефибрилляция [69, 93, 119, 133, 214, 316, 385, 444, 498, 503, 534, 535], инженерия динамических систем [405], и других [238, 260, 400, 483, 492, 496].

Современное состояние проблемы контроля и подавления хаоса дано в обзорах [238, 400, 486, 487]. Развитые сейчас методы дают возможность подавлять пространственно-временной хаос и стабилизировать динамику распределенных систем [93, 214, 385] (см. также литературу к «Предисловию ко второму изданию»). Обобщение данных результатов может оказаться чрезвычайно важным в приложениях, когда требуется *мягкое* выведение систем из состояния пространственно-временного хаоса, например, в кардиологии, где его развитие оказывается фатальным для организма. Современные методы стабилизации сильно хаотических сердечных ритмов (фибрилляции) являются очень жесткими: подача короткого электрического импульса большого напряжения и тока. Однако, как показывают предварительные

исследования, по-видимому, возможно слабым воздействием осуществить стабилизацию и восстановить ритм [93, 133, 214, 385, 444, 498, 503, 534, 535].

Теория пространственно-временного хаоса в распределенных системах в настоящее время далека от завершения. Существование конечномерных аттракторов в таких системах показано во многих работах, как экспериментальных (см., например, [71, 183, 241, 414]), так и численных и теоретических [43, 109, 275, 297, 366, 367, 441]. Современные методы позволяют непосредственно из экспериментальных данных, когда число степеней свободы исследуемой системы неизвестно, определить размерность вложения [3, 135, 245, 447], корреляционную размерность [326, 327], энтропию [268, 328], ляпуновские показатели [527] и обобщенную размерность [451]. В работах [244, 310, 480] описаны практические методы измерения различных характеристик динамического хаоса. Оптимизация метода Грассбергера–Прокаччиа проведена в [209, 512].

Математические аспекты пространственно-временного хаоса рассматриваются в статьях [106, 253, 289], а некоторые физические концепции — в [11, 15, 18, 42, 111, 210, 259, 299, 376, 509].

Ряд проблем, связанных с зарождением турбулентности, анализируется в [18, 64, 259, 266, 272, 506].

Численные и аналитические исследования свойств пространственно-временного хаоса в распределенных средах в настоящее время испытывают бурное развитие. Наряду с такими исследованиями для уравнения Гинзбурга–Ландау [15–17, 41, 62, 63, 79] проводился также подробный анализ развития конечномерной турбулентности в гидродинамических потоковых системах [8, 16, 18, 62]. Кроме того, имеется ряд работ [105–108, 253], в которых получены строгие результаты о характеристиках пространственно-временного хаоса для сцепленных отображений, служащих аппроксимациями распределенных активных сред.

Исследование таких отображений началось относительно недавно, когда для их моделирования стало возможным использовать достаточно мощные компьютеры. В публикациях [258, 376–378, 434, 513] рассмотрены различные аспекты динамики таких систем и представлены методы их анализа. Однако для решеток связанных отображений знание величин, определяющих их глобальное поведение на асимптотически больших временах, не представляет большого интереса. Более важно найти характеристики, описывающие локальные свойства таких систем и их эволюцию во времени. Такой критерий, позволяющей визуализировать поведение отдельных элементов и динамику всей системы в целом, недавно был предложен в работе [128]. Математические вопросы теории решеток сцепленных отображений описаны в работах [251, 253, 319, 370, 373].

Исследования, посвященные анализу временных рядов, в последнее время бурно развиваются [126, 141, 493]. Во многих случаях основное внимание уделяется не изучению свойств самой системы, а прогнозу динамики порожденного ею временного ряда [31, 309, 384]. Здесь выделяется два основных метода: сингулярный спектральный анализ [68] и локальная аппроксимация [299, 300]. Их сочетание, однако, дает более точный прогноз [126]. В работе [495] рассматривается перемещение донных отложений в бассейне реки Миссисипи. Прогноз строится с помощью метода LA. Полученные результаты свидетельствуют о высокой точности краткосрочных (до 5-и дней) прогнозов. Другое направление применения этих методов связано с анализом и прогнозированием природного явления Эль-Ниньо — теплого сезонного течения поверхностных вод пониженной солености в восточной части Тихого океана. В отдельные годы оно усиливается и, проникая далеко на юг, оттесняет от побережья холодные воды Перуанского течения. Это губительно действует на планктон, кораллы и рыб, а сопровождающие Эль-Ниньо сильнейшие ливни вызывают катастрофические наводнения на побережье. Некоторые археологи считают, что Эль-Ниньо в далеком прошлом могло стать одной из причин гибели высокоразвитых культур народов Южной Америки. При исследовании Эль-Ниньо в качестве наблюдаемой выступает индекс «южных колебаний» (Southern Oscillation Index). Он определяется как разность между среднемесячными значениями давления на уровне моря на острове Таити и в порту Дарвин (Австралия).

Применение метода ССА для анализа динамики индекса «южных колебаний» подробно изложено в работе [318]. Там же продемонстрирована процедура выделения наиболее значимых составляющих временного ряда. Прогноз индекса на основе ССА-разложения представлен в работе [384].

Рассмотрение вопроса о выделении основных составляющих системы и возможности фильтрации шума можно найти в [125, 141, 318, 519]. Математические аспекты анализа временных рядов, обоснование возможности прогноза представлены в статьях [68, 519]. Различные решения задачи реконструкции породившей временной ряд динамической системы описаны в [9, 47, 140, 141, 153, 208, 284, 462]. Обоснование выбора того или иного варианта обработки временного ряда можно найти в [126, 141, 256]. Качественные методы прогноза описаны в [153].

Книги [24, 25, 27, 29, 190, 204] можно рекомендовать в качестве введения в теорию обыкновенных дифференциальных уравнений, методы которой широко применяются при исследовании динамических систем.

К главе 3

Рассмотрение широкого круга вопросов, связанных с принципами обработки информации в распределенных активных средах и нейронных сетях, можно найти в книгах [3, 4, 12–14, 23, 29, 109, 115, 179, 180] и статьях [15, 31, 38, 51, 57, 64, 65, 70, 113, 134, 162, 186, 189, 205, 233]. В [16, 24, 52, 59, 61, 62, 166] обсуждаются перспективы создания и возможные механизмы функционирования молекулярных вычислительных устройств. В последнее время возрос интерес к оптическим компьютерам, где в качестве элементарной единицы обработки информации выступают целые двумерные изображения (см. [36, 37, 47, 48, 82, 102, 212, 224, 236]). Построение нейрокompьютеров с использованием традиционных электронных элементов рассматривается в [181, 216]. Сравнение возможностей компьютеров разных типов проведено в [120]. Реализация моделей нейронных сетей с помощью имеющихся матричных процессоров описана в [88]. Нетрадиционные компьютерные архитектуры предлагались в [39, 86, 87, 96, 127, 141, 162, 217].

Понятие клеточного автомата было введено Нейманом [185]. Различные свойства детерминированных клеточных автоматов рассматривались в [105, 106, 146, 163, 188, 230, 232]. Вероятностные клеточные автоматы изучаются в [7, 67, 68, 74, 107, 108, 145, 151, 156, 173]. Использование клеточных автоматов в качестве моделей различных физических процессов обсуждается в [54, 56, 77, 90, 117, 118, 152, 169, 199, 217, 218, 221–223, 231, 234].

В течение долгого времени схемы ассоциативной памяти исследовались главным образом в работах нейрофизиологов и специалистов по теории нейронных сетей (см. [9, 10, 13, 119, 167, 168]). Задача привлекла к себе внимание физиков-теоретиков после того, как Хопфилд показал [136], что она может быть решена аналоговым методом — с помощью искусственно построенной системы взаимодействующих изинговских спинов. Поскольку в рамках теории спиновых стекол развиты весьма эффективные математические методы для анализа свойств неупорядоченных спиновых систем, это позволило добиться быстрого прогресса в изучении статистических свойств модели Хопфилда и предложить ее многочисленные усовершенствования [4, 5, 8, 11, 32–35, 66, 71–73, 75, 76, 85, 89, 97, 98, 122–126, 150, 155, 161, 170, 172, 177, 183, 184, 191, 193–198, 213, 227]. На основе схемы Хопфилда с асимметричными взаимодействиями предложены также модели ассоциативной памяти для временных последовательностей сменяющих друг друга картин [49, 50, 69, 111, 160, 214]. Другие физические подходы к созданию устройств ассоциативной памяти обсуждаются в [6, 92, 93, 113–116]. Реализация модели Хопфилда с помощью сверхбольшой интегральной полупро-

водниковой схемы осуществлена в [103]. Оптоэлектронная реализация этой модели ассоциативной памяти предложена в [82].

Метод имитации отжига для решения задач комбинаторной оптимизации был предложен в [158] (см. также [157, 201, 202]). Статистическая механика искусственных физических систем, ассоциированных с различными сложными задачами комбинаторной оптимизации, изучалась в [40, 44, 91, 159, 175, 176, 182, 187, 215, 220, 235]. Связь между задачами оптимизации и проблемой распознавания образов обсуждается в [112].

Общие теоретические вопросы, связанные с созданием распределенных обучающихся систем, рассматривались в [128, 130, 131, 207]. Метод обратного распространения ошибок [164, 203] был использован для создания обучающихся систем, способных, например, определять пространственную форму объектов по их теневому плоскому изображению [165] или распознавать цели по данным звуковой локации [104]. Подробное описание деталей этого метода, а также дальнейшие примеры можно найти в [207, 210, 225]. Оптическая реализация метода описана в [224]. Обучающаяся «машина Больцмана» [25, 81, 129] использовалась в таких задачах, как восприятие речи [200] или поиск фигуры на имеющемся фоне [154, 209]. Существуют специализированные электронные интегральные схемы для реализации «машины Больцмана» [26]. Некоторые другие подходы к созданию распределенных обучающихся систем излагаются в [27, 43, 94, 95, 101, 132, 133, 135, 219].

Использование систем с естественным отбором (с нелинейной конкуренцией мод) для задач распознавания образов рассматривалось в [6, 92, 93, 113–116]. В [147–149] было отмечено, что иммунную систему человека можно рассматривать в качестве примера обучающейся эволюционной системы. Независимо были предложены [135] математические схемы управляющих устройств, обучение которых осуществляется в процессе конкуренции и естественного отбора. В статье [83] показано, что эти математические схемы очень близки к моделям иммунной системы (см. также [226]). Самоконструирующиеся системы с реакциями и диффузией рассматривались в [83, 84]. Применение эволюционных систем для аналогового решения сложных задач комбинаторной оптимизации исследовалось в [46, 79, 80]. Роль реакционно-диффузионных явлений в процессах обработки информации живой клеткой обсуждалась в [17, 18].

В работах [53, 58, 143, 144, 178, 192] обсуждались вопросы, связанные с использованием графов при аналоговой обработке информации.

Литература

К главе 1

1. *Абрамычев А. Ю., Давыдов В. А., Михайлов А. С.* // Биофизика. 1989. Т. 35. С. 100.
2. *Агладзе К. И., Давыдов В. А., Михайлов А. С.* // Письма ЖЭТФ. 1987. Т. 45. С. 601.
3. *Алдушин А. Л., Зельдович Я. Б., Маломед Б. А.* // ДАН СССР, 1980. Т. 251. С. 1102–1106.
4. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И.* // ЖЭТФ. 1985. Т. 89. С. 92–105.
5. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И., Розальский А. В., Сагдеев Р. З.* Решеточные модели в нелинейной динамике неравновесных сред: Препринт ИПФ АН СССР. №163. – Горький, 1987. 24 с.
6. *Арансон И. С., Рабинович М. И.* // Изв. вузов. Сер. «Радиофизика». 1986. Т. 29. С. 1514–1517.
7. *Ахромеева Т. С., Курдюмов С. П., Малинецкий Г. Г., Самарский А. А.* // ДАН СССР. 1984. Т. 274. С. 1091–1095.
8. *Балаховский И. С.* // Биофизика. 1965. Т. 10. С. 1063–1069.
9. *Балкарей Ю. И., Голик Л. Л., Елинсон М. И.* Автоволновые среды (использование в электронике). – М.: Знание, 1985. 64 с.
10. *Балкарей Ю. И., Елинсон М. И., Никулин М. Г.* // Проблемы современной радиотехники и электроники. – М.: Наука, 1980. С. 431.
11. *Балкарей Ю. И., Елинсон М. И., Никулин М. Г.* // Автоволновые процессы в системах с диффузией. – Горький: ИПФ АН СССР, 1981. С. 117–134.
12. *Балкарей Ю. И., Евтихов М. Г., Елинсон М. И.* // Микроэлектроника. 1980. Т. 9. С. 144–156.

13. Балкарей Ю. И., Никулин М. Г. // ЖТФ. 1979. Т. 49. 231–239.
14. Балкарей Ю. И., Никулин М. Г. // ЖТФ. 1979. Т. 49. С. 224–229.
15. Балкарей Ю. И., Ржанов Ю. А., Голик Л. Л., Елинсон М. И. // ФТП. 1982. Т. 16. С. 1558–1565.
16. Барелко В. В., Бейбутян В. М., Володин Ю. В., Зельдович Я. Б. // ДАН, СССР. 1981. Т. 257, С. 339–344.
17. Барелко В. В., Бейбутян В. М., Володин Ю. В., Зельдович Я. Б. // Автоволновые процессы в системах с диффузией. – Горький: ИПФ АН СССР, 1981. С. 135–148.
18. Барелко В. В., Володин Ю. И. // ДАН СССР. 1975. Т. 223. С. 112–115.
19. Баренблатт Г. И., Зельдович Я. Б. // ПММ. 1957. Т. 21. С. 856–859.
20. Белинцев Б. Н. // УФН, 1983. Т. 141. С. 55–101.
21. Белинцев Б. Н., Дибров Б. Ф., Ливищ М. А., Волькенштейн М. В. // Биофизика. 1978. Т. 23. С. 864–869.
22. Белинцев Б. Н., Ливищ М. А., Волькенштейн М. В. // Биофизика. 1978. Т. 23. С. 1056–1062.
23. Белинцев Б. Н., Ливищ М. А., Волькенштейн М. В. // ДАН СССР. 1981. Т. 257. С. 487–491.
24. Белинцев Б. Н., Ливищ М. А., Волькенштейн М. В. // Биофизика. 1982. Т. 27. С. 293–296.
25. Белинцев Б. Н., Савич Д. // Биофизика. 1985. Т. 30. С. 284–291.
26. Белинцев Б. Н., Юдин И. Д., Волькенштейн М. В. // ДАН СССР. 1985. Т. 285. С. 470–474.
27. Белоусов Б. П. // Сб. рефератов по радиационной медицине за 1958 г. – М.: Медгиз, 1959. С. 145.
28. Белоусов Б. П. // Автоволновые процессы в системах с диффузией. – Горький: ИПФ АН СССР, 1981. С. 176–189.
29. Бражник П. К., Давыдов В. А., Зыков В. С., Михайлов А. С. // ЖЭТФ. 1987. Т. 93. С. 1725–1736.

30. Бражник П. К., Давыдов В. А., Зыков В. С., Михайлов А. С. // Изв. вузов. Сер. «Радиофизика». 1988. Т. 31, №5. С. 574–584.
31. Бражник П. К., Давыдов В. А., Михайлов А. С. // Кинетика и горение. – Филиал Института химической физики АН СССР, Черноголовка. 1986. С. 39–43.
32. Бражник П. К., Давыдов В. А., Михайлов А. С. // Теор. мат. физ. 1987. Т. 74. С. 440–447.
33. Васильев В. А. Термодинамика биологических процессов. – М.: Наука, 1980. С. 186–192.
34. Васильев В. А., Полякова М. С. // Вестн. Моск. ун-та. Сер. физ. и астроном. 1975. Т. 16. С. 99–104.
35. Васильев В. А., Романовский Ю. М., Чернавский Д. С. Математическая биология развития. – М.: Наука, 1982. С. 82–101.
36. Васильев В. А., Романовский Ю. М., Яхно В. Г. // УФН. 1979. Т. 128. С. 625–666.
37. Васильев В. А., Романовский Ю. М., Яхно В. Г. Автоволновые процессы. – М.: Наука, 1987. 240 с.
38. Винер Н., Розенблют А. // Кибернетический сборник, вып. 3, М.: ИЛ, 1961. С. 3–56.
39. Гапонов-Грехов А. В., Рабинович М. И. // Нелинейные волны. Структуры и бифуркации. – М.: Наука, 1987. С. 37–44.
40. Гапонов-Грехов А. В., Рабинович М. И. // Изв. вузов. Сер. «Радиофизика». 1987. Т. 30. С. 131–143.
41. Гельфанд И. М., Цетлин М. Л. // ДАН СССР. 1960. Т. 131. С. 1242–1245.
42. Голубев В. Н., Рабинович М. И., Таланов В. И. и др. // Письма ЖЭТФ. 1985. Т. 42. С. 84–87.
43. Гренадер А. К., Панфилов А. В. // Биофизика. 1981. Т. 26. С. 1107–1108.
44. Гуляев Ю. В., Балкарей Ю. И. и др. // ДАН СССР. 1981. Т. 260. С. 82–85.
45. Гуревич А. В., Минц Р. Г. // УФН. 1984. Т. 142. С. 61–98.

46. Давыдов В. А., Михайлов А. С. // Нелинейные волны. Структуры и бифуркации. – М.: Наука, 1987. С. 261–279.
47. Давыдов В. А., Михайлов А. С., Бражник П. К. // Изв. вузов. Сер. «Радиофизика». 1989. Т. 32. №3. С. 289–293.
48. Езерский А. Б., Коротин П. И., Рабинович М. И. // Письма ЖЭТФ. 1985. Т. 41.
49. Еленин Г. Г., Крылов В. В., Полежаев А. А., Чернавский Д. С. // ДАН СССР. 1983. Т. 271. С. 84–88.
50. Елюхин В. А. // Биофизика. 1979. Т. 24. С. 1085–1089.
51. Ермакова Е. А., Кринский В. И., Панфилов А. В., Перцов А. М. // Биофизика. 1986. Т. 31. С. 855–861.
52. Ермакова Е. А., Перцов А. М. // Биофизика. 1986. Т. 31. С. 855–861.
53. Жаботинский А. М. // Колебательные процессы в биологических и химических системах. – М.: Наука, 1967. С. 252–254.
54. Жаботинский А. М. Концентрационные автоколебания. – М.: Наука, 1974. 178 с.
55. Жаботинский А. М. // Нелинейные волны. Самоорганизация. – М.: Наука, 1983. С. 16–25.
56. Жаботинский А. М., Заикин А. Н. // Колебательные процессы в биологических и химических системах. – НЦБИ АН СССР. Пущино, 1971. С. 314–317.
57. Жуков С. А., Барелко В. В., Мержанов А. Г. // ДАН СССР. 1979. Т. 245. С. 94–98.
58. Зельдович Я. Б., Баренблатт Г. И., Либрович В. Б., Махвиладзе Г. М. Математическая теория горения и взрыва. – М.: Наука, 1980. 478 с.
59. Зельдович Я. Б., Маломед Б. А. // Изв. вузов. Сер. «Радиофизика». 1982. Т. 25. С. 591–618.
60. Зельдович Я. Б., Маломед Б. А. // ДАН СССР. 1980. Т. 254. С. 92–94.
61. Зельдович Я. Б., Михайлов А. С. // УФН. 1987. Т. 153. С. 470–496.

62. Зельдович Я. Б., Франк-Каменецкий Д. А. // ЖЭХ. 1938. Т. 12. С. 100–105.
63. Зыков В. С. // Биофизика. 1980. Т. 25. С. 888–892.
64. Зыков В. С. // Биофизика. 1980. Т. 25. С. 319–322.
65. Зыков В. С. Моделирование волновых процессов в возбудимых средах. – М.: Наука, 1984. 165 с.
66. Зыков В. С. // Биофизика. 1986. Т. 31. С. 862–865.
67. Зыков В. С., Михайлов А. С. // ДАН СССР. 1986. Т. 286. С. 341–344.
68. Зыков В. С., Морозова О. Л. // Биофизика. 1980. Т. 25. С. 1071–1076.
69. Иваницкй Г. Р., Кринский В. И., Сельков К. Е. Математическая биофизика клетки. – М.: Наука, 1978. 300 с.
70. Ивлева Т. П., Мержанов А. Г., Шкадинский К. Г. // Физика горения и взрыва. 1980. Т. 16. С. 3–14.
71. Калафати Ю. Д., Сербинов И. А., Рябова Л. А. // ДАН СССР. 1982. Т. 263. С. 862–864.
72. Кернер Б. С., Кузнецов Е. М., Осипов В. В. // ДАН, 1984. Т. 277. С. 1114–1118.
73. Кернер Б. С., Кузнецов Е. М., Осипов В. В. // Микроэлектроника. 1984. Т. 13. С. 407–426.
74. Кернер Б. С., Осипов В. В. // ЖЭТФ. 1978. Т. 74. С. 1675–1697.
75. Кернер Б. С., Осипов В. В. // ЖЭТФ. 1980. Т. 79. С. 2218–2238.
76. Кернер Б. С., Осипов В. В. // Микроэлектроника. 1981. Т. 10. С. 407–432.
77. Кернер Б. С., Осипов В. В. // ЖЭТФ. 1982. Т. 83. С. 2201–2214.
78. Кернер Б. С., Осипов В. В. // ДАН СССР. 1982. Т. 264. С. 1366–1370.
79. Кернер Б. С., Осипов В. В. // ДАН СССР. 1983. Т. 270. С. 1104–1108.
80. Кернер Б. С., Осипов В. В. // Микроэлектроника. 1983. Т. 12. С. 512–529.
81. Кернер Б. С., Осипов В. В. // Письма ЖЭТФ. 1985. Т. 41. С. 386–388.

82. Кернер Б. С., Осипов В. В. // Микроэлектроника. 1985. Т. 14. С. 389–407.
83. Кернер Б. С., Осипов В. В. // ЖЭТФ. 1985. Т. 89. С. 589–607.
84. Колмогоров А. Н., Петровский И. Г., Пискунов Н. С. // Бюл. МГУ. Секция «Математика и механика». 1937. Т. 1. С. 1–26.
85. Кринский В. И. // Проблемы кибернетики. 1968. Вып. 20. С. 59–80.
86. Кринский В. И., Агладзе К. И. // ДАН СССР. 1982. Т. 263. С. 335–337.
87. Кринский В. И., Медвинский А. Б., Панфилов А. В. Эволюция автоволновых вихрей. – М.: Знание, 1986. 48 с.
88. Кринский В. И., Михайлов А. С. и др. // Изв. вузов. Сер. «Радиофизика». 1984. Т. 27. С. 1116–1129.
89. Кринский В. И., Михайлов А. С. Автоволны. – М.: Знание, 1984. 64 с.
90. Кринский В. И., Холопов А. В. // Биофизика. 1967. Т. 12. С. 224–529.
91. Кузнецов Ю. А., Панфилов А. В. Стохастические волны в системе Фитц–Хью–Нагумо. – Препринт НЦБИ АН СССР. Пущино, 1981. 8 с.
92. Кузнецов Ю. А. Существование и устойчивость бегущих волн в системах «реакция–диффузия» с одной пространственной переменной. – Препринт НЦБИ АН СССР. Пущино, 1982. 40 с.
93. Курдюмов С. П., Малинецкий Г. Г. Синергетика — теория самоорганизации. Идеи, методы, перспективы. – М.: Знание, 1983. 64 с.
94. Ландау Л. Д., Лифшиц Е. М. Квантовая механика. М.: Наука, 1974.
95. Лившиц М. А. // Биофизика. 1984. Т. 29. С. 699–705.
96. Маломед Б. А. // ЖТФ. 1981. Т. 51. С. 2476–2482.
97. Маломед Б. А. // ДАН СССР. 1986. Т. 291. С. 327–332.
98. Маломед Б. А. // Нелинейные волны. Структуры и бифуркации. – М.: Наука, 1987. С. 251–261.
99. Маломед Б. А., Руманов Э. Н. // ДАН СССР. 1985. Т. 284. С. 1355–1359.
100. Маломед Б. А., Трибельский М. И. // ДАН СССР. 1984. Т. 275. С. 1363–1367.

101. Маркин В. С., Пастушенко В. Ф., Чизмаджев Ю. А. Теория возбудимых сред. – М.: Наука, 1981. 273 с.
102. Маркман Г. С. // Биофизика. 1980. Т. 25. С. 713–715.
103. Маркман Г. С. // Биофизика. 1980. Т. 25. С. 697–699.
104. Маркман Г. С., Уринцев А. Л. // Биофизика. 1980. Т. 25. С. 697–699.
105. Марри Д. Нелинейные дифференциальные уравнения в биологии: Пер. с англ. – М.: Мир, 1982. 212 с.
106. Мержанов А. Г., Руманов Э. Н. // УФН. 1987. Т. 151. С. 553–593.
107. Михайлов А. С., Упоров И. В. // ДАН СССР. 1979. Т. 249. С. 733–736.
108. Николис Г., Пригожин И. Самоорганизация в неравновесных системах: Пер. с англ. – М.: Мир, 1979. 308 с.
109. Николис Г. Динамика иерархических систем. – М.: Мир, 1989. 486 с.
110. Островский Л. А., Яхно В. Г. // Биофизика. 1975. Т. 20. С. 489–496.
111. Панфилов А. В., Морнев О. Л., Цыганов М. А., Кринский В. И. // Биофизика. 1982. Т. 27. С. 1064–1069.
112. Панфилов А. В., Перцов А. М. // Биофизика. 1982. Т. 27. С. 886–889.
113. Панфилов А. В., Перцов А. М. // ДАН СССР. 1984. Т. 274. С. 1500–1503.
114. Панфилов А. В., Руденко А. В., Кринский В. И. // Биофизика. 1982. Т. 27. С. 886.
115. Панфилов А. В., Руденко А. В., Винфри А. Т. // Биофизика. 1985. Т. 30. С. 464–466.
116. Перцов А. М., Панфилов А. В. // Автоволновые процессы в системах с диффузией. – Горький: ИПФ АН СССР, 1981. С. 77–84.
117. Перцов А. М., Панфилов А. В., Медведева Ф. У. // Биофизика. 1983. Т. 28. С. 100–102.
118. Петвиашвили В. И., Сергеев А. М. // ДАН СССР. 1984. Т. 276. С. 1380–1384.

119. *Полак Л. С., Михайлов А. С.* Самоорганизация в неравновесных физико-химических системах. – М.: Наука, 1983. 286 с.
120. *Полежаев А. А.* // Биофизика, 1983.
121. *Романовский Ю. М.* // Изв. вузов. Сер. «Радиофизика», 1972. Т. 15. С. 718–722.
122. *Романовский Ю. М., Степанова Н. В., Чернавский Д. С.* Математическая биофизика. – М.: Наука, 1984. 304 с.
123. *Самарский А. А., Еленин Г. Г. и др.* // ДАН СССР. 1977. Т. 237. С. 1330–1337.
124. *Самарский А. А., Змитренко Н. В. и др.* // ДАН СССР. 1976. Т. 227. С. 321–324.
125. *Свирижев Ю. М., Гигаури А. А., Разжевайкин В. Н.* // Нелинейные волны. Самоорганизация. – М.: Наука, 1983. С. 32–47.
126. *Скотт Э.* Волны в активных и нелинейных средах в приложении к электронике. – М.: Сов. Радио, 1977. 322 с.
127. *Хакен Г.* Синергетика: Пер. с англ. – М.: Мир, 1980. 404 с.
128. *Эбелинг В.* Образование структур при необратимых процессах. – М.: Мир, 1979. 278 с.
129. *Юдин И. Д., Белинцев Б. Н., Лившиц М. А. и др.* // ДАН СССР. 1984. Т. 277. С. 987–991.
130. *Яхно В. Г.* // Биофизика. 1975. Т. 20. С. 669–675.
131. *Яхно В. Г.* // Биофизика. 1976. Т. 21. С. 547–552.
132. *Яхно В. Г.* // Автоволновые процессы в системах с диффузией. – Горький: ИПФ АН СССР, 1981. С. 46–76.
133. *Agladze K. I., Krinsky V. I.* // Nature. 1982. V. 296. P. 424–426.
134. *Agladze K. I., Krinsky V. I.* // Self-Organization: Autowaves and Structures far from Equilibrium / Ed. V. I. Krinsky. – Berlin: Springer, 1984. P. 147–149.
135. *Agladze K. I., Krinsky V. I., Pertsov A. M.* // Nature. 1984. V. 308. P. 834–835.

136. *Belintsev B. N., Livshits M. A., Volkenstein M. V.* // Z. Phys. 1978. V. 30B. P. 211–218.
137. *Belintsev B. N., Livshits M. A., Volkenstein M. V.* // Z. Phys. 1981. V. 44B. P. 345–351.
138. *Belintsev B. N., Livshits M. A., Volkenstein M. V.* // Phys. Lett. 1981. V. 82A. P. 375–377.
139. *Burton W. K., Cabrera N., Frank F. C.* // Phil. Trans. Roy. Soc. 1951. V. 243. P. 299–312.
140. *Casten R. C., Cohen H., Lagerstrom P. A.* // Quart. Appl. Math. 1975. V. 32. P. 365–381.
141. *Chernavskii D. S., Ruijgrok T. W.* // J. Theor. Biol. 1978. V. 73. P. 585–607.
142. *Chernavskii D. S., Ruijgrok T. W.* // BioSystems 1982. V. 15. P. 75–81.
143. *Cohen D. S., Neu J. C., Rosales R. R.* // SIAM J. Appl. Math. 1978. V. 35. P. 536–549.
144. *Duffy M. R., Britton N. F., Murray J. D.* // SIAM J. Appl. Math. 1980. V. 38. P. 8–13.
145. *Engel A.* // Phys. Lett. 1985. V. 113A. P. 139–142.
146. *Engel A., Ebeling W., Feistel R., Schimansky-Geier L.* // Selforganization by Nonlinear Irreversible Processes / Eds W. Ebeling, H. Ulbricht. – Berlin: Springer, 1986. P. 110–117.
147. *Erneux T., Herchkowitz-Kaufman M.* // Bull. Math. Biol. 1979. V. 41. P. 767–782.
148. *Feeney R., Schmidt S. L., Ortoleva P.* // Physica. 1981. V. 2D. P. 536–547.
149. *Feroe J. A.* // Biophys. J. 1978. V. 21. P. 103–110.
150. *Field R. J., Troy W. C.* // SIAM J. Appl. Math. 1979. V. 37. P. 561–589.
151. *Fife P. C.* // SIAM-AMS Proceedings. 1976. V. 10. P. 23.
152. *Fife P. C.* // J. Chem. Phys. 1976. V. 14. P. 554–561.
153. *Fife P. C.* Mathematical Aspects of Reacting and Diffusing Systems. – Berlin: Springer, 1979 (Lecture Notes in Biomathematics, №28).

154. *Fife P. C.* // Non-Equilibrium Dynamics in Chemical Systems // Eds C. Vidal, A. Pacault. – Berlin: Springer, 1984. P. 76–88.
155. *Fisher P.* // Ann. Eugenics. 1937. V. 7. P. 335–367.
156. *Fujii H., Sawada Y.* // J. Chem. Phys. 1978. V. 69. P. 3830–3842.
157. *Gibbs R. G.* // SIAM J. Appl. Math. 1980. V. 38. P. 422–444.
158. *Glendinning P.* // Phys. Lett. 1987. V. 121A. P. 411–413.
159. *Gomatam J.* // J. Phys. 1982. V. 15A. P. 1463–1471.
160. *Greenberg J. M.* // SIAM J. Appl. Math. 1976. V. 30. P. 199–215.
161. *Greenberg J. M.* // SIAM J. Appl. Math. 1976. V. 30. P. 421–433.
162. *Greenberg J. M., Hastings S. P.* // SIAM J. Appl. Math. 1978. V. 34. P. 515–523.
163. *Greenberg J. M., Hassard B. D., Hastings S. P.* // Bull. Amer. Math. Soc. 1978. V. 84. P. 1296–1327.
164. *Greenberg J. M.* // SIAM J. Appl. Math. 1980. V. 39. P. 301–309.
165. *Greenberg J. M.* // Adv. Appl. Math. 1981. V. 2. P. 450–462.
166. *Guckenheimer J.* // J. Math. Biol. 1975. V. 1. P. 259–271.
167. *Gurija G. T., Livshits M. A.* // Z. Phys. 1982. V. 47B. P. 71–77.
168. *Gurija G. T., Livshits M. A.* // Phys. Lett. 1983. V. 97A. P. 175–177.
169. *Hagan P. S.* // SIAM J. Appl. Math. 1982. V. 42. P. 762–781.
170. *Hastings S. P.* // Stud. Appl. Math. 1976. V. 55. P. 293–312.
171. *Hastings S. P.* // J. Math. Biol. 1981. V. 11. P. 105–117.
172. *Howard L. N., Koppell N.* // Stud. Appl. Math. 1977. V. 56. P. 95–111.
173. *Hramov R. N., Rudenko A. N., Panfilov A. V., Krinsky V. I.* // Studia Biophys. 1984. V. 102. P. 69–74.
174. *Karfunkel H. R., Seelig F. F.* // J. Math. Biol. 1975. V. 2. P. 123–135.
175. *Karfunkel H. R., Kahlert C.* // J. Math. Biol. 1977. V. 4. P. 183–192.

176. *Katon K.* // J. Phys. 1986. V. 19A. P. L1063–L1064.
177. *Keener J. P.* // SIAM J. Appl. Math. 1980. V. 39. P. 528–548.
178. *Keener J. P.* // SIAM J. Appl. Math. 1986. V. 46. P. 1039–1056.
179. *Keener J. P., Tyson J. J.* // Physica. 1986. V. 21D. P. 300–324.
180. *Koga S.* // Prog. Theor. Phys. 1982. V. 67. P. 164–178.
181. *Koga S.* // Prog. Theor. Phys. 1982. V. 67. P. 454–463.
182. *Kopell N., Howard L. N.* // Science. 1973. V. 180. P. 1171–1173.
183. *Kopell N., Howard L. N.* // Stud. Appl. Math. 1973. V. 52. P. 291–312.
184. *Kopell N., Howard L. N.* // Adv. Appl. Math. 1981. V. 2. P. 417–449.
185. *Kopell N., Howard L. N.* // Stud. Appl. Math. 1981. V. 64. P. 1–56.
186. *Krinsky V. I., Malomed B. A.* // Physica. 1983. V. 9D. P. 81–95.
187. *Krinsky V. I., Agladze K. I.* // Physica. 1983. V. 8D. P. 50–56.
188. *Kuhnert L.* // Nature. 1986. V. 6. P. 393–395.
189. *Kuramoto Y.* // Prog. Theor. Phys. Suppl. 1978. V. 64. P. 346–359.
190. *Kuramoto Y.* // Prog. Theor. Phys. 1980. V. 63. P. 1885–1895.
191. *Kuramoto Y.* // Physica. 1981. V. 106A. P. 128–133.
192. *Kuramoto Y.* Chemical Oscillations, Waves and Turbulence. – Berlin: Springer, 1984. 424 p.
193. *Kuramoto Y., Koga S.* // Prog. Theor. Phys. 1981. V. 66. P. 1081–1093.
194. *Kuramoto Y., Yamada T.* // Prog. Theor. Phys. 1976. V. 55. P. 2035–2042.
195. *Langer J. C.* // Rev. Mod. Phys. 1980. V. 52. P. 1–28.
196. *Lefever R.* // J. Chem. Phys. 1968. V. 49. P. 4977–4978.
197. *Livshits M. A.* // Z. Phys. 1983. V. 53B. P. 83–88.
198. *Madore B. F., Freedman W. L.* // Science. 1983. V. 222. P. 615–616.

199. *Maginu K.* // J. Math. Biol. 1978. V. 6. P. 49–57.
200. *Malomed B. A.* // Physica. 1983. V. 8D. P. 353–359.
201. *Malomed B. A.* // Z. Phys. 1984. V. 55B. P. 241–248; P. 249–256.
202. *Malomed B. A., Staroselsky I. E.* // Phys. Lett. 1983. V. 99A. P. 145–146.
203. *Malomed B. A., Tribelsky M. I.* // Physica. 1984. V. 14D. P. 67–87.
204. *Markstein G. H.* // J. Aeronaut. Sci. 1951. V. 18. P. 199–207.
205. *Medvinsky A. B., Panfilov A. V., Pertsov A. M.* // Self-Organization: Autowaves and Structures far from Equilibrium / Ed. V.I. Krinsky. – Berlin: Springer, 1984. P. 195–199.
206. *Mikhailov A. S., Engel A.* // Phys. Lett. 1986. V. 117A. P. 257–260.
207. *Mikhailov A. S., Krinsky V. I.* // Physica. 1983. V. 9D. P. 346–376.
208. *Mikhailov A. S., Panfilov A. V., Rudenko A. N.* // Phys. Lett. 1985. V. 109A. P. 246–250.
209. *Mikhailov A. S., Schimansky-Geier L., Ebeling W.* // Phys. Lett. 1983. V. 96A. P. 453–456.
210. *Miller R. N., Rinzel J.* // Biophys. J. 1981. V. 34. P. 227–259.
211. *Muller S. L., Plessner T., Hess B.* // Science. 1985. V. 230. P. 661–663.
212. *Muller S. L., Plessner T., Hess B.* // Naturwiss. 1986. V. 73. P. 165–179.
213. *Muller S. L., Plessner T., Hess B.* // Physica. 1987. V. 24D. P. 71–86.
214. *Muller S. L., Plessner T., Hess B.* // Physica. 1987. V. 24D. P. 87–96.
215. *Nandapurkar P. J., Winfree A. T.* // Physica. 1987. V. 29D. P. 69–83.
216. *Neu J. C.* // SIAM J. Appl. Math. 1979. V. 36. P. 509–514.
217. *Neu J. C.* // SIAM J. Appl. Math. 1979. V. 37. P. 307–317.
218. *Neu J. C.* // SIAM J. Appl. Math. 1980. V. 38. P. 305–309.
219. *Ortoleva P.* // J. Chem. Phys. 1978. V. 69. P. 300–307.
220. *Ortoleva P., Ross J.* // J. Chem. Phys. 1973. V. 58. P. 5673–5691.

221. *Ortoleva P., Ross J.* // J. Chem. Phys. 1974. V. 60. P. 5090–5099.
222. *Ortoleva P., Ross J.* // J. Chem. Phys. 1975. V. 63. P. 3398–3431.
223. *Panfilov A. V., Rudenko A. N., Pertsov A. M.* // Self-Organization: Autowaves and Structures far from Equilibrium / Ed. V.I. Krinsky. – Berlin: Springer, 1984. P. 103–105.
224. *Panfilov A. V., Rudenko A. N.* // Physica, 1987. V. 28D. P. 215–218.
225. *Panfilov A. V., Winfree A. T.* // Physica. 1985. V. 17D. P. 323–330.
226. *Pertsov A. M., Ermakova E. A., Panfilov A. V.* // Physica. 1984. V. 14D. P. 117–124.
227. *Pertsov A. M., Panfilov A. V., Ermakova E. A.* // Physica. 1984. V. 14D. P. 311–328.
228. *Prigogine I., Lefever R.* // J. Chem Phys. 1968. V. 49. P. 283–292.
229. *Rajagopal K.* // Phys. Lett. 1983. V. 99A. P. 261–264.
230. *Rajagopal K.* // Phys. Lett. 1984. V. 100A. P. 269–272.
231. *Rajagopal K.* // Phys. Lett. 1984. V. 100A. P. 49–51.
232. *Reshodko L. V.* // J. Gen. Biol. 1973. V. 1. P. 80–87.
233. *Reshodko L. V., Bures J.* // Biol. Cybern. 1975. V. 18. P. 181–189.
234. *Richter P. H., Procaccia I., Ross J.* // Adv. Chem. Phys. 1980. V. 43. P. 217–285.
235. *Rinzel J.* // J. Math. Biol. 1975. V. 2. P. 205–224.
236. *Rinzel J., Elmentrout G. B.* // J. Chem. Phys. 1982. V. 86. P. 2954–2967.
237. *Rinzel J., Keller J. B.* // Biophys. J. 1973. V. 13. P. 1313–1337.
238. *Rinzel J., Maginu K.* // Non-Equilibrium Dynamics in Chemical Systems. / Eds. C. Vidal, A. Pacault. – Berlin: Springer, 1984. P. 107–113.
239. *Rosenblueth A.* // Amer. J. Physiol. 1958. V. 194. P. 491–508.
240. *Rossler O. E., Kahlert C.* // Z. Naturforsch. 1979. Bd 34. S. 565–573.

241. *Rovinsky A. B.* // J. Phys. Chem. 1986. V. 90. P. 217–219.
242. *Rudenko A. N., Panfilov A. V.* // Studia Biophys. 1983. V. 98. P. 183–188.
243. *Schlogl F. Z.* // Z. Phys. 1971. Bd. 243. S. 303–313.
244. *Schmidt S., Ortoleva P.* // J. Chem. Phys. 1980. V. 72. P. 2733–2742.
245. *Selfridge O.* // Arch. Inst. Cardiol. Mech. 1948. V. 53. P. 113–121.
246. *Shinomoto S., Kuramoto Y.* // Prog. Theor. Phys. 1986. V. 75. P. 1105–1110.
247. *Shyldkrot H., Ross J.* // J. Chem. Phys. 1985. V. 82. P. 113–122.
248. *Sivashinsky G. I.* // Ann. Rev. Fluid Mech. 1983. V. 15. P. 179–199.
249. *Sultan R., Ortoleva P.* // J. Chem. Phys. 1986, V. 84. P. 6781–6789.
250. *Turing A. M.* // Phil. Trans. Roy. Soc. 1952. V. 237. P. 37–72.
251. *Tyson J. J.* The Belousov–Zhabotinsky Reaction. – Berlin: Springer, 1976. 150 p. (Lecture Notes in Biomathematics, V. 10).
252. *Tyson J. J.* // Ann. N.Y. Acad. Sci. 1979. V. 316. P. 279–295.
253. *Tyson J. J.* // J. Chem. Phys. 1984. V. 80. P. 6079–6982.
254. *Tyson J. J., Fife P. C.* // J. Chem. Phys. 1980. V. 73. P. 2224–2236.
255. *Tyson J. J., Keener J. P.* // Physica D. 1987. V. 29D. P. 215–222.
256. *Walgraef D., Dewel G., Borckmans P.* // J. Chem. Phys. 1983. V. 78. P. 3043–3051.
257. *Welsh B. J., Gomatam J., Burgess A. E.* // Nature. 1983. V. 304. P. 611–614.
258. *Winfree A. T.* // J. Theor. Biol. 1967. V. 16. P. 15–31.
259. *Winfree A. T.* // Science. 1972. V. 175. P. 634–640.
260. *Winfree A. T.* // Farad. Symp. Chem. Soc. 1974. V. 9. P. 38–41.
261. *Winfree A. T.* // Sci. Amer. 1974. V. 230. P. 82–95.
262. *Winfree A. T.* // The Geometry of Biological Time. – Berlin: Springer, 1980.
263. *Winfree A. T.* // Sci. Amer. 1983. V. 248. P. 144–161.

264. *Winfree A. T.* // Physica. 1984. V. 12D. P. 321–332.
265. *Winfree A. T.* The Timing of Biological Clocks. – N.Y.: Sci. Amer. Books, 1986.
266. *Winfree A. T.* When Time Breaks Down. – N.Y.: Princeton Univ. Press, 1987.
267. *Winfree A. T., Strogatz S. H.* // Physica. 1983. V. 9D. P. 35–49.
268. *Winfree A. T., Strogatz S. H.* // Physica. 1983. V. 9D. P. 65–80.
269. *Winfree A. T., Strogatz S. H.* // Physica. 1983. V. 9D. P. 333–345.
270. *Winfree A. T., Strogatz S. H.* // Physica. 1983. V. 13D. P. 221–233.
271. *Winfree A. T., Winfree E. M., Seifert H.* // Physica. 1985. V. 17D. P. 109–131.
272. *Wood P. M., Ross J.* // J. Chem. Phys. 1985. V. 82. P. 1924–1936.
273. *Yakhov V.* // Phys. Rev. 1981. V. 24A. P. 642–652.
274. *Yakushevich L. V.* // Studia Biophys. 1984. V. 100. P. 195–200.
275. *Zaikin A. N., Zhabotinsky A. M.* // Nature. 1970. V. 225. P. 535–538.
276. *Zhabotinsky A. M., Zaikin A. N.* // J. Theor. Biol. 1973. V. 40. P. 45–61.

К главе 2

1. *Автоволновые процессы в системах с диффузией. Сб. научн. трудов.* – ИПФ АН СССР, 1981.
2. *Айвазян С. А., Мхитарян В. С.* Прикладная статистика и основы эконометрики. – М.: ЮНИТИ, 1998. 1022 с.
3. *Александров Б. П., Ланда П. С., Швилкин Б. Н.* // Физика плазмы. 1986. Т. 12. С.120–123.
4. *Алексеев В. В., Лоскутов А. Ю.* // Вестник Моск. ун-та, сер. Физ.-Астр. 1985. Т.26. №3. С.40–44.
5. *Алексеев В. В., Лоскутов А. Ю.* // ДАН СССР. 1987. Т.293. Вып.6. С.1346–1348.
6. *Анищенко В. С.* Сложные колебания в простых системах. – М.: Наука, 1990. – 312 с.

7. *Анищенко В. С.* // ЖТФ. 1986. Т. 56. С. 225–237.
8. *Анищенко В. С., Арансон И. С., Постнов Д. Э., Рабинович М. И.* // ДАН СССР. 1986. Т. 286. С. 1120–1124.
9. *Анищенко В. С., Астахов В. В., Вадивасова Т. Е. и др.* Нелинейные эффекты в хаотических и стохастических системах. — Москва–Ижевск: Ин-т компьт. иссл-ий, 2003. 530 с.
10. *Анищенко В. С., Летчфорд Т. Е., Сафонова М. А.* // Изв. вузов. Сер. «Радиофизика». 1984. Т. 27. С. 565–575.
11. *Анищенко В. С., Постнов Д. Э.* // ЖТФ. 1985. Т. 55. С. 162–167.
12. *Андронов А. А., Витт А. А.* // ЖЭТФ. 1933. Т. 3. С. 373–374.
13. *Андронов А. А., Леонтович Е. А., Гордон И. И., Майер А. Г.* Теория бифуркаций динамических систем на плоскости. — М.: Наука, 1967. 488 с.
14. *Андронов А. А., Понтрягин Л. С.* // ДАН СССР. 1937. Т. 14. С. 247–250.
15. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И.* // ЖЭТФ. 1985. Т. 89. С. 92–105.
16. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И., Старобинец И. М.* // ЖЭТФ. 1986. Т. 90. С. 1707–1718.
17. *Арансон И. С., Рабинович М. И.* // Изв. вузов. Сер. «Радиофизика». 1986. Т. 29. С. 1514–1517.
18. *Арансон И. С., Рабинович М. И.* Ренормгрупповое описание пространственного развития турбулентности: Препринт ИПФ АН СССР. №152. — Горький, 1987.
19. *Арансон И. С., Рульков Н. Ф.* // ЖТФ. 1988. Т. 58. С. 1656–1669.
20. *Арнольд В. И.* // Успехи матем. наук. 1963. Т. 18. С. 13–40.
21. *Арнольд В. И.* // Успехи матем. наук. 1963. Т. 18. С. 91–192.
22. *Арнольд В. И.* // ДАН СССР. 1964. Т. 156. С. 9–12.
23. *Арнольд В. И.* Математические методы классической механики. — М.: Наука, 1974. 432 с.

24. *Арнольд В. И.* Дополнительные главы теории обыкновенных дифференциальных уравнений. — М.: Наука, 1978. 304 с.
25. *Арнольд В. И.* Обыкновенные дифференциальные уравнения. — М.: Наука, 1984. 272 с.
26. *Арнольд В. И., Авец А.* Эргодические проблемы классической механики. — Изд. «Рег. и хаотич. динамика», 1999. 284 с.
27. *Арнольд В. И.* Геометрические методы в теории обыкновенных дифференциальных уравнений. — М-Ижевск: РХД, 2000. — 400 с..
28. *Арнольд В. И., Афраймович В. С., Ильяшенко Ю. С., Шильников Л. П.* // Современные проблемы математики. Фундаментальные направления. Т. 5. — М.: Изд-во ВИНТИ, 1986. С. 5–218.
29. *Арнольд В. И., Ильяшенко Ю. С.* // Современные проблемы математики. Фундаментальные направления, Т. 1. — М.: Изд-во ВИНТИ, 1985. С. 7–149.
30. *Арнольд В. И., Козлов В. В., Нейштадт А. И.* Математические аспекты классической и небесной механики. — УРСС, Москва, 2002. 416 с.
31. *Афанасьев В. Н., Юзбашев М. М.* Анализ временных рядов и прогнозирование. — М.: Финансы и статистика, 2001. 226 с.
32. *Афраймович В. С.* // Проблемы нелинейных и турбулентных процессов в физике. — Киев: Наук. думка, 1985. ч. 2. С. 21–24.
33. *Афраймович В. С.* // Нелинейные волны. Структуры и бифуркации./ Ред. А.В.Гапонов-Грехов, М.И.Рабинович. — М.: Наука, 1987, с.189–213.
34. *Афраймович В. С.* // Нелинейные волны. Динамика и эволюция./ Ред. А.В.Гапонов-Грехов, М.И.Рабинович. — М.: Наука, 1989. С.16–29.
35. *Афраймович В. С., Быков В. В., Шильников Л. П.* // Тр. Моск. Матем. об-ва, 1982. Т. 44. С.152–212.
36. *Афраймович В. С., Возовой Л. П.* // ДАН СССР. 1988. Т. 302. С. 823–826.
37. *Афраймович В. С., Шильников Л. П.* // ДАН СССР. 1974. Т. 219. С. 1281–1285.
38. *Афраймович В. С., Быков В. В., Шильников Л. П.* // УМН, 1980. Т.35. С.164–165.

39. *Афраймович В. С., Некоркин В. И., Осипов Г. В., Шалфеев В. Д.* Устойчивость, структуры и хаос в нелинейных сетях синхронизации. / Ред. А. В. Гапонов-Грехов, М. И. Рабинович. — Изд-во ИПФ АН, Горький, 1989.
40. *Афраймович В. С., Шильников Л. П.* // Методы качественной теории дифференциальных уравнений. — Горький: ГГУ, 1983. С.3–26.
41. *Ахромеева Т. С., Курдюмов С. П., Малинецкий Г. Г., Самарский А. А.* // ДАН СССР. 1984. Т. 279. С. 1091–1096.
42. *Ахромеева Т. С., Курдюмов С. П., Малинецкий Г. Г., Самарский А. А.* // Математическое моделирование. Процессы в нелинейных средах. — М.: Наука, 1986. С. 7–59.
43. *Бабин А. В., Вишик М. И.* // УМН. 1983. Т. 38. С. 132–187.
44. *Балеску Р.* Равновесная и неравновесная статистическая механика. Т. 2. — М.: Мир, 1978. 400 с.
45. *Баутин Н. Н.* Поведение динамических систем вблизи границ области устойчивости. — М.: Наука, 1984. 176 с.
46. *Баутин Н. Н., Леонтович Е. А.* Методы и приемы качественного исследования динамических систем на плоскости. — М.: Наука, 1976. 496 с.
47. *Безручко Б. П., Смирнов Д. А.* Математическое моделирование и хаотические временные ряды. — Саратов: ГосУНЦ «Колледж», 2005. — 320 с.
48. *Биркгоф Д.* Динамические системы. — Ижевск, Изд. дом «Удмуртский ун-т», 1999. 408 с.
49. *Боголюбов Н. Н., Митропольский Ю. А.* Асимптотические методы в теории нелинейных колебаний. — М., Физматгиз, 1958. 412 с.
50. *Бокс Дж., Дженкинс Г.* Анализ временных рядов. — М.: Мир, 1974. 405 с.
51. *Бонч-Бруевич В. Л.* Стохастические автоколебания в конденсированной среде. Часть I. — М.: Изд-во МГУ, 1986. 68 с.
52. *Бунимович Л. А.* // Матем. сб. 1974. Т.94. С.49–73.

53. *Бунимович Л. А.* // Динамические системы — 2. Серия «Современные проблемы математики. Фундаментальные направления.» — ВИНТИ, 1985. С.173–204.
54. *Бунимович Л. А., Песин Я. Б., Синай Я. Г., Якобсон М. В.* // Современные проблемы математики. Фундаментальные направления. Т. 2. — М.: Изд-во ВИНТИ. 1985. С. 113–231.
55. *Бунимович Л. А., Синай Я. Г.* // Нелинейные волны. — М.: Наука, 1979. С. 212–226.
56. *Бутенин Н. В., Неймарк Ю. И., Фуфаев Н. А.* Введение в теорию нелинейных колебаний. — М.: Наука, 1987. 384 с.
57. *Быков В. В.* // Методы качественной теории дифференциальных уравнений. — Горький, ГГУ, 1980. С.44–72.
58. *Вайнберг С.* Гравитация и космология. — М.: Мир, 1975, 696 с.
59. *Васильев В. А., Романовский Ю. М., Яхно В. Г.* Автоволновые процессы. / Ред. Д. С. Чернавский. — М., Наука, 1987. 240 с.
60. *Вершик А. М., Корнфельд И. П., Синай Я. Г.* // Современные проблемы математики. Фундаментальные направления. Т. 2. — М.: Изд-во ВИНТИ, 1985. С. 5–111.
61. *Видаль К.* // Синергетика. — М.: Мир, 1984. С. 109–125.
62. *Гапонов-Грехов А. В., Рабинович М. И.* // Нелинейные волны. Структуры и бифуркации. — М.: Наука, 1987. С. 7–44.
63. *Гапонов-Грехов А. В., Рабинович М. И.* // Изв. вузов. Сер. «Радиофизика». 1987. Т. 30. С. 131–143.
64. *Гапонов-Грехов А. В., Рабинович М. И., Старобинец И. М.* // Письма в ЖЭТФ, 1984. Т. 39. С.561–565.
65. *Гарел Д., Гарел О.* Колебательные химические реакции. — М.: Мир, 1986. 148 с.
66. *Гельфрейх В. Г., Лазуткин В. Ф.* // Успехи матем. наук. 2001. Т.56. С.79–142.
67. *Гилмор Р.* Прикладная теория катастроф. Кн. 2. — М.: Мир, 1984. 285 с.

68. *Главные компоненты временных рядов: метод «Гусеница»*. Сб. статей. Ред. Д. Л. Данилов и А. А. Жиглявский. СПбГУ, 1997. 308 с.
69. *Гласс Л., Мэки М.* От часов к хаосу. Ритмы жизни. — М., Мир, 1991. 248 с.
70. *Гольберг А. И., Синай Я. Г., Ханин К. М.* // УМН. 1983. Т. 38. С. 159–160.
71. *Громов П. Р., Зобнин А. Б., Рабинович М. И., Рейман А. М., Суццик М. М.* // ДАН СССР. 1987. Т. 292. С. 284–287.
72. *Гукенхеймер Дж., Холмс Ф.* Нелинейный колебания, динамические системы и бифуркации векторных полей. — Москва–Ижевск, Ин–т комп. иссл., 2002. 560 с.
73. *Дайлюденко В. Ф., Крот А. М.* // Журнал вычислит. матем. и матем. физики. 1997. Т.27. С.315–324.
74. *Данилов Ю. А., Кадоццев Б. Б.* // Нелинейные волны. Самоорганизация. М.: Наука, 1983. С. 5–16.
75. *Диакун Ф., Холмс Ф.* Небесные встречи. Истоки хаоса и нелинейности. — РХД, Москва–Ижевск, 2004. 304 с.
76. *Добрушин Р. Л., Маслова Н. Б., Синай Я. Г., Сухов Ю. М.* // Современные проблемы математики. Фундаментальные направления. Т. 2. — М.: Изд-во ВИНТИ, 1985. С. 233–307.
77. *Дуади А.* // Множества Жюлиа и множество Мандельброта. В кн. Х.–О.Пайтген, П.Х.Рихтер. Красота фракталов. Образы комплексных динамических систем. — М.: Мир, 1993. С.141–153.
78. *Дубров А. М., Мхитарян В. С., Трошин Л. И.* Многомерные статистические методы. — М.: Финансы и статистика, 2000. 352 с.
79. *Езерский А. Б., Рабинович М. И., Реутов В. П., Старобинец И. М.* // ЖЭТФ. 1986. Т. 91. С. 2070–2083.
80. *Заславский Г. М.* Стохастическая необратимость в нелинейных системах. — М., Наука, 1970.
81. *Заславский Г. М.* Стохастичность динамических систем. — М., Наука, 1984. 272 с.

82. *Заславский Г. М.* Физика хаоса в гамильтоновых системах. — Москва–Ижевск, Ин–т компьютерных иссл–ий, 2004. 288 с.
83. *Заславский Г. М., Сагдеев Р. З.* Введение в нелинейную физику. — М., Наука, 1988. 368 с.
84. *Заславский Г. М., Сагдеев Р. З., Усиков Д. А., Черников А. А.* Слабый хаос и квазирегулярные структуры. — М., Наука, 1991. 237 с.
85. *Заславский Г. М., Чириков Б. В.* // УФН. 1971. Т. 105. С. 3–40.
86. *Зельдович Я. Б., Баренблатт Г. И., Либрович В. Б., Махвиладзе Г. И.* Математическая теория горения и взрыва. — М., Наука, 1980. 478 с.
87. *Зельдович Я. Б., Соколов Д. Д.* // УФН. 1985. Т. 146. С. 493–506.
88. *Ильин В. А., Позняк Э. Г.* Основы математического анализа. Часть II. — М.: Наука, 1973. 448 с.
89. *Йорке Дж., Йорке Е.* Странные аттракторы. — М.: Мир, 1981. С. 193–212.
90. *Йосс Ж., Джозеф Д.* Элементарная теория устойчивости и бифуркаций. — М.: Мир, 1983. 301 с.
91. *Каханер Д., Моулер К., Нэш С.* Численные методы и программное обеспечение. — М.: Мир, 1998. 575 с.
92. *Кац В. А.* // Письма в ЖТФ. 1984. Т. 10. С. 684–689.
93. *Ковалева Н. А., Лоскутов А. Ю.* // Доклады РАН. 2004. Т.396. С. 105–107.
94. *Козлов В. В.* Методы качественного анализа в динамике твердого тела. — М.: МГУ, 1980. 232 с.
95. *Козлов В. В.* // УМН. 1983. Т. 38. С. 3–67.
96. *Колмогоров А. Н.* // Докл. АН СССР. 1954. Т.98. С.527–530.
97. *Колмогоров А. Н.* // ДАН СССР. 1958. Т. 119. С. 861–864.
98. *Колмогоров А. Н.* // ДАН СССР. 1959. Т. 124. С. 754–755.
99. *Колмогоров А. Н., Тихомиров В. М.* // УМН. 1959. Т. 14. С. 3–86.

100. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. – М.: Наука, 1981. 544 с.
101. Корнфельд И. П., Синай Я. Г., Фомин С. В. Эргодическая теория. – М., Наука, 1980. 382 с.
102. Кроновер Р. М. Фракталы и хаос в динамических системах. Основы теории. – М.: Постмаркет, 2000. 352 с.
103. Крылов Н. С. Работы по обоснованию статистической физики. – М.-Л., АН СССР, 1950.
104. Кузнецов С. П. Динамический хаос. – М.: Физматлит, 2001. 296 с.
105. Кузнецов С. П. // Изв. вузов. Сер. «Радиофизика». 1985. Т. 28. С. 991–1007.
106. Кузнецов С. П. // Изв. вузов. Сер. «Радиофизика». 1986. Т. 29. С. 887–896.
107. Кузнецов С. П., Пиковский А. С. // Изв. вузов. Сер. «Радиофизика». 1985. Т. 28. С. 308–319.
108. Кузнецов С. П., Пиковский А. С. // Нелинейные волны. Структуры и бифуркации. – М.: Наука, 1987. С. 237–250.
109. Ладыженская О. А. // Зап. науч. сем. ЛОМИ. 1982. Т. 115. С. 137–155.
110. Ланда П. С. Автоколебания в системах с конечным числом степеней свободы. М.: Наука, 1980. 359 с.
111. Ланда П. С. Автоколебания в распределенных системах. – М.: Наука, 1983. 320 с.
112. Ландау Л. Д. // ДАН СССР. 1944. Т. 44. С. 339–342.
113. Ландау Л. Д., Лифшиц Е. М. Гидродинамика. – М.: Наука, 1988. 736 с.
114. Ланфорд О. Е. // Странные аттракторы. – М.: Мир, 1981 – С. 73.
115. Ланфорд О. Е. // Гидродинамические неустойчивости и переход к турбулентности. – М.: Мир, 1984. С. 22–46.
116. Лихтенберг А., Либерман М. Регулярная и стохастическая динамика. – М., Мир, 1984. 528 с.

117. Лоренц Э. // Странные аттракторы. – М.: Мир, 1981. С. 88–116.
118. Лоскутов А. Ю. Параметрическая дестохастизация систем со странным аттрактором спирального типа и механизмы перехода к динамической стохастичности. // Изд-во ВИНТИ. М.: 1987. – Деп. в ВИНТИ 7.07.87 №4802-В 87.
119. Лоскутов А. Ю. // Прикладная нелинейная динамика. 1994. Т. 2. С. 14–25.
120. Лоскутов А. Ю. // Вестн. Моск. ун-та. Сер. Физ.-Астр. 2001. No 3. С. 3–21.
121. Лоскутов А. Ю., Джаноев А. Р. // ЖЭТФ. 2004. Т.125. С. 1194–1203.
122. Лоскутов А. Ю., Томас Г. Э. // Вестн. Моск. ун-та. Сер. Физ.-Астр. 1993. Т. 34. С. 3–11.
123. Лоскутов А. Ю., Шишмарев А. И. // Успехи матем. наук. 1993. Т. 48. С. 169–170.
124. Лоскутов А. Ю., Журавлев Д. И., Котляров О. Л. // Синергетика-5. Москва–Ижевск. 2003. С. 115–128.
125. Лоскутов А. Ю., Истомин И. А., Котляров О. Л., Кузаян К. М. // Письма в Астр. журнал. 2001. Т. 27. С. 867–876.
126. Лоскутов А. Ю., Котляров О. Л., Истомин И. А., Журавлев Д. И. // Вестн. Моск. ун-та. Сер. Физ.-Астр. 2002. № 6. С. 3–21.
127. Лоскутов А. Ю., Миценко Ю. В., Рыбалко С. Д. // Физ. мысль России. 1997. С. 53–66.
128. Лоскутов А. Ю., Прохоров А. К., Рыбалко С. Д. // Теор. и матем. физика. 2002. Т. 132. С. 105–125.
129. Лоскутов А. Ю., Рыбалко С. Д., Удин Д. Н., Васильев К. А. // Теор. и матем. физика. 2000. Т. 124. С. 506–519.
130. Лоскутов А. Ю., Рыбалко С. Д. // Радиотехника и электроника. 2005. Т. 50. С. 1466–1475.
131. Лоскутов А. Ю., Рыбалко С. Д., Чураев А. А. // Письма в ЖТФ. 2004. Т. 30. С. 1–7.

132. Лоскутов А. Ю., Рябов А. Б., Акиншин Л. Г. // ЖЭТФ. 1999. Т. 116. С. 1781–1797.
133. Лоскутов А. Ю., Черемин Р. В., Высоцкий С. А. // Доклады РАН. 2005. Т. 404. С. 477–480.
134. Лукашук С. Н., Львов В. С., Предтеченский А. А., Черных А. И. // Проблемы нелинейных и турбулентных процессов в физике. – Киев: Наук. думка, 1985. Ч. II. С. 80–84.
135. Лукашук С. Н., Предтеченский А. А., Фалькович Г. Е., Черных А. И. Препринт Ин-та автоматизации и электрометрии СО АН СССР. № 280. – Новосибирск, 1985.
136. Лукьянов В. И., Шильников Л. П. // ДАН СССР. 1978. Т. 243. С. 26–29.
137. Львов В. С., Предтеченский А. А., Черных А. И. // ЖЭТФ. 1980. Т. 80. С. 1099–1121.
138. Любенски Т. С., Пинкус Ф. А. // Физика за рубежом. Сер. А. – М.: Мир, 1986. С. 147–164.
139. Любимов Д. В., Путин Г. Ф., Чернатынский В. И. // ДАН СССР. 1977. Т. 235. С. 554–556.
140. Макаренко Н. Г. // Нелинейные волны 2004. Ред. А. В. Гапонов–Грехов, В. И. Некоркин. Нижний Новгород: ИПФ РАН, 2005. С. 398–410.
141. Малинецкий Г. Г., Потапов А. Б. Современные проблемы нелинейной динамики. М.: УРСС, 2000. 336 с.
142. Мандельброт Б. // Странные аттракторы. – М.: Мир, 1981. С. 47–57.
143. Мандельброт Б. Фрактальная геометрия природы. – М.: Ин-т комп. иссл-ий, 2002. 656 с.
144. Мандельброт Б. Фракталы, случай и финансы. – R&C Dynamics, Москва, 2004. 256 с.
145. Марсден Дж., Мак-Кракен М. Бифуркация рождения цикла и ее приложения. – М.: Мир, 1980. 368 с.
146. Мельников В. К. // Труды Моск. матем. об-ва. 1963. Т. 12. С. 3–52.

147. Милнор Дж. Голоморфная динамика. – Ижевск: Регулярная и хаотич. динамика, 2000. 320 с.
148. Мозер Ю. // Математика. Период. сб. перев. ин. статей. 1962. Т. 6. С. 51–67.
149. Мозер Ю. Лекции о гамильтоновых системах. – М.: Мир, 1973. 168 с.
150. Мозер Ю. // УМН. 1981. Т. 36. С. 109–151.
151. Монин А. С. // УФН. 1978. Т. 125. С. 97–122.
152. Монин А. С. // УФН. 1986. Т. 150. С. 61–105.
153. Мольков Я. И., Фейгин А. М. // Нелинейные волны 2002. Ред. А. В. Гапонов–Грехов, В. И. Некоркин. Нижний Новгород: ИПФ РАН, 2003. С. 34–52.
154. Морозов А. Д. Введение в теорию фракталов. – Москва–Ижевск, Ин-т компьют. иссл-ий, 2002. 160 с.
155. Мун Ф. Хаотические колебания. – М., Мир, 1990. 312 с.
156. Неймарк Ю. И. Методы точечных отображений в теории нелинейных колебаний. – М.: Наука, 1972. 471 с.
157. Неймарк Ю. И. // Качественные методы исследования дифференциальных уравнений и нелинейных колебаний. – Киев: Ин-т матем. АН УССР, 1981. С. 99–115.
158. Неймарк Ю. И., Ланда П. С. Стохастические и хаотические колебания. // М.: Наука, 1987. 424 с.
159. Нейштадт А. И. // Успехи матем. наук. 1985. Т. 40. С. 190–191.
160. Нехорошев Н. Н. // Успехи матем. наук. 1977. Т. 32. С. 5–66.
161. Огнев А. И. // Матем. заметки. 1981. Т. 30. С. 723–736.
162. Отнес Р., Энноксон Л. Прикладной анализ временных рядов. – М.: Мир, 1982. 428 с.
163. Пайтген Х.-О., Рихтер П. Х. Красота фракталов. Образы комплексных динамических систем. – М.: Мир, 1993. 176 с.

164. *Пиковский А. С.* // Изв. вузов. Сер. «Радиофизика». 1986. Т. 29. С. 1438–1446.
165. *Постников М. М.* Гладкие многообразия. – М.: Наука, 1987. 480 с.
166. *Пригожин И.* От существующего к возникающему. – М.: Наука, 1985. 328 с.
167. *Пуанкаре А.* Новые методы небесной механики. – М.: Наука, 1971. 772 с.
168. *Пустьльников Л. Д.* // Докл. Акад. наук СССР. 1987. Т. 292. С. 549–553.
169. *Пустьльников Л. Д.* // Матем. сб. 1994. Т. 85. С. 113–124.
170. *Рабинович М. И., Фабрикант А. Л.* // ЖЭТФ. 1979. Т. 77. С. 617–629.
171. *Рожков В. А., Трапезников Ю. А.* Вероятностные модели океанологических процессов. – Л.: Гидрометеоздат, 1990. 272 с.
172. *Рюэль Д.* Случайность и хаос. – Москва–Ижевск: РХД, 2001. 192 с.
173. *Рюэль Д., Такенс Ф.* // Странные аттракторы. М.: Мир, 1981. С. 117–151.
174. *Самойленко А. М.* Элементы математической теории многочастотных колебаний. Инвариантные торы. – М.: Наука, 1987. 304 с.
175. *Сатаев Е. А.* // Успехи матем. наук. 1992. Т. 47. С. 147–202.
176. *Синай Я. Г.* // Докл. АН СССР. 1959. Т. 124. С. 768–771.
177. *Синай Я. Г.* // Докл. АН СССР. 1963. Т. 153. С. 1261–1264.
178. *Синай Я. Г.* // Успехи матем. наук. 1970. Т. 25. С. 141–192.
179. *Синай Я. Г.* // Нелинейные волны. / Ред. А. В. Гапонов-Грехов. – М.: Наука, 1979. С. 192–212.
180. *Синай Я. Г.* Современные проблемы эргодической теории. – М., Наука, 1995. 202 с.
181. *Смейл С.* // УМН. 1970. Т. 25. С. 113–185.
182. *Смирнов А. И.* // ЖЭТФ. 1986. Т. 90. С. 385–397.
183. *Смирнов А. И.* // ЖЭТФ. 1988. Т. 94. С. 185–193.

184. *Трепцев Д. В.* Введение в теорию возмущений гамильтоновых систем. – Фазис, Москва, 1998. 184 с.
185. *Труженникова В. М.* // Изв. вузов. Сер. «Радиофизика». 1962. Т. 5. С. 356–361.
186. *Федер Е.* Фракталы. – М.: Мир, 1991. 254 с.
187. *Фейгенбаум М.* // Успехи физ. наук. 1983. Т. 141. С. 343–374.
188. *Фове С., Либхабер А.* // Синергетика. – М.: Мир, 1984. С. 234–247.
189. *Хакен Г.* Синергетика. Иерархии неустойчивостей в самоорганизующихся системах и устройствах. – М.: Мир, 1985. 423 с.
190. *Хартман Ф.* Обыкновенные дифференциальные уравнения. Мир, М., 1970. 720 с.
191. *Хорн Р., Джонсон Ч.* Матричный анализ. – М., Мир, 1989. 656 с.
192. *Хэссард Б., Казаринов Н., Вэн И.* Теория и приложения бифуркации рождения цикла. – М.: Мир, 1985. 280 с.
193. *Четаев А. Н.* Нейронные сети и цепи Маркова. – М.: Наука, 1985. 128 с.
194. *Чириков Б. В.* // Атом. энергия. 1959. Т. 6. С. 630–633.
195. *Чириков Б. В.* Исследования по теории нелинейного резонанса. – Новосибирск, Препринт ИЯФ № 267, 1969.
196. *Шарковский А. Н.* // Укр. матем. журн. 1964. № 1. С. 61–71.
197. *Шарковский А. Н., Майстренко Ю. Л., Романенко Е. Ю.* Разностные уравнения и их приложения. – Киев: Наукова думка, 1986. 280 с.
198. *Шильников Л. П.* Теория бифуркаций и модель Лоренца // Кн.: Марсен Дж., Мак-Кракен М. Бифуркация рождения цикла и ее приложения. Добавление II – М.: Мир, 1980. С. 317–335.
199. *Шишкова М. А.* // Докл. АН СССР. 1973. Т. 209. С. 576–579.
200. *Шредер М.* Фракталы, хаос, степенные законы. – Москва–Ижевск: РХД, 2001. 528 с.
201. *Шустер Г.* Детерминированный хаос. Введение. – М.: Мир, 1988. 240 с.

202. Эрман Ж. П. // Синергетика. – М.: Мир, 1984. С. 190–219.
203. Эно М. // Странные аттракторы. – М.: Мир, 1981. С. 152–163.
204. Эрроусмит Д., Плейс К. Обыкновенные дифференциальные уравнения. Качественная теория с приложениями. – М.: Мир, 1986. 243 с.
205. де ла Яве Р. Введение в КАМ–теорию. – РХД, Москва–Ижевск, 2003. 176 с.
206. Ababrbanel H. D. I., Kennel M. B. // Phys. Rev. E. 1993. V. 47. P. 3057–3068.
207. Abarbanel H. D. I., Lindsay P. S. // IEEE Trans. Circuits Syst. 1993. V. 40. P. 643–645.
208. Abarbanel H. D. I., Brown R., Sidorowich J. J., Tsimring L. S. // Rev. Mod. Phys. 1993. V. 65. P. 1331–1393.
209. Abraham N. B., Albano A. M., Das B., et al. // Phys. Lett. 1986. V. 114A. P. 217–221.
210. Aceves A., Adachihara H., Jones C., et al. // Physica. 1986. V. 18D. P. 85–112.
211. Afraimovich V. S. // Nonlinear and Turbulent Processes in Physics. – N.Y.: Gordon and Breach, 1984. V. 3. P. 1133–1138.
212. Afraimovich V. S., Shilnikov L. P. // Nonlinear Dynamics and Turbulence. / Ed. G. I. Barenblatt, G. Iooss, D. D. Joseph. – New York: Pitman, 1983. P. 1–34.
213. Afraimovich V. S., Hsu S. Lectures on Chaotic Dynamical Systems. – AMS–Press, N.Y., 2003.
214. Alonso S., Sagues F., Mikhailov A. S. // Science. 2003. V. 299. P. 1722–1725.
215. Arneodo A., Collet P. H., Spiegel E. A. // Phys. Lett. 1983. V. 94A. P. 1–4.
216. Andreyev Yu. V., Dmitriev A. S., Starkov S. O. // IEEE Trans. 1997. V. CS-44. P. 21–28.
217. Arecchi F. T. // Instabilities and Chaos in Quantum Optics / Eds F. T. Arecchi, R. G. Harrison. – Berlin: Springer, 1987.

218. Argoul F., Arneodo A., Roux J. C. // J. Chem. Phys., 1987. V. 86. P. 3325–3338.
219. Arnold V. I. Geometrical Methods in the Theory of Ordinary Differential Equations. – Berlin: Springer, 1983. [Русский перевод: № 27].
220. Arnold V. I., Avez A. Ergodic Problems of Classical Mechanics. – N.Y.: Benjamin, 1968. 286 p. [Русский перевод: № 26].
221. Aronson D. G., Weinberger H. F. // Lect. Notes in Math. 1975. V. 446. P. 5–49.
222. Babloyantz A., Salazar J. N., Nicolis G. // Phys. Lett. 1985. V. 111A. P. 152–156.
223. Bak P. // Physics Today. 1986. V. 39. P. 38–45.
224. Bakker L. E., Diacu F. N. // Romanian Astron. J. 1993. V. 3. P. 139–155.
225. Baldwin P. R. // J. Phys. A. 1991. V. 24. P. L941–L947.
226. Bär M., Eiswirth M. // Phys. Rev. E. 1993. V. 48. P. 1635–1637.
227. Benettin G., Froschle C., Scheidecker J.-P. // Phys. Rev. 1979. V. 19A. P. 2454–2460.
228. Benettin G., Galgani L., Giorgilli A., Strelcyn J. M. // Meccanica. 1980. V. 15. P. 9–31.
229. Benettin G., Galgani L., Strelcyn J. M. // Phys. Rev. 1976. V. 14A. P. 2338–2345.
230. Ben–Mizrachi A., Procaccia I., Grassberger P. // Phys. Rev. 1984. V. 29A. P. 975–977.
231. Berge P. // CLEFS CEA (France). 1987. № 4. P. 2–12.
232. Berge P., Dubois M., Manneville P., Pomeau Y. // J. Phys. Lett. (Paris). 1980. V. 41. P. L341–L345.
233. Berge P., Pomeau Y., Vidal C. Order Within Chaos. – N.Y.: Wiley, 1986. [Русский перевод: П. Берге, И. Помо, К. Вудаль. Порядок в хаосе. – М.: Мир, 1991. – 368 с.]

234. *Berry M. V.* Regular and irregular motion. — AIP Conf. Proc. No46. Topics on Nonlinear Dynamics. AIP, New York, 1978. P. 16–120.
235. Bibliography of Dynamical Systems, compiled by K. Shiraiva. — Nagoya Univ.: Preprint №1. 1985. — 390 p.
236. *Biswas D. J., Vas Dev., Chatterjee U. K.* // Phys. Rev. 1988. V. 38A. P. 555–558.
237. *Biswas D. J., Vas Dev., Chatterjee U. K.* // Phys. Rev. 1987. V. 35A. P. 456–458.
238. *Boccaletti S., Grebogi C., Lai Y.-C., Mancini H., Maza D.* // Phys. Rep. 2000. V. 329. P. 103–197.
239. *Bountis T. C.* // Physica. 1981. V. 3D. P. 577–589.
240. *Brahic A.* // Astron. Astrophys. 1971. V. 12. P. 98–110.
241. *Brandstater A., Swift J., Swinney H. L., Wolf A.* // Phys. Rev. Lett. 1983. V. 51. P. 1442–1445.
242. *Bresler L., Metcalfe G., Ottino J. M., Shinbrot T.* // Chem. Eng. Sci. 1996. V. 58. P. 1671–1679.
243. *Broomhead D. S., Jones R.* // Proc. Roy. Soc. London. 1989. V. 423. P. 103–121.
244. *Broomhead D. S., Jones R., King G. P.* // J. Phys. 1987. V. 20A. P. L563–L569.
245. *Broomhead D. S., King G. P.* // Physica D. 1986. V. 20. P. 217–236.
246. *Broomhead D. S., King G. P.* // Nonlinear Phenomena and Chaos. Eds. S. Sarkar. Bristol, Adam Hilger. 1986. P. 113–144.
247. *Brown R. G.* Smoothing, Forecasting and Prediction of Discrete Time-Series. — New Jersey, Prentice-Hall, 1962.
248. *Bunde A., Havlin S.* (eds.) Fractals in Science. — Springer, Berlin, 1994. 304 p.
249. *Bunimovich L. A.* // Commun. Math. Phys. 1979. V. 65. P. 295–312.
250. *Bunimovich L. A.* // Chaos. 1991. V. 1. P. 187–193.

251. *Bunimovich L. A.* // Physica D. 1995. V. 86. P. 248–255.
252. *Bunimovich L. A., Sinai Ya. G.* // Commun. Math. Phys. 1981. V. 78. P. 479–497.
253. *Bunimovich L. A., Sinai Ya. G.* // Nonlinearity. 1988. V. 1. P. 491–504.
254. *Carrido R., Simó L.* // Springer Lect. Notes in Phys. — Springer, Berlin, 1983. V. 179. P. 1–28.
255. *Casati G., Chirikov B. V., Ford J.* // Phys. Lett. 1980. V. 77A. P. 91–94.
256. *Casdagli M.* // Physica D. 1989. V. 35. P. 335–356.
257. *Chacón R.* // Phys. Rev. Lett. 2001. V. 86. P. 1737–1740.
258. *Chaos.* // 1992. V. 2. No 3. Номер, посвященный анализу сцепленных отображений.
259. *Chate H., Manneville P.* // Phys. Rev. Lett. 1987. V. 58. P. 112–115.
260. *Chen G., Dong X.* // Int. J. Bifurcation and Chaos. 1993. V. 3. P. 1363–1409.
261. *Chernov N.* // J. Stat. Phys. 1997. V. 88. P. 1–29.
262. *Chierchia L., Gallavotti G.* // Annales de l'Institut Poincaré B. 1994. V. 60. P. 1–144.
263. *Chirikov B. V.* // Phys. Rep. 1979. V. 52. P. 263–379.
264. *Chirikov B. V., Izrailev F. M.* // Physica. 1981. V. 2D. P. 30–37.
265. *Chow S. N., Hale J. K.* Methods of Bifurcation Theory. — Berlin: Springer, 1982. 515 p.
266. *Ciliberto S., Bigazzi P.* // Phys. Rev. Lett., 1988. V. 60. P. 286–289.
267. *Collet P., Eckmann J.-P., Koch H.* // Physica D. 1981. V. 3. P. 457–467.
268. *Cohen A., Procaccia I.* // Phys. Rev. 1984. V. 31A. P. 1872–1882.
269. *Collet P., Eckman J.-P., Lanford O. E.* // Commun. Math. Phys. 1980. V. 76. P. 211–254.
270. *Collet P.* // Phys. Rep. 1984. V. 103. P. 121–123.

271. *Collet P., Eckmann J.-P.* Iterated Maps on the Interval as Dynamical Systems. – Boston: Birkhauser, 1980. 248 p.
272. *Collet P., Elphick C., Repaux D.* // Phys. Rev. Lett. 1987. V. 58. P. 431–434.
273. *Constantin P., Foias C.* // Commun. Pure Appl. Math. 1985. V. 38. P. 1–27.
274. *Constantin P., Foias C., Manley O. P., Temam R.* // J. Fluid Mech., 1985. V. 15. P. 427–440.
275. *Constantin P., Foias C., Temam R.* // Mem. Amer. Math. Soc. 1985. V. 53. 67 p.
276. *Coste J., Peyrand N.* // Physica. 1982. V. 5D. P. 415–420.
277. *Crutchfield J. P., Farmer J. D., Packard N., Shaw R., Jones G., Donnelly R. J.* // Phys. Lett. 1980. V. 76A. P. 1–4.
278. *Crutchfield J., Nauenberg M., Rudnick J.* // Phys. Rev. Lett. 1981. V. 46. P. 933–935.
279. *Cuadros F., Chacón R.* // Phys. Rev. E. 1993. V. 47. P. 4628–4629.
280. *Cumming A., Lindsay P. S.* // Phys. Rev. Lett. 1988. V. 60. P. 2719–2722.
281. *Curry J. H.* // J. Stat. Phys. 1981. V. 26. P. 683–695.
282. *Cvitanovic P., Gunaratne G. H., Procaccia I.* // Phys. Rev. 1988. V. 38A. P. 1503–1520.
283. *Dadio H.* // Progr. Theor. Phys. 1983. V. 70. P. 879–882.
284. *Davies M. E.* // Physica D. 1997. V. 101. P. 195–206.
285. *Deissler R. J., Ecke R. E., Haucke H.* // Phys. Rev. 1987. V. 36A. P. 4390–4401.
286. *Destexhe A., Sepulchre J. A., Babloyantz A.* // Phys. Lett. 1988. V. 132A. P. 101–106.
287. *Devaney R. L.* An Introduction to Chaotic Dynamical Systems. 2nd edition. – Mass.: Addison-Wesley, 1989. 336 p.
288. *Dmitriev A. S., Kassian G., Khilinsky A.* // Int. J. Bifurcation and Chaos. 2000. V. 10. P. 749–761.

289. *Doering C. R., Gibbon J. D., Holm D. D., Nicolaenko B.* // Nonlinearity. 1988. V. 1. P. 279–309.
290. *Douady A., Hubbard J. L.* // Compt. Rend. Acad. Sci. Paris. 1982. V. 294. P. 123–126.
291. *Dubois M.* Experimental aspects of the transition to turbulence in Rayleigh–Benard convection. – Berlin: Springer, 1982. V. 164. P. 177–191.
292. *Duffing G.* // Erzwungene Schwingungen bei Veränderlicher Eigenfrequenz. – F. Vieweg und Sohn, Braunschweig, 1918.
293. *Dvorak I., Siska J.* // Phys. Lett. 1986. V. 118A. P. 63–66.
294. *Eckmann J. -P., Ruelle D.* // Rev. Mod. Phys. 1985. V. 57. P. 617–656.
295. *Eiswirth M., Krischer K., Ertl G.* // Surface Sci. 1988. V. 202. P. 565–591.
296. *Escande D. F.* // Phys. Rep. 1985. V. 121. P. 165–261.
297. *Farmer J. D.* // Physica. 1982. V. 4D. P. 366–393.
298. *Farmer J. D., Ott E., Yorke J. A.* // Physica. 1983. V. 7D. P. 153–180.
299. *Farmer J. D., Sidorowich J. J.* // Phys. Rev. Lett. 1987. V. 59. P. 845–849.
300. *Farmer J. D., Sidorowich J. J.* // Evolution, Learning, and Cognition, ed. Y. C. Lee, Singapore, World Scientific Press. 1988. P. 277–302.
301. *Fatou P.* // Bull. Soc. Math. Fr. 1919. V. 47. P. 161–271. 1920. V. 48. P. 33–94; 208–314.
302. *Feigenbaum M. J.* // J. Stat. Phys. 1978. V. 19. P. 25–52.
303. *Feigenbaum M. J.* // J. Stat. Phys. 1979. V. 21. P. 669–706.
304. *Feigenbaum M. J.* // Phys. Lett. A. 1979. V. 74. P. 375–378.
305. *Feigenbaum M. J.* // Commun. Math. Phys. 1980. V. 77. P. 65–86.
306. *Feigenbaum M. J., Hasslacher B.* // Phys. Rev. Lett. 1982. V. 49. P. 605–609.
307. *Feigenbaum M. J., Kadanoff L. P., Shenker S. J.* // Physica D. 1982. V. 5. P. 370–386.
308. *Fermi E.* // Phys. Rev. 1949. V. 75. P. 1169–1174.

309. *Franses P. H.* Time series models for business and economic forecasting. — Cambridge Univ. Press, 1998. 280 p.
310. *Fraser A. M., Swinney H. L.* // Phys. Rev. A. 1986. V. 33. P. 1134–1140.
311. *Frederickson P., Kaplan J. L., Yorke E. D., Yorke J. A.* // J. Diff. Equ., 1983. V. 49. P. 185–201.
312. *Froeschle C., Benettin G., Scheidecker J.-P.* // Intrinsic Stochasticity in Plasmas. — Cargese, 1979. P. 93–114.
313. *Fronzoni L., Geocondo M., Pettini M.* // Phys. Rev. A. 1991. V. 43. P. 6483–6487.
314. *Fundamental Problems in Statistical Mechanics, vol. 3.* Ed. R. H. Cohen. — Elsevier, Amsterdam, 1975.
315. *Galgani L., Giorgilli A., Strelcyn J.-M.* // Nuovo Chimento. 1981. V. 61B, Ser. 2. P. 1–20.
316. *Garfinkel A., Spano M. L., Ditto W. L.* // Science. 1992. V. 257. P. 1230–1235.
317. *Garrido P. L.* // J. Stat. Phys. 1997. V. 88. P. 807–824.
318. *Ghil M., Allen R. M., Dettinger M. D., et al.* // Rev. Geophys. 2002. V. 40. P. 1–41.
319. *Gielis G., MacKay R. S.* // Nonlinearity. 2000. V. 13. P. 867–888.
320. *Gioggia R. S., Abraham N. B.* // Phys. Rev. Lett. 1983. V. 51. P. 650–653.
321. *Giglio G., Musazzi S., Perini U.* // Phys. Rev. Lett. 1981. V. 47. P. 243–246.
322. *Glans J.* // Science. 1994. V. 265. P. 474–475.
323. *Gollub J. P., Benson S. V., Steinman J.* // Nonlinear Dynamics / Ed. R. H. G. Helleman. — N. Y.: Ann. N. Y. Acad. Sci., 1980. V. 357. P. 22–27.
324. *Gollub J. P., Swinney H. L.* // Phys. Rev. Lett. 1975. V. 35. P. 927–930.
325. *Golyandina N., Nekrutkin V., Solntsev V.* «Caterpillar»-SSA technique for analysis of time series in economics. — <http://vega.math.spbu.ru/caterpillar>
326. *Grassberger P., Procaccia I.* // Phys. Rev. Lett. 1983. V. 50. P. 346–349.

327. *Grassberger P., Procaccia I.* // Physica D. 1983. V. 9. P. 189–208.
328. *Grassberger P., Procaccia I.* // Phys. Rev. 1983. V. 29A. P. 2591–2593.
329. *Grebogi C., Ott E., Yorke J. A.* // Phys. Rev. Lett. 1982. V. 48. P. 1507–1510.
330. *Grebogi C., Ott E., Yorke J. A.* // Physica. 1983. V. 7D. P. 181–200.
331. *Grebogi C., Ott E., Yorke J. A.* // Phys. Rev. Lett. 1983. V. 51. P. 339–342.
332. *Grebogi C., Ott E., Yorke J. A.* // Physica D. 1985. V. 15. P. 354–373.
333. *Grebogi C., Ott E., Pelikan S., Yorke J. A.* // Physica D. 1984. V. 13. P. 261–268.
334. *Greene J. M., Mackay R. S., Vivaldi F., Feigenbaum M. J.* // Physica D. 1981. V. 3. P. 468–486.
335. *Greenside H. S., Wolf A., Swift J., Pignataro T.* // Phys. Rev. 1982. V. 25A. P. 3453–3456.
336. *Gouyet J.-F.* Physics of Fractal Structures. — Springer-Verlag, Berlin, 1996. 234 p.
337. *Grossmann S., Thomae S.* // Z. Naturforsch A. 1977. Bd. 32. S. 1353–1363.
338. *Guckenheimer J., Holmes P.* Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields. — Berlin: Springer, 1983. 453 p. [Русский перевод: № 72]
339. *Gumowski I., Mira C.* Dynamique Chaotique. — Toulouse: Cepadues, 1980.
340. *Gurel O.* // Phys. Lett. 1977. V. 61A. P. 219–223.
341. *Gustavson F.* // Astron. J. 1966. V. 71. P. 670–686.
342. *Hadarnard J.* // J. Math. Pures et Appl. 1898. V. 4. P. 27–73.
343. *Halsey T. S., Jensen M. H., Kadanoff L. P., Procaccia I., Shraiman B. I.* // Phys. Rev. 1986. V. 33A. P. 1141–1151.
344. *Hayes S., Grebogi C., Ott E.* // Phys. Rev. Lett. 1993. V. 70. P. 3031–3034.
345. *Hayes S., Grebogi C., Ott E., Mark A.* // Phys. Rev. Lett. 1994. V. 73. P. 1781–1784.

346. *Helleman R. H. G.* // Fundamental Problems in Statistical Mechanics. Ed. E. G. D. Cohen. 1980. V. 5. P. 165–233.
347. *Helleman R. H. G., Bountis T.* // Stochastic Behaviour in Classical and Quantum Hamiltonian Systems / Eds G. Casati, J. Ford. – N. Y.: Springer, 1979. V. 93. P. 353.
348. *Henon M.* // Physica. 1982. V. 5D. P. 412–414.
349. *Henon M., Heiles C.* // Astron. J. 1964. V. 69. P. 73–79.
350. *Hentschel H. G. E., Procaccia I.* // Physica D. 1983. V. 8. P. 435–444.
351. *Hirsch J. E., Huberman B. A., Scalapino D. J.* // Phys. Rev. 1982. V. 25A. P. 519–532.
352. *Holmes P. J.* // SIAM J. Appl. Math. 1980. V. 38. P. 65–80.
353. *Holmes P.* // Physica. 1982. V. 5D. P. 335–347.
354. *Holmes P. J., Marsden J. E.* // J. Math. Phys. 1982. V. 23. P. 669–675.
355. *Holmes P. J., Marsden J. E.* // Commun. Math. Phys. 1982. V. 82. P. 523–544.
356. *Hopf E.* // Commun. Pure and Appl. Math. 1948. V. 1. P. 303–322.
357. *Hu B.* // Chaos and Statistical Methods. / Ed. Y. Kuramoto. — Berlin: Springer, 1984. P. 72–82.
358. *Hu B., Mao J. M.* // Phys. Rev. 1982. V. 25A. P. 3259–3261.
359. *Hu B., Rudnick J.* // Phys. Rev. Lett. 1982. V. 48. P. 1645–1648.
360. *Hu B., Satija I. I.* // Phys. Lett. 1983. V. 98A. P. 143–146.
361. *Huang J. -Y., Kim J. -J.* // Phys. Rev. 1987. V. 36A. P. 1495–1497.
362. *Huberman B. A., Crutchfield J. P.* // Phys. Rev. Lett. 1979. V. 43. P. 1743–1747.
363. *Huberman B. A., Crutchfield J. P., Packard N. H.* // Appl. Phys. Lett. 1980. V. 37. P. 750–752.
364. *Hübler A.* // Helv. Phys. Acta. 1989. V. 62. P. 343–346.

365. *Hübler A., Georgii R., Kuckler M., Stelzl W., Lüsher E.* // Helv. Phys. Acta. 1988. V. 61. P. 897–900.
366. *Hyman J. M., Nicolaenko B.* // Physica. 1986. V. 18D. P. 113–126.
367. *Hyman J. M., Nicolaenko B., Zaleski S.* // Physica. 1986. V. 23D. P. 265–292.
368. *Jensen R. V.* // Am. Sci. 1987. V. 75. P. 168–181.
369. *Jensen M. H., Bak P., Bohr T.* // Phys. Rev. Lett. 1983. V. 50. P. 1637–1639.
370. *Jiang M., Pesin Ya. B.* // Commun. Math. Phys. 1998. V. 193. P. 677–711.
371. *Johnston J., DiNardo J.* Econometric methods. 4th Edition. McGraw-Hill, 1997. 531 p.
372. *Julia G.* // J. Math. Pures et Appl. 1918. V. 8. P. 47–245
373. *Just W.* // J. Stat. Phys. 1995. V. 79. P. 429–449.
374. *Kaneko K.* // Progr. Theor. Phys. 1984. V. 72. P. 202–215.
375. *Kaneko K.* Collapse of Tori and Genesis of Chaos in Dissipative Systems. — World Sci., Singapore, 1986.
376. *Kaneko K.* // Physica D. 1989. V. 34. P. 1–41.
377. *Kaneko K.* // Physica D. 1989. V. 37. P. 60–82.
378. *Kaneko K.* // Physica D. 1990. V. 41. P. 137–172.
379. *Kaneko K.* // Physica D. 1991. V. 54. P. 5–19.
380. *Kaneko K., Willeboords F.* Bifurcation and Spatial Chaos in Open Flow Model. — <http://xxx.lanl.gov/abs/chao-dyn/9312007>.
381. *Kaneko K., Willeboords F.* Self-Organized Periodic Lattices of Chaotic Defects. — <http://xxx.lanl.gov/abs/chao-dyn/9405007>.
382. *Kantz H., Grassberger P.* // J. Phys. A. 1988. V. 21. P. L127–L133.
383. *Kaplan J. L., Yorke J. A.* // Lect. Notes in Math. 1979. №730. P. 228–237.
384. *Keppenne C. L., Ghil M.* // J. Geophys. Res. 1992. V. 97. P. 20449–20554.
385. *Kim M., Bertran M., Pollmann M. et al.* // Science. 2001. V. 292. P. 1357–1360.

386. *Kivshar Yu. S., Rödelberger B., Benner H.* // Phys. Rev. E. 1994. V.49. P. 319–324.
387. *Kolumbán G.* // IEEE Trans. 2000. V. CS-47. P. 1692–1701.
388. *Kostelich E., Grebogi C., Ott E., Yorke J. A.* // Phys. Rev. E. 1993. V.47. P. 305–310.
389. *Krüger T., Pustyl'nikov L. D., Troubetzkoy S. E.* // Nonlinearity. 1995. V. 8. P. 397–410.
390. *Kugiumtzis D., Lingjoerde O. C., Christophersen N.* // Physica D. 1998. V. 112. P. 344–360.
391. *Ledrappier F.* // Commun. Math. Phys. 1981. V. 81. P. 229–238.
392. *Li T. -Y., Yorke J.* // Am. Math. Monthly. 1975. V. 82. P. 985–992.
393. *Libhaber A., Maurer J.* // J. Phys. Colloq. (France). 1980. V. 41. P. 51–56.
394. *Lichtenberg A. J., Lieberman M. A., Cohen R. H.* // Physica D. 1980. V. 1. P. 291–305.
395. *Lieberman M. A.* // Nonlinear Dynamics / Ed. R. H. G. Helleman. – N. Y.: Ann. N. Y. Acad. Sci., 1980. V. 357. P. 119–142.
396. *Lieberman M. A., Tennyson J. L.* // Long-Time Prediction in Dynamics / Eds Horton et al. – N. Y.: Wiley, 1983. P. 179–211.
397. *Lima R., Pettini M.* // Phys. Rev. A. 1990. V. 41. P. 726–733.
398. *Linder J. F., Ditto W. L.* // Appl. Mech. Rev. 1995. V. 48. P. 795–807.
399. *Linsay P. S.* // Phys. Rev. Lett. 1981. V. 47. P. 1349–1352.
400. *Loskutov A.* // Comput. Math. and Modeling. 2001. V. 12. P. 314–352.
401. *Loskutov A., Andrievsky D., Ivanov V. and Ryabov A.* // Emergent Nature. Ed. M. M. Novak. – World Scientific, 2002. P. 263–272.
402. *Loskutov A., Istomin I. A., Kuzanyan K. M., Kotlyarov O. L.* // Nonlin. Phenom. in Complex Syst. 2001. V. 4. P. 47–57.
403. *Loskutov A., Tereshko V. M.* // Artificial Neural Networks. Eds. I. Alexander and J. Taylor. – Elsevier, North-Holland, 1992. P. 449–452.

404. *Loskutov A., Tereshko V. M.* // SPIE. 1993. V. 2038. P. 263–272.
405. *Loskutov A., Tereshko V. M., Vasiliev K. A.* // Int. J. Neural Systems. 1995. V. 6. P. 175–182.
406. *Loskutov A., Thomas G. E.* // SPIE. 1993. V. 2037. P. 238–249.
407. *Loskutov A., Ryabov A.* // J. Stat. Phys. 2002. V. 108. P. 995–1014.
408. *Loskutov A., Shishmarev A. I.* // Chaos. 1994. V. 4. P. 351–355.
409. *Loskutov A., Ryabov A. B., Akinshin L. G.* // J. Phys. A. 2000. V. 33. P. 7973–7986.
410. *Loskutov A., Schwalger T., Dzhanoev A.* // Chaos. 2006. V.16. P. 023109.
411. *Lüscher E., Hübler A.* // Helv. Phys. Acta. 1989. V. 62. P. 544–551.
412. *Mackey M. C., Glass L.* // Science. 1977. V. 197. P. 287–289.
413. *Mac-Kay R. S., Meiss J. D., Pervival I. C.* Transport in Hamiltonian Systems. – L.: Queen Mary College, 1983. 81 p.
414. *Malraison B., Atten P., Berge P., Dubois M.* // J. Phys. Lett. (Paris). 1983. V. 44. P. L897–L902.
415. *Mandelbrot B.* Les objets fractals: forme, hasard et dimension. – Paris: Flammarion, 1975. 192 p.
416. *Mandelbrot B.* Fractals: Form, Chance and Dimension. – San Francisco, CA: W. H. Freeman and Company, 1977. 265 p.
417. *Mandelbrot B. B.* // Ann. N. Y. Acad. Sci. 1980. V. 357. P. 249–259.
418. *Mane R.* // Lect. Notes in Math. Berlin, Springer. 1981. V. 898. P. 230–242.
419. *Manneville P.* // Phys. Rep. 1984. V. 103. P. 107–119.
420. *Manneville P., Pomeau Y.* // Phys. Lett. 1979. V. 75A. P. 1–2.
421. *Manneville P., Pomeau Y.* // Physica 1980. V. 1D. P. 219–226.
422. *Maurer J., Libhaber A. J.* // J. Phys. Lett (Paris). 1980. V. 41. P. L515–L518.
423. *Meakin P.* // Phys. Rev. Lett. 1983. V. 51. P. 1119–1122.

424. *Meakin P., Coniglio A., Stanley H.E., Witten T.A.* // Phys. Rev. A. 1986. V. 34. P. 3325–3340.
425. *Meakin P., Stanley H.E., Coniglio A., Witten T.A.* // Phys. Rev. A. 1985. V. 32. P. 2364–2369.
426. *de Melo W., van Strien S.* One-Dimensional Dynamics. — Springer: Berlin, 1993. 630 p.
427. *Meyer-Kress G., Haken H.* // Phys. Lett. 1981. V. 82A. P. 151–155.
428. *Mills T. C.* The econometric modeling of financial time series. — Cambridge Univ. Press, 1999. 380 p.
429. *Milnor J.* // Commun. Math. Physics, 1985. V. 99. P. 177–196.
430. *Misiurewicz M.* // Publ. Math. I. H. E. S. 1981. V. 53. P. 17–51.
431. *Miura T., Kai T.* // J. Phys. Soc. Japan. 1986. V. 55. P. 2562–2571.
432. *Moon F. C.* Chaotic Vibrations. — N. Y.: Wiley, 1987. 309 p. [Русский перевод: Ф. Мун. *Хаотические колебания*. — М.: Мир, 1990. — 312 с.]
433. *Morita S.* // Phys. Lett. A. 1997. V. 226. P. 172–178.
434. *Muñuzuri A. P. et al.* // Int. J. Bif. and Chaos. 1995. V. 5. P. 17–50.
435. *Namajunas A., Pozela J., Tamasevicius A.* // Phys. Lett. 1988. V. 131A. P. 85–90.
436. *Nauenberg M., Rudnick J.* // Phys. Rev. 1981. V. 24B. P. 493–495.
437. *Nese J. M.* // Lect. Notes in Phys. 1987. V. 271. P. 412–443.
438. *Newhouse S.* // Nonlinear Dynamics / Ed. R. H. G. Helleman. — N. Y.: Ann. N. Y. Acad. Sci. 1980. V. 357. P. 292–299.
439. *Newhouse S.* // Dynamical Systems. — Boston: Birkhauser, 1980. V. 8. P. 2–114.
440. *Newhouse S., Ruelle D., Takens F.* // Commun. Math. Phys. 1978. V. 64. P. 35–40.
441. *Nicolaenko B., Scheurer B., Temam R.* // Physica. 1985. V. 16D. P. 155–183.
442. *Niederman L.* // Nonlinearity. 1996. V. 9. P. 1703–1751.

443. *Nonnenmacher T.F., Losa G.A., Weibel E.R.* (ed.) Fractals in Biology and Medicine. — Springer, Berlin, 1994. G. A. Losa, D. Merlini, T. F. Nonnenmacher, E. R. Weibel (eds). Fractals in Biology and Medicine. Vol. I–VII. — Basel, Birkhauser, 1996–2003.
444. *Osipov G. V., Collins J. J.* // Phys. Rev. E. 1999. V. 60. P. 54–57.
445. *Ott E., Spano M. L.* // Physics Today. 1995. V. 48. P. 34–40.
446. *Ott E., Grebogi C., Yorke J. A.* // Phys. Rev. Lett. 1990. V. 64. P. 1196–1199.
447. *Packard N. H., Crutchfield J. P., Farmer J. D., Shaw R. S.* // Phys. Rev. Lett. 1980. V. 45. P. 712–716.
448. *Paladin G., Vulpiani A.* // Phys. Rep. 1987. V. 156. P. 148–225.
449. *Paramio M.* // Phys. Lett. 1988. V. 132A. P. 98–100.
450. *Parlitz U., Lauterborn W.* // Phys. Rev. 1987. V. 36A. P. 1428–1434.
451. *Pawelzik K., Schuster H. G.* // Phys. Rev. 1987. V. 35A. P. 481–484.
452. *Peitgen H.-O., Prüfer M., Richter P. H.* // Lotka–Volterra Approach to Cooperation and Competition in Dynamical Systems. / Eds. W. Ebeling, M. Peschel. — Akademie-Verlag, Berlin 1985. P. 81–102.
453. *Perez G., Sinha S., Cerdeira H.* // Phys. Rev. E. 1996. V. 54. P. 6936–6939.
454. *Peters E. E.* Fractal Market Analysis: Applying Chaos Theory to Investment and Economics. — John Wiley and Sons, 1994. 336 p.
455. *Peters E. E.* Patterns in the Dark: Understanding Risk and Financial Crisis with Complexity Theory. — John Wiley and Sons, 1999. 240 p.
456. *Peterson I.* Newton's Clock — Chaos in the Solar System. — Freeman, New York, 1993. 317 p.
457. *Pettini M.* // Dynamics and Stochastic Processes. Ed. R. Lima, L. Streit, R. Vilela Mendes. — Springer, Berlin, 1990. P. 242–250.
458. *Poincaré H. J.* // Rendicont del Circolo Matematico di Palermo. 1912. V. 33. P. 375–407.
459. *Pomeau Y.* // USMG NATO ASI. Les Nouches Session XXXVI on Chaotic Behaviour of Deterministic Systems. — Amsterdam, 1983. P. 609–619.

460. *Pomeau Y., Manneville P.* // Commun. Math. Phys. 1980. V. 74. P. 189–197.
461. *Proc. of the SPIE 1993 Annual Meeting «Chaos in Communications».* // San Diego, California, 11-16 July, 1993. V. 2038.
462. *Prokhorov M. D., Ponomarenko V. I., Karavaev A. S., Bezruchko B. P.* // Physica D. 2005. V. 203. P. 209–223.
463. *Rand D.* // «New Directions in Dynamical Systems.» / Eds. T. Bedford, J. Swift. — Cambridge, Cambridge Univ. Press., 1987. P. 1–56.
464. *Rand D., Ostlund S., Sethna J., Siggia E. D.* // Physica D. 1983. V. 8. P. 303–312.
465. *Richetti P., De Kepper P., Roux J. C., Swinney H. L.* // J. Stat. Phys. 1987. V. 48. P. 977–990.
466. *Richetti P., Roux J. C., Argoul F., Arneodo A.* // J. Chem. Phys. 1987. V. 86. P. 3339–3356.
467. *Robbins K. A.* // SIAM J. Appl. Math. 1979. V. 36. P. 457–472.
468. *Romeiras F. J., Ott E., Grebogi C., Dayawansa W. P.* // Physica D. 1992. V. 58. P. 165–192.
469. *Rossler O. E.* // Phys. Lett. 1976. V. 57A. P. 397–398.
470. *Roux J.-C.* // Physica. 1983. V. 7D. P. 57–68.
471. *Roux J.-C., Simoyi R. H., Swinney H. L.* // Physica. 1983. V. 8D. P. 257–266.
472. *Ruelle D.* // Math. Intelligencer. 1980. V. 2. P. 126–137.
473. *Ruelle D.* // Physica. 1983. V. 7D. P. 40–44.
474. *Ruelle D.* // Phys. Rep. 1984. V. 103. P. 81–85.
475. *Ruelle D.* Chaotic Evolution and Strange Attractors. — London: Cambridge Univ. Press. 1988. 120 p.
476. *Russel D. A., Hanson J. D., Ott E.* // Phys. Rev. Lett. 1980. V. 45. P. 1175–1178.
477. *Saito N., Hirooka H., Ford J. et al.* // Physica. 1982. V. 5D. P. 273–286.
478. *Salzman B.* // Journ. Atmos. Sci. 1962. V. 19. P. 329–341.

479. *Sano M., Savada Y.* // Chaos and Statistical Methods. / Ed. Y. Kuramoto. — Berlin: Springer, 1984. P. 226–231.
480. *Sato S., Sano M., Savada Y.* // Progr. Theor. Phys. 1987. V. 77. P. 1–5.
481. *Shenker S. J.* // Physica D. 1982. V. 5. P. 405–411.
482. *Shaw R.* // Z. Naturforsch A. 1981. V. 36. P. 80–112.
483. *Schiff S. J., Jerger K., Duong D. H., Chang T., Spano M. L., Ditto W. L.* // Nature. 1994. V. 370. P. 615–620.
484. *Shilnikov L.* // Proceedings of the International Conference on Nonlinear Dynamics, Chaotic and Complex Systems, Zakopane, 1995. / Eds. E. Infeld, R. Zelazny, and A. Galkowski. — Cambridge Univ. Press, Cambridge, 1997. Vol. 363. P. 39–63.
485. *Shimada I., Nagascima T.* // Progr. Theor. Phys. 1979. V. 61. P. 1605–1616.
486. *Shinbrot T.* // Nonlinear Sci. Today. 1993. V. 3. P. 1–8.
487. *Shinbrot T., Grebogi C., Ott E., Yorke J. A.* // Nature. 1993. V. 363. P. 411–417.
488. *Shinbrot T., Ott E., Grebogi C., Yorke J. A.* // Phys. Rev. Lett. 1990. V. 65. P. 3215–3218.
489. *Shinbrot T., Grebogi C., Ott E., Yorke J. A.* // Phys. Lett. A. 1992. V. 169. P. 349–354.
490. *Shinbrot T., Ott E., Grebogi C., Yorke J. A.* // Phys. Rev. A. 1992. V. 45. P. 4165–4168.
491. *Shinbrot T., Ottino J. M.* // Phys. Rev. Lett. 1993. V. 71. P. 843–847.
492. *Shinbrot T.* // Adv. Phys. 1995. V. 44. P. 73–111.
493. *Shumway R., Stoffer D. S.* Time Series Analysis and its Applications. — Springer, 2000. 549 p.
494. *Sieber M.* // Phys. Lett. 1987. V. 122A. P. 467–470.
495. *Sivakumar B., Jayawardena A. W.* // Environ. Modelling and Software. 2003. V. 18. P. 831–838.

496. *Solé R. V., Menéndez de la Prida L.* // Phys. Lett. A. 1995. V. 199. P. 65–69.
497. *Sparrow C.* // Appl. Math. Sci. 1982. V. 41. 269 p.
498. *Stamp A. T., Osipov G. V., Collins J. J.* // Chaos. 2002. V. 12. P. 931–940.
499. *Stauffer D.* Introduction to Percolation Theory. — London: Taylor and Francis, 1994 (second printing). 181 p.
500. *Swinney H. L., Gollub J. P.* // Phys. Today. 1978. V. 31, №8. P. 41–49.
501. *Swinney H. L., Simoyi R. H., Roux J. C.* // Chaos and Statistical Methods / Ed. Y. Kuramoto. — Berlin: Springer, 1984. P. 244–248.
502. *Tabachnikov A.* Billiards. — France Mathematical Soc. Press, 1995. 142 p.
503. *Takagi S., Pumir A., Pazo D., et al.* // Phys. Rev. Lett. 2004. V. 93. P. 058101.
504. *Takens F.* // Lect. Notes in Math. Berlin, Springer. 1981. V. 898. P. 366–381.
505. *Takens F.* // Nonlinear Dynamical Systems and Chaos. Eds. H. W. Broer, S. A. van Gills, I. Hoveijn, F. Takens. In: Progress in Nonlinear Differential Equations and Their Applications, V. 19. Birkhauser, Basel. 1996.
506. *Tam W. Y., Vastano J. A., Swinney H. L., Horsthemke W.* // Phys. Rev. Lett. 1988. V. 61. P. 2163–2166.
507. *Tavakol R. K., Tworkowski A. S.* // Phys. Lett. A. 1984. V. 100. P. 65–67.
508. *Tavakol R. K., Tworkowski A. S.* // Phys. Lett. A. 1984. V. 100, P. 273–276.
509. *Temam R.* Infinite Dimensional Dynamical Systems in Mechanics and Physics. — Berlin: Springer, 1988. 615 p.
510. *Tennison J. L., Lieberman M. A., Lichtenberg A. J.* // Nonlinear Dynamics and the Beam–Beam Interaction / Eds J. C. Herrera, M. Month. N. Y.: AIP Conference Proc. 1979. №57. P. 272–301.
511. *Testa J., Perez J., Jeffries C.* // Phys. Rev. Lett. 1982. V. 48. P. 714–717.
512. *Theiler J.* // Phys. Rev. 1987. V. 36A. P. 4456–4462.
513. *Theory and Applications of Coupled Map Lattices.* Ed. K. Kaneko. — Wiley, 1993.

514. *Thompson J. M. T., Stewart H. B.* Nonlinear Dynamics and Chaos. — N. Y.: Wiley, 1986.
515. *Turner J. S., Roux J. -C., McCormick W. D., Swinney H. L.* // Phys. Lett. 1981. V. 85A. P. 9–12.
516. *Ulam S. M.* // Proc. of the 4th Berkeley Symp. on Math. Stat. and Probability. — California Univ. Press, 1961. V. 3. P. 315.
517. *Ulam S. M., von Neuman J.* // Bull. Amer. Math. Soc., 1947. V. 53. P. 1120.
518. *Umberger D. K., Mayer-Kress J., Jen J.* // Dimensions and Entropies in Chaotic Systems. / Ed. G. Mayer-Kress. — Springer: Berlin, 1986. P. 42–53.
519. *Vautard R., Yiou P., Ghil M.* // Physica D. 1992. V. 58. P. 95–126.
520. *Vicsek T.* // J. Phys. A. 1983. V. 16. P. L647–L652.
521. *Walden R. W., Kolodner P., Rensner A., Surko C. M.* // Phys. Rev. Lett. 1984. V. 53. P. 242–245.
522. *Walker G. H., Ford J.* // Phys. Rev. 1969. V. 188. P. 416–432.
523. *Whiteman K. J.* // Rep. Progr. Phys. 1977. V. 40. P. 1033–1099.
524. *Williams R. F.* // Proc. Symp. in Pure Math.: Global Analysis Amer. Math. Soc. 1970. V. 14. P. 361–393.
525. *Witten T. A., Sander L. M.* // Phys. Rev. Lett. 1983. V. 47. P. 1400–1403.
526. *Witten T. A., Sander L. M.* // Phys. Rev. B. 1983. V. 27. P. 5686–5697.
527. *Wolf A., Swift J., Swinney H. L., Vastano J.* // Physica. 1985. V. 16D. P. 285–317.
528. *Wright J.* // Phys. Rev. 1984. V. 29A. P. 2924–2927.
529. *Xia Z.* // J. Dynamics and Diff. Eqs. 1993. V. 5. P. 219–240.
530. *Xia Z.* // J. Diff. Eqs. 1994. V. 110. P. 289–321.
531. *Yamazaki H., Mino M., Nagashima H., Warden M.* // J. Phys. Soc. Japan. 1987. V. 56. P. 742–750.
532. *Yeh W. J., Kao Y. H.* // Appl. Phys. Lett. 1983. V. 42. P. 299–301.

533. *Young L. -S.* // *Ergod. Theory and Dyn. Syst.* 1982. V. 2. P. 109–124.
534. *Zhang H., Hu B., Hu G.* // *Phys. Rev. E.* 2003. V. 68. P. 026134.
535. *Zhang H., Cao Zh., Ning-Jie Wu, He-Ping Ying, Hu G.* // *Phys. Rev. Lett.* 2005. V. 94. P. 188301.
536. *Zhilin Qu, Weiss J. N., Garfinkel A.* // *Phys. Rev. Lett.* 1997. V. 58. P. 1378–1390.

К главе 3

1. *Адамар Ж.* Исследования психологии процесса изобретения в области математики. – М.: Мир, 1970. С. 31.
2. *Вайнцвайг М. Н., Либерман Е. А.* // *Биофизика*, 1973. Т. 18. С. 939–942.
3. *Веденов А. А.* Моделирование элементов мышления. – М.: Наука, 1988.
4. *Веденов А. А., Ежов А. А., Левченко Е. Б.* // *Нелинейные волны. Структуры и бифуркации* // Под ред. А. В. Гапонова-Грехова и М. И. Рабиновича. – М.: Наука, 1987. С. 53–67.
5. *Веденов А. А., Левченко Е. Б.* // *Письма ЖЭТФ*. 1985. Т. 41. С. 328–331.
6. *Герзон С. А., Михайлов А. С.* // *ДАН СССР*. 1986. Т. 291. С. 228–231.
7. *Дружинин О. А., Михайлов А. С.* // *Изв. вузов. Сер. «Радиофизика»*. 1989. Т. 32. №4. С. 444–450.
8. *Доценко В. С.* // *Письма ЖЭТФ*. 1986. Т. 44. С. 151–153.
9. *Дунин-Барковский В. Д.* Информационные процессы в нейронных структурах. – М.: Наука, 1978.
10. *Дунин-Барковский В. Д.* // *Моделирование возбудимых структур*. – Пуццино, 1975. С. 90–141.
11. *Иоффе Л. В., Фейгельман М. В.* // *Письма в ЖЭТФ*. 1986. Т. 44. С. 148–150.
12. *Куффер С., Николис Дж.* От нейрона к мозгу. – М.: Мир, 1979. 439 с.
13. *Кохонен Т.* Ассоциативная память. – М.: Мир, 1982. 239 с.

14. *Клацки Р.* Память человека: структуры и процессы. – М.: Мир, 1978. 319 с.
15. *Кадомцев Б. Б.* // *Нелинейные волны. Структуры и бифуркации.* / Под ред. А. В. Гапонова-Грехова и М. И. Рабиновича. – М.: Наука, 1987. С. 45–53.
16. *Либерман Е. А.* // *Биофизика*. 1972. Т. 17. С. 932–943.
17. *Либерман Е. А., Минина С. В., Шкловский-Корди Н. Б.* // *Биофизика*. 1980. Т. 25. С. 455–461.
18. *Либерман Е. А., Минина С. В., Шкловский-Корди Н. Б., Конрад М.* // *Биофизика*. 1982. Т. 27. С. 863–870.
19. *Либерман Е. А., Минина С. В., Шкловский-Корди Н. Б.* // *Биофизика*. 1986. Т. 31. С. 298–303.
20. *Ландау Л. Д., Лифшиц Е. М.* *Гидродинамика*. – М.: Наука, 1986. 733 с.
21. *Михайлов А. С.* // *Природа*. 1987. № 3. С. 15–26.
22. *Минский М.* // *Реальность и прогнозы искусственного интеллекта*. – М.: Мир, 1987. С. 231–244.
23. *Минский М.* *Вычисления и автоматы*. – М.: Мир, 1971. 364 с.
24. *Рамбиди Н. Г., Замалин В. М., Сандлер Ю. М. и др.* Молекулярная элементная база перспективных информационно-логических устройств. // *Итоги науки и техники. Сер. «Электроника»*. Т. 22. – М.: Изд-во ВИНТИ, 1987. 168 с.
25. *Aoke K., Mugebayasche N.* // *Phys. Lett.* 1986. V. 114A. P. 425.
26. *Almeida L. B.* // *Neural Computers* / Eds R. Eckmiller and Ch. van den Malsburg. – Berlin: Springer, 1988. P. 199–208.
27. *Alspector J., Allen R. B.* // *Advanced Research in VLSI. Proc. of the 1987 Stanford Conf.* – Cambridge: MIT Press, 1987. P. 313–349.
28. *Amari S.* // *Neural and Synergetic Computers* / Ed. H. Haken. – Berlin: Springer, 1988.
29. *Amari S., Arbib M. A.* *Competition and Cooperation in Neural Nets*. – Berlin: Springer, 1982.

30. *Amari S., Maginu K.* // Neural Networks. 1988. V. 1. P. 63–73.
31. *Amari S., Yoshida K., Kanatani K.* // SIAM J. Appl. Math. 1977. V. 33. P. 95–126.
32. *Amit D. J., Gutfreund H., Sompolinsky H.* // Phys. Rev. 1985. V. 32A. P. 1007.
33. *Amit D. J., Gutfreund H., Sompolinsky H.* // Phys. Rev. Lett. 1985. V. 55. P. 1530–1538.
34. *Amit D. J., Gutfreund H., Sompolinsky H.* // Ann. Phys. (N.Y.). 1987. V. 173. P. 30–67.
35. *Amit D. J., Gutfreund H., Sompolinsky H.* // Phys. Rev. 1987. V. 35A. P. 2293.
36. *Anderson D. Z.* // Opt. Lett. 1986. V. 11. P. 56.
37. *Anderson D. Z.* // Neural Computers / Eds. R. Eckmiller and Ch. van den Malsburg. – Berlin: Springer, 1988. P. 417–424.
38. *Anderson J. A.* // Disordered Systems and Biological Organization / Eds. E. Bienenstock et al. – Berlin: Springer, 1986. P. 209–226.
39. *Ballard D. H., Hinton G. E., Sejnowski T. J.* // Nature. 1983. V. 306. P. 21–26.
40. *Banavar J. R., Sherrington D., Sorlas N.* // J. Phys. 1987. V. 20A. P. L1–L8.
41. *Banzhaf W.* // J. Physique. 1987. V. 48. P. 2027–2035.
42. *Banzhaf W.* // Neural and Synergetic Computers / Ed. H. Haken. – Berlin: Springer, 1988. P. 155–167.
43. *Barto A. G.* // Human Neurobiology. 1985. V. 4. P. 229–256.
44. *Baskaran G., Fu Y., Anderson P. W.* // J. Stat. Phys. 1986. V. 45. P. 1.
45. *Berlekamp E. R., Conway J. H., Guy R. K.* Winning Ways. V. 2. – N.Y.: Academic Press, 1982.
46. *Boseniuk T., Ebeling W., Engel A.* // Phys. Lett. 1987. V. 125A. P. 307–311.
47. *Brenner K.-H.* // Appl. Opt. 1986. V. 25. P. 3061.
48. *Brenner K.-H.* // Appl. Opt. 1988. V. 27. P. 1687–1691.
49. *Buhmann J., Schulten K.* // Europhys. Lett. 1987. V. 4. P. 1205–1209.

50. *Buhmann J., Schulten K.* // Neural Computers / Eds. R. Eckmiller and Ch. van den Malsburg. – Berlin: Springer, 1988. P. 231–242.
51. *Carpenter G. A., Grossberg S.* // Neural and Synergetic Computers / Ed. H. Haken. – Berlin: Springer, 1988. P. 42–74.
52. *Carter F. L.* // Physica. 1984. V. 10D. P. 175–194.
53. *Ceccatto H. A., Huberman B. A.* // Physica Scripta. 1988. V. 37. P. 145–150.
54. *Chen H., Matthaeus W. H.* // Phys. Rev. Lett. 1987. V. 58. P. 1845.
55. *Chilingaryan A. A., Khudonyan N. O., Saakyan O. B., Zazyan G. Z.* Recognition of correlated patterns with spin-glass-like models. – Печать ЕФИ–992 (42)–87. Ереванский физ. институт. 1987. 10 с.
56. *Chopard B., Droz M.* // J. Phys. 1988. V. 21A. P. 205–212.
57. *Clark J. W., Rafelski J., Winston J. W.* // Phys. Rep. 1985. V. 123. P. 215–273.
58. *Collins A. M., Lotrus E. F.* // Psychol. Rev. 1975. V. 82. P. 407–428.
59. *Conrad M.* // Biosystems. 1972. V. 5. P. 1–14.
60. *Conrad M.* // J. Neurosci. Res. 1976. V. 2. P. 233–254.
61. *Conrad M.* Adaptability. – The Significance of Variability from Molecule to Ecosystem – N.Y.: Plenum Press, 1983.
62. *Conrad M.* // Comm. ACM. 1985. V. 28. P. 464–480.
63. *Cottrell M.* // Biol. Cybern. 1988. V. 58. P. 129–139.
64. *Crick F. C.* // Sci. Amer. 1979. No. 3. P. 181–193.
65. *Crick F. C., Mitchison G.* // Nature. 1983. V. 304. P. 111–114.
66. *Crisanti A., Amit D. J., Gutfreund H.* // Europhys. Lett. 1986. V. 2. P. 337.
67. *De Archangelis L.* // J. Phys. 1987. V. 20A. P. L369–L373.
68. *De Archangelis L., Stauffer D.* // J. Physique. 1987. V. 48. P. 1881.
69. *Dehaene S., Changeux J. P., Nadal J. P.* // Proc. Nat. Acad. Sci. USA. 1987. V. 84. P. 2727.

70. *Denker J. S.* // *Physica*. 1986. V. 22D. P. 216–232.
71. *Derrida B., Gardner E., Zippelius A.* // *Europhys. Lett.* 1987. V. 4. P. 167.
72. *Derrida B., Nadal J. P.* // *J. Stat. Phys.* 1987. V. 49. P. 993.
73. *Diederich S., Oppen M.* // *Phys. Rev. Lett.* 1987. V. 58. P. 949.
74. *Domany E., Kinzel W.* // *Phys. Rev. Lett.* 1984. V. 53. P. 311–314.
75. *Domany E., Meiz R., Kinzel W.* // *Europhys. Lett.* 1986. V. 2. P. 175.
76. *Dotsenko V. S.* // *J. Phys.* 1985. V. 18C. P. L1017–L1022.
77. *Duong-Van M., Feit M. O., Keller P., Pound M.* // *Physica*. 1986. V. 23D. P. 448.
78. *Durbin R., Willshaw D. J.* // *Nature*. 1987. V. 326. P. 689.
79. *Ebeling W.* // *J. Stat. Phys.* 1986. V. 45. P. 891–903.
80. *Ebeling W., Engel A.* // *Syst. Ana. Model. Simul.* 1986. V. 5. P. 377–385.
81. *Fahlman S. E., Hinton G. E., Sejnowski T. J.* // *Proc. Nat. Conf. Artificial Intelligence*, Washington, 1983. P. 109–113.
82. *Farhat N., Psaltis D., Prata A., Paek E.* // *Appl. Opt.* 1985. V. 24. P. 1469–1475.
83. *Farmer J. D., Packard N. H.* // *Physica*. 1986. V. 22D. P. 187–204.
84. *Farmer J. D., Kauffman S. A., Packard N. H.* // *Physica*. 1986. V. 22D. P. 50–67.
85. *Feigelman M. V., Ioffe L. B.* // *Int. J. Mod. Phys. B*. 1987. V. 1. P. 51–68.
86. *Feynman P. P.* // *Int. J. Theor. Phys.* 1982. V. 21. P. 467–488.
87. *Feynman P. P.* // *Foundations of Physics*. 1986. V. 16. P. 507–531.
88. *Forrest B. M., Roweth D., Stroud N. et al.* // *The Computer J.* 1987. V. 30. P. 413.
89. *Forrest B. M.* // *J. Phys.* 1988. V. 21A. P. 245–256.
90. *Frisch U., Hasslacher B., Pomeau Y.* // *Phys. Rev. Lett.* 1986. V. 56. P. 1505.

91. *Fu Y., Anderson P. W.* // *J. Phys.* 1986. V. 19A. P. 1605.
92. *Fuchs A., Haken H.* // *Neural and Synergetic Computers* / ed. H. Haken. – Berlin: Springer, 1988. P. 16–28.
93. *Fuchs A., Haken H.* // *Biol. Cybern.* 1988. V. 60. P. 17–22.
94. *Fukushima K.* // *Biol. Cybern.* 1986. V. 55. P. 515.
95. *Fukushima K., Miyake S.* // *Pattern Recognition*. 1982. V. 15. P. 455–469.
96. *Gabriel P. P.* // *Science*. 1986. V. 231. P. 975.
97. *Gardner E.* // *J. Phys. A* 1988. V. 21. P. 256–270.
98. *Gardner E., Forrest B. M.* // *J. Phys.* 1988. V. 21A. P. 271–284.
99. *Gardner M.* *Wheels, Life and Other Mathematical Amusements*. – San Francisco: Freeman, 1982.
100. *Garey M. R., Johnson D. S.* *A Guide to the Theory of NP-Completeness*. – San-Francisco: Freeman, 1979.
101. *Geman S., Geman D.* // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. PAMI-6. 1984. P. 721–741.
102. *Goodman S. D., Rhodes W. T.* // *Appl. Opt.* 1988. V. 27. P. 1708–1714.
103. *Graf H. P., Jackel L. D., Hubbard W. E.* // *IEEE Computer*. 1988. V. 21. P. 41–49.
104. *Gorman R. P., Sejnowski T. J.* // *Neural Networks*. 1988. V. 1. P. 75–89.
105. *Grassberger P.* // *Physica*. 1984. V. 10D. P. 52–58.
106. *Grassberger P.* // *J. Stat. Phys.* 1986. V. 45. P. 27.
107. *Grassberger P., Krause F., von der Twer T.* // *J. Phys.* 1984. V. 17A. P. L105–L109.
108. *Grinstein G., Jayaprakash C., He Yu.* // *Phys. Rev. Lett.* 1985. V. 55. P. 2527–2530.
109. *Grossberg S.* *Studies of Mind and Brain*. – Norwell: Reidel, 1982.
110. *Grossberg S.* // *Neural Networks*. 1988. V. 1. P. 17–61.

111. *Guyon I., Personnaz L., Dreyfus G.* // Neural Computers / Eds R. Eckmiller and Ch. van den Malsburg. – Berlin: Springer, 1988. P. 261–269.
112. *Guttinger W., Dangelmayr G.* // Neural and Synergetic Computers / Ed. H. Haken. – Berlin: Springer, 1988 P. 29–41.
113. *Haken H.* // Pattern Formation and Pattern Recognition // Ed. H. Haken. – Berlin: Springer, 1979.
114. *Haken H.* // Z. Phys. 1988. Bd 70B. S. 121–123.
115. *Haken H.* Information and Self-organization. – Berlin: Springer, 1988. 196 p.
116. *Haken H.* // Neural and Synergetic Computers / Ed. H. Haken. – Berlin: Springer, 1988. P. 2–15.
117. *Hardy J., Pomeau Y.* // J. Math. Phys. 1972. V. 13. P. 1042.
118. *Hardy J., de Pazzis O., Pomeau Y.* // Phys. Rev. 1976. V. 13A. P. 1949.
119. *Hebb D. C.* The Organization of Behavior: A Neuropsychological Theory. – N.Y.: Wiley, 1957.
120. *Hecht-Nielsen R.* // Proc. Soc. Photo-Opt. Instrum. Eng. 1986. V. 634. P. 277.
121. *Hecht-Nielsen P.* // IEEE Spectrum. 1988. V. 25. P. 36–41.
122. *Hemmen, van J. L., Grensing D., Huber A., Kuhn P.* // J. Stat. Phys. 1988. V. 50. P. 231.
123. *Hemmen, van J. L., Grensing D., Huber A., Kuhn P.* // J. Stat. Phys. 1988. V. 50. P. 259.
124. *Hemmen, van J. L., Keller G., Kuhn R.* // Europhys. Lett. 1988. V. 5. P. 663–668.
125. *Hemmen, van J. L., Kuhn R.* // Phys. Rev Lett. 1986. V. 57. P. 913.
126. *Hemmen, van J. L., Zagrebnov V. A.* // J. Phys. 1987. V. 20A. P. 3989–3999.
127. *Hillis D.* The Connection Machine. – Cambridge: MIT Press, 1985.
128. *Hinton G. E., Sejnowski T. J.* // Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition, Washington, 1983. P. 448–453.

129. *Hinton G. E., Sejnowski T. J., Ackley D. H.* // Carnegie-Mellon Univ. Tech. Rep. CMU–CS–84–119. 1984.
130. *Hinton G. E., McLelland J. L., Rumelhart D. E.* // Parallel Distributed Processing: Explorations in the Microstructure of Cognition / Eds D. E. Rumelhart et al. – Cambridge: MIT Press, 1986. P. 77–109.
131. *Hinton G. E.* // Proc. Eighth Annual Conf. of the Cognitive Science Society. – N.Y.: Laurence Erlbaum, 1986. P. 1–12.
132. *Hogg T., Huberman B. A.* // Proc. Nat. Acad. Sci. USA. 1984. V. 81. P. 6871–6875.
133. *Hogg T., Huberman B. A.* // J. Stat. Phys. 1985. V. 41. P. 115–123.
134. *Hogg T., Huberman B. A.* // Phys. Rep. 1987. V. 156. P. 227–310.
135. *Holland J. H.* // Physica. 1986. V.22D. P. 307–317.
136. *Hopfield J. J.* Neural networks and physical systems with emergent collective computation abilities // Proc. Nat. Acad. Sci. USA. 1982. V. 79. P. 2554–2558.
137. *Hopfield J. J., Feinstein D. J., Palmer R. G.* // Nature. 1983. V. 304. P. 158–159.
138. *Hopfield J. J., Tank D. W.* // Biol. Cybern. 1985. V. 52. P. 141–152.
139. *Hopfield J. J., Tank D. W.* // Science. 1986. V. 233. P. 625–633.
140. *Hopfield J. J., Tank D. W.* // Disordered Systems and Biological Organization / Eds E. Bienenstock et al. – Berlin: Springer, 1986.
141. *Huberman B. A.* // Neural Computers / Eds R. Eckmiller and Ch. van den Malsburg. – Berlin: Springer, 1988. P. 455–465.
142. *Huberman B. A., Hogg T.* // Phys. Rev. Lett. 1984. V. 52. P. 1048–1051.
143. *Huberman B. A., Hogg T.* // Physica. 1986. V. 22D. P. 376.
144. *Huberman B. A., Hogg T.* // Artificial Intelligence. 1987. V. 33. P. 155–171.
145. *Ingerson T. E., Buvel R. L.* // Physica. 1984. V. 10D. P. 59.
146. *Jen E.* // J. Stat. Phys. 1986. V. 43. P. 243.

147. *Jerne N. K.* // *Sci. Amer.* 1973. №7. P. 52–60.
148. *Jerne N. K.* // *Science.* 1985. V. 229. P. 1057–1059.
149. *Jerne N. K.* // *Stability and Origin of Biological Information* / Ed. I. R. Miller. – N.Y.: Wiley, 1975. P. 201–204.
150. *Kanter I., Sompolinsky H.* // *Phys. Rev.* 1987. V. 35A. P. 380.
151. *Kauffman S. A.* // *Physica.* 1984. V.10D. P. 145–156.
152. *Keirstead W., Huberman B. A.* // *Phys. Rev. Lett.* 1986. V. 56. P. 1094–1097.
153. *Kelso J. A. S., deGuzman G. C.* // *Neural and Synergetic Computers* / Ed. H. Haken. – Berlin: Springer, 1988. P. 180–196.
154. *Kienker P. K., Sejnowski T. J., Hinton G. E., Schumacher L. E.* // *Perception.* 1986. V. 15. P. 197–216.
155. *Kinzel W.* // *Z. Phys.* 1985. V. 60B. P. 205–213.
156. *Kinzel W.* // *Z. Phys.* 1985. V. 58B. P. 229–244.
157. *Kirkpatrick S.* // *J. Stat. Phys.* 1984. V. 34. P. 975.
158. *Kirkpatrick S., Gelatt C. D., Vecchi M. P.* // *Science.* 1983. V. 220. P. 671.
159. *Kirkpatrick S., Toulouse G.* // *J. Physique.* 1985. V. 46. P. 1277.
160. *Kleinfield D.* // *Proc. Nat. Acad. Sci. USA.* 1986. V. 83. P. 9469–9473.
161. *Krauth W., Mezard M. J.* // *J. Phys.* 1987. V. 20. A. P. L745.
162. *Landauer R.* // *Physica Scripta* 1987. V. 35. P. 88–95.
163. *Langton C. G.* // *Physica.* 1984. V. 10D. P. 135–144.
164. *Le Cun Y.* // *Proc. Int. Conf. Cognitiva'85.* – Paris, 1985. P. 599–604.
165. *Lehky S., Sejnowski T. J.* // *Nature.* 1988. V. 333. P. 452–454.
166. *Lieberman E. A.* // *Biosystems.* 1979. V. 11. P. 111–124.
167. *Little W. A.* // *Math. Biosci.* 1974. V. 19. P. 101.
168. *Little W. A., Shaw G. L.* // *Math Biosci.* 1978. V. 39. P. 281.

169. *Margolus N. H.* // *Physica.* 1984. V. 10D. P. 85–95.
170. *Marlsburg, von der C., Bienenstock E.* // *Europhys. Lett.* 1987. V. 3. P. 1243–1249.
171. *McCulloch W. C., Pitts W.* // *Bull. Math. Biophys.* 1943. V. 5. P. 115–133.
172. *Meir R., Domany E.* // *Phys. Rev. Lett.* 1987. V. 59. P. 359–362.
173. *Menyhard N.* // *J. Phys.* 1988. V. 21A. P. 1283–1292.
174. *Metropolis N. et al.* // *J. Chem. Phys.* 1953. V. 6. P. 1087–1092.
175. *Mezard M., Parisi G.* // *Europhys. Lett.* 1986. V. 2. P. 913.
176. *Mezard M., Parisi G.* // *J. Physique.* 1986. V. 47. P. 1285.
177. *Mezard M., Nadal J. P., Toulouse G.* // *J. Physique.* 1986. V. 47. P. 1457–1462.
178. *Mikhailov A. S.* // *J. Phys.* 1988. V.21A. P. L487–L490.
179. *Minsky M.* *The Society of Mind.* – N.Y.: Simon and Shuster, 1986.
180. *Minsky M., Papert S.* *Perceptrons.* – Cambridge: MIT Press, 1969.
181. *Moopenn A., Lambe J., Thakoor A. P.* // *IEEE Trans. on systems, Man, and Cybern.* 1987. V. 17. №2. P. 325–331.
182. *Moore M. A.* // *Phys. Rev. Lett.* 1987. V. 58. P. 1703.
183. *Nadal J. P., Toulouse G., Changeux J. P., Dehane S.* // *Europhys. Lett.* 1986. V. 1. P. 535–542.
184. *Nemeth R.* // *J. Phys.* 1987. V.20A. P. L85–L88.
185. *Neuman, von J.* *Theory of Self-Reproducing Automata.* – Univ. Illinois Press, 1966.
186. *Nicolis J. S.* // *Rep. Progr. Phys.* 1986. V. 49. P. 1109–1196.
187. *Orland H.* // *J. Physique Lett.* 1985. V. 46. P. L763–L770.
188. *Packard N. H., Wolfram S.* // *J. Stat. Phys.* 1985 V. 38. P. 901–946.
189. *Palm G.* // *Science.* 1987. V. 235. P. 1227.

190. *Palm G.* // Neural and Synergetic Computers / Ed. H. Haken. – Berlin: Springer, 1988. P. 100–110.
191. *Parisi G.* // J. Phys. 1986. V. 19A. P. L675–L680.
192. *Parisi G.* // Physica Scripta. 1987. V. 35. P. 123.
193. *Parga N., Virasoro M. A.* // J. Physique. 1986. V. 47. P. 1857–1864.
194. *Peretto P.* // Biol. Cybern. 1984. V. 5. P. 51–62.
195. *Peretto P., Niez J. J.* // Disordered Systems and Biological Organization / Eds E. Bienenstock et al. – Berlin: Springer, 1986. P. 171–185.
196. *Personnaz L., Guyon I., Dreyfus G., Toulouse G.* // J. Stat. Phys. 1986. V. 43. P. 411–422.
197. *Personnaz L., Guyon I., Dreyfus G.* // J. Physique Lett. 1985. V. 46. P. L359–L365.
198. *Personnaz L., Guyon I., Dreyfus G.* // Phys. Rev. 1987. V. 34A. P. 4217.
199. *Pomeau Y.* // J. Phys. 1984. V. 17A. P. L415–L418.
200. *Prager R. W., Harrison T. O., Fallside F.* // Computer Speech and Language, 1986. V. 1. P. 3–27.
201. *Randelman R. E., Crest G. S.* // J. Stat. Phys. 1986. V. 45. P. 885.
202. *Rees S., Ball R. C.* // J. Phys. 1987. V. 20A. P. 1239–1249.
203. *Rumelhart D. E., Hinton G. E., Williams R. J.* // Nature. 1986. V. 323. P. 533.
204. *Rumelhart D. E., McClelland J. L.* // Parallel Distributed Processing: Explorations in the Microstructure of Cognition / Eds D. E. Rumelhart et al. – Cambridge: MIT Press, 1986. P. 216–271.
205. *Schastri L.* // Appl. Opt. 1987. V. 26. P. 1829–1844.
206. *Schuster P.* // Physica. 1986. V. 22D. P. 100–119.
207. *Sejnowski T. J.* // Neural Computers / Eds R. Eckmiller and Ch. van den Malsburg. – Berlin: Springer, 1988. P. 291–300.
208. *Sejnowski T. J., Rosenberg C. R.* // Complex Systems. 1987. V. 1. P. 145–168.

209. *Sejnowski T. J., Hinton G. E.* // Vision. Brain and Cooperative Computation / Eds M. A. Arbib, A. R. Hanson, S. Sejnowski. – Cambridge: MIT Press, 1985. P. 201.
210. *Sejnowski T. J., Kienker P. K., Hinton G. E.* // Physica. 1986. V. 22D. P. 260–275.
211. *Sejnowski T. J., Rosenberg C. R.* // John Hopkins Univ. Electrical Eng. and Computer Sci. Tech. Report JHU/EECS–86/01. 1986.
212. *Soffer B. H., Dunning G. J., Owechko Y., Marom E.* // Opt. Lett. 1986. V. 11. P. 118.
213. *Sompolinsky H.* // Phys. Rev. 1986. V. 34A. P. 2571–2574.
214. *Sompolinsky H., Kanter I.* // Phys. Rev. Lett. 1986. V. 57. P. 2861–2864.
215. *Sourlas N.* // Europhys. Lett. 1986. V. 2. P. 919.
216. *Thakoor A. P., Moopenn A., Lambe J., Khanna S. K.* // Appl. Opt. 1987. V. 26. P. 5085–5092.
217. *Toffoli T.* // Physica. 1984. V. 10D. P. 195–204.
218. *Toffoli T.* // Physica. 1984. V. 10D. P. 117–127.
219. *Toulouse G., Dehaene S., Changeaux J. P.* // Proc. Nat. Acad. Sci. USA. 1986. V. 83. P. 1695–1698.
220. *Vannimenius J., Mezard M.* // J. Physique Lett. 1984. V. 45. P. L1145.
221. *Vichniac G. Y.* // Physica. 1984. V. 10D. P. 96–115.
222. *Vichniac G. Y.* // Disordered Systems and Biological Organization / Eds E. Bienenstock et al. – Berlin: Springer, 1986. P. 3–20.
223. *Vichniac G. Y., Tamayo P., Hartman H.* // J. Stat. Phys. 1986. V. 45. P. 875.
224. *Wagner K., Psaltis D.* // Appl. Opt. 1987. V. 26. P. 5061–5076.
225. *Weick M.* // Neural and Synergetic Computers / Ed. H. Haken. – Berlin: Springer, 1988. P. 122.
226. *Weisbuch G., Atlan H.* // J. Phys. 1988 V. 21A. P. L189–192.

227. *Weisbuch G., Fogelman-Soulie F.* // *J. Physique Lett.* 1985. V. 46. P. L623–L630.
228. *Willson S. J.* // *Physica.* 1987. V. 24D. P. 179–189.
229. *Withege W., Sherrington D.* // *J. Phys.* 1987. V. 20A. P. L9–L11.
230. *Wolfram S.* // *Rev. Mod. Phys.* 1983. V. 55. P. 601.
231. *Wolfram S.* // *Nature.* 1984. V. 311. P. 419–424.
232. *Wolfram S.* // *Physica.* 1986. V. 10D. P. 1–35.
233. *Wolfram S.* // *Physica.* 1986. V. 22D. P. 385–399.
234. *Wolfram S.* // *J. Stat. Phys.* 1986. V. 45. P. 471.
235. *Wong K. J. M., Sherrington D., Mottishaw P. et al.* // *J. Phys.* 1988. V. 21A. P. L99–L106.
236. *Yariv A., Kwong S.* // *Opt. Lett.* V. 11. P. 186.

Предметный указатель

- Автокорреляционная функция 242–244
- Авторегрессия 429–430
- Автомат 20
- клеточный 47–58, 475–488
 - вероятностный 476
 - детерминированный 476
 - класса I 478
 - класса II 478
 - классов III и IV 479–480
 - модели гидродинамических течений 482–487
- Агрегация, ограниченная диффузией 272–275
- Аксон 462, 463
- Активатор 48, 113
- Архимедова спираль 63, 71
- Аттрактор 208, 210, 211, 216, 221, 222, 223, 234, 241, 329, 499, 505
- гиперболический 233
 - квазистохастический 233, 234
 - Лоренца 224–231
 - область притяжения 211
 - простой 210–211, 309
 - размерность вложения 386
 - Смейла–Вильямса 231–232
 - стохастический 233
 - странный 222–224, 241, 258, 259, 261, 265, 266, 267, 329, 347–349, 357, 360
 - нехаотический 350
 - модель 223–224
- Фейгенбаума 306
 - хаотический 233, 268, 374
 - Эно 258–259, 260
- Бареттор 111
- Бегущий импульс 39–46
- Бильярд 177–185
- компоненты 180
 - нейтральные 180
 - рассеивающие 180
 - фокусирующие 180, 184
 - рассеивающий 180
 - с возмущаемыми границами 185–186
 - замедление частиц 185, 207–208
 - ускорение частиц 185–186, 194–196
 - типа «стадион» 180, 182–184, 196–201
 - фокусирующий 180
- Бистабильность 343
- Бифуркация 229–233, 286
- Андронова–Хопфа 330
 - внутренняя 329
 - касательная 336, 361
 - обратная 335
 - рождения тора 336
 - удвоения 333, 335, 356
- Вихрь 81
- скрученный 82–84, 89–90
- Вихревое кольцо 83–88
- плотное 84

- резонанс 89
- рыхлое 85–88
- Возмущение стабилизирующее 376
- Волна
 - возбуждения 20
 - гашения 39, 40
 - горения 40
 - заселения 31
 - переключения 19, 21–32, 25, 37
 - быстрая 30
 - медленная 29
 - спиральная 51–66, 101–108
 - дрейф 74–81
 - резонанс 78–79
 - частота вращения 60, 61–65, 71–74
 - ядро 51, 61–65
 - фазовая 20, 96–98
- Временной ряд 386, 429, 436,
 - прогноз 429–458
 - глобальные методы 429–433
 - локальные методы 433–435
- Время предсказуемости 236
- Вязкие пальцы 270, 271

- Газ Лоренца 178–179
 - с возмущаемыми границами 188–196
- Гамильтониан 117, 118, 119, 131, 136, 142, 148
 - интегрируемый 119
 - невырожденный 149
 - нелинейного резонанса 138
- Гетероклиническая
 - точка 160, 161
 - траектория 160
- Гетероклинические сплетения 164
- Гипотеза Каплана–Йорка 265–267
- Главные компоненты 431–432
- Глобальная устойчивость 150
- Гомоклиническая
 - точка 160, 161
 - кривая 160, 250, 251
 - петля 160, 226, 250
- Гомоклинические сплетения 164, 250–251
- Грассбергера–Прокаччия метод 389–390
- Грубость 348

- Движение
 - квазипериодическое 124, 125, 221–222, 329
 - периодическое 124, 329
 - условно периодическое 124
 - финитное 118
 - эргодическое 169, 223
- Дендрит 313, 462, 463
- Дефект 397, 401–411
- Диаграмма Ламерея 283, 284
- Дисперсии закон 44
- Диссипативные структуры 108–114
- Диффузионная длина 8, 75
- Диффузия Арнольда 150–152

- Емкость 263, 279
- Жесткое возбуждение 332

- Задача
 - Кеплера 346
 - запоминания родственных связей 527
 - комбинаторной оптимизации 506–507, 513
 - неполиномиальной сложности 507
 - о коммивояжере 506, 512
 - прогноза 443–449
- Зародыш 35
 - критический 36

- фазы 35
- Затягивание потери устойчивости 332
- Золотое среднее 351

- Игра «Жизнь» 480
- Инвариантность аттрактора 210
 - геометрическая 258–259
 - масштабная 258
- Ингибитор 39, 113
- Интеграл
 - движения 118
 - корреляционный 390
- Информация 462–465
 - аналоговая обработка 462, 471–472
 - сжатие 469

- Канонические переменные 117
- Картина развития хаоса 346
- Квазиаттрактор 234, 375
- Ковер Серпиньского 261, 262
- Конвекция Рэлея–Бенара 245, 247, 355
- Контролирование хаоса 367
 - без обратной связи 367
 - параметрическое 367, 368
 - с обратной связью 367
 - силовое 369
- Коррелятор 175
- Корреляции временные, расщепление 175–177
- Коэффициент растяжения фаз 183
- Кривая фазовая 117
- Кризис 329, 331, 332
- Критерий хаотичности 234–258
 - Чирикова 145, 146
- Кручение вихря 83

- Локальная аппроксимация 433, 435–438
 - алгоритм 435–443
 - второго порядка 436, 438–440
 - нулевого порядка 436, 438–440
 - первого порядка 436, 438–440

- Матрица
 - задержек 431, 436
 - линеаризации 213
 - монодромии 218, 220
 - соседей 446
- Машина Больцмана 528–529, 532, 533
- Метод «Гусеница» 432
- Мера 170
 - информационно–теоретическая 530
- Метаразмерность 278
- Метод
 - главных компонент 432
 - имитации отжига 507, 511–512, 537
 - Мельникова 250–254
 - обратного распространения ошибок 521
 - Отта–Гребоджи–Йорка 373–376
 - упругой петли 514–517
 - эволюционный 538–539
- Многообразие 223
 - неустойчивое 159, 221
 - устойчивое 159, 221
- Множество
 - Жюлиа 309–315
 - заполненное 310, 313, 314
 - Кантора 259–262
 - Мандельброта 315–321
 - связное 314
 - вполне несвязное 314
 - средних третей 260–261
 - Фату 310
- Модель

- базовая 17
- Винера–Розенблюта 47–48
- DLA 272–277
- решеточного газа 482–485
- Ринцеля–Келлера 45
- Хопфилда 490–506, 517
- Шлёгля 23
- эволюционная 533–542
- Эно–Эейлеса 128–130
- $\lambda - \omega$ 93–94
- Мозг 462–464, 472
- работа 463–465
- Морфогенез 113
- Мультистабильность 417
- Мультипликатор 218, 282
- Мутации 512, 535
- Мышление 469

- Наблюдаемая 428–429
- Нейрон 462, 463, 492
- Нелинейность, параметр 141
- умеренная, условие 141
- Нелинейный маятник 120–121
- резонанс 133–144
- — ширина 140–141
- Необратимость 174
- Непредсказуемость 174

- Образ 469, 470
- распознавание 470, 472, 502–506, 513
- Обучение 470
- Окружность
- инвариантная 155
- резонансная 154
- Осциллятор Дуффинга–Холмса 254
- Отображение 281
- взаимно однозначное 285–288
- закручивающее 154
- квадратичное 290, 297, 312, 319
- кольца 344–346
- логистическое 289, 290, 295, 296, 297–300, 312
- Мисюревича 295
- не взаимно однозначное 288–290
- окружности 338–344
- — стандартное 339
- поворота 153–155, 201
- Пуанкаре 125–127, 248–250, 338
- растягивающее 291–293
- типа палатки 291
- треугольное 291
- унимодальное 294–297
- Эно 258, 282–283
- Отображения топологически эквивалентные 291
- сцепленные 393–398

- Память 464
- ассоциативная 488–496, 503, 413
- Параметр нелинейности 141
- перекрытия резонансов 145
- Пейсмекер 55, 98
- конкуренция 54, 55, 101
- Перекрытие резонансов 144–145
- Переमेжаемость 346, 358–363
- I рода 337, 361
- II рода 337, 362
- III рода 337, 362
- Переменные действие–угол 119
- канонические 117
- резонансные 136–138
- Перемешивание 171–174
- Период три рождает хаос 293–294
- Перцептрон 519–521
- Подмногобразия 223
- Показатели
- Ляпунова 237–242, 257, 292
- Флоке 218
- Поле векторное 167

- структурно устойчивое 348
- грубое 348
- Положение равновесия 209
- Порядок
- аппроксимации 438–440
- Шарковского 293–294
- Потеря устойчивости
- жесткая 332
- мягкая 330
- Поток 209
- перемешивающий 173
- фазовый 167
- эргодический 169
- Правило Хебба 490, 517
- Преобразование удвоения 304
- Проблема
- NP 507
- N тел
- Прогноз 429, 443–449
- величина ошибки 453–461
- итеративный метод 441–443, 450–451
- на один шаг вперед 437–438
- на несколько шагов вперед 440–443
- прямой 442–443, 451
- с пересчетом 441–442, 451–452
- сравнение методов 440, 443, 452
- Производная Шварца 294, 295
- Пространство фазовое 117, 209
- Процессор 464
- матричный 465–466
- Пыль Фату 314, 315

- Равновесие
- устойчивое 211
- неустойчивое 211
- Размерность
- вероятностная 262
- вложения 385–386
- аттрактора 386
- информационная 279
- корреляционная 280, 390
- Ляпуновская 265, 392
- метрическая 262
- обобщенная 278–279
- фрактальная 262–265, 267–268, 279, 392
- хаусдорфова 263
- энтропийная 263
- Резонанс 132, 135
- внутренний 132, 142
- порядка s 134
- Распознавание образов 472, 497–505, 518
- Расщепление сепаратрис 160–165, 250–251
- Реакция Белоусова–Жаботинского 41, 58, 245, 246
- Ренормализация 297
- в теории Фейгенбаума 301–306
- в теории перемежаемости 364–366
- при переходе от квазипериодичности к хаосу 350–354
- Решетки сцепленных отображений 394
- диффузионная связь 396, 420–424
- критерий динамики 411–413
- неоднородные 398, 401–411, 424–428
- однородные 398–401
- потоковая модель 396, 413–420
- Решеточный газ 482–486
- и уравнения Навье–Стокса 485–487

- Седло 214
- Седло–узел 215–216, 226
- Седло–фокус 215–216, 226

- Секущая поверхность 126, 220
 Семантическая структура 472
 Сепаратриса 120
 — расщепление 160–164, 250
 — седловой точки 27, 214, 226
 — ширина 140–141
 Сеть
 — вычислительная 467
 — нейронная 463
 Силовое резонансное воздействие 373
 Синапс 463, 492
 Синаптическая связь 492
 Сингулярное разложение 438
 Сингулярный спектральный анализ 430–433
 Синхронизация колебаний 95, 420, 424, 427, 428
 Система диссипативная 210
 — близкая к интегрируемой 130–132
 — Дуффинга 255
 — Дуффинга–Холмса 255
 — консервативная 118, 210
 — Лоренца 225
 — обучающаяся 474
 — полностью интегрируемая 119
 — распределенная 11–13, 383–384
 — Реслера 356
 — эргодическая 169
 Скейлинговая структура 258
 Смещение перигелия 147
 Соотношение резонанса 134, 155
 Соседи стартового вектора 437
 Спектр
 — дискретный 244, 247
 — непрерывный 247
 — универсальный 306–308
 Спектральная плотность 242–248
 Спин Изинга 489
 Спиновые стекла 488–493
 Среда
 — возбудимая 37
 — — неоднородная 90
 — бистабильная 32
 Среднее
 — вдоль траектории 169
 — по времени 168
 — фазовое 169
 Стабилизация
 — высокочастотная 379
 — резонансная 378–379
 Стохастический слой 198
 Структурная устойчивость 348
 Сценарий развития хаоса 346
 — Ландау–Хопфа 347
 — Рюэля–Такенса 347–349
 — Фейгенбаума 356–357
 Теорема
 — Гершгорина 406
 — Колмогорова 148–150
 — Лиувилля 119
 — Ляпунова об устойчивости 212–213
 — о неподвижной точке (Пуанкаре–Биркгофа) 155–158
 — о разрушении тора 346
 Теория
 — Колмогорова–Арнольда–Мозера (КАМ) 146–152
 — сложных систем 5
 — Флоке 216–221
 Тор 123, 221, 241
 — бифуркации 336–346
 — обмотка 123
 — — всюду плотная 124
 — разрушение 337, 346–349
 — резонансный 125
 — рождение 333, 337
 Точка

- гиперболическая 120, 158–160, 250
 — критическая 294, 353
 — неподвижная 281, 282
 — — неустойчивая 282
 — — седловая 282
 — — устойчивая 282
 — особая 211
 — стационарная 211, 241
 — — асимптотически устойчивая 212
 — — устойчивая 211
 — фазовая 117
 — эллиптическая 120, 158–160
 Траектория фазовая 117, 209
 Турбулентность 7, 95, 96, 391, 421
 — амплитудная 8
 — фазовая 8
 Узел
 — неустойчивый 213, 214
 — устойчивый 213, 214, 225
 Универсальность Мандельброта 316
 — Фейгенбаума 297–301
 Управление динамикой хаотических систем 366
 Уравнение
 — Гинзбурга–Ландау 95, 392
 — Дуффинга 378
 — кинематики волновых фронтов 69
 — Колмогорова–Петровского–Пискунова 399
 — Маккея–Гласса 442
 — фазовой динамики 92–94
 — Фейгенбаума–Цветановича 304
 — Фейгенбаума–Каданова–Шенкера 353
 — характеристическое 213
 Уравнения Гамильтона 116
 Ускорение Ферми 186
 Условие умеренной нелинейности 141
 Устойчивость
 — асимптотическая 212
 — по Ляпунову 211
 — структурная 316
 Фаза 35
 — метастабильная 35
 — зародыш 35
 Фазовая точка 117
 — кривая 117, 166
 Фарей (Фэри) дерево 342
 Фейгенбаума
 — постоянные 300
 — — для консервативных систем 308
 — универсальность 300–305
 Фигура Лихтенберга 271, 272
 Фигуры Лиссажу 124–125
 Фокус
 — неустойчивый 214
 — устойчивый 214
 Фрактал 268–277
 — Вичека 269
 — жирный 277–278
 — толстый 278
 Фрактальная геометрия 268
 Функция
 — Гамильтона 117, 119, 131
 — Мельникова 253
 — универсальная 302–304
 — последования 126
 Хаос 145, 148, 149, 150, 152, 164, 165, 171, 179, 184, 196, 198, 224, 234, 236, 247, 250, 257, 290, 293, 294, 306, 309, 315, 333, 337, 340, 343, 346, 348, 350, 355, 357, 358, 363, 366, 367, 394, 396, 405, 428, 433, 441

- глобальный 150
- контроль 367
- метастабильный 227
- подавление 367, 376–377
- природа 164–165
- пространственно–временной 8–9, 384, 421,
- пространственный 417

- Центр 214
- Цикл
 - отображения 284
 - суперустойчивый 301
 - предельный 216, 241
 - бифуркации 333–336
 - устойчивый 216–221, 241
 - седловой 221, 282
 - абсолютно неустойчивый 216
- Чертова лестница 341, 342

- полная 343
- Число
 - вращения 339
 - степеней свободы 117
- Числа Фибоначчи 351

- Шаг спирали 71
- Шварциан 294

- Элемент
 - автоколебательный 18, 19
 - бистабильный 18
 - возбудимый 18
- Энтропия Колмогорова–Синяя (КС) 235–236
- Эргодичность 168–171

- Языки Арнольда 340–342
 - структура 343
- Ячейка Хеле–Шоу 270, 271, 384–385

*Александр Юрьевич Лоскутов,
Александр Сергеевич Михайлов*

ОСНОВЫ ТЕОРИИ СЛОЖНЫХ СИСТЕМ

*Дизайнер М. В. Ботя
Технический редактор А. В. Ширококов
Компьютерный набор и верстка Д. В. Панкратов
Корректор Г. Г. Тетерина*

Подписано в печать 23.04.2007. Формат 60 × 84¹/₁₆.
Печать офсетная. Усл. печ. л. ????. Уч. изд. л. ?????.
Гарнитура ????. Бумага офсетная №1.
Тираж ???? экз. Заказ №

АНО «Институт компьютерных исследований»
426034, г. Ижевск, ул. Университетская, 1.
<http://rcd.ru> E-mail: mail@rcd.ru Тел./факс: (+73412) 500–295
