

ПОЗНАНИЕ СЛОЖНОГО. ВВЕДЕНИЕ

Книга написана бельгийскими авторами. Один из них, И.Пригожин — ученый с мировым именем, лауреат Нобелевской премии, хорошо известен советскому читателю по многим переведенным на русский язык трудам, в частности монографиям «Термодинамическая теория структуры, устойчивости и флуктуации» (М.: Мир, 1973), «Самоорганизация в неравновесных системах» (М.: Мир, 1979). В предлагаемой книге дан общедоступный обзор методов, разработанных в последние годы в области нелинейной динамики, для изучения сложных систем и процессов, таких, как эволюция, самоорганизация и т. д. Приводятся конкретные примеры из разных областей науки — от химии, физики, биологии до социологии и климатологии.

Для студентов естественнонаучных специальностей, для специалистов широкого профиля.

ОГЛАВЛЕНИЕ

| | |
|---|----|
| От переводчика | 5 |
| Предисловие | 7 |
| Глава 1. Сложность в природе | 10 |
| 1.1. Что такое сложность? | 10 |
| 1.2. Самоорганизация в физико-химических системах: рождение сложного | 12 |
| 1.3. Тепловая конвекция как прототип явлений самоорганизации в физике | 13 |
| 1.4. Явления самоорганизации в химии | 21 |
| А. Реакция БЖ в системе с перемешиванием: химические часы и хаос | 24 |
| Б. Реакция БЖ в неоднородной системе: пространственные фигуры | 29 |
| 1.5. Физико-химическая сложность и алгоритмическая сложность | 35 |
| 1.6. Некоторые дополнительные примеры сложного поведения в обычном масштабе | 37 |
| А. Явления, вызванные поперечным натяжением. Наука о материалах | 37 |
| Б. Кооперативные явления, обусловленные электромагнитными полями. Электрические цепи, лазеры, оптическая бистабильность | 39 |
| 1.7. Снова биологические системы | 42 |
| 1.8. Сложность в планетарном и космическом масштабах | 47 |
| 1.9. Связь между силами и корреляциями. Подведение итогов | 53 |
| Глава 2. Словарь сложного | 58 |
| 2.1. Консервативные системы | 58 |
| 2.2. Диссипативные системы | 63 |
| 2.3. Механическое и термодинамическое равновесия. Неравновесные ограничения | 67 |
| 2.4. Нелинейность и обратные связи | 70 |
| 2.5. Многогранность второго закона термодинамики | 75 |
| 2.6. Устойчивость | 80 |
| 2.7. Бифуркация и нарушение симметрии | 88 |

| | |
|--|-----|
| 2.8. Упорядоченность и корреляции | 92 |
| Глава 3. Динамические системы и сложность | 97 |
| 3.1. Геометрия фазового пространства | 97 |
| 3.2. Меры в фазовом пространстве | 100 |
| 3.3. Интегрируемые консервативные системы | 106 |
| 3.4. Бифуркация в простой диссипативной системе: поиск прототипов сложного | 112 |
| 3.5. Диссипативные системы в двумерном фазовом пространстве: предельные циклы | 117 |
| 3.6. Сведение к системам меньшей размерности: параметры порядка и нормальные формы | 122 |
| А. Бифуркации типа острия и предельной точки | 123 |
| Б. Бифуркация Хопфа и предельные циклы | 123 |
| 3.7. Снова фазовое пространство: топологические многообразия и фракталы | 130 |
| А. Периодические аттракторы: циклы порядка k | 132 |
| Б. Квазипериодические аттракторы: инвариантные торы | 133 |
| В. Непериодические аттракторы: фракталы | 133 |
| 3.8. Неинтегрируемые консервативные системы: новая механика | 137 |
| А. Возмущение квазипериодических движений | 140 |
| Б. Возмущение периодических движений | 141 |
| 3.9. Модель неустойчивого движения: подкова | 143 |
| 3.10. Диссипативные системы в многомерных фазовых пространствах. Хаос и странные аттракторы | 146 |
| А. Некоторые модели-прототипы, приводящие к хаотическому поведению | 147 |
| Б. Некоторые «сценарии» становления хаотического поведения | 151 |
| 3.11. Пространственно распределенные системы. Бифуркации с нарушением симметрии и морфогенез | 156 |
| 3.12. Дискретные динамические системы. Клеточные автоматы | 162 |
| 3.13. Асимметрия, отбор и информация | 167 |
| Глава 4. Случайное и сложное | 173 |
| 4.1. Флуктуации и вероятностное описание | 173 |
| 4.2. Марковские процессы. Основное уравнение | 179 |
| 4.3. Марковские процессы и необратимость. Информационная энтропия и физическая энтропия | 186 |
| 4.4. Пространственные корреляции и критическое поведение | 191 |
| 4.5. Поведение флуктуации во времени. Кинетика и временные масштабы самоорганизации | 198 |
| 4.6. Чувствительность и отбор | 207 |
| 4.7. Символическая динамика и информация | 212 |
| 4.8. Генерация асимметричных, информационно-насыщенных структур | 216 |
| 4.9. Снова алгоритмическая сложность | 223 |

| | |
|--|-----|
| Глава 5. На пути к единой формулировке понятия сложного | 225 |
| 5.1. Общие свойства консервативных динамических систем | 226 |
| 5.2. Общие свойства диссипативных динамических систем | 229 |
| 5.3. Поиски унификации | 230 |
| 5.4. Вероятность и динамика | 232 |
| 5.5. Преобразование пекаря | 233 |
| 5.6. Многообразия с нарушенной временной симметрией | 237 |
| 5.7. Нарушающее симметрию преобразование Λ | 239 |
| 5.8. Ансамбли Гиббса и Больцмана | 243 |
| 5.9. Кинетическая теория | 244 |
| 5.10. Резонанс и взаимодействие света с веществом | 247 |
| 5.11. Заключительные замечания | 249 |
| Глава 6. Сложное и перенос знаний | 251 |
| 6.1. Нелинейная динамика вдали от равновесия и моделирование сложного | 252 |
| 6.2. Наука о материалах | 253 |
| 6.3. Пороговые явления в клеточной динамике | 257 |
| 6.4. Моделирование климатической изменчивости | 261 |
| 6.5. Вероятностное поведение и адаптивные стратегии у общественных насекомых | 268 |
| 6.6. Самоорганизация в человеческих сообществах | 275 |
| Приложение I. Линейный анализ устойчивости | 281 |
| I.1. Основные уравнения | 281 |
| I.2. Принцип устойчивости линеаризованной системы | 284 |
| I.3. Характеристическое уравнение | 285 |
| I.4. Иллюстрации | 288 |
| I.5. Системы с хаотической динамикой | 291 |
| Приложение II. Анализ бифуркаций | 294 |
| II.1. Общие свойства | 294 |
| II.2. Разложение решений в ряд по теории возмущений | 296 |
| II.3. Бифуркационные уравнения | 298 |
| Приложение III. Возмущение резонансных движений в неинтегрируемых консервативных системах | 301 |
| III.1. Отображение закручивания | 301 |
| III.2. Влияние возмущения в случае рациональных вращательных чисел | 303 |
| III.3. Гомоклинические точки | 306 |
| Приложение IV. Реконструкция динамики сложных систем по временной последовательности данных. Применение к климатической изменчивости | 310 |
| IV.1. Вводные замечания | 310 |
| IV.2. Теоретические основы анализа данных | 313 |
| IV.3. Климатический аттрактор | 314 |
| IV.4. Выводы и перспективы | 316 |

| | |
|--|-----|
| Приложение V. Первичные необратимые процессы | 317 |
| V.1. Введение | 317 |
| V.2. Стандартная космологическая модель | 318 |
| V.3. Черные дыры | 319 |
| V.4. Роль необратимости | 320 |
| Литература | 326 |
| Предметный указатель | 336 |

Предметный указатель

| | |
|--|--|
| Автокатализ 29, 33, 52 | <i>Бернулли</i> сдвиг 235 |
| — реакция БЖ 24—34 | Бимодальность переходная 206 |
| — явления саморегуляции 27, 39 | Биологические системы 42 |
| Автономные системы 97, 285 | Биомолекулы 170 |
| Алгоритмическая сложность 36, 37, 223 | Бистабильность 32, 40 |
| Альбедо 52, 265 | — в иммунном ответе 261 |
| Амебы 42—47 | — оптическая 40 |
| Аминокислоты 170 | Бифуркация 88, 112 |
| Антивещество 169 | — анализ 88, 294 |
| <i>Аррениуса</i> закон 33 | — и асимметричные решения 156—160 |
| Архимедова сила 17 | — в диссипативных системах 112 |
| Асимметрия 167 | — зависимость от параметров 114—116 |
| — первичная 156 | — и нарушение симметрии 156 |
| — пространственная 159 | — и пространственные корреляции 196 |
| Аттракторы 84 | — надкритическая 113 |
| — временные последовательности 275 | — поведение во времени 205—206 |
| — глобальные 86 | — предельная точка 123 |
| — климатические 314 | — типа камертона 113 |
| — непериодические 133 | — уравнения 128—130, 298 |
| — периодические 120, 132 | — <i>Хопфа</i> 121, 124 <i>Больцмана</i> |
| — случайность 152—154 | — ансамбли 243 |
| — странные 146 | — кинетическая теория 244 |
| —хаотические 152 | — постоянная 34 |
| Белки 170 | Больших чисел закон 175 |
| <i>Белоусова—Жаботинского</i> реакция 24—34, 89, 125 | Большой взрыв, теория 56, 319 |
| — — и петля гистерезиса 33 | Броуновское движение 244 |
| — — и распространение волны 30 | Брюсселятор, модель 123, 289 |
| — — и хаотические аттракторы 152 | Булева логика 164 |
| — — и химические часы 25 | Возвратные множества 143 |
| <i>Бенара</i> задача 18—21, 29 | Возмущение 140, 296 |

— квазипериодического движения
140

— неинтегрируемое 138—140

— периодических (резонансных)
движений 141

Время нахождения 28

— возвращения 187

Гамильтоновы системы 72

Гауссово распределение 176

Генетический код 171

Гиббса ансамбль 243

Гистерезис 33, 116

Гомоклиническая

— орбита 154

— точка 306

Границы межфазные 37

Грюнайзена параметр 255

Дезоксприбонуклеиновая кислота
(ДНК) 171

Де Суммера пространство 323

Dictyostelium discoideum 44, 71

Дислокации 255

Диссипативная структура 20

Диссипативные системы 26, 63, 117,
229

— — и необратимость 26, 320

Дифференциация внутренняя 91

Диффузия 79

— и нарушающие симметрию
неустойчивости 156—159

— и основное уравнение со многими
переменными 179

— пространственные корреляции
191—194

Зародышеобразование 199

Изолированные системы 69

Иммунные системы, нелинейное
поведение 261

Инвариантность 91

Инвариантные множества 99

Инерциальное многообразие 159

Инерции принцип 63

Информация 92, 156, 167

— позиционная 47

— символическая динамика 212

— хаотическая динамика 219—222

Кантора множество 135

Катастроф теория 117

Квантовомеханический вакуум 249

Кварки 56

Кинетическая

— теория 186, 244

— энергия 61

Кинетический потенциал 199

Кинетическое уравнение 245

Клеточная динамика 257

Клеточные автоматы 162

Климатические изменения 147, 261

Климатический потенциал 266

Колебания

— в диссипативных системах 30—33

— в консервативных системах 62

Консервативные системы 28, 102,
106, 226

— — интегрируемые 106

— — неинтегрируемые 137

— — сравнение с диссипативными
системами 84—86

Константа равновесия 21

Корреляции 92, 191

— крупномасштабные 42—43

— и неравновесные ограничения
196—198

Кортвега—Де Фриза уравнение 108

Космология 317—325

Ландау—Гинзбурга потенциал 198

Лиувилля

— теорема 102, 105, 110, 226

— уравнение 105

Лоренца модель 244

Лотки — Вольтерра механизм 67

Ляпунова показатели 234, 292

— функция 87, 186, 230

Ляпунова—Шмидта процедура 129

- Марковская
— полугруппа 190
— цепь 213
- Марковские процессы 179, 186
- Материаловедение 37, 253
- Машина бифуркационная 168
- Маятник 62
— обратимость 28
— фазовое пространство 62
- Межфазные границы 37
- Мезофаза 254
- Метастабильность 202
- Механика классическая 112
— небесных тел 137
- Множество меры нуль 135
- Молекулярный хаос 20—22
- Морфогенез 156
- Муравьи, поиск пищи 268—275
- Натяжение поверхностное 38
- Нарушение симметрии 18, 91, 167, 237
- Нелинейность 70, 252
- Необратимость 193
— и второй закон термодинамики 75-79
— и марковские процессы 188
- Неоднородность пространственная 156—159
- Неподвижная точка 98
— — гиперболическая 111
— — эллиптическая 110
- Неравновесные состояния 70
- Нормальная форма 122, 129
- Нуклеиновые кислоты 171
- Обратная связь 70
- Ограничение внешнее 15
- Оледенение 49
- Оператор столкновений 245
- Оптика 39—40
- Орбитальная
— неустойчивость 85
— устойчивость 84, 86
- Основное уравнение 179
- Особые точки 99, 117—122
- Открытые системы, химические 28—30
- Отображение типа подковы 143—146
— закручивания 303
- Параметры порядка 122, 129
- Петля обратной связи 46
- Планковское время 324
- Помо—Манвила* пульсация 154
- Предельные циклы 120; *см. также*
Бифуркация
- Преобразование пекаря 233
- Пуанкаре—Биркгофа* теорема 303
- Пуанкаре* секущая поверхность 131
- Пуассона* распределение 93
- Равновесие
— детальное 69
— механическое 67
— термодинамическое 67
— химическое 21
- Разбиение 213
- Разрушение материалов 254
- Реакционно-диффузионные модели 160
- Резонанс 138, 247
— и взаимодействие вещества со светом 247
- Ренорм-группа 198
- Рёсслера* модель 155
- Рибонуклеиновая кислота (РНК) 171
- Самовоспроизведение 24
- Самоорганизация 12, 55
— в геологии 39—40
— в материаловедении 253—257
— в механике жидких сред 18—21
— в химии 24—34
— в человеческих сообществах 275—280
- Сгорание 204
- Седло—фокус 148, 154
- Сепаратриса 111

Снежинка 38
Собственные значения 157
Сродство 79
Субаддитивность 187
Теория относительности 225, 317
Тепловая конвекция 13; *см. Бенара*
задача
Теплопроводность 15
Узлы устойчивые и неустойчивые
118—119
Устойчивость 80
— анализ 281
— асимптотическая 80
— маргинальная 128
— орбитальная 84
Фазовое пространство 99, 100, 130
Фазовые переходы 56—57
Ферменты 46
— кооперативные 46
Фика закон 65
Флуктуации 82
— вероятностное описание 173
— критическое поведение 193—195
— локальный подход 185
— поведение во времени 198
Фокусы устойчивые и неустойчива
118
Фракталы 130, 134
Фредгольма альтернатива 299
Фурье интегралы 246

Хаббла функция 323
Хаотическая динамика 152
Хемотаксис 44
Химический потенциал 69
Хиральность 30
Центральная предельная теорема 175
Центральные многообразия, теория
129
Циклический аденозинмонофосфат
(цАМФ) 44, 46
Часы химические 25
Черные дыры 319
Чувствительность 207
Шредингера уравнение 230
Шум, внешний 178
Эволюции законы 71
Экзотермические реакции 204
Электрические цепи, колебания 39
Эмбриональное развитие 42
Энергия активации 34, 171
Энтропия
— возрастание 78, 80
— и вероятность 180—190
— и информация 186
— и космология 320—323
— как стрела времени 193
— производство, и второй закон
термодинамики 78, 80
Эргодические состояния 187

ОТ ПЕРЕВОДЧИКА

Читателю хорошо известны авторы новой книги своими трудами по теории систем, находящихся вдали от равновесия. Это в значительной мере снимает необходимость составления подробного предисловия. Чем же отличается данная книга от предыдущих, написанных на тему о сложных явлениях? Авторы скромно отмечают, что их цель — всего лишь дать определение понятия сложного. Однако если это понятие считать интуитивно ясным, то по существу их цель — показать, каковы те принципы, которые позволяют свести сложное к простому, т. е. решить основную задачу науки. В качестве примеров сложных явлений в книге рассматриваются различные процессы самоорганизации в физико-химических и биологических системах. Анализ этих примеров позволяет установить то общее, что характерно для сложных явлений, и сформулировать какие-то принципы, достаточно абстрагированные от конкретной природы системы. Особое внимание уделено теории динамических систем, в которой прогресс в понимании природы сложного поведения наиболее заметен. В результате становится значительно более понятным принципиальное различие между консервативными и диссипативными системами, между обратимыми и необратимыми процессами.

Что же ожидает читателя, рискнувшего углубиться в изучение текста? Во всяком случае, среди его впечатлений вряд ли встретится разочарование. В самом деле, разве не приятно познакомиться с такой проблемой, как научное прогнозирование погодных и климатических изменений, из рук наиболее авторитетных ученых? И наступит ли на Земле новый ледниковый период? Каковы современные взгляды на происхождение Вселенной, вещества и пространства-времени? Можно ли методами естественных наук изучать процессы организации в биологических и гуманитарных сообществах?

В связи с обобщающим, натурфилософским характером содержания стиль изложения материала приобретает особый оттенок, напоминающий путешествие по горным хребтам, — крупный план, перспектива и даже непосредственно эмоциональное восприятие — вот особенности этого стиля. Написана она с боль-

шим темпераментом, что придает ей своеобразную эмоциональную окраску и страхует читателя от скуки даже в таких местах, которые представляют собой, казалось бы, традиционно-математические ареалы. Таковы, например, разделы, в которых рассматриваются преобразования типа подковы или пекаря.

На кого же все-таки рассчитана эта книга? На мой взгляд, минимум формально-математического материала, сохраненный в книге, делает ее чрезвычайно широко доступной, по крайней мере для представителей естественнонаучных специальностей, получивших необходимые знания по физике и математике. С другой стороны, некоторые новые ответвления современной математики изложены столь четко, что могут быть рекомендованы даже для первоначального ознакомления с основными идеями и представлениями, как, например, в случае теории фракталов. В целом же, учитывая как доступность, с которой подается материал, так и то обстоятельство, что этот материал представляет собой результат бурного развития ряда направлений в современной физике, химии и математике в последние годы и десятилетия, предлагаемую книгу можно рассматривать как практически необходимый (для естественнонаучных специальностей) и, во всяком случае, весьма желательный элемент культуры современной личности.

В. Пастушенко

ПРЕДИСЛОВИЕ

Символом нашего физического мира не может быть устойчивое и периодическое движение планет, что лежит в основе классической механики. Это мир неустойчивостей и флуктуаций, в конечном счете ответственных за поразительное разнообразие и богатство форм и структур, которые мы видим в окружающей нас природе. Очевидно, для описания такой ситуации, когда эволюция и плюрализм становятся ключевыми словами, необходимы новые понятия и новые средства. Данная монография представляет собой краткое введение в методы, разработанные в течение последних десятилетий для исследования *сложных систем* независимо от того, идет ли речь о молекулах, биологических или социальных системах. Особо подчеркивается роль двух дисциплин, в корне изменивших наши взгляды на представления о сложном. Первая из них — физика неравновесных состояний. Здесь одним из наиболее неожиданных выводов является открытие новых фундаментальных свойств вещества в условиях сильного отклонения от равновесия. Вторая дисциплина — это современная теория динамических систем, в которой центральное место занимает открытие превалирования неустойчивостей. Короче, это означает, что изначально малые изменения с течением времени могут существенно усиливаться.

Развиваемые в этой связи новые методы позволяют лучше понять нашу среду обитания, в которой мы обнаруживаем как неожиданные закономерности, так и неожиданные крупномасштабные флуктуации. Существование материи связано с подавляющим доминированием частиц над античастицами, существование жизни — с доминированием хиральных, асимметричных биомолекул над их антиподами. Каким мог быть селекционный механизм, приведший к столь крупномасштабным закономерностям? Опять же, можно было бы ожидать, что наши климатические условия будут однородными и стабильными. Однако, вопреки таким ожиданиям, климат на Земле сильно флуктуировал на протяжении периодов, малых по сравнению с характерным временем эволюции Солнца. Как такое возможно? Сейчас мы начинаем получать методы, позволяющие непосредственно приблизиться к решению этих вопросов.

В гл. 1 данной монографии представлены избранные примеры сложных явлений, возникающих в физико-химических и биологических системах, равно как и в окружающей среде в целом. Такое описание приводит к ряду встречающихся и в других явлениях понятий, таких, как неравновесность, устойчивость, бифуркация, нарушение симметрии и дальний порядок. В гл. 2 мы возвращаемся к этим понятиям и анализируем их более детально. Они становятся основными элементами того, что мы можем назвать новым научным словарем — *словарем сложного*.

Вслед за этими двумя чисто описательными главами в гл. 3 мы обращаемся к проблеме сложного с точки зрения современной теории динамических систем. В ней обсуждаются некоторые механизмы, посредством которых нелинейная система, отклоненная от равновесия, может порождать неустойчивости, приводящие к бифуркациям и нарушению симметрии. Особое место в нашем анализе отводится внезапному появлению *хаотической динамики* — естественной тенденции широкого класса систем к переходу в состояния, в которых обнаруживаются как детерминистическое поведение, так и непредсказуемость.

В гл. 4 мы пытаемся дать более подробное описание сложных явлений, выходящее за рамки чисто феноменологического уровня предыдущих глав. Здесь излагаются основы вероятностного анализа нелинейных неравновесных систем и строится микроскопическая модель бифуркации и эволюции. Кроме того, обсуждаются некоторые возможности использования понятия информации при описании динамических систем.

С классической точки зрения существовало резкое различие между возможностью и необходимостью, между стохастическим и детерминистическим поведением. Изучение гл. 3 и 4 показывает, что в действительности наблюдается значительно более тонкая ситуация. Имеются различные формы случайности, причем некоторые из них связаны с хаотическим поведением решений простых детерминистических уравнений. В гл. 5 мы обращаемся к вопросам возникновения случайности и необратимости. С ними тесно связана проблема понимания смысла энергии, а на самом деле — само понятие времени. Нам представляется, что мы начинаем приближаться к расшифровке содержания знаменитого второго закона термодинамики. Мы существуем в мире неустойчивых процессов, и это позволяет нам определить функцию энтропии. Более того, мы живем в мире, где симметрия между прошедшим и будущим нарушена, в мире, где необратимые процессы ведут к равновесию в нашем будущем. Это универсальное превалирование нарушения временной симметрии лежит в основе второго закона термодинамики.

Выше мы выразили уверенность в том, что наука призвана играть все более важную роль в попытках понять нашу глобаль-

ную окружающую среду. По этой причине одна из главных целей изложения методов анализа сложных систем, осуществленного в данной монографии, является возможность преодоления междисциплинарных барьеров и попытка взглянуть по-новому на проблемы, из которых некоторые уже «с бородой». В заключительной гл. 6 мы показываем, как может выглядеть такой перенос знаний из одной области в другую. Значительная ее часть посвящена вопросам, находящимся за пределами круга традиционных проблем физических наук, таким, как динамика изменения климата, поведение общественных насекомых и человеческие популяции. Разумеется, каждая из этих проблем имеет свою собственную специфику, поэтому не следует ожидать возможности широкого обобщения соответствующих результатов. Но все-таки и в этих примерах роль нелинейностей и флуктуаций проявляется очень четко. Это наводит на мысль: при моделировании таких систем нужно использовать преимущество новых перспектив, открытых перед наукой исследованиями сложных явлений в нелинейных динамических системах.

Наконец, приложения I—V посвящены количественному обзору некоторых методов, использованных в основных разделах этой монографии.

При подготовке книги мы широко пользовались обсуждениями с многочисленными коллегами и сотрудниками. Более всего мы признательны Э. Ребану за критические замечания по оригиналу рукописи и множество предложений по улучшению текста. Нам приятно также выразить признательность за помощь и предложения, которые высказали К. Базенс, Ф. Барас, Дж. Л. Деннойбург, Й. Элскенс, Р. Файстел, Х. Фриш, Р. Мазо, М. Малек-Мансур, К. Николис, Дж. С. Николис, Г. Рао и С. Рао.

Исследования авторов в области термодинамики необратимых процессов и нелинейной динамики проводились при поддержке Международного сольвеевского института физики и химии, министерства образования и научных исследований Бельгии, комиссии ЕЭС, фонда Уэлча (Хьюстон, шт. Техас) и министерства энергетики США.

*Г. Николис
И. Пригожин*

СЛОЖНОСТЬ В ПРИРОДЕ

В чем состоит различие между маятником и сокращающимся сердцем, между кристаллом воды и снежинкой? Является ли мир физических и химических явлений, где всем наблюдаемым фактам можно дать адекватную интерпретацию на основе небольшого числа фундаментальных взаимодействий, простым и предсказуемым миром? С другой стороны, можно ли сказать, что сложность встречается только в биологии? Эти и другие вопросы будут поставлены в настоящей главе, являющейся вводной по своему характеру. Анализ экспериментальных данных позволит заключить, что различие между физико-химическими и биологическими явлениями, между «простым» и «сложным» поведением не столь резко, как нам это интуитивно представляется. Отсюда в свою очередь вытекает *плюралистический взгляд* на физический мир, где бок о бок сосуществуют различные типы явлений при изменении наложенных на систему условий. Представление о таком открытом мире и есть то главное, что авторам этой книги хотелось бы сообщить читателю.

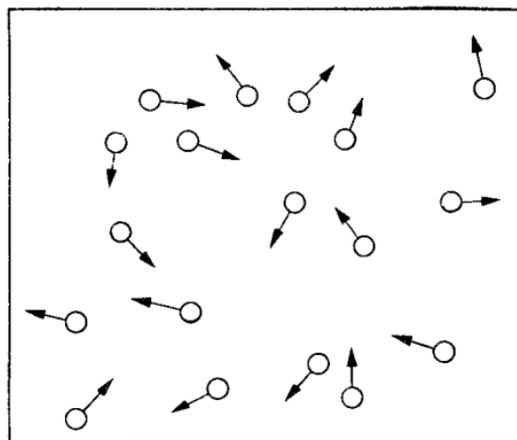
1.1. Что такое сложность!

Сложность представляет собой понятие, принадлежащее нашему каждодневному словарю. С детства мы пользовались этим словом в самых разнообразных контекстах, хотя всегда существовало инстинктивное ощущение, что сложность представляет собой нечто, относящееся к разнообразным проявлениям жизни. Таким образом, для большинства из нас явления, рассматриваемые в руководствах по физике, типа свободного падения тела под действием силы тяжести или колебания маятника являются в своей основе «простыми». Напротив, мы все готовы доказывать, что наша экономическая система, язык, на котором мы разговариваем, мозг млекопитающих или даже простейшая из бактерий являются «сложными системами».

А что мы можем сказать об 1 см^3 газа или жидкости, скажем, воды? Как и в предыдущих примерах, здесь мы имеем дело с системами, в которых содержится огромное число взаимодействующих элементов — молекул. Огромность этого числа не под-

дается воображению. Из элементарной физической химии известно, что при давлении 760 мм рт. ст. и температуре 0°C в одном моле любого газа содержится количество молекул, равное числу Авогадро ($6,023 \cdot 10^{23}$). Это означает, что в 1 см^3 скопилось около 10^{19} , или 10 миллиардов миллиардов, молекул, движущихся во всевозможных направлениях и непрерывно сталкивающихся друг с другом.

Достаточно ли этого, чтобы считать такую систему «сложной»? Несмотря на вышеупомянутые впечатляющие числа, интуиция все-таки подсказывает отрицательный ответ на этот вопрос, поскольку здесь нет каких-либо координированных активности, формы или динамики. По существу такая система выглядит как прототип неупорядоченного, неправильного движения, которое физики часто называют *молекулярным хаосом*, — при



а



б

Рис. 1.1. а — схематическое представление теплового движения молекул в жидкости. Стрелки указывают величину и направление случайно распределенных отдельных скоростей; б — снежинка характерной дендритной формы.

этом молекулы движутся несогласованно, поскольку они не распознают друг друга на расстояниях, превышающих несколько ангстрем (\AA), т. е. несколько стомиллионных сантиметра. В свою очередь этот факт следует отнести на счет основного свойства межмолекулярных сил взаимодействия в данной и многих других встречающихся в природе явлений, а именно *короткодействующего характера сил*.

Представим теперь, что мы поместили 1 см^3 воды в условия, характерные для зимней погоды во время снегопада. Тогда могла бы случиться великолепная снежинка (рис. 1.1, б) с ее характерной дендритной формой. Теперь, созерцая это созданное природой неповторимое произведение искусства, мы смогли бы по-

говорить о сложности. Мы видим, что одна и та же система в разных условиях может выглядеть совершенно по-разному, что поочередно вызывает у нас впечатления «простоты» и «сложности».

Из этого примера следует важный урок: естественнее или по меньшей мере менее двусмысленно говорить о *сложном поведении*, нежели о сложных системах. Можно надеяться, что изучение такого поведения позволит установить то общее, что имеется между различными классами систем, и даст нам лучшее понимание сложного. Эта задача является главной целью настоящей монографии.

1.2. Самоорганизация в физико-химических системах: рождение сложного

Основное, чем занимается наука,— это эксперимент, так что давайте оглянемся и составим некоторое представление об общности и важности сложных явлений.

Как уже отмечалось, то очарование, которое все мы испытываем от биологии, ответственно за несколько расплывчатое отождествление сложности с жизнью в нашем сознании. Хотя это и удивительно, но именно эти представления будут разрушены первыми в нашей погоне за пониманием сложного. В самом деле, с 60-х годов мы были свидетелями революционных достижений как в математике, так и в физике, что ставит *в особое положение работу типа описания природы*. Бывшие в течение многих лет параллельными пути развития термодинамической теории необратимых явлений, теории динамических систем и классической механики в конце концов сошлись. Это убедительно доказывает, что брешь между «простым» и «сложным», между «упорядоченностью» и «разупорядоченностью» гораздо уже, чем думалось раньше. Как известно, простые примеры из вузовской программы по механике не могут продемонстрировать сложного поведения. Маятник, к которому приложена периодическая возмущающая сила, на границе между вибрацией и вращением приводит к богатому разнообразию типов движения, включая возможность случайных квазитурбулентных отклонений от равновесного положения. В таких обычных системах, как слой жидкости или смесь химических продуктов, при определенных условиях могут возникать макроскопические явления *самоорганизации* в виде ритмически изменяющихся во времени пространственных картин. Короче, ясно, что сложность присуща не только биологии. Она вторгается в физические науки, и похоже, что ее корни уходят глубоко в законы природы.

В результате этих открытий интерес к макроскопической физике, т. е. физике явлений, протекающих в привычных нам мас-

штабах, возрастает чрезвычайно. В последующих разделах этой главы будут описаны некоторые конкретные явления такого рода, и мы будем возвращаться к ним на протяжении всей книги.

1.3. Тепловая конвекция как прототип явлений самоорганизации в физике

Прежде всего мы рассмотрим объемные движения текучих сред (жидкостей или газов) под действием температурных неоднородностей. Такие типы движения известны под названием *тепловой конвекции*.

Исследование этих видов движения, даже вне рамок того специального предмета, которому мы будем уделять внимание, далеко не носит чисто академического характера. Даже не перечисляя многочисленных технологических применений, можно сразу отметить, что тепловая конвекция лежит в основе нескольких важных и впечатляющих явлений, наблюдаемых на нашей планете. Одним из таких примеров является циркуляция атмосферы и океанов, в значительной мере определяющая погодные изменения на короткие и средние сроки. Другой пример — дрейф континентов, т. е. движение континентальных платформ под действием крупномасштабных движений мантии. Хотя это и несколько дальше от нас, на Солнце тепловая конвекция лежит в основе переноса тепла и вещества, что в свою очередь в значительной мере определяет солнечную активность. В лабораторных условиях можно наблюдать, как тепловая конвекция приводит к некоторой организации более скромных размеров. Следующий простой эксперимент, впервые реализованный Бенаром, приводит к наблюдению целого ряда удивительных свойств.

Представим себе слой жидкости (например, воды) между двумя горизонтальными параллельными плоскостями, латеральные размеры которых значительно превосходят толщину слоя. Предоставленная самой себе, жидкость быстро устремится к однородному состоянию, в котором, выражаясь языком статистики, все ее части будут тождественны между собой. Например, помещенный в такой слой миниатюрный наблюдатель, который строит свои суждения лишь на наблюдениях над окружающей средой, не сможет определить, в каком из двух малых объемов (V_A или V_B) он находится (рис. 1.2). Таким образом, все объемы, которые можно выделить в жидкости, будут неразличимы между собой. Соответственно, чтобы знать состояние всех таких объемов, достаточно знать состояние одного из них независимо от их формы и размера. Иными словами, с точки зрения нашего наблюдателя ему безразлично занимаемое им положение. Или по-другому — нет внутренних возможностей, которые позволили бы ему воспринять понятие пространства.

Разумеется, однородность этой системы распространяется на все ее свойства, и в частности на ее температуру, которая будет одинаковой во всех частях жидкости и равной температуре ограничивающих ее плоскостей, т. е. температуре «внешней среды».

Все эти свойства характерны для системы в особом состоянии, в котором нет ни объемных движений, ни разности температур со внешним миром. Можно дать следующее, более количественное выражение этого свойства. Обозначим через T_1 и T_2 температуры плоскостей 1 и 2 соответственно. Тогда при равновесии мы имеем

$$\Delta T_e = T_2 - T_1 = 0 \quad (1.1)$$

Представим теперь, что некто на короткое время касается пальцем одной из плоскостей. Температура этой части плоскости моментально изменится (например, от 20°C до температуры человеческого тела $36,5^\circ\text{C}$). Подобное событие, происходящее в

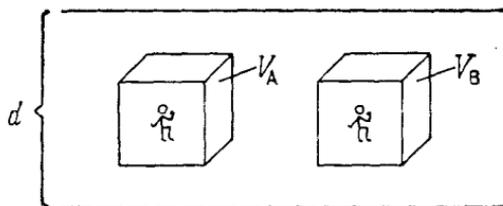


Рис. 1.2. Миниатюрный наблюдатель, изучающий состояния элементов объема V_A и V_B в тонком слое жидкости, находящейся в равновесии, выявляет неразличимость этих состояний и приходит к выводу о трансляционной инвариантности вдоль горизонтального направления.

системе случайно и локально изменяющее (в общем случае слабо) некоторые из ее характеристик или свойств, называется *возмущением*. На нашу систему, находящуюся в равновесии, это температурное возмущение не окажет влияния, поскольку температура быстро станет однородной и равной ее первоначальному значению (в нашем примере 20°C). Иными словами, возмущение затухает, не оставляя следов в системе. Когда система находится в состоянии, в котором действующие на нее возмущения затухают во времени, мы говорим, что состояние является *асимптотически устойчивым*.

Вернемся к нашему миниатюрному наблюдателю, которому однородность жидкости не позволяет выработать собственное представление о пространстве. Выясняется, что дополнительно устойчивость равновесного состояния делает в конечном счете тождественными еще и все моменты времени. Поэтому наш наблюдатель не может получить собственное представление и о времени. Вряд ли можно говорить о поведении применительно к

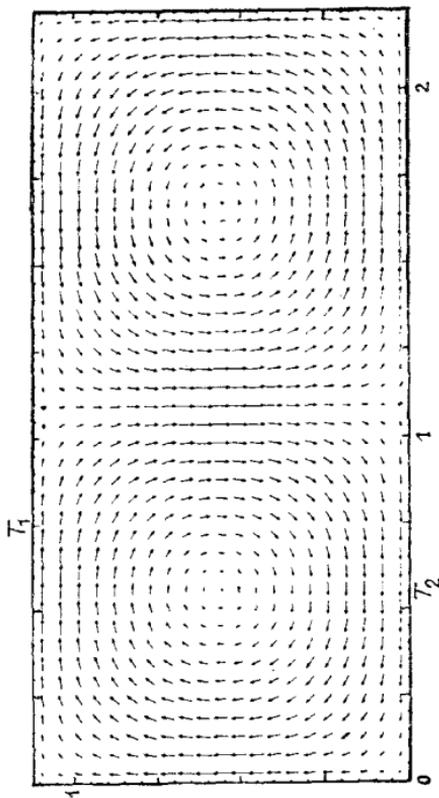
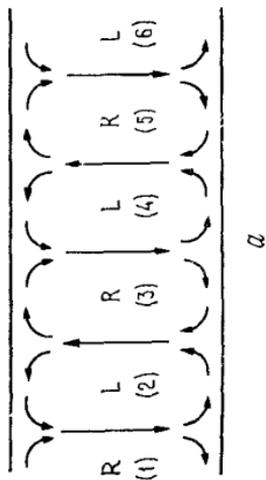
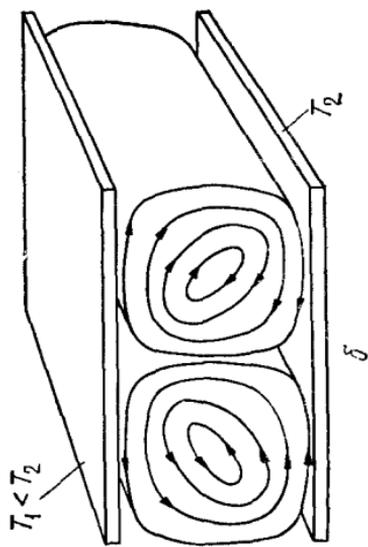
системе в такой *простой* ситуации. Поэтому давайте несколько отклонимся от нее.

Сделать это можно, например, нагреванием жидкого слоя снизу. При этом мы сообщаем энергию системе в виде тепла. Более того, поскольку температура нижней плоскости T_2 выше, чем T_1 , условие равновесия (1.1) нарушается ($\Delta T > 0$). Иными словами, мы не позволяем системе достичь равновесия, налагая на нее *внешнее ограничение*. Заметим, что в данном примере внешнее ограничение подразумевает поток энергии и наоборот.

Допустим вначале, что ограничение слабое (ΔT мало). В системе снова установится простое и единственное состояние, в котором единственный протекающий процесс сводится к переносу тепла от слоя жидкости к верхней плоскости, от которой тепло будет передаваться во внешнюю среду для обеспечения постоянства T_1 . Единственное отличие этого состояния от равновесного будет состоять в том, что температура, а с ней и плотность и давление не будут более однородными. Они будут практически линейно изменяться от теплой области (внизу) к холодной (вверху). Это явление известно под названием *теплопроводности*. В этом новом состоянии, достигаемом системой в ответ на внешнее ограничение, опять будет превалировать стабильность, и поведение системы в конечном счете будет столь же простым, как и в равновесии.

Все дальше отклоняя систему от равновесия путем увеличения ΔT , мы увидим, что внезапно, при некотором значении ΔT , которое мы назовем *критическим* (ΔT_c), объем вещества приходит в движение. Более того, это движение далеко не случайное: жидкость структурируется в виде небольших ячеек, называемых ячейками Бенара (рис. 1.3, а). Это и есть режим тепловой конвекции, определение которой было дано в начале настоящего раздела.

На рис. 1.4 иллюстрируется качественное объяснение этого явления. Вследствие теплового расширения жидкость расслаивается, причем часть жидкости, находящаяся ближе к нижней плоскости, характеризуется пониженной плотностью по сравнению с верхними слоями. Это приводит к градиенту плотности, направленному противоположно силе тяжести. Легко понять, что такая конфигурация потенциально неустойчива. Рассмотрим, например, малый объем жидкости вблизи нижней плоскости. Вообразим теперь, что этот элемент объема немного смещается вверх вследствие возмущения. Находясь теперь в более холодной и, следовательно, в более плотной области, этот элемент будет испытывать направленную вверх архимедову силу, которая будет стремиться усилить восходящее движение. С другой стороны, если находящаяся вначале у верхней плоскости малая капля смещается вниз, то она проникнет в область пониженной



б

Рис. 1.3. Два изображения конвективных (бенаровских) ячеек. Обратите внимание на противоположные направления вращения в двух соседних ячейках, дополнительно показанные на рис. в.

плотности, и архимедова сила будет ускорять нисходящее движение. Поэтому в принципе ясно, что в жидкости могут возникать восходящие и нисходящие потоки, как это и наблюдается в эксперименте. Изложенные выше соображения вроде бы позволяют наблюдать такие потоки при любом отклонении от нуля температурного скачка ΔT . Причина, по которой такие потоки не наблюдаются при малых ΔT , связана со стабилизирующим влиянием вязкости жидкости: в результате этого в жидкости возникают внутренние силы трения, направленные против движения. Стабилизирующим фактором оказывается и теплопроводность, вследствие которой разность температур между смещенной каплей и ее окружением стремится исчезнуть. Это объясняет существование критического значения перепада температуры ΔT_c , наблюдаемого экспериментально. На рис. 1.3, б показано, насколько сложно возникающее движение: в некоторой

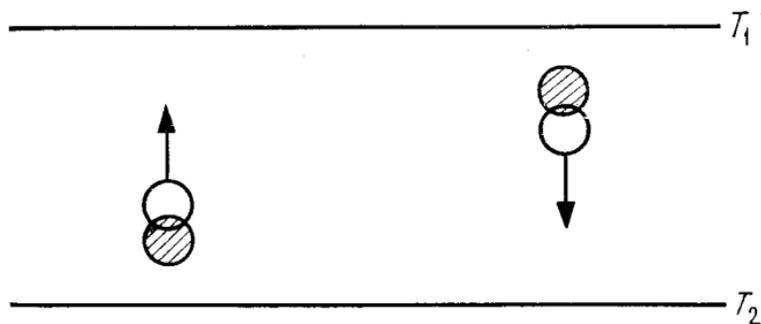


Рис. 1.4. Качественное объяснение природы тепловой конвекции.

точке жидкость движется вверх, проходит вдоль плоскости 1, затем идет вниз, движется мимо плоскости 2, идет опять вверх и т. д. Ячейки выстраиваются вдоль горизонтальной оси, причем жидкость в ячейках приходит последовательно во вращение то по (R), то против (L) часовой стрелки.

Вы, пожалуй, скажете, что эта сложность выглядит довольно скромно по сравнению со сложностью, присущей простейшей из бактерий. Но вернемся еще раз к нашему миниатюрному наблюдателю!

Окружающий его мир совершенно изменился. Например, он теперь может решить, где он находится, наблюдая характер вращения в той ячейке, где он находится в настоящий момент. Более того, подсчитывая количество ячеек, которые он минует, перемещаясь по жидкости, он вполне может получить достаточно полное представление о пространстве. В таких случаях, когда в некоторой системе возникает возможность выработать представление о пространстве (а до того такой возможности не было),

мы будем говорить о *нарушении симметрии*. В некотором смысле нарушение симметрии приводит нас от статического, геометрического представления о пространстве к аристотелевским взглядам, согласно которым свойства пространства определяются событиями, совершающимися в данной системе координат.

По-видимому, наиболее примечательной чертой, которую следует отметить в таком внезапном переходе от простого поведения к сложному, являются *упорядоченность* и *согласованность* системы. Когда ΔT было ниже критического значения ΔT_c , однородность жидкости в горизонтальном направлении делала независимыми друг от друга различные ее части. Так, объемы V_A и V_B можно было бы поменять местами (рис. 1.2) без каких-либо последствий. В частности, наблюдаемые свойства объема V_C , помещенного между объемами V_A и V_B , никоим образом не изменятся вследствие того, что объем А будет теперь справа, а не слева. Напротив, выше порогового значения ΔT_c все происходит так, как если бы каждый элемент объема следил за поведением своих соседей и учитывал его с тем, чтобы играть нужную роль в общем процессе. Такая картина предполагает наличие *корреляций*, т. е. статистически воспроизводимых соотношений между удаленными частями системы. К этому ключевому выводу мы приходим здесь, непосредственно наблюдая данное явление. К этому вопросу в целом и к его детальному анализу мы вернемся в гл. 2 и 4. Тем не менее уже на этой ранней стадии нашего анализа следует обратить внимание на различие между дальнедействующим характером этих флуктуаций и короткодействующим характером межмолекулярных сил. Характерные размеры ячеек Бенара в обычных лабораторных условиях находятся в миллиметровом диапазоне (10^{-1} см), в то время как характерный пространственный масштаб межмолекулярных сил приходится на ангстремный диапазон (10^{-8} см). Иначе говоря, отдельная ячейка Бенара содержит что-то около $(10^7)^3 \sim \sim 10^{21}$ молекул. Тот факт, что такое огромное число частиц может демонстрировать когерентное поведение, несмотря на случайное тепловое движение каждой из частиц, является одним из основных свойств, характеризующих возникновение сложного поведения.

Отметим, что в используемый нами на протяжении двух последних страниц словарь входили такие понятия (когерентность, сложность, упорядоченность), которые уже давно являются составной частью биологии и до сравнительно недавнего времени находились за пределами основного русла физики. Возможность описывать с помощью этих фундаментальных понятий поведение как живых существ, так и обычных физических систем является одним из основных достижений, которые, по-видимому, наука не смогла бы предсказать еще несколько лет тому назад.

Однако этим не исчерпывается все то удивительное, что связано с ячейками Бенара. С одной стороны, такой эксперимент характеризуется идеальной воспроизводимостью, поскольку при одних и тех же условиях превышение некоторого критического значения ΔT_c всегда приводит к возникновению конвекционной картины. С другой же стороны, как видно из рис. 1.3, вещество структурируется в ячейки с попеременно право- и левовращательным движением. Однажды установившись, направление вращения в дальнейшем сохраняется. И все же, сколько бы мы ни изошрялись в воспроизведении условий эксперимента, по достижении критического порогового значения ΔT_c возможны две качественно различные ситуации*. В ячейке 1 на рис. 1.3, а имеет место вращение по часовой стрелке (и тем самым в ячейке 2 — против часовой стрелки, в ячейке 3 — по часовой стрелке и т. д.). Однако могло бы быть и наоборот, так что в ячейке 2 было бы вращение по часовой стрелке с соответствующими изменениями в других ячейках. Как только ΔT слегка превысит ΔT_c , мы совершенно точно знаем: появляется ячеистая структура течения. Таким образом, это явление подвержено строгому детерминизму. Напротив, направление вращения в ячейках непредсказуемо и неуправляемо. Лишь случай в виде тех или иных возмущений, доминирующих в момент проведения эксперимента, решает, каким будет вращение в данной ячейке — право- или левовращательным. Таким образом, мы приходим к удивительному сотрудничеству между случайностью и определенностью, наводящему на мысли об аналогичном дуализме, известном в биологии со времен Дарвина (мутация — естественный отбор). В области физики такой дуализм до сих пор наблюдался лишь при квантовомеханическом описании микроскопических явлений. Высказанные соображения суммируются на рис. 1.5.

Таким образом, мы видим, что вдали от равновесия, т. е. при существенном ограничении степеней свободы, система может приспособляться к своему окружению несколькими различными способами или, выражаясь менее антропоморфически, при одних и тех же значениях параметров возможно несколько различных решений. И лишь случай решает, какое из этих решений будет реализовано. Тот факт, что из многих возможных вариантов был выбран лишь один, придает системе историческую размерность, своего рода «память» о прошлом событии, произошедшем в критический момент и оказавшем влияние на дальнейшую эволюцию.

* На самом деле следовало бы говорить о бесконечном числе возможностей, поскольку в большой системе структура допускает трансляции на любую длину (нарушение непрерывной группы симметрии). Однако наиболее ярким проявлением этой множественности остается направление вращения в ячейке, и именно это обстоятельство выбрано в качестве критерия в данном разделе.

Что же произойдет в том случае, когда тепловое ограничение степеней свободы превысит первый порог структурирования? В некоторой области значений ΔT ячейки Бенара будут существовать как таковые, однако некоторые их характеристики изменятся. Затем, после перехода через другое критическое значение ΔT_c , мы внезапно окажемся свидетелями нового мощного проявления случайности: структура течения станет размытой и возникнет новый режим, характеризующийся неупорядоченной зависимостью переменных от времени. Этот режим предшествует тому, что инженеры и специалисты по гидродинамике в течение последнего столетия называли *турбулентностью*. В более общем случае можно сказать: сейчас ясно, что турбулентность — это лишь один из аспектов общей тенденции целого класса систем к переходу в *хаотический* режим при определенных условиях. Подробнее этот вопрос будет обсуждаться в гл. 3.

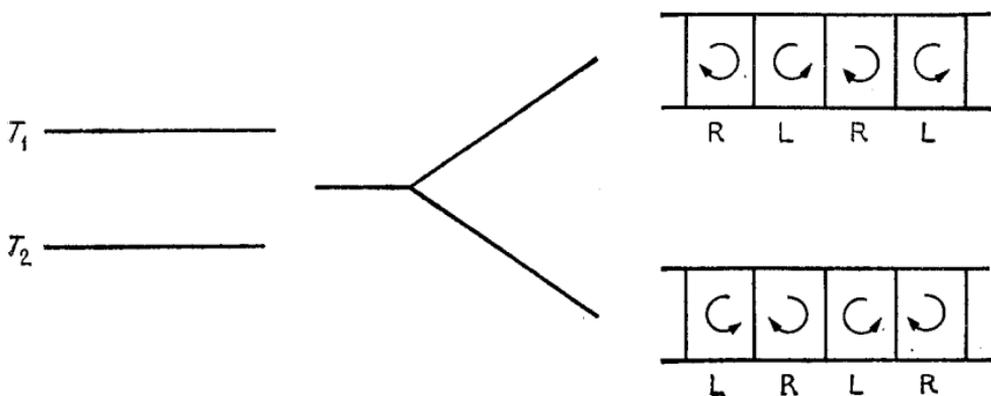


Рис. 1.5. Множественность решений выше порога неустойчивости тепловой конвекции. В зависимости от конкретных условий данная пространственная область может оказаться частью ячейки с левой или правой ориентацией вращения.

Итак, мы видели, что неравновесность позволила системе избежать тепловой разупорядоченности, изображенной на рис. 1.1, а, и трансформировать часть энергии, сообщаемой внешней средой, в упорядоченное поведение нового типа — *диссипативную структуру*, характеризующуюся нарушением симметрии, множественными выборами и корреляциями в макроскопических масштабах. Поэтому можно сказать, что мы буквально оказались свидетелями рождения сложного. Правда, достигнутая сложность имеет довольно скромный характер, но тем не менее она наделена такими особенностями, которые обычно приписываются исключительно биологическим системам. Важнее то обстоятельство, что сложность отнюдь не противоречит законам физики; более того, она является неизбежным следствием этих законов при выполнении определенных условий.

1.4. Явления самоорганизации в химии

По традиции мы рассматривали ячейки Бенара с точки зрения физики главным образом по той причине, что химическая природа вещества, образующего жидкий слой, не изменяется за время наблюдения. Рассмотрим теперь химические реакции, т. е. такие процессы, которые приводят к превращениям образующих данное вещество молекул. Интерес к таким явлениям представляется довольно общим. Значительная часть химической промышленности основана на гетерогенном катализе, ускоряющем некоторые стадии синтеза в присутствии соответствующей поверхности, на которой происходит реакция. Например, реакцию окисления аммиака обычно проводят в присутствии платинового катализатора, в то время как для разложения N_2O используются каталитические свойства поверхности оксида меди. Другим классом важных химических превращений является сгорание, например углеводородов, благодаря которому работают тепловые машины. Кроме того, некоторые проявления биологической упорядоченности также связаны с химическими реакциями, протекающими в присутствии специальных катализаторов — ферментов.

В некоторой типичной химической реакции молекула типа А (скажем, молекула гидроксила OH)* может соединиться с молекулой типа В (скажем, молекулой водорода H_2), в результате чего возникают одна молекула типа С и одна молекула типа D (соответственно H_2O и атомарный водород в нашем примере). Символически этот процесс можно записать в виде



где k — константа скорости, в общем случае зависящая от температуры и давления. Фигурирующие в левой части уравнения реакции *реагенты* А и В соединяются и со временем исчезают, а *продукты* С и D (правая часть уравнения) образуются по мере протекания реакции. Тем не менее опытным путем установлено, что даже после длительного ожидания в изолированной системе реагенты А и В не исчезают полностью никогда. Точнее, количества сосуществующих компонент А, В, С и D характеризуются определенным значением отношения $c_{CCD}/(c_{ACB})$, устанавливающимся по истечении достаточно большого времени (c с соответствующим индексом обозначает концентрацию данного компонента). Это отношение называется *константой равновесия*, причем, когда оно достигается в системе, мы говорим о состоянии *химического равновесия*. Это состояние является аналогом однородного состояния покоя в задаче Бенара. Разумеет-

* Очевидно, выбранный авторами пример служит лишь для формальной иллюстрации рассуждений. — Прим. перев.

ся, при этом значения физических параметров типа давления, температуры и т. д. предполагаются постоянными.

Как этот результат можно согласовать с уравнением (1.2а), согласно которому компоненты А и В исчезают при связывании? Экспериментально выяснено существование превращения, обратного (1.2а):



В равновесии обе реакции протекают с совершенно одинаковыми скоростями. Это фундаментальное свойство природы, известное под названием *детального равновесия*, ответственно за многие особенности равновесного состояния вещества. Обратимую реакцию можно представить следующим образом:



В задаче Бенара отклонение от равновесия достигалось путем сообщения системе энергии в виде непрерывного потока. Наиболее очевидный химический аналог такой процедуры состоит в установлении потока массы из окружающей среды к системе (или наоборот), реализуя тем самым то, что впоследствии будет называться *открытой системой*. Этого можно достичь, например, путем удаления компонентов С или D из реакционного объема, когда их концентрации превышают некоторые заранее заданные значения, или же путем пропускания через реакционный объем смеси, обогащенной компонентами А и В, с последующей эвакуацией или возвращением в повторный цикл. На рис. 1.6 показана схема функционирования открытой системы в лабораторных условиях.

Комбинируя соответствующим образом скорости притока и оттока вещества, можно создать такие условия, когда в системе устанавливаются постоянные концентрации А, В, С и D, причем их отношение уже не определяется константой равновесия. Математически такое состояние описывается условием обращения в нуль скоростей изменения концентраций, т. е. их производных по времени:

$$\frac{dc_A}{dt} = \frac{dc_B}{dt} = \dots = 0 \quad (1.3)$$

В таких случаях мы будем говорить о *стационарном неравновесном состоянии*.

Можно ли ожидать, что поведение открытой системы вблизи такого неравновесного состояния будет подобно тому, которое наблюдается в случае задачи Бенара? Заметим, что в такой системе детальное равновесие уже не имеет места. Предположим

на мгновение, что какой-то конкретный процесс, скажем прямая реакция, ускоряется по сравнению с обратной реакцией. В случае равновесия система стремилась бы восстановить начальное положение дел. Однако вдали от равновесия ситуация изменяется. Более того, если бы некоторые части химического механизма позволяли системе улавливать и далее усиливать упомянутое ускорение, то возникла бы потенциально неустойчивая ситуация, сходная с изображенной на рис. 1.4 для случая задачи Бенара. Известно, что такие механизмы в химии существуют, и наиболее ярким их проявлением служит *автокатализ*. Например, присутствие продукта может увеличивать скорость его собственного производства. По существу, это кажущееся экзотическим явление довольно обычно в любом процессе горения благодаря присутствию свободных радикалов — чрезвычайно ак-

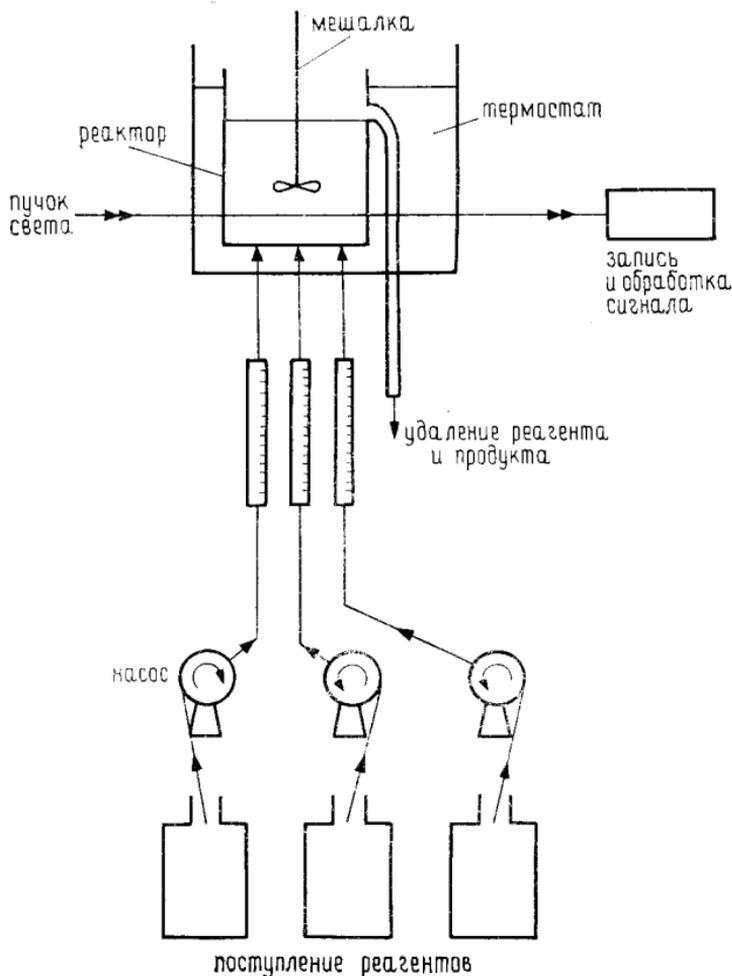


Рис. 1.6. Экспериментальная реализация открытой химической системы.

тивных молекул с неспаренным электроном, которые, реагируя с другими молекулами, приводят к дальнейшему увеличению количества свободных радикалов и тем самым к самоускоряющемуся процессу. Другим примером является *самовоспроизведение* — одно из наиболее характерных свойств жизни. В своей основе оно является результатом некоторого автокаталитического цикла, в котором репликация генетического материала осуществляется благодаря участию специфических белков. В свою очередь сами эти белки синтезируются благодаря инструкциям, содержащимся в генетическом материале.

Довольно долго химики полагали, что какое бы химическое превращение ни рассматривалось, в конечном счете в системе должно установиться однородное и независящее от времени состояние, подобное равновесному. Любое отклонение от этого правила считалось возможным лишь в результате какой-то хитрости, какой-то забавной диковины, возникающей вследствие влияния некоторого явления, не относящегося собственно к области химии.

Рассмотрим теперь некоторую химическую реакцию, известную под названием реакции Белоусова — Жаботинского (сокращенно БЖ), которая при определенных неравновесных условиях демонстрирует целый спектр очаровательных и неожиданных режимов поведения.

Собственно реагенты, участвующие в реакции БЖ, не представляют собой ничего особенного. Типичный препарат состоит из сульфата церия $\text{Ce}_2(\text{SO}_4)_3$, малоновой кислоты $\text{CH}_2(\text{COOH})_2$ и бромата калия KBrO_3 , растворенного в серной кислоте. Происходящие в системе процессы можно наблюдать невооруженным глазом благодаря окрашивающему веществу типа, например, феррона, дающего красный цвет при избытке ионов Fe^{2+} (или Ce^{3+}) и голубой при избытке Fe^{3+} (или Ce^{4+}). Более изощренные методы наблюдения реализуются при помощи специфических электродов или спектроскопическими методами, основанными на измерении поглощения каким-то конкретным веществом.

Рассмотрим теперь различные типы поведения, демонстрируемые этой системой в различных экспериментальных условиях при комнатной температуре.

А. Реакция БЖ в системе с перемешиванием: химические часы и хаос

Допустим сначала, что реакция протекает при условиях, показанных на рис. 1.6. Благодаря интенсивному переносу вещества вследствие перемешивания в каждый момент времени система остается практически однородной в пространстве. Такое

экспериментальное устройство позволяет также довольно легко управлять удаленностью системы от равновесия: достаточно изменить скорости, с которыми химические вещества поступают в систему (или выбывают из нее), меняя тем самым *время пребывания* этих веществ в реакционном объеме. Очень большие времена пребывания реагентов приводят по существу к реализации замкнутой системы, и в таких условиях можно ожидать, что поведение системы будет подобно равновесному, характеризующемуся детальным равновесием. Уменьшая время пребывания, мы не допускаем полного выравнивания скоростей прямой и обратной реакций. При этом можно ожидать, что поведение системы будет неравновесным. Именно это и показывает эксперимент. В случае очень больших времен пребывания в системе достигается однородное стационарное состояние — концентрации остаются постоянными во времени. Это типичное состояние, весьма привычное химикам, наделено всеми качественными свойствами химического равновесия. Оно является аналогом режима теплопроводности, реализуемого в системе Бенара при небольшой разности температур между пластинами.

Если теперь уменьшить время пребывания, мы встретимся с совершенно иным типом поведения. В какой-то момент времени вся система внезапно окрашивается в голубой цвет (если в качестве красящего вещества используется ферроин), что указывает на избыток ионов Fe^{3+} (или Ce^{4+}). Спустя несколько минут (или в зависимости от условий — долю минуты) голубой цвет сменяется красным, указывая на избыток ионов Fe^{2+} (или Ce^{3+}). Этот процесс так и продолжается: голубой, красный, голубой, красный и т. д. — ритмическая смена цвета с идеально регулярными периодом и амплитудой, зависящими лишь от параметров и тем самым являющимися собственными характеристиками системы. Эти колебания можно рассматривать как *химические часы* — устройство для измерения времени с помощью внутренней динамики системы. Типичное поведение такого рода показано на рис. 1.7.

На этой стадии трезвый наблюдатель может задать естественный вопрос: с чего бы это ритмическое поведение должно приводить нас в такое замешательство? Разве мы не наблюдаем такие же явления ежедневно, например в случае маятника? В самом деле, любой курс механики начинается с колебаний идеального маятника, предваряющих изучение таких фундаментальных принципов, как второй закон Ньютона и сохранение энергии в момент количества движения!

Глубокое различие между этими двумя типами колебаний иллюстрируется рис. 1.8. Верхняя часть рисунка показывает колебания маятника без трения относительно вертикального направления с максимальным углом отклонения θ_1 [амплитуда перио-

дической функции $\theta(t)$, представляющей мгновенное значение угла] и периодом T_1 . Там же показано изменение во времени концентрации некоторого реагента в реакции БЖ, характеризующее амплитудой A и периодом T .

Рассмотрим теперь мгновенные возмущения в обеих системах, соответственно отклоняя маятник на больший угол от вертикали θ_2 ($\theta_2 > \theta_1$) и накладывая небольшой концентрационный или температурный импульс на реакцию БЖ (например, инжестируя несколько миллимолей бромата калия или прикасаясь к сосуду в течение нескольких секунд). В нижней части рис. 1.8 показаны ответы обеих систем на эти возмущающие воздействия. Маятник будет снова совершать колебания, однако их амплитуда будет теперь равна θ_2 вместо θ_1 , а период несколько увеличится. Иными словами, эта система сохранит память о возмущении навсег-

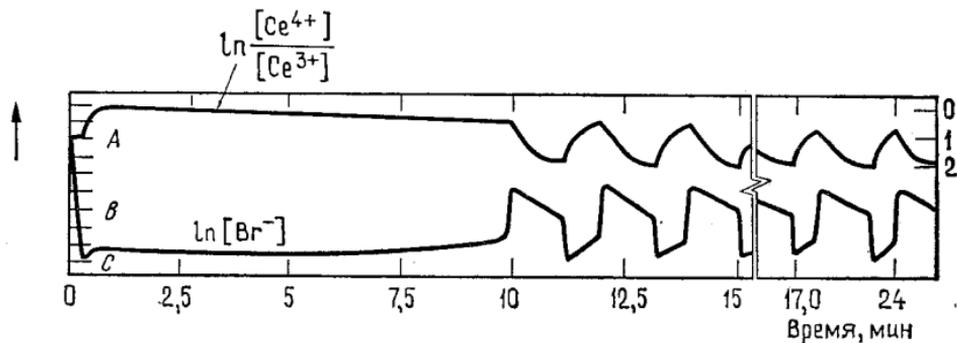


Рис. 1.7. Зависимость $\ln[\text{Br}^-]$ и $\ln\left(\frac{[\text{Ce}^{4+}]}{[\text{Ce}^{3+}]}\right)$ от времени, снятая потенциометрическим методом в ходе реакции Белоусова — Жаботинского. Начальные значения концентраций: $\text{CH}_2(\text{COOH})_2 = 0,032 \text{ M}$, $\text{KBrO}_3 = 0,063 \text{ M}$, $\text{Ce}(\text{NH}_4)_2(\text{NO}_3)_6 = 0,01 \text{ M}$, $\text{H}_2\text{SO}_4 = 0,8 \text{ M}$ и $\text{KBr} = 1,5 \cdot 10^{-5} \text{ M}$.

да. Напротив, в случае реакции БЖ рассматриваемый реагент после некоторого переходного процесса будет совершать колебания с точно теми же амплитудой и периодом, что и до возмущения. Это свойство системы, называемое *асимптотической устойчивостью*, уже затрагивалось в предыдущем разделе. Оно непосредственно связано с вездесущим свойством большинства явлений, наблюдаемых в природе, а именно *необратимостью*, и существенным образом ответственно за воспроизводимость событий. С другой стороны, системы типа маятника не обладают этим свойством, так как их динамика инвариантна относительно обращения времени. Поэтому такие системы всецело зависят от возможных возмущений, и, поскольку эти возмущения существенно непредсказуемы, поведение таких систем становится раньше или позже разупорядоченным.

Более систематическое изучение различий между *диссипативными системами*, демонстрирующими необратимость подобно

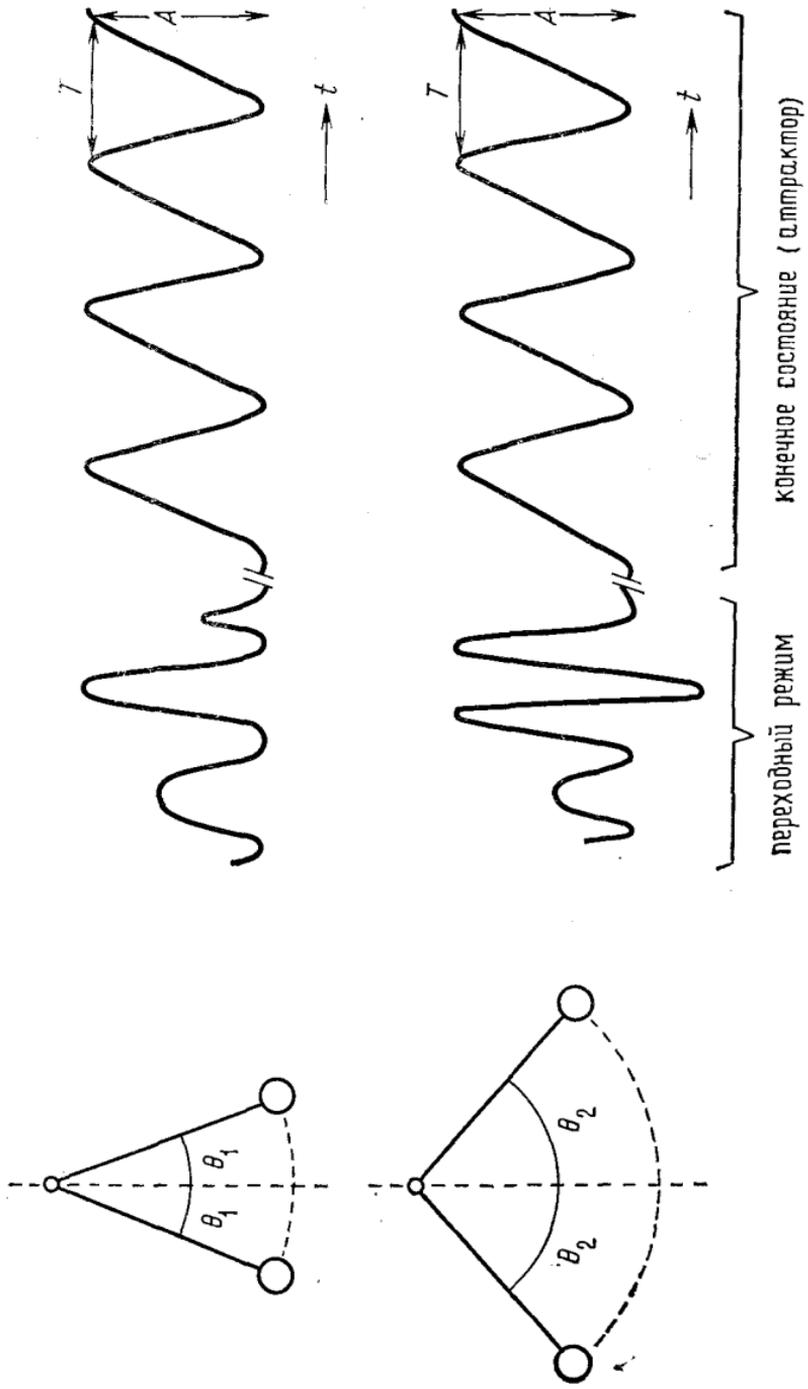


Рис. 1.8. Иллюстрация различия между незатухающими колебаниями в консервативных системах типа маятника (слева) и диссипативных системах типа реакции Белоусова — Жаботинского (справа).

течению Бенара или реакции БЖ, и *консервативными системами* типа маятника, динамика которых инвариантна относительно обращения времени, мы отложим до гл. 3. Эти вводные слова хочется заключить еще одним замечанием. Всем нам хорошо знаком осциллятор, с которым мы не расстанемся уже с эмбрионального периода, — это наше сердце. Его биение более или менее регулярно (в среднем для человека около 70 или 80 ударов в минуту). Однако эта регулярность может быть нарушена. Особо драматичной разновидностью такого нарушения является фибрилляция — неспособность совершать колебания, захватывающие все сердце в целом. В самом деле, какой бы ни была в конечном счете причина смерти, она всегда сопровождается состоянием фибрилляции сердца. Теперь нам ясно, что, если бы наше сердце работало подобно маятнику, фибрилляция наступила бы задолго до нашего рождения. Тот факт, что биение сердца (а на самом деле любые другие воспроизводимые ритмические явления, наблюдаемые в природе, — от циркадных ритмов и цикла клеточного деления до изменения яркости переменных звезд — цефеид) должно относиться к тому же кругу явлений, что и колебания в реакции БЖ, показывает чрезвычайно важную роль, которую играют необратимость и диссипативные системы.

Вернемся теперь к химическим часам. Пользуясь терминами, введенными в разд. 1.3, можно сказать, что в режиме однородного стационарного состояния (которое также асимптотически устойчиво!) система игнорирует время. Однако в периодическом режиме она вдруг «открывает» для себя время в виде фазы периодического движения и соответственно в заранее заданной последовательности максимумов различных концентраций, сменяющих друг друга. Такой переход мы будем называть *нарушением временной симметрии*.

С еще более фундаментальной точки зрения существование незатухающих колебаний в системе как в целом предполагает, что ее отдельные части действуют согласованным образом путем строгого поддержания между ними определенных соотношений между фазами — в противном случае внешние возмущения сделали бы колебания невозможными. Иными словами, как и в задаче Бенара, мы снова ожидаем, что неравновесность должна привести к *крупномасштабным корреляциям*. Появление таких корреляций и их характеристики будут значительно более подробно рассмотрены в гл. 2 и 4.

Однако все то удивительное, что содержится в реакции Белоусова — Жаботинского, еще далеко не исчерпано! Детальные эксперименты показывают, что в случае, когда время нахождения реагентов в реакционном объеме имеет промежуточное значение по отношению ко временам, характерным для двух осцил-

ляторных режимов, система демонстрирует сложное непериодическое поведение. Появление такой *химической турбулентности* снова иллюстрирует тенденцию многих естественных систем к хаотической эволюции при соблюдении определенных условий. Кроме того, оно выявляет следующее важное свойство химических систем: если в гидродинамике и в большинстве других физических примеров сложное поведение неизменно связано с пространственной неоднородностью, то в химии даже пространственно однородные системы могут иметь сложное поведение во времени. Причина этого уже упоминалась в начале этого раздела: химические системы наделены механизмами типа автокатализа, которые связаны с особыми молекулярными структурами и особой реакционной способностью определенных компонентов, что и позволяет таким системам переходить в новые состояния путем усиления (или ослабления) влияния слабых возмущений.

Б. Реакция БЖ в неоднородной системе: пространственные фигуры

Допустим теперь, что реакция БЖ протекает без перемешивания, что допускает развитие пространственных неоднородностей. Оказывается, что при этом можно наблюдать регулярные пространственно-временные картины в виде распространяющихся волновых фронтов. По необходимости плоское изображение этих картин (рис. 1.9) дает лишь слабое представление об их красоте и динамике развития. Показанные на рисунке волны создаются в тонком слое реагирующего вещества. Они бывают двух разных видов: кольцевые фронты (*а*), имеющие почти цилиндрическую симметрию относительно оси, перпендикулярной плоскости слоя, и напоминающие *мишень для пулевой стрельбы*, и спиральные фронты (*б*), вращающиеся в пространстве по или против часовой стрелки. Можно также получить, хотя и при довольно исключительных условиях, многозаходные спирали, показанные на рис. 1.9, *в*. В любом из этих случаев волновые фронты распространяются на макроскопические расстояния без искажений и с заданной скоростью, что эквивалентно некоторому химическому сообщению, посылаемому из центра излучения волн. Здесь мы снова являемся свидетелями рождения сложного.

Как и в случае задачи Бенара, образование волновых фронтов можно связывать с нарушением пространственной симметрии. Нарушение симметрии в случае картины типа мишени, показанной на рис. 1.9, *а*, весьма сходно с нарушением симметрии в задаче Бенара: по существу, и там и здесь система теряет трансляционную инвариантность вдоль определенного направления в пространстве. С другой стороны, нарушение симметрии,

связанное с образованием спиралей (рис. 1.9, б, в), носит совершенно иной характер, поскольку оно связано с понятием *хиральности*, или вращения. Этот тип асимметрии вещества всегда обладал особой прелестью. Например, Луи Пастер — основатель современной биохимии — не устал все снова и снова выражать свое восхищение оптической асимметрией биомолекул, которая проявляется в виде вращения плоскости поляризации света в выделенном направлении. Он считал это свойство одним из фундаментальных аспектов жизни. Более того, наблюдение над морфологической асимметрией взрослых организмов позволило человечеству ввести понятия правого и левого, что отмечалось философами и писателями еще со времен Платона. Поэтому удивительно наблюдать, как эти понятия довольно естественным образом возникают из внутренней динамики скромных, обыкновенных физико-химических систем.

Описанное выше поведение не является исключительно особенностью системы БЖ. Оно присуще также множеству других реакций в гомогенной фазе с участием столь же простых ве-

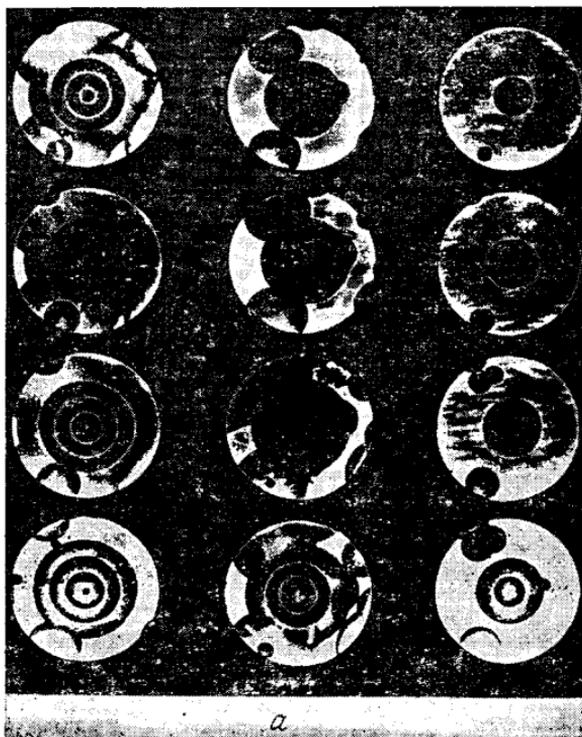


Рис. 1.9. Распространение химической волны в двумерной системе Белоусова — Жаботинского. а — ведущий центр (картинка напоминает мишень для стрельбы); б — спиральные волны; в — многозаходная спираль.

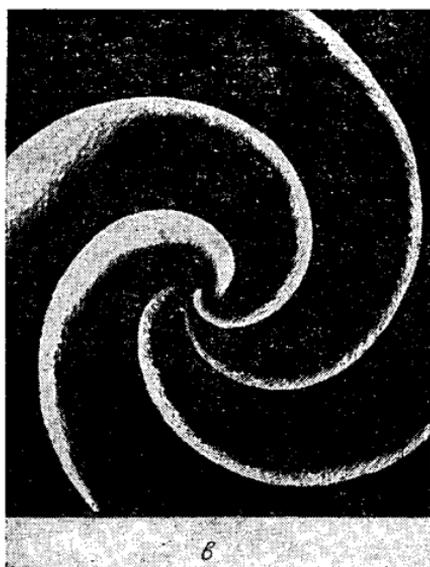
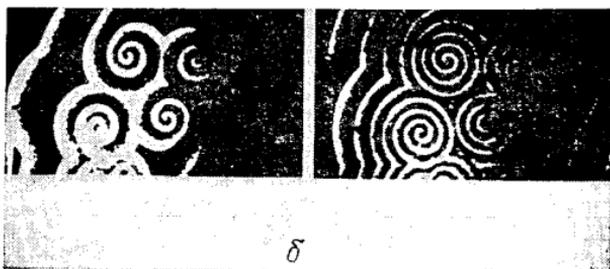


Таблица 1.1. Химические осцилляторы

| Основные компоненты | Дополнительные компоненты |
|---|--|
| <i>В реакторах с перемешиванием</i> | |
| I ⁻ | IO ₃ ⁻ , MnO ₄ ⁻ или Cr ₂ O ₇ |
| I ⁻ | Малоновая кислота |
| IO ₃ ⁻ | H ₃ AsO ₃ |
| IO ₃ ⁻ | Fe(CN) ₆ ⁴⁻ , SO ₃ ²⁻ , аскорбиновая кислота или CH ₂ O·SO ₂ |
| I ₂ | Fe(CN) ₆ ⁴⁻ , SO ₃ ²⁻ |
| IO ₃ ⁻ | I ⁻ , малоновая кислота |
| IO ₃ ⁻ | I ⁻ , H ₃ AsO ₃ |
| I ⁻ | BrO ₃ ⁻ |
| BrO ₃ ⁻ | SO ₃ ⁻ , Fe(CN) ₆ ⁴⁻ , H ₃ AsO ₃ или Sn ²⁺ |
| I ⁻ | I ₂ , S ₂ O ₃ ²⁻ |
| Основные компоненты | Дополнительные компоненты (катализаторы) |
| <i>В случае поверхностного катализа</i> | |
| CO + O ₂ | Pt, Pd, CuO, Ir |
| H ₂ + O ₂ | Pt, Pd, Ni |
| NH ₃ + O ₂ | Pt |
| C ₂ H ₄ + O ₂ | Pt |
| C ₃ H ₆ + O ₂ | Pt |
| C ₆ H ₁₂ + O ₂ | NaY (цеолит) |
| N ₂ O (разложение) | CuO |

шесть, частичный список которых приводится в табл. 1.1. Характерно, что во всех таких случаях область параметров, при которых наблюдаются осцилляции, находится близко к области, где наблюдается другое интересное явление — *бистабильность*. Точнее, при одних и тех же граничных условиях одновременно сосуществуют два (иногда несколько) устойчивых стационарных состояния. Выбор системой одного из этих состояний зависит от экспериментальных условий. Это положение иллюстрирует рис. 1.10, на котором показана зависимость некоторой переменной (например, концентрации какого-то соединения) от характерного управляющего параметра (например, времени пребывания). Допустим, что эксперимент проводится при некотором значении этого параметра $\lambda = \lambda'$, при котором существует лишь одно устойчивое состояние (ветвь *a*). Постепенно увеличивая λ , мы входим в область множественных состояний ($\lambda_1 < \lambda < \lambda_2$). Однако система остается на ветви *a* до тех пор, пока λ не превысит значения λ_2 , после чего происходит скачкообразный переход на ветвь *b*, на которой система и остается. Если теперь изменять значения λ в противоположном направлении, начиная, скажем,

от значения λ'' и доходя до λ' , то система будет находиться на ветви b вплоть до значения λ_1 параметра λ . Иными словами, поведение системы зависит от ее предыстории. Такое явление называется *гистерезисом*.

Во всех известных до сих пор случаях оказывается, что все три типа поведения — бистабильность, осцилляции или генерация волн — могут быть объяснены одним и тем же химическим механизмом. Как уже отмечалось, хотя бы одна стадия при этом должна быть автокаталитической. В системе БЖ на такой стадии производятся две молекулы промежуточного продукта

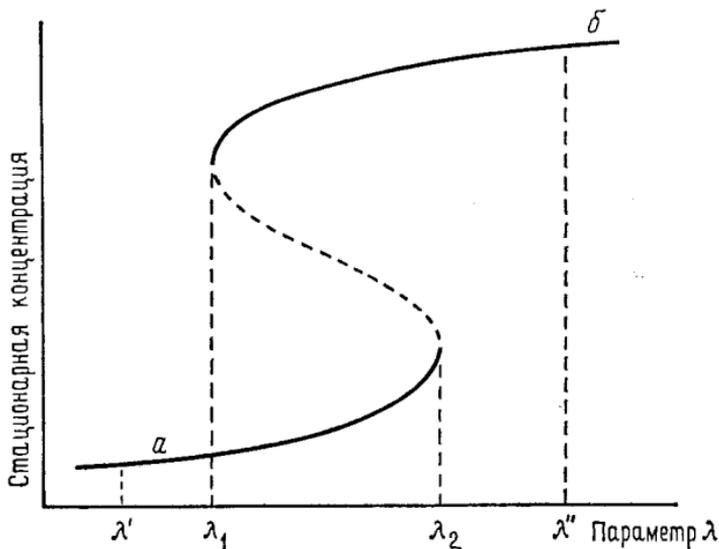
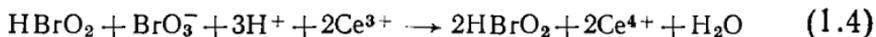


Рис. 1.10. Иллюстрация явлений бистабильности и гистерезиса.

(бромистой кислоты HBrO_2) из одной молекулы того же вещества:



Однако прямой автокатализ не является единственным механизмом, способным породить сложное поведение в химии. Например, в явлении гетерогенного катализа, уже упоминавшегося в начале этого раздела, некоторые химические стадии протекают с высвобождением энергии (по этой причине они называются экзотермическими реакциями). Среда при этом нагревается. А как уже отмечалось, константа скорости химической реакции зависит от температуры. С ростом температуры скорость реакции возрастает, что достаточно хорошо описывается законом Аррениуса:

$$k(T) = k_0 \exp\left(-\frac{E_0}{k_B T}\right) \quad (1.5)$$

где E_0 — энергия активации, а k_B — универсальная постоянная, называемая постоянной Больцмана. Качественно эту зависимость можно понять следующим образом (см. рис. 1.11). В процессе химического превращения происходит разрыв химических связей, или, иными словами, приходится преодолевать «энергетический барьер», соответствующий энергии данной связи. Этот барьер преодолевается за счет кинетической энергии поступательного движения сталкивающихся молекул. При нагревании среды средняя кинетическая энергия молекул возрастает, и тем самым возрастает число пар сталкивающихся молекул с кинетической энергией, достаточной для преодоления барьера. Другими словами, реакция ускоряется, что и описывается равенством (1.5).

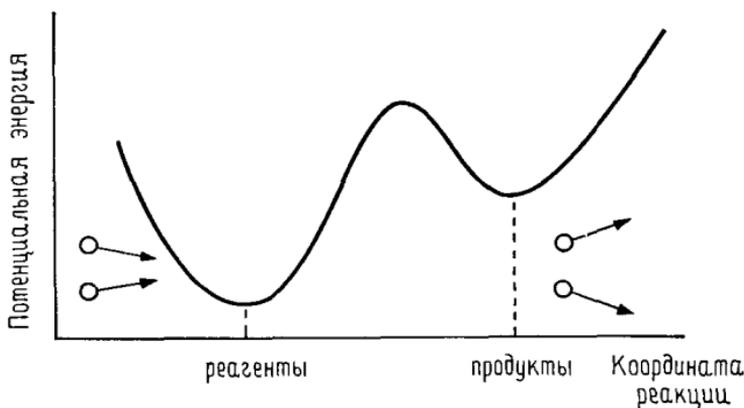


Рис. 1.11. Схематическое представление химической реакции в виде движения в потенциальном поле с двумя ямами. Оба минимума соответствуют устойчивым конфигурациям участвующих в процессе компонентов, рассматривающихся соответственно как «реагенты» и «продукты».

Допустим теперь, что рассматриваемая реакция является экзотермической. При ускорении такого процесса выделение тепла усиливается, так что температура среды будет возрастать и далее, и соответственно реакция будет продолжать ускоряться. Таким образом, вырисовывается потенциально дестабилизирующая конструкция, способная индуцировать переходы к новым типам поведения. И в самом деле, во многих каталитических реакциях (включая используемые в промышленности) наблюдаются типы поведения, описанные на примере системы БЖ. Такие процессы существенно влияют на локальное состояние катализатора и тем самым на течение всей реакции, а следовательно, на результативность и эффективность соответствующего химического производства.

1.5. Физико-химическая сложность и алгоритмическая сложность

До настоящего момента мы встречались с тремя основными типами самоорганизации вещества, приводящими к сложному поведению: бистабильностью и гистерезисом, колебаниями (как периодическими, так и непериодическими) и пространственными фигурами. Возникновение таких явлений мы видели на примерах двух систем — объемных течений при некоторых температурных условиях и открытых систем с каталитическими химическими реакциями. Можно ли с помощью этих примеров дать четкое определение сложного? Это оказало бы существенную помощь в дальнейшем анализе механизмов, ответственных за возникновение сложного поведения.

Один из способов достичь определения, ухватывающего сущность сложного, состоит в абстрагировании от деталей экспериментов, рассмотренных в разд. 1.3 и 1.4, и в попытке передачи другому наблюдателю такого набора инструкций, который позволил бы ему воспроизвести нужное поведение, не повторяя самих экспериментов. Допустим теперь, что любое сообщение такому наблюдателю стоит очень дорого. Это накладывает дополнительные ограничения — такие инструкции должны быть по мере возможности краткими, лишь бы не терялась заложенная в них информация.

Как бы мы поступили при таких обстоятельствах в случае эксперимента Бенара? Для описания равновесной ситуации необходимы и достаточны следующие инструкции (см. рис. 1.2)*:

- 1) $\Delta T = T_2 - T_1 = 0$;
- 2) $v = 0$, т.е. скорость конвекции v обращается в нуль в любой точке (горизонтальной) плоскости, параллельной металлическим пластинам.

Попытаемся теперь передать информацию о картинах, возникающих выше бенаровской неустойчивости. Из рис. 1.3 и 1.5 ясно, что соответствующее сообщение должно быть следующего вида:

- 1) $\Delta T = T_2 - T_1 > \Delta T_c$;
- 2) $v = v_0$ в конкретной точке P , выбранной в качестве точки отсчета;
- 3) эта конкретная точка принадлежит ячейке с конкретным направлением вращения (L или R);
- 4) в любой другой точке значение v можно получить из v_0 , исходя из периодичности последовательности LRLR и определенной длины волны рассматриваемой картины.

* Для краткости мы сконцентрируем свое внимание на данных, связанных со скоростью v , оставляя в стороне температуру и давление.

Мы видим, что при переходе через порог Бенара для определения состояния, возникающего в системе, количество инструкций увеличивается. Отсюда непосредственно вытекает определение *алгоритмической сложности*, предложенной в 1965 г. независимо советским математиком А. Колмогоровым и американским математиком Г. Чейтином: алгоритмическая сложность некоторой последовательности данных определяется как *минимальная длина вычислительного алгоритма* (измеряемая, например, числом бит, если алгоритм предназначен для передачи на компьютер), который мог бы воспроизвести заданную последовательность.

Равносильна ли определенная таким образом алгоритмическая сложность сложности, наблюдаемой в физических науках и биологии? Рассмотрим последовательность из N данных, выраженных в двоичной форме, причем полностью регулярных (например, 100 100 100 ...). Очевидно, содержащееся в этой последовательности сообщение можно существенным образом сжать. Например, его можно передать на компьютер с помощью очень простого алгоритма: «напечатать число 100 десять, сто, миллион... раз». Количество двоичных цифр в таком алгоритме составляет малую долю от количества двоичных цифр в исходном ряду, и по мере удлинения ряда длина алгоритма возрастает с гораздо меньшей скоростью. Следовательно, в соответствии с определением это означает ограниченную алгоритмическую сложность. Рассмотрим теперь набор всех случайных последовательностей длиной N двоичных цифр. Число таких последовательностей равно 2^N . Допустим теперь, что по мере возрастания N некоторые из этих последовательностей могут быть сжаты подобно тому, как это было сделано в предыдущем примере, т. е. выражены в более компактной форме последовательностей длиной в 1, 2, ... и до K цифр, так что отношение K/N стремится к нулю при неограниченном возрастании N . Полное число таких регулярных последовательностей равно

$$2 + 2^2 + \dots + 2^K = 2^{K+1} - 2$$

При больших N это число становится чрезвычайно малым по сравнению с 2^N . Иными словами, большинство случайных последовательностей будет несжимаемым. В соответствии с определением это означает максимальную алгоритмическую сложность, по существу равную длине последовательности N . Такой тип сложности реализуется в типичном эксперименте по подбрасыванию монеты, разумеется, при условии соблюдения должных мер предосторожностей по устранению возможных корреляций между последовательными подбрасываниями, т. е. при условии «честной» игры.

Все это пока еще далеко от интуитивного представления, вытекающего из экспериментальных результатов разд. 1.3 и 1.4, где была показана связь сложного с крупномасштабной когерентностью и согласованным поведением. В самом деле, полная упорядоченность в виде полного отсутствия изменчивости представляет собой предельный случай когерентности, когда поведение объекта настолько примитивно, что едва ли можно говорить о сложном. С другой стороны, сильная изменчивость, представляемая случайным шумом, и сопутствующее такой изменчивости отсутствие корреляций являются другой, столь же непредставительной формой организации. Сложность реальных физико-химических систем должна находиться где-то между этими двумя пределами и, следовательно, не должна полностью отождествляться с алгоритмической сложностью.

Как мы видели в разд. 1.3 и 1.4, поведение физико-химических систем при определенных условиях может принять случайные формы в виде хаотической динамики. В разд. 3.10, 4.7 и 4.8 мы покажем, что имеются существенные различия между этой формой случайности и случайностью, генерируемой шумом. Тем не менее хаотическая динамика наводит на мысль о существовании глубоких связей между физической и алгоритмической сложностью. На данном этапе мы еще не готовы к детальному рассмотрению этого вопроса, которое будет отложено до гл. 4.

1.6. Некоторые дополнительные примеры сложного поведения в обычном масштабе

Как выясняется, три типа организации, обсуждавшиеся в разд. 1.3—1.5, встречаются во множестве других макроскопических задач, возникающих в самых разнообразных контекстах. Не пытаясь нарисовать исчерпывающую картину, мы в дальнейшем приведем несколько характерных примеров, в каждом случае отмечая основные причины, вызывающие соответствующие переходы.

А. Явления, вызванные поверхностным натяжением. Наука о материалах

При самых разных обстоятельствах можно столкнуться с обменом энергией и веществом в неравновесных условиях, осуществляемым между одинаковыми фазами вещества (твердыми или жидкими), между разными фазами или между биологическими мембранами и жидкостями или газами. Все такие процессы характеризуются наличием *межфазных границ*, т. е. поверхностей раздела между сосуществующими фазами. Подобно тому

как в выражении для работы, совершаемой при изменении объема тела, давление является коэффициентом пропорциональности при изменении объема, точно так же в выражении для работы, совершаемой при изменении площади элемента поверхности, коэффициентом пропорциональности при изменении площади такого элемента служит *поверхностное натяжение*. Как правило, поверхностное натяжение зависит от кривизны, температуры и состава граничащих фаз. Таким образом, оно осуществляет сопряжение между этими свойствами и механическими свойствами данного вещества или материала.

Одно из проявлений такого сопряжения — спонтанная деформация поверхности, приводящая к возникновению на ней регулярных пространственных картин, либо к возникновению потоков в соседних жидкостях, покоившихся в начальном состоянии, либо к формированию капель. Эти явления важны в таких областях, как фармакология, пищевая промышленность или добыча нефти.

Другим поразительным проявлением такого сопряжения служит рост твердого тела (чистого или сплава) из расплава, когда продвижение границы раздела фаз в направлении расплава обеспечивается внешним охлаждением. Помимо того, что этот процесс часто встречается в промышленности, он также имеет место при образовании снежинок из капель воды. Во всех этих случаях наблюдается большое разнообразие пространственных картин типа ламеллярных или дендритоподобных структур, характеризующихся масштабами, на несколько порядков превышающими типичный кристаллографический размер.

Недавно идеи неравновесной физики нашли многообещающие приложения в области геологии, что также связано с двумя обсуждаемыми темами. Дело в том, что во многочисленных геологических отложениях для целого ряда пространственных масштабов наблюдается занятная регулярность структур, возникших путем минерализации (см. рис. 1.12): метаморфные слои (мм—м), граниты (см), агаты (мм—см) и т. д. Согласно традиционным взглядам, эти структуры объясняются «последовательными» явлениями, обусловленными сменами времен года или климата. Однако выясняется, что более удовлетворительной является интерпретация, основанная на представлениях о нарушении симметрии за счет переходов, вызванных неравновесностью системы. Если такая точка зрения в дальнейшем подтвердится, то это очень сильно повлияет на интерпретацию происхождения многочисленных геологических отложений.

В дополнение к этим примерам мы рассмотрим также проблемы образования и распространения дефектов и изломов, которые так или иначе связаны с переходами, опосредованными наличием межфазных границ.

**Б. Кооперативные явления,
обусловленные электромагнитными полями.
Электрические цепи, лазеры, оптическая бистабильность**

В электромашиностроении и вычислительной технике широко используются кооперативные явления, возникающие в системах с активными элементами типа диодов Ганна, с резонансными цепями, построенными на триодах, или с полупроводниками. В конце концов работа всех современных компьютеров основана на использовании цепей бистабильных элементов, способных переключаться между двумя различными состояниями в зависимости от значений определенных параметров.

Перемены, произошедшие с 60-х годов в оптике, никак нельзя назвать иначе, кроме как настоящей революцией, причем это в оптике — науке, считавшейся до тех пор одной из самых класси-



Рис. 1.12. Вверху: образец скарна из Сан-Леоне, Сардиния. Светлые полосы толщиной 1—2 мм состоят из граната — андрадита (силикат кальция и железа). Темные полосы толщиной 5—8 мм состоят из магнетита и кварца. Белый прямоугольник имеет в длину 1 см. (Любезно предоставлено Б. Гаем.) Внизу: концентрически-зональный диорит из Эпо, Финляндия. Концентрические слои попеременно обогащены биотитом (темные полосы) и плагиоклазом (светлые полосы). Радиус образца 10 см. (Любезно предоставлено Э. Мерино.)

ческих и устоявшихся областей физики. Генерация когерентных пучков света с помощью лазера стимулировала впечатляющее развитие как теории, так и эксперимента, в результате чего было показано, что это явление можно рассматривать как своего рода процесс самоорганизации, возникающий при неравновесных условиях. Более того, оптическая бистабильность — возможность двух устойчивых режимов работы резонансной полости — есть не что иное, как заманчивый вызов современной технологии: разработать компьютеры на основе бистабильных электрооптических цепей, которые были бы более выгодны, нежели электронные, вследствие более коротких времен переключения и меньшего потребления энергии. Рассмотрим вкратце механизм, ответственный за кооперативность и переходы в частном случае оптической бистабильности.

Допустим, что в некоторую полость (рис. 1.13, а) поступает когерентное электрическое поле E_1 с четко определенной частотой. С помощью современных лазеров подобные поля создаются довольно легко. Устройство полости таково, что она почти резонансна по отношению к падающему пучку света. Очевидно, если полость пуста, то интенсивность прошедшего пучка I_t пропорциональна интенсивности падающего пучка I_i (рис. 1.13, б). Однако, если заполнить полость средой, способной поглощать свет и находиться почти в резонансе с падающим пучком, то возникают совершенно неожиданные эффекты. Для начала заметим, что взаимодействия вещества со светом существенно нелинейно. Например, скорость изменения электрической поляризации данного образца зависит как от интенсивности света, так и от числа атомов, уже находящихся в возбужденном состоянии. Не менее важно то, что за счет определенного расположения зеркал, показанных на рис. 1.13, а, часть света, которая иначе ушла бы во внешнее пространство, возвращается в среду. Это эквивалентно положительной обратной связи, весьма сходной с химическим автокатализом. Наличие этой обратной связи при определенных условиях приводит к тому, что на некоторой части передаточной кривой (рис. 1.13, в) возникает «дифференциальное усиление» в том смысле, что наклон зависимости I_t от I_i становится больше единицы. При этом система приобретает свойства *оптического транзистора*, поскольку слабая модуляция падающего пучка света усиливается на выходе. При других условиях стационарная передаточная кривая приобретает S-образный вид (рис. 1.13, г). При этом существует область значений интенсивности падающего света, в которой поведение системы характеризуется наличием бистабильности. Два имеющихся уровня передачи можно привести в соответствие логическим состояниям «0» и «1», что подтверждает возможность построения на такой основе элемента *оптической памяти*.

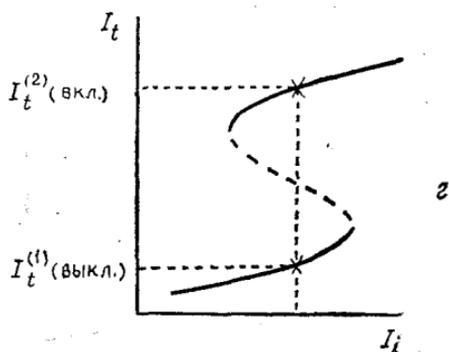
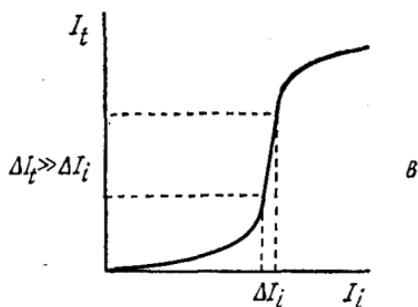
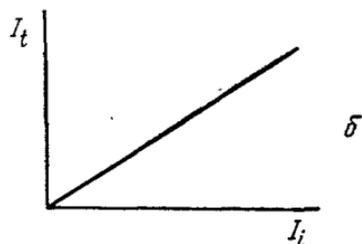
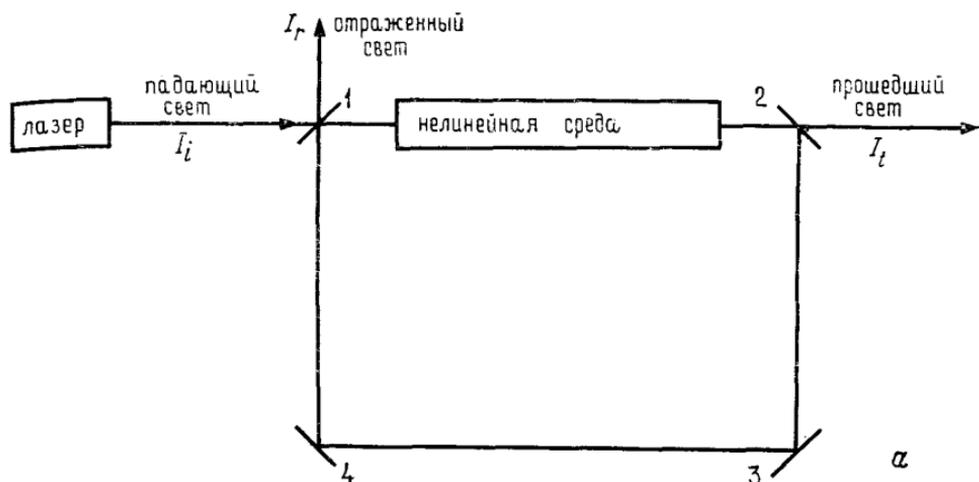


Рис. 1.13. *а* — схематическое представление экспериментального устройства для наблюдения оптической бистабильности; *б* — отношение интенсивностей прошедшего и падающего света в пустой полости; *в*, *г* — полость, заполненная оптически активной средой, может служить в качестве оптического транзистора или оптического переключателя.

1.7. Снова биологические системы

Прежде чем продолжить обзор различных типов сложного поведения, нам нужно провести несколько более систематическое обсуждение биологических систем, выходя за пределы нескольких интуитивных замечаний, которые были сделаны ранее в этой главе. Из всех природных объектов живые существа, несомненно, и функционально, и морфологически являются наиболее сложными и высокоорганизованными. Как уже отмечалось, они служат в качестве прототипов, дающих импульс физическим наукам для лучшего понимания сложного. Живые существа являются в буквальном смысле историческими структурами, поскольку они способны сохранять память о формах и функциях, приобретенных в прошлом за длительный период биологической эволюции. Более того, живые системы совершенно определенно функционируют вдали от равновесия. Организм как целое непрерывно получает потоки энергии (например, в виде солнечной энергии, используемой растениями при фотосинтезе) и вещества (в виде питательных веществ) и преобразует их в самые разные отходы, выделяемые во внешнюю среду. На клеточном уровне также наблюдаются сильные неоднородности. Например, концентрация иона калия внутри клеток нервной системы — нейронов — много выше, чем во внеклеточной среде, в то время как для ионов натрия имеет место обратная ситуация. Эти различия, подразумевающие сильную неравновесность состояния, лежат в основе процессов типа распространения нервного возбуждения, играющего важную роль в жизни. Такая неравновесность поддерживается за счет работы активного транспорта и биоэнергетических реакций типа гликолиза и дыхания.

Итак, мы уже убедились, что обычные физико-химические системы могут обладать сложным поведением со многими особенностями, обычно приписываемыми живым системам. В связи с этим вполне уместно задаться следующим вопросом: можно ли некоторые из таких особенностей биологических систем объяснить за счет переходов, индуцированных неравновесностью и соответствующими механизмами дестабилизации типа химического автокатализа? По-видимому, это один из наиболее фундаментальных научных вопросов. Хотя исчерпывающего ответа в настоящее время нет, тем не менее можно отметить несколько примеров, в которых связь между физико-химической организацией и биологической упорядоченностью просматривается особенно четко. Конкретная проблема, на которой мы здесь сосредоточимся, — это управление эмбриональным развитием.

Эмбриональное развитие представляет собой последовательность событий, ведущих от отдельной клетки — оплодотворенного яйца — к целому организму. Природа поставляет нам огром-

ное множество примеров такого рода процессов. Один из простейших случаев относится к бактерии, развитие которой представляет собой ряд клеточных делений. На другом конце спектра возможных примеров находятся развитые организмы типа млекопитающих. В таких случаях в результате эмбрионального развития возникает многоклеточное тело; в нем клетки образуют специализированные ткани и органы, которые могут содержать 10^{12} и более клеток.

В настоящее время вопрос отнюдь не ставится таким образом, чтобы детально понять, как протекают такие процессы и, в частности, как им удается координироваться столь фантастическим образом, что каждая клетка получает возможность сыграть свою роль в нужное время и на нужном месте. Вместо этого мы обсудим живые системы, развитие которых характеризуется промежуточным уровнем сложности, как, например, амебы вида *Dictyostelium discoideum*. Здесь развитие по существу сводится к переходному явлению, весьма сходному с тем, что наблюдается в реакции БЖ, и отмечающему переход от одноклеточной к многоклеточной стадии жизни.

На рис. 1.14 показан жизненный цикл этого вида. На рис. 1.14, а амебы находятся в одноклеточной стадии. Они движутся во внешней среде, питаются бактериями и размножаются путем клеточного деления. В целом среда с амебами представляет собой однородную систему, поскольку их плотность (число клеток на квадратный сантиметр) весьма постоянна. Допустим теперь, что амебы начинают голодать (в лаборатории это легко сделать искусственно, а в природе такое может случиться вследствие неблагоприятных изменений среды обитания). Это условие является аналогом некоторого ограничения в физическом или химическом эксперименте. При этом можно видеть, что отдельные клетки не умирают. Вместо этого в ответ на данное ограничение они начинают агрегироваться по направлению к некоторому центру притяжения (рис. 1.14, б). Изначальная однородность распределения клеток нарушается — пространство структурируется. Возникающее при этом многоклеточное тело способно двигаться главным образом с целью отыскания более благоприятных условий по температуре и влажности (рис. 1.14, в). После такой миграции в результате дифференциации из многоклеточного тела возникают два типа клеток, причем клетки одного типа образуют своеобразную ножку, а клетки другого типа — плодовое тело, внутри которого образуются споры (рис. 1.14, г). В конечном счете споры рассеиваются в среде обитания (рис. 1.14, д), и при благоприятных условиях они прорастают, с тем чтобы превратиться в амеб. После этого начинается новый жизненный цикл.

Рассмотрим несколько подробнее стадию агрегации (рис. 1.15).

В течение этого процесса наблюдаются следующие явления. Прежде всего после установления условий нехватки питательных веществ некоторые клетки начинают синтезировать и выделять во внеклеточную среду химическое вещество, известное под названием циклического аденозинмонофосфата (цАМФ) и играющее роль своего рода сигнала. Его синтез и выделение происходят периодически, подобно химическим часам в системе БЖ,

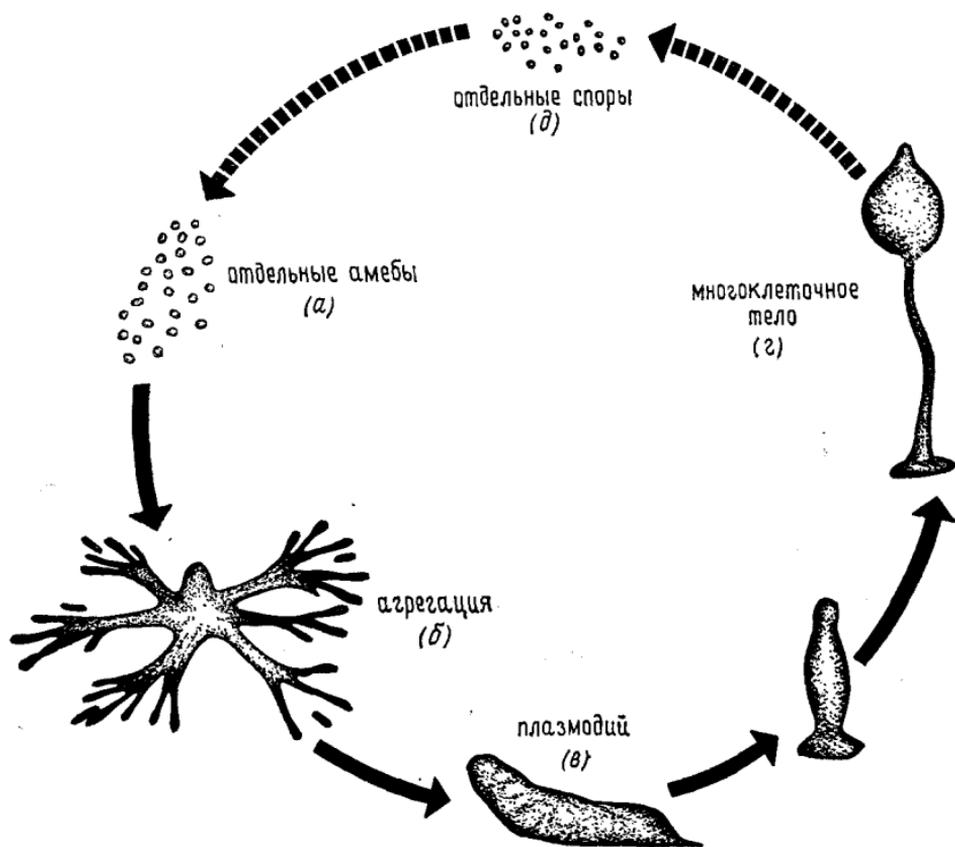


Рис. 1.14. Жизненный цикл амебы *Dictyostelium discoideum*.

с четким периодом при данных экспериментальных условиях. Диффундируя во внеклеточной среде, выделенный клетками-«пионерами» цАМФ достигает поверхности соседних клеток. После этого происходят события двух типов. Во-первых, эти клетки направленно движутся в область повышенной концентрации цАМФ (так называемый хемотаксис). В результате возникает распределение плотности клеток, сильно напоминающее волновые картины в реакции БЖ (рис. 1.9). Во-вторых, процесс

агрегации ускоряется за счет способности клеток усиливать сигнал и передавать его дальше по среде. Это позволяет организму контролировать большую территорию и формировать многоклеточное тело, содержащее около 10^5 клеток.

Суммируя сказанное, можно отметить, что ответ системы на нехватку пищи приводит к новому уровню организации, характеризующему согласованным поведением большого числа клеток и позволяющему организму гибко реагировать на неблагоприятные изменения среды обитания. Какими механизмами опосредуется этот переход? Прежде всего заметим, что хемотактический процесс усиливает исходно созданную неоднородность, возник-



Рис. 1.15. Концентрические и спиральные волны агрегирующих популяций клеток *Dictyostelium discoideum* на поверхности агара. Движущиеся к центру амёбы выглядят как блестящие полосы, а неподвижные — как темные полосы.

шую в результате импульсного выделения цАМФ клетками-«пионерами». В самом деле, этот процесс увеличивает плотность клеток вблизи центра выделения цАМФ и благодаря эффекту реле увеличивает выделение цАМФ в этой области, что в свою очередь усиливает хемотактическое движение других клеток к этому центру. В результате возникает то, что обычно называется *петлей обратной связи* и весьма сходно с химическим автокатализом или самоускорением экзотермической реакции, с которыми мы встречались в разд. 1.4.

Оказывается, в жизненном цикле *Dictyostelium discoideum* имеется и второй механизм обратной связи. Этот механизм работает на субклеточном уровне и отвечает как за периодическое выделение цАМФ, так и за передачу хемотактического сигнала. Этот механизм связан с синтезом цАМФ клеткой. Это вещество возникает в результате превращения другого важного клеточного компонента — аденозинтрифосфата (АТФ), являющегося одним из основных носителей энергии в живой клетке благодаря высокой энергии фосфатной связи. Однако превращение АТФ→цАМФ не происходит самопроизвольно — для его осуществления необходим соответствующий катализатор, ускоряющий это превращение до уровня, совместимого с жизненными потребностями. В биологических системах задача катализа выполняется специальными молекулами — *ферментами*. Эти молекулы обычно содержат довольно много атомов и имеют высокую молекулярную массу, за что они и получили название *макромокул*. Некоторые ферменты имеют единственный активный центр, с которым реагенты должны столкнуться для того, чтобы произошло их превращение в продукты. Однако во многих случаях приходится иметь дело с *кооперативными ферментами*, обладающими несколькими центрами. Одни из этих центров являются каталитическими, другие — регуляторными. Когда к последним присоединяются специальные молекулы — *эффекторы*, каталитическая функция фермента претерпевает серьезные изменения. В некоторых случаях такими эффекторами могут служить сами молекулы, реагирующие на каталитическом центре или создаваемые на нем. При этом снова замыкается петля обратной связи, которая будет положительной (активация), если в результате катализ ускоряется, или отрицательной (ингибирование) в противном случае.

Вернемся к превращению АТФ→цАМФ. Катализирующий этот процесс фермент называется аденилатциклазой. Он прикреплен к мембране клетки. Этот фермент является кооперативным, однако подробности его работы пока еще не выяснены. Естественно допустить, что внутри клетки имеется некоторый активный центр, катализирующий реакцию. Получающийся в результате цАМФ диффундирует во внеклеточную среду через

мембрану клетки и с некоторой вероятностью может связываться с внешней (регуляторной) частью аденилатциклазы, тем самым активируя ее (рис. 1.16). В результате производство цАМФ ускоряется, что и приводит к замыканию петли обратной связи, способной усиливать сигналы и индуцировать осцилляторное поведение.

Будучи настолько простыми, насколько это возможно, лежащие в основе развития *Dictyostelium discoideum* процессы являются прототипическими для широкого класса более изощренных феноменов развития. В самом деле, нарушение пространственной симметрии и сопровождающее ее неоднородное распределение клеток, по-видимому, объясняют многие экспериментальные данные по морфогенезу насекомых. В общем случае в развивающихся тканях часто наблюдаются градиенты многих веществ типа ионов или метаболитов. Естественно предполо-

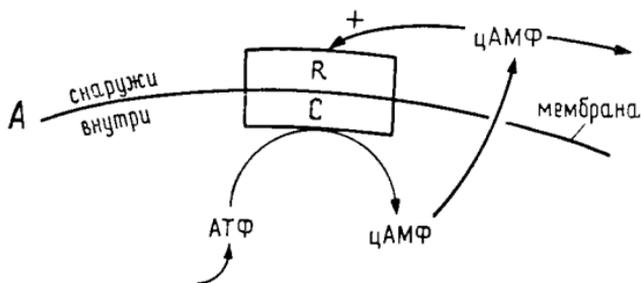


Рис. 1.16. Механизм колебательного синтеза цАМФ у амёбы *Dictyostelium discoideum*.

жить, что эти градиенты служат для ткани своеобразной «координатной системой», поставляющей *позиционную информацию* отдельным клеткам. Благодаря этой информации клетки могут различать свое положение по отношению к своим партнерам. Поэтому, вероятнее всего, переходы, опосредованные химическими веществами и приводящие к нарушению симметрии, являются одним из ключевых свойств жизни, ответственных также за явления физико-химической самоорганизации. Эта поразительная идея была впервые провозглашена в 1952 г. английским математиком Аланом Тьюрингом и с тех пор служила постоянным источником вдохновения для физиков и биологов в одинаковой мере.

1.8. Сложность в планетарном и космическом масштабах

Мы все еще далеки от того, чтобы дать определение сложного, и маловероятно, что эта цель будет достигнута до конца этой книги. Сложность является одним из понятий, которые вхо-

дят составной частью в порождаемые ими проблемы. Благодаря многочисленным примерам, рассмотренным в предыдущих разделах, мы знаем, что одна из существенных особенностей сложного поведения — это способность осуществлять *переходы* между различными режимами. Иначе говоря, понятие сложности относится к таким системам, в которых наблюдаемое поведение в значительной мере связано с их эволюцией, т. е. предысторией.

Насколько общим является этот класс систем? В течение длительного времени в науке существовал статический взгляд, согласно которому мир представляет собой «хорошо смазанную машину», работающую с точно такой же непогрешимой точностью, что и в незапамятные времена, и которая будет так же работать в любом обозримом будущем. Сегодня такой взгляд вряд ли сможет выдержать критику. Биологическая революция, произведенная Дарвином в XIX в., и революция, происходящая в геофизике и космологии с начала 50-х годов, показывают — куда ни посмотри на окружающую нас среду в широком смысле этого слова, мы всюду обнаруживаем следы прошлого. Эти следы отнюдь не являются чем-то вроде окаменелостей; напротив, они активно готовят еще предстоящую историю. Земная атмосфера, насколько нам сегодня известно, является результатом развития жизни на планете. Живые существа — это результат длительной эволюции, химическая или предбиологическая стадия которой протекала в условиях, имевших место на Земле около 4 млрд. лет назад. Гренландские или атлантические ледники, положение континентов, дно океана, наш климат — все это претерпело длительную эволюцию, характеризующую рядом крупномасштабных переходов. И наконец, последнее по счету, но не по важности: микроволновое фоновое излучение (черного тела при 2,7 К), обнаруженное астрономами, является серьезным указанием на то, что Вселенная образовалась около 20 млрд. лет назад в результате первичного взрыва.

Эти примеры дают представление о том, как далеко могут простираться переходные явления, аналогичные тем, которые происходят в скромной кастрюле с водой или в химическом растворе. Разумеется, мы не можем исчерпывающим образом рассмотреть все эти примеры, поэтому для определенности остановимся на проблеме климатических изменений, к которой мы еще вернемся в одной из последующих глав. Дополнительные замечания относительно процессов космического масштаба приводятся в этом разделе.

Примерно с 60-х годов все возрастающее количество наблюдений приводит к выводу о том, что земной климат способен проявлять весьма выраженную внутреннюю изменчивость. Этот факт в сопоставлении с аномально благоприятным и стабильным климатическим периодом начала нашего столетия одновре-

менно и удивляет и озабочивает специалистов, политиков и общественность. Впервые человек осознает глобальный, планетарный характер климатической системы, а также тот факт, что его собственная деятельность может также повлиять на работу впечатляющей климатической машины.

Рассмотрим ряд фактов, относящихся к этому новому положению, и начать можно с проблемы изменений на *больших временах*. Большинство имеющихся наблюдений сделаны с помощью мощных методов анализа изотопного состава реликтовых органических остатков, давших непосредственную информацию о температуре в древние времена.

Прежде всего из таких данных стало ясно, что климатические условия, имевшие место в течение последних 200—300 млн. лет, чрезвычайно разнятся с современными условиями. Все это время, за исключением четвертичного периода (начавшегося около 2 млн. лет назад) на континентах практически не было льда, и уровень моря был выше современного примерно на 80 м. Климат был чрезвычайно мягким, с относительно небольшой разницей температур на экваторе (25—30 °С) и полюсах (8—10 °С).

Более заметный контраст между экваториальной и полярной температурами начал развиваться в третичном периоде, около 40 млн. лет назад. Так, за относительно короткое время в 100 000 лет температура морской воды южнее Новой Зеландии упала на несколько градусов. Вероятно, именно в это время возникло Антарктическое течение, ослабляющее теплообмен между высокими и низкими широтами, что способствует дальнейшему охлаждению масс воды, «захваченных» таким образом вблизи полярных областей. Здесь мы опять наблюдаем механизм обратной связи в действии.

В начале четвертичного периода это различие в температурах было достаточным для образования и сохранения континентального льда. В Северном полушарии произошел ряд последовательных *оледенений*, продвигавших иногда ледники до средних широт. Эти климатические события происходили со средней периодичностью около 100 000 лет. Впрочем, они носят также значительный отпечаток случайности (см. рис. 1.17).

Последнее продвижение континентального льда в Северном полушарии достигло максимума около 18 000 лет назад, и его следы видны до сих пор. В то время как сейчас объем континентального льда (сосредоточенного в основном в Антарктике и Гренландии) составляет около 30 млн. кубических километров, тогда его было от 70 до 80 млн. кубических километров, причем льдом была покрыта значительная часть севера Америки и Европы. Поскольку ледники захватили огромное количество воды, уровень моря был на 120 м ниже сегодняшнего. С тех пор зна-

чительная часть льда растаяла, что и определило контуры бережий и прочие черты современного ландшафта.

Обсудим теперь климатические изменения, происходившие на *более коротком временном масштабе*. Последние следы ледяных шапок континентов (за исключением Гренландии и Антарктики) исчезли около 7000 лет назад, и последовавший за этим событием период принято называть *климатическим оптимумом*. Самое удивительное в этой эпохе — появление влажных периодов в области между Мавританией (17° западной долготы) и Раджастханом (78° восточной долготы), являющейся одной из самых засушливых зон планеты. Даже в центре Сахары, где в

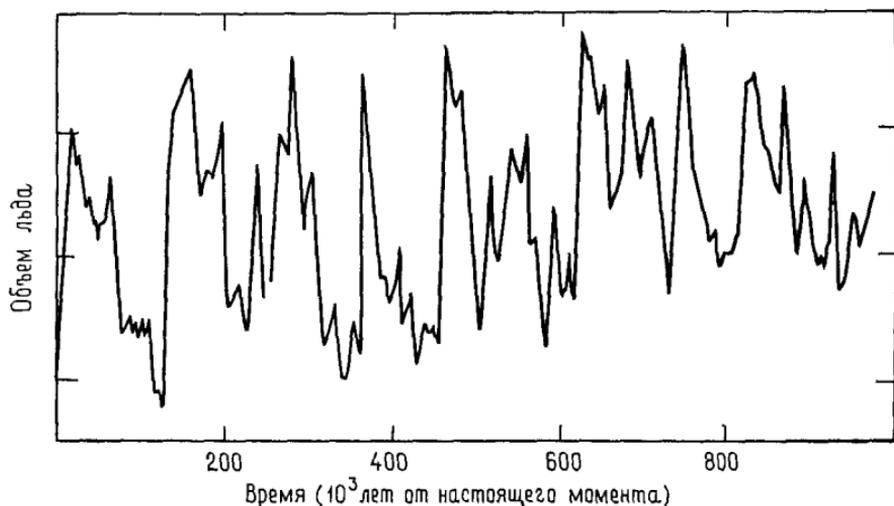


Рис. 1.17. Изменение объема льда на Земле за последние 800 000 лет. Данные получены по изотопным исследованиям глубоководных осадочных пород.

наше время выпадает едва ли 5 мм осадков в год, тогда имелись непересыхающие реки и развитое земледелие, что указывает на наличие ежегодных осадков по меньшей мере 300—400 мм. По всей видимости, эта благоприятная ситуация длилась до начала железного века. Образ жизни людей восточного бассейна Средиземного моря, открывающийся нам в поэмах Гомера или в творениях Шлимана, вызывает представления о нескончаемой весне, благоухающей природе и плодородной земле.

Примерно около 3000 лет назад началось глобальное похолодание. Однако эта общая тенденция прерывалась многочисленными флуктуациями длительностью порядка сотни лет и амплитудой около 1°C . Хорошо известным примером случайного потепления является период 800—1200 гг. н. э., когда викинги открыли Гренландию (дословно Зеленую землю) и, возможно, Северную Америку. Противоположная тенденция наблюдалась

между 1550 и 1700 гг. н.э. Этот «малый ледниковый период» обернулся Европе и Северной Америке нехваткой продуктов питания и множеством других катастроф.

Теперь понятно, почему климат первой половины нашего столетия представляется скорее аномалией, нежели типичным образцом длительной и бурной истории климата. Человечество наблюдало относительно предсказуемую погоду и отступление льдов в Северном полушарии. Сельское хозяйство и производство продуктов питания получили значительные дополнительные возможности, и это привело к увеличению земного населения. С другой стороны, такое кажущееся постоянство благоприятных условий позволило установиться ложному представлению о том, что «нормально» и что «аномально» в климатологии.

С середины 70-х годов наблюдается возвращение переменчивости климата. Примером может служить аномально суровая зима, поразившая восточную часть Северной Америки в 1976—1977 гг., и одновременно длительная засуха в западной части этого континента. По-видимому, это явление было связано с ослаблением или даже с локальным блокированием атмосферной циркуляции, что отразилось, например, в значительном смещении части полярного течения к югу (на востоке американского континента) и части — к северу (на западе этого континента).

Убедившись, что длинная история климатической системы насыщена сложными переходными явлениями, рассмотрим теперь некоторые факторы, лежащие в основе этого сложного поведения. Как и в предыдущих разделах настоящей главы, наше обсуждение останется качественным. Более количественный подход будет изложен в гл. 6*.

Основным фактором, управляющим земным климатом, является, несомненно, приток солнечной энергии. Величина энергии, поступающей в верхние слои атмосферы в единицу времени на единицу площади, перпендикулярной солнечным лучам, называется солнечной постоянной Q . Астрофизические расчеты показывают, что яркость Солнца (как и яркость прочих звезд так называемой главной последовательности) со временем возрастает. Отсюда следует, что несколько сот миллионов лет назад величина Q имела значение, на несколько процентов меньше современного. И все же, как мы видели, климат в то время был очень мягким и ледников не было. Этот парадокс «тусклого раннего Солнца» показывает, что климатическая система не следует пассивно за изменениями солнечной энергии. Иными словами, чрезвычайно нетривиальная модуляция внешних эффектов осуще-

* Для читателя, интересующегося вопросами климатологии, несомненный интерес представит книга Р. Гилмора «Прикладная теория катастроф» (М.: Мир, 1984, т. 2, с. 67—87). — *Прим. перев.*

ствляется внутренними процессами климатической системы, и в первую очередь такими, как состав атмосферы, облака и интенсивность атмосферной циркуляции.

Сделанные выше замечания справедливы также применительно к ответу климатической системы на другое внешнее воздействие, обусловленное изменением характеристик земной орбиты. Астрономические расчеты показывают, что благодаря возмущающему действию прочих тел солнечной системы имеют место три основных вида отклонения земной орбиты от расчетов, проводимых в рамках элементарной механики. Во-первых, ось земного вращения участвует в прецессионном движении подобно верхушке веретена, так что точка на небе, в которую направлена ось, описывает окружность. Во-вторых, угол между осью земного вращения и направлением, перпендикулярным плоскости эклиптики (составляющий сейчас около $23,5^\circ$), изменяется примерно на $\pm 1,5^\circ$. И в-третьих, эксцентриситет земной орбиты также изменяется. Эти изменения происходят с периодами порядка 22 000, 41 000 и 100 000 лет соответственно. Теория предсказывает, что первые два эффекта ответственны за распределение энергии Солнца по временам года, третий же влияет на количество энергии, получаемое Землей ежегодно. Во всех трех случаях непосредственные эффекты малы. И все же они оставили свой отпечаток в основной периодичности ледниковых циклов, упоминавшейся ранее в настоящем разделе. Опять роль внутренней динамики должна быть исключительно нетривиальной — за ее счет должны усиливаться слабые внешние воздействия, вызывая тем самым общие климатические сдвиги.

Приведем пример такого механизма усиления. Рассмотрим взаимодействие между радиационными процессами и ледовыми покрытиями. Хорошо известно, что отражательная способность льда — *альbedo* — очень высока. Допустим теперь, что за счет некоторого возмущения, природу которого здесь можно не уточнять, одно из ледовых покрытий несколько продвинулось к экватору. Поскольку это приводит к возрастанию глобального альbedo, Земля будет в результате поглощать меньше энергии, и это приведет к дальнейшему охлаждению системы. Вследствие этого ледовое покрытие продвинется еще сильнее и температура еще более понизится. Таким образом, перед нами — петля положительной обратной связи, весьма сходная с химическим автокатализом. Заметим, что это лишь одно из многочисленных сложных взаимодействий, происходящих в климатической системе.

Следует также иметь в виду, что основную часть солнечной энергии Земля получает в экваториальных областях. В полярных же областях, получающих значительно меньше энергии, за счет более высокого отражения гораздо большая доля энергии возвращается в космос. Таким образом формируется и поддержи-

вается меридиональный перепад температур между полюсами и экватором. Эта ситуация аналогична рассмотренной в разд. 1.3 системе с тепловыми ограничениями, что указывает на *неравно-весность климатической системы*.

Неравновесность, обратные связи, переходные явления, эволюция... Мы снова встречаемся с теми же ингредиентами сложного поведения, что и в предыдущих примерах, но только на этот раз в гораздо более впечатляющем планетарном масштабе.

1.9. Связь между силами и корреляциями. Подведение итогов

Кусок железа при температуре выше 1044 К не проявляет сколько-нибудь заметных магнитных свойств. Однако при охлаждении ниже этой «критической» температуры его можно намагнитить. Это типичный пример широкого класса естественных явлений, известных под названием *фазовых переходов*. Выше критической температуры вещество изотропно, т. е. ни одно из его свойств не характеризуется каким-либо выделенным направлением. Однако ниже этой температуры, в определенной для данного материала области (так называемой области Вайса), возникает намагниченность, которая является векторным свойством, связанным с определенным направлением в пространстве. Вещество становится анизотропным, или, выражаясь более научно, нарушается *вращательная симметрия*, характеризующая состояние ненамагниченности.

Имеется также множество других примеров. Так, жидкость — это такое состояние вещества, в котором молекулы движутся во всех возможных направлениях и не чувствуют друг друга на расстояниях, превышающих несколько стоимиллионных долей сантиметра (см. рис. 1.1, а). Поэтому ее можно рассматривать как однородное вещество, в котором все точки пространства эквивалентны. Будем теперь медленно и однородно охлаждать эту систему. Ниже определенной температуры (например, 0°C для чистой воды при обычном давлении) мы получим кристаллическую решетку — новую, твердую фазу вещества. Различные свойства этой фазы, такие, например, как плотность, уже не остаются одними и теми же при движении в различных направлениях. Иными словами, происходит нарушение *трансляционной симметрии*, характеризующей жидкое состояние.

В обоих случаях нарушение симметрии сопровождается возникновением новых свойств вещества, позволяющих говорить об *упорядоченном состоянии* вещества. Например, в кристаллической решетке молекулы совершают малые колебания около регулярно расположенных в пространстве положений, которые в разных случаях могут находиться в узлах кубической или гексаго-

нальной решетки, или по центрам соответствующих граней и т. д.

Что здесь происходит? Может быть, мы просто забыли включить эти явления в наш длинный список переходов к самоорганизации и сложному поведению?

В отсутствие взаимодействий молекулы вещества движутся свободно во всех направлениях, и это движение полностью характеризуется их кинетической энергией $(1/2)m_i v_i^2$ (m_i и v_i — массы и скорости молекул соответственно), не меняющейся со временем. Мы знаем, однако, что в любой физической системе межмолекулярное взаимодействие имеет место, причем оно обусловлено главным образом силами электромагнитной приро-

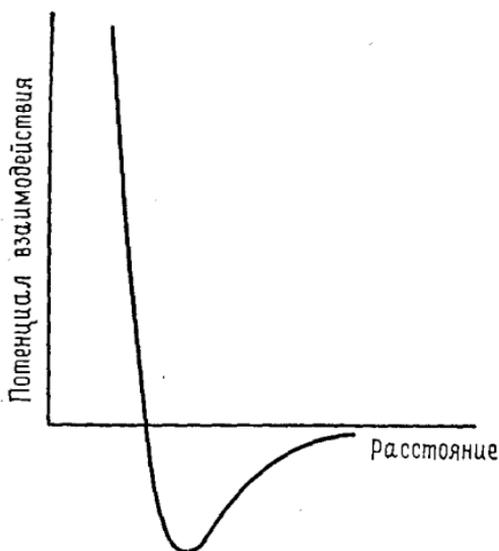


Рис. 1.18. Типичный вид потенциала межмолекулярного взаимодействия в нейтральном газе. Соответствующее минимуму этой кривой межмолекулярное расстояние составляет несколько ангстрем, а значение потенциальной энергии в этой точке порядка $k_B T_c$, где k_B — постоянная Больцмана, T_c — критическая температура перехода жидкость — пар.

ды. Типичным примером таких сил служат вандерваальсовы силы, являющиеся основными в электрически нейтральной жидкости (рис. 1.18). Эти силы — *короткодействующие*, т. е. они быстро исчезают, как только расстояние между частицами превысит несколько молекулярных диаметров.

В случаях когда плотность в системе низка или температура очень высока, межмолекулярные силы не проявляются заметным образом: доминирование кинетической энергии приводит к неупорядоченному поведению вещества. Однако при понижении температуры (или при повышении давления) кинетическая и по-

тенциальная энергии стремятся поменяться ролями. В конце концов доминировать начинают межмолекулярные взаимодействия и система принимает такую конфигурацию, в которой соответствующая потенциальная энергия имеет минимальное возможное значение. Это очень напоминает движение тела, перемещающегося под действием силы тяжести в самую глубокую впадину и затем колеблющегося около этого положения *механического равновесия*. Именно такого рода пространственные структуры возникают в результате фазового перехода. Будучи всецело опосредованной межмолекулярными взаимодействиями, такая структура характеризуется определенной микроскопической длиной порядка нескольких ангстрем (пример — период кристаллической решетки), сравнимой с радиусом этих взаимодействий. Кроме того, поведение вещества не зависит от времени, если внешняя среда находится в стационарном состоянии. В свете этих замечаний можно сказать, что обусловленная фазовым переходом упорядоченность приводит к объектам типа «окаменелостей».

Такие свойства резко контрастируют со свойствами переходных явлений, описанными в предыдущих разделах настоящей главы. В самом деле, пространственные масштабы системы после таких переходов являются *макроскопическими*, сравнимыми с характерными масштабами биологических систем или даже превышающими биомасштабы. Кроме того, в результате таких переходов помимо пространственных картин могут возникать и длительно существовать разнообразные явления, зависящие от времени, причем опять-таки с макроскопическими масштабами.

Причина столь фундаментального различия состоит в том, что организующим фактором теперь является неравновесность системы, не сводящаяся к межмолекулярным взаимодействиям. Поскольку понятие неравновесности подразумевает возможность выделения частей системы, содержащих громадное число молекул, возникают новые явления, обусловленные состоянием вещества и характеризующиеся крупномасштабными корреляциями. Следует отметить, что в дендритоподобных структурах типа снежинки (см., например, рис. 1.1, б или 1.12) эти два типа упорядоченности накладываются друг на друга и могут быть легко сопоставлены. В самом деле, лежащая в основе такой структуры кристаллическая решетка имеет весьма малое отношение к возникновению дендритов, их размеру и расстоянию между ветвлениями, которые на несколько порядков величины превышают любую кристаллографическую длину. Короче говоря, об интересующем нас здесь сложном поведении можно думать как о фазовом переходе нового типа, в котором понижение температуры заменено созданием и усилением подходящих неравновесных условий. По этой причине, говоря о явлениях самоорганизации,

приводящих к сложным типам поведения, мы будем иногда пользоваться выражением «неравновесный фазовый переход».

Опосредованные силами явления нарушения симметрии, в некоторых отношениях аналогичные фазовым переходам, возникают также на уровне фундаментальных взаимодействий. Одним из наиболее поразительных примеров такого рода является нарушение симметрии между веществом и антивеществом. Хорошо известно, что в релятивистской квантовой механике каждой заряженной частице соответствует античастица с той же массой, но с противоположным зарядом — мы имеем электроны и позитроны, протоны и антипротоны и т. д. Существование античастиц подтверждено специальными экспериментами. И все же, несмотря на симметрию фундаментальных уравнений физики относительно перехода от вещества к антивеществу, доля антивещества в наблюдаемой части Вселенной пренебрежимо мала по сравнению с долей вещества. Как могла произойти такая асимметрия?

Один из возможных ответов дает глубокая единая теория сильного, электромагнитного и слабого взаимодействий элементарных частиц, которая предсказывает поведение вещества при чрезвычайно высоких температурах, имевших место спустя примерно 10^{-35} с после изначального взрыва Вселенной. В частности, не исключено, что вещество претерпело один или несколько фазовых переходов при некоторых критических температурах по мере его охлаждения вследствие расширения Вселенной. Согласно единой теории, выше температуры первого перехода вещество находилось в совершенно симметричном состоянии. Поведение кварков — составных частей протонов и нейтронов, а также лептонов, включая электроны, нейтрино и их античастицы, было тогда совершенно идентичным. Однако ниже точки перехода вследствие нарушения симметрии начали проявляться различия. В конечном счете эти дифференцированные частицы превратились в исходный материал, из которого возникают звезды, планеты, все живые существа. Следовательно, в этом смысле переход, опосредованный упомянутыми тремя чрезвычайно короткодействующими взаимодействиями, достиг макроскопического масштаба, охватывая современную Вселенную, которая сама представляется своеобразным реликтом события, произошедшего в чрезвычайно далеком прошлом. Однако между этим утверждением и предшествовавшими попытками противопоставить обычные фазовые переходы и неравновесные фазовые переходы на основе сильных различий масштабов нет никакого противоречия. В самом деле, ранняя история Вселенной и, в частности, ее расширение отмечены переходом от состояния теплового равновесия к состоянию, в котором нарушено равновесие между различными компонентами вещества, а также между веществом и излучением. В таком неравновесном мире могут иметь место

крупномасштабные переходы с нарушением симметрии типа тех, что уже упоминались выше. Следовательно, в этом отношении мы можем рассматривать современное вещество как результат некоторого первичного неравновесия.

С другой стороны, для адекватного описания ранней истории Вселенной необходимо воспользоваться общей теорией относительности. Это выдвигает на арену тяготение — четвертую фундаментальную силу природы, хорошо известную благодаря своему *дальнодействию*.

По этой причине тяготение следует рассматривать как основной организующий фактор, опосредующий во Вселенной переход от равновесия к неравновесию и тем самым позволяющий микроскопическим событиям проявляться в глобальном масштабе. В известном смысле именно благодаря гравитации и расширению переходы в ранней Вселенной наделены чертами одновременно равновесных и неравновесных фазовых переходов типа тех, что мы наблюдаем ежедневно в наших более скромных масштабах.

СЛОВАРЬ СЛОЖНОГО

В предыдущей главе мы руководствовались точкой зрения естествоиспытателя, стараясь из наблюдений над внешним миром получить интуитивное представление о сложном на различных примерах переходных явлений, приводящих к самоорганизации. При этом мы встретились с рядом новых понятий, с которыми мы сталкивались и далее в рассмотренных примерах. Сейчас наступает самый подходящий момент для более глубокого и систематического рассмотрения этих понятий, что позволит разработать соответствующий *словарь сложного*. Тем не менее мы не станем утверждать, что к концу настоящей главы читатель получит возможность конструктивного использования этих понятий — такую задачу мы ставим в гл. 3—5.

2.1. Консервативные системы

С самого момента зарождения абстрактного мышления, по существу совпавшего с началом развития науки и современной цивилизации, ключевую роль играло понятие *сохранения*. Великий греческий мыслитель Фалес, традиционно считающийся одним из первых философов, придерживался той точки зрения, что все многообразие природы можно свести к единственному фундаментальному веществу — воде. Следовательно, говоря современным языком, он рассматривал мир как *консервативную систему*, в которой, невзирая на разнообразные обменные процессы между разными частями, существует некий первичный элемент, не подверженный изменениям. Аналогичные соображения, должно быть, оказали сильное влияние на молодого Платона, чья теория идей по существу представляет собой блестящую попытку отыскания перманентности на более абстрактном уровне.

Теперь мы совершим скачок вперед на два тысячелетия, к другому великому периоду истории человечества — периоду Ньютона и Лейбница. С появлением классической механики идея консервативности существенно укрепилась: если до того ее единственной опорой была некоторая непроверяемая аксиома, то теперь она стала следствием ряда законов, дающих количественное объяснение движению небесных тел и движению различ-

ных предметов под действием земного притяжения. В самом деле, из второго и третьего законов механики Ньютона, связывающих соответственно силы с ускорениями и действие с противодействием, следует, что в системе взаимодействующих точечных масс в отсутствие внешнего влияния остаются постоянными следующие три величины (см. рис. 2.1):

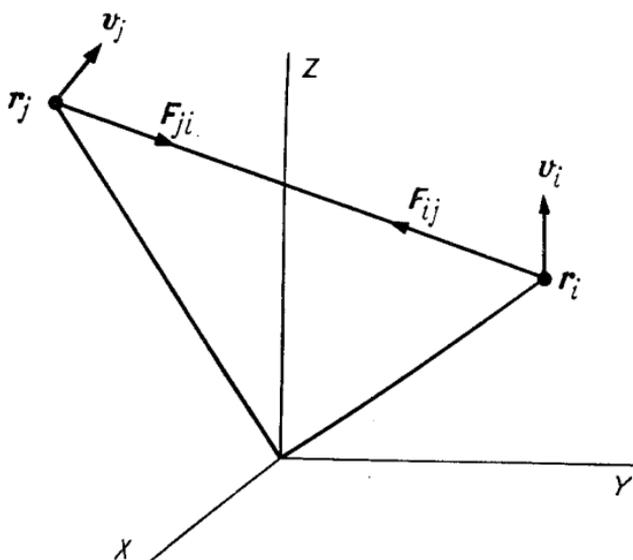


Рис. 2.1. Центральные силы взаимодействия между двумя частицами, движущимися в точках r_i и r_j со скоростями v_i и v_j . Согласно третьему закону Ньютона $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$.

1. Полная энергия:

$$\begin{aligned}
 E &= \text{кинетическая энергия} + \text{потенциальная энергия} = \\
 &= \frac{1}{2} \sum_i m_i v_i^2 + \sum_{i \neq j} V_{ij} \quad (2.1a)
 \end{aligned}$$

где m_i и v_i — масса и скорость i -й материальной точки соответственно, а V_{ij} — потенциальная энергия взаимодействия i -й и j -й точек, связанная с действующей между ними силой \mathbf{F}_{ij} соотношением

работа силы = изменение потенциальной энергии

Если сила зависит лишь от расстояния между материальными точками $\mathbf{r} \equiv \mathbf{r}_j - \mathbf{r}_i$, то это соотношение принимает вид

$$\mathbf{F}_{ij}(\mathbf{r}) = -\nabla_{\mathbf{r}} V_{ij}(\mathbf{r}) \quad (2.1б)$$

где символом $\nabla_{\mathbf{r}}$ обозначена производная от потенциальной энергии по расстоянию.

2. Полное количество движения:

$$P = \sum_i p_i = \sum_i m_i v_i \quad (2.2)$$

3. Полный момент количества движения:

$$L = \sum l_i = \sum_i m_i r_i \times v_i \quad (2.3)$$

где l_i означает момент количества движения i -й материальной точки относительно определенной точки пространства, выбранной для удобства в качестве начала координат.

Все эти законы сохранения тесно связаны с ключевым свойством основных уравнений механики, а именно с тем обстоятельством, что динамически связанной с силой величиной является ускорение, т. е. *вторая производная* координаты (но не скорость, как думал Аристотель!):

$$m_i \frac{d^2 r_i}{dt^2} = \sum_j F_{ij} (r_i - r_j) \quad (2.4)$$

Это означает, что, наблюдая обратное во времени движение ($t' = -t$), можно убедиться в справедливости тех же свойств, которые установлены обычным земным наблюдателем для положительного направления времени. Иными словами, с учетом $d^2 r/dt^2 = dv/dt$ и $v = dr/dt$ можно видеть, что преобразование $t' = -t$ вызывает преобразование $v' = -v$, оставляя ускорение неизменным — движение материальных точек происходит по тем же траекториям, но в обратном направлении. В этом нет ничего особенного, поскольку по тем же траекториям, но в обратном направлении можно пройти и без обращения времени, если сообщить нашим материальным точкам нужные начальные скорости. Другими словами, в самих уравнениях движения нет ничего такого, что позволило бы различить прямое и обращенное движение.

Таким образом, классическая механика представляется наукой о системах, являющихся одновременно консервативными и обратимыми во времени. Эти особенности особенно наглядны в рамках элегантного подхода, основанного на гамильтоновом формализме, в котором структура уравнений движения непосредственно связана с энергией — самой важной из трех сохраняющихся величин. Более конкретно, вводя гамильтониан H

$$H = H(p, q) = E = \text{const}$$

где p и q являются обобщенными импульсом ($p = mv$) и координатой ($q = r$) из ньютоновского формализма, имеем

$$H = \sum_i \frac{p_i^2}{2m_i} + \frac{1}{2} \sum_{i,j} V_{ij}(q_i - q_j) \quad (2.5a)$$

Легко видеть, что уравнение (2.4) и соотношение $v = dr/dt$ эквивалентны следующим выражениям:

$$\frac{d}{dt} q_i = \frac{\partial H}{\partial p_i}, \quad \frac{d}{dt} p_i = -\frac{\partial H}{\partial q_i} \quad (2.5b)$$

Сегодня динамика в гамильтоновой форме представляет собой важную часть науки. В гл. 3 мы кратко рассмотрим революционное развитие, имевшее место в этой области с 50-х годов. Здесь же в завершение настоящего раздела мы приведем несколько примеров консервативных систем.

Одной из наиболее известных консервативных систем является маятник (см. рис. 1.8). Обозначим расстояние до точки подвеса l , ускорение свободного падения g и угол отклонения маятника от вертикали θ . Кроме того, в нижней точке потенциальная энергия равна нулю. Тогда полную энергию — гамильтониан — можно записать в виде

$$H = \text{кинетическая энергия} + \text{потенциальная энергия} = \\ = \frac{1}{2} mv^2 + mgl(1 - \cos \theta)$$

Поскольку скорость v направлена тангенциально к траектории материальной точки, имеем $v = l(d\theta/dt)$ и тем самым

$$H = ml \left[\frac{1}{2} l \left(\frac{d\theta}{dt} \right)^2 + g(1 - \cos \theta) \right] = \text{const} \quad (2.6a)$$

Очень полезно мысленно представлять эволюцию системы в некотором абстрактном пространстве, известном под названием *фазового* (в гл. 3 этот вопрос рассматривается подробнее), в котором в качестве координат служат обычные пространственные координаты и импульсы, фигурирующие в уравнениях Гамильтона (2.5b). В случае маятника $q = \theta$, $p = ml(d\theta/dt) \equiv ml\dot{\theta}$ и фазовым пространством является плоскость $(\theta, \dot{\theta})$. Таким образом, мгновенное состояние маятника в этом пространстве соответствует точке, называемой *изображающей*. Величины θ и $\dot{\theta}$ являются однозначными функциями времени. Это означает, что изображающая точка движется по некоторой кривой в фазовом пространстве — фазовой траектории. Соответствующей фазовой скоростью будет служить тангенс угла наклона касательной к этой кривой (рис. 2.2). Не следует путать фазовую траекторию и скорость с обычными траекторией и скоростью тела, движущегося в физическом пространстве (несмотря на свою полезность, последние менее удобны для детального описания движения).

Частным случаем фазовой траектории является *положение равновесия*, в котором как скорость, так и ускорение равны

нулю. Эта траектория является вырожденной, поскольку она представлена единственной точкой в фазовом пространстве. В случае маятника имеются два таких положения равновесия (рис. 2.2), соответствующие положению материальной точки ниже ($\theta=0$) и выше ($\theta=\pm\pi$) точки подвеса по вертикали. Как мы вскоре увидим, эти точки имеют совершенно разный статус.

Из элементарных курсов механики и опытных наблюдений известно, что в случае малых начальных скоростей и отклонений от вертикального луча, исходящего вниз от точки подвеса, маятник совершает во времени незатухающие колебания. Это значит, что и θ и $\dot{\theta}$ пробегают одни и те же значения с каждым периодом колебаний. В представлении фазового пространства это соответствует фазовой траектории в виде замкнутой кривой (a или a' на рис. 2.2), вдоль которой сохраняется энергия (2.6а).

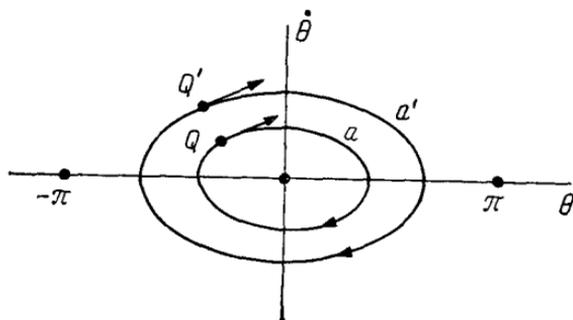


Рис. 2.2. Фазовое пространство $(\dot{\theta}, \theta)$ маятника ($-\pi < \theta < \pi$); $\theta = \dot{\theta} = 0$ — устойчивое равновесие; $\theta = 0$, $\theta = \pm\pi$ — неустойчивое равновесие; a , a' — периодические траектории, соответствующие двум различным значениям энергии; Q , Q' — изображающие точки на этих траекториях.

Другим известным примером консервативной системы является гармонический осциллятор, который моделируется движением без трения частицы с массой m , прикрепленной к горизонтально расположенной идеально упругой пружине, другой конец которой жестко фиксирован. Действующая на такую частицу сила пропорциональна отклонению x длины пружины от ее равновесного значения. Обозначая соответствующий коэффициент пропорциональности через k , гамильтониан системы можно записать в виде суммы кинетической энергии частицы и упругой энергии пружины:

$$H = \frac{1}{2} m \dot{x}^2 + k \frac{x^2}{2} \quad (2.6б)$$

Еще один класс консервативных систем возникает в задаче о двух телах, взаимодействующих за счет сил тяготения или любых других сил, причем движение этих тел зависит лишь от аб-

солютного значения расстояния между ними. Однако осложнения в механике возникают очень быстро. Уже в случае двух маятников, соединенных пружиной, или в случае маятника, на который действует периодическая внешняя сила, или в случае трех взаимодействующих масс (знаменитая задача о трех телах) достигается предел современного знания. Несколько подробнее эти задачи мы обсудим в гл. 3.

Из сказанного выше отнюдь не следует, что консервативные системы встречаются только в механике. Вот еще один важный пример консервативных систем из области электромагнитных явлений: распространение света в вакууме, описываемое знаменитыми уравнениями Максвелла. Консервативные системы возникают также в квантовой механике, в которых изучаются свойства отдельных атомов или молекул в отсутствие внешних полей.

2.2. Диссипативные системы

Кроме консервативных систем, изучаемых в классической механике, нам нужно рассмотреть также системы, приводящие к необратимым процессам. Простейшим примером такого рода могут служить системы с трением.

Важная роль трения, представляющего собой особую форму диссипативного процесса, была осознана задолго до создания классической механики. Когда Аристотель высказал предположение, что все подлунные динамические системы в общем случае стремятся к равновесию, на самом деле он выражал идею о том, что нечто вроде «трения» должно замедлять движение. В этом плане классический принцип инерции, отражающий основную роль ускорения, а не скорости, соответствует некоторой идеализации, возникающей в результате пренебрежения трением.

Начиная с работ Фурье и Клаузиуса, в XIX в. возрос интерес к диссипативным системам, приводящим к необратимым процессам. Это было довольно естественно с учетом происшедшей тогда промышленной революции. Однако по той же причине диссипацию рассматривали тогда лишь в связи с исчерпанием доступной энергии.

Интересно, что один из великих древнегреческих философов, Платон, был глубоко убежден в том, что как постоянство, так и изменчивость являются составными частями реальности. Однако в XIX в. возникла конфликтная ситуация. Так, в физике необратимость и диссипация воспринимались как некоторая деградация, а, с другой стороны, биологическая эволюция, очевидно, также являющаяся необратимым процессом, ассоциировалась с возрастанием сложности. Возможно, благодаря своему техно-

логическому значению механика жидких сред исторически оказалась первой областью, в которой была полностью осознана решающая роль диссипативных процессов. Однако, по мере того как постепенно утверждалась молекулярная концепция строения вещества, аналогичная тенденция получила развитие в науке, химической кинетике, теории броуновского движения и различных типах транспортных явлений. Сегодня уже общепризнано, что диссипативные системы представляют собой весьма широкий и важный класс естественных систем. В большинстве рассмотренных в гл. 1 примеров по существу рассматривались явления, происходящие именно в таких системах, среди которых далеко не последнее место занимает сама жизнь.

Ярче всего различие между консервативными и диссипативными системами проявляется при попытке *макроскопического описания* последних, когда для определения мгновенного состояния системы используются такие коллективные переменные, как температура, концентрация, давление, конвективная скорость и т. д. При рассмотрении уравнений, управляющих поведением этих переменных, выясняется следующая их важная особенность: они не инвариантны относительно операции обращения времени в отличие от уравнений (2.4) и (2.5б). На этой основе можно ожидать, что чередование соответствующих событий будет необратимым.

Эта ситуация удивительно ярко иллюстрируется на примере химической реакции. Рассмотрим процесс, описываемый уравнением (1.2а). Скорость расходования частиц типа А пропорциональна частоте встреч молекул типов А и В, которая в случае разбавленных систем пропорциональна произведению их концентраций. Таким образом, имеем

$$\frac{dc_A}{dt} = -kc_Ac_B \quad (2.7a)$$

Очевидно, при обращении времени ($t' = -t$) и введении обозначений c'_A , c'_B для концентраций в зависимости от времени уравнение (2.7а) принимает вид

$$\frac{dc'_A}{dt'} = kc'_Ac'_B$$

Теперь это уравнение описывает процесс, в котором вещество типа А не расходуется, а производится. Разумеется, такой процесс неэквивалентен описываемому уравнением (2.7а).

В качестве дальнейших примеров диссипативных процессов можно рассмотреть теплопроводность и диффузию. Как показывает эксперимент, если в однородной жидкости возникает небольшая неоднородность, то такое возмущение со временем расплывается и постепенно исчезает. Аналогичная односторонне на-

правленная эволюция наблюдается в случае небольшого изменения температуры, внесенного быстро и локально в изотермическую жидкость. Количественное описание этих явлений, блестяще согласующееся с опытными данными, дается следующими уравнениями, называемыми соответственно уравнением Фика и уравнением Фурье:

$$\frac{\partial c}{\partial t} = D \nabla^2 c, \quad D > 0 \quad (2.76)$$

$$\frac{\partial T}{\partial t} = \kappa \nabla^2 T, \quad \kappa > 0 \quad (2.7в)$$

где c — концентрация некоторого вещества, растворенного в жидкости, T — температура, D — массовый коэффициент диффузии и κ — коэффициент температуропроводности. При обращении времени мы опять получаем совершенно другие законы:

$$\frac{\partial c'}{\partial t'} = -D \nabla^2 c', \quad \frac{\partial T'}{\partial t'} = -\kappa \nabla^2 T'$$

Согласно этим уравнениям, начальное возмущение температуры или концентрации будет не затухать, а возрастать.

Как концентрация, так и температура являются примерами «четных» переменных, поскольку знак этих переменных при обращении времени не меняется. Напротив, импульс частиц или конвективная скорость жидкости являются «нечетными» переменными, поскольку они являются производными по времени от переменных типа координаты и меняют знак при обращении времени. Это приводит к следующему общему свойству уравнения эволюции диссипативной системы. Обозначим полный набор макроскопических переменных такой системы через X_1, \dots, X_n . Эволюция этих переменных во времени будет описываться системой уравнений

$$\frac{\partial X_i}{\partial t} = F_i(X_1, \dots, X_n, \mathbf{r}, t, \dots) \quad (2.8)$$

Здесь функции F_i могут сколь угодно сложным образом зависеть от переменных X и их пространственных производных и явным образом — от пространственных координат \mathbf{r} и времени t . Тогда, если мы совершим операцию обращения времени $t' = -t$ в диссипативной системе, то по меньшей мере одна из функций скоростей F_i , соответствующая четной переменной X_i , должна содержать инвариантную часть, в то время как функция скорости F_i , соответствующая нечетной переменной X_i , должна содержать часть, меняющую знак при обращении времени. Примеры функций скоростей первого класса дают правые части уравнений (2.7а) — (2.7в), примером же второго класса является вклад вязкости в уравнение баланса импульса жидкости, участвующей в конвективном движении.

Дополнительное существенное продвижение на пути к различению между консервативными и диссипативными системами совершим в разд. 2.5, где будут обсуждаться второй закон термодинамики и его многочисленные следствия.

Как и в случае консервативных систем, для диссипативных систем также можно ввести удобное фазовое пространство. Оно включает в себя ансамбль имеющихся переменных и поэтому становится бесконечномерным пространством в случае непрерывной среды, где различные характеристики являются пространственно распределенными величинами [см. уравнения (2.7б) и (2.7в)]. Поэтому удобнее всего работать с фазовым пространством

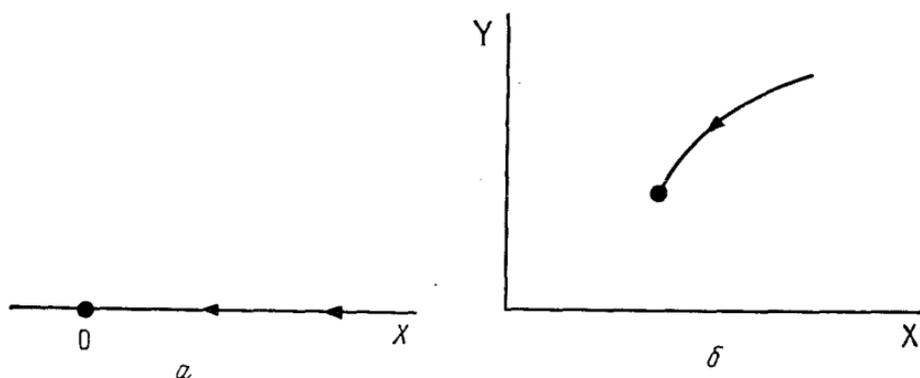


Рис. 2.3. Представление диссипативной системы в фазовом пространстве. *a* — система, описываемая одной переменной в соответствии с уравнением (2.7а); *б* — система с двумя переменными, уравнение (2.9б).

ством, когда оно содержит дискретное число переменных, и в особенности когда это число конечно и, желательно, невелико. Например, в случае уравнения (2.7а) фазовое пространство сводится к линии, на которой и находится фазовая траектория (рис. 2.3, *a*). Менее тривиальным примером является химическая реакция, описываемая следующей кинетической схемой:



Соответствующие кинетические уравнения имеют вид

$$\frac{dc_X}{dt} = k_1 c_A - k_2 c_X; \quad \frac{dc_Y}{dt} = k_2 c_X - k_3 c_Y \quad (2.9б)$$

Фазовые траектории для такой системы показаны на рис. 2.3, *б*.

Полезно иметь в виду, что некоторые диссипативные системы можно преобразовать к консервативному виду и привести их к

гамильтоновой форме. Примером может служить знаменитый механизм Лотки — Вольтерра



В этой системе имеется некоторый нетривиальный интеграл движения, играющий роль «гамильтониана». И все же, несмотря на свой кажущийся консервативный характер, эта система неинвариантна относительно обращения времени, поскольку обе переменные X и Y являются положительными. Поэтому нет смысла приписывать им свойства, аналогичные импульсу в классической механике, что необходимо для такой инвариантности.

Пока еще не рассматривался вопрос о связи между диссипативными и консервативными системами, а также вопрос о возможности перехода от одного описания к другому. Этому вопросу посвящена основная часть гл. 5.

2.3. Механическое и термодинамическое равновесия. Неравновесные ограничения

В механике равновесие представляет собой особое состояние, в котором как скорости, так и ускорения всех материальных точек равны нулю. Такого рода «состояние покоя» уже встречалось нам в разд. 1.1, 1.4 и 2.1. По определению результирующая сила, действующая на каждую точку в любой момент времени, равна нулю. Нарушение такого баланса сил приводит к нарушению равновесия. Это происходит, например, при разрушении металлического образца под нагрузкой.

Некоторые известные механические системы функционируют таким образом, что равновесие в них никогда не достигается. Например, земная орбита свидетельствует о наличии постоянного отклонения от равновесия в системе Солнце — Земля. Такое равновесие было бы возможным лишь при бесконечном удалении одного тела от другого.

Совершенно иной характер имеет понятие *термодинамического равновесия*, с которым мы неоднократно встречались в разд. 1.3, 1.4, 1.6 и 1.7. В отличие от случая механического равновесия входящие в систему молекулы подвержены действию несбалансированных сил и движутся во всех возможных направлениях, если только температура не становится чрезвычайно низкой. Поэтому здесь понятие равновесия относится к некоторым коллективным характеристикам системы в целом, таким, как температура, давление или концентрация химического ком-

понента. Попытаемся дать общее определение этого понятия, для которого рассмотренные в гл. 1 примеры будут служить частными случаями.

Рассмотрим некоторую систему, помещенную в определенную внешнюю среду (рис. 2.4). Каждый из этих объектов будем характеризовать некоторым набором характеристик, X_i и X_{ie} соответственно. Главное, что нам нужно сейчас знать об этих характеристиках,— это их отношение к различным *обменам*, происходящим между системой и внешней средой. Например, если система находится в сосуде с идеально жесткими стенками, проницаемыми для тепла и непроницаемыми для вещества, то одна из этих величин, скажем X , будет совпадать с *температурой* и будет управлять обменом энергии в виде тепла между системой и внешней средой.

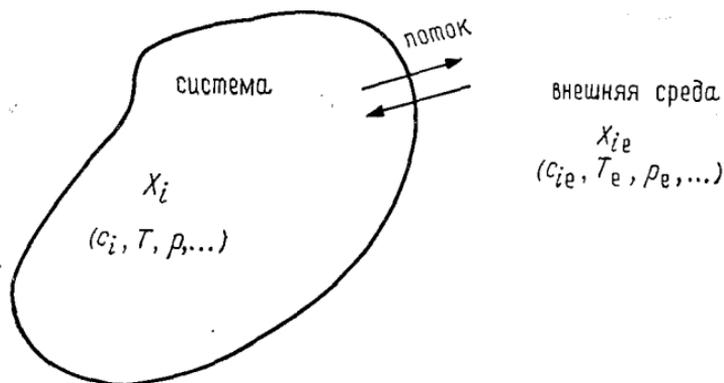


Рис. 2.4. Схематическое представление открытой системы.

Мы будем говорить, что система находится в термодинамическом равновесии, если ответственные за обмены характеристики имеют одинаковые значения в системе и во внешней среде. В предыдущем примере термодинамическое равновесие между системой и внешней средой равносильно соотношению $T = T_e$, справедливому в любой момент времени в любой точке системы. И все же, поскольку стенки сосуда для вещества непроницаемы, данная система может сильно отличаться от внешней среды по своему химическому составу. С другой стороны, если сделать стенки проницаемыми для некоторых химических веществ i , то термодинамическое равновесие между системой и внешней средой наступит тогда, когда обмены этими веществами прекратятся. В простейшем случае это означает, что соответствующие переменные, характеризующие состав системы и внешней среды, будут совпадать ($c_i = c_{ie}$), однако в более общем случае равновесие характеризуется равенством величин, известных под названием *химических потенциалов* ($\mu_i = \mu_{ie}$). Совершенно анало-

гично, если стенки сосуда не являются жесткими, то система может обмениваться с внешней средой механической энергией. Тогда в равновесии будет иметь место также равенство *давлений* ($p = p_e$). При наличии внешних сил могут возникать дополнительные переменные, характеризующие термодинамическое равновесие. В качестве таких сил могут выступать, например, электрическое поле, созданное в поляризуемой среде, механическое напряжение в деформируемом материале и т. д.

Отдельно следует упомянуть *изолированные системы*, в которых нет обмена ни энергией, ни веществом с внешней средой. По определению нельзя сказать, что в равновесии изолированная система находится в том же состоянии, что и внешняя среда. Но и в этом случае значительная часть высказанных соображений справедлива, поскольку некоторая небольшая часть изолированной системы испытывает действие внешней среды, которой для нее является остальная часть системы. По этой причине в равновесии различные части системы между собой термодинамически неразличимы.

В соответствии с данными определениями равновесие автоматически является стационарным состоянием ($\partial X_i / \partial t = 0$). Однако это стационарное состояние специального типа, поскольку ввиду тождественности по характеристикам X_i системе нечем обмениваться с внешней средой. Например, если равновесие обусловлено равенством температур, химических составов и давлений, то не будет систематических обменов тепловой или механической энергией, а также веществом. Это условие формулируется в виде равенства нулю *полных потоков* в системе:

$$J_i^e = 0 \quad (2.11)$$

Это, разумеется, не означает, что равновесие представляет собой некоторое совершенно застывшее состояние. Это условие следует понимать в динамическом смысле, как объяснялось в начале разд. 1.4, посвященного химическим реакциям. В частности, для каждого процесса, вызывающего изменение J_i (скажем, небольшое увеличение T по сравнению с T_e в малой части системы), найдется обратный процесс, дающий противоположный вклад в J_i . Это основное свойство *детального равновесия* является не чем иным, как проявлением обратимости во времени элементарных процессов в системе, описываемой консервативными уравнениями (разд. 2.1). Именно оно ответственно за большинство фундаментальных свойств равновесия типа, например, единственности и устойчивости, как это будет обсуждаться в разд. 2.4—2.6.

Обратимся теперь к *неравновесным состояниям*. Такие состояния будут связаны с неисчезающими потоками между системой и внешней средой, а также с различием в некоторых переменных

состояния X_i и X_{ie} . Эти различия могут иметь *переходный* характер в том смысле, что они могут мгновенно возникать благодаря некоторому начальному условию и постепенно релаксировать по мере установления равновесия между системой и внешней средой. Однако такие различия могут быть и постоянными, если создать и поддерживать соответствующие условия, которые мы будем называть *ограничениями*. Так, постоянная разность температур между двумя сечениями пластины или подкачка исходного реагента с одновременной эвакуацией конечного продукта в системе с химическими реакциями приведут к неравновесным ситуациям, в которых система не сможет прийти в состояние, тождественное состоянию внешней среды. Разумеется, из этих примеров отнюдь не следует делать вывод о том, что неравновесность представляет собой некоторое условие, искусственно наложенное умным экспериментатором. Как отмечалось в разд. 1.7, в окружающей нас среде мы часто являемся свидетелями неравновесных состояний типа, например, состояния биосферы, которая находится под действием потока энергии, возникающего благодаря равновесию между излучениями Солнца и Земли.

Вследствие постоянного или временного действия ограничений в неравновесном состоянии детального равновесия уже нет. В результате режим неравновесного состояния приобретает способность к изменениям: небольшие локальные отклонения от него не обязательно разрушаются постоянно возникающим противодействием — они могут быть усвоены или даже усилены системой, становясь тем самым источником новообразований и разнообразия. Как мы увидим в дальнейшем, это свойство лежит в основе способности неравновесных систем осуществлять бифуркационные переходы к новым состояниям, и эта способность делает неравновесное состояние непохожим на любое состояние равновесия.

2.4. Нелинейность и обратные связи

В предыдущем разделе мы дали «операционное» определение равновесных состояний и неравновесных ограничений. Теперь мы примем динамическую точку зрения, поставив себе целью выяснение связи между этими понятиями и законами эволюции переменных состояния. По сути дела, наши рассуждения в конце разд. 2.1 уже установили эту связь в случае механического равновесия. По этой причине в большей части настоящего раздела мы сосредоточим свое внимание на диссипативных системах, хотя многие понятия, которые нам предстоит ввести, можно легко перенести на консервативные системы.

В разд. 2.2 уже приводилось математическое выражение, описывающее временную эволюцию набора переменных состояния $\{X_i\}$ [уравнение (2.8)]. Мы перепишем эти уравнения, введя в них новую идею, возникшую в разд. 2.3 по поводу того, что эволюция обусловлена соответствующими неравновесными ограничениями. В более общем виде нам бы хотелось выразить ту идею, что на эволюцию влияет изменение некоторых характерных для данной задачи параметров, которые могут быть изменены внешним миром. Эти параметры мы будем называть *управляющими*. Обозначая их посредством λ , запишем уравнение (2.8) в виде

$$\frac{\partial X_i}{\partial t} = F_i(\{X\}, \lambda) \quad (2.12)$$

В диссипативной системе эти уравнения должны рассматриваться совместно с условиями, обсуждавшимися в конце разд. 2.2. Типичным примером таких диссипативных *законов эволюции* является уравнение (2.9б).

В силу разнообразия естественных процессов можно ожидать, что структура функций $\{F_i\}$ будет весьма специфически зависеть от характера рассматриваемой системы, а также от типа протекающих в ней процессов. В самом деле, на первый взгляд нет абсолютно никакой причины, по которой между течением жидкости между двух пластин в задаче Бенара и агрегацией амёб *Dictyostelium discoideum* или обледенением в четвертичном периоде могла бы существовать связь в каком-либо смысле. И все-таки, несмотря на такое поразительное разнообразие, можно выделить некоторые основные черты, позволяющие приступить к систематическому изучению сложных явлений.

Одним из таких свойств является соответствие решения уравнений (2.12) равновесию в отсутствие внешних ограничений при любом виде функций F_i . Поскольку равновесие является стационарным состоянием, отсюда следует

$$F_i(\{X_{\text{eq}}\}, \lambda_{\text{eq}}) = 0 \quad (2.13а)$$

В более общем случае для неравновесного стационарного состояния можно написать

$$F_i(\{X_s\}, \lambda) = 0 \quad (2.13б)$$

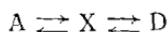
Эти соотношения налагают определенные ограничения. Например, законы эволюции должны быть таковы, чтобы выполнялось требование положительности температуры или химической концентрации, полученных как решение соответствующих уравнений.

Этот важный момент показывает, что анализ физических систем нельзя свести к некоторой математической игре. Во многих

отношениях с математической точки зрения физические системы можно рассматривать как весьма нетипические, или, в более современной терминологии, нехарактерные. Взять, например, те же консервативные гамильтоновы системы, построенные на исключительно строгом условии, согласно которому определенная комбинация координат и скоростей — энергия — остается постоянной, т. е. является интегралом движения.

Допустим теперь, что эта особенность физических систем обеспечена. Что еще можно сказать о возможных решениях уравнений (2.13а) или (2.13б)? Как раз здесь и начинает играть свою особую роль представление о нелинейности.

Рассмотрим следующий пример. Пусть X будет единственной переменной состояния, k — параметр и λ представляет собой внешнее ограничение. Представим теперь, что имеет место следующая реакция (символ X используется для обозначения не только концентрации, но и самого вещества):



Согласно этому механизму, переменная X удовлетворяет уравнению

$$\frac{dX}{dt} = \lambda - kX$$

Отсюда следует, что стационарное состояние определяется из уравнения

$$\lambda - kX_s = 0$$

или

$$X_s = \frac{\lambda}{k} \quad (2.14)$$

Это решение представлено на рис. 2.5, а. Строя график зависимости стационарного значения X_s от характеризующего внешнее ограничение параметра λ (в частности, этот параметр может соответствовать степени удаленности от равновесия), мы получаем прямую линию, проходящую через начало координат. В результате мы получаем типичный пример *линейного закона*. Для каждого значения λ , соответствующего данному внешнему ограничению, имеется одно и только одно стационарное состояние X_s . Более того, поскольку для определения прямой линии достаточно двух точек, можно предсказывать совершенно однозначно стационарное значение X при любом значении параметра λ , если иметь экспериментально определенную пару точек (λ_1, X_{1s} ; λ_2, X_{2s}). Иными словами, поведение такой линейной системы вынужденно похоже на равновесие даже при наличии сильно неравновесных ограничений.

Посмотрим теперь на рис. 2.5, б, соответствующий *нелинейному закону*, связывающему стационарное значение X_s с управляющим параметром λ . Разумеется, приведенная на этом рисунке кривая служит лишь для иллюстративных целей и приводится в качестве некоторого примера, поскольку имеется бесчисленное количество различных нелинейных зависимостей. Когда λ меньше λ_1 или больше λ_2 , решение однозначно, как и на рис. 2.5, а. Однако для $\lambda_1 < \lambda < \lambda_2$ в системе может иметься несколько различных решений. Это качественное изменение наступает после перехода через определенное критическое значение управляющего параметра и уже позволяет понять, каким образом оказывается возможной левая или правая ориентация вращения в ячейках Бенара, равно как и множество других экспериментальных фактов, рассмотренных в гл. 1.

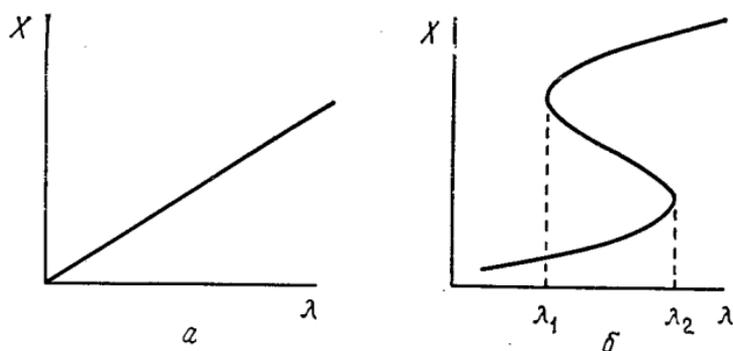
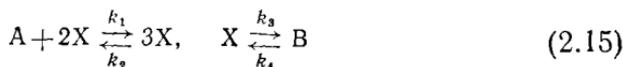


Рис. 2.5. Влияние параметра λ на переменную состояния X . а — линейный закон; б — нелинейный закон. В случае б следует отметить чувствительность ответа X к изменению λ в окрестности значений λ_1 и λ_2 .

Поразительное различие между линейными и нелинейными законами проявляется в применимости или нарушении принципа суперпозиции. Так, в линейных системах результатом совместного действия двух различных факторов является простая суперпозиция результатов каждого из них, взятого отдельно. Однако в нелинейных системах небольшое увеличение внешнего воздействия может привести к очень сильным эффектам, несоизмеримым по амплитуде с исходным воздействием. Примером могут служить скачки, совершаемые системой при изменении параметра λ вблизи одного из критических значений (см. рис. 2.5, б).

Заметим, что при недостаточной удаленности от равновесия сама по себе нелинейность не может привести к множественным решениям. В самом деле, при равновесии возникает дополнительное условие в виде детального равновесия. Это условие ограничивает и даже фиксирует единственным образом переменную состояния. Более того, в силу непрерывности эта един-

ственность должна распространяться на некоторую область состояний вблизи равновесия. Полезно проиллюстрировать такое положение на простом примере двух сопряженных реакций



Здесь в качестве единственной переменной состояния принята концентрация x промежуточного продукта X . При этом подразумевается, что вещества A и B непрерывно поступают в систему или удаляются из нее, обеспечивая тем самым постоянство концентраций a и b . При равновесии из принципа детального равновесия следует

$$k_1 a x^2 = k_2 x^3 \quad (\text{равенство скоростей прямой и обратной реакций для первого из уравнений (2.15)})$$

$$k_3 x = k_4 b \quad (\text{равновесие во второй реакции})$$

Эти соотношения единственным образом определяют x_{eq} и, кроме того, налагают некоторое условие на концентрации a и b :

$$x_{\text{eq}} = \frac{k_4 b_{\text{eq}}}{k_3} = \frac{k_1 a_{\text{eq}}}{k_2}$$

$$\left(\frac{b}{a}\right)_{\text{eq}} = \frac{k_1 k_3}{k_2 k_4} \quad (2.16)$$

С другой стороны, в стационарном состоянии вдали от равновесия должно выполняться лишь требование сбалансированности суммарного влияния прямых реакций (2.15) суммарным влиянием обратных реакций, откуда следует уравнение

$$-k_2 x_s^3 + k_1 a x_s^2 - k_3 x_s + k_4 b = 0 \quad (2.17)$$

Это кубическое уравнение относительно x_s , и при определенных значениях a и b оно может иметь до трех решений, что аналогично случаю, изображенному на рис. 2.5, б, и отличается от случая равновесия. Поэтому можно сказать, что неравновесность выявляет потенциальные возможности, содержащиеся в нелинейностях и как бы «дремлющие» в равновесии или вблизи него.

Возможность множественных решений в нелинейных системах автоматически ставит вопрос о *выборе* между различными исходами — ведь очевидно, что в данной ячейке Бенара течение не может быть одновременно право- и левонаправленным! Чтобы толком разобраться в этом решающем вопросе, необходимо познакомиться с некоторыми дополнительными понятиями, обсуждаемыми в последующих разделах. В заключение же этого раздела мы приведем список некоторых важных типов нелинейностей, постоянно встречающихся в различных областях.

1. Механика жидких сред. Нелинейных задач в механике жидких сред достаточно много главным образом по той причи-

не, что скорость изменения той или иной характеристики среды помимо прочих факторов зависит от переноса этой характеристики за счет течения среды, скорость которого v является одной из переменных данной задачи. Так, в уравнении сохранения импульса явно фигурирует нелинейный член $(v\nabla)v$. Кроме того, во многих задачах неизвестными переменными являются также температура T и плотность ρ . Дополнительные нелинейности возникают за счет сопряжения между этими переменными и конвективной скоростью, представленного в уравнениях баланса массы и внутренней энергии соответственно членами $\nabla(\rho v)$ и $(v\nabla)T$.

2. Химия и биология. В химии и биологии нелинейности могут возникать самопроизвольно независимо от наличия пространственных неоднородностей. Большинство химических реакций происходит в результате столкновений. Например, $X + Y \rightarrow D$ означает, что для данного превращения должны взаимодействовать молекулы типов X и Y , в результате чего скорость реакции оказывается пропорциональной произведению $c_X c_Y$, т. е. нелинейной функции. Однако, как уже подчеркивалось в разд. 1.4 и 1.6, особый интерес представляют нелинейности, связанные с *регуляцией*. В таком случае вещество X может ускорять или замедлять собственное производство или производство другого компонента, в свою очередь замкнутого на первое вещество. Характерными примерами такого рода могут служить автокатализ, иллюстрируемый прямой реакцией (2.15), и тепловая обратная связь, обсуждавшаяся в разд. 1.4 [см. уравнение (1.5) и рис. 1.11]. Эти примеры взяты из области классической химии. В биологии же регуляция тесно связана с особой структурой и реакционной способностью ферментов. Во многих случаях это приводит к нелинейностям вида

$$\left(\frac{dc}{dt}\right)_{\text{act}} \sim \frac{ac^m}{1+bc^n}$$

При достаточно больших концентрациях это приводит к активации в случае $m \geq n$ и ингибированию в случае $m < n$. Аналогичные математические выражения встречаются и в контексте других задач, например в задачах об электрических цепях с активными элементами или в случае излучательных процессов в задаче о климатической динамике.

2.5. Многогранность второго закона термодинамики

Мы видели, что нелинейность в совокупности с неравновесными ограничениями ответственна за множественность решений и, следовательно, за диверсификацию типов поведения системы. До сих пор в настоящей главе мы принимали эти решения как

нечто данное, не интересуясь особо причинами их появления, и лишь анализировали некоторые их свойства. Теперь настало время задать вопрос: *как* устанавливаются с течением времени эти решения уравнений эволюции типа (2.12)? В случае консервативных систем ответ на этот вопрос весьма прост, по крайней мере в принципе: эволюция целиком диктуется начальными условиями (см., например, рис. 1.8 и 2.2), фиксирующими различные интегралы движения, а значит, и сами траектории с помощью соотношений типа (2.6). На практике, однако, этот рецепт оказывается полезным лишь в сравнительно простых ситуациях. Обсуждение остальных случаев мы отложим до гл. 3.

В случае диссипативных систем поиск критериев эволюции приводит к некоторым глубоким и неожиданным результатам. В самом деле, из неинвариантности уравнений для переменных состояния относительно обращения времени следует, что для таких систем характерно *необратимое* приближение к конечному состоянию, которое, как обычно считается, должно соответствовать независящему от времени режиму.

В случае изолированных систем, в которых нет никаких обменов с внешней средой, эта тенденция к необратимости выражается знаменитым вторым законом термодинамики. В соответствии с этим законом существует функция переменных состояния системы (которую мы в дальнейшем будем называть *функцией состояния*), изменяющаяся монотонно в процессе приближения к единственному конечному состоянию термодинамического равновесия. Обычно в качестве такой функции выбирается энтропия S — величина, введенная феноменологически при изучении тепловых машин и в дальнейшем изученная на более микроскопическом уровне Больцманом, Гиббсом и другими. Формулировка второго закона имеет вид

$$\frac{dS}{dt} \geq 0 \quad (\text{для изолированных систем}) \quad (2.18)$$

Обычно этот закон интерпретируется как тенденция к возрастанию разупорядоченности, поскольку в микроскопической теории энтропия ассоциируется с числом допустимых состояний при определенных условиях, наложенных на систему. Отсюда вытекает, что чем меньше допустимых состояний имеет система, тем более она упорядочена. Например, энтропия твердого тела вблизи абсолютного нуля температуры мала, а энтропия газа при комнатной температуре и атмосферном давлении, когда частицы сталкиваются и движутся во всех возможных направлениях, велика.

Такая интерпретация энтропии и второго закона термодинамики довольно очевидна для систем, состоящих из слабо взаимодействующих частиц. Именно к таким системам относились

соображения, исходно развитые Больцманом. Например, как эксперимент, так и здравый смысл подсказывают, что газ, занимающий в начальном состоянии лишь половину ящика (рис. 2.6, *a*), со временем перейдет к состоянию с более или менее равномерным распределением вещества по всему объему ящика (рис. 2.6, *б*). Ясно, что при этом неупорядоченность увеличится, так как, вообще говоря, однородное состояние *б* менее структурировано, нежели исходное состояние *a*. Соответственно возрастет и число макроскопически эквивалентных микросостояний, поскольку объем газа в состоянии *a* составляет половину объема в состоянии *б*. Таким образом, если определить энтропию как меру числа доступных микросостояний, становится ясно, что изображенный на рис. 2.6 процесс превосходно согласуется с уравнением (2.18), отражающим тенденцию к необратимому росту разупорядоченности.



Рис. 2.6. Спонтанный переход в однородное состояние газа, в начальном состоянии *a* содержавшегося в левой половине ящика. Однородное состояние *б* устанавливается после того, как газ получает доступ ко всему объему. В этом процессе возрастают как энтропия, так и неупорядоченность.

Для систем из сильно взаимодействующих частиц такая интерпретация непосредственно не применима прежде всего по той причине, что для подобных систем существует возможность перехода к более упорядоченному состоянию за счет механизма фазовых переходов.

Каким бы ни было детальное микроскопическое определение энтропии для таких систем, можно ожидать, что в отсутствие внешних ограничений термодинамически равновесное состояние системы будет целиком определяться межмолекулярными взаимодействиями. Отсюда вытекает, что в изолированной физической системе, в которой гравитационные силы не играют заметной роли в динамике, после переходного процесса, описываемого соотношением (2.18) и осуществляющегося необратимо, в системе установится конечное состояние равновесия. Характерный масштаб пространственной упорядоченности будет при этом обязательно микроскопическим, порядка нескольких ангстрем. Типичным примером такого рода является кристалл, в котором постоянная решетки близка по порядку величины к молекулярному диаметру. Очевидно, в таком состоянии никогда не смогли бы возникнуть ячейка Бенара или химические часы!

Обратимся теперь к неизолированным системам, обменивающимся с внешней средой энергией или веществом. Изменение энтропии будет теперь представлено в виде суммы двух членов. Один из них обусловлен происходящими обмeнами, обозначим его $d_e S$ (поток энтропии). Другой обусловлен процессами внутри системы, обозначим его $d_i S$ (производство энтропии):

$$\frac{dS}{dt} = \frac{d_i S}{dt} + \frac{d_e S}{dt} \quad (2.19a)$$

Для изолированной системы $d_e S = 0$, так что уравнение (2.19a) с учетом (2.18) сводится к $dS = d_i S \geq 0$, т. е. к обычной формулировке второго закона. Однако даже в случае неизолированной системы член $d_i S$ будет описывать те (необратимые) процессы, которые будут по-прежнему иметь место даже в отсутствие потокового члена $d_e S$. Таким образом, мы приходим к следующей обобщенной форме второго закона термодинамики:

$$\frac{d_i S}{dt} \geq 0 \quad (\text{для неизолированных систем}) \quad (2.19б)$$

До тех пор пока $d_i S$ строго положительно, в системе будут протекать необратимые процессы, причем непрерывно. Таким образом, неравенство $d_i S > 0$ эквивалентно условию диссипативности, обсуждавшемуся в разд. 2.2. С другой стороны, если $d_i S$ обращается в нуль, то процесс будет обратимым и будет всего лишь захватывать соседние состояния равновесия путем медленного изменения потокового члена $d_e S$. Именно эта последняя ситуация обычно изучается во вводных термодинамических курсах, где рассматриваются циклы Карно и различные типы тепловых машин.

В число наиболее общих необратимых процессов, дающих вклад в $d_i S$, входят химические реакции, теплопередача, диффузия, вязкая диссипация и релаксационные явления в электрически или магнитно-поляризуемых системах. Для каждого из этих явлений можно определить соответствующий внутренний поток J_k , по существу отражающий скорость течения процесса, а также движущую силу X_k , отражающую имеющееся неравновесное ограничение. Замечательная особенность члена $d_i S$ состоит в том, что он является квадратичной формой по J_k и X_k , по крайней мере для того диапазона явлений, в котором можно пользоваться локальным описанием, аналогично тому, как это имеет место в гидродинамике и химической кинетике. Эта квадратичная форма имеет вид

$$\frac{d_i S}{dt} = \sum_k J_k X_k \quad (2.20)$$

В табл. 2.1 суммированы потоки и силы, связанные с некоторыми часто наблюдаемыми необратимыми явлениями.

Таблица 2.1. Термодинамические потоки и силы в некоторых часто наблюдаемых необратимых процессах^a

| Явление | Поток | Сила | Векторный характер |
|--------------------|---|---|--------------------|
| Теплоперенос | Тепловой поток, J_{th} | $\nabla(1/T)$ | Вектор |
| Диффузия | Поток массы компонента i , $J_{d,i}$ | $-\nabla(\mu_i/T) - F_i$ | Вектор |
| Вязкое течение | Диссипативная часть тензора давления, P | $\nabla v(1/T)$ | Тензор 2-го ранга |
| Химическая реакция | Скорость реакции ρ , ω_ρ | Средство реакции деленное на T , A_ρ/T | Скаляр |

^a T — температура, μ_i — химический потенциал компонента i , F_i — внешняя сила, действующая на единицу массы компонента i , v — гидродинамическая скорость. Средство A_ρ связано с μ_i соотношением $A_\rho = -\sum_i \nu_{i\rho} \mu_i$, где стехиометрические коэффициенты $\nu_{i\rho}$ дают полное число молекул, образующихся ($\nu > 0$) или исчезающих ($\nu < 0$) в реакции.

В общем случае потоки J_k являются очень сложными функциями сил. Особенно простая ситуация возникает в случае, когда эта связь линейна:

$$J_k = \sum_l L_{kl} X_l \quad (2.21)$$

В этой формуле посредством L_{kl} обозначен набор *феноменологических коэффициентов*. Такая ситуация имеет место вблизи равновесия, где, кроме того, как показано Онзагером на основе соображений, связанных с принципом детального равновесия, выполняется дополнительное соотношение $L_{kl} = L_{lk}$. Заметим, однако, что в некоторых далеких от равновесия состояниях линейная зависимость вида (2.21) все-таки может иметь место. Это возможно либо благодаря случайности, либо из-за наличия регуляторных процессов специального типа. Так или иначе, путем подстановки вместо фигурирующих в эволюционных (динамических) уравнениях [см., например, (2.12)] потоков их выражений через силы и с учетом того, что сами силы линейно выражаются через переменные состояния, можно еще одним способом прийти к сформулированному уже в разд. 2.4 выводу о неэффективности возможных нелинейностей вблизи равновесия.

Вернемся теперь к общему виду выражения для скорости изменения энтропии (2.19а). Как уже отмечалось, согласно вто-

рому закону термодинамики, $d_i S \geq 0$. Наоборот, нет такого физического закона, который определял бы знак изменения энтропии $d_e S$ в неизолированной системе ($d_e S \neq 0$). Иными словами, в зависимости от рассматриваемой системы величина $d_e S$ может быть как положительной, так и отрицательной. Тогда понятно, что если $d_e S$ отрицательно и превышает по абсолютной величине $d_i S$, то определенные стадии эволюции могут происходить при общем понижении энтропии:

$$\frac{dS}{dt} < 0 \quad (\text{в неизолированных системах}) \quad (2.22)$$

Согласно традиционной интерпретации энтропии, это будет означать, что в ходе эволюции разупорядоченность уменьшается за счет оттока энтропии. Это замечание полезно иметь в виду при выяснении природы сложных явлений, рассматривавшихся в гл. 1 на термодинамической основе.

Вопрос, однако, не так прост. Как уже подчеркивалось в первой части данного раздела, связь между энтропией и упорядоченностью довольно прозрачна в идеальных системах (в отсутствие взаимодействия между частицами), а также в случае равновесия в системах более общего характера, когда характерный пространственный масштаб возможных упорядоченных структур сравним с масштабом межмолекулярных сил. С другой стороны, та упорядоченность, которая обусловлена возникновением диссипативных структур в неравновесных условиях, совсем не обязательно должна быть напрямую связана с убылью энтропии. По этой причине утверждение (2.22) следует рассматривать всего лишь как мотивировку поиска более четкого и недвусмысленного способа исследования сложного.

2.6. Устойчивость

Приведем еще одну точку зрения на содержание второго закона термодинамики применительно к изолированным системам. С одной стороны, доказано, что при любом начальном состоянии системы с течением времени в ней установится единственное состояние — равновесное. С другой стороны, такая тенденция подразумевает монотонное возрастание энтропии со временем, или, иными словами, монотонное возрастание разности $\Delta S = S - S_{\text{eq}}$ от отрицательных значений до нуля (рис. 2.7).

В действительности оба этих утверждения эквивалентны, и, пользуясь более математическим языком, они выражают тот факт, что равновесие является *глобальным асимптотически устойчивым* состоянием, а энтропия — *функцией Ляпунова*, отражающей эту исключительно сильную форму устойчивости. Мы не будем здесь более точно определять эти термины, по-

сколько эта задача является важной частью дальнейшего анализа (см. также разд. 1.3 и 1.4, где уже происходило обращение к асимптотической устойчивости на интуитивном уровне).

Рассмотрим некоторую систему, неважно — консервативную или диссипативную, эволюционирующую в соответствии с определенными динамическими законами, выражаемыми, например, уравнениями (2.12). Пусть за счет какого-то механизма, который здесь можно не уточнять, система оказалась в некотором определенном состоянии X_s , в котором она и остается впоследствии. В случае консервативной системы это состояние может быть состоянием механического равновесия, в то время как в диссипативной системе оно может быть либо состоянием термодинамического равновесия, либо стационарным неравновесным состоянием.

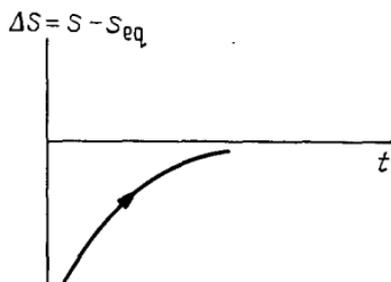


Рис. 2.7. Альтернативный взгляд на второй закон термодинамики для изолированной системы: избыток энтропии отрицателен ($\Delta S < 0$) и монотонно возрастает со временем, тем самым играя роль функции Ляпунова. Равновесное состояние асимптотически устойчиво и по существу является глобальным аттрактором.

На самом деле состояние реальных систем никогда не остается постоянным. Начнем с того, что большинство систем контактирует со сложным и даже непредсказуемым окружением. Это окружение непрерывно сообщает системе небольшие (реже значительные) количества вещества, импульса или энергии, что делает невозможным контроль любой из переменных состояния со сколь угодно высокой степенью точности. Любой экспериментатор в той или иной мере знаком с этими трудностями, и освященные веками выражения «экспериментальная погрешность» или «доверительный интервал» являются просто отражением вмешательства внешней среды во внутреннюю динамику системы. Этот вездесущий факт выражается в том, что мгновенное состояние $X(t)$ системы не совпадает со стационарным состоянием X_s — оно представляет собой некоторое соседнее состояние:

$$X(t) = X_s + x(t) \quad (2.23)$$

Величина $x(t)$ называется *возмущением*.

Имеется другая, более фундаментальная точка зрения на ту же проблему. Большинство систем, с которыми мы имеем дело, состоит из огромного числа взаимодействующих объектов. Для таких систем переменные состояния, которые упоминались на протяжении этой главы, следует понимать в статистическом смысле. Это означает, что такие переменные представляют собой либо средние значения по мгновенным состояниям на длинном временном интервале, либо, возможно, наиболее вероятные значения, которые могут приниматься этими переменными. Поэтому в действительности регистрация мгновенного состояния показала бы наличие отклонений от таких «стандартных» значений макроскопических переменных состояния. Такие отклонения генерируются системой спонтанно независимо от внешней среды. Эти внутренние отклонения мы будем называть *флуктуациями*. Их влияние по-прежнему можно выразить равенством (2.23) при том понимании, что $x(t)$ теперь обусловлено внутренней динамикой рассматриваемой системы.

Обычно флуктуация возникает в виде локализованного, маломасштабного события. Например, в некоторой малой части макроскопической системы определенные частицы могут иметь большую тепловую скорость по сравнению с их соседями. Потом плотность частиц в каком-то другом малом объеме может оказаться пониженной и т. д. Легко понять, что такие события являются случайными. Это свойство отражается терминами типа «фоновый шум», часто используемыми для обозначения флуктуаций.

Мы объяснили причины, по которым состояние системы будет непрерывно отклоняться от *стандартного состояния* X_s . Обсудим теперь, каким образом система отвечает на такие отклонения. Полный список всех возможных вариантов представлен в виде следующих случаев.

Случай 1. Состояние $X(t)$ остается в некоторой окрестности стандартного состояния X_s на всех временах, превышающих некоторое начальное время t_0 . Точнее, при любом наперед заданном $\epsilon > 0$ всегда найдется такое δ , зависящее в общем случае от ϵ и t_0 , что при начальном возмущении, не превышающем δ , отклонение $x(t)$ остается меньше ϵ на всех временах, превышающих t_0 . Мы будем тогда говорить, что состояние X_s *устойчиво по Ляпунову*. Иллюстрация устойчивости по Ляпунову уже приводилась слева в нижней части рис. 1.8. Это понятие иллюстрируется также тремя альтернативными способами на рис. 2.8.

Во многих случаях нас интересует не ответ какого-то данного состояния на возмущения, но скорее ответ целой последовательности состояний, определяющей некоторую траекторию. В этом случае вводится подходящая мера расстояния между стандартной и возмущенной траекториями, после чего исполь-

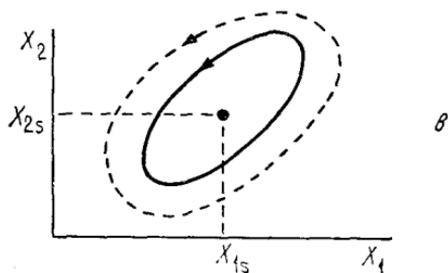
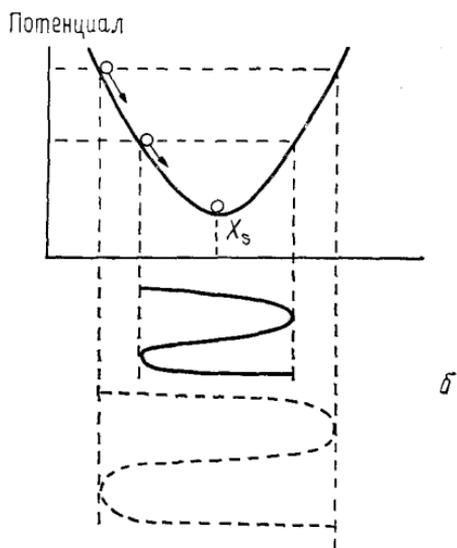
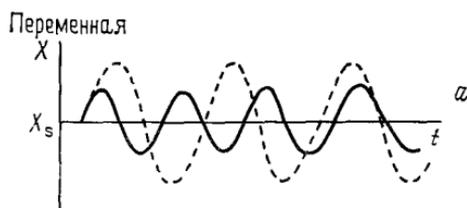


Рис. 2.8. Три взгляда на устойчивость по Ляпунову. *a* — отклонения от стандартного состояния X_s остаются ограниченными на всех временах; *b* — механический аналог случая *a*, когда движение материальной точки в потенциальной яме представлено для двух различных начальных условий; *b* — орбиты в фазовом пространстве, соответствующие двум различным возмущениям стандартного состояния (X_{1s} , X_{2s}). Последний рисунок иллюстрирует также понятие орбитальной устойчивости.

зуется упомянутое выше определение устойчивости. Это приводит к понятию *орбитальной устойчивости*.

Случай 2. Состояние $X(t)$ стремится к X_s по мере того, как время стремится к бесконечности, или, иными словами, возмущение $x(t)$ со временем стремится к нулю. Тогда мы говорим, что состояние X_s является *асимптотически устойчивым*. Асимптотическая устойчивость иллюстрируется справа в нижней части рис. 1.8; некоторые дополнительные аспекты суммированы на рис. 2.9. Это определение можно обобщить на *асимптотическую орбитальную устойчивость*. Ясно, что асимптотическая устойчивость обязательно подразумевает необратимость, так что это понятие неприменимо к консервативным системам. Такие системы в лучшем случае могут быть устойчивыми по Ляпунову. С другой стороны, диссипативные системы могут устранять

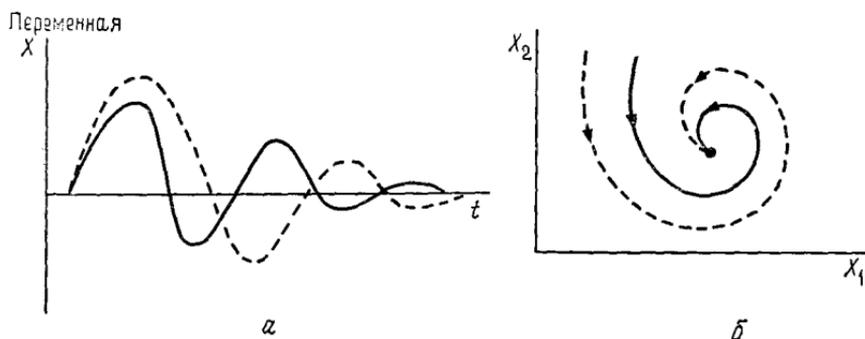


Рис. 2.9. Два представления об асимптотической устойчивости. *а* — временная эволюция переменной состояния X ; *б* — фазовые траектории, сходящиеся к аттрактору.

влияние возмущений, действующих на них, и тем самым восстанавливать стандартное состояние. Это обеспечивает предсказуемость и воспроизводимость такого режима, который в дальнейшем будем называть *аттрактором*. Например, ячейки Бенара, колебания в реакции БЖ или состояние термодинамического равновесия в изолированной системе являются аттракторами соответствующим образом определенных динамических систем. Асимптотическая устойчивость, несомненно, является одним из наиболее ярких проявлений конструктивной роли необратимости в природе.

Случай 3. Состояние $X(t)$ не остается в окрестности X_s . Точнее, для каждой окрестности X_s существует такое начальное возмущение, для которого амплитуда $x(t)$ не может оставаться все время меньше некоторой заранее, но произвольным образом заданной величины t_0 . Тогда мы говорим, что стандартное состояние X_s *неустойчиво*. Такая ситуация проявляется в наличии

начальной стадии быстрого (обычно экспоненциального) роста возмущения. Все эти свойства легко обобщить и дать определение *орбитальной неустойчивости*.

Одну из разновидностей неустойчивости можно представить с помощью графика, аналогичного рис. 2.9, б, с тем отличием, что стрелки на различных кривых должны поменять направление на противоположное. Другой вид неустойчивости изображен на рис. 2.10.

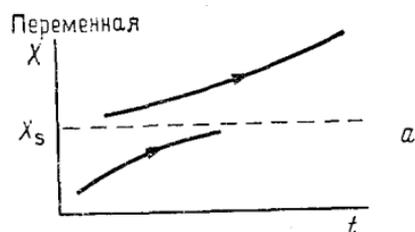
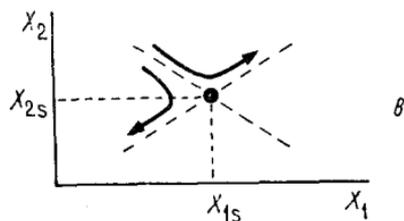


Рис. 2.10. Иллюстрация неустойчивости стандартного состояния. а — эволюция переменной состояния X во времени; б — механическая иллюстрация с помощью движения материальной точки, помещенной в начальный момент времени на вершину потенциального барьера; в — представление неустойчивости в фазовом пространстве.



Неустойчивые стандартные состояния могут встретиться как в консервативных, так и в диссипативных системах. Примерами такого рода служат верхнее положение равновесия маятника и состояние покоя в задаче Бенара выше порога возникновения конвекции. Однако, в то время как в случае диссипативной системы особых ограничений нет, состояние консервативной системы должно остаться неустойчивым и после обращения времени. Это замечание находится в согласии с рис. 2.10, а, в, но противоречит рис. 2.9, б с обращенными стрелками.

Случай 4. Состояние $X(t)$ остается в некоторой окрестности стандартного состояния X_s , если начальные возмущения не превышают некоторого порогового значения, и удаляется от него в случае возмущений, превышающих пороговое. Тогда мы говорим, что состояние X_s *локально устойчиво*, но *глобально неустойчиво*. С другой стороны, если устойчивость имеет место при любых по амплитуде начальных возмущениях, мы говорим о *глобальной устойчивости*, причем состояние X_s в таком случае является *глобальным аттрактором*. Например, термодинамическое равновесие в изолированной системе является глобальным аттрактором, однако состояние конвекции в задаче Бенара, для которого течение в определенной ячейке (например, в ячейке 1 на рис. 1.3) ориентировано по часовой стрелке, глобально неустойчиво по отношению к конечному возмущению,

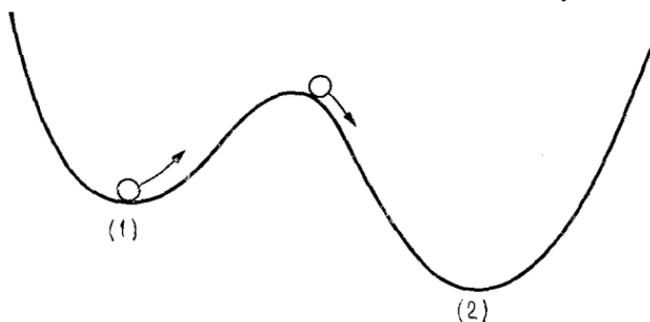


Рис. 2.11. Иллюстрация глобальной и локальной устойчивости. Шарик, начинающий движение во впадине (1), может остаться в ней или перескочить во впадину (2) в зависимости от того, мала его начальная кинетическая энергия или превосходит конечный пороговый уровень.

меняющему ориентацию течения в той же ячейке. Механический пример локальной устойчивости и глобальной неустойчивости представлен на рис. 2.11. Подвижное тело с малой кинетической энергией, начинающее движение (без трения) вблизи нижней точки впадины (1) остается вблизи этой точки, однако если тому же телу сообщить достаточно большую начальную кинетическую энергию, то оно преодолет возвышение и в дальнейшем будет двигаться по пути, охватывающему обе впадины, (1) и (2). Иными словами, равновесные положения во впадинах (1) и (2) являются устойчивыми лишь для таких возмущений, которые не выводят за возвышение.

В свете обсуждения устойчивости теперь легче понять концептуальный прогресс, достигнутый при переходе от изолированных систем, для которых справедливо соотношение (2.18), к открытым системам вдали от равновесия, для которых справедливы соотношения (2.19) — (2.20).

В изолированной системе равновесие глобально устойчиво, и это отражается в существовании некоторой функции — энтропии, или, точнее, избыточной энтропии $\Delta S = S - S_{\text{eq}}$, обладающей следующими свойствами (см. рис. 2.7):

$$\Delta S \leq 0, \quad \frac{d}{dt} \Delta S \geq 0 \quad (\text{в изолированных системах}) \quad (2.24)$$

Величину ΔS мы называем *функцией Ляпунова*. В математической литературе показано, что соотношения (2.24) равносильны глобальной устойчивости.

В неизолированной системе, не находящейся в состоянии равновесия, вследствие наличия двух членов в выражении для баланса энтропии ($d_e S$ и $d_i S$) второй закон термодинамики ($d_i S/dt \geq 0$) больше не определяет знак изменения энтропии (то же справедливо и для любой другой функции состояния). Таким образом, универсальной функции Ляпунова не существует, и это порождает проблему устойчивости состояний, далеких от равновесия. Возможность потери устойчивости при определенных условиях открывает путь переходным явлениям, что может привести к внутренней дифференциации и сложному поведению. Следует, однако, отметить одно исключение — открытые системы вблизи равновесия. Поскольку в этой области феноменологические соотношения линейны и кинетические коэффициенты симметричны [см. уравнение (2.21)], выражаемое равенством (2.20) производство энтропии становится положительно определенной квадратичной формой переменных состояния X_k , отражающих внешние ограничения. Более того, можно показать, что любое отклонение от стационарного состояния приводит к развитию таких процессов, что величина $d_i S/dt$ стремится уменьшиться со временем. Иными словами, ситуация здесь аналогична той, которая выражается соотношениями (2.24), причем в этой ситуации производство энтропии играет роль функции Ляпунова, гарантируя глобальную устойчивость стационарных состояний, близких к равновесному.

Необходимо отметить, что, несмотря на отсутствие общего закона, гарантирующего устойчивость состояний, далеких от равновесия, тем не менее можно найти некоторые полезные критерии, позволяющие сделать более слабые выводы типа, например, *достаточных условий* устойчивости. Один такой критерий дается условием положительности избыточного производства энтропии, вычисленного вблизи неравновесного стандартного состояния. Этот критерий подробно обсуждается в книге Гленсдорфа и Пригожина.

Следует предупредить читателя, что развитые брюссельской школой термодинамические критерии устойчивости часто неверно цитируются в литературе. Минимум производства энтропии

справедлив при специальных условиях, соответствующих ситуациям, близким к равновесию. Что касается критериев, содержащих избыточное производство энтропии, то, во-первых, не следует смешивать эту величину с самим производством энтропии, а, во-вторых, следует отдавать себе отчет в отсутствии как *необходимости*, так и *достаточности* одновременно. В дальнейшем постепенно выяснится, что сложность в природе невозможно свести к некоторому принципу глобальной оптимальности. Нам представляется, что в своей погоне за сложностью природа занимает более «прагматическую» позицию, в которой существенную роль играет поиск устойчивости.

2.7. Бифуркация и нарушение симметрии

Рассмотрим еще раз задачи типа бенаровской под несколько иным углом зрения. Мы уже отмечали важную роль нелинейностей и внешних ограничений, управляющих поведением системы

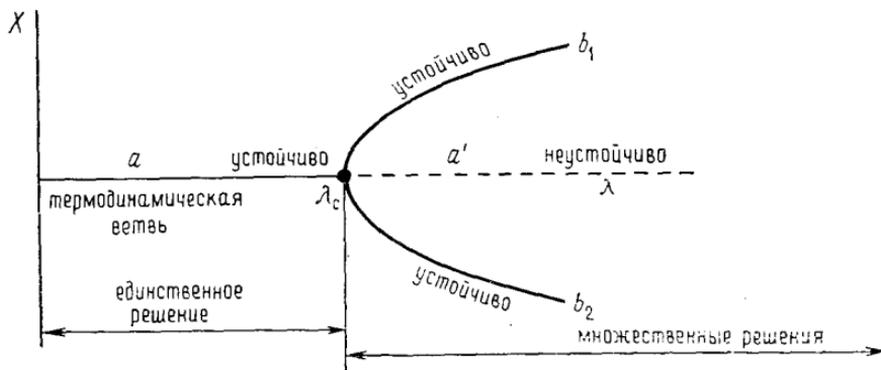


Рис. 2.12. Диаграмма бифуркации, показывающая влияние изменения управляющего параметра на переменную состояния X . Единственное решение (a), соответствующее термодинамической ветви, теряет устойчивость при λ_c . При этом значении управляющего параметра появляются новые ветви решения (b_1, b_2), которые в данном случае устойчивы.

вдали от равновесия. Нам известно также, что при изменении управляющих параметров в системе наблюдаются разнообразные переходные явления. Выделим теперь из этих наблюдений определенные общие черты, характерные для большого числа других переходов в физико-химических системах. С этой целью представим графически (рис. 2.12) зависимость вертикальной компоненты скорости течения жидкости в некоторой определенной точке от внешнего ограничения, или, в более общем виде, зависимость переменной состояния системы X (или $x = X - X_s$) от управляющего параметра λ . Таким образом мы получим график, известный под названием *бифуркационной диаграммы*.

При малых значениях λ возможно лишь одно решение, соответствующее состоянию покоя в бенаровском эксперименте. Оно представляет собой непосредственную экстраполяцию термодинамического равновесия и, подобно равновесию, характеризуется важным свойством — асимптотической устойчивостью, поскольку в этой области система способна гасить внутренние флуктуации или внешние возмущения. По этой причине такую ветвь состояний мы будем называть *термодинамической ветвью*. При переходе через критическое значение параметра λ , обозначенное λ_c на рис. 2.12, состояния на этой ветви становятся неустойчивыми, так как флуктуации или малые внешние возмущения уже не гасятся. Действуя подобно усилителю, система отклоняется от стандартного состояния и переходит к новому режиму, в случае бенаровского эксперимента соответствующему состоянию стационарной конвекции. Оба этих режима сливаются при $\lambda = \lambda_c$ и различаются при $\lambda > \lambda_c$. Это явление называется *бифуркацией*. Легко понять причины, по которым это явление следует ассоциировать с катастрофическими изменениями и конфликтами. В самом деле, в решающий момент перехода система должна совершить критический выбор (в окрестности $\lambda = \lambda_c$), что в задаче Бенара связано с возникновением право- или левовращательных ячеек в определенной области пространства (рис. 2.12, ветви b_1 или b_2). В устройстве экспериментальной установки нет ничего такого, что позволило бы наблюдателю определить заранее, какое состояние будет выбрано. Решать будет лишь случай через динамику флуктуаций. Просканировав «фон», система совершит несколько попыток, поначалу, возможно, безуспешных, и наконец, какая-то особенная флуктуация победит. Стабилизировав ее, система превращается в своеобразный исторический объект в том смысле, что его дальнейшая эволюция будет зависеть от этого критического выбора. Таким образом, нам удалось на абстрактном языке сформулировать замечательное взаимодействие между случаем и внешним ограничением, между флуктуациями и необратимостью, лежащее в основе большинства явлений, рассмотренных в гл. 1.

Можно отметить сходство между этими представлениями и понятиями мутации и отбора, известными из теории биологической эволюции. На самом деле можно сказать, что флуктуации являются физическим аналогом мутантов, в то время как поиск устойчивости играет роль естественного отбора. Даже сама диаграмма бифуркации по своему устройству напоминает филогенетические деревья, столь широко используемые в биологии!

Теперь достаточно совсем небольших мысленных усилий для того, чтобы убедиться в следующем. Если бы мы интересовались не бенаровским течением, а реакцией Белоусова — Жаботинского, или агрегацией *Dictyostelium discoideum*, или в дей-

ствительности любым другим переходным явлением, то, построив для этих систем бифуркационные диаграммы, мы увидели бы, что все они устроены качественно одинаково. Единственное различие состояло бы в терминологии, которая по необходимости является спецификой каждой рассматриваемой задачи. Таким образом, все такого рода явления имеют глубокое сходство. Поэтому можно утверждать, что переход к сложному тесно связан с появлением новых ветвей решений в результате бифуркации, происходящей вследствие *потери устойчивости* стандартного состояния, вызванной нелинейностями и внешними ограничениями в открытой системе.

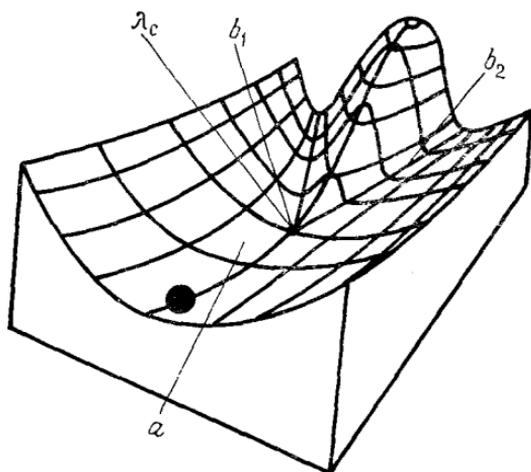


Рис. 2.13. Механическая иллюстрация явления бифуркации.

Механическая иллюстрация этого явления представлена на рис. 2.13. Шарик движется по впадине (рис. 2.12, ветвь *a*), которая в некоторой конкретной точке λ_c разветвляется и образует две новые впадины (ветви b_1 и b_2), разделенные возвышением. Хотя на этой стадии еще слишком рано приводить какие-то аналогии или экстраполяции, довольно заманчиво на мгновение вообразить, что вместо шарика на рис. 2.13 мог оказаться динозавр конца мезозойской эры или группа наших предков, собирающихся принять идеографическую или символическую письменность!

Мы неоднократно подчеркивали, что источником инноваций и диверсификации является бифуркация, поскольку именно благодаря ей в системе появляются новые решения. Но каковы же характеристики этих решений? Основную часть ответа на этот вопрос мы отложим до гл. 3. В заключение же настоящего раздела мы отметим тот особый момент, который уже возникал

при обсуждении экспериментальных фактов в гл. 1: появляющиеся в результате бифуркаций решения характеризуются *нарушением симметрии*.

Путем наблюдений установлено, что определенные физические явления допускают совершенно одинаковое описание даже в том случае, когда экспериментальные наблюдения проводятся в различных условиях. Например, как мы видели в разд. 1.3 и 1.8, ниже порога бенаровской конвекции сосуд с жидкостью имеет идентичные свойства для всех наблюдателей, смещенных друг относительно друга вдоль определенного направления (трансляционная инвариантность). Другой известный пример из механики: ускорение свободно падающего вблизи поверхности Земли тела направлено по вертикали не только для наблюдателя, зафиксированного на Земле, но и для всех наблюдателей, движущихся по отношению к нему с постоянной скоростью (галилеевская инвариантность). Важность этих фактов состоит в том, что они налагают определенные ограничения, которые должны учитываться любой разумной математической моделью. В частности, законы, описывающие эти явления, должны характеризоваться *определенной симметрией*, допускающей соответствующую *инвариантность* этих явлений. Эти виды симметрии должны сохраняться при всех пространственных и временных операциях, меняющих места наблюдателей, для которых данный вид инвариантности возможен. Довольно часто эти операции образуют некоторый математический объект, называемый группой, поскольку последовательное выполнение нескольких таких операций эквивалентно некоторой новой операции, что уже само по себе сходно с инвариантностью. Так, динамические уравнения, описывающие жидкость, нагреваемую снизу и содержащуюся между двумя пластинами, бесконечно протяженными в горизонтальном направлении, инвариантны относительно группы трансляций вдоль любой горизонтальной плоскости. Аналогичным образом уравнения химической кинетики в бесконечно протяженной среде инвариантны относительно групп трансляций, вращений и отражений.

И все же, несмотря на исключительную симметрию этих динамических законов, экспериментально наблюдается возникновение таких состояний, как ячейка Бенара или химическая волна, причем каждое из этих состояний менее симметрично, нежели сами законы. Этот переход с *нарушением симметрии* возник благодаря бифуркации. Что же конкретно достигнуто в результате такого перехода?

Начнем с того, что нарушение симметрии является проявлением *внутренней дифференциации* между различными частями системы или между системой и ее окружением. В этом плане нарушение симметрии воплощает одну из первейших пред-

посылок сложного поведения, которая должна была реализоваться во время таких событий, как сгущение первичной материи при образовании галактик или возникновение первых живых клеток. Кроме того, как только такая дифференциация обеспечена, могут включаться другие процессы, невозможные в недифференцированной среде. Например, как подчеркивалось в разд. 1.6, возникновение пространственных неоднородностей в биологической среде может дать недифференцированным клеткам в популяции возможность распознавать окружающую их среду и дифференцироваться в специализированные клетки. В свою очередь это может позволить генетическому материалу проявить свои потенциальные возможности. Следовательно, в этом плане нарушение симметрии оказывается также предпосылкой *информации*.

И все же не следует забывать, что между предпосылкой и достаточным условием довольно большое расстояние. В самом деле, переходы с нарушением симметрии имели бы большое значение лишь в том случае, если бы они могли привести к асимптотически устойчивым решениям. Помимо прочего, необходимо иметь возможность тем или иным способом *отбирать* конкретную асимметричную форму материи из всего множества решений, становящихся одновременно возможными выше точки бифуркации. Дальнейшим обсуждением этих вопросов мы займемся в последующих главах.

2.8. Упорядоченность и корреляции

Как уже неоднократно отмечалось, термодинамическое равновесие — это состояние с микроскопическими характерными пространственными масштабами, а возникающие вдали от равновесия переходные явления связаны с состояниями, характеризующимися корреляциями в макроскопическом диапазоне. И если первое представляется вполне естественным с точки зрения короткодействующих межмолекулярных сил, то второе, наоборот, весьма поражает с той же точки зрения. Этот кажущийся парадокс еще больше усиливается, если учесть, что неравновесные переходы, когерентно охватывающие макроскопические области пространства, могут иметь место даже в идеальных системах, в которых можно полностью пренебречь межмолекулярными взаимодействиями!

Для того чтобы хотя бы интуитивно разобраться, в чем здесь дело, придется воспользоваться флуктуацией в качестве зонда для исследования состояния вещества. Представим себе вначале некоторую идеальную систему (например, газ) в равновесии. Пусть V и N — полный объем системы и полное число частиц в ней соответственно, а ΔV и n — объем некоторой малой под-

системы и число содержащихся в ней частиц. Поскольку система однородна, вероятность обнаружения любой частицы в выделенной подсистеме равна $\Delta V/V$, а вероятность обнаружения в той же подсистеме одновременно n частиц равна соответственно $(\Delta V/V)^n$. Аналогично вероятность того, что некоторая частица не содержится в выделенной подсистеме, равна $(V - \Delta V)/V$, а вероятность того же для m частиц равна $[(V - \Delta V)/V]^m$. Таким образом, вероятность того, что в объеме ΔV содержится n частиц, равна P_n :

$$P_n = (\text{число способов выбора } n \text{ частиц из } N) \times \left(\frac{\Delta V}{V}\right)^n \left(1 - \frac{\Delta V}{V}\right)^{N-n} \quad (2.25)$$

В интересующем нас случае ΔV много меньше, чем V , и n также мало по сравнению с N . По этой причине комбинаторный множитель в (2.25) можно приближенно представить в виде $N^n/n!$, где $n! = 1 \cdot 2 \cdot \dots \cdot n$.

Пренебрегая также величиной n по сравнению с N в показателе степени в (2.25), перепишем это равенство в виде

$$P_n = \frac{1}{n!} \left(\frac{N\Delta V}{V}\right)^n \left(1 - \frac{\Delta V}{V}\right)^N$$

Однако величина $N\Delta V/V$ представляет собой не что иное, как среднее число частиц в объеме ΔV , обозначаемое \bar{n} . Следовательно,

$$P_n = \frac{\bar{n}^n}{n!} \left(1 - \frac{\bar{n}}{N}\right)^N$$

Наконец, с учетом того, что N настолько велико, что можно воспользоваться формулой

$$\lim_{N \rightarrow \infty} \left(1 - \frac{x}{N}\right)^N = e^{-x}$$

мы получаем

$$P_n = \frac{\bar{n}^n}{n!} e^{-\bar{n}} \quad (2.26)$$

Это знаменитое *распределение Пуассона*, играющее важную роль в теории вероятностей. Из рис. 2.14 видно, что такое распределение имеет острый максимум вблизи \bar{n} . Чтобы охарактеризовать ширину соответствующего пика, введем *моменты* числа частиц, определенные следующим образом:

$$\langle n^k \rangle = \sum_{n=0}^{\infty} n^k P_n$$

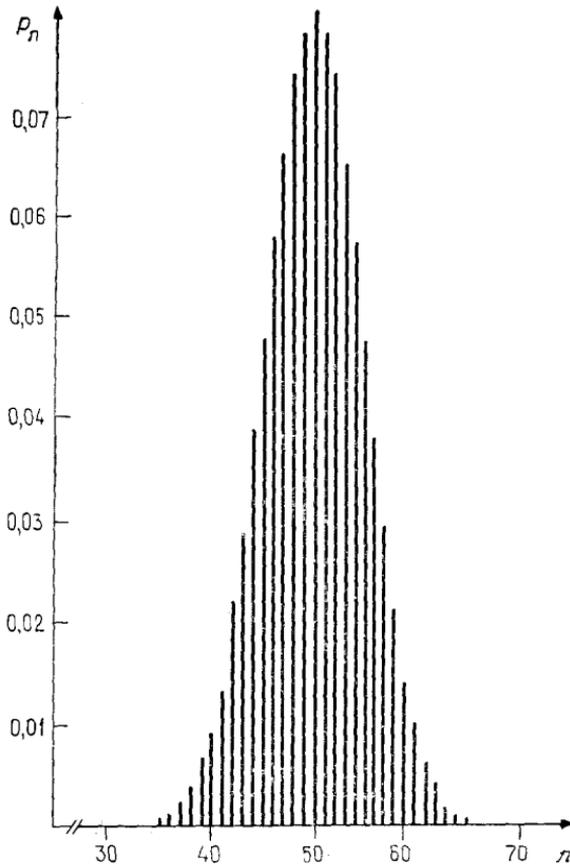


Рис. 2.14. Пуассоновское распределение вероятностей в зависимости от n , построенное для $\bar{n}=50$. Заметим, что это распределение практически симметрично относительно своего острого максимума при $n=\bar{n}$.

В частности, при $k=0$ мы видим, что $\sum_{n=0}^{\infty} P_n = 1$ в соответствии со смыслом вероятности (условие нормировки), а при $k=1$, как и ожидалось, имеем $\langle n \rangle = \bar{n}$. Если бы функции P_n были совершенно не размыты вблизи \bar{n} , то все последующие моменты выразились бы в виде степеней \bar{n} . Напротив, если эти моменты не будут выражаться в виде степеней \bar{n} , то это может служить указанием на важную роль конечной ширины распределения, а также флуктуаций. Из равенства (2.26) имеем

$$\langle n^2 \rangle = \bar{n}^2 + \bar{n}$$

Следовательно, вводя дисперсию P_n с помощью соотношения

$$\langle \delta n^2 \rangle = \langle n^2 \rangle - \langle n \rangle^2 \quad (2.27)$$

мы получаем, что для пуассоновского распределения имеет место

$$\langle \delta n^2 \rangle_p = \bar{n} = \langle n \rangle \quad (2.28)$$

Кроме того, из вывода (2.26) ясно, что вероятность обнаружения n_1 частиц в подсистеме с объемом ΔV_1 и одновременно n_2 частиц в подсистеме с объемом ΔV_2 (вторая подсистема не должна пересекаться с первой) равна произведению отдельных вероятностей P_n :

$$P_{n_1 n_2} = P_{n_1} P_{n_2} \quad (2.29)$$

Таким образом, вычисляя второй момент $\langle n_1 n_2 \rangle$ по вероятностям $P_{n_1 n_2}$, находим

$$\langle n_1 n_2 \rangle = \langle n_1 \rangle \langle n_2 \rangle = \bar{n}_1 \bar{n}_2$$

Вводя ковариацию распределения $P_{n_1 n_2}$

$$\langle \delta n_1 \delta n_2 \rangle = \langle n_1 n_2 \rangle - \langle n_1 \rangle \langle n_2 \rangle \quad (2.30)$$

получаем

$$\langle \delta n_1 \delta n_2 \rangle_p = 0 \quad (2.31)$$

Итак, мы получили следующие две важные статистические характеристики:

1. Дисперсию, служащую мерой интенсивности флуктуаций относительно среднего значения.

2. Ковариацию, характеризующую степень коррелированности поведения различных пространственных областей системы.

Обращение в нуль ковариации, имеющее место в случае пуассоновского распределения при равновесии в идеальной системе, означает полное отсутствие пространственных корреляций. Поэтому такое состояние может служить в качестве прототипа разупорядоченности. Это, однако, не означает, что флуктуации велики. Напротив, как следует из (2.28), величина $\sqrt{\langle \delta n^2 \rangle_p / \bar{n}^2}$, которую можно рассматривать как относительную характеристику величины флуктуаций, равна $1/\sqrt{\bar{n}}$ и тем самым является малой величиной, если ΔV велико по сравнению с молекулярными размерами. Например, если ΔV равно 1 см^3 , то $1/\sqrt{\bar{n}}$ составляет величину порядка $10^{-8} \%$ в случае газа при комнатной температуре и атмосферном давлении. Однако в объеме порядка 1 мкм флуктуации будут достигать значений порядка сотых долей процента.

С другой стороны, если возникает такое состояние вещества, которое характеризуется макроскопическими пространственным и временным масштабами, то это означает, что в системе создаются и поддерживаются воспроизводимые взаимоотношения между удаленными частями. Это будет проявляться в конечной амплитуде и макроскопической длине корреляции пространственной корреляционной функции — ковариации $\langle \delta n(\mathbf{r}_1) \delta n(\mathbf{r}_2) \rangle$. Это означает, что такая функция будет заметно отличаться от нуля даже для макроскопических расстояний между подсистемами $|\mathbf{r}_2 - \mathbf{r}_1|$. Обе эти особенности корреляционной функции свидетельствуют о том, что система смогла преодолеть пуассоновскую «тиранию». В гл. 4 мы обсудим вопрос о том, каким образом это становится возможным даже в тех случаях, когда межмолекулярные взаимодействия играют лишь незначительную роль.

В заключение этой главы подчеркнем, что, хотя мы еще и не пришли к четкому определению сложного, мы начинаем постигать некоторые его существенные ингредиенты: возникновение бифуркационных переходов вдали от равновесия и при наличии подходящих нелинейностей; нарушение симметрии выше точки бифуркации, а также образование и поддержка корреляций макроскопического масштаба. В следующих двух главах эти понятия постепенно приобретут более количественный смысл, в результате чего представления о сложном станут более четкими.

ДИНАМИЧЕСКИЕ СИСТЕМЫ И СЛОЖНОСТЬ

Сложность — это неотъемлемая черта мира динамических систем. Анри Пуанкаре, великий математический гений конца прошлого — начала нынешнего веков, разработал современную теорию динамических систем, цель которой — исследовать типы поведения систем, описываемых взаимосвязанными нелинейными уравнениями. В данной главе мы изложим некоторые аспекты этой теории, имеющие отношение к проблеме сложного, обращая особое внимание на минимальные условия возникновения сложного поведения, а также на основные особенности решений соответствующих уравнений.

3.1. Геометрия фазового пространства

В разд. 2.1 и 2.2 мы уже говорили о двух основных семействах динамических систем, встречающихся в природе, а именно о *консервативных* и *диссипативных системах*. Там же мы рассматривали, каким образом можно представить эволюцию этих систем во времени с помощью представления о движении изображающей точки в *фазовом пространстве*. В настоящем разделе мы продолжим анализ этих представлений, изучая возможные типы движения в фазовом пространстве. Поскольку между этими типами движения и поведением исходных систем имеется взаимно однозначное соответствие, достигается первая качественная классификация явлений, которые могут наблюдаться в динамических системах.

Начнем с рассмотрения динамических систем с конечным числом переменных. Тем самым на какое-то время мы исключим из рассмотрения пространственно распределенные системы, характеризующиеся непрерывной зависимостью переменных от пространственных координат. Такие системы будут рассмотрены в последних двух разделах настоящей главы. Динамические уравнения для систем с конечным числом степеней свободы имеют вид [ср. с (2.12)]

$$\frac{dX_i}{dt} = F_i(\{X_j\}, \lambda), \quad i = 1, \dots, n \quad (3.1)$$

При этом подразумевается, что в операторе F_i нет явной зависимости от пространственной координаты. Более того, мы ограничимся случаем автономных систем, для которых F_i не содержит явной зависимости от времени, вследствие чего траектории в фазовом пространстве инвариантны. Заметим, что в гамильтоновой системе уравнений n должно быть четным, а выражения для F_i должны иметь специальный вид в соответствии с (2.56).

Первый тип траектории в фазовом пространстве, совместимой с (3.1), соответствует нулевым скоростям:

$$\frac{dX_1}{dt} = \dots = \frac{dX_n}{dt} = 0 \quad (3.2)$$

Как частный случай к этому типу относится состояние механического равновесия, с которым мы встречались в случае консервативных систем, и стационарные состояния диссипативных систем. Такие траектории в фазовом пространстве являются вырожденными, поскольку они получаются в результате решения n алгебраических уравнений $F_i=0$, содержащих ровно n неизвестных. Таким образом, эти решения представляются точками (см., например, точку P на рис. 3.1), которые будут называться *неподвижными точками*.

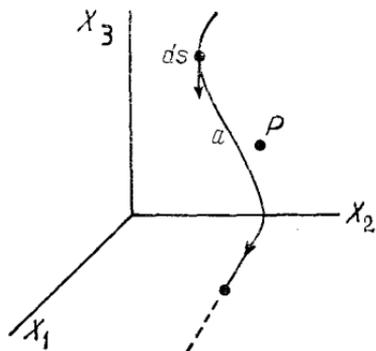


Рис. 3.1. Описание динамической системы в фазовом пространстве. P — неподвижная точка; a — фазовая траектория.

Если уравнения (3.2) выполняются, то изображающая точка будет двигаться вдоль некоторой траектории в фазовом пространстве (рис. 3.1, кривая a). Элемент пути вдоль этой траектории, соответствующий смещениям по отдельным осям (dX_1, \dots, dX_n), определяется соотношением

$$ds = \sqrt{\sum_k dX_k^2} = \sqrt{\sum_k F_k^2} dt$$

Проекция этого элемента на каждую из осей определены с помощью соотношения

$$\frac{dX_\alpha}{ds} = F_\alpha / \sqrt{\sum_k F_k^2}$$

Это соотношение определено везде, где $\sum_k F_k^2 \neq 0$, и такие точки называются *регулярными*. Напротив, в неподвижных точках ка-

сательная к траектории не определена в силу одновременного обращения в нуль всех F_i . По этой причине неподвижные точки можно также называть *особыми точками* решения системы уравнений (3.1). Множество неподвижных точек и траектории в фазовом пространстве образуют *фазовый портрет* динамической системы.

Решающую роль в определении структуры фазового портрета играет теорема единственности решений системы обыкновенных дифференциальных уравнений. Этот важный результат, восходящий к работам французского математика Коши, состоит в доказательстве того, что при довольно мягких условиях на функции F_i существует единственное в любой момент времени из интервала $(0, T_m)$ решение при условии, что в начальном состоянии система не находится в неподвижной точке. При этом верхняя граница временного интервала T_m зависит от специфики функций F_i . Если говорить о представлении системы в фазовом пространстве, то эта теорема автоматически исключает пересечение двух траекторий в любой регулярной точке.

Другая важная особенность фазового портрета — существование и структура *инвариантных множеств*. Этим термином обозначаются объекты, встроенные в фазовое пространство, которые, будучи ограничены, отображаются сами на себя в ходе эволюции системы, описываемой уравнениями (3.1). Тривиальным примером инвариантного множества является совокупность неподвижных точек. Другой пример — замкнутая кривая в фазовом пространстве типа кривых a и a' на рис. 2.2, представляющих движение маятника. Находясь на такой траектории, система вновь и вновь проходит через одни и те же состояния. Иными словами, замкнутая траектория в фазовом пространстве соответствует периодическому движению.

Проиллюстрируем теперь эти две особенности на примере динамических систем в двумерном фазовом пространстве. Допустим, что в части рассматриваемого фазового пространства имеются лишь два инвариантных множества — неподвижная точка и замкнутая кривая, окружающая эту точку (рис. 3.2, точка P и кривая a). Заметим, что тем самым мы автоматически ограничиваемся диссипативными системами. По теореме единственности траектория, начинающаяся внутри кривой a , останется там навсегда, так как в противном случае она пересекла бы кривую a . С другой стороны, самопересечения такой траектории запрещены той же теоремой. Тогда единственно возможными выходами остаются:

1. Движение по кривой, стремящейся к неподвижной точке P (рис. 3.2, а).
2. Движение по кривой, стремящейся к замкнутой кривой a (рис. 3.2, б).

Две последние возможности исключены в случае консервативных систем, поскольку эти варианты нарушают симметрию при обращении времени.

Мы видим, что невозможность самопересечений на траекториях и существование инвариантного множества в значительной мере определяют структуру фазового портрета в двумерных пространствах, а тем самым и возможные типы поведения системы. Для трехмерных и более высокого порядка систем налагаемые этими особенностями ограничения оказываются значительно более слабыми, поскольку траектории в многомерных пространствах имеют гораздо больше возможностей избегать друг друга, проскальзывая в зазорах, остаются между инвариантными множествами. Как будет показано в дальнейшем, благодаря этой дополнительной гибкости становятся возможными

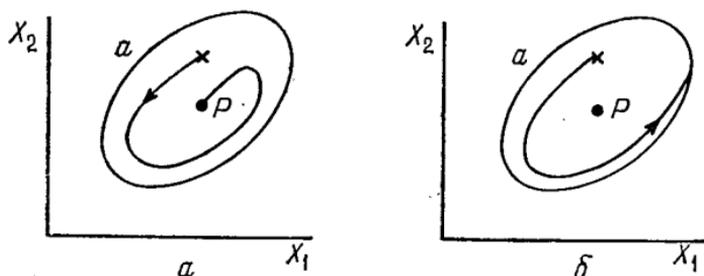


Рис. 3.2. Роль инвариантных множеств и непересечения траекторий в структуре фазового портрета. Предполагается, что единственным инвариантным множеством являются замкнутая траектория (a) и неподвижная точка P . Отсюда следует, что фазовая траектория, выходящая из точки внутри кривой a , обязана соответствовать одному из двух типов движения, представленных на рис. 3.2, a и b .

ми новые типы динамического поведения, из числа которых далеко не на последнем месте стоит возникновение хаотической, турбулентной динамики.

3.2. Меры в фазовом пространстве

В соответствии с изложенными соображениями эволюцию динамической системы можно рассматривать как преобразование, отображающее фазовое пространство само на себя. В процессе этого преобразования траектории, выходящие из определенной части фазового пространства, могут заходить в соседние области, достигать удаленных частей фазового пространства и, возможно, оставаться там навсегда или же возвращаться в исходную часть фазового пространства. Как правило, при этом различные части фазового пространства будут посещаться с различными частотами. Возникает вопрос: как может выглядеть

наиболее компактное и информативное описание этого сложного движения?

Одна из возможностей состоит в том, чтобы приписать определенный «вес» каждой части фазового пространства и посмотреть дальше, как этот вес меняется в процессе эволюции системы. В качестве такого веса представляется естественным взять объем рассматриваемой части фазового пространства. Чтобы увидеть, как этот вес изменяется в процессе эволюции системы, необходимо рассмотреть все траектории, начинающиеся в нулевой момент времени в данной части фазового пространства, и проследить их вплоть до момента времени t , а затем рассмотреть множество всех изображающих точек в момент t (рис. 3.3).

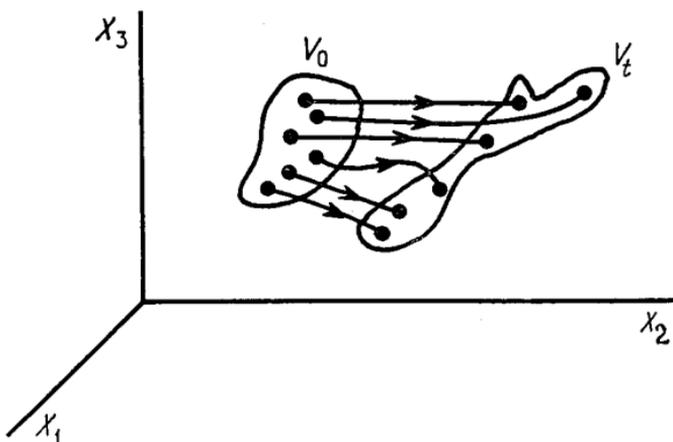


Рис. 3.3. Изменение в фазовом пространстве начального объема V_0 , обусловленное движением входящих в этот объем изображающих точек в соответствии с динамическими законами.

Таким образом, если в начальный момент времени изображающие точки занимали объем V_0 , то в момент t они займут объем V_t . Три различных характерных случая, приводящих к разным типам изменения V_t относительно V_0 , показаны на рис. 3.4.

Простейший случай равномерного движения по окружности показан на рис. 3.4, а. Уравнения движения в этом случае получаются из гамильтониана $H = \frac{1}{2}(X_1^2 + X_2^2)$. При этом фазовый объем не только сохраняется (роль объема здесь играет площадь), $|V_0| = |V_t|$, но также и не подвергается какой-либо деформации вследствие изохронности движения по каждой из траекторий.

Случай б соответствует движению по гиперболе, так что асимптотами траекторий являются оси X_1 и X_2 . Соответствующий гамильтониан имеет вид $H = X_1 X_2$. При этом проекция дви-

жения на ось X_1 приводит к экспоненциальному росту, а проекция на ось X_2 — к экспоненциальному затуханию с тем же абсолютным значением характерного времени. Благодаря этому фазовый объем опять сохраняется, $|V_t| = |V_0|$, однако в отличие от предыдущего случая он сильно деформируется. Например, прямоугольник, перемещающийся в фазовом пространстве вследствие движения соответствующих изображающих точек, будет удлиняться вдоль оси X_1 и сокращаться вдоль другой оси. Эта особенность характерна для класса *неустойчивых движений*, который будет изучаться подробнее позднее в настоящей главе и в гл. 5.

Случай *в* также описывает неустойчивое движение. Однако, поскольку имеется растяжение по обеим осям (X_1 и X_2), фазовый объем не только деформируется, но и систематически воз-

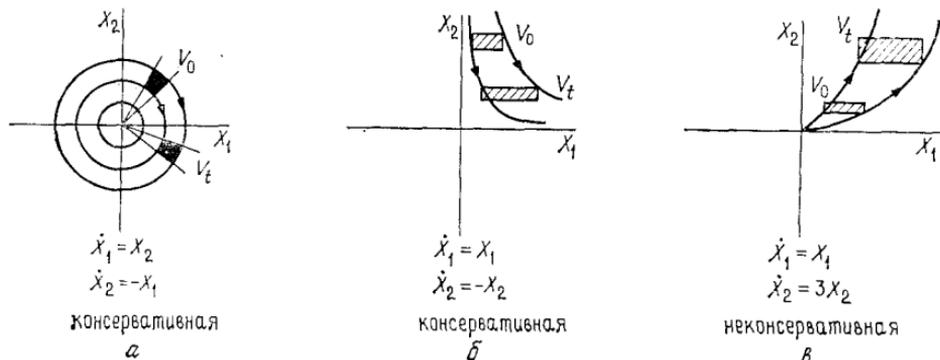


Рис. 3.4. Разные виды изменений, претерпеваемых начальным объемом V_0 для трех характерных примеров динамических систем.

растает. Если бы знаки в правых частях уравнений эволюции изменились на противоположные, то траектории остались бы те же, а направление движения поменялось бы на противоположное. В этом случае движение устойчиво, и фазовый объем будет также деформироваться, однако в отличие от прямого движения объем будет уменьшаться.

В то время как для неконсервативных систем нет общего правила изменения V_t , для консервативных систем, как оказывается, имевший место для случаев на рис. 3.4, а, б результат $|V_t| = |V_0|$ не случаен. В самом деле, согласно чрезвычайно важной теореме французского математика Лиувилля, фазовый объем является инвариантной величиной в *любых* консервативных системах.

Прежде чем обсудить смысл этого результата, рассмотрим вопрос о деформации фазового объема с несколько иной точки зрения. В принципе решение динамических уравнений (3.1)

однозначно в том смысле, что полное задание состояния системы (X_1, \dots, X_n) в любой момент времени позволяет предсказывать ее состояние во все последующие времена. Однако во многих случаях такая полная спецификация состояния системы не имеет смысла. Например, в случае гамильтоновой системы, число частиц в которой порядка числа Авогадро, не существует такого физического эксперимента, который позволил бы определить одновременно координаты и скорости всех частиц. Еще менее осмыслен такой подход, когда движение в фазовом пространстве становится чрезвычайно сложным, как в случае хаотического режима (см. также разд. 3.7—3.9). Здесь рассуждать на языке отдельных траекторий вообще бессмысленно. По этой причине необходимы новые способы описания, из которых одним из наиболее распространенных является вероятностный подход. С целью перехода к такому описанию вместо довольно специального случая отдельной системы мы рассмотрим некоторый ансамбль очень большого числа идентичных систем, находящихся в одних и тех же макроскопических условиях. Элементы этого ансамбля, называемого в честь американского ученого Дж. У. Гиббса, впервые введенного в науку это понятие, *ансамблем Гиббса*, в общем случае будут находиться в различных состояниях. Поэтому можно считать, что в начальный момент времени они принадлежали ансамблю, изображающие точки которого содержались в определенном объеме фазового пространства V_0 (рис. 3.3). В таком определении не содержится какого-либо ограничения числа систем в ансамбле, так что удобно рассматривать предельный случай, в котором любая, сколь угодно малая область содержит большое число систем. Основной вопрос, возникающий при таком подходе, состоит в следующем: какая доля элементов ансамбля характеризуется значениями переменных в интервалах $\{X_i, X_i + dX_i\}$? В предельном случае непрерывного описания ответ на этот вопрос выражается через *плотность вероятности* $\rho(X_1, \dots, X_n, t)$, смысл которого состоит в том, что произведение

$$\rho dX_1 \dots dX_n \quad (3.3)$$

представляет собой вероятность обнаружения в момент времени t некоторого члена ансамбля в элементе объема фазового пространства $dX_1 \dots dX_n$, содержащего внутри себя состояние $\{X_1, \dots, X_n\}$. Величина ρ должна быть неотрицательной и удовлетворять следующему условию нормировки:

$$\int \rho dX_1 \dots dX_n = 1 \quad (3.4)$$

Для того чтобы иметь возможность предсказывать вероятность появления в ансамбле тех или иных значений определенных

переменных, необходимо установить вид уравнения для ρ . С этой целью мы воспользуемся некоторыми рассуждениями, весьма близкими к тем, которые используются в механике жидких сред. Это можно обосновать, рассматривая величину ρ как плотность жидкости, линии тока которой совпадают с траекториями, определяемыми уравнениями (3.1).

Рассмотрим некоторый элемент объема фазового пространства, содержащийся между X_1 и X_1+dX_1 , X_2 и X_2+dX_2 и т. д. (рис. 3.5). Поскольку переменные X_i изменяются в соответствии с уравнениями (3.1), число систем, для которых изображающие точки содержатся в элементе объема $dX_1 dX_2 \dots$, будет со временем изменяться, так как точки входят в этот объем и покидают его, пересекая его границы. Например, изображающие точки, входящие в этот объем через поверхность $X_1 = \text{const}$ за интервал

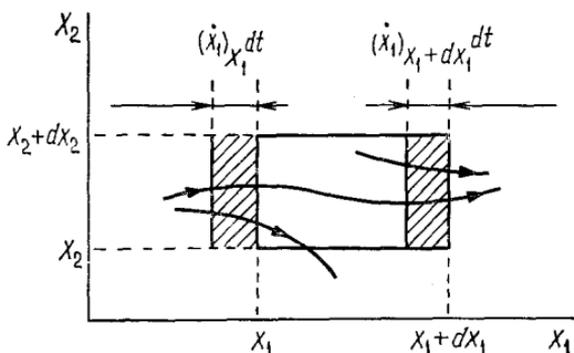


Рис. 3.5. К вопросу о балансе между числом систем, являющихся элементами ансамбля Гиббса, поступающих в определенный элемент фазового объема (левая сторона прямоугольника) и затем покидающих этот объем (правая сторона прямоугольника).

времени dt , содержатся в объеме $X_1 dt dX_2 \dots dX_n = F_1 dt dX_2 \dots dX_n$, а их число тем самым равно

$$\rho(X_1, \dots, X_n, t) F_1 dt dX_2 \dots dX_n$$

Число точек, покидающих этот объем через поверхность $X_1 + dX_1 = \text{const}$ за то же время, дается совершенно таким же выражением, с той разницей, что ρ и F_1 должны вычисляться при $X_1 + dX_1$ вместо X_1 :

$$\begin{aligned} \rho(X_1 + dX_1, \dots, X_n, t) F_1(X_1 + dX_1, \dots, X_n) dt dX_2 \dots dX_n = \\ = \rho(X_1, \dots, X_n, t) F_1 dt dX_2 \dots dX_n + \\ + \frac{\partial \rho}{\partial X_1} F_1 dt dX_1 \dots dX_n + \rho \frac{\partial F_1}{\partial X_1} dt dX_1 \dots dX_n \end{aligned}$$

Правое выражение здесь получено из левого путем разложения в ряд по степеням dX_1 величин $\rho(X_1 + dX_1)$ и $F_1(X_1 + dX_1)$ в ок-

рестности точки X_1 и пренебрежения членами разложения выше первой степени.

С помощью таких же соображений, но только применительно к другим границам этого объема, можно получить следующее уравнение:

(изменение ρ за время dt) $dX_1 \dots dX_n =$ число точек, поступающих в объем—число точек, выходящих из объема

Поделив обе части этого равенства на $dX_1 \dots dX_n$ и воспользовавшись предыдущим равенством, получаем

$$\frac{\partial \rho}{\partial t} = - \sum_{i=1}^n \left(\frac{\partial \rho}{\partial X_i} F_i + \rho \frac{\partial F_i}{\partial X_i} \right) \quad (3.5)$$

Заметим, что величину $\sum_i \partial F_i / \partial X_i$ можно рассматривать как дивергенцию вектора \mathbf{F} , компоненты которого отражают скорость изменения соответствующих переменных X_i [см. уравнение (3.1)]. Кроме того, если проследить во времени за течением «газа» изображающих точек, то можно составить выражение для «гидродинамической» производной $d\rho/dt$, в котором учтены как явная зависимость ρ от времени, так и неявная зависимость от времени через переменные $X_1(t) \dots X_n(t)$:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_i \frac{\partial \rho}{\partial X_i} \dot{X}_i = \frac{\partial \rho}{\partial t} + \sum_i \frac{\partial \rho}{\partial X_i} F_i$$

Таким образом, из (3.5) получается следующий важный результат:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^n \frac{\partial \rho}{\partial X_i} F_i = -\rho \operatorname{div} \mathbf{F} \quad (3.6)$$

Если вычислить $\operatorname{div} \mathbf{F}$ для гамильтоновой системы, то из (2.56) видно, что

$$\operatorname{div} \mathbf{F} = \sum_{i=1}^n \left(\frac{\partial}{\partial q_i} \dot{q}_i + \frac{\partial}{\partial p_i} \dot{p}_i \right) = \sum_{i=1}^n \left(\frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0$$

Смысл уравнения (3.6) состоит в том, что плотность вероятности ρ при движении в фазовом пространстве сохраняется. Это и есть теорема Лиувилля, однако сформулированная несколько иначе. С учетом (3.6) получаем уравнение Лиувилля:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^n \frac{\partial \rho}{\partial X_i} F_i = 0 \quad (\text{для консервативных систем}) \quad (3.7)$$

На самом деле уравнение (3.7) может выполняться не только для гамильтоновых систем, но и в более общем случае. Для этого достаточно, чтобы жидкость, образованная изображающими точками в фазовом пространстве, была несжимаемой, т. е. выполнялось условие $\operatorname{div} \mathbf{F} = 0$. В действительности с помощью этого условия можно *определить* класс консервативных систем. Соответственно уравнение (3.6) с $\operatorname{div} \mathbf{F} \neq 0$ будет необходимым (хотя и не всегда достаточным) условием диссипативности системы.

Уравнение Лиувилля для консервативных систем (3.7) играет весьма важную роль в статистической физике, поскольку оно служит отправной точкой как для эргодической теории, так и для кинетической теории необратимых процессов. Детальнее эти вопросы мы обсудим в гл. 5.

Для диссипативных систем использование статистических представлений, и в частности уравнения (3.6), оказывается необходимым в случае хаотического поведения, когда детерминистическое движение проявляет в значительной мере случайный характер. Эти вопросы будут рассмотрены в разд. 3.10. В следующих нескольких разделах мы последовательно обсудим консервативные и диссипативные системы, динамика которых поддается описанию в фазовом пространстве низкой размерности. При этом мы увидим, что, несмотря на свою кажущуюся простоту, некоторые из этих систем могут демонстрировать весьма нетривиальные типы поведения, напоминающие явления, которые рассматривались в гл. 1 и 2.

3.3. Интегрируемые консервативные системы

Рассмотрим гамильтонову систему, описываемую уравнениями (2.5) или (3.7). Пусть $n = 2N$, где N — число независимых координат или импульсов для данной задачи, называемое также числом *степеней свободы*. Как известно, интегрирование системы $2N$ обыкновенных дифференциальных уравнений первого порядка сводится к отысканию $2N$ независимых первых интегралов движения, один из которых дает связь между переменными в данной системе. Однако гамильтоновы системы образуют особый класс, для которого при довольно общих условиях достаточно знать лишь N первых интегралов. Как впервые было показано Лиувиллем, если эти интегралы достаточно регулярны (в математическом смысле), то решение системы уравнений можно свести к квадратурам.

Содержание теоремы Лиувилля этим не исчерпывается. Пусть I_i , $i = 1, \dots, N$ — подходящие комбинации N интегралов движения. Тогда гамильтоново течение можно свести к системе уравнений вида

$$\frac{dI_i}{dt} = 0, \quad \frac{d\varphi_i}{dt} = \omega_i(\{I_j\}), \quad i = 1, \dots, N \quad (3.8)$$

Эти уравнения также имеют гамильтонову структуру, благодаря чему величины ω_i связаны с новым (не зависящим от φ) гамильтонианом H посредством $\omega_i = \partial H / \partial I_i$. Величины I_i и φ_i называются фазовыми интегралами и угловыми переменными соответственно. Доказательство этого важного результата можно найти в классических руководствах по механике, например в книге Арнольда.

Уравнения (3.8) дают определение частного класса *интегрируемых систем*. Согласно теореме Лиувилля, такие системы могут характеризоваться любым из следующих двух свойств: иметь N достаточно регулярных первых интегралов или сводиться к системе (3.8) путем соответствующих преобразований, сохраняющих гамильтонову структуру (такие преобразования называются каноническими).

Как следствие этого, любая гамильтонова система с одной степенью свободы ($n=2$, $N=1$) является интегрируемой, поскольку она имеет один интеграл движения, а именно полную энергию $H=E=\text{const}$. Таким образом, системы уравнений для маятника [см. уравнение (2.6) и рис. 2.2] и для гармонического осциллятора являются интегрируемыми системами. Система уравнений гамильтонова типа с двумя степенями свободы интегрируема в том случае, когда существует достаточно регулярный первый интеграл, независимый от гамильтониана H . В случае трех степеней свободы нужно иметь три первых интеграла. В случае задачи о волчке, одной из наиболее популярных во вводных курсах механики, существование этих интегралов вытекает из законов сохранения энергии и момента количества движения. Следовательно, волчок также представляет собой интегрируемую систему. То же самое относится и к задаче о двух телах в случае центральных сил взаимодействия. Значение этой задачи обусловлено прежде всего ее связью с движением небесных тел.

В более общем случае можно сказать, что интегрируемы все системы, которые расцепляются на системы с одной степенью свободы. Очевидным примером такого рода являются линейные уравнения движения, поскольку преобразование к нормальным колебаниям, диагонализующее матрицу коэффициентов, сводит их к расцепленным уравнениям с одной степенью свободы. Это обстоятельство лежит в основе обширной литературы по малым колебаниям вблизи положения равновесия. Иллюстрацией может служить физика твердого тела в гармоническом приближении. Следует отметить, что интегрируемыми могут быть и некоторые нелинейные системы. Наиболее интересными приме-

рами такого рода являются пространственно распределенные системы типа решетки Тода или уравнения Кортевега — де Фриза.

Перейдем теперь к изучению возможных типов поведения интегрируемых систем. Ответ выглядит просто в представлении, описывающем эволюцию системы на языке интегралов движения (действий) и соответствующих углов. В случае одной степени свободы имеем

$$I_1 = I_{10} = \text{const}, \quad \varphi_1 = \omega_1 t + \varphi_{10} \quad (3.9)$$

Если рассматривать (I_1, φ_1) в качестве полярных координат на плоскости с ортогональными координатами P, Q , то задача будет изоморфной задаче о гармоническом осцилляторе. Фазовое пространство в данном случае является плоскостью, и для каждого значения I_1 движение представляет собой равномерное вращение по окружности (рис. 3.6).

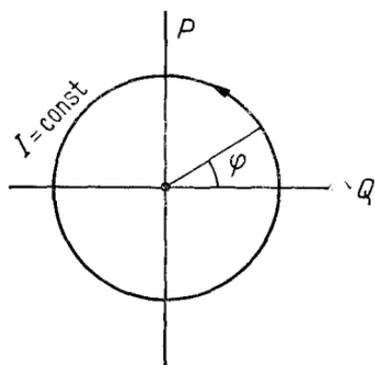


Рис. 3.6. Фазовая траектория консервативной системы с одной степенью свободы. Такая система автоматически является интегрируемой.

В случае многих степеней свободы движение есть суперпозиция N независимых движений, каждое из которых описывается уравнением типа (3.9). Размерность фазового пространства равна $2N$, однако ввиду постоянства величин I_i движение происходит на множестве размерности N , проекции которого на отдельные фазовые плоскости, соответствующие различным степеням свободы, представляют собой окружности. Такой объект, называемый тором, показан на рис. 3.7, а, б в случае $N=2$. Заметим, что в силу сохранения полной энергии фазовые траектории системы с двумя степенями свободы содержатся в трехмерном подпространстве. Таким образом, тор, показанный на рис. 3.7, является фигурой в трехмерном пространстве.

Отсюда непосредственно следует зависимость переменных P_i, Q_i от времени. При $N=1$ эта зависимость имеет периодический характер, причем период равен $2\pi/\omega_1$. В случае $N>1$ имеет место суперпозиция периодических функций с различными

периодами. Если эти периоды соизмеримы, т. е. если существует набор отличных от нуля целых чисел k_1, k_2, \dots , таких, что

$$k_1\omega_1 + k_2\omega_2 + \dots = 0$$

то, очевидно, мы имеем дело с *периодическим* движением, которому соответствует замкнутая кривая. Однако, если такого набора целых чисел не существует, то движение является *квазипериодическим*. Такому движению соответствует спиральная траектория, навитая на тор и не имеющая самопересечений. При $N=2$ такой случай может иметь место, например, если $\omega_1=1$, а ω_2 представлено иррациональным числом, например $\omega_2=\sqrt{2}$.

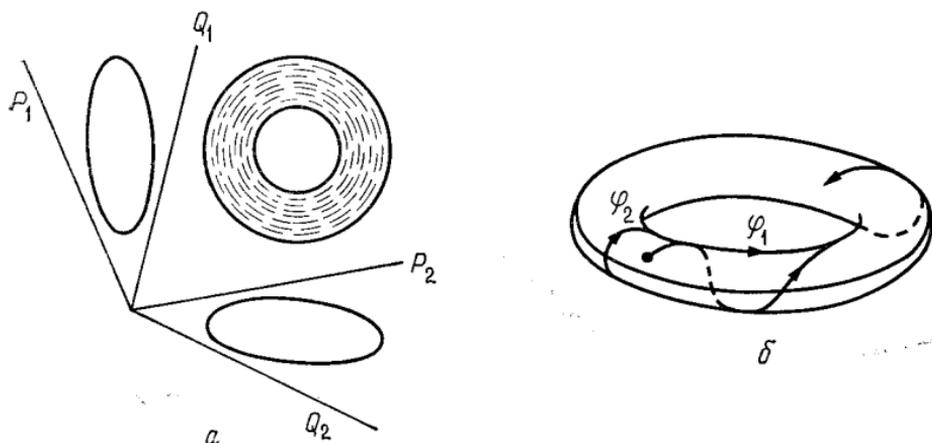


Рис. 3.7. Представление фазовых траекторий интегрируемой динамической системы с двумя степенями свободы. Движение происходит на двумерной поверхности тора, размер которого зависит от комбинации переменных действия. Сохранение полной энергии уменьшает число независимых переменных с 4 до 3.

Поэтому тор на рис. б находится в трехмерном пространстве.

Траектория в случае квазипериодического движения может иметь довольно сложный вид. Поскольку такая траектория незамкнута и благодаря теореме о единственности, обсуждавшейся в разд. 3.1, не имеет самопересечений, она постепенно заполняет весь тор в том смысле, что рано или поздно она пройдет через любую заранее заданную окрестность любой точки тороидальной поверхности. Движение, удовлетворяющее этому условию, называется «всюду плотным».

Как неоднократно подчеркивалось в гл. 2 и в особенности в разд. 2.6, асимптотическая устойчивость не имеет места в консервативной системе. Тем самым ни периодическое, ни квазипериодическое движения, обсуждавшиеся выше, не могут служить в качестве моделей аттракторов. И это понятно: благодаря первому из уравнений (3.9) имеется континуум допустимых

значений интегралов движения I_i . Таким образом, фазовый портрет не соответствует ни рис. 3.6, ни рис. 3.7, б — на самом деле он содержит континуум замкнутых орбит, или торсидальных поверхностей, вложенных друг в друга. Поэтому малое возмущение, сдвигающее систему с данной траектории, переводит ее на другую инвариантную поверхность, на которой система и остается до тех пор, пока новое возмущение не переведет ее на еще одну инвариантную поверхность и т. д. При воздействии возмущений вместо четко воспроизводимого поведения в лучшем случае следует ожидать простую или орбитальную устойчивость и «дрожащую» траекторию.

Такие виды движения, когда точка равновесия окружена континуумом замкнутых инвариантных множеств, называются *эллиптическими*, а сама точка равновесия — *центром*. В соответствии с разд. 3.2 (особое внимание следует обратить на рис. 3.4, а) можно ожидать, что эти виды движения приведут к умеренным деформациям фазовых объемов. Разумеется, согласно теореме Лиувилля, сами фазовые объемы будут сохраняться.

До сих пор мы занимались описанием поведения интегрируемых систем в представлении переменных действие — угол. Как объяснялось в начале настоящего раздела, эти переменные связаны с исходными с помощью некоторого преобразования. В ряде исключительных случаев это преобразование имеет довольно простой вид. Например, в задаче о гармоническом осцилляторе

$$H = \frac{1}{2} m \dot{x}^2 + k \frac{x^2}{2}$$

переменные I и φ неявно задаются с помощью соотношений

$$x = \left(\frac{2I}{\sqrt{mk}} \right)^{1/2} \sin \varphi, \quad \dot{x} = \left(2I \sqrt{\frac{k}{m^3}} \right)^{1/2} \cos \varphi$$

а гамильтониан сводится к

$$H = \sqrt{\frac{k}{m}} I = \omega I$$

Здесь ω — угловая частота осциллятора. Однако в общем случае это преобразование выглядит довольно сложно. Поэтому можно ожидать, что в рамках исходных переменных траектории будут представлять собой весьма искаженные варианты картинок, показанных на рис. 3.6 и 3.7, б. Более того, для некоторых типов движения траектории в исходном фазовом пространстве будут выглядеть весьма искусственно в представлении переменных действие — угол по той простой причине, что «угловая» пе-

ременная не будет удовлетворять типичному для угла условию $0 \leq \varphi \leq 2\pi$. Это характерно для неустойчивых движений типа гиперболического движения, показанного на рис. 3.4, б. В случае таких движений некоторые переменные становятся неограниченными и соответственно периоды определенных траекторий стремятся к бесконечности. В случае маятника (см. рис. 2.2) это происходит вблизи точки неустойчивого равновесия $\theta = \pm\pi$. Область ограниченных колебаний (вибраций) отделена от области неограниченного движения (вращения) траекторией особого типа, называемой *сепаратрисой* (рис. 3.8). Эта траектория со-

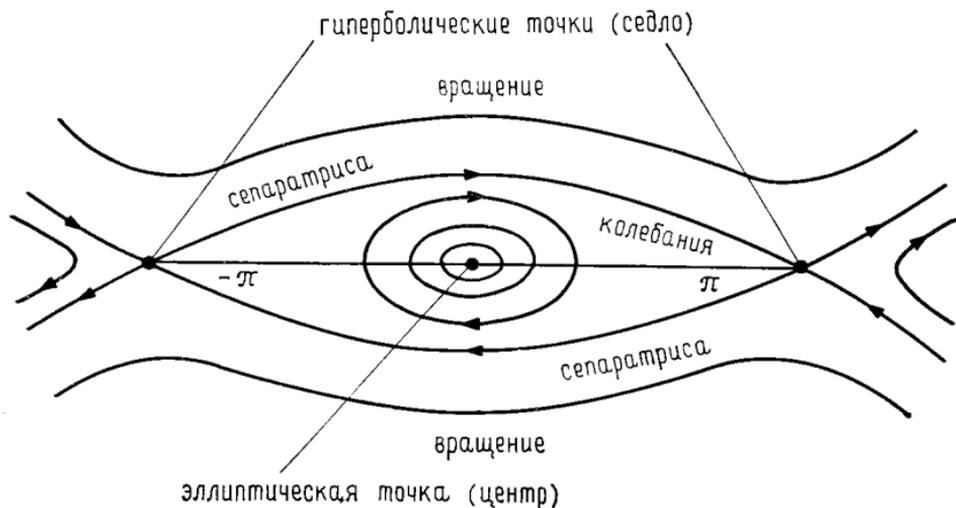


Рис. 3.8. Фазовый портрет маятника. Около эллиптической неподвижной точки в нуле, соответствующей единственному состоянию устойчивого равновесия, возможен континуум периодических колебательных движений, период и амплитуда которых зависят от начальных условий. Гиперболические неподвижные точки $\pm\pi$ соответствуют состояниям неустойчивого равновесия. Они соединяются замкнутой петлей, состоящей из двух сегментов сепаратрис. Начальные условия вне этой петли приводят к свободному вращательному движению маятника.

единяет точки $\theta = \pi$ и $\theta = -\pi$. Поскольку вблизи неподвижной точки скорость изменения переменной стремится к нулю, время, необходимое для перемещения от любой точки сепаратрисы до неподвижной точки, равно бесконечности. Особая конфигурация траекторий вблизи сепаратрисы, похожая на линии постоянного уровня вблизи горного перевала, породила термин *седловая точка*, который используется для описания некоторых неустойчивых равновесий в консервативных системах. Движение вблизи седловой точки приводит к значительным деформациям элемента фазового пространства при сохранении его объема. Короче, за исключением, быть может, сепаратрисы и седловых точек, таящих в себе общий механизм сложного поведения (в дальней-

шем мы в этом сможем убедиться), можно утверждать, что мир интегрируемых систем живет простой, без особых происшествий жизнью. И это не удивительно, поскольку в интегрируемой системе мы эффективно избавились от взаимодействий и имеем дело с независимыми объектами. Однако, как будет видно из следующего раздела, такую ситуацию можно считать довольно патологической. Примечательно, что, несмотря на свой исключительный характер, интегрируемые системы доминировали в науке до 50-х годов, и сегодня они все еще составляют основной предмет большинства книг по механике. Эта парадоксальная ситуация частично объясняется их огромной исторической ролью и бесспорной педагогической ценностью. Однако в заключение данного раздела следует признать, что мир интегрируемых систем не может быть местом рождения сложного. Это место надо еще поискать, и такой поиск будет нашей основной целью в последующих разделах настоящей главы.

3.4. Бифуркация в простой диссипативной системе: поиск прототипов сложного

Обратимся прежде всего к диссипативным системам. Для этого выберем сначала простую абстрактную математическую модель, на которой легче представить себе некоторые механизмы, ответственные за качественное изменение поведения, а затем свяжем эту модель с реальными физико-химическими системами.

Рассмотрим единственную переменную x , изменяющуюся в соответствии с уравнением

$$\frac{dx}{dt} = f(x, \lambda) = -x^3 + \lambda x \quad (3.10)$$

В этой простой системе имеется единственный управляющий параметр λ , причем скорость изменения x , обозначенная как функция f , зависит от этого параметра строго линейно. Прежде всего рассмотрим неподвижные точки (стационарные состояния):

$$-x_s^3 + \lambda x_s = 0$$

Это простое алгебраическое уравнение всегда имеет одно тривиальное решение:

$$x_0 = 0 \quad (3.11a)$$

Однако могут быть и нетривиальные уравнения. Они удовлетворяют уравнению, получаемому из кубического после сокращения на x_s :

$$-x_s^2 + \lambda = 0$$

При отрицательном λ это уравнение имеет мнимое решение, что не может соответствовать какой-либо физически возможной ситуации. Однако при положительных λ это уравнение допускает следующие два решения:

$$x_{\pm} = \pm \sqrt{\lambda} \quad (3.116)$$

Эти решения сливаются с x_0 при $\lambda=0$ и ответвляются от него при $\lambda>0$. Это так называемое явление *бифуркации*, обсуждавшееся в разд. 2.7 и иллюстрированное рис. 2.12. На рис. 3.9, а приведен аналогичный график применительно к нашей конкретной задаче. Сплошной и штриховой линиями обозначены соответственно асимптотически устойчивое и неустойчивое решения.

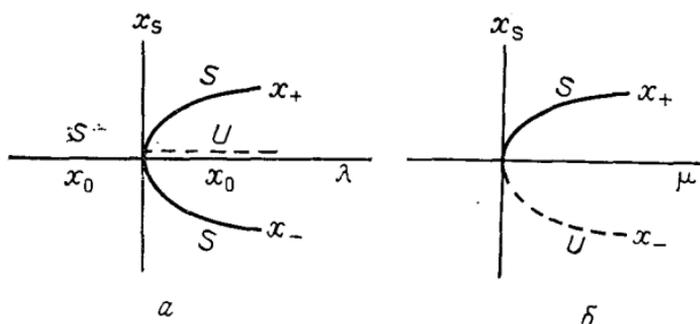


Рис. 3.9. Два примера элементарных бифуркаций в диссипативной динамической системе. *а* — симметричная бифуркация (типа камертона), при которой стандартное состояние x_0 теряет устойчивость при $\lambda>0$ и приводит к двум ветвям решения x_+ и x_- , сливающимся с x_0 при $\lambda=0$; *б* — бифуркация типа предельной точки, при которой устойчивая (x_+) и неустойчивая (x_-) ветви аннигилируют при встрече в точке $\mu=0$.

В случае рассматриваемой здесь простой модели проблема устойчивости довольно тривиальна, поскольку уравнение (3.10) допускает точное решение с помощью стандартных методов. Оказывается, что решение x_0 является глобально асимптотически устойчивым при $\lambda<0$ и неустойчивым при $\lambda>0$, а решения x_+ и x_- асимптотически (но не глобально) устойчивы. Иными словами, ветви x_{\pm} появляются в результате бифуркации в тот момент, когда «стандартное состояние» $x_0=0$ теряет устойчивость, причем сами эти ветви устойчивы. Такую бифуркацию мы будем называть *надкритической*. На самом деле связь между надкритическим ветвлением и устойчивостью отнюдь не случайна. Согласно одному из общих результатов теории бифуркаций, обсуждающейся в какой-то мере в приложении II, при определенных условиях, налагаемых на функцию f , часто выполняющихся в практических приложениях, надкритические ветви устойчивы, а подкритические неустойчивы.

Однако наиболее поразительный результат нашей простой

модели связан с уравнением (3.11б). Вспомним, что исходно мы имели простое кубическое уравнение, плавно зависящее от параметра λ . Теперь же видно, что эта плавная динамика приводит к *особенности*. В самом деле, в окрестности точки бифуркации $\lambda=0$ решения x_{\pm} не могут быть разложены в ряд по степеням параметра λ . Следовательно, они зависят от λ *не аналитически*. Это является математическим отражением качественного изменения поведения, обусловленного явлением бифуркации.

Рассмотрим еще более простой пример диссипативной системы:

$$\frac{dx}{dt} = -x^2 + \mu \quad (3.12)$$

Управляющий параметр здесь обозначен через μ . Выражения для неподвижных точек x_s имеют вид

$$x_{\pm} = \pm \sqrt{\mu} \quad (3.13)$$

Эта зависимость представлена на рис. 3.9, б. В этом случае тоже можно найти точное решение динамического уравнения и показать, что ветвь x_- неустойчива, в то время как ветвь x_+ асимптотически устойчива. Заметим, что при уменьшении положительных μ в точке $\mu=0$ происходит столкновение и аннигиляция устойчивой и неустойчивой ветвей. По этой причине точка $\mu=0$ называется *предельной точкой*, или складкой. Ее наличие снова указывает на появление особенностей, как это видно из уравнения (3.13).

Объединим теперь два предыдущих примера, рассматривая диссипативную систему, описываемую уравнением

$$\frac{dx}{dt} = -x^3 + \lambda x + \mu \quad (3.14)$$

Здесь нет квадратичного члена, фигурирующего в уравнении (3.12), поскольку его всегда можно устранить подходящей заменой переменных. Таким образом мы приходим к задаче с двумя параметрами (λ и μ). Неподвижные точки теперь определяются канонической формой кубического уравнения общего вида:

$$-x_s^3 + \lambda x_s + \mu = 0$$

Как известно из элементарной алгебры, это уравнение может иметь до трех реальных решений. Более того, при изменении параметров происходит слияние трех решений, в результате чего остается единственное реальное решение. Можно определить кривые в *параметрическом пространстве*, разделяющие эти два режима:

$$4\lambda^3 + 27\mu^2 = 0 \quad (3.15)$$

Эти кривые представлены на рис. 3.10, а. Область существования трех реальных корней заканчивается в точке (начало координат на рис. 3.10, а), в которой зависимость λ от μ имеет особенность. Это известная особенность типа острья.

На рис. 3.10, б, в приведены зависимости решений от параметров. На рис. 3.10, б представлена зависимость x_s от μ при фиксированном λ . S-образная форма кривой обусловлена множественностью решений в соответствующей области значений па-

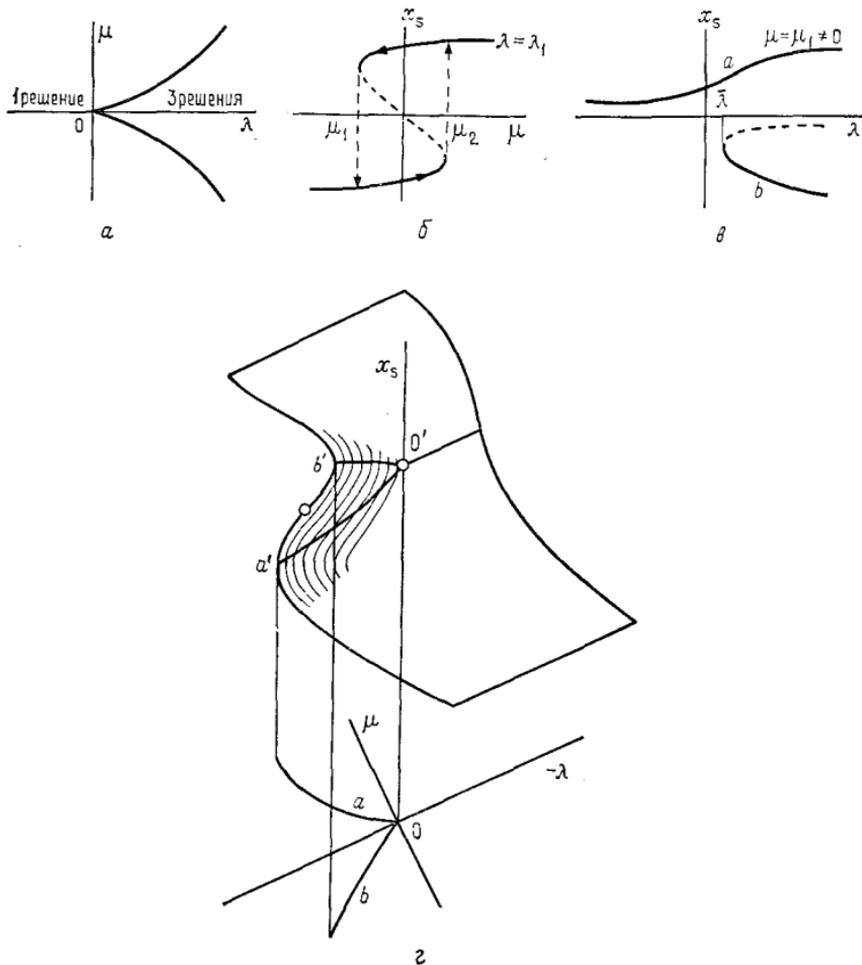


Рис. 3.10. Влияние параметров на бифуркацию стационарных решений уравнения (3.14). а — кривая в пространстве параметров, ограничивающая область существования трех реальных решений; б — гистерезисное поведение решения при изменении параметра μ . Параметр λ фиксирован. Бифуркация типа предельной точки остается грубой; в — разрушение бифуркации типа камертона в случае, когда выступающий в роли возмущения параметр μ не равен строго нулю; г — трехмерная картина, объединяющая в себе информацию, представленную в случаях а—в.

раметра μ . Более того, две из ветвей оказываются устойчивыми одновременно. Область бистабильности оканчивается в двух предельных точках μ_1 и μ_2 , в окрестности которых имеет место поведение, изображенное на рис. 3.9, б. Легко убедиться в том, что при этих условиях возрастание μ выше μ_1 и далее выше μ_2 с последующим изменением μ в обратном направлении приведет к *циклу гистерезиса*. Это явление уже упоминалось в разд. 1.4 в связи с экспериментами по реакции БЖ.

На рис. 3.10, в построена зависимость решения x_s от λ при фиксированном значении μ . Теперь мы имеем две отдельные кривые, одна из которых (кривая *a*) определена для всех значений λ , а другая (кривая *b*) определена лишь при $\lambda \geq \bar{\lambda}$ и имеет особенность типа предельной точки при $\bar{\lambda}$. При $\lambda < \bar{\lambda}$ имеется лишь устойчивое решение, а при $\lambda > \bar{\lambda}$, как и прежде, имеет место бистабильность. Наконец, рис. 3.10, г объединяет информацию, содержащуюся на рис. 3.10, а — в в виде трехмерной зависимости решения x_s от параметров λ и μ .

Важно иметь в виду, что изображенная на рис. 3.9, а симметричная бифуркация не имеет места при любом, сколь угодно малом, но отличном от нуля значении $|\mu|$. В этом смысле параметр μ уподобляется своего рода *дефекту*, разрушающему бифуркацию. С другой стороны, предельная точка (рис. 3.9, б) оказывается «прочной» в том смысле, что присутствует как на рис. 3.10, б, так и на рис. 3.10, в. Однако если и λ и μ изменяются одновременно, то всегда найдется такая комбинация их значений (в нашем случае $\mu = 0$ и λ переменны), при которой бифуркация станет возможной, поскольку система сможет пересечь точку возврата (см. рис. 3.10, а) симметричным образом.

В действительности проведенное выше довольно подробное обсуждение иллюстрирует глубокую концепцию *структурной устойчивости*, играющую исключительно важную роль в теории динамических систем. Из него следует, что определенные явления вроде обсуждавшейся бифуркации типа острия имеют место лишь в том случае, когда параметры удовлетворяют по меньшей мере одному равенству. Поскольку в реальной физической системе строгое выполнение этого требования трудно достичь, можно ожидать, что подобные явления исчезнут при небольших изменениях значений параметров. Такие явления мы будем называть *структурно-неустойчивыми*. С другой стороны, имеются и другие явления типа, например, бифуркации в предельной точке (рис. 3.9, б), которые продолжают существовать (хотя и могут оказаться сдвинутыми) при изменении управляющих параметров, влияющих на структуру динамических уравнений. Такие явления называются *структурно-устойчивыми*. Заметим, что в соответствии с этим определением *все консервативные си-*

стемы структурно-неустойчивы, поскольку наличие малых диссипативных членов (типа слабого трения в случае маятника) качественно изменяет фазовый портрет, сообщая свойство асимптотической устойчивости определенным решениям — аттракторам.

Чего же мы достигли в данном разделе с точки зрения физических приложений? Мы установили некоторые «минимальные» математические модели, воспроизводящие одно из экспериментально наблюдаемых переходных явлений, сопутствующих сложному поведению, а именно возникновение множественных одновременно устойчивых решений. Это явление напрямую связано со способностью к переключениям и тем самым к выполнению регуляторных задач. Правда, мы еще должны убедиться в том, что такие математические модели могут реализоваться в виде физических систем, представляющих определенный интерес. Эту задачу мы отложим до разд. 3.6. Завершая обсуждение, добавим, что в системах с одной переменной, содержащих нелинейности более высокого порядка, возможны более сложные переходные явления. В теории катастроф, представляющей собой одно из ответвлений теории динамических систем, проводится классификация таких явлений и устанавливается связь между параметрами, позволяющая выделить области качественно различающихся типов поведения.

3.5. Диссипативные системы в двумерном фазовом пространстве: предельные циклы

Эволюция одномерных систем, рассматривавшихся в предыдущем разделе, происходила в одномерном фазовом пространстве. Пользуясь терминологией разд. 3.1, в таком пространстве возможен лишь один вид инвариантного множества — неподвижные точки, соответствующие стационарным состояниям. Фазовые траектории представляют собой отрезки прямых, сходящихся к неподвижным точкам или расходящихся от них в зависимости от устойчивости последних. Примеры такого рода представлены на рис. 3.11, *а, б*, построенном для моделей, описываемых уравнениями (3.10) и (3.12) соответственно.

Из топологических соображений ясно, что в двумерном фазовом пространстве имеется большая гибкость, однако в разд. 3.1 показано, что различные возможности ограничены имеющимися жесткими условиями. Именно в таком фазовом пространстве происходит эволюция системы с двумя переменными. Естественно, мы ожидаем, что и в двумерном случае останется инвариантное множество одномерных систем, т. е. неподвижные точки. Более того, поскольку консервативные системы с одной степенью свободы составляют особый класс систем с двумя пере-

менными, при определенных условиях можно рассчитывать встретить фазовые портреты интегрируемых систем, обсуждавшиеся в разд. 3.3. В частности, неподвижные точки могут оказаться седловыми точками или центрами; вокруг центров могут существовать континуумы замкнутых траекторий, а сепаратрисы, проходящие через седловые точки, могут изгибаться, образуя петли, соединяющие неустойчивые состояния (см. рис. 3.8). Вопрос в том, можно ли ожидать, что благодаря лишней размерности фазового пространства в дополнение к указанным появятся новые возможности, специфически связанные с диссипативным характером системы?

Одна такая возможность, связанная со способом приближения к неподвижной точке P , показана на рис. 3.12. В обоих случаях, изображенных на этом рисунке, точка P асимптотически устойчива, однако фазовые портреты вблизи ее окрестности совершенно различны (ясно, что это было невозможно в случае рис. 3.11!). В случае рис. 3.12, *а* траектории стремятся к точке P

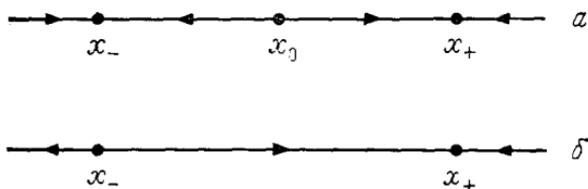


Рис. 3.11. Фазовые траектории диссипативной динамической системы с одной переменной. *а* — одно неустойчивое и два асимптотически устойчивых решения, как на рис. 3.9, *а*; *б* — одно неустойчивое и одно асимптотически устойчивое решение, как на рис. 3.9, *б*.

с возрастанием времени примерно вдоль радиуса, что соответствует экспоненциальному затуханию во времени возмущений вблизи этой точки. Такую точку мы будем называть (устойчивым) *узлом*. С другой стороны, в случае рис. 3.12, *б*, прежде чем сойтись в точке P , траектории совершают вокруг нее несколько спиральных витков. На временной зависимости соответствующей переменной это проявляется в виде затухающих колебаний. Такую точку мы называем (устойчивым) *фокусом*. Заметим, что возможны и такие фазовые портреты, которые получаются изменением направления траектории на противоположное. В этом случае мы имели бы дело с неустойчивыми узлом или фокусом. Независимо от их устойчивости узлы или фокусы могут возникать только в диссипативных системах, поскольку совершенно ясно, что они нарушают инвариантность относительно обращения времени.

Несомненно, однако, что наиболее интересная возможность показана на рис. 3.13. В рассматриваемой области фазового

пространства система имеет единственную неподвижную точку P , которая при выбранных значениях параметров неустойчива. Поэтому траектории, начинающиеся вблизи P , удаляются от этой точки. Однако в физически разумном случае траектория не может удаляться бесконечно, поскольку это означало бы, что система может поглотить или произвести бесконечное количество вещества или энергии. Поэтому можно ожидать, что траектории останутся в ограниченной области фазового пространства и тем не менее будут удовлетворять топологическим требованиям, обсуждавшимся в разд. 3.1 (отсутствие самопересечений и т. д.). В двумерном пространстве единственный выход состоит в том, что траектории закончатся в виде замкнутой кривой (на рис. 3.13 она обозначена C). Такие кривые, впервые открытые Пуанкаре, мы будем называть *предельными циклами*. В конкретном примере, показанном на рис. 3.13, предельный цикл

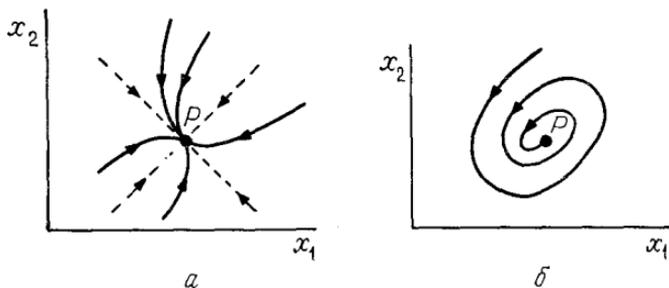


Рис. 3.12. Возможные конфигурации в фазовом пространстве вблизи неподвижной точки P диссипативной динамической системы с двумя переменными. Эта конфигурация дополняет возможности (седло и центр) консервативных систем. a — асимптотически устойчивый узел. Решения приходят в точку P монотонно; b — асимптотически устойчивый фокус. Решения приближаются к P на фоне затухающих колебаний. В обоих случаях точка P является аттрактором нулевой размерности (точка).

является асимптотически устойчивым. Однако, как только реализуется возможность существования замкнутых кривых на конечном расстоянии друг от друга (в противоположность консервативным системам!), становятся возможными и другие конфигурации. Например, кривая C может оказаться орбитально неустойчивой, и вытолкнутые с нее траектории могут устремиться к другому предельному циклу или устойчивой неподвижной точке. Так или иначе, в соответствии с разд. 3.1, находясь в состоянии предельного цикла, система всегда проходит одну и ту же последовательность состояний, т. е. демонстрирует периодическое поведение. Таким образом, асимптотически устойчивый предельный цикл представляет собой *периодический аттрактор*. Вследствие своей асимптотической устойчивости он не поддается разрушительному действию возмущений и тем самым

может служить естественным прототипом наблюдаемых в природе ритмических явлений, примеры которых приводились в гл. 1.

Перейдем теперь к изучению механизмов образования предельного цикла. Чтобы дать необходимые представления о таких механизмах и при этом по мере возможности избежать специальной терминологии, прежде всего рассмотрим модель, имеющую точное решение. Представим некоторую консервативную систему с двумя переменными типа, скажем, гармонического осциллятора (в соответствии с разд. 3.3 такая система является автоматически интегрируемой). Эволюция системы в перемен-

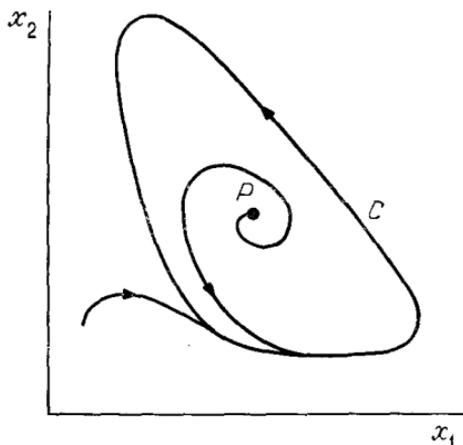


Рис. 3.13. Периодический аттрактор: фазовые траектории приближаются к предельному циклу C , окружающему неустойчивую неподвижную точку P .

ных действие — угол описывается уравнениями, аналогичными (3.8):

$$\frac{dr}{dt} = 0, \quad \frac{d\varphi}{dt} = \omega$$

По причинам, которые вскоре прояснятся, в этой системе действие обозначено буквой r . При данном значении r эта система движется по замкнутой траектории типа изображенной на рис. 3.6. Эта траектория напоминает предельный цикл с тем отличием, что эта траектория принадлежит континууму замкнутых траекторий. Итак, модифицируем эту систему, введя в нее *диссипативное возмущение*, нарушающее симметрию при обращении времени. Выберем это возмущение таким образом, чтобы существовала также и неподвижная точка, а возникающий при этом предельный цикл был бы асимптотически устойчивым, по крайней мере при определенных условиях. Очевидно, нет смысла накладывать возмущение на угловую переменную φ , которая «пассивно» возрастает на 2π при каждом обороте вдоль периодической траектории. Поэтому нам остается ввести возмущение

в радиальную часть путем добавления нетривиального члена к правой части первого уравнения. В соответствии с разд. 3.4 [см., например, уравнение (3.10)] простейшим возмущением, удовлетворяющим этим требованиям, является кубический член. Таким образом, можно написать

$$\frac{dr}{dt} = \lambda r - r^3, \quad \frac{d\varphi}{dt} = \omega \quad (3.16)$$

Первое из этих уравнений совпадает с (3.10), с тем дополнительным условием, что радиус r должен быть положительным. Свойства его решения, показанного на рис. 3.14, непосредственно следуют из рис. 3.9, а. При $\lambda < 0$ существует лишь тривиаль-

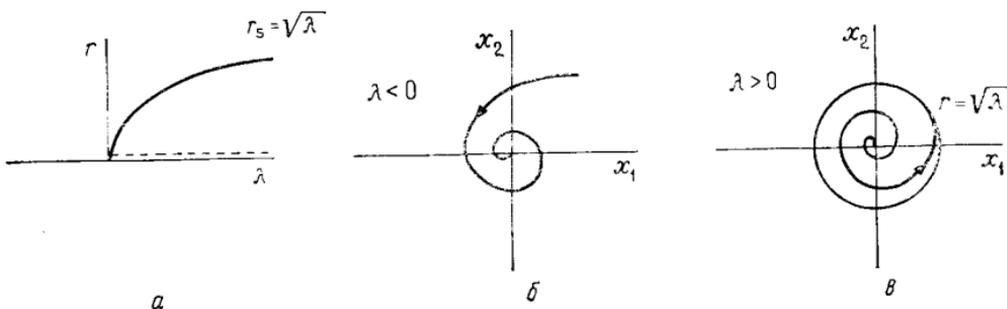


Рис. 3.14. Бифуркация Хопфа, иллюстрируемая с помощью простой модели (3.16). а — радиус предельного цикла как функция управляющего параметра λ ; б, в — фазовые портреты при $\lambda < 0$ и $\lambda > 0$ соответственно.

ное решение $r=0$, являющееся устойчивым. В фазовом пространстве, представленном ортогональными координатами

$$x_1 = r \cos \varphi, \quad x_2 = r \sin \varphi$$

траектории будут сходиться к неподвижной точке $x_1 = x_2 = 0$. Эта точка должна быть фокусом в связи с тем, что угловая переменная φ (фаза) будет изменяться в соответствии с

$$\varphi = \varphi_0 + \omega t \quad (3.17)$$

Здесь величина φ_0 представляет собой начальное значение фазы (см. рис. 3.14, б). Однако при $\lambda > 0$ первое из уравнений (3.16) допускает нетривиальное решение $r_s = \sqrt{\lambda}$. Оно асимптотически устойчиво, в то время как неподвижная точка $r=0$ неустойчива. В фазовом пространстве (рис. 3.14, в) такая конфигурация будет представлена траекториями, удаляющимися по спирали от неподвижной точки и сходящимися к замкнутой кривой $r_s = \sqrt{\lambda}$ — предельному циклу. Таким образом, $\lambda=0$ представляет собой точку бифуркации нового типа, называемую *бифуркацией Хопфа*. В отличие от статических бифуркаций, с которыми мы встречались в разд. 3.4 (все они в конечном счете приводили

к решениям, не зависящим от времени), бифуркация Хопфа представляет собой динамическое явление, приводящее к периодическому поведению. По мере отклонения параметра λ от его значения при бифуркации амплитуда периодического решения плавно возрастает, начиная с нулевого значения (см. рис. 3.14, а). Более того, она зависит от расстояния до точки бифуркации не аналитически. С другой стороны, период колебаний в общем случае оказывается конечным уже в самой точке бифуркации, поскольку в ней величина ω в равенстве (3.16) не обращается тождественно в нуль. В любом случае важно, что как период, так и амплитуда колебаний являются собственными свойствами системы и зависят лишь от значений параметров λ и ω , фигурирующих в уравнениях эволюции. Это резко отличается от консервативных систем, в которых характеристики колебаний зависят от начальных условий. Таким образом, мы имеем простую математическую модель, воспроизводящую существенные черты широкого класса ритмических явлений из числа тех, с которыми мы встречались в гл. 1.

До тех пор пока r и φ в уравнениях (3.16) обладают свойствами радиальной и угловой переменной соответственно, можно убедиться в том, что эти уравнения не страдают от избытка чувствительности к возмущениям, которая была характерной для решений уравнения (3.10). Сделать это можно с помощью теоремы единственности решения. Таким образом, можно заключить, что *бифуркация Хопфа является структурно-устойчивым явлением.*

В двумерных диссипативных системах могут возникать и более сложные переходные явления. Однако соответствующие бифуркации структурно-неустойчивы в том смысле, что они имеют место лишь тогда, когда между входящими в уравнения параметрами существует определенная связь. Следует отметить, что в теории динамических систем разработана полная классификация всех структурно-устойчивых фазовых портретов систем, эволюционирующих в двумерном фазовом пространстве.

3.6. Сведение к системам меньшей размерности: параметры порядка и нормальные формы

В двух последних разделах мы видели, что некоторые явления, имеющие место в фазовых пространствах низкой размерности, сохраняются при модификации структуры уравнений эволюции путем изменения параметров. Поэтому можно ожидать, что по крайней мере на множестве одно- и двумерных систем эти явления *типичны* в том смысле, что для широких классов таких систем они должны наблюдаться в определенном диапазоне значений параметров. Убедимся в этом на нескольких про-

стных моделях, которые теперь уже не будут абстрактными математическими примерами — мы возьмем их непосредственно из физики и химии.

А. Бифуркации типа острия и предельной точки

Эти явления можно прекрасно проиллюстрировать с помощью моделей, взятых из области химической кинетики. В соответствии с разд. 1.4 и 2.4, как правило, в химической реакции происходят столкновения между различными молекулами, причем частоты столкновений с высокой точностью пропорциональны произведению концентраций рассматриваемых компонентов. Это сразу же порождает степенные нелинейности, лежащие в основе результатов, полученных в разд. 3.4.

Такие возможности полезно рассмотреть на одномерной автокаталитической модели [уравнение (2.15)]. Как показано в (2.17), неподвижные точки являются решениями кубического уравнения. В условиях равновесия это уравнение допускает лишь одно решение, вдали же от равновесия попеременное изменение знаков различных членов приводит к существованию до трех действительных решений. Как известно из алгебры, любое кубическое уравнение можно привести к «канонической форме» типа (3.14) путем линейной замены переменных. Таким образом получается, что модель (2.15), совместимая с требованиями термодинамики и химической кинетики, приводит ко всем статическим бифуркациям, обсуждавшимся в разд. 3.4.

Б. Бифуркация Хопфа и предельные циклы

В химической кинетике имеется еще несколько превосходных иллюстративных примеров. Одной из наиболее интенсивно изучавшихся является тримолекулярная модель, или *Брюсселятор*. В ней фигурируют два сопряженных промежуточных продукта и одна автокаталитическая стадия, а для простоты ее обычно рассматривают как предельный случай необратимой реакции:



Предполагается также, что концентрации A и B контролируются извне. В предельном случае пространственно однородной (перемешиваемой) системы при соответственном обезразмеривании переменных кинетические уравнения имеют вид

$$\frac{dX}{dt} = A - BX + X^2Y - X, \quad \frac{dY}{dt} = BX - X^2Y \quad (3.19)$$

Приравнивая производные по времени нулю, легко найти, что в системе имеется единственная неподвижная точка, $X_s = A$ и $Y_s = B/A$. Более того, стандартными аналитическими методами можно показать, что характер этой точки меняется при $B = A^2 + 1$: если B несколько меньше, чем $A^2 + 1$, то эта точка представляет собой устойчивый фокус; если же B несколько превышает величину $A^2 + 1$, то эта точка становится неустойчивым фокусом (рис. 3.12, б). В последнем случае имеется режим предельного цикла, изображенный на рис. 3.15. Его характеристики

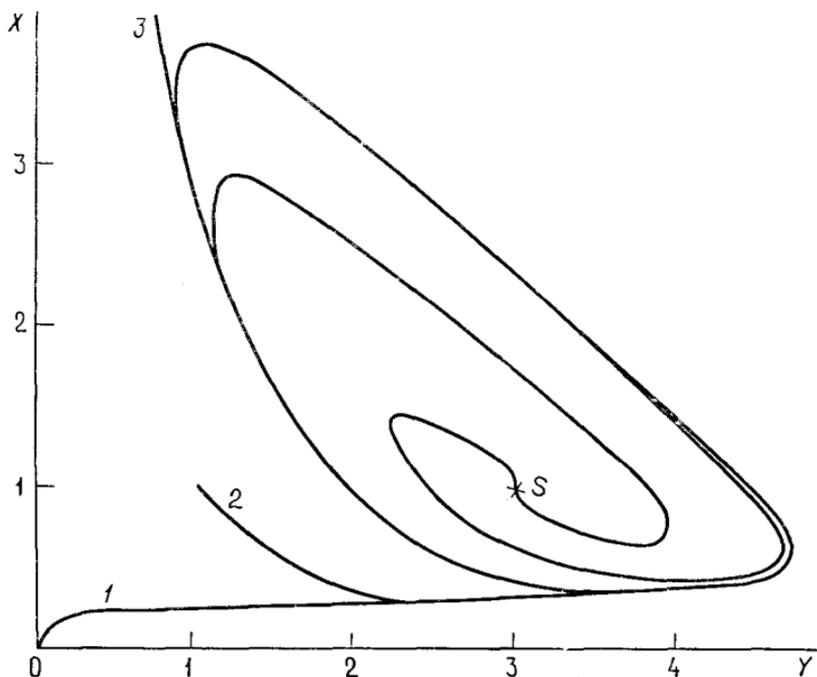


Рис. 3.15. Предельный цикл для модели Брюсселятора (3.19), построенный численными методами. Значения параметров: $A=1$, $B=3$.

совпадают с характеристиками абстрактной математической модели, использовавшейся в разд. 3.5. Иными словами, критическое значение

$$B_c = A^2 + 1 \quad (3.20)$$

представляет собой точку бифуркации Хопфа. На самом деле аналогия между уравнением (3.19) и абстрактной математической моделью простирается гораздо дальше, однако этот вопрос будет рассмотрен в настоящем разделе несколько позже.

Другой класс систем с двумя переменными, приводящих к предельным циклам, представляют электрические цепи. Здесь

нелинейности обусловлены наличием *активных элементов* (электронные лампы, транзисторы и т. д.), имеющих нелинейные вольт-амперные характеристики. Классическим примером служит индуктивное взаимодействие триода с резонансной цепью, впервые подробно изученное нидерландским ученым Ван-дер-По-лем. Применение известных законов сохранения заряда для электрических цепей приводит к замкнутому уравнению для на-пряжения на индуктивности резонансной цепи. В безразмерной форме это уравнение имеет вид

$$\frac{d^2x}{dt^2} - \varepsilon(1 - x^2) \frac{dx}{dt} + x = 0 \quad (3.21)$$

где параметр ε связан со свойствами как триода, так и резонансной цепи. Анализ этого уравнения в классической работе Ван-дер-Поля показал существование незатухающих колебаний типа предельного цикла. В электронике возникновение таких колебаний представляет собой хорошо известное и иногда нежелательное явление, называемое самовозбуждением.

Однако значение бифуркаций, описанных в разд. 3.4 и 3.5, выходит далеко за пределы их применимости к одно- или двумерным системам. В большинстве физических задач с самого начала приходится иметь дело со множеством переменных (около 10 в реакции Белоусова — Жаботинского!). Поэтому если бы анализ бифуркаций оказался полезным лишь для ограниченного круга систем, рассмотренных выше, то он оказался бы скорее занятой игрушкой, нежели мощным средством анализа естественных систем.

К счастью, это не так. Существуют два довольно общих (хотя и не вполне независимых) механизма, благодаря которым формулировки бифуркаций типа острия, предельной точки или Хопфа применимы к чрезвычайно широким классам систем независимо от числа описывающих их переменных и степени нелинейности уравнений.

Начнем с более легкого для восприятия механизма, который тут же приведет к понятиям, знакомым читателям с некоторым химическим багажом. Во многих случаях анализ динамики физико-химических систем приводит к целому набору временных масштабов в связи с тем, что между значениями параметров и/или переменных состояния имеется различие в несколько порядков. Возьмем пример из химии — каталитическая реакция в лабораторных условиях. В таких случаях обычно концентрация катализатора много меньше, чем концентрации реагентов или продуктов реакции, а скорость реакции намного превышает собственную, т. е. в отсутствие катализатора. В результате некоторые промежуточные стадии с каталитическими комплексами протекают очень быстро. В процессах горения энергия акти-

вации некоторых экзотермических реакций очень высока, так что эти реакции, по крайней мере на начальных стадиях, протекают гораздо медленнее, нежели процессы переноса энергии и импульса. Аналогичные примеры можно найти также в биологии, оптике и т. д.

Рассмотрим систему с сильно различающимися временными масштабами, в результате чего можно выделить «быстрые переменные» Y и «медленные переменные» X . Интуитивно ясно, что существуют такие временные интервалы, в течение которых «медленные переменные» изменяются мало, а «быстрые переменные» претерпевают значительную эволюцию и достигают значений, соответствующих стационарным решениям уравнений для Y . Важно отметить, что те значения «медленных переменных», которые ощущаются на этом этапе «быстрыми переменными», соответствуют не окончательному стационарному уровню, а тем мгновенным значениям, которые предсказываются уравнениями для X . Таким образом, для «быстрых переменных» можно ввести «квазистационарное» соотношение вида

$$Y = h(X) \quad (3.22a)$$

По существу это соотношение играет роль «уравнения состояния». Подставляя это выражение в уравнения для «медленных переменных», в которых фигурируют как X , так и Y , получим

$$\frac{dX}{dt} = F(X, Y) = F(X, h(X)) \equiv f(X) \quad (3.22b)$$

Эта запись соответствует замкнутой системе уравнений для X . Таким образом, путем исключения «быстрых переменных» нам удалось значительно понизить число переменных. Такой подход аналогичен представлениям, лежащим в основе выделения «скоростьопределяющей стадии» и «квазистационарного приближения» в химии и, в частности, в кинетике ферментативных реакций. Если число оставшихся «медленных переменных» равно одному или двум, то можно воспользоваться результатами разд. 3.4 и 3.5 — они применимы к системам гораздо более общим, нежели рассмотренные там абстрактные математические модели. Конкретный пример такого рода будет рассмотрен в разд. 6.3.

Выше мы неоднократно обращались к интуитивным представлениям, однако все основные выводы в действительности можно строго доказать с помощью фундаментальной теоремы советского математика Тихонова. В ней явно сформулированы все условия, которым должны удовлетворять правые части кинетических уравнений и фигурирующие в них параметры.

Обсудим теперь другой, гораздо более тонкий механизм, позволяющий свести исходную систему со многими переменными

ми к нескольким переменным. Пусть нам дана система динамических уравнений вида (3.1), в которой число переменных n может быть очень велико. С помощью стандартного метода, известного под названием *линейного анализа устойчивости*, можно найти такие значения параметра λ , при переходе через которые асимптотически устойчивое стандартное состояние становится неустойчивым. Интересующиеся техническими подробностями могут найти в приложении I компактное изложение этого метода, а также иллюстрации на примерах простых моделей. Здесь же мы остановимся лишь на основных идеях.

Как обсуждалось в разд. 2.6, устойчивость по существу определяется ответом системы на возмущения или флуктуации, возникающие в некотором стандартном состоянии. Поэтому естественно представить динамические уравнения (3.1) в таком виде, чтобы возмущения фигурировали в них явно. Полагая

$$X_i(t) = X_{si} + x_i(t) \quad (3.23)$$

путем подстановки этого выражения в (3.1) и с учетом того, что X_{si} также является решением этих уравнений, получаем

$$\frac{dx_i}{dt} = F_i(\{X_{si} + x_i\}, \lambda) - F_i(\{X_{si}\}, \lambda)$$

Эти уравнения однородны в том смысле, что их правые части обращаются в нуль, если все $x_i = 0$. Чтобы сделать эту однородную систему более наглядной, разложим все $F_i(\{X_{si} + x_i\}, \lambda)$ вблизи $\{X_{si}\}$ и выделим из полученного результата линейные по $\{x_j\}$ члены, а также нелинейную поправку, структуру которой на данном этапе можно не уточнять:

$$\frac{dx_i}{dt} = \sum_j L_{ij}(\lambda) x_j + h_i(\{x_j\}, \lambda), \quad i = 1, \dots, n \quad (3.24)$$

Здесь L_{ij} — коэффициенты при линейных членах, а h_i — нелинейный вклад. Набор коэффициентов L_{ij} определяет некоторый *оператор* (в данном случае матрицу $n \times n$), зависящий от стандартного состояния X_s и параметров λ .

Основной результат теории, подробнее обсуждающийся в приложении I, состоит в том, что стандартное состояние $X = X_s$ (или $x = 0$) системы (3.24) асимптотически устойчиво (или неустойчиво) тогда и только тогда, когда устойчиво (или неустойчиво) решение линейного варианта этой системы:

$$\frac{dx_i}{dt} = \sum_j L_{ij}(\lambda) x_j, \quad i = 1, \dots, n \quad (3.25)$$

Заметим, что такая линейаризация может оказаться лишь в граничном случае, когда $X = X_s$ (или $x = 0$) устойчиво по Ляпунову,

но не асимптотически. Во всех же остальных случаях проблема устойчивости сводится к линейной задаче, которую можно решить элементарными методами теории дифференциальных уравнений.

На рис. 3.16 представлен типичный результат анализа устойчивости, выполненный в соответствии с такой процедурой. Основной вычисленной величиной является скорость роста возмущений γ в зависимости от одного (или нескольких) управляющего параметра (или параметров). Если $\gamma < 0$ (как это имеет мес-

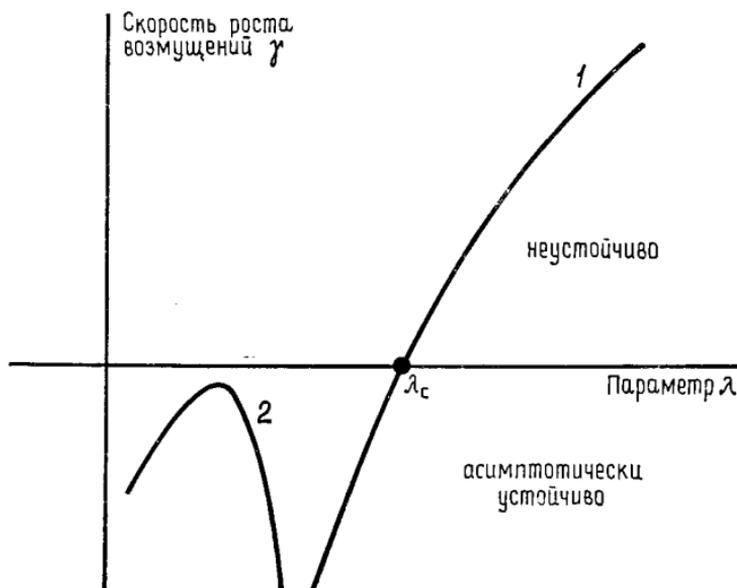


Рис. 3.16. Скорость роста возмущений γ как функция управляющего параметра λ , полученная путем линейного анализа устойчивости [см. уравнение (3.25)]. Кривая 1 — стандартное состояние асимптотически устойчиво при $\lambda < \lambda_c$ и становится неустойчивым при $\lambda > \lambda_c$, где λ_c — критическое значение управляющего параметра, соответствующее маргинальной устойчивости. Кривая 2 — стандартное состояние остается асимптотически устойчивым при всех значениях λ .

то на рис. 3.16, ветвь 1 при $\lambda < \lambda_c$), то стандартное состояние асимптотически устойчиво. Наоборот, при $\gamma > 0$ ($\lambda > \lambda_c$ на ветви 1, рис. 3.16) оно неустойчиво. При $\lambda = \lambda_c$ мы имеем состояние, характеризующееся *маргинальной устойчивостью*, которая представляет собой границу между асимптотической устойчивостью и неустойчивостью.

В общем случае система со многими переменными приводит к целому спектру значений γ подобно тому, как в кристалле имеется множество колебательных мод. Поэтому на рисунке

типа 3.16 возможно несколько кривых зависимости γ от λ . Допустим сначала, что лишь одна из этих кривых (кривая 1 на рис. 3.16) пересекает ось λ , а остальные кривые лежат ниже этой осн. Тогда можно показать, что в точке $\lambda = \lambda_c$ при вполне определенных и довольно мягких условиях, обсуждающихся подробнее в приложении II, *имеет место бифуркация с появлением новых ветвей решения*. При этом можно различать два случая.

1. Если при $\lambda = \lambda_c$ возмущения не имеют колебательного характера (как, например, в случае рис. 3.12, *a*), то возникшие в результате бифуркации ветви будут представлять собой стационарные решения.

2. Если при $\lambda = \lambda_c$ возмущения представляют собой колебания (см., например, рис. 3.13), то возникающие ветви будут соответствовать периодическим решениям в виде предельных циклов.

В любом случае, если параметры близки к их критическим значениям, для описания возникающих в результате бифуркации новых ветвей (или для описания переходного поведения) можно определить соответствующий набор величин, удовлетворяющий замкнутой системе уравнений. В случае 1 имеется лишь *одна* такая величина, служащая мерой амплитуды ветвей, образованных бифуркацией. С другой стороны, в случае 2 имеются *две* такие величины, которые характеризуют как амплитуду, так и фазу колебаний. Следовательно, исходная система динамических уравнений эффективно расцепляется на одно или два уравнения, из которых находится сама точка бифуркации, и систему из $n - 1$ или $n - 2$ уравнений для величин, напрямую не связанных с точкой бифуркации. Величины, удовлетворяющие уравнениям для точки бифуркации, называются *параметрами порядка*.

Удивительнее всего то, что бифуркационные уравнения, как оказывается, имеют совершенно такой же вид, как и в случае абстрактных математических моделей, приведших в разд. 3.4 и 3.5 к бифуркациям типа острия и бифуркациям Хопфа. Иными словами, любой тип динамики, удовлетворяющий условиям случаев 1 или 2, вблизи точки бифуркации можно выразить в универсальной, т. е. *нормальной*, форме. Этот сильный результат в действительности был известен еще Пуанкаре, однако лишь американскому математику Дж. Биркгофу удалось значительно развить его применительно к изучению бифуркаций в консервативных системах. В случае диссипативных систем идею сведения динамики к нескольким уравнениям для параметра порядка часто связывают с именем советского физика Л. Ландау, построившего теорию фазовых переходов. В математической литературе такое сведение часто называют процедурой Ляпунова — Шмидта, а сравнительно недавно ее начали называть *теорией центральных многообразий*. Тот, кто интересуется, смо-

жет найти компактный вывод бифуркационных уравнений в приложении II.

Возможны и более сложные ситуации, когда на рисунках типа 3.16 ось λ пересекается сразу несколькими ветвями. В итоге взаимодействие между ответвившимися решениями порождает явления *вторичной* или даже *третичной* бифуркации. В этом случае также справедливы сформулированные выше результаты в том смысле, что можно с уверенностью утверждать: та часть динамики, которая поставляет информацию о новых ветвях решения, реализуется в фазовом пространстве пониженной размерности. Однако явное построение нормальной формы уравнений представляет собой гораздо более сложную задачу.

Суммируя, можно утверждать, что обсуждавшиеся в разд. 3.4 и 3.5 статические и динамические бифуркации являются подлинными *прототипами сложного* в том смысле, что они имеют место в широких классах систем и при весьма различных условиях. В следующем разделе мы покажем, что, несмотря на свою общность, они составляют лишь часть удивительного репертуара типов поведения динамических систем.

3.7. Снова фазовое пространство: топологические многообразия и фракталы

В предыдущих разделах мы выяснили, что поведение всех систем, сводимых к одно- или двумерному фазовому пространству, жестко ограничено топологическими условиями, налагаемыми на движение в любом пространстве низкой размерности. В частности, аттракторами могут быть только неподвижные точки и предельные циклы.

Теперь нас будут интересовать качественные особенности поведения динамических систем в фазовых пространствах с размерностью выше двух. В конце разд. 3.1 мы уже отмечали, что в основе совершенно новых типов поведения лежит та свобода, которая привносится дополнительными размерностями. Помимо этих новых аспектов мы ожидаем, что в таких системах сохранятся все особенности одно- и двумерных систем, и в частности аттракторы типа неподвижных точек и предельных циклов. Естественно, что топологическая структура траекторий будет при этом более сложной, но если достаточно долго подождать, то в конце концов они начнут повторяться. Поэтому оставим пока эти вопросы и обратимся к совершенно новым возможностям.

Чтобы получить по возможности более ясную картину, проведем наши рассуждения для систем, динамика которых протекает в трехмерном фазовом пространстве. Разумеется, нам хотелось бы воспользоваться опытом, приобретенным при ана-

лизе двумерных задач. Попытаемся обрисовать наилучшую стратегию для этого. Одна из возможностей вытекает непосредственно из знаменитой пещерной истории Платона, изложенной в его «Государстве», — взгляните на тень от целого предмета и попытайтесь восстановить его по этой двумерной картине. Сам Платон полагал, что эта задача очень трудна, и лишь философия может помочь приблизиться к ее решению. Выражаясь более современным языком, одна частная проекция, конечно, не дает взаимно однозначного соответствия с предметом. Для полного описания этого предмета необходимо иметь несколько проекций на поверхности с различной ориентацией. Однако ясно, что в нашем случае подобные соображения автоматически приведут нас к исходному многомерному пространству, т. е. к тому, от чего мы хотели уйти.

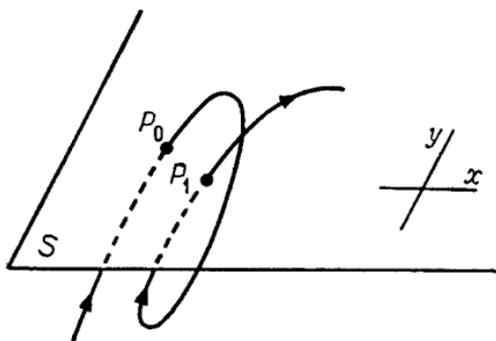


Рис. 3.17. Иллюстрация поверхности Пуанкаре. Изучаются пересечения траекторий динамической системы, непрерывно эволюционирующей во времени, с поверхностью S , трансверсальной к траекториям. Дискретная временная динамика, описывающая положение на S последовательных точек пересечения P_0, P_1, \dots , дает богатую информацию о соответствующем непрерывном потоке.

К счастью, остается открытой еще одна возможность, найденная другим гением — Анри Пуанкаре. Рассмотрим плоскость S , трансверсально пересекающую фазовые траектории (рис. 3.17), и изучим последовательность точек P_n пересечения ее с траекториями. С течением времени эти точки начнут сгущаться, и по тому, где это произойдет, мы сможем получить представление о трехмерном аттракторе, поскольку по существу в нашем распоряжении будет сечение такого аттрактора.

Плоскость S называется секущей поверхностью Пуанкаре, а характер расположения точек P_n соответствует так называемому отображению Пуанкаре. Заметим, что в таком представлении динамика системы выглядит как некоторое рекуррентное соотношение с дискретным временем, поскольку интервалы времени между последовательными пересечениями поверхности S траекторией конечны. Вводя на плоскости S соответствующую

систему координат, это рекуррентное соотношение можно записать в виде

$$x_{n+1} = f(x_n, y_n), \quad y_{n+1} = g(x_n, y_n) \quad (3.26)$$

Теперь, когда динамика системы уже сведена к двумерному случаю, можно воспользоваться нашими предыдущими результатами и сделать возможные выводы о динамике. Попытаемся дать предварительную классификацию различных возможностей, возникающих при таком подходе.

А. Периодические аттракторы: циклы порядка k

На рис. 3.18, *а* представлена одна из наиболее очевидных возможностей такого рода — точки пересечения P_0, P_1, \dots сходятся к неподвижной точке P . Очевидно, это означает, что аттрактор,

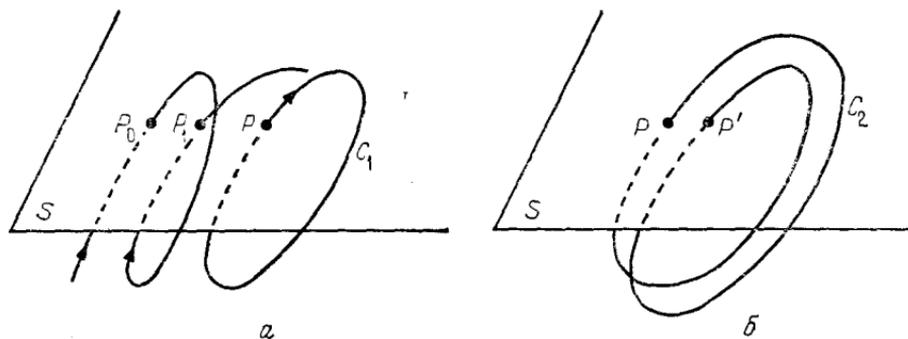


Рис. 3.18. Сходимость траекторий динамической системы в непрерывном времени к предельному циклу C_1 (*а*) и предельный цикл C_2 с периодом, в два раза превышающим период C_1 (*б*). На поверхности Пуанкаре сходимость к этим предельным циклам проявляется в виде последовательности P_0, P_1, \dots , сходящейся соответственно к неподвижной точке P (*а*) и паре неподвижных точек P, P' (*б*).

если он ограничен, представляет собой предельный цикл, C_1 . В рекуррентных соотношениях (3.26) существование точки P проявляется как возможность одновременного выполнения равенств $x_{n+1} = x_n, y_{n+1} = y_n$, или, что равносильно, $x_n = f(x_n, y_n), y_n = g(x_n, y_n)$.

Допускаются, однако, и другие возможности. Так, вместо только что описанного цикла «порядка один» может оказаться, что x_{n+1}, x_{n+2} и т. д. не совпадают с x_n (и то же самое происходит с переменными y) вплоть до k -й итерации, для которой такое совпадение будет достигнуто: $x_{n+k} = x_n, y_{n+k} = y_n$. В этом случае мы будем говорить о цикле порядка k . На рис. 3.18, *б* показано, как может выглядеть цикл порядка два. На поверхности Пуанкаре аттрактор представлен набором из двух точек (P и

P'), последовательно посещаемых системой, в то время как соответствующим трехмерным объектом является кривая C_2 , которая, перед тем как замкнуться, совершает два оборота. (В двумерном фазовом пространстве такое было бы невозможно, поскольку это означало бы наличие самопересечений на траекториях.) Очевидно, здесь речь идет о естественной модели для описания периодических явлений, когда происходит резкое изменение периода при изменении одного из параметров. Такие переходы в большом количестве наблюдались экспериментально в гидродинамике, химии и оптике. Подробнее они будут обсуждаться в разд. 3.10.

Отметим, что, как и в разд. 3.5, неподвижные точки отображения Пуанкаре также можно классифицировать как седловые точки, узлы и т. д.

Б. Квазипериодические аттракторы: инвариантные торы

Еще одна топологически допустимая возможность представлена на рис. 3.19, где точки пересечения P_0, P_1, \dots сходятся к замкнутой кривой C на поверхности Пуанкаре S . Поэтому соответствующим притягивающим объектом в трехмерном пространстве будет двумерная поверхность. Если из физических соображений потребовать, чтобы траектории не уходили на бесконечность, то такая поверхность будет ограниченной. Кроме того, структура этой поверхности должна допускать отсутствие точек пересечения траекторий в соответствии с описанной в разд. 3.1 теоремой единственности. Это условие исключает возможность существования поверхности типа сферы. Таким образом, остается лишь возможность существования поверхности типа *тора*, с которым мы уже встречались в разд. 3.3, где отмечалось, что движение по тороидальной поверхности в общем случае является квазипериодическим благодаря наличию несоизмеримых частот. Новым элементом здесь можно считать то, что тор притягивает траектории в отличие от обсуждавшегося в разд. 3.3 случая консервативных систем, где имелся континуум торов, окружающих периодическую орбиту.

Таким образом, притягивающий инвариантный тор является естественной моделью воспроизводимых квазипериодических явлений, наблюдаемых в природе.

В. Непериодические аттракторы: фракталы

Возможны ли в трехмерном фазовом пространстве аттракторы, отличные от неподвижных точек, замкнутых кривых или ограниченных двумерных поверхностей? Прежде всего заметим, что размерность множества, представляющего собой аттрактор,

обязательно меньше размерности самого фазового пространства, поскольку это множество состояний остается инвариантным в процессе движения. Поэтому поставленный вопрос сводится к следующему: существуют ли объекты с размерностью, промежуточной между точкой и линией, линией и поверхностью или даже поверхностью и объемом? Такие объекты, если они существуют, не будут ни точками, ни кривыми или поверхностями,

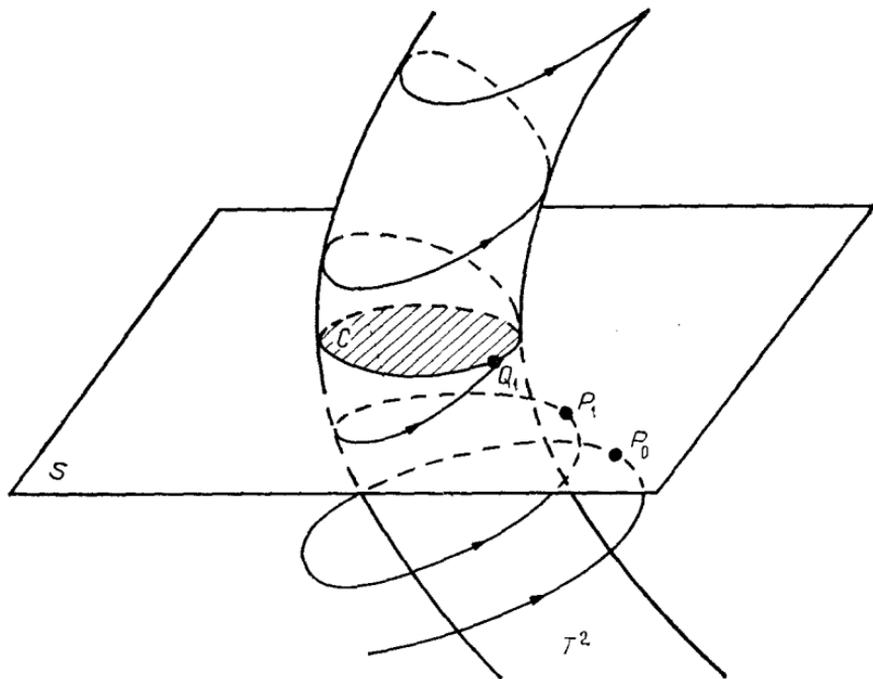


Рис. 3.19. Квазипериодический аттрактор в виде тороидальной поверхности T^2 , помещенной в трехмерное фазовое пространство. На переходном этапе, прежде чем сойтись к поверхности тора, траектория перескакает поверхность Пуанкаре S в точках P_0, P_1, \dots . В дальнейшем она бесконечно навивается на поверхность тора. При каждом прохождении через S траектория дает точку пересечения типа Q_1 . Множество всех таких точек образует кривую C , представляющую собой аттрактор на поверхности Пуанкаре. Последовательность переходных точек P_0, P_1, \dots сходится к этой кривой.

ни вообще топологическими многообразиями. Французский математик Бенуа Мандельброт ввел для них специальное название — *фракталы*.

Теория чисел, считающаяся одной из наиболее глубоко развитых областей математики, позволяет предвидеть подобные возможности, поскольку поставленный вопрос можно свести к следующему: существуют ли множества, промежуточные между счетными множествами (типа множества натуральных чи-

сел) и континуумом (типа, например, совокупности точек на прямой)? Ответ на этот вопрос будет положительным. Этот ответ, тесно связанный с понятием бесконечности, был найден немецким математиком Георгом Кантором. Мы не можем здесь вникать в тонкости математических доказательств, поэтому просто проиллюстрируем результат на конкретном примере *множества Кантора* (рис. 3.20).

Рассмотрим отрезок $[0, 1]$. Разделим его на три равные части и выбросим средний интервал. В каждом из двух оставшихся отрезков, которые мы тоже разделим на три равные части, снова выбросим средний интервал и т. д. По построению множество точек, полученных в пределе таким путем, является несчетным множеством изолированных точек. Оно не имеет собственной длины (как говорят математики, это *множество меры нуль*), поскольку длина его дополнения до отрезка $[0, 1]$ равна единице:

$$1/3 + 2/9 + 4/27 + \dots = 1$$

Однако удивительно, что размерность этого множества не равна нулю! Рассмотрим, как это может получиться.

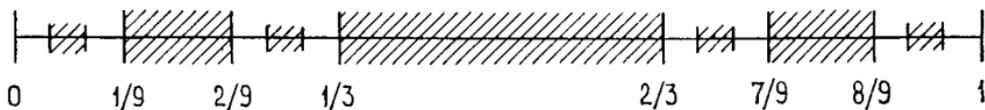


Рис. 3.20. Построение множества Кантора.

Каким образом мы формулируем то обстоятельство, что кусок бумаги, на котором мы пишем, является двумерным, или что пространство, в котором мы движемся, трехмерно? Обозначим посредством D любое интересующее нас множество и попытаемся закрыть его отрезками, или квадратиками, или вообще гиперкубами со стороной, равной ε (рис. 3.21). Заметим, что размерность d всех известных «топологических» многообразий вытекает из соотношения

$$d = \lim_{\varepsilon \rightarrow 0} \frac{\ln N_\varepsilon}{\ln (1/\varepsilon)} \quad (3.27)$$

Пояснить это соотношение можно следующим образом. Разобьем множество D на ячейки со стороной ε и сосчитаем минимальное число таких ячеек, чтобы покрыть все множество, разделим логарифм этого числа на логарифм $1/\varepsilon$, перейдем к пределу при очень малых ε и в результате получим искомую размерность.

Применим эти соображения к плоскости, точнее, к квадрату со стороной 1. Для того чтобы закрыть этот квадрат, нам нуж-

но иметь $(1/\varepsilon)^2$ квадратов со стороной ε . Следовательно, $d = [\ln(1/\varepsilon)^2 / \ln(1/\varepsilon)] = 2[\ln(1/\varepsilon) / \ln(1/\varepsilon)]$, т. е. $d=2$, как это и ожидалось. Но что же получается, когда мы распространяем эти представления на множество Кантора?

При первом разбиении, для того чтобы покрыть все имеющиеся подмножества, достаточно иметь 2 отрезка длиной $1/3$. При втором разбиении нужно уже 4 отрезка длиной $1/9$, и вообще, при n -м разбиении нужно иметь 2^n отрезков длиной $1/3^n$. Итак, $\ln N_\varepsilon / \ln \frac{1}{3} = \ln 2^n / \ln 3^n = (n \ln 2) / (n \ln 3)$, или

$$d_{\text{Cantor}} = \frac{\ln 2}{\ln 3} \approx 0,63 \quad (3.28)$$

Таким образом, в этом плане множество Кантора — промежуточное между точкой ($d=0$) и линией ($d=1$), т. е. оно является фракталом.

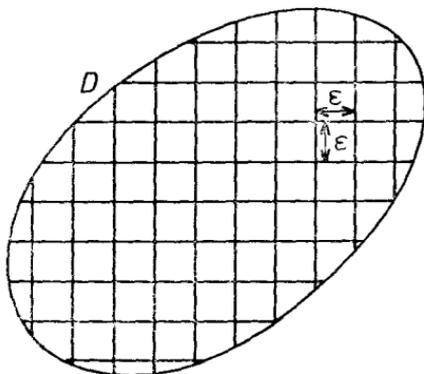


Рис. 3.21. Иллюстрация понятия размерности d множества D . Подсчитывается минимальное число гиперкубов размером ε , необходимое для полного покрытия данного множества, и затем вычисляется предел (3.27) при бесконечном убывании размера ε .

Установив существование подобных множеств, теперь мы можем вообразить такие аттракторы в трехмерном фазовом пространстве, которые отображаются на поверхности Пуанкаре в виде бесконечного несчетного множества исключительно извилистых отрезков, последние с течением времени посещаются траекторией невоспроизводимым образом. Соответствующим трехмерным объектом будет бесконечный набор листов, такой, что их трансверсальное сечение образует канторовское множество. Можно ожидать, что такие объекты, часто именуемые *странными аттракторами*, окажутся естественной моделью возникновения хаотического поведения, напоминающего турбулентность.

Мы здесь сосредоточились на аттракторах, которые могут существовать лишь в диссипативных системах. Однако ясно, что те же соображения можно привести и для описания инвариантных множеств в консервативных системах. До сих пор упомянутые выше возможности являлись всего лишь умозрительными заключениями, основанными на некоторых топологических соображениях. В трех последующих разделах мы постараемся убедиться в том, что эти заключения действительно соответствуют физической реальности.

3.8. Неинтегрируемые консервативные системы: новая механика

Рассматривавшиеся в разд. 3.3 интегрируемые консервативные системы в чистом виде в природе практически не встречаются. Обратимся на некоторое время к движению планет вокруг Солнца. Масса планет примерно в тысячу раз меньше массы Солнца, так что в первом приближении их взаимодействие можно рассматривать лишь как притяжение планет к Солнцу. Таким образом, мы приходим к задаче о двух телах — классическому примеру интегрируемой (и точно решаемой) системы. Анализ этой системы приводит к периодическому движению каждой планеты и к квазипериодическому движению планетной системы в целом.

Однако планеты всегда взаимодействуют, хотя и слабо, что оказывает *возмущающее влияние* на кеплеровские траектории. Весь вопрос в том, как изменится от этого движение — происходит ли слабое изменение траекторий или же их характер качественно изменяется.

Это основная проблема механики небесных тел. Аналогичные вопросы можно сформулировать и для ряда других механических систем.

Очевидно, влияние такого возмущения может быть значительным, если интересующие нас масштабы времени, грубо говоря, обратно пропорциональны интенсивности возмущения (для планетной системы это время составит около 1000 лет). Для того чтобы найти приближенное решение полной системы уравнений, достаточно к решению в «нулевом» приближении, соответствующему интегрируемой «стандартной» системе, прибавить малую поправку, пропорциональную интенсивности возмущения. Это делается стандартными методами. Однако если мы попытаемся найти решения, пригодные на очень больших временах, и, в частности, если мы захотим понять качественный характер точных решений, то перед нами встанет целый ряд непреодолимых трудностей. В рассматриваемом случае планетной системы такие трудности возникнут на временах порядка

миллиарда лет при попытке решения таких вопросов: покинут ли планеты Солнечную систему; упадут ли они на Солнце; столкнутся ли они между собой. Менее драматические, но технически столь же сложные вопросы возникают при анализе движения быстрых заряженных частиц в мощных магнитных полях типа тех, что имеются в кольцевых ловушках гигантских ускорителей. Такая же ситуация возникает при анализе орбитальной устойчивости космического корабля.

Упомянутые выше трудности проявляются главным образом в виде различных *расходимостей*. Это означает, что при вычислении традиционными методами теории возмущений поправки к движению интегрируемой системы выясняется, что при определенных условиях эта поправка становится очень большой или даже стремится к бесконечности. Чаще всего это происходит при возникновении *резонанса*, т. е. такого случая, когда собственные частоты интегрируемой стандартной системы соизмеримы в том смысле, что выполняется соотношение $k_1\omega_1 + k_2\omega_2 + \dots = 0$ (см. разд. 3.3). Такие катастрофы свидетельствуют о неприменимости используемого метода и служат четким указанием на то, что поведение системы при таких условиях должно качественно измениться.

Попытаемся, хотя бы интуитивно, разобраться в следующих двух вопросах. Во-первых, почему резонансы встречаются не как исключение, а скорее как правило? И во-вторых, чем они опасны?

Из теории чисел известно, что среди всех действительных чисел иррациональные числа образуют подавляющее большинство, поскольку, во-первых, они образуют несчетное множество (в отличие от рациональных чисел), а во-вторых, они заполняют всю «длину» или «меру» любого отрезка реальной линии. Это означает, что [см., например, (3.8)] в большинстве точек фазового пространства соответствующие действию переменные $\{I_j\}$ будут иметь такие значения, что собственные частоты системы $\{\omega_j\}$ будут связаны иррациональным образом, и по этой причине они будут несоизмеримы. Поэтому упомянутые выше трудности попросту не возникнут, и, как правило, движение будет происходить на *нерезонансных торах* (см. рис. 3.7, б), обусловленных иррациональностью связи между частотами. Однако не все так просто: хотя они и кажутся в некотором смысле «редкими», тем не менее рациональные числа всегда имеются в избытке (выражаясь математическим языком, множество рациональных чисел всюду плотно), поскольку в любой заданной окрестности любого действительного числа всегда существуют рациональные числа! Таким образом, любой, *сколь угодно малый объем* фазового пространства содержит бесконечное число изображающих точек, для которых выполняется усло-

вие резонанса. В частности, можно ожидать, что резонансы играют важную роль в эволюции гиббсова ансамбля (см. разд. 3.2) и тем самым в статистическом описании системы.

Рассмотрим теперь, почему резонансы могут оказаться «опасными». Обратимся снова к условию резонанса в связи со вторым из уравнений (3.8). Из них следует, что существует определенная комбинация угловых переменных $\varphi = k_1\varphi_1 + \dots + k_n\varphi_n$, остающаяся инвариантной во время движения и потому практически не участвующая в описании динамики системы. Иными словами, наш исходный выбор n независимых углов оказался неверным, поскольку $n-1$ из них являются на самом деле независимыми. Применяя более специальную терминологию, можно сказать, что n -мерный тор стягивается в $(n-1)$ -мерный. Та-

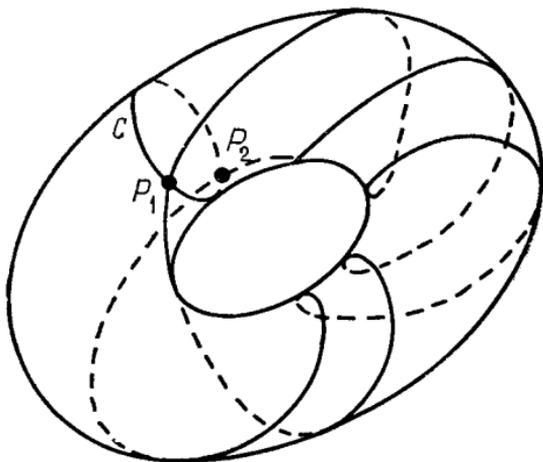


Рис. 3.22. Неподвижные точки P_1 и P_2 , порождаемые на секущей поверхности Пуанкаре S периодической траекторией, навивающейся на резонансный тор.

кие инвариантные поверхности более низкой размерности называются *резонансными торами*. В случае двух степеней свободы это будет означать, что инвариантная поверхность покрыта семейством замкнутых (одномерных) орбит, соответствующих периодическому (но не квазипериодическому) движению. Одна из этих орбит будет порождать на поверхности Пуанкаре S (рис. 3.22) ряд неподвижных точек (P_1 и P_2 на рис. 3.22), число которых равно числу оборотов, совершаемых орбитой вокруг тора.

Разумеется, любой другой член того же семейства породит другой набор из того же числа неподвижных точек на S . Таким образом, вся кривая C будет состоять из неподвижных точек! Трудно отделаться от ощущения, что такая ситуация представляет собой нечто исключительное, вроде трех точек на од-

ной и той же прямой. По этой причине можно ожидать, что наиболее типичные возмущения (неинтегрируемые!), воздействующие на систему, разрушат эту конфигурацию траекторий в фазовом пространстве. Это можно видеть из того факта, что поправка к невозмущенному движению оказывается столь же важной, как и само движение.

Поиск решения, позволяющего преодолеть эти трудности, связанные с неинтегрируемыми системами, привел к общей научной революции, в результате чего в начале 50-х годов возникла новая область в механике. Многие важные результаты, легшие в основу этой науки, были известны уже Пуанкаре и Биркгофу, однако главную роль в этом прорыве сыграл советский математик А. Колмогоров. Из числа других ученых впоследствии важный вклад был сделан В. Арнольдом и Дж. Мозером. Рассмотрим вкратце наиболее яркие положения новой теории.

А. Возмущение квазипериодических движений

Использованные ранее в настоящем разделе соображения в пользу того, что резонансные торы должны быть чувствительны к малым неинтегрируемым возмущениям, позволяют ожидать, что в нерезонансном случае движение должно оставаться «грубым». Иными словами, типичное квазипериодическое поведение системы не должно сильно изменяться слабым возмущением. Этот важный результат был строго доказан Колмогоровым; к тому же он сумел дать четкое описание тех видов квазипериодического движения, которые продолжают существовать после наложения возмущения. Точнее, было установлено, что, для того чтобы квазипериодическое движение продолжало существовать, связь между собственными частотами должна быть достаточно «иррациональной». В случае двух степеней свободы (рассматриваемом здесь для иллюстрации) это означает, что отношение частот ω_1/ω_2 должно достаточно отличаться от рационального значения:

$$\left| \frac{\omega_1}{\omega_2} - \frac{r}{s} \right| > \frac{c}{s^\nu} \quad (\text{для всех целых } r \text{ и } s) \quad (3.29)$$

Здесь c — малое число, обращающееся в нуль в отсутствие возмущения, а ν достаточно велико (например, $\nu=2,5$). Можно показать, что квазипериодические виды движения, допускаемые соотношением (3.29), составляют большинство видов движения, реализуемых в фазовом пространстве.

Короче говоря, следует ожидать, что в обширных областях фазового пространства неинтегрируемая система, близкая к не-

которой интегрируемой, должна вести себя примерно так, как показано на рис. 3.7, т. е. должно наблюдаться устойчивое движение вокруг периодических орбит.

Б. Возмущение периодических движений

В качестве удобного подхода к этой проблеме можно воспользоваться методом секущей поверхности Пуанкаре (рис. 3.22). Прежде всего целесообразно задаться вопросом о судьбе континуума неподвижных точек на S в связи с воздействием возмущения. Пусть число витков, совершаемых орбитами, равно s . Можно показать (краткое пояснение см. в приложении III), что под действием возмущения появляется $2ks$ неподвижных точек (k — определенное целое число), из которых ks точек являются эллиптическими, а ks — гиперболическими (седлами). Эллиптические неподвижные точки соответствуют устойчивым периодическим орбитам, каждая из которых окружена (нерезонансными) инвариантными торами, проявляющимися на отображении Пуанкаре в виде замкнутых кривых. Соответствующую часть фазового пространства часто называют «островом» (такой остров четко виден на рис. 3.23, область a).

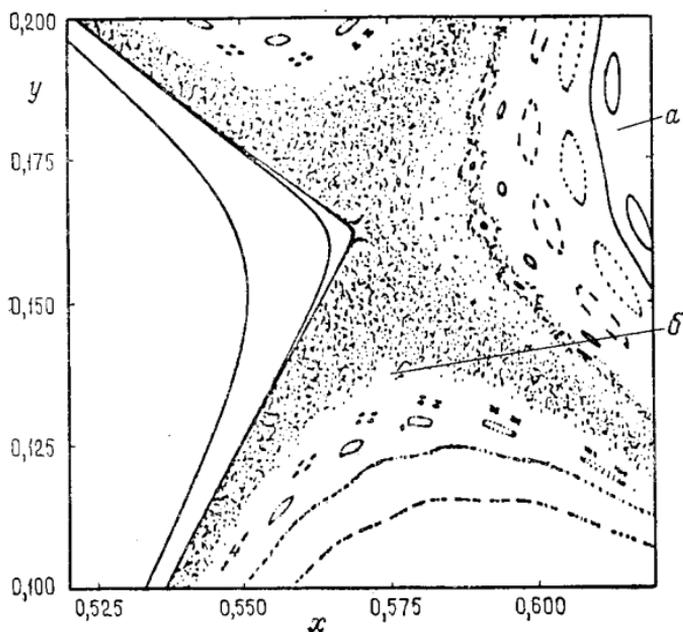


Рис. 3.23. Представление регулярных и стохастических траекторий неинтегрируемой гамильтоновой системы на поверхности Пуанкаре. Область a : устойчивое движение вблизи эллиптических неподвижных точек; область b : неустойчивое движение вблизи гиперболических неподвижных точек [из статьи: Hénon M., Q. Appl. Math., 27, 291 (1969)].

Совершенно иная ситуация возникает в случае гиперболических неподвижных точек. Вследствие неустойчивости соответствующих периодических орбит в окрестности таких точек развивается исключительно сложная структура, что по существу обусловлено особой конфигурацией их сепаратрис (рис. 3.23, область б). Точнее, на отображении Пуанкаре эти сепаратрисы изгибаются и пересекаются в бесконечном числе точек, известных под названием гомоклинических точек (см. приложение III). В этих областях возникает режим чрезвычайно *иррегулярного движения* — траектории очень чувствительны к малым изменениям начальных условий и замысловато бродят по широким областям фазового пространства. На отображении Пуанкаре это проявляется в виде облака, по всей видимости, случайно разбросанных точек, несмотря на то что соответствующие траектории, естественно, абсолютно детерминированы. Это хаотическое поведение указывает на возможную глубокую связь между неинтегрируемыми консервативными системами и стохастическим поведением, которое обычно связывают с необратимостью и диссипативными системами. Этот важнейший вопрос будет детально обсуждаться в гл. 5.

Отметим, что структура отображения Пуанкаре повторяется на всех масштабах вблизи всех эллиптических точек, порождаемых различными резонансными орбитами. Это замечательное свойство «самоподобия» показывает, насколько большая сложность может достигаться даже в таких скромных на вид задачах, как задача о двух слабо взаимодействующих степенях свободы!

Такое поведение исследовалось численными методами на многих простых моделях. В случае малых возмущений вычисления согласуются с общей картиной, представленной на рис. 3.23. Однако с ростом энергии системы наблюдается тенденция к отступлению островов и хаотические траектории заполняют все более обширные участки фазового пространства. Какие-либо качественные выводы из этих наблюдений сделать затруднительно, поскольку этот режим трудно поддается аналитическому описанию и, более того, нелегко провести какие-либо общие топологические рассуждения. И тем не менее удается ввести меру неустойчивости движения и чувствительности к начальным условиям. Это обеспечивается *показателями Ляпунова*, описывающими среднюю скорость расхождения двух изначально близких траекторий. Следовательно, ляпуновские показатели представляют собой обобщение идей линейной теории устойчивости, изложенных в разд. 3.6, на случай хаотических систем. Более подробное изложение этих вопросов можно найти в приложении I.

Сложная структура регулярных и иррегулярных траекторий,

обсуждавшихся выше, не является исключительно свойством неинтегрируемых гамильтоновых систем. Такое поведение типично и для других широких классов консервативных систем, среди которых важную роль играют так называемые колмогоровские системы (К-системы). Их обсуждение мы отложим до гл. 5, поскольку переход от консервативной к диссипативной динамике будет иллюстрироваться в основном на таких системах. Другим полезным для изучения динамических систем подходом служат отображения в дискретном времени, что обусловлено их тесной связью с динамическими системами в непрерывном времени, опосредованной поверхностью Пуанкаре [см. уравнение (3.26)]. Более того, за несколькими редкими исключениями, дискретные отображения — единственный известный до сих пор тип динамических систем, для которого удается дать явное и строгое доказательство возникновения иррегулярных видов движения, соответствующих хаотическому поведению. Известным примером такого рода является *отображение типа подковы*, которое можно представить как в виде абстрактной консервативной, так и в виде диссипативной динамической системы. Хотя ни одну из известных гамильтоновых систем не удастся явно свести к этому отображению, ниже мы приведем его описание и в дальнейшем воспользуемся им в качестве прототипа этого нового поведения, в том же смысле, в каком до сих пор при анализе более простых типов поведения в качестве прототипов использовались циклы или торы.

3.9. Модель неустойчивого движения: подкова

Рассмотрим отображение T плоскости на себя, действующее на квадрате R способом, показанным на рис. 3.24. Квадрат сжимается по вертикали и растягивается по горизонтали, а затем изгибается в виде подковы, так что изогнутая часть и концы «заготовки для подковы» выходят за пределы исходного квадрата. Таким образом, квадрат R не отображается сам на себя — в нем содержатся лишь принадлежащие подкове полосы I и J . Их прообразами (т. е. теми частями R , из которых эти полосы получаются в результате отображения) в действительности являются две вертикальные полосы K и L исходного прямоугольника.

Очевидно, инвариантное множество точек этого преобразования и обратного ему преобразования T^{-1} , обозначаемое Λ , содержится в обеих парах прямоугольников (I, J и K, L). Следовательно, эти точки являются подмножествами четырех малых квадратов $C_1—C_4$, показанных на рис. 3.24. Мы будем называть их *возвратными множествами*.

На рис. 3.25 показан результат повторного воздействия этого отображения на квадрат R , т. е. результат отображения $T^2 = T \cdot T$. В частности, на рисунке видны полосы T^2R , содержащиеся в исходном квадрате, и их прообразы, представляющие собой часть множества, порожденного обратным преобразованием T^{-2} квадрата R . Так получается возвратное множество из 16 квадратов, по четыре в каждом из квадратов $C_1—C_4$ (рис. 3.24).

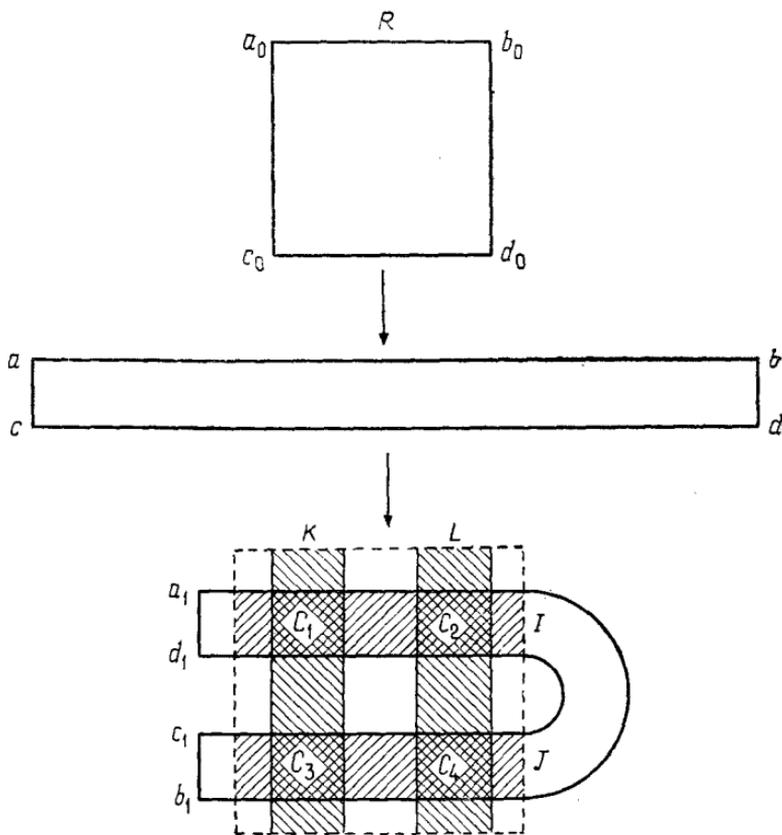


Рис. 3.24. Отображение «подкова». Начальный квадрат $a_0b_0c_0d_0$ растягивается по меньшей мере в два раза в горизонтальном направлении и сжимается более чем в два раза — в вертикальном. Полученная полоса сгибается в форме подковы. При этом преобразовании квадраты $C_1—C_4$ остаются инвариантными.

Совершив n итераций, можно видеть, что точки, не покидающие квадрата, расположены по малым квадратам, число которых равно 4^n , а сторона экспоненциально стремится к нулю по мере того, как n стремится к бесконечности. Результирующее множество Λ возвратных точек состоит из несчетного множества точек, суммарная площадь которых равна нулю. Следова-

тельно, это множество обладает свойствами множества Кантора (см. разд. 3.7), или, точнее, произведения двух канторовских множеств.

Детальный анализ динамики на множестве Λ можно выполнить лишь в рамках специального метода — *символической динамики*, основные черты которой будут описаны в дальнейшем. Таким образом удастся показать, что в любой окрестности любой точки Λ содержатся периодические орбиты типа

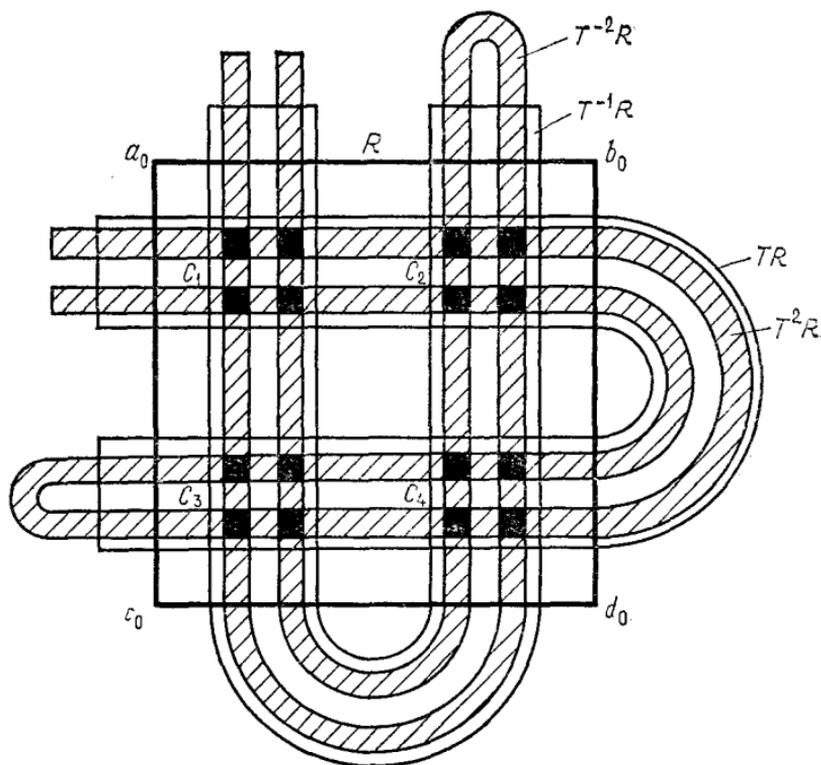


Рис. 3.25. Преобразование «подкова» после двукратной итерации. Иллюстрируется образование множества Λ рекуррентных точек при последующей фрагментации квадратов C_1 — C_4 .

рассмотренных в разд. 3.7А. Кроме того, существуют непериодические орбиты, подходящие сколь угодно близко к любой точке Λ . Заметим, что каждая точка Λ по построению является гиперболической, поскольку она лежит на пересечении направлений растяжения (горизонталь) и сжатия (вертикаль). Таким образом, две орбиты, начинающиеся в двух соседних точках Λ , будут экспоненциально расходиться в процессе последовательных итераций. Это означает чрезвычайную чувствительность к начальным условиям — одно из основных свойств ирре-

гулярных траекторий, встречавшихся в предыдущем разделе в связи с неинтегрируемыми гамильтоновыми системами. Это же свойство присуще и странным аттракторам, обсуждаемым в следующем разделе.

В действительности отмеченная аналогия простирается гораздо дальше: можно показать, что, когда сепаратрисы неустойчивых точек на отображении Пуанкаре изгибаются и пересекаются, как на рис. 3.23, исходное движение порождает на поверхности Пуанкаре динамику, сходную с той, что мы наблюдали в случае подковы. Это обстоятельство оправдывает использование подковы как прототипа иррегулярного движения и соответственно хаотического поведения.

3.10. Диссипативные системы в многомерных фазовых пространствах. Хаос и странные аттракторы

Рассмотрим теперь, что же скрывается за различными возможными типами поведения диссипативных систем, обсуждавшимися в разд. 3.7—3.9. Как мы видели, в случае консервативных систем возможно сосуществование регулярных и хаотических фазовых траекторий. Однако оценить их относительный вес трудно, особенно когда отклонение от интегрируемой системы уже не мало. С другой стороны, в случае диссипативных систем наличие асимптотической устойчивости подразумевает возможность *притягивающего хаоса*. В системе, характеризуемой таким свойством, все траектории, выходящие из определенной, конечной части фазового пространства, будут раньше или позже стремиться к аттрактору. Иными словами, появление хаоса в таких системах будет закономерно, причем он должен определенным образом проявляться в наблюдаемом поведении системы. Поэтому основной вопрос, который следует задать в связи с хаотическим поведением диссипативных систем, состоит в следующем: действительно ли в этих системах возникают хаотические аттракторы при изменении некоторых из управляющих параметров?

Возникновение неперiodического, сходного с турбулентным, поведения во многих из рассмотренных в гл. 1 экспериментов позволяет предположить, что ответ на этот вопрос должен быть утвердительным. Поэтому очень важно найти модели, объясняющие эти наблюдения, а также отыскать минимально необходимые условия возникновения хаотического поведения. Существовавшие до 60-х годов представления не позволяли понять, каким образом возникает турбулентность, по крайней мере в жидкости. Причина этого состояла в том, что турбулентность считалась проявлением динамики *бесконечного числа* сцеплен-

ных мод, каждой из которых приписывалась определенная частота. Более того, для расщепления всех этих мод представлялось необходимым бесконечное число переходов, происходящих выше точки потери устойчивости регулярного стационарного (ламинарного) течения. Поэтому, когда в дальнейшем ученые убедились в том, что в динамической системе лишь с несколькими переменными поведение может быть хаотическим уже после нескольких бифуркаций по мере удаления от стандартного состояния, это сильно удивило и возбудило их.

Первый пример такого рода был открыт в 1963 г. американским метеорологом Э. Лоренцем. Его модель дает упрощенное описание задачи о тепловой конвекции, рассматривавшейся в разд. 1.3, поскольку в ней удерживались лишь три моды, каждая из которых была связана с конкретной длиной волны в пространстве. Она приводит к аттрактору, наделенному всеми основными чертами турбулентной конвекции. Такое явление обычно для земной атмосферы, что создает определенные трудности в предсказании погоды. Более того, отсюда следует, что погода и климат в своей основе непредсказуемы, поскольку им присуще основное свойство хаотической динамики — существенная зависимость от начальных условий. Чтобы подчеркнуть это важнейшее положение, часто говорят — и вряд ли это слишком большое преувеличение, — что полет мухи в Кембридже (шт. Массачусетс) может привести к общему климатическому изменению в Индии!

Работа Лоренца оставалась незамеченной до 70-х годов, однако с тех пор в области хаотической динамики произошел настоящий взрыв, приведший к новым результатам. Идеи произросли одновременно на базе многих дисциплин, из которых отметим чистую математику, широкие модельные исследования с помощью компьютерной техники, теоретическую и экспериментальную физику и химию, а также популяционную биологию. Они привели к интересным «прототипическим» уравнениям и далее к хаотическому поведению, а также к ряду «сценариев» достижения хаотического аттрактора при изменении параметров. Мы приведем несколько выборочных иллюстраций по каждому из этих аспектов.

А. Некоторые модели-прототипы, приводящие к хаотическому поведению

Очень интересная математическая модель хаотического поведения была предложена О. Рёсслером. В ней имеются три переменные и лишь одна квадратичная нелинейность:

$$\frac{dx}{dt} = -y - z$$

$$\frac{dy}{dt} = x + ay \quad (3.30)$$

$$\frac{dz}{dt} = bx - cz + xz$$

Здесь a , b и c — положительные постоянные. Эти динамические уравнения имеют две неподвижные точки, одна из которых соответствует тривиальному решению $x_s = y_s = z_s = 0$, а другая — точке $x_s = c - ab$, $y_s = b - c/a$, $z_s = c/a - b$. В дальнейшем мы обсудим лишь явления, происходящие вблизи первой точки, которую обозначим P_0 .

Как показывает линейный анализ устойчивости, рассматривавшийся в разд. 3.6 и приложении I, в широком диапазоне значений параметров точка P_0 неустойчива. В ее окрестности поведение системы имеет следующие особенности: траектории отталкиваются от P_0 вдоль двумерной поверхности в фазовом пространстве, в которой неподвижная точка выглядит как неустойчивый фокус, и притягиваются вдоль одномерной кривой подобно тому, как это показано на рис. 3.29. Такая неподвижная точка называется *седло — фокус*. Такая конфигурация приводит к неустойчивости движения — основной особенности хаотического поведения. Одновременно не исключается и *возврат* неустойчивых траекторий в окрестность P_0 , что в конечном счете и обуславливает формирование устойчивого аттрактора.

На рис. 3.26, *а, б* показаны хаотические аттракторы, достигаемые при двух различных значениях параметров. На них видны обе только что отмеченные особенности. Однако в случае *а* все траектории возвращаются к P_0 с одной стороны плоскости xz , в то время как в случае *б* некоторые траектории возвращаются и с другой стороны. Эти два вида движения мы называем *спиральным* и *винтовым* типами хаоса соответственно.

Другая важная особенность, четко прослеживаемая на обоих рис. 3.26, *а, б*, — образование *складки* на поверхности, вдоль которой происходит неустойчивое движение. Это видно на поверхности Пуанкаре, полученной трансверсальным сечением фазового потока плоскостью $y=0$, $x < 0$, $z < 1$ (рис. 3.27). В случае рис. 3.27, *а* видно, что расхождение траекторий в неустойчивом направлении — x и дальнейшее их складывание приводит к отображению типа подковы на плоскости Пуанкаре. С учетом результатов разд. 3.9 это дает качественное объяснение случайного, непериодического характера траекторий на этом аттракторе.

На рис. 3.27, *б* приводится другое представление этого явления. Здесь изображена зависимость $(n+1)$ -й точки пересечения определенной выше поверхности Пуанкаре с потоком от положения n -й точки пересечения. Численное построение этой

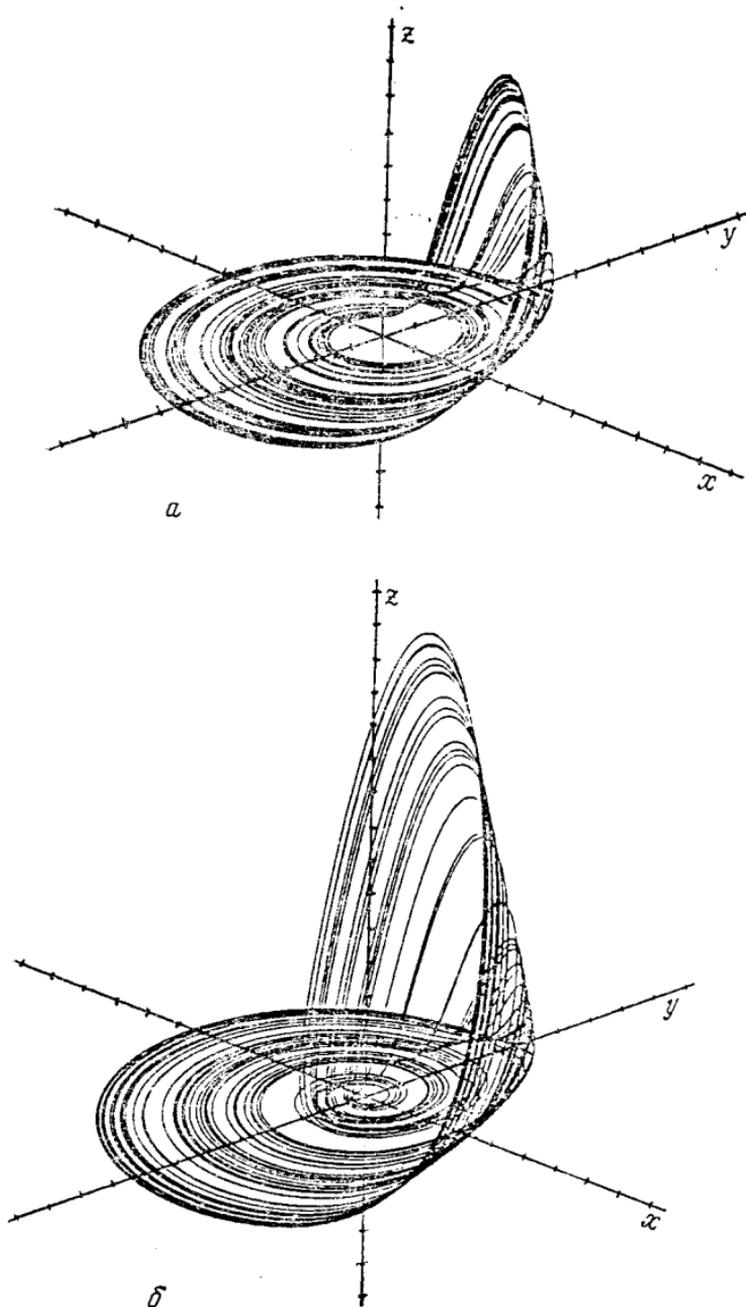
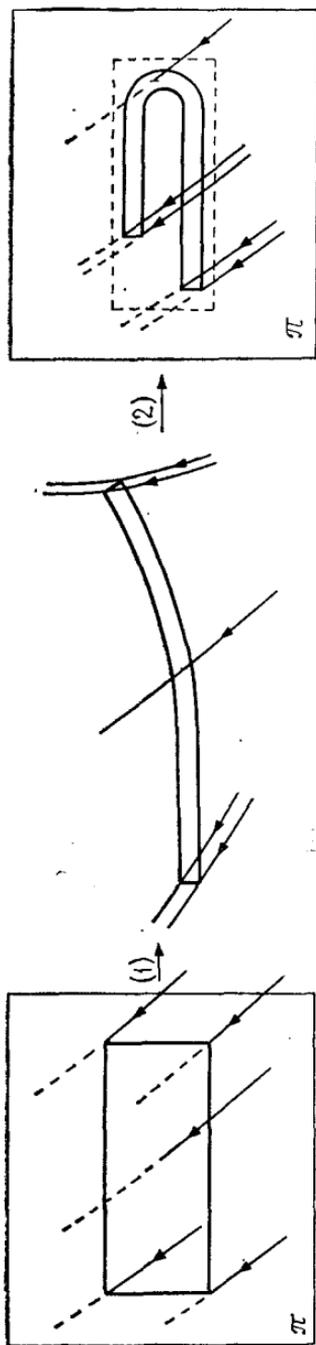


Рис. 3.26. Хаотические аттракторы, полученные численным интегрированием модели Рёсслера [уравнения (3.30)]. *a* — спиральный хаос при $a=0,32$, $b=0,3$, $c=4,5$. Траектории начинаются по одну и ту же сторону от неустойчивой неподвижной точки; *б* — винтовой хаос при $a=0,38$, $b=0,3$, $c=4,5$. Траектории начинаются по обе стороны от неустойчивой неподвижной точки.



a

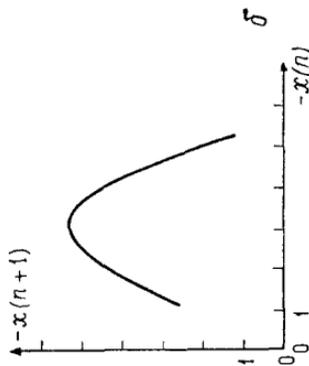


Рис. 3.27. *a* — отображение «подкова», индуцированное на секущей поверхности Пуанкаре динамической системой (3.30). Показанный на рис. 3.26, *a* поток пересекается плоскостью $y=0$, $x < 0$, $z < 1$. Траектории, выходящие из маленького прямоугольника в этой плоскости, вращаются вокруг одномерного устойчивого многообразия неподвижных точек, и каждый раз, отследив складку неустойчивого многообразия, пересекаются с поверхностью Пуанкаре; *b* — одномерное отображение, полученное путем построения зависимости $-x$, найденного при $(n+1)$ -м пересечении траектории с поверхностью Пуанкаре, от такого же значения при n -м пересечении.

зависимости приводит к плавной колоколообразной кривой. Заметим, что положения последовательных точек пересечения на этой кривой не соответствуют последовательным точкам этой кривой, а распределены как бы случайно.

Рис. 3.27, б устанавливает естественную связь между рассматриваемой моделью в непрерывном времени [уравнение (3.30)] и другим весьма важным классом моделей, демонстрирующих хаотическое поведение. Это модели с дискретным временем, определяемые рекуррентным соотношением

$$x_{n+1} = f(x_n) \quad (3.31a)$$

Типичная колоколообразная кривая, аналогичная представленной на рис. 3.27, б, получается при следующем выборе функции f :

$$f = 1 - \lambda x^2 \quad (3.31б)$$

Помимо своей связи с системами в непрерывном времени через отображение Пуанкаре системы этого типа могут также непосредственно моделировать некоторые ситуации из области биологии, когда рассматриваются популяции, в которых возрастные группы особей определяются возрастными группами особей предшествующих поколений.

Несмотря на свою кажущуюся простоту, уравнения (3.31) поражают невообразимым многообразием типов поведения, варьирующего от простых точек равновесия до множественных периодических или хаотических решений. Кроме того, из этих уравнений проистекает большая часть информации о «сценариях» становления хаотического поведения. В заключение этого обзора отметим, что имеется довольно много экспериментальных примеров хаотических аттракторов, которые при использовании подходящей секущей поверхности Пуанкаре приводят к дискретной временной динамике, аналогичной представленному на рис. 3.27, б примеру. Яркий пример приводится на рис. 3.28, построенном по экспериментам с реакцией Белоусова — Жаботинского в условиях полного перемешивания.

Б. Некоторые «сценарии» становления хаотического поведения

В отличие от статических бифуркаций и бифуркаций типа предельного цикла, обсуждавшихся в разд. 3.4 и 3.5, механизмы возникновения хаотических аттракторов невозможно представить в некоторой универсальной, канонической форме, которая позволила бы прояснить роль соответствующих параметров. Вместо этого удается лишь выявить некоторые конкретные пути установления хаоса на основе закономерностей, получен-

ных объединением некоторых общих рассуждений и опыта, приобретенного на многочисленных примерах моделирования. Не ясно, является ли это неизбежным следствием самой природы хаоса, или же это лишь временное затруднение, обусловленное сложностью задачи. Ситуация здесь сильно напоминает ту, которая имела место в науках о живой природе до появления биохимии и молекулярной биологии, когда основным подходом была таксономия — искусство классифицирования в группы и семейства.

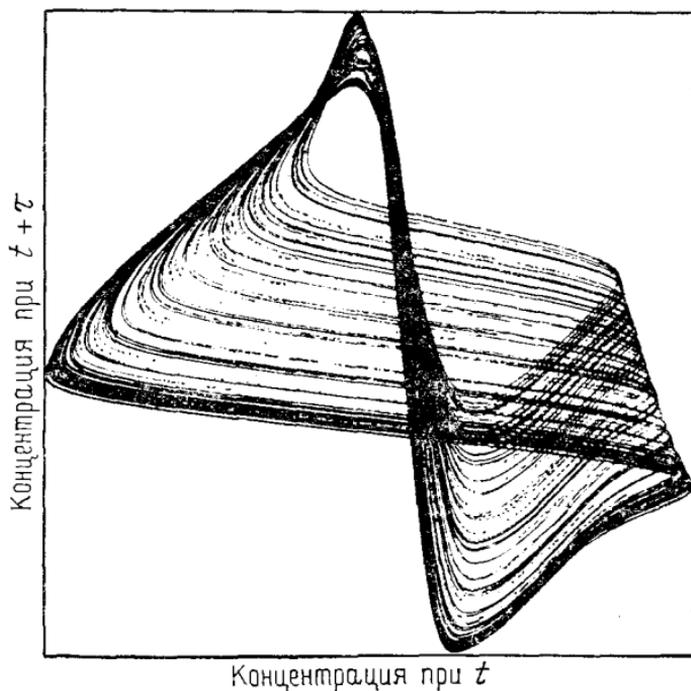


Рис. 3.28. Хаотический аттрактор реакции Белоусова — Жаботинского в открытой системе. Измеряемый сигнал пропорционален логарифму концентрации иона Br^- . По осям отложены значения сигнала в момент времени t и в момент времени $t + \tau$ при разных значениях t . Значение временной задержки для этого рисунка составляет $\tau = 53$ с [из работы: Roux J. C., Simoyi R., Swinney H., *Physica*, 8D, 257 (1983)].

Важный результат, определяющий некий частный «сценарий» установления хаоса, был получен Рюэлем и Такенсом. Как мы видели в разд. 3.7, одним из аттракторов, которые становятся возможными в фазовых пространствах с размерностью не ниже трех, является тор, соответствующий квазипериодическому движению. Такие торы часто возникают из периодических орбит в результате бифуркации второго или более высокого порядка. Рюэль и Такенс показали, что двумерный тор являет-

ся «грубым» в смысле структурной устойчивости, определенной в разд. 3.4. Иными словами, при наложении малого возмущения в виде изменения параметров в общем случае такой аттрактор продолжает существовать. Однако в случае торов с размерностью три и более это не так при наличии по меньшей мере трех несоизмеримых частот. Это хрупкие объекты, которые, исчезая под действием малых возмущений, приводят к структурно-устойчивым странным аттракторам, проявляющим свойства, обсуждавшиеся после разд. 3.7. Такие переходы наблюдались в модельных системах и в эксперименте, главным образом в механике жидких сред.

Второй «сценарий», приводящий к более количественным предсказаниям, возникает из изучения отображений в дискретном времени, определенных уравнениями (3.31). Он соответствует бесконечной последовательности бифуркаций при определенных значениях параметра λ : $\lambda_1 < \lambda_2 < \dots < \lambda_n < \dots$, каждое из которых приводит к циклам более высокого порядка с периодом, удваивающимся при каждой последовательной бифуркации. Значения λ_n скапливаются возле некоторого особого значения λ_∞ , после чего получаются орбиты с «бесконечным периодом», т. е. с ярко выраженным хаотическим поведением. В конечном счете все пространство состояний этой дискретной динамической системы (за исключением некоторого множества меры нуль) оказывается принадлежащим единственному хаотическому аттрактору, характеризующему неустойчивостью движения и чувствительностью к начальным условиям.

Важнейший результат, согласно М. Фейгенбауму, состоит в существовании закономерностей приближения к области хаотического поведения. Этот результат устанавливает, что приближение к λ_∞ происходит в соответствии с законами геометрической прогрессии, причем отношение предыдущего шага к последующему составляет $\delta \approx 4,669$. Такое поведение зависит лишь от того, как устроена функция $f(x)$ в окрестности максимума. Поэтому оно структурно-устойчиво и имеет место при совершенно тех же количественных характеристиках для всех семейств отображений, включая простую параболу типа (3.31б), а также синусоиду, окружность, эллипс и т. д. Это называется *метрической универсальностью*. Описанная ситуация иллюстрируется табл. 3.1. Читателю настоятельно рекомендуется поиграть в «экспериментальную математику» и самостоятельно проверить эту таблицу. Для этого достаточно выполнить последовательные итерации, соответствующие уравнениям (3.31), на карманном калькуляторе!

Многие эксперименты, рассматривавшиеся в гл. 1, выявляют существование каскадов с удвоением периода, в конечном счете приводящих к хаотическому поведению и подтверждаю-

Таблица 3.1. Каскад периодического удвоения в преобразовании
 $x_{n+1} = 1 - \lambda x_n^2$

| Период орбиты, n | Точка бифуркации | $\lambda_n - \lambda_{n-1}$ | $\delta = \frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n}$ |
|--------------------|------------------|-----------------------------|--|
| 0 | 0,75 | — | — |
| 1 | 1,25 | 0,5 | 4,233738275 |
| 2 | 1,3680989394 | 0,1180989394 | 4,551506949 |
| 3 | 1,3940461566 | 0,0259472172 | 4,645807493 |
| 4 | 1,3996312384 | 0,0055850823 | 4,663938185 |
| 5 | 1,4008287424 | 0,0011975035 | 4,668103672 |
| · | · | · | · |
| · | · | · | · |
| · | · | · | · |
| ∞ | 1,4011552... | — | 4,669... |

(непериодическое поведение)

щих теоретическое значение отношения δ , характеризующего скорость приближения к λ_∞ . Поэтому несомненно, что главный прорыв в области нелинейной динамики связан именно с открытием этой закономерности.

Еще один «сценарий» хаотического поведения в отображениях с дискретным временем, получивший известность под названием пульсация Пома — Манвиля (Pomeau — Manneville intermittency), возникает при потере устойчивости периодического решения. Возникающие при этом траектории характеризуются чередованием участков с практически периодическим поведением и участков с иррегулярным поведением в виде «турбулентных» всплесков, длительность которых также случайна. Такая картина наблюдалась и в большинстве экспериментов, обсуждавшихся в гл. 1.

Что касается непрерывного времени, то можно отметить, что возвращение ушедшей из окрестности неустойчивой неподвижной точки типа седло — фокус (см. рис. 3.26) часто связано с появлением орбит довольно редкого типа, называемых *гомоклическими*. Это те траектории, которые покидают неподвижную точку, но затем возвращаются к ней. Иными словами, эти траектории имеют один и тот же предел при $t \rightarrow +\infty$ и $t \rightarrow -\infty$ (рис. 3.29). Они очень чувствительны к изменению значений параметров и, вообще говоря, отсутствуют, если параметры не удовлетворяют некоторому строгому равенству (по терминологии разд. 3.4, они структурно-неустойчивы). Если это равенство нарушено слабо, то вместо таких траекторий возникает богатая структура орбит в фазовом пространстве, причем некоторые из этих орбит ведут себя хаотически. Имеются аргументы

в пользу того, что некоторые из вышеупомянутых путей к хаосу сопряжены с возникновением гомоклинических орбит. Таким образом, «гомоклинический сценарий», по всей вероятности, играет важную роль в хаотической динамике.

В заключение этого раздела сформулируем еще два вопроса. Во-первых, можно ли что-то сказать об относительной важности хаотического и регулярного поведения в диссипативных системах со многими переменными? Полученный при изучении модельных и реальных систем опыт подсказывает, что оба этих типа поведения возможны не только по отдельности, но и одновременно. Поэтому не приходится говорить о том, что системы

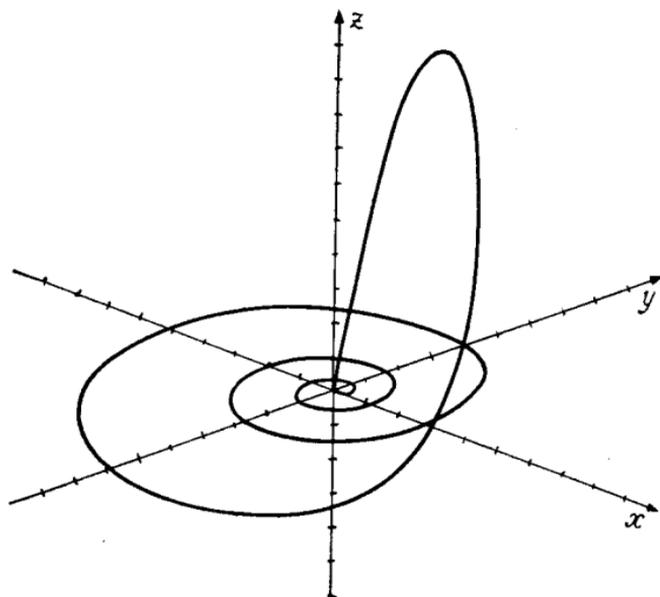


Рис. 3.29. Гомоклиническая орбита, связанная с неподвижной точкой $(0, 0, 0)$ модели Рёсслера, найденная при следующих значениях параметров: $a=0,38$, $b=0,30$, $c \simeq 4,82$.

со многими переменными проявляют тенденцию к соскальзыванию в хаотическое состояние сразу же после первой бифуркации. Кроме того, для определенных типов кинетики задача установления области хаотического поведения в пространстве параметров оказывается довольно сложной. Видимо, не случайно характерным примером служат биохимические регуляторные процессы, поскольку воспроизводимость должна быть существенной чертой биологических ритмов. В этом плане полезно воспроизвести сделанные в разд. 2.4 замечания относительно того, что следует считать «общим» или «типичным». Поведение физических систем отнюдь не обязательно отражается случайно выбранными абстрактными моделями. Особенно четко это

видно в биологии, где каждая функция живой клетки есть результат длительной эволюции. Здесь мы имеем другую иллюстрацию специфичности неравновесных систем в отличие от равновесных. В самом деле, если понижать температуру любого материала в условиях равновесия, то в конце концов мы достигнем точки фазового перехода. Однако если мы будем изменять значение параметра, влияющего на поведение неравновесной системы, то можно будет наблюдать предельные циклы, стационарные состояния, аналогичные равновесию, хаос или что-нибудь еще. Можно наблюдать переходы от теплового беспорядка типа равновесного к крупномасштабной упорядоченности и далее к особому состоянию мелкомасштабной разупорядоченности и крупномасштабной упорядоченности, представляющему собой хаос. Однако можно также оказаться свидетелями эволюции в обратном направлении и даже вообще пропустить стадию хаоса.

Второй вопрос (позднее мы еще вернемся к нему) касается роли хаоса в природе. Мы уже отмечали тот факт, что хаотические аттракторы позволяют интерпретировать некоторые экспериментальные результаты. Кроме этого успешного сопоставления, хаос открывает нам целый мир новых форм и картин. Выясняется, что разупорядоченность в определенном диапазоне отлично совмещается с упорядоченностью в другом диапазоне, как это следует из самого факта существования аттрактора. Изучение хаоса показывает также, что случайность не является следствием несовершенства эксперимента или сложности внешней среды, которой мы не можем управлять, — она лежит в самой основе динамики идеально детерминистических систем с несколькими переменными. Помимо прочего, сосуществование случайности и упорядоченности приводит к новому, до сих пор не обсуждавшемуся понятию *информации*. Этого нового пришельца в мир сложного мы вкратце обсудим в разд. 3.12 и более подробно в гл. 4.

3.11. Пространственно распределенные системы. Бифуркации с нарушением симметрии и морфогенез

До сих пор в этой главе рассматривались динамические системы с конечным числом переменных. Обратимся теперь к пространственно распределенным системам. Здесь переменные состояния (скажем, температура в пластине при систематическом нагревании одной из сторон) представляют собой поля в том смысле, что имеется непрерывная зависимость их от пространственных координат. Следовательно, это означает, что в задаче имеется бесконечное множество переменных. Вообще говоря, эти переменные будут сцепленными, поскольку пространствен-

ные неоднородности приведут к транспортным явлениям типа теплопроводности, диффузии или конвекции. Математически это обстоятельство выражается в присутствии *пространственных производных* в динамических уравнениях, таких, как лапласиан температуры или концентрации $\nabla^2 T(\mathbf{r}, t)$ или $\nabla^2 c(\mathbf{r}, t)$, градиент скорости $\nabla v(\mathbf{r}, t)$ и т. д. Можно ожидать, что работать с такими системами будет гораздо труднее. Однако во многих случаях удастся получить интересные результаты путем применения методов теории возмущений, позволяющих свести задачу к конечному числу переменных.

Этот механизм очень похож на обсуждавшийся в разд. 3.6 в связи со сведением динамики системы со многими переменными к нормальному виду. При этом исходным пунктом опять служит линейный анализ устойчивости (см. приложение I). В присутствии пространственных неоднородностей в оператор L из уравнения (3.25) войдут пространственные производные. Однако во многих случаях эти производные появляются аддитивным образом [см., например, уравнение (I.14)]. Тогда решение ищется в виде произведения функции времени $c(t)$ и функции пространственных координат $u(\mathbf{r})$ [см. уравнение (I.19)]. Это позволяет разделить переменные и получить две более простые системы уравнений, содержащие лишь $u(\mathbf{r})$ и $c(t)$. Часто удастся найти точное решение для $u(\mathbf{r})$. Например, в одномерной химической системе с реакцией и диффузией функция $u(\mathbf{r})$ представляется суперпозицией синуса и косинуса — собственных функций лапласиана в рассматриваемой пространственной области. После этого с зависящей от времени функцией $c(t)$ можно работать в точности так же, как и с уравнением (3.25), получая в итоге аналогичный показанному на рис. 3.16 график, описывающий зависимость скорости роста возмущения от управляющего параметра λ . Однако теперь в игру вступает новый элемент: некоторые из коэффициентов $L_{ij}(\lambda)$ содержат «следы» пространственных неоднородностей, возникающие в результате действия операторов пространственных производных на функцию $u(\mathbf{r})$. Эти следы выражаются в наличии собственного значения пространственно зависимой части оператора L . Оно зависит от транспортных коэффициентов и характерной пространственной длины данного режима, описываемого функцией $u(\mathbf{r})$. В системе с реакцией и диффузией это будет просто произведение $-D_i k^2$, где D_i — коэффициент диффузии i -го компонента и $k = 2\pi/l$, а l — характерная пространственная длина [ср. с равенством (I.20)].

Имея эту дополнительную свободу, можно задать следующий вопрос. Допустим, что ось λ пересекается лишь одной из кривых зависимости γ от λ (рис. 3.16). В маргинальном состоянии $\lambda = \lambda_c$ характеристики системы будут все еще зависеть от

дополнительного параметра k . Можно ли установить наиболее «опасное» значение k , т. е. такое значение, которое будет определять переход к неустойчивости? Чтобы ответить на этот вопрос, полезно построить зависимость критического значения параметра λ_c от k . На рис. 3.30 показаны два типичных результата, построенные по уравнениям (I.23а) и (I.23б) соответственно.

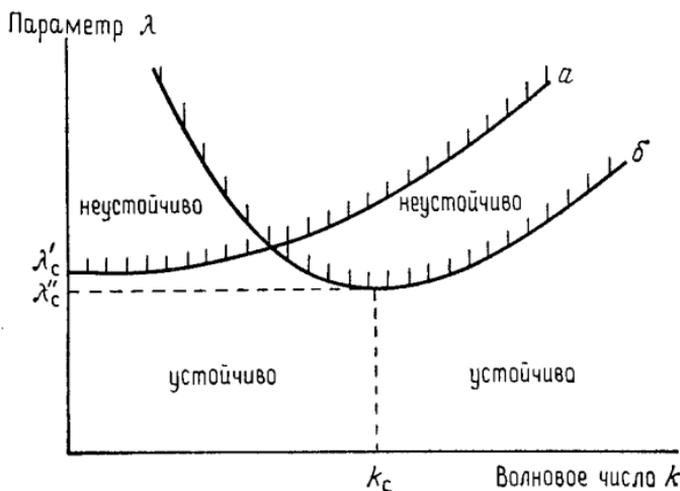


Рис. 3.30. Два типичных исхода линейного анализа устойчивости однородного стационарного состояния пространственно распределенной системы. a — кривая маргинальной устойчивости имеет экстремум при $k=0$. На этом основании можно ожидать, что доминирующая мода выше точки неустойчивости будет оставаться пространственно однородной; b — кривая маргинальной устойчивости имеет экстремум при $k=k_c \neq 0$. Возникающие выше точки бифуркации ветви могут характеризоваться нетривиальной пространственной зависимостью. Эта особенность лежит в основе явления нарушения пространственной симметрии.

Как в случае a , так и в случае b кривая зависимости λ_c от k имеет экстремум, обозначенный λ'_c и λ''_c соответственно. Это означает, что при изменении λ от значений, соответствующих асимптотически устойчивому режиму (ниже кривых a и b соответственно), до значений, приводящих к неустойчивости (выше кривых a и b) первый переход будет происходить при λ'_c или λ''_c соответственно, что и будет определять поведение системы при близких значениях λ . В случае a экстремум λ'_c имеет место при $k=0$, т. е. при характерной длине $l=\infty$. Очевидно, это соответствует пространственно однородному случаю. Иными словами, в окрестности первой точки бифуркации будет доминировать однородная мода. Тем самым анализ сводится к случаю, рассмотренному в предыдущих разделах.

Новая удивительная возможность возникает в случае б (рис. 3.30). Здесь экстремум λ''_c имеет место при нетривиальном значении $k=k_c$, т. е. при вполне определенной характерной длине $l=l_c$. Отсюда следует, что в данном случае доминирующая мода в окрестности точки бифуркации будет пространственно неоднородной. Более того, эта мода будет иметь *собственные* характеристики в том смысле, что они будут целиком определяться системными параметрами. В такой ситуации мы будем говорить о *нарушении пространственной симметрии*.

Как в случае а, так и в случае б (рис. 3.30) можно снова определить подходящий набор параметров порядка, удовлетворяющий замкнутой системе уравнений в окрестности первой точки бифуркации. Будучи приведены к *нормальной форме*, эти уравнения имеют ту же структуру, что и в разд. 3.4—3.6, а также в приложении II. Более того, когда неустойчивыми становятся несколько мод одновременно, формализм допускает обобщение в виде большего числа уравнений, описывающих взаимодействие между несколькими параметрами порядка. Это приводит к бифуркациям второго и более высоких порядков и, кроме прочего, к сосуществованию нескольких устойчивых пространственных режимов. Наконец, в системах с большими размерами число взаимодействующих мод стремится к бесконечности, вследствие чего нормальной формой уравнений для параметров порядка становятся дифференциальные уравнения в частных производных. И все же можно показать, что даже для таких систем траектории стремятся к вполне конечномерному аттрактору, расположенному в конечномерном же фазовом пространстве, называемом *инерциальным многообразием*. Размерность такого многообразия может стать довольно большой, если система находится вдали от порога неустойчивости.

Во многих областях науки возникновение пространственной структуры в среде, до тех пор бывшей однородной, представляет собой центральную проблему. Например, в гидродинамике пространственные режимы часто сопровождают переход к неустойчивости, как это было в задаче Бенара, описанной в гл. I. В науке о поверхностных и объемных свойствах материалов они приводят к различным загадочным формам, и в качестве яркого примера можно отметить дендриты. В случае эмбрионального развития пространственные структуры поставляют необходимую «позиционную информацию» для клеточной дифференциации, как это было кратко отмечено в разд. 2.7. В последнем случае (а в действительности во всех задачах при наличии сцепления между переменными, как в задаче о диффузии) интересной особенностью является существование *минимального размера системы*, ниже которого развитие пространственных режимов оказывается невозможным. Интуитивно это легко по-

нять. В пространственно распределенной химической системе реакции и диффузия в случае бифуркации с нарушением симметрии взаимодействуют следующим образом. За счет внутренних обратных связей химической кинетики могут локально возникать лавинообразные явления, усиливающие эффект малых возмущений или флуктуаций. Диффузия стремится сгладить возникающие при этом неоднородности, однако, когда скорость диффузии сравнима со скоростью реакции, ей это не вполне удастся. В результате возникает некоторый пространственный режим с характерным расстоянием, отражающим среднее расстояние, на которое могут продиффундировать реагирующие молекулы, прежде чем они примут участие в реакции. Заметим, что в системах малых размеров реакция протекает с той же скоростью, что и в больших системах, а диффузия сильно ускоряется (скорость диффузии пропорциональна D/l^2 , где D — коэффициент диффузии и l — размер системы). Поэтому система как бы находится в условиях перемешивания, осуществляемого диффузией, и в силу имеющихся граничных условий любые локальные попытки дестабилизации немедленно подавляются.

Возникающие в результате бифуркаций с нарушением симметрии пространственные режимы могут быть как стационарными, так и зависящими от времени. Один из возможных примеров стационарных режимов приведен на рис. 3.31. Было проведено численное моделирование сопряжения реакции с диффузией с помощью двух уравнений, рассматривавшихся на круглом диске. Четко видно возникновение *полярной оси* в системе с дифференциацией вдоль выделенного направления, в качестве которого в реальной системе может служить, например, вертикальная или дорсовентральная ось.

Поразительный экспериментальный пример пространственных режимов, зависящих от времени, дают волны в реакции Белоусова — Жаботинского и в популяции *Dictyostelium discoideum*, показанные на рис. 1.9 и 1.15. Такие режимы получают-ся и в модельных системах. На рис. 3.32 (см. с. 162—163) показаны два примера такого рода, приводящие к простым вращающимся волнам и вращающимся спиральям соответственно.

Как для стационарных, так и для зависящих от времени пространственных режимов, упоминавшихся уже в разд. 2.7, характерно появление в виде *мультиплетов*, каждый участник которых в симметричной среде имеет одну и ту же структуру и одинаковые свойства устойчивости. Например, на рис. 3.31 полярной осью может служить любой из диаметров окружности. Аналогичным образом в бесконечно протяженной колеблющейся среде спирали возникают парами, причем каждая из спиралей вращается в противоположном направлении (по и против

часовой стрелки). Таким образом, каждая из спиралей соответствует асимметричному режиму с выделенной *хиральностью*. Однако, когда в макроскопической системе имеется множество спиралей, симметрия восстанавливается благодаря одновременному существованию одинакового количества правых и левых форм. Проблема выбора между этими формами обсуждается в разд. 3.12.

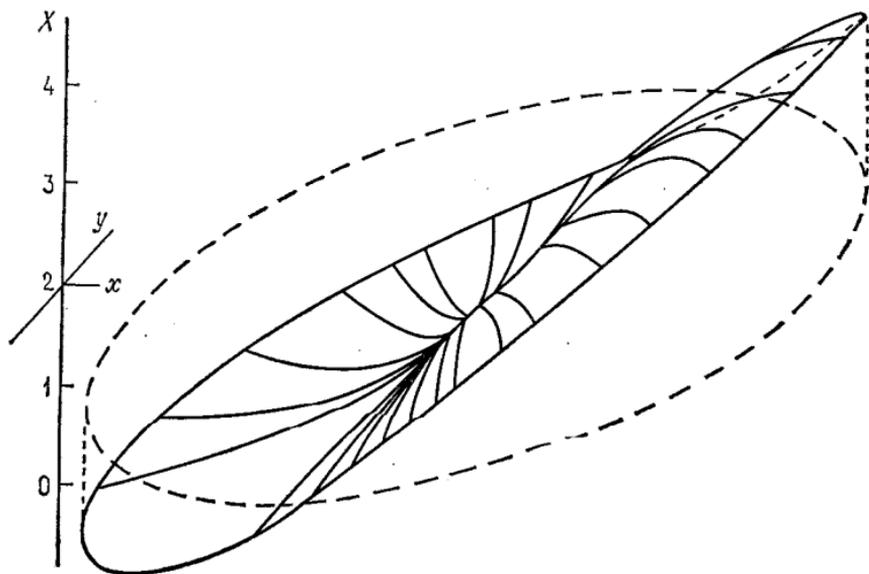
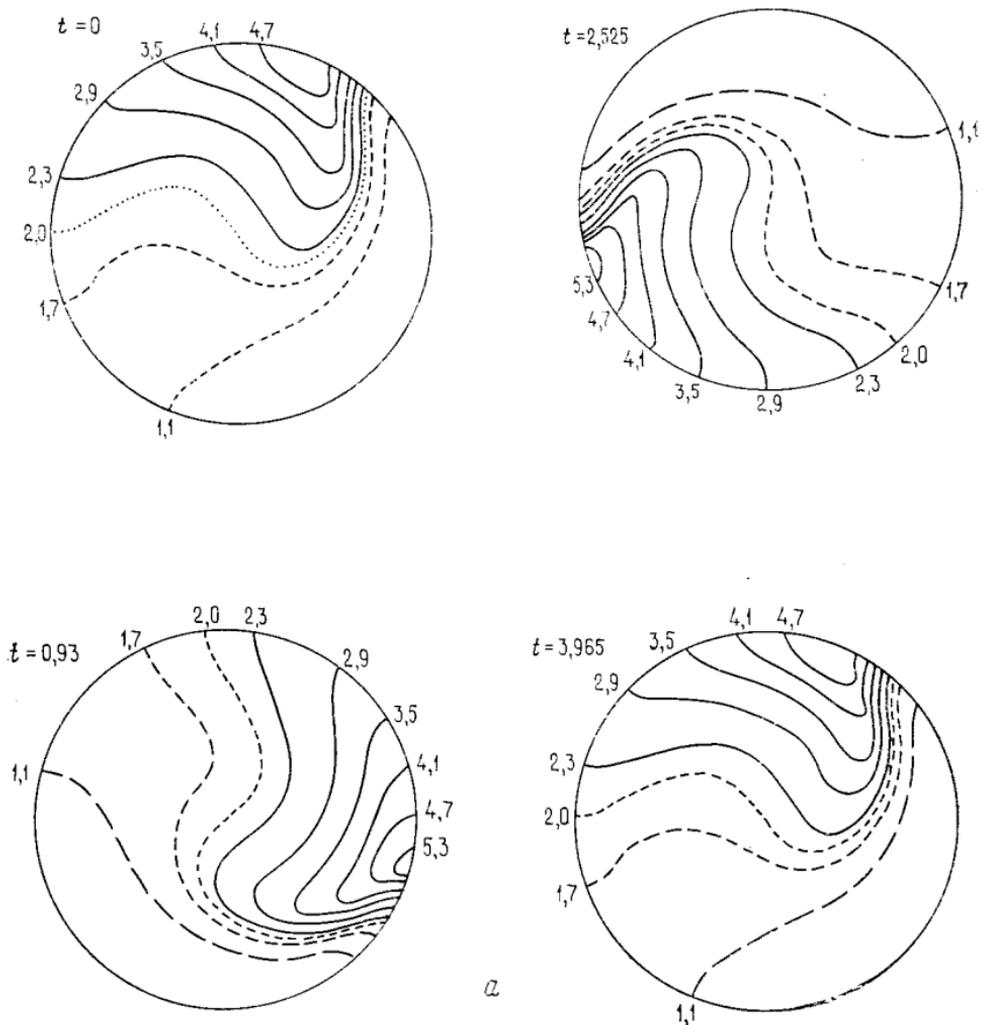


Рис. 3.31. Вид в перспективе на распределение пространственной концентрации X , порождаемое в круглой области (в плоскости x, y). Это распределение описывается произведением радиальной части [функции Бесселя первого порядка $J_1(r)$] и угловой части $\cos \varphi$.

Интригующий класс пространственно-временных режимов возникает в том случае, когда диффузия или какие-то другие транспортные явления порождают хаотическую динамику в системе, поведение которой в обычных условиях характеризуется однородностью. На рис. 3.33 показана типичная конфигурация, полученная численными методами в модели, названной Брюсселятором. Похожее поведение характерно для связанных нелинейных осцилляторов, когда путем изменения интенсивности связи удастся перевести систему из режима постоянного сдвига фаз в хаотический режим, характеризуемый выраженной фазовой десинхронизацией.

3.12. Дискретные динамические системы. Клеточные автоматы

Во многих случаях физические явления характеризуются наличием четкого порога. Выше такого порога скорость некоторого процесса заметно отличается от ее значения ниже порога, причем часто переход от одного значения к другому бывает довольно крутым. Например, электронная цепь с активными эле-



ментами типа транзистора переключается между двумя значениями тока в очень узком диапазоне значений напряжения. В компьютерах такие цепи используются как основные блоки. Пример из биологии: синтез продукта, скажем какого-либо белка, управляется таким образом, что концентрация последнего практически равна нулю при инактивации гена. В свою оче-

редь на этот факт может повлиять уровень концентрации некоторого эффектора, сопряженного с активностью белка и т. д. В настоящем разделе будет дано краткое описание динамического поведения систем, в которых составные элементы характеризуются наличием упомянутых пороговых явлений.

Рассмотрим вначале простой элемент, описываемый единственной переменной X . Типичной бывает такая ситуация: при $0 < X < X_0$ скорость изменения некоторой функции $F(X)$ очень мала, а при $X > X_0$ она внезапно становится значительной и

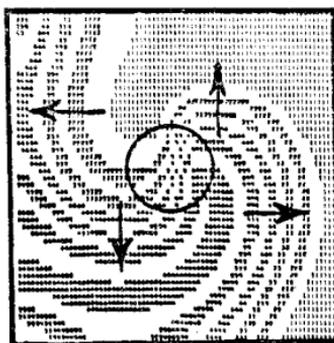
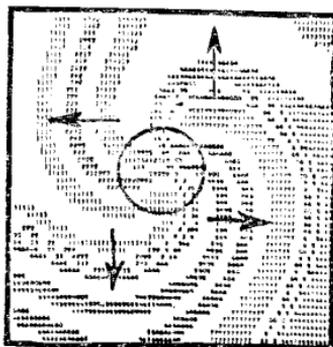
Рис. 3.32. *a* — вращающиеся волны, полученные путем численного решения модели Брюсселятора (3.19) в круге радиусом 0,5861 при условии отсутствия потока на границе. Сплошные и штриховые линии относятся к постоянным концентрациям соответственно выше и ниже значений в однородном (неустойчивом) стационарном состоянии. Значения параметров: $A=2$, $B=5,8$. Коэффициенты диффузии компонентов X и Y равны соответственно 0,008 и 0,004; *б* — вращающиеся спирали, полученные численным интегрированием уравнений:

$$\frac{\partial A}{\partial t} = \nabla^2 A - A - B + A \cdot \eta(A - 0,05)$$

где $\eta(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$

$$\frac{\partial B}{\partial t} = \nabla^2 B + A/2$$

[из работы: Winfree A., SIAM—AMS Proceedings, 8, 13 (1974)].



б

вскоре опять исчезает, после того как устанавливается некоторое значение F_{\max} (рис. 3.34). Чтобы ухватить суть такого отклика, будем рассматривать F как дискретную функцию дискретной переменной X . Иными словами, будем считать, что при $0 < X < X_0$ как X , так и F равны «нулю», а при $X > X_0$ обе величины (X и F) равны «единице». Такие величины известны под названием *булевых переменных* в честь английского математика XIX в. Джорджа Буля, считающегося основателем формальной логики.

Изменение величины X во времени будет диктоваться соответствующими логическими уравнениями, связывающими X и F . Эти уравнения выступают в качестве замены обычных дифференциальных уравнений, соответствующих непрерывному формализму [см., например, уравнения (3.1)]. Простейшим примером служит уравнение вида

$$X_{t+1} = F(X_t) \quad (3.32)$$

Фигурирующие здесь переменные изменяются совершенно синхронно, но в дискретные моменты времени. Вид функции F за-



Рис. 3.33. Схематическое представление индуцированного диффузией хаоса, полученное численным решением уравнений, описывающих модель Брюсселя-тора [из работы: Kuramoto Y., *Physica*, 106A, 128 (1981)].

висит от того, какая конкретно задача рассматривается. Пусть рассматривается петля положительной обратной связи (примером может служить альbedo ледовой поверхности, обсуждавшаяся в разд. 1.8). На языке логических переменных это означает, что если X (в нашем конкретном случае — температура) велико, то за счет динамики системы оно будет стремиться к дальнейшему увеличению, т. е. значение $F(X)$ будет также велико. Справедливо и обратное утверждение, если рассматривается понижение X . Вспоминая, что F и X — булевы переменные, эту зависимость можно выразить логическим равенством $F =$

$= T^*$. Решения этого уравнения наглядно представляются в виде таблицы состояний, в которой последовательно выписаны логические значения переменных F и X :

| F | X |
|-----|-----|
| 0 | 0 |
| 1 | 1 |

Отсюда видно, что имеются два состояния, в которые переменные F и X выключаются и включаются одновременно. Такие состояния, обозначаемые $(0, 0)$ и $(1, 1)$, не эволюционируют во

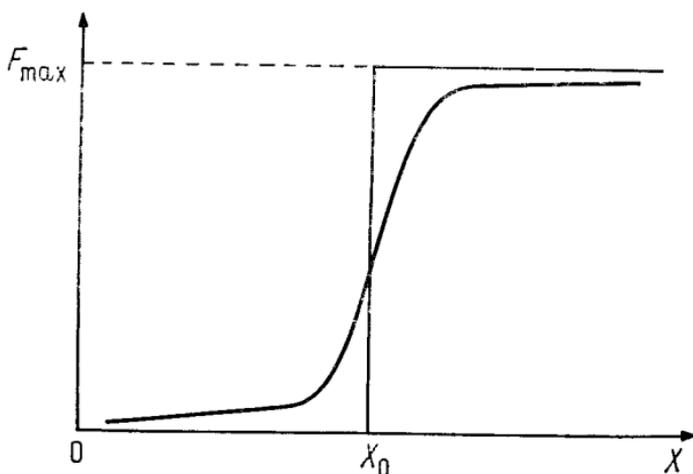


Рис. 3.34. Схематическое представление булевой функции F , зависящей от булевой переменной X .

времени и являются аналогами неподвижных точек непрерывных динамических систем. Таким образом, мы получили очень простую модель перехода между различными стационарными состояниями в системах с положительной обратной связью. Можно проанализировать совершенно таким же образом и более сложные ситуации, возникающие при наличии систем со многими переменными, а также положительных и отрицательных обратных связей.

Очень интересная ситуация возникает в системах, состоящих из решетки идентичных булевых компонент и взаимодействующих

* Заметим, что, несмотря на свою кажущуюся линейность, это уравнение в действительности отражает факт сильно нелинейной зависимости, изображенной на рис. 3.34.

щих в пространстве локально. Такие системы называются *клеточными автоматами*. Типичный одномерный клеточный автомат описывается уравнением

$$X_{i,t+1} = F_i(X_{i-r,t}, X_{i-r+1,t}, \dots, X_{i,t}, \dots, X_{i+r,t}) \quad (3.33)$$

В соответствии с этим уравнением булева переменная X , соответствующая конкретному центру i , определяется предыдущими значениями переменной X из окрестности центра i , содержащей $2r+1$ центров. В зависимости от правил построения функций F_i динамические системы этого типа и их двумерные аналоги могут порождать множество поразительных пространственных режимов, очень похожих на те, что были описаны в разд. 3.11.

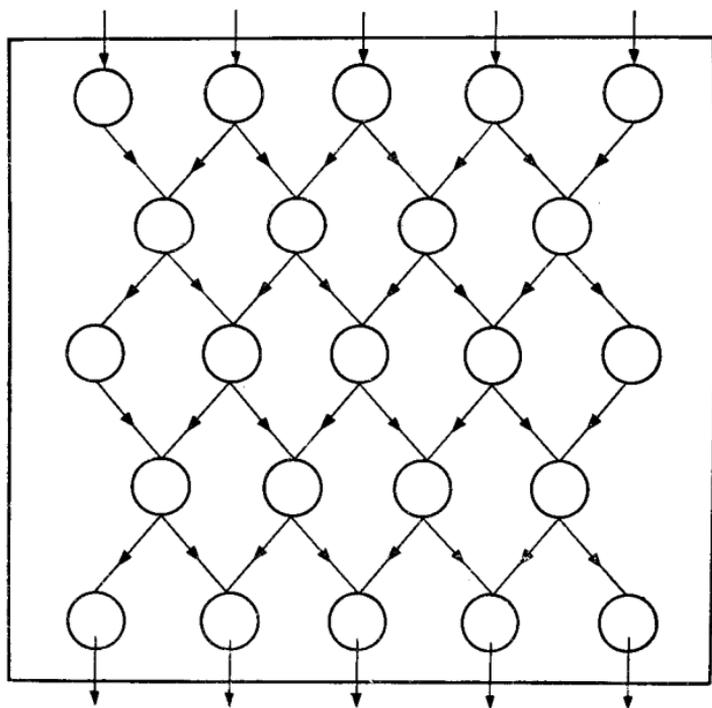


Рис. 3.35. Ряд процессоров, каждый из которых обрабатывает данные, получаемые локально от своих соседей.

Зададимся теперь вопросом другого рода. Мы имеем ряд процессоров, работающих с булевыми данными, получаемыми локально от соседей (см., например, рис. 3.35). Каждый из этих элементов, состояние которых также представлено булевой функцией, «вычисляет» свой выход, который затем передается соседям. Может ли сеть такого типа породить глобальный «вычислительный» режим как функцию времени, который характе-

ризовался бы некоторыми новыми, *возникшими свойствами*, не содержащимися в исходно запрограммированных элементах? В принципе, на этот вопрос можно дать утвердительный ответ. Он основан на существовании аттракторов в диссипативной динамической системе, лежащей в основе такой сети. Аттрактор, соответствующий режиму самоорганизации, возникающему вследствие наличия неустойчивости в условиях нарушения симметрии, обязательно наделяет систему некоторыми коллективными свойствами, которые выходят за пределы свойств отдельных субъединиц. В действительности в системе с большой пространственной протяженностью будет иметься много аттракторов. В зависимости от характера начального стимула система оказывается в итоге в различных зонах притяжения. Это позволяет предвидеть возможность обучения машин надежно различать различные входы, даже слегка искаженные.

В связи с тем что аттракторы играют большую роль в вычислительных или мыслительных процессах, возникает множество серьезных проблем, находящихся на грани между теорией динамических систем и физикой неравновесных состояний с одной стороны, и наукой о вычислительных и распознающих устройствах — с другой. По всей вероятности, результаты, получаемые из этого неожиданного синтеза, будут иметь далеко идущие последствия.

3.13. Асимметрия, отбор и информация

Мы выяснили, что пространственно асимметричные решения, возникающие в результате бифуркации в исходно идеально симметричной среде, имеют некоторый особый статус: в достаточно малой области пространства может реализоваться лишь одно такое решение. Например, как видно из рис. 1.9, некоторая точка в сосуде, где протекает реакция Белоусова — Жаботинского, становится центром, из которого исходят спирали лишь с данной хиральностью, скажем левосторонней. Следовательно, о *нарушении симметрии* можно говорить именно в этом смысле. Однако, если ту же часть системы рассматривать на протяжении большого числа независимых экспериментов, то можно будет с одинаковой частотой наблюдать решения с противоположной симметрией. Совершенно аналогично для данного эксперимента, выполняемого в системе с большой протяженностью, в различных частях системы могут сосуществовать решения с противоположной симметрией. Таким образом, в этом смысле симметрия исходной среды восстанавливается.

Этот вывод представляется естественным в обычных лабораторных условиях. В конце концов, с какой стати должна появиться ячейка Бенара, в которой одно из двух возможных

направлений вращения — по или против часовой стрелки — стало бы предпочтительным? Соответственно в экспериментах по реакции Белоусова — Жаботинского всегда наблюдалось появление спиральных волн парами до тех пор, пока экспериментатор не предпримет специальных мер в пользу конкретной конфигурации.

Во многих других случаях невозможность выбора конкретного асимметричного состояния вещества из множества имеющихся форм приводит к некоторым фундаментальным вопросам. Чтобы разобраться в сути проблемы, проведем следующий мысленный эксперимент, проиллюстрированный рис. 3.36. Допустим,

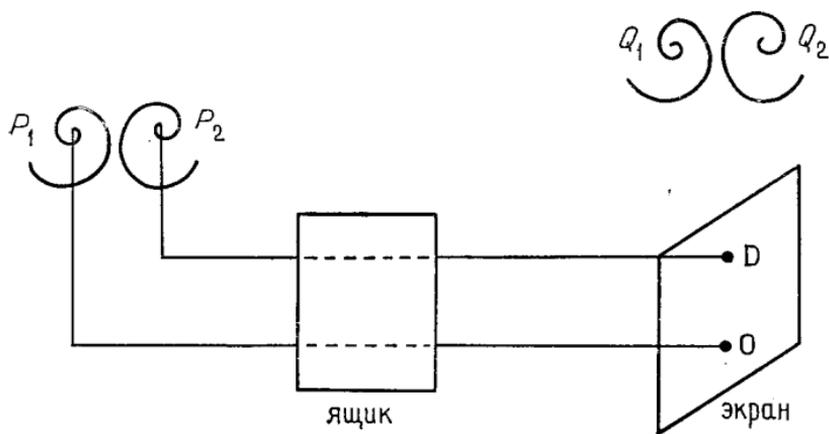


Рис. 3.36. Наделяя систему свойством множественности решений, бифуркация потенциально способна создавать информацию. Это иллюстрируется приведенной выше «бифуркационной машиной», выдающей символы D и O в зависимости от того, какая спираль появляется в результате бифуркации — левая или правая. В отсутствие селекции между волнами, реализуемыми при последовательных действиях такого устройства, воспроизводимость ориентации не наблюдается. В результате последовательность двух символов D и O становится случайной, что исключает возможность обработки информации.

что в подходящей химической реакции реализовано определенное состояние с вращающимися спиральями. Обозначим центры правых и левых спиралей в этом конкретном состоянии через $P_1, Q_1, P_2, Q_2, \dots$ соответственно. Допустим дальше, что левая спираль с помощью подходящей для этой цели схемы порождает на экране S некоторый символ, например «D», а правая спираль с помощью такого же устройства порождает на экране символ «O». Одна из функций, которые могут быть построены с помощью этих символов, состоит в том, что сигнал, посылаемый P_2 , высвечивается в верхней части экрана, а P_1 — в нижней. При чтении текста сверху вниз получающееся сообщение должно восприниматься как специальная команда, предписывающая выполнить какую-то конкретную задачу. Эта команда

имеет вид «DO». Естественно, при этом предполагается, что «читатель» этого текста уже знает, в чем именно состоит его задача, и, кроме того, способен прочесть текст, для чего достаточно в рассматриваемом примере рудиментарных знаний английского языка.

Проследим теперь за работой этой машины, когда в системе запущены решения в виде вращающихся волн. Точки P_1 и P_2 (рис. 3.36) мы выбрали таким образом, чтобы в первом эксперименте P_2 была центром левой спирали, а P_1 — центром правой спирали. На экране сразу же появится команда «DO», она будет адекватно интерпретирована, и наш «читатель» приступит к выполнению предписанной ему задачи. После нескольких раз успешного срабатывания эта машина выключится. Посмотрим, что же произойдет, когда ею воспользуются снова.

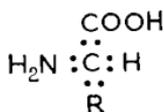
Вследствие множественности имеющихся режимов точки P_1 и P_2 становятся центрами как левых, так и правых спиралей. Таким образом, наша машина отнюдь не обязательно будет генерировать одно и то же сообщение «DO», как это было отмечено выше. Действительно, велика вероятность того, что наш «читатель» увидит на экране сообщение бессодержательное, типа «OD», «OO» или «DD». В результате он не сможет выполнить поставленной перед ним задачи. Иными словами, изобретенная нами «бифуркационная машина» оказалась ненадежным *информационным процессором*.

Допустим теперь, что мы проделали следующий трюк. После первого запуска машины мы установим два небольших мотора под точками P_1 и P_2 , придающих вращение среде в правом и левом направлениях соответственно. Вследствие такой предрасположенности среды волна будет неизменно закручиваться в окрестности P_1 и P_2 влево и вправо соответственно, в результате чего при каждом срабатывании системы будет передаваться одно и то же сообщение «DO». Очевидно, теперь наша машина станет надежной! Читателю должно быть ясно, что введенные в нашем простом примере моторы представляют собой всего лишь иллюстрацию более общего понятия *отбора*. Иными словами, нарушение пространственной симметрии является необходимой предпосылкой, без которой вопрос о построении информационного процессора попросту не возник бы. Однако в конечном счете именно отбор позволяет детектировать, интерпретировать и передавать «сообщение», спрятанное в нелинейной, неравновесной динамике системы. *Отбор позволяет декодировать информацию* и тем самым допускает переход сложного с одного уровня на другой.

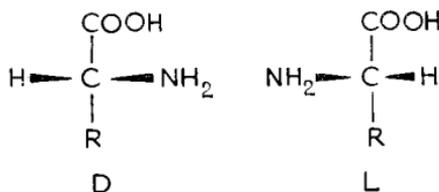
Однако необходимость отбора выходит далеко за пределы нашего простого примера. В разд. 1.9 упоминался вопрос о нарушении симметрии между веществом и антивеществом. Имен-

но благодаря тому, что веществу было отдано предпочтение перед антивеществом, в существующей Вселенной смогли образоваться ядра, галактики и живые организмы вместо того, чтобы происходила аннигиляция вещества с антивеществом с образованием фотонов. Если вернуться к более привычным масштабам, то можно утверждать, что в основе структуры и функционирования современных организмов лежит именно отбор некоторых предпочтительных асимметричных форм. Рассмотрим некоторые характерные примеры, иллюстрирующие вездесущий характер асимметрии в этой области.

Из элементарных учебников по биологии мы знаем, что белки являются одними из основных строительных блоков жизни. Независимо от их происхождения и функционирования все удивительное разнообразие существующих белков возникает благодаря различным способам соединения одних и тех же 20 видов *аминокислот*. Найденные в белках аминокислоты имеют четыре группы, связанные с центральным атомом углерода, называемым α -углеродом:



Эти группы таковы: 1) COOH — карбоксильная группа, 2) NH₂ — аминогруппа, 3) H — атом водорода и 4) R — боковая группа, разная для каждой аминокислоты. Во всех аминокислотах, за исключением глицина, для которого R совпадает с H, α -углерод является асимметричным, поскольку к нему присоединены различные группы! Таким образом, имеются две возможные конфигурации аминокислот: D-изомеры и L-изомеры:



Добиться совпадения изомеров D и L путем вращений невозможно. Они подобны правой и левой рукам и в этом смысле аналогичны паре спиралей, порождаемых реакцией Белоусова — Жаботинского. Их асимметрия проявляется в способности вращать плоскость поляризации поляризованного света в заданном направлении — D-изомеры вращают ее вправо, а L-изомеры — влево.

Аминокислоты в белках всегда соответствуют L-конфигурации. Поэтому здесь мы имеем поразительный пример абсолют-

ного отбора асимметрии. Великий французский ученый Луи Пастер, основатель современной биохимии, считал это явление одним из основных свойств жизни.

Однако мы еще далеко не исчерпали список примеров асимметрии в биологии. Нуклеиновые кислоты, образующие, как и белки, еще один класс строительных блоков жизни, состоят из субъединиц, называемых нуклеотидами. Составной частью в этих молекулы входит пуриновое или пиримидиновое основание, представляющее собой по существу плоскую молекулу. Кроме того, в них входит тот или иной сахар, обладающий трехмерной хиральной структурой: D-дезоксирибоза в случае дезоксирибонуклеиновой кислоты (ДНК) и D-рибоза в случае рибонуклеиновой кислоты (РНК). Опять же все нуклеиновые кислоты живых организмов представляются результатом абсолютного отбора, оставившего только правые изомеры D.

Еще один столь же поразительный пример предпочтительной асимметрии представляет собой механизм считывания генетического кода. Современной наукой установлено, что наследственная информация шифруется триплетным кодом, в котором каждый символ соответствует набору из трех нуклеотидов (кодонов). Для нас здесь важнее всего то, что существуют *стартовые сигналы*, позволяющие читать генетическую информацию в виде троек нуклеотидов в *заданном направлении* от точки старта. Эта асимметрия первична по своей важности, поскольку ее отсутствие привело бы к эффектам, аналогичным чтению текста на английском языке в произвольном направлении — как слева направо, так и справа налево! В этом плане следует отметить, что информационные свойства синтетических нуклеотидов, не обладающих такой полярностью, весьма невелики.

Теперь уже конструктивная роль отбора не вызывает у нас сомнений. Поэтому на передний план выходит вопрос о механизме его возникновения.

Наиболее четкий пример отбора представляет собой случай, когда бифуркация с нарушением симметрии происходит в виде единственного, неповторяемого события. В качестве такого примера может служить выбор вещества вместо антивещества, который, осуществившись, ведет Вселенную к невозвратному состоянию: трудно даже вообразить какую-либо другую реализацию, по крайней мере в рамках временного масштаба в 20 млрд. лет! Аналогичный случай возникает и тогда, когда переход к асимметричным формам происходит не спонтанно, а через стадию нуклеации, в которой система сначала «активируется» и преодолевает некоторый «барьер». Если требуемая «энергия активации» очень велика, то переход к конкретному асимметричному состоянию происходит очень мед-

ленно. Поэтому ясно, что к тому времени, когда состояние с противоположной симметрией будет готово появиться, первое состояние уже победит и модифицирует внешнюю среду, запрещая тем самым в дальнейшем появление любой формы с отличной симметрией.

Другим общим механизмом отбора является передача некоторой внешней асимметрии конкретной рассматриваемой системе. Здесь сразу же возникают два момента. Во-первых, такая первичная асимметрия должна всегда действовать одинаково. Один из примеров, когда это условие выполняется, — тяготение, которое неизбежно влияет на любой проводимый на Земле эксперимент с массивными объектами за счет введения выделенного вертикального направления. Во-вторых, система должна быть достаточно чувствительной для того, чтобы уловить эти эффекты, которые обычно очень малы. Именно здесь бифуркации начинают снова играть важную роль. В самом деле, как мы видели в разд. 3.4, даже малые изменения параметров в динамических уравнениях сильно влияют на бифуркацию и даже могут привести к ее устранению при определенных условиях (см., например, рис. 3.10, *в*). Поэтому, если в истории развития некоторой системы имеются достаточно большие промежутки времени, когда она находится близко к точке бифуркации, она может стать чувствительной к малым эффектам. По-видимому, такие эффекты могут привести к отбору конкретной формы асимметрии. К обсуждению этого механизма мы еще вернемся в гл. 4, где будут развиты представления, необходимые для работы со стохастическими возмущениями.

Ясно, что в обоих рассмотренных механизмах имеет место необратимость, обеспечивающая, как это неоднократно отмечалось, асимптотическую устойчивость различных состояний. Однако можно рассмотреть еще один механизм отбора, в котором необратимость проявляется даже более непосредственно. В частности, можно поставить такой вопрос: может ли связанное с необратимостью универсальное нарушение симметрии относительно обращения времени само служить источником отбора, приводящего к информационно-насыщенным, *пространственно* асимметричным формам материи? В настоящий момент ответить на этот вопрос не представляется возможным. Сначала необходимо должным образом определить понятие информации, которое до сих пор рассматривалось очень приблизительно в расчете на интуицию читателя. Этого мы сможем достичь лишь после того, как будет изложена стохастическая теория динамических систем, которой мы сейчас и займемся.

СЛУЧАЙНОЕ И СЛОЖНОЕ

Ранее уже упоминалась роль элементов случайности в возникновении сложного поведения. Такие элементы всегда имеются в любой физико-химической системе в виде внутренних *термодинамических флуктуаций*. Далее, во многих случаях действующая на систему внешняя среда сама содержит значительные шумы. Примером могут служить сезонные или иные флуктуации климата, влияющие на некоторую экосистему. В данной главе рассматривается вопрос о динамике флуктуаций. Специальное внимание будет уделено механизму их влияния на макроскопическое поведение, а также новым понятиям, обусловленным вероятностным подходом к описанию динамических систем. Как будет ясно из дальнейшего, особо важную роль стохастические эффекты играют вблизи точек бифуркации.

4.1. Флуктуации и вероятностное описание

Наблюдение состояния какой-либо физико-химической системы, как правило, сопряжено с усреднением мгновенных значений соответствующих переменных либо по времени, либо по объему надмолекулярных размеров. Например, если поместить около $0,33 \cdot 10^{23}$ молекул воды в сосуд объемом 1 см^3 при нормальных температуре и давлении, то получится жидкость с плотностью $0,33 \cdot 10^{23}$ молекула/ см^3 , или, пользуясь для выражения плотности массой, $0,33 \cdot 10^{23} \times$ (масса одной молекулы H_2O) = 1 г/см^3 . Именно с такими переменными мы работали большей частью в предыдущих главах. С другой стороны, отношение числа молекул в некотором малом объеме к этому объему (например, такой объем может быть порядка 10 \AA^3) будет непрерывно отклоняться от этого макроскопического значения. Эти флуктуации обусловлены тем, что в процессе хаотического движения молекулы пересекают границы заданного объема (случайный характер молекулярного движения отмечался уже в первых главах этой книги), в силу чего число молекул в каждом небольшом объеме является существенно непредсказуемой величиной. Порождаемые этим механизмом от-

клонения мы называем *флуктуациями*. Благодаря таким флуктуациям физико-химические системы приобретают способность «прощупывать» фазовое пространство и блуждать около состояния, предсказываемого решением феноменологического, детерминистического уравнения (системы уравнений).

При анализе проблемы флуктуаций естественно воспользоваться подходами теории вероятностей. Например, в системе с химическими реакциями нас может интересовать вопрос о том, какова вероятность $P(X_\alpha, \Delta V, t)$ того, что в момент времени t в элементе объема ΔV содержится X_α частиц компонента α . Более полную информацию дают зависящие от многих переменных вероятности $P(X_{\alpha j}, \Delta V_j; X_{\beta k}, \Delta V_k; \dots, t)$ одновременного присутствия в непересекающихся объемах j, k количества молекул $X_{\alpha j}, X_{\beta k}$ химических компонентов α, β, \dots

Удобный способ описания распределения вероятностей дают его моменты, определенные следующими уравнениями:

$$\langle X_{\alpha j}^{m_{\alpha j}} X_{\beta k}^{m_{\beta k}}, \dots \rangle = \sum_{X_{\alpha j}, X_{\beta k}, \dots} X_{\alpha j}^{m_{\alpha j}} X_{\beta k}^{m_{\beta k}} \dots P(X_{\alpha j}, \Delta V_j; X_{\beta k}, \Delta V_k; \dots) \quad (4.1)$$

Здесь $m_{\alpha j}, \dots$ — целые числа. Особый интерес представляют первые два момента, дающие среднее значение $\langle X_{\alpha j} \rangle$, и второй момент $\langle X_{\alpha j} X_{\beta k} \rangle$. В большинстве случаев удобнее работать с некоторой избыточной величиной, называемой *ковариацией* и определенной соотношением

$$\langle \delta X_{\alpha j} \delta X_{\beta k} \rangle = \langle X_{\alpha j} X_{\beta k} \rangle - \langle X_{\alpha j} \rangle \langle X_{\beta k} \rangle \quad (4.2)$$

Проиллюстрируем важную роль этих величин в частном случае единственной стохастической переменной X , для которой выражение (4.2) сводится к *дисперсии* $\langle \delta X^2 \rangle$ соответствующего распределения вероятностей. Введем прежде всего меру статистического разброса с помощью *стандартного отклонения*

$$\delta = (\langle \delta X^2 \rangle / \langle X \rangle^2)^{1/2} \quad (4.3)$$

Если δ мало, то это означает, что распределение имеет острый пик вблизи своего среднего значения $\langle X \rangle$, которое в свою очередь близко к наиболее вероятному значению \bar{X} , при котором $P(X)$ имеет единственный максимум. В предельном случае такая ситуация соответствует распределению вероятностей в виде бесконечно острого пика $P(X) = \delta^{hr}_{X, \bar{X}}$, т. е.

$$\begin{aligned} P(X) &= 1 \quad \text{при } X = \bar{X} \\ &= 0 \quad \text{при } X \neq \bar{X} \end{aligned} \quad (4.4)$$

Иными словами, все состояния пусты, за исключением одного, $X = \bar{X}$. Очевидно, среднее значение и дисперсия равны

$$\langle X \rangle = \sum_X X \delta_{X, \bar{X}}^{kr} = \bar{X}$$

$$\langle X^2 \rangle = \sum_X X^2 \delta_{X, \bar{X}}^{kr} = \bar{X}^2$$

$$\text{и } \langle \delta X^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2 = 0$$

или, записывая то же самое короче, $\delta = 0$. Более общий вид распределения, удовлетворяющий требованию малости δ , показан на рис. 4.1, а. По сути дела, распределение Пуассона, с которым мы встречались в разд. 2.8, принадлежит этому классу. Кроме того, выполненные в конце разд. 2.8 явные вычисления δ показывают, что $\delta = 1/N^{1/2}$, где N — большое число, измеряющее размер системы (например, в виде среднего числа частиц,

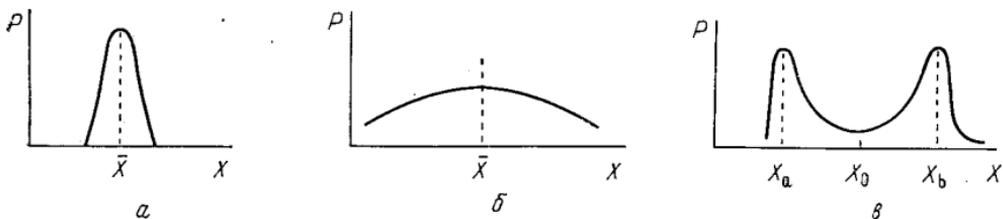


Рис. 4.1. Стохастический аналог бифуркации. При переходе управляющего параметра через критическое значение λ_c вероятностная функция переключается от одномодальной формы с резким пиком на единственном аттракторе ($\lambda < \lambda_c$, рис. 4.1, а) к плоскому распределению ($\lambda = \lambda_c$, рис. 4.1, б) и, наконец, ко многогорбому распределению, максимумы которого совпадают с новыми аттракторами, возникающими выше точки бифуркации ($\lambda > \lambda_c$, рис. 4.1, в).

содержащихся в объеме V , или через сам объем V). Далее, в любой системе, которая представляет практический интерес, линейные размеры значительно превышают среднее расстояние между молекулами. Формально это выражается в виде *термодинамического предела* $V \rightarrow \infty$. Видно, что в этом случае δ стремится к нулю. Иными словами, между макроскопическим поведением и флуктуациями имеется четкое различие. Множество других фундаментальных распределений, встречающихся в теории вероятностей, типа, например, биномиального или гауссова распределения также обладают этим важным свойством.

Отмеченные выше особенности допускают элегантную и общую формулировку в виде двух фундаментальных результатов теории вероятностей, известных под названием *закона больших чисел* и *центральной предельной теоремы*. В обоих случаях нас интересуют свойства переменной X , представляющей собой сумму статистически независимых случайных переменных, распре-

деленных по одному и тому же закону:

$$X = X_1 + X_2 + \dots + X_n \quad (4.5)$$

При очень мягком предположении о том, что существует среднее значение $\langle X_h \rangle \equiv m$ отдельно взятого распределения, согласно закону больших чисел, вероятность того, что арифметическое среднее $(X_1 + \dots + X_n)/n$ отличается от статистического среднего m на величину, меньшую любого наперед заданного ϵ , стремится к единице при $n \rightarrow \infty$ и $\epsilon \rightarrow 0$. Допуская дополнительно, что существует дисперсия отдельного распределения $\langle \delta X^2_n \rangle \equiv \sigma^2$, можно доказать центральную предельную теорему, согласно которой в пределе $n \rightarrow \infty$ распределение вероятностей величины $(X_1 + \dots + X_n)/n$ становится сколь угодно близким к *нормальному распределению*, представляющему собой интеграл от *гауссовой плотности распределения*. Кроме того, дисперсия этого распределения по порядку величины равна $\sigma^2 n^{-1}$, откуда следует, что дисперсия самого X равна $\sigma^2 n$. Последнее в точности совпадает с результатом, полученным путем непосредственного вычисления дисперсии пуассоновского распределения, в котором роль n играет представляющий размер системы параметр N , выражающий тот факт, что переменная X есть сумма конкретного набора наблюдаемых частиц, относящаяся к малому элементу объема системы (экстенсивная переменная).

Результаты такого рода легко найти в литературе по теории вероятностей. Однако нетрудно построить и противоположные примеры. Рассмотрим, например, распределение вероятностей некоторой микроскопической переменной (переменная может быть экстенсивной или интенсивной — последняя получается делением экстенсивной переменной на размер системы), вид которого представлен на рис. 4.1, в. Ясно, что это распределение не является гауссовым! Это распределение с двумя максимумами, расположенными при $X = X_a$ и $X = X_b$, разделенными точкой минимума при $X = X_0$. Нагляднее всего такое распределение представляется двумя бесконечно узкими пиками:

$$P = \frac{1}{2} [\delta_{X, X_a}^{kr} + \delta_{X, X_b}^{kr}] \quad (4.6)$$

Здесь множитель $1/2$ обеспечивает нормировку $P(X)$ на единицу: $\sum_{X=0}^{\infty} P(X) = 1$. Найдем среднее значение и дисперсию этого распределения. Имеем

$$\begin{aligned} \langle X \rangle &= \sum_X X P(X) = \frac{1}{2} (X_a + X_b) \\ \langle X^2 \rangle &= \sum_X X^2 P(X) = \frac{1}{2} (X_a^2 + X_b^2) \end{aligned}$$

$$\langle \delta X^2 \rangle = \frac{1}{2} (X_a^2 + X_b^2) - \frac{1}{4} (X_a + X_b)^2 = \frac{1}{4} (X_a - X_b)^2$$

Это дает следующую статистическую дисперсию:

$$\delta = \left(\frac{\langle \delta X^2 \rangle}{\langle X \rangle^2} \right)^{1/2} = \frac{(X_b - X_a)}{(X_b + X_a)}$$

Допустим, что X — экстенсивная переменная, т. е. X пропорциональна размеру системы N :

$$X = N \cdot x \quad (4.7)$$

Здесь интенсивная переменная x конечна. Отсюда следует такая оценка δ :

$$\delta = \frac{(x_b - x_a)}{(x_b + x_a)} = \text{конечная величина порядка } 1 \quad (4.8)$$

Иными словами, дисперсия будет того же порядка, что и квадрат среднего, в результате чего флуктуации проявляются на макроскопическом уровне. Ясно, что предпосылки центральной предельной теоремы здесь не должны выполняться, так что переменная, распределение которой показано на рис. 4.1, в, никоим образом не может быть суммой статистически независимых случайных переменных. Уже можно догадаться, что это должно отражать некоторую *согласованность поведения* системы. Из рис. 4.1, а и 4.1, в следует, что эту согласованность следует отнести на счет перехода из состояния с единственным наиболее вероятным значением \bar{X} в состояние с двумя такими значениями, X_a и X_b . Это не что иное, как *стохастический аналог бифуркации* — явления, которому было уделено основное внимание в гл. 3. Ясно, что в рамках детерминистического описания проблему возникновения и существования согласованности поведения не удастся даже сформулировать. Это одна из основных причин того, что исследование флуктуаций входит составной частью в динамику сложных явлений. Другая причина состоит в том, что в случае «многогорбого» режима после бифуркации важно оценить относительную устойчивость наиболее вероятных состояний, а также время нахождения системы в их окрестности до совершения перехода на одну из параллельно существующих устойчивых ветвей. Этот вопрос также удастся поставить лишь в рамках расширенного описания, учитывающего влияние флуктуаций.

На рис. 4.1, б показан вид распределения, которое можно ожидать в промежуточном случае между одномодальным и многомодальным режимами. Мы имеем плоское распределение, которое «готовит путь» для бифуркации, прощупывая с конечной

вероятностью области фазового пространства, довольно удаленные от наиболее вероятного значения. Этот *вероятностный аналог маргинальной устойчивости* обсуждается далее в разд. 4.4, где будет показано, что в этом случае центральная предельная теорема не выполняется.

До сих пор в настоящем разделе мы уделяли основное внимание роли флуктуаций, генерируемых спонтанно за счет молекулярной динамики системы. Однако, как было видно из разд. 3.7—3.10, в широких классах динамических систем случайное поведение может возникать также вследствие бифуркации неустойчивых режимов, приводящей, например, к хаотическим аттракторам. По этой причине вероятностное описание становится ценным дополнением детерминистического подхода, позволяющим, например, определять частоту посещения различных областей хаотического аттрактора, а также время, после которого начинают исчезать воспоминания о конкретном исходном состоянии.

Основное различие между флуктуациями и случайностью, обусловленной хаосом, состоит в *интенсивности и временном масштабе* отклонений от среднего. Как было показано ранее в настоящем разделе, в случае единственного асимптотически устойчивого аттрактора в физико-химической системе флуктуации вокруг среднего значения чрезвычайно малы — примерно в $N^{1/2}$ раз меньше, нежели детерминистические значения. Даже в многомодальном режиме, когда разброс вокруг среднего велик, дисперсия вблизи каждого из пиков остается малой. Кроме того, значения вероятности того, что изображающая точка окажется в окрестности неустойчивого состояния, малы примерно как e^{-N} , т. е. убывают экспоненциально с ростом размера системы N . Это означает, что большую часть времени система проводит в непосредственной окрестности вероятностного пика и что *время перехода* между пиками чрезвычайно высоко — порядка e^N . Ни одним из этих свойств типичный хаотический аттрактор не обладает. В этом случае система посещает удаленные друг от друга состояния с временным интервалом, сравнимым с детерминистическими временными масштабами, что указывает на следующее обстоятельство: амплитуда отклонения от среднего сравнима с самим средним. Такое поведение напоминает турбулентность, что еще раз подчеркивает роль хаоса как естественной модели турбулентности в физико-химических системах.

Аналогичное свойство характерно для динамики систем, помещенных в случайную внешнюю среду, изменчивость которой воспринимается системой как *внешний шум*. Более детальный обзор этой важной темы читатель сможет найти в книге Хорстэмке и Лефевра.

4.2. Марковские процессы. Основное уравнение

В предыдущем разделе мы исходили из того, что распределение вероятностей, описывающее рассматриваемую систему, известно по крайней мере в общих чертах. В некоторых случаях так и бывает. Например, как показано в разд. 2.8, в состоянии термодинамического равновесия для идеального газа вероятность обнаружения заданного числа частиц в малом объеме определяется пуассоновским распределением. Аналогично если в какой-то системе выполняются условия применимости центральной предельной теоремы, упоминавшейся в разд. 4.1, то соответствующие переменные будут описываться нормальным распределением. Однако в общем случае известны лишь процессы, происходящие в динамической системе: теплопроводность и конвекция в задаче Бенара, реакции и диффузия в изотермической химической системе и т. д. Каким образом на основе этой информации можно установить вид соответствующего распределения вероятностей?

Чтобы выяснить этот вопрос, напишем соответствующее уравнение на основе баланса процессов, приводящих систему к определенному состоянию Q , и процессов, уводящих систему из этого состояния. Очевидно,

$$\frac{d}{dt} \text{Prob}(Q, t) = (\text{член, соответствующий переходам в состояние } Q \text{ в единицу времени}) - (\text{член, соответствующий уходам из состояния } Q \text{ в единицу времени}) \equiv R_+(Q) - R_-(Q) \quad (4.9)$$

Теперь задача сводится к определению скоростей прихода в состояние Q (R_+) и ухода из этого состояния (R_-). Какими бы ни были явные выражения для этих величин, они должны удовлетворять определенным условиям. Известно, например, что в состоянии равновесия физико-химические системы удовлетворяют условию *детального равновесия*, неоднократно упоминавшемуся в гл. 1 и 2. Так, если представить R_+ и R_- в виде сумм, соответствующих элементарным процессам, происходящим в системе,

$$R_{\pm} = \sum_k r_{k, \pm}$$

то должно выполняться следующее условие:

$$(r_{k, +})_{\text{eq}} = (r_{k, -})_{\text{eq}} \quad (4.10)$$

В свою очередь эти соотношения должны согласоваться с равновесным распределением вероятностей, известным из статистической механики. Примером такого рода служит пуассоновское распределение. Эйнштейну первому удалось показать в общем

виде, что в состоянии равновесия вероятность флуктуации целиком определяется термодинамическими величинами. Обращая известную формулу Больцмана $S = k_B \ln$ (число молекулярных состояний, совместимых с данным значением энергии), справедливую для изолированной системы, получаем

$$P_{\text{eq}} \sim \exp\left(\frac{1}{k_B} \Delta S\right) \quad (4.11)$$

где ΔS — изменение энтропии вследствие флуктуации, $\Delta S = S(Q) - S(Q_{\text{eq}})$.

Другое условие, которому должно удовлетворять уравнение (4.9), состоит в том, что оно должно сводиться к динамическому уравнению, с которым приходится работать при детерминистическом описании, типа уравнений механики жидких сред или химической кинетики. Можно ожидать, что макроскопическое наблюдение над физико-химическими системами позволит получить значения, характерные для наиболее вероятного состояния. Поэтому можно потребовать, чтобы положения пиков $P(Q, t)$ являлись решениями детерминистических уравнений. Для одномодальных распределений, как на рис. 4.1, а, это в свою очередь означает, что уравнение для среднего значения близко к детерминистическому уравнению, причем поправка обратно пропорциональна некоторой степени размера системы.

Хотя эти условия и представляют определенный интерес, тем не менее их недостаточно для однозначного установления вида уравнения (4.9). Поэтому сейчас мы обратимся к некоторым интуитивным аргументам. Как уже подчеркивалось, флуктуации обусловлены случайным характером движения молекул, который связан с межмолекулярными взаимодействиями и большим числом частиц, образующих большинство рассматриваемых физико-химических систем. Поэтому флуктуации представляют собой локализованные события, имеющие весьма малые пространственные и временные масштабы. В этих условиях можно ожидать, что фигурирующие в уравнении (4.9) скорости переходов будут зависеть лишь от состояния Q и состояний, которые могут быть связаны с Q единственной флуктуацией. Иными словами, мы считаем, что в процессе последовательных переходов, обусловленных последовательными флуктуациями, сохраняется память лишь о последнем переходе. Этим условием дается определение чрезвычайно важного класса явлений, известного под названием *марковских процессов*. Этим процессам отводится значительное место в современных руководствах по теории вероятности.

На рис. 4.2 показана простая иллюстрация марковского процесса. Рассматриваемая система имеет пять состояний (I—V). Состояния I и V представляют собой «поглощающие границы»

в том смысле, что по достижении этих состояний система остается в них навсегда. С другой стороны, находясь в состояниях II, III или IV, система совершает быстрые скачки налево или направо, с вероятностями q или $1-q$ соответственно. В предельном случае, когда число состояний очень велико, эти «случайные блуждания» дают весьма реалистические представления о движении какой-либо частицы в жидкости и приводят к известному диффузионному закону Фика.

Разумеется, в системах со сложным поведением приходится иметь дело с гораздо более сложными ситуациями, обусловленными нелинейностями кинетики. Тем не менее можно выписать *основное уравнение* (4.9) и в явном виде. В нем скорость

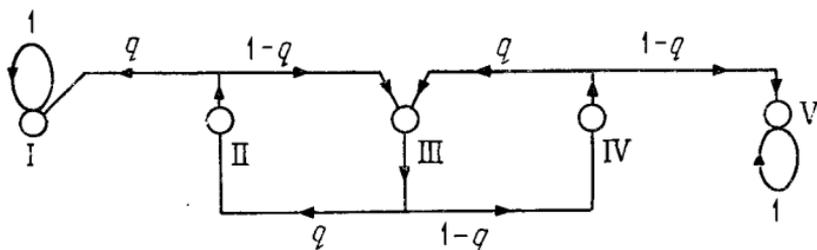


Рис. 4.2. Иллюстрация марковского процесса. Стрелки между состояниями I и V показывают направления возможных переходов.

перехода $R_+(Q)$ будет представлена в виде произведения вероятности перехода в единицу времени из состояния Q' в состояние Q и вероятности найти систему в состоянии Q' в момент времени t , причем это произведение суммируется по всем состояниям Q' , из которых можно перейти в Q за счет одного из элементарных динамических процессов, происходящих в системе. Аналогично величина $R_-(Q)$ представляет собой просуммированное по всем возможным состояниям Q' , достижимым из состояния Q , произведение вероятности нахождения в состоянии Q в момент времени t и вероятности перехода в единицу времени из состояния Q в состояние Q' . Это соотношение баланса можно записать в виде

$$\frac{dP(Q, t)}{dt} = \sum_{Q' \neq Q} [W(Q|Q')P(Q', t) - W(Q'|Q)P(Q, t)] \quad (4.12)$$

Здесь вероятность перехода в единицу времени $W(Q|Q')$ представляет собой неотрицательную величину для любых $Q' \neq Q$ и считается независимой от времени (стационарный марковский процесс).

Хотя уравнение (4.12) и линейно относительно неизвестной функции $P(Q, t)$, однако это не означает, что оно легко решается. В действительности сложность динамических законов

эволюции (например, обратные связи, имеющиеся в химических или биологических системах) отражается в наличии нелинейностей скоростей перехода W . Вблизи равновесия благодаря принципу детального равновесия эти нелинейности становятся несущественными, однако вдали от равновесия они проявляются в полной мере. Вследствие этого полный анализ уравнения (4.12) на сегодняшний день еще представляет нерешенную задачу. Тем не менее, пользуясь теорией возмущений, отыскивая простые случаи, допускающие точные решения, и прибегая к численному моделированию, все-таки удастся дать общее описание законов, управляющих флуктуациями в нелинейных неравновесных системах. Основные результаты таких исследований представлены в последующих разделах настоящей главы. В заключение этого раздела мы приведем несколько примеров, иллюстрирующих структуру основного уравнения.

Существует некоторый класс явлений, которые можно описать с помощью уравнения (4.12) — это *процессы рождения и гибели*. Сюда относятся явления, полностью характеризуемые набором целых чисел $\{X_\alpha\}$, которые могут либо оставаться неизменными, либо изменяться на единицу. Очевидно, примером такого процесса могут служить химические и биохимические реакции, в которых $\{X_\alpha\}$ представляет набор чисел молекул типа α . Рассматривая для простоты конкретный компонент X , можно написать

$$W(X|X') = \sum_{\rho} W_{\rho}(X|X-r_{\rho}) \delta_{X', X-r_{\rho}}, \quad r_{\rho} = 0, \pm 1 \quad (4.13a)$$

где ρ соответствует конкретной стадии реакции, а r_{ρ} представляет собой число молекул, потребляемых ($r_{\rho} < 0$) или возникающих ($r_{\rho} > 0$) в данном процессе. Скорости рождения и гибели, обозначенные соответственно λ и μ , определим следующим образом:

$$\begin{aligned} \lambda_X &= \sum_{\rho} W_{\rho}(X|X-1) \equiv N\lambda(X/N), \quad X \geq 0 \\ \mu_X &= \sum_{\rho} W_{\rho}(X|X+1) \equiv N\mu(X/N), \quad X \geq 1 \end{aligned} \quad (4.13b)$$

Стоящие после знака тождества выражения подчеркивают важную особенность стохастической динамики: ввиду короткодействующего характера межмолекулярных взаимодействий число переходов, совершающихся в единицу времени в системе размера N , пропорционально этому размеру. Иначе говоря, каждый малый элемент системы может чувствовать лишь своих соседей, так что, суммируя по всем элементам, мы получаем произведение N и некоторой функции интенсивной переменной X/N . Естественно, здесь подразумевается, что мы имеем дело со

взаимодействиями физико-химической природы. Иная ситуация может оказаться в случае динамики социальных систем. Благодаря различным видам телекоммуникации, средствам массовой информации и т. д. каждый элемент системы может одновременно ощущать все остальные элементы системы. Это могло бы привести к скоростям перехода, пропорциональным некоторым степеням N^a с показателем степени a , превышающим единицу.

Другую иллюстрацию структуры вероятностей перехода в единицу времени можно привести на примере автокаталитической модели (2.15). Нам нужно вычислить четыре величины W_1 , W_{-1} , W_2 , W_{-2} , описывающие соответственно скорости прямых и обратных реакций, имеющих в данной модели. Рассмотрим вначале величину W_1 . Реакция $A + 2X \xrightarrow{k_1} 3X$ происходит при встрече пары частиц X с частицей типа A . Поэтому, воспользовавшись комбинаторными соображениями для числа столкновений, имеем

$$W_1(X+1|X) = k_1 \frac{AX(X-1)}{N^2} \quad (\text{стадия } A + 2X \xrightarrow{k_1} 3X)$$

Здесь множитель $1/N^2$ выделен для того, чтобы все выражение соответствовало экстенсивной переменной. Аналогично

$$W_{-1}(X-1|X) = k_2 \frac{X(X-1)(X-2)}{N^2} \quad (\text{стадия } 3X \xrightarrow{k_2} A + 2X)$$

$$W_2(X-1|X) = k_3 X \quad (\text{стадия } X \xrightarrow{k_3} B)$$

$$W_{-2}(X+1|X) = k_4 B \quad (\text{стадия } B \xrightarrow{k_4} X)$$

Скорости рождения и гибели определяются соотношениями

$$\lambda_X = W_1 + W_{-2}, \quad \mu_X = W_{-1} + W_2$$

С учетом (4.13) и (4.12) основное уравнение, описывающее процессы рождения и гибели, можно записать в следующем виде:

$$\frac{dP(X, t)}{dt} = \lambda_{X-1} P(X-1, t) + \mu_{X+1} P(X+1, t) - (\lambda_X + \mu_X) P(X, t) \quad (4.14)$$

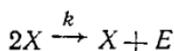
На примере этого явного уравнения полезно убедиться в наличии связи между стохастическим и макроскопическим описаниями. Умножая обе части (4.14) на X и суммируя по всем X , получаем

$$\frac{d\langle X \rangle}{dt} = \langle \lambda_X \rangle - \langle \mu_X \rangle$$

В то же время макроскопическое описание этой системы да-
ло бы

$$\frac{d\bar{X}}{dt} = \text{главная часть разности } (\lambda_{\bar{X}} - \mu_{\bar{X}}) \text{ в пределе больших значений } \bar{X}$$

Отсюда видно, что стохастическое и макроскопическое описания различаются в том случае, когда средние значения λ или μ отличаются от значений этих функций, вычисленных при $X = \langle X \rangle$. Именно это и происходит в нелинейных системах. Чтобы выяснить природу поправки к макроскопическому описанию, рассмотрим простой пример чистого процесса гибели:



Скорость гибели представляет собой величину, пропорциональную частоте парных встреч частиц X . Поэтому она может быть записана в виде

$$\mu_X = k \frac{1}{N} X(X-1)$$

Таким образом, имеем

$$\frac{d\langle X \rangle}{dt} = -k \frac{1}{N} \langle X(X-1) \rangle = -k \frac{1}{N} (\langle X \rangle^2 + \langle \delta X^2 \rangle - \langle X \rangle)$$

Переходя к интенсивной переменной $x = X/N$, это соотношение можно записать в виде

$$\frac{d\langle x \rangle}{dt} = -k \langle x \rangle^2 - \frac{k}{N} \frac{\langle \delta X^2 \rangle - \langle X \rangle}{N} \quad (4.15)$$

макроскопиче- поправка, обусловленная
ский закон флуктуациями

Как видно, отклонение от макроскопического закона связано с отличием действующего распределения вероятностей от пуассоновского, поскольку для последнего справедливо $\langle \delta X^2 \rangle = \langle X \rangle$ [см. уравнение (2.28)]. Этот вывод носит весьма общий характер и по существу не зависит от деталей кинетики. Возвращаясь к нашему примеру с одномодальным распределением, мы видим, что $\langle \delta X^2 \rangle - \langle X \rangle$ представляет собой величину порядка N , как это следует из центральной предельной теоремы. Тогда поправка к макроскопическому закону обратно пропорциональна размеру N и тем самым мала при больших N . Однако N не всегда бывает велико: в отличие от упрощающих предположений, приведших к выражениям (4.13) для переходных вероятностей, большинство флуктуаций в действительности представляют собой локализованные события, так что типичное значение N будет соответствовать области с характерными размерами по-

рядка масштаба флуктуаций. Это соображение показывает, что эффект флуктуаций не является универсальным — он существенно зависит от их масштаба.

В действительности потребность в *локальном подходе* к описанию флуктуаций имеет гораздо более глубокие корни. Для того чтобы можно было применять марковское описание, важно выбрать переменные состояния таким образом, чтобы каждый отдельный переход между состояниями можно было отнести непосредственно за счет отдельной «элементарной» флуктуации. В глобальном описании это будет не обязательно так, и соответствующий стохастический процесс может оказаться не марковским. Поэтому при написании уравнения (4.12) имеется в

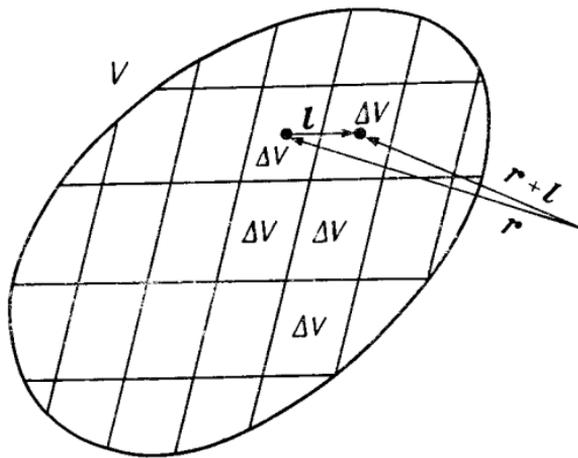


Рис. 4.3. Разбиение макроскопической системы на малые объемы ΔV с центрами в точках r , взаимодействующие со своими соседями $r+l$ за счет переноса массы, энергии или импульса. Благодаря локальному характеру флуктуаций можно ожидать, что при должном выборе ΔV процесс, состоящий в отслеживании эволюции переменных состояния на множестве всех ΔV , будет марковским.

виду, что отдельные процессы типа химической реакции происходят в каждом малом объеме системы ΔV (см. рис. 4.3), а результат затем суммируется по всему объему V . Разумеется, различные элементы объема будут статистически связаны между собой за счет переноса частиц, импульса или энергии. Поэтому в нашем примере процессов рождения и гибели глобальное уравнение (4.14) должно быть дополнено членами, соответствующими этим процессам переноса.

До сих пор мы рассуждали на языке дискретных переменных и дискретных наборов состояний. В физической системе это и на самом деле соответствует фундаментальному уровню описания, поскольку, в конце концов, флуктуации возникают вследствие дискретного строения материи. Однако во многих случаях

дискретные стохастические процессы можно свести (в некотором асимптотическом смысле) к стохастическому процессу с непрерывными реализациями. В этом пределе описание становится очень похожим на макроскопическое, с той разницей, что теперь к макроскопической скорости перехода добавляется *стохастическая сила*, отражающая влияние флуктуаций. Это описание, восходящее к классической работе Ланжевена по броуновскому движению, во многих случаях оказывается плодотворным подходом к стохастическому анализу широких классов динамических систем.

4.3. Марковские процессы и необратимость. Информационная энтропия и физическая энтропия

Читателю может показаться поразительным сходство между разд. 4.1—4.2, с одной стороны, и разд. 3.2 — с другой. В самом деле, в обоих случаях используется вероятностный язык, и эволюция соответствующего распределения вероятностей описывается линейным уравнением первого порядка по времени — достаточно взглянуть на основное уравнение (4.12) и уравнение Лиувилля (3.7). Хотя между этими двумя способами описания и имеется аналогия, мы покажем, что статус марковских процессов совершенно отличен. Это обстоятельство можно было заметить уже из установленного в предыдущем разделе соотношения между основным уравнением и макроскопическими выражениями для скоростей переходов в диссипативных системах — свойство, прямо противоположное инвариантности относительно обращения времени, присущей уравнению Лиувилля. Теперь ясно, что связь между марковскими процессами и диссипативными системами является общей особенностью основного уравнения, коренящейся в самой его структуре. Более конкретно, мы покажем, что по мере стремления времени к бесконечности решения основного уравнения стремятся к единственному конечному стационарному состоянию, обусловленному *функцией Ляпунова* весьма сходным образом с термодинамическими функциями, рассмотренными в разд. 2.6.

Для начала мы введем *энтропию* марковского процесса. Потребность в такой величине возникла в теории вероятностей уже довольно давно. В особенности это стало ясно с появлением допускающих точное решение вероятностных моделей, имеющих целью примирить необратимую релаксацию к равновесию, предсказываемую кинетической теорией Больцмана, с обратимостью законов движения консервативных систем. С появлением теории информации энтропия стала также важным понятием математической теории сообщений. Этот информационно-теоретический аспект энтропии будет рассмотрен полнее в дальнейшем

(разд. 4.7). Здесь мы заметим лишь, что как в физико-химическом плане, так и в информационно-теоретическом контексте обычно ищется величина, удовлетворяющая ряду требований, которые следуют как из интуитивных соображений, так и по аналогии с термодинамикой. Эти требования выглядят следующим образом:

1. При данном числе состояний N и заданных вероятностях $P(Q) \geq 0$, $\sum_Q P(Q) = 1$ соответствующая энтропии функция S принимает максимальное значение при $P(Q) = 1/N$.

2. Энтропия составной системы AB равна энтропии подсистемы A плюс условная энтропия $S_A(B)$ подсистемы B при условии, что подсистема A находится в определенном состоянии:

$$S(AB) = S(A) + S_A(B)$$

Это свойство обычно называется *субаддитивностью*.

3. Формальное добавление невозможного состояния Q_α ($P(Q_\alpha) = 0$) не изменяет энтропии рассматриваемого процесса.

Если потребовать, чтобы S была непрерывной функцией вероятностей P , то можно показать, что при выполнении перечисленных условий функция S определяется единственным образом с точностью до постоянного множителя:

$$S(P(Q), t) = - \sum_Q P(Q, t) \ln P(Q, t) \quad (4.16)$$

В таком виде это выражение не соответствует функции Ляпунова в наиболее общем случае марковского процесса, описывающего неравновесную систему. Чтобы получить такую функцию, надо ввести дополнительное предположение о существовании стационарного распределения вероятностей для рассматриваемого марковского процесса, $P_s(Q)$, тождественно удовлетворяющего основному уравнению (4.12) при $dP_s/dt = 0$. Далеко не все марковские процессы удовлетворяют этому условию. В самом деле, из теории вероятностей известно, что это свойство представляет собой замечательную особенность стохастических процессов с *эргодическими состояниями*. Эргодичность состояний означает, что во всех состояниях система бывает с вероятностью единица, и *время возвращения* в любое конкретное состояние, отсчитываемое с момента ухода из него, конечно.

Воспользовавшись вероятностями $P_s(Q)$, можно построить следующую величину:

$$K(P(Q), t) = \sum_Q P(Q, t) \ln \frac{P(Q, t)}{P_s(Q)} \quad (4.17)$$

Очевидно, эта величина ограничена снизу и обращается в нуль

при $P = P_s$. Изменение ее со временем можно рассчитать с помощью уравнения (4.12):

$$\frac{dK}{dt} = \sum_{\substack{Q, Q' \\ (Q \neq Q')}} [W(Q|Q')P(Q', t) - W(Q'|Q)P(Q, t)] \ln [P(Q, t)/P_s(Q)]$$

Меняя местами индексы Q и Q' , а также с учетом положительности переходных вероятностей $W(Q|Q')$ при $Q \neq Q'$ и известного соотношения $\ln x \leq x - 1$, справедливого для любого действительного числа x , можно получить

$$\frac{dK}{dt} \leq 0 \quad (4.18)$$

Этот результат, иллюстрируемый рис. 4.4, очень похож на результаты разд. 2.5 и 2.6 (см. в особенности рис. 2.7) относительно монотонного поведения энтропии изолированной системы, предсказываемого вторым законом термодинамики. Следует, однако, отметить, что функция K не совпадает с энтропией.

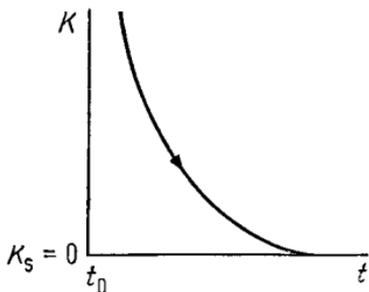


Рис. 4.4. Необратимость марковского процесса, описываемого уравнением (4.12). Избыточная энтропия, описываемая функцией K [уравнение (4.17)], монотонно уменьшается со временем.

Более того, неравенство (4.18) с тем же успехом применимо к неизолированным системам, подверженным постоянному воздействию неравновесных условий. И все же, подобно тому, как второй закон термодинамики выражает универсальное стремление изолированной системы к единственному равновесному состоянию, неравенство (4.18) также выражает единственность и глобальную устойчивость P_s в том смысле, что из любого начального состояния величина $P(Q, t)$ непременно стремится к P_s при $t \rightarrow \infty$. Иными словами, проблема неустойчивости и множественности решений, с которой мы знакомы по обсуждениям бифуркаций в нелинейных динамических системах, попросту не возникает на вероятностном уровне. С первого взгляда на рис. 4.1 становится ясно, почему так и должно быть: бифуркация отражается в качественном изменении структуры соответствующего распределения вероятностей, например в виде появления множественных максимумов, но никак не в том, что распределение вероятностей само становится множественным. Сле-

дует, однако, отметить одно исключение, возникающее в предельном случае, когда размер системы стремится к бесконечности. Тогда ширина каждого вероятностного пика обращается в нуль, а распределение вероятностей распадается на ряд бесконечно узких распределений [мы уже имели аналогичный пример в виде уравнения (4.6)]. Теперь уже различные состояния такого марковского процесса не образуют эргодического множества, а различные начальные условия приводят к различным конечным распределениям вероятностей. Можно показать, что такая ситуация адекватно соответствует как равновесным, так и неравновесным переходам в бесконечно протяженных системах.

Другое важное следствие неравенства (4.18) состоит в следующем: на его основе можно показать, что уравнение (4.12) равносильно своеобразному *сокращению* пространства распределения вероятностей, поскольку при любых особенностях динамики «расстояние» от этих распределений до $P_s(Q)$ непременно уменьшается. Это свойство наряду с линейностью основного уравнения лежит в основе глубокой связи между марковскими процессами и важным математическим понятием о *диссипативных подгруппах*. Чтобы получить интуитивное представление об этой связи, сравним следующие два динамических уравнения:

$$\frac{dy}{dt} = -\lambda y \quad (t > 0, \lambda > 0) \quad (4.19a)$$

$$\frac{d^2x}{dt^2} = -\omega_0^2 x \quad (4.19b)$$

Эти уравнения описывают соответственно необратимое стремление диссипативной системы к конечному асимптотически устойчивому состоянию $y=0$ (см., например, разд. 2.2) и типичную обратимую во времени динамику, характерную, например, для гармонического осциллятора (см. разд. 3.3). Вместо второго из уравнений (4.19) можно также написать соответствующее уравнение Лиувилля на языке переменных действие — угол [см. равенство (3.7) и разд. 3.3]:

$$\frac{\partial \rho}{\partial t} = -\omega \frac{\partial \rho}{\partial \Phi} \quad (4.20)$$

Решения кажущихся похожими уравнений (4.19a) и (4.20) в действительности совершенно различны. Для (4.19a) имеем попросту

$$y(t) = e^{-\lambda t} y(0) \quad (4.21a)$$

В то же время решение уравнения (4.20) должно учитывать то

обстоятельство, что переменная φ представляет собой угол. По этой причине приходится ввести мнимую единицу $i = \sqrt{-1}$:

$$\rho(t) = e^{i(\omega t - \varphi)} \rho(0) \quad (4.21б)$$

Такое решение действительно инвариантно при изменении угла φ на $\pm 2\pi$, $\pm 4\pi$, ...

В обоих уравнениях (4.21) имеется некоторый «эволюционный оператор», дающий значение функции в данный момент времени в результате действия на начальное значение. Этот оператор имеет вид

$$U_t = \exp(tL) \quad (4.22а)$$

Он удовлетворяет так называемому свойству группы

$$U_t U_{t'} = \exp((t + t')L) = U_{t+t'} \quad (4.22б)$$

Однако в (4.21а) $L = -\lambda$ представляет собой отрицательное действительное число. Это означает, что поведение U_t регулярно лишь на положительной полуоси, где переменная состояния монотонно стремится к единственному стационарному состоянию. Выражаясь более абстрактным языком, можно сказать, что эволюционный оператор U_t определяет диссипативную полугруппу. С другой стороны, в уравнении (4.21б) абсолютное значение ρ (или, более абстрактно, его «мера») остается неизменным: $|\rho(t)| = |\rho(0)|$. В таком случае мы будем говорить, что соответствующий оператор U_t определяет *унитарную группу*. Заметим, что в этом, последнем случае инвариантность динамических уравнений при обращении времени выражается следующим образом: $U_t = U_{-t}^*$. Иными словами, при обращении времени необходимо брать комплексно-сопряженный оператор U_t .

Теперь уже по недолгом размышлении можно убедиться в том, что приведенные выше две модели на идейном уровне сходны с основным уравнением и уравнением Лиувилля для консервативной системы. Единственное различие сводится к тому факту, что соответствующие величины L в экспонентах в (4.22а) теперь уже не числа, а операторы. Таким образом, можно утверждать, что *признаком необратимости* является возникновение диссипативных полугрупп, описывающих соответствующим образом определенный марковский процесс. В свою очередь это приводит к одному из наиболее глубоких вопросов в физике: может ли такой марковский процесс появиться при обратимой во времени динамике? Если ответ на этот вопрос окажется положительным, то переход от унитарной группы к диссипативной полугруппе будет эквивалентен *нарушению временной симметрии*. Тем самым консервативные системы оказались бы «предшественниками» диссипативных систем, проявляющих эту

специальную разновидность нарушения симметрии, подобно тому, как однородное стационарное состояние системы с химической реакцией является «предшественником» диссипативной структуры, соответствующей какому-то режиму или ритму выше подходящей точки бифуркации. Обсуждению этой фундаментальной проблемы посвящается гл. 5.

4.4. Пространственные корреляции и критическое поведение

Обсудив основные особенности вероятностного описания динамических систем, теперь можно рассмотреть, как используются соответствующие понятия при решении различных вопросов, относящихся к возникновению сложного поведения.

Как уже неоднократно отмечалось, переходные явления вдали от равновесия связаны с состояниями, характеризующимися корреляциями макроскопических масштабов. В разд. 2.8 мы отмечали, что такие флуктуации могут возникать лишь в системах, в которых флуктуации не носят характер пуассоновских вследствие поддержания соответствующих неравновесных условий. Более того, в разд. 2.8 мы пошли еще дальше, установив связь между согласованностью, вызванной бифуркацией, и нарушением центральной предельной теоремы наряду с соответствующим усилением флуктуаций. Так или иначе, следует четко понимать, что макроскопическая согласованность и корреляции, о которых идет речь, имеют мало общего с силами межмолекулярного взаимодействия, поскольку эти макроскопические понятия применимы и в случае разбавленных растворов, в которых молекулы не способны распознавать друг друга. Каким образом такое довольно необычное явление оказывается возможным? В данном разделе мы постараемся ответить на этот вопрос, воспользовавшись свойствами основного уравнения и апеллируя как к аналогиям, так и к различиям между этим явлением и равновесными фазовыми переходами.

Как подчеркивалось в разд. 1.8, равновесный фазовый переход типа замерзания или спонтанного намагничивания является результатом конкуренции между силами межмолекулярного взаимодействия, стремящимися упорядочить систему, и случайным тепловым движением молекул, оказывающим противоположное воздействие. Вот почему обычно равновесные переходы вызываются повышением давления (усиливающим влияние межмолекулярных взаимодействий) или понижением температуры (ослабляющим влияние теплового шума). По существу даже вдали от фазового перехода межмолекулярные силы сообщают сравнительно плотной системе определенную степень упорядоченности. Правда, происходит это лишь в масштабах короткодействия, поскольку этот эффект целиком обусловлен

масштабом действия межмолекулярных сил (см. рис. 1.18). На рис. 4.5 приведена полезная иллюстрация такой ситуации.

Рассмотрим некоторую «стандартную систему» в состоянии термодинамического равновесия, когда можно пренебречь межмолекулярным взаимодействием. Примерами таких систем могут служить идеальный газ или идеальный раствор (рис. 4.5, б). Как было показано в разд. 2.8, в такой системе флуктуации числа частиц в заданном элементе объема носят пуассоновский характер, что является отражением полной разупорядоченности, обусловленной тепловым движением соответствующих частиц. Допустим теперь, что в какой-то момент времени включаются силы притяжения. Тогда частицы будут стремиться к образованию кластеров (рис. 4.5, в), что приведет к увеличению

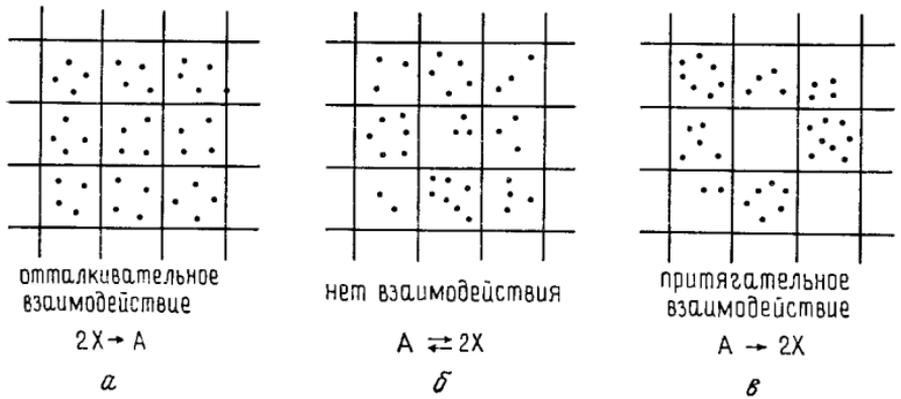


Рис. 4.5. К вопросу о природе пространственных корреляций макроскопического масштаба в идеальной неравновесной системе.

отклонения их пространственного распределения от среднего значения, т. е. мы получим $\langle \delta X^2 \rangle > \langle X \rangle$. С другой стороны, если взаимодействие будет отталкивательным, то частицы распределятся по объему системы более равномерно (рис. 4.5, а), что уменьшит отклонение от среднего значения. В обоих случаях пространственная «упорядоченность», обусловленная непуассоновским поведением флуктуаций, носит короткодействующий характер, если только система не находится вблизи точки фазового перехода, где хвост пространственных корреляций тянется довольно далеко. Удлинение масштаба корреляций отражает способность к совершению фазового перехода всей системой в целом.

Рассмотрим теперь некую идеальную систему, способную функционировать и вдали, и вблизи от равновесия. На первый взгляд можно было бы ожидать, что описанные выше эффекты кластеризации и критического поведения не будут наблюдаться в ней ни при каких условиях. В случае равновесия в этом легко

убедиться непосредственно, так как распределение вероятностей является пуассоновским и по этой причине пространственные корреляции отсутствуют. Но что же происходит с системой, которая за счет подбора соответствующих условий значительно отклоняется от равновесия? Для конкретности рассмотрим реакцию $A \rightleftharpoons 2X$ и допустим вначале, что A имеется в значительном избытке, так что стадией $2X \rightarrow A$ можно пренебречь. Во избежание бесконечного накопления вещества X введем также реакцию удаления этого вещества из среды. Таким образом, получаем следующие реакции:



Очевидно, первая из этих реакций равносильна образованию в среде кластеров частиц типа X . Иными словами, нелинейная химическая кинетика приводит к эффекту, напоминающему притяжение между молекулами. Согласно принципу детального равновесия, любая реакция, в частности такая, как $A \rightarrow 2X$, в среднем происходит столь же часто, как и противоположная ей, т. е. $2X \rightarrow A$. Поэтому описанная выше кластеризация будет происходить в условиях противодействия реакции $2X \rightarrow A$, устраняющей пары частиц в системе и тем самым играющей роль аналога отталкивания. Это приводит к восстановлению известного пуассоновского поведения. Аналогично, если бы все реакции были линейными, частицы создавались бы и разрушались в произвольных точках пространства, и в результате поведение системы снова было бы пуассоновским. Наличие одновременно *нелинейной кинетики* и *неравновесных условий* нарушает этот баланс, и в результате возникают систематические отклонения от закона Пуассона. Как связать эти отклонения с существованием пространственных корреляций? В реакции $A \rightarrow 2X$ обе частицы X , очевидно, коррелированы, поскольку они образуются вместе. Благодаря диффузии эти корреляции приобретают пространственную протяженность, растущую до тех пор, пока по крайней мере одна из частиц не выключается из игры за счет участия во второй из реакций (4.23). Таким образом, результирующая длина корреляции может быть оценена следующим образом: скорость убыли X за счет участия во второй из реакций (4.23) приблизительно равна диффузионной скорости на расстоянии порядка корреляционной длины, т. е.

$$kx \sim \frac{D}{l_{\text{corr}}^2} x \quad (4.24)$$

или

$$l_{\text{corr}} = \sqrt{\frac{D}{k}}$$

Здесь D — коэффициент диффузии компонента X .

Заметим, что если обозначить скорость изменения X , вытекающую из схемы (4.23), посредством $F(X)$, т. е. $F(X) = k_1A - k_2X$, то величина k_2 просто равна $|F'(X_s)|$, что в соответствии с разд. 3.6 и приложением I дает скорость затухания возмущения γ вблизи стационарного состояния $X_s = k_1A/k_2$. Этот вывод носит довольно общий характер: в любой системе, где имеют место транспортные процессы диффузионного типа и локальные процессы типа химических, для которых самая медленная мода описывается скоростью затухания γ , длина корреляции определяется по формуле

$$l_{\text{corr}} \simeq \sqrt{\frac{D}{\gamma}} \quad (4.25)$$

Как неоднократно отмечалось в гл. 3, в окрестности точки бифуркации коэффициент γ стремится к нулю, в результате чего радиус корреляции неограниченно возрастает. Такое возрастание указывает на тенденцию к установлению согласованности во всем объеме системы, которая тем самым приобретает способность к коллективному переходу в новое состояние. Заметим, однако, что вследствие такого «замедления» динамики вблизи точки бифуркации теряется возможность описания системы на языке отдельных мод и их скоростей релаксации. Для этого необходима более полная теория, учитывающая *нелинейное сопряжение* между модами. Такой подход был разработан с использованием как основного уравнения в полном виде, так и предельного случая, соответствующего непрерывным процессам. Основные качественные выводы, сделанные на этом пути, состоят в следующем: во-первых, вследствие взаимодействия между отдельными модами меняется закон возрастания радиуса корреляции по сравнению с формулой (4.25) и, во-вторых, переход может стать вообще невозможным.

Чтобы лучше понять этот самый удивительный момент, следует ввести понятие *критической размерности*. Вспомним, что мы имеем дело с системами, в которых сосуществуют локальные процессы (например, типа автокаталитической химической реакции) и транспортные процессы, осуществляющие связь соседних элементов объема (см., например, рис. 4.3). За счет флуктуаций процесс химического типа будет стремиться отклонить систему от пространственной однородности. В предельном случае отсутствия связи между элементами объема это привело бы к случайному набору комбинаций локализованных состояний. При этом, даже если в каждом элементе объема параметры близки к своим точкам бифуркации, система просто не «замечит» бифуркации ввиду преобладания в ней шума. В противоположном же случае, когда пространственные ячейки между собой связаны, но не процессами химического типа, простран-

ственная однородность сохранится, однако упорядоченность не возникнет, поскольку попросту не будет того «сообщения», которое требуется передать и сохранить. Отсюда следует вывод, что оба процесса необходимы и оба одинаково важны, точнее, оба должны иметь близкие характерные времена. Как было показано раньше, характерное время выравнивания неоднородности за счет транспорта в пространственной ячейке с линейным размером L дается формулой

$$\tau_D \simeq L^2 D^{-1}$$

или (воспользовавшись соотношением $L^d = \Delta V$, где d — размерность пространства, в котором существует рассматриваемая система)

$$\frac{1}{\tau_D} = D \Delta V^{-2/d} \quad (4.26)$$

Если пространственная когерентность устанавливается на макроскопических масштабах, то ΔV велико, и, следовательно, диффузия будет медленной. Более того, такое замедление будет более значительным по мере уменьшения размерности пространства. С другой стороны, как отмечалось ранее, вблизи точки бифуркации химические процессы будут также замедляться, и соответственно их характерное время τ_{ch} станет большим. Однако в отличие от диффузионного это замедление не будет зависеть от размерности в силу локального характера процесса. Фактически при фиксированном объеме пространственной ячейки ΔV это время будет равно обратной скорости релаксации самой медленной моды и тем самым будет зависеть лишь от параметров и степеней основных нелинейностей. Исходя из условия, что τ_D и τ_{ch} должны быть одного и того же порядка величины, мы получаем возможность связать размерность с внутренними характеристиками системы — так и определяется критическая размерность d_c . При $d \geq d_c$ диффузия оказывается действенной, и в результате будут существовать крупномасштабные пространственные корреляции. Однако, если d значительно меньше, чем d_c , диффузия окажется неспособной достаточно хорошо коррелировать различные пространственные ячейки и бифуркация будет «пропущена».

Конкретная структура системы довольно слабо влияет на значение d_c , поскольку наиболее существенным фактором представляются некоторые особенности довольно общего характера типа степени доминирующей нелинейности. На этом пути можно дать универсальную классификацию совершенно различных систем, основываясь на качественно аналогичных особенностях поведения. Например, как в автокаталитической модели (2.19), так и в Брюсселяторе (3.18) для некоторых бифуркаций крити-

ческая размерность оказывается равной четырем ($d_c=4$). Заметим, однако, что вдали от точки бифуркации размерность перестает играть столь важную роль. Например, вычисление радиуса корреляции по формуле (4.25) будет допустимо при любых значениях d .

Для того чтобы наше изложение могло выйти за пределы этих качественных результатов, необходимо найти решения основного уравнения (4.12) в явном виде. Мы не будем обсуждать здесь техническую сторону этой проблемы и ограничимся некоторыми, наиболее яркими моментами.

1. Коль скоро система не находится строго в состоянии равновесия, корреляции макроскопического масштаба возникают внезапно и могут существовать неопределенно долго. В системах, где химическая стадия сопряжена с транспортной, масштаб корреляции определяется внутренними параметрами и по существу описывается формулой (4.25). Точный вид зависимости корреляции от пространственной координаты зависит как от геометрии, так и от внутренних параметров. В трехмерной системе с большими линейными размерами корреляция между точками с координатами \mathbf{r} и \mathbf{r}' [тесно связанная с ковариацией (4.2)] определяется соотношением

$$g(\mathbf{r}-\mathbf{r}') = \frac{J_s}{2\pi D} \frac{1}{|\mathbf{r}-\mathbf{r}'|} \exp \left[- \left(\frac{|F'(X_s)|}{D} \right)^{1/2} |\mathbf{r}-\mathbf{r}'| \right] \quad (4.27)$$

Амплитуда корреляции J_s прямо связана со степенью отклонения от равновесия. Поэтому в некотором смысле амплитуду J_s можно рассматривать как «параметр порядка», характеризующий «переход» от некоррелированного состояния вещества (равновесия) к коррелированному состоянию (неравновесию). Заметим, что сам по себе радиус корреляции остается конечным в равновесии. С другой стороны, в системах, где нет отмеченной выше связи между различными типами процессов, радиус корреляции может определяться внешними условиями. Например, если в теплопроводной среде создать разность температур, то можно будет видеть, что радиус корреляции существенно определяется размером системы. Однако амплитуда корреляции будет по-прежнему связана со степенью отклонения от равновесия (в данном случае — от потока энергии через систему) и опять будет играть роль параметра порядка. Качественно эта ситуация описана на рис. 4.6.

2. Как уже отмечалось, если только размерность не слишком низка, по мере приближения к точке бифуркации радиус корреляции неограниченно возрастает, что соответствует распространению корреляции на всю систему. Это приводит к качественному изменению особенностей соответствующего распределения вероятностей, несколько напоминающему переход от рис. 4.1, а

к рис. 4.1, б, с той разницей, что теперь мы имеем дело с многомерными распределениями, зависящими от значений стохастических переменных X в каждой точке пространства r . Например, в бесконечно протяженной системе типа автокаталитической модели, описываемой уравнением (2.15), совершающей бифуркационный переход к множественным стационарным состояниям без нарушения пространственной симметрии, оказывается, что вблизи особой точки типа острья (рис. 3.10, а) стационарное распределение вероятности имеет вид

$$P_S(\{x_r\}) = \exp \left\{ \frac{2\Delta V}{Q_S} \sum_r \left[\frac{(\lambda - \lambda_c)}{2} (x_r - x_S)^2 - \frac{1}{4} (x_r - x_S)^4 + \frac{\mathcal{D}}{8d} \sum_l (x_{r+l} - x_r)^2 \right] \right\} \quad (4.28)$$

Здесь $\lambda - \lambda_c$ — расстояние от точки бифуркации, \mathcal{D} — константа скорости перескока между соседними элементами объема ΔV ,

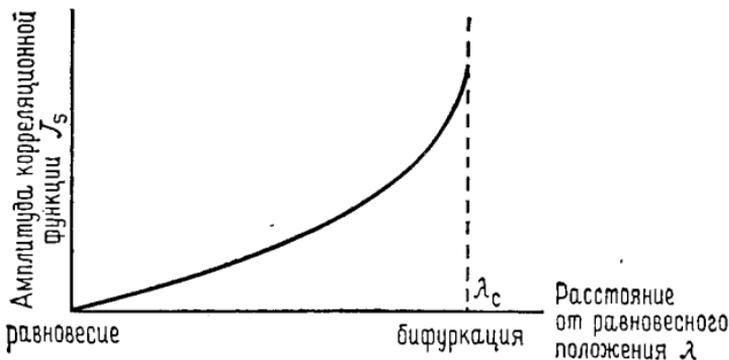


Рис. 4.6. В неравновесной системе возникают пространственные корреляции макроскопического масштаба, причем амплитуда этих корреляций возрастает по мере удаления от равновесия.

связанная с коэффициентом диффузии D , Q_S — некоторая положительная комбинация параметров системы, x_r — локальное значение переменной состояния X_r , приведенной к величине данного элемента объема ΔV ($x_r = X_r/\Delta V$). Индексом l отмечены элементы, соседние с ячейкой r , с которыми обмен веществом X осуществляется лишь в соответствии с имеющимися в данной системе транспортными возможностями.

Тот факт, что это распределение не является гауссовым, отражает влияние бифуркации — систему уже невозможно разбить на множество слабо взаимодействующих подсистем. Аналогичные результаты имеются для бифуркаций, приводящих к предельным циклам и пространственным структурам, связанным с нарушением симметрии.

Интересно рассмотреть пример неприменимости центральной предельной теоремы, как это имеет место в случае зависимости дисперсии флуктуаций числа частиц в некотором конечном объеме от величины этого объема V . Вычисление дисперсии с помощью (4.28) в режиме маргинальной устойчивости, соответствующей переходу в новое состояние, дает следующий результат:

$$\langle \delta X^2 \rangle_V \sim V^{3/2}$$

Этот результат следует сопоставить с зависимостью $\langle \delta X^2 \rangle_V \sim V$, характерной для гауссова режима, который имеет место ниже точки бифуркации.

Следует отметить, что экспонента в (4.28), известная как потенциал Ландау — Гинзбурга, широко распространена в теории равновесных фазовых переходов. Воспользовавшись этой аналогией, можно вычислить многие характеристики системы, такие, как радиус корреляции, дисперсию флуктуаций и т. д. Для этого полезно применить метод *ренорм-группы*, предложенный Уилсоном.

В итоге мы приходим к простой и понятной картине возникновения упорядоченности в системе. Выражаясь несколько антропоморфически, можно сказать, что упорядоченность выглядит как своего рода компромисс между двумя антагонистическими факторами — нелинейным процессом типа химического, непрерывно и несогласованно посылающим инновационные сигналы в виде флуктуаций, и процессом типа транспортного, который улавливает, передает и стабилизирует эти сигналы. Нарушение тонкого равновесия между этими двумя «действующими лицами» приводит к таким качественным изменениям, как хаотическое состояние, в котором каждый элемент системы действует сам по себе, или, наоборот, «гомеостатическое», застывшее состояние, в котором флуктуации подавлены и установилась полная однородность.

Таким образом, сложное представляется с двух сторон ограниченным двумя различными видами разупорядоченности. Трудно избежать впечатления, что этот вывод, следующий из весьма общих законов динамики физико-химических систем, должен послужить ценным ориентиром при моделировании систем, выходящих за рамки физики и химии.

Этой проблеме «переноса знаний» будет посвящена гл. 6.

4.5. Поведение флуктуаций во времени.

Кинетика и временные масштабы самоорганизации

К концу разд. 4.1 мы установили две основные причины, мотивирующие исследование флуктуаций в нелинейных динамических системах. Первая из них, проясняющая механизм возник-

новения и поддержания согласованности, привела к изложенному в разд. 4.4 анализу. Вторая причина, а именно кинетика переходов между различными состояниями, составит предмет данного раздела.

Отправным пунктом для нас служит уравнение (4.15), которое полезно переписать в более общем виде:

$$\frac{d\langle x \rangle}{dt} = (\text{макроскопическое выражение для скорости}) + \quad (4.29) \\ + (\text{отклонение от пуассоновского поведения})$$

Как было показано выше, вблизи точки бифуркации поведение корреляционной функции становится аномальным и отклонение от пуассоновского распределения усиливается вследствие возрастания характерной длины корреляции. Поэтому уже на этой стадии следует рассмотреть зависимость критической моды от времени в окрестности точки неустойчивости и выше этой точки. По мере роста радиуса корреляции отклонение дисперсии от значения, соответствующего пуассоновскому распределению, становится все более значительным. Таким образом, второй член в правой части уравнения (4.29) в малых элементах объема ΔV может оказаться заметным, и по истечении достаточно большого (возможно, очень большого) времени он может привести к значительному отклонению среднего значения, фактически к новому макроскопическому режиму.

На рис. 4.7 приводится иллюстрация такого теоретического представления о неравновесных переходах, сходных с *зародышеобразованием* и предложенных впервые авторами этой книги. Напомним, что мы сейчас изучаем автокаталитическую модель (2.15). Исходно мы имеем распределение вероятностей с максимумом вблизи одного из устойчивых (на детерминистическом уровне) решений, возникающих после бифуркации. На рисунке показаны три различных способа представления такой ситуации.

На рис. 4.7, *а*, который следует сравнить с рис. 3.10, *б*, множественные стационарные состояния изображены в зависимости от характеристического параметра μ , представляющего собой некоторую комбинацию из констант скоростей и управляющих химических переменных A и B . Верхняя и нижняя ветви (x_+ и x_- соответственно) асимптотически устойчивы, средняя же ветвь x_0 неустойчива. На рис. 4.7, *б* представлен график некоторой полезной величины, с которой читатель имел недолгую встречу в разд. 2.6. Эта величина, называемая *кинетическим потенциалом*, определяется соотношением

$$\frac{dx}{dt} = - \frac{\partial U(x)}{\partial x} \quad (4.30)$$

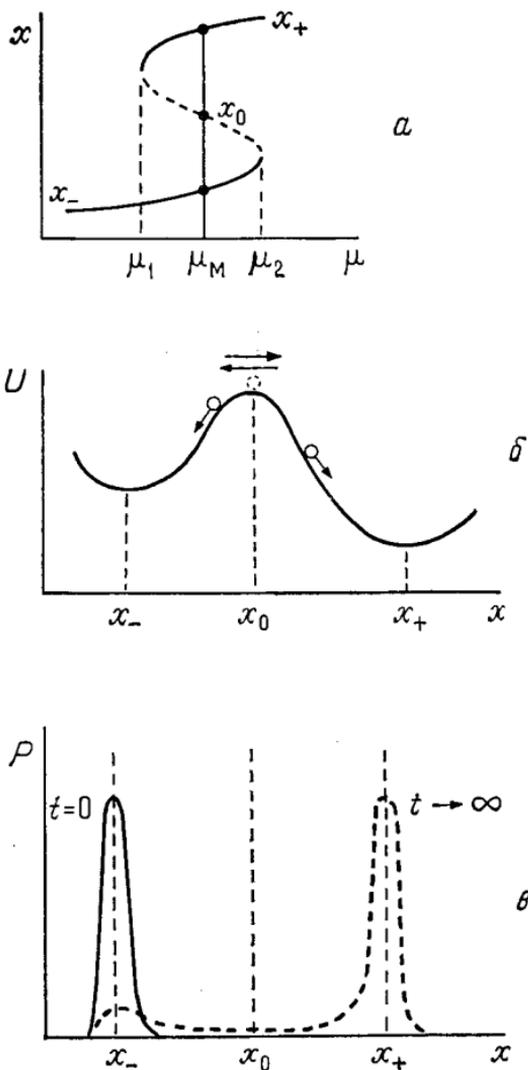


Рис. 4.7. Индуцированные флуктуациями переходы между одновременно устойчивыми стационарными состояниями. *a* — зависимость решений детерминистических уравнений от параметра μ . Предполагается, что при $\mu = \mu_M$ состояние x_- менее устойчиво, нежели состояние x_+ ; *б* — кинетический потенциал U (4.30) в зависимости от переменной состояния x . Случайные флуктуации вокруг каждого из устойчивых аттракторов x_+ , x_- затухают, если только они не переводят систему в другой бассейн притяжения через потенциальный барьер в неустойчивом состоянии $x = x_0$. Большая глубина U при $x = x_+$ соответствует предположению о большей устойчивости x_+ по сравнению с x_- ; *в* — распределение вероятностей, исходно сосредоточенное около $x = x_-$, в конце концов принимает бимодальный вид, причем горб, соответствующий x_+ , гораздо выше горба при x_- . Временной масштаб такого перехода определяется соотношением (4.31).

В системах с единственной переменной величина U всегда существует и равна взятому с обратным знаком интегралу от выражения для скорости изменения x :

$$U = - \int F(x) dx$$

С другой стороны, в системах со многими переменными в общем случае гарантировать существование кинетического потенциала невозможно. Так или иначе, если только кинетический потенциал удастся построить, он будет весьма полезным, поскольку его минимумы и максимумы дают соответственно устойчивые и неустойчивые решения макроскопических динамических уравнений. И наконец, сплошной линией на рис. 4.7, *в* представлено исходное распределение вероятностей с центром в состоянии x_- .

Займемся теперь описанием основных результатов. Допустим, что в системе можно поддерживать пространственную однородность в элементе объема ΔV , содержащем N частиц химически активных компонентов, ответственных за бистабильность. Как объяснялось выше, следует ожидать, что характерный размер ΔV будет связан с радиусом корреляции. Упрощенное описание такой ситуации можно получить с помощью решения основного уравнения для однородной системы размера N , пренебрегая диффузией. Это даст возможность получить переход на вторую, устойчивую ветвь x_+ (рис. 4.7, *а*), для которого переходное время определяется величиной N и разностью значений U между исходным состоянием x_- и неустойчивым состоянием x_0 . Конечное распределение вероятностей представлено на рис. 4.7, *в* штриховой линией для такого случая, когда в начальном состоянии U имеет гораздо более мелкий минимум, нежели x_+ .

Механизм описанного выше явления тесно связан с теорией энергии активации, с которой мы уже слегка соприкасались в конце разд. 1.4. Благодаря флуктуациям различные части системы испытывают спонтанные отклонения от исходного состояния x_- , соответствующего наиболее вероятному состоянию. Большинство из этих флуктуаций малы по амплитуде и масштабу, однако если размер рассматриваемой подсистемы не слишком велик, то флуктуации конечного масштаба будут также возникать с некоторой вероятностью. Некоторые из них выведут систему за «потенциальный барьер», образованный неустойчивым состоянием x_0 . Из этой точки подсистема совершает перескок в другое притягивающее состояние x_+ , образуя тем самым «зародыш» неустойчивости. Поскольку по предположению состояние x_+ более устойчиво, чем состояние x_- , вероятность ухода из него будет мала, а тем временем новые элементы системы будут совершать такие же переходы, и в конечном счете вся система окажется засеянной такими зародышами.

Детальные вычисления показывают, что типичное время перехода из x_- в x_+ по порядку величины составляет

$$\tau_{\text{trans}} \simeq \exp(N \Delta U) \quad (4.31)$$

где $\Delta U = U(x_0) - U(x_-)$ — потенциальный барьер, который должна преодолеть флуктуация. В физико-химических системах больших размеров, в которых $N \sim 10^{23}$ искусственно поддерживается в однородном состоянии, например, за счет очень интенсивного перемешивания, значение τ_{trans} будет чрезвычайно велико, гораздо больше любого воображаемого геологического или космологического времени. Это означает, что в реальных системах переход должен начинаться как локализованное событие, охватывающее элемент малого размера. Даже при этом условии характерное время τ_{trans} будет все еще гораздо больше характерных времен релаксации, вытекающих из детерминистических динамических уравнений, если только система не слишком близка к точкам поворота μ_1 и μ_2 (рис. 4.7, а), в которых потенциальный барьер становится очень низким.

Отмеченная выше потребность в локальном описании означает, что следует обратиться к более детальной теории, основанной на фундаментальном уравнении со многими переменными, учитывающем диффузионные эффекты (см. рис. 4.3). Здесь имеются некоторые аналитические результаты. Однако численное моделирование показывает, что характерное время перехода очень интересным образом зависит от коэффициента диффузии. Если диффузия замедлена по сравнению со скоростью химической реакции, то оказывается, что переходы из x_- в x_+ , обуславливающие появление зародышей состояния x_- в состоянии x_+ , происходят с высокой скоростью. В дальнейшем эти зародыши медленно растут и за счет диффузии увлекают остальную часть системы на новую ветвь решения x_+ . С другой стороны, в случае большого коэффициента диффузии начальное состояние x_- существует довольно долго. Такая ситуация напоминает *метастабильность*, хорошо известную в теории равновесных фазовых переходов. В экспериментах с хлорид-йодидной реакцией Ру и др. показали, что аналогичное поведение характерно для системы под действием *внешних* неоднородных возмущений, обусловленных, например, не слишком тщательным перемешиванием.

На рис. 4.8 представлена кинетика среднего значения (часть б) и дисперсии (часть а) флуктуаций. После относительно длительного ожидания возникает первый зародыш (точка А на кривой рис. 4.8, б), а затем появляется второй зародыш в точке В. В точке В оба зародыша почти одновременно достигают границ системы, после чего образуются сходящиеся волновые фронты. В дальнейшем, после достижения состояния Г,

соответствующего ветви x_+ , эти фронты исчезают. Следует отметить значительное усиление флуктуаций вблизи точки перегиба кривой зависимости среднего значения от времени, о чем свидетельствует четкий максимум дисперсии.

Кроме описанных выше явлений, разнообразные типы поведения возможны также в системах, в которых стадия медленного движения сменяется динамикой на много меньших временных масштабах. Одна из возможных реализаций такого поведения связана с выбором начального состояния, которое мо-

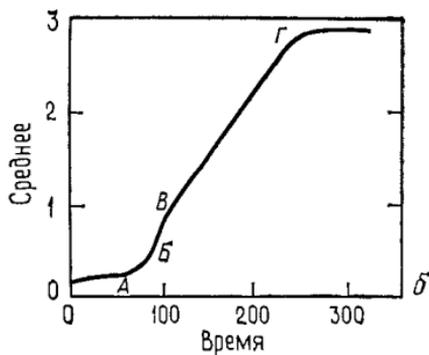
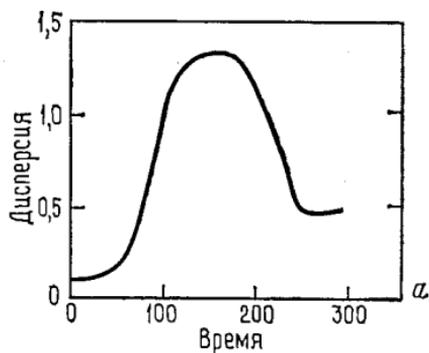


Рис. 4.8. Опосредованный зародышеобразованием переход между множественными стационарными состояниями в модельной системе. *а* — дисперсия флуктуации. На рис. *б* показана временная эволюция среднего: его резкий рост начинается по мере того, как зародыш состояния x_+ , возникающий благодаря флуктуации, начинает занимать значительную часть системы, исходно находящейся в состоянии x_- .

жет быть либо неустойчивым, либо маргинально устойчивым. Другая возможность, на которой мы здесь сосредоточимся, возникает в том случае, когда начальная фаза индукции, характеризующаяся очень малой скоростью изменения соответствующей переменной, внезапно сменяется взрывоподобным поведением, инициируемым в некоторый характерный момент времени t_e (рис. 4.9). В конечном счете в системе устанавливается состояние, которое здесь для простоты считается единственным устойчивым состоянием в диапазоне значений параметров, представляющих интерес. Типичным примером явления, в котором эти условия реализуются, служит горение. Другие примеры — автокаталитические химические реакции или явления переключения в лазерах.

Проиллюстрируем основные особенности динамики такого типа на примере адиабатического взрыва, или, иначе говоря, сгорания в замкнутом сосуде, когда высвобождающееся в химических реакциях тепло расходуется на нагревание горючей смеси. Простейший нетривиальный пример такого рода — необратимая экзотермическая реакция



в которой зависимость константы скорости $k(T)$ от температуры T определяется законом Аррениуса [см. равенство (1.5)]:

$$k(T) \simeq k_0 e^{-E_0/k_B T}$$

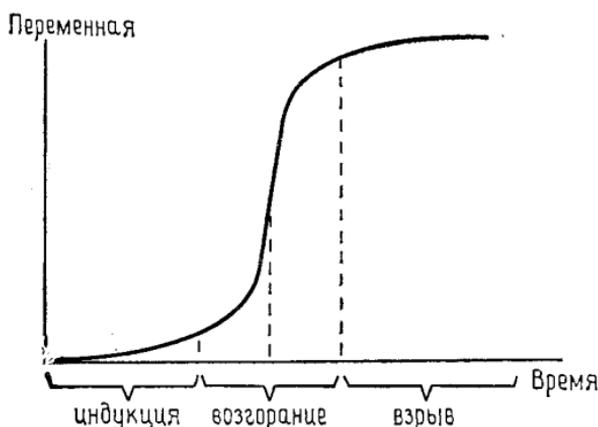


Рис. 4.9. Иллюстрация взрывоподобного поведения.

При этом уравнения баланса массы и энергии имеют вид

$$\begin{aligned} \frac{dx}{dt} &= -k(T)x \\ C_v \frac{dT}{dt} &= -r_v \frac{dx}{dt} = r_v k(T)x \end{aligned} \quad (4.33a)$$

где C_v и r_v — соответственно теплоемкость и теплота реакции при постоянном объеме. Умножая первое из этих уравнений на r_v и складывая их, легко убедиться в том, что имеет место следующий закон сохранения:

$$C_v T_0 + r_v x_0 = C_v T + r_v x \equiv C_v T_{\max} \quad (4.33б)$$

Здесь T_0 и x_0 — начальные значения T и x , а T_{\max} — конечная температура, достигаемая по завершении реакции. Благодаря этому закону сохранения уравнение (4.33a) становится замкну-

тым. Например, уравнение теплового баланса приобретает следующий простой вид:

$$\frac{dT}{dt} = (T_{\max} - T) \cdot k(T) \quad (4.33\text{в})$$

Решение уравнений (4.33) представлено на рис. 4.10. Видно, что реакция резко достигает максимальной скорости в момент времени t_e , которое мы будем называть «временем взрыва», в соответствии с общей картинкой, представленной на рис. 4.9.

Введем теперь в наше рассмотрение флуктуации. Очевидно, мы здесь имеем дело с процессом типа гибели, поскольку концентрация вещества X может лишь убывать. Поэтому основное

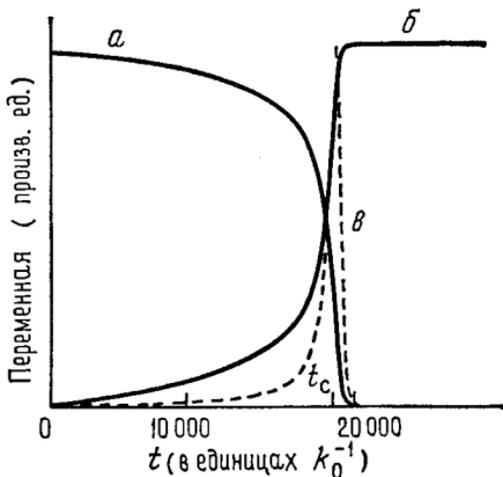


Рис. 4.10. Концентрация x (кривая a), температура T (кривая $б$) и скорость реакции (кривая $в$), полученные путем решения уравнений (4.33). Значения параметров: $E_0/R = 10\,000$; $T_{\max} = 2000^\circ$; $r_v/C_v = 1200$.

уравнение можно записать в виде (4.14), где константа скорости рождения λ принята равной нулю. Обозначая посредством X число частиц реагента, имеем

$$\frac{dP(X, t)}{dt} = \mu_{X+1} P(X+1, t) - \mu_X P(X, t) \quad (4.34a)$$

Скорость гибели легко найти, воспользовавшись законом сохранения (4.33б):

$$\mu(X) = k(T) X = k_0 X \exp \left[- \frac{E_0}{k_B \left(T_{\max} - \frac{r_v X}{C_v N} \right)} \right] \quad (4.34б)$$

Здесь N — параметр, пропорциональный размеру системы.

На рис. 4.11 показаны основные стадии эволюции распределения вероятностей, предсказываемого уравнениями (4.34).

В качестве начального состояния выбрано распределение в виде острого пика, находящегося на пологой (соответствующей фазе индукции) части кривой рис. 4.9, где скорость процесса должна быть очень малой. Тогда можно наблюдать явление *переходной бимодальности*: в то время как один из максимумов распределения вероятности остается вблизи начального состояния, вблизи состояния, соответствующего полному сгоранию, образуется второй максимум. В конце концов первый пик исчезает и система переходит в свое единственное устойчивое состояние. Поэтому переходный процесс в данной системе по своей сути представляет развитие бифуркации во времени.

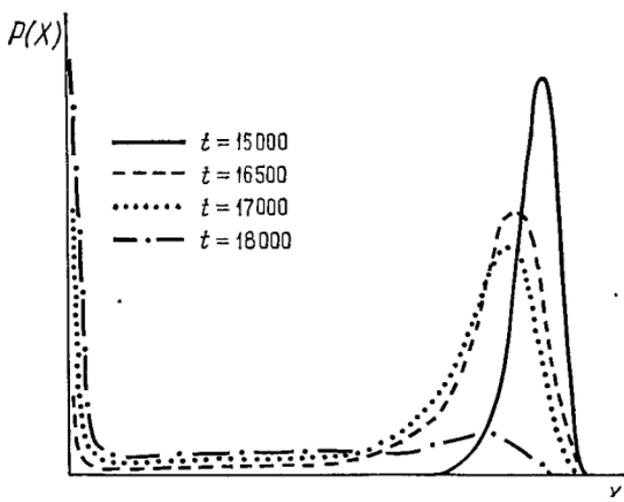


Рис. 4.11. Переходная бимодальность, возникающая при адиабатическом взрыве. При $t = 15000$ (время измеряется в специальных единицах) вероятность носит одномодальный характер. При $t > 16500$ на ней развивается все более выраженный второй пик в области низких значений X . При этом амплитуда исходного пика, динамика которого развивается медленно (индукция), непрерывно уменьшается. Значения параметров те же, что и на рис. 4.10.

Приведем качественное объяснение этого неожиданного явления. Напомним, что мы имеем дело с процессом, характеризуемым двумя значительно различающимися временными масштабами, причем в начальном состоянии скорость детерминистического процесса весьма мала. Поэтому максимум соответствующего распределения вероятностей, движение которого примерно следует движению детерминистического максимума, будет медленно смещаться в область более низких значений X . При этом вследствие флуктуаций ширина распределения будет пропорциональна длительности индукционного периода и обратно пропорциональна квадратному корню из N (в соответствии с центральной предельной теоремой, см. разд. 4.1). Достаточно

большие длительности индукционного периода могут преодолеть влияние размера системы, и ширина распределения может стать значительной. В результате точка возгорания будет достигнута с очень большой вероятностью еще до того, как ее достигнет максимум распределения. В этот момент система быстро вовлекается в движение в сторону низких значений X . Такая утечка вероятности будет происходить непрерывно, однако, поскольку отрицательные значения X не достигаются, возникает своего рода затор, в результате которого в области малых X образуется новый пик вероятности. В конце концов первичный пик, к этому моменту сильно «подсевший», достигает точки возгорания, и на этом переходная бимодальность закончится.

Очевидно, в состоянии переходной бимодальности поведение системы будет в значительной мере случайным — расхождение между средним и наиболее вероятным значением станет значительным, а дисперсия достигнет макроскопического уровня. Это проявится в виде флуктуаций самого времени возгорания, которое тем самым становится случайным событием.

Помимо специфического применения в проблемах, связанных с горением, развитие бифуркации во времени, основываясь исключительно на собственных, внутренних причинах, представляет собой новую модель дифференциации и эволюции в неравновесных системах. В самом деле, такая система не требует для развития бифуркации постоянных возмущений извне, как в дарвиновской картине эволюции через мутации и естественный отбор. Возникающих на какое-то время из равновесия динамических процессов оказывается достаточно, чтобы индуцировать временную дифференциацию — в ходе медленной стадии эволюции флуктуации привносят в систему множество состояний, посещаемых с некоторой вероятностью, хотя начальное состояние по-прежнему доминирует. С появлением быстрой стадии эти состояния захватываются с такой высокой скоростью, что они скоро забывают дорогу домой, к своим общим «предкам», и образуют новый «клон». Аналогичные представления могли бы быть полезны в различных задачах, связанных с динамикой популяций и эпидемиологией. Возвращаясь к физико-химическим системам, можно также отметить проблемы, связанные с кластеризацией, коллапсированием и пролиферацией дефектов в процессе дезорганизации или ломки различных материалов.

4.6. Чувствительность и отбор

В предыдущем разделе мы получили некоторые средства решения проблемы выбора между различными состояниями, неоднократно поднимавшейся в гл. 3. В самом деле, допустим, что в некоторой системе имеется несколько одновременно устой-

чивых состояний. Учитывая с помощью основного уравнения влияние флуктуаций, можно найти высоты пиков распределения вероятности, а также время, необходимое в среднем для того, чтобы покинуть область притяжения каждого из этих устойчивых состояний [см. уравнение (4.31)]. Тогда доминирующим (т. е. тем, которое окажется избранным) будет то состояние, в котором время жизни наибольшее и пик распределения вероятности самый высокий. Такой подход в целом удовлетворителен, однако он не дает ответа на следующий важный вопрос: что будет, если вероятностные пики и переходные времена совпадают хотя бы для двух состояний?

Не следует считать, что такая возможность является искусственно рожденной. Некоторые из наиболее удивительных переходных явлений, рассматривавшихся до сих пор в этой книге, соответствуют разнообразным нарушениям симметрии и с неизбежностью приводят к множественным состояниям, неразличимым между собой с точки зрения флуктуаций или устойчивости. Один из примеров — нарушение пространственной симметрии в макроскопическом переходе, обсуждавшемся в разд. 3.11. Другой пример — нарушение хиральной симметрии на молекулярном уровне, различные аспекты которого были затронуты в разд. 3.12. Очевидно, в каждом случае требуется новый механизм отбора. В этом разделе будет рассмотрена типичная возможность отбора за счет взаимодействий с внешними полями, осуществляющихся с нарушением симметрии.

Основа этого механизма уже была заложена в разд. 3.4, посвященном параметрической зависимости бифуркации и соответствующему представлению о структурной устойчивости. Как отмечалось в связи с уравнением (3.14), приводящая к двум неразличимым состояниям [рис. 3.9, *a* и уравнение (3.14) при $\mu=0$] бифуркация претерпевает качественное изменение, если должным образом изменить некоторый дополнительный параметр [рис. 3.10, *b* и уравнение (3.14) при $\mu \neq 0$]. Такая ситуация еще раз показана на рис. 4.12. Соответствующее уравнение для переменной x имеет вид

$$\frac{dx}{dt} = -x^3 + \lambda x + \mu \quad (4.35a)$$

Его стационарное решение определяется уравнением

$$-x_s^3 + \lambda x_s + \mu = 0 \quad (4.35b)$$

Отсюда видно, что параметр μ играет роль некоторой внешней асимметрии. Даже если этот параметр мал, его влияние может значительно усилиться за счет нелинейности системы. Устойчивая ветвь x_- , которая при $\mu=0$ была неотличима от x_+ в смысле устойчивости, существует по-прежнему, но отделена от ветви x_+

при $\mu \neq 0$. Следовательно, если параметр λ возрастает от значений, лежащих левее предельной точки, выше которой ветвь x_- существует, система будет стремиться оставаться на верхней ветви (рис. 4.12, б), если только не появится достаточно сильная флуктуация, которая сбросит ее на нижнюю ветвь. Поскольку в каждой системе имеются тепловые флуктуации, такой пере-склок возможен при малых расстояниях Δ , показанных на рис. 4.12, б. Напротив, если «внешняя асимметрия» μ должна служить «хорошим селектором», то зазор Δ должен быть доста-точно большим, чтобы можно было противостоять флуктуа-циям. Прежде чем продолжить эти рассуждения, остановимся вкратце на возможной природе «нарушителя симметрии» — па-раметра μ .

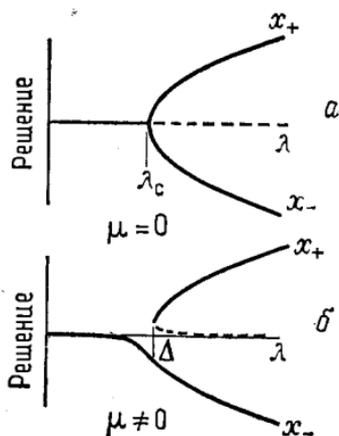


Рис. 4.12. Отбор, реализуемый возмущением бифуркации внешним полем μ , нарушающим симметрию. Минимальное расстояние между возмущенными ветвями Δ должно превышать уровень шума, обусловленный флуктуациями.

Как подчеркивалось в разд. 3.6, уравнения типа (4.35а) не просто относятся к некой конкретной модели — скорее они представляют собой типичную, *нормальную форму*, описывающую целые классы динамических систем, функционирующих вблизи точки бифуркации. Поэтому в связи с параметром μ следовало бы задаться вопросом идентификации некоторых важных, нарушающих симметрию полей, которые в сочетании с химическими, гидродинамическими и другими системами приводят к нормальной форме, соответствующей (4.35а). Здесь мы отметим лишь два универсальных случая — поле тяготения, которое действует на массивный предмет всегда в одном и том же направлении, и взаимодействия, обусловленные слабыми нейтральными токами. Последние, несмотря на свою субъядерную природу, слегка изменяют константы скоростей реакций стереоизомеров, причем также всегда в одном и том же направлении. В обоих этих случаях можно показать, что динамика сводится к уравнениям, сходным с (4.35а), при условии, что в отсутствие внешнего воздействия система функционирует в окрестности

точки бифуркации, подобной той, что показана на рис. 4.12, а.

Перейдем теперь к количественной оценке чувствительности к малому параметру μ , ответственному за нарушение симметрии. Из (4.35б) легко видеть, что при $\mu \neq 0$ невозмущенная точка бифуркации $\lambda_c = 0$ принадлежит области, где решение единственно, $x_s = \mu^{1/3}$. С другой стороны, величина минимального зазора Δ может быть найдена из условия существования трех действительных корней (4.35б) в этой точке, два из которых совпадают. Тогда легко получить [см. уравнение (3.15)]

$$\begin{aligned}\lambda_\Delta &= 3(\mu^2/4)^{1/3} \\ \Delta &= \frac{3}{2}(4\mu)^{1/3}\end{aligned}\quad (4.36)$$

Наличие множителя $\mu^{1/3}$ в этом результате имеет решающее значение. Если система функционирует вблизи равновесия (λ заметно ниже нуля), то эффект внешнего поля будет линейным, поскольку кубическими членами в (4.35а) можно пренебречь. Поэтому поле создает чрезвычайно слабое смещение порядка $x_s \sim \mu$. Однако, когда система находится вдали от равновесия и ее нелинейность способна проявиться полностью, тогда то же самое взаимодействие может создать гораздо более значительное смещение $x_s \sim \mu^{1/3}$, что приводит к существенному усилению сигнала. Например, в случае нарушения хиральной симметрии, когда x_s соответствует разности концентраций правых и левых изомеров, параметр μ представляет собой разность энергий активации (измеряемых в характерных тепловых единицах), обусловленных слабыми взаимодействиями. Согласно надежным оценкам, $\mu \sim 10^{-15}$. Поэтому в равновесии отклонение ($x_s \sim 10^{-15}$) макроскопически неощутимо. Однако вдали от равновесия ($x_s \sim 10^{-5}$) оно вполне способно конкурировать с макроскопическими эффектами.

Как отмечалось выше, жесткость отбора оценить невозможно до тех пор, пока Δ не сравнится с внутренними флуктуациями x в окрестности точки бифуркации. Свойства этих флуктуаций обсуждались в разд. 4.1 и 4.4. Для наших целей достаточно лишь нескольких общих особенностей этих флуктуаций.

Прежде всего необходимо найти, насколько значение λ_Δ , соответствующее точке минимального расстояния между решениями, отличается от критического значения $\lambda = 0$, соответствующего бифуркации. Как видно из (4.28), если λ_Δ велико по сравнению с N^{-1} , где N — безразмерная величина, характеризующая размеры системы, то флуктуации около ветви x_+ будут существенно гауссовыми, и их дисперсия будет равна

$$\sigma_{\text{Gauss}}^2 \simeq N^{-1} \quad (4.37а)$$

Однако если λ_Δ очень мало, то дисперсию следует вычислять, пользуясь полиномом четвертой степени, фигурирующим в (4.28). Тогда дисперсия по порядку величины должна быть

$$\sigma^2 \approx N^{-1/2} \quad (4.376)$$

Поскольку из жесткости отбора следует $\Delta \gg \sigma$, из уравнений (4.36) и (4.37) мы приходим к условию

$$\frac{3}{2} (4\mu)^{1/3} \gg N^{-1/2} \quad (4.38)$$

Из этого неравенства автоматически следует $\lambda_\Delta > N^{-1}$. В задаче о нарушении хиральной симметрии левая часть (4.38) составляет по порядку величины 10^{-5} . Чтобы оценить правую часть, примем, что размер системы близок к радиусу корреляции (4.25), на котором система остается пространственно однородной, и рассмотрим некоторый разбавленный раствор, содержащий активные стереоизомеры в миллимолярных концентрациях. Если можно достичь радиуса корреляции $l_{\text{корр}} \sim 1$ см вблизи бифуркации, то $N = 10^{18}$ или $N^{-1/2} = 10^{-9}$ и неравенство (4.38) выполняется. Заметим, что вблизи равновесия это неравенство стало бы противоположным, поскольку левую часть пришлось бы заменить членом порядка $\mu \sim 10^{-15}$.

Установив, что вдали от равновесия отбор принципиально возможен, отметим следующие предписания, определяющие хорошую стратегию отбора:

1. В начальный момент $t=0$ система находится вблизи $\lambda = \lambda_c = 0$ — точки бифуркации в отсутствие поля ($\mu=0$). Распределение считается симметричным и центрированным при $x=0$. Как отмечалось ранее, в этой области имеется лишь одно устойчивое состояние, $x_s = \mu^{1/3}$. Поэтому вероятностный пик будет постепенно дрейфовать к этому состоянию. Характерное время этого процесса можно оценить, решая зависящее от времени детерминистическое уравнение (4.35а) при $\lambda = \lambda_c = 0$. В результате прямого интегрирования получается выражение

$$\Delta t_{0 \rightarrow x} \approx \frac{1}{3\mu^{2/3}} \left[\ln \frac{\mu^{1/3}}{|x - x_s|} + \text{члены порядка единицы} \right] \quad (4.39)$$

При x , не слишком близких к x_s , порядок величины Δt существенно определяется множителем $\mu^{-2/3}$. Приняв для примера $\mu \sim 10^{-15}$, получаем $\Delta t \sim 10^{10}$ с $\sim 10^3$ лет! Отсюда следует, что характерный детерминистический масштаб времени для отбора чрезвычайно велик.

2. Во избежание трудности, сопряженной с таким очень большим временем отбора, допустим, что система эволюционирует лишь до тех пор, пока пик вероятности не сдвинется к x_{max} ,

вблизи которого уровень шума (измеряемый в единицах дисперсии σ) меньше, чем само x_{\max} . С учетом предыдущей оценки при $\mu \sim 10^{-15}$ достаточно приблизиться к x_{\max} на расстояние, слегка превышающее 10^{-9} . В этой точке в системе скачком устанавливается новое значение параметра λ , гораздо большее, чем λ_{Δ} . Поскольку теперь система далека от критической точки, она будет экспоненциально приближаться к ветви x_+ , и это произойдет гораздо раньше заметного заселения состояния x_- за счет флуктуаций. Тем самым можно ожидать, что селекция будет почти идеальной. Как показывают численные методы, такая процедура уменьшает временной масштаб до нескольких дней вместо 10^3 лет, предсказываемых детерминистическим уравнением.

Столь удивительные возможности материи при наличии как нелинейной динамики, так и неравновесных условий снова вызывают в нас чувство глубокого изумления.

4.7. Символическая динамика и информация

До сих пор мы занимались описанием стохастической динамики на языке распределений вероятностей и их различных моментов. Альтернативный подход, для многих целей оказывающийся особенно наглядным, состоит в изучении отдельных реализаций рассматриваемого стохастического процесса. В этом разделе будут суммированы некоторые основные элементы такого подхода; в дальнейшем он будет использоваться в разд. 4.8 и гл. 5.

Обозначим набор допустимых состояний системы* символом $\{Q_i\}$ ($i=1, 2, \dots$). В соответствии с замечаниями, сделанными в разд. 4.2, предполагается, что набор $\{Q_i\}$ выбран таким образом, что временная эволюция системы определяет некоторый марковский процесс. Кроме того, отказавшись от непрерывного времени, мы будем считать, что переходы между различными состояниями совершаются в моменты времени $t_m = m\Delta t$, где m — целое число и Δt — некоторая фиксированная величина, значение которой зависит от того, какой конкретно процесс рассматривается. Это предположение не является слишком жестким, поскольку последовательные переходы в природе должны разделяться некоторым временным интервалом. Естественно, Δt должно быть гораздо меньше характерного времени эволюции системы, описываемой феноменологическими уравнениями баланса.

* Вначале мы рассмотрим системы с дискретным пространством состояний. Случай непрерывных переменных будет вкратце рассмотрен в конце настоящего раздела.

При этих условиях результат p последовательных шагов процесса можно записать в виде последовательности

$$Q_{i_1} Q_{i_2} \dots Q_{i_p} \quad (4.40)$$

Набор всех последовательностей вида (4.40) определяет *марковскую цепь*. Следовательно, представление о системе как о марковской цепи равносильно описанию ее эволюции на языке «букв» $\{i_p\}$ соответствующего «алфавита». На первый взгляд может показаться, что для большинства рассматриваемых систем количество букв в алфавите должно быть бесконечным. Однако это не обязательно так. В качестве примера рассмотрим систему с двумя устойчивыми стационарными состояниями. Вследствие индуцированных флуктуациями переходов между этими состояниями (см. рис. 4.7) система будет проводить значительное время в каждом из них, причем соответствующее время жизни определяется равенством (4.31), после чего осуществляется быстрый переход в новое состояние. Можно ввести более абстрактное описание, обозначая каждое мгновенное состояние «0» или «1» в зависимости от того, к какому из двух состояний система ближе — к первому или второму. Если наблюдатель будет следить за системой, отмечая через правильные промежутки времени, в какой области притяжения оказывается система, то составленная им последовательность может иметь следующий вид:

$$111001100010110\dots \quad (4.41)$$

Такая запись соответствует алфавиту, содержащему только два символа: 0 и 1. Вообще говоря, последовательности типа (4.41) не определяют марковскую цепь. Однако при конкретных условиях можно показать, что разбиение исходного пространства состояний, в типичном случае содержащего несчетное множество или даже континуум состояний, на два класса, соответствующие указанным выше состояниям, сохраняет марковский характер процесса. В таком случае говорят о *марковском разбиении*. Описание, основанное на последовательностях (4.41), мы будем называть *символической динамикой* — теперь уже приходится иметь дело с последовательностью не состояний системы в обычном смысле слова, а с последовательностями символов — букв алфавита конечной длины.

Попытаемся охарактеризовать множество различных последовательностей типа (4.40) или (4.41). Из определения марковского процесса вытекает, что вероятность события C , описываемого последовательностью (4.40), равна

$$P(C) = P_{i_1} p_{i_1 i_2} \dots p_{i_{p-1} i_p} \quad (4.42)$$

где P_{i_1} представляет исходное распределение вероятности состояния Q_{i_1} , а $p_{i_k i_{k+1}}$ — вероятность перехода из состояния Q_{i_k} в состояние $Q_{i_{k+1}}$ за время Δt . Заметим, что в типичном случае некоторое выделенное состояние в течение этого процесса посещается системой многократно.

Теперь можно доказать следующую теорему: пусть описываемый рассматриваемыми последовательностями процесс является марковским и удовлетворяет закону больших чисел (см. разд. 4.1). Очевидно, число всех последовательностей длины ρ , содержащих N состояний, равно $N_\rho = N^\rho$. Расположим эти последовательности в порядке убывания их вероятностей $P(C)$ и в этом порядке выберем первые n_ρ последовательностей, сумма вероятностей которых как раз превосходит некоторое заранее заданное число между 0 и 1. Тогда, в пределе больших ρ , число n_ρ выбранных таким образом последовательностей слабо зависит от упомянутого заранее заданного числа (его, в частности, можно выбрать близким к единице). Для этого числа последовательностей имеется следующая оценка:

$$n_\rho \sim e^{\rho I} \quad (4.43a)$$

где

$$I = - \sum_{ij} P_i p_{ij} \ln p_{ij} \quad (4.43b)$$

Величина I обычно рассматривается как мера информации, получаемой при продвижении вдоль марковской цепи на один шаг от начального состояния Q_i . Связь между этой величиной и энтропией, использованной в (4.16), самая непосредственная. Чтобы увидеть ее, достаточно рассмотреть «детерминистические» начальные вероятности ($P_i = 1$ при $i = i_1$ и $P_i = 0$ в остальных случаях), и тогда $P_{i_1 j} = P_j$ представляет собой вероятность нахождения в состоянии Q_j .

В огромной ценности сформулированной выше теоремы можно убедиться с помощью следующих соображений. В то время как число всех последовательностей типа C есть $N_\rho = e^{\rho \ln N}$, число n_ρ избранных последовательностей составляет примерно $e^{\rho I}$. Напомним (разд. 4.3), что энтропия марковской цепи достигает максимального значения, когда все состояния равновероятны, $p_i = 1/N$. Это дает $I_{\max} = \ln N$. Таким образом, всегда справедливо неравенство $I < \ln N$ (за исключением тривиального случая), в результате чего при больших ρ имеем

$$e^{\rho I} \ll e^{\rho \ln N} \quad (\rho \rightarrow \infty)$$

Иными словами, в число тех последовательностей типа C , сумма вероятностей которых дает близкую к единице величину,

входит лишь малая доля от всех возможных последовательностей. Насколько эта доля мала, зависит от энтропии марковской цепи. Таким образом, мы приходим к важному заключению о том, что отклонение от равновероятного распределения выглядит как весьма эффективное средство отбора выделенных последовательностей из огромного количества всевозможных случайных последовательностей. Такое отклонение можно реализовать лишь в том случае, если рассматриваемая система взаимодействует с внешней средой, т. е. не является изолированной. Как неоднократно отмечалось, такие условия могут привести к *неравновесным состояниям*, которые за счет обмена веществом и энергией с окружающей средой могут существовать неопределенно длительное время. По этой причине неравновесность представляется естественным условием осуществления отбора.

В качестве примера рассмотрим процесс образования некоторого биополимера, скажем белка длиной в 100 аминокислот ($\rho=100$). Хорошо известно, что в природе существует 20 различных аминокислот ($N=20$). Если бы биополимеризация происходила в изолированной системе, в которой все последовательности априори равновероятны, мы имели бы $N_\rho = e^{100 \ln 20} \sim \sim e^{300}$ равновероятных последовательностей. Следовательно, любая конкретная последовательность, например такая, которая необходима для выполнения желаемой биологической функции, встречалась бы с чрезвычайно низкой вероятностью около e^{-300} ! С другой стороны, если биополимеризация происходит в неравновесных условиях, соответствующих энтропии, например $I=0,1I_{\max}$, то лишь $n_\rho \sim e^{30}$ последовательностей длиной в 100 аминокислот будут реализоваться с заметной вероятностью. Это число уже гораздо ближе к нашим масштабам. Таким образом, становится понятно, что на таком ограниченном поле возможностей в неравновесных условиях эволюция способна при наличии достаточного времени создать образования, наделенные такими специальными свойствами, как саморепликация, эффективное преобразование энергии и т. д.

До сих пор в этом разделе предполагалось, что динамические системы, приводящие к рассматриваемым стохастическим процессам, определены в дискретном пространстве состояний, в котором эволюция соответствует некоторому марковскому процессу. Как объяснялось в первых разделах настоящей главы, этим условиям должны удовлетворять внутренние флуктуации, спонтанно возникающие за счет молекулярной динамики в любой макроскопической физико-химической системе. А как же насчет хаотической динамики — второго универсального источника стохастичности в природе? Начнем с того, что пространство состояний системы, приводящей к такой динамике, обычно

непрерывно. Более того, априори нет никаких причин к тому, чтобы процесс был марковским, поскольку в отличие от флуктуаций отклонения от среднего движения здесь не являются ни локализованными, ни маломасштабными событиями.

Примечательно то, что, несмотря на эти обстоятельства, хаотической динамике все-таки можно придать вид, очень сходный с марковской цепью. Причина этого состоит в том, что часто удается разбить пространство состояний на конечное число ячеек, таких, что с течением времени за счет динамики между ними индуцируются переходы, и наоборот: любая заданная последовательность ячеек соответствует одному и только одному варианту истории соответствующей непрерывной системы. Разбиения, обладающие этим замечательным свойством, мы будем называть *производящими*. Благодаря таким разбиениям можно достичь нового уровня абстрагирования, когда естественным способом описания снова становится *символическая динамика*, основанная на работе с последовательностями букв алфавита, соответствующих различным ячейкам. В следующем разделе мы обсудим конкретный пример такого символического описания в контексте проблемы отбора асимметричных структур.

4.8. Генерация асимметричных, информационно-насыщенных структур

В предыдущем разделе было показано, что количество последовательностей, реализующихся с заметной вероятностью, определяется энтропией марковской цепи. Кроме того, символическая динамика позволила рассматривать эти избранные последовательности как набор «сообщений» или «текстов», содержащих определенную информацию. Здесь нам хотелось бы задаться следующим вопросом: каким образом динамическая система может самопроизвольно генерировать такие информационно-насыщенные последовательности?

Когда перед нами лист бумаги с английским текстом, наше восприятие содержащегося в нем сообщения связано с тем, что текст читается слева направо, в нем содержатся точки и запятые и вообще соблюдаются определенные синтаксические правила. Текст на арабском языке будет осмысленным лишь в том случае, если его читать справа налево, причем в тексте также соблюдаются определенные синтаксические правила, регулирующие чтение (в общем случае эти правила отличаются от английских). Аналогично этому, как отмечалось в разд. 3.13, генетический код характеризуется определенным стартовым сигналом и инструкциями, благодаря которым текст может быть прочтен в определенном направлении от начальной точки.

С восприятием сообщения связан и другой, столь же важный аспект — наличие определенного набора символов, развертывающегося во времени, или, в более общем случае, вдоль направления чтения. Однако даже тогда, когда такая последовательность символов идеально определена и легко воспроизводится, по существу она непредсказуема в том смысле, что во всех нетривиальных случаях ее глобальная структура не может быть установлена по структуре части этой последовательности, сколь бы большой эта часть ни была. В этом плане ее можно рассматривать как стохастический процесс, и именно это обстоятельство позволяет нам говорить об информации. Например, читая «Элементы» Евклида или «Принципы» Ньютона, мы воспринимаем определенную информацию. Однако что бы это была за информация, если бы существовал простой алгоритм, позволяющий восстановить вторые части этих произведений по первым!

Короче говоря, информация сопряжена с двумя фундаментальными условиями.

1. Резкое нарушение пространственной симметрии, непрерывно отсеивающее другие возможные результаты чтения (здесь мы опять приходим к вопросам, поднятым в разд. 3.13 и 4.6, т. е. к вопросам о том, как такие асимметричные формы могут возникать и доминировать над своими зеркальными изображениями).

2. Элемент непредсказуемости, связанный с раскрытием содержания данного объекта или с сообщением, заранее неизвестным читателю.

Обрисует теперь вкратце механизм, позволяющий автоматически удовлетворить обоим этим условиям.

Рассмотрим некоторую диссипативную систему типа совокупности химических реакций, в которых синтезируются компоненты с концентрациями X_i . Допустим, что мы регистрируем набор значений некоторых представительных переменных X_i , получаемых при достаточно больших временах. Эту последовательность можно рассматривать как одномерную цепочку цифр. Вследствие необратимости, присущей эволюции любой диссипативной системы, полученная таким образом «прямая» последовательность будет асимметричной в том смысле, что она будет отличаться от последовательности, получаемой при ее чтении в обратном направлении, поскольку такая обращенная последовательность могла бы генерироваться лишь некоторой физически неприемлемой динамической системой, получаемой путем обращения времени. Поэтому ясно, что необратимость способна индуцировать некоторую абсолютную асимметрию в абстрактном пространстве одномерных последовательностей символов. Теперь мы зададимся следующим вопросом: как перенести эту

асимметрию символической динамики на некоторую одномерную систему, находящуюся в *физическом пространстве*?

Предложим некоторый частный алгоритм достижения этой цели. Допустим, что рассматриваемая химическая система способна генерировать незатухающие колебания, хотя и не обязательно периодические. Далее допустим, что, когда некоторая переменная X_i пересекает определенный уровень L_i с положительным наклоном (в точках, выделенных на рис. 4.13), включается новый процесс, в результате которого вещество i быстро осаждается или диффундирует за пределы реакционного пространства и затем собирается на некоторой «ленте». Получаемая таким образом картина будет носить отпечаток различных порогов, пересекаемых различными концентрациями. Как уже от-

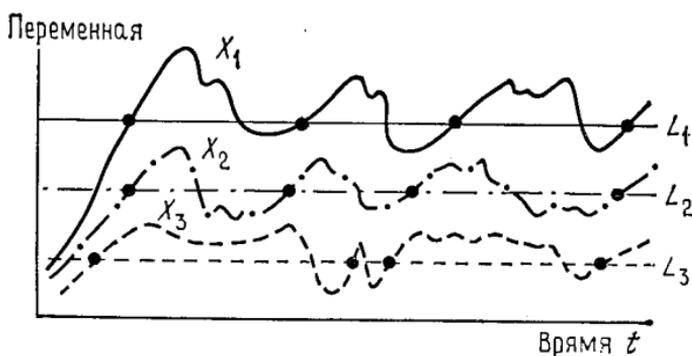


Рис. 4.13. Генерация последовательности символов динамической системой. Когда возрастающие переменные X_1, X_2, \dots пересекают уровни L_1, L_2, \dots , соответствующие соединения начинают выпадать в осадок, что позволяет перенести данную временную последовательность на некоторую пространственную координату.

мечалось, в типичном случае такая последовательность будет асимметричной в силу нарушения временной симметрии, обусловленной необратимостью. Таким образом можно произвести отображение некоторого необратимого процесса в одномерную пространственную картину. В более общем случае можно показать, что получение пространственных картин, содержащих информацию о последовательности значений, необратимо достигаемых концентрационными переменными, возможно и в случае реакционно-диффузионных систем, сопряженных с некоторым *полярным* вектором (типа напряженности гравитационного или электростатического поля). Такое пространственное отображение необратимости становится невозможным при сопряжении с аксиальным вектором типа магнитного поля, поскольку сопряжение этого типа невозможно в уравнении баланса для скалярной переменной типа концентрации. По этой причине обсуждаемый здесь механизм не может дать информацию о механизме

селекции хиральности — второй главной формы асимметрии, наблюдаемой в природе.

Перейдем теперь к классификации картин, которые могут генерироваться различными химическими процессами. Допустим сначала, что переменные X_i асимптотически стремятся к некоторому предельному циклу. Из гл. 3 мы знаем, что для такого типа поведения нужны по меньшей мере две сопряженные переменные и что разность фаз для любой пары таких переменных будет постоянной во времени. Отсюда следует, что для двух переменных пересечение уровней может в лучшем случае дать следующую картину символической динамики:

$$\dots X Y X Y X Y \dots \quad (a)$$

или, если заданные уровни пересекаются тремя переменными,

$$\dots X Y Z X Y Z X Y Z \dots \quad (b)$$

Последовательность (a) обратима, т. е. она читается одинаково в обоих направлениях, последовательность же (b) описывает некоторый абсолютно асимметричный объект. Однако, вводя символическое сокращение (аналог кодона)

$$\alpha = \{X Y Z\}$$

последовательность (b) можно записать в виде

$$\dots \alpha \alpha \alpha \dots \quad (в)$$

и в таком виде она уже теряет пространственную асимметрию.

Если рассматриваемая структура соответствует последовательностям (b) или (в), то вряд ли можно говорить об информации, поскольку такая система полностью предсказуема. Например, пользуясь данным Клодом Шэнноном определением информации (4.436), $I = -\sum_i P_i \ln P_i$, мы получаем, что I равно нулю.

Допустим теперь, что рассматриваемая система находится в режиме хаотического аттрактора. Как подчеркивалось в предыдущем разделе и как будет видно из гл. 5, при выполнении соответствующих условий и использовании подходящей символической динамики временную эволюцию неустойчивой динамической системы можно отобразить на некоторый стохастический процесс. В этом смысле запись хаотической динамики в соответствии с представленным в настоящем разделе механизмом потенциально может привести к структурам, которые 1) асимметричны в силу необратимости; 2) информационно насыщены, поскольку теоретико-информационная энтропия стохастического процесса в общем случае отлична от нуля.

триплета YXY исключает возможность того, что данная последовательность определяет некоторый простой марковский процесс, поскольку для такого процесса

$$P(YXY) = P(Y) P(X|Y) P(Y|X)$$

и все вероятности, входящие в правую часть, отличны от нуля. Поэтому следует поискать марковский процесс более высокого порядка, для которого память о некотором событии длится на протяжении нескольких шагов. Как выясняется, хорошее согласие с численно найденными частотами септулетов символов достигается в предположении марковского процесса пятого порядка*

$$\begin{aligned} P(ABCDEFG) &= P(ABCDEF) P(G|BCDEF) = \\ &= P(ABCDE) P(F|ABCDE) P(G|BCDEF) \end{aligned}$$

где $P(G|BCDEF)$ представляет собой численно найденную условную вероятность встречи символа G при условии наличия квинтуплета $BCDEF$.

При ближайшем рассмотрении статистических характеристик таких последовательностей выясняются некоторые удивительные особенности. Например, из всех возможных 3^7 последовательностей семисимвольной длины, которые можно построить из X , Y и Z , в динамике реализуется лишь 21. Более того, примерно для половины из них условная вероятность некоторого символа при условии, что заданы пять предыдущих символов, оказывается равной единице. Следовательно, все выглядит так, как если бы в систему были встроены «грамматические правила», автоматически выполняемые в результате динамики.

Описанную процедуру можно повторить при другом значении параметра $a=0,40$ модели (3.30). Результаты получаются качественно аналогичными, включая возможность выражения последовательности на языке гиперсимволов. При этом разве что меняются численные значения вероятностей различных мультиплетов. Подобным образом основные результаты не изменяются при малых изменениях пороговых значений. Однако большие изменения пороговых значений приводят к значительным последствиям. Например, используя в качестве порогов для величин X , Y и Z их средние во времени значения, мы находим последовательности, которые невозможно полностью выразить на языке гиперсимволов. Более того, адекватное описание дан-

* То, что цепь должна быть по крайней мере четвертого порядка, ясно уже из существования самого длинного гиперсимвола β , а также из того факта, что символу Z всегда предшествует символ X .

ных не удастся получить даже с привлечением модели марковской цепи вплоть до шестого порядка.

Количество информации, содержащейся в последовательности символов X , Y и Z , можно вычислить с помощью расширенного шэнноновского определения (4.436), в котором состояние i теперь представлено квинтуплетом символов. При этом было установлено, что количество информации конечно, хотя оно и значительно ниже максимального значения, соответствующего равновероятному распределению. Путем генерации таких последовательностей, как естественный результат собственной неравновесной динамики, причем последовательностей, представляющих собой одну из весьма многочисленных возможностей, система тем самым становится генератором информации. Заметим, однако, что каждое индивидуальное сообщение остается непредсказуемым, о чем свидетельствует стохастический характер данной последовательности. С другой стороны, степень непредсказуемости ослабляется существованием корреляций, простирающихся за пределы пяти ближайших соседей в цепи. Заимствуя термин из теории сообщений, можно сказать, что это позволяет *сжать* информацию, содержащуюся в последовательности X , Y , Z , воспроизводя ее более компактно с помощью гиперсимволов α , β , γ . Таким образом, наша модель позволяет установить один из наиболее существенных аспектов информации, отмечавшихся в начале настоящего раздела.

Сумируя сказанное, отметим следующее: с самого начала перед нами стояла та трудность, что само существование информации автоматически порождает два взаимно противоречащих фактора. С одной стороны, информация должна быть связана со своего рода случайностью — отсюда и многочисленные попытки в литературе связать ее с «незнанием» или с «неполным описанием» системы. А с другой стороны, астрономически большое число случайных последовательностей длины $\rho \gg 1$ делает чрезвычайно маловероятным выбор на априорной основе конкретного класса последовательностей, который мог бы играть главную роль в наблюдаемом явлении.

Мы предложили механизм, позволяющий достичь поставленной цели при следующих ограничениях: информационно осмысленные структуры могут генерироваться соответствующим нелинейным, необратимым во времени механизмом, действующим в удаленной от равновесия хаотической области. Таким образом, случайность и асимметрия — две предпосылки информации — с самого начала оказываются включенными в создаваемую структуру. Кроме того, поскольку такие структуры являются результатом действия некоего механизма (который в конечном счете должен работать в присутствии молекул, наделенных соответствующими каталитическими свойствами), упоминавшаяся

трудность, связанная с чудовищной «термодинамической маловероятностью» случайных последовательностей, оказывается преодоленной автоматически.

Один из наиболее существенных аспектов концепции, обсуждавшейся в настоящем и предыдущем разделах, состоит в появлении *нового уровня описания*, обусловленного соответствующей динамикой. Подобно тому как состояние неравновесной системы оказалось удобно описывать на языке корреляций между макроскопически разделенными элементами, а явление бифуркации описывать с помощью параметра порядка вместо исходно имевшихся переменных состояния, так и теперь выясняется, что для определенного класса стохастических динамических систем естественно ввести еще более высокий уровень абстракции и говорить о символах и информации. Именно эту общую возможность, более чем любую другую особенность из всех, упоминавшихся и анализировавшихся на протяжении четырех глав, следует рассматривать как неотъемлемый признак сложного, своего рода «отпечатки пальцев».

4.9. Снова алгоритмическая сложность

На этой стадии целесообразно еще раз вернуться к связи между физико-химической сложностью и концепцией *алгоритмической сложности*, введенной в разд. 1.5. Как мы уже видели, алгоритмическая сложность представляет собой длину самого короткого способа описания данной (конечной) последовательности знаков. В этом смысле некоторая совершенно случайная последовательность, порождаемая шумовым процессом, имеет максимально возможную алгоритмическую сложность, по существу равную длине этой последовательности. В то же время последовательность типа XXX... имеет минимальную возможную сложность, поскольку для ее воспроизведения достаточно единственной инструкции. Аналогичным образом следует считать максимальной информацию, содержащуюся в некоторой конкретной случайной последовательности, так как реализация этой конкретной последовательности из громадного числа случайных последовательностей равносильна локализации системы в очень узкой области пространства состояний. В последовательности с минимальной сложностью информация равна нулю, поскольку с ее помощью можно построить лишь одинаковые последовательности.

Мы уже отмечали, что сложность естественных объектов лежит где-то между этими двумя пределами, поскольку помимо случайности такие объекты содержат также некоторые типы крупномасштабной упорядоченности. Особенно четко и ярко этот дуализм ухвачен в знаменитом афоризме Эрвина Шрёдин-

гера, согласно которому ДНК представляет собой «аперiodический кристалл».

Заметим теперь, что самоорганизованные состояния материи, допускаемые законами физики неравновесных процессов, представляют собой образцы моделей со сложностью именно такого типа. Из этих состояний важнейшей для наших целей является хаотическая динамика. В самом деле, связанная с разупорядоченностью неустойчивость движения позволяет системе непрерывно прощупывать собственное пространство состояний, создавая тем самым информацию и сложность. С другой стороны, являясь результатом некоторого физического механизма, эти состояния создаются с вероятностью единица, так что проблема выбора конкретной последовательности из очень большого числа априорно равновероятных последовательностей попросту не возникает. В некотором отношении динамическая система, порождающая хаос, действует как своего рода селектор, отбрасывающий огромное большинство случайных последовательностей и сохраняющий лишь те из них, которые совместимы с соответствующими динамическими законами. По-видимому, не менее важно то обстоятельство, что присущая этим законам необратимость допускает существование асимптотически устойчивых и тем самым воспроизводимых аттракторов. В заключение этой главы, пожалуй, уместно отметить, что мы, наконец, установили физическое происхождение и некоторые правдоподобные математические прототипы сложного.

НА ПУТИ К ЕДИНОЙ ФОРМУЛИРОВКЕ ПОНЯТИЯ СЛОЖНОГО

Одна из ведущих тем этой книги — *выяснение динамических основ сложных явлений*. Сюда относятся как детерминистическая теория динамических систем, так и понятия теории вероятностей (и марковские цепи в особенности). Мы уже отмечали основное различие между консервативными и диссипативными системами (гл. 2). Тот факт, что для понимания динамических основ сложного приходится привлекать самые разнообразные средства, является отражением широкого набора механизмов возникновения сложных явлений, таких, как хаотическое поведение неинтегрируемых консервативных систем, нарушение симметрии в диссипативных системах или появление крупномасштабных корреляций в неравновесных условиях.

Каковы же соотношения между различными средствами, привлекаемыми в этой области? Возможен ли переход от консервативных систем к диссипативным? Какова роль вероятности? Мы собираемся рассмотреть эти вопросы в данной главе. Поскольку понятие диссипативности тесно связано со вторым законом термодинамики, а энтропия — с вероятностью (разд. 4.3), мы возвращаемся к одному из фундаментальных вопросов физики о смысле и роли необратимости.

Этот вопрос интриговал физиков с момента формулировки второго закона термодинамики, т. е. с середины XIX в. Второй закон — результат многовекового опыта, накопленного главным образом в области техники и физической химии. И тем не менее он длительное время оставался в стороне от основного русла физических исследований. Второй закон позволяет ввести новую физическую величину — энтропию, придающую времени выделенное направление, или, по выражению Эддингтона, «стрелу времени». С другой стороны, еще со времен Ньютона в физике основной задачей стало постижение фундаментальных свойств материи, и в этой задаче время не фигурирует. Эта программа не изменилась даже после таких великих революций в нашем сознании, какие были совершены теорией относительности и квантовой механикой.

Ввиду вопиющего противоречия между законами консервативных динамических систем и вторым законом термодинамики большинство физиков полагали, что второй закон обуслов-

лен какими-то приближениями, которые привносятся в законы динамики. Например, Макс Борн утверждал, что «необратимость является следствием того, что недостаток наших знаний явным образом привносится в фундаментальные (динамические) законы».

Однако многие примеры, рассмотренные в этой книге, четко показывают, что в природе необратимость играет фундаментальную роль. Мы попросту не поймем большинство наблюдаемых вокруг явлений, будь то рост кристаллов, регуляция биологических систем или история климата, если не будем пользоваться представлениями о диссипативных системах. Роль необратимости постоянно возрастает в самых различных областях физики. Кто мог бы ожидать, что большинство, если не все, элементарных частиц окажется неустойчивым? Или что диссипативные процессы сыграли столь важную роль на ранних стадиях космологической эволюции? Таким образом, сейчас уже невозможно просто отделаться от необратимости, как, казалось бы, того требует ее противоречие общепринятым взглядам в физике. Прежде чем мы приступим к динамической формулировке второго закона термодинамики, представляется целесообразным подытожить основные свойства консервативных и диссипативных динамических систем, обсуждавшиеся в предыдущих главах.

5.1. Общие свойства консервативных динамических систем

Динамика какой-либо консервативной системы соответствует точечному преобразованию S_t в соответствующем фазовом пространстве Γ :

$$\omega_t = S_t \omega_0 \quad (5.1a)$$

где ω_t и ω_0 — точки фазового пространства в моменты времени t и 0 соответственно. Как следует из теоремы Лиувилля, фундаментальная особенность таких систем состоит в том, что соответствующее движение точек в фазовом пространстве можно рассматривать как течение «несжимаемой жидкости». Иными словами, «мера» фазового пространства (см. разд. 3.2) сохраняется. Динамику таких систем можно также представить в виде унитарного оператора U_t , действующего на функцию распределения (см. разд. 4.3):

$$\rho_t(\omega) = U_t \rho_0(\omega) \quad (5.1b)$$

Здесь $\rho_t(\omega)$ — функция распределения в момент времени t , а $\rho_0(\omega)$ — начальная функция распределения*. Как отмечалось

* Поскольку движение можно представить с помощью траекторий, описываемых (5.1a), уравнение (5.1b) эквивалентно $\rho_t(\omega_t) = \rho_0(\omega_0)$ или $\rho_t(\omega) = \rho_0(S_{-t}\omega)$.

выше [см. уравнение (4.22б)], операторы U_t образуют динамическую группу

$$U_t U_s = U_{t+s} \quad (t, s \text{ положительны или отрицательны}) \quad (5.2)$$

В рамках такого формального описания можно рассматривать целые классы динамических систем с самыми различными типами поведения. В классической физике в центре внимания были интегрируемые системы типа свободно колеблющегося маятника или задачи двух тел. Теперь мы знаем, что такие системы относятся к исключительно простому классу, дающему неадекватное представление о сложности динамических систем. Таким образом, открытие целого ряда типов динамического поведения неинтегрируемых систем знаменует собой один из главных прорывов в современной физике и математике (см. разд. 3.8).

Что было бы, если бы в мире существовали одни лишь консервативные динамические системы? * Главное в ответе на этот вопрос состоит в том, что такой мир был бы неустойчивым, поскольку свойство асимптотической устойчивости, позволяющее системе забыть случайные возмущения, не имеет места в случае консервативных динамических систем (см. разд. 2.6). В лучшем случае мы имели бы «орбитальную устойчивость» вблизи некоей эллиптической траектории, когда возмущения сдвигают систему с одной орбиты на другую. Хуже дело обстояло бы с гиперболическими точками, поскольку здесь возмущения приводят к неустойчивости. Кроме того, в общем случае имеется сколько угодно эллиптических и гиперболических точек. Поэтому в таком мире можно ожидать исключительно высокой чувствительности к начальным условиям. Один пример поведения такого типа мы уже видели в случае подковы, а в дальнейшем в качестве еще более простой иллюстрации орбитальной случайности будет рассмотрено преобразование пекаря (разд. 5.5).

Еще один пример неустойчивой системы дает движение с участием жестких шаров. Эта многократно изучавшаяся система основана на модели Лоренца, названной так в честь нидерландского физика Хендрика Антона Лоренца. В этой модели рассматривается движение маленького шарика, сталкивающегося со случайно расположенными в пространстве большими шарами, фиксированными на своих местах (см. рис. 5.1). Траектория такого маленького подвижного шарика всегда четко определена. Однако если внести малейшую неопределенность в начальное условие, то при последовательных столкновениях эта

* В том смысле, что макроскопические переменные подчинялись бы обрабатываемой во времени консервативной динамике.

неопределенность будет очень быстро возрастать. Заметим, что химические реакции часто можно рассматривать как результат столкновения шаров, когда кинетическая энергия относительно движения превышает определенный порог (так называемую энергию активации). Можно снова ожидать, что такое движение будет сильно неустойчивым.

Как уже отмечалось, в процессе движения объем («мера») фазового пространства сохраняется, в то время как форма соот-

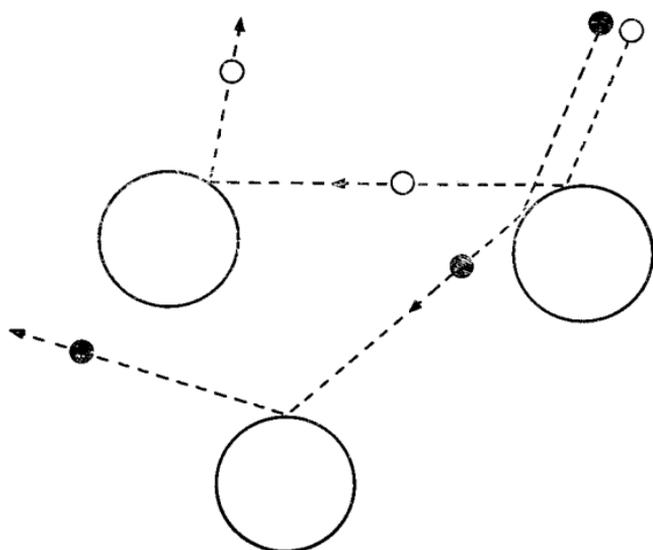


Рис. 5.1. Схематическое представление неустойчивости траектории маленького шарика, отражающегося от больших шаров. Малейшая неточность в определении исходного положения маленького шарика делает невозможным предсказание того, с каким из больших шаров он столкнется уже после нескольких первых столкновений.

ветствующего элемента пространства вследствие неустойчивости сильно изменяется. Как будет показано в разд. 5.5, начальный элемент объема в процессе движения может расщепляться на все более и более мелкие кусочки. Именно эта деформация, или фрагментирование, создает впечатление приближения к равновесию, в котором все точки должны равномерно распределиться по фазовому пространству, точнее, по его доступным областям. В течение почти трех веков со времен Ньютона классическая динамика представлялась некоторой завершенной наукой, позволяющей находить любые траектории из первых принципов и определенных начальных условий. Теперь мы видим, что это справедливо лишь для ограниченного класса динамических систем. В случае достаточно неустойчивых динамических систем в каждой области фазового пространства независимо от ее раз-

меров содержатся расходящиеся траектории. В этом случае, для того чтобы можно было говорить об отдельной четко определенной траектории, нам потребовалось бы задать начальные условия с бесконечно высокой точностью. Иными словами, нам нужна была бы бесконечная информация, обусловленная бесконечным количеством цифр, требующихся для задания начальных данных. Как будет показано в дальнейшем, именно устранение этой бесконечной информации приводит к необратимости. Разумеется, это относится лишь к классической механике, поскольку в квантовой теории распад неустойчивых частиц представляет собой дополнительный источник необратимости.

Находясь в мире неустойчивых динамических систем, мы можем рассматривать внешние события лишь «через окошко». Таким образом, здесь наблюдается крушение идеала «полного знания», царившего в западной науке в течение трех столетий.

5.2. Общие свойства диссипативных динамических систем

По своему разнообразию диссипативные динамические системы совершенно экстраординарны. В этой книге рассмотрены многие примеры из химии, механики жидких сред и других естественных наук. Нам известно, что в отличие от консервативных систем диссипативные системы не сохраняют меру в фазовом пространстве (разд. 3.2). Кроме того, существенным общим моментом для диссипативных систем является их неинвариантность относительно обращения времени. Отсюда следует, что асимптотически, т. е. при $t \rightarrow +\infty$, будут достигаться некоторые выделенные ситуации. Если не налагать никаких дополнительных условий, то такая выделенная ситуация будет соответствовать равновесию. Например, в соответствии с уравнениями (2.7б) и (2.7в) концентрация и температура характеризуются тенденцией к выравниванию. Аналогичным образом концентрации, соответствующие некоторой смеси реагентов, будут стремиться к значениям, соответствующим закону действующих масс. Если система находится в неравновесных условиях, то изображающая точка в фазовом пространстве будет стремиться к некоторому притягивающему множеству, будь то точечный аттрактор, предельный цикл или «странный» аттрактор типа того, что обсуждался в разд. 3.10. Существенно лишь то, что для всех диссипативных систем характерно наличие выделенного направления времени. В принципе можно представить себе такой мир, в котором некоторые биологические системы стареют, а некоторые молодеют, т. е. мир, в котором одни диссипативные системы стремятся к равновесию при $t \rightarrow +\infty$, а другие — при $t \rightarrow -\infty$. Ясно лишь, что в нашем мире такое невозможно, по-

сколькx в нем нарушена универсальная временная симметрия. И это, как заметил Норберт Винер, хорошо, поскольку в любом мире, для того чтобы можно было обмениваться сообщениями, направление времени должно быть одно и то же. В отличие от орбитальной случайности, характерной для консервативных систем, здесь мы обнаруживаем асимптотическую устойчивость. Это обеспечивает возможность появления тонко сбалансированных процессов, таких, какие мы наблюдаем, например, в биологии. Однако на уровне аттракторов случайность все-таки еще может сохраниться.

В разд. 3.10 мы рассматривали различные модели, позволяющие наблюдать хаотическое поведение. Однако теперь показатели Ляпунова, в соответствии с приложением I характеризующие неустойчивость движения, определяются не микроскопической динамикой, а микроскопическими величинами типа скоростей химических реакций и коэффициентов диффузии.

С нарушением универсальной симметрии времени связана не только диссипативность — такое нарушение позволяет связать с диссипативными процессами функцию Ляпунова, как, например, энтропия, монотонно изменяющуюся во времени. В самом деле, все до сих пор изучавшиеся диссипативные процессы приводят к положительному производству энтропии (разд. 2.5). Само существование диссипативных систем есть проявление второго закона термодинамики.

5.3. Поиски унификации

Имеется еще одна точка зрения, позволяющая различать консервативные и диссипативные динамические системы. Представляется невозможным найти более «фундаментальный» уровень описания, с которого консервативные динамические системы выглядели бы как результат какого-то приближения. Конечно, если ввести в рассмотрение квантовую теорию, то классические траектории становятся результатом усреднения по волновым пакетам. Однако это не приводит к новой точке зрения, поскольку уравнение Шрёдингера — основное уравнение квантовой теории — само по себе является детерминистическим и обратимым во времени, как и уравнения классической механики*.

Совершенно другая ситуация имеет место в случае диссипативных систем. В качестве примера рассмотрим химическую

* Аналогичным образом не помогла бы и замена динамики частиц уравнениями поля типа тех, что используются в электродинамике или теории относительности, поскольку и здесь уравнения поля детерминистичны и обратимы во времени.

кинетику. Кинетические уравнения не соответствуют «фундаментальному» уровню, поскольку, как было показано ранее, они являются результатом приближенного подхода, пренебрегающего *флуктуациями* [см. особо уравнение (4.15)]. Мы наблюдаем также *корреляции*, отражающие согласованное поведение, характерное для неравновесных ситуаций. Следовательно, для диссипативных систем фундаментальный уровень следует ассоциировать с *вероятностным подходом*, примером которого может служить теория марковских цепей, рассматривавшаяся в гл. 4.

Можно ли обобщить такое марковское описание с пространства состояний, выражающихся через макроскопические переменные, на полное фазовое пространство Γ , в котором до сих пор было возможно лишь описание на языке обратимых процессов? Мы попытаемся сделать такое обобщение, вводя в Γ функцию распределения $\tilde{\rho}(\omega)$, которая в отличие от классической механики претерпевает эволюцию в соответствии с законами марковских процессов. По аналогии с (5.1) для $\tilde{\rho}_t$ в момент времени t можно написать

$$\tilde{\rho}_t(\omega) = W_t \tilde{\rho}_0(\omega) \quad (5.3)$$

Здесь $\tilde{\rho}_t(\omega)$ — решение основного уравнения, описывающего марковский процесс. Как было показано, в случае описания обратимых процессов с помощью уравнения Лиувилля выполняется условие группы (5.2). Оператор же W_t удовлетворяет условию полугруппы:

$$W_t W_s = W_{t+s} \quad (t, s \geq 0) \quad (5.4)$$

Марковский процесс (5.3) при условии (5.4) соответствует движению системы к равновесию или к некоторому неравновесному состоянию, характеризующему тем или иным аттрактором.

В заключение отметим, что в динамике консервативных систем, по-видимому, нет места истинно вероятностным эволюционным процессам — такие процессы составляют основу описания диссипативных систем. Следовательно, чтобы установить соотношение между консервативными и диссипативными динамическими системами, необходимо выяснить связь между детерминистической динамикой и вероятностным подходом. После этого будет установлена связь между динамикой и вторым законом термодинамики на основе соотношения между энтропией и вероятностью. Поскольку основные законы динамики консервативных систем выражаются в виде (5.1), а истинно вероятностные процессы удовлетворяют соотношениям (5.3) — (5.4), нам нужно также обсудить, как может выглядеть связь между (5.1) и (5.3) — (5.4).

5.4. Вероятность и динамика

Основную роль в теории марковских процессов играет переходная вероятность. Поскольку нам желательно установить истинно вероятностное описание в фазовом пространстве Γ , рассмотрим переходную вероятность $P_t(\omega|\Delta)$, характеризующую переход из точки ω в некоторую область Δ в момент времени t и равную некоторому неотрицательному числу меньше единицы. Здесь на первый план выступает основное различие между марковским описанием и описанием на языке траекторий.

Допустим, что траектория проходит из ω в ω_t . Тогда для системы, движущейся вдоль этой траектории, имеем (см. рис. 5.2)

$$P_t(\omega|\Delta) \begin{cases} = 1, & \text{если } \omega_t \text{ содержится в } \Delta \\ = 0, & \text{если } \omega_t \text{ не содержится в } \Delta \end{cases} \quad (5.5)$$

Напротив, в аналогичном случае для марковского процесса по крайней мере некоторые переходные вероятности не равны ни нулю, ни единице.

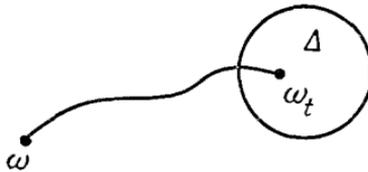


Рис. 5.2. Переходная вероятность $P_t(\omega|\Delta) = 1$, характерная для детерминистического описания, основанного на траекториях.

У нас нет оснований считать неверным динамическое описание (5.1). Однако оно может оказаться неполным, поскольку в нем не содержится элемента нелокальности, который, очевидно, необходим для заполнения брешы между консервативными и диссипативными системами путем перехода от вырожденных переходных вероятностей (5.5) к собственно марковским вероятностям. В связи с этим напрашивается возможность установления связи между описаниями (5.1) и (5.3) с помощью преобразования

$$\tilde{\rho} = \Lambda \rho \quad (5.6)$$

Недостающая нам нелокальность, разумеется, зависящая от динамических законов, должна содержаться в операторе Λ , действующем на ρ . После того как оператор Λ будет определен, уравнение (5.6) установит связь между функцией распределения ρ , эволюционирующей в соответствии с законами обратимой динамики, и функцией распределения $\tilde{\rho}$, эволюционирующей в соответствии с некоторым марковским процессом.

Соотношение (5.6) позволяет установить замечательную связь между операторами U_t и W_t . В самом деле, подставляя (5.16) и (5.3) в (5.6), получаем

$$\wedge U_t \rho_0 = W_t \wedge \rho_0 \quad (5.7)$$

Поскольку это равенство справедливо при любых ρ_0 , можно прийти к выводу о равенстве соответствующих операторов:

$$\wedge U_t = W_t \wedge \quad (5.8)$$

Если оператор \wedge допускает обращение, соотношение (5.8) можно переписать в виде

$$W_t = \wedge U_t \wedge^{-1}, \quad U_t = \wedge^{-1} W_t \wedge \quad (5.9)$$

С этой точки зрения теория вероятностей представляется неким «образом» динамики, опосредованным преобразованием \wedge .

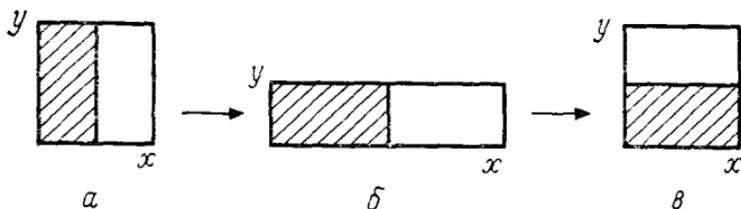


Рис. 5.3. Преобразование пекаря. Показанный на рис. а единичный квадрат расплющивается в прямоугольник $(1/2) \times 2$ (рис. б). После этого правая половина прямоугольника налагается на левую с образованием нового квадрата (рис. в).

И наоборот, всегда можно восстановить динамику, лежащую в основе некоторого марковского процесса, если только можно установить вид оператора \wedge . Таким образом, центральная проблема при переброске моста между детерминистической динамикой и вероятностным подходом — это построение оператора \wedge . Перейдем теперь к рассмотрению некоторого класса консервативных динамических систем, для которых удастся построить этот оператор явным образом.

5.5. Преобразование пекаря

Рассмотрим некоторый класс чрезвычайно неустойчивых динамических систем, называемых К-струями (обозначение К принято в честь А. Н. Колмогорова). Пример такого рода — подкову — мы уже изучали в разд. 3.9. Здесь мы рассмотрим другой знаменитый пример — так называемое преобразование пекаря. Определим следующее отображение (см. рис. 5.3) в фазовом

пространстве, представляющем собой единичный квадрат $0 \leq x \leq 1, 0 \leq y \leq 1$:

$$\begin{aligned} x' &= 2x, & y' &= \frac{y}{2} & \text{при } 0 \leq x < \frac{1}{2} \\ x' &= 2x - 1, & y' &= \frac{y+1}{2} & \text{при } \frac{1}{2} \leq x \leq 1 \end{aligned} \quad (5.10)$$

Это преобразование соответствует дискретному динамическому процессу, сохраняющему меру (в данном случае — площадь поверхности) и реализующемуся через регулярные временные интервалы, скажем каждую секунду. Следовательно, этот процесс определяет некоторую консервативную динамическую систему. Очевидно, это преобразование допускает обращение, поскольку единичный квадрат отображается сам на себя (в отличие от преобразования типа подковы).

Несколько труднее установить обратимость такой динамики во времени. Выпишем решение системы уравнений (5.10) относительно x и y :

$$\begin{aligned} x &= \frac{x'}{2}, & y &= 2y', & 0 \leq y < \frac{1}{2} \\ x &= \frac{x'+1}{2}, & y &= 2y' - 1, & \frac{1}{2} \leq y \leq 1 \end{aligned} \quad (5.11)$$

Легко видеть, что равенства (5.11) получаются из (5.10) путем обращения времени с заменой x на y и y на x . Такая взаимозаменяемость аналогична обращению импульсов (при неизменных пространственных координатах) при обращении времени в уравнениях механики в гамильтоновой форме (см. разд. 2.1).

Возвращаясь к уравнениям (5.10), можно видеть, что в направлении x происходит растяжение, а в направлении y — сжатие. После n последовательных преобразований расстояние между двумя точками на оси x может увеличиться в $2^n = e^{n \ln 2}$ раз. В соответствии с определениями, данными в приложении I, мы имеем здесь положительный показатель Ляпунова

$$\sigma_1 = \ln 2 \quad (5.12)$$

Это свидетельствует о хаотическом характере динамики. Естественно, что второй показатель Ляпунова отрицателен, $\sigma_2 = -\ln 2$, что обусловлено консервативностью процесса [см. равенство (I.27)].

Легко также видеть, что при повторении процесса, показанного на рис. 5.3, с течением времени каждая конечная область будет разбиваться на все более и более тонкие полоски, и это постепенно приведет к практически равномерному распределению по всему единичному квадрату (см. рис. 5.4).

Механизм действия преобразования пекаря легче понять, если представить его в виде процесса, известного под названием *сдвига Бернулли*. Для этого поставим в соответствие каждой точке (x, y) единичного квадрата бесконечную в обе стороны последовательность чисел $\{u_n\}$, определенную следующим двоичным представлением чисел x и y :

$$x = \sum_{n=-\infty}^0 \frac{u_n}{2^{-n+1}}, \quad y = \sum_{n=1}^{\infty} \frac{u_n}{2^n} \quad (5.13)$$

где каждое из u_n может принимать значения либо 0, либо 1. Подставляя эти определения в (5.10), получаем

$$\begin{aligned} \frac{u_0'}{2} + \frac{u_{-1}'}{2^2} + \dots &= x' = \frac{u_{-1}}{2} + \frac{u_{-2}}{2^2} + \dots \\ \frac{u_1'}{2} + \frac{u_2'}{2^2} + \dots &= y' = \frac{u_0}{2} + \frac{u_1}{2^2} + \dots \end{aligned}$$

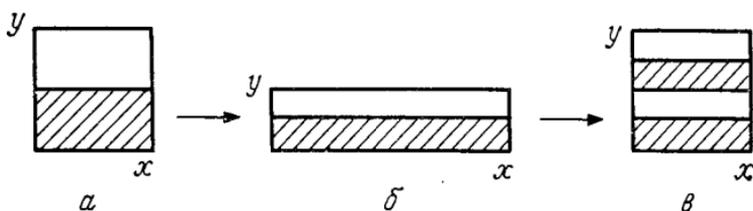


Рис. 5.4. Последовательные итерации преобразования пекаря приводят к фрагментированию заштрихованных и светлых полосок на все большее и большее число разъединенных областей.

Следовательно, эволюция во времени, представляемая преобразованием пекаря, соответствует сдвигу

$$u_n' = u_{n-1} \quad (5.14)$$

называемому сдвигом Бернулли. Отсюда видно, что содержащаяся в начальных условиях информация включает в себя всю прошлую и будущую историю системы, как того и следовало ожидать.

Заметим, что число u_0 определяет, к какой из двух половин единичного квадрата относится изображающая точка в фазовом пространстве — к левой ($u_0=0$) или правой ($u_0=1$). Поскольку цифры, соответствующие $u_0, u_{-1}, u_{-2}, \dots$, можно определить путем подбрасывания монеты, последовательные во времени значения $u_0, u_0' = u_{-1}, u_0'' = u_{-2}$ будут обладать теми же случайными свойствами, что и результат подбрасывания монеты. Отсюда следует, что событие, состоящее в принадлежности системы левой или правой половине квадрата, можно рас-

смаивривать как процесс Бернулли — один из наиболее известных примеров случайных процессов в теории вероятностей.

Преобразование пекаря обладает также важным свойством *возвратности* состояний. Рассмотрим некоторую точку (x, y) , для которой последовательность $\{u_n\}$ в двоичном представлении (5.13) является периодической. В этом случае числа x и y являются рациональными (заметим, однако, что отнюдь не все рациональные числа имеют периодическое представление, хотя подавляющее их большинство такое представление имеет). Поскольку динамика сдвигает все u_n одинаковым образом, все такого рода состояния будут циклически повторяться с определенным периодом. Еще более важно, что в некотором смысле, который мы сейчас уточним, тот же результат справедлив практически и для всех остальных состояний (за исключением множества меры нуль)! Чтобы убедиться в этом, рассмотрим двоичное представление некоторой иррациональной точки, соответствующее бесконечной непериодической последовательности из единиц и нулей. Можно показать, что практически во всех иррациональных числах (за исключением множества меры нуль) любая конечная последовательность цифр встречается бесконечно часто. Таким образом, данная последовательность, скажем из $2m$ цифр, ближайших от 0, определяющая состояние с погрешностью 2^{-m} , будет повторяться за счет сдвига бесконечно часто. Поскольку число m можно выбрать сколь угодно большим, хотя и конечным, это означает, что почти каждое состояние будет возвращаться сколь угодно близко к любой точке, включая, разумеется, и начальную. Иными словами, большинство траекторий будет «эргодическим» и «закроет» все фазовое пространство. В этом состоит содержание знаменитой теоремы Пуанкаре о возвращении, исходно доказанной для весьма общего класса консервативных динамических систем. Долгое время это свойство рекуррентности наряду с обратимостью во времени рассматривалось как важный аргумент против существования истинно диссипативных процессов. Как будет вскоре показано, эта точка зрения оказывается несостоятельной.

Суммируя, отметим свойства динамики, описываемой преобразованием пекаря: эта динамика консервативна, допускает обращение, обратима во времени, возвратна и хаотична. Особую ценность преобразованию пекаря придает то обстоятельство, что те же свойства характерны и для реальных систем со сложным поведением. Поразительно то, что, несмотря на обратимость и возвратность, свойство хаотичности позволяет установить истинную необратимость, воспользовавшись более адекватным описанием без каких-либо приближений.

Динамика консервативных систем определяется законами движения и начальными условиями. Законы движения здесь

просты. Рассмотрим несколько подробнее вопрос о начальных данных. Такие данные для отдельной траектории соответствуют бесконечному набору чисел $\{u_n\}$ (n изменяется от $-\infty$ до $+\infty$). Однако реальный мир нам виден лишь через окошко конечной ширины. В данном случае это означает, что нашему контролю поддается некоторое произвольное, но конечное число цифр u_n . Допустим, что такое окошко соответствует набору

$$u_{-3}u_{-2}u_{-1} \cdot u_0u_1u_2u_3 \quad (5.15)$$

а все остальные цифры неизвестны (здесь точка указывает границу раздела между цифрами, соответствующими x и y). Наличие сдвига Бернулли означает, что на следующем шаге набор (5.15) будет заменен набором

$$u_{-4}u_{-3}u_{-2}u_{-1} \cdot u_0u_1u_2 \quad (5.16)$$

в котором уже содержится неизвестная цифра u_{-4} . Продолжая эти рассуждения, можно прийти к более строгой формулировке результата: благодаря существованию положительного показателя Ляпунова для того чтобы иметь возможность определить положение с точностью до N цифр после n итераций, необходимо знать начальное положение с точностью до $N+n$ цифр.

Для того чтобы справиться с такой ситуацией, статистическая механика по традиции предложила бы ввести «мозаичное» распределение вероятностей, определенное не на отдельных точках, а сразу на целых областях единичного квадрата, как это и было сделано впервые Эренфестом. Однако на расширяющемся многообразии можно найти две выделенные точки даже в том случае, когда их нельзя различить в начальный момент времени с помощью измерений с заданной конечной точностью. Поэтому традиционный прием с мозаичным распределением не заменяет динамической эволюции. По этой причине нам нужен какой-то более тонкий метод, который и будет развит в двух последующих разделах.

5.6. Многообразия с нарушенной временной симметрией

Несмотря на тот факт, что К-струи обратимы во времени, в них содержатся объекты, характеризующиеся нарушением временной симметрии. Размерность этих объектов меньше, чем у полного фазового пространства. В случае преобразования пекаря такими объектами являются «сокращающиеся волокна» и «удлиняющиеся волокна» (см. рис. 5.5). Сокращающиеся волокна — это вертикальные линии; они сокращаются в 2 раза при каждом преобразовании пекаря. Удлиняющиеся волокна — горизонтальные линии — растягиваются во столько же раз. Каждая точка фазового пространства находится на пересечении со-

сжимающегося и удлиняющегося волокон. Сокращающиеся и удлиняющиеся многообразия переходят друг в друга при обращении времени и, очевидно, являются объектами, асимметричными во времени. Все точки сжимающегося многообразия стремятся к одному и тому же будущему. Напротив, точки удлиняющегося волокна равномерно покроют фазовое пространство при $t \rightarrow +\infty$. На вероятностном языке это соответствует приближению к однородному равновесному распределению в фазовом пространстве.

С какой бы точностью ни определялось состояние, всегда будет существовать такое характерное расстояние ϵ , ниже которого мы не сможем различать точки на сокращающихся волокнах.

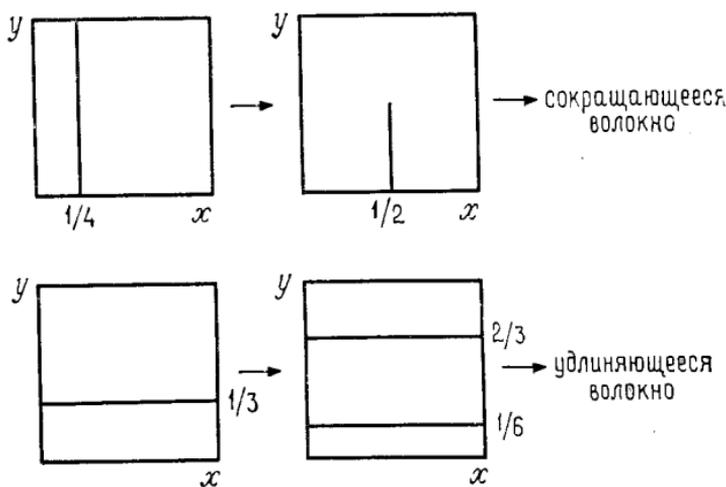


Рис. 5.5. Сокращающиеся и удлиняющиеся волокна в преобразовании пекаря.

Поскольку точки на таких волокнах принадлежат сходящимся траекториям, следует считать, что все траектории, начинающиеся внутри интервала ϵ , в будущем неразличимы. Это сразу же делает неприменимой теорему Лиувилля, выражающую сохранение меры в фазовом пространстве. К такому выводу можно прийти с помощью следующих простых соображений. Рассмотрим некоторую динамическую систему и допустим, что при $t=0$ функция распределения ρ сосредоточена на некоторой площади Δ_0 . С течением времени удлиняющийся размер будет возрастать экспоненциально, в то время как размер сокращающихся волокон уменьшится до ϵ и в дальнейшем должен считаться постоянным (см. рис. 5.6). В результате с этого момента мера Δ будет возрастать до тех пор, пока не достигнет меры всего фазового пространства. Заметим, что этот вывод не является результатом некоторого приближения — это результат того фундамен-

тального факта, что измерения (или, в более общем случае, эффекты физических взаимодействий) относятся не к математическим точкам, а к фазовым областям конечной меры. Конечно же, всегда понятно, что точные начальные условия соответствуют некоторой идеализации. Новым здесь является то, что отказ от этой идеализации приводит к столь важному следствию в чрезвычайно неустойчивых системах, какими являются К-струи.

5.7. Нарушающее симметрию преобразование Λ

Покажем теперь, как изложенные в предыдущем разделе представления приводят к построению нарушающего симметрию преобразования Λ [уравнение (5.6)], которое служит связующим звеном между динамическим и вероятностным процессами.

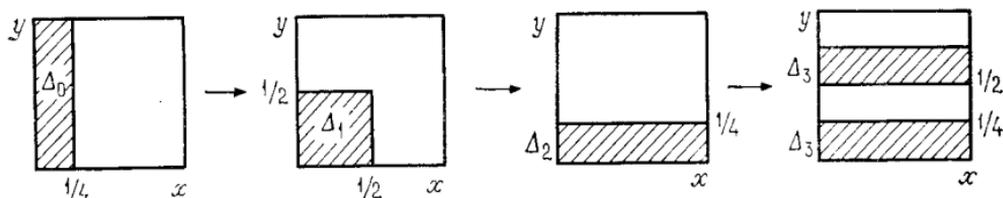


Рис. 5.6. Неадекватность теоремы Лиувилля в случае, когда наблюдение над сокращающимися волокнами происходит через окошко конечной ширины (здесь $\varepsilon = 1/4$). Мера начального объема Δ_0 остается инвариантной лишь в течение нескольких первых итераций, а затем возрастает, пока не достигнет полной меры фазового пространства.

С этой целью состояние рассматриваемой неустойчивой динамической системы удобно представить с помощью объектов, называемых разбиениями $\{\chi_n\}$. В случае преобразования пекаря такие разбиения построить очень легко. В самом деле, пусть χ_0 — функция, которая в левой половине единичного квадрата принимает значение -1 , а в правой — значение $+1$. Функцию χ_n определим как результат последовательных воздействий оператора преобразования (n — положительное или отрицательное число) на функцию χ_0 :

$$\chi_n = U^n \chi_0 \tag{5.17}$$

Здесь U — эволюционный оператор для функций распределения (5.16). Некоторые из этих разбиений представлены на рис. 5.7. Заметим, что χ_0 полностью определена цифрой u_0 в представлении (5.13), в то время как посредством (5.17) функция χ_n определена цифрой u_n .

Рассмотрим теперь, каким образом можно описать динамику на языке разбиений. Для этого полезно ввести еще одну характеристику разбиений — их «возраст» n . Это понятие, довольно

естественное с учетом (5.17), можно формализовать и обобщить с помощью важного результата Мисра, показавшего существование оператора T , для которого χ_n и n являются собственной функцией и собственным значением соответственно, $T\chi_n = n\chi_n$. В действительности каждое собственное значение n этого оператора является бесконечно вырожденным, поскольку кроме χ_n собственными функциями оператора T являются также все произведения

$$\varphi_n = \chi_{j_1} \chi_{j_2} \cdots \chi_{j_n}, \quad j_1 < j_2 < \cdots < n$$

Поэтому можно написать

$$T\varphi_n = n\varphi_n \quad (-\infty < n < \infty) \quad (5.18a)$$

Чтобы завершить определение оператора T , потребуем, чтобы функция $\varphi \equiv 1$ также была собственной, с собственным значением 1:

$$T \cdot 1 = 1 \quad (5.18b)$$

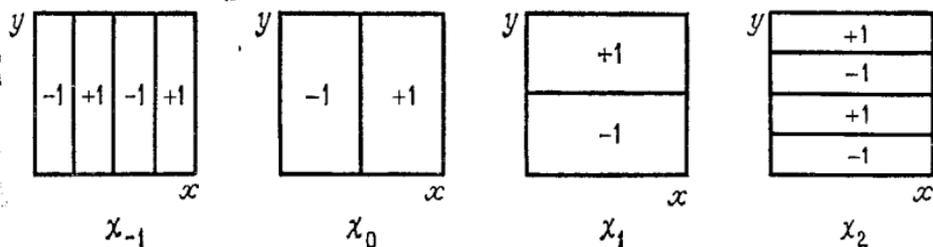


Рис. 5.7. Представление динамики преобразования пекаря с помощью разбиений.

а также чтобы T был линейным эрмитовым оператором. Можно показать, что функции φ_n вместе с $\varphi = 1$ образуют полный набор.

Заметим, что действие оператора T на разбиение χ_n не совпадает с действием динамического оператора U , поскольку в соответствии с (5.17)

$$U\chi_n = \chi_{n+1} \quad (5.19)$$

Благодаря этому различию операторы T и U не коммутируют друг с другом. Из (5.18) — (5.19) следует

$$TU\chi_n = T\chi_{n+1} = (n+1)\chi_{n+1}$$

С другой стороны,

$$UT\chi_n = Un\chi_n = n\chi_{n+1}$$

Следовательно,

$$(TU - UT)\chi_n = \chi_n \neq 0$$

Примечательно, что таким образом классическая динамика приводит к некоммутативной алгебре (в чем имеется сходство с квантовой механикой).

Сформулируем теперь следующее положение: нарушающее симметрию преобразование \wedge является линейным эрмитовым оператором, представляющим собой неотрицательную убывающую функцию T в том смысле, что

$$\begin{aligned} \wedge(T) \varphi_n &= \lambda_n \varphi_n \\ \wedge(T) 1 &= 1 \end{aligned} \quad (5.20)$$

где

$$1 \geq \dots \lambda_n \geq \lambda_{n+1} \geq \dots \geq 0$$

и

$$\lambda_n \rightarrow \begin{cases} 1 & \text{при } n \rightarrow -\infty \\ 0 & \text{при } n \rightarrow +\infty \end{cases}$$

Поскольку разбиения φ_n вместе с однородным распределением 1 образуют полный функциональный набор, любую функцию распределения ρ можно представить в виде

$$\rho = 1 + \sum_{n=-\infty}^{+\infty} c_n \varphi_n \quad (5.21)$$

В такой записи для простоты опущен дополнительный индекс, указывающий кратность вырождения собственных значений оператора T . С учетом (5.20) и соотношения $\tilde{\rho} = \wedge \rho$ отсюда следует

$$\tilde{\rho} = 1 + \sum_{n=-\infty}^{+\infty} c_n \lambda_n \varphi_n \quad (5.22)$$

Заметим, что при переходе от ρ к $\tilde{\rho}$ в данной точке фазового пространства задействованы все фазовые точки. Тем самым устанавливается ожидаемая нелокальность оператора \wedge .

Далее, все «состояния» φ_n входят в ρ линейно с весом c_n , зависящим от начальных условий. В этом плане здесь нет различия между прошлым ($n < 0$) и будущим ($n > 0$). Напротив, при данных начальных условиях вклад в $\tilde{\rho}$ функций φ_n , относящихся к будущему, затухает ($\lambda_n \rightarrow 0$ при $n \rightarrow \infty$). Вспомним теперь, что разбиение χ_n определяется цифрой u_n . Кроме того, цифры u_n при $n > 0$ целиком задают сокращающиеся волокна [см. (5.13)]. Следовательно, затухание $\lambda_n \rightarrow 0$ выражает тот факт, что происходит потеря информации относительно локализации на этих волокнах. Это в точности соответствует представлению о неразличимости соседних точек. Например, при $\lambda_n = 1$ для $n \leq 2$ и $\lambda_n = 0$ для $n > 2$ естественным расстоянием ε , введенным в разд. 5.6, будет $\varepsilon = 1/8$. Интуитивные соображения о том,

каким образом неразличимость может привести к неприменимости теоремы Лиувилля, были изложены в разд. 5.6.

Покажем теперь, что для новой величины $\tilde{\rho}$ в соответствии со вторым законом термодинамики можно определить некоторую монотонную функцию (см. разд. 4.3). Рассмотрим так называемую \mathcal{H} -функцию:

$$\mathcal{H}_{\tilde{\rho}} = \int \tilde{\rho}^2 d\omega \quad (5.23)$$

Заметим, что к аналогичному результату привело бы использование какой-либо другой выпуклой функции, например типа $\int \tilde{\rho} \ln \tilde{\rho} d\omega$. Допустим, что при $t=0$ функции ρ и $\tilde{\rho}$ определены соотношениями (5.21) и (5.22). Тогда, воспользовавшись (5.19) и очевидным равенством $U \cdot 1 = 1$, имеем

$$\rho_1 = U\rho_0 = 1 + \sum_n c_n \lambda_{n+1} \quad (5.24)$$

$$\tilde{\rho}_1 = \wedge \rho_1 = 1 + \sum_n c_n \lambda_{n+1} \lambda_{n+1} \quad (5.25)$$

В результате из (5.23) получаем

$$\mathcal{H}_{\tilde{\rho}_0} - \mathcal{H}_{\tilde{\rho}_1} = \sum_n (\lambda_{n+1}^2 - \lambda_n^2) c_n^2 \leq 0 \quad (5.26)$$

т. е. со временем $\mathcal{H}_{\tilde{\rho}}$ монотонно убывает. Напротив, величина $\mathcal{H}_{\rho} = \int \rho^2 d\omega$, выраженная через исходную функцию распределения ρ , остается постоянной. Только что полученный результат подтверждает предположение о монотонности, сделанное при введении оператора \wedge [см. равенство (5.20)]. В действительности монотонность собственных значений λ'_n является как необходимым, так и достаточным условием справедливости неравенства (5.26).

В заключение отметим, что наша цель достигнута — построено преобразование \wedge и получена микроскопическая формулировка второго закона термодинамики. Это ярко демонстрирует важную роль функции распределения $\tilde{\rho}$. На самом деле подразумевается, что не только энтропия (через $\mathcal{H}_{\tilde{\rho}}$), но и любая другая наблюдаемая величина должна выражаться через $\tilde{\rho}$.

До сих пор предполагалось лишь то, что оператор преобразования $\wedge(T)$ является убывающей функцией T — в остальных отношениях это была произвольная функция. Здесь у нас нет возможности подробно обсуждать вид зависимости $\wedge(T)$ или, что то же самое, зависимости λ_n от n . Поскольку единственным внутренним параметром нашей модели является показатель Ля-

пунова (5.12), естественно ожидать, что по крайней мере при больших положительных n

$$\lambda_n = e^{-n/\tau} \quad (5.27)$$

где $1/\tau = \ln 2$ — показатель Ляпунова. Как бы конкретно ни выглядели собственные значения λ_n , важно то, что нелокальность и нарушение временной симметрии являются фундаментальными свойствами оператора \wedge . Именно эти свойства обеспечивают переход от неустойчивых консервативных динамических систем к диссипативным системам.

5.8. Ансамбли Гиббса и Больцмана

Классическая динамика приводит к некоторому точечному преобразованию. В случае ансамбля Гиббса каждой точке фазового пространства соответствует один ансамбль, представленный с помощью δ -функции. При этом движение $\omega_0 \rightarrow \omega_t$ преобразует одну δ -функцию в другую:

$$\delta_{\omega_t} = U_t \delta_{\omega_0} \quad (5.28)$$

где $\delta_{\omega_0} = \delta(\omega - \omega_0)$ и т. д. По этой причине представление δ -функции инвариантно при действии унитарного оператора U_t . Применим преобразование \wedge к δ_{ω_0} :

$$\tilde{\delta}_{\omega_0} = \wedge \delta_{\omega_0} \quad (5.29)$$

В общем случае функция $\tilde{\delta}_{\omega_0}$ не является строго локализованной. Однако в силу монотонности собственных значений λ_n она сильно сконцентрирована на сокращающемся волокне, проходящем через точку ω_0 , причем вдоль этого волокна функция представлена в виде более или менее широкого распределения с максимумом вблизи точки ω_0 . На рис. 5.8 представлен пример, когда $\tilde{\delta}_{\omega_0}$ целиком сконцентрирована на сокращающемся волокне. Разумеется, эволюция $\tilde{\delta}_{\omega_0}$ во времени описывается уравнением (5.3), $\tilde{\delta}_{\omega_t} = W_t \tilde{\delta}_{\omega_0}$. В представленном на рис. 5.8 случае можно показать, что действие оператора W_t равносильно жесткому смещению пика $\tilde{\delta}$ -распределения.

В традиционном представлении динамики основными элементами являются точки. Теперь роль таких элементов играют нелокальные образования $\tilde{\delta}_{\omega_0}$, которые по существу связаны с отрезками сокращающихся волокон. Иными словами, вместо отдельных траекторий, представленных δ -функциями, мы имеем теперь дело с *пучками траекторий*, представленных $\tilde{\delta}$ -функциями.

Ансамбли Гиббса можно рассматривать как суперпозицию точек в фазовом пространстве с помощью представления

$$\rho(\omega) = \int_{\Gamma} \rho(\omega_0) \delta(\omega - \omega_0) d\omega_0 \quad (5.30)$$

В случае неустойчивых динамических систем в этом соотношении следует заменить ρ на $\tilde{\rho}$. Действуя на обе стороны (5.30) оператором \wedge , можно видеть, что ансамбли $\tilde{\rho}$ являются суперпозициями δ -функций:

$$\tilde{\rho}(\omega) = \wedge \rho(\omega) = \int_{\Gamma} \rho(\omega_0) \tilde{\delta}(\omega - \omega_0) d\omega_0 \quad (5.31)$$

Эти ансамбли можно назвать *больцмановскими*, поскольку их эволюция описывается марковскими полугруппами и их использование приводит к монотонному возрастанию энтропии. Переход от больцмановских ансамблей к гиббсовым соответствует переходу от произвольной конечной точности к бесконечно точ-

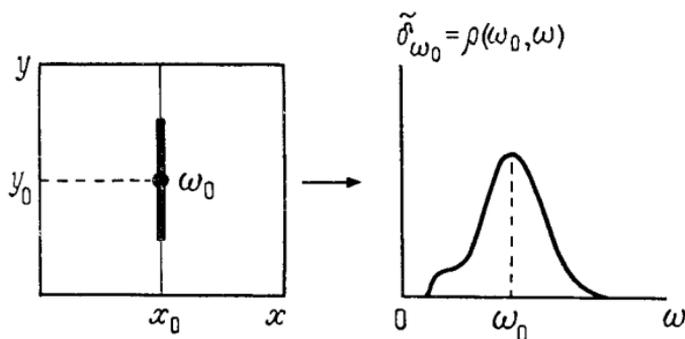


Рис. 5.8. Типичный вид плотности $\tilde{\delta}_{\omega_0}$ на сокращающемся волокне.

ному определению состояния. По этой причине *гиббсов ансамбль представляет собой особый, нефизический предел больцмановского ансамбля*. Еще раз подчеркнем огромное различие между этим положением и утверждением о том, что необратимость обусловлена некоторой приближенностью описания обратимых процессов.

5.9. Кинетическая теория

До сих пор в этой главе рассматривался некоторый специальный класс неустойчивых динамических систем — К-струи. Особенно подробно был рассмотрен простой пример — преобразование пекаря. С помощью развитых при этом методов можно рассмотреть целый ряд более реалистических ситуаций, таких, как броуновское движение или модель Лоренца (см. разд. 5.1).

Этот раздел нам хотелось бы посвятить другому вопросу:

почему необратимость так характерна для больших систем, например газов или жидкостей, и в чем состоит природа неустойчивости в таких системах? Это далеко не простой вопрос. Модели жестких шаров типа лоренцевой не являются гамильтоновыми системами, поскольку потенциальная энергия двух жестких шаров либо равна нулю, либо обращается в бесконечность и тем самым не является непрерывной функцией расстояния. Поэтому здесь неприменимы гамильтоновы уравнения движения (2.56). Напротив, в реальных системах, содержащих много тел, взаимодействия достаточно гладки, а значит, имеется и гамильтониан с соответствующими кинетической и потенциальной энергиями. Как нам представляется, мы начинаем понимать причины того, что подобные большие системы могут быть близки к К-струям.

Вернемся к гл. 3. В разд. 3.8 мы уже подчеркивали важную роль резонансных условий. В рассматривавшихся там системах с малой размерностью на большинстве торов частоты оказывались несоизмеримыми (нерезонансные торы). В такой ситуации возникает квазипериодическое движение вокруг эллиптических точек. Однако имеется также и плотное распределение гиперболических точек, возникающих при разрушении резонансных торов. Здесь нам хотелось бы показать, что в пределе больших систем ситуация, по всей вероятности, резко изменяется, поскольку резонансные условия должны удовлетворяться почти всюду.

Для начала заметим, что в таких системах динамика отражает свободное движение частиц между последовательными столкновениями. Это взаимодействие лучше всего видно вблизи предельного случая свободных частиц при анализе решения уравнения Лиувилля методом теории возмущений с удержанием в уравнении лишь главных членов. Путем дополнительного предположения о малости радиуса корреляции отсюда получается так называемое *кинетическое уравнение*, описывающее релаксацию системы к равновесию. В это уравнение входит *оператор столкновений*, играющий центральную роль в современной кинетической теории. В главном порядке теории возмущений этот оператор содержит множитель вида

$$\left(\frac{2\pi}{L}\right)^3 \sum_l |V_l|^2 l^2 \delta(l \cdot (p_j - p_k)) \quad (5.32)$$

Здесь p_j и p_k — импульсы частиц j и k , δ — дельта-функция Дирака, а V_l — фурье-образ потенциала (V) взаимодействия частиц j и k :

$$V(|x_j - x_k|) = \left(\frac{2\pi}{L}\right)^3 \sum_l V_l e^{il(x_j - x_k)}, \quad l = \frac{2\pi}{L} n \quad (5.33)$$

причем \mathbf{n} — вектор с целочисленными компонентами. В такой записи предполагается, что граничные условия периодические, а потенциал V зависит лишь от расстояния между частицами $|\mathbf{x}_j - \mathbf{x}_k|$. Множитель $(2\pi/L)^3$, где L^3 — объем системы, обеспечивает переход от ряда к интегралу Фурье в предельном случае больших систем ($L \rightarrow \infty$).

Для применимости кинетической теории требуется, чтобы выражение в (5.32) существовало и было конечным во всем фазовом пространстве. Рассмотрим, к каким динамическим следствиям приводит это свойство. Благодаря присутствию дельта-функции в (5.32) имеется ненулевой вклад для таких \mathbf{l} (при $|\mathbf{l}| \neq 0$), для которых

$$\mathbf{l}(\mathbf{p}_j - \mathbf{p}_k) = 0 \quad (5.34)$$

Каков смысл этого соотношения? Из второго из уравнений (5.33) следует, что с точностью до постоянного множителя величина \mathbf{l} равна вектору с целочисленными компонентами. Кроме того, в соответствии с основами теории возмущений величины \mathbf{p}_j и \mathbf{p}_k связаны со свободным движением частиц и потому обратно пропорциональны времени, необходимому для того, чтобы частица пересекла расстояние порядка характерного размера системы. Иными словами, эти величины можно рассматривать как аналоги частот в представлении угол — действие. Таким образом, структура (5.34) напоминает резонансное условие, рассматривавшееся в гл. 3. Как мы видели в разд. 3.8, резонансы «опасны» в том смысле, что малые возмущения «стандартного» состояния интегрируемой системы (в данном случае такое состояние соответствует свободному движению частиц) могут привести к гиперболическим точкам и неустойчивому движению.

Поскольку \mathbf{l} принимает лишь дискретные значения, (5.34) будет удовлетворяться только при определенных \mathbf{p} . Таким образом, оператор столкновений обращается в нуль почти всюду в фазовом пространстве. Однако в предельном случае больших систем \mathbf{l} становится непрерывной переменной, а (5.32) — (5.33) переходят в интегралы Фурье. Поэтому можно ожидать, что резонансы возникнут почти всюду, и соответственно оператор столкновений существует и конечен. Таким образом, напрашивается вывод о том, что в системах с ненулевым оператором столкновений почти каждая точка фазового пространства является гиперболической, как в случае преобразования пекаря. Изложенные соображения снова позволяют проследить связь между необратимостью, присущей кинетической теории, и неустойчивостью движения при динамическом подходе к описанию системы.

5.10. Резонанс и взаимодействие света с веществом

В предыдущем разделе мы установили, что в кинетической теории резонансы играют фундаментальную роль. Оператор столкновений, лежащий в основе современной теории неравновесных процессов, можно определить лишь в том случае, если резонансные условия выполняются почти всюду в фазовом пространстве. Таким образом, резонансы ответственны за приближение к равновесию, а также за необратимость поведения многочастичных систем.

В настоящем разделе мы покажем, что резонансы играют ключевую роль и во взаимодействии электромагнитного излучения с веществом — одним из наиболее фундаментальных взаимодействий в природе. Однако в отличие от предыдущего раздела, где резонансы появляются автоматически с самого начала эволюции системы, здесь они будут создаваться постепенно, в результате процесса самоорганизации нового типа.

Рассмотрим гармонический осциллятор с собственной частотой ω . Соответствующий гамильтониан описывается равенством (2.6б). Допустим, что осциллятор заряжен и находится в электромагнитном поле. Как это обычно делается в теории электромагнетизма, будем описывать поле с помощью бесконечного набора гармонических осцилляторов и примем, что их собственные частоты распределены в узкой полосе около некоторой величины ω_e .

Вследствие взаимодействия с электромагнитным полем наш осциллятор будет испытывать силу, пропорциональную его заряду и зависящую от его скорости, а также от интенсивностей и ориентации электрического и магнитного полей. Эта сила вызовет движение с ускорением, даже если в начальном состоянии осциллятор покоился. Как известно из классической теории электромагнитных явлений, ускоренно движущаяся заряженная частица излучает и, следовательно, теряет энергию. Описывающая такое «излучательное затухание» теория не вполне удовлетворительна, так что более детальное описание такого излучения энергии веществом возможно лишь с помощью квантовой механики. Тем не менее в первом приближении представление о роли резонанса во взаимодействии света с веществом дает следующее уравнение:

$$\frac{d^2x}{dt^2} + \frac{\gamma}{m} \frac{dx}{dt} + \omega^2 x = qF_e \cos \omega_e t \quad (5.35)$$

где x — координата осциллятора, m и q — его масса и заряд соответственно, γ — константа затухания и F_e — амплитуда напряженности поля.

Решение этого уравнения с помощью стандартных методов можно представить в виде

$$x = \rho F_e \cos(\omega_e t + \theta) \quad (5.36)$$

где

$$\rho^2 = \frac{q^2}{(\omega^2 - \omega_e^2)^2 + \frac{\gamma^2}{m^2} \omega_e^2} \quad (5.37a)$$

$$\operatorname{tg} \theta = \frac{\gamma \omega_e}{m(\omega^2 - \omega_e^2)} \quad (5.37b)$$

В том случае, когда собственная частота осциллятора ω приближается к частоте поля ω_e , наблюдается очень большое усиление ответа. Поэтому в предельном случае слабого затухания γ осциллятор будет давать заметное излучение лишь при частоте поля. Это и есть явление резонанса.

Обсудим теперь этот вопрос с несколько более фундаментальной точки зрения. Мы рассматриваем полный гамильтониан заряженного осциллятора вместе с полевыми осцилляторами, учитывающий их взаимодействие. Начальное состояние соответствует движущемуся осциллятору и покоящимся полевым осцилляторам, или, что то же самое, полевым осцилляторам с нулевой энергией (это соответствует классическому вакууму — с учетом квантовомеханических поправок его энергия не равна нулю). Далее с помощью компьютера мы решаем полные (обратимые) уравнения движения, не делая феноменологического допущения о существовании константы затухания γ . Тогда такое численное моделирование показывает, что заряженный осциллятор начинает излучать с некоторой временной задержкой, которая не описывается уравнением (5.35). Очень интересно понять происхождение этой задержки. Мы только что видели, что процесс излучения сопряжен с наличием резонанса между материальным осциллятором и полевыми осцилляторами. Однако в начальном состоянии полевые осцилляторы отсутствуют. Поэтому в качестве первого шага материальный осциллятор должен создать полевые осцилляторы, или, иными словами, создать среду, обеспечивающую возможность излучения. При любых начальных условиях осциллятор начинает излучать через короткое время (порядка 10^{-14} — 10^{-18} с для атома водорода). Короче говоря, механизм излучения представляет собой необратимый «автокаталитический» процесс, осуществляющийся на границе между вакуумом и веществом.

Сравним такой тип поведения с необратимыми процессами в жидкостях. Там также при любых начальных условиях невозможно запретить установление теплового равновесия. Однако в случае жидкостей характерный временной масштаб опре-

деляется временем релаксации (в типичном случае порядка 10^{-9} с), существенно превышающим характерное время необратимых процессов, предваряющих излучение.

Существование релаксационных процессов на очень коротких временах, связанных со взаимодействием излучения с веществом, еще раз иллюстрирует универсальную роль необратимости в природе.

5.11. Заключительные замечания

Мы живем в плюралистическом мире. Простые динамические явления типа периодического движения Земли вокруг Солнца кажутся нам обратимыми во времени и детерминистическими. Однако мы наблюдаем также необратимые процессы типа химических реакций, а также стохастические процессы типа выбора между разветвившимися решениями.

Почему такой плюрализм оказывается возможным? Согласно классическим представлениям, необратимость и случайность на макроскопическом уровне рассматривались как своего рода артефакты, обусловленные сложностью коллективного поведения совокупности простых объектов. Почти три столетия вера в простоту мироздания на фундаментальном уровне являлась одной из движущих сил классической науки. Сегодня мы приходим к необходимости пересмотра наших представлений. Фундаментальный уровень оказывается непростым даже в классической физике, так долго служившей оплотом детерминизма и обратимости. Классическая точка зрения сформировалась при изучении систем, в которых периодическое движение и эллиптические точки играли важную роль. Теперь же мы видим, что имеется множество систем, в которых центральную роль играют гиперболические точки. В этой главе были затронуты два типа систем — К-струи и системы, изучаемые в современной кинетической теории, — в которых почти все точки фазового пространства являются гиперболическими. Поскольку в таких системах траектории не поддаются наблюдению, для описания динамики приходится использовать расширенный формализм, включающий стохастичность и необратимость. Такой подход, ведущий от групп к полугруппам, обусловлен не столько нашим незнанием, сколько самой структурой динамических систем.

В физике часто оказывается полезным провести различие между состояниями и законами. Информацию о состояниях мы можем получить лишь путем соответствующего препарирования системы и наблюдения над ней. Спецификация этих процессов является неотъемлемой частью любого объективного описания системы, что неоднократно подчеркивалось Бором и Розенфельдом. Такие препарирование и наблюдение могут привести лишь

к «окошку конечной ширины». Механизм наших измерений непременно должен опираться на одно из естественных взаимодействий между динамическими системами. Поэтому имеющееся в таком механизме «окошко конечной ширины» обусловлено ограниченной чувствительностью динамических систем.

Осознание того факта, что «окошко» имеет конечную ширину, приводит к новому соотношению между состояниями и законами. С классической точки зрения (включая квантовую механику) состояния симметричны во времени и развиваются по законам, которые также обладают такой временной симметрией. Теперь же нам приходится рассматривать состояния типа ансамблей Больцмана, введенных в разд. 5.8, которые характеризуются нарушением временной симметрии и развиваются также по законам с нарушенной временной симметрией.

СЛОЖНОЕ И ПЕРЕНОС ЗНАНИЙ

Для того чтобы понять происходящие вокруг нас события, необходимо строить различные модели и сопоставлять их с нашими наблюдениями. Для физика или химика это утверждение может выглядеть чем-то вроде трюизма, однако на самом деле его сфера простирается далеко за пределы физики или химии. Каждый данный момент в нашем мозгу происходит сканирование впечатлений, сопоставление наблюдений с уже сформировавшимися образами, что в конечном счете приводит к некоторому предварительному заключению. Одна из отличительных особенностей этой процедуры состоит в широком использовании *аналогий и прототипов*.

Цель данной главы состоит в том, чтобы показать, что физико-химические системы, приводящие к переходным явлениям, крупномасштабной упорядоченности, нарушению симметрии вдали от равновесия могут служить в качестве прототипов для понимания других систем со сложным поведением, в которых законы эволюции соответствующих переменных достаточно подробно неизвестны. Не менее важно и то обстоятельство, что во многих из этих систем сам выбор подходящих переменных может оказаться частью задачи, которую требуется решить.

Во всех рассматриваемых примерах анализ осуществляется в две стадии. Сначала проводятся определенные аналогии между наблюдениями и поведением «стандартной» физико-химической системы. Это позволяет установить тип модели, которая выглядит наиболее адекватной для представления рассматриваемой системы. Далее делается попытка выйти за пределы простых аналогий, чтобы установить *в рамках принятой модели* специфические особенности каждой задачи и затем включить их в описание системы. В конечном счете результаты такого анализа сопоставляются с опытом, и если качественное согласие получается, то они используются для дальнейших предсказаний. При этом можно надеяться, что в результате такого процесса должны появиться конкретные предложения о том, как можно справиться со сложностью.

6.1. Нелинейная динамика вдали от равновесия и моделирование сложного

На протяжении всей этой книги сложное рассматривалось как способность к переключению между различными типами поведения при изменении внешних условий. В свою очередь вытекающие отсюда гибкость и приспособляемость приводят к понятию о *выборе* между различными имеющимися возможностями. Было подчеркнуто, что выбор опосредуется динамикой флуктуаций, для чего требуется вмешательство двух их антагонистических проявлений — случайности на малых масштабах и упорядоченности на крупных масштабах. При этом случайность выступает в роли некоторого инновационного элемента, нужного для прощупывания пространства состояний, а упорядоченность позволяет системе поддерживать коллективный режим, охватывающий макроскопические пространственные области и временные интервалы.

Необходимой предпосылкой всех этих явлений служит нелинейная динамика, при определенных условиях приводящая к *неустойчивости движения* и *бифуркациям*. Поэтому первый шаг в моделировании сложного поведения состоит в установлении нелинейного характера соответствующей динамики и в идентификации набора переменных, способных продемонстрировать неустойчивости и бифуркации. Эта процедура существенно облегчается благодаря результатам теории динамических систем, изложенным в гл. 3 (см., в частности, разд. 3.6). В соответствии с этими результатами сущность, качественные особенности поведения часто содержатся в простом универсальном динамическом уравнении, записанном в *нормальной форме*. Этому уравнению удовлетворяет ограниченный набор коллективных переменных — *параметров порядка*. Поэтому потенциально системы со многими переменными, к которым относится большинство встречающихся на практике систем, вблизи точки бифуркации могут допускать упрощенное описание. Например, если по тем или иным причинам нам известно, что в рассматриваемой системе могут совершаться переходы к двум одновременно устойчивым стационарным состояниям без нарушения пространственной симметрии, то в таком случае появляется возможность использования кубического уравнения (3.14), как описания в нулевом приближении. С другой стороны, если бы мы имели дело с переходом к незатухающим колебаниям, то в качестве разумной исходной точки мы могли бы рассматривать систему уравнений (3.16). Наконец, при рассмотрении хаотической динамики следовало бы исходить из того, что в непрерывном описании должно участвовать по меньшей мере три переменных и что в случае дискретного описания, представляемого, напри-

мер, отображением Пуанкаре, можно пытаться отыскать итеративную динамику, аналогичную представленной на рис. 3.27 и уравнениями (3.31).

Разумеется, пользоваться этими соображениями (а их полезность не следует недооценивать) нужно достаточно осторожно — в противном случае они могут заслонить то, что является важной частью рассматриваемой задачи и должно оставаться таковой, а именно специфику рассматриваемой задачи. Поэтому на каждой стадии анализа конструктор модели должен иметь возможность увязывать параметры порядка с исходными переменными, придавать механизмам протекающих в системе процессов нормальную форму, а также удовлетворять условиям, которые могут накладываться уже самой природой рассматриваемой задачи. В физико-химических системах в роли таких условий могут выступать положительность производства энтропии, температуры или концентрации. Аналогичным образом в популяционных задачах таким ограничением будет служить требование неотрицательности числа особей того или иного вида. В геофизических задачах будут выбраны такие модели, в которых пространственные и временные масштабы используемых данных адекватно воспроизводятся правой частью кинетических уравнений, включая значения применяемых параметров.

Имея в виду эти моменты, мы сейчас приступим к изложению ряда примеров успешного применения моделирования с использованием в качестве соответствующего прототипа нелинейных динамических систем в условиях, далеких от равновесия. Будут последовательно рассмотрены наука о материалах, биология, геофизика и социальное поведение популяций животных и человека.

6.2. Наука о материалах

Одной из областей, в которых быстро возрастает влияние достижений нелинейной динамики сильно неравновесных систем, является наука о материалах. Создание новых материалов представляется решающим элементом развития новой технологии. Примечательно, что во многих случаях соответствующие процессы протекают в условиях значительного удаления от равновесия.

Например, в настоящее время достигнуто понимание того, что транспортные явления, происходящие в жидкой фазе во время роста кристаллов или отвердевания сплавов, оказывают глубокое влияние на структуру и свойства твердой фазы. Как и во многих проблемах с участием жидкостей в неравновесных условиях, главную роль здесь играет конвекция, поскольку она влияет на температуру и состав на границе раздела между жид-

кой и твердой фазами. В свою очередь конвективное движение жидкости определяется борьбой между внешними условиями (вынужденная конвекция) и внутренними процессами (естественная конвекция). Последние обусловлены перепадами концентрации или температуры (см., например, разд. 1.3), но могут быть также вызваны и поверхностным натяжением (см. разд. 1.5). Следовательно, рост кристаллов и отвердевание сплавов — сложные процессы, включающие массо- и теплоперенос, течение жидкости, химические реакции и фазовые переходы. Как отмечалось в разд. 1.3 и в гл. 3, при определенных условиях конвекция может стать нерегулярной. Такое хаотическое поведение индуцирует случайные изменения температуры и концентраций, приводя к дефектам кристаллической структуры.

Разумеется, главную роль в этих явлениях играют свойства границы раздела между жидкостью и твердым телом. В этом контексте ярким примером неравновесия является наблюдение, согласно которому во время роста кристалла существует однородная фаза между твердой и жидкой областями, причем свойства этой фазы отличны от свойств и расплава, и твердого тела. Эксперименты по рассеянию света на этом слое, размеры которого значительно превышают постоянную решетки, позволили установить, что коэффициент температуропроводности в нем по порядку величины составляет 10^{-8} см²/с, в то время как его типичное значение для воды равно 10^{-3} см²/с. Иными словами, температурные флуктуации затухают здесь чрезвычайно медленно, и это можно рассматривать как указание на критическое состояние всего слоя, подобное тому, что имеет место вблизи точки бифуркации. По-видимому, этот переход опосредован потоком тепла через границу раздела, обусловленным продвижением фронта затвердевания. В результате в промежуточном слое в буквальном смысле слова поддерживается *неравновесная фаза вещества* (называемая также мезофазой).

Другой класс неравновесных явлений возникает при наблюдении излучения мощного лазера или источника частиц, позволяющего за короткий интервал времени инжектировать большое количество энергии в поверхностный слой различных материалов. Возникающие при этом явления плавления и вторичного затвердевания характеризуются высокими температурными градиентами (от 10^6 до 10^8 К/см) и скоростями движения границ раздела (1 м/с). Среди различных применений этих предельно неравновесных условий особенно многообещающим выглядит синтез новых материалов для электронной промышленности.

Возможно, самым ярким примером вмешательства неравновесности в область науки о материалах служит исследование разрушения материалов. Такие явления, как пластичность, де-

формация или усталость, обычно опосредуются дефектами, которые создаются напряжениями и в дальнейшем усиливаются и мигрируют, образуя структуры, предшествующие излому. Для достижения прогресса в этой области нужно понять механизм образования дефектов и дислокаций и их взаимодействия в системе, находящейся при определенных внешних условиях. Моделирование соответствующей нелинейной динамики должно привести к неустойчивостям, создающим переходы к множественным стационарным состояниям и к пространственным структурам, образуемым концентрацией дефектов. Остановимся вкратце на важнейших моментах такого подхода.

Прежде всего нужно установить набор переменных и параметров. Кроме деформации ϵ , которая часто встречается в технических приложениях, по-видимому, определенную роль играет также температура T . В самом деле, во время испытания на растяжение в упругой области образец обычно охлаждается. Снижение температуры пропорционально величине деформации (для простоты считающейся однородной по объему образца) и зависит от так называемого *параметра Грюнайзена* γ . Этот параметр отражает отклонение межмолекулярных сил от закона, характерного для гармонических колебаний, и тем самым служит мерой термомеханического сопряжения. При комнатной температуре γ находится в диапазоне от 1,5 до 2,5 для широкого набора материалов. Однако этот параметр очень сильно зависит от относительной деформации, приближаясь практически к нулю при критических напряжениях.

С другой стороны, только что отмеченному понижению температуры противостоят эффекты нагревания, обусловленные вязкостью материала. С возрастанием приложенного напряжения изначально статические дислокации и другие дефекты становятся все более и более подвижными, и их движение в образце приводит к увеличению тепловой энергии. В свою очередь это все больше увеличивает численность дефектов, что очень напоминает течение экзотермической химической реакции (см. разд. 1.4). Таким образом, здесь имеется механизм дестабилизации с положительной обратной связью, противодействующий стабилизирующему влиянию охлаждения, вызванного деформацией. Оба механизма уравниваются в некоторой критической точке, которую можно рассматривать как границу между упругим и пластическим режимами деформации. Такая точка баланса представлена на рис. 6.1. Можно ожидать, что пороговое значение σ_0 напряжения, играющего роль внешнего фактора, отклоняющего систему от равновесия, в этом критическом положении определяет точку бифуркации нового состояния.

Для более количественного описания затронутых выше явлений необходимо привлечь динамические уравнения для клю-

чевых переменных ϵ и T , в которых можно было бы явным образом проследить за влиянием параметров σ и γ . Не вдаваясь особо в подробности, отметим лишь, что при некоторых разумных предположениях эти два уравнения имеют нелинейную связь в виде члена

$$-\frac{1}{2} C_v \left(\frac{\partial \gamma}{\partial \epsilon} \right) \left(\frac{T - T_0}{T_0} \right) \epsilon^2 \quad (6.1)$$

Здесь C_v — удельная теплоемкость при постоянном объеме и T_0 — некая стандартная температура. Вследствие отмечавшейся выше зависимости параметра Грюнайна от деформации этот член положителен, благодаря чему возникает механизм

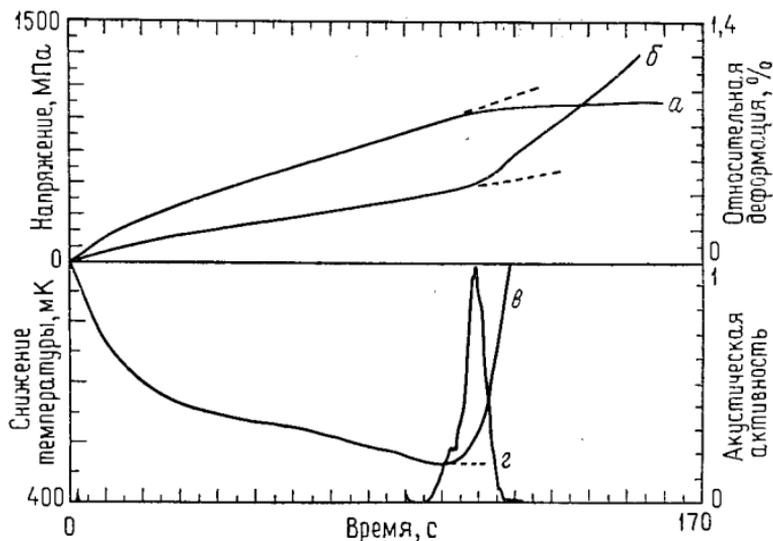


Рис. 6.1. Напряжение (a), деформация (b), температура ($в$) и акустическая излучательная способность ($г$) стального образца в зависимости от времени. Следует отметить резкое возрастание температуры и акустической активности в критической области.

усиления флуктуаций, приводящий к неустойчивости. Следуя результатам разд. 3.6, вблизи точки бифуркации σ_0 можно еще больше упростить описание, исключив все переменные, за исключением параметра порядка. Последний, являясь соответствующей линейной комбинацией ϵ и $T - T_0$, удовлетворяет динамическому уравнению, подобному (3.14) или первому из уравнений (3.16):

$$\frac{dx}{dt} = (\sigma - \sigma_0) x - bx^3 \quad (6.2)$$

В зависимости от условий бифуркация может привести либо к новому стационарному состоянию, как в уравнении (3.14),

либо к колебаниям в режиме предельного цикла, как в случае уравнения (3.16).

Для того чтобы развить более микроскопический подход к изложенной выше проблеме, необходимо рассмотреть соотношения между различными элементарными процессами, такими, как создание, аннигиляция, блокирование или диффузия дислокаций. Это приводит к диффузионным моделям с учетом химических реакций, очень похожих на те, что использовались при изучении химических неустойчивостей (см., например, разд. 3.11). Будучи подвержены действию внешних анизотропных факторов, такие модели предсказывают некоторую иерархию переходов с нарушением симметрии, приводящих к дислокационным картинам, весьма напоминающим экспериментальные (см. рис. 6.2).

Итак, оказывается, что столь важные и широко распространенные механические явления, как пластичность и текучесть, невозможно исследовать на чисто механической основе! Вместо этого их следует рассматривать как часть общей проблематики нелинейных динамических систем, работающих вдали от равновесия. Нам представляется, что уже само осознание этого обстоятельства есть существенное продвижение в области науки о материалах.

6.3. Пороговые явления в клеточной динамике

Прыжок из области материалов под нагрузкой к живым клеткам может показаться довольно рискованным предприятием. Однако к этому времени наш читатель уже должен был понять, что такие скачки вообще характерны для моделирования по самой его сути. Мы только что видели, что пластичность и текучесть можно рассматривать как пороговые явления, опосредованные конкурентным взаимодействием популяции дефектов и объема данного материала. При возникновении текучести мы должны, следовательно, заключить, что данный материал не смог справиться с деятельностью дефектов — нарушителей спокойствия, в результате чего их популяция достигла катастрофических размеров. Цель настоящего раздела — показать, что весьма похожая конкуренция наблюдается и в биологических явлениях, например в таких жизненно важных, как рост опухолей. В этом случае исходно малая популяция «зачинщиков беспорядка», например определенные клетки некоторой ткани, по тем или иным причинам, выясняемым молекулярной биологией, потеряла свои физиологические функции и стала злокачественной. В дальнейшем они норовят захватить весь организм за счет быстрой пролиферации, однако организм старается оказать им противодействие, посылая на «поле боя» спе-

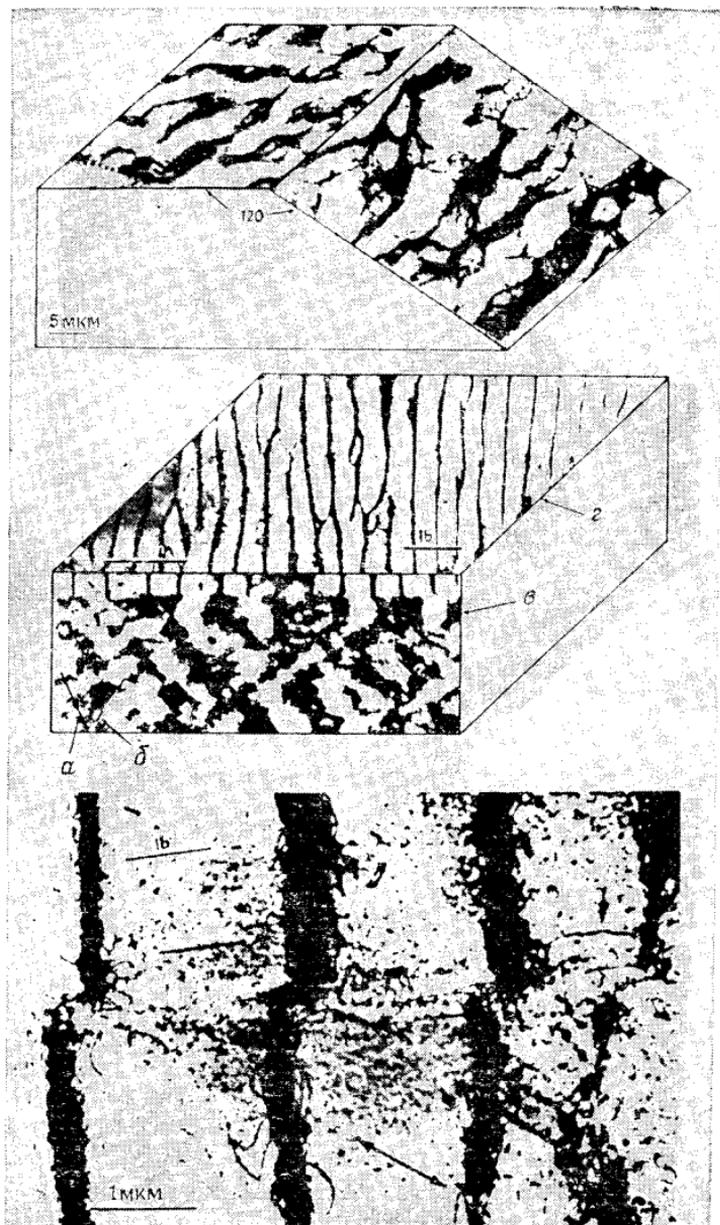
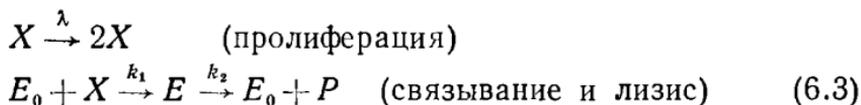


Рис. 6.2. Трехмерная структура дислокации в периодически растягиваемом монокристалле меди. *a* — критическая; *b* — сопряженная; *c* — плоскость поперечного скольжения; *d* — первичная плоскость скольжения.

циализированные клетки-убийцы. Дальнейшая судьба организма решится именно в результате возникающей конкуренции между злокачественными клетками и клетками-убийцами — то ли «зачинщики беспорядка» будут отброшены, то ли, наоборот, они победят. Анализ этой конкуренции позволяет прийти к выводу, на наш взгляд имеющему исключительную важность, а именно: отбрасывание «зачинщиков» или их победу можно рассматривать как пороговое явление.

Прежде чем приступить к более количественному описанию конкуренции, остановимся вкратце на основных свойствах клеток-убийц. Иммунная система позвоночных, поставляющая организму средства противодействия чужеродным патогенным веществам, поступающим в организм, производит в костном мозге недифференцированные исходные зародышевые клетки. В результате дальнейшей дифференциации эти клетки превращаются в В- или Т-лимфоциты, поступающие в организм. Встречаясь с чужеродным веществом, называемым обычно *антигеном*, В-клетки подвергаются дальнейшей дифференциации в соответствующие большие клетки, которые пролиферируют и секретируют способные нейтрализовать антиген химические вещества, называемые *антителами*. С другой стороны, Т-клетки, после дальнейшей дифференциации, происходящей в тимусе, регулируют действие В-клеток, усиливая или подавляя их деятельность. Кроме того, они участвуют в иммунных ответах, вызванных непосредственно клетками, в особенности раковыми или чужеродными, как это бывает, например, при пересадке тканей или органов. Другие клеточные разновидности иммунной системы типа макрофагов имеют сходную функцию, называемую цитотоксической активностью. В настоящем разделе мы будем иметь дело прежде всего именно с клеточно-опосредованным ответом.

Пусть X — плотность популяции пролиферирующих клеток-мишеней (например, злокачественной опухоли), а E_0 — плотность свободных цитотоксических клеток. Согласно экспериментальным данным, клетки E_0 распознают клетки X , фиксируют их в виде комплекса $E \equiv E_0X$ и в дальнейшем осуществляют их лизис при диссоциации комплекса E . Эту последовательность стадий можно записать в виде



Примем, что во время рассматриваемого процесса конкуренции между E_0 и X состояние иммунной системы можно считать стационарным практически во всех отношениях. Это означает, что полная плотность популяции свободных и связанных иммунных

клеток ($E_t = E_0 + E$) остается постоянной во времени. Далее примем упрощенное описание, согласно которому кинетика обеих стадий во втором из уравнений (6.3) может считаться некооперативной (без обратных связей), по крайней мере в области рассматриваемых значений параметров. Подчеркнем, однако, что от этого предположения можно легко отказаться без качественного изменения результатов.

При перечисленных условиях можно написать следующие уравнения баланса:

$$\begin{aligned}\frac{dX}{dt} &= \lambda X (1 - (X/N)) - k_1 E_0 X \\ \frac{dE_0}{dt} &= -k_1 E_0 X + k_2 E\end{aligned}$$

В первом из этих уравнений множитель $(1 - X/N)$ выражает существование верхнего предела N для X внутри элемента объема, в котором происходит конкурентная борьба. Поскольку лизис должен протекать гораздо быстрее других стадий, для второго уравнения можно воспользоваться квазистационарным приближением, рассматривавшимся в разд. 3.6. Поэтому, полагая $dE_0/dt \sim 0$ и используя условие сохранения для клеток E -типа, можно выразить E_0 через X и получить первое уравнение в виде

$$\frac{dX}{dt} = \left(1 - \frac{X}{N}\right) \lambda X - \frac{k_1 E_t X}{1 + \frac{k_1}{k_2} X} \quad (6.4)$$

Это уравнение можно переписать в безразмерном виде:

$$\frac{dx}{d\tau} = (1 - \theta x) x - \frac{\beta x}{1 + x} \quad (6.5)$$

где

$$\begin{aligned}\beta &= \frac{k_1 E_t}{\lambda}, & \theta &= \frac{k_2}{k_1 N} \\ \tau &= \lambda t, & x &= \frac{k_1}{k_2} X\end{aligned} \quad (6.6)$$

Очевидно, уравнение (6.5) всегда допускает стационарное решение $x=0$. Линейный анализ устойчивости показывает, что (см. приложение I) независимо от значения θ это решение неустойчиво при $\beta < 1$ и устойчиво при $\beta > 1$. Существование ненулевого стационарного решения для x соответствует опухолевому состоянию ткани. В зависимости от значения θ здесь могут возникнуть две качественно различные ситуации. Эти случаи представлены на рис. 6.3. При $\theta > 1$ опухоль может существовать лишь при условии $\beta < 1$. Поэтому условие полного по-

давления опухоли сводится к тому, что в процессе иммунного ответа величина β должна превысить единицу. При $\theta < 1$ переход между опухолевым и безопухолевым состояниями связан с явлением бистабильности, свойства которого напоминают неравновесные переходы, обсуждавшиеся в этой книге. В области $1 < \beta < \beta_c = (1 + \theta)^2 / 4\theta$ нулевое (нормальное) и единичное (опухолевое) состояния устойчивы одновременно. Можно ожидать, что переход между этими состояниями связан с процессами нуклеации. Если начальное состояние соответствует опухоли, то ее подавление при возрастании β возможно лишь при условии $\beta > \beta_c$.

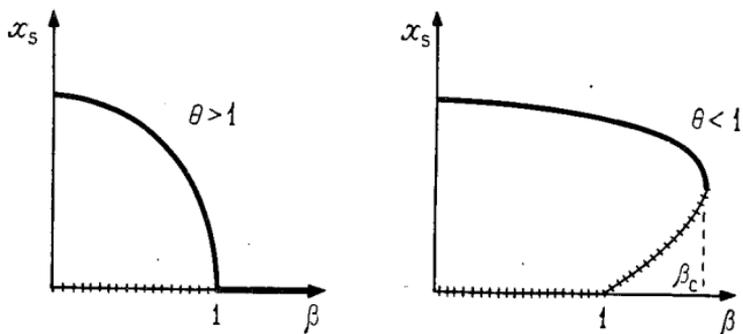


Рис. 6.3. Стационарная концентрация опухолевых клеток в зависимости от параметра β , характеризующего активность иммунной системы. На правом рисунке показана возможность сосуществования нормального ($x_s = 0$) и злокачественного ($x_s > 0$) состояний, разделенных определенным порогом.

Анализируя литературные данные по различным параметрам, можно прийти к выводу о реальной возможности обоих типов стационарных состояний. Вспоминая с помощью первого из соотношений (6.6), что параметр β по существу описывает эффективность иммунной системы, легко понять, что для данного состояния иммунной системы в области $1 < \beta < \beta_c$, $\theta < 1$, в зависимости от предыстории ткани возможна как нормальная, так и патологическая ситуация.

Можно ожидать, что принятый в этом разделе динамический подход окажется важным для понимания множества других биологических явлений на надмолекулярном уровне, например типа агрегации клеток крови при артериосклерозе.

6.4. Моделирование климатической изменчивости

В дальнейшем в этой главе мы еще вернемся к популяциям взаимодействующих особей в связи с моделированием социального поведения. Однако прежде чем перейти к этой проблеме,

полезно изложить некоторые представления о моделировании климатической изменчивости — эта проблема не только имеет большое научное значение, но и чрезвычайно важна в плане экономики и социологии. И все же, несмотря на столь глобальный характер, климатическая изменчивость остается в основном проблемой, которая должна поддаваться физико-химическому описанию, что в значительной мере роднит ее с задачами, уже обсуждавшимися в этой главе.

Основные факты, свидетельствующие об изменчивости климатической системы, уже излагались в разд. 1.7. Арена, на которой будет разыгрываться основное действие настоящего раздела — моделирование климатической изменчивости, представлена на рис. 6.4.

Правый рисунок соответствует Северному полушарию во время пика последнего общего оледенения, т. е. около 18 000 лет назад. В то время слой льда толщиной в несколько километров покрывал значительную часть Северной Америки вплоть до средней части Соединенных Штатов, а также значительную часть Северной и Западной Европы вплоть до широт Парижа и Берлина. На левом рисунке показана современная ситуация. Видно, что континентальный лед сосредоточен главным образом в Гренландии. Далее, относительно распространения континентального льда известно, что современная ситуация установилась уже около 10 000 лет назад. Иными словами, мы вынуждены сделать вывод, что такая физическая система, как планета Земля, в течение нескольких тысячелетий (т. е. за короткое в геологическом масштабе время) совершила переход между двумя сильно различающимися состояниями, причем различия в этих состояниях наблюдаются на масштабах, сравнимых с размерами самой Земли.

Из этих замечаний становится ясно, что любая разумная модель климатической системы должна учитывать возможность таких крупномасштабных переходов. Какие же возможности имеются в нашем распоряжении для построения адекватной модели такого рода?

Система Земля — атмосфера — криосфера является физической и как таковая удовлетворяет уравнениям баланса энергии, массы и импульса. Энергетический баланс описывает, что происходит в данной системе с поступающим солнечным излучением — какая часть его поглощается и какая часть снова излучается, как эти части распределяются по широтам и долготам и каковы их соотношения между поверхностями материков и океанов. С другой стороны, при анализе баланса массы важнее всего учесть баланс континентального и морского льда, а также некоторые небольшие, но важные компоненты атмосферы типа диоксида углерода и водяных паров. Наконец, ба-

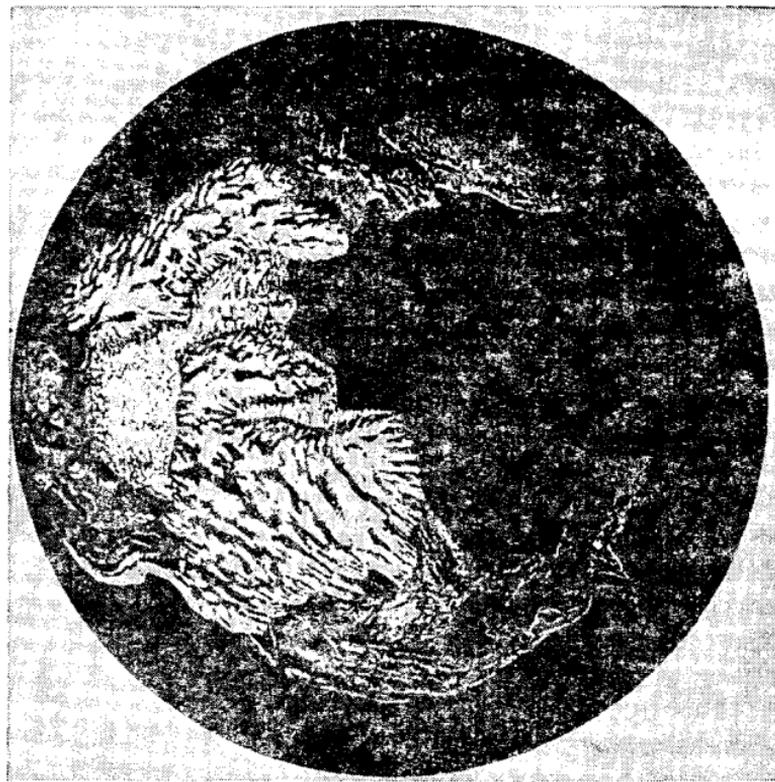


Рис. 6.4. Схематическое представление Северного полушария на пике прошлого оледенения (справа) и в современных условиях (слева).

ланс импульса определяет атмосферную циркуляцию, которая в конечном счете устанавливает локальную температуру и картину распределения осадков. Проблема состоит в том, что все эти виды баланса тесно переплетаются между собой. Кроме того, недостаточно хорошо известны точные феноменологические законы, выражающие различные потоки в такой системе через ограниченный набор макроскопических переменных (см. разд. 2.2). По сути дела, один из немногих неоспоримых фактов состоит в существовании атмосферной и океанической турбулентности. Если отсюда что и следует, так это то, что (см. разд. 3.10) климат должен обладать такой же внутренней непредсказуемостью, как и системы с хаотической динамикой!

В связи с этим можно высказать следующие два положения. С одной стороны, ясно, что полный набор уравнений баланса, в которых отдельные потоки параметризованы наилучшим образом в соответствии с имеющимися данными, должен решаться численно с помощью компьютера. Эта тенденция привела к *общим циркуляционным моделям* и стимулировала развитие сверхбыстрых компьютеров, способных перерабатывать такое огромное количество входных данных.

С другой стороны, можно возразить, что именно благодаря чрезвычайной сложности климатической системы необходим *качественный* анализ, ставящий целью выделение долговременных тенденций и идентификацию ключевых параметров. Читателю этой книги такой подход должен показаться совершенно естественным способом обращения со сложными явлениями! В конце концов в значительной мере своим прогрессом в понимании явлений в лабораторном масштабе, таких, как гидродинамические и химические неустойчивости, проследить за которыми гораздо легче, чем за глобальным климатом, мы обязаны широкому использованию понятий и методов теории динамических систем. По этой причине в оставшейся части настоящего раздела мы приведем некоторые примеры на тему о том, какую информацию и какое понимание может дать качественный анализ простых математических моделей климата. Мы сосредоточимся на проблемах оледенения в четвертичном периоде и, в частности, на механизмах, позволяющих климатической системе улавливать и усиливать слабые изменения в поступающей солнечной энергии. Как мы видели в разд. 1.7, несмотря на свою малость, эти изменения самым впечатляющим образом коррелируют с характерными временными масштабами ледниковых периодов.

Рассмотрим теперь в самом общем виде поступающее на Землю солнечное излучение и возвращаемое в космос инфракрасное излучение, считая, что наша планета представляет собой объект с нулевой пространственной размерностью. В таком подходе имеется лишь одна существенная переменная состоя-

ния — температура T , описываемая уравнением теплового баланса:

$$\frac{dT}{dt} = \frac{1}{C} [Q(1-a(T)) - \epsilon_{\text{в}}\sigma T^4] \quad (6.7)$$

Здесь C — теплоемкость системы, Q — солнечная постоянная, и a — альbedo, представляющее собой долю солнечного излучения, отражаемого в космос. Температурная зависимость этой величины учитывает обратную связь между альbedo и поверхностью, описанную в конце разд. 1.8. Последний член в (6.7) описывает излучаемое тепло в виде модифицированного закона Стефана — Больцмана, которому подчиняется излучение черного тела. Величина σ — постоянная Стефана, $\epsilon_{\text{в}}$ — эмиссионная способность, учитывающая отличие Земли от черного тела.

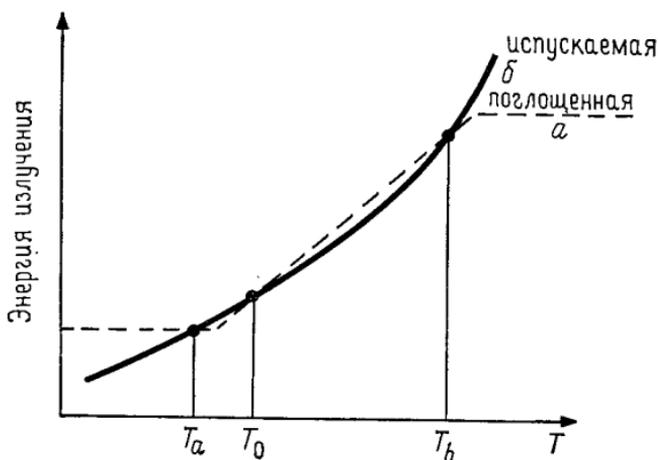


Рис. 6.5. Глобальный энергетический баланс на Земле. При соответствующей параметризации функции альbedo $a(T)$ кривые a и β описывают температурную зависимость поглощаемой доли падающей солнечной энергии и инфракрасного излучения Земли в космос. Эти зависимости допускают три стационарных режима, два из которых (T_a , T_b) устойчивы и соответствуют ледниковому климату и климату, сходному с современным.

При анализе уравнения (6.7), графически представленном на рис. 6.5, выясняется, что система допускает до трех стационарных состояний. Два из них, T_b и T_a , являются устойчивыми — они соответствуют современному климату и холодному климату, напоминающему глобальное оледенение. Третье состояние, T_0 , неустойчиво — оно разделяет два отмеченных выше устойчивых режима.

Как мы видели в разд. 4.5, эволюцию системы с одной переменной всегда можно представить как движение «частицы» в

потенциальном поле U , которое в данном случае определено соотношением [см. (4.30)]

$$U(T) = -\frac{1}{c} \int dT [Q(1 - a(T)) - \epsilon_{\text{вс}} \sigma T^4]$$

Величина $U(T)$ называется *климатическим потенциалом*. В случае, представленном на рис. 6.5, этот потенциал характеризуется наличием бистабильности, как и на рис. 4.7, б.

Как и другие физические системы, климатическая система непрерывно подвергается воздействию статистических *флуктуаций* — случайным отклонениям от детерминистического поведения, довольно детально обсуждавшимся в гл. 4. Поскольку нам недостает информации, на основе которой можно было бы написать уравнение для соответствующего распределения вероятностей, мы будем относить эффект флуктуаций на счет некоторой *случайной силы* $F(t)$ в соответствии с тем, что говорилось в конце разд. 4.2. Теперь уравнение энергетического баланса (6.7) становится стохастическим дифференциальным уравнением вида

$$\frac{dT}{dt} = -\frac{\partial U(T)}{\partial T} + F(t) \quad (6.8)$$

Такое расширенное описание вносит новый важный элемент — между различными состояниями теперь осуществляется *связь* с помощью флуктуаций. Иначе говоря, каким бы ни было начальное состояние системы, рано или поздно она достигнет любого другого состояния. Это справедливо и для устойчивых состояний T_a и T_b в качестве исходных; при этом, будучи с детерминистической точки зрения устойчивыми, они становятся *метастабильными состояниями*. Как обсуждалось в разд. 4.5, временной масштаб этого явления определяется двумя факторами — *высотой потенциального барьера*

$$\Delta U_{a,b} = U(T_0) - U(T_{a,b})$$

и интенсивностью флуктуаций, обычно характеризуемой *дисперсией* q^2 функции $F(t)$. Оценка этого времени по порядку величины, уже приводившаяся в разд. 4.5 [уравнение (4.31)], дает

$$\tau_{a,b} \sim \exp \frac{\Delta U_{a,b}}{q^2} \quad (6.9)$$

где $\tau_{a,b}$ — среднее время перехода из состояния T_a в состояние T_b через неустойчивое состояние T_0 . Если эти флуктуации достаточно малы, то, очевидно, при любых разумных высотах потенциального барьера это время будет весьма велико. Получаемые при этом типичные значения оказываются порядка

10^4 — 10^5 лет, что попадает в диапазон характерных времен ледников в четвертичном периоде. И все же утверждать, что уравнение (6.9) дает объяснение ледниковым циклам, нельзя, поскольку переход между двумя климатическими состояниями T_a и T_b остается процессом случайным, не имеющим четкой периодичности. Именно здесь влияние слабых внешних силовых факторов становится решающим.

Повторим изложенный только что анализ, прибавив к солнечной постоянной Q малый член вида $Q\epsilon \sin \omega t$, период которого $\Delta = 2\pi/\omega$ представляет влияние изменений, обуславливающих колебания средней ежегодной величины солнечной энергии, поглощаемой Землей, наподобие одного из орбитальных возмущений, обсуждавшихся в разд. 1.7. Как оказывается, если Δ гораздо меньше переходного времени $\tau_{a,b}$, то внешний силовой фактор практически не оказывает влияния. Однако если Δ сравнимо с $\tau_{a,b}$, то ответ системы на внешнее силовое воздействие усиливается чрезвычайно. По существу, это воздействие приводит к снижению потенциального барьера и тем самым к облегчению перехода между климатическими состояниями. Более того, вследствие периодического характера такого сигнала система оказывается вынужденной совершать ритмические переходы через пониженный таким образом барьер между двумя разными климатами. Следовательно, начинает просматриваться качественный механизм наступления ледниковых периодов, в котором важную роль играют как внутренние процессы, так и внешние факторы. Обсуждавшаяся выше простая модель позволяет также выявить параметры, которые с наибольшей вероятностью влияют на поведение системы.

Естественно, что ввиду схематичности рассматриваемой модели требуется также разработать более тонкое описание, которое, поддаваясь аналитическому исследованию, содержало бы также некоторые дополнительные ключевые переменные. Особо интересный класс таких моделей образуется при учете взаимодействия между средней температурой поверхности и континентальным или морским льдом. При вполне допустимых значениях параметров эти модели предсказывают периодические решения типа предельного цикла, причем период сопоставим с характерными временами оледенений. Однако при включении в описание флуктуаций оказывается, что по истечении достаточно длительного времени в колебательном поведении стираются все следы согласованности, что обусловлено слабой устойчивостью фазы колебаний рассматриваемого осциллятора. Сопряжение с периодическими внешними воздействиями позволяет стабилизировать фазу и тем самым снова зафиксировать разность фаз между колебаниями климатической системы и внешних воздействий.

Дальнейший прогресс в захватывающей области климатической динамики должен уменьшить расхождение между простыми качественными моделями и детальным количественным описанием. Полезным представляется также сопоставление зарегистрированных климатических изменений с теорией динамических систем — можно ожидать, что это приведет к лучшему пониманию климатической системы, что в свою очередь позволит выдвинуть новое описание сложности окружающего нас мира в крупных масштабах. Некоторые попытки в этом направлении представлены в приложении IV.

6.5. Вероятностное поведение и адаптивные стратегии у общественных насекомых

Обратимся теперь к еще одной грани сложного. Нас будут интересовать явления, в которых переменными состояниями уже не будут такие традиционные для физико-химических наук величины, как температура или химические концентрации. Здесь в игру вступают новые элементы, для описания которых нужен язык, в котором ключевыми словами являются стратегия, предвидение, символы и ритуализация. Однако и в этой области основным отличительным признаком сложного будут переходы между различными типами поведения. Это позволит пользоваться при анализе понятиями и методами теории нелинейных динамических систем, функционирующих вдали от равновесного состояния. Начать наше рассмотрение удобно с некоторых аспектов поведения общественных насекомых.

Общественные насекомые (пчелы, муравьи, термиты) достигли огромных успехов в процессе биологической эволюции. Ими занято большинство известных ниш, особенно в тропиках, где их биомасса может достигать нескольких центнеров на один акр (0,4 га).

Согласно обычным представлениям, постройка гнезда, поиски пищи и другие виды коллективной деятельности служат прототипами детерминистического мира, в котором отдельные насекомые являются маленькими, но надежными автоматами, подчиняющимися строго установленным генетическим программам. В настоящее время эта картина абсолютной жесткости постепенно стирается и возникает новая парадигма, в которой наряду с пластичностью индивидуального поведения важную роль начинают играть случайные элементы, присущие внешней среде.

Наиболее поразительным в сообществе насекомых представляется наличие двух масштабов: один на уровне индивидуума, характеризующегося исключительно высокой степенью случайности поведения, другой на уровне всего сообщества в целом. Во втором случае, несмотря на невысокую эффективность и не-

предсказуемость поведения отдельных особей, характерные для данного вида согласованные картины возникают в масштабе колонии в целом. Рассмотрим, какая же связь между этими двумя аспектами обеспечивает организацию всего сообщества. Возьмем, например, поиск пищи у муравьев и сравним эту деятельность в двух разных ситуациях: единственный предсказуемый источник пищи типа колонии более мелких насекомых (скажем, тлей) с продолжительностью жизни порядка нескольких месяцев и некоторый неожиданный источник типа мертвой птицы или другой колонии вида-жертвы, которая вдруг становится доступной.

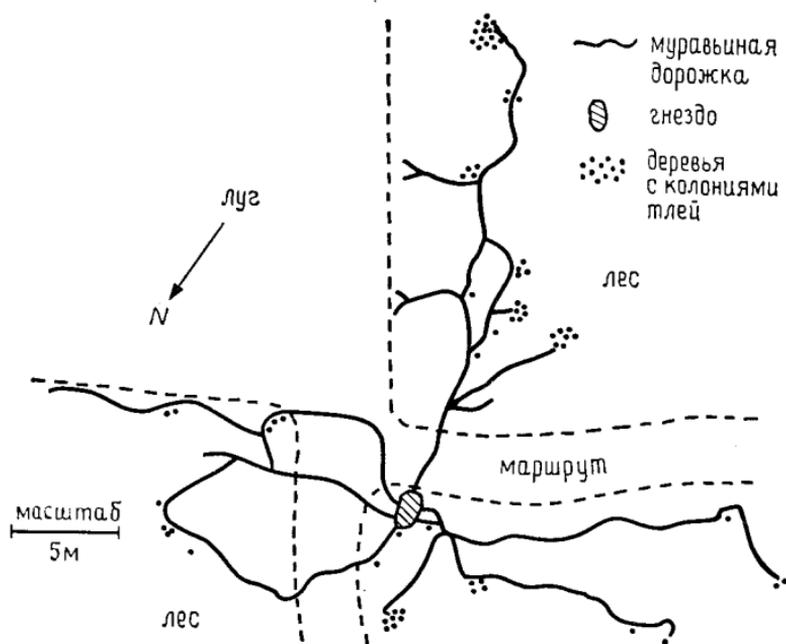


Рис. 6.6. Схематическое представление гнезда и сети дорожек сообщества *Lasius fuliginosus*.

В первом случае наблюдается образование стабильных путей от гнезда до соседней колонии, по каждой из которых ходят свои специальные пользователи (рис. 6.6). Более того, вне этих основных путей находится лишь малое число муравьев. Очевидно, в таких условиях сообществу выгоднее образовывать постоянные устойчивые структуры с низким уровнем шума.

Напротив, в случае непредсказуемой внешней среды образование постоянных структур может повредить пластичности колонии и даже привести к нарушению оптимального режима жизни. Поэтому в случае такой внешней среды возможная реакция состоит в том, чтобы поддерживать высокий уровень ин-

Таблица 6.1. Сравнение точности вербовки у двух видов муравьев^a

| | <i>Tetramorium impurum</i> | <i>Tapinoma erraticum</i> |
|--|----------------------------|---------------------------|
| Длина (в %) отдельных дорожек, фактически проходимых рекрутами | 17(40) | 67,7(47) |
| Процент рекрутов, достигающих источника пищи | Одиночки 8,9(45) | 73,6(216) |
| | Группы 60(10) | |
| | Всего 18,2(55) | 73,6(216) |

^a Числа в скобках указывают количество фактически наблюдавшихся муравьев.

тенсивности исследований, а также способность быстро создавать временные структуры, позволяющие получать выгоду из любых благоприятных возможностей, которые только могут возникнуть. Иными словами, в организации насекомых представляется, что случайность имеет определенную ценность с точки зрения способности к адаптации. В пользу этого, высказанного между делом утверждения свидетельствуют экспериментальные данные, к которым мы сейчас и обратимся.

Чтобы обеспечить эффективную эксплуатацию нового источника пищи, в колонии муравьев происходит мобилизация. Механизмы, с помощью которых муравьи собирают большое число особей возле источников пищи, можно назвать пищевым рекрутированием. Рассмотрим, как выглядит пищевое рекрутирование у двух видов муравьев — *Tetramorium impurum* и *Tapinoma erraticum*. Первый вид способен проводить групповое рекрутирование: муравей-вербовщик, возвращаясь в гнездо от источника пищи, прокладывает след, по которому в дальнейшем он проводит небольшое число рекрутов к источнику пищи. Вместе с тем оба вида способны проводить массовое рекрутирование, когда достаточно лишь проложить след, а дальше рекруты пойдут по нему уже независимо от вербовщика.

Предположим теперь, что оба вида поставлены в одинаковые экспериментальные условия. Для конкретности допустим, что пищевое рекрутирование инициировано одномолярным раствором сахара, расположенным на расстоянии 10 см от гнезда, а муравьи 4 дня голодали. Поведение вербовщиков и рекрутов фиксируется с помощью видеотехники. Точность рекрутирования измеряется по доле завербованных муравьев, фактически достигающих источника пищи, а также средней долей муравьев, следующих по данной вербовочной дорожке в одиночку. Результаты приведены в табл. 6.1. Очевидно, для муравьев *Tetramorium* вербовочная дорожка служит гораздо более слабым сти-

мулом, нежели для *Taripota*. Более того, у *Taripota* массовое рекрутирование происходит с гораздо более высокой точностью. При одинаковых условиях у этого вида до источника пищи доходит целых 73% всех рекрутов, покинувших гнездо, причем движутся они по дорожке, проложенной единственным вербовщиком. Иными словами, ошибки в пищевом рекрутировании значительно варьируют от вида к виду. Покажем теперь, что в зависимости от условий они могут изменяться также и внутри одного и того же вида.

Рассмотрим теперь другой случай — два контейнера (размеры $5 \times 20 \times 30$ см) соединены мостиком. В одном контейнере помещается гнездо, в другом — источники пищи. Рекрутирование на доставку сахарного раствора инициируется после 8 дней голодания. Источники пищи находятся в 11 см от мостика и в 14 см друг от друга. Они выставляются одновременно или с интервалом в 1 ч; производится видеорегистрация движения муравьев *Tetramorium* по мостику. Выясняется, что рекрутирование в этом случае характеризуется совсем невысокой точностью — в сборе пищи участвует лишь от 20 до 40% рекрутов. Большинство муравьев-фуражиров покидает след вербовщика и изучает окрестности. Они быстро обнаруживают второй источник пищи даже в том случае, когда первый источник уже интенсивно используется. Если качество обоих источников одинаково, то обычно один из них используется с максимальной интенсивностью вплоть до полного его исчерпания. Затем муравьи переключаются на второй источник, в результате чего интенсифицируется и его эксплуатация. Таким образом, скорость сбора пищи не снижается даже временно. Если концентрация глюкозы во втором источнике выше, чем в первом, то муравьи поначалу сосредоточивают свои усилия на более выгодном источнике, не отказываясь, впрочем, и от первого. Такой тип поведения сильно напоминает бифуркацию двух одновременно устойчивых состояний (см. рис. 3.9) и ее искажение за счет влияния небольших дефектов (см. рис. 3.10). Кроме того, отсюда четко видно, как случайность позволяет системе переключаться с одного типа поведения на другое.

Зададимся теперь следующим вопросом. Те муравьи, которые теряют по ошибке след, получают возможность открыть новые источники пищи. Однако они не участвуют в эксплуатации уже известных источников. Каково же наилучшее соотношение между флуктуациями, позволяющими делать открытия и инновации, и точным детерминизмом, допускающим непосредственную эксплуатацию? В какой-то мере ответ на этот вопрос помогает получить следующая модель.

Сначала рассмотрим модель глобального рекрутирования при наличии единственного источника пищи. Разумно счи-

тять, что средний поток муравьев, прибывающих к данному источнику (J_+), пропорционален числу встреч между X рабочими особями, использующими этот источник, и остальными $N-X$ особями, где N — число муравьев, способных принять участие в рекрутировании. Таким образом,

$$J_+ = aX(N-X) \quad (6.10a)$$

где a — удельная скорость рекрутирования (в пересчете на одну особь). Поток насекомых от источника (J_-) имеет простой вид:

$$J_- = -bX \quad (6.10б)$$

где b — обратное среднее время, проведенное насекомым вблизи источника и по дороге в гнездо. После исчерпания источника поведением X управляет лишь последний член. Однако пока источник пищи существует, полная скорость изменения X задается равенством

$$\frac{dX}{dt} = aX(N-X) - bX \quad (\text{источник пищи не исчерпан}) \quad (6.11)$$

Это так называемое *логистическое уравнение* широко используется в самых разных популяционных задачах. Предельное значение N называется предельной плотностью насыщения. При $X < N$ поток J_+ положителен, что соответствует положительной обратной связи популяции X на саму себя, хотя эта связь и ослабляется множителем $N-X$, ответственным за насыщение.

Допустим теперь, что перед колонией муравьев на равных расстояниях от гнезда регулярным образом расположено k тождественных источников пищи. Иницируется рекрутирование к источнику, расположенному посередине. Кроме того, предполагается, что муравьи размещены вокруг источника в соответствии с нормальным распределением. Предполагается, что муравей, обнаруживший источник пищи, тотчас начнет мобилизацию своих собратьев. При этих предположениях непосредственное обобщение уравнений (6.10) — (6.11) дает

$$\left. \begin{aligned} \frac{dX_i}{dt} = \sum_{j=1}^k a_{ji} X_j (N - X_1 - \dots - X_k) - bX_i \end{aligned} \right\} \quad (i=1, \dots, k) \quad (6.12a)$$

$$\left. \begin{aligned} (i\text{-й источник пищи не исчерпан}) \\ \frac{dX_i}{dt} = -bX_i \quad (i\text{-й источник пищи исчерпан}) \end{aligned} \right\} \quad (6.12б)$$

Здесь X_1, \dots, X_k — число муравьев, рекрутированных соответственно источниками 1, ..., k . Коэффициенты a_{ji} ($j \neq i$) соответствуют скоростям перерекрутирования i -м источником муравьев, рекрутированных уже j -м источником. Такое перерекрутирование

обусловлено элементом случайности в поведении насекомых. Оценить эти коэффициенты можно с помощью формулы

$$a_{ji} = \frac{1}{S} \int_{|r_i - r| \leq d} \exp \left[-\frac{(r_j - r)^2}{2\sigma^2} \right] dr \quad (6.13)$$

Здесь S введено для должной нормировки, $2d$ — диаметр источника пищи, σ — дисперсия гауссова распределения вероятностей, r_i — координаты i -го источника. Ясно, что чем меньше величина σ , тем более высокой точностью характеризуется рекрутирование.

На рис. 6.7 представлены результаты численного решения уравнений: время, необходимое для полного использования заданного количества пищи, изображено в зависимости от дисперсии статистического распределения σ , характеризующего ошибки поиска. Расчеты выполнены для различных количеств источников пищи. Как и ожидалось, если имеется единственный источник пищи, то наилучшей стратегией рекрутирования будет та, которая функционирует без ошибок. Однако, если пища дробна, то существует некоторое «оптимальное» значение ошибки, минимизирующее время разработки источника. Это определяет уровень случайности, который может оказаться выгодным за счет увеличения возможности открытий новых источников, что позволяет сообществу сосредоточить свою деятельность на наиболее выгодных источниках и способствует колонизации источников, которые будут эксплуатироваться в дальнейшем.

До сих пор у нас различие между муравьиными особями состояло лишь в том, что некоторые аспекты их деятельности подчинялись определенным статистическим распределениям. Однако внутри сообществ насекомых существует разделение труда. Муравьиное гнездо представляет собой женское сообщество, в котором самцы появляются лишь на короткий период, связанный с оплодотворением. Во многих популярных книгах описано разделение труда между матками, обеспечивающими воспроизводство муравьиной колонии, и рабочими, выполняющими необходимые для существования сообщества функции.

У некоторых видов группа рабочих состоит из многих каст, различающихся по внешним признакам; у таких видов наблюдается четкое разделение труда, основанное на возрастных различиях. Хотя это и удивительно, но, по-видимому, внутри данной касты и для данной возрастной группы существует некое дифференцирование между особями по выполняемой деятельности, причем неясно, насколько эти различия важны для сообщества.

При наступлении неблагоприятных условий муравьи перемещают свои гнезда. При этом сначала исследуется местность, затем отыскивается благоприятное место, после чего туда переносится содержимое из предыдущего гнезда. Был поставлен эксперимент, в котором колония муравьев была вынуждена

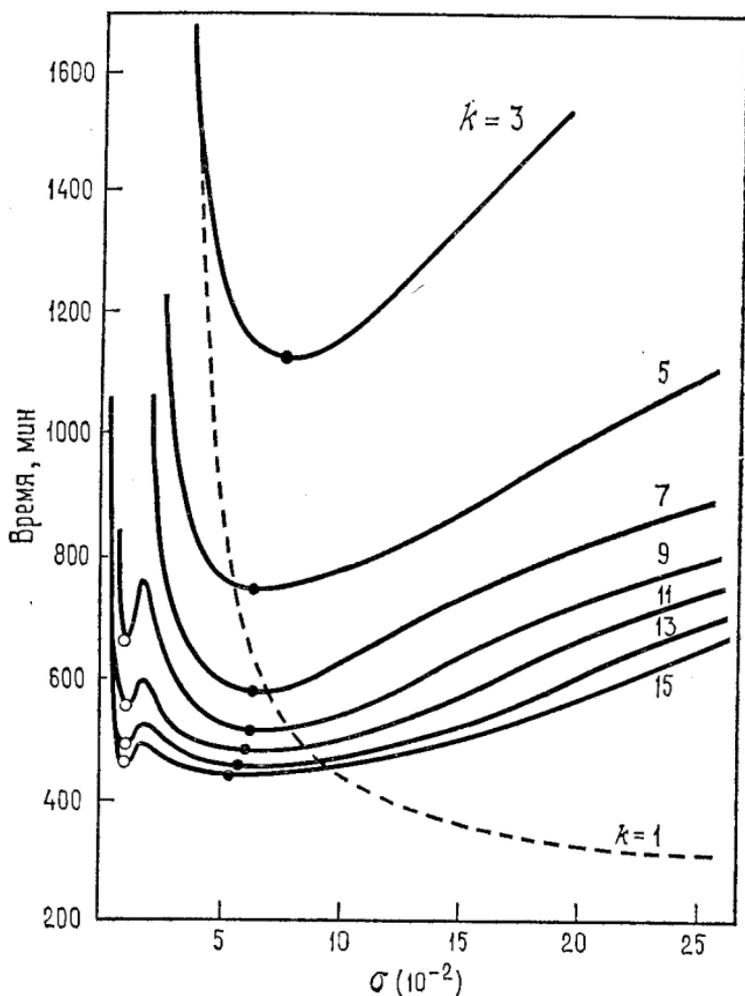


Рис. 6.7. Время полного потребления запасов пищи при разных числа пищевых источников (k). Значения параметров: $N=650$, $a=10^{-3}$, $b=0,016$, $2d=1,4$ см.

совершить несколько переездов в течение 8 мес, причем регистрировалось количество груза, переносимого каждым членом муравьиной семьи. Удивительнее всего оказалось то, что уровень активности характеризуется определенной стабильностью: те муравьи, которые были активны в начальном опыте, проявляют тенденцию к активности в течение нескольких месяцев.

Аналогичная закономерность имеет место и для «ленивых» муравьев. Эти различия зависят не только от данной особи — они определяются также и внешней средой, и предысторией особи. В самом деле, если разделить эти две группы, то ленивые муравьи превращаются в старательных. Если теперь соединить две группы, то на некоторое время колония сохраняет память о произошедшем «возмущении».

В заключение можно отметить, что любое сообщество насекомых проявляет замечательную пластичность, позволяющую при минимальной генетической запрограммированности выполнять ряд сложных задач. И снова первичным механизмом, лежащим в основе столь высокой эффективности, является переход между различными типами поведения, вызываемый внешними условиями.

6.6. Самоорганизация в человеческих сообществах

Ежедневный опыт учит нас, что приспособляемость и пластичность поведения — два основных свойства нелинейных динамических систем, способных совершать переходы вдали от равновесия, относятся к числу наиболее заметных особенностей человеческих сообществ. Поэтому естественно ожидать, что наиболее адекватными для социальных систем будут динамические модели, учитывающие эволюцию и изменчивость.

При построении динамической модели сообщества людей прежде всего следует четко уяснить, что помимо определенной внутренней структуры нужно учитывать довольно жестко заданное внешнее окружение, с которым рассматриваемая система обменивается веществом, энергией и информацией. Возьмем, например, некий город, в который непрерывно поступает сырье и сельскохозяйственная продукция и из которого вывозятся произведенные товары. Одновременно средства связи и массовой информации держат различные группы в курсе текущих событий и современных тенденций.

Эволюция таких систем определяется поведением действующих лиц во взаимодействии с условиями, накладываемыми внешней средой. Именно в этом состоит уникальная специфика гуманитарных систем. В отличие от молекул — основных «актеров» в физико-химических системах и даже в отличие от муравьев или членов других животных сообществ человек всегда строит свои собственные *проекты*, имеет свои собственные *желания*. Некоторые из них основаны на предвидении разумного возможного будущего, а также на догадках относительно желаний других действующих лиц. По этой причине различие между желательным и действительным поведением выступает как внешнее условие нового типа, определяющее

контуры динамики наряду с внешней средой. Основной вопрос, который здесь можно поставить, таков: способна ли при таких условиях эволюция в целом привести к своего рода глобальному оптимуму, или же, напротив, каждая гуманитарная система представляет собой уникальную реализацию некоторого сложного стохастического процесса, для которого никоим образом невозможно установить правила заранее? Иными словами, достаточен ли опыт прошлого для предсказания будущего, или же высокая степень непредсказуемости будущего составляет саму суть человеческих поступков, будь то на уровне изучения индивидуума или на уровне коллективного сотворения истории? Развитые в предыдущих главах представления свидетельствуют о том, что ответ на этот вопрос должен скорее склониться ко второй альтернативе. Посмотрим, позволяет ли математическое моделирование подтвердить это интуитивное ощущение и в то же время более четко установить природу непредсказуемости гуманитарных систем.

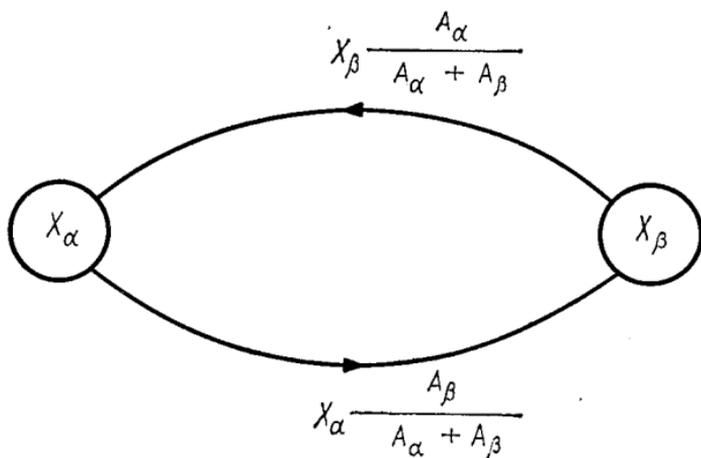


Рис. 6.8. Петля обратной связи, характеризующая реализацию выбора при условии двух одновременно существующих возможностей α и β , степень привлекательности которых характеризуют числа A_α и A_β . Размеры популяций, сделавших выбор α или β , обозначены X_α и X_β соответственно.

Прежде всего попытаемся оценить тот выигрыш, который связан с данным выбором. Можно предположить, что желательность принятия конкретного i -го решения из K альтернативных возможностей, приходящаяся на единицу времени, пропорциональна относительной привлекательности i -го варианта. Однако по мере того как происходит принятие решения относительно данного варианта, в общем случае выигрыш изменяется, так что на картине предпочтительных вариантов на уровне популяции это отразится в виде увеличения или уменьшения привлекательности соответствующих выборов. Эта петля обрат-

ной связи показана на рис. 6.8 в частном случае двух выборов: α и β . Здесь A_α представляет собой привлекательность выбора α , а X_α — число людей, склонившихся к выбору α к настоящему моменту времени. Ясно, что относительное число лиц, желающих поменять свой выбор на выбор β , пропорционально числу тех, кто уже сделал некоторый другой выбор типа α , и относительной привлекательности выбора β , определенной как $A_\beta/(A_\alpha + A_\beta)$. Аналогичным образом число лиц, желающих поменять выбор β на выбор α , пропорционально X_β , умноженному на относительную привлекательность α , равную $A_\alpha/(A_\alpha + A_\beta)$. Это приводит к системе уравнений для X_α балансного типа, очень похожих на экологические уравнения (6.11) — (6.12), полученные в предыдущем разделе:

$$\frac{dX_\alpha}{dt} = aX_\alpha \left(\frac{X_\beta A_\alpha}{A_\alpha + A_\beta} - \frac{X_\alpha A_\beta}{A_\alpha + A_\beta} \right)$$

или, с учетом того, что $X_\beta = N - X_\alpha$, где N — полный размер популяции,

$$\frac{dX_\alpha}{dt} = aX_\alpha \left(\frac{NA_\alpha}{A_\alpha + A_\beta} - X_\alpha \right) \quad (6.14)$$

Аналогичное уравнение можно выписать для X_β . Сопоставление этого уравнения с (6.11) позволяет сделать вывод, что существование различных вариантов выбора влияет на несущую способность системы, которая становится функцией мгновенного состояния системы за счет зависимости привлекательностей от переменных X .

Изложенные соображения легко обобщить на случай произвольного числа выборов K с учетом более реальной ситуации, когда привлекательность i -го варианта зависит от номера j популяции, которая должна сделать выбор. Таким образом получается следующая система уравнений:

$$\frac{dX_i}{dt} = CX_i \left(1 - \frac{X_i}{\sum_{j=1}^K \frac{N_j A_{ij}}{\sum_l A_{lj}}} \right) \quad (i=1, \dots, K) \quad (6.15)$$

Разумеется, здесь предполагается, что популяция неоднородна — она распадается на несколько различных субпопуляций N_j , каждая из которых имеет свое представление об относительной привлекательности данного выбора. Величины N_j удовлетворяют системе уравнений, похожей на уравнения популяционной динамики [см. (6.12)].

Поведение описываемой уравнениями (6.15) динамической системы решающим образом зависит от того, каков характер зависимости привлекательностей A_{ij} и N_j от переменных X_i ,

характеризующих мгновенное состояние популяции. В свою очередь эта зависимость определяется тем, какая система рассматривается в данный момент. В дальнейшем этот вопрос будет обсуждаться в связи с эволюцией городских структур.

Рассмотрим развитие городского центра, в котором определенные территории специализируются на конкретной экономической деятельности и где соседние районы различаются не только по жизненным условиям, но и по доступности рабочих мест и различных видов услуг. Мы будем выделять две основные группы населения, состоящие из трудящихся типа так называемых синих и белых воротничков. Что касается рабочих мест, то они могут относиться к производству товаров для местного потребления или на экспорт, к так называемым третичным функциям (основные и специальные виды обслуживания населения), а также к области финансовой деятельности. Основными переменными, определяющими состояние системы, являются количества представителей k -й группы в точке i , обозначаемые X_i^k ($k=1, 2$), а также соответствующие рабочие места J_i^k . Различные варианты выбора, предлагаемые каждому человеку, находятся в точках i , которые можно считать центрами некоторой решетки, представляющей имеющееся в нашем распоряжении пространство. Механизмы взаимодействия различных переменных суммированы на рис. 6.9.

Чтобы записать видоизмененные уравнения (6.15) для переменных X_i^k , надо выразить величины N_i и A_{ij} через J_i^k и X_j^k . Разумно считать, что основная характеристика субпопуляции N_i — это количество рабочих мест типа k , имеющих в точке i . Что касается привлекательностей A_{ij} , то они являются сложными функциями X_j^k , J_j^k и расстояния между точками i и j . Нет необходимости рассматривать эти величины явным образом, однако можно отметить, что они существенно зависят от цены и времени проезда на работу, цены на землю, характера окрестностей и т. д.

Аналогичные уравнения можно написать и для других категорий «действующих лиц». Например, дополнительно можно описать необходимость развития промышленности в центрах пересечения торговых путей и по возможности недалеко от мест с уже развитой промышленностью.

Только что описанная модель рассматривает эволюцию как автономный процесс, течение которого в каждый момент времени определено механизмами взаимодействия между различными действующими лицами. Внешние ограничения учитываются с помощью различных параметров, а начальное условие можно рассматривать как выражение влияния случайности или систематического внешнего вмешательства вроде «планирования». В качестве альтернативного сценария, более близкого

к реальности, можно позволить системе эволюционировать в течение некоторого времени, затем грубо изменить ее состояние путем внедрения нового вида деятельности или какой-либо иной инновации, затем снова позволить системе следовать ее собственной, автономной динамике до момента введения новой инновации и т. д. Ввиду сильной нелинейности уравнений можно ожидать, что возникнет несколько ветвей решений со сложным набором бифуркационных явлений. Различие в начальных

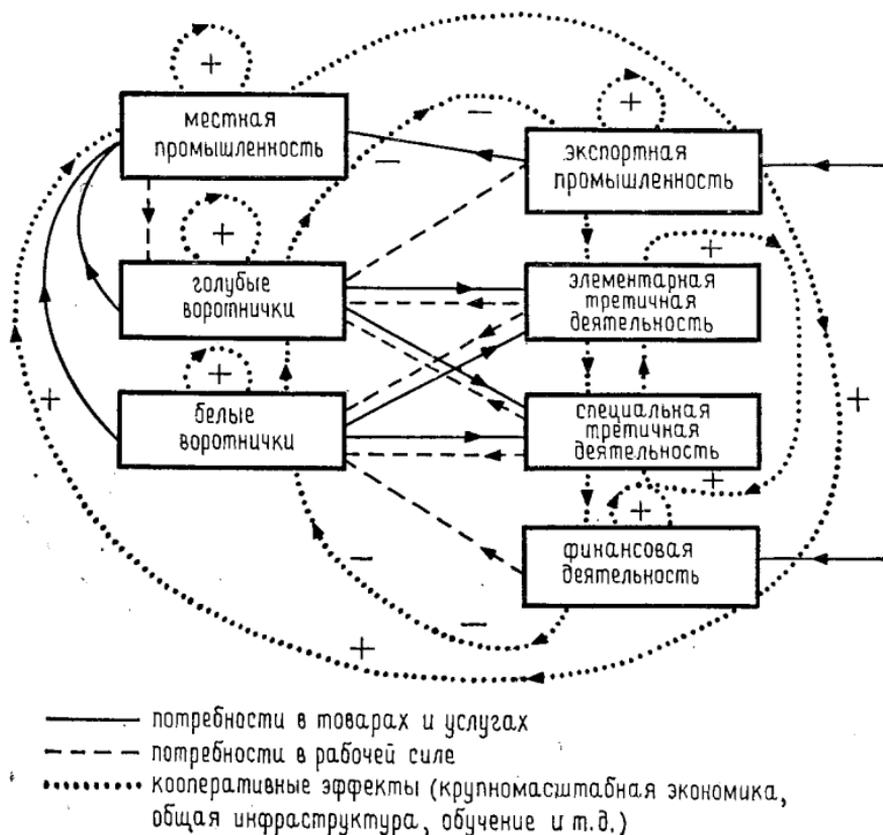


Рис. 6.9. Схематическое представление взаимодействий и потоков информации в упрощенной модели города.

условиях приведет к тому, что система может оказаться в областях притяжения различных режимов, что равносильно включению различных типов эволюции, различных вариантов истории. Тот факт, что из многих возможных регистрируется некоторый конкретный исторический вариант, совсем не обязательно является отражением усилий некоторого «составителя глобального плана», пытающегося оптимизировать какую-то всеобщую функцию — это может быть простым следствием ус-

тойчивости и жизнеспособности данного конкретного типа поведения. Теперь в нашем распоряжении достаточно средств, чтобы ответить на поднятый в начале этого раздела вопрос.

Численный анализ уравнений, описывающих рассматриваемую модель, позволил четко установить существование большого числа решений и сложных бифуркационных явлений. Если в начальном состоянии переменные распределены в пространстве случайным образом, то постепенно наблюдается возникновение некоторой организованной картины с собственными административными и деловыми центрами, промышленными зонами, торговыми центрами и «спальными районами» различного качества. В отсутствие каких-либо сильных возмущений картина остается устойчивой неопределенно долгое время. Это спонтанное нарушение симметрии очень похоже на образование пространственных структур и в гидродинамике и химической кинетике, обсуждавшееся в разд. 1.3 и 1.4.

Из этой модели вытекает также следующий очень интересный результат. Если в какой-то момент времени в систему вводится новый вид деятельности, то в дальнейшем он будет расширяться и стабилизироваться. Если место выбрано удачно, то в дальнейшем это может сделать безуспешными аналогичные попытки, совершенные поблизости. Однако если тот же вид деятельности попытаться внедрить в какое-то другое время, то успех наблюдается отнюдь не всегда — инновация может полностью регрессировать и тем самым оказаться сугубо убыточной. Этот результат иллюстрирует опасности краткосрочного узкого планирования, основанного на непосредственной экстраполяции прошлого опыта. Подобные статические, неперспективные методы грозят обществу застоєм и через какое-то время катастрофой. С другой стороны, главный вывод, следующий из рассмотренного в этом разделе динамического моделирования, состоит в том, что основным источником, позволяющим обществу существовать длительное время, обновляться и находить самобытные пути развития, являются его адаптационные возможности.

ЛИНЕЙНЫЙ АНАЛИЗ УСТОЙЧИВОСТИ

I.1. Основные уравнения

В разд. 2.6 и 2.7 этой книги были даны понятия об устойчивости и бифуркации, а в гл. 3 обсуждался целый ряд свойств ветвления решений и их устойчивости. В настоящем приложении, рассчитанном на математически более подготовленных читателей, излагается формализм, позволяющий определять устойчивость решений уравнений, описывающих динамические системы. В приложении II этот формализм дополняется схемой построения решений, ветвящихся за точкой неустойчивости.

Как уже неоднократно подчеркивалось, свойство устойчивости связано с ответом системы на возмущения различного типа. В связи с этим вообразим некоторое «стандартное состояние» $X_{s1}, \dots, X_{si}, \dots$, где $\{X_{si}\}$ представляет собой набор переменных состояния, непрерывно зондируемого внутренними флуктуациями или внешними возмущениями. Если мы имеем дело с совершенно неискаженной системой, т. е. если в системе не действуют внешние силы и окружающая ее среда совершенно однородна и не зависит от времени, то естественно, что $\{X_{si}\}$ будет соответствовать независящему от времени и пространственно однородному решению, как и следует ожидать при этих условиях. Такой результат означает, что в системе нет самоорганизации и сложных типов поведения. В более общем случае в качестве $\{X_{si}\}$ можно выбрать решение, обладающее наивысшей степенью симметрии, допускаемой внешними ограничениями. Пользуясь терминологией разд. 2.7, можно сказать, что решение $\{X_{si}\}$ лежит на *термодинамической ветви*, которая служит продолжением близкого к равновесию поведения по мере удаления от равновесного состояния. Таким образом, самоорганизацию и возникновение сложного поведения следует рассматривать как переход от $\{X_{si}\}$ к решениям нового типа.

Рассмотрим теперь, каким должен быть формализм, позволяющий выяснить устойчивость рассматриваемого стандартного состояния. Основная трудность теории устойчивости связана с большим числом переменных, описывающих данную задачу, а также с тем обстоятельством, что во многих случаях эти переменные могут оказаться функциями пространственных координат. Поэтому имеет смысл ввести краткое обозначение пол-

ного набора переменных в виде вектора-столбца X с компонентами X_1, \dots, X_i, \dots . В пространственно однородном случае обычно полное число переменных n (например, химических концентраций) бывает конечным. В этом случае X будет определен в обычном векторном пространстве, известном из курса алгебры. Однако если некоторые из переменных зависят от пространственной координаты r , то в дополнение к отмеченному векторному характеру переменная X будет также относиться к некоторому функциональному пространству, например типа функционального пространства с интегрируемым квадратом. Такие пространства нам известны из элементарных курсов квантовой механики.

Воспользовавшись этим обозначением, изменение X можно записать в следующем виде [см. уравнения (2.12) или (3.1)]:

$$\frac{\partial X}{\partial t} = F(X, \lambda) \quad (1.1)$$

Здесь F — оператор, действующий в пространстве, в котором определен вектор X . Вследствие действующих в системе обратных связей этот оператор обычно нелинеен. Посредством λ обозначен набор управляющих эволюцией параметров, таких, как константы скоростей, коэффициенты диффузии и т. д.

Введенное выше стандартное состояние X_s является частным решением уравнения (1.1). Следовательно,

$$\frac{\partial X_s}{\partial t} \equiv F(X_s, \lambda) \quad (1.2)$$

Роль возмущений можно учесть, полагая

$$X = X_s + x \quad (1.3)$$

Здесь x — возмущение. С помощью этого представления из (1.1) можно получить уравнение для x :

$$\frac{\partial x}{\partial t} = F(X_s + x, \lambda) - F(X_s, \lambda) \quad (1.4)$$

Чтобы читателю было легче понять смысл x и оператора F , полезно привести некоторые примеры. В простейшем случае, разумеется, можно рассмотреть задачу с единственной переменной, независимой от пространственной координаты, как в случае модели (2.15) в среде с интенсивным перемешиванием. Тогда оператор F сводится к скалярной величине:

$$F = \text{скорость реакции} = -k_2 X^3 + k_1 a X^2 - k_3 X + k_4 b$$

Тогда правая часть (1.4) принимает вид

$$\begin{aligned} & F(X_s + x, \lambda) - F(X_s, \lambda) = \\ & = -k_2 (X_s + x)^3 + k_1 a (X_s + x)^2 - k_3 (X_s + x) + k_4 b - \\ & - [k_2 X_s^3 + k_1 a X_s^2 - k_3 X_s + k_4 b] \end{aligned}$$

или, после разложения кубического и квадратного биномов,

$$\frac{\partial x}{\partial t} = (-3k_2 X_s^2 + 2k_1 a X_s - k_3) x + (-3k_2 X_s + k_1 a) x^2 - k_2 x^3 \quad (I.5)$$

С другой стороны, если в качестве такой модели рассматривать Брюсселятор в хорошо перемешиваемой среде [уравнения (3.18)], то переменные состояния и скорости процесса получают следующее представление в виде двухкомпонентного вектора:

$$\mathbf{X} = \begin{pmatrix} X \\ Y \end{pmatrix}, \quad X_s = \begin{pmatrix} A \\ B/A \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} A - BX + X^2 Y - X \\ BX - X^2 Y \end{pmatrix}$$

Непосредственные алгебраические выкладки приводят к явному виду уравнения (I.4):

$$\frac{\partial}{\partial t} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (B-1)x + A^2 y + \frac{B}{A} x^2 + 2Axy + x^2 y \\ -Bx - A^2 y - \frac{B}{A} x^2 - 2Axy - x^2 y \end{pmatrix} \quad (I.6a)$$

Если реакции происходят в неперемешиваемой среде, то в \mathbf{F} будет содержаться дополнительный вклад, обусловленный переносом вещества за счет диффузии или конвекции. Для конкретного класса химических систем с диффузией, подробно рассмотренного в разд. 1.4, можно учитывать только диффузию. Как подчеркивалось в разд. 2.2, превосходное представление диффузии часто можно получить с помощью закона Фика, выражаемого уравнением (2.76). Тогда соответствующий оператор \mathbf{F} , например в случае Брюсселятора, будет иметь вид

$$\mathbf{F}_{\text{dif}}(\mathbf{X}) = \begin{pmatrix} D_1 \nabla^2 X \\ D_2 \nabla^2 Y \end{pmatrix}$$

В силу линейности по x и y это дает

$$\frac{\partial}{\partial t} \begin{pmatrix} x \\ y \end{pmatrix}_{\text{dif}} = \begin{pmatrix} D_1 \nabla^2 x \\ D_2 \nabla^2 y \end{pmatrix} \quad (I.6b)$$

Заметим, что оператор \mathbf{F}_{dif} действует как в пространстве двумерных векторов, так и в пространстве функций, зависящих от координаты \mathbf{r} .

Вернемся к более общему уравнению (I.4). Как следует из рассмотренных выше явных примеров, естественно разложить его правую часть вблизи стандартного состояния X_s . Если \mathbf{F} имеет вид полинома по \mathbf{X} , то это всегда возможно, что приводит к конечному числу членов. Однако в более сложных случаях \mathbf{F} может зависеть от \mathbf{X} каким-либо другим способом. В этом случае будем предполагать, что 1) \mathbf{F} по-прежнему допускает разложение в ряд по степеням \mathbf{x} и 2) разложение можно обор-

вать на члене конечного порядка. В силу последнего предположения мы вынуждены ограничиться изучением *инфинитезимальной устойчивости*, т. е. изучением ответа системы на малые возмущения, такие, что $|x|/|X_s| \ll 1$. Это ограничение малосущественно, поскольку инфинитезимальная устойчивость дает необходимое условие неустойчивости в том смысле, что если X_s неустойчиво относительно малых x , то оно будет неустойчиво относительно любых x .

Формально описанное выше разложение можно представить в виде

$$F(X_s + x, \lambda) = F(X_s) + \left(\frac{\delta F}{\delta X}\right)_{X_s} \cdot x + \frac{1}{2} \left(\frac{\delta^2 F}{\delta X \delta X}\right)_{X_s} \cdot xx + \frac{1}{6} \left(\frac{\delta^3 F}{\delta X \delta X \delta X}\right)_{X_s} \cdot xxx + \dots \quad (1.7)$$

Объекты $(\delta F/\delta X)_{X_s}$ и т. п. представляют собой обобщения производных. В математической литературе они называются производными Фреше. Для примера отметим, что сопоставление этой записи с уравнениями (I.6a) и (I.6б) приводит к следующему определению первой производной Фреше для Брюсселятора:

$$\left(\frac{\delta F}{\delta X}\right)_{X_s} = \begin{pmatrix} B-1 + D_1 \nabla^2 & A^2 \\ -B & -A^2 + D_2 \nabla^2 \end{pmatrix} \quad (1.8)$$

Отсюда видно, что $(\delta F/\delta X)_{X_s}$ действует как матрица в пространстве двумерных векторов и как дифференциальный оператор в пространстве функций, зависящих от координат.

1.2. Принцип устойчивости линеаризованной системы

Уравнения (I.4) вместе с (I.7) составляют весьма нелинейную задачу, которая, как правило, столь же сложна, как и исходная задача, описываемая уравнением (I.1). Для успешного дальнейшего продвижения здесь удобно привлечь один очень важный результат анализа. Это теорема, известная под названием принципа линейной устойчивости. С ее помощью можно привести в соответствие свойства устойчивости следующих двух задач: 1) исходной, совершенно нелинейной задачи и 2) линеаризованной задачи, в которой опущены члены выше первого порядка.

Вводя сокращенные обозначения

$$\begin{aligned} \mathcal{L}(\lambda) &= (\delta F/\delta X)_{X_s} \\ h(x, \lambda) &= \frac{1}{2} \left(\frac{\delta^2 F}{\delta X \delta \lambda}\right)_{X_s} \cdot xx + \dots \end{aligned} \quad (1.9)$$

для линейного и нелинейного вкладов соответственно, эти две задачи можно свести к поиску решений следующих систем уравнений:

$$\frac{\partial \mathbf{x}}{\partial t} = \mathcal{L}(\lambda) \cdot \mathbf{x} + \mathbf{h}(\mathbf{x}, \lambda) \quad (\text{нелинейная задача}) \quad (\text{I.10})$$

и

$$\frac{\partial \mathbf{x}}{\partial t} = \mathcal{L}(\lambda) \cdot \mathbf{x} \quad (\text{вспомогательная линейная задача}) \quad (\text{I.11})$$

Заметим, что обе эти системы однородны, поскольку им удовлетворяет тривиальное решение $\mathbf{x} = 0$.

Теорема о линейной устойчивости сводится к следующим утверждениям:

1. Если тривиальное решение $\mathbf{x} = 0$ линеаризованной задачи (I.11) асимптотически устойчиво, то $\mathbf{x} = 0$ (или, что то же самое, $\mathbf{X} = \mathbf{X}_s$) представляет собой асимптотически устойчивое решение нелинейной задачи [уравнения (I.10) или (I.1)].

2. Если тривиальное решение $\mathbf{x} = 0$ линеаризованной задачи (I.11) неустойчиво, то $\mathbf{x} = 0$ (или, что то же самое, $\mathbf{X} = \mathbf{X}_s$) представляет собой неустойчивое решение нелинейной задачи.

Мы не приводим здесь доказательства, поскольку замечания, сделанные в конце разд. I.1, позволяют читателю интуитивно ощутить разумность этих утверждений. Важно иметь в виду, что благодаря этой теореме анализ устойчивости — одного из фундаментальных свойств динамических систем — сводится к решению линейной задачи, т. е. к гораздо более привычной и легкой задаче анализа.

1.3. Характеристическое уравнение

Для дальнейшего продвижения имеет смысл сосредоточиться на конкретном классе *автономных систем*, для которых внешние ограничения не зависят от времени. Кроме того, предположим, что стандартное состояние \mathbf{X}_s соответствует стационарному решению. Отсюда следует, что оператор линейной устойчивости $\mathcal{L}(\lambda)$ не зависит от времени. При этих условиях легко убедиться в том, что уравнения (I.11) допускают решения вида

$$\mathbf{x} = \mathbf{u}e^{\omega t} \quad (\text{I.12})$$

Иными словами, зависимость решения от времени описывается экспоненциальной функцией (заметим, что в общем случае величина ω принимает комплексные значения), в то время как вектор \mathbf{u} учитывает векторный характер \mathbf{x} (в пространстве переменных состояния), а также зависимость \mathbf{x} от пространственной координаты \mathbf{r} .

Подставляя (I.12) в (I.11) и вспоминая, что $\mathcal{L}(\lambda)$ не зависит от времени, после сокращения на множитель $e^{\omega t}$ получаем

$$\mathcal{L}(\lambda) \cdot \mathbf{u} = \omega \mathbf{i} \quad (\text{I.13})$$

Это уравнение, дополненное соответствующими краевыми условиями, определяет задачу на *собственные значения* — классическую задачу математического анализа. Такие задачи часто возникают в квантовой механике или в задачах о тепло- и массопереносе. Примечательно здесь то, что проблема устойчивости разрешается, если удастся найти собственные значения ω независимо от того, какими свойствами при этом будут обладать собственные векторы \mathbf{u} . В самом деле, поскольку зависимость \mathbf{x} от времени описывается соотношением (I.12), имеем

$$\mathbf{x} \sim e^{\omega t} = e^{(\text{Re } \omega) t} \cdot e^{i(\text{Im } \omega) t}$$

где $\text{Re } \omega$ и $\text{Im } \omega$ — соответственно действительная и мнимая части комплексной величины ω . Отсюда следует, что при $\text{Re } \omega < 0$ значение \mathbf{x} экспоненциально убывает со временем. Это затухание может модулироваться колебаниями, если $\text{Im } \omega$ не обращается в нуль. Следовательно, в пределе $t \rightarrow \infty$ достигается стандартное решение $\mathbf{x} = 0$, что означает асимптотическую устойчивость этого решения. С другой стороны, если $\text{Re } \omega > 0$, то возмущения будут экспоненциально расти, так что соответствующее стандартное решение $\mathbf{x} = 0$ будет неустойчивым.

Между этими двумя режимами, к которым относится принцип линейной устойчивости, находится так называемый режим *маргинальной устойчивости*, $\text{Re } \omega = 0$, соответствующий порогу неустойчивости стандартного состояния.

Теперь нам легче понять, насколько важную роль играет управляющий параметр λ . В самом деле, изменение λ приводит к изменению \mathcal{L} и тем самым вызывает изменение собственных значений ω . Тогда наличие перехода между двумя качественно различными режимами будет отражаться в том, что по крайней мере одно из собственных значений ω в зависимости от λ будет изменяться так, как показано на рис. I.1 (см. также рис. 3.16). Значение λ_c , т. е. то значение λ , при котором $\text{Re } \omega$ изменяет свой знак, является *критическим*. Выше этой точки система обязательно неустойчива.

Конкретный вид зависимости собственных значений ω от управляющего параметра λ зависит от характера рассматриваемой задачи. На протяжении этой книги мы уже неоднократно встречались с системами, образующими весьма важный класс, когда оператор \mathbf{F} можно разложить на две части, одна из которых соответствует скорости локального процесса $\mathbf{v}(\mathbf{x}, \lambda)$ типа химической реакции, а другая отражает транспорт вещества за счет диффузии. Конкретным примером таких *реакционно-диф-*

фузионных систем является Брюсселятор, описываемый уравнениями (I.6a) и (I.6б).

В более общем случае реакционно-диффузионной системы со многими переменными можно написать

$$F = v(X, \lambda) + D \cdot \nabla^2 X \quad (I.14)$$

Здесь линеаризованный оператор \mathcal{L} определяется соотношением

$$\mathcal{L} = (\partial v / \partial X)_{x_s} + D \cdot \nabla^2 \quad (I.15)$$

Во многих случаях функции скорости v зависят от пространственных координат. В результате этого полная пространственная зависимость решений линеаризованного уравнения [см. (I.11)]

$$\frac{\partial x}{\partial t} = \left(\frac{\partial v}{\partial X} \right)_{x_s} \cdot x + D \cdot \nabla^2 x \quad (I.16)$$

будет определяться наличием оператора Лапласа ∇^2 . Классический путь определения возможных видов этой пространствен-

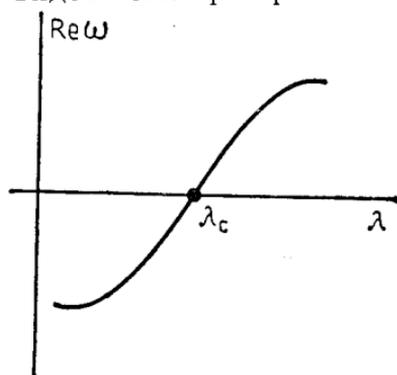


Рис. I.1. Зависимость действительной части показателя ω , характеризующего устойчивость, от управляющего параметра λ . Значение $\lambda = \lambda_c$ соответствует переходу от асимптотической устойчивости ($\lambda < \lambda_c$) к неустойчивости ($\lambda > \lambda_c$).

ной зависимости, при сколь угодно сложной кинетике, отражаемой функцией скорости v , состоит в отыскании собственных значений лапласиана:

$$\nabla^2 \varphi_m(\mathbf{r}) = -k_m^2 \varphi_m \quad (I.17)$$

Разумеется, собственные функции и собственные значения $(\varphi_m, -k_m^2)$ будут зависеть от геометрии, размерности системы и типа краевых условий. Например, если решение (I.17) ищется на отрезке длиной L с краевыми условиями, состоящими в том, что на концах отрезка $r=0$ и $r=L$ функция φ_m обращается в нуль, то легко получить следующие собственные значения и собственные функции:

$$k_m = \frac{m\pi}{L}, \quad \varphi_m(r) = \sin \frac{m\pi r}{L} \quad (I.18)$$

Здесь m — любое отличное от нуля целое. С другой стороны, если решение уравнения (I.10) ищется внутри круга, то функций φ_m будут представлены в виде произведения функций Бесселя и тригонометрических функций.

Какой бы конкретный вид ни имели функции φ_m , решение линейного уравнения (I.11) должно выражаться формулой [см. уравнение (I.12)]

$$\mathbf{x} = c\varphi_m(r) e^{\omega_m t} \quad (\text{I.19})$$

где множитель c учитывает статистический вес различных переменных состояния, а собственное значение ω линеаризованного оператора \mathcal{L} снабжено тем же индексом m , что и соответствующая собственная функция лапласиана φ_m .

Подставляя выражение (I.19) в уравнение (I.16), видим, что в результате действия операторов $\partial/\partial t$ и ∇^2 на \mathbf{x} получается произведение \mathbf{x} и величин ω_m и $-k_m^2$ соответственно. При этом как φ_m , так и $e^{\omega_m t}$, выступающие в роли сомножителей при всех компонентах вектора \mathbf{x} , сокращаются. Таким образом, проблема сводится к уравнению

$$\omega_m \mathbf{c} = \left(\frac{\partial v}{\partial X} \right)_{X_s} \cdot \mathbf{c} - D k_m^2 \cdot \mathbf{c}$$

или, если матрица D имеет диагональный вид,

$$(\omega_m + D_i k_m^2) c_i = \sum_j \left(\frac{\partial v_i}{\partial X_j} \right)_{X_s} \cdot c_j$$

Это система однородных алгебраических уравнений относительно $\{c_i\}$. Она допускает нетривиальные решения лишь в том случае, если определитель матрицы коэффициентов при c_i обращается в нуль. Таким образом, условие разрешимости этой системы приводит к следующему *характеристическому уравнению*:

$$\det \left| \left(\frac{\partial v_i}{\partial X_j} \right)_{X_s} - \delta_{ij}^{kr} (D_i k_m^2 + \omega_m) \right| = 0 \quad (\text{I.20})$$

где δ_{ij}^{kr} — символ Кронекера: $\delta_{ij}^{kr} = 0$ при $i \neq j$ и $\delta_{ij}^{kr} = 1$ при $i = j$. Отсюда можно выразить собственные значения ω_m через собственные значения лапласиана k_m^2 , коэффициенты диффузии D_i и другие параметры, содержащиеся в $(\partial v_i / \partial X_j)_{X_s}$.

I.4. Иллюстрации

Рассмотрим на нескольких характерных примерах, какую же информацию позволяет получить уравнение (I.20).

1. Случай одной переменной. Хорошим примером может служить модель, описываемая уравнением (2.15) с функцией

скорости (I.5). При наличии диффузии линеаризованный оператор представляет собой сумму скаляра и лапласиана:

$$\mathcal{L} = (-3k_2 X_s^2 + 2k_1 a X_s - k_3) + D \nabla^2$$

В результате характеристическое уравнение сводится к

$$\omega_m = -3k_2 X_s^2 + 2k_1 a X_s - k_3 - D k_m^2$$

или, в более общем случае системы с одной переменной,

$$\omega_m = \left(\frac{\partial v}{\partial X} \right)_{X_s} - D k_m^2 \quad (\text{I.21})$$

Отсюда следует, что ω_m — действительное число, так что возмущения стандартного состояния развиваются во времени монотонно. Более того, вклад диффузии в ω_m либо отрицателен, либо равен нулю, т. е. в любом случае диффузия не может дестабилизировать такую систему. Поэтому неустойчивость стационарных решений X_s обусловлена только кинетикой. Отсюда следует, что если после точки неустойчивости новое решение существует, то оно не будет иметь собственной пространственной зависимости.

2. Случай двух переменных. Типичным примером такого рода служит Брюсселятор, для которого явная форма оператора \mathcal{L} дается уравнением (I.8). Соответствующее характеристическое уравнение получается из условия обращения в нуль определителя матрицы 2×2 :

$$\begin{vmatrix} B - 1 - (D_1 k_m^2 + \omega_m) & A^2 \\ -B & -A^2 - (D_2 k_m^2 + \omega_m) \end{vmatrix} = 0$$

или, в явном виде,

$$\omega_m^2 - \omega_m [B - A^2 - 1 - (D_1 + D_2) k_m^2] - A^2 + [(B - 1) D_2 k_m^2 + A^2 D_1 k_m^2] + D_1 D_2 k_m^4 = 0 \quad (\text{I.22})$$

В более общем случае можно также утверждать, что для систем с двумя переменными характеристическое уравнение будет квадратично по ω_m и поэтому может допускать комплексно-сопряженные корни. Полагая $\omega_m = \text{Re} \omega_m + i \text{Im} \omega_m$, из (I.20) находим, что $\text{Re} \omega_m$ обращается в нуль, в то время как $\text{Im} \omega_m$ остается конечной величиной при выполнении следующего условия маргинальной устойчивости:

$$\left(\frac{\partial v_1}{\partial X_1} \right)_{X_s} + \left(\frac{\partial v_2}{\partial X_2} \right)_{X_s} = (D_1 + D_2) k_m^2 \quad (\text{I.23a})$$

В случае Брюсселятора это дает

$$B \equiv B_m = A^2 + 1 + (D_1 + D_2) k_m^2$$

Это означает, что параметр B выражается через собственные значения лапласиана k_m^2 . Для произвольной системы с двумя переменными содержащийся в $(\partial v_i / \partial X_j)_{X_s}$ управляющий параметр λ всегда может быть выбран таким образом, чтобы уравнение (I.23) можно было представить в виде кривой a на

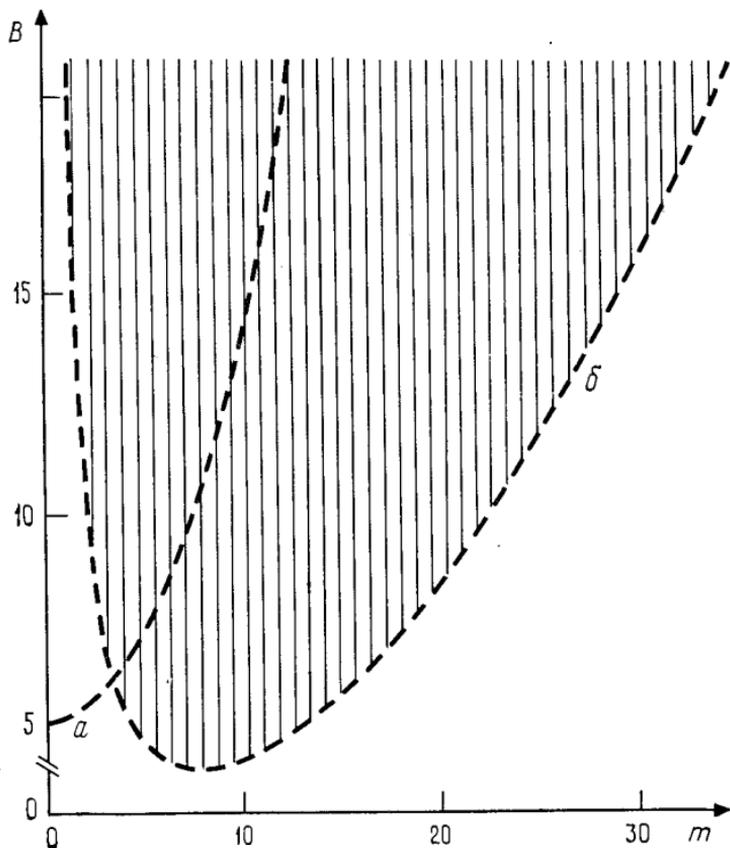


Рис. I.2. Диаграмма линейной устойчивости для Брюсселятора в одномерной среде. Бифуркационный параметр B отложен в зависимости от волнового числа m ; a и b — кривые маргинальной устойчивости для двух типов неустойчивости, характеризуемой комплексно-сопряженными и действительными собственными значениями соответственно. Значения параметров: $A=2$, $L=1$, $D_1=0,0016$, $D_2=0,008$. Величины m и k_m связаны соотношением $k_m = m\pi/L$.

рис. I.2 (см. также рис. 3.29). Соответствующую мнимую часть можно вычислить с помощью уравнения (I.20) после подстановки в него (I.23a). Для случая Брюсселятора это дает

$$\Omega_m^2 = (\text{Im } \omega_m)^2 = A^2 + [A^2 D_1 - (B_m - 1) D_2] k_m^2 + D_1 D_2 k_m^4$$

Если выше кривой маргинальной устойчивости должны существовать новые решения, то на основании изложенных сооб-

ражений можно ожидать, что эти решения будут колебательными с *собственной частотой*, по существу определяемой Ω_m . С другой стороны, эти решения будут все еще характеризоваться тривиальной пространственной зависимостью, поскольку и в этом случае первая неустойчивая мода будет возбуждена при $k_m=0$.

Совершенно другая картина возникает в том случае, если уравнение (I.20) допускает реальные корни ω_m . Тогда условие маргинальной устойчивости принимает простой вид $\omega_m=0$, или

$$\det \left| \left(\frac{\partial v_i}{\partial X_j} \right)_{X_s} - \delta_{ij}^k D_i k_m^2 \right| = 0 \quad (\text{I.236})$$

Слева в этом соотношении — полином четвертого порядка относительно k_m . Если построить график зависимости управляющего параметра λ [он содержится в $(\partial v_i / \partial X_j)_{X_s}$] от k_m , то можно ожидать, что при некотором ненулевом k_{m_c} имеется экстремум λ_c'' , как это показано на кривой *б* рис. I.2 (см. также рис. 3.29). Тем самым первая неустойчивая мода будет иметь пространственную зависимость. Важно отметить, что ввиду (I.236) k_{m_c} целиком определяется системными параметрами независимо от размера, геометрии и даже размерности. Кроме того, в соответствии с (I.19) и (I.17) k_{m_c} определяет пространственную длину волны возмущения стационарного решения X_s , и при $\lambda \geq \lambda_c$, т. е. выше точки неустойчивости, можно ожидать, что это возмущение будет определять основные свойства системы. Таким образом, мы имеем здесь механизм генезиса *собственной длины волны* в системе, до сих пор считавшейся однородной. В таких случаях говорят о *нарушении пространственной симметрии*.

I.5. Системы с хаотической динамикой

Когда динамика становится хаотической, соответствующее термодинамической ветви стационарное состояние уже не может быть стандартным. И все же, как это кратко отмечалось в разд. 3.8, имеется возможность характеризовать неустойчивость движения и соответствующую случайность поведения с помощью *показателей Ляпунова*. Рассмотрим две фазовые траектории, для которых в качестве начальных условий взяты две близко расположенные точки X_{0i} и $X_{0i} + \Delta x_{0i}$ ($i=1, \dots, N$) (рис. I.3). Эволюция расстояния между траекториями $\Delta x_i \equiv \Delta x_i(X_0, t)$ в ближайшей окрестности точки X_{0i} существенно определяется вектором касательной к «стандартной» траектории, скажем φ , выходящей из точки X_{0i} . Таким образом, динамику расхождения между траекториями можно установить, линеаризуя уравнения для переменных X_i вблизи их мгновенных

значений $X_i(t)$. По аналогии с разд. 3.6 [см. уравнение (3.25)] получаем

$$\frac{d\Delta x_i}{dt} = \sum_j L_{ij}(\{\varphi_m(t)\}) \Delta x_j \quad (i=1, \dots, N) \quad (I.24)$$

Введем теперь среднюю экспоненциальную скорость расхождения двух исходно близких траекторий следующим образом:

$$\lambda_L(X_0, \Delta x_0) = \lim_{\substack{t \rightarrow \infty \\ |\Delta x_0| \rightarrow 0}} \frac{1}{t} \ln \frac{|\Delta x(X_0, t)|}{|\Delta x_0|} \quad (I.25)$$

где посредством Δx обозначен вектор в фазовом пространстве с компонентами $\{\Delta x_i\}$, а $|\Delta x|$ — его норма. При этом можно показать, что λ_L существует и конечно. Далее, поскольку в фазовом пространстве имеется N независимых направле-

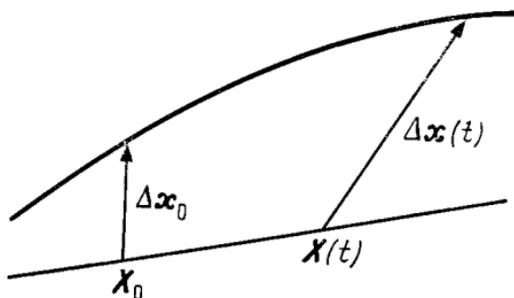


Рис. I.3. Определение характеристических показателей Ляпунова. Две исходно близкие траектории экспоненциально расходятся с течением времени.

ний, λ_L принимает N значений (не обязательно различных), которые и являются показателями Ляпунова. Они обычно упорядочиваются по величине:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$$

Можно показать также, что они не зависят от принятого определения нормы. Заметим, что для любой струи, порождаемой автономной системой, по крайней мере один показатель Ляпунова должен обратиться в нуль. В самом деле, при движении вдоль струи величина $\Delta x(t)$ возрастает в среднем почти линейно, так что при больших t из (I.25) получаем $\lambda_L \sim (1/t) \ln t \sim \sim 0$. С другой стороны, в хаотической струе экспоненциальная расходимость траекторий будет отражаться в положительности по крайней мере одного из показателей Ляпунова, поскольку $|\Delta x(t)|/|\Delta x_0| \sim \exp(\lambda_\alpha t)$ и $(1/t) \ln |\Delta x(t)|/|\Delta x_0| \sim (1/t) \lambda_\alpha t \sim \sim \lambda_\alpha > 0$.

Показатели Ляпунова полезны также при описании средней скорости расширения (или сжатия) $\lambda^{(p)}$ фазового объема $\Delta V = \Delta x_{01} \dots \Delta x_{0p}$, $p \leq N$. При $p = N$ это приводит к соотношению

$$\lambda^{(N)} = \sum_{i=1}^N \lambda_i(X_0) \quad (1.26)$$

В случае консервативной системы фазовый объем сохраняется (см. разд. 3.2), следовательно, $\lambda^{(N)} = 0$. С другой стороны, в случае диссипативных систем наличие аттракторов подразумевает глобальное сокращение фазового объема и, следовательно, $\lambda^{(N)} < 0$. Таким образом, хаотическое движение в консервативной или диссипативной системах с необходимостью подразумевает расширение фазового объема вдоль одних направлений и сокращение — вдоль других. Напрашивается естественный вывод: в случае хаотического движения разнообразие и возможности выбора непрерывно генерируются вдоль тех направлений движения, которые характеризуются сжатием фазового объема. Соответственно предсказуемость становится возможной, когда речь идет о движении вдоль направлений, характеризующихся сжатием фазового объема. Иными словами, хаотические аттракторы (помимо перечисленных свойств они должны еще быть асимптотически устойчивыми) являются потенциальными генераторами и устройствами для переработки информации. Более полное обсуждение этого вопроса можно найти в разд. 4.8.

Приложение II

АНАЛИЗ БИФУРКАЦИЙ

II.1. Общие свойства

Важная роль линейного анализа устойчивости состоит в том, что с его помощью удастся продемонстрировать непрерывное качественное изменение поведения системы выше критического значения λ_c управляющего параметра. Однако нельзя считать доказанным существование новых, физически разумных решений выше порога неустойчивости до тех пор, пока не будут проанализированы полные нелинейные уравнения. В данном приложении будут описаны некоторые шаги в этом направлении.

В качестве исходной точки воспользуемся уравнением (I.10)

$$\frac{\partial \mathbf{x}}{\partial t} = \mathcal{L}(\lambda) \cdot \mathbf{x} + \mathbf{h}(\mathbf{x}, \lambda) \quad (\text{II.1})$$

Фигурирующая здесь величина \mathbf{h} учитывает нелинейность правой части (I.4). Прежде чем мы опишем способ построения ветвящихся решений, полезно затронуть некоторые общие свойства, связанные с вопросами существования и устойчивости.

В разд. 3.4 и 3.5 мы приводили примеры [см. уравнения (3.10), (3.12) и (3.16)], которые действительно подтвердили, что в ряде случаев можно установить существование новых ветвей решения, возникающих при таких значениях параметров, при которых теряется устойчивость стандартного состояния. Однако отсюда не следует делать вывод, что это справедливо всегда.

Чтобы убедиться в том, что возможны и более сложные ситуации, рассмотрим следующую нелинейную систему уравнений:

$$\frac{dX_1}{dt} = X_1 + X_2^3 - \lambda X_1, \quad \frac{dX_2}{dt} = X_2 - X_1^3 - \lambda X_2 \quad (\text{II.2})$$

Стандартным здесь является состояние $X_{s1} = X_{s2} = 0$, а линеаризованные вблизи этого решения уравнения имеют вид

$$\omega X_1 = (1 - \lambda) X_1, \quad \omega X_2 = (1 - \lambda) X_2$$

Эта линейная система однородных алгебраических уравнений имеет нетривиальное решение лишь при условии $\omega = 1 - \lambda$. При

$\lambda=1$ существует состояние маргинальной устойчивости. По сути дела, в рассматриваемой модели $1-\lambda$ представляет собой двукратный корень характеристического уравнения, поскольку оба линеаризованных уравнения содержат лишь по одной переменной. Теперь нам бы хотелось поставить следующий вопрос: будут ли в критической точке возникать новые стационарные решения в результате бифуркации? Критическая точка определяется значением параметра $\lambda=1$, для которого корнем характеристического уравнения служит $\omega=0$. Чтобы ответить на этот вопрос, умножим первое уравнение на X_2 , второе на X_1 и вычтем одно из другого. Таким образом получаем

$$X_1^4 + X_2^4 = 0$$

Очевидно, это уравнение не имеет нетривиальных действительных корней (и тем самым физически приемлемых). Иными словами, несмотря на то что линеаризованная задача удовлетворяет всем требованиям, приведенным в приложении I, ветвящихся стационарных решений не существует.

Рассмотрим несколько внимательнее различие между моделями, обсуждавшимися, с одной стороны, в разд. 3.5 и 3.6, а с другой — в данном разделе [уравнения (II.2)]. При этом выясняется, что в первой из этих моделей корень характеристического уравнения, по крайней мере в точке маргинальной устойчивости, простой, а во второй модели этот корень всегда двукратный. Этот вывод следует рассматривать не как некоторую случайность, но скорее как иллюстрацию следующей общей теоремы:

Если собственное значение, найденное из линеаризованных вблизи ветви X_s уравнений, характеризуется *нечетной кратностью* при $\lambda \approx \lambda_c$ (т. е. если ω — кратный корень характеристического уравнения с нечетной кратностью), и обращается в нуль при $\lambda = \lambda_c$, то существует по крайней мере одна новая ветвь решения, выходящая из (X_s, λ_c) . Более того, эта ветвь либо тянется до бесконечности, либо претерпевает другую бифуркацию.

Значительно более подробную информацию можно получить в случае простых собственных значений в предположении, что при λ_c величина $\text{Re } \omega$ не просто обращается в нуль, но еще и вся ветвь $\text{Re } \omega$ как функция λ пересекает ось λ . Это *условие трансверсальности*, записываемое в виде

$$\frac{d}{d\lambda} \text{Re } \omega(\lambda) |_{\lambda = \lambda_c} \neq 0 \quad (\text{II.3})$$

иллюстрируется рис. I.1. При таких условиях можно получить следующий результат:

Если ветвь (X_s, λ) , асимптотически устойчивая при $\lambda < \lambda_c$, становится неустойчивой при $\lambda = \lambda_c$, имея простое собственное значение, то при выполнении условия трансверсальности в этой точке справедливы утверждения:

1. Решения, ответвляющиеся в надкритическую область, устойчивы, а решения, ответвляющиеся в подкритическую область, неустойчивы.

2. Ответвляющиеся от стационарного, решения будут стационарными, если $\text{Im} \omega(\lambda_c) = 0$, и периодическими, если $\text{Im} \omega(\lambda_c) \neq 0$.

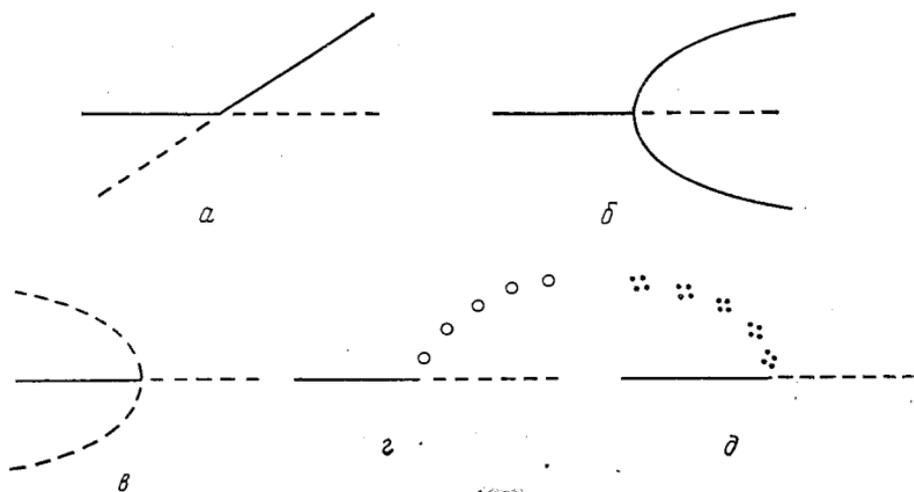


Рис. II.1. Бифуркационные диаграммы вблизи стационарного состояния в случае простого собственного значения оператора линейной устойчивости. а — закритическая бифуркация стационарных решений (II.12); б, в — надкритическая и подкритическая симметричные бифуркации стационарных решений (II.13); г — надкритическая бифуркация Хопфа; д — подкритическая бифуркация Хопфа.

Различные возможные варианты бифуркаций представлены на рис. II.1. Мы не будем в дальнейшем на них останавливаться, поскольку все они довольно подробно обсуждались в гл. 3 и приложении I.

II.2. Разложение решений в ряд по теории возмущений

Предположим теперь, что условия применимости сформулированной общей теоремы выполнены, и опишем процедуру построения ответвляющихся решений в явном виде. Для простоты ограничимся случаем бифуркации стационарных решений при простых собственных значениях ($\text{Im} \omega(\lambda_c) = 0$), возникаю-

щих в первой точке неустойчивости стандартного состояния. Это означает, что нам нужно решить уравнение [см. (II.1)]:

$$\mathcal{L}(\lambda) \cdot \mathbf{x} + \mathbf{h}(\mathbf{x}, \lambda) = 0 \quad (\text{II.4})$$

Мы уже неоднократно отмечали, что при решении нелинейных задач возникают значительные трудности. По этой причине нам приходится отказаться от намерения получить точные результаты общего характера и ограничиться анализом *локального поведения* решений в окрестности точки бифуркации λ_c . Далее, предположим, что новые решения возникают в точке λ_c непрерывно, что позволяет исключить из рассмотрения вертикальные ответвления или скачки.

При этих допущениях можно разложить \mathbf{x} в окрестности λ_c в ряд по степеням некоторого малого параметра. Ясно, что такой параметр должен быть связан с разностью $\lambda - \lambda_c$, поскольку при $\lambda = \lambda_c$ норма $|\mathbf{x}|$ решения обращается в нуль. Однако у нас нет оснований считать, что таким малым параметром должна быть сама разность $\lambda - \lambda_c$, так как у нас нет принципиальных гарантий аналитичности решений по переменной $\lambda - \lambda_c$. В действительности в гл. 3 мы уже встречались с примерами выраженной неаналитичности [см., например, уравнение (3.116)].

Чтобы не задаваться определенной зависимостью решения от $\lambda - \lambda_c$, введем вспомогательный параметр малости ε , по которому разложим как \mathbf{x} , так и $\lambda - \lambda_c$:

$$\begin{aligned} \mathbf{x} &= \varepsilon \mathbf{x}_1 + \varepsilon^2 \mathbf{x}_2 + \dots \\ \lambda - \lambda_c &= \varepsilon \gamma_1 + \varepsilon^2 \gamma_2 + \dots \end{aligned} \quad (\text{II.5})$$

Подставляя эти разложения в уравнения (II.4), в первом порядке по ε получаем

$$\mathcal{L}(\lambda_c) \cdot \mathbf{x}_1 = 0 \quad (\text{II.6})$$

Эта однородная система уравнений подробно обсуждалась в приложении I в связи с линейным анализом устойчивости. Ее решение имеет вид [см. равенство (I.12) при $\omega = 0$, что соответствует маргинальной устойчивости при $\lambda = \lambda_c$]:

$$\mathbf{x}_1 = c \cdot \mathbf{u} \quad (\text{II.7})$$

Для дальнейшего в решении было полезно выделить множитель c , который на данном этапе является неопределенным в связи с однородностью уравнений (II.6). Поэтому будем считать, что \mathbf{u} полностью определяется из решения линеаризованной задачи. Например, для Брюсселятора из (I.8) имеем

$$\begin{aligned} (B_c - 1)u_1 + A^2 u_2 - D_1 k_c^2 u_1 &= 0 \\ -B_c u_1 - A^2 u_2 - D_2 k_c^2 u_2 &= 0 \end{aligned}$$

Отсюда вектор u получается в следующем простом виде:

$$u = \begin{pmatrix} 1 \\ -\frac{B_c}{A^2 + D_2 k_c^2} \end{pmatrix}$$

Перейдем теперь к членам более высокого порядка теории возмущений, что должно нам, в частности, позволить определить амплитуду c главной части решения, ϵx_1 . С учетом членов порядка ϵ^2 из (II.4) получаются вклады трех различных типов. Во-первых, оператор $\mathcal{L}(\lambda_c)$ может действовать на член второго порядка x_2 разложения x . Во-вторых, можно вычислить $\mathcal{L}(\lambda)$ в некоторой точке λ , близкой к λ_c , и рассмотреть его действие на x_1 . Такой сдвиг будет формально выражаться первым членом разложения $\mathcal{L}(\lambda)$ в ряд Тейлора вблизи λ_c . Это даст вклад типа $\gamma_1 \mathcal{L}_\lambda(\lambda_c) \cdot x_1$, где производные отмечены индексами. Например, если в качестве бифуркационного параметра в случае Брюсселятора используется параметр B , то, как видно из уравнений (I.8),

$$\mathcal{L}_\lambda(\lambda_c) \cdot x_1 |_{\text{Brusselator}} = \begin{pmatrix} x_1 \\ -x_1 \end{pmatrix}$$

В-третьих, вклад порядка ϵ^2 дает квадратичная часть h , которая снова будет формально обозначаться как второй член разложения в ряд Тейлора вблизи нуля. В случае Брюсселятора это будет просто квадратичная часть (I.6a):

$$\begin{pmatrix} \frac{B}{A} x^2 + 2Axy \\ -\frac{B}{A} x^2 - 2Axy \end{pmatrix}$$

Снова обозначая порядок производных индексами, получаем наконец полное уравнение с учетом членов порядка ϵ^2 :

$$\mathcal{L}(\lambda_c) \cdot x_2 = -\gamma_1 \mathcal{L}_\lambda(\lambda_c) \cdot x_1 - \frac{1}{2} h_{xx}(\lambda_c) \cdot x_1 x_1 \quad (\text{II.8})$$

Продолжая в том же духе, для членов более высокого порядка получим

$$\begin{aligned} \mathcal{L}(\lambda_c) \cdot x_3 = & -\gamma_1 \mathcal{L}_{\lambda\lambda}(\lambda_c) \cdot x_2 - \gamma_2 \mathcal{L}_\lambda(\lambda_c) \cdot x_1 - \\ & - \frac{\gamma_1}{2} h_{xx\lambda}(\lambda_c) \cdot x_1 x_1 - \frac{\gamma_1^2}{2} \mathcal{L}_{\lambda\lambda}(\lambda_c) \cdot x_1 - \\ & - \frac{1}{6} h_{xxx}(\lambda_c) \cdot x_1 x_1 x_1 - h_{xx}(\lambda_c) \cdot x_1 x_2 \end{aligned} \quad (\text{II.9})$$

и т. д.

II.3. Бифуркационные уравнения

Мы показали, что путем разложения в ряд по теории возмущений исходная нелинейная задача [уравнения (II.4)] сво-

дится к бесконечной последовательности *линейных* уравнений. Первое из этих уравнений [формула (II.6)] линейно и однородно, и его решение с точностью до множителя дается равенством (II.7). Остальные уравнения неоднородны, поскольку их правые части зависят от решения уравнений более низкого порядка. Вспоминая, что $\mathcal{L}(\lambda_c)$ имеет вид матрицы, действующей в пространстве векторов $\{x_i\}$, попытаемся записать их решения в виде

$$x_i \underset{\text{(формально)}}{\sim} \mathcal{L}^{-1}(\lambda_c) \cdot q_i, \quad i \geq 2 \quad (\text{II.10})$$

где посредством q_i обозначены правые части уравнений (II.8), (II.9) и т. д. Однако дело в том, что $\mathcal{L}(\lambda_c)$ не допускает обращения, поскольку в силу (II.6) существуют нетривиальные решения, которые после воздействия на них этой матрицы дают нуль. Похожая ситуация встречается в алгебре, когда определитель системы линейных неоднородных уравнений обращается в нуль. Интуитивно можно ожидать, что в таких случаях действие обратного оператора (II.10) приведет к расходимостям, если не удастся элиминировать части q_i , принадлежащие нуль-пространству оператора $\mathcal{L}(\lambda_c)$ [пространству функций, удовлетворяющих (II.6)]. Достичь этого позволяет следующий важный результат математического анализа, известный под названием *теоремы об альтернативе Фредгольма*. Вкратце он сводится к тому, что правая часть q_i должна быть *ортогональна* векторам u^+ , удовлетворяющим уравнению

$$\mathcal{L}^+(\lambda_c) \cdot u^+ = 0 \quad (\text{II.11})$$

где \mathcal{L}^+ — оператор, сопряженный \mathcal{L} . Для систем с реакцией и диффузией \mathcal{L}^+ получается из \mathcal{L} простым транспонированием, поскольку для симметричных краевых условий лапласиан ∇^2 является самосопряженным оператором. Например, для Брюсселятора имеем

$$\mathcal{L}^+(\lambda_c) = \begin{pmatrix} B_c - 1 + D_1 \nabla^2 & -B_c \\ A^2 & -A^2 + D_2 \nabla^2 \end{pmatrix}$$

Как следует понимать ортогональность в упомянутой теореме? Мы уже отмечали двойственный характер вектора x , который является одновременно вектором в обычном конечномерном пространстве и функцией в бесконечномерном пространстве. Ортогональность двух векторов в обычном конечномерном пространстве сводится к обращению в нуль их скалярного произведения, обозначаемого точкой: $u^+ \cdot q_i$. Скалярное произведение в пространстве функций — вещь несколько более тонкая и здесь не будет обсуждаться подробно. Достаточно отметить, что во многих задачах квантовой механики и в других

физических областях она сводится к вычислению пространственного интеграла от произведения функций (или скалярного произведения векторов). Обозначая это обобщенное произведение угловыми скобками $\langle \rangle$, альтернативу Фредгольма можно записать в виде

$$\langle \mathbf{u}^+, \mathbf{q}_i \rangle = 0, \quad i \geq 2 \quad (\text{II.12})$$

Заметим, что точно так же, как и \mathbf{u} , вектор \mathbf{u}^+ определен однозначно. Более того, помимо своей зависимости от пространственных координат, в правых частях содержатся лишь неопределенные амплитуды c . Следовательно, уравнения (II.12), в которых интегрирование по пространству предполагается выполненным, сводятся к алгебраическим уравнениям, позволяющим найти амплитуду c главного вклада в ответившееся решение, x_1 . Вводя полную амплитуду $z = \varepsilon c$ решения εx_1 , с помощью непосредственных алгебраических выкладок из (II.12) можно получить следующие *бифуркационные уравнения*:

$$-(\lambda - \lambda_c) Q_1 z + Q_2 z^2 = 0 \quad \text{при } Q_2 \neq 0 \quad (\text{II.13})$$

$$-(\lambda - \lambda_c) Q_1 z + Q_3 z^3 = 0 \quad \text{при } Q_2 = 0, Q_3 \neq 0 \quad (\text{II.14})$$

где коэффициенты Q_i получаются непосредственно при выполнении указанных действий.

Заметим, что мы начинали с задачи, в которой фигурировало большое число переменных, которые к тому же могут сложным образом зависеть от пространственных координат. И все-таки в окрестности точки бифуркации всю динамику удалось свести к очень простым *нормальным формам*, аналогичным тем, что обсуждались в разд. 3.4 и 3.5, и содержащим лишь одну неизвестную величину — *параметр порядка* z . Это служит обоснованием для соображений, изложенных во второй части разд. 3.6.

Более сложные бифуркационные уравнения возникают в том случае, когда рассматриваются периодические решения или бифуркации вблизи кратных собственных значений. Краткий обзор этого круга вопросов читатель может найти в книге Гукенхаймера и Холмса.

ВОЗМУЩЕНИЕ РЕЗОНАНСНЫХ ДВИЖЕНИЙ В НЕИНТЕГРИРУЕМЫХ КОНСЕРВАТИВНЫХ СИСТЕМАХ

III.1. Отображение закручивания

В разд. 3.3 мы видели, что движение интегрируемой консервативной системы с N степенями свободы изображается как движение точки по поверхности N -мерного тора, находящегося в $2N$ -мерном фазовом пространстве (см. рис. 3.7, б при $N=2$). Хотелось бы интуитивно понять влияние неинтегрируемых возмущений на этот тип движения и, в частности, разобраться в качественных изменениях, возникающих вблизи резонансов, с которыми мы уже встречались в подразд. 3.8Б. Для большей наглядности в этом приложении мы ограничимся случаем двух степеней свободы, однако следовало бы отметить, что основные соображения применимы также и к системам более общего характера.

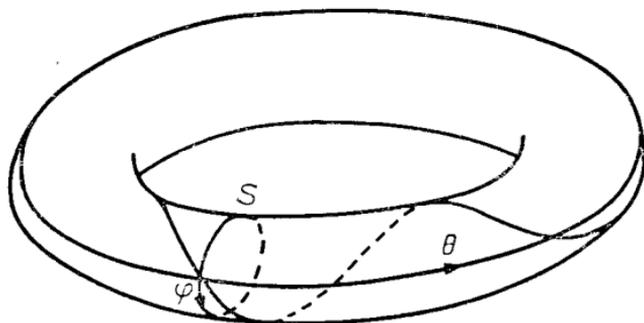


Рис. III.1. Секущая поверхность Пуанкаре S , удобная для изучения резонансных движений на торе.

Пусть θ и φ — два угла, определяющие соответственно «долготу» и «широту» точки на торе (рис. III.1), а ω_1 и ω_2 — угловые частоты движения вдоль этих направлений. На языке переменных угол — действие уравнения движения имеют простой вид:

$$I = I_0 = \text{const}$$

$$\varphi = \omega_1(I)t + \varphi_0, \quad \theta = \omega_2(I)t + \theta_0 \quad (\text{III.1})$$

Мы выписали лишь одну переменную типа действия I , поскольку с учетом сохранения энергии вторая такая переменная

полностью выражается через I и полную энергию. Кроме того, мы допускаем здесь, что угловые частоты зависят от действия.

Удобно рассматривать движение с помощью секущей поверхности Пуанкаре S (см. также рис. 3.22), расположенной меридионально. Фазовые траектории будут пересекаться с этой плоскостью каждый раз, когда они будут совершать полный оборот в θ -направлении, т. е. при $\theta - \theta_0 = 2\pi n$ ($n=1, 2, \dots$). Следовательно, такие пересечения будут происходить в моменты времени $t_n = 2\pi n / \omega_2$. Подставляя эти значения t в уравнение для φ (III.1), получаем

$$\varphi_n - \varphi_0 = 2\pi \frac{\omega_1}{\omega_2} n$$

или, воспользовавшись характерными для рекуррентных соотношений обозначениями,

$$\begin{aligned} I_{n+1} &= I_n \\ \varphi_{n+1} &= \varphi_n + 2\pi\alpha(I_n) \end{aligned} \quad (\text{III.2})$$

где для удобства вращательное число α определяется как

$$\alpha = \frac{\omega_1}{\omega_2} \quad (\text{III.3})$$

Очевидно, определяемое равенствами (III.2) отображение преобразует окружности в окружности. Однако в общем случае

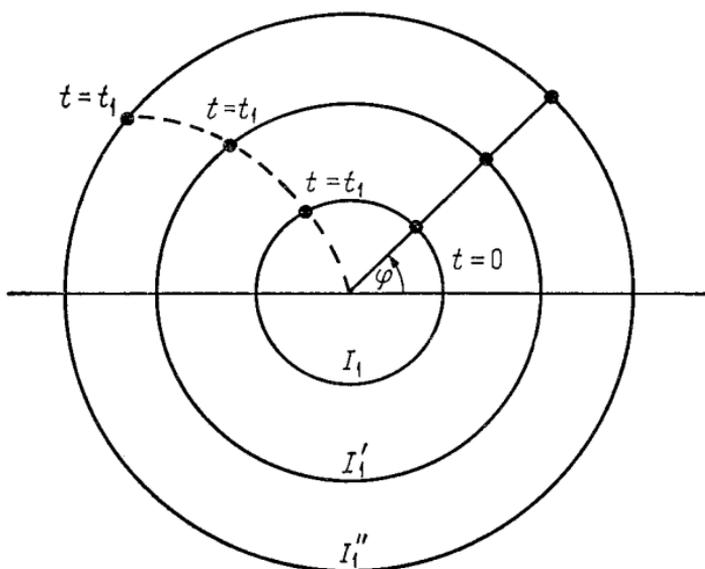


Рис. III.2. Отображение закручивания. Образы трех точек, находящихся вначале (при $t=0$) на одном и том же радиусе, приобретают сдвиг по фазе при $t=t_1$.

вращательное число будет изменяться от одной окружности к другой, поскольку оно зависит от действия, которое служит как бы меткой для разных окружностей. Иными словами, начиная с одинакового значения φ_0 угла φ при $t=0$, изображающие точки на разных окружностях в какой-то другой момент пересечения $t=t_1$ окажутся сдвинутыми друг относительно друга по фазе (см., например, рис. III.2). По этой причине отображение (III.2) называется *отображением закручивания*. Как отмечалось в разд. 3.8 в связи с рис. 3.22, каждая периодическая траектория на торе (α — рациональное число, равное r/s , где r и s — целые числа) пересекает поверхность Пуанкаре в фиксированных точках, число которых равно числу оборотов s вдоль направления θ .

Рассмотрим теперь слабое возмущение соответствующей гамильтоновой системы. Его можно представить в виде возмущенного отображения закручивания:

$$\begin{aligned} I_{n+1} &= I_n + \varepsilon f(I_n, \varphi_n) \\ \varphi_{n+1} &= \varphi_n + 2\pi\alpha(I_n) + \varepsilon g(I_n, \varphi_n), \quad \varepsilon \ll 1 \end{aligned} \quad (\text{III.4})$$

Можно показать, что, как следствие теоремы Лиувилля (см. разд. 3.2), это отображение сохраняет площадь в фазовом пространстве.

III.2. Влияние возмущения в случае рациональных вращательных чисел

Когда речь идет о невозмущенном отображении закручивания (III.2), как показывают изложенные соображения, для рационального $\alpha(I) = r/s$ любая точка на окружности $I = \text{const}$ является неподвижной точкой с периодом s , поскольку в ней периодически должны происходить пересечения поверхности Пуанкаре с некоторой траекторией, совершившей s оборотов вдоль направления изменения долготы. Главный результат, относящийся к влиянию возмущения на эту картину, восходит ко времени Пуанкаре и Биркгофа. Суть его сводится к следующему: из всего континуума неподвижных точек после наложения возмущения неподвижными остаются лишь $2ks$ точек, где k — целое. При этом половина их будет гиперболическими, а половина — эллиптическими. Рассмотрим доказательство, для чего нам придется часто обращаться к рис. III.3, $a-g$.

Допустим для определенности, что число α возрастает по направлению внутрь тора. Тогда интегрируемая система будет обязательно иметь две квазипериодические траектории — одну снаружи и одну внутри резонансного тора, причем после совершения s оборотов эти траектории отобразятся на поверх-

ности Пуанкаре со сдвигом соответственно по и против часовой стрелки относительно их начальных пересечений (см. рис. III.3, *a*).

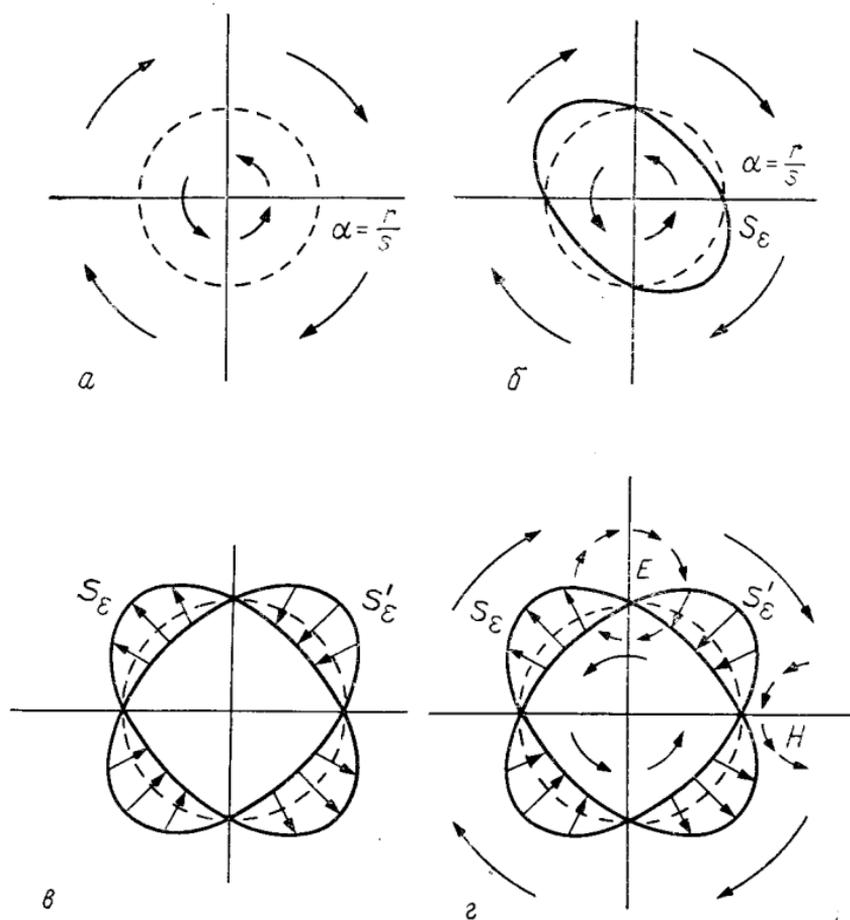


Рис. III.3. Иллюстрация теоремы Пуанкаре — Биркгофа о сохранении некоторых неподвижных точек на поверхности Пуанкаре S при слабом возмущении резонансного движения. *a* — стрелками по обе стороны штриховой кривой указаны направления фазовых сдвигов, претерпеваемых двумя квазипериодическими траекториями при последовательных итерациях отображения; *b* — геометрическое место точек пересечения траекторий возмущенной системы, возвращающихся к одной и той же угловой координате φ после s итераций, с поверхностью Пуанкаре (кривая S_ε). Величина s соответствует числу оборотов при изменении угла θ ; *в* — образ S_ε (кривая S'_ε) после s итераций возмущенного отображения. Следует отметить наличие четного числа неподвижных точек; *г* — устойчивость неподвижных точек возмущенного отображения закручивания. E — эллиптические точки, H — гиперболические точки.

Поскольку квазипериодическое движение при малых возмущениях не теряет свойства грубости, в возмущенной системе будут по-прежнему иметься две траектории, обладающие эти-

ми свойствами после s итераций возмущенного отображения закручивания. Между инвариантными кривыми, порожденными на поверхности Пуанкаре этими траекториями, должна существовать некоторая кривая, охватывающая начало координат. На рис. III.3, б эта кривая обозначена S_e . Она должна состоять из точек пересечения тех траекторий, для которых угловая координата φ остается неизменной после s итераций возмущенного отображения закручивания. И все-таки изменения в *радиальном* направлении остаются возможными. Следовательно, в типичном случае после s итераций кривая S_e будет смещена. Обозначим через S_e' создаваемый тем самым образ S_e (см. рис. III.3, в), который снова охватывает начало координат. Как отмечалось ранее, ввиду своей консервативности рассматриваемое преобразование сохраняет площадь на поверхности Пуанкаре. Отсюда следует, что кривые S_e' и S_e должны охватывать одинаковые площади. Однако это возможно лишь в том случае, если они пересекаются друг с другом четное число раз. По построению каждая из этих точек после s итераций возвращается к исходному положению и потому является неподвижной точкой с периодом s . Каждая из s ее итераций, вместе образующих период s , разумеется, также будет неподвижной точкой. Таким образом, мы установили существование неподвижных точек, заключенных между двумя вложенными инвариантными кривыми на поверхности Пуанкаре, причем число этих точек четно и кратно s . Принимая без доказательства, что эллиптические точки чередуются с гиперболическими, а также воспользовавшись тем очевидным фактом, что эллиптические точки не могут отображаться в гиперболические и наоборот, мы приходим к выводу, что всего остается $2ks$ неподвижных точек. При этом целое число k остается неизвестным.

Перейдем теперь к обсуждению устойчивости этих неподвижных точек. Рассмотрим окрестность точки E , показанной на рис. III.3, а. Отклонение от этой точки вниз приведет сначала ко вращению по часовой стрелке, что обусловлено радиальным изменением вращательного числа α , затем последует движение по направлению внутрь (в соответствии с направлением стрелок на рисунке) и вращение против часовой стрелки и, наконец, движение будет направлено наружу. В результате мы получаем вращение вокруг точки E , которая, следовательно, является эллиптической (устойчивой) неподвижной точкой. С другой стороны, отклонение за счет возмущения от неподвижной точки H (рис. III.3, а) приводит ко все большему удалению изображающей точки от H . Следовательно, точка H является гиперболической (неустойчивой) точкой. Такое чередование устойчивости неподвижных точек дает качественное

объяснение основных особенностей фазового портрета, представленного на рис. 3.23.

III.3. Гомоклинические точки

Что будет в том случае, когда на отображении Пуанкаре появляется неустойчивая неподвижная точка гиперболического типа (седло)? По аналогии с результатами, полученными в разд. 3.3—3.5, в окрестности любой такой точки существуют две особые кривые — сепаратрисы, пересекающиеся в этой точке (рис. III.4, кривые S и U). Это не противоречит теореме о единственности решений соответствующей системы динамических уравнений (см. разд. 3.1), если только система не дости-

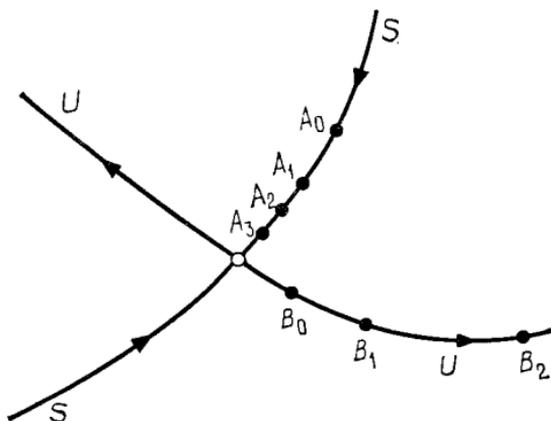


Рис. III.4. Последовательные итерации точек A_0 и B_0 , лежащих соответственно на устойчивой и неустойчивой сепаратрисах гиперболической неподвижной точки.

гает неподвижной точки H из любой начальной точки на кривых S и U за конечное время. Так, если рассмотреть некоторое начальное состояние A_0 на сходящейся сепаратрисе S , то ее первая итерация за счет отображения приведет к состоянию A_1 на кривой S , более близкому к точке H . Однако последовательность A_0, A_1, \dots, A_n будет сходиться к H лишь при $n \rightarrow \infty$. Аналогично начальное состояние B_0 на расходящейся сепаратрисе U приведет с ростом времени к последовательности состояний B_1, B_2, \dots , все более и более удаляющихся от H , в то время как при обращении времени оно приведет к последовательности, аналогичной A_n , т. е. сходящейся к H при $n \rightarrow -\infty$. Теперь спрашивается, что происходит с сепаратрисами вдали от неподвижной точки H , и в особенности когда вместе с H имеется и много других гиперболических неподвижных точек, — та-

кая возможность была показана ранее в настоящем приложении.

В связи с этим вопросом на рис. III.5 представлены три различные возможности. В первом случае сепаратрисы соединят две гиперболические точки таким образом, что устойчивая сепаратриса одной точки превращается в неустойчивую сепаратрису другой и наоборот. Такие *гетероклинические кривые* возникают уже при анализе интегрируемых гамильтоновых систем типа маятника (см. рис. 3.8). Во втором случае обе кривые S и U совпадают, образуя *гомоклиническую кривую*. Такая возможность уже упоминалась в разд. 3.10 (см. рис. 3.29). Третья возможность, представленная на рис. III.5, в, сводится к тому, что неустойчивая сепаратриса уходит на бесконечность, а устойчивая — приходит из бесконечности.

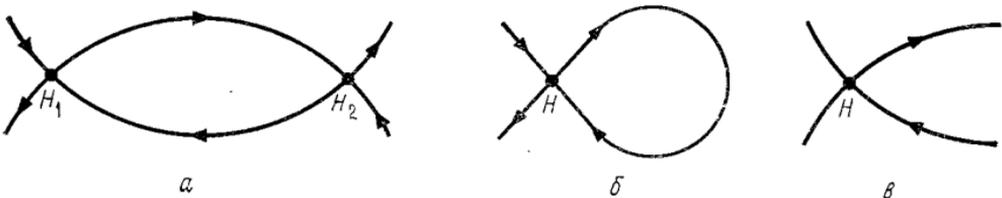


Рис. III.5. Возможные конфигурации сепаратрис гиперболических неподвижных точек. *а* — гетероклинические кривые; *б* — гомоклиническая кривая (петля сепаратрисы); *в* — сепаратрисы уходят на бесконечность.

Ясно что две первые ситуации носят довольно исключительный характер, третья же физически нереальна, поскольку она подразумевает (в отсутствие других притягивающих неподвижных точек) взрывоподобное поведение одной из переменных, описывающих состояние системы. На рис. III.6 показана гораздо более естественная ситуация, открытая Пуанкаре: устойчивая и неустойчивая сепаратрисы неподвижной точки H пересекаются в точке P , которую Пуанкаре назвал *гомоклинической точкой*. Еще раз подчеркнем, что такое пересечение не противоречит единственности решения соответствующей непрерывной динамической системы, хотя P и не является неподвижной точкой отображения, поскольку инвариантные множества U и S этого отображения *не* представляют фазовых траекторий такой системы.

Итак, в показанной на рис. III.6 картинке таятся богатые возможности невообразимо сложного поведения. Поскольку точка P принадлежит устойчивой сепаратрисе S , все ее итерации P_1, P_2, \dots, P_n на положительных временах лежат на S и будут сходиться к неподвижной точке H при $n \rightarrow \infty$. С другой стороны, точка P содержится и на неустойчивой сепаратрисе U , и, следовательно, она является образом определенных

точек этой сепаратрисы при последовательных итерациях, соответствующих эволюции системы в отображении Пуанкаре. Иначе эти точки можно рассматривать как итерации P_{-1} , P_{-2} , ..., P_{-n} точки P при обращении времени, причем эти точки

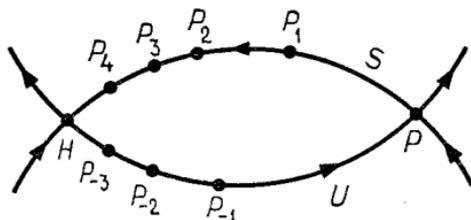


Рис. III.6. Гомоклиническая точка P на секущей поверхности Пуанкаре: устойчивая и неустойчивая сепаратрисы пересекаются в точке P трансверсально. Точки P_1, P_2, \dots представляют собой последовательные итерации точки P при положительном ходе времени и лежат на устойчивом многообразии S , в то время как прообразы P_{-1}, P_{-2}, \dots являются результатами итераций точки P при отрицательном ходе времени и лежат на неустойчивом многообразии U .

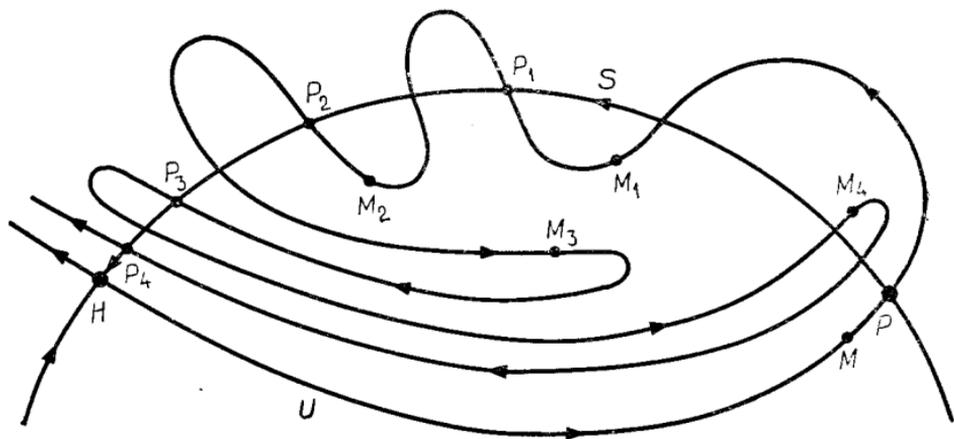


Рис. III.7. Влияние гомоклинической точки на возникновение сложного движения вблизи сепаратрисы. Устойчивое и неустойчивое многообразия гиперболической точки пересекаются бесконечно много раз. Кроме того, в случае консервативной системы сохранение площади подразумевает, что эти сепаратрисы сильно изгибаются в окрестности H . После достаточно большого числа итераций точка M (не являющаяся прообразом точки P), находящаяся исходно по одну сторону устойчивой сепаратрисы, переходит на другую сторону (точка M_4 на рисунке). Таким образом выполняется одна из предпосылок возникновения случайности.

будут сходиться к H при $n \rightarrow \infty$. Однако, поскольку в дальнейшем точка P отображается на P_1, P_2, \dots , точки P_{-1}, P_{-2}, \dots также рано или поздно отобразятся на P_1, P_2, \dots . Таким образом, мы приходим к замечательному выводу о том, что кривая U проходит через точки P_1, P_2, \dots кривой S . Иными словами, точ-

ки P_1, \dots, P_n, \dots также являются гомоклиническими точками, разумеется, как и точки P_{-1}, \dots, P_{-n} . Итак, из единственной гомоклинической точки P мы получили в обе стороны бесконечную последовательность таких точек!

Остается выяснить, как должны выглядеть кривые U и S в такой ситуации. Понятно, что они должны постоянно образовывать волны, изгибы, причем со все возрастающей частотой вблизи гиперболических неподвижных точек. Можно убедиться в том, что образование таких волн сопряжено со все более значительным удалением от неподвижной точки с последующим возвращением в ее окрестность, как это показано на рис. III.7. Это качественно объясняет утверждение, сделанное в конце разд. 3.9, согласно которому при наличии гомоклинических точек динамика приобретает черты сходства с отображением типа подковы (см. разд. 3.8). В свою очередь отсюда следует возможность неустойчивых хаотических движений, что согласуется с результатами численного решения полных нелинейных динамических уравнений, описывающих неинтегрируемые консервативные или диссипативные системы.

РЕКОНСТРУКЦИЯ ДИНАМИКИ СЛОЖНЫХ СИСТЕМ ПО ВРЕМЕННОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ДАННЫХ. ПРИМЕНЕНИЕ К КЛИМАТИЧЕСКОЙ ИЗМЕНЧИВОСТИ

IV.1. Вводные замечания

Экспериментально поведение сложной системы зондируется путем наблюдений в течение какого-то интервала времени над некоторой «существенной» переменной. Пример такой *временной последовательности* приведен на рис. 1.17, где показаны изменения земного объема льда, полученные изотопными измерениями отношения $^{18}\text{O}/^{16}\text{O}$, выполненными на глубоководных осадочных породах, как функция времени. Другой пример — электроэнцефалограмма (ЭЭГ), получающаяся при регистрации суммарного сигнала, обусловленного непрерывной деятельностью нейронов, расположенных в небольших частях коры головного мозга вблизи макушки. Этот сигнал характеризуется сравнительно невысокими частотами порядка 0,5—40 Гц.

На первый взгляд кажется, что временная последовательность по единственной переменной дает довольно ограниченную информацию. В частности, можно было бы возразить в том духе, что это дает как бы «одномерный» взгляд на систему, которая в действительности содержит большое число взаимосвязанных переменных. В данном приложении будет показано, что временная последовательность содержит гораздо более богатую информацию — она носит следы всех других переменных, участвующих в описании динамики системы, а также позволяет безмодельным способом идентифицировать некоторые ключевые особенности соответствующей системы.

Пусть $X_0(t)$ — временная последовательность экспериментально измеренных величин. Разумеется, на самом деле в динамическом процессе должны принимать участие и другие переменные $\{X_k(t)\}$, $k=1, \dots, n-1$. Однако, поскольку мы не опираемся ни на какую конкретную модель, нам хотелось бы восстановить эту динамику исключительно на основе известной последовательности $X_0(t)$. С этой целью рассмотрим фазовое пространство, образованное переменными $\{X_k\}$, $k=0, \dots, n-1$. Как мы видели в гл. 3, некоторое мгновенное состояние системы в этом пространстве соответствует точке, например P , а последовательность таких состояний, проходимая системой во времени, определяет некоторую кривую — фазовую траекторию. С течением времени в системе устанавливается какой-то постоянный режим, если только ее динамика сводится к системе детерминистических уравнений, описывающих диссипативные процессы. Это находит свое отражение в сходимости семейств фа-

зовых траекторий к некоторому подмножеству фазового пространства. Как неоднократно отмечалось ранее, это инвариантное подмножество называется *аттрактором*. Здесь наша основная цель — найти ответы на следующие вопросы:

1. Можно ли идентифицировать аттрактор по данной временной последовательности? Иными словами, могут ли свойства системы, зондируемые с помощью этой временной последовательности, рассматриваться как проявление детерминистической динамики (возможно, очень сложной), или же в ней содержится некоторый неустранимый элемент стохастичности?

2. Если аттрактор существует, то какова его размерность d ? Из гл. 3 мы знаем, что размерность дает нам ценную информацию о динамике системы. Например, если $d=1$, то мы имеем дело с незатухающими *периодическими* колебаниями; если $d=2$, то налицо *квазипериодические* колебания с двумя несоизмеримыми частотами, а если d не целое и превышает 2 (случай фрактального аттрактора, см. разд. 3.10), то можно ожидать, что в системе будут наблюдаться хаотические колебания, характеризующиеся высокой чувствительностью к начальным условиям и существенной непредсказуемостью.

3. Какова минимальная размерность n фазового пространства, содержащего упомянутый аттрактор? От этого зависит минимальное число переменных, которые необходимо использовать для описания соответствующей динамики. Заметим, что d всегда меньше, чем n .

Первый наш шаг будет состоять в определении подходящего набора переменных, образующих фазовое пространство. Для этого удобно развернуть исходную временную последовательность $X_0(t)$ в ряд наборов с последовательно возрастающими сдвигами, определенными как величины, кратные некоторой фиксированной задержке τ ($\tau = m\Delta t$, где m — целое и Δt — интервал между последовательными выборками). Кроме того, выбирая из набора экспериментальных данных N эквидистантных точек, мы приходим к следующему набору дискретных переменных:

$$\begin{aligned} X_0: & X_0(t_1), \dots, X_0(t_N) \\ X_1: & X_0(t_1 + \tau), \dots, X_0(t_N + \tau) \\ & \cdot \\ X_{n-1}: & X_0(t_1 + (n-1)\tau), \dots, X_0(t_N + (n-1)\tau) \end{aligned} \quad (\text{IV.1})$$

При должном выборе τ можно ожидать, что эти переменные будут линейно независимыми, а это все, что требуется для определения фазового пространства. И все эти переменные можно получить из единственной временной последовательности, относящейся к $X_0(t)$, определенной экспериментально. Таким образом, мы видим, что в принципе имеющейся в нашем распоряжении информации достаточно для того, чтобы выйти за рамки одномерного пространства исходной временной последовательности и развернуть динамику системы в многомер-

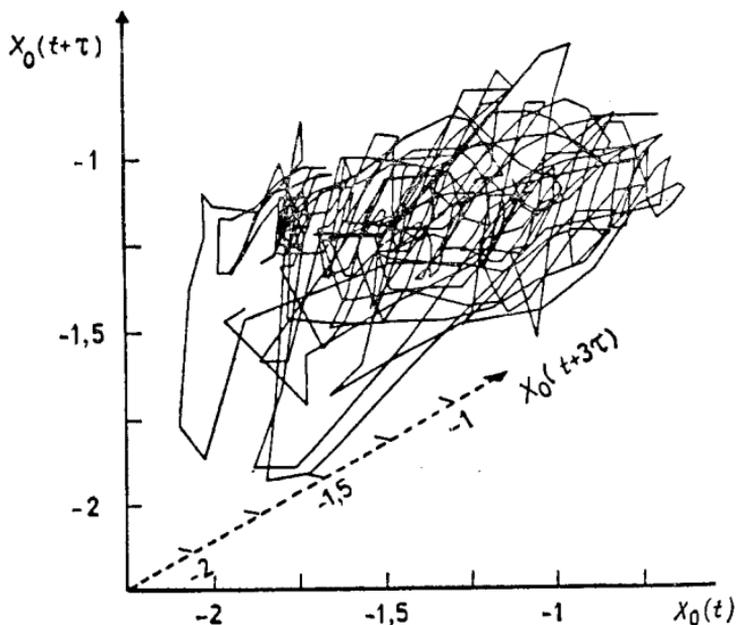


Рис. IV.1. Фазовый портрет, представляющий в трехмерном пространстве эволюцию климата за последний миллион лет. Этот результат получен с помощью около 500 эквидистантных значений X_0 — переменной, изображенной на рис. 1.17. Данные были получены путем анализа изотопа кислорода в глубоководном керне V28—238 из Шеклтона и Опдайка с последующей интерполяцией на основе банка данных из Университета в Лувене. При построении этого портрета принято значение сдвига $\tau=2000$ лет.

ном фазовом пространстве. Уже эта информация позволяет нарисовать *фазовый портрет* системы, или, точнее, его проекцию на малоразмерное подпространство полного фазового пространства. На рис. IV.1 показан трехмерный фазовый портрет климатической эволюции за последний миллион лет. Этот портрет построен на основе данных, представленных на рис. 1.17. Его вид четко свидетельствует о сложном характере движений, лежащих в основе климатической динамики. Однако имеющихся данных слишком мало, чтобы на основе одного лишь этого

рисунка можно было сделать какие-либо выводы. Поэтому нашей следующей задачей будет более четкое исследование сложности динамики с помощью методов теории динамических систем.

IV.2. Теоретические основы анализа данных

Введем следующее векторное обозначение: пусть X_i обозначает точку фазового пространства с координатами $\{X_0(t_i), \dots, X_0(t_i + (n-1)\tau)\}$. Таким образом устанавливается начало отсчета X_i для всех имеющихся данных, и можно вычислить расстояние от этой точки до остающихся $N-1$ точек; $|X_i - X_j|$. Это позволяет сосчитать число точек в фазовом пространстве, отстоящих от X_i на расстояние, не превышающее некоторую заданную величину r . Повторяя этот процесс для всех значений i , можно вычислить следующую величину:

$$C(r) = \frac{1}{N^2} \sum_{\substack{i, j=1 \\ i \neq j}}^N \theta(r - |X_i - X_j|) \quad (IV.2)$$

где θ — функция Хевисайда: $\theta(x) = 0$ при $x < 0$ и $\theta(x) = 1$ при $x > 0$. Отклонение $C(r)$ от нуля служит мерой влияния точки X_i на положение других точек. Поэтому функцию $C(r)$ можно рассматривать как интегральную *корреляционную функцию* аттрактора.

Зафиксируем некоторое малое ε и воспользуемся им в качестве своеобразного метра для зондирования структуры аттрактора. Если последний представляет собой линию, то, очевидно, число пробных точек, расстояние которых до заданной точки не превышает r , должно быть пропорционально r/ε . Если же аттрактор представляет собой поверхность, то число таких точек должно быть пропорционально $(r/\varepsilon)^2$. В более общем случае (см. разд. 3.7), если аттрактор представляет собой d -мерное многообразие, то число точек должно быть пропорционально $(r/\varepsilon)^d$. Поэтому можно ожидать, что при сравнительно малых r функция $C(r)$ должна изменяться как

$$C(r) = r^d \quad (IV.3)$$

Иными словами, размерность аттрактора d дается наклоном зависимости $\ln C(r)$ от $\ln r$ в определенном диапазоне r :

$$\ln C(r) = d \ln r \quad (IV.4)$$

Из этих результатов вытекает следующий алгоритм:

1. Исходя из данной временной последовательности, построить корреляционную функцию (IV.2), рассматривая последо-

вательно возрастающие значения размерности фазового пространства n .

2. Получить наклон d вблизи начала координат, согласно (IV.4), и посмотреть, как этот результат изменяется при возрастании n .

3. Если величина d в зависимости от n выходит на плато выше некоторого относительно небольшого n , то представленная данной временной последовательностью система должна иметь аттрактор. Вышедшую на насыщение величину d следует рассматривать как размерность аттрактора, представленного

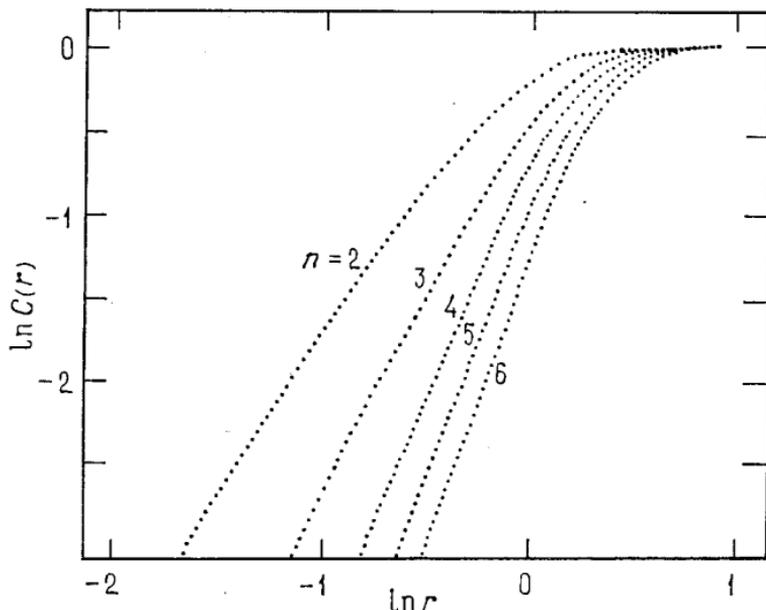


Рис. IV.2. Зависимость корреляционной функции климатического аттрактора от расстояния между точками. Следует отметить наличие протяженной области линейности, из которой можно получить размерность аттрактора. Здесь $\tau = 8000$ лет.

временной последовательностью. Значение n , выше которого наблюдается насыщение, представляет собой минимальное число переменных, необходимых для моделирования поведения, соответствующего данному аттрактору.

IV.3. Климатический аттрактор

Эта процедура была применена к анализу климатических данных, показанных на рис. 1.17. На рис. IV.2 представлена зависимость $\ln C(r)$ от $\ln r$ для значений n от 2 до 6. Из рисунка видно, что, действительно, имеется протяженная область, в которой эта зависимость линейна в соответствии с (IV.4).

На рис. IV.3 кружками показано, что наклон этой зависимости достигает насыщения при $n=4$, что происходит при $d \approx 3,1$. На том же рисунке крестиками показано изменение d с n в случае, когда рассматриваемый сигнал соответствует гауссовому белому шуму, — здесь тенденция к насыщению не наблюдается. На самом деле в этом случае d оказывается равным n . Это устанавливает четкое различие между динамикой климата и случайным сигналом-шумом и показывает, что эволюция

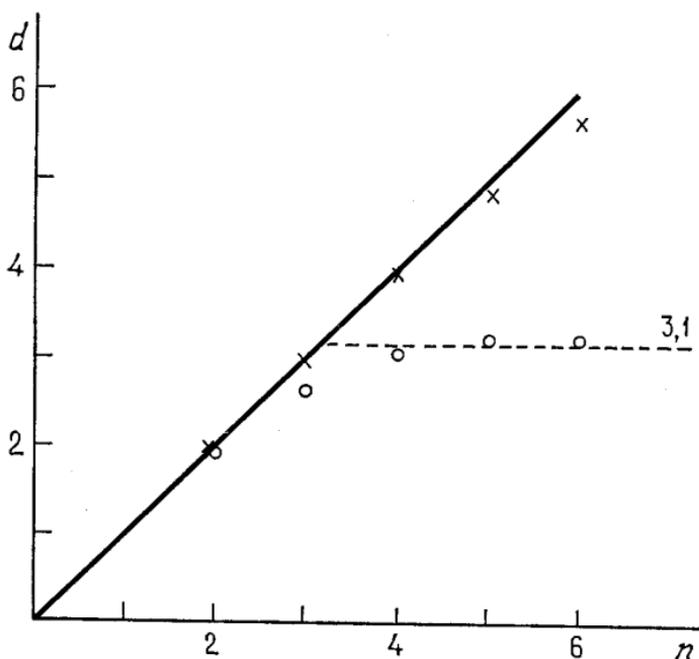


Рис. IV.3. Зависимость размерности d от числа фазовых переменных n для климатического аттрактора (O) и для сигнала, соответствующего белому шуму (X) при $\tau=8000$ лет. Заметим, что в первом случае эта зависимость выходит на насыщение при $d \sim 3,1$, во втором же случае $d \sim n$.

климата за последний миллион лет может рассматриваться как проявление детерминистической динамики, обладающей аттрактором малой размерности. Тот факт, что размерность аттрактора выражается нецелым числом, служит естественным ключом к пониманию внутренней изменчивости и непредсказуемости климатической системы, поскольку обе эти особенности относятся к основным свойствам хаотической динамики. Примечателен тот факт, что эта динамика сводится к ограниченному набору ключевых переменных, хотя на данном этапе установить их конкретную природу не представляется возможным.

Этот вывод, в некотором смысле с неизбежностью следующий из имеющихся данных, полезно сопоставить с результатами моделирования, кратко описанными в разд. 6.4. При этом возникает два основных момента. Во-первых, вследствие неустойчивости движения в режиме хаотического аттрактора становятся возможными усиления слабых внешних сигналов типа малых изменений потока солнечной энергии, обусловленных изменчивостью земной орбиты. И в самом деле, возможность таких усилений была одним из выводов разд. 6.4, хотя это явление приписывалось стохастическим эффектам, непрерывно индуцирующим переходы между сосуществующими состояниями.

Второй момент касается природы стохастического воздействия, которое следует добавить к уравнению энергетического баланса (6.8). Как мы видели в разд. 4.7 и 4.8, хаотический аттрактор способен порождать множество стохастических процессов. По этой причине флуктуации климатической динамики можно рассматривать как проявления хаотического характера самого аттрактора.

IV.4. Выводы и перспективы

Описанный выше подход был также применен к анализу данных электроэнцефалографии, соответствующих различным фазам сна. Здесь снова можно с уверенностью утверждать наличие аттрактора, по крайней мере во время фазы глубокого сна. В зависимости от субъекта размерность аттрактора оказывается где-то между 4 и 4,3. С другой стороны, на определенных стадиях патологической активности (типа эпилепсии) обнаруживается гораздо более низкая размерность, близкая к 2. В связи с замечаниями, сделанными в разд. 4.7 и 4.8, напрашивается сопоставление этого результата с самой способностью мозга генерировать и обрабатывать информацию. С той же точки зрения были проанализированы другие сложные системы, встречающиеся в природе, например быстрые изменения погоды, а также экономическая деятельность.

Не имея возможности подробно обсудить этот вопрос, мы отсылаем читателя к оригинальным публикациям.

ПЕРВИЧНЫЕ НЕОБРАТИМЫЕ ПРОЦЕССЫ

V.1. Введение

Как подчеркивалось в гл. 1, теорию самоорганизации невозможно завершить без учета силы гравитации, самой универсальной из всех известных сил. Основой современной теории гравитации является общая теория относительности, которую Л. Д. Ландау и Е. М. Лифшиц назвали «самой красивой из всех существующих теорий». Эта теория, несомненно, представляет собой существенный отход от ньютоновского описания нашей Вселенной. В самом деле, ньютоновские представления были основаны на дуализме: с одной стороны, пространство-время, с другой — материя (а в действительности, пространство и время считались не связанными друг с другом, равно как и с материей). Основная новизна эйнштейновской общей теории относительности состояла в устранении этого дуализма и в установлении связи между пространством-временем, с одной стороны, и материей — с другой.

Эйнштейновские уравнения поля связывают свойства пространства-времени со свойствами вещества; первые из них описывают кривизну, вторые — давление и плотность энергии. По свойству кривизны возможны различные типы пространства-времени: плоское или искривленное. Само пространство-время становится физическим явлением, и мы вправе задаться вопросом о его происхождении и эволюции. Это — одна из основных целей современной космологии. Поразительно, что еще в 1917 г., всего лишь через год после создания общей теории относительности, Эйнштейн выдвинул первую теоретическую модель космологии. Ему хотелось, чтобы в описании Вселенной время не присутствовало явным образом. Необратимость и стрела времени воспринимались им как нечто субъективное, подлежащее устранению наукой. В соответствии с такими взглядами Эйнштейн представил статическую модель Вселенной.

Последующее развитие космологии — один из наиболее ярких примеров вторжения времени в современную физику. В 1923 г. А. А. Фридман показал, что эйнштейновская Вселенная неустойчива, и предложил теорию расширяющейся Вселенной, впоследствии подтвержденную наблюдениями Эдвина Хаб-

бла (1929 г.). Последний установил, что наблюдаемые с Земли галактики разбегаются со скоростями, пропорциональными расстоянию до них. Так родилась концепция эволюционирующей Вселенной. При этом все же можно было бы предположить, что это чисто геометрическое свойство, не имеющее особого значения для физики.

Однако в 1965 г. А. Пензиас и Р. Вильсон обнаружили существование остаточного излучения черного тела — наша Вселенная полна фотонов, находящихся в тепловом равновесии с черным телом при температуре около 2,7 К (именно из-за малости этой температуры обнаружить такое излучение экспериментально было столь затруднительно). Это излучение возникло чрезвычайно давно, еще на первых стадиях существования Вселенной (отсюда и термин «остаточное»). Оно возникло вследствие космологического расширения, когда температура понизилась достаточно для того, чтобы могли образоваться атомы. На этой стадии излучение отделилось от материи. С момента открытия остаточного излучения черного тела стало известно, что структуре Вселенной присущ определенный дуализм: с одной стороны, имеются массивные частицы (барионы), с другой — частицы, не обладающие массой, т. е. фотоны. По числу фотоны многократно (в сотни миллионов раз) превосходят барионы. Кроме того, поражает то обстоятельство, что связанная с остаточным излучением энтропия составляет подавляющую часть энтропии всей Вселенной. Так было, подсчитано, что при испарении всего вещества Вселенной ее полная энтропия изменилась бы всего лишь на 1%. Отсюда вытекают важные термодинамические следствия. Мы видели, что, согласно классической формулировке второго закона термодинамики, в будущем нас ожидает тепловая смерть (максимум энтропии). Однако теперь стало ясно, что при рождении Вселенной возникло столько энтропии, что можно сказать: тепловая смерть осталась уже позади. Поэтому следует внимательнее проанализировать необратимые процессы, протекавшие на начальных этапах существования Вселенной.

V.2. Стандартная космологическая модель

Как уже отмечалось, современная космология основана на уравнениях Эйнштейна. Вселенная считается однородной и изотропной (это составляет суть так называемого космологического принципа). Для такого предельного случая уравнения Эйнштейна принимают исключительно простую форму. Мы приведем эти уравнения без доказательства для простейшего случая нулевого давления и евклидова пространства, в котором имеются лишь две переменные. Одна из них, иногда назы-

ваемая «радиусом» Вселенной R , определяет пределы, доступные астрономическим наблюдениям. Вторая переменная ρ соответствует плотности энергии. Уравнения Эйнштейна устанавливают связь между ρ и R :

$$\kappa\rho = 3H^2, \quad \dot{\rho} = -3H\rho \quad (\text{V.1, V.2})$$

где $H = \dot{R}/R$ — функция Хаббла. Точкой обозначена производная по космологическому времени. Второе из этих уравнений имеет особенно простой смысл — эволюция Вселенной происходит адиабатически и обратимо, т. е. энтропия сохраняется. В более общем случае имеются три переменные: радиус Вселенной, плотность энергии ρ и давление p . Поэтому необходимо знать еще уравнение состояния, определяемое зависимостью $p(\rho)$. Кроме того, нужно учесть кривизну пространства.

Для однородных и изотропных вселенных возможны лишь три случая, соответствующие отрицательной, нулевой или положительной кривизне. Они приводят соответственно к открытому пространству (отрицательная кривизна), евклидову пространству (нулевая кривизна) и замкнутому пространству (положительная кривизна). Все эти решения начинаются со вселенной, сосредоточенной в точке. Поэтому стандартная модель приводит к первичной особенности — Большому взрыву. Этот вывод был назван Джоном Уилером «величайшим кризисом физики». В самом деле, в чем мог бы быть смысл такой особенности? Если проследить за историей Вселенной в обратном направлении, то придем ли мы к точке, за которую уже невозможно проводить экстраполяцию, за которой прекращается действие законов физики? Действительно ли мы имеем дело с первичной особенностью, или же рождение Вселенной следует рассматривать как результат некоторой неустойчивости, связанной с явлением типа фазового перехода?

V.3. Черные дыры

Каким бы ни был ответ, в физике принято считать, что эти события происходили в так называемую планковскую эру, определяемую тремя фундаментальными константами: скоростью света c , постоянной всемирного тяготения κ и постоянной Планка h . Из этих констант можно скомбинировать планковское время (около 10^{-44} с, очень короткое время), планковскую массу ($\sim 10^{-5}$ г, очень большая величина по сравнению с массой протона) и планковскую длину ($\sim 10^{-33}$ см, очень небольшая). До сих пор у нас нет фундаментальной теории планковской эры. В такой теории основой должна быть квантовая теория гравитации, построение которой встречает большие

трудности. Интересно, однако, что некоторые аспекты квантовой теории гравитации можно установить, не решая основных задач квантовой теории полей. Примером служит теория черных дыр, построенная Стефеном Хокингом и приводящая к ряду замечательных предсказаний.

Черную дыру можно представить как своего рода «мембрану», окружающую пространственно-временную область и не позволяющую свету покинуть эту область. Границы черной дыры называются «горизонтом событий». Свет не может пересечь горизонт событий и покинуть черную дыру (поэтому она и называется «черной»). Однако на самом деле, как показано Стефеном Хокингом и другими, черные дыры не столь уж и «черны»: они излучают, находясь в вакууме, т. е. в таком состоянии пространства-времени, в котором плотность вещества и энергии равны нулю. Более того, они излучают как черное тело при температуре T , зависящей лишь от массы черной дыры. Природа этого теплового излучения связана со взаимодействием между черными дырами и квантовомеханическим вакуумом. Как отмечалось в разд. 5.10, квантовомеханический вакуум всегда флуктуирует. Эти флуктуации приводят к образованию пар частиц. Однако эти частицы являются виртуальными — они существуют лишь короткое время. Одна из этих частиц имеет положительную энергию, другая — отрицательную. Излучение Хокинга обусловлено тем обстоятельством, что черная дыра поглощает частицы с отрицательной энергией. В результате имеется поток положительной энергии от черной дыры, т. е. черная дыра «испаряется». Таким образом, черные дыры приводят к необратимым процессам.

V.4. Роль необратимости

С чем же связано происхождение Вселенной — с особенностью или с неустойчивостью? Часто высказывается идея, согласно которой наша Вселенная существует благодаря тому, что вся имеющаяся энергия принимает одну из двух фундаментальных форм, компенсирующих друг друга, — положительная энергия, связанная с массой, и отрицательная энергия, связанная с гравитацией. Таким образом, с энергетической точки зрения между вакуумом (вселенной Минковского) и материальной вселенной нет различий.

Вспомним случай бенаровской неустойчивости. Цена, уплачиваемая за существование бенаровских ячеек, не выражается в энергетических единицах — в режиме стационарной конвекции энергия сохраняется. Другое дело с производством энтропии. Поэтому напрашивается мысль о том, что рождение Все-

ленной также было «оплачено» производством энтропии. Чтобы выяснить этот вопрос, нужно более тщательно изучить влияние рождения материи на оба принципа термодинамики, равно как и на уравнения Эйнштейна.

Рассмотрим некоторый объем V , содержащий N частиц. В случае замкнутой системы N постоянно. Сохранение внутренней энергии E выражается следующим образом:

$$dE = dQ - pdV \quad (V.3)$$

где dQ — количество тепла, получаемое системой за время dt . Это соотношение можно переписать в виде

$$d\left(\frac{\rho}{n}\right) = dq - pd\left(\frac{1}{n}\right) \quad (V.4)$$

где $\rho = E/V$, $n = N/V$ и $dq = dQ/N$. Поскольку соотношение (V.4) локально, оно справедливо и для открытых систем, в которых N зависит от времени. В этом случае уравнение (V.3) модифицируется — в нем нужно учесть изменение числа частиц:

$$d(\rho V) = dQ - pdV + \frac{h}{n} d(nV) \quad (V.5)$$

где $h = \rho + p$ — энтальпия единицы объема. Для замкнутых систем адиабатические превращения ($dQ = 0$) определены соотношением

$$d(\rho V) + pdV = 0 \quad (V.6)$$

Обобщение на открытые системы дается уравнением

$$d(\rho V) + pdV - \frac{h}{n} d(nV) = 0 \quad (V.7)$$

При таких превращениях «тепло», получаемое системой, целиком обусловлено изменением числа частиц. Применительно к нашему космологическому примеру это изменение связано с передачей энергии от гравитации к веществу. Следовательно, образование вещества действует как источник внутренней энергии.

Обратимся теперь ко второму закону термодинамики (см. разд. 2.5). Как обычно, разложим изменение энтропии на ее приток $d_e S$ и образование $d_i S$:

$$dS = d_e S + d_i S \quad \text{при } d_i S \geq 0 \quad (V.8)$$

Для оценки притока энтропии и ее производства будем исходить из полного дифференциала энтропии:

$$Td(sV) = d(\rho V) + pdV - \mu d(nV) \quad (V.9)$$

где $s = S/V$ и μ — химический потенциал:

$$\mu n = h - Ts \quad (\text{V.10})$$

где $s \geq 0$. Для замкнутых систем и адиабатических превращений соотношение (V.9) приводит к

$$dS = 0 \text{ и } d_i S = 0 \quad (\text{V.11})$$

Рассмотрим влияние образования вещества. Наша система однородна, и, следовательно, можно ожидать, что $d_e S = 0$. В противоположность этому мы должны принять, что создание вещества дает вклад в производство энтропии. Тогда имеем

$$T d_i S = T dS = \frac{h}{n} d(nV) - \mu d(nV) = T \frac{s}{n} d(nV) \geq 0 \quad (\text{V.12})$$

В космологическом контексте это неравенство подразумевает, что пространство-время может создавать вещество, в то время как обратный процесс термодинамически запрещен. Связь между пространством-временем и веществом несимметрична, поскольку образование частиц, происходящее за счет гравитационной энергии, оказывается необратимым процессом. Микроскопическая интерпретация этого процесса, в котором пространство-время преобразуется в вещество, будет вкратце обсуждена в конце приложения.

Соотношение (V.7) можно записать несколькими эквивалентными способами, например

$$\dot{\rho} = \frac{h}{n} \dot{n}, \quad p = \frac{n\dot{\rho} - \rho\dot{n}}{\dot{n}} \quad (\text{V.13, V.14})$$

Интересно, что давление p определяется скоростями образования энергии ($\dot{\rho}$) и частиц (\dot{n}). В качестве примера отметим, что из $\rho = Mn$ следует $p = 0$ и далее из $\rho = \alpha T^4$, $n = bT^3$ следует $p = \rho/3$.

Чтобы обсудить уравнения Эйнштейна для открытых систем, полезно переписать соотношение (V.7) в форме, близкой к (V.6), а именно:

$$d(\rho V) = -(\rho + p_c) dV \quad (\text{V.15})$$

где

$$p_c = -\frac{h}{n} \frac{d(nV)}{dV} = -\frac{\rho + p}{n} \frac{d(nV)}{dV} \quad (\text{V.16})$$

соответствует дополнительному давлению, обусловленному образованием вещества. Согласно соотношению (V.12), величина p_c всегда отрицательна. Теперь можно выяснить, что же изменилось в уравнениях Эйнштейна в связи с образованием вещества. Как уже отмечалось, уравнения Эйнштейна связывают

свойства пространства-времени (кривизну) с плотностью энергии и давлением. Чтобы учесть образование вещества, в уравнения Эйнштейна нужно ввести соответствующее давление \tilde{p} , представленное в виде суммы

$$\tilde{p} = p + p_c \quad (\text{V.17})$$

где p — обычное давление и p_c — давление, обусловленное образованием вещества.

Таким образом мы получаем целый класс новых космологических моделей, в которых переменными являются R , r , ρ и n . Рассмотрим следующий простой случай:

$$\rho = Mn \quad (\text{V.18})$$

из которого для открытых систем следует $p=0$. У нас остаются две переменные: R и n . Одно из уравнений соответствует (V.1). Нужно еще одно уравнение, связывающее образование частиц с функцией Хаббла H . Примем в качестве объема V величину

$$V \sim R^3 \quad (\text{V.19})$$

Простейшее кинетическое уравнение, удовлетворяющее неравенству (V.12), имеет вид

$$\frac{1}{R^3} \frac{d(nR^3)}{dt} = \alpha H^2 \geq 0 \quad (\text{V.20})$$

где $\alpha \geq 0$. При $\alpha=0$ мы возвращаемся к обычной, стандартной модели, содержащей особенность типа Большого взрыва. Однако при $\alpha \neq 0$ ситуация резко меняется. Можно решить уравнения (V.1) и (V.20) и показать, что при любых начальных условиях система стремится к *пространству де Ситтера*, характеризующему постоянными во времени величинами n и H . Этот случай соответствует экспоненциальному расширению Вселенной, поскольку $H = \dot{R}/R$. Полученный результат совершенно не зависит от вида уравнения (V.20). Таким образом, учет образования вещества делает Большой взрыв структурно-неустойчивым, и даже малейшая скорость образования вещества приводит к экспоненциально расширяющейся Вселенной. На более микроскопическом уровне образование вещества из вакуума можно рассматривать как проявление «неустойчивости» вакуума по отношению к флуктуациям. Как уже упоминалось, квантовомеханический вакуум связан с возникновением виртуальных частиц. Если в данном объеме появляется достаточно большое число этих частиц (благодаря флуктуациям), то это может привести к образованию временно существующей черной дыры, которая затем будет разрушаться по механизму Хокинга. Та-

ким путем виртуальные частицы превращаются в реальные. Следовательно, космологическую историю можно считать как бы состоящей из трех этапов. Первый этап — нестабильность вакуума, с неизбежностью порождающая вселенную де Ситтера. Второй этап — процесс «испарения» черной дыры, которая, по всей видимости, тормозит образование новой черной дыры в том же месте. Наконец, третий этап — переход от фазы де Ситтера к фазе современной Вселенной, адиабатически расширяющейся в соответствии со стандартной моделью.

Качественный космологический анализ этой картины позволяет в виде основного параметра ввести массу черной дыры m_{bh} , действующей как промежуточное звено между виртуальными и реальными частицами. При $m_{bh} \sim 100 m_{\text{планск}}$ мы получаем превосходное согласие с известными основными термодинамическими характеристиками Вселенной, такими, как отношение числа барионов к числу фотонов. Поскольку время жизни черной дыры очень мало, два первых этапа также соответствуют чрезвычайно малому времени порядка 10^{-37} с, что в 10^7 раз больше времени Планка.

Интересно отметить, что квантовомеханический анализ этого вопроса дает аналогичную по порядку величины массу черных дыр в процессе образования Вселенной. На языке нашей модели мы имели бы своего рода фазовый переход, отделяющий вещество от гравитации. В первичном вакууме эти две составляющие Вселенной очень тесно переплетены. В современной космологии сильно увязываются микроскопические и макроскопические аспекты. Мы не смогли бы описать ранние стадии формирования Вселенной без постоянной Планка, «управляющей» микромиром. Эта ситуация ярко отражает так называемый холистический характер Вселенной. Разумеется, остается множество вопросов, и создание материи из ничего — еще не самый острый из них: уже вакуум Минковского является вполне определенной средой, характеризующейся значениями трех фундаментальных постоянных — скоростью света c , гравитационной постоянной κ и постоянной Планка h . Можно еще спросить: откуда взялись эти постоянные? Действительно ли эти величины постоянны во времени, или же они сами изменяются? В настоящий момент найти теоретические ответы на эти вопросы не удается.

В начале данного приложения мы отмечали дуализм ньютоновского мира. Была подчеркнута объединяющая роль эйнштейновской общей теории относительности. Уравнения Эйнштейна можно рассматривать примерно так же, как инженер смотрит на эквивалентность между теплом и механической работой: уравнения Эйнштейна выражают своего рода эквивалентность между пространством-временем и веществом. Одна-

ко, как и в случае тепла и работы, здесь нет полной эквивалентности. Пространство-время выступает в роли некоторой «благородной формы», из которой материя возникает в необратимом процессе, сопровождаемом производством громадного количества энтропии. И одна из тем, периодически рассматриваемых в этой книге, — связь производства энтропии как с упорядоченностью, так и с беспорядком в соответствии с ярко выраженным дуализмом нашей Вселенной.

Общие вопросы

Физические аспекты

Изложение физических представлений, лежащих в основе термодинамики необратимых процессов, самоорганизации и сложных явлений, можно найти в следующих работах:

- Babloyantz A.*, *Molecules, dynamics and life*, Wiley, N. Y. (1986).
Careri G., *Order and disorder in matter*, Benjamin Cummings, Menlo Park, Calif. (1984).
Ebeling W., Feistel R., *Physik der selbstorganisation und evolution*, Akademie-Verlag, Berlin (1982).
Гленсдорф П., Пригожин И. Термодинамическая теория структуры, устойчивости и флуктуаций. — М.: Мир, 1973.
Хакен Г. Синергетика. — М.: Мир, 1980.
Nicolis J. S., *Dynamics of hierarchical systems*, Springer, Berlin (1986).
Николис Г., Пригожин И. Самоорганизация в неравновесных системах. — М.: Мир, 1979.
Peacocke A., *The physical chemistry of biological organization*, Clarendon, Oxford (1983).
Пригожин И. Введение в термодинамику необратимых процессов. — М.: ИЛ, 1960.
Пригожин И. От существующего к возникающему. — М.: Наука, 1985.
Thompson J. M. T., *Instabilities and catastrophes in science and engineering*, Wiley, Chichester (1982).

Математические аспекты

Изложение фрактальной геометрии, теории нелинейных динамических систем и хаотической динамики можно найти в следующих работах:

- Арнольд В.* Дополнительные главы теории обыкновенных дифференциальных уравнений. — М.: Мир, 1980.
Bergé P., Pomeau Y., Vidal C., *L'ordre dans le chaos*, Herman, Paris (1984).
Guckenheimer J., Holmes Ph., *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Springer, Berlin (1983).
Lichtenberg A., Lieberman M., *Regular and stochastic motion*, Springer, Berlin (1983).
Mandelbrot B., *Fractals: form, chance, dimension*, Freeman, San Francisco (1977).
Sattinger D., *Topics in stability and bifurcation theory*, Springer, Berlin (1973).
Шустер Г. Детерминированный хаос. — М.: Мир, 1988.

Глава 1

Тепловая конвекция

- Chandrasekar S.*, Hydrodynamic and hydromagnetic stability. Oxford University Press, Oxford (1961).
Koschmieder E. L., in: Order and fluctuations in equilibrium and nonequilibrium statistical mechanics (G. Nocolis, G. Dewel and J. W. Turner, eds.), Wiley, New York (1981).
Velarde M., Normant C., Scientific American, **243** (1), 78 (1980).

Явления самоорганизации в химии, на границах раздела фаз, в науке о материалах и в геологии

- Колесания и бегущие волны в химических системах. Под ред. Р. Филда, М. Бургер. — М.: Мир, 1988.
Müller S., Plesser T., Hess B., Naturwissenschaften, **73**, 165 (1986).
Nocolis G., Baras F. (eds.), Chemical instabilities, D. Reidel, Dordrecht (1984).
Nocolis G., Portnow J., Chem. Rev., **73**, 365 (1973).
Pacault A., Vidal C. (eds.), Far from equilibrium, Springer, Berlin (1978).
Vidal C., Pacault A. (eds.), Nonlinear phenomena in chemical dynamics, Springer, Berlin (1981).
Vidal C., Pacault A. (eds.), Nonequilibrium dynamics in chemical systems, Springer, Berlin (1984).
Winfree A., Science, **175**, 634 (1972).
Жаботинский А. М. — Биофизика, **9**, 306 (1964). (Жаботинский А. М. Концентрационные автоколебания. — М.: Наука, 1974. — доп. перев.).

Квантовая оптика

- Abraham E., Smith S.*, Rep. Progr. Phys., **45**, 815 (1982).
Bonifacio R., Lugiato L., Opt. Commun., **19**, 172 (1976).
Gibbs H., McCall S., Venkatesan T., Phys. Rev. Lett., **36**, 113 (1976).
Хакен Г. Синергетика. — М.: Мир, 1980.

Управление процессами развития

- Goldbeter A., Segel L.*, Proc. Nat. Acad. Sci. USA, **74**, 1543 (1977).
Kauffman S., Shymko R., Trabert E., Science, **199**, 259 (1978).
Segel L., Modeling dynamical phenomena in molecular and cellular biology, Cambridge University Press, Cambridge, England (1984).
Sussman M., Growth and development, Prentice-Hall, Englewood Cliffs, New Jersey (1964).
Turing A., Phil. Trans. Roy. Soc., **B237**, 37 (1952).
Walpert L., Scientific American, **239**, 124 (1978).

Климатическая изменчивость

- Berger A.* (ed.), Climatic variations and variability: facts and theories, Reidel, Dordrecht (1981).
Lamb H., Climate: present, past and future, Methuen, London (1977).
Schneider S., The genesis strategy, Plenum, New York (1976).

Фазовые переходы, космология и нарушение симметрии

Georgi H., Pour la Science, 44, 68 (1981).

Guth A., Steinhardt P., L'univers inflatoire, Pour la Science, 81, 86 (1984).

Hawkins S., A brief history of time, from the big bang to black holes, Bantam, New York (1988).

Ландау Л. Д., Лифшиц Е. М. Статистическая физика. — М.: Наука, 1964.

Ма Ш. Современная теория критических явлений. — М.: Мир, 1980.

Stanley H., Introduction to phase transitions and critical phenomena, Clarendon, Oxford (1971).

Weinberg S., Relativity and cosmology, Wiley, New York (1972).

Weinberg S., The first three minutes of the universe, Basic Books, New York (1977).

Алгоритмическая сложность

Chaitin G., Scientific American, 233 (5), 47 (1975).

Nicolis J. S., Rep. Progr. Phys., 49, 1109 (1986).

Глава 2*Ньютоновский и гамильтоновский формализмы в классической механике*

Goldstein H., Classical mechanics, Addison-Wesley, Reading, Mass. (1959).

Термодинамическое описание диссипативных систем

Де Гроот С., Мазур П. Неравновесная термодинамика. — М.: Мир, 1964.

Katchalsky A., Curran P., Nonequilibrium thermodynamics in biophysics, Harvard U. P., Cambridge, Mass. (1965).

Onsager L., Phys. Rev., 37, 405 (1931); 38, 2265 (1931).

Пригожин И. Введение в термодинамику необратимых процессов. — М.: ИЛ, 1960.

Простые модели нелинейных систем

Prigogine I., Lefever R., J. Chem. Phys., 48, 1695 (1968).

Schlögl F., Z. Phys., 253, 147 (1972).

Volterra V., Leçons sur la théorie mathématique de la lutte pour la vie, Gauthier-Villars, Paris (1936).

Основные определения теории устойчивости и бифуркаций

Андронов А. А., Витт А. А., Хайкин С. Э. Теория колебаний. — М.: Физматгиз, 1959.

Распределение Пуассона

Феллер В. Введение в теорию вероятностей и ее приложения. В 2-х томах. — М.: Мир, 1984.

Глава 3

Описание систем с помощью фазового пространства, уравнение Лиувилля

Гиббс Дж. У. Термодинамика. Статистическая механика. — М.: Наука, 1982.
Пригожин И. Неравновесная статистическая механика. — М.: Мир, 1964.

Устойчивые и неустойчивые движения в консервативных системах

Арнольд В. Математические методы классической механики. — М.: Мир, 1976.
Arnold V., Avez A., Ergodic problems of classical mechanics, Benjamin, New York (1968).
Birkhoff G., Dynamical systems, AMS Publ., Providence, Rhode Island (1927).
Колмогоров А. Н. — Докл. АН СССР, 98, 527 (1954).
Lichtenberg A., Lieberman M., Regular and stochastic motion, Springer, Berlin (1983).
Moser J., Stable and random motions in dynamical systems, Princeton Univ. Press, Princeton, N. J. (1973).
Poincaré H., Les méthodes nouvelles de la mécanique céleste, Gauthier-Villars, Paris (1899).

Неустойчивости и бифуркации в диссипативных системах

Андронов А. А., Витт А. А., Хайкин С. Э. Теория колебаний. — М.: Физматгиз, 1959.
Арнольд В. Дополнительные главы теории обыкновенных дифференциальных уравнений. — М.: Мир, 1980.
Guckenheimer J., Holmes Ph., Nonlinear oscillations, dynamical systems, and bifurcations of vector fields, Springer, Berlin (1983).
Хакем Г. Синергетика. — М.: Мир, 1980.
Helleman R., Rössler O. (eds.), Bifurcation theory and applications in scientific disciplines, Ann. N. Y. Acad. Sci., 316 (1979).
Iooss G., Joseph D., Elementary stability and bifurcation theory, Springer, Berlin (1981).
Николис Г., Пригожин И. Самоорганизация в неравновесных системах. — М.: Мир, 1979.
Sattinger D., Topics in stability and bifurcation theory, Springer, Berlin (1973).
Thom R., Stabilité structurelle et morphogénèse, Benjamin, Reading, Mass. (1972).
Wasow W., Asymptotic expansions for ordinary differential equations, Wiley-Interscience, New York (1965).

Хаотическая динамика, фракталы

Arneodo A., Couillet P., Tresser C., Commun. Math. Phys., 79, 573 (1981); J. Stat. Phys., 27, 171 (1982).
Bergé P., Pomeau Y., Vidal C., L'ordre dans le chaos, Herman, Paris (1984).
Collet P., Eckmann J. P., Iterated maps on the interval as dynamical systems, Birkhäuser, Basel (1980).
Feigenbaum M., J. Stat. Phys., 19, 25 (1978).
Ford J., Physics today, April issue, 40 (1983).
Gaspard P., Nicolis G., J. Stat. Phys., 31, 499 (1983).
Gumowski I., Mira C., Recurrences and discrete dynamical systems, Springer, Berlin (1980).
Lorenz E., J. Atmos. Sci., 20, 130 (1963).
Mandelbrot B., Fractals: form, chance, dimension, Freeman, San Francisco (1977).

- Nicolis G., Nicolis C.*, Phys. Rev., **A38**, 437 (1988).
Peitgen H., Richter P., The beauty of fractals, Springer, Berlin (1986).
Pomeau Y., Manneville P., Commun. Math. Phys., **74**, 189 (1980).
Roux J. C., Simoyi R., Swinney H., Physica, **8D**, 257 (1983).
Rössler O., Ann. N. Y. Acad. Sci., **316**, 376 (1979).
Ruelle D., Takens F., Commun. Math. Phys., **20**, 167 (1971).
Шустер Г. Детерминированный хаос. — М.: Мир, 1988.
Сильников Л. — Мат. докл., **6**, 163 (1965); Мат. сборник СССР, **10**, 91 (1970).
Smale S., Bull. Am. Math. Soc., **73**, 747 (1967).
Sparrow C., The Lorenz equations, Springer, Berlin (1982).

Образование упорядоченных картин и асимметрия

- Auchmuty J. F. G., Nicolis G.*, Bull. Math. Biol., **37**, 323 (1975).
Cailliois R., La dissymétrie, Gallimard, Paris (1973).
Gardner M., The ambidextrous universe, Charles Scribner's Sons, New York (1979).
Kondepudi D., Nelson G., Nature, **314**, 438 (1985).
Kuramoto Y., Chemical oscillations, waves, and turbulence, Springer, Berlin (1984).
Luria S., Thirty-six lectures in biology, MIT Press, Cambridge, Mass. (1975).
Nicolis G., Auchmuty J. F. G., Proc. Nat. Acad. Sci. USA, **71**, 2748 (1974).
Nicolis G., Prigogine I., Proc. Nat. Acad. Sci. USA, **78**, 659 (1981).
Turing A., Phil. Trans. Roy. Soc., **B237**, 37 (1952).

Клеточные автоматы

- Huberman B., Hogg T.*, Phys. Rev. Lett., **52**, 1048 (1984).
Thomas R., Adv. Chem. Phys., **55**, 247 (1984).
Wolfram S., Rev. Mod. Phys., **55**, 601 (1983).
Wolfram S. (ed.), Theory and applications of cellular automata, World Scientific, Singapore (1987).

Глава 4

Теория вероятностей и стохастические процессы

- Феллер В.* Введение в теорию вероятностей и ее приложения. В 2-х томах. — М.: Мир, 1984, т. 1.

Флуктуации вблизи термодинамического равновесия

- Ландау Л. Д., Лифшиц Е. М.* Статистическая физика. — М.: Наука, 1976.

Внутренние флуктуации и внешний шум в физике, химии и биологии

- Gardiner C.*, Handbook of stochastic methods, Springer, Berlin (1983).
Хакен Г. Синергетика. — М.: Мир, 1980.
Хорстэмке В., Лефевр Р. Индуцированные шумом переходы: теория и применение в физике, химии и биологии. — М.: Мир, 1987.
Николис Г., Пригожин И. Самоорганизация в неравновесных системах. — М.: Мир, 1979.
Van Kampen N., Stochastic processes in physics and chemistry, North-Holland, Amsterdam (1981).

Энтропия и необратимость в марковских процессах. Полугруппы

- Феллер В. Введение в теорию вероятностей и ее применения. В 2-х томах. — М.: Мир, 1984, т. 2.
 Schlögl F., in: Cooperative effects (H. Haken, ed.), North-Holland, Amsterdam (1974).
 Schnakenberg J., Rev. Mod. Phys., **48**, 571 (1976).

Пространственные корреляции и критическое поведение флуктуаций

- Gardiner C., McNeil K., Walls D., Matheson I., J. Stat. Phys., **14**, 307 (1976).
 Lemarchand H., Nicolis G., J. Stat. Phys., **37**, 609 (1984).
 Ма III. Современная теория критических явлений. — М.: Мир, 1980.
 Malek-Mansour M., Van den Broeck C., Nicolis G., Turner J. W., Ann. Phys., **131**, 283 (1981).
 Mareschal M., Kestemont E., Phys. Rev., **A30**, 1158 (1984).
 Nicolis G., Malek-Mansour M., Phys. Rev., **A29**, 2845 (1984).
 Nicolis G., Turner J. W., Physica, **89A**, 326 (1977).
 Spohn H., J. Phys., **A16**, 4275 (1983).
 Tremblay A. M., Arai M., Siggia E., Phys. Rev., **A23**, 1451 (1981).
 Walgraef D., Dewel G., Borckmans P., Adv. Chem. Phys., **49**, 311 (1982).

Поведение флуктуаций во времени. Зародышеобразование

- Borgis D., Moreau M., J. Stat. Phys., **37**, 631 (1984).
 Frankowicz M., Gudowska-Nowak E., Physica, **116A**, 331 (1982).
 Hanusse P., Doctorat ès Sciences, University of Bordeaux I (1976).
 Nicolis G., Prigogine I., Proc. Nat. Acad. Sci. USA, **68**, 2102 (1971).
 Suzuki M., Adv. Chem. Phys., **46**, 195 (1981).
 Van Kampen N., J. Stat. Phys., **17**, 71 (1977).

Влияние неоднородных флуктуаций, возникающих при неполном смешивании

- Kumpinsky E., Epstein I., J. Chem. Phys., **82**, 53 (1985).
 Menzinger M., Giraudi A., J. Phys. Chem., **91**, 4391 (1987).
 Puhl A., Nicolis G., Chem. Eng. Sci., **41**, 3111 (1986).
 Puhl A., Altarès V., Nicolis G., Phys. Rev., **A37**, 3039 (1988).
 Roux J. C., De Kepper P., Boissonade J., Phys. Lett., **97A**, 168 (1983).
 Zwietering Th., Chem. Eng. Sci., **39**, 1765 (1984).

Бимодальность в переходных процессах

- Arimondo E., Dangoisse D., Fronzoni L., Europh. Lett., **4**, 287 (1987).
 Baras F., Nicolis G., Malek-Mansour M., Turner J. W., J. Stat. Phys., **32**, 1 (1983).
 Broggi G., Lugiato L., Phil. Trans. Roy. Soc., **A313**, 425 (1984).
 Frankowicz M., Nicolis G., J. Stat. Phys., **33**, 595 (1983).
 Lange W., Mitschke F., Deserno R., Mlynek J., Phys. Rev., **A32**, 1271 (1985).
 Nicolis G., Baras F., Malek-Mansour M., in: Nonequilibrium dynamics of chemical systems (C. Vidal and A. Pacault, eds.), Springer, Berlin (1984).
 Nicolis G., Baras F., J. Stat. Phys., **48**, 1070 (1987).

Чувствительность и отбор

- Kondepudi D., Prigogine I., Physica, **107A**, 1 (1981).
 Mathowsky B., Reiss E., SIAM J. Appl. Math., **33**, 230 (1977).
 Nicolis G., Prigogine I., Proc. Nat. Acad. Sci. USA, **78**, 659 (1981).

Молекулярная хиральность и слабые взаимодействия

- Kondepudi D., Nelson G.*, Nature, 314, 438 (1985).
Mason S., Nature, 311, 19 (1984).

Информация и сложные явления

- Chaitin G.*, Scientific American, 233(5), 47 (1975).
Chaitin G., Algorithmic information theory, Cambridge University Press, Cambridge, England (1987).
Айген М., Шустер П. Гиперцикл. Принципы самоорганизации макромолекул. — М.: Мир, 1982.
Khinchin A., Mathematical foundations of information theory, Dover, New York (1957). [На русском языке см.: Об основных теоремах теории информации. — УМН, 11:1 (67) (1956).]
Nicolis G., Nicolis C., Phys. Rev., A38, 437 (1988).
Nicolis G., Rao G., Rao S., Nicolis C., in: Structure, coherence and chaos in dynamical systems (P. Christiansen, R. Parmentier, eds.), Manchester Univ. Press (1989).
Nicolis J. S., Dynamics of hierarchical systems, Springer, Berlin (1986).
Shaw R., Z. Naturf., 30a, 80 (1981).
Шредингер Э. Что такое жизнь? С точки зрения физика. — М.: Атомиздат, 1972.

Глава 5

Модели твердых шариков

- Vukobratovich L., Sinai Y.*, Commun. Math. Phys., 78, 247 (1980); 79, 479 (1981).
Синай Я. Г. — В кн.: Крылов Н. С. Работы по обоснованию статистической физики. — М. — Л., Изд-во Акад. Наук СССР, 1950.

Преобразование пекаря и сдвиги Бернулли

- Arnold V., Avez A.*, Ergodic problems of classical mechanics, Benjamin, New York (1968).
Lebowitz J., Penrose O., Physics today, February issue, 23 (1973).
Ornstein D., in: Dynamical systems: theory and applications (J. Moser, ed.), Springer, Berlin (1975).
Penrose O., Foundations of statistical mechanics, Pergamon, Oxford (1970); Rep. Prog. Phys., 42, 1937 (1979).
Ehrenfest P., Ehrenfest T., The conceptual foundations of the statistical approach in mechanics, Cornell Univ. Press, Ithaca, New York (1959).
Гиббс Дж. У. Термодинамика. Статистическая механика. — М.: Наука, 1982.

Временной оператор и \wedge -преобразование

- Elskens Y., Prigogine I.*, Proc. Nat. Acad. Sci. USA, 83, 5756 (1986).
Martinez S., Tirapegui E., J. Stat. Phys., 37, 173 (1984); Phys. Lett., 110A, 81 (1985).
Misra B., Proc. Nat. Acad. Sci. USA, 75, 1627 (1978).
Misra B., Prigogine I., Courbage M., Physica, A98, 1 (1979).

Взаимодействие света с веществом

- Panofsky W., Phillips M.*, Classical electricity and magnetism, Addison-Wesley, Reading, Mass. (1956).
Petrosky T., Prigogine I., Physica, **147A**, 439 (1988).
Prigogine I., Petrosky T., Physica, **147A**, 461 (1988).

Кинетическая теория. Столкновительный оператор

- Балеску Р.* Равновесная и неравновесная статистическая механика. В 2-х томах. — М.: Мир, 1978.
Пригожин И. Неравновесная статистическая механика. — М.: Мир, 1964.
Résibois P., De Leener M., Kinetic theory of dense fluids, Wiley, New York (1977).

Теоретико-познавательные аспекты

- Born M.*, Natural philosophy of cause and chance, Clarendon, Oxford (1949).
Prigogine I., Stengers I., La nouvelle alliance, Gallimard, Paris (1979).
Rosenfeld L., Selected papers (R. Cohen and J. Stachel, eds.), Reidel, Dordrecht (1979).
Синай Я. Г. Случайность неслучайного. — Природа, № 3, 72 (1981).

Глава 6

Наука о материалах

- Balakrishnan V., Bottani C.* (eds.), Mechanical properties and behavior of solids—plastic instabilities, World Scientific, Scientific, Singapore (1986).
Bass M. (ed.), Laser materials processing, North-Holland, Amsterdam (1983).
Bilgram J., Böni P., in: Chemical instabilities (G. Nicolis and F. Baras, eds.), Reidel, Dordrecht (1984).
Bottani C., Caglioti G., Physica Scripta, **T1**, 65 (1982).
Langer J., Rev. Mod. Phys., **52**, 1 (1980).
Picraux S., Choyke W. (eds.), Metastable materials formation by ion implantation, North-Holland, Amsterdam (1982).
Walgraef D., Aifantis E., J. Appl. Phys., **58**, 688 (1985).
Walgraef D. (ed.), Patterns, defects and microstructures, M. Nijhoff, Dordrecht (1987).

Клеточная динамика

- Les défenses du corps humain, La Recherche, Special issue, May 1986.
Garay R., Lefever R., J. Theor. Biol., **73**, 417 (1978).
Hiernaux J., Lefever R., Uyttenhove G., Boon T., in: Paradoxes in Immunology (G. Hoffman, J. Levy and G. Nepom, eds.), CRC Press, Boca Raton, Florida (1986).

Моделирование климатической изменчивости

- Benzi R., Parisi G., Sutera A., Vulpiani A.*, Tellus, **34**, 10 (1982).
Berger A. (ed.), Climatic variations and variability: facts and theories, Reidel, Dordrecht (1981).

- Berger A. et al.* (eds.), *Milankovitch and climate*, Reidel, Dordrecht (1984).
Budyko M., *Climate and life*, Academic Press, New York (1974).
Gates W., in: *Climatic variations and variability: facts and theories* (A. Berger, ed.), Reidel, Dordrecht (1981).
Ghil et al. (eds.), *Turbulence and predictability in geophysical fluid dynamics and climate dynamics*, North-Holland, Amsterdam (1985).
Haltiner G., *Numerical weather prediction*, Wiley, New York (1971).
Imbrie J., Imbrie K., *Ice ages*, Enslow Publ., Hillside, N. J. (1979).
Källén E., Crafoord C., Ghil M., *J. Atmos. Sci.*, **36**, 2292 (1979).
Lorenz E., *Tellus*, **16**, 1 (1964).
Nicolis C., *Tellus*, **34**, 1 (1982).
Nicolis C., *Tellus*, **36A**, 1 (1984); **36A**, 217 (1984).
Nicolis C., Nicolis G. (eds.), *Irreversible phenomena and dynamical systems analysis in geosciences*, Reidel, Dordrecht (1986).
Saltzman B., *Adv. Geophys.*, **20**, 184 (1978); **25**, 173 (1983).
Schneider S., Dickinson R., *Rev. Geophys. Space Phys.*, **12**, 447 (1974).

Социальные явления

- Allen P.*, *Environment and planning*, **B9**, 95 (1982).
Allen P., Engelen G., Sanglier M., in: *The praxis and management of complexity*. Un. Nations Univ. Press, Tokyo (1986).
Allen P., McGlade J., *Can J. Fish Aqu. Sci.*, **43**, 1187 (1986).
Deneubourg J. L., Pasteels J., Verhaege J. C., *J. Theor. Biol.*, **105**, 259 (1983).
Grassé P., *Insectes sociaux*, **6**, 127 (1959).
May R., *Model ecosystems*, Princeton Univ. Press, Princeton, N. J. (1973).
Montroll E., Badger W., *Quantitative aspects of social phenomena*, Gordon and Breach, London (1974).
Prigogine I., Herman R., *Kinetic theory of vehicular traffic*, Elsevier, New York (1971).
Prigogine I., Allen P., Herman R., in: *Goals for a global community* (E. Laszlo and J. Bierman, eds.), Pergamon Press, New York (1977).

Приложения

Линейный анализ устойчивости и бифуркационный анализ

- Guckenheimer J., Holmes Ph.*, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Springer, Berlin (1983).
Iooss G., Joseph D., *Elementary stability and bifurcation theory*, Springer, Berlin (1981).
Николис Г., Пригожин И. *Самоорганизация в неравновесных системах*. — М.: Мир, 1979.
Nicolis G., in: *Nonlinear problems in physics and biology* (R. Enns et al., eds.), Plenum, New York (1981).
Sattinger D., *Topics in stability and bifurcation theory*, Springer, Berlin (1973).

Ляпуновские показатели

- Collet P., Eckmann J. P.*, *Iterated maps on the interval as dynamical systems*, Birkhäuser, Basel (1980).
Lichtenberg A., Leiberman M., *Regular and stochastic motion*, Springer, Berlin (1983).
Оселедец В. — *Тр. Моск. матем. общ-ва*, **19**, 197 (1968).
Ruelle D., *Ann. New York Acad. Sci.*, **316**, 408 (1979).

Возмущение резонансных движений

Birkhoff G., Mem. Pont. Acad. Sci. Novi Lyncaei, **1**, 85 (1935).

Chirikov B., Phys. Rev., **52**, 463 (1979).

Moser J., Stable and random motions in dynamical systems, Princeton Univ. Press, Princeton, N. J. (1973).

Гомоклинические точки

Ekeland I., Le calcul, l'imprévu, Seuil, Paris (1984).

Newhouse S., Topology, **13**, 9 (1974); Publ. Math. IHES, **50**, 101 (1979).

Newhouse S., in: Dynamical systems (J. Guckenheimer, J. Moser and Newhouse S., eds.), Birkhäuser, Basel (1980).

Nitecki Z., Differentiable dynamics, MIT Press, Cambridge, Mass. (1971).

Poincaré H., Les méthodes nouvelles de la mécanique céleste, Gauthier-Villars, Paris (1899).

Smale S., Bull. Am. Math. Soc., **73**, 747 (1967).

Восстановление динамики по временной последовательности данных. Основные ссылки

Eckmann J. P., Ruelle D., Rev. Mod. Phys., **57**, 617 (1985).

Grassberger P., Procaccia I., Phys. Rev. Lett., **50**, 346 (1983).

Packard N., Crutchfield J., Farmer J., Shaw R., Phys. Rev. Lett., **45**, 712 (1980).

Takens F., in: Lecture Notes in Math., Vol. 898, Springer, Berlin (1981).

Wolf A., Swift J., Swinney H., Vastano H., Physica, **D16**, 285 (1985).

Применения в гидродинамике и химии

Brandstätter A., Swift J., Swinney H., Wolf A., Farmer J., Jen E., Crutchfield J., Phys. Rev. Lett., **51**, 1442 (1983).

Malraison B., Atten P., Bergé P., Dubois M., Phys. Lett., **44**, 897 (1983).

Swinney H., Roux J. C., in: Nonequilibrium dynamics in chemical systems (C. Vidal and A. Pacault, eds.), Springer, Berlin (1984).

Применения к анализу естественных временных последовательностей

Babloyantz A., Salazar M., Nicolis C., Phys. Lett., **111A**, 152 (1985).

Babloyantz A., Destexhe A., Proc. Nat. Acad. Sci. USA, **83**, 3513 (1986).

Berger A., Pestiaux P., Techn. Rep. n°28, Inst. Astron. and Geophys., Univ. of Louvain, Belgium (1982).

Dvorak I., Siska J., Phys. Lett., **118A**, 63 (1986).

Fraedrich K., J. Atmos. Sci., **43**, 419 (1986).

Mayer-Kress G. (ed.), Dimensions and entropies in chaotic systems, Springer, Berlin (1986).

Nicolis C., Nicolis G., Nature, **311**, 529 (1984).

Nicolis C., Nicolis G., Proc. Nat. Acad. Sci. USA, **83**, 536 (1986).

Rapp P., Zimmerman I., Albano A., Deguzman G., Greenbaum N., Phys. Lett., **110A**, 335 (1985).

Saltzman B., in: Irreversible phenomena and dynamical systems analysis in geosciences (C. Nicolis and G. Nicolis, eds.), Reidel, Dordrecht (1986).

Shackleton N., Opdyke N., Quat. Res., **3**, 39 (1973).

Предметный указатель

- Автокатализ** 29, 33, 52
— реакция БЖ 24—34
— явления саморегуляции 27, 39
Автономные системы 97, 285
Алгоритмическая сложность 36, 37, 223
Альbedo 52, 265
Амебы 42—47
Аминокислоты 170
Антивещество 169
Аррениуса закон 33
Архимедова сила 17
Асимметрия 167
— первичная 156
— пространственная 159
Аттракторы 84
— временные последовательности 275
— глобальные 86
— климатические 314
— непериодические 133
— периодические 120, 132
— случайность 152—154
— странные 146
— хаотические 152
- Белки** 170
Белоусова — Жаботинского реакция 24—34, 89, 125
— — и петля гистерезиса 33
— — и распространение волны 30
— — и хаотические аттракторы 152
— — и химические часы 25
Бенара задача 18—21, 29
Бернулли сдвиг 235
Бимодальность переходная 206
Биологические системы 42
Биомолекулы 170
Бистабильность 32, 40
— в иммунном ответе 261
— оптическая 40
Бифуркация 88, 112
— анализ 88, 294
— и асимметричные решения 156—160
- в диссипативных системах 112
— зависимость от параметров 114—116
— и нарушение симметрии 156
— и пространственные корреляции 196
— надкритическая 113
— поведение во времени 205—206
— предельная точка 123
— типа камертона 113
— уравнения 128—130, 298
— *Хопфа* 121, 124
Больцмана
— ансамбли 243
— кинетическая теория 244
— постоянная 34
Больших чисел закон 175
Большой взрыв, теория 56, 319
Броуновское движение 244
Брюсселятор, модель 123, 289
Булева логика 164
- Возвратные множества** 143
Возмущение 140, 296
— квазипериодического движения 140
— неинтегрируемое 138—140
— периодических (резонансных) движений 141
Время нахождения 28
— возвращения 187
- Гамильтоновы системы** 72
Гауссово распределение 176
Генетический код 171
Гиббса ансамбль 243
Гистерезис 33, 116
Гомоклиническая
— орбита 154
— точка 306
Границы межфазные 37
Грюнайзена параметр 255

- Дезоксирибонуклеиновая кислота — — сравнение с диссипативными системами 84—86
 (ДНК) 171
Де Ситтера пространство 323
Dictyostelium discoideum 44, 71
 Дислокации 255
 Диссипативная структура 20
 Диссипативные системы 26, 63, 117, 229
 — — и необратимость 26, 320
 Дифференциация внутренняя 91
 Диффузия 79
 — и нарушающие симметрию неустойчивости 156—159
 — и основное уравнение со многими переменными 179
 — пространственные корреляции 191—194
- Зародышеобразование 199
- Изолированные системы 69
 Иммунные системы, нелинейное поведение 261
 Инвариантность 91
 Инвариантные множества 99
 Инерциальное многообразие 159
 Инерции принцип 63
 Информация 92, 156, 167
 — позиционная 47
 — символическая динамика 212
 — хаотическая динамика 219—222
- Кантора* множество 135
 Катастроф теория 117
 Квантовомеханический вакуум 249
 Кварки 56
 Кинетическая
 — теория 186, 244
 — энергия 61
 Кинетический потенциал 199
 Кинетическое уравнение 245
 Клеточная динамика 257
 Клеточные автоматы 162
 Климатические изменения 147, 261
 Климатический потенциал 266
 Колебания
 — в диссипативных системах 30—33
 — в консервативных системах 62
 Консервативные системы 28, 102, 106, 226
 — — интегрируемые 106
 — — неинтегрируемые 137
- Ландау — Гинзбурга* потенциал 198
Лиувилля
 — теорема 102, 105, 110, 226
 — уравнение 105
Лоренца модель 244
Лотки — Вольтерра механизм 67
Ляпунова показатели 234, 292
 — функция 87, 186, 230
Ляпунова — Шмидта процедура 129
- Марковская
 — подгруппа 190
 — цепь 213
 Марковские процессы 179, 186
 Материаловедение 37, 253
 Машина бифуркационная 168
 Маятник 62
 — обратимость 28
 — фазовое пространство 62
 Межфазные границы 37
 Мезофаза 254
 Метастабильность 202
 Механика классическая 112
 — небесных тел 137
 Множество меры нуль 135
 Молекулярный хаос 20—22
 Морфогенез 156
 Муравьи, поиск пищи 268—275
- Натяжение поверхностное 38
 Нарушение симметрии 18, 91, 167, 237
 Нелинейность 70, 252
 Необратимость 193
 — и второй закон термодинамики 75—79
 — и марковские процессы 188
 Неоднородность пространственная 156—159
 неподвижная точка 98
 — — гиперболическая 111
 — — эллиптическая 110
 Неравновесные состояния 70
 Нормальная форма 122, 129
 Нуклеиновые кислоты 171

- Обратная связь 70
 Ограничение внешнее 15
 Оледенение 49
 Оператор столкновений 245
 Оптика 39—40
 Орбитальная
 — неустойчивость 85
 — устойчивость 84, 86
 Основное уравнение 179
 Особые точки 99, 117—122
 Открытые системы, химические 28—30
 Отображение типа подковы 143—146
 — закручивания 303
- Параметры порядка 122, 129
 Петля обратной связи 46
 Планковское время 324
Помо — *Манвиля* пульсация 154
 Предельные циклы 120; *см. также* Би-
 фуркация
 Преобразование пекаря 233
Пуанкаре — *Биркгофа* теорема 303
Пуанкаре секущая поверхность 131
Пуассона распределение 93
- Равновесие
 — детальное 69
 — механическое 67
 — термодинамическое 67
 — химическое 21
 Разбиение 213
 Разрушение материалов 254
 Реакционно-диффузионные модели
 160
 Резонанс 138, 247
 — и взаимодействие вещества со све-
 том 247
 Ренорм-группа 198
Рёсслера модель 155
 Рибонуклеиновая кислота (РНК) 171
- Самовоспроизведение 24
 Самоорганизация 12, 55
 — в геологии 39—40
 — в материаловедении 253—257
 — в механике жидких сред 18—21
 — в химии 24—34
 — в человеческих сообществах 275—
 280
 Сгорание 204
 Седло — фокус 148, 154
 Сепаратриса 111
 Снежинка 38
 Собственные значения 157
- Сродство 79
 Субаддитивность 187
- Теория относительности 225, 317
 Тепловая конвекция 13; *см. Бенара*
 задача
 Теплопроводность 15
- Узлы устойчивые и неустойчивые
 118—119
 Устойчивость 80
 — анализ 281
 — асимптотическая 80
 — маргинальная 128
 — орбитальная 84
- Фазовое пространство 99, 100, 130
 Фазовые переходы 56—57
 Ферменты 46
 — кооперативные 46
Фика закон 65
 Флуктуации 82
 — вероятностное описание 173
 — критическое поведение 193—195
 — локальный подход 185
 — поведение во времени 198
 Фокусы устойчивые и неустойчивые
 118
 Фракталы 130, 134
Фредгольма альтернатива 299
Фурье интегралы 246
- Хаббла* функция 323
 Хаотическая динамика 152
 Хемотаксис 44
 Химический потенциал 69
 Хиральность 30
- Центральная предельная теорема 175
 Центральные многообразия, теория
 129
 Циклический аденозинмонофосфат
 (цАМФ) 44, 46
- Часы химические 25
 Черные дыры 319
 Чувствительность 207

Шрёдингера уравнение 230

Шум, внешний 178

Эволюции законы 71

Экзотермические реакции 204

Электрические цепи, колебания 39

Эмбриональное развитие 42

Энергия активации 34, 171

Энтропия

— возрастание 78, 80

— и вероятность 180—190

— и информация 186

— и космология 320—323

— как стрела времени 193

— производство, и второй закон термодинамики 78, 80

Эргодические состояния 187

ОГЛАВЛЕНИЕ

| | |
|---|-----------|
| От переводчика | 5 |
| Предисловие | 7 |
| Глава 1. Сложность в природе | 10 |
| 1.1. Что такое сложность? | 10 |
| 1.2. Самоорганизация в физико-химических системах: рождение сложного | 12 |
| 1.3. Тепловая конвекция как прототип явлений самоорганизации в физике | 13 |
| 1.4. Явления самоорганизации в химии | 21 |
| А. Реакция БЖ в системе с перемешиванием: химические часы и хаос | 24 |
| Б. Реакция БЖ в неоднородной системе: пространственные фигуры | 29 |
| 1.5. Физико-химическая сложность и алгоритмическая сложность | 35 |
| 1.6. Некоторые дополнительные примеры сложного поведения в обычном масштабе | 37 |
| А. Явления, вызванные поверхностным натяжением. Наука о материалах | 37 |
| Б. Кооперативные явления, обусловленные электромагнитными полями. Электрические цепи, лазеры, оптическая бистабильность | 39 |
| 1.7. Снова биологические системы | 42 |
| 1.8. Сложность в планетарном и космическом масштабах | 47 |
| 1.9. Связь между силами и корреляциями. Подведение итогов | 53 |
| Глава 2. Словарь сложного | 58 |
| 2.1. Консервативные системы | 58 |
| 2.2. Диссипативные системы | 63 |
| 2.3. Механическое и термодинамическое равновесия. Неравновесные ограничения | 67 |
| 2.4. Нелинейность и обратные связи | 70 |
| 2.5. Многогранность второго закона термодинамики | 75 |
| 2.6. Устойчивость | 80 |
| 2.7. Бифуркация и нарушение симметрии | 88 |
| 2.8. Упорядоченность и корреляции | 92 |
| Глава 3. Динамические системы и сложность | 97 |
| 3.1. Геометрия фазового пространства | 97 |
| 3.2. Меры в фазовом пространстве | 100 |
| 3.3. Интегрируемые консервативные системы | 106 |
| 3.4. Бифуркация в простой диссипативной системе: поиск прототипов сложного | 112 |

| | | |
|---|--|-----|
| 3.5. | Диссипативные системы в двумерном фазовом пространстве: предельные циклы | 117 |
| 3.6. | Сведение к системам меньшей размерности: параметры порядка и нормальные формы | 122 |
| | А. Бифуркации типа острия и предельной точки | 123 |
| | Б. Бифуркация Хопфа и предельные циклы | 123 |
| 3.7. | Снова фазовое пространство: топологические многообразия и фракталы | 130 |
| | А. Периодические аттракторы: циклы порядка k | 132 |
| | Б. Квазипериодические аттракторы: инвариантные торы | 133 |
| | В. Непериодические аттракторы: фракталы | 133 |
| 3.8. | Неинтегрируемые консервативные системы: новая механика | 137 |
| | А. Возмущение квазипериодических движений | 140 |
| | Б. Возмущение периодических движений | 141 |
| 3.9. | Модель неустойчивого движения: подкова | 143 |
| 3.10. | Диссипативные системы в многомерных фазовых пространствах. Хаос и странные аттракторы | 146 |
| | А. Некоторые модели-прототипы, приводящие к хаотическому поведению | 147 |
| | Б. Некоторые «сценарии» становления хаотического поведения | 151 |
| 3.11. | Пространственно распределенные системы. Бифуркации с нарушением симметрии и морфогенез | 156 |
| 3.12. | Дискретные динамические системы. Клеточные автоматы | 162 |
| 3.13. | Асимметрия, отбор и информация | 167 |
| Глава 4. Случайное и сложное | | 173 |
| 4.1. | Флуктуации и вероятностное описание | 173 |
| 4.2. | Марковские процессы. Основное уравнение | 179 |
| 4.3. | Марковские процессы и необратимость. Информационная энтропия и физическая энтропия. | 186 |
| 4.4. | Пространственные корреляции и критическое поведение | 191 |
| 4.5. | Поведение флуктуаций во времени. Кинетика и временные масштабы самоорганизации | 198 |
| 4.6. | Чувствительность и отбор | 207 |
| 4.7. | Символическая динамика и информация | 212 |
| 4.8. | Генерация асимметричных, информационно-насыщенных структур | 216 |
| 4.9. | Снова алгоритмическая сложность | 223 |
| Глава 5. На пути к единой формулировке понятия сложного | | 225 |
| 5.1. | Общие свойства консервативных динамических систем | 226 |
| 5.2. | Общие свойства диссипативных динамических систем | 229 |
| 5.3. | Поиски унификации | 230 |
| 5.4. | Вероятность и динамика | 232 |
| 5.5. | Преобразование пекаря | 233 |
| 5.6. | Многообразия с нарушенной временной симметрией | 237 |
| 5.7. | Нарушающее симметрию преобразование \wedge | 239 |
| 5.8. | Ансамбли Гиббса и Больцмана | 243 |
| 5.9. | Кинетическая теория | 244 |
| 5.10. | Резонанс и взаимодействие света с веществом | 247 |
| 5.11. | Заключительные замечания | 249 |
| Глава 6. Сложное и перенос знаний | | 251 |
| 6.1. | Нелинейная динамика вдали от равновесия и моделирование сложного | 252 |
| 6.2. | Наука о материалах | 253 |
| 6.3. | Пороговые явления в клеточной динамике | 257 |

| | |
|--|-----|
| 6.4. Моделирование климатической изменчивости | 261 |
| 6.5. Вероятностное поведение и адаптивные стратегии у общественных насекомых | 268 |
| 6.6. Самоорганизация в человеческих сообществах | 275 |
| Приложение I. Линейный анализ устойчивости | 281 |
| I.1. Основные уравнения | 281 |
| I.2. Принцип устойчивости линеаризованной системы | 284 |
| I.3. Характеристическое уравнение | 285 |
| I.4. Иллюстрации | 288 |
| I.5. Системы с хаотической динамикой | 291 |
| Приложение II. Анализ бифуркаций | 294 |
| II.1. Общие свойства | 294 |
| II.2. Разложение решений в ряд по теории возмущений | 296 |
| II.3. Бифуркационные уравнения | 298 |
| Приложение III. Возмущение резонансных движений в неинтегрируемых консервативных системах | 301 |
| III.1. Отображение закручивания | 301 |
| III.2. Влияние возмущения в случае рациональных вращательных чисел | 303 |
| III.3. Гомоклинические точки | 306 |
| Приложение IV. Реконструкция динамики сложных систем по временной последовательности данных. Применение к климатической изменчивости | 310 |
| IV.1. Вводные замечания | 310 |
| IV.2. Теоретические основы анализа данных | 313 |
| IV.3. Климатический аттрактор | 314 |
| IV.4. Выводы и перспективы | 316 |
| Приложение V. Первичные необратимые процессы | 317 |
| V.1. Введение | 317 |
| V.2. Стандартная космологическая модель | 318 |
| V.3. Черные дыры | 319 |
| V.4. Роль необратимости | 320 |
| Литература | 326 |
| Предметный указатель | 336 |