Р.Гилмор

ПРИКЛАДНАЯ ТЕОРИЯ КАТАСТРОФ. Т.1

Монография известного американского ученого посвящена теории катастроф я ее приложениям в различных областях науки и техники. В русском переводе монография выходит в двух книгах. В книге 1 рассматриваются основные понятия теории катастроф и применение ее методов в физике, механике, математике, термодинамике и аэродинамике.

Для специалистов, занимающихся моделированием сложных систем управления, физиков, механиков, математиков, а также для аспирантов и студентов соответствующих специальностей.

Содержание

Предисловие редакторов перевода 5
Предисловие 7
Часть I. Элементарная теория катастроф
Глава 1. ТЕОРИЯ КАТАСТРОФ КАК ИССЛЕДОВАТЕЛЬСКАЯ ПРОГРАММА 9
Глава 2. ЛОКАЛЬНЫЙ ХАРАКТЕР ПОТЕНЦИАЛЬНЫХ ФУНКЦИЙ 13
1. Форма теоремы о неявной функции 14
2. Морсовские формы 15
3. Формы Тома 16
4. Канонические формы в окрестности критической точки 17
5. Функции катастроф 18
6. Выводы 20
Глава 3. ЗАМЕНА ПЕРЕМЕННЫХ. КАНОНИЧЕСКИЕ ФОРМЫ 21
1. Замена переменных 21
2. Приложения 24
3. Выводы 40
Глава 4. ЗАМЕНА ПЕРЕМЕННЫХ. ВОЗМУЩЕНИЯ 42
1. Возмущения 42
2. Приложения 43
3. Выводы 57
Глава 5. "ПРИНЦИП ЛОМА" 60
1. Потенциальная функция 60
2. Семейство потенциальных функций 64
3. Элементарные катастрофы 67
4. Бифуркационное множество и множество Максвелла 96
5. Выводы 99
Глава 6. ГЕОМЕТРИЯ СКЛАДКИ И СБОРКИ 102
1. Складка 102
2. Сборка 103
3. Двойственная сборка 109
4. Выводы 110
Глава 7. ОРГАНИЗАЦИЯ КАТАСТРОФ 111

1. $A_{k+1} \to A_k$ 112
2. $D_{k+1} \to D_k$ 113
3. Диаграммные представления 115
4. Контурные представления 129
5. Примыкания 131
6. Выводы 135
Глава 8. ПРИНЦИП МАКСИМАЛЬНОГО ПРОМЕДЛЕНИЯ И ПРИНЦИП МАКСВЕЛЛА **137**
1. Необходимость принятия соглашений 137
2. Соглашения 138
3. Какие соглашения использовать 141
4. Неадекватность принципов 150
5. Выводы 151
Глава 9. ФЛАГИ КАТАСТРОФ **153**
1. Модальность 154
2. Недостижимость 154
3. Катастрофические скачки 154
4. Расходимость 155
5. Гистерезис 156
6. Расходимость линейного отклика 157
7. Критическое замедление/смягчение моды 158
8. Аномальная дисперсия 163
9. Выводы 171
Часть II. Приложения элементарной теории катастроф в науке и технике
Глава 10. ТЕРМОДИНАМИКА **173**
1. Общее описание фазовых переходов 173
2. Фазовые переходы второго рода 175
3. Топологические соображения 179
4. Критическая точка жидкости 185
5. Уравнение состояния жидкости 189
6. Уравнение Ван-дер-Ваальса: количественный анализ 191
7. Фазовые переходы первого рода 196
8. Трикритические точки 200
9. Максвелловское множество для катастроф типа A_[+5] 203
10. Тройная точка 206
11. Топология катастроф и термодинамика 208
12. Метрическая геометрия и термодинамика 214
13. Частные производные термодинамических величин 220
14. Возможные вариационные представления 225
15. Флуктуации 231
16. Дополнительные соображения 232
17. Выводы 236

Глава 11. МЕХАНИКА КОНСТРУКЦИЙ **239**
1. Системы, описываемые потенциальной функцией 239
2. Выпучивание эйлерова стержня 243
3. Прощелкивание пологой арки 250
4. Смена типа устойчивости 256
5. Составные системы 261
6. Инженерная оптимизация 266
7. Нежелательные последствия инженерной оптимизации 267
8. Поддерживающие кронштейны 268
9. Смягчение моды 277
10. Выводы 279
Глава 12. АЭРОДИНАМИКА **281**
1. Описание состояния динамической системы 281
2. Описание состояния усеченной динамической системы 284
3. Линейный анализ устойчивости 286
4. Уравнения движения летательного аппарата 287
5. Многообразия стационарных состояний 291
6. Приложение к конкретному летательному аппарату 292
7. Выводы 303
Глава 13. КАУСТИКИ И ДИФРАКЦИОННЫЕ КАРТИНЫ **306**
1. Геометрическая оптика 306
2. Метод стационарной фазы 310
3. Степень сингулярности 312
4. Каустики и ростки катастрофы 316
5. Дифракционные картины и функции катастрофы 319
6. Выводы 325
Глава 14. КАНОНИЧЕСКИЕ ФОРМЫ ЖОРДАНА — АРНОЛЬДА. **327**
1. Исследовательская программа теории катастроф в случае систем линейных уравнений 327
2. Возмущения 330
3. Приложения 336
4. Вырожденности в семействах линейных операторов 339
5. Бифуркационные множества канонических форм Жордана—Арнольда 342
6. Выводы 344

ПРЕДИСЛОВИЕ РЕДАКТОРОВ ПЕРЕВОДА

Как показывает опыт, серьезные успехи в прикладной механике, физике, химии, а также технических дисциплинах часто являются результатом применения новых математических идей и методов. Можно с уверенностью утверждать, что и активно развивающаяся в последние годы прикладная математическая теория — теория катастроф — в сочетании с современными методами системного анализа станет полезным и эффективным средством качественного анализа различных реальных процессов.

Теория катастроф родилась на стыке двух дисциплин — топологии и математического анализа, ее источниками являются теория особенностей гладких отображений Х. Уитни и теория устойчивости и бифуркаций динамических систем А. Пуанкаре, А. Ляпунова и А. Андронова. Оба эти направления слились благодаря усилиям французского математика Р. Тома в единую стройную теорию, которая получила столь броское название — теория катастроф.

Пожалуй, именно своему названию новая теория в значительной мере обязана столь громкой известностью, довольно необычной для математических работ. Появление десятков, а затем и сотен публикаций, в которых сообщалось о применении теории катастроф не только в традиционных областях приложений математических методов, но и в таких сферах человеческой деятельности, как психология, социология, экономика, биология, медицина, геология, породило не менее мощный поток публикаций, относящихся скорее к жанру научной фантастики.

Некоторые приверженцы теории катастроф всерьез заявили об универсальности и всемогуществе новой теории. К этому можно добавить еще и склонность отдельных популяризаторов к туманному философствованию с налетом мистицизма, от которого, к сожалению, не были свободны и некоторые труды самого Р. Тома. Вполне естественно, что подобные необоснованные претензии вызывали резкую критику. Однако в пылу полемики критики часто впадали в другую крайность и вместе с мутной водой выплескивали и младенца, т. е. вообще отрицали какую-либо роль теории катастроф в развитии науки и техники на современном этапе.

Из сказанного ясно, насколько сложной и ответственной является задача создания руководства по теории катастроф для массового читателя, поскольку необходимо не только тщательно систематизировать трудную для чтения специальную литературу, но и освободить здоровое ядро новой теории от околонаучной и мистической шелухи.

Без преувеличения можно сказать, что автор книги, известный американский физик Р. Гилмор, хорошо справился с этой задачей. Правда, что касается приложений теории катастроф, то здесь он ограничился лишь сферой точных, или, как теперь часто говорят, «жестких» наук, оставив в стороне биологию, экономику, социологию и другие «мягкие» науки, где ощущается недостаток в моделях, основанных на системе разумных гипотез, на которые теория катастроф могла бы опереться. С этого обширного поля деятельности теория катастроф уже принесла отдельные ценные плоды, однако массовый

сбор урожая еще впереди, и любая поспешность может привести лишь к появлению незрелой и недоброкачественной продукции.

Главная цель книги — научить читателя применять мощный аналитический аппарат теории катастроф при решении проблем, возникающих в различных областях науки и техники. Этой цели автор достигает тщательным разбором примеров приложений элементарной теории катастроф в термодинамике, гидромеханике, строительной механике, геометрической и волновой оптике, квантовой механике, где часть результатов получена непосредственно автором.

Хотя ряд классических явлений и фактов из разных областей знаний был хорошо изучен и понят теоретиками и инженерами задолго до появления теории катастроф, она позволяет взглянуть на эти, казалось бы, совершенно различные реальные ситуации с единых позиций. Важно отметить, что автор при этом никоим образом не пытается дать готовые однозначные рецепты получения решений, а, напротив, стремится показать «кухню» научного поиска, вскрывая недостатки различных подходов и откровенно делясь с читателем своими сомнениями. Существенно также, что теория катастроф предстает не как замена традиционных методов, а как новое направление в рамках анализа, пригодное не только для выявления качественных закономерностей, но и для получения количественных результатов.

Автор обработал громадный материал по теории катастроф и теории динамических систем, представил их современное состояние и убедительно продемонстрировал плодотворность их приложений. Книга удачно дополняет недавно изданный в русском переводе обширный труд Т. Постона и И. Стюарта [1]), одна из глав которого, кстати, принадлежит Р. Гилмору, как в отношении охвата различных аспектов и перспективных направлений развития теории катастроф, так и в отношении распространения ее результатов и методов на более широкий класс динамических систем. В целом книга представляет несомненный интерес для читателей, деятельность которых так или иначе связана с применением математических методов в проводимых ими исследованиях.

Перевод книги выполнен И. И. Чижиковым (гл. 1—9, 14, 21—23), С. П. Чеботаревым (гл. 10, 15, 17—19), Д. А. Казениным (гл. 11—13, 20) и О. П. Бердышевой (гл. 16). Существенную помощь при редактировании гл. 14 оказал проф. В. И. Арнольд.

Ю. Гупало
А. Пионтковский

[1]) Постон Т., Стюарт И. Теория катастроф и ее приложения. — М.: Мир, 1980.

ПРЕДИСЛОВИЕ

Согласно традиционному подходу к изучению свойств решений дифференциальных уравнений, сначала необходимо явно определить полное множество решений и лишь только потом анализировать их свойства. Именно так поступали Лежандр, Лагер, Бессель, Эрмит, Гегенбауэр, Якоби и Чебышев при изучении дифференциальных уравнений второго порядка. Однако помимо уравнений данного типа возможны нелинейные или линейные дифференциальные уравнения выше второго порядка. Как следует поступать в этом случае и так ли уж необходим и целесообразен полный набор собственных значений функций для качественного описания поведения решений уравнений, моделирующих интересующую нас систему?

Одним из первых, кто осознал ограниченность и неадекватность традиционного подхода, который, кстати, неизбежно должен был привести к «опустошению» источника классических специальных функций математической физики, был американский астроном Г. Хилл. Выдающийся французский ученый А. Пуанкаре убедительно показал, что во многих случаях необходим лишь ограниченный объем информации качественного характера, которая, собственно, и представляет интерес при изучении конкретных систем уравнений. В таких случаях полный набор решений уравнений, полученных в результате кропотливой работы, можно считать скорее препятствием, чем подспорьем в деле понимания качественных изменений в поведении решений уравнений или систем уравнений.

Основы современного подхода к определению качественных изменений в поведении решений обыкновенных дифференциальных уравнений были заложены почти сто лет назад Пуанкаре, который впервые ввел такие понятия, как структурная устойчивость, динамическая устойчивость и критические множества. Особенно интересовало Пуанкаре, как качественно меняется поведение системы при изменении описывающих ее параметров. Эта работа Пуанкаре, по всей видимости, значительно опередила свое время. Сам Пуанкаре не смог реализовать намеченную им исследовательскую программу, поскольку был уже тяжело болен, а из его современников только А. Ляпунов следовал этой программе при изучении критических решений уравнений. После Ляпунова работы по теории бифуркаций практически прекратились, хотя независимо от программы Пуанкаре по изучению динамических систем на основе тех же работ зародилась топология, которая не только получила право на самостоятельное существование, но и успешно развивалась. Такая ситуация сохранялась вплоть до 30-х годов, пока советские математики А. Андронов и Л. Понтрягин, разрабатывая концепции структурной устойчивости, вновь не обратились к идеям Пуанкаре. Особое оживление в этой области наблюдалось в 1950—1966 гг. В 1967 г. дифференциальная топология и качественная динамика были синтезированы Смейлом в топологическую теорию динамических систем, которая по существу представляла собой не что иное, как исследовательскую программу Пуанкаре, сформулированную на современном математическом языке.

Приблизительно в 1930 г. М. Морс исследовал структуру канонических форм функции в окрестности изолированной точки, а Х. Уитни описал канонические формы отображений в особых точках. В 1950 г. Р. Том ввел важное понятие трансверсальности, которое стало основным при описании

структурной устойчивости Позднее Том использовал это понятие при описании канонических форм определенных особенностей отображений $f: \mathbb{R}^n \to \mathbb{R}^1$ (функций), которые он назвал «катастрофами».

Теория катастроф Р. Тома возникла на основе данных исследований, начатых Пуанкаре, и результатов дифференциального исчисления по каноническим формам функций и отображений, включая теорему о неявной функции и работы Морса и Уитни. В частности, элементарную теорию катастроф Р. Тома можно рассматривать как реализацию программы А. Пуанкаре применительно к изучению состояний равновесия динамических систем, описываемых потенциальной функцией.

Я вторгся в эту область совершенно случайно. Несколько лет назад я занимался исследованием задачи из области лазерной физики. Мне потребовалось почти три недели, чтобы осознать, что в этой задаче я имею дело с бифуркацией. Будучи по образованию физиком, я, естественно, был знаком с методами решений большого числа линейных задач, но практически ничего не знал о теории и методах решения нелинейных задач. Изучая этот математический аппарат, я встретил несколько загадочных ссылок на новую область математики, называемую «теорией катастроф». К сожалению, написанное по этому предмету оказалось еще более загадочным, чем сами ссылки. Наконец, не без помощи моих друзей, мне удалось связаться непосредственно с Т. Постоном. Именно он наглядно продемонстрировал мне, что элементарные катастрофы Р. Тома принадлежат к доступной области математики, а не к недоступной области философии.

Книга состоит из четырех частей. В ч. I и IV излагаются математические основы теории катастроф, причем ч. I написана на элементарном уровне и знакомство с ней требует лишь знания дифференциального исчисления. В ч. IV приводятся сведения из области математики, необходимые для того, чтобы можно было по крайней мере сформулировать теорему Тома, доказательство теоремы можно найти в сборнике трудов Зимана. Эта часть книги с избытком удовлетворяет потребности и научного работника, и инженера в необходимой информации. Именно на таком уровне я сам познакомился с теорией катастроф.

Две последние главы ч. I являются своего рода преамбулой к изучению приложений элементарной теории катастроф. В книге рассматриваются особенно характерные приложения теории катастроф в «жестких» науках. Однако они никоим образом не исчерпывают всех возможных приложений этой теории в математике, физике, химии и инженерных дисциплинах.

В ч. III обсуждаются различные пути распространения элементарной теории катастроф и проводится обсуждение функций катастроф более сложных, чем те, что изучались Р. Томом, и дается качественное описание поведения динамических систем.

Я искренне благодарен всем тем, кто помогал мне в осуществлении моего замысла — написании данной книги. Среди них Т. Постон, который познакомил меня с основами теории катастроф и предоставил возможность ознакомиться с его фундаментальной работой по теории катастроф до ее опубликования. И. Стюард и Н. Базли снабдили меня полезной информацией о теории катастроф и нелинейной математике.

Идея написания настоящей книги принадлежит С. Динзу, который оказался моим соавтором в одном весьма рискованном предприятии — написании статей по математической физике. Как он, так и Е. Ратинган, П. Драйпер, Н. Поп, Дж. Росс оказали мне неоценимую помощь в проведении некоторых вычислений. Нелегкий труд по подготовке прекрасных рисунков взял на себя Л. Миллер и его отдел. И, наконец, я благодарен моей жене Кларе, которая проявила завидные стойкость и терпение в процессе моей работы над книгой.

Арлингтон, Виргиния
Март 1981

Роберт Гилмор

Часть I.
ЭЛЕМЕНТАРНАЯ ТЕОРИЯ КАТАСТРОФ

1

ТЕОРИЯ КАТАСТРОФ КАК ИССЛЕДОВАТЕЛЬСКАЯ ПРОГРАММА

По аналогии с Эрлангенской программой Ф. Клейна теорию катастроф можно рассматривать как своего рода исследовательскую программу. Однако если суть первой состояла в том, чтобы свести классификацию важнейших геометрических систем к классификации групп преобразований, при действии которых теоремы рассматриваемых геометрий остаются инвариантными, то предметом теории катастроф является изучение зависимости качественной природы решений уравнений от значений параметров, присутствующих в заданных уравнениях.

Для того чтобы уточнить, что же конкретно изучает теория катастроф, рассмотрим решение $\psi_1(t, x; c_\alpha)$, $\psi_2(t, x; c_\alpha)$, ... системы n уравнений, определенной в пространстве $\mathbb{R}^N$ с координатами $x = (x_1, x_2, \ldots, x_N)$,

$$F_i\left(\psi_j;\ c_\alpha;\ t,\ \frac{d\psi_j}{dt},\ \frac{d^2\psi_j}{dt^2},\ \ldots;\ x_l;\right.$$

$$\left.\frac{\partial\psi_j}{\partial x_l},\ \frac{\partial^2\psi_j}{\partial x_l\,\partial x_m},\ \ldots;\ \int dx_l,\ \ldots\right)=0, \qquad (1.1)$$

где

$$1\leqslant i\leqslant n,\quad 1\leqslant l,\quad m\leqslant N,\quad 1\leqslant\alpha\leqslant k,$$

причем переменные x_l и t можно считать соответственно пространственными и временными координатами. Решения ψ_j описывают состояние некоторой системы, поэтому будем называть их *переменными состояния*. Предполагается, что уравнения $F_i = 0$ зависят от k параметров c_α (числа Рейнольдса, структурной константы, напряженности магнитного поля и т. д.), т. е. последние могут качественно влиять на свойства решений ψ_j, и естественно назвать их *управляющими параметрами*.

Проблема исследования решений системы уравнений (1.1), даже если речь идет лишь о том, как зависят эти решения от управляющих параметров c_α, является исключительно сложной.

Однако ее можно упростить, сделав ряд последовательных предположений.

1. Предположим, что выражение (1.1), которое в самом общем виде будет интегро-дифференциальным уравнением (или значительно хуже), в действительности не содержит интегралов. Фактически это означает, что система уравнений (1.1) есть не что иное, как множество (нелинейных) уравнений в частных производных.

2. В целях дальнейшего упрощения предположим, что система уравнений (1.1) не содержит пространственных производных любого порядка, т. е.

$$F_i = F_i\left(\psi_j;\ c_\alpha;\ t,\ \frac{d\psi_j}{dt},\ \frac{d^2\psi_j}{dt^2},\ \ldots;\ x_l;\ —;\ —\right). \tag{1.2}$$

3. Поскольку решение данной системы уравнений вызывает существенные затруднения, предположим, что она полностью не зависит от пространственных координат x_l:

$$F_i = F_i\left(\psi_j;\ c_\alpha;\ t,\ \frac{d\psi_j}{dt},\ \frac{d^2\psi_j}{dt^2},\ \ldots;\ —;\ —;\ —\right). \tag{1.3}$$

4. Следующее предположение сводится к тому, что система уравнений (1.3) содержит производные по времени не выше первого порядка и, кроме того, эти производные входят в упрощенную функцию F_i специальным («каноническим») образом:

$$F_i = \frac{d\psi_i}{dt} - f_i(\psi_j;\ c_\alpha;\ t). \tag{1.4}$$

Систему уравнений данного типа ($F_i = 0$) называют *динамической системой*. И опять же она слишком трудна для исследования.

5. Для упрощения динамической системы предположим, что функции f_i [выражение (1.4)] полностью не зависят от времени. Тогда получим так называемую *автономную динамическую систему* уравнений

$$F_i = \frac{d\psi_i}{dt} - f_i(\psi_j;\ c_\alpha;\ ——) = 0. \tag{1.5}$$

Относительно автономных динамических систем, зависящих от малого числа управляющих параметров ($k \leqslant 4$), уже может быть высказано несколько полезных и сильных утверждений.

6. Наконец, заметим, что функции f_i во многом аналогичны компонентам силы в классической механике. В последней существенное упрощение возможно тогда, когда сила является консервативной. Если все функции f_i могут быть заданы

Таблица 1.1. **Предмет изучения элементарной теории катастроф**

Упрощающие предположения	Структурный вид уравнений $F_i=0$	Примечания
0	$F_i\left(\psi_j;\ c_\alpha;\ t,\ \frac{d\psi_j}{dt},\ \frac{d^2\psi_j}{dt^2},\ \ldots;\ x_l;\ \frac{\partial\psi_j}{\partial x_l},\ \frac{\partial^2\psi_j}{\partial x_l\,\partial x_m},\ \ldots;\ \int dx_l,\ \ldots\right)$	
1	$F_i\left(\psi_j;\ c_\alpha;\ t,\ \frac{d\psi_j}{dt},\ \frac{d^2\psi_j}{dt^2},\ \ldots;\ x_l;\ \frac{\partial\psi_j}{\partial x_l},\ \frac{\partial^2\psi_j}{\partial x_l\,\partial x_m},\ \ldots;\ \text{———}\right)$	
2	$F_i\left(\psi_j;\ c_\alpha;\ t,\ \frac{d\psi_j}{dt},\ \frac{d^2\psi_j}{dt^2},\ \ldots;\ x_l;\ \text{———};\ \text{———}\right)$	
3	$F_i\left(\psi_j;\ c_\alpha;\ t,\ \frac{d\psi_j}{dt},\ \frac{d^2\psi_j}{dt^2},\ \ldots;\ —;\ \text{———};\ \text{———}\right)$	
4	$F_i=\frac{d\psi_i}{dt}-f_i\left(\psi_j;\ c_\alpha;\ t\right)$	Динамические системы
5	$F_i=\frac{d\psi_i}{dt}-f_i\left(\psi_j;\ c_\alpha;\ —\right)$	Автономные динамические системы
6	$F_i=\frac{d\psi_i}{dt}+\frac{\partial V\left(\psi_j;\ c_\alpha\right)}{\partial\psi_i}$	Градиентные системы
7	$0=0+\frac{\partial V\left(\psi_j;\ c_\alpha\right)}{\partial\psi_i}$	Состояния равновесия градиентных систем

антиградиентом (по отношению к ψ_i) некоторой потенциальной функции, то получаем систему уравнений

$$f_i = -\frac{\partial V(\psi_j; c_\alpha)}{\partial \psi_i},$$
$$F_i = \frac{d\psi_i}{dt} + \frac{\partial V(\psi_j; c_\alpha)}{\partial \psi_i} = 0, \tag{1.6}$$

которую называют *градиентной системой* ($\dot{\psi} = -\nabla_\psi V$). О свойствах таких систем может быть доказано довольно много глубоких теорем.

Особый интерес представляет изучение состояния равновесия $d\psi_i/dt = 0$ градиентных динамических систем, которое может быть описано с помощью следующей системы уравнений:

$$\frac{\partial V(\psi_j; c_\alpha)}{\partial \psi_i} = 0. \tag{1.7}$$

(Эти уравнения могут не иметь решений $[V(\psi) = \psi]$, иметь одно $[V(\psi) = \psi^2]$ или более чем одно решение $[V(\psi; c) = \psi^4 + c\psi^2$, одно решение, если $c > 0$, и три решения, если $c < 0]$.) В этом случае может быть доказано большое число полезных и сильных утверждений как о состоянии равновесия градиентных систем, так и о том, как эти состояния зависят от управляющих параметров c_α.

Таким образом, можно сделать вывод, согласно которому элементарная теория катастроф — это наука о том, каким образом состояния равновесия $\psi_j(c_\alpha)$ потенциальной функции $V(\psi_j, c_\alpha)$ изменяются при изменении управляющих параметров c_α (табл. 1.1).

2

ЛОКАЛЬНЫЙ ХАРАКТЕР ПОТЕНЦИАЛЬНЫХ ФУНКЦИЙ

Переменные состояния ψ_j, от которых зависит потенциальная функция $V(\psi_j; c_\alpha)$ [выражение (1.7)], по существу являются обобщенными координатами рассматриваемой системы. В соответствии с общепринятыми математическими обозначениями заменим ψ_j на x_j [следует иметь в виду, что x_j не имеют ничего общего с пространственными координатами в выражении (1.1)]. В случае точечной покоящейся частицы в поле, описываемом потенциальной функцией $V(x; c)$, $x \in \mathbb{R}^n$, $c \in \mathbb{R}^k$, состояние последней полностью описывается ее расположением в пространстве. В этом частном случае решения $x_j(c_\alpha)$ системы уравнений (1.7) являются координатами частицы, находящейся в состоянии равновесия. Однако в более общем случае переменные x_j могут быть аналогами коэффициентов Фурье или параметра порядка.

Данная глава посвящена исследованию локальных свойств потенциальных функций $V(x)$ и семейств таких функций $V(x; c)$; при этом первые можно рассматривать как отображения $\mathbb{R}^n \to \mathbb{R}^1$, а вторые — как отображения $\mathbb{R}^n \otimes \mathbb{R}^k \to \mathbb{R}^1$. Локальные свойства потенциальных функций и семейств потенциальных функций определяются рядом теорем функционального анализа (табл. 2.1), для понимания которых требуются определенные сведения из топологии.

Таблица 2.1. **Результаты функционального анализа**

Условия	$V \doteq$	Уравнение	Теорема/Лемма	Литературный источник
1. $\nabla V \neq 0$	y_1	(2.1)	Теорема о неявной функции [1])	
2. $\nabla V = 0$, но $\det V_{ij} \neq 0$	$\sum_{i=1}^{n} \lambda_i y_i^2$	(2.2а)	Лемма Морса	[2]
	$M_i^n(\tilde{y})$	(2.2б)		
3. $\nabla V = 0$ и $\det V_{ij} = 0$	$f_{NM}(y_1, \ldots, y_l) + M_i^{n-l}$	(2.3а)	Лемма расщепления	[3]
V «общая»	$CG(l) + M_i^{n-l}$	(2.3б)	Теорема Тома	[4]
$k \leqslant 5$	$\mathrm{Cat}(l, k) + M_i^{n-l}$	(2.4)	» »	[4]

[1]) До конца 19 столетия математиков мало заботило строгое доказательство данной теоремы, поскольку знание теоремы предшествовало эре математической строгости. В действительности оно предшествовало дифференциальному и интегральному исчислению, так как эта теорема, несомненно, была известна Галилею [1].

1. ФОРМА ТЕОРЕМЫ О НЕЯВНОЙ ФУНКЦИИ

Известно, что обобщенная сила, действующая на систему, поведение которой описывается потенциальной функцией, равна антиградиенту этой функции. Поэтому если в рассматриваемой точке пространства состояний градиент потенциальной функции отличен от нуля, то и сила, действующая в точке, также отлична от нуля: $F = -\nabla V \neq 0$ (рис. 2.1). В этом случае в некоторой

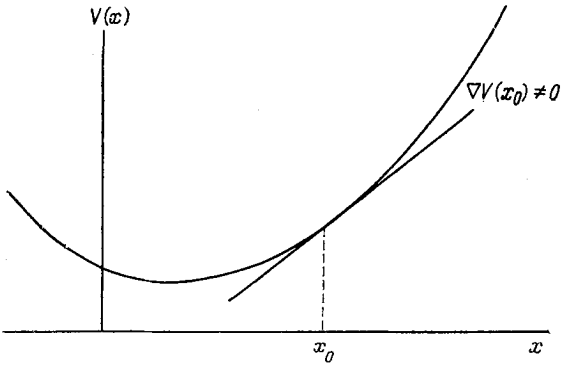


Рис. 2.1. Преобразование функции V в локально линейную функцию $V \to a + (y - y_0)b$ с помощью гладкой замены координат в точке x_0, в которой градиент $\nabla V \neq 0$.

окрестности заданной точки можно выбрать новую систему координат, такую, что сила в этих новых координатах будет иметь единственную отличную от нуля компоненту (без ограничения общности можно считать, что ее первая компонента равна 1). Для того чтобы сделать все эти рассуждения математически строгими, необходимо использовать теорему о неявной функции [1], согласно которой возможна гладкая (т. е. имеющая производные любого порядка) замена координат:

$$\begin{aligned} y_1 &= y_1(x_1, x_2, \ldots, x_n), \\ y_2 &= y_2(x_1, x_2, \ldots, x_n), \\ &\vdots \\ y_n &= y_n(x_1, x_2, \ldots, x_n), \end{aligned} \tag{2.0}$$

в результате которой в новой системе координат[1]) имеем

$$V \doteq y_1 (+ \text{Константа}). \tag{2.1}$$

[1]) Знак $\doteq$ означает «... равно после гладкой замены переменных ...».

При исследовании локальных свойств потенциальной функции постоянную в формуле (2.1) можно не учитывать. (От нее можно также избавиться при помощи соответствующего сдвига начала системы координат.)

2. МОРСОВСКИЕ ФОРМЫ

Если рассматриваемая физическая система находится в состоянии равновесия (устойчивого или неустойчивого), то $\nabla V = 0$. При этом тип равновесия определяется собственными зна-

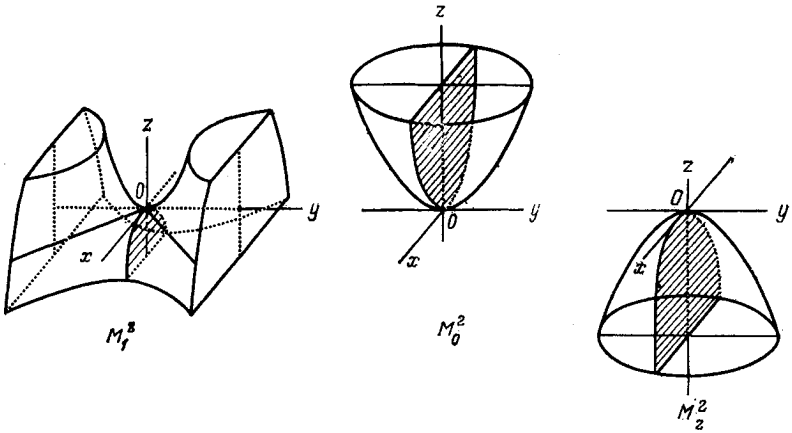


Рис. 2.2. Морсовское седло в пространстве $\mathbb{R}^2$, имеющее локальный минимум и локальный максимум.

чениями матрицы устойчивости, или гессиана, $V_{ij} = \partial^2 V/\partial x_i\, \partial x_j$. Поскольку $\nabla V = 0$ — условие, необходимое для применимости теоремы о неявной функции ($\nabla V \neq 0$), не выполняется, потенциальная функция не может быть представлена в канонической форме (2.1). Однако, если $\det V_{ij} \neq 0$, теорема Морса [2] гарантирует существование гладкой замены переменных, такой, что потенциальная функция локально может быть представлена квадратичной формой

$$V \doteq \sum_{i=1}^{n} \lambda_i y_i^2. \qquad (2.2а)$$

Здесь λ_i — собственные значения матрицы устойчивости V_{ij}, вычисленные для состояния равновесия. С учетом новой замены координат в соответствии с $\tilde{y}_i = |\lambda_i|^{1/2} y_i$ квадратичная форма (2.2а) может быть приведена к морсовской канонической форме

$$V \doteq -\tilde{y}_1^2 - \ldots - \tilde{y}_i^2 + \tilde{y}_{i+1}^2 + \ldots + \tilde{y}_n^2 = M_i^n(\tilde{y}). \qquad (2.2б)$$

Функцию $M_i^n(\tilde{y})$ называют *морсовским i-седлом*. Исключительно морсовские 0-седла имеют локальный минимум в точке равновесия, так что только такие седла локально устойчивы. На рис. 2.2 изображены три морсовские функции, заданные в пространстве $\mathbb{R}^2$.

Определение. Точки, в которых $\nabla V = 0$, *являются точками равновесия, или критическими точками, гладкой функции* $V(x_1, x_2, \ldots, x_n)$. *Критические точки, в которых* $\det V_{ij} \neq 0$, *называют изолированными, невырожденными или морсовскими критическими точками.*

3. ФОРМЫ ТОМА

Определение. Критические точки функции $V(x_1, x_2, \ldots, x_n)$, *в которых* $\det V_{ij} = 0$, *являются неизолированными, вырожденными или неморсовскими критическими точками.*

Если потенциальная функция зависит от одного или более управляющих параметров $c_1, c_2, \ldots, c_k$, то матрица устойчивости V_{ij} и ее собственные значения также зависят от этих параметров. В этом случае вполне возможно, что при некоторых значениях управляющих параметров одно (или несколько) собственное значение матрицы устойчивости может (могут) обратиться в нуль. Если это так, то $\det V_{ij} = 0$ и, следовательно, условия, необходимые для применимости леммы Морса ($\nabla V = 0$, $\det V_{ij} \neq 0$), не выполняются, и в точке равновесия потенциальная функция не может быть представлена в канонической форме (2.2). Однако можно найти каноническую форму потенциальной функции в неморсовской критической точке. Пусть только l собственных значений $\lambda_1(c), \ldots, \lambda_l(c)$ обращаются в нуль в точке $c = c^0$, тогда для расщепления потенциальной функции на неморсовскую и морсовскую составляющие можно использовать лемму расщепления Тома:

$$V(x; c) \doteq f_{NM}(y_1(x; c), \ldots, y_l(x; c); c) + \sum_{i=l+1}^{n} \lambda_i(c)(y_i(x))^2. \quad (2.3a)$$

В таком представлении l «плохих» координат $y_1(x; c), \ldots, y_l(x; c)$, соответствующих l обращающимся в нуль собственным значениям $\lambda_1(c), \ldots, \lambda_l(c)$, являются гладкими функциями как n переменных состояния, так и k управляющих параметров $c_1, \ldots, c_k$. Что касается «хороших» координат $y_{l+1}(x), \ldots, y_n(x)$, которые соответствуют отличным от нуля собственным значениям $\lambda_{l+1}(c), \ldots, \lambda_n(c)$, то они являются гладкими функциями исключительно исходных переменных состояния x. В точке $(x^0; c^0)$ матрица устойчивости $\partial^2 f_{NM}/\partial y_i \partial y_j (1 \leqslant i, j \leqslant l)$ обращается в нуль (все элементы матрицы — нули), в то время как матрица

устойчивости морсовской функции в представлении (2.3а) не вырождена. При соответствующих условиях [$k \leqslant 5$ и нет особых или симметричных ограничений на семейство потенциальных функций $V(x_1, \ldots, x_n; c_1, \ldots, c_k)$] теорема Тома [4] гарантирует существование гладкой замены переменных, такой, что потенциальная функция может быть записана в канонической форме

$$V \doteq CG(l) + \sum_{j=l+1}^{n} \lambda_j y_j^2. \qquad (2.3б)$$

Функцию $CG(l)$ называют *ростком катастрофы.* Все канонические ростки катастроф при $k \leqslant 5$ перечислены в табл. 2.2. Они соответствуют лишь одному ($l = 1$) или двум ($l = 2$) нулевым собственным значениям матрицы устойчивости.

Таблица 2.2. **Элементарные катастрофы Тома**

Тип катастрофы	k	Росток	Возмущение
A_2	1	x^3	a_1x
$A_{\pm 3}$	2	$\pm x^4$	$a_1x + a_2x^2$
A_4	3	x^5	$a_1x + a_2x^2 + a_3x^3$
$A_{\pm 5}$	4	$\pm x^6$	$a_1x + a_2x^2 + a_3x^3 + a_4x^4$
A_6	5	x^7	$a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5$
D_{-4}	3	$x^2y - y^3$	$a_1x + a_2y + a_3y^2$
D_{+4}	3	$x^2y + y^3$	$a_1x + a_2y + a_3y^2$
D_5	4	$x^2y + y^4$	$a_1x + a_2y + a_3x^2 + a_4y^2$
D_{-6}	5	$x^2y - y^5$	$a_1x + a_2y + a_3x^2 + a_4y^2 + a_5y^3$
D_{+6}	5	$x^2y + y^5$	$a_1x + a_2y + a_3x^2 + a_4y^2 + a_5y^3$
$E_{\pm 6}$	5	$x^3 \pm y^4$	$a_1x + a_2y + a_3xy + a_4y^2 + a_5xy^2$

4. КАНОНИЧЕСКИЕ ФОРМЫ В ОКРЕСТНОСТИ КРИТИЧЕСКОЙ ТОЧКИ

На первый взгляд канонические формы (2.3а) и (2.3б) могут показаться аналогичными, однако в действительности это не так. При выводе канонической формы (2.3а) используется лемма расщепления Тома, в основе которой лежат положения, ничуть не более сложные, чем теоремы курса математического анализа повышенной трудности, в частности теорема о неявной функции. Разложение (2.3а) имеет место в некоторой открытой окрест-

ности точки $(x^0; c^0)$ пространства $\mathbb{R}^n \otimes \mathbb{R}^k$; здесь x^0 — неморсовская критическая точка, соответствующая $c = c^0$ при заданном значении управляющих параметров. Однако лемма расщепления ничего не говорит нам о виде функции $f_{NM}(y_1, \ldots, y_l)$, за исключением того, что ее разложение в ряд Тейлора в вырожденной критической точке при заданном $c = c^0$ не содержит членов, в которые входят первая и вторая степени разностей $(y_i - y_i^0)$.

При выводе же канонической формы (2.3б) существенно используется теорема Тома [1]). Разложение (2.3б) потенциальной функции V имеет место в открытой окрестности точки x^0 пространства $\mathbb{R}^n$ при фиксированном значении $c = c^0$. В этом случае можно перечислить ряд определенных канонических форм (получаемых после подходящей замены координат), с помощью которых можно получить все способы появления «плохих» неморсовских переменных в рассматриваемой вырожденной критической точке.

Таким образом, первое разложение (2.3а) справедливо в окрестности точки $(x^0; c^0)$ пространства $\mathbb{R}^n \otimes \mathbb{R}^k$, но в нем f_{NM} не имеет какого-нибудь определенного вида; второе разложение (2.3б) имеет место в окрестности точки x^0 пространства $\mathbb{R}^n$, и в нем функция f_{NM} имеет вполне определенный вид, называемый *ростком катастрофы* этой функции.

Каноническое разложение потенциальной функции, сочетающее сильные стороны обоих разложений (2.3а) и (2.3б) и не имеющее их слабых сторон, также было получено Р. Томом. Так, если x^0 — неморсовская критическая точка потенциальной функции семейства $V(x; c)$ при $c = c^0$, то в открытой окрестности точки $(x^0; c^0)$ пространства $\mathbb{R}^n \otimes \mathbb{R}^k$

$$V \doteq \mathrm{Cat}(l, k) + \sum_{j=l+1}^{n} \lambda_j(c)\, y_j^2. \qquad (2.4)$$

Функцию $\mathrm{Cat}(l, k)$ называют *функцией катастрофы* или просто *катастрофой*. Здесь l — размерность нулевого пространства V_{ij} в неморсовской критической точке, а k — число управляющих параметров.

5. ФУНКЦИИ КАТАСТРОФ

Канонические формы катастроф с одной и двумя переменными состояния перечислены в табл. 2.2. Эта таблица, впервые

[1]) Теорема Тома исключительно глубока по содержанию, и ее доказательство стимулировало развитие ряда новых областей математики.

предложенная Р. Томом, по существу является таблицей элементарных катастроф и содержит в каждой своей строке две функции: росток катастрофы $CG(l)$ и ее возмущение $\mathrm{Pert}(l, k)$. Функция катастрофы $\mathrm{Cat}(l, k)$ равна сумме этих двух функций, взятых из любой строки.

Определение.

$$\mathrm{Cat}(l, k) = CG(l) + \mathrm{Pert}(l, k). \tag{2.5}$$

Таким образом, росток катастрофы $CG(l)$ есть неморсовская функция l переменных. Именно эта функция присутствует в канонической форме (2.3б), имеющей место в окрестности неморсовской критической точки x^0 (или y^0) при фиксированном значении $c = c^0$. Функция катастрофы $\mathrm{Cat}(l, k)$ представляет собой функцию l переменных (состояния) и k (управляющих) параметров, которая появляется в канонической форме (2.4), имеющей место в окрестности неморсовской критической точки x^0 «вокруг» значения c^0. Функция катастрофы $\mathrm{Cat}(l, k)$ сводится к ростку катастрофы $CG(l)$ только тогда, когда в пространстве $\mathbb{R}^k$ физические c_α или математические a_α управляющие параметры соответственно равны c_α^0 и 0. В общем случае математические $a_1, \ldots, a_k$ и физические $c_1, \ldots, c_k$ управляющие параметры связаны посредством гладкого преобразования с невырожденным якобианом

$$\det \frac{\partial a_\alpha(c)}{\partial c_\beta}\Big|_{c^0} \neq 0, \quad 1 \leqslant \alpha, \beta \leqslant k. \tag{2.6}$$

Пример 1. Функция катастрофы, обозначенная в табл. 2.2 D_{-4}, имеет вид

$$D_{-4}\colon\ f(x, y; a_1, a_2, a_3) = (x^2y - y^3) + (a_1x + a_2y + a_3y^2), \tag{2.7}$$

где для удобства предполагается, что $x_1 = x$, $x_2 = y$.

Пример 2. Предположим, что потенциальная функция $V(x; c)$ зависит от десяти переменных состояния $x_1, \ldots, x_{10}$ и трех управляющих параметров c_1, c_2, c_3. Пусть точка $(x^0; c^0)$ — неморсовская критическая точка, в которой квадратная матрица V_{ij} порядка 10 имеет два собственных значения, равные нулю (т. е. дефект 2), три отрицательных и пять положительных собственных значений. Определим локальный характер этой функции в точке $(x^0; c^0)$.

Согласно формуле (2.4),

$$V \doteq \mathrm{Cat}(l, k) + f_M(y_3, \ldots, y_{10}). \tag{2.8}$$

Поскольку $l = 2$, $k = 3$, непосредственно из табл. 2.2 находим, что $\mathrm{Cat}(2, 3) = D_{\pm 4}$, т. е.

$$V \doteq (y_1^2y_2 \pm y_2^3) + (a_1y_1 + a_2y_2 + a_3y_2^2) + \sum_{j=3}^{10} \lambda_j(c)\, y_j^2. \tag{2.9}$$

Эта каноническая форма имеет место в открытой окрестности точки $(x^0; c^0)$ пространства $\mathbb{R}^n \otimes \mathbb{R}^k$, в которой $\lambda_j(c^0) = -1$, $j = 3, 4, 5$ и $\lambda_j(c^0) = +1$, $j = 6, 7, 8, 9, 10$.

6. ВЫВОДЫ

Локально поведение потенциальной функции описывается рядом теорем функционального анализа:

— если $\nabla V \neq 0$, то для приведения потенциальной функции к канонической форме (2.1) может быть использована теорема о неявной функции;

— если $\nabla V = 0$, то теорема о неявной функции неприменима. Однако если $\det V_{ij} \neq 0$, то для приведения V к каноническому виду может быть использована лемма Морса;

— если $\nabla V = 0$ и $\det V_{ij} = 0$, то лемма Морса уже неприемлема. Однако если $V(x; c)$, $x \in \mathbb{R}^n$, $c \in \mathbb{R}^k$ — типичное k-параметрическое семейство потенциальных функций, то для приведения V к канонической форме (2.3) может быть использована теорема Тома. При значениях $k \leqslant 5$ теорема Тома может быть также использована для приведения V к более сильной канонической форме (2.4).

Термины «критическая точка», «точка равновесия», «стационарная точка» имеют различный смысл в математике, физике и других науках (табл. 2.3), поэтому следует иметь в виду, что здесь и далее используется чисто математическая терминология, за исключением случаев, когда явно оговорено противное.

Таблица 2.3. **Принятые обозначения**

Свойство	Математика	Физика
$\nabla V \neq 0$	Некритическая точка	Неравновесная точка
$\nabla V = 0$, $\det V_{ij} \neq 0$	Критическая точка (морсовская или обычная критическая точка)	Точка равновесия
$\nabla V = 0$, $\det V_{ij} = 0$	Неморсовская критическая точка, вырожденная критическая точка	Критическая точка

Литература

1. Galilei G., Dialogues Concerning Two New World Sciences (transl. H. Crew and A. De Salvio, Evanston. Northwestern Univ. Press, 1950.
2. Morse M., The Critical Points of a Function of n Variables, *Trans. Am. Math. Soc.*, 33, 72—91 (1931).
3. Gromoll D., Meyer W., On Differentiable Functions with Isolated Critical Points, *Topology*, 8, 361—370 (1969).
4. Thom R., Stabilité Structurelle et Morphogénèse, New York, Benjamin, 1972; transl. Structural Stability and Morphogenesis, Reading, Benjamin, 1975.

3

ЗАМЕНА ПЕРЕМЕННЫХ. КАНОНИЧЕСКИЕ ФОРМЫ

Канонические формы (2.1)—(2.4) потенциальных функций были получены в результате некоторых гладких замен переменных. Попытаемся теперь показать, руководствуясь в основном интуицией, почему канонические формы потенциальных функций задаются именно формулами (2.1) для некритической точки, (2.2) — для морсовской критической точки и (2.4) — для неморсовской критической точки семейств функций, зависящих от менее чем шести управляющих параметров.

Свойства потенциальной функции в рассматриваемой точке определяются начальными (наименьшей степени) членами разложения ее в ряд Тейлора в этой точке. Если функция зависит от управляющих параметров, то от них также зависят и коэффициенты разложения ее в ряд Тейлора. Эти ведущие члены могут исчезнуть при некоторых значениях управляющих параметров, при этом свойства функции, как правило, изменяются. В этом случае возможно такое преобразование координат, которое позволит удалить последние (более высокой степени) члены разложения функции в ряд Тейлора, а оставшиеся члены будут определять свойства функции в рассматриваемой точке. Именно эти члены и называют *ростком функции.*

Следовательно, росток находится между начальными членами, которые исключаются при помощи управляющих параметров, и последними членами, которые исключаются в результате преобразования координат. Если росток линеен, применима теорема о неявной функции, если росток является невырожденной квадратичной формой, применима лемма Морса, если росток содержит вырожденную квадратичную форму, то применима лемма расщепления.

1. ЗАМЕНА ПЕРЕМЕННЫХ

При описании физического процесса в пространстве $\mathbb{R}^n$ удобно выделить некоторую произвольную систему координат $(x_1, x_2, \ldots, x_n)$. Если рассмотреть новую систему координат $(x'_1, x'_2, \ldots, x'_n)$, то в любой точке $p \in \mathbb{R}^n$, в которой якобиан $\det \left| \partial x'_i / \partial x_j \right|_{p \in \mathbb{R}^n}$ отличен от нуля, исходная координатная система преобразуется в новую систему, т. е.

$$\begin{aligned} x'_1 &= x'_1(x_1, x_2, \ldots, x_n), \\ x'_2 &= x'_2(x_1, x_2, \ldots, x_n), \\ &\vdots \\ x'_n &= x'_n(x_1, x_2, \ldots, x_n). \end{aligned} \tag{3.1}$$

В такой точке пространства $\mathbb{R}^n$ возможно и обратное преобразование $x_j = x_j(x_i')$. Существование этого обратного преобразования следует из теоремы о неявной функции.

Во многих ситуациях бывает полезно использовать разложение координат $x_i' = x_i'(x_j)$ в ряд Тейлора:

$$x_i' = A_i + A_{i,\,j}x_j + A_{i,\,jk}x_jx_k + A_{i,\,jkl}x_jx_kx_l + \ldots . \qquad (3.2)$$

Преобразование (3.2) обратимо в точке $(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0)$, если $\det|A_{i,\,j}| = 0$. Факториальные коэффициенты, которые иногда приводятся в разложениях типа (3.2), здесь для простоты опущены.

С целью упрощения вычислений как в этой, так и в следующей главе общее нелинейное преобразование (3.2) будем выполнять путем последовательного осуществления следующих преобразований:

$$x_i \to x_i' = x_i + A_i \qquad (3.3\text{ih})$$

— неоднородное линейное преобразование, которое означает перемещение начала координат и параллельный перенос координатных осей;

$$x_i' \to x_i'' = A_{i,\,j}x_j' \qquad (3.3\text{hl})$$

— однородное линейное преобразование, которое геометрически можно интерпретировать как вращение и сжатие координатных осей. Оно не вырождено, если $\det|A_{i,\,j}| \neq 0$;

$$x_i'' \to x_i''' = \delta_{ij}x_j'' + A_{i,\,jk}x_j''x_k'' + A_{i,\,jkl}x_j''x_k''x_l'' + \ldots \qquad (3.3\text{nl})$$

— «осесохраняющее» нелинейное преобразование. То, что оно нелинейно, очевидно, а сохранение начала координат и координатных осей следует из того, что $\partial x_i'''/\partial x_j'' = \delta_{ij}$.

Неоднородное и однородное линейные преобразования легко обратить, однако этого нельзя сказать о нелинейном преобразовании.

Заметим, что, поскольку в данной главе свойства потенциальной функции рассматриваются только в точке (которую будем далее везде считать началом координат), нет необходимости в переносе начала координат посредством неоднородного преобразования (3.3ih).

Так как при использовании штрихов весьма легко допустить ошибку, то исходную систему координат в преобразованиях координат будем обозначать без штрихов, а новую систему координат — одним штрихом.

В пространстве $\mathbb{R}^n$ может быть выбрано много различных систем координат, поэтому будем представлять физический процесс или его параметры (например, потенциал V) в функциональной форме:

$$\begin{array}{ccc} & V & \\ \swarrow & & \searrow \\ f(x) & = & f'(x') \end{array} \qquad (3.4)$$

Очевидно, что вид функционального выражения зависит от выбранной системы координат. Равенство $f'(x') = f(x)$ (3.4) означает, что

$$f'(x'(x)) = f(x), \qquad (3.4')$$

т. е. один и тот же физический процесс может быть описан двумя различными функциями в двух разных системах координат.

Пример. Потенциальная функция, определенная в $\mathbb{R}^2$ и изображенная на рис. 3.1 в виде седла, вложенного в пространство $\mathbb{R}^2 \oplus \mathbb{R}^1$, может быть функционально представлена как

$$\begin{array}{ccc} & V\,(\textit{седло}) & \\ \swarrow & & \searrow \\ -2x^2 + y^2 & = & 2x'y' \end{array}$$

Системы координат (x, y) и (x', y') связаны между собой линейным преобразованием:

$$x' = (x + y)/\sqrt{2}, \quad y' = (-x + y)/\sqrt{2}.$$

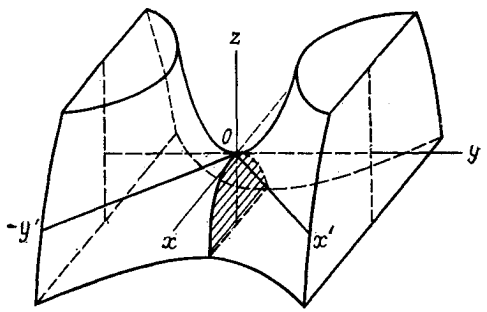


Рис. 3.1. Изображение седла в пространстве $\mathbb{R}^2 \otimes \mathbb{R}^1$ зависит от выбранной системы координат.

Два различных физических процесса в одной и той же системе координат (x, y) описываются функциями

$$\begin{array}{ccc} V & & V' \\ \downarrow & & \downarrow \\ f(x, y) & \neq & f'(x, y) \end{array}$$

Например, два различных седла на рис. 3.2 представляются двумя различными функциями в системе координат (x, y): $V_1 = -x^2 + y^2$, $V = 2xy$. Однако качественно эти седла одинаковы, так как любое из них может быть совмещено с другим посредством вращения системы координат.

В общем случае два различных процесса V и V' имеют различные функциональные формы в одной и той же системе координат: $f(x) \neq f'(x)$. Однако эти два процесса качественно подобны, если можно найти такую гладкую замену координат (3.1), что функциональная форма для V', записанная в новых координатах, в точности совпадает с функциональной формой V, записанной в исходной системе координат:

$$f'(x') = f(x). \tag{3.5}$$

В частности, это означает, что

$$f'[x'(x)] = f(x), \tag{3.6а}$$

$$f'(x') = f[x(x')]. \tag{3.6б}$$

Если f и f' качественно подобны, то все их свойства полностью совпадают [1]).

2. ПРИЛОЖЕНИЯ

2.1. Теорема о неявной функции

В качестве примера, раскрывающего смысл формулы (3.5), рассмотрим два физических процесса, которые являются результатом действия отличных от нуля сил в точке $(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0) \in \mathbb{R}^n$:

$$\begin{array}{lll} V = f(x_1, x_2, \ldots, x_n), & & \mathbf{F} = -\nabla V, \\ V' = \quad x_1, & & \mathbf{F}' = -\hat{\mathbf{e}}_1, \end{array} \tag{3.7}$$

где $\hat{\mathbf{e}}_1$ — единичный вектор, имеющий то же направление, что и первая координатная ось.

[1]) Разница между (3.4) и (3.5) аналогична разнице между «пассивной» и «активной» интерпретациями приложений теории групп в квантовой механике.

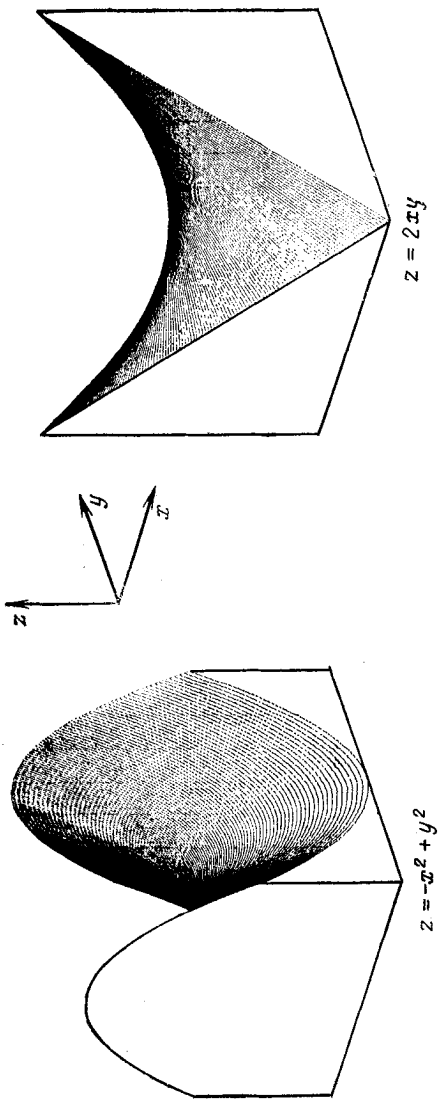


Рис. 3.2. В данной системе координат различные функции изображаются различными поверхностями. В одной и той же системе координат изображены два седла $f(x, y) = -x^2+y^2$ и $f'(x, y) = 2xy$ и оба на «пьедестале» $(x, y) \in [-1, 1] \times [-1, 1]$ с $z \in [-2, +2]$.

Поскольку в результате действия отличных от нуля сил возникает движение, любые две потенциальные функции, описывающие отличные от нуля силы в некоторой точке, должны быть качественно одинаковыми в этой точке. Для того чтобы показать, что V и V' качественно подобны, необходимо найти такое гладкое преобразование, при котором будет выполняться равенство (3.5). Одно из таких преобразований имеет вид

$$\begin{aligned} x_1' &= f(x_1, x_2, \ldots, x_n), \\ x_2' &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\vdots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n. \end{aligned} \tag{3.8}$$

Матрица Якоби этого преобразования не вырождена при условии, что

$$\det \left| \frac{\partial x_i'}{\partial x_j} \right|_{x=0} = \det \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \ldots & \frac{\partial f}{\partial x_n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \neq 0. \tag{3.9}$$

Если верхняя строка матрицы Якоби не является нулевой, то вещественные числа a_{ij} всегда можно выбрать таким образом, чтобы якобиан был отличен от нуля. В этом случае преобразование (3.8) обратимо и

$$V'(x') = x_1' = f(x_1, x_2, \ldots, x_n) = V(x). \tag{3.10}$$

Так как $V'[x'(x)] = V(x)$, а $V'(x') = V[x(x')]$, имеем

$$V[x(x')] = V'(x') = x_1'. \tag{3.11}$$

Выражения (3.11) и (2.1) в точности совпадают.

2.2. *Лемма Морса*

Если $\nabla f = 0$, то первая строка матрицы (3.9) становится нулевой и преобразование (3.8) уже не определяет новой координатной системы, оно — необратимо. Это означает, что теорема о неявной функции не может быть применена, и $f(x) \neq x_1'$. В данном случае необходимо рассмотреть в разложении функции $f(x)$ в ряд Тейлора квадратичные члены и члены более высокой степени. Если окажется, что $\det f_{ij} \neq 0$, то применима лемма Мор-

са, и функция может быть представлена в канонической форме (2.2).

Рассмотрим сначала функцию двух переменных, не имеющую постоянного и линейного членов:

$$f(x, y) = ax^2 + bxy + cy^2 + \text{Члены третьей степени} + \ldots. \quad (3.12)$$

Посредством линейного преобразования

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3.13\text{hl})$$

нетрудно привести квадратичные члены разложения к диагональному виду

$$f(\tilde{x}, \tilde{y}) = \lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + (d\tilde{x}^3 + e\tilde{x}^2\tilde{y} + f\tilde{x}\tilde{y}^2 + g\tilde{y}^3) + \\ + (h\tilde{x}^4 + \ldots), \quad (3.14)$$

где $2\lambda_1 = (a + c) + \sqrt{(a - c)^2 + b^2}$, а $2\lambda_2 = (a + c) - \sqrt{(a - c)^2 + b^2}$.

Следующий этап состоит в определении нелинейного осесохраняющего преобразования

$$\begin{aligned} x' &= x + \Delta_1 = x + (A_{20}x^2 + A_{11}xy + A_{02}y^2) + (A_{30}x^3 + \ldots), \\ y' &= y + \Delta_2 = y + (B_{20}x^2 + B_{11}xy + B_{02}y^2) + (B_{30}x^3 + \ldots). \end{aligned} \quad (3.13\text{nl})$$

Для этого необходимо так выбрать коэффициенты A_{pq}, B_{pq} $(p + q \geqslant 2)$ преобразования, задаваемого формулой (3.13nl), чтобы стало возможным привести функцию, определяемую формулой (3.14), к морсовской канонической форме. Кроме того, полагаем, что функция $f'(x', y') = \lambda_1 x'^2 + \lambda_2 y'^2$, и приравняем ее к функции $f(x, y)$ (3.14). Далее введем некоторые ограничения на коэффициенты A_{pq}, B_{pq}:

$$\lambda_1 x'^2 \lambda_2 y'^2 = \lambda_1 x^2 + \lambda_2 y^2 + (dx^3 + ex^2y + fxy^2 + gy^3) + (hx^4 + \ldots), \\ \| \\ \lambda_1 \Big(x + \sum_{p+q \geqslant 2} A_{pq} x^p y^q\Big)^2 + \lambda_2 \Big(y + \sum_{p+q \geqslant 2} B_{pq} x^p y^q\Big)^2. \quad (3.15)$$

Раскрывая круглые скобки во второй строке формулы (3.15) и приравнивая коэффициенты при соответствующих одночленах $x^p y^q$ правой и левой частей этого уравнения, для членов третьей степени получаем следующую систему уравнений:

$$\begin{aligned} x^3: &\quad \lambda_1 2A_{20} &&= d, \\ x^2y: &\quad \lambda_1 2A_{11} + \lambda_2 2B_{20} &&= e, \\ xy^2: &\quad \lambda_1 2A_{02} + \lambda_2 2B_{11} &&= f, \\ y^3: &\quad \lambda_2 2B_{02} &&= g. \end{aligned} \quad (3.16)$$

Из этой системы уравнений следует, что коэффициенты A и B (3.14) определяются не однозначно, так как для шести таких коэффициентов имеются лишь четыре коэффициента при членах третьей степени. Если перейти к рассмотрению членов четвертой степени (3.15), то детали изменятся $[h=\lambda_1(2A_{30}+A_{20}^2)]$, но результат будет тот же. Члены четвертой степени в правой части выражения (3.15) определяются ровно пятью новыми коэффициентами (при x^4, x^3y, x^2y^2, xy^3, y^4); в левой же части при проведении координатного преобразования (3.13nl) должны сохраниться члены вплоть до третьей степени. В результате возникает необходимость ввести в рассмотрение восемь новых коэффициентов ($A_{30}, A_{21}, A_{12}, A_{03}, B_{30}, B_{21}, B_{12}, B_{03}$). И далее при учете членов более высокой степени разложения (3.14) следует учитывать члены более высокой степени из преобразования (3.13nl). Число коэффициентов A_{pq}, B_{pq} при этом возрастает значительно быстрее, чем число коэффициентов, которые необходимо учитывать в разложении $f(x, y)$. Поэтому при переходе к рассмотрению членов более высокой степени всегда можно найти гладкое преобразование (3.13nl), такое, что

$$f'(x', y')=\lambda_1 x'^2+\lambda_2 y'^2=f(x, y). \tag{3.17}$$

Так как преобразование (3.13) обратимо, то с учетом выражения (3.6б) имеем

$$f[x(x', y'), y(x', y')]=f'(x', y')=\lambda_1 x'^2+\lambda_2 y'^2. \tag{3.18}$$

Аналогично обстоит дело и в случае функции n переменных. Посредством линейного преобразования (3.3hl) всегда можно привести разложение функции f в ряд Тейлора к следующему виду:

$$f(x)=\sum_{i=1}^{n}\lambda_i x_i^2+\underset{\text{степени}}{\overset{\text{Члены третьей}}{}}+\underset{\text{степени}}{\overset{\text{Члены четвертой}}{}}+\ldots. \tag{3.19}$$

Число членов степени $k>2$ в точности равно $(n+k-1)!/k!(n-1)!$. Нелинейное преобразование (3.3nl) имеет следующий вид:

$$x_i'=x_i+\Delta_i=x_i+\sum_{p+q+\ldots+r\geqslant 2}A_{i;\,pq\ldots r}x_1^p x_2^q\ldots x_n^r. \tag{3.20}$$

При любом i число членов степени k в преобразовании (3.20) равно $(n+k-1)!/k!(n-1)!$, $k>1$. На коэффициенты $A_{i;\,pq\ldots r}$ (3.20) налагаются ограничения известными нам коэффициентами разложения (3.19) и соотношением

$$f'(x')=f(x), \tag{3.21}$$

где

$$f'(x') = \sum_{i=1}^{n} \lambda_i x_i'^2 = \sum \lambda_i x_i^2 + \sum 2\lambda_i x_i \Delta_i + \sum \Delta_i^2. \quad (3.22)$$

Общее число коэффициентов при членах степени 3, 4, ..., k в разложении (3.19) равно [1])

$$N_{\mathrm{Ts}} = N_{\mathrm{Ts}}(n, k) = \binom{n+k}{k} - \binom{n+2}{2},$$

где $\binom{p}{q}$ — биномиальный коэффициент $p!/q!(p-q)!$. Члены степени 3, 4, ..., k (3.22) обусловлены наличием членов степени 2, ..., $k-1$ в разложении (3.20). Для каждого x_i' существует в точности

$$\binom{n+k-1}{k-1} - \binom{n+1}{1}$$

таких членов, а так как в данном случае имеется точно n координат, то число находящихся в нашем распоряжении коэффициентов в формуле замены переменных (3.20), входящих в члены степени $\leqslant n$ разложения (3.22), равно

$$N_{\mathrm{cv}} = N_{\mathrm{cv}}(n, k) = n\binom{n+k-1}{k-1} - n\binom{n+1}{1}.$$

Теперь легко проверить, что

$$N_{\mathrm{cv}}(n, k) \geqslant N_{\mathrm{Ts}}(n, k), \quad n \geqslant 1, \quad k \geqslant 3. \quad (3.23)$$

Из этого неравенства следует, что всегда имеется достаточная «свобода» в выборе находящихся в нашем распоряжении коэффициентов гладкой замены переменных с целью привести рассматриваемую функцию к морсовской канонической форме. Так как преобразование (3.20) обратимо, то с учетом формулы (3.6б), как и раньше, имеем

$$f(x) = f[x(x')] = f'(x') = \sum_{i=1}^{n} \lambda_i x_i'^2. \quad (3.24)$$

Выражения (3.24), (2.2а) в точности совпадают.

2.3. *Лемма расщепления*

Покажем, что если $\det f_{ij} = 0$ и, следовательно, одно или несколько собственных значений (3.22) равны нулю, то в этом случае лемма Морса неприменима и $f(x) \neq \Sigma \lambda_i x'^2$.

[1]) Ts — ряд Тейлора, cv — замена переменных.

Для этого рассмотрим функцию двух переменных, для которой $\lambda_1 = 0$ (3.15). Из формулы (3.16) следует, что коэффициенты B_{20}, B_{11}, B_{02} преобразования y' определяются однозначно, если исключить члены x^2y, xy^2, y^3 из разложения в ряд Тейлора функции $f(x, y)$. Однако при этом не удается исключить из разложения член x^3. Написав систему уравнений, аналогичную системе (3.16), для членов четвертой степени разложения в ряд Тейлора функции $f(x, y)$ найдем, что четыре находящихся в нашем распоряжении коэффициента B_{30}, B_{21}, B_{12}, B_{03} (умноженные на λ_2) линейным образом входят в выражения для коэффициентов при одночленах x^3y, x^2y^2, xy^3, y^4 и, следовательно, определяются однозначно, если исключить эти одночлены из разложения. Однако в этом случае не удается исключить член x^4. Аналогичная ситуация наблюдается и для членов степени r разложения в ряд Тейлора функции $f(x, y)$; r находящихся в нашем распоряжении коэффициентов $B_{r-1,0}$, ..., $B_{0,r-1}$ (умноженные на λ_2) входят линейным образом в выражения коэффициентов r одночленов $x^{r-1}y$, ..., y^r, так что от этих одночленов можно избавиться. Таким образом, существует замена переменных $y \to y'$ (3.13nl), позволяющая исключить все одночлены вида x^py^q, $p + q > 2$, $q \geqslant 1$. Следовательно, при $\lambda_1 = 0$

$$f(x, y) \doteq px'^3 + qx'^4 + \ldots + \lambda_2 y'^2 = f_{NM}(x') + \lambda_2 y'^2. \quad (3.25)$$

Заметим, что, приводя функцию $f(x, y)$ к виду (3.25), мы совсем не использовали имеющуюся возможность выбора нелинейной замены координат $x \to x'$, позволяющей привести рассматриваемую функцию к канонической форме.

Аналогично можно поступить и при поиске канонической формы функции n переменных в неморсовской критической точке, в которой имеется в точности l «плохих» переменных.

Очевидно, что в разложении в ряд Тейлора функции n переменных состояния в неморсовской критической точке $x_1 = x_2 = \ldots = x_n = 0$ отсутствуют все члены первой степени ($\nabla f = 0$). При исследовании свойств функции в этой точке постоянные члены разложения не играют существенной роли. Поэтому можно считать, что разложение в ряд Тейлора начинается с квадратичных членов. Тогда посредством линейного преобразования (3.3hl) можно привести квадратичные члены разложения к каноническому диагональному виду $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2$.

Теперь можно осуществить нелинейное преобразование $x_i \to x_i'$ (3.3nl) и посмотреть, насколько вид функции в новой системе координат близок к канонической форме. Для этого

сравним соответствующие члены в обеих частях уравнения:

$$\sum_{i=1}^{n} \lambda_i x_i'^2 \qquad = f(x_1, x_2, \ldots, x_n)$$

$$\downarrow \qquad\qquad \downarrow$$

$$\sum_{i=1}^{n} \lambda_i x_i^2 + 2\sum_{i=1}^{n} \lambda_i x_i \Delta_i + \sum_{i=1}^{n} \lambda_i \Delta_i^2 = \sum_{i=1}^{n} \lambda_i x_i^2 + \frac{1}{3!} \sum_{ijk} f_{ijk} x_i x_j x_k + \ldots . \tag{3.26}$$

Если все собственные значения функции отличны от нуля, то нелинейную часть преобразования (Δ_i) можно выбрать таким образом, чтобы исключить все члены третьей степени и выше. (Как это сделать, уже было показано в разд. 3.) Если же хотя бы одно собственное значение функции равно нулю (скажем, λ_1), то член x_1^2 отсутствует в обеих частях уравнения (3.26). Член разложения вида $\lambda_i x_i \Delta_i$ порождает лишь члены вида $x_1^{p_1} \ldots x_i^{p_i} \ldots$ $\ldots x_n^{p_n}$, где $p_i \geqslant 1$, $\sum_{j=1}^{n} p_j > 3$ (так как Δ_i начинается с квадратичных членов). Следовательно, все члены вида $x_1^{p_1} \ldots x_n^{p_n}$, исключая члены, у которых $p_2 = p_3 = \ldots = p_n = 0$, могут быть удалены с помощью нелинейной замены координат и

$$f(x_1, x_2, \ldots, x_n) \doteq (t_3 x_1'^3 + t_4 x_1'^4 + \ldots) + \sum_{i=2}^{n} \lambda_i x_i'^2 =$$

$$= f_{NM}(x_1') + \sum_{i=2}^{n} \lambda_i x_i'^2. \tag{3.27i}$$

Подобные рассуждения применимы и тогда, когда $l (l > 1)$ собственных значений обращаются в нуль. Если $\lambda_1 = \lambda_2 = \ldots$ $\ldots = \lambda_l = 0$ (3.26), то нелинейную замену переменных $x_i \to$ $\to x_i' = x_i + \Delta_i$, $i = l + 1, \ldots, n$ можно выбрать таким образом, чтобы удалить в формуле (3.26) все одночлены вида $x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n}$, за исключением тех, у которых $p_{l+1} = \ldots = p_n = 0$. Как следствие, получим следующее представление функции:

$$f(x_1, x_2, \ldots, x_n) \doteq \left(\sum_{p_1 + \ldots + p_l \geqslant 3} t_{p1 \ldots pl} x_1'^{p_1} \ldots x_l'^{p_l} \right) + \sum_{i=l+1}^{n} \lambda_i x_i'^2 =$$

$$= f_{NM}(x_1', \ldots, x_l') + \sum_{i=l+1}^{n} \lambda_i x_i'^2. \tag{3.27l}$$

Выражения (3.27l), (2.3а) в точности совпадают.

◇◇◇ В процессе доказательства леммы расщепления (3.271), для того чтобы исключить все «смешанные» члены [1]) из разложения в ряд Тейлора, использовалось нелинейное преобразование лишь «хороших» переменных. Аналогичным образом можно исключить из разложения все члены, состоящие из «хороших» переменных степени больше двух. Однако при этом мы нигде не использовали возможность осуществить нелинейное преобразование «плохих» переменных [2]). По существу представление (3.271) может быть получено в предположении, что $\Delta_1 = \Delta_2 = \ldots = \Delta_l = 0$, т. е. $x_1' = x_1$, $x_2' = x_2$, ..., $x_l' = x_l$.

В связи с этим напомним, что неморсовские критические точки в общем случае могут типично встречаться лишь при рассмотрении семейств функции или функций $f(x; c)$, зависящих от управляющих параметров $c \in \mathbb{R}$, поскольку в критической точке собственные значения $\lambda_i(c)$ матрицы устойчивости $f_{ij} = \partial^2 f(x; c)/\partial x_i\, \partial x_j$ являются функциями управляющих параметров. Поэтому важно знать, сколько собственных значений могут одновременно обращаться в нуль в типичном k-параметрическом семействе функций.

Согласно лемме расщепления (2.3а), в окрестности неморсовской критической точки, в которой одновременно обращаются в нуль l собственных значений, всегда имеется возможность выбора l «плохих» переменных и $n - l$ «хороших». В такой координатной системе матрица устойчивости будет иметь блочно-диагональную структуру

$$V_{ij} \to \begin{bmatrix} \dfrac{\partial^2 f_{NM}}{\partial x_i\, \partial x_j} & \bigcirc \\ \bigcirc & \dfrac{\partial^2 f_M}{\partial x_r\, \partial x_s} \end{bmatrix} \begin{matrix} 1 \leqslant i,\ j \leqslant l, \\ \\ l+1 \leqslant r,\ s \leqslant n. \end{matrix} \tag{3.28}$$

Вещественная симметрическая $l \times l$-матрица $\partial^2 f_{NM}/\partial x_i \partial x_j$ имеет в точности $l(l+1)/2$ независимых элементов, которые являются функциями k управляющих параметров. Следовательно, естественно ожидать, что все $l(l+1)/2$ элементов матрицы $\partial^2 f_{NM}/\partial x_i\, \partial x_j$ не могут одновременно обратиться в нуль, если число их не меньше, чем число управляющих параметров. Поэтому при исследовании k-параметрического семейства функций можно ожидать, что существует l одновременно обращающихся в нуль

[1]) «Смешанные» члены — это члены разложения, являющиеся произведением как «хороших», так и «плохих» переменных.

[2]) В следующих разделах мы существенно используем эту возможность при поиске канонических форм для неморсовских функций $f_{NM}(x_1, \ldots, x_l)$.

собственных значений только при условии, что l и k связаны следующим соотношением:

$$k \geqslant l(l+1)/2. \tag{3.29}$$

2.4. $l = 1$ (одно вырожденное собственное значение)

В типичном k-параметрическом семействе функций ($k \geqslant 1$) в общем случае можно встретить только неморсовскую критическую точку, для описания которой требуется всего лишь одна «плохая» переменная, которую мы обозначим x. Вблизи такой точки неморсовская функция в вырожденной критической точке $x = 0$ (2.3а) может быть разложена в ряд Тейлора

$$f_{NM}(x;\ c) = t_0 + t_1 x + t_2 x^2 + t_3 x^3 + t_4 x^4 + \ldots . \tag{3.30}$$

В данном случае постоянный член разложения не играет существенной роли, линейный член обращается в нуль, так как $x = 0$ — критическая точка (т. е. $t_0 = t_1 = 0$), а оставшиеся коэффициенты ряда Тейлора $t_2(c) = \lambda_2(c)$, $t_3(c)$, ..., $t_n(c)$, ... являются функциями k управляющих параметров c_α. Поэтому можно предположить, что имеются такие значения управляющих параметров, при которых в точности k коэффициентов $t_j(c)$ обращаются в нуль. Если выбрать точку $c \in \mathbb{R}^k$ таким образом, что $t_2 = t_3 = \ldots = t_{k+1} = 0$ в точке $c = c^0$, то первым отличным от нуля членом разложения (3.30) в ней будет x^{k+2}, т. е.

$$f_{NM}(x;\ c^0) = t_{k+2}x^{k+2} + \text{Члены более высокой степени.} \tag{3.31}$$

Покажем, что посредством гладкой замены переменных можно удалить все члены более высокой степени и привести функцию f_{NM} к канонической форме $\pm x^{k+2}$.

Допустим, что

$$f(x) = t_n x^n + t_{n+1}x^{n+1} + \ldots, \tag{3.32}$$

$$x' = A_1 x + A_2 x^2 + A_3 x^3 + \ldots . \tag{3.33}$$

Поскольку необходимо показать, что функция $f(x)$ качественно аналогична функциям $\pm x^n$, пусть $f'(x') = \pm(x')^n$ и $f'(x') = f(x)$:

$$\pm(A_1 x + A_2 x^2 + A_3 x^3 + \ldots)^n = t_n x^n + t_{n+1}x^{n+1} + \ldots . \tag{3.34}$$

Повторяя предыдущий ход рассуждений [ср. с формулой (3.16)], найдем

$$x^n\colon \quad \pm A_1^n = t_n, \tag{3.35n}$$

если n четно, то знак $+$ или $-$ выбирается в зависимости от того, положительно или отрицательно t_n; если n нечетно, то все-

гда можно выбрать знак плюс:

$$x^{n+1}: \pm nA_1^{n-1}A_2 = t_{n+1}, \qquad (3.35\,\mathrm{n}+1)$$

$$x^{n+2}: \pm\left(nA_1^{n-2}A_3 + \binom{n}{2}A_1^{n-2}A_2^2\right) = t_{n+2}. \qquad (3.35\,\mathrm{n}+2)$$

Как только коэффициент A_1 определен из уравнения (3.35n), A_2 может быть однозначно найден из уравнения (3.35 n + 1), которое линейно по этому неизвестному. По существу как только коэффициенты $A_1, A_2, \ldots, A_j$ определены из (3.35 n → n + j — — 1), коэффициент A_{j+1} однозначно определяется из уравнения (3.35 n + j), куда единственное неизвестное A_{j+1} входит линейно.

Таким образом, преобразование (3.33) может быть найдено. Тогда из формулы (3.6б) имеем

$$f_{NM}[x(x')] = f'_{NM}(x') = \pm (x')^n. \qquad (3.36)$$

Следовательно, при k управляющих параметрах и лишь одном обращающемся в нуль собственном значении ведущий член разложения в ряд Тейлора функции $f_{NM}(x;\ c)$ в общем случае не хуже, чем $\pm x^{k+2}$, а неморсовская функция может быть приведена в вырожденной критической точке к канонической форме $\pm x^{k+2}$.

◇◇◇ Иногда бывает полезно брать за каноническую форму $\pm x^n/n$. Это может быть легко достигнуто посредством преобразования $x' = n^{1/n}x$.

Вывод. Для исключения начальных членов разложения в ряд Тейлора функции в критической точке могут быть использованы управляющие параметры, а для удаления последних членов разложения в ряд Тейлора — гладкая замена переменных. Участок разложения в ряд Тейлора, на котором эти операции выполняются одновременно, называется *ростком функции в неморсовской критической точке.* Ростками k-параметрических семейств функций с единственной переменной состояния являются $\pm x^{j+2} (j \leqslant k)$ (табл. 2.2).

2.5. $l = 2$ (два вырожденных собственных значения)

В k-параметрическом семействе функций ($k \geqslant 3$) в общем случае возможно обнаружить лишь неморсовские критические точки с двумя «плохими» переменными x и y.

Определяющая часть разложения функции в критической точке $(x, y) = (0, 0)$ будет следующая:

$$f(x, y) = t_{20}x^2 + t_{11}xy + t_{02}y^2 + t_{30}x^3 + t_{21}x^2y + \\ + t_{12}xy^2 + t_{03}y^3 + t_{40}x^4 + \ldots . \qquad (3.37)$$

Чтобы исключить квадратичные члены с коэффициентами (t_{20}, t_{11}, t_{02}), используем три управляющих параметра, а чтобы привести кубические члены разложения (3.37) к какой-либо канонической форме и исключить оставшиеся члены (четвертой степени и выше) разложения в ряд Тейлора — гладкую замену переменных. Прежде чем это сделать, целесообразно привести возможные канонические формы кубических членов:

$$f'(x', y') = x'^3 + ay'^3, \tag{3.38а}$$

$$f'(x', y') = x'^2y' + by'^3. \tag{3.38б}$$

Теперь, используя каноническую форму (3.38б), поступим точно так же, как при рассмотрении морсовского случая для двух переменных. Сначала осуществим линейное преобразование (3.13hl) и сравним $\tilde{x}^2\tilde{y} + b\tilde{y}^3$ с кубическими членами разложения (3.37). Получаемая при этом система из четырех уравнений с пятью неизвестными

$$\begin{aligned}
x^3\colon\ & \alpha^2\gamma + b\gamma^3 = t_{30},\\
x^2y\colon\ & 2\alpha\beta\gamma + \alpha^2\delta + 3b\gamma^2\delta = t_{21},\\
xy^2\colon\ & \beta^2\gamma + 2\alpha\beta\delta + 3b\gamma\delta^2 = t_{12},\\
y^3\colon\ & \beta^2\delta + b\delta^3 = t_{03}
\end{aligned} \tag{3.39}$$

всегда разрешима относительно неизвестных α, β, γ, δ. Разложение (3.37) при этом в данной вырожденной критической точке имеет вид

$$\tilde{f}(\tilde{x}, \tilde{y}) = \tilde{x}^2\tilde{y} + b\tilde{y}^3 + (t_{40}\tilde{x}^4 + t_{31}\tilde{x}^3\tilde{y} + \ldots) + \ldots. \tag{3.40}$$

Можно показать, что если $b \neq 0$, то члены четвертой степени и выше могут быть исключены посредством гладкой замены переменных. Преобразование (3.13nl) произведем в каноническом виде $f'(x', y') = x'^2y' + by'^3$, а затем, раскрыв скобки, сравним соответствующие члены. В результате для находящихся в нашем распоряжении коэффициентов замены переменных (3.13nl) получим ряд простых уравнений

$$\begin{aligned}
x^4\colon\ & B_{20} = t_{40},\\
x^3y\colon\ & 2A_{20} + B_{11} = t_{31},\\
x^2y^2\colon\ & 2A_{11} + B_{02} + 3bB_{20} = t_{22},\\
xy^3\colon\ & 2A_{02} + 3bB_{11} = t_{13},\\
y^4\colon\ & + 3bB_{02} = t_{04}.
\end{aligned} \tag{3.41}$$

Поскольку имеются точно шесть неопределенных коэффициентов и лишь пять коэффициентов ряда Тейлора, то все квадра-

тичные члены могут быть исключены. Аналогичные рассуждения могут быть проведены и в случае членов степени 5, 6, ...

Наконец, если $b \neq 0$, то для переменных x', y' может быть введен новый масштаб, такой, чтобы получить канонические ростки

$$\begin{aligned} D_{+4}&: \ x'^2y' + y'^3, \ b > 0, \\ D_{-4}&: \ x'^2y' - y'^3, \ b < 0. \end{aligned} \tag{3.42}$$

В случае 4-параметрического семейства функций три степени свободы могут быть использованы для обращения в нуль трех квадратичных коэффициентов, а четвертая степень свободы — для исключения параметра b из канонической формы (3.38б). В этом случае гладкая замена переменных (3.13nl) не позволяет удалить все члены выше третьей степени. Из анализа последней строки системы уравнений (3.41) следует, что члены x^4, x^3y, x^2y^2 и xy^3 можно исключить из разложения, а y^4 нет. Поэтому в случае 4-параметрического семейства получим функции вида $x'^2y' + b'y'^4$, где в общем случае $b' \neq 0$, так как все четыре степени свободы в пространстве управляющих параметров были уже использованы. Когда $b' \neq 0$, для переменных x' и y' всегда можно ввести новый масштаб, такой, чтобы получить канонический росток

$$\pm D_5: \ \pm (x'^2y' + y'^4). \tag{3.43}$$

В 5-параметрическом семействе функций имеется еще одна дополнительная степень свободы, которая может быть использована для обращения в нуль коэффициента b'. Тогда с помощью уже известных рассуждений, получим, что член $b'y'^5$ не может быть удален, и будем иметь следующие канонические ростки:

$$\begin{aligned} D_{+6}&: \ x'^2y' + y'^5, \ b'' > 0, \\ D_{-6}&: \ x'^2y' - y'^5, \ b'' < 0. \end{aligned} \tag{3.44}$$

Продолжая подобным образом, получим, что в k-параметрических семействах функций можно встретить следующие канонические ростки:

$$D_{\pm(k+1)}: \ x'^2y' \pm y'^k. \tag{3.45}$$

Совершенно аналогичные рассуждения могут быть проведены при выводе канонической формы (3.38а). В 3-параметрическом семействе функций в общем случае $a \neq 0$. Вводя новый масштаб, координату y' можно преобразовать так, что получим каноническую форму

$$D_{+4}: \ x'^3 + y'^3. \tag{3.46}$$

Функции вида (3.38а) при $a \neq 0$ эквивалентны функциям (3.38б) при $b > 0$. Так как при этом имеется соотношение между a и b ($a^2 \sim b$), то это дает в общем случае два дополнительных управляющих параметра для обращения в нуль a (только один дополнительный параметр для b), так что соответствующий канонический росток 5-параметрического семейства функций имеет вид

$$E_{\pm 6}: \quad x'^3 \pm y'^4. \tag{3.47}$$

2.6. $l = 2$, $k \geqslant 7$

Если пространство управляющих параметров имеет размерность, превышающую шесть, то имеющиеся семь степеней свободы могут быть использованы для удаления членов второй и третьей степени из разложения в ряд Тейлора (3.37). И в этом случае можно выполнить

— линейное преобразование (3.13hl), приводящее члены четвертой степени к некоторой канонической форме;

— нелинейное преобразование (3.13nl), чтобы избавиться от членов более высокой степени.

Однако линейное преобразование (3.13hl) имеет четыре параметра (α, β, γ, δ), в то время как члены четвертой степени определяются пятью коэффициентами (t_{40}, t_{31}, t_{22}, t_{13}, t_{04}). Поэтому невозможно привести члены четвертой степени разложения к канонической форме, в которой все члены имеют канонические коэффициенты ± 1 или 0. Лучшее, что может быть достигнуто, — это каноническая форма, зависящая от одного ($1 = 5-4$) параметра,

$$\pm x'^4 + a x'^2 y'^2 \pm y'^4, \quad a^2 \neq 4. \tag{3.48}$$

◇◇◇ Каноническая форма кубических коэффициентов совершенно не зависит от параметров (для b может быть назначено ± 1), так как имеется достаточно свободы в проведении линейного преобразования (четыре степени), для того чтобы задать однозначно четыре кубических коэффициента.

2.7. $l = 3$, $k \geqslant 6$

Если число управляющих параметров больше пяти, то в k-параметрическом семействе функций в общем случае можно встретить лишь неморсовские критические точки с тремя «плохими» переменными x, y, z. Разложение в ряд Тейлора функции $f_{NM}(x, y, z; c)$ в критической точке $(x, y, z) = (0, 0, 0)$ имеет шесть квадратичных членов, десять кубических членов, ... Шесть из k имеющихся степеней свободы могут быть использо-

ваны для того, чтобы обратить в нуль шесть квадратичных коэффициентов, а нелинейное преобразование можно попытаться использовать для приведения оставшейся части разложения в ряд Тейлора к стандартному виду. Рассуждения при этом совершенно аналогичны рассуждениям в случае двух переменных:

— найти линейное преобразование $(x, y, z) \to (\tilde{x}, \tilde{y}, \tilde{z})$, приводящее кубические члены к стандартному виду;

— найти нелинейное осесохраняющее преобразование, удаляющее члены четвертой степени и выше.

Однако найденное линейное преобразование имеет девять степеней свободы, так как оно полностью определяется действительной квадратной матрицей порядка три, в то время как коэффициентов, параметризующих кубические члены, десять. Следовательно, любая каноническая форма для кубических членов разложения должна зависеть от одного действительного параметра. Одна из канонических форм имеет вид

$$T_{3,3,3}: \quad x'^3 + y'^3 + z'^3 + ax'y'z', \quad a^3 \neq -27. \tag{3.49}$$

2.8. $l \geqslant 4$

Если $k \geqslant l(l+1)/2$, то в k-параметрическом семействе функций можем ожидать лишь неморсовские критические точки, имеющие b «плохих» переменных [сравните с (3.29)]. В разложении в ряд Тейлора такой функции в неморсовской критической точке имеется $l(l+1)/2$ квадратичных членов, $l(l+1)\times$ $\times(l+2)/6$ кубических членов и т. д. После того как с помощью управляющих параметров удалим квадратичные члены, можно попытаться найти каноническую форму кубических членов. Линейная часть нелинейного преобразования имеет размерность l^2, так как оно задается квадратной матрицей порядка l. Это дает возможность получить каноническую форму, зависящую от

$$\frac{l(l+1)(l+2)}{6} - l^2 = \frac{l(l-1)(l-2)}{6} \tag{3.50}$$

действительных параметров. Эти параметры будем называть *модулями*; росток, зависящий от t модулей, назовем *t-модальным*, а росток, не зависящий от модулей, — *нуль-модальным* или *простым*.

Если известно, от скольких модулей зависит росток, то иногда легко догадаться, и какова его каноническая форма. Например, в случае k-параметрического семейства, если $k \geqslant$ $\geqslant 4(4+1)/2 = 10$, в общем случае невозможно встретить неморсовскую критическую точку с четырьмя вырожденными собственными значениями. В 10-параметрическом семействе все 10 квадратичных коэффициентов могут одновременно обра-

щаться в нуль. Поэтому каноническая форма кубических коэффициентов будет обязательно зависеть от четырех модулей (3.50). Это дает право предложить следующую каноническую форму для кубических членов:

$$f_{NM}(x, y, z, w) =$$

$$= x^3 + y^3 + z^3 + w^3 + a_1yzw + a_2xzw + a_3xyw + a_4xyz. \quad (3.51)$$

Теперь можно проверить, что гладкая замена переменных вида (3.2) с $A_i = 0$ $(i = 1, 2, 3, 4)$ позволяет исключить все члены четвертой степени и выше, если только a_i (3.51) случайно не подчиняются другим обязательным соотношениям.

В более общем случае l «плохих» переменных состояния в k-параметрическом семействе $k \geqslant l(l+1)/2$ можно также найти каноническую форму для кубических членов. Она, конечно, будет зависеть от $\binom{l}{3}$ модулей. Биномиальная природа этой числовой зависимости и вид формы (3.49) при $l = 3$ и (3.51) при $l = 4$ позволяют предположить, что общая каноническая форма при $l \geqslant 3$ имеет следующее представление:

$$f_{NM}(x_1, x_2, \ldots, x_l) = \sum_{i=1}^{l} x_i^3 + \sum_{1 \leqslant i < j < k \leqslant l} a_{ijk} x_i x_j x_k. \quad (3.52)$$

Здесь точно $\binom{l}{3}$ различных коэффициентов a_{ijk}. Используя предполагаемый вид канонической формы и простые алгебраические преобразования, легко проверить, что гладкое преобразование переменных вида (3.2nl) может быть с успехом использовано для удаления всех членов четвертой степени и выше, если коэффициенты a_{ijk} не подчиняются другим обязательным соотношениям.

Таким образом, в канонической форме кубических членов модули встречаются лишь тогда, когда количество «плохих» переменных состояния слишком велико [$l > 2$ (3.50)] или если $l = 2$, но количество управляющих параметров слишком велико [$k \geqslant 7$ (3.48)]. Модули совсем не встречаются в канонических формах при $l = 1$ независимо от значения k.

Простые ростки (т. е. ростки, не зависящие от модулей) встречаются, когда $l = 1$, k-произвольно или $l = 2$, $k < 7$, и полностью отсутствуют при значениях $l > 3$. Однако при $l = 3$ они могут встречаться только при $k = 6$. Таким образом, при $k \leqslant 5$ все возможные ростки будут простые. Поэтому для списка элементарных (простых) катастроф Тома $k = 5$ является границей. При $k \geqslant 6$ существуют как устойчивые простые, так и непростые ростки (табл. 3.1). Простые ростки

встречаются только при $l = 1$ и $l = 2$:

$$\begin{aligned} A_{\pm k}&: \pm x^{k+1}, && k = 2, 3, \ldots, \\ D_{\pm k}&: x^2 y \pm y^{k-1}, && k = 4, 5, \ldots, \\ E_{\pm 6}&: x^3 \pm y^4, \\ E_{\pm 7}&: \pm (x^3 y + y^4), \\ E_8&: x^3 + y^5. \end{aligned} \tag{3.53}$$

Таблица 3.1. **Простые и непростые ростки катастроф** [1])

k	1	2	3	4	5	6	7	8	9	10	11
l											
1	S	S	S	S	S	S	S	S	S	S	S
2			S	S	S	S	S, M	S, M	S, M	S, M	S, M
3						M	M	M	M	M	M
4										M	M

[1]) В k-параметрическом семействе функций с неморсовской критической точкой, имеющем l «плохих» направлений, канонические ростки могут быть простыми (S) или могут зависеть от модуля (M). При $k \leqslant 5$ встречаются лишь простые ростки.

3. ВЫВОДЫ

Локальное поведение потенциальной функции определяется начальными членами разложения ее в ряд Тейлора. Две функции качественно подобны, если они связаны с помощью гладкой замены переменных (3.6). Такое преобразование координат весьма часто можно использовать и для удаления последних членов разложения в ряд Тейлора. В этом случае функция имеет такие же свойства, что и основные члены разложения ее в ряд Тейлора, т. е. обычный многочлен.

Управляющие параметры могут быть использованы для «аннигиляции» начальных членов разложения функции в ряд Тейлора и тем самым для изменения ее свойств, а замена переменных — для сокращения ряда Тейлора. Члены, располагающиеся между членами, удаляемыми при помощи управляющих параметров, и членами, удаляемыми при помощи замены координат, приводятся к некоторой стандартной, *канонической форме.*

Отдельная функция, не зависящая от управляющих параметров, в общем случае будет иметь лишь некритические и изолированные критические точки и соответствующие канонические формы (2.1) и (2.2).

Неморсовские критические точки могут встречаться в семействах функций. Для таких семейств существует общая каноническая форма (2.3а).

Литература

1. Arnol'd V. I., Critical Points of Smooth Functions, Proc. Int. Cong. Math., Vancouver, 1974, pp. 19—75.
2. Арнольд В. И. Критические точки гладких функций и их нормальные формы. — УМН, 1975, 30 : 5, 3—65.
3. Poston T., Stewart N. I., Catastrophe Theory and Application, London: Pitman, 1978. [Имеется перевод: Постон Т., Стюарт И. Теория катастроф и ее приложения. — М.: Мир, 1980.]

4

ЗАМЕНА ПЕРЕМЕННЫХ. ВОЗМУЩЕНИЯ

Неморсовские функции, присутствующие в канонической форме (2.3) (левый столбец табл. 2.1), встречаются вместе с возмущающими функциями (правый столбец табл. 2.1). Однако если для канонической (2.1) и морсовской канонической (2.2) форм никаких дополнительных членов не требуется, то иначе обстоит дело в случае канонической формы Тома. Объясняется это тем, что возмущения не влияют на свойства функции в некритической или в морсовской критической точке, но могут влиять на свойства функции в вырожденной критической точке. По существу список возмущений требуется именно для того, чтобы описать все изменения свойств, которые возмущение может вызвать в окрестности вырожденной критической точки.

Очень часто бывает полезно знать свойства функции или семейства функций в окрестности некоторой точки. Если $f(x; c)$ — семейство функций, имеющее некоторую некритическую точку, морсовскую критическую точку или вырожденную критическую точку в $x^0 \in \mathbb{R}^n$ при $c = c^0 \in \mathbb{R}^k$, то разность

$$p(x_1, \ldots, x_n; c_1, \ldots, c_k) = f(x; c) - f(x; c^0) \tag{4.1}$$

можно рассматривать как возмущение функции $f(x; c^0)$ в окрестности точки (x^0, c^0).

Для того чтобы выяснить, как возмущения влияют на свойства функции в окрестности точки x^0, вполне достаточно изучить лишь возмущение канонической формы функции $f(x; c^0)$ в окрестности точки x^0. Канонические формы возмущений функций находятся точно так же, как канонические формы самих функций. Для удобства будем помещать точку $x^0 \in \mathbb{R}^n$ в начало координат пространства $\mathbb{R}^n$. Удобно также брать точку $c^0 = 0$ в пространстве $\mathbb{R}^k$.

Начнем рассмотрение с канонической формы, в которой уже отсутствуют несколько ведущих членов разложения функции в ряд Тейлора и все члены «хвоста» разложения, и, выбрав произвольное возмущение, выясним, какие из возмущающих членов могут быть удалены посредством некоторой гладкой замены переменных.

1. ВОЗМУЩЕНИЯ

Действие возмущающей функции $p(x; c)$ вида (4.1) на функцию $f(x; 0)$ наиболее просто изучить, используя разложения функции $p(x; c)$ в ряд Тейлора в точке $x = 0 \in \mathbb{R}^n$:

$$p(x; c) = p(0; c) + p_i x_i + p_{ij} x_i x_j + \ldots . \tag{4.2}$$

Поскольку свойства функции в точке независимы от ее значения в этой точке, член $p(0; c)$ разложения (4.2) можно считать равным нулю. Все производные берутся в точке $x = 0$;

суммирование предполагается по всем двойным индексам. Если $p(x; c)$ зависит от k управляющих параметров, то и коэффициенты ряда Тейлора p_i, p_{ij}, ... также зависят от этих параметров.

2. ПРИЛОЖЕНИЯ

2.1. Форма теоремы о неявной функции

Рассмотрим возмущение в некритической точке. В окрестности такой точки функция $f(x)$ имеет каноническую форму $f(x) = x_1$, так что возмущенная функция имеет вид

$$F = f + p = x_1 + p. \tag{4.3}$$

В общем случае $\nabla F \neq 0$ в точке $x = 0$ и теорема о неявной функции применима к функции $F(x)$ с таким же успехом, как

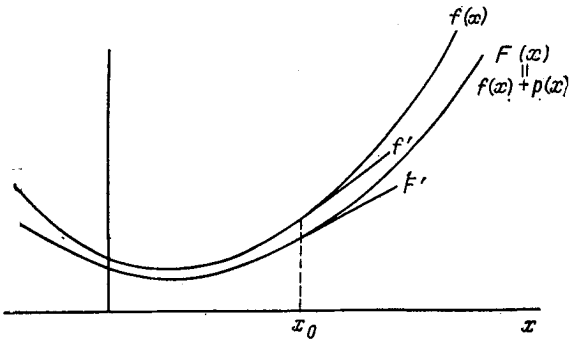


Рис. 4.1. Возмущение функции в некритической точке.

Значение функции и ее градиента слегка изменяется; локальных качественных изменений не наблюдается.

и к функции $f(x)$. Это означает, что возмущенная функция $F(x)$ может быть приведена к канонической форме $F(x) \doteq x_1$ с помощью процедуры, описанной в гл. 3.

Действие возмущения на функцию f в точке, где $\nabla f \neq 0$, двояко (рис. 4.1): с одной стороны, значение возмущенной функции слегка изменяется [на $O(p(0; c))$], а с другой стороны, градиент $\nabla(f + p)$ также слегка изменяется [на $O(p_i)$]. В результате возмущение функции $f(x)$ в некритической точке не вызывает качественных изменений в поведении $f(x)$ в окрестности этой точки.

2.2. *Морсовские формы*

Рассмотрим, что происходит с функцией, если возмущение наблюдается в окрестности морсовской критической точки. В качестве основной канонической формы функции f в окрестности изолированной критической точки $x = 0$ выберем форму (2.2а). Тогда для возмущенной функции $F = f + p$ имеем следующее разложение в ряд Тейлора:

$$F = p(0;\ c) + p_i x_i + \\ + (\lambda_i \delta_{ij} + p_{ij}) x_i x_j + \text{Члены третьей степени} + \ldots . \quad (4.4)$$

В общем случае в точке $x = 0$ градиент $\nabla F = \nabla(f + p) = \\ = \nabla p \neq 0$, так что на первый взгляд кажется, что возмущение влияет на поведение функции $f(x)$ — равновесное состояние сменяется неравновесным в точке $x = 0$.

Попытаемся, однако, проанализировать действие возмущения на поведение функции $f(x)$ под другим углом зрения. Предположим, что мы имеем дело с функцией единственной переменной состояния ($n = 1$). В этом случае разложение $F(x)$ в ряд Тейлора (4.4) будет иметь вид ($\lambda \neq 0$)

$$F(x) = p(0;\ c) + p_1 x + (\lambda + p_2) x^2 + p_3 x^3 + \ldots, \quad (4.5)$$

где все коэффициенты возмущения p_j можно считать достаточно малыми. Поскольку $p_1 \neq 0$, можно ввести новую независимую переменную y (гл. 3):

$$y = p_1 x + (\lambda + p_2) x^2 + p_3 x^3 + \ldots . \quad (4.6)$$

Обратное преобразование существует и может быть записано в виде

$$x = A_1 y + A_2 y^2 + A_3 y^3 + \ldots . \quad (4.7)$$

Производя подстановку (4.6) в (4.7) и принимая во внимание тождество $x = x$, легко найти коэффициенты A_1, A_2, ... Для этого приравниваем коэффициенты при членах, содержащих x в одинаковой степени в обеих частях уравнения, и получаем следующую последовательность уравнений со все возрастающим числом коэффициентов A_i:

$$\begin{aligned} x^1&: \ 1 = A_1 p_1, \\ x^2&: \ 0 = A_1(\lambda + p_2) + A_2 p_1^2, \\ x^3&: \ 0 = A_1 p_3 + 2A_2 p_1(\lambda + p_2) + A_3 p_1^3, \\ &\vdots \end{aligned} \quad (4.8)$$

Отсюда можно однозначно определить коэффициенты

$$\begin{aligned} A_1 &= 1/p_1, \\ A_2 &= -(\lambda + p_2)/p_1^3, \\ A_3 &= 2(\lambda + p_2)^2 p_1^{-5} - p_3 p_1^{-4}, \\ &\vdots \end{aligned} \tag{4.9}$$

И тем не менее данный подход неприемлем, так как чрезвычайно трудно построить обратное преобразование $x(y)$ (4.7) даже при малом возмущении. Более того, обратное преобразование (4.7) разрывно в окрестности точки $p_1 = 0$. Поскольку все трудности проистекают из-за того, что при поиске преобразования, позволяющего исключить квадратичные члены и члены более высокой степени, коэффициент p_1 входит в знаменатель формул (4.9), необходимо найти преобразование, обладающее обратным преобразованием, непрерывным при $p_1 \to 0$, и позволяющее исключить все члены разложения, начиная с кубических.

Для того чтобы выяснить, так ли это на самом деле, рассмотрим преобразование вида (3.3n1)

$$x = y + B_2 y^2 + B_3 y^3 + \ldots, \tag{4.10}$$

в результате которого исключаются члены выше второй степени из разложения (4.5) функции $F(x)$. Сделав подстановку (4.10) в (4.5) и приравняв коэффициенты y^3, y^4, ... нулю, получим следующую систему уравнений:

$$\begin{aligned} y^1&: \text{ Произвольная константа} = p_1, \\ y^2&: \text{ Произвольная константа} = p_1 B_2 + (\lambda + p_2), \\ y^3&: \quad 0 \quad = p_1 B_3 + (\lambda + p_2) \cdot 2B_2 + p_3, \\ y^4&: \quad 0 \quad = p_1 B_4 + (\lambda + p_2) \cdot (2B_3 + B_2^2) + \\ &\qquad\qquad + p_3 \cdot 3B_2 + p_4. \end{aligned} \tag{4.11}$$

Если $p_1 = 0$, то система уравнений, определяемая коэффициентами при членах выше второй степени y^3, y^4, ..., является линейной и может быть легко решена:

$$\begin{aligned} B_2 &= -\frac{p_3}{2(\lambda + p_2)}, \\ B_3 &= \frac{5p_3^2}{8(\lambda + p_2)^2} - \frac{p_4}{2(\lambda + p_2)}. \end{aligned} \tag{4.12}$$

В этом случае все коэффициенты B_2, B_3, ... определены корректно, так как единственный множитель, который встречается

в знаменателях, — это отличная от нуля сумма $(\lambda + p_2)$. Если же $p_1 \neq 0$, система уравнений может быть также разрешена и коэффициенты $B_2(p_1)$, $B_3(p_1)$, ... зависят от p_1 непрерывным образом (степенные ряды). Существенным моментом является то, что, если $\lambda_1 \neq 0$ (4.5), всегда может быть найдено гладкое обратимое преобразование (4.10), удаляющее все члены выше второй степени.

◇◇◇ Это соображение может быть распространено на более общие случаи. Если первый член с отличным от нуля коэффициентом разложения в ряд Тейлора функции $f(x)$ равен x^n, то возмущение x^n будет иметь следующий вид:

$$F(x) = f + p = p_1 x + \ldots + p_{n-1}x^{n-1} + \\ + (1 + p_n)\, x^n + p_{n+1}x^{n+1} + \ldots . \qquad (4.13)$$

Если принять во внимание знаменатели дробей в формулах (4.9), (4.12), то становится очевидным, что всегда может быть найдена гладкая замена переменных (4.10), позволяющая исключить члены степени больше n. Эти же соображения распространяются и на случай n переменных состояния. Все члены более высокой степени, которые исключаются вследствие гладкой замены переменных, приводящей функцию $f(x_1, \ldots, x_n)$ к некоторой канонической форме, также могут быть удалены и в возмущенной функции $F = f + p$. Это преобразование является гладкой обратимой функцией коэффициентов ряда Тейлора, задаваемых разложением (4.4).

Вернемся к возмущенной функции $F(x) = \lambda x^2 + p(x)$, задаваемой формулой (4.5). В новой системе координат $0y$, связанной с системой координат $0x$ посредством преобразования (4.10), $F(y)$ записывается как

$$F[x(y)] = p(0;\, c) + p_1 y + \{(\lambda + p_2) + p_1 B_2\}\, y^2 = \\ = p(0;\, c) + p_1 y + \lambda' y^2. \qquad (4.14)$$

По координате y можно ввести новый масштаб $(y \to \tilde{y} = |\lambda'|^{1/2} y)$ и в результате получить квадратичный член $\tilde{y}^2$ с каноническим коэффициентом ± 1. Кроме того, можно перенести начало координат $(y' = y + p_1/2\lambda')$, что даст каноническую форму

$$F(y') = \left\{ p(0;\, c) - \frac{p_1^2}{4\lambda'} \right\} + \lambda' y'^2. \qquad (4.15)$$

Таким образом, возмущение функции одной переменной в морсовской критической точке не влияет на качественную природу этой функции, и хотя при этом критическая точка сдвигается на $\mathcal{O}(p_1)$, а значение функции в критической точке из-

меняется на $\mathcal{O}(p(0;\ c))$, тип критической точки остается без изменений.

Теперь проанализируем, как влияет возмущение на функцию с морсовской критической точкой в $0 \in \mathbb{R}^n$. Разложение в ряд Тейлора такой функции имеет вид (4.4). Проведем подобный анализ с (1) научной (или инженерной) и с (2) чисто математической точки зрения.

1. Если необходимо определить свойства функции (4.4) в окрестности начала координат, то следует учесть, что все x_i малы. Тогда могут быть выполнены следующие операции:

— усечение рядов до членов второй степени, так что

$$F \simeq p(0;\ c) + p_i x_i + (\lambda_i \delta_{ij} + p_{ij})\, x_i x_j; \tag{4.16а}$$

— приведение квадратичных членов к диагональному виду

$$F \simeq p(0,\ c) + p_i' x_i' + \sum_{i=1}^{n} \lambda_i' x_i'^2; \tag{4.16б}$$

— перенос начала координат $\left(x_i' \to x_i'' = x_i' + (p_i/2\lambda_i')\right)$, что приводит к

$$F \simeq \left\{ p(0;\ c) - \sum_{i=1}^{n} \frac{p_i'^2}{4\lambda_i'} \right\} + \sum_{i=1}^{n} \lambda_i' x_i''^2. \tag{4.16в}$$

2. В этом случае можно

— привести квадратичные члены в формуле (4.4) к диагональному виду посредством однородного линейного преобразования типа (3.3h1)

$$F' = p(0;\ c) + p_i' x_i' + \sum_{i=1}^{n} \lambda_i'' x_i'^2 + \\ + \text{Члены третьей степени} + \ldots; \tag{4.17а}$$

— исключить члены третьей степени и выше путем нелинейного преобразования типа (3.3n1), найденного с помощью процедур, описанных в гл. 3, принимая во внимание соображения непрерывности:

$$F'' = p(0;\ c) + p_i' x_i'' + \sum_{i=1}^{n} \lambda_i''' x_i''^2; \tag{4.17б}$$

— удалить линейные члены посредством переноса начала координат в соответствии с преобразованием $x_i'' \to x_i''' = x_i'' + p_i'/2\lambda_i'''$. В результате этих преобразований получаем следующую каноническую форму:

$$F''' = \left\{ p(0;\ c) - \sum_{i=1}^{n} \frac{p_i'^2}{4\lambda_i'''} \right\} + \sum_{i=1}^{n} \lambda_i''' x_i'''^2. \tag{4.17в}$$

Так как собственные значения λ_i''' непрерывно зависят от коэффициентов ряда Тейлора p_i, p_{ij}, ..., присутствующих в разложении (4.2), то возмущенная функция $F = f + p$ имеет морсовское i-седло вблизи $0 \in \mathbb{R}^n$, если исходная функция имеет i-седло в точке $0 \in \mathbb{R}^n$.

◇◇◇ Внимательный читатель не мог не заметить, что и интуитивный (4.16), и формальный (4.17) подходы предусматривают сходный процесс рассуждений, приводящий в конечном счете к одной и той же канонической форме. Естественно, возникает вопрос: нужны ли оба подхода, если вполне достаточно только интуитивного? Однако несмотря на все достоинства интуитивного подхода, он при помощи ряда преобразований, даваемых (4.17а), (4.17б) и (4.17в), в отличие от формального не позволяет прямо привести возмущение любых ростков катастроф, рассмотренных в гл. 3, к канонической форме. Кроме того, последовательность координатных преобразований (3.3), выполняемых при интуитивном подходе, такова, что за (3.3nl) следует (3.3hl) и далее следует (3.3ih), а при формальном подходе за (3.3hl) следует (3.3nl) и далее следует (3.3ih).

◇◇◇ На основе идентичности канонических форм (4.16в) и (4.17в) можно предположить, что последовательность выполняемых преобразований не столь уже существенна, если не принимать во внимание целей вычислений. Действительно, коэффициенты p_i, p_{ij}, ... в выражении (4.2) достаточно малы, и, следовательно, необходимые преобразования (3.3) будут близки к тождественным. Это по существу и означает, что последовательность выполняемых преобразований не играет существенной роли. Суперпозиция этих преобразований в любой последовательности вызывает в целом совершенно одинаковые преобразования коэффициентов возмущения p_i, p_{ij}, ... меньших степеней. Если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения канонической формы для $f(x_1, \ldots, x_n)$, то

$$\begin{aligned} &\lambda_i' = \lambda_i + p_{ii} + \mathcal{O}(p^2) \quad \text{в (4.16в)}, \\ &\lambda_i''' = \lambda_i + p_{ii} + \mathcal{O}(p^2) \quad \text{в (4.17в)}, \\ &\lambda_i''' - \lambda_i' = \mathcal{O}(p^2). \end{aligned} \tag{4.18}$$

Этот результат непосредственно следует из формулы (4.14).

Таким образом, в результате возмущения функции f в точке, где $\nabla f = 0$, а $\det f_{ij} \neq 0$ (рис. 4.2), слегка изменяются, во-первых, локализация точки равновесия [на $\mathcal{O}(p_i)$], во-вторых, значение возмущенной функции в новой точке равновесия, если $p(0; c) = 0$ [на $\mathcal{O}(p_i^2)$], и, в-третьих, собственные значения [на $\mathcal{O}(p_{ii})$], однако тип морсовской критической точки остается тем же.

Возмущение функции $f(x)$ в морсовской критической точке не влияет на качественный характер поведения этой функции в окрестности этой точки.

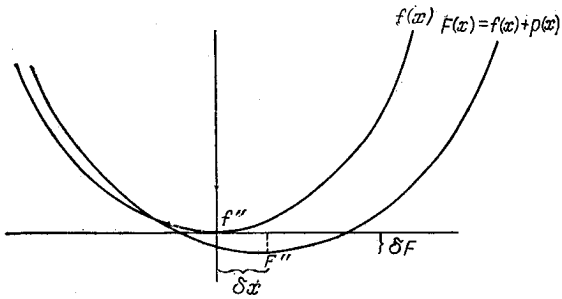


Рис. 4.2. Возмущение функции в морсовской критической точке.

Местонахождение критической точки, критическое и собственные значения функции слегка изменяются; локальных качественных изменений не наблюдается.

2.3. *Формы леммы расщепления*

Рассмотрим, что происходит, если возмущение действует в окрестности неморсовской критической точки. Из леммы расщепления (2.3а) следует, что невозмущенная функция имеет вид

$$f(x_1, \ldots, x_n) = f_{NM}(x_1, \ldots, x_l) + \sum_{i=l+1}^{n} \lambda_i x_i^2, \qquad (4.19)$$

где $\lambda_i \neq 0$ $(l+1 \leqslant i \leqslant n)$, а функция f_{NM} от l «плохих» переменных не имеет линейных и квадратичных членов в своем разложении в ряд Тейлора в окрестности вырожденной критической точки $x = 0 \in \mathbb{R}^n$. Самое общее возмущение функции f задается формулой (4.1), так что, полагая $p(0; c) = 0$, для возмущенной функции $F = f + p$ имеем следующее разложение в ряд Тейлора:

$$\begin{aligned} F = f + p = (2.3\text{а}) + (4.1) = \\ = p_i x_i + (\lambda_i \delta_{ij} + p_{ij}) x_i x_j + f_{NM}(x_1, \ldots, x_l) + \\ + \text{Члены третьей степени и выше} \end{aligned} \qquad (4.20)$$

(здесь $\lambda_1 = \lambda_2 = \ldots = \lambda_l = 0$). Теперь найдем каноническую форму для возмущенной функции (4.20).

Сначала с помощью однородного линейного преобразования приведем квадратичные члены функции (4.20) к диагональному

виду. В новой координатной системе имеем

$$F' = p'_i x'_i + \sum_{i=1}^{n} \lambda'_i x'^2_i + f'_{NM} + \text{Члены третьей степени и выше.} \quad (4.21)$$

Разности собственных значений $\lambda'_i - \lambda_i$ малы. Отсюда следует, что $\lambda'_1, \ldots, \lambda'_l$ также малы, а $\lambda'_{l+1}, \ldots, \lambda'_n$ нет. Однородное линейное преобразование типа (3.3hl) переводит члены третьей степени в члены третьей степени и, в более общем виде, члены n-й степени в члены n-й степени, что является следствием однородности и линейности преобразования.

Попытаемся путем осесохраняющего нелинейного преобразования типа (3.3hl) исключить члены более высокой степени. В силу соображений непрерывности нельзя исключить все члены выше линейных, поскольку все линейные коэффициенты малы. Аналогично с помощью непрерывного преобразования $x'_i \to x''_i$ $(i = 1, 2, \ldots, l)$ невозможно исключить члены степени >2, так как коэффициенты при x_i^2 малы. Однако коэффициенты при $x'^2_j (j = l+1, \ldots, n)$ уже не являются малыми, и поэтому для $n - l$ переменных можно найти осесохраняющее непрерывное нелинейное преобразование типа (3.3nl). Можно попытаться уточнить, какие именно члены степени > 2 могут быть исключены в результате таких преобразований. Как можно видеть, линейные члены вносят в рассматриваемую алгебру беспорядок. Если опустить детали вычислений, то получим результаты, неявно описанные в гл. 3:

— все члены вида $x_1^{r_1} \ldots x_n^{r_n}$, у которых $r_1 + r_2 + \ldots + r_n \geqslant 3$, за исключением тех, у которых $r'_{l+1} = \ldots = r_n = 0$, могут быть удалены из разложения посредством нелинейного преобразования «хороших» переменных. Это означает, что в разложении в ряд Тейлора (4.21) все члены третьей степени и выше, включая перекрестные произведения между l «плохими» переменными $x'_1, \ldots, x'_l$ и $n - l$ «хорошими» переменными $x'_{l+1}, \ldots, x'_n$, могут быть удалены посредством непрерывного нелинейного преобразования типа (3.2);

— все члены выше второй степени, включающие лишь «хорошие» переменные, также могут быть удалены путем такого преобразования.

Следовательно, в результате ряда таких преобразований возмущенная функция F (4.20) может быть представлена в виде суммы двух функций

$$F \doteq F_{NM} + F_M, \quad (4.22)$$

$$F_{NM} = f_{NM}(x''_1, \ldots, x''_l) + p(x''_1, \ldots, x''_l),$$

$$p = p''_i x''_i + p''_{ij} x''_i x''_j + \ldots, \quad 1 \leqslant i, j, \ldots \leqslant l, \quad (4.22\text{NM})$$

$$F_M = \sum_{i=l+1}^{n} p_i'' x_i'' + \sum_{i=l+1}^{n} \lambda_i'' x_i''^2. \qquad (4.22\text{M})$$

2.4. *Возмущение* $\pm x^n$

Важность вывода, следующего из принципиального результата (4.22), трудно переоценить:

— если в семействе функций $f(x;\ c)$ встречается функция $f(x;\ c^0)$, имеющая неморсовскую критическую точку в x_0, то, согласно (4.22), для любой другой функции $f(x;\ c)$ $(c \neq c^0)$, близкой к $f(x;\ c^0)$, можно найти такую координатную систему, что возмущенная функция $f(x;\ c) = f(x;\ c^0) + p(x)$ может быть получена отдельным возмущением морсовской $\sum_{i=l+1}^{n} \lambda_i x_i^2$ и неморсовской f_{NM} частей функции f. Заметим, что возмущение морсовской части в отличие от возмущения неморсовской части не вызывает качественных изменений в поведении этой функции, как возмущение f_{NM}.

Начнем исследование канонических форм неморсовских функций с изучения возмущений ростков катастроф от одной переменной: $\pm x^n$. Из формулы (4.22) и (4.1) следует, что наиболее общее возмущение $\pm x^n$ имеет вид

$$F(x) = p(0;\ c) + p_1 x + \ldots + p_{n-1} x^{n+1} + \\ + (\pm 1 + p_n) x^n + p_{n+1} x^{n+1} + \ldots . \qquad (4.23)$$

Из соображений непрерывности следует, что невозможно найти непрерывное преобразование, в результате которого исключаются все члены степени больше j при $j = 1, 2, \ldots, n-1$, так как соответствующие коэффициенты при x^j малы. Однако коэффициент при x^n уже не является малым и поэтому возможно удалить члены степени больше n. Это можно сделать с помощью нелинейного преобразования типа (3.3nl). Однородное линейное преобразование типа (3.3hl) может быть использовано для введения нового масштаба по переменной x и приведения коэффициента при x^n к каноническому виду.

После преобразований (3.3nl) и (3.3hl) будем иметь

$$F''(x'') = p_1'' x'' + \ldots + p_{n-1}'' x''^{n-1} \pm x''^n. \qquad (4.23'')$$

Если выполнить еще и неоднородное линейное преобразование $x'' \to x''' = x'' - s$ (3.3ih), то в результате получим

$$F'''(x''') = p_1''(s + x''') + p_2''(s^2 + 2sx''' + x'''^2) + \ldots \\ \ldots + p_{n-1}''(s^{n-1} + \ldots + x'''^{n-1}) \pm \\ \pm (s^n + \ldots + nsx'''^{n-1} + x'''^n). \qquad (4.23''')$$

Теперь приведем подобные члены и выберем величину (сдвиг) s так, чтобы исключить x какой-либо степени. Целесообразно исключить x в степени $(n-1)$, так как коэффициент при этом члене наиболее прост и определяется как $p''_{n-1} \pm ns$. Он будет равен нулю, если $s = \mp p''_{n-1}/n$. Каноническая форма возмущенной функции будет иметь вид

$$F'''(x''') = p'''_1 x''' + \ldots + p'''_{n-2} x'''^{n-2} \pm x'''^n. \qquad (4.24)$$

Удалив штрихи и приравняв преобразованные коэффициенты ряда Тейлора p'''_j (4.24) к каноническим возмущающим параметрам a_j, получим

$$\pm x^n + \text{Произвольное возмущение} \doteq \pm x^n + \sum_{\alpha=1}^{n-2} a_\alpha x^\alpha. \qquad (4.25)$$

Таким образом, в общем k-параметрическом семействе функций с одной «плохой» переменной каноническая форма для самой «плохой» неморсовской критической точки, с которой мы можем встретиться, имеет вид $\pm x^{k+2}$, а канонической формой для любой функции, близкой к этой функции, является $\pm x^{k+2} + \sum_{\alpha=1}^{k} a_\alpha x^\alpha$.

◇◇◇ Все k канонических коэффициентов a_α являются одновременно функциями всех коэффициентов ряда Тейлора p_i, p_{ij}, ... возмущенной функции и функциями физических управляющих параметров c_1, c_2, ..., c_k. Можно использовать как канонические математические коэффициенты, так и физические управляющие параметры c_β, так как они связаны между собой невырожденным преобразованием. В общем случае это так, и условие (2.6) удовлетворяется.

Пример. В 1-параметрическом семействе функций в общем случае можно встретить лишь одну «плохую» переменную в неморсовской критической точке (3.28). В подходящей системе координат «хорошие» и «плохие» переменные можно расщепить (3.27) Неморсовская функция «плохой» переменной x имеет каноническую форму x^3:

$$f_{NM}(x) = x^3. \qquad (4.26)$$

Возмущения исходной функции, такие, как изменения значений управляющих параметров, возмущают раздельно морсовскую и неморсовскую части исходной функции в специально подобранной системе координат (4.22). Согласно (4.26), возмущение неморсовской функции может быть усечено сверху до членов третьей степени с помощью соответствующего нелинейного преобразования. В действительности возмущение может быть усечено до членов второй степени, так как для возмущения кубического члена можно ввести новую шкалу перенормализации x. Следовательно,

$$f_{NM}(x) + \text{Возмущение} = p_0 + p_1 x + p_2 x^2 + x^3. \qquad (4.27)$$

Сдвиг начала координат ($x = x' + s$) дает

$$f_{NM}(x') + p(x') = a_0 + a_1 x' + a_2 x'^2 + x'^3,$$
$$a_0 = p_0 + p_1 s + p_2 s^2 + s^3, \qquad (4.28)$$
$$a_1 = p_1 + 2p_2 s + 3s^2,$$
$$a_2 = p_2 + 3s.$$

В этом выражении постоянный член не существен и можно считать его равным 0. Коэффициент a_2 всегда можно сделать равным нулю с помощью соответствующего выбора величины s, однако к коэффициенту a_1 это уже не относится, если только $p_2^2 - 3p_1 < 0$. Так как желательно иметь одинаковую каноническую форму для всех возмущений, то можно взять

$$x^3 \xrightarrow{\text{Каноническое возмущение}} x^3 + a_1 x. \qquad (4.29)$$

Проанализируем свойства этого 1-параметрического семейства функций $F(x; a_1)$ (рис. 4.3):

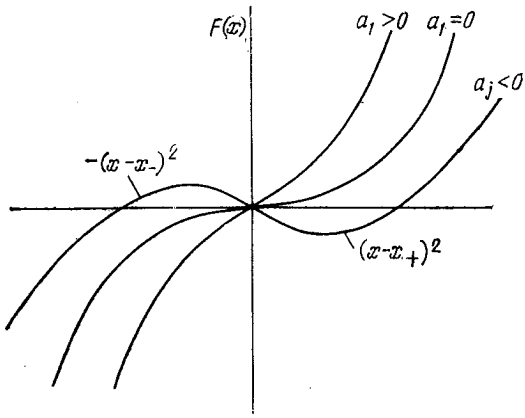


Рис. 4.3. 1-параметрическое семейство функций $F(x; a) = x^3 + a_1 x$, содержащее функции с двумя изолированными критическими точками и без критических точек. Функции этого семейства разделяются функцией $F(x; 0)$, имеющей дважды вырожденную критическую точку.

— при $a_1 = 0$ $F(x; 0)$ имеет вырожденную критическую точку в $x = 0$;
— при $a_1 < 0$ $F(x; a_1)$ имеет в точности две изолированные морсовские критические точки. При возрастании a_1 эти две критические точки стремятся друг к другу и становятся вырожденной точкой при $a_1 = 0$;
— при $a_1 > 0$ $F(x; a_1)$ не имеет критических точек.

В общем случае возмущение ростка катастрофы, имеющего вырожденную критическую точку, вызывает расщепление вырожденной точки на ряд невырожденных критических точек («морсификация»). (Максимальное число изолированных критических точек, получаемых при возмущении ростка катастрофы, указывается в нижнем индексе ростка.) Можно показать,

что для катастроф от одной переменной состояния A_k: x^{k+1}, так как

$$\frac{d}{dx}\left(x^{k+1}+\sum_{j=1}^{k-1} a_j x^j\right)=0 \tag{4.30}$$

является полиномиальным уравнением степени k, которое может иметь самое большее k вещественных корней.

Следовательно, в результате возмущения функции f в точке, где $\nabla f=0$ и $\det f_{ij}=0$ (рис. 4.3), изменяется:

— число точек равновесия. Вырожденная критическая точка кратности k расщепляется на не более чем k изолированных критических точек;

— локализация критических точек;

— значение функции в критических точках (критические значения), причем типы морсовских седел изолированных точек равновесия связаны между собой способом, характеристическим для рассматриваемой катастрофы.

Возмущение функции $f(x)$ в неморсовской критической точке вызывает качественное изменение в поведении $f(x)$ в окрестности критической точки. Число критических точек, их локализация и критические значения зависят от коэффициентов возмущения p_i, p_{ij}, ... значительно более сложным образом, чем соответствующие зависимости формы теоремы о неявной функции (разд. 2) или морсовских форм (разд. 3).

2.5. *Возмущение $x^2y \pm y^k$*

Рассмотрим возмущение ростков $D_{\pm(k+1)}$ катастроф (табл. 2.2). Наиболее общее возмущение этих ростков имеет вид (пренебрегая постоянным членом)

$$p(x, y)=\sum_{r+s\geqslant 1} p_{rs}x^r y^s. \tag{4.31}$$

Для простоты вначале изучим росток

$$D_4(x, y)=x^2y+y^3. \tag{4.32}$$

Наиболее общее возмущение $D_4(x, y)$ имеет вид

$$\begin{aligned} F=(4.32)+(4.31)=&\text{ Линейные члены }+\text{ Квадратичные члены }+\\ &+p_{30}x^3+(1+p_{21})x^2y+p_{12}xy^2+(1+p_{03})y^3+\\ &+\text{ Члены четвертой степени }+\ \ldots\ . \end{aligned} \tag{4.33}$$

В силу соображений непрерывности (разд. 3) можно ограничиться двумя членами ростка D_4, которые будем считать «центральными», и попытаться удалить окружающие их члены более

высокой степени с помощью однородного нелинейного преобразования типа (3.2). Это действительно возможно, поскольку лишь при x^2y и y^3 в формуле (4.33) стоят большие коэффициенты.

Как было показано (гл. 3, разд. 6), кубические члены ростка можно привести к канонической форме D_{+4} путем однородного линейного преобразования. Тогда

$$F' = (\text{Линейные члены})' + (\text{Квадратичные члены})' + \\ + (x^2y + y^3)' + (\text{Члены четвертой степени})' + \ldots . \quad (4.34)$$

Кроме того, для удаления всех членов четвертой степени и выше можно использовать нелинейное преобразование типа (3.3nl). В данном случае вычисления оказываются более сложными, так как приходится учитывать вклад линейных и квадратичных членов. Однако можно удалить все члены четвертой степени и выше, оставляя при этом кубические члены в канонической форме (4.32), если применить однородное нелинейное преобразование типа (3.2) к формуле (4.33) либо к формуле (4.34). Приведем лишь конечные результаты. Опуская штрихи, получим

$$F \doteq (\text{Линейные члены}) + (\text{Квадратичные члены}) + \\ + (x^2y + y^3). \quad (4.35)$$

Наконец, при помощи неоднородного преобразования типа (3.3ih) $x' = x + s_1$, $y' = y + s_2$ можно перенести начало координат. Тогда в случае квадратичных членов вида

$$p_{20}x^2 + p_{11}xy + p_{02}y^2 \quad (4.36)$$

путем соответствующего выбора коэффициентов $(p_{11}/2, p_{20})$ или $(p_{11}/2, p_{02}/3)$ можно добиться равенства нулю соответственно коэффициентов при x^2 и xy либо при xy и y^2. Для наших целей удобно удалить члены x^2 и xy. Поэтому общее возмущение $D_4(x, y)$ и наиболее общее возмущение ростка катастроф $D_{k+1}(x, y) = x^2y + y^k$ будут иметь следующий вид:

$$D_4(x, y) + \text{Произвольное возмущение} \doteq \\ \doteq x^2y + y^3 + a_1x + a_2y + a_3y^2, \quad (4.37)$$

$$F \doteq \text{Линейные члены} + \text{Квадратичные члены} + \\ + (1 + p_{21})x^2y + \ldots + (1 + p_{0k})y^k + \ldots . \quad (4.38)$$

Все члены, которые не присутствуют явно, являются малыми величинами. При помощи линейного преобразования кубические члены могут быть приведены к каноническому виду

$$x^2y + by^3, \quad (4.39)$$

где коэффициент b мал при малых возмущениях, так как при отсутствии возмущения $b = 0$ $(k > 3)$. В силу соображений непрерывности (разд. 3) должно существовать непрерывное преобразование, позволяющее исключить только те члены, которые могут быть получены из одночленов (4.38), имеющих большие коэффициенты, а именно x^2y и y^k. Следуя стандартным алгоритмам и проводя вычисления, подобные вычислениям, выполненным в разд. 6, найдем, что

$$F = \text{Линейные члены} + \text{Квадратичные члены} + \\ + x^2y + p_{03}y^3 + \ldots + p_{0,k-1}y^{k-1} + y^k. \qquad (4.40)$$

Если полученные квадратичные члены такие же, как в выражении (4.36), то перенос начала координат на $(s_1, s_2) = (p_{11}/2; p_{20})$ может быть использован для того, чтобы избавиться от членов x^2 и xy. При этом получим следующую каноническую форму для функции, близкой $D_{k+1}(x, y)$:

$$F \doteq (x^2y + y^k) + a_1x + \sum_{j=2}^{k} a_jy^{j-1}. \qquad (4.41)$$

Канонический вид возмущений ростков $D_{-(k+1)}$ в точности таков же, как и в случае возмущений $D_{+(k+1)}$.

2.6. *Возмущение* $x^3 \pm y^4$

Анализ влияния возмущения данного вида на функцию предусматривает выполнение операций, которые уже, наверно, набили оскомину у читателя. Поэтому постараемся сократить ход наших рассуждений, срезая где можно углы, и тем самым избежать монотонности и повторяемости изложения, навевающих на читателя скуку. При этом, однако, следует иметь в виду, что подобные «углосрезающие» рассуждения основываются на строгих математических положениях, и, следовательно, эти «интуитивные» представления в действительности являются полностью обоснованными. Схематически общее возмущение ростка функции двух переменных состояния можно представлять так, как показано на рис. 4.4. Необходимо только каждому одночлену (например, с xy^2) поставить в соответствие определенный коэффициент ряда Тейлора (p_{12}). Так как все коэффициенты ряда Тейлора малы, то, естественно, возникает вопрос, какие из этих коэффициентов могут быть обращены в нуль посредством преобразования типа (3.2), лежащего вблизи тождественного преобразования

$$x \to x' = x + \Delta_1, \quad y \to y' = y + \Delta_2. \qquad (4.42)$$

Здесь Δ_i получены с помощью замены конечных коэффициентов A, присутствующих в (3.2) в виде инфинитезимальных коэффициентов δA. В случае ростка x^3+y^4 это преобразование дает

$$(x')^3+(y')^4 \quad = x^3+y^4+(4.1),$$

$$\downarrow$$

$$x^3+3x^3\Delta_1+3x\Delta_1^2+\Delta_1^3+y^4+4y^3\Delta_2+6y^2\Delta_2^2+4y\Delta_2^3+\Delta_2^4. \qquad (4.43)$$

Так как Δ_1 и Δ_2 являются «инфинитезимальными» величинами, то представление (4.43) можно «линеаризовать». Однако путем замены переменных (4.42) из этого представления могут быть

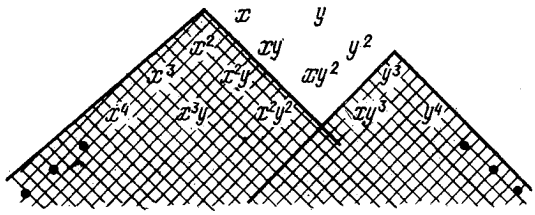


Рис. 4.4. Представление общего возмущения функции двух переменных с помощью вектора линейного векторного пространства, базисные векторы x^py^q которого могут быть представлены в виде треугольника Паскаля.

удалены только члены вида $3x^2\Delta_1$ и $4y^3\Delta_2$. (Эти члены показаны в заштрихованных областях рис. 4.4; по существу все члены, находящиеся в этих заштрихованных областях, могут быть удалены.) Единственные члены, которые не могут быть удалены из возмущения $p(x, y)$ в (4.43), следующие: x, y, xy, y^2, xy^2. Следовательно, для канонического вида любого возмущения ростка x^3+y^4 имеем

$$(x^3+y^4)+(a_1x+a_2y+a_3xy+a_4y^2+a_5xy^2). \qquad (4.44)$$

Канонические возмущения ростка $E_{-6}=x^3-y^4$ в точности таковы, как и возмущения ростка $E_{+6}=x^3+y^4$.

3. ВЫВОДЫ

Возмущения не влияют на качественный характер поведения функции в окрестности некритической или морсовской критической точек. В первом случае незначительно изменяется величина и направление градиента функции; в последнем слегка смещается критическая точка и изменяется критическое значение функции, но тип морсовского седла в этой точке остается без изменения. Именно отсутствие качественных изменений, вносимых возмущениями в окрестностях таких точек, и является

причиной того, что теорема о неявной функции и лемма Морса не дополняются списком возмущений.

В тех случаях, когда семейство функций содержит члены с неморсовской критической точкой (гл. 3, разд. 4), возможно найти такую координатную систему, которая расщепляет эту функцию на «плохую» неморсовскую и «хорошую» морсовскую части. Этот результат верен и для семейств функций, близких к рассматриваемой неморсовской функции. Другими словами, всегда можно найти некоторую координатную систему, в которой возмущенная функция расщепляется на две части, причем каждая из них может быть изучена отдельно. Возмущенная морсовская функция не имеет качественно новых свойств, в то время как возмущенная неморсовская функция может их иметь.

Изучение возмущенных ростков катастроф проводилось методами, изложенными в гл. 3. Была введена наиболее общая форма возмущения (4.1) и показано, какая часть «хвоста» ряда Тейлора может быть удалена путем гладкой замены переменных. Самые первые члены разложения в ряд Тейлора возмущенной функции не могут быть исключены из-за соображений непрерывности. Для сохранения непрерывности эти члены необходимо рассматривать лишь с конечными коэффициентами. Было установлено, и это вовсе неудивительно, что те члены, которые исключались при нахождении канонического ростка (гл. 3), также могут быть исключены посредством непрерывной замены переменных. Некоторые из первых коэффициентов могут быть исключены при помощи неоднородного преобразования (при изучении семейств функций местонахождение начала координат не является столь важным, поэтому возможно рассмотрение и неоднородных преобразований).

Возмущение функции в неморсовской критической точке оказывает значительное влияние на свойства функции в окрестности этой точки. В общем случае k-кратная вырожденная критическая точка расщепляется самое большее на k-изолированных критических точек. К счастью, наиболее общая форма, которую может иметь возмущение, можно свести гладкой заменой переменных к каноническому виду.

Существенное качественное влияние возмущений на ростки катастроф явилось причиной того, что список элементарных катастроф Тома содержит перечень простых ростков катастроф и дополнительный перечень канонических возмущений. Полный список канонических возмущений для всех простых ростков катастроф (3.53) содержится в табл. 4.1.

Методы, изложенные в этой главе, могут быть с успехом использованы и при определении канонических форм непростых ростков, рассмотренных в гл. 3.

Таблица 4.1. **Функции элементарных катастроф (нуль-модальные)**

Тип катастрофы	Росток катастрофы [1]	Возмущение
$A_{\pm k}$ [2]	$\pm x^{k+1}$	$\sum_{j=1}^{k-1} a_j x^j$
$D_{\pm k}$	$x^2y \pm y^{k-1}$ k четное $\pm(x^2y + y^{k-1})$ k нечетное	$\sum_{j=1}^{k-3} a_j y^j + \sum_{j=k-2}^{k-1} a_j x^{j-(k-3)}$ [3]
$E_{\pm 6}$	$\pm(x^3 + y^4)$	$\sum_{j=1}^{2} a_j y^j + \sum_{j=3}^{5} a_j x y^{j-3}$
E_7	$x^3 + xy^3$	$\sum_{j=1}^{4} a_j y^j + \sum_{j=5}^{6} a_j x y^{j-5}$ [3]
E_8	$x^3 + y^5$	$\sum_{j=1}^{3} a_j y^j + \sum_{j=4}^{7} a_j x y^{j-4}$

[1] Для удобства используем $y_1 = x$, $y_2 = y$.
[2] $A_{+k} = A_{-k}$, если k четное.
[3] Представление данного возмущения посредством одночленов не единственно.

5

«ПРИНЦИП ЛОМА»

Лом — это довольно мощный инструмент, используемый для разбивания льда, камней и т. д. Обычно его вбивают в уже имеющуюся в глыбе льда трещину («меры нуль») и «раскачивают» до тех пор, пока смежные куски не расколятся. В теории катастроф роль такого инструмента могут выполнять некоторые теоремы, сформулированные при рассмотрении отдельных изолированных функций и семейств функций (гл. 2). Для отдельных изолированных функций большинство точек $x \in \mathbb{R}^n$ являются некритическими, и тем не менее качественно глобальное поведение рассматриваемой функции полностью определяется изолированными критическими точками (разд. 1). Для семейства функций большинство точек $x \in \mathbb{R}^k$ параметризует морсовские функции, и все-таки глобальное качественное поведение семейства функций полностью определяется множеством меры нуль в пространстве $\mathbb{R}^k$, точки которого параметризуют функции с вырожденными точками. Подобное множество меры нуль, или сепаратриса, называется *множеством бифуркации* (или бифуркационным множеством) и обозначается $\mathscr{B}$.

Множество точек $c \in \mathbb{R}^k$, параметризующих неморсовские функции, разбивает пространство управляющих параметров $\mathbb{R}^k$ на открытые области, каждая из которых параметризует лишь качественно подобные функции. Определим эти сепаратрисы и качественный тип функций, ассоциируемых с каждой открытой областью, для всех элементарных катастроф, перечисленных в табл. 2.2, пространство управляющих параметров которых может быть представлено наглядно (т. е. $k \leqslant 3$).

Сепаратриса, параметризующая функции с вырожденными критическими значениями в изолированных критических точках, называется *множеством Максвелла* и обозначается $\mathscr{D}_M$; множество Максвелла определяется с помощью уравнений Клаузиуса — Клапейрона.

1. ПОТЕНЦИАЛЬНАЯ ФУНКЦИЯ

Предположим, что мы случайно выбираем функции n переменных (состояния) или более частные потенциальные функции $V(x_1, x_2, \ldots, x_n)$ из всего множества таких функций. Тогда большинство функций, которые мы встретим, вообще не имеют критических точек или имеют изолированные (невырожденные) критические точки. Функции, имеющие только изолированные критические точки, назовем *морсовскими*. В действительности почти все [1]) функции, которые мы при этом встретим, будут морсовскими функциями, и поэтому при рассмотрении типичных потенциальных функций нет необходимости в использовании теоремы Тома.

[1]) Понятие «почти все» будет строго обосновано в гл. 21.

Рассмотрим типичную потенциальную функцию $V(x_1, \ldots, x_n)$. Если точка x^0 выбрана случайно или является типичной точкой в $\mathbb{R}^n$, то $\nabla V \neq 0$ в x^0, так как типичная потенциальная функция имеет лишь изолированные критические точки. Поскольку $\nabla V \neq 0$ и почти везде в $\mathbb{R}^n$ справедлива теорема о неявной функции, то практически снимается вопрос о применимости леммы Морса, так как точки, в которых $\nabla V = 0$, встречаются редко. Однако критические точки имеют большую ценность, чем некритические точки, и именно они в основном характеризуют глобальные качественные изменения в поведении потенциальной

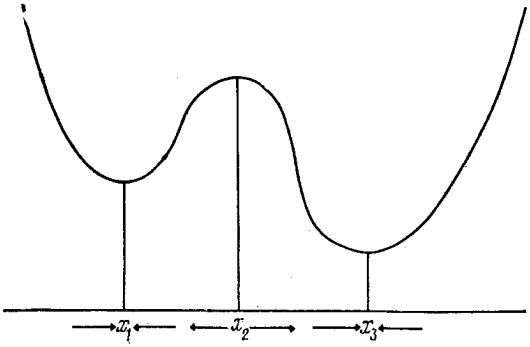
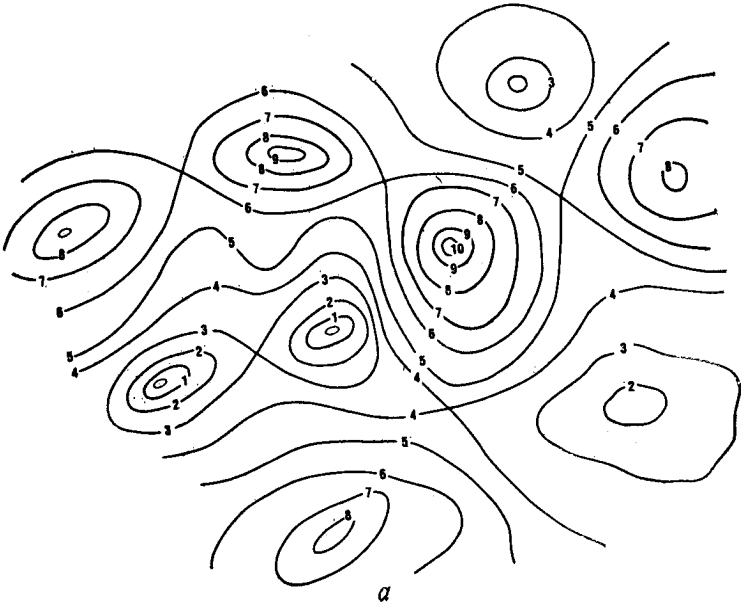


Рис. 5.1. Положение изолированных критических точек на прямой и их типы полностью определяют качественную природу потенциальной функции.

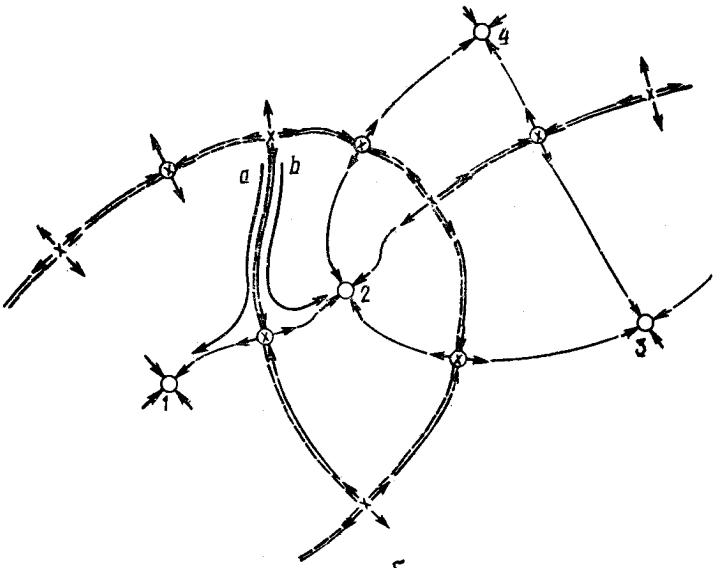
Единственное не 0-седло в точке x_2 делит $\mathbb{R}^1$ на левосторонний и правосторонний бассейны с аттракторами x_1 и x_2 соответственно.

функции $V(x)$. Это легко показать для одномерного пространства (рис. 5.1). Если $V(x)$ имеет критические точки x_1, x_2, x_3 (как показано на этом рисунке), то в любой точке слева от x_1 сила ($F = -dV/dx$) направлена вправо, а в любой точке между x_1, x_2 — влево. Направление силы F меняется при переходе через невырожденную критическую точку. Следовательно, если $V(x)$ имеет лишь изолированные критические точки и координаты всех этих точек известны, то можно определить все качественные изменения в поведении функции $V(x)$ при условии, что известны направление F в любой промежуточной открытой области между критическими точками или морсовский тип каждой критической точки.

Аналогичные рассуждения имеют место и для пространств размерности два и более. Например, если $V(x_1, x_2)$ имеет два изолированных минимума (0-седла), то должен быть отрезок прямой (или некоторый путь), соединяющий эти два 0-седла и проходящий через 1-седло. Определяя координаты и тип седла каждой изолированной точки в $\mathbb{R}^2$, можно сделать выводы о качественном изменении в поведении $V(x_1, x_2)$.



а



б

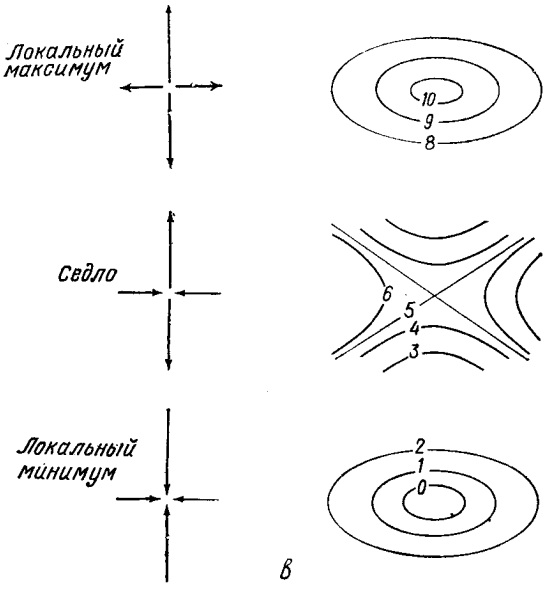


Рис. 5.2.

а — представление потенциальной функции двух переменных посредством линий уровня. *б* — представление той же функции в теории динамических систем; × вершина горы; ⊗ перевал; ○ дно озера; ——— сепаратриса; — — — ребро, бассейн притяжения; *в* — в окрестности изолированной критической точки легко устанавливается связь между представлениями (*а*) и (*в*).

Геометрически качественная природа функции $V(x_1, x_2)$ может быть представлена двояко (рис. 5.2): в случае рис. 5.2, *а* для представления используются контуры или линии уровня $V(x_1, x_2) = \text{const}$ вместе с соответствующими значениями констант на них. В окрестности локального максимума или минимума линии уровня имеют эллиптическую форму, а в окрестности седла — гиперболическую. Вершины гор и седла связаны посредством хребтов, озерных впадин, седлообразных долин. Подобное топографическое представление чрезвычайно удобно для путешественников и альпинистов, так как оно содержит информацию не только о местонахождении горных вершин, озер и перевалов.

В случае рис. 5.2, *б* для представления, которое оказывается очень удобным для наших целей, для обозначения каждой критической точки и ее типа используются соответственно кружки или стрелки. Направления стрелок определяются матрицей устойчивости V_{ij} в рассматриваемой точке. Диагонализация V_{ij} в рассматриваемой критической точке позволяет найти направления главных осей [направления максимума/минимума,

градиента/антиградиента]. Собственные значения, связанные с главными осями, указывают на величину и направления сил вдоль главных осей. Согласно стандартному соглашению, стрелка, направленная к критической точке, указывает силу в данном направлении, и чем длиннее стрелка, тем больше сила. Очевидно, что главные оси (рис. 5.2, *в*) могут быть связаны таким образом, чтобы получились долины и хребты, о наличии которых свидетельствует контурная карта, представленная на рис. 5.2, *а*.

Если для представления функции $V(x_1, x_2)$ используются координаты главных осей и собственные значения критических точек (как это сделано на рис. 5.2, *в*), то соответствующее контурное представление можно считать беглым наброском, по крайней мере в окрестности любой критической точки.

Если контурную карту (рис. 5.2, *а*) наполнить водой, то последняя соберется в озера, расположенные на дне долин. Минимум, который притягивает воду, называют *аттрактором*. Каждый аттрактор лежит во впадине водораздела, или бассейна, называемого *бассейном притяжения* (*аттракции*). Внутри каждого бассейна притяжения существует один аттрактор (локальный минимум). Аттракторы разделяются седлами, хребтами и максимумами, которые образуют границу между различными бассейнами притяжения.

Представление $V(x)$, приведенное на рис. 5.2, *б*, более просто обобщить на случай n (>2) переменных состояния, чем контурное представление (рис. 5.2, *а*), так как для каждого морсовского 0-седла имеется всего лишь один бассейн притяжения. Очертание бассейна притяжения может быть определено при помощи связывания i-седла со смежными $(i+1)$-седлами $(i>0)$. Этот процесс для случая $\mathbb{R}^2$ проиллюстрирован на рис. 5.2, *б*.

2. СЕМЕЙСТВО ПОТЕНЦИАЛЬНЫХ ФУНКЦИЙ

Рассмотрим семейство типичных потенциальных функций $V(x_1, \ldots, x_n; c_1, \ldots, c_k)$. Если точка c^0 в пространстве управляющих параметров $\mathbb{R}^k$ выбрана случайно или является его типичной точкой, то потенциальная функция, которую параметризует эта точка, будет обязательно морсовской функцией, имеющей только изолированные критические точки. Так как $V(x; c^0)$ — морсовская функция, то для нее в любой точке $x^0 \in \mathbb{R}^n$ почти для всех точек $c^0 \in \mathbb{R}^k$ справедлива либо теорема о неявной функции, либо лемма Морса. Тогда возникает вопрос: как быть с теоремой Тома, которая верна для совсем маловероятных точек (меры нуль, но в общем случае не изолированных) $c^0 \in \mathbb{R}^k$, в которых потенциальная функция V не

является морсовской? Оказывается, что в основном именно неморсовские функции $V(x;\ c^0)$ ответственны за качественную природу семейства потенциальных функций $V(x;\ c)$.

Чтобы показать, что это так, выберем семейство типичных потенциальных функций $V(x;\ c)$, и пусть $c^0 \in \mathbb{R}^k$ — точка, в которой $V(x;\ c^0)$ является морсовской функцией. Тогда потенциальная функция $V(x;\ c^0)$ имеет только изолированные критические точки. Исследуем, как изменяется положение этих критических точек при изменении управляющих параметров $c^0 \in \mathbb{R}^k$. Поскольку изолированные критические точки задают аттракторы и бассейны притяжения потенциала V, то по существу достаточно изучить влияние значения c^0 на «поверхностную топографию» потенциальной функции $V(x;\ c^0)$. Предположим, что при $c = c^0$ существует некоторая критическая точка $x = x^0$. Если c^0 изменяется на $c^0 + \delta c^0$, то положение критической точки также изменится и станет $x^0 + \delta x^0$. Определим вариации δx через значения δc^0_α. Для этого используем разложение функции в ряд Тейлора в окрестности точки $(x^0;\ c^0)$ из $\mathbb{R}^n \oplus \mathbb{R}^k$:

$$V(x;\ c) = V(x^0;\ c^0) + (x - x^0)_i V_i + (c - c^0)_\alpha V_\alpha,$$

$$\frac{1}{2!}(x - x^0)_i (x - x^0)_j V_{ij} + (x - x^0)_i (c - c^0)_\alpha V_{i\alpha} +$$

$$+ \frac{1}{2!}(c - c^0)_\alpha (c - c^0)_\beta V_{\alpha\beta} + \mathcal{O}(3). \qquad (5.1)$$

Все производные в разложении берутся в точке $(x^0;\ c^0)$. Так как x^0 — критическая точка, $V_i = 0$.

Теперь положим $c = c^0 + \delta c^0$ и найдем точку равновесия $V(x; c^0 + \delta c^0)$, расположенную вблизи точки x^0. Для этого приравняем первые производные $\partial V/\partial x_i$ нулю:

$$\frac{\partial V}{\partial x_i} = V_{ij}\,\delta x^0_j + V_{i\alpha}\,\delta c^0_\alpha + \mathcal{O}(2) = 0.$$

Опуская члены второй степени, выразим, если, конечно, матрица V_{ij} не вырождена в точке $(x^0;\ c^0)$, вариации δx^0_j через δc^0_α следующим образом:

$$\delta x^0_j = -(V^{-1})_{ji} V_{i\alpha}\,\delta c^0_\alpha$$

или

$$\frac{\partial x^0_j}{\partial c^0_\alpha} = -(V^{-1})_{ji} V_{i\alpha}. \qquad (5.2)$$

Таким образом, при изменении управляющих параметров меняется как положение критических точек, так и значения функций в этих точках. Изменение критических значений

соответствует изменению δc управляющих параметров и определяется путем подстановки (5.2) в разложение (5.1), т. е.

$$V(x^0+\delta x;\ c^0+\delta c)=V(x^0;\ c^0)+\delta^{(1)}V+\delta^{(2)}V+\ \ldots, \qquad (5.3)$$

$$\delta^{(1)}V=V_\alpha\,\delta c^\alpha, \qquad (5.3;\ 1)$$

$$\delta^{(2)}V=\frac{1}{2}\,\delta c^\alpha\left[V_{\alpha\beta}-V_{\alpha i}\left(V^{-1}\right)^{ij}V_{j\beta}\right]\delta c^\beta=\frac{1}{2}\,g_{\alpha\beta}\,\delta c^\alpha\,\delta c^\beta. \qquad (5.3;\ 2)$$

Уравнения (5.3) отражают «линейный отклик» положения изолированных критических точек на малые изменения управляющих параметров, а уравнения (5.3; 1) — «линейный отклик» критических значений на малые изменения управляющих параметров. «Квадратичный отклик» удобно представлять в терминах «метрического тензора» $g_{\alpha\beta}$, определяемого формулой (5.3; 2).

Анализ соотношений (5.2) позволяет сделать следующий важный вывод: малые изменения управляющих параметров приводят лишь к незначительным изменениям положения изолированных критических точек. Кроме того, поскольку собственные значения и собственные векторы V_{ij} в изолированных критических точках являются гладкими функциями управляющего параметра c, малые изменения управляющих параметров могут вызывать лишь малые изменения собственных значений и собственных векторов. Из рассмотрения, проведенного в разд. 5.1, следует, что аттракторы и бассейны притяжения $V(x;\ c^0+\delta c^0)$ незначительно отличаются от аттракторов и бассейнов притяжения $V(x;\ c^0)$. Поэтому эти две функции качественно (топографически) подобны.

Однако если $V(x;\ c^0)$ не является морсовской функцией, то рассматриваемая потенциальная функция должна иметь по крайней мере одну вырожденную критическую точку x^0. Это в свою очередь означает, что соотношения (5.2) в данной критической точке не выполняются, так как матрица V_{ij} в этой точке не обратима. При $c\to c^0$ две (или более) изолированные точки стремятся слиться в одну точку $x=x^0$. Как только это произойдет, их главные оси становятся взаимно параллельными, а собственные значения, связанные с одним (или более) главным направлением, стремятся к нулю.

Следовательно, наличие вырожденных критических точек соответствует качественному изменению в топографии семейства потенциальных функций, что вполне естественно, поскольку топография определяется распределением и типом изолированных критических точек (когда одна или более критических точек становятся вырожденными, топография должна измениться).

Если функцию, поведение которой качественно (число и тип критических точек, бассейны притяжения и т. д.) не изменяется при достаточно малых возмущениях, считать структурно устойчивой, то из соотношений (5.2) следует, что морсовские функ-

ции семейства потенциальных функций — структурно устойчивы. Множество точек $c \in \mathbb{R}^k$, в которых $V(x; c)$ не является морсовской функцией, образует сепаратрису, разделяющую $\mathbb{R}^k$ на открытые области, параметризующие структурно устойчивые функции качественно различных типов. Найдем сепаратрисы, параметризующие структурно неустойчивые функции для всех катастроф, перечисленных в табл. 2.2, размерность пространства управляющих параметров которых не превышает трех ($k \leqslant 3$).

3. ЭЛЕМЕНТАРНЫЕ КАТАСТРОФЫ

3.1. Катастрофы типа A_2

Предположим, что $V(x_1, \ldots, x_n; c)$ — общее 1-параметрическое семейство потенциальных функций. Тогда при исследовании этого семейства можно встретить отдельные функции, которые имеют неморсовские критические точки. Для таких функций разложение (2.3б) имеет место в некоторой окрестности критической точки (гл. 3, разд. 4). В действительности для любой функции, расположенной вблизи неморсовской функции, имеет место разложение (2.4) (гл. 4, разд. 4). Поскольку на морсовскую часть этого разложения возмущение качественно не влияет, можно ограничиться лишь изучением зависимости качественных изменений в поведении функции катастрофы от управляющих параметров [1]).

Катастрофа A_2 задается формулой

$$A_2: \quad F(x; a) = \frac{1}{3}x^3 + ax, \qquad (5.4; 0)$$

при этом коэффициенты в простых ростках катастроф могут быть выбраны равными каноническим значениям, подобным ± 1. В тех случаях, когда берутся производные, могут быть выбраны другие канонические значения с помощью изменения масштабов. Для удобства такие же множители могут быть введены и в возмущение.

Критические и дважды вырожденные критические точки $F(x; a)$ определяются соответственно из условий равенства нулю градиента $F(x; a)$ и $d^2F/dx^2 = 0$:

$$x^2 + a = 0, \qquad (5.4; 1)$$

$$2x = 0. \qquad (5.4; 2)$$

[1]) Вышеизложенные соображения справедливы при рассмотрении катастрофы любого типа, поэтому мы не будем повторять их при анализе других катастроф.

Дважды вырожденная критическая точка расположена в точке пространства $\mathbb{R}^2$ с координатами $x = 0$ из (5.4; 2) и $a = 0$ из (5.4; 1). Таким образом, изолированная точка $a = 0$ является сепаратрисой пространства управляющих параметров, разделяющей функции двух качественно различных видов. Для определения этих видов требуется выбрать некоторое $a > 0$ и определить качественные изменения в поведении функции $F(x; a)$.

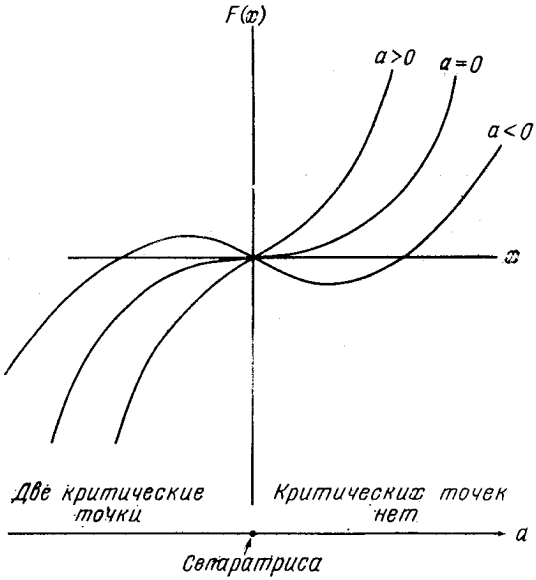


Рис. 5.3. Все функции $F(x; a) = \frac{1}{3}x^3 + ax$ с $a > 0$ являются функциями качественно одного и того же типа, не имеющими критических точек. Все функции с $a < 0$ также качественно подобны. Два качественно различных типа кривых имеют общую границу, сепаратрису $a = 0$.

Тогда из (5.4) следует, что все функции $F(x; a)$ с $a > 0$ качественно идентичны (не имеют критических точек).

Аналогичное рассмотрение может быть выполнено и в случае функции $F(x; a)$ с $a < 0$ (рис. 5.3).

3.2. Катастрофы типа A_3

Критические, дважды вырожденные критические и трижды вырожденные критические точки катастрофы A_3 определяются приравниванием соответственно первой, второй и третьей произ-

водных $F(x; a, b)$ нулю:

$$A_3\colon\ F(x; a, b) = \frac{1}{4}x^4 + \frac{1}{2}ax^2 + bx, \qquad (5.5; 0)$$

$$x^3 + ax + b = 0, \qquad (5.5; 1)$$

$$3x^2 + a = 0, \qquad (5.5; 2)$$

$$6x = 0. \qquad (5.5; 3)$$

Условие (5.5; 1) выполняется в критических точках; условия (5.5; 1) и (5.5; 2) — в дважды вырожденных критических точках, а условия (5.5; 1) — (5.5; 3) — в трижды вырожденных критических точках. Положение в пространстве параметров точки, которая описывает функцию с трижды вырожденной критической точкой, определяется как

$$(5.5\ 3;) \Rightarrow x = 0 \overset{(5.5;\ 2)}{\Rightarrow} a = 0 \overset{(5.5;\ 1)}{\Rightarrow} b = 0. \qquad (5.6)$$

Соответствующая функция $F(x; 0, 0) = x^4/4$ имеет трижды вырожденную критическую точку в начале координат.

Точки пространства управляющих параметров, которые параметризуют функции с дважды вырожденными критическими точками, определяются из уравнений (5.5; 2) и (5.5; 1):

$$(5.5;\ 2) \Rightarrow a = -3x^2 \overset{(5.5;\ 1)}{\Rightarrow} b = 2x^3. \qquad (5.7)$$

Если положение дважды вырожденной критической точки обозначить через x_c, то формула (5.7) дает значения управляющих параметров a и b, которые описывают функцию с дважды вырожденной критической точкой x_c.

Уравнения (5.7) определяют параметрическое представление связи между a и b. Более прямое выражение для связи между a и b может быть получено, если исключить x_c из (5.7):

$$\left(-\frac{a}{3}\right)^{1/2} = x = \left(\frac{b}{2}\right)^{1/3},$$

$$\left(-\frac{a}{3}\right)^3 = \left(\frac{b}{2}\right)^2,$$

$$\left(\frac{a}{3}\right)^3 + \left(\frac{b}{2}\right)^2 = 0. \qquad (5.8)$$

Кривые, задаваемые уравнением (5.8), изображены на рис. 5.4.

Сепаратриса пространства управляющих параметров состоит из точки $(a, b) = (0, 0)$, координаты которой определяются с помощью формулы (5.6), и «линии складки», описываемой уравнением (5.8). Чтобы определить качественное поведение функций $F(x; a, b)$, параметризуемых точками области I, достаточно выбрать любую точку из этой области и изучить качественные

изменения в поведении соответствующей функции. Возьмем для простоты точку $(a, b) = (+1, 0)$. Тогда уравнение (5.5; 1) имеет лишь одно действительное решение $x = 0$ и это решение определяет минимум. Поэтому в силу (5.3) все функции, параметризуемые управляющими параметрами из области I, имеют единственный минимум. Аналогично определяется качественное поведение функций, параметризуемых точками из области III.

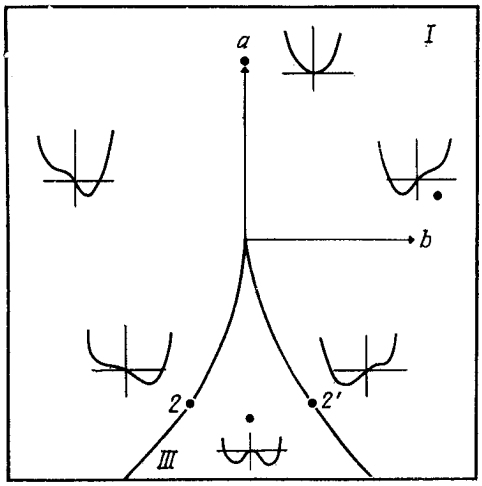


Рис. 5.4. Сепаратриса катастрофы сборки, определяемая уравнениями $dF/dx = 0$, $dF^2/dx^2 = 0$, разделяет пространство управляющих параметров на две открытые области, представляющие функции с одной критической точкой или функции с тремя критическими точками.

Точки на линиях складок (5.8) представляют неморсовские функции, а точка в начале координат $\mathbb{R}^2$ – функцию $x^4/4$, имеющую трижды вырожденную критическую точку в начале координат $\mathbb{R}^1$. Сепаратриса имеет компоненты размерности 1 (линии складки, при $ab \neq 0$) и размерности 0. На рисунке изображены также потенциальные функции, соответствующие некоторым точкам плоскости $\mathbb{R}^2$ управляющих параметров.

Удобной для нас точкой в этой области является $(a, b) = (-1, 0)$: в этой точке уравнение (5.5; 1) имеет три решения — локальный максимум в $x = 0$ и два локальных минимума в $x = +1$, $x = -1$. Следовательно, все функции, параметризуемые точками области III, должны иметь два локальных минимума и один локальный максимум.

Что касается функций, параметризуемых точками линии складки, то в точке 2 (рис. 5.4) управляющие параметры a и b имеют отрицательные значения. Из (5.7) следует, что соответствующая дважды вырожденная критическая точка x_c должна также иметь отрицательное значение. Так как сумма трех корней уравнения (5.5; 1) равна нулю (в силу того что коэффициент

при квадратичном члене нуль), а два вырожденных корня отрицательны, то третий корень должен быть положительным. Следовательно, вдоль левой кривой сборки локальный максимум и левосторонний минимум должны быть вырожденными. Аналогично в точке 2 локальный максимум и правосторонний минимум становятся вырожденными (рис. 5.4). На пересечении двух линий складки в точке $(a, b) = (0, 0)$ локальный максимум и два минимума соответствующей функции вырождены.

3.3. Катастрофы типа A_4

Критические точки вырожденности j (изолированные критические точки имеют $j = 1$) могут быть получены приравниванием первых производных $F(x; a, b, c)$ нулю:

$$A_4\colon\ F(x; a, b, c) = \tfrac{1}{5}x^5 + \tfrac{1}{3}ax^3 + \tfrac{1}{2}bx^2 + cx. \qquad (5.9;\ 0)$$

1. Критические точки: $x^4 + ax^2 + bx + c = 0.$ (5.9; 1)
2. Дважды вырожденные: $4x^3 + 2ax + b = 0.$ (5.9; 2)
3. Трижды вырожденные: $12x^2 + 2a = 0.$ (5.9; 3)
4. Четырежды вырожденные: $24x = 0.$ (5.9; 4)

Четырежды вырожденные критические точки определяются так же точно, как в (5.6):

$$(5.9;\ 4) \Rightarrow x = 0 \overset{(5.9;\ 3)}{\Rightarrow} a = 0 \overset{(5.9;\ 2)}{\Rightarrow} b = 0 \overset{(5.9;\ 1)}{\Rightarrow} c = 0. \qquad (5.10)$$

Это означает, что функция $F(x; 0, 0, 0)$ имеет четырежды вырожденную критическую точку $x = 0$.

Линии, соединяющие точки, которые характеризуют поведение функции с трижды вырожденными критическими точками, имеют следующее параметрическое представление в пространстве управляющих параметров $\mathbb{R}^3$ (рис. 5.6):

$$(5.9;\ 3) \Rightarrow a = -6x^2 \overset{(5.9;\ 2)}{\Rightarrow} b = 8x^3 \overset{(5.9;\ 1)}{\Rightarrow} c = -3x^4. \qquad (5.11)$$

Такое представление может быть получено следующим образом.

Точки, характеризующие функции с дважды вырожденными критическими точками, образуют поверхность, которая в пространстве параметров $\mathbb{R}^3$ может быть представлена как

$$(5.9;\ 2) \Rightarrow b = -4x^3 - 2ax \overset{(5.9;\ 1)}{\Rightarrow} c = 3x^4 + ax^2, \qquad (5.12)$$

т. е. в параметрическое представление двумерной поверхности, определяемой (5.12), входят как координата дважды вырожденной критической точки x, так и значение управляющего

параметра a. Для изучения этой поверхности используем следующее наблюдение: если вводить новый масштаб по x с помощью множителя λ, а по a — с помощью множителя λ^2, то при этом b умножается на множитель λ^3, а c — на λ^4:

$$\text{Если } \begin{matrix} x \to \lambda x \\ a \to \lambda^2 a, \end{matrix} \quad \text{тогда} \quad \begin{matrix} b \to \lambda^3 b \\ c \to \lambda^4 c. \end{matrix} \tag{5.13}$$

Таким образом, необходимо лишь определить поперечное сечение b-c поверхности (5.12) в трех плоскостях, скажем, $a = +1$, $a = 0$, $a = -1$. Затем, изменяя масштаб в этих попе-

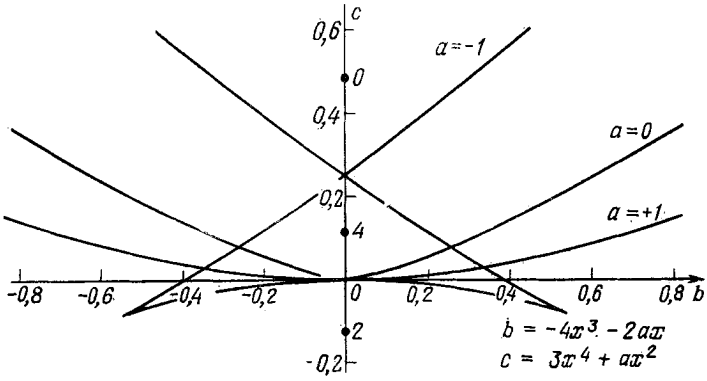


Рис. 5.5. Поперечные сечения $a = +1$, $a = 0$, $a = -1$ сепаратрисы A_4 в пространстве $\mathbb{R}^3$, дающие параметрическое представление (5.12) для b и c.
Сечения находятся фиксацией соответствующего значения a (x меняется).

речных сечениях и используя формулы (5.13), получим всю поверхность. Поперечные сечения, изображенные на рис. 5.5, получены при фиксированных значениях a. Если собрать эти поперечные сечения вместе, то, изменяя масштаб, получим поверхность, изображенную на рис. 5.6. Данная поверхность делит $\mathbb{R}^3$ на три открытые области. Качественное поведение функций, параметризуемых точками любой одной области, одинаково и изменяется только при прохождении сквозь эту поверхность.

Чтобы определить качественный тип любой области, достаточно рассмотреть любую ее точку, например точки 0, 2, 4 (рис. 5.5). Уравнение $(x^2)^2 + ax^2 + c = 0$ имеет решение

$$x = \pm\left[-(a/2) \pm \sqrt{(a/2)^2 - c}\right]^{1/2}$$

при $a > 0$: два вещественных корня, если $c < 0$, и ни одного вещественного корня, если $c > 0$; *при* $a < 0$: два вещественных корня, если $c < 0$, четыре вещественных корня, если $0 < c <$

$< (a/2)^2$, ни одного вещественного решения, если $0 < c < (a/2)^2$.

Интерес может также представлять тип функций, соответствующих точкам поверхности (5.12) на ребрах (5.11) или на

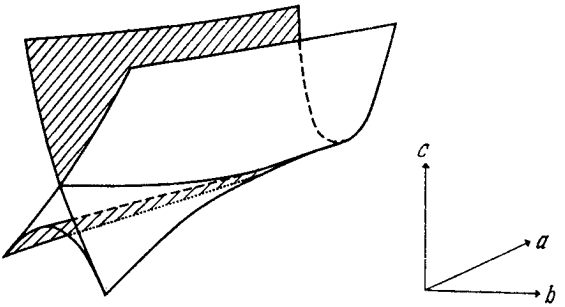


Рис. 5.6. При построении сепаратрисы катастрофы A_4 в пространстве $\mathbb{R}^3$ три поперечных сечения, изображенные на рис. 5.5, могут быть соединены вместе посредством пересчетных соотношений (5.13).

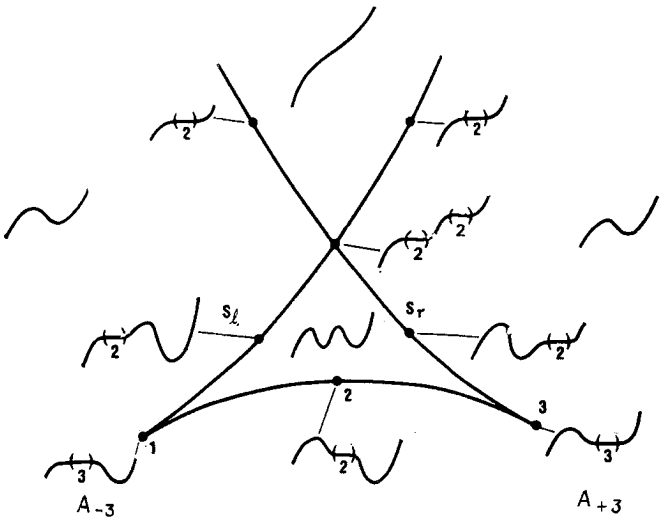


Рис. 5.7. Любая точка сепаратрисы, изображенной на рис. 5.6, представляет неморсовскую функцию.

Здесь изображены некоторые морсовские функции, ассоциированные с различными компонентами сепаратрисы.

линии самопересечения. Из (5.12) следует, что в точке 2 (рис. 5.7) при $x = 0$ имеется дважды вырожденная критическая точка, так что изолированные критические точки встречаются как справа, так и слева от нее. Согласно (5.11), получаем, что

в точке 1 трижды вырожденная критическая точка существует при $x < 0$, а в точке 3— при $x > 0$. Обе левосторонние критические точки вырождены на кривой S_l (рис. 5.7), а обе правосторонние критические точки —на кривой S_r. Линия пересечения левосторонней и правосторонней поверхностей (рис. 5.6) описывает функции с двумя дважды вырожденными критическими

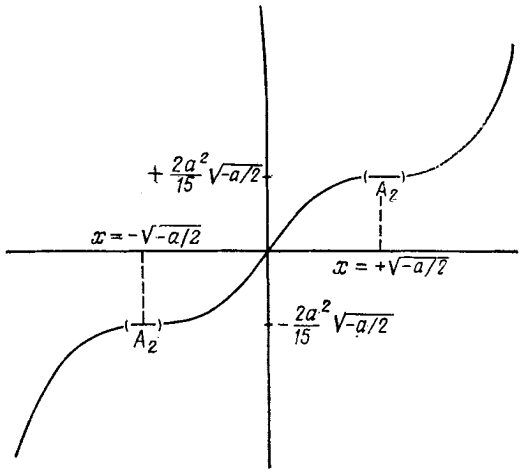


Рис. 5.8. Точки, расположенные вдоль самопересечения двух крыльев «складки» сепаратрисы, изображенной на рис. 5.6, представляют функции с парой дважды вырожденных критических точек.

точками. На этой линии $b = 0$, а $c = (a/2)^2$, так что уравнение (5.9; 1) принимает вид

$$x^4 + ax^2 + \left(\frac{a}{2}\right)^2 = \left(x^2 + \frac{a}{2}\right)^2 = 0$$

и имеет кратные корни $x = -\sqrt{-a/2}$ и $x = +\sqrt{-a/2}$ при $a < 0$; соответствующие значения потенциальной функции (5.9; 0) изображены на рис. 5.8.

Итак, множество точек, в которых $F(x; a, b, c)$ имеет неморсовские критические точки, разделяет пространство управляющих параметров $\mathbb{R}^3$ на три открытые области. Любая точка $\mathbb{R}^3$ может быть аппроксимирована с любой наперед заданной точностью последовательности точек, полностью лежащих в одной из этих областей. Это значит, что неморсовские функции могут быть приближены с любой необходимой точностью морсовскими функциями. Сепаратриса состоит из точек, характеризующих функции с дважды вырожденными критическими точками, двух кривых, описывающих функции с трижды вырожденными критическими точками, кривой, описывающей функции с двумя два-

жды вырожденными критическими точками, и трех поверхностей, описывающих функции с дважды вырожденными критическими точками.

3.4. Катастрофы типа D_{+4}

Как было показано (гл. 4, разд. 6), для катастроф, имеющих две переменные состояния и более трех управляющих параметров, всегда можно выбрать возмущения так, чтобы исключить два или три квадратичных одночлена. Более общим образом можно выбрать возмущение так, чтобы удалить две из трех линейно независимых комбинаций одночленов второй степени, оставляя только одну такую комбинацию вместе с линейными членами, как и в случае наиболее общего возмущения [1]).

Функция катастрофы D_{+4} имеет вид

$$D_{+4}: F(x,y;\ a,\ b,\ c) = x^2y + \tfrac{1}{3}y^3 + a(y^2 - x^2) + bx + cy. \quad (5.14;\ 0)$$

Следует заметить, что эта функция имеет несколько симметрий, из которых наиболее полезна для нас $F \to +F$, если $(x, y) \to$ $\to (-x, +y)$ и $(a, b, c) \to (+a, -b, +c)$.

Критические точки F определяются из уравнений

$$\partial F/\partial x = 2xy - 2ax + b = 0, \quad (5.14;\ 1а)$$

$$\partial F/\partial y = x^2 + y^2 + 2ay + c = 0. \quad (5.14;\ 1б)$$

Матрица устойчивости D_{+4} имеет вид

$$F_{ij} = 2\begin{bmatrix} y-a & x \\ x & y+a \end{bmatrix}. \quad (5.14;\ 2)$$

Когда все элементы матрицы устойчивости обращаются в нуль, появляются четырежды вырожденные критические точки:

$$F_{ij} = 0 \overset{(5.14;\ 2)}{\Rightarrow} x = y = a = 0 \overset{(5.14;\ а)}{\Rightarrow} b = 0, \overset{(5.14;\ б)}{\Rightarrow} c = 0. \quad (5.15)$$

Отсюда следует, что точка $(a, b, c) = (0, 0, 0) \in \mathbb{R}^3$, принадлежащая пространству управляющих параметров, параметризует функцию

$$F(x,\ y;\ 0,\ 0,\ 0) = x^2y + \tfrac{1}{3}y^3, \quad (5.16)$$

---

[1]) В этом и последующих разделах будем выбирать члены второй степени в возмущении таким образом, чтобы минимизировать сложность вычислений.

которая имеет четырежды вырожденную критическую точку $(x, y) = (0, 0) \in \mathbb{R}^2$ в пространстве переменных состояния. Функция (5.16) изображена на рис. 5.9.

Дважды и трижды вырожденные критические точки могут быть найдены из условия, согласно которому по крайней мере одно собственное значение матрицы устойчивости обращается в

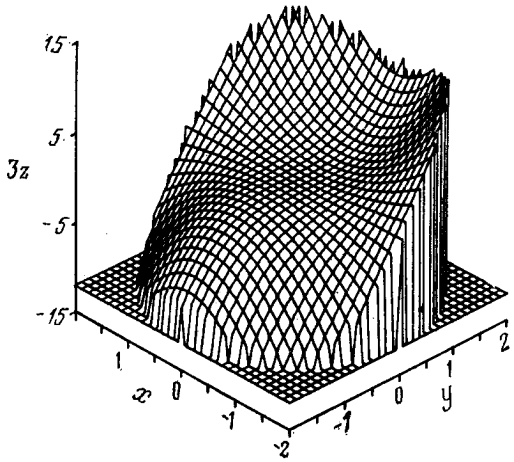


Рис. 5.9. Росток катастрофы D_{+4}. (Воспроизводится с разрешения Н. А. Поупа.)

нуль. Если это имеет место, то определитель матрицы (5.14; 2) равен нулю

$$\det F_{ij} = 4(y^2 - a^2 - x^2) = 0. \tag{5.17}$$

Таким образом, «плохое множество» образует гиперболу

$$y^2 - x^2 = a^2. \tag{5.18}$$

Это означает, что если критические точки лежат на этой гиперболе, то они дважды или трижды вырожденные.

Теперь, используя переменные x и y, получим параметрическое представление сепаратрисы в пространстве управляющих параметров: параметрическое представление $a(x, y)$ дается (5.18), а параметрическое представление для b и c может быть найдено из уравнений (5.14; 1а, б):

$$\begin{aligned} -b &= 2xy - 2ax, \\ -c &= x^2 + y^2 + 2ay. \end{aligned} \tag{5.19}$$

Заметим, что эту же задачу можно решить значительно проще. Так, если по переменным x и y ввести новый масштаб λ,

это будет означать и новый масштаб для a, т. е. ($x \to \lambda x$, $y \to \lambda y \Rightarrow a \to \lambda a$). Тогда для b и c надо ввести масштаб λ^2:

$$\left.\begin{matrix} x \to \lambda x \\ y \to \lambda y \end{matrix}\right\} \Rightarrow \begin{matrix} a \to \lambda a \\ b \to \lambda^2 b, \\ c \to \lambda^2 c \end{matrix} \quad V \to \lambda^3 V. \tag{5.20}$$

Следовательно, достаточно определить поперечное сечение b-c сепаратрисы в плоскостях $a = +1$, $a = 0$, $a = -1$, а затем, используя соотношение (5.20), восстановить всю сепаратрису. Сделав замену при $a > 0$:

$$\begin{aligned} y &= -a \operatorname{ch} \theta, \\ x &= -a \operatorname{sh} \theta, \end{aligned} \tag{5.21, верхняя}$$

можно определить, как верхняя ветвь гиперболы параметризует b-c сепаратрису:

$$\begin{aligned} -\frac{b}{a^2} &= \operatorname{sh} 2\theta + 2 \operatorname{sh} \theta, \\ -\frac{c}{a^2} &= \operatorname{ch} 2\theta + 2 \operatorname{ch} \theta. \end{aligned} \tag{5.22, верхняя}$$

Оставшуюся часть этой сепаратрисы параметризует нижняя ветвь гиперболы. Осуществляя замену

$$\begin{aligned} y &= -a \operatorname{ch} \theta, \\ x &= -a \operatorname{sh} \theta, \end{aligned} \tag{5.21, нижняя}$$

в этом случае получаем

$$\begin{aligned} -\frac{b}{a^2} &= \operatorname{sh} 2\theta + 2 \operatorname{sh} \theta, \\ -\frac{c}{a^2} &= \operatorname{ch} 2\theta - 2 \operatorname{ch} \theta. \end{aligned} \tag{5.22, нижняя}$$

Как уже отмечалось, $\theta \to -\theta \Rightarrow (x, y) \to (-x, +y)$ и $(a, b, c) \to \to (+a, -b, +c)$. Зависимости (5.22) легко представить графически (рис. 5.10).

Все три поперечных сечения, изображенные на рис. 5.10, могут быть связаны вместе использованием пересчетных соотношений (5.20). При этом получим сепаратрису, или бифуркационное множество катастрофы D_{+4}, которое изображено на рис. 5.11. Сепаратрису можно представить как две пересекающиеся в плоскости $a = 0$ поверхности. Одна из поверхностей имеет складку перед плоскостью $a = 0$ и не имеет ее сзади этой плоскости. Другая поверхность является зеркальным отражением первой. Эти поверхности делят пространство управляющих параметров

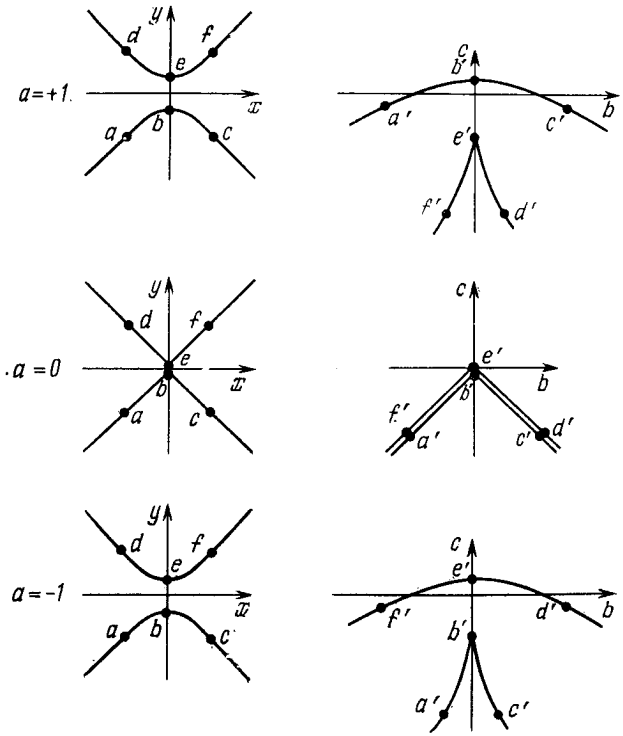


Рис. 5.10. Взаимосвязь между гиперболическими «плохими множествами» и сепаратрисой катастрофы D_{+4} в пространстве $\mathbb{R}^3$.

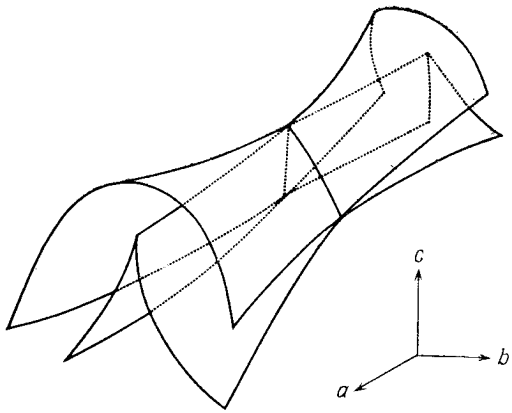


Рис. 5.11. Сепаратриса для D_{+4} в пространстве $\mathbb{R}^3$ может быть построена из трех поперечных сечений, изображенных на рис. 5.10, используя пересчетные соотношения (5.20).

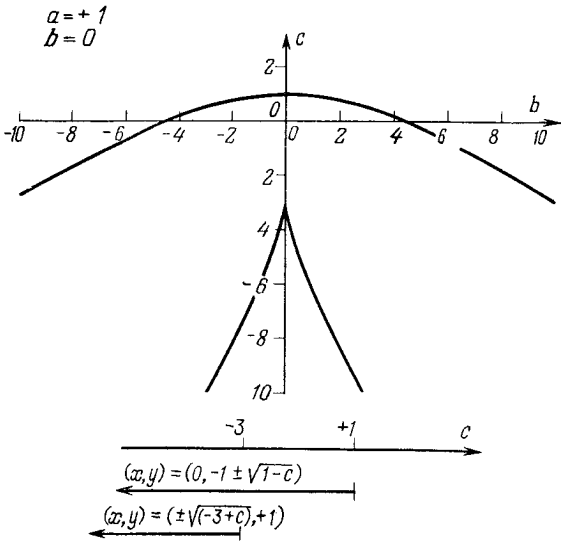


Рис. 5.12. В плоскости $a = +1$ структурно устойчивые морсовские функции, параметризуемые любой из трех открытых областей, могут быть определены, если $b = 0$, а c произвольно меняется.

$\mathbb{R}^3$ на четыре открытые области, которые описывают функции четырех качественно различных типов.

Чтобы определить качественно поведение функций, параметризуемых точками каждой из этих открытых областей, исследуем свойства функций, параметризуемых точками вдоль прямой $a = +1$, $b = 0$ в $\mathbb{R}^3$ (рис. 5.12). Вдоль этой прямой уравнения (5.14), определяющие критические точки, становятся

$$2xy - 2x = 0, \tag{5.23а}$$

$$x^2 + y^2 + 2y + c = 0. \tag{5.23б}$$

Из (5.23а) следует, что критические точки должны иметь координатами либо $x = 0$, либо $y = +1$. Если $x = 0$, то из (5.23б) получим

$$x = 0 \overset{(5.23б)}{\Rightarrow} y = -1 \pm \sqrt{1 - c}. \tag{5.24}$$

Существует пара вещественных корней при $c < +1$. Если $y = +1$, то из (5.23б) имеем

$$y = 1 \overset{(5.23б)}{\Rightarrow} x = \pm \sqrt{-(3 + c)}. \tag{5.25}$$

Следовательно, только при $c < 3$ имеется пара вещественных критических точек. Положение и область значений c, для

которых существуют критические точки, — (5.24) и (5.25), изображены на рис. 5.12.

Как только положение (координаты) критических точек найдено, их свойства устойчивости (морсовский тип седла) могут быть определены из матрицы устойчивости (5.14; 2).

В критических точках (5.24) матрица устойчивости имеет следующий вид:

$$F_{ij} = 2\begin{bmatrix} y-1 & 0 \\ 0 & y+1 \end{bmatrix}, \quad x=0, \quad y=-1\pm\sqrt{1-c}. \qquad (5.26)$$

Следовательно, оси x и y являются главными направлениями, а собственные значения в этих главных направлениях равны со-

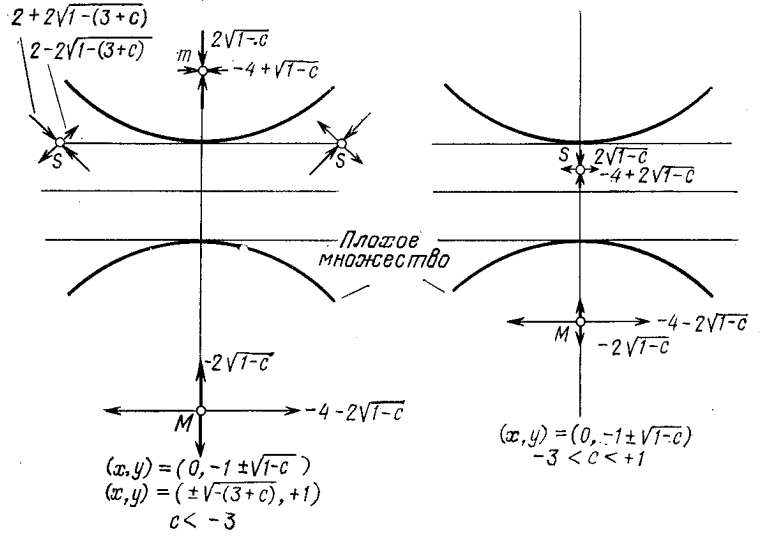


Рис. 5.13. Критические точки катастрофы D_{+4} в плоскости (x, y) как функции управляющего параметра c при $a = +1$, $b = 0$.

ответственно $2(y-1)$ и $2(y+1)$. Собственное значение в направлении оси x верхней критической точки меняет знак при прохождении y через $+1$, в то время как собственное значение в направлении оси y обращается в нуль, когда $y = -1$. Это и есть те точки, где критические точки (5.24) пересекают «плохое» множество (5.18) (рис. 5.13).

Аналогично определяются свойства устойчивости и главные оси критических точек (5.25). Из матрицы устойчивости нахо-

дим

$$F_{ij} \overset{(5.25)}{\to} 2\begin{bmatrix} 0 & \pm\sqrt{-(3+c)} \\ \pm\sqrt{-(3+c)} & 2 \end{bmatrix} \Rightarrow$$
$$\Rightarrow \begin{matrix} \lambda_1 = 2\left(1+\sqrt{1-(3+c)}\right), \\ \lambda_2 = 2\left(1-\sqrt{1-(3+c)}\right). \end{matrix} \tag{5.27}$$

В этом случае главные направления уже не совпадают с направлениями осей x и y и определяются при помощи простого процесса диагонализации матрицы. Результаты подобных вычислений приведены на рис. 5.13: два корня пересекают «плохое» множество (5.18) в точке $(x, y) = (0, 1)$ при $c = -3$. В этой точке собственное значение $\lambda_2 = 0$, и она, несомненно, представляет особый интерес. Однако, прежде чем обсуждать, что же может произойти в этой точке, когда $c = -3$ или $c = +1$, заметим, что при $c > 1$ нет критических точек, при $-3 < c < +1$ имеется один минимум и одно седло, а при $c < -3$ — один минимум, один максимум и два седла. Это полностью определяет качественные изменения в поведении всех функций, параметризуемых тремя из четырех открытых областей, изображенных на рис. 5.11. В четвертой области («средняя» область при $a < 0$) имеется один минимум и одно седло.

Если c стремится к $+1$ снизу, то критические точки $(x, y) = (0, -1 \pm \sqrt{1-c})$ стремятся друг к другу и становятся вырожденной точкой в $(0, -1)$ при $c = +1$, которая исчезает при $c > 1$. Собственное значение в направлении оси x стремится к -4, в то время как собственное значение в направлении оси y становится исчезающе малым и обращается в нуль в пределе при $c = +1$. Это означает, что мы имеем дело с катастрофой, так как число изолированных критических точек изменяется, когда c проходит через $+1$, где $\det F_{ij} = 0$. Поскольку в ней присутствуют две критические точки, то вполне естественно ожидать, что это катастрофа складки или A_2. Так как собственное значение отлично от нуля по оси x и равно нулю по оси y, то направление x считается «хорошим», а направление y — «плохим».

Эти соображения можно сделать строгими, если найти каноническую форму для неморсовской функции $F(x, y; a, b, c)$ при $(a, b, c) = (1, 0, 1)$ в дважды вырожденной критической точке $(x, y) = (0, -1)$:

$$F(x, y; 1, 0, 1) = x^2 y + \tfrac{1}{3}y^2 + (y^2 - x^2) + y. \tag{5.28}$$

Полезно разложить эту функцию в окрестности вырожденной критической точки, сделав следующее неоднородное линейное

преобразование $x = X$, $y = (y + 1) - 1 = Y - 1$. Тогда

$$F(X, Y) = X^2Y + \frac{1}{3}Y^3 + 2X^2 - \frac{1}{3}. \tag{5.29}$$

Постоянный член не играет существенной роли и его можно не принимать во внимание. Матрица устойчивости $F(X, Y)$ в дважды вырожденной критической точке $(X, Y) = (0, 0)$ имеет вид

$$F_{ij}|_{(0,\,0)} = \begin{bmatrix} -4 & 0 \\ 0 & 0 \end{bmatrix}. \tag{5.30}$$

Следовательно, направление по оси X «хорошее», а направление по оси Y «плохое».

Теперь попытаемся привести функцию (5.29) к каноническому виду (2.3б), сделав следующую гладкую замену переменных:

$$\begin{aligned} x' &= X + \Delta_1 = X + A_{20}X^2 + A_{11}XY + A_{02}Y^2 + \ldots, \\ y' &= Y + \Delta_2 = Y + B_{20}X^2 + B_{11}XY + B_{02}Y^2 + \ldots. \end{aligned} \tag{5.31}$$

Поскольку есть все основания полагать, что мы имеем дело с ростком катастрофы типа A_2, а коэффициент при x'^2 должен быть равен -2, замену переменных (5.31) надо выбрать так, чтобы выполнялось равенство [1])

$$\begin{aligned} &-2X^2 + X^2Y + \frac{1}{3}Y^3 \stackrel{?}{=} -2x'^2 + ky'^3 = \\ &= -2\underset{2}{\left(X^2 + 2X\Delta_1 + \Delta_1^2\right)} + k\underset{3}{\left(Y^3 + 3Y^2\Delta_2 + 3Y\Delta_2^2 + \Delta_2^3\right)}. \end{aligned} \tag{5.32}$$

Это равенство выполняется для членов второй степени. Для членов третьей степени находим, что

$$X^2Y + \frac{1}{3}Y^3 \stackrel{?}{=} -4X(A_{20}X^2 + A_{11}XY + A_{02}Y^2) + kY^3. \tag{5.33}$$

Данное уравнение удовлетворяется при выборе

$$A_{20} = 0, \quad A_{11} = -\frac{1}{4}, \quad A_{02} = 0, \quad k = \frac{1}{3}. \tag{5.34}$$

Для членов четвертой степени и выше все коэффициенты в левой части (5.32) равны нулю. Соответствующие члены в правой части формулы зависят от находящихся в нашем распоряжении коэффициентов A_{ij}, B_{ij} $(i + j \geqslant 2)$, которые мы используем в замене (5.31). Если для данной степени в нашем распоряжении в правой части (5.32) имеется больше коэффициентов, чем в левой части, то $F(X, Y)$ можно привести к каноническому виду

[1]) Целые числа под каждым членом указывают одночлены наименьшей степени в рассматриваемом произведении.

$-2x'^2+\frac{1}{3}y'^3$. Например, для членов пятой степени в левой части формулы (5.32) имеется ровно шесть фиксированных коэффициентов, и они все равны нулю. Члены степеней вплоть до четвертой возникают из Δ_1 (тех членов четвертой степени, которые входят линейно), а члены степеней вплоть до третьей возникают из Δ_2 (тех членов третьей степени, которые входят линейно). Количество новых (не представленных в рассмотрении четвертой степени) коэффициентов, находящихся в нашем распоряжении, равно $5+4=9$, так что мы обладаем достаточной свободой выбора этих коэффициентов с целью исключить все члены пятой степени. Для членов малых степеней эти соображения подсчета суммированы в табл. 5.1. Из этой таблицы совершенно

Таблица 5.1. **Соображения непрерывности, используемые при приведении функции $F(x, y; a, b, c)$ к канонической форме (2.3б) в точке $(a, b, c)=(1, 0, 1)$**

Левая сторона		Правая сторона			
		Δ_1		Δ_2	
степень	число новых фиксированных коэффициентов	степень	число новых свободных коэффициентов	степень	число новых свободных коэффициентов
4	5	3	4	2	3
5	6	4	5	3	4
6	7	5	6	4	5
Возрастает на 1 ↓		Возрастает на 1 ↓		Возрастает на 1 ↓	

ясно, что функция $F(X, Y)$ в действительности может быть приведена к предполагаемому каноническому виду. Кроме того, при помощи простого преобразования масштабов по x' и y' получим каноническую форму

$$F(X, Y) \doteq -x''^2+y''^3, \tag{5.35}$$

где $x''=(2)^{1/2}x'$, $y''=(3)^{1/3}y'$.

Теперь исследуем другую интересную точку прямой $[a=1, b=0]$, а именно $c=-3$. Когда c стремится к -3 снизу, то три критические точки $(x, y)=(\pm\sqrt{-(3+c)}, 1)$ и $(0, -1+\sqrt{1-c})$ стремятся друг к другу, образуя при $c=-3$ одну трижды вырожденную точку $(x, y)=(0, 1)$, которая становится простой критической точкой, когда c стремится к -3 сверху. При стремлении c к -3 снизу главные оси этих трех критических точек вращаются в направлении осей x и y и остаются

направленными по этим осям, если $c > -3$. Собственное значение в направлении оси y стремится к 4 вблизи $c = -3$, в то время как собственное значение по оси x становится равным нулю при $c = -3$. Когда c проходит через -3, знак собственного значения по оси x в критической точке $(x, y) = (0, -1 + \sqrt{1-c})$ меняется на противоположный. Это дает возможность предположить наличие катастрофы типа $A_{\pm 3}$ с «хорошим» направлением по оси y и «плохим» — по оси x.

Теперь проведем более детальный анализ точки $c = -3$ прямой $[a = 1, b = 0]$. Для этого сначала разложим функцию $F(x, y; a, b, c)$, где $(a, b, c) = (1, 0, -3)$, в трижды вырожденной критической точке $(x, y) = (0, 1)$, используя преобразование $x = X$, $y = (y - 1) + 1 = Y + 1$:

$$F(X, Y) = X^2Y + \frac{1}{3}Y^3 + 2Y^2 - \frac{5}{3}. \tag{5.36}$$

Постоянный член для нас не важен, и мы его исключим из дальнейших рассмотрений. Матрица устойчивости $F(X, Y)$ в точке $(0, 0)$ равна

$$F_{ij}|_{(0,\,0)} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}. \tag{5.37}$$

Ясно, что X — это «плохая» переменная, а Y — «хорошая».

Теперь попытаемся привести (5.36) к каноническому виду путем осесохраняющего нелинейного преобразования типа (5.31). Предполагаемая при этом каноническая форма имеет вид $k'x'^4 + 2y'^2$. В результате получим равенство

$$2Y^2 + X^2Y + \frac{1}{3}Y^3 \stackrel{?}{=} 2y'^2 + kx'^4 =$$

$$= 2\underset{2}{\left(Y^2 + 2Y\Delta_2 + \Delta_2^2\right)} + k\underset{4}{\left(X^4 + 4X^3\Delta_1 + 6X^2\Delta_1^2 + 4X\Delta_1^3 + \Delta_1^4\right)}, \tag{5.38}$$

которое выполняется для членов второй степени. Для членов третьей и четвертой степени соответственно имеем

$$X^2Y + \frac{1}{3}Y^3 \stackrel{?}{=} 4Y(B_{20}X^2 + B_{11}XY + B_{02}Y^2), \tag{5.39}$$

$$0 \stackrel{?}{=} 4Y(B_{30}X^3 + B_{21}X^2Y + B_{12}XY^2 + B_{03}Y^3) + 2\left(\frac{1}{4}X^2 + \frac{1}{12}Y^2\right)^2 + kX^4, \tag{5.40}$$

причем (5.39) справедливо только при $B_{20} = \frac{1}{4}$, $B_{11} = 0$, $B_{02} = \frac{1}{12}$, а (5.40) выполняется при условии

$$\frac{1}{8}+k=0, \qquad B_{30}=0,$$
$$4B_{21}+\frac{1}{12}=0, \qquad B_{12}=0. \tag{5.41}$$
$$4B_{03}+\frac{1}{72}=0,$$

Используя табл. 5.1, можно показать, что (5.38) удовлетворяется для всех членов больше четвертой степени.

Для того чтобы читатель с неослабевающим вниманием следил за ходом наших рассуждений, используем на этот раз свойство симметрии функции (5.36): $F(X, Y)=F(-X, Y)$. В силу симметрии коэффициенты одночленов X^pY^q обязательно должны быть нулями при условии, что p — нечетно. Потребуем, чтобы и каноническая форма для $F(X, Y)$ имела тот же тип симметрии, т. е. $(X, Y)\to(-X, Y)\Rightarrow(x', y')\to(-x', y')$. Тогда не все коэффициенты в Δ_1, Δ_2 свободны и находятся в нашем распоряжении; некоторые из них обязательно должны быть нулями. Теперь построим табл. 5.2, аналогичную табл. 5.1, которая будет

Таблица 5.2. **Соображения непрерывности, включая рассмотрение симметрии, используемые при приведении функции $F(x, y; a, b, c)$ к канонической форме (2.3б) в точке $(a, b, c)=(1, 0, -3)$**

Левая сторона		Правая сторона				
		Δ_1		Δ_2		Общее число свободных коэффициентов
степень	число новых фиксированных коэффициентов	степень	число новых свободных коэффициентов	степень	число новых свободных коэффициентов	
5	3	2	1	4	3	4
6	4	3	2	5	3	5
7	4	4	2	6	4	6
8	5	5	3	7	4	7
9	5	6	3	8	5	8
	Возрастает на $1/2$ ↓					Возрастает на 1 ↓

показывать число коэффициентов в левой и правой частях формулы (5.38), не равных нулю в силу требуемой симметрии. Из табл. 5.2 видно, что

$$F(X, Y)\doteq 2y'^2-\frac{1}{8}x'^4. \tag{5.42}$$

Путем изменения масштаба вводим новые координаты $x'' = (8)^{-1/4}x'$ и $y'' = (2)^{1/2}y'$, что дает каноническую форму

$$F(X, Y) \doteq -x''^4 + y''^2. \tag{5.43}$$

◇◇◇ Ввиду аналогии между формулами (5.29) и (5.38) может показаться весьма заманчивым предположить в (5.38) канони-

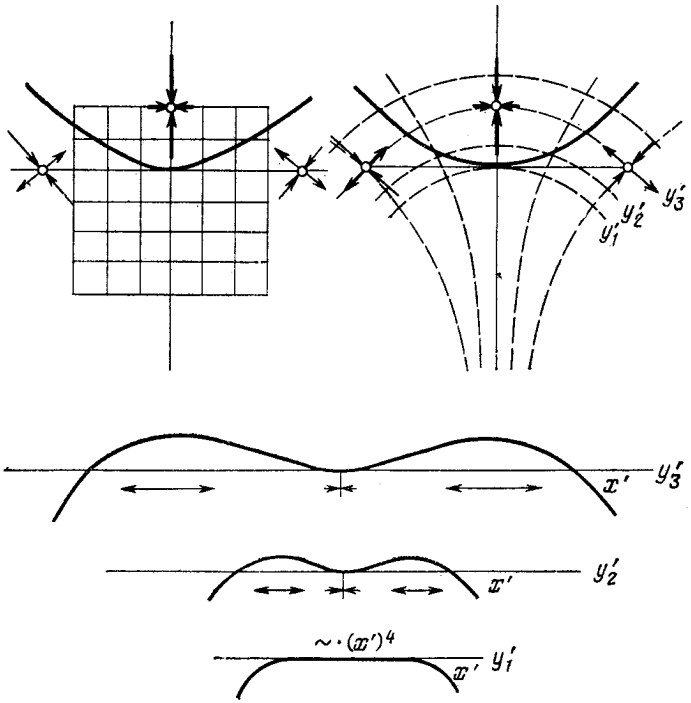


Рис. 5.14. Гладкое преобразование из декартовой в криволинейную систему координат, показывающее, каким образом двойственная сборка ассоциируется с катастрофой D_{+4}.

ческую форму $2y'^2 + kx'^3$. Однако при этом предположении равенство в (5.38) не может быть удовлетворено. Это легко показать, поскольку член kX^3 должен бы тогда появиться в правой части (5.39). Но в этом случае этот член *не мог бы быть* удален никаким выбором коэффициентов, находящихся в нашем распоряжении.

Преобразование (5.31) декартовых координат (X, Y) в криволинейные координаты (x', y') иллюстрируется на рис. 5.14. Ли-

нии уровня $F(x', y')$ при различных значениях $y' = \text{const}$ указывают на присутствие катастрофы «двойственной сборки» $-x^4 = A - 3$.

Подобный анализ может быть проведен и при $a = +1$. Кроме того, можно упростить вычисления, если предположить существование симметрии в плоскости $b = 0$. Когда же $b \neq 0$, симметрия разрушается (рис. 5.13), и в этом случае для фиксированных $a = 1$ и $b \neq 0$ линия с возрастающим c должна пробить одну или другую сторону складчатых слоев, изображенных на рис. 5.11. Такие слои представляют катастрофы складки; при этом минимум, изображенный на рис. 5.13, совпадает либо с левым седлом и, таким образом, уничтожает его, либо с правым седлом, так как они встречаются на верхней части гиперболического «плохого» множества (5.18). Оставшиеся верхние седла и максимум совпадают и уничтожаются, когда они встречаются на нижней ветви гиперболического «плохого» множества, при достижении c верхней криволинейной поверхности (рис. 5.11). При $b < 0$ минимум совпадает с правосторонним, а максимум — с левосторонним седлом.

Весь этот анализ можно целиком осуществить и в плоскости $a = -1$. Однако в этом нет необходимости, так как два седла лежат между верхней и нижней ветвями гиперболического «плохого» множества; минимум лежит выше верхней ветви, а максимум — ниже нижней ветви; когда c возрастает, максимум совпадает с левосторонним седлом (если $b < 0$) на левой гиперболической ветви и уничтожает его, а для больших значений c минимум и правостороннее седло уничтожают друг друга на верхней ветви.

В плоскости $a = 0$ самопересечение двух поверхностей описывает присутствие двух катастроф складки, т. е. минимум и седло уничтожают друг друга в некоторой точке плоскости (x, y), в то время как максимум и седло одновременно уничтожают друг друга в некоторой другой точке «плохого» множества.

3.5. Катастрофы типа D_{-4}

В этом случае катастрофы описываются функцией вида

$$D_{-4}: \quad F(x, y; a, b, c) = x^2y - \frac{1}{3}y^3 + a(y^2 + x^2) + bx + cy. \quad (5.44; 0)$$

Прежде всего заметим, что эта функция имеет несколько симметрий, из которых наиболее полезной для нас будет симметрия $F \to +F$, когда $(x, y) \to (-x, +y)$ и $(a, b, c) \to (+a, -b, +c)$. Критические точки F определяются из уравнений

$$\partial F/\partial x = 2xy + 2ax + b = 0, \quad (5.44; 1а)$$

$$\partial F/\partial y = x^2 - y^2 + 2ay + c = 0. \quad (5.44; 1б)$$

Матрица устойчивости F равна

$$F_{ij} = 2\begin{bmatrix} y+a & x \\ x & -y+a \end{bmatrix}. \qquad (5.44;\ 2)$$

Четырежды вырожденная критическая точка определяется из условия обращения в нуль матрицы устойчивости:

$$F_{ij} = 0 \overset{(5.44;\ 2)}{\Rightarrow} x = y = a = 0 \overset{(5.44;\ 1а)}{\Rightarrow} b = 0, \overset{(5.44;\ 1б)}{\Rightarrow} c = 0. \qquad (5.45)$$

Поэтому точка $(a, b, c) = (0, 0, 0) \in \mathbb{R}^2$ в пространстве управляющих параметров параметризует функцию

$$F(x, y; 0, 0, 0) = x^2 y - \frac{1}{3} y^3, \qquad (5.46)$$

которая имеет четырежды вырожденную критическую точку $(x, y) = (0, 0) \in \mathbb{R}^2$ в пространстве переменных состояния. Функ-

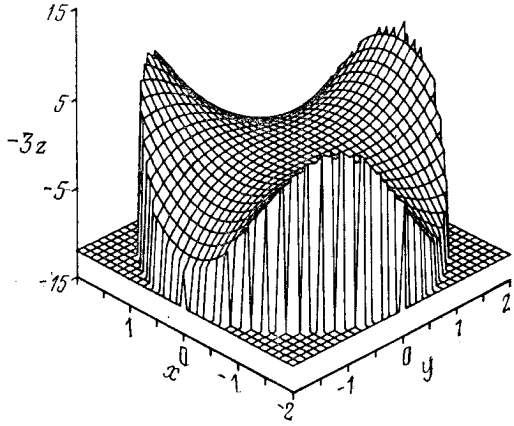


Рис. 5.15. Изображение ростка катастрофы D_{-4}. (Воспроизводится с разрешения Н. А. Поупа.)

ция (5.46) изображена на рис. 5.15. Для нахождения дважды и трижды вырожденных критических точек используют условие, согласно которому одно из собственных значений матрицы устойчивости должно обращаться в нуль. Если это так, то определитель матрицы (5.42) становится равным нулю:

$$\det F_{ij} = 4(a^2 - y^2 - x^2) = 0. \qquad (5.47)$$

Таким образом, «плохое множество» является окружностью

$$y^2 + x^2 = a^2. \qquad (5.48)$$

Это означает, что если критическая точка лежит на этой окружности, то она дважды или трижды вырождена (четырежды вырождена, если $a = 0$).

Чтобы получить параметрическое представление сепаратрисы в пространстве управляющих параметров, можно использовать две переменные x и y. Параметрическое представление $a(x, y)$ дается формулой (5.48), а параметрическое представление для b, c может быть определено из формул (5.44; 1а) и (5.44; 1б):

$$-b = 2xy + 2ax, \qquad -c = x^2 - y^2 + 2ay. \tag{5.49}$$

И в этом случае существует простой путь решения поставленной задачи, и он состоит в изменении масштабов. Соотношения пересчета для формул (5.44) совершенно идентичны соотношениям пересчета для (5.14):

$$\left.\begin{array}{l} x \to \lambda x \\ y \to \lambda y \end{array}\right\} \Rightarrow \begin{array}{l} a \to \lambda a \\ b \to \lambda^2 b \\ c \to \lambda^2 c. \end{array} \tag{5.50}$$

Следовательно, достаточно определить поперечное сечение $b - c$ сепаратрисы в плоскостях $a = +1$, $a = 0$, $a = -1$, а затем с помощью этих соотношений восстановить всю сепаратрису. Если $a > 0$, то, производя замену [см. (5.21), (5.22)]

$$y = -a\cos\theta, \qquad x = a\sin\theta, \tag{5.51}$$

можно определить, каким образом «плохое» множество окружности параметризует b-c сепаратрису. Тогда

$$\frac{b}{a^2} = \sin 2\theta - 2\sin\theta, \qquad \frac{c}{a^2} = \cos 2\theta + 2\cos\theta. \tag{5.52}$$

Функцию (5.52) легко изобразить графически (рис. 5.16). Все три поперечных сечения, показанные на рис. 5.16, можно связать воедино, используя соотношения (5.50). В результате получаем сепаратрису, или бифуркационное множество, катастрофы D_{-4}, изображенное на рис. 5.17. Эта поверхность разделяет пространство управляющих параметров $\mathbb{R}^3$ на три открытые области, которые описывают функции трех качественно различных типов.

Чтобы определить качественные изменения в поведении функций, параметризуемых точками каждой из этих открытых областей, исследуем свойства функций, параметризуемых точками вдоль прямой $a = +1$, $b = 0$ в $\mathbb{R}^3$ (рис. 5.18). Вдоль этой прямой уравнения (5.44), определяющие критические точки, в точности совпадают с уравнениями

$$2xy + 2x = 0, \tag{5.53а}$$

$$x^2 - y^2 + 2y + c = 0. \tag{5.53б}$$

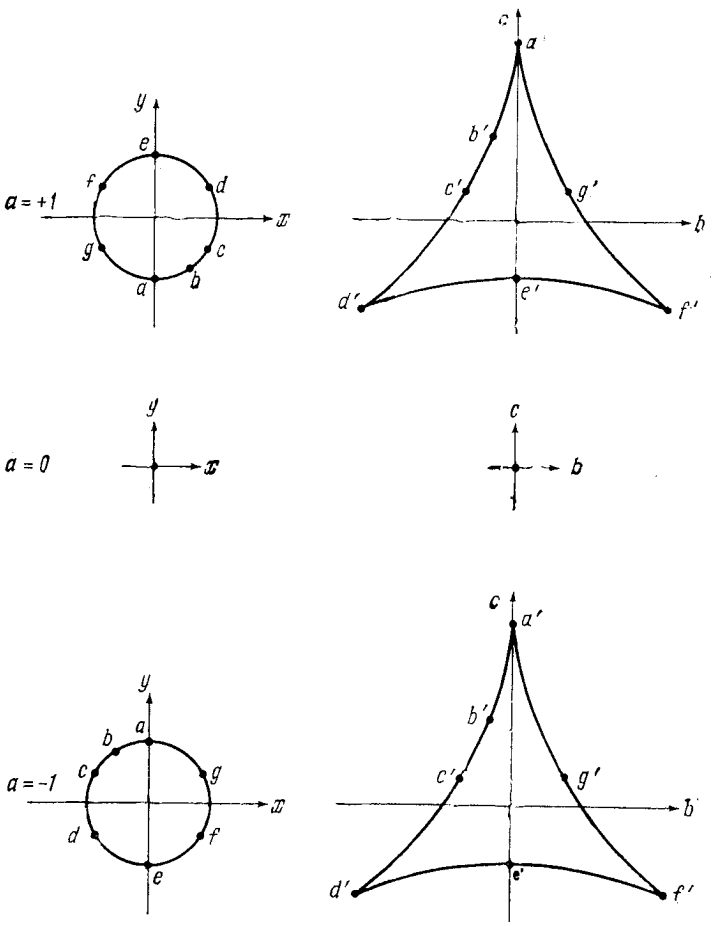


Рис. 5.16. Изображение взаимосвязи между «плохими» множествами формы окружности и сепаратрисой D_{-4} в пространстве $\mathbb{R}^3$.

Из (5.53) следует, что критические точки должны иметь либо $x = 0$, либо $y = -1$. Если $x = 0$, то из (5.53б) получаем

$$x = 0 \overset{(5.53б)}{\Rightarrow} y = +1 \pm \sqrt{1 + c}. \tag{5.54}$$

Существуют две вещественные точки, если $c > 1$. Если $y = -1$, то из (5.53б) имеем

$$y = -1 \overset{(5.53б)}{\Rightarrow} x = \pm \sqrt{3 - c}. \tag{5.55}$$

Эта пара вещественных критических точек существует, если $c < +3$. Положение и область значений c, для которых суще-

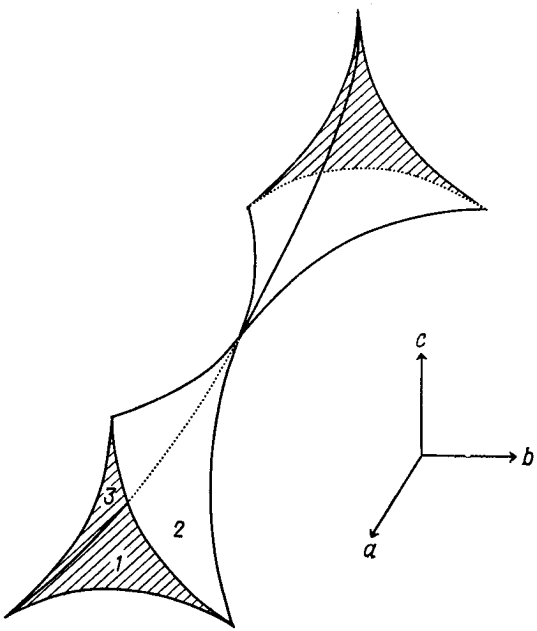


Рис. 5.17. Сепаратриса катастрофы D_{-4} в пространстве может быть построена из трех поперечных сечений, изображенных на рис. 5.16, с помощью пересчетных соотношений (5.50).

ствуют критические точки (5.54) и (5.55), изображены на рис. 5.18.

Если положение критических точек известно, их свойства устойчивости могут быть определены из матрицы устойчивости (5.44; 2). В критических точках (5.34) матрица устойчивости имеет вид

$$F_{ij} = 2\begin{bmatrix} y+1 & 0 \\ 0 & -y+1 \end{bmatrix}, \quad x=0, \quad y=+1\pm\sqrt{1+c}. \qquad (5.56)$$

Таким образом, оси x и y являются главными направлениями с собственными значениями по главным направлениям, равными соответственно $2(1+y)$ и $2(1-y)$. Собственное значение по направлению x нижней критической точки меняет знак, когда y проходит через -1, в то время как собственное значение по направлению y меняет знак, когда $y=+1$. Это и есть те точки, в которых критические точки (5.54) пересекают «плохое» множество (5.48); положение и свойства устойчивости этих критических точек отражены на рис. 5.19.

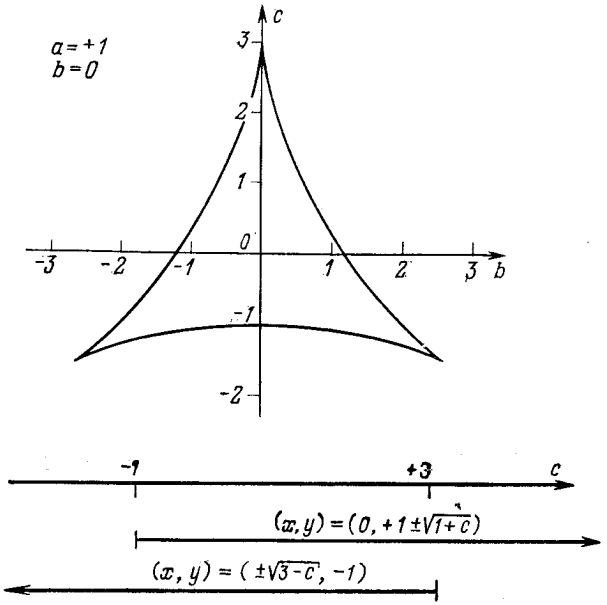


Рис. 5.18. В плоскости $a = +1$ структурно устойчивые морсовские функции, параметризуемые двумя открытыми областями, могут быть определены, если $b = 0$, а c меняется.

Аналогично определяются свойства устойчивости и главные оси критических точек (5.55). Матрица устойчивости дает

$$F_{ij} \overset{(5.55)}{\to} 2\begin{bmatrix} 0 & \pm\sqrt{3-c} \\ \pm\sqrt{3-c} & 2 \end{bmatrix} \Rightarrow$$
$$\Rightarrow \begin{aligned} \lambda_1 &= 2\left(1+\sqrt{1+(3-c)}\right), \\ \lambda_2 &= 2\left(1-\sqrt{1+(3-c)}\right). \end{aligned} \tag{5.57}$$

Теперь главные направления не совпадают с осями x и y, но их легко найти путем диагонализации матрицы F_{ij}. Результаты вычислений изображены на рис. 5.19: эти два корня пересекают «плохое» множество (5.48) в точке $(x, y) = (0, -1)$, когда $c = +3$; в этой точке собственное значение $\lambda_2 = 0$, и мы можем ожидать много интересного.

Принимая во внимание соображения, изложенные в разд. 6, можно сразу же установить, что происходит при возрастании c вдоль прямой $a = +1$, $b = 0$. При $c < -1$ имеется в точности одна пара седловых точек. Когда c возрастает до -1, в точке $(x, y) = (0, 1)$ имеет место катастрофа складки с седлом и ло-

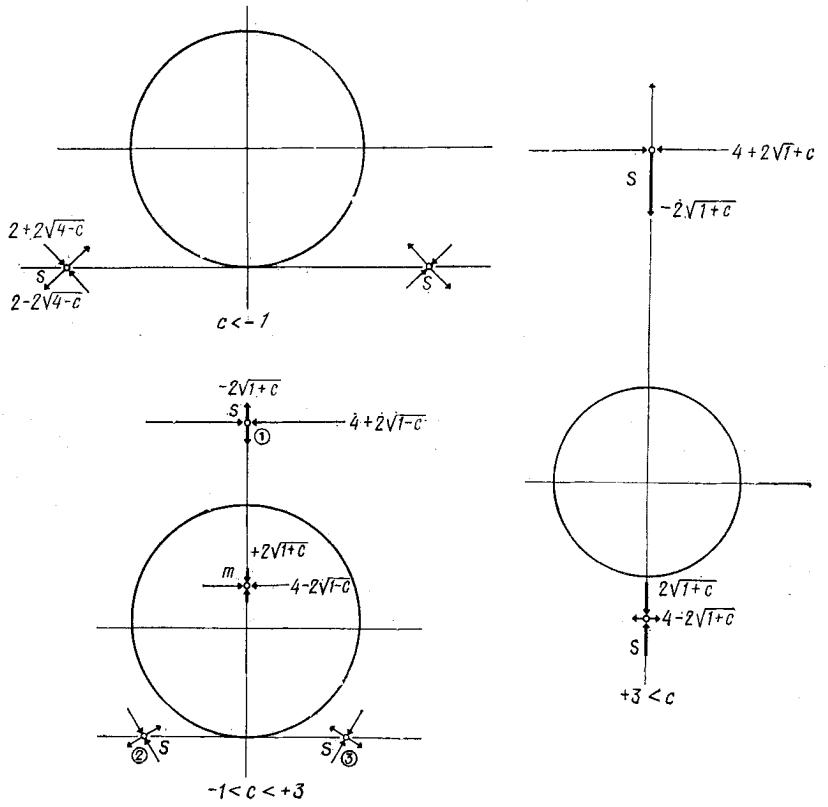


Рис. 5.19. Критические точки катастрофы D_{-4} в плоскости (x, y) как функции управляющего параметра c при $a = +1$, $b = 0$.

кальным минимумом (внутри окружности). Направление по оси x в рассматриваемой катастрофе складки является «хорошим», а направление по оси y — «плохим». При увеличении c до $+3$ новое седло на оси x перемещается на север, локальный минимум движется через окружность по направлению к точке $(x, y) = (0, -1)$, а два исходных седла движутся к одной и той же точке вдоль прямой $y = -1$. Когда три критические точки приближаются к точке $(0, -1)$, главные направления двух седел вращаются в направлении осей x и y. Собственное значение по оси y всех этих точек стремится к $+4$, а собственное значение по оси x — к нулю. Когда c проходит через $+3$, два седла и локальный минимум совпадают в $(x, y) = (0, -1)$, образуя катастрофу двойственной сборки. При $c > +3$ два седла и локальный

минимум объединяются, образуя одно седло (вдоль $x = 0$, $y < -1$). Второе седло существует лишь вдоль $x = 0$, $y > +1$.

В результате проведенного анализа становится очевидным, какие типы функций параметризуются каждой из трех открытых областей, изображенных на рис. 5.17. Область, ограниченная треугольником, параметризует функции, имеющие три седла и один локальный минимум (если $a > 0$) или локальный максимум (если $a < 0$). Седла при этом располагаются вне «плохого множества», представляемого окружностью (5.48), а минимум и максимум — внутри «плохого множества». Область вне треугольника (рис. 5.17) параметризует функции, имеющие два седла, расположенных вне «плохого множества», представляемого окружностью.

Подобный анализ можно провести, используя симметрию в плоскости $b = 0$. (В этом случае вычисления значительно упрощаются, а графические представления становятся более изящными.) При перемещении точки пространства управляющих параметров через сепаратрису (рис. 5.17) возникновение катастрофы складки или сборки зависит от того, движется ли точка через одну из трех поверхностей или через одно из трех ребер. При движении через поверхность (в предположении $a > 0$) минимум совпадает с одним из трех имеющихся седел и уничтожает его. Какое из этих седел уничтожается, зависит от того, какую поверхность пересечет точка при своем перемещении (см. рис. 5.16).

Канонические формы катастроф, имеющих место при $(a, b, c) = (1, 0, -1)$ в точке $(x, y) = (0, +1)$, могут быть определены так же, как в разд. 6. В этой точке

$$F(x, y; 1, 0, -1) = x^2 y - \frac{1}{3} y^3 + (x^2 + y^2) - y. \qquad (5.58)$$

Разложение этой функции в вырожденной критической точке $(0, 1)$ при использовании замены $x = X$, $y = (y - 1) + 1 = Y + 1$ дает

$$F(X, Y) = X^2 Y - \frac{1}{3} Y^3 + 2X^2 - \frac{1}{3}. \qquad (5.59)$$

Очевидно, что X — «хорошее» а Y — «плохое» направление. Можно попытаться найти каноническую форму при помощи нелинейного осесохраняющего преобразования (5.31). Предположим, что $2x'^2 + k'y'^2$ является канонической формой. Теперь попытаемся удовлетворить следующее уравнение:

$$2X^2 + X^2 Y - \frac{1}{3} Y^3 \stackrel{?}{=} + 2x'^2 + k'y'^3. \qquad (5.60)$$

В нашем случае уравнение (5.60) идентично уравнению (5.32). Это в свою очередь означает, что, во-первых, результат, за

исключением знака $\left(\text{т. е. } k' = -\frac{1}{3}\right)$, должен быть таким же, во-вторых, дальнейший вид рассуждений будет идентичен тому, которого мы придерживались после вывода уравнения, а соответ-

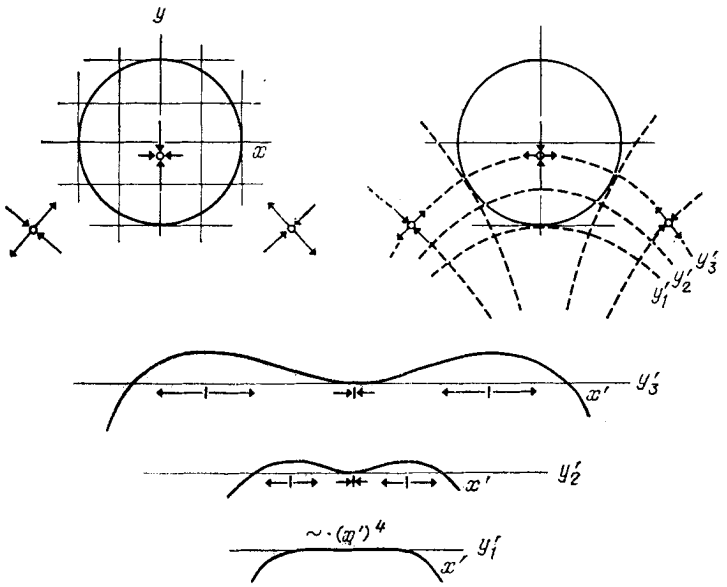


Рис. 5.20. Гладкая замена переменных от декартовой к криволинейной системе координат, показывающая, каким образом катастрофа двойственной сборки связана с катастрофой D_{-4}.

ствующая каноническая форма, аналогичная (5.35), будет иметь вид

$$F(X, Y) \doteq +x''^2 - y''^3. \tag{5.61}$$

В точке $(a, b, c) = (1, 0, +3)$ можно ожидать появления катастрофы сборки в точке $(x, y) = (0, -1)$. В этой точке

$$F(x, y; 1, 0, +3) = x^2y - \frac{1}{3}y^3 + (y^2 + x^2) + 3y. \tag{5.62}$$

Если разложить эту функцию в окрестности данной критической точки, используя замену $x = X$, $y = (y + 1) - 1 = Y - 1$, то

$$F(X, Y) = X^2Y - \frac{1}{3}Y^3 + 2Y^2 - \frac{5}{3}. \tag{5.63}$$

Полученное выражение почти полностью совпадает с формулой (5.36), за исключением знака. Таким образом, ясно, что Y является «хорошим», а X — «плохим» направлением. Поэтому

можно ожидать, что канонической формой $F(X, Y)$ будет $k'x'^4 + 2y'^2$. И, следовательно, можно попытаться удовлетворить уравнению вида

$$2Y^2 + X^2Y - \frac{1}{3}Y^3 \stackrel{?}{=} 2y'^2 + k'x'^4. \tag{5.64}$$

Используя соображения, изложенные при выводе формулы (5.38), получим, что $k' = -\frac{1}{8}$, т. е. имеет место катастрофа двойственной сборки. Соображения подсчета суммированы в табл. 5.2, и, наконец, каноническая форма аналогична (5.43) и равна

$$F(X, Y) \doteq -x''^4 + y''^2. \tag{5.65}$$

Результат перехода от декартовой системы координат X, Y к криволинейной x', y' демонстрируется на рис. 5.20.

4. БИФУРКАЦИОННОЕ МНОЖЕСТВО И МНОЖЕСТВО МАКСВЕЛЛА

Помимо бифуркационных множеств для элементарных катастроф размерности $k \leqslant 3$ существует другое множество точек в $\mathbb{R}^k$, которое представляет структурно неустойчивые функции и называется *множеством Максвелла* $\mathcal{S}_M$. На этом множестве критические значения функции совпадают в двух или более критических точках. Такая вырожденность в критическом значении в общем случае может быть удалена при помощи произвольного возмущения. Множество Максвелла состоит из компонент размерности менее чем k в $\mathbb{R}^k$ и используется для моделирования физических систем, состояния которых описываются с помощью глобального минимума потенциальной функции. В тех случаях, когда управляющие параметры изменяются так, что глобальное минимальное значение перескакивает из одного локального минимума в другой, в поведении системы наблюдаются качественные изменения. Имеющие первостепенный интерес в приложениях компоненты множества Максвелла описывают вырождение критического значения глобального минимума.

4.1. Уравнения Клаузиуса — Клапейрона

Уравнения, определяющие множество Максвелла $\mathcal{S}_M$, называются *уравнениями Клаузиуса — Клапейрона* и могут быть получены из (5.3). Вначале предположим, что при $c = c^0 \in \mathbb{R}^k$ имеется ровно t изолированных критических точек $x^{(1)}, \ldots, x^{(t)}$,

в которых потенциал имеет вырожденные критические значения:

$$V^{(1)} = V^{(2)} = \ldots = V^{(t)},$$
$$V^{(p)} = V(x^{(p)}, c^0), \quad 1 \leqslant p \leqslant t. \tag{5.66}$$

При малом изменении значений управляющих параметров p-е критическое значение изменяется на

$$\delta^{(1)} V^{(p)} = V^{(p)}_{\alpha}\, \delta c^{\alpha}. \tag{5.67}$$

Требование, чтобы все t критических значений изменялись на одно и то же число, накладывает $t-1$ ограничений на вариации δc^{α} управляющих параметров. При этом получаем следующую систему из $t-1$ уравнений:

$$\left[V^{(p)} - V^{(p+1)}\right]_{\alpha} \delta c^{\alpha} = 0, \quad p = 1, 2, \ldots, t-1, \tag{5.68}$$

или в матричной форме

$$\begin{bmatrix} V^{(1)}_1 - V^{(2)}_1 & V^{(1)}_2 - V^{(2)}_2 & \ldots & V^{(1)}_k - V^{(2)}_k \\ V^{(2)}_1 - V^{(3)}_1 & V^{(2)}_2 - V^{(3)}_2 & \ldots & V^{(2)}_k - V^{(3)}_k \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ V^{(t-1)}_1 - V^{(t)}_1 & V^{(t-1)}_2 - V^{(t)}_2 & \ldots & V^{(t-1)}_k - V^{(t)}_k \end{bmatrix} \begin{bmatrix} \delta c^1 \\ \delta c^2 \\ \cdot \\ \cdot \\ \cdot \\ \delta c^k \end{bmatrix} = 0. \tag{5.68'}$$

Эти уравнения Клаузиуса — Клапейрона определяют $k - (t-1)$-мерную касательную плоскость к множеству Максвелла для t вырожденных критических значений в k-параметрическом семействе потенциалов.

Множество из $k+1-t$ базисных векторов данной касательной плоскости в точке c^0 может быть найдено посредством элементарного алгоритма, который имеет место почти везде на множестве Максвелла $\mathcal{S}_M$. Введем следующие обозначения:

$$B_j = \begin{bmatrix} \delta c^1 \\ \cdot \\ \cdot \\ \cdot \\ \delta c^{t-1} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} + C_j\, \delta s_j, \quad C_j = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad t \leqslant j \leqslant k. \tag{5.69}$$

Вектор-столбец C_j имеет нули во всех строках, за исключением j-й строки, а δs_j — множество $k-t+1$ независимых параметров. Тогда $k+1-t$ векторов B_j должны удовлетворять уравнению вида

$$MB_j = C_j\,\delta s_j, \tag{5.70}$$

где M — квадратная матрица порядка k, получаемая заменой $(t-1)\times k$-матрицы, появляющейся в (5.68′), на $(k+1-t)\times$ $\times k$-матрицу, состоящую из нулей и единиц, так что

$$\begin{aligned} &M_{ij} = V_j^{(i)} - V_j^{(i+1)}, && 1 \leqslant i \leqslant t-1, \\ &M_{ij} = \delta_{ij}, && t \leqslant i \leqslant k, \quad 1 \leqslant j \leqslant k. \end{aligned} \tag{5.71}$$

Если положить

$$N_{ij} = (-)^{i+j} \det(\hat{M}_{ji}), \tag{5.72}$$

где $\hat{M}_{ji}$ — квадратная матрица порядка $k-1$, получаемая удалением j-й строки и i-го столбца из матрицы M, то

$$MN = I_k(\det M). \tag{5.73}$$

Если $\det M \neq 0$, то соотношение (5.73) может быть обращено, и тогда получаем

$$B_j \simeq NC_j\,\delta s_j. \tag{5.74}$$

Короче говоря, элементы вектор-столбца B_j пропорциональны кофакторам матрицы M:

$$B_j \simeq \begin{bmatrix} N_{1j} \\ N_{2j} \\ \cdot \\ \cdot \\ N_{kj} \end{bmatrix} = \begin{bmatrix} (-)^{j+1} & \det & (\hat{M}_{j1}) \\ (-)^{j+2} & \det & (\hat{M}_{j2}) \\ & \cdot & \\ & \cdot & \\ (-)^{j+k} & \det & (\hat{M}_{jk}) \end{bmatrix}. \tag{5.75}$$

Эти векторы порождают касательную плоскость, если $\det M \neq$ $\neq 0$. Если же $\det M = 0$, то порождаемое множество может быть определено путем замены $(t-1)\times k$-матрицы в (5.68′) совершенно другим способом. Уравнения (5.75) эквивалентны множеству уравнений

$$\frac{\partial c_\alpha}{\partial s_j} = (-)^{\alpha-j} \det(\hat{M}_{j\alpha}), \quad 1 \leqslant \alpha \leqslant k, \quad t \leqslant j \leqslant k, \tag{5.76}$$

для сепаратрисы в $\mathbb{R}^k$.

Пример 1. Катастрофа складки A_2 не имеет множества Максвелла, так как изолированные точки, если они существуют, имеют различные критические значения. Если условится считать точку $a = 0$ множеством Максвелла, поскольку критические значения становятся вырожденными в вырожденной критической точке, то множество Максвелла для высших катастроф будет включать это бифуркационное множество в качестве своего подмножества. ***Пример 2.*** Точка $a < 0$, $b = 0$ лежит в множестве Максвелла катастрофы сборки

$$V(x;\ a,\ b) = \frac{1}{4}x^4 + \frac{1}{2}ax^2 + bx. \tag{5.77}$$

В данной точке

$$x^{(1)} = -\sqrt{-a}, \qquad x^{(2)} = +\sqrt{-a}. \tag{5.78}$$

В этих двух точках $\widetilde{V}^{(1)} = V^{(2)} = -a^2/4$, $(a,\ b)$ — это то же, что и $(c_1,\ c_2)$ выше, так что

$$\begin{aligned} V_1^{(1)} &= -\frac{a}{2}, & V_2^{(1)} &= -\sqrt{-a}, \\ V_1^{(2)} &= -\frac{a}{2}, & V_2^{(2)} &= +\sqrt{-a}. \end{aligned} \tag{5.79}$$

В этом случае уравнения Клаузиуса — Клапейрона (5.68) сводятся к

$$(0,\ -2\sqrt{-a})\begin{pmatrix} \delta c^1 \\ \delta c^2 \end{pmatrix} = 0, \tag{5.80}$$

а решение имеет вид

$$B_2 \simeq \begin{pmatrix} -2\sqrt{-a} \\ 0 \end{pmatrix} \delta s_2 \tag{5.81}$$

или δa произвольно, $\delta b = 0$.

Подобный путь решения оказывается, с одной стороны, довольно сложным, когда речь идет об определении, состоит ли множество Максвелла $\mathscr{P}_M$ для A_3 из луча $a \leqslant 0$, $b = 0$, а с другой стороны, является достаточно простым, если речь идет об определении $\mathscr{P}_M$ для высших катастроф.

◇◇◇ С уравнениями Клаузиуса — Клапейрона (5.68′) легко работать в частном случае, когда $k = t$; система уравнений в частных производных (5.76) сводится к системе обыкновенных дифференциальных уравнений.

5. ВЫВОДЫ

В данной главе была показана практическая эффективность «принципа лома», с помощью которого можно получить много полезной информации, рассматривая точки с «плохим» поведением.

Для функции, у которой большинство точек $x \in \mathbb{R}^n$ не являются критическими, найдено, что именно критические точки организуют глобальную качественную топографию.

Для случая семейств функций, где большинство точек $c \in \mathbb{R}^k$ параметризует морсовские функции, установлено, что именно сепаратриса, параметризующая неморсовские функции, организует качественные изменения семейства. Внутри открытой области $\mathbb{R}^k$ на некотором расстоянии от сепаратрисы малые изменения в управляющих параметрах вызывают незначительные изменения положения критических точек, их главных направлений и соответствующих сил (собственных значений); эти возмущения не вызывают качественных изменений в поведении функций, параметризуемых рассматриваемой областью $\mathbb{R}^k$.

В случае неморсовской критической точки $(x^0; c^0)$ центральной точкой для «лома» является соответствующий росток катастрофы. Возмущение функции в неморсовской критической точке не преподносит качественных сюрпризов до тех пор, пока рассматривается морсовская часть канонической формы. Качественные изменения возникают только при возмущении ростка катастрофы. Поскольку ростки и их возмущения являются каноническими, изучение сепаратрисы достаточно провести лишь один раз.

Построены сепаратрисы пространства управляющих параметров для элементарных катастроф типа A_2, A_3, A_4, D_{+4} и D_{-4}. Сепаратриса в $\mathbb{R}^k$ состоит из многообразий размерностей $k-1$, $k-2$, ... В каждом случае (и для остальных элементарных катастроф также) сепаратриса разбивает пространство $\mathbb{R}^k$ на некоторое число открытых областей. Каждая открытая область является связной и параметризует функции одного и того же качественного типа. Открытые области всюду плотны в $\mathbb{R}^k$, в том смысле, что их замыкание совпадает со всем пространством. Это значит, что любая функция из семейства катастроф (независимо от того, имеет ли она вырожденные критические точки или нет) может быть аппроксимирована с любой наперед заданной точностью морсовскими функциями.

Проведено фактическое сведение функции, параметризуемой точкой сепаратрисы, к канонической форме (2.3б). В случае ростка $D_{\pm 4}$ очевидно присутствие катастроф складки, так как $D_{\pm 4}$ содержит члены третьей степени, однако совершенно не ясно, каким образом появляются катастрофы сборки (каким образом мы можем получить квартики из кубиков). Но фактическое сведение к канонической форме может потребовать и нелинейной замены переменных. Это и рассеивает наше недоумение.

Ради завершенности изложения была рассмотрена сепаратриса в $\mathbb{R}^k$ другого типа — так называемое множество Максвелла. Его точки параметризуют функции с вырожденными критическими значениями. Бифуркационное множество $\mathscr{S}_B$ и множество Максвелла $\mathscr{S}_M$ определяются соответственно посредством следующих уравнений бифуркации и уравнений Клаузиуса —

Клапейрона:

$$\mathscr{P}_E: \quad \nabla V = 0, \quad \det V_{ij} = 0,$$

$$\mathscr{P}_M: \quad \nabla V = 0, \quad V^{(p)} - V^{(p+1)} = 0, \quad \left[V_\alpha^{(p)} - V_\alpha^{(p+1)}\right] \delta c^\alpha = 0.$$

В приложениях важность сепаратрисы зависит от принимаемых допущений: если применяется принцип максимального промедления, то важно $\mathscr{P}_E$, если же принцип Максвелла, то $\mathscr{P}_M$.

В заключение отметим, что преобразования с изменением масштаба явились чрезвычайно полезным средством при построении сепаратрис и изучении их свойств.

6

ГЕОМЕТРИЯ СКЛАДКИ И СБОРКИ

Данная глава посвящена изучению геометрии катастроф типа A_2 и A_3. Поскольку ростки рассматриваемых катастроф и их возмущения являются каноническими, то и геометрические свойства катастроф будут также каноническими. Это в свою очередь означает, что нет необходимости проводить изучение подобных свойств катастрофы каждого типа, вполне достаточно выполнить такое исследование применительно к катастрофе какого-то одного типа.

1. СКЛАДКА

Катастрофа A_2 задается следующим семейством функций, зависящих от одного управляющего параметра a:

$$A_2: \quad F(x; a) = \frac{1}{3}x^3 + ax. \tag{6.1а}$$

Функция, принадлежащая семейству $F(x; a)$, при $a < 0$ имеет две критические точки, при $a = 0$ — одну дважды вырожденную критическую точку, а при $a > 0$ не имеет ни одной критической точки (рис. 6.1). Бифуркационное множество состоит из единственной точки $x = 0$. Положение критических точек опре-

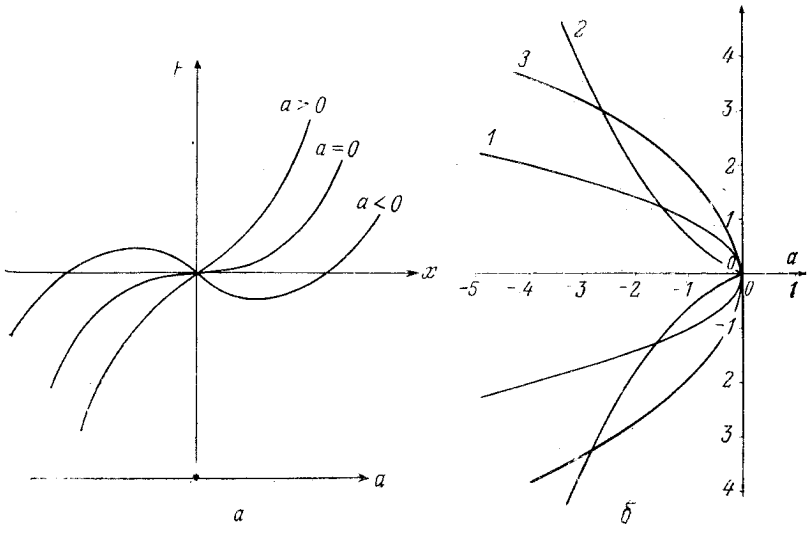


Рис. 6.1.

а — функция семейства $F(x; a)$ при $a > 0$, $a = 0$, $a < 0$; *б* — (*1*) положение критических точек $F(x; a)$ как функция a; (*2*) критические значения функции $F(x; a)$ в зависимости от величины a; (*3*) критическая кривизна функции $F(x; a)$ в зависимости от величины a.

деляется решением квадратного уравнения вида

$$\nabla F = x^2 + a = 0. \tag{6.1б}$$

Критическое многообразие, определяемое уравнением (6.1б), изображено на рис. 6.1.

Значения функции $F(x; a)$ в критических точках (рис. 6.1) определяются из формул (6.1а) и (6.1б):

$$\begin{aligned} &\text{в точке } x = -\sqrt{-a} \qquad F(x; a) = \tfrac{2}{3}|a|^{3/2}, \\ &\text{в точке } x = +\sqrt{-a} \qquad F(x; a) = -\tfrac{2}{3}|a|^{3/2}. \end{aligned} \tag{6.1в}$$

Собственные значения матрицы устойчивости (т. е. квадратной матрицы d^2F/dx^2 порядка 1) в этих точках равны

$$\begin{aligned} &\text{в точке } x = -\sqrt{-a} \qquad d^2F/dx^2 = -2|a|^{1/2}, \\ &\text{в точке } x = +\sqrt{-a} \qquad d^2F/dx^2 = +2|a|^{1/2}. \end{aligned} \tag{6.1г}$$

Соотношения (6.1г) изображены на рис. 6.1, *г*.

2. СБОРКА

Катастрофа типа A_3 задается следующим семейством функций, зависящих от двух управляющих параметров a и b:

$$A_{+3}: \quad F(x; a, b) = \tfrac{1}{4}x^4 + \tfrac{1}{2}ax^2 + bx. \tag{6.2а}$$

Любая точка пространства параметров $(a, b) \in \mathbb{R}^2$, за исключением точек сепаратрисы (рис. 5.4), параметризует функции с одной или тремя изолированными критическими точками. График функции (6.2а) изображен на рис. 6.2 при различных значениях управляющих параметров (a, b): внутри области, имеющей форму сборки, $F(x; a, b)$ имеет три изолированные критические точки, а вне этой области — всего одну; на границе функция семейства имеет изолированную критическую точку и дважды вырожденную критическую точку, а в начале координат — трижды вырожденную критическую точку. Положение критических точек находится путем решения кубического уравнения вида

$$\nabla F = x^3 + ax + b = 0. \tag{6.2б}$$

Уравнение (6.2б) задает двумерное многообразие, расположенное в трехмерном пространстве с координатными осями x — a — b (рис. 6.2, *б*).

Критические значения функции F определяются путем решения кубического уравнения (6.2б) для критических точек

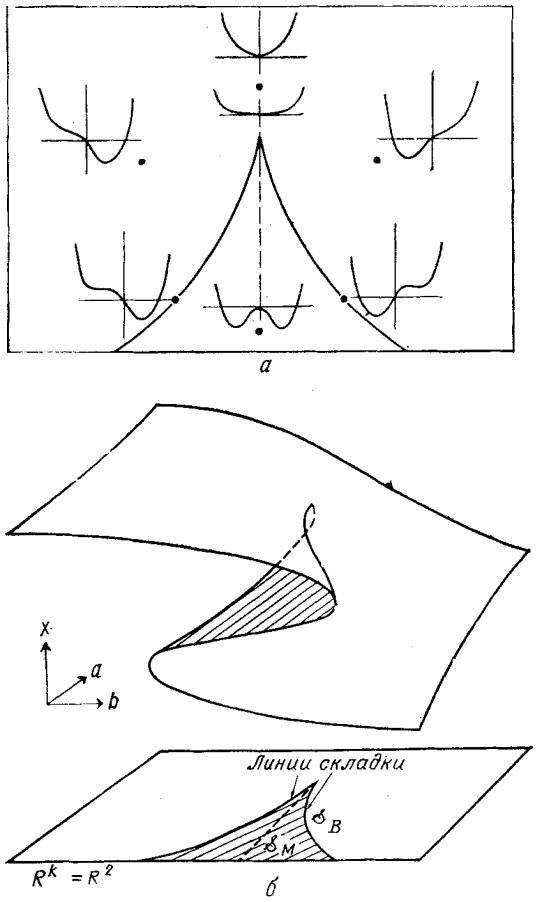
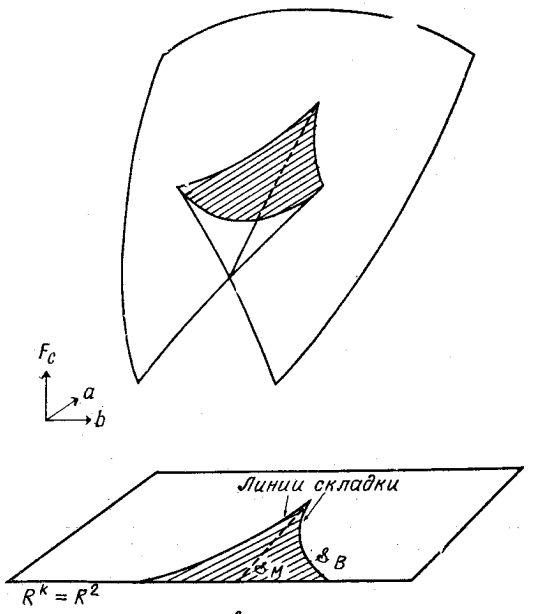
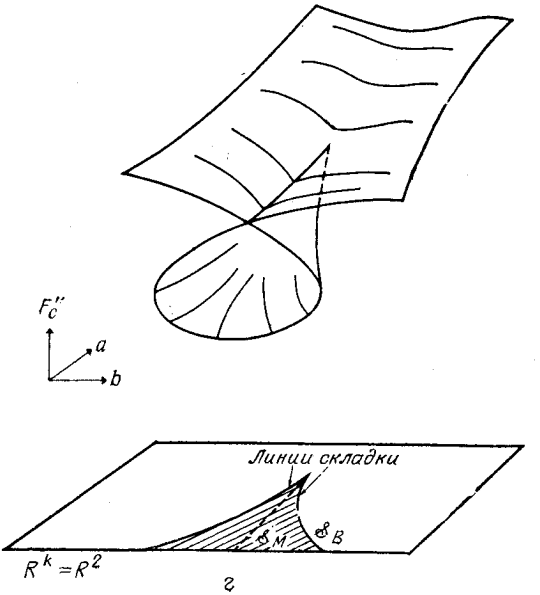


Рис. 6.2.

а — функции семейства $F(x; a, b)$ для различных значений управляющих параметров (a, b); *б* — двумерное многообразие катастрофы сборки, решение уравнения $\nabla F = 0$ изображено вложенным в пространство $R^3 = R^1 \otimes R^2$. Проекция этого многообразия вниз на плоскость управляющих параметров R^2 представляет линию складки. Искаженная линия складки $(a/3)^3 + (b/2)^2 = 0$ в плоскости управляющих параметров является тенью надскладочной части многообразия. Штриховая линия $a < 0$, $b = 0$ в плоскости управляющих параметров представляет множество Максвелла или нелокальную сепаратрису и служит границей раздела между функциями, имеющими более глубокий левосторонний минимум, и функциями, имеющими более глубокий правосторонний минимум; *в* — изображение поверхности критических значений в плоскости управляющих параметров; *г* — изображение поверхности критической кривизны.



в



г

и оценивания (6.2а) в этих критических точках. Получаемое «критическое множество» (не являющееся многообразием) изображено на рис. 6.2, *в*. Проекция этого критического множества на плоскость управляющих параметров позволяет сделать вывод о свойствах локальных и нелокальных сепаратрис, описанных выше. Собственные значения матрицы устойчивости F находятся путем вычисления $d^2F/dx^2 = 3x^2 + a$ в критических точках [уравнения (6.2в)]. Получаемое множество изображено на рис. 6.2, *г*.

Заметим, что геометрия сборки (рис. 6.2, *а — г*) полностью аналогична геометрии складки (рис. 6.1, *а — г*). Однако если кривые, изображенные на рис. 6.1, *б — г*, являются геометрическим представлением точных выражений (6.1б—6.1в), то соответствующие им кривые на рис. 6.2, *б — г* являются по своей природе качественными, т. е. в последнем случае мы располагаем лишь способом построения изображенных поверхностей. Вместе с тем ощущается острая необходимость в количественных выражениях для поверхностей, изображенных на рис. 6.2, *б — г*. Оказывается это можно сделать, используя пересчетные соотношения. Основываясь на наблюдениях, сделанных в гл. 5, можно отметить следующее преобразование подобия:

$$x \to \lambda x, \quad a \to \lambda^2 a, \quad b \to \lambda^3 b, \quad F \to \lambda^4 F. \tag{6.3}$$

Это позволяет сразу же получить количественные выражения для пересечений плоскостей $a = +1$, $a = 0$, $a = -1$ с тремя поверхностями, изображенными на рис. 6.2, *б — г*. Если количественные выражения для каждой плоскости найдены, то численные значения любой точки на этой поверхности могут быть определены из соотношений (6.3). Продемонстрируем, как практически реализуется данный метод.

В плоскостях $a = +1$, $a = 0$, $a = -1$ критические точки задаются как функции одного параметра b:

$$x^3 + x + b = 0, \qquad (6.4,\ a = +1)$$

$$x^3 + b = 0, \qquad (6.4,\ a = 0)$$

$$x^3 - x + b = 0. \qquad (6.4,\ a = -1)$$

Чтобы получить выражение x через b, нет необходимости решать уравнения (6.4); достаточно изобразить их графически, представив b как функцию x. Если затем эти графические представления собрать воедино, то в результате получим не что иное, как аналог рис. 6.2б, причем оси x, a и b будут иметь масштаб, указанный в преобразованиях (6.3).

Пример 1. Найти критические точки, если $(a, b) = (-3, 1)$.

Решение. Пусть x_0, b_0 — точки на плоскости $a = -1$, которые мы будем пересчитывать до соответствующих точек на плоскости $a = -3$. Тогда имеем

$$\begin{gathered} -3 = \lambda^2 (-1) \Rightarrow |\lambda| = \sqrt{3}, \\ 1 = \lambda^3 b_0 \Rightarrow b_0 = 3^{-3/2} = 0{,}19. \end{gathered} \tag{6.5}$$

Для $b_0 = 0{,}19$ три критических значения x могут быть взяты из рис. 6.3, а именно $x_{10} = 0{,}88$, $x_{20} = 0{,}20$, $x_{30} = -1{,}08$.

Соотношения пересчета дают

$$\begin{gathered} x_1 = \lambda x_{10} = 1{,}53, \quad x_2 = \lambda x_{20} = 0{,}35, \\ x_3 = \lambda x_{30} = -1{,}88. \end{gathered} \tag{6.6}$$

Эти критические значения могут быть проверены посредством подстановки их в уравнение (6.2б).

В плоскостях $a = +1$, $a = 0$, $a = -1$ критические значения b могут быть определены из соотношений (6.4) и рис. 6.3. В частности, для любого значения b критические точки определяются из рис. 6.3, а значения в этих точках определяются из (6.2а) при $a = 1, 0, -1$ (рис. 6.4).

Пример 2. Найти критические значения, если $(a, b) = (-3, 1)$.

Решение. При $(a, b) = (-1, 0{,}19)$ критические значения равны 0,02, —0,07 и —0,45. Так как критические значения пересчитываются в соответствии с $\lambda^4 = (\sqrt{3})^4$, то критические значения при $(a, b) = (-3, 1)$ будут равны 0,18, —0,63 и —4,05; они могут быть также вычислены непосредственно из соотношений (6.6).

Аналогично может быть определена кривизна функции в критической точке. Критические точки находим из рис. 6.3, а кривизну в любой из них определяем с помощью формулы $d^2F/dx^2 = 3x^2 + a$. Результаты вычислений в плоскостях $a = +1$, $a = 0$, $a = -1$ изображены на рис. 6.5.

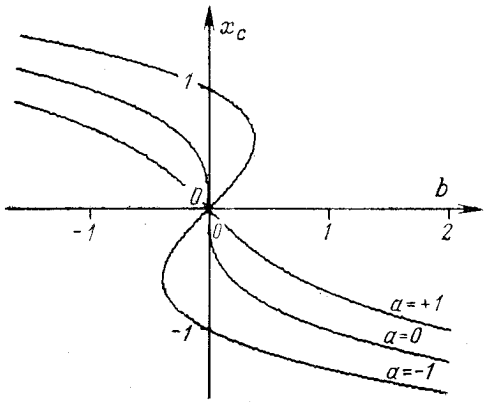


Рис. 6.3. Критические точки F как функция параметра b.

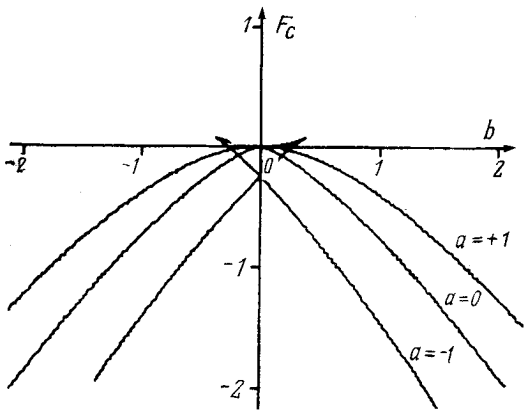


Рис. 6.4. Критические значения F как функция параметра b.

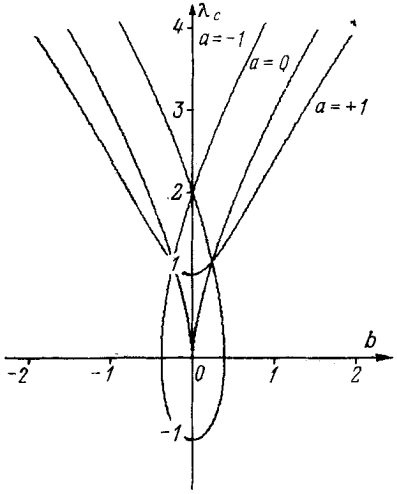


Рис. 6.5. Критическая кривизна F как функция параметра b.

◇◇◇ Заметим, что на основе кривых, приведенных на рис. 6.3—6.5, и пересчетных соотношений (6.3) могут быть построены поверхности, изображенные на рис. 6.2.

Пример 3. Найти кривизну в критических точках функций F, задаваемых точками прямолинейной траектории (рис. 6.6, *а*):

$$b(t) = -b_0 + b_1 t, \quad a(t) = -a_1 t \tag{6.7}$$

в пространстве управляющих параметров.

Решение. Для каждого $t > 0$ можно определить пересчетный множитель $\lambda(t) = (a_1 t)^{1/2}$, а затем вычислить собственные значения функции в соответ-

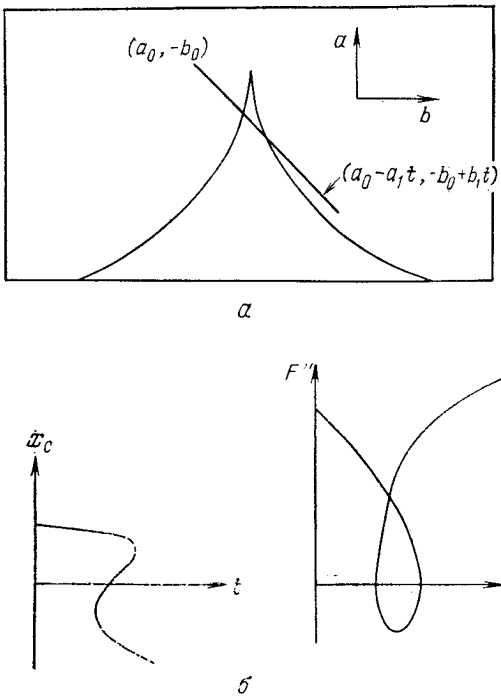


Рис. 6.6. Определение критической кривизны в пространстве управляющих параметров для прямолинейной траектории в пространстве управляющих параметров с помощью соотношений (6.3) и кривизны в соответствующих плоскостях $a = +1$, $a = 0$ или $a = -1$.

ствующей точке $b_0 = b(t)/\lambda^3(t)$ на плоскости $a = -1$. Эти собственные значения пересчитываются с учетом множителя $\lambda^2(t)$. Процесс может быть повторен для $t = 0$, $t < 0$. Собственные значения, вычисленные этим способом, изображены на рис. 6.6, *б* для $b_0 = 1$, $a_1 = 2$, $b_1 = 3$.

3. ДВОЙСТВЕННАЯ СБОРКА

Катастрофы типа A_{-3} задаются следующим семейством функций, зависящих от двух параметров a и b:

$$A_{-3}: \quad F(x; a', b') = -\frac{1}{4}x^4 - \frac{1}{2}a'x^2 - b'x. \qquad (6.8а)$$

Заметим, что при $(a', b') = (a, b)$ формула (6.8а) отличается от (6.2а) лишь знаком и, следовательно, (6.8б) полностью совпадает с (6.2б). Критические точки сборки и двойственной сборки определяются одним и тем же уравнением, поэтому геометрическое представление рис. 6.2, *б* верно в обоих случаях. Так как семейство функций (6.8а) — это в точности взятое со знаком

минус семейство (6.2а), то геометрическое представление двойной сборки получают отражением геометрии рис. 6.2, *а*, *в* и *г*.

Хотя сборка и двойственная сборка математически связаны простым соотношением, их физические приложения существенно различны. У сборки всегда имеется по крайней мере один глобальный минимум, и когда управляющие параметры изменяются, исчезает один минимум, но возникает другой.

При переходе от сборки к двойственной сборке максимумы и минимумы меняются ролями. Двойственная сборка имеет локальный минимум, когда значения управляющих параметров лежат внутри области пространства управляющих параметров, имеющей форму сборки, и не имеет глобального минимума. Если управляющие параметры перемещаются за линией складки, то система, описываемая такой потенциальной функцией, вообще не имеет минимума. Результатом будет катастрофа.

4. ВЫВОДЫ

В данной главе были рассмотрены геометрические свойства катастроф складки и сборки и, в частности, качественная природа функций, параметризуемых различными областями пространства управляющих параметров. Было дано количественное описание представляющих особый интерес геометрических свойств, таких, как положение критических точек и кривизна функции в этих точках. Поскольку эти свойства являются каноническими, они являются инвариантами всех складок и сборок, в частности складок и сборок, возникающих при катастрофах с более высокой размерностью пространства управляющих параметров.

7

ОРГАНИЗАЦИЯ КАТАСТРОФ

В данной главе описываются вычислительные методы, позволяющие найти уравнения одномерных кривых ростков катастроф, имеющих размерность k.

Одновременно рассматривается другой возможный метод определения структуры сепаратрис в пространстве $\mathbb{R}^k$ управляющих параметров, который может быть использован при нахождении компонент сепаратрис размерности больше 1.

Этот метод позволяет определить как полный «спектр» (набор) компонент сепаратрисы, так и размерность каждой из них. Кроме того, он дает возможность легко установить тип структурно устойчивых морсовских функций, которые могут быть встречены в возмущениях любого ростка катастрофы, и выяснить, имеют ли открытые множества, представляющие эти функции, общие компоненты сепаратрисы.

Приводится также схема наглядного представления свойств катастроф, «двойственная» по отношению к диаграммному представлению, которая оказывается очень полезной при контурном представлении катастрофы.

Обсуждается вопрос о том, какие из катастроф, имеющих размерность k, связаны («примыкают») с катастрофами, размерность которых на единицу меньше. Высшие катастрофы организуют низшие катастрофы, и иерархия связанных катастроф организуется вполне определенным способом. Это взгляд «сверху вниз». Вместе с тем возможен и взгляд «снизу вверх»: присутствие в соседстве двух или более конкретных катастроф, скажем A_{+3} и A_{-3}, с определенностью предполагает, что они организованы некоторой высшей катастрофой, которой является A_4 (рис. 7.1). Последнее обстоятельство побуждает нас искать (и находить) ростки этих высших катастроф, что в свою очередь приводит к нетривиальным соотношениям между исходными катастрофами сборки и двойственной сборки A_{+3} и A_{-3}.

1. $A_{k+1} \to A_k$

Из начала координат пространства $\mathbb{R}^k$ управляющих параметров ростка катастрофы $A_{k+1} = x^{k+2}$ исходят две линии, параметризующие ростки катастроф $A_k = \tilde{x}^{k+1}$. Наиболее легкий путь определения уравнений этих линий в пространстве управляющих параметров — это анализ размерностей и, в частности, использование пересчетных соображений. Вначале продемонстрируем эффективность применения подобных соображений в хорошо знакомом случае: линий складки, исходящие из начала координат $(0, 0) \in \mathbb{R}^2$ пространства управляющих параметров сборки.

В окрестности ростка катастрофы сборки общая функция имеет вид $F(x; a, b) = \frac{1}{4}x^4 + \frac{1}{2}ax^2 + bx$.

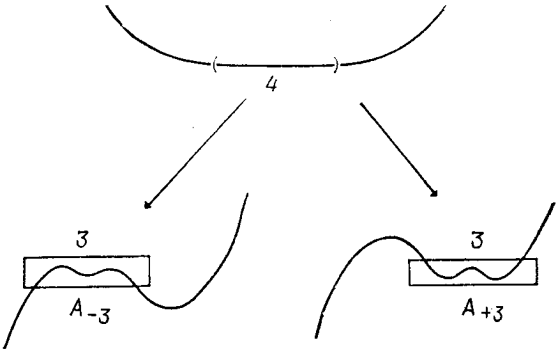


Рис. 7.1. Близкое соседство катастроф типа A_{+3} и A_{-3} позволяет сделать вывод о том, что они определенным образом организуются катастрофой, которая имеет большую размерность пространства управляющих параметров и которая должна быть по крайней мере A_{+4}.

Если предположить, что в точке $x = -\lambda$ уравнение $dF/dx = 0$ имеет два корня (в данном случае третий корень нас не интересует), то

$$A_3 \to A_2,$$

$$\frac{1}{4}x^4 + \frac{1}{2}ax^2 + bx \to \frac{1}{4}(x+\lambda)^3(x+\alpha) + \text{Константа},$$

$$\frac{1}{4}x^4 + \frac{1}{4}(3\lambda+\alpha)x^3 + \frac{1}{4}(3\lambda^2+3\lambda\alpha)x^2 + \tag{7.1}$$

$$+\frac{1}{4}(\lambda^3+3\lambda^2\alpha)x + \text{Константа}.$$

Следовательно, в окрестности точки $x = -\lambda$ функция ведет себя подобно $(\alpha-\lambda)\tilde{x}^3$, $\tilde{x} = x+\lambda$. Сравнивая коэффициенты при x^j, получаем следующее параметрическое представление управляющих параметров вдоль линии складки:

$$\begin{aligned} a &= -3\lambda^2 \\ b &= -2\lambda^3 \end{aligned} \Rightarrow \left(\frac{a}{3}\right)^3 + \left(\frac{b}{2}\right)^2 = 0. \tag{7.2}$$

Общая процедура нахождения 1-мерных кривых $A_{k+1} \to A_k$ совершенно аналогична, т. е.

$$A_{k+1} \to A_k,$$

$$\frac{1}{k+2}x^{k+2} + \sum_{j=1}^{k}\frac{1}{j}a_j x^j = \frac{1}{k+2}(x+\lambda)^{k+1}(x+\alpha) + \text{Константа}. \tag{7.3}$$

Сравнение коэффициентов при различных степенях x дает

$$\frac{k+2}{j}a_j=\binom{k+1}{j-1}\lambda^{k+1-(j-1)}+\binom{k+1}{j}\lambda^{k+1-j}\alpha. \qquad (7.4)$$

Условие $a_{k+1}=0$ определяет взаимосвязь между k-кратным вырожденным корнем в точке $x=-\lambda$ и изолированным корнем в $x=-((k+1)\alpha+\lambda)/(k+2)=k\lambda$:

$$a_{k+1}=0=(k+1)\lambda+a\Rightarrow\alpha=-(k+1)\lambda. \qquad (7.5)$$

Линии катастроф A_k, исходящие из ростка x^{k+2} катастрофы A_{k+1}, имеют следующее параметрическое представление:

$$a_j=(-)\binom{k+1}{j-1}(k+1-j)\lambda^{k-j+2}, \quad j=1, 2, \ldots, k. \qquad (7.6)$$

Если исключить параметр λ, то получим уравнения, содержащие лишь управляющие параметры, т. е.

$$\lambda=\left[\frac{-a_j}{\binom{k+1}{j-1}(k+1-j)}\right]^{1/(k-j+2)}. \qquad (7.7)$$

Выражение (7.7) сводится к формуле (7.2) при $k=2$.

Линии катастроф A_{k-1} исходят из каждого ростка A_k точно так же, как и линии A_k исходят из каждого ростка A_{k+1}. Следовательно, двумерные поверхности A_{k-1} исходят из каждого ростка A_{k+1}. Рассуждая подобным образом, можно построить всю сепаратрису ростка A_{k+1} в пространстве $\mathbb{R}^k$.

2. $D_{k+1}\to D_k$

Изложенные выше соображения верны и в случае ростков катастроф типа D_k и могут быть использованы для нахождения уравнений линии катастрофы. В частности, росток

$$D_{k+1}\colon\ x^2y+y^k \qquad (7.8)$$

представляется началом координат пространства $\mathbb{R}^k$. Он имеет $(k+1)$-кратную вырожденную критическую точку в начале координат $(0, 0)\in\mathbb{R}^2$ пространства переменных состояния. Одномерные кривые исходят из начала координат пространства управляющих параметров и параметризуют ростки катастроф, имеющих k-кратные вырожденные критические точки где-то в $\mathbb{R}^2$. Эти вырожденные критические точки могут иметь либо одно «плохое» направление, и тогда соответствующим ростком является $A_k(\tilde{x}^{k+1})$, либо два «плохих» направления, и тогда ростком является D_k.

Линии ростков D_k, исходящие из D_{k+1}, получают на основе пересчетных соображений (разд. 1). Предположим, что k-квадратная вырожденная критическая точка находится в $(x, y) = (-\lambda_1, -\lambda_2)$. Тогда можно приравнять общее возмущение ростка D_{k+1} (слева) ростку D_k в окрестности точки $(-\lambda_1, -\lambda_2)$ (справа):

$$x^2y + y^k + \sum_{j=1}^{k-2} a_j y^j + a_{k-1}x + a_k x^2 =$$
$$= (x + \lambda_1)^2 (y + \lambda_2) + (y + \lambda_2)^{k-1} (y + \alpha) + \text{Константа}. \quad (7.9)$$

Раскрывая скобки и сравнивая коэффициенты, находим, что

$$\begin{aligned} x&: \ a_{k-1} = 2\lambda_1\lambda_2, \\ x^2&: \ a_k \ = \lambda_2, \\ xy&: \ 0 = 2\lambda_1, \\ y^j&: \ a_j = \binom{k-1}{j-1} \lambda_2^{k-j} + \binom{k-1}{j} \lambda_2^{k-j-1}\alpha. \end{aligned} \quad (7.10)$$

Из анализа (7.10) очевидно, что $\lambda_1 = 0$. Поэтому величина $\lambda_2 = \lambda$ может быть использована в качестве пересчетного множителя. Поскольку в качестве универсального возмущающего члена использовался x^2, а не y^{k-1}, коэффициент при y^{k-1} должен равняться нулю. Это в свою очередь означает, что должно выполняться соотношение

$$0 = (k-1)\lambda + \alpha, \quad (7.11)$$

из которого легко определяется

$$a_j = -\binom{k}{j} (k - 1 - j) \lambda^{k-j}. \quad (7.12)$$

В результате получаем, что 1-параметрическая кривая в пространстве $\mathbb{R}^k$ управляющих параметров ростка D_{k+1}, координаты $(a_1, \ldots, a_k)$ которой представляют возмущение

$$a_1 y + a_2 y^2 + \ \ldots \ + a_{k-2} y^{k-2} + a_{k-1} x + a_k x^2 \quad (7.13)$$

с k-кратной вырожденной критической точкой типа D_k, имеет следующее параметрическое представление:

$$a_j = -\binom{k}{j} (k - 1 - j) \lambda^{k-j}, \quad j = 1, 2, \ldots, k-2,$$
$$a_{k-1} = 0, \quad a_k = \lambda. \quad (7.14)$$

Эта критическая точка имеет в пространстве $\mathbb{R}^2$ координаты $(x, y) = (0, -\lambda)$.

Таким образом, для катастрофы D_k исходящими из начала координат пространства управляющих параметров $\mathbb{R}^k$ могут быть одномерные кривые катастроф типа D_k и, кроме того, одномерные кривые катастроф типа A_k. Это верно в случае, когда $k+1$ нечетно, а также, когда рассматриваются катастрофы типа D_{-p}, где p четно. Однако это неверно для катастроф типа D_p, p четно. Поэтому, для того чтобы видеть, каким образом катастрофы меньшей размерности «извергаются» некоторой конкретной катастрофой высшей размерности, необходим какой-то простой вычислительный алгоритм.

3. ДИАГРАММНЫЕ ПРЕДСТАВЛЕНИЯ

Подобные представления катастроф позволяют выяснить, как выглядят и что разделяют сепаратрисы в пространстве управляющих параметров, какими будут неморсовские функции на каждой компоненте сепаратрисы и, наконец, какова качественная природа морсовских функций, параметризуемых точками любой из открытых областей пространства управляющих параметров при условии, что $k > 3$.

3.1. Катастрофы, имеющие единственную «плохую» переменную состояния ($l = 1$)

Рассмотрим сначала катастрофы типа A_k. При универсальной деформации k-кратная вырожденная критическая точка может расщепиться самое большее на k изолированных точек. Так как катастрофы типа A_k имеют всего лишь одну переменную состояния, то положение этих критических точек может быть отмечено кружочками вдоль оси x. Поскольку кривизна в локальном минимуме положительна, то поставим внутри кружочка

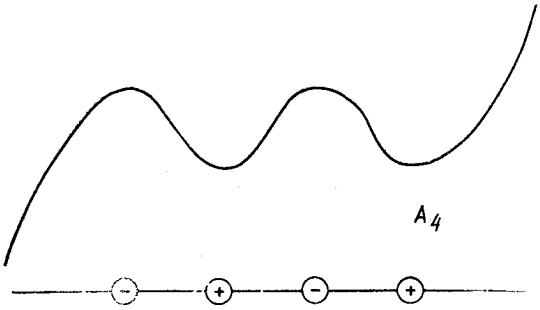


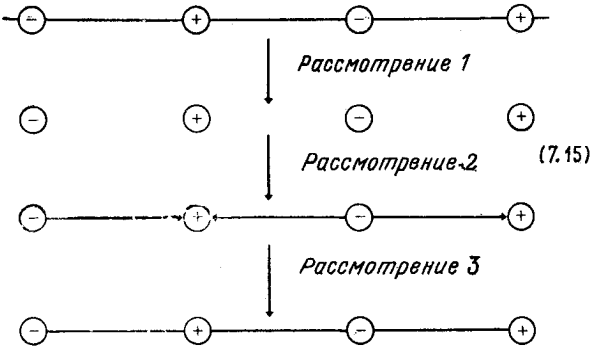
Рис. 7.2. Диаграммное представление катастрофы A_k, имеющей одну переменную состояния. Как видно, в случае $l = 0$ могут иметь место только максимумы $\ominus$ и минимумы $\oplus$.

знак $\oplus$; для изображения локальных максимумов используем символ $\ominus$ (рис. 7.2).

Прежде чем получить диаграммное представление катастрофы A_k (рис. 7.2), примем следующие соглашения:

1. На диаграммном представлении не указывается ось x.
2. С помощью стрелок показывают, какой из максимумов в какой из минимумов может «втекать».
3. Поскольку максимумы и минимумы выделены специальными символами, нет необходимости указывать направление потока.

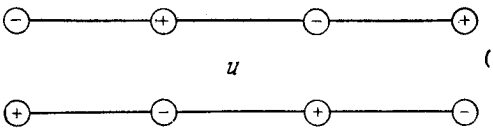
С учетом этих соглашений построены диаграммы



(7.15)

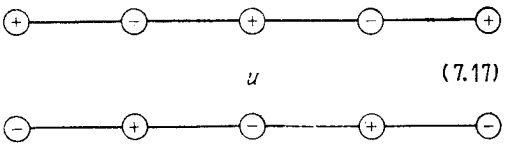
последняя из которых является диаграммным представлением катастрофы A_k ($k = 4$).

Для четного k диаграммы



(7.16)

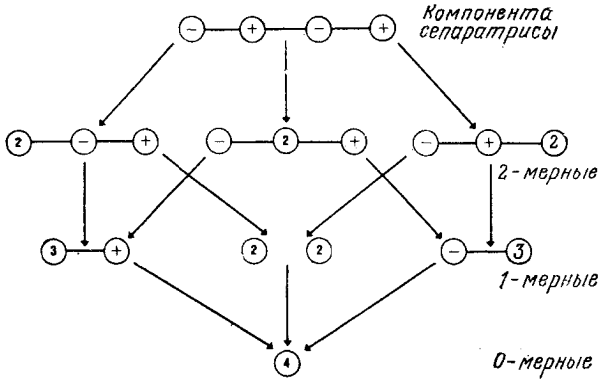
представляют качественно подобные функции (относительно замены $x \to -x$). Иначе говоря, одна диаграмма может быть преобразована в другую путем вращения в плоскости (x, y). Для нечетного k диаграммы



(7.17)

представляют качественно различные функции ($A_{\pm k} = \pm x^{k+1} + \ldots$).

Если теперь рассмотреть 1-параметрическую кривую в пространстве управляющих параметров, то при ее пересечении с некоторой $(k-1)$-мерной компонентой сепаратрисы типичным образом будут встречаться лишь неморсовские критические точки типа складки. В месте пересечения локальные максимум и минимум совпадают и уничтожают друг друга. Подобное может случиться $k-1$ различными способами, как показано ниже для A_4:



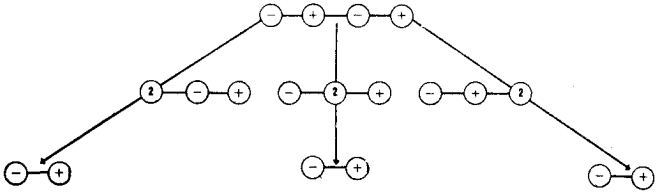
(7.18)

Если мы имеем двумерные поверхности в пространстве $\mathbb{R}^k$, то они будут пересекать $(k-1)$-мерные компоненты сепаратрисы по одномерным кривым складок, а $(k-2)$-мерные компоненты сепаратрисы в изолированных точках. Последние представляют катастрофы сборки либо пары катастроф складки. Именно эти возможности изображены для катастрофы A_4 (7.18).

Аналогичным образом можно определить все компоненты сепаратрисы и установить, каким образом они соединяются вместе. Из изображенной выше диаграммы для катастрофы A_4 можно видеть, что существует 0-мерная компонента (начало координат пространства $\mathbb{R}^3$), параметризующая функцию x^5, имеющую четырехкратно вырожденную критическую точку в начале координат $x=0$. Одномерные кривые катастрофы сборки и одна одномерная кривая, представляющая пару катастроф складки, исходят из начала координат пространства $\mathbb{R}^3$. Имеется в точности три двумерные компоненты сепаратрисы, представляющие катастрофы складки. Пересечение «внешних» поверхностей дает линию из пары складок, и это ясно видно как на диаграмме (7.18), так и на рис. 5.6.

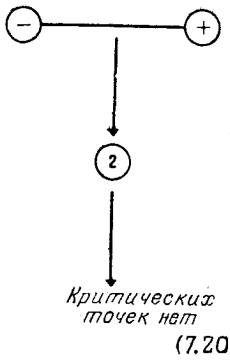
С помощью диаграммных методов могут быть определены открытые области в пространстве управляющих параметров, описывающие различные типы структурно устойчивых семейств

морсовских функций и разделенные одна от другой компонентами сепаратрисы. Рассмотрим в качестве примера катастрофу A_4. Когда 1-параметрическая кривая в пространстве управляющих параметров начинается в открытой области, содержащей максимальное количество изолированных критических точек, и проходит через некоторую двумерную компоненту сепаратрисы, то соответствующая катастрофа складки может включать совпадение и исчезновение двух критических точек. На диаграмме это можно представить следующим образом:



(7.19)

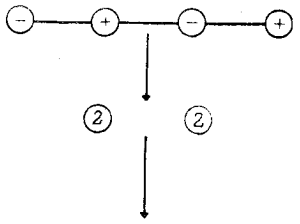
Следовательно, переход сквозь любую двумерную компоненту сепаратрисы ведет в открытую область пространства $\mathbb{R}^3$, описывающую функции с двумя изолированными критическими точками. Из такой области можно пройти через другую катастрофу складки в открытую область, описывающую функции, совсем не имеющие критических точек. Это может быть изображено следующей диаграммой:



(7.20)

В такую область можно попасть также и из области, содержащей четыре изолированные точки, с помощью перехода через кривую, представляющую пару катастроф складки, как это изображено на диаграмме (7.21).

Короче говоря, пространство $\mathbb{R}^3$ разбивается сепаратрисой на три открытые области, и эти области описывают устойчивые



Критических точек нет

(7.21)

семейства функций, имеющие четыре, две и нуль критических точек. Представление пространства управляющих параметров $\mathbb{R}^3$ для катастрофы A_4 дано на рис. 5.6, а в этом разделе оно синтезировано на рис. 7.3 и 7.4.

◇◇◇ Эту диаграммную схему для определения открытых областей пространства управляющих параметров $\mathbb{R}^k$ можно было бы назвать *методом стягивания*, так как при этом происходит «стягивание вместе смежных» критических точек в катастрофах складки.

3.2. Катастрофы, имеющие две «плохие» переменные состояния ($l = 2$)

С небольшими модификациями диаграммное представление может быть распространено на катастрофы, имеющие две «плохие» переменные. Перечислим основные правила построения диаграмм:

— изолированные критические точки распределяются в плоскости (x, y);

— локальные минимумы, седла и локальные максимумы изображаются при помощи следующих символов: ⊕ — локальный минимум, ○ — седло, ⊖ — локальный максимум;

— критические точки связываются посредством линии, если линия потока градиентной системы связывает их. Максимум может втекать в седло или минимум, а седло может втекать в некоторый минимум; седла не могут быть связаны [1]);

— вырожденность неморсовской критической точки представляется числом, т. е. 2 — дважды вырожденная, 3 — трижды вырожденная;

— катастрофы складки могут встречаться лишь при «стягивании» седла с одним из максимумов или минимумов;

— все возможные диаграммные представления катастрофы могут быть получены путем выбора возмущения, приводящего к максимальному числу изолированных критических точек;

[1]) Связанность седел является структурно неустойчивой.

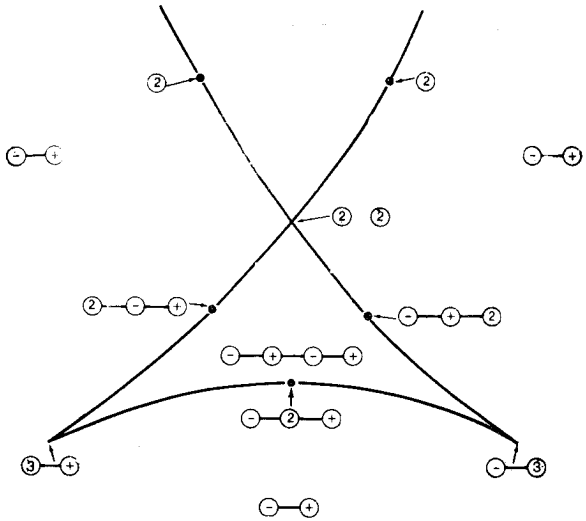


Рис. 7.3. Диаграммный метод позволяет осуществить организацию пространства управляющих параметров $\mathbb{R}^3$ катастрофы A_4 в открытые области, представляющие структурно устойчивые функции, и в компоненты сепаратрисы различных размерностей.

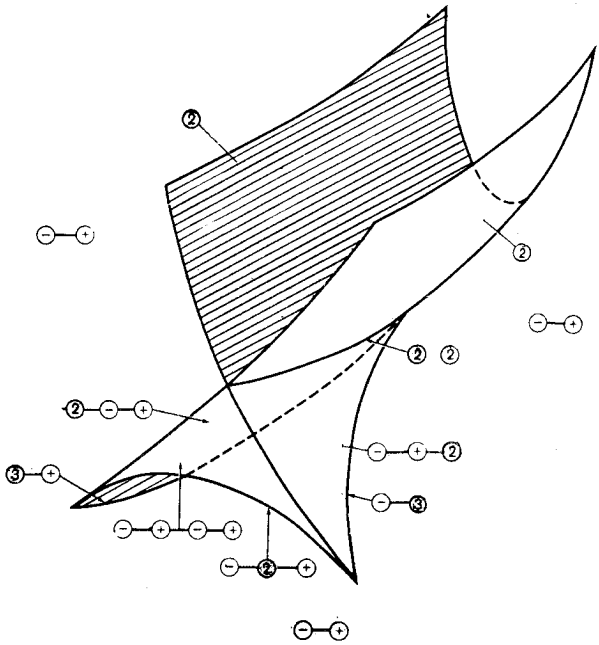
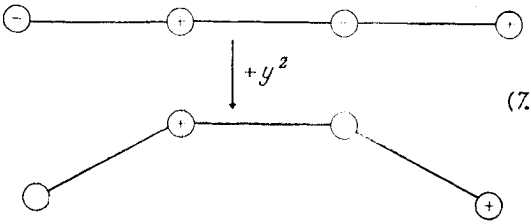


Рис. 7.4. Взаимосвязь между диаграммным методом и трехмерной организацией пространства для катастрофы A_{+4}.

— компоненты сепаратрисы и структурно устойчивые морсовские семейства, представляемые открытыми множествами, могут быть определены при помощи метода стягивания.

◇◇◇ Катастрофы типа A_k могут быть представлены как функции двух переменных состояния, к которым прибавлен квадратичный член от y: $F(x, y) = x^{k+1} +$ Возмущение $x \pm y^2$.

Если мы выберем сложение с $+y^2$, то одномерные минимумы на оси x становятся минимумами на плоскости $x - y$, в то время как локальные максимумы становятся седлами на плоскости $x - y$, так что имеет место диаграмма

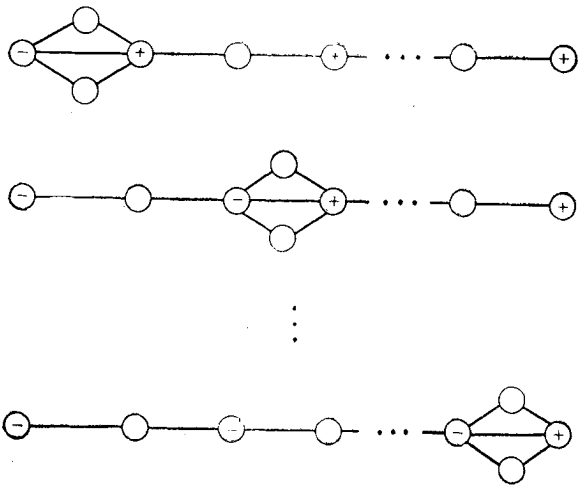


(7.22)

Одномерная цепь из k изолированных точек (чередующихся минимумов и седел $(A_k + y^2)$ или чередующихся максимумов и седел $-(A_k + y^2)$) представляет катастрофу типа A_k.

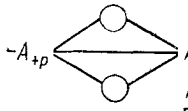
Для серии катастроф D_k было показано, что $D_k \neq D_{-k}$, если k четно, хотя $D_{-k} \doteq -D_k$, если k нечетно. Рассмотрим каждый из этих случаев:

D_{2k}:



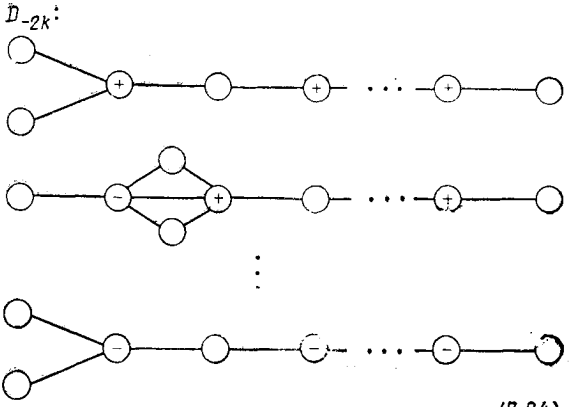
(7.23)

Общая диаграмма имеет структуру



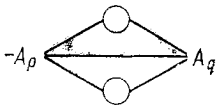
$p+q+2=2k$
p, q четные (7.23g)

D_{-2k}:



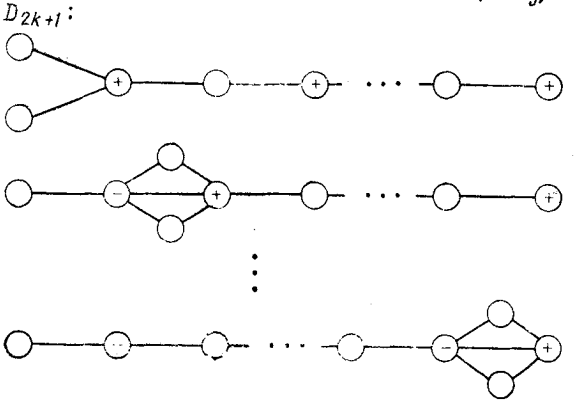
(7.24)

Общая диаграмма имеет структуру



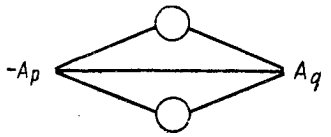
$p+q+2=2k$
p, q нечетные
(7.24g)

D_{2k+1}:



(7.25)

Общая диаграмма имеет структуру



$p+q+2=2k+1$
p нечетное
q четное

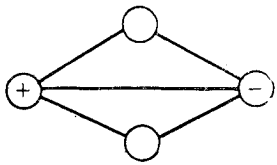
(7.25g)

Класс диаграмм с нечетным p и четным q представляет $-D_{2k+1} \doteq D_{-(2k+1)}$, как это может быть показано вращением общей диаграммы на 180°.

Диаграммные представления особых катастроф $E_{\pm 6}$, E_7, E_8 общеизвестны и перечислены в табл. 7.1, в которой также собраны диаграммные представления катастроф типа D_k и A_k. Последние представлены на основе их диаграмм в плоскости (x, y), следуя соглашению (7.22).

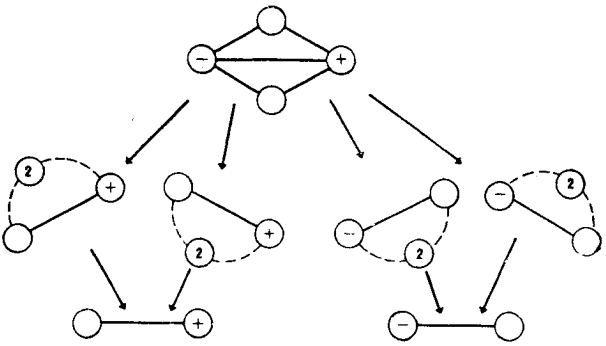
Используя информацию, содержащуюся в табл. 7.1, можно легко осуществить разбиение пространства управляющих параметров на открытые множества и определить компоненты сепаратрисы различных размерностей. Размерность компоненты сепаратрисы равна $k - \sum(\mu_i - 1)$ [суммирование производится по вырожденным критическим точкам, а μ_i есть вырожденность i-й вырожденной критической точки [ср. с (7.18)]. Открытые множества параметризуют классы структурно устойчивых функций, которые могут быть получены стягиванием. Два различных открытых множества являются смежными, т. е. разделяются компонентой сепаратрисы, если обе невырожденные диаграммы могут быть стянуты к вырожденной диаграмме, описывающей эту компоненту сепаратрисы.

***Пример 1*: D_{+4}.** Из табл. 7.1 следует, что открытая область $\mathbb{R}^3$, описывающая функции с максимальным числом изолированных критических точек, имеет следующее диаграммное представление:



(7.26)

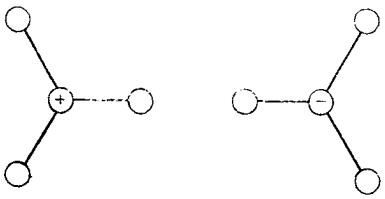
Представления других открытых областей могут быть получены посредством стягивания максимума и седла или минимума и седла:



(7.27)

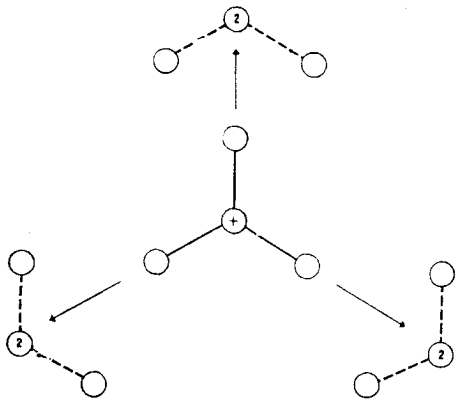
Аналогично могут быть стянуты две оставшиеся точки, давая открытое семейство функций, не имеющих критических точек. Кривая сборок A_{+3} может быть получена посредством стягивания двух седел и максимума. Кривые других сборок, например A_{-3}, получаются стягиванием двух седел и минимума. Две кривые в сепаратрисе, соответствующей компоненте ②②, представляют самопересечение. Эти элементарные рассмотрения приводят к разбиению пространства $\mathbb{R}^3$, топологически эквивалентному разбиению, изображенному на рис. 5.11 и полученному значительно более трудоемким методом.

Пример 2: D_{-4}. Из табл. 7.1 находим, что открытые области пространства $\mathbb{R}^3$, описывающие функции с максимальным числом изолированных точек, имеют следующие диаграммные представления:



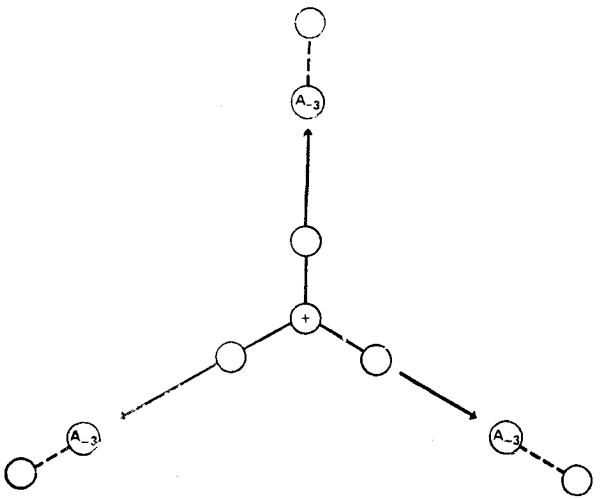
(7.28)

Вид этих диаграмм сразу же подсказывает, что должна быть симметрия вращения ранее описанного вида (гл. 5, разд. 7). 2-мерные компоненты сепаратрисы типа складки определяются посредством стягивания экстремума и седла:



(7.29)

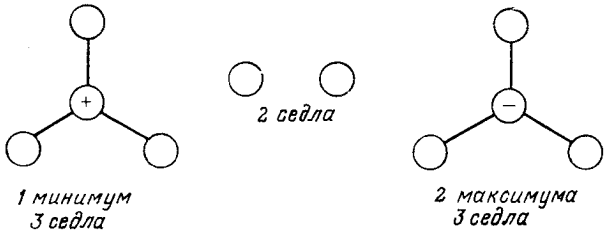
а 1-мерные компоненты сборки посредством стягивания трех критических точек:



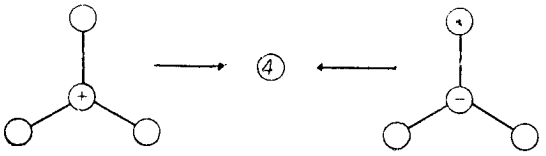
(7.30)

Открытые множества, представляющие структурно устойчивые функции и имеющие менее чем четыре критические точки, получаются продолжением стягивания складок [см. (7.25) и (7.30)].

Короче говоря, существуют три открытые области, характеризующие функции следующих трех типов:



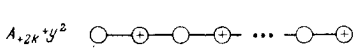
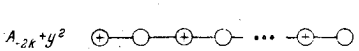
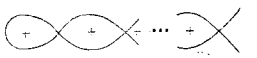
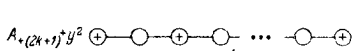
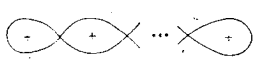
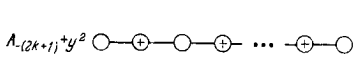
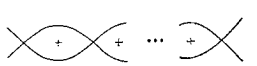
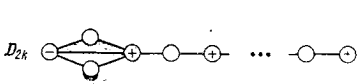
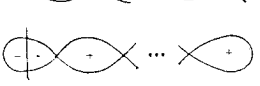
Открытая область, описывающая функции с двумя седлами, граничит с двумя другими открытыми областями через 2 (1 и 0)-мерные компоненты сепаратрисы. Однако два различных открытых множества, описывающие функции с четырьмя критическими точками, не являются примыкающими, за исключением начала координат, так как единственное стягивание, которое у них общее — это ④:



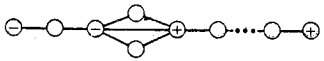
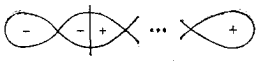
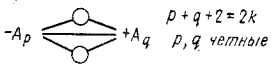
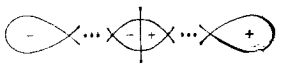
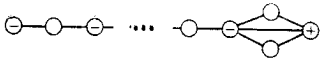
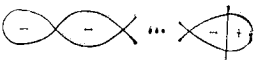
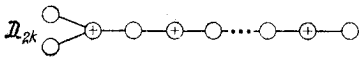
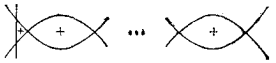
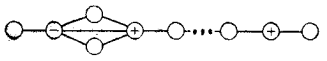
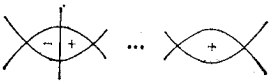
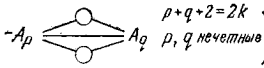
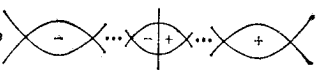
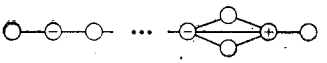
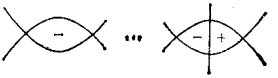
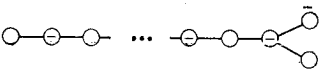
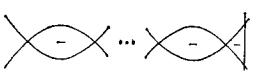
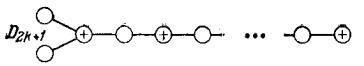
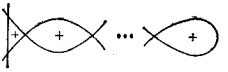
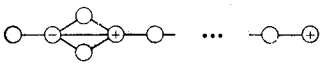
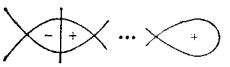
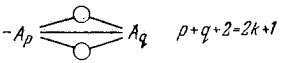
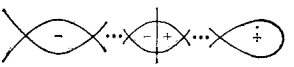
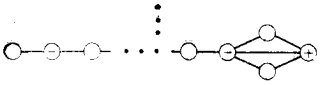
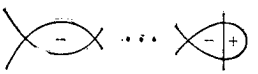
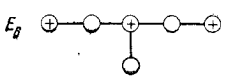
(7.31)

Проведенное рассмотрение непосредственно приводит к разбиению пространства $\mathbb{R}^3$ управляющих параметров катастрофы D_{-4}, топологически эквивалентному разбиению, изображенному на рис. 5.17.

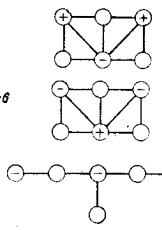
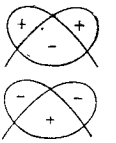
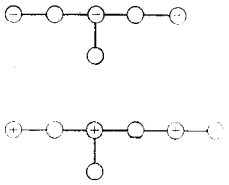
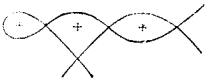
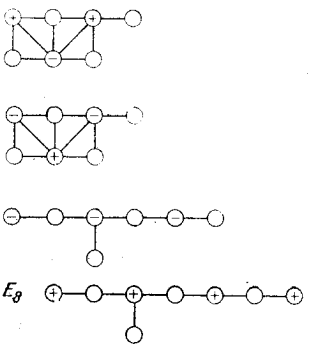
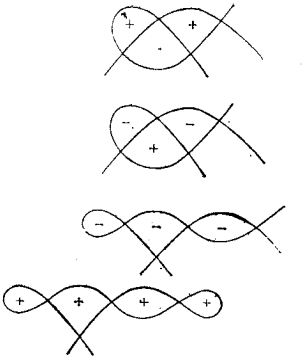
Таблица 7.1. **Диаграммное и контурное представления элементарных катастроф**

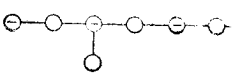
Тип катастрофы	Диаграммное представление	Контурное представление
$A_{+2k}+y^2$		
$A_{-2k}+y^2$		
$A_{+(2k+1)}+y^2$		
$A_{-(2k+1)}+y^2$		
D_{2k}		

Продолжение

Тип катастрофы	Диаграммное представление	Контурное представление
		
		
		
D_{2k}		
		
		
		
		
D_{2k+1}		
		
		
		
E_8		

Продолжение

Тип катастрофы	Диаграммное представление	Контурное представление
E_{-6}		
E_7		
E_8		

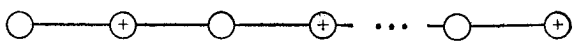


4. КОНТУРНЫЕ ПРЕДСТАВЛЕНИЯ

Одним из возможных графических представлений возмущаемых ростков катастроф является контурное представление. Для того чтобы получить такое представление, удобно выбрать начало координат и масштаб на прямой $\mathbb{R}^1(\mathbb{R}^n \otimes \mathbb{R}^k \to \mathbb{R}^1)$ и управляющие параметры так, что все максимумы имеют критическое значение1, все седла имеют критическое значение 0 и все минимумы имеют критическое значение —1. Это всегда можно сделать для катастроф, зависящих от одной или двух переменных состояния.

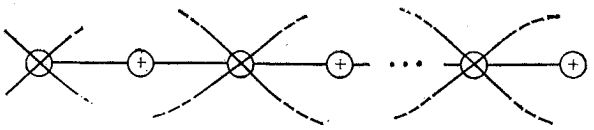
Такой выбор очень удобен, так как пересечение седла $f(x, y) = xy$ с плоскостью $z = 0$ имеет вид x. Контур нулевого уровня в этом случае дает удобный остов, вокруг которого могут быть построены остальные контуры. (Мы этого делать не будем.) Контурное представление может быть сконструировано непосредственно на основе диаграммного представления.

***Пример 1*: A_{2k}.** Вначале нарисуем диаграммное представление, которое в пространстве $\mathbb{R}^2$ имеет вид



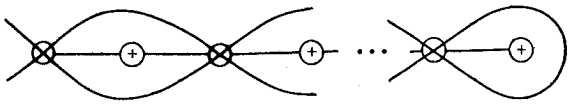
(7.32a)

Пересечение функций, представляемых диаграммой (7.32а), с плоскостью $z = 0$ имеет следующий вид:



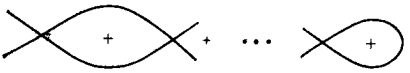
(7.32 б)

Локальные минимумы обязательно должны быть окружены контурами нулевого уровня, так что контуры могут быть соединены



(7.32 в)

Наконец, нет больше смысла оставлять как кружочки, представляющие локализацию критических точек, так и потоковые линии, связывающие «смежные» точки. Таким образом, контурное представление катастрофы A_{2k} упростится до



(7.32 г)

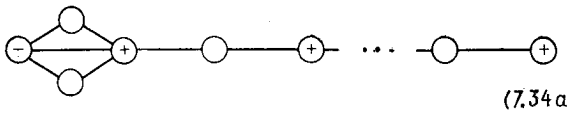
Связи седел в этих контурах структурно неустойчивы.
Пример 2: A_{2k+1}. Диаграмма для этой функции катастрофы имеет вид

(7.33a)

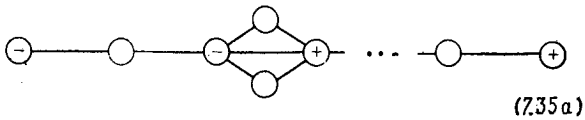
Следуя шагам $a-d$ вышеизложенного примера, получим следующее контурное представление A_{2k+1}:

(7.33d)

Пример 3. Два открытых множества для катастрофы D_{2k}, описывающие семейство функций с $2k$ изолированными критическими точками, имеют следующее диаграммное представление:

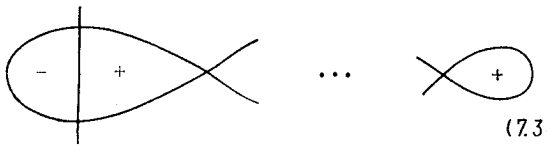


(7.34a)



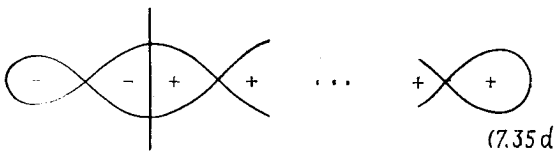
(7.35a)

Соответствующие контурные представления имеют вид



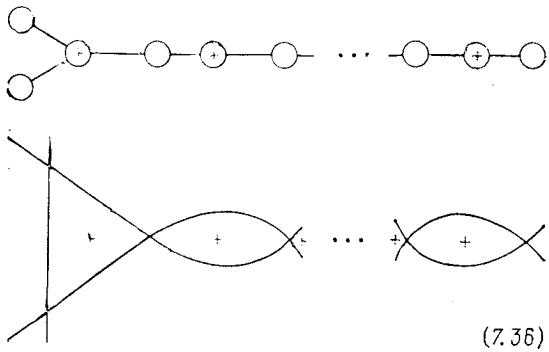
(7.34d)

и



(7.35d)

Пример 4. Одним из возможных диаграммных представлений D_{-2k} и соответствующим ему контурным представлением являются



(7.36)

В табл. 7.1 также приведен контур нулевого уровня, соответствующий каждому типу структурно устойчивого семейства функций для каждой из элементарных катастроф.

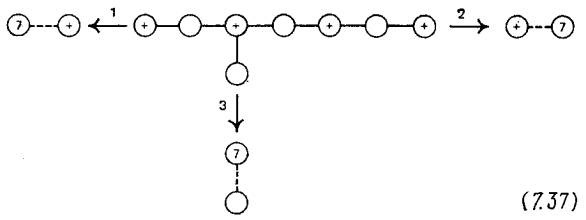
5. ПРИМЫКАНИЯ

Было показано, что для катастрофы A_{k+1} существуют одномерные кривые катастроф типа A_k, исходящие из начала координат пространства управляющих параметров $\mathbb{R}^k$, а для катастрофы D_{k+1} существуют 1-мерные кривые катастроф типа D_k ($k \geqslant 4$) и типа A_k, исходящие из начала координат пространства управляющих параметров $\mathbb{R}^k$. Про эти катастрофы меньшей размерности говорят, что они примыкают к катастрофам высшей размерности.

Теперь можно попытаться ответить на вопрос: какие катастрофы с пространством управляющих параметров размерности $k-1$ могут примыкать к данной элементарной катастрофе размерности k? Для того чтобы проиллюстрировать, как на этот вопрос можно отвечать, используя диаграммное представление исходной катастрофы, и какое количество информации может быть выжато из диаграммного метода, рассмотрим, какие катастрофы примыкают к E_8.

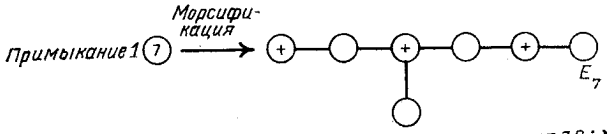
В пространстве управляющих параметров $\mathbb{R}^7$ катастрофы E_8 существуют многочисленные открытые множества, описывающие структурно устойчивые функции с четным числом критических точек. Однако существует всего лишь пять открытых множеств, параметризующих функции с максимальным числом восемь изолированных критических точек (ср. с табл. 7.1). Сначала рассмотрим одно конкретное открытое множество, диаграммное представление которого включает четыре минимума и четыре седла. Умело обращаясь с управляющими параметрами, можно заставить семь из восьми критических точек стать вырожденными.

Следовательно, для данного конкретного открытого множества E_8 имеем три примыкания:



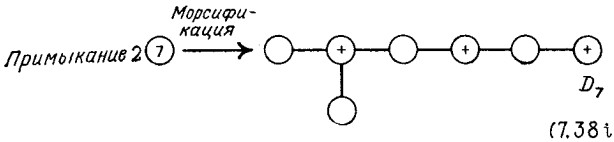
(7.37)

Для примыкания 1 очевидно, что в случае одной из морсификаций вырожденной точки кратности семь имеем следующее диаграммное представление:



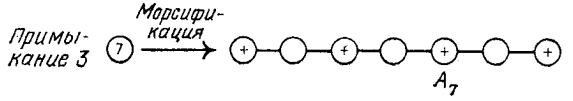
(7.38 i)

Морсификацией, связанной с примыканием 2, является



(7.38 ii)

И наконец, морсификацией, связанной с примыканием 3, является

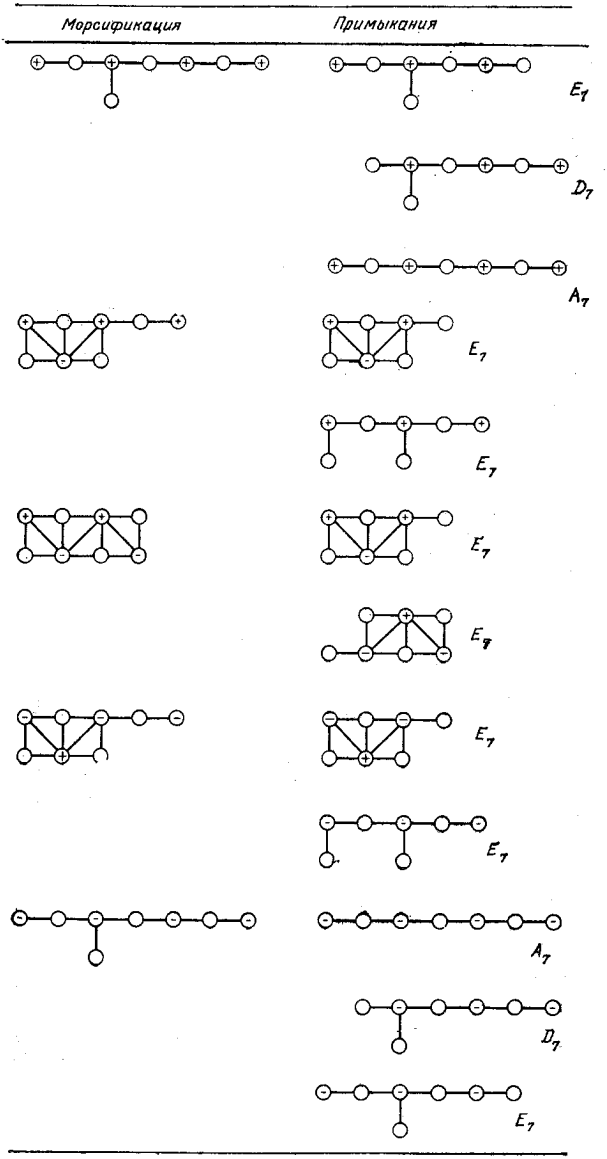


(7.38 iii)

Подобные диаграммные методы позволяют не только установить, какие из катастроф меньших размерностей примыкают к E_8, но и также выяснить, какие открытые множества в пространстве управляющих параметров этих катастроф примыкают к определенным открытым множествам в пространстве управления E_8.

Аналогично могут быть обработаны остальные четыре открытых множества E_8, параметризующие функции типа E_8 с максимальным числом изолированных критических точек. (Результаты представлены в табл. 7.2.)

Таблица 7.2. **Максимальные морсификации E_8 и их примыкания**



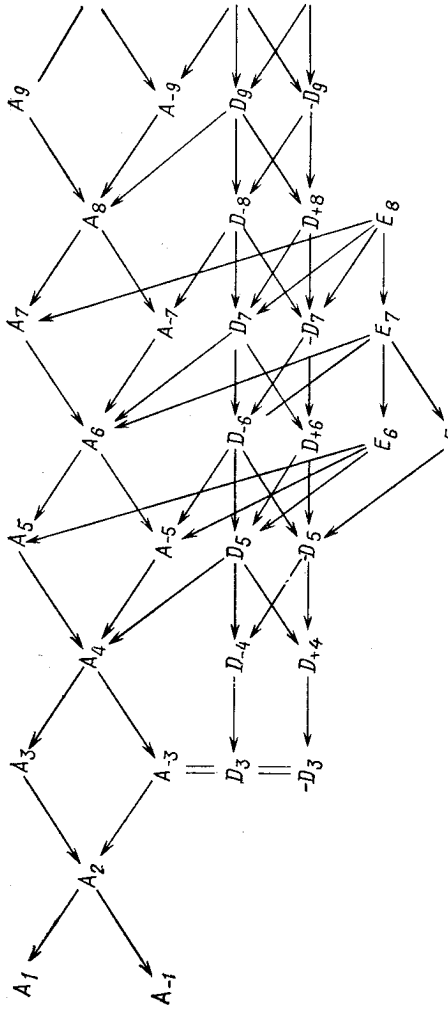


Рис. 7.5. Диаграмма, содержащая информацию относительно того, какие из катастроф, пространство управляющих параметров которых имеет размерность $k-1$, примыкают к любой из элементарных катастроф, пространство управляющих параметров которой имеет размерность k.

Примыкания для других элементарных катастроф могут быть определены тем же самым способом. Мы представили на рис. 7.5 диаграмму примыканий для всех элементарных ката-

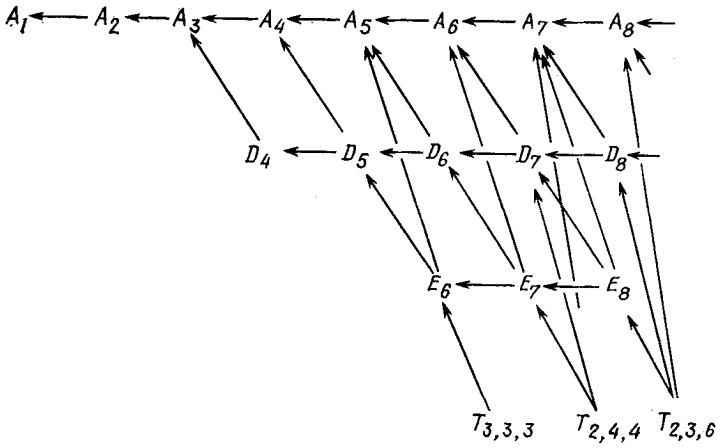


Рис. 7.6. Диаграммы примыканий комплексных элементарных катастроф (x, y — комплексные числа) соответствуют диаграммам Коксетера — Дынкина некоторых простых групп Ли над полем комплексных чисел (по данным Арнольда). Катастрофы $T_{p,q,r}$ не являются элементарными (ср. с гл. 17, разд. 1).

строф. Для некоторых целей очень полезно иметь несколько более простую диаграмму примыканий, которая содержит несколько меньше информации (рис. 7.6).

6. ВЫВОДЫ

Если F является функцией катастрофы, то уравнения $\nabla F = 0$, $\det F_{ij} = 0$ определяют сепаратрису в пространстве управляющих параметров. Точки, расположенные на этой сепаратрисе, представляют структурно неустойчивые функции с одной или более неморсовскими критическими точками. Сепаратриса состоит из компонент размерности 0, 1, 2, ..., $k-1$, лежащих в пространстве $\mathbb{R}^k$, и разбивает пространство $\mathbb{R}^k$ на конечное число открытых областей, представляющих структурно устойчивые функции. Если F есть элементарная катастрофа, то объединение этих открытых множеств всюду плотно в пространстве $\mathbb{R}^k$.

Уравнения одномерных компонент сепаратрисы и любой из ее компонент можно определить, используя пересчетные соображения. Так как даже в лучшем случае это довольно утомительный процесс, который к тому же не позволяет получить ответ на вопрос: как организованы катастрофы, предлагается более простой и менее утомительный диаграммный метод.

С помощью этого метода можно не только определить все компоненты сепаратрисы и то, каким образом они собраны вместе, но и найти все различные открытые области, параметризующие структурно устойчивые функции. В действительности, можно установить, какие из компонент сепаратрисы будут общей границей между парой открытых областей (так называемый *метод стягивания*).

Для любой из катастроф также возможно контурное представление.

Было показано, что, используя диаграммное представление функций катастроф, можно легко определить, какие кривые катастроф размерности $k-1$ исходят из начала координат $\mathbb{R}^k$ рассматриваемой катастрофы размерности управления k.

Поскольку катастрофы являются каноническими, то такова же и их организация: любой росток катастрофы действует как организующий центр, вокруг которого по каноническим орбитам вращаются низшие катастрофы,которые привносят свой «канонический вклад».

◇◇◇ Диаграммные и контурные методы, описанные в этой главе, были развиты В. И. Арнольдом [1, 2, 3] для случая комплексных форм элементарных катастроф. Ограничение этих приемов для случая вещественных форм было проведено Гусейн-Заде [4] и А. Кампо [5, 6]. Прекрасный обзор указанных работ можно найти у Каллагана [7], и я благодарен проф. Дж. Эдвину Кларку за то, что он обратил мое внимание на эту монографию.

Литература

1. Арнольд В. И. Нормальные формы для функций вблизи вырожденных критических точек; группы Вейля A_k, D_k, E_k и лагранжевы особенности. — Функ. анализ и его прилож., 1972, 6 : 4, 3—25.
2. Арнольд В. И. Замечания о методе стационарной фазы и числах Коксетера. — УМН, 1973, 28 : 5, 17—44.
3. Arnol'd V. I., Critical Points of Smooth Functions, Proc. Int. Cong. Math., Vancouver, 1974, pp. 19—39.
4. Гусейн-Заде С. М. Диаграммы Дынкина для особенностей функций двух переменных. — Функц. анализ и его прилож. 1974, 8 : 4, 23—30.
5. A'Campo N. A., Le Groupe de Monodromie du Déploiement des Singularités Isolées de Courbes Planes, I, *Math. Ann.*, **213**, 1—32 (1975).
6. A'Campo N. A., Le Groupe de Monodromie de Déploiement des Singularités Isolées de Courbes Planes, II, Proc. Int. Cong. Math., Vancouver, 1974, pp. 395—404.
7. Callahan J., Singularities and Plane Maps II: Sketching Catastrophes, *Math. Monthly*, **84**, 765—803 (1977).

8

ПРИНЦИП МАКСИМАЛЬНОГО ПРОМЕДЛЕНИЯ И ПРИНЦИП МАКСВЕЛЛА

Для того чтобы сделать математический аппарат элементарной теории катастроф доступной специалистам в различных областях науки и техники, принимаются ряд соглашений, выполняющих роль dei ex machina. К числу таких соглашений относятся принцип максимального промедления и принцип Максвелла, которые и являются предметом обсуждения данной главы. Эти принципы можно рассматривать как своего рода «экстремали» всех возможных допущений теории катастроф; по существу они являются временной мерой, позволяющей расширить область применения результатов элементарной теории катастроф.

1. НЕОБХОДИМОСТЬ ПРИНЯТИЯ СОГЛАШЕНИЙ

Как было показано (гл. 1), описание качественных изменений в поведении системы в зависимости от управляющих параметров не может быть получено в общем виде. Подобная задача может быть полностью решена лишь в простейшем случае, а именно при описании состояний равновесия градиентных систем. В физике довольно часто приходится иметь дело с такими системами (т. е. с системами, состояние которых может быть описано потенциальными функциями, зависящими от управляющих параметров), и только в случае таких систем результаты элементарной теории катастроф могут быть использованы непосредственно. Поскольку, однако, мы исключили все изменения состояния системы динамического характера (т. е. все производные по времени, гл. 1), то указанные системы обязательно должны быть статическими, и, следовательно, анализ эволюции состояния равновесия градиентных систем выходит за пределы возможностей теории катастроф. Поэтому в случае изучения динамики системы необходимо сделать о ней некоторые предположения физической или интуитивной природы (например, эволюция носит квазистатический или адиабатический характер, т. е. все производные по времени очень малы). Даже если состояние системы может быть описано потенциальной функцией, совсем не обязательно, чтобы оно полностью характеризовалось точкой, в которой функция минимальна. Типичным примером этого является одномерный гармонический осциллятор, поведение которого описывается потенциальной функцией

$V(x) = \frac{1}{2}k(x-a)^2$. Точка $x = a$ минимизирует потенциал и полностью описывает классический (статический) гармонический осциллятор. Однако квантовомеханический гармонический осциллятор описывается посредством вероятностного распределения, которое концентрируется в точке a основного состояния, а его ширина определяется коэффициентом члена диффузии $(-\hbar^2/2m)(\partial^2/\partial x^2)$ в уравнении Шредингера. Многие представляющие интерес физические системы также могут быть описаны функциями распределения $P(x)$, которые, как правило, довольно просто связаны с потенциальными функциями $V(x)$ (например, $P(x) \simeq e^{-V(x)/D}$). По существу если дифференциальное уравнение, описывающее систему, содержит вторые производные, функции распределения будут присутствовать в общем случае. Поскольку, однако, мы уже достигли шага 2 (гл. 1) и тем самым исключили пространственные производные любого порядка, для изучения систем, описываемых функциями распределения, необходимо наложить на изучаемую систему некоторые внешние ограничения физической или интуитивной природы. Например, можно предположить, что по сравнению с высотой внутреннего барьера уровень шума в системе либо очень мал, либо слишком велик.

Короче говоря, математические методы элементарной теории катастроф, описанные в гл. 2—7, формально применимы лишь на уровне шага 7 (табл. 1.1). Для того чтобы иметь возможность применять их на более раннем уровне (табл. 1.1), необходимы какие-то внешние (неприсущие самой теории катастроф) соглашения (принципы) более высоко структурированной природы.

2. СОГЛАШЕНИЯ

Состояние физической системы, управляемой потенциалом $V(x; c)$, описывается (по крайней мере частично) такой точкой $x \in \mathbb{R}^n$, в которой потенциал достигает минимального значения. Изменения внешних условий приводят к изменениям управляющих параметров, что в свою очередь влияет на вид потенциальной функции $V(x; c)$; при изменении последней первоначальный глобальный минимум, который определял состояние системы, может стать метастабильным локальным минимумом (так как некоторый удаленный минимум стал иметь меньшее значение) или даже совсем исчезнуть. В этом случае система должна перескочить из одного локального минимума в другой. Момент перехода и минимум, в котором состояние системы будет устойчивым, определяются в соответствии со следующими принципами:

Принцип максимального промедления — состояние системы определяется устойчивым (стабильным) или метастабильным минимумом до тех пор, пока он существует (рис. 8.1). Данный принцип обычно применяется при описании градиентных динамических систем

$$\frac{dx_i}{dt} = -\frac{\partial}{\partial x_i} V(x; c), \tag{8.1}$$

как только $dc_\alpha/dt \ll 1$. Это объясняется тем, что потенциальная функция изменяется очень медленно, и мы не располагаем способом, позволяющим определить, когда некоторый отдаленный минимум станет глобально устойчивым. Если существует ло-

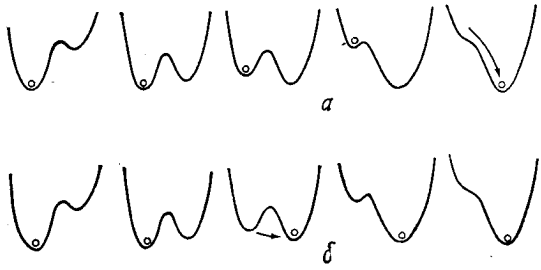


Рис. 8.1.
а — принцип максимального промедления; *б* — принцип Максвелла.

кально устойчивый минимум равновесия в точке $(x^0(c(t)), c(t))$, то $dx^0/dt = 0$ остается условием равновесия до тех пор, пока потенциальная функция устойчива в точке x^0. Только тогда система может достигнуть нового локального минимума.

Принцип Максвелла — состояние системы определяется глобальным минимумом потенциальной функции (рис. 8.1).

Если система описывается функцией распределения, то условия, при которых система выходит из одного состояния метастабильного равновесия и начинает двигаться к другому метастабильному равновесию или устойчивому равновесию, зависят от уровня шума в системе (рис. 8.2).

Высота барьера ΔE может быть определена непосредственно из потенциала $V(x; c)$, а удовлетворительное описание шума может быть получено на основе уравнений в частных производных. Тогда $\mathcal{N}$ определяется через коэффициенты вторых частных производных соответствующего уравнения. Малый уровень шума вызывает малые флуктуации, которые в свою очередь могут влиять на вид потенциала только локально, так что отдаленные минимумы не обнаруживаются (принцип максимального промедления). Большие же флуктуации могут приводить к отдален-

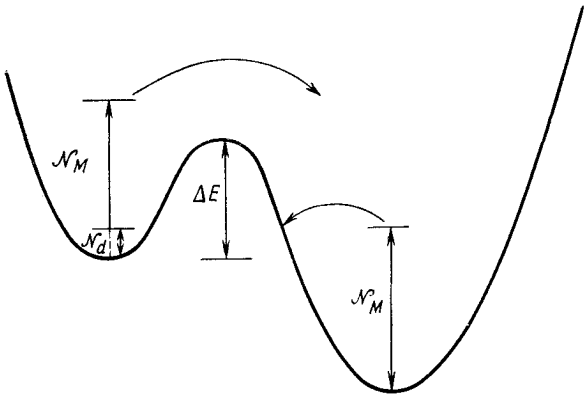


Рис. 8.2. Отношение уровня шума $\mathcal{N}$ к высоте барьера ΔE позволяет определять, какой из двух принципов следует принять: $\mathcal{N}_d/\Delta E \ll 1$ — принцип максимального промедления или $\mathcal{N}_M/\Delta E \simeq 1$ — принцип Максвелла.

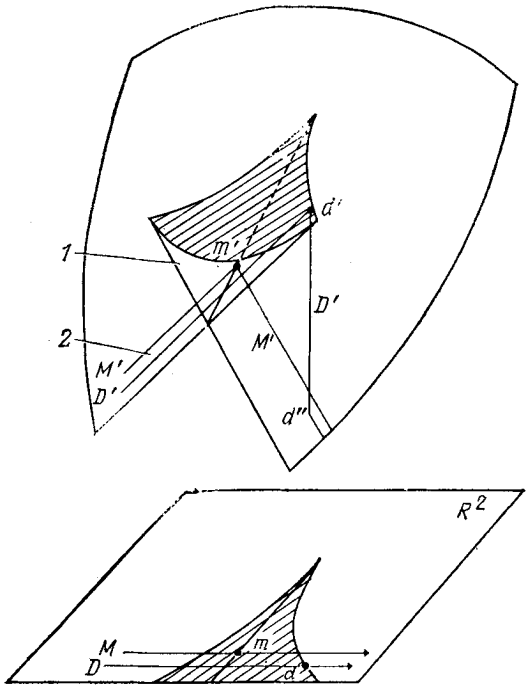


Рис. 8.3. Интерпретация соглашений в терминах поверхности критических значений катастрофы.

1 — левостороннее минимальное значение; *2* — правостороннее минимальное значение.

ным глубоким минимумам и соответствуют принципу Максвелла.

В случае катастрофы сборки эти два принципа могут быть интерпретированы в терминах поверхности критических значений (рис. 8.3.).

Принцип максимального промедления — состояние системы изменяется вдоль листа на поверхности до тех пор, пока этот лист не исчезнет. В этом случае система переходит из одного состояния в другое, что соответствует переходу на нижерасположенный лист.

Принцип Максвелла — состояние системы непрерывно изменяется вдоль самого нижнего листа поверхности, соответствующего рассматриваемой точке пространства управляющих параметров, перемещение с одного листа на другой происходит всякий раз, как только управляющий параметр пересекает полупрямую $a < 0$, $b = 0$.

Определение. Бифуркационным множеством называется множество точек в пространстве управляющих параметров, в которых происходит переход из одного локального минимума в другой.

◇◇◇ Бифуркационное множество существенно зависит от используемого соглашения (принципа). Для катастрофы сборки A_3 при использовании принципа максимального промедления бифуркационное множество состоит из двух линий складки, а при использовании принципа Максвелла из полупрямой $a < 0$, $b = 0$. Бифуркационное множество при использовании принципа максимального промедления в общем случае может быть определено из условия $\det|\partial^2 V/\partial x_i \partial x_j| = 0$. Так как это условие зависит от производных, то оно является локальным условием, и соответствующее бифуркационное множество называют *локальным бифуркационным множеством.* При использовании принципа Максвелла бифуркационное множество может быть определено посредством уравнений Клаузиуса — Клапейрона, такое множество называют *нелокальным бифуркационным множеством.*

◇◇◇ Переход из одного локального минимума в другой называют *фазовым переходом.* Фазовые переходы соответствуют качественным изменениям свойств системы.

3. КАКИЕ СОГЛАШЕНИЯ ИСПОЛЬЗОВАТЬ

Между двумя «экстремалями» возможных соглашений теории катастроф — принципом максимального промедления и принципом Максвелла — «лежат» соглашения, которые могут быть описаны некоторым уравнением в частных производных. Последнее может быть решено с учетом изменения управляющих параметров при условии, что ни один из вышеуказанных

принципов не применяется. В этом случае бифуркационное множество представляет собой *размытое множество*. Более точно, если не используется ни одно из возможных соглашений, то в общем случае невозможно определить бифуркационное множество рассматриваемого физического процесса.

Теперь наполним эти веские замечания реальным содержанием и рассмотрим ситуации, которые весьма часто возникают в физических системах. В таких случаях система обычно описывается вероятностной функцией распределения P, которая связана с потенциальной функцией V посредством уравнения Фоккера — Планка [2]:

$$\frac{\partial}{\partial t} P = \nabla \cdot (P \nabla V) + \nabla^2 (DP). \tag{8.2}$$

Потенциал V является функцией переменной состояния $x \in \mathbb{R}^n$ и k управляющих параметров $c \in \mathbb{R}^k$. В то время как управляющие параметры зависят от времени, потенциал V от него не зависит. Функция распределения P зависит как от x, так и от c и t. Найдем решения уравнения (8.2) в случае (*А*) асимптотического предела, не зависящего от времени, и (*Б*) предела, свободного от диффузии.

А. Управляющие параметры зафиксированы, и потенциальная функция не зависит от времени. Поэтому можно искать стационарную ($t \to \infty$) вероятностную функцию распределения, положив $\partial P/\partial t = 0$. В результате уравнение (8.2) сводится к

$$0 = \nabla \cdot (P \nabla V + \nabla (DP)), \tag{8.3}$$

решением которого является

$$P(x;\ c) = Ne^{-V(x;\ c)/D}, \tag{8.4}$$

где D — некоторая константа. Если коэффициент диффузии не постоянен, то решение может быть записано в виде

$$P(x;\ c) = ND^{-1}(x) \exp - \left\{ \int^{x} D^{-1}(x') \nabla V(x';\ c)\, dx' \right\}. \tag{8.5}$$

Член N в формуле (8.5) представляет собой соответствующую константу нормализации.

Б. Так как коэффициент диффузии D равен нулю и уравнение (8.2) принимает вид

$$\frac{\partial P}{\partial t} + \nabla \cdot j = 0, \qquad j(x;\ t) = - P \nabla V. \tag{8.6}$$

Теперь легко проверить [3], что решением уравнения (8.6) является

$$P(x,\ t) = \delta (x - x_c(t)), \tag{8.7}$$

где $x_c(t)$ — некоторая критическая точка потенциальной функции V (т. е. $\nabla V = 0$, если $x = x_c$).

Вернемся к обсуждению качественных свойств уравнения Фоккера — Планка (8.2). Правая часть этого уравнения состоит из двух членов — «дрейфа» и «диффузии». Грубо говоря, дрейф $\nabla(P\nabla V)$ заставляет функцию распределения двигаться по направлению к ближайшему локальному минимуму. Роль диффузии $\nabla^2(DP)$ двояка: она описывает (1) размах функции распределения, которая концентрируется вокруг локального минимума, и (2) вероятность, с которой флуктуация может перевести систему из метастабильного минимума в некоторый отдаленный глобальный минимум.

Наличие двух членов в правой части уравнения Фоккера — Планка (8.2) свидетельствует о том, что оно является уравнением с двумя временными масштабами. Это значит, что явления, описываемые качественно уравнением (8.2), происходят в двух совершенно различных временны́х шкалах: быстрой шкале времени T_1, связанной с обратной релаксацией к локальному минимуму после возмущения, и медленной шкале времени T_2, связанной с переходом из метастабильного минимума к глобальному минимуму.

Релаксация. Для простоты рассмотрим потенциальную функцию одной переменной $V(x) = \frac{1}{2} kx^2$. С этой функцией связана не зависящая от времени функция вероятностного распределения (8.4) $P(x) = Ne^{-kx^2/2D}$. Если подвергнуть эту систему возмущению таким образом, чтобы начальная вероятность сконцентрировалась вокруг точки $x = a$ $(a \neq 0)$, а не вокруг точки $x = 0$, так что

$$P(x,\ t=0) = Ne^{-k(x-a)^2/2D},$$

то функция вероятностного распределения будет эволюционировать во времени, подобно функции

$$P(x,\ t) = Ne^{-k[x-a(t)]^2/2D}, \tag{8.8}$$

где

$$a(t) = ae^{-kt} = ae^{-t/T_1}. \tag{8.9}$$

Время релаксации $T_1 = 1/k = (d^2V/dx^2)^{-1}$.

В более общем случае, когда потенциал V имеет в точке $x = x^0$ локальный минимум, а начальное распределение вероятности концентрируется вокруг точки, расположенной вблизи,

$$P(x,\ t=0) = Ne^{-V(x^0+\delta x)/D}, \tag{8.10}$$

функция распределения будет релаксировать к не зависящей от времени функции распределения вероятности $P(x;\ t \to \infty) =$

$= Ne^{-V(x_0)/D}$, имеющей масштабы времени, соразмерные с обратными величинами собственных значений λ_i матрицы устойчивости $V_{ij}(x^0)$. Следовательно, масштаб быстрой шкалы времени T_1, связанный с релаксацией, задается формулой

$$T_1 = \operatorname*{Max}_{i=1,\ldots} 1/\lambda_i. \tag{8.11}$$

Время релаксации расходится по мере того, как приближается росток катастрофы (критическое замедление).

Время первого перехода. Уравнения диффузии типа Фоккера — Планка имеют очень глубокую связь с теорией вероятностей [4]. Эта связь подробно рассматривается в теории марковских процессов. Здесь же мы исследуем ее для того, чтобы определить некоторые качественные характеристики уравнения Фоккера — Планка, не решая его.

Рассмотрим случайный путь, описываемый уравнением Фоккера — Планка, которое содержит потенциальную функцию $V(x)$ и член «хаотичности» D. Для конкретности ограничимся одномерным случаем, когда граница состоит из точек $x = a$, $x = b$, а случайный путь начинается в некоторой внутренней точке x: $a < x < b$. Любой случайный путь, начинающийся в точке x, должен закончиться спустя некоторое время τ, которое является случайной переменной. Среднее значение времени до «остановки» $T(x) = \langle\tau\rangle_{\text{ens}}$ случайного пути, начинающегося из x, может быть вычислено с помощью методов теории вероятностей. Если случайный путь начинается в точке x с вероятностью $P_0(x)$, то среднее время до его «остановки» определяется формулой

$$T_{fp} = \int_a^b P_0(x)\,T(x)\,dx. \tag{8.12}$$

Время первого перехода T_{fp} задает масштаб медленной шкалы времени T_2 уравнения Фоккера — Планка (8.2) для катастрофы сборки.

Хотя функция распределения вероятностей $P(x, t)$ уравнения (8.2) определяется уравнением в частных производных, функция $T(x)$ [функция времени первого перехода для дельта-функции распределения $P_0(y) = \delta(y - x)$] удовлетворяет обыкновенному дифференциальному уравнению вида

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x}\left(\frac{dV}{dx}P\right) + \frac{\partial^2}{\partial x^2}(DP), \tag{8.13FPE}$$

$$-1 = -\frac{dV}{dx}\frac{dT}{dx} + D\frac{d^2T}{dx^2}. \tag{8.13FPT}$$

Поскольку случайный путь, начинающийся на границе $x = a$ или $x = b$, должен окончиться немедленно, функция времени

первого перехода $T(x)$ подчиняется следующим граничным условиям:

$$T(a) = T(b) = 0. \tag{8.14}$$

Эти два граничных условия определяют в дифференциальном уравнении второго порядка (8.13 FPT) две постоянных интегрирования.

Прежде чем решать задачу о времени первого перехода для катастрофы сборки, попытаемся пояснить ее важность на примере двух простых ситуаций

1. $V = 0$. В этом случае уравнение для $T(x)$ сводится к уравнению

$$D\frac{d^2T}{dx^2} = -1. \tag{8.15}$$

Решением уравнения (8.15), удовлетворяющим граничным условиям (8.14), является

$$T(x) = \frac{(b-x)(x-a)}{2D}. \tag{8.16}$$

Случайный путь имеет наибольшую длительность до «остановки», когда $x = (a+b)/2$ («в середине»).

2. $V = \alpha x$. В этом случае уравнение для $T(x)$ на первый взгляд кажется более сложным, чем оно есть на самом деле. В действительности же оно сводится к двум простым обыкновенным дифференциальным уравнениям первого порядка:

$$D\frac{dT'}{dx} - \alpha T' = -1, \tag{8.17а}$$

$$T' = \frac{dT}{dx}. \tag{8.17б}$$

Уравнение (8.17а) имеет решение

$$T'(x) = Ce^{\alpha x/D} + \frac{1}{\alpha}. \tag{8.18}$$

Используя граничные условия, можно определить постоянную интегрирования C:

$$\begin{array}{ccc} \displaystyle\int_a^b \frac{dT}{dx}\,dx & & \\ \swarrow \quad \searrow & & \\ T(b) - T(a) = & \displaystyle\int_a^b \left(Ce^{\alpha x'/D} + \frac{1}{\alpha}\right) dx' & \\ \| \qquad\qquad \| & & \\ 0 \qquad = C\dfrac{e^{\alpha b/D} - e^{\alpha a/D}}{\alpha/D} + \dfrac{b-a}{\alpha}. & & (8.19) \end{array}$$

Полученное уравнение определяет константу C.

Аналогично может быть найдена функция $T(x)$: заменяя верхний предел b на внутреннюю точку x, получим

$$T(x) - T(a) = C\frac{e^{\alpha x/D} - e^{\alpha a/D}}{\alpha/D} + \frac{x-a}{\alpha},$$

$$T(x) = \frac{1}{\alpha}\left\{(x-a) - (b-a)\frac{e^{\alpha x/D} - e^{\alpha a/D}}{e^{\alpha b/D} - e^{\alpha a/D}}\right\}. \qquad (8.20)$$

Анализ формулы (8.20) позволяет сделать вывод о существовании конкуренции между дрейфом ($V' = \alpha$) и диффузией (D). Для того чтобы пояснить действие этой конкуренции, рассмотрим два частных предела:

1. $\alpha \to 0$, $D =$ Константа, $\zeta = \alpha/D \to 0$. В этом случае член, описывающий дрейф, менее важен, чем член, описывающий диффузию, и

$$DT(x) = \zeta^{-1}\left\{(x-a) - (b-a)\frac{e^{\zeta x} - e^{\zeta a}}{e^{\zeta b} - e^{\zeta a}}\right\} \overset{\zeta \to 0}{\to} \frac{1}{2}(b-x)(x-a). \qquad (8.21)$$

2. $\alpha =$ Константа, $D \to 0$, $\zeta = \alpha/D \to \infty$. В этом случае член, описывающий диффузию, менее важен, чем член, описывающий дрейф, и

$$\alpha T(x) = \left\{(x-a) - (b-a)\frac{e^{\zeta x} - e^{\zeta a}}{e^{\zeta b} - e^{\zeta a}}\right\} \overset{\zeta \to \infty}{\to} (x-a) - (b-a)e^{-\zeta(b-x)}. \qquad (8.22)$$

Вероятность того, что случайный путь кончится в точке $x = b$, весьма мала, если только не $\zeta(b-x) \sim 1$. Случайный путь закончится в точке $x = a$, и время $T(x)$ связано с расстоянием $(x-a)$ соотношением $\alpha T = x - a$. Поэтому резонно интерпретировать α как скорость дрейфа.

Вернемся к задаче решения уравнения (8.13 FPT), позволяющего найти время первого перехода $T(x)$ для системы, задаваемой общей потенциальной функцией одной переменной $V(x)$. Для простоты предположим, что коэффициент диффузии D постоянен. Уравнение (8.13 FPT) является обыкновенным дифференциальным уравнением первого порядка от функции $T'(x) = dT/dx$. Поэтому оно может быть решено введением интегрирующего множителя $\Lambda(x)$. В результате имеем

$$\Lambda T'(y) = -\int^{y} D^{-1}\Lambda(z)\,dz + C_1, \qquad (8.23)$$

где $\Lambda(z) = e^{-V(z)/D}$. Постоянная C_1 может быть определена интегрированием подобно тому, как это было сделано в случае

(8.19). При этом получим

$$0 = T(b) - T(a) = \int_a^b T'(y)\,dy = \tag{8.24}$$

$$= C_1 \int_a^b \Lambda^{-1}(y)\,dy - \int_a^b dy \Lambda^{-1}(y) \int^y D^{-1}\Lambda(z)\,dz.$$

Теперь остается лишь выполнить интегрирование. Для этого предположим, что $V(x)$ является функцией с двумя локальными минимумами (рис. 8.4). Так как нас в основном интересует про-

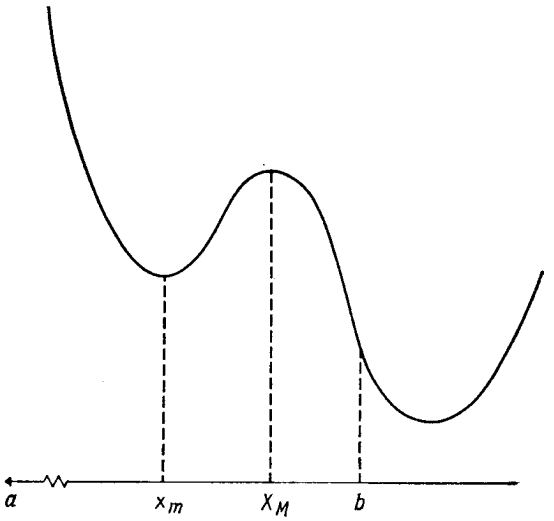


Рис. 8.4. Интеграл (8.24) может быть оценен разложением области интегрирования, как это показано на рисунке, и последующим применением метода Лапласа.

цесс диффузии из метастабильного минимума в устойчивый минимум, то определим границы случайного пути a, b согласно рис. 8.4.

Если D мало, то первый член в правой части формулы (8.24) может быть легко оценен:

$$C_1 \int_a^b e^{V(y)/D}\,dy = C_1 \int_0^{b-a} e^{[V(a)+(y+a)V'(a)+\ldots]/D}\,d(y+a) \simeq$$

$$\simeq C_1 e^{V(a)/D} \int_0^{b-a} e^{(y+a)V'(a)/D} d(y+a). \tag{8.25}$$

Верхний же предел может быть распространен до ∞ без значительного изменения значения этого интеграла.

Аналогично может быть оценен второй интеграл. Максимальный вклад в интеграл

$$I = \int_a^b dy \int^y dz D^{-1} e^{V(y)-V(z)/D} \tag{8.26}$$

получается при значениях y, z, $y > x$, которые максимизируют $V(y) - V(z)$. Этот максимум достигается при $y = X_m$, $z = x_m$. Таким образом, получаем

$$I = D^{-1} e^{V(X_M)-V(x_m)/D} \iint e^{(1/2)(y-X_M)^2 V''(X_M)-(1/2)(z-x_m)^2 V''(x_m)/D} dy\, dz \times$$

$$\times \underset{D \to 0}{} e^{V(X_M)-V(x_m)/D} \frac{2\pi}{|V''(X_M)\, V''(x_m)|^{1/2}}. \tag{8.27}$$

Время первого перехода $T(x)$ определяется интегрированием

$$T(x) - T(a) = \int_a^x T'(y)\, dy = C_1 \int_a^x e^{V(y)/D} dy - \int_a^x dy \int^y dz D^{-1} e^{V(y)-V(z)/D}. \tag{8.28}$$

Первый член формулы (8.28) может быть оценен так же, как и в случае формулы (8.25). В действительности результаты совершенно не различаются. Используя элементарную формулу математического анализа $\int_a^x + \int_x^b = \int_a^b$, найдем

$$T(x) = \int_x^b dy \int^y dz D^{-1} e^{V(y)-V(z)/D}. \tag{8.29}$$

Этот интеграл также может быть оценен с помощью методов, использованных при выводе (8.27). В результате получаем

$$T(x) = \frac{2\pi}{|V''(X_M)\, V''(x_m)|^{1/2}} e^{V(X_M)-V(x_m)/D}, \quad x < x_m,$$

$$= \frac{(2\pi D)^{1/2}}{V'(x)\, |V''(X_M)|^{1/2}} e^{V(X_M)-V(x)/D}, \quad x_m < x < X_M. \tag{8.30}$$

Ясно, что эти оценки не имеют места в окрестности ростка катастрофы. Аналогичные результаты могут быть получены и в n-мерном случае.

Для сравнения масштаб времени T_1 (релаксация в локальный минимум) и масштаб времени T_2 (диффузия из метаста-

бильного минимума в глобальный минимум) даются формулами

$$T_1 = 1/\lambda_1, \quad T_2 = \frac{2\pi}{|\lambda_1\lambda_2|^{1/2}} e^{\Delta V/D}, \tag{8.31}$$

где λ_1, λ_2 — кривизна (d^2V/dx^2) функции в локальном минимуме и локальном максимуме соответственно, а $\Delta V = V(X_M) - V(x_m)$. В общем случае, когда $\Delta V/D \geqslant 1$, $T_2 \gg T_1$. Эволюция вероятностного распределения, управляемая уравнением Фоккера — Планка, схематически изображена на рис. 8.5. Очевидно, что

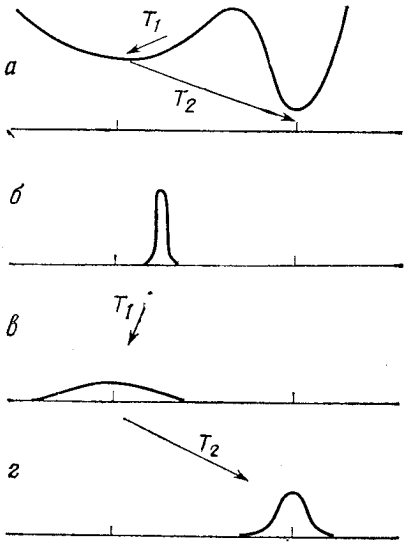


Рис. 8.5. Вероятностное распределение, концентрирующееся в окрестности одного из локальных минимумов, будет в шкале времени T_1 «асимптотически» сходиться к распределению Гаусса с соответствующей кривизной, концентрирующемуся в этом локальном минимуме. В шкале времени $T_2 \gg T_1$ вероятностное распределение будет асимптотически сходиться к ее не зависящей от времени форме $P(x) \simeq e^{-V(x)/D}$.

можно отождествить высоту барьера ΔV с ΔE (рис. 8.2), а D с $\mathcal{N}$; в этом смысле шум D является обобщенной температурой.

Теперь обсудим условия применимости соглашений для систем, описываемых уравнениями диффузии [5]. Скорость, с которой движутся критические точки $x_c(t)$ потенциальной функции $V(x; c(t))$, сравнима с производной dc/dt, поэтому условия применимости соглашения могут быть выражены через

производные по времени от управляющих параметров. В результате имеем

$$\textit{Принцип максимального промедления: } T_1^{-1} \gg \frac{dc}{dt} \gg T_2^{-1}.$$

$$\textit{Принцип Максвелла: } T_2^{-1} \gg \frac{dc}{dt}. \qquad (8.32)$$

Ни один из принципов не может быть применен, если $dc/dt \gg T_1^{-1}$ или если dc/dt сравнимо по величине с T_1^{-1} или с T_2^{-1}.

◇◇◇ Если $D \to 0$, то $T_2 \to \infty$ и $T_2^{-1} \to 0$. Если управляющие параметры изменились так, что $dc/dt > T_2^{-1}$, то принцип Максвелла совсем неприменим. В случае отсутствия диффузии (нет флуктуаций) может быть применим лишь принцип максимального промедления. Следовательно, *выбираемая критическая точка* должна быть одной из точек, в которых система может находиться в начальный момент (в момент $t = 0$). При изменении значений управляющих параметров система будет оставаться в этой точке до тех пор, пока эта критическая точка не исчезнет совсем. Новая критическая точка, в которой система будет эволюционировать, может быть определена из решения уравнения движения, которым в этом случае является (1.6) или (8.1).

4. НЕАДЕКВАТНОСТЬ ПРИНЦИПОВ

Продемонстрируем на простом примере, почему принцип максимального промедления и принцип Максвелла являются вынужденной мерой, действующей до тех пор, пока не будет достигнут определенный прогресс в реализации исследовательской программы теории катастроф.

Классическая механическая система удовлетворяет следующему уравнению движения

$$\frac{d^2x}{dt^2} = \frac{F(x)}{m} - \gamma\dot{x}, \qquad (8.33)$$

в котором трение зависит от скорости движения, а сила определяется из соотношения $F = -\nabla V$. Потенциал задается формулой

$$V(x) = \frac{x^6}{6} + \frac{a_1x^4}{4} + \frac{a_2x^3}{3} + \frac{a_3x^2}{2} + a_4x. \qquad (8.34)$$

Предположим, что в начале движения система находится в левом минимуме (рис. 8.6) и остается в нем до тех пор, пока управляющие параметры не изменятся настолько, что этот минимум исчезнет. К какому из минимумов будет двигаться система?

Принцип максимального промедления в этой точке нарушен. Если бы уравнение движения имело вид

$$\frac{dx}{dt} = \frac{F(x)}{m} - \gamma\dot{x}, \tag{8.35}$$

то система должна была бы двигаться к среднему минимуму, так как в нем $F = 0$ и, следовательно, $dx/dt = 0$. Но уравнение движения (8.33) имеет второй порядок, так что инерция системы протащит ее через эту точку и, вероятно, через локальный максимум, разделяющий средний и правосторонний мини-

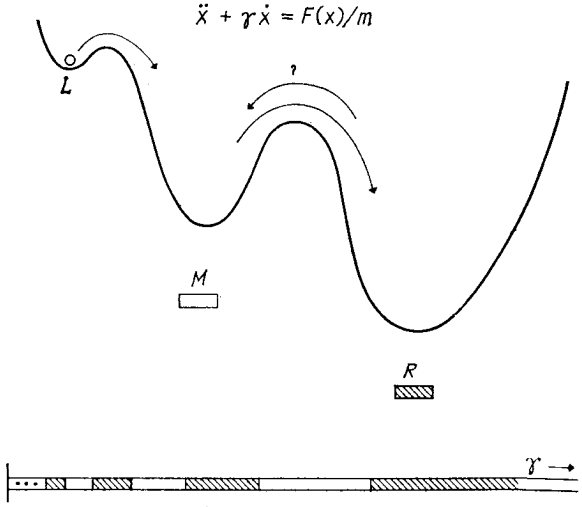


Рис. 8.6. Конечное состояние системы (8.33), метастабильный минимум которой исчезает в момент времени $t = 0$, может соответствовать как центральному, так и правому минимуму в зависимости от значения константы демпфирования.

мумы. В случае же «вязкого предела» $\gamma \to \infty$ уравнение (8.33) ведет себя, подобно уравнению (8.35), и система втягивается в средний минимум. В случае слегка «демпфированного предела» (γ — мало) система может стабилизироваться в любом из этих двух минимумов. В действительности положение минимума, в котором система в конечном счете окажется, зависит от значения демпфирующего параметра γ (рис. 8.6).

5. ВЫВОДЫ

Для большинства систем, представляющих практический интерес, предположения, лежащие в основе элементарной теории катастроф, не выполняются в полном объеме. Действительно,

в общем случае или управляющие параметры зависят от времени, или переменные состояния подвержены флуктуациям, или то и другое одновременно. Однако в ряде случаев следование определенным принципам позволяет использовать результаты элементарной теории катастроф более высокого уровня (например, выше шага 7, табл. 1.1) и таким образом получить более содержательное описание исследуемой системы (или явления). Естественно, что при этом мы жертвуем полнотой описания, но игра стоит свеч.

◇◇◇ В теорию катастроф эти два принципа впервые ввел Том [1]. Использование уравнения Фоккера — Планка в качестве механизма интерполяции между этими двумя принципами впервые было предложено Гилмором [5].

Литература

1. Thom R., Stabilité Structurelle et Morphogénèse, New York: Benjamin, 1972.
2. Wang M. C., Uhlenbeck G. E., On the Theory of the Brownian Motion II, *Rev. Mod. Phys.*, **17**, 323—342 (1945).
3. Jackson J. D., *Classical Electrodynamics*, New York: Wiley, 1962.
4. Stratonovich R. L., Topics in the Theory of Random Noise, New York: Gordon and Breach, 1963.
5. Gilmore R., Catastrophe Time Scales and Conventions, *Phys. Rev.*, **A20**, 2510—2515 (1979).

9

ФЛАГИ КАТАСТРОФ

Я уверен, что элементарная теория катастроф в конце концов станет полезным инструментом в наборе методов, используемых большинством ученых и инженеров. О наличии элементарной катастрофы свидетельствуют неморсовские критические точки в семействе потенциальных функций, описывающих данную систему или данное явление. Однако по некоторым причинам такие точки не могут быть распознаны сразу. Например, (1) потенциальная функция является слишком сложной или (2) точно не известна. Еще хуже, когда (3) система не является даже градиентной системой, или совсем плохо, когда (4) у нас нет даже туманных соображений относительно вида уравнения, надлежащим образом описывающего систему. Тем не менее катастрофы встречаются в реальных ситуациях, и поэтому важно уметь их распознавать. Это в общем-то и не так уж сложно, поскольку катастрофы имеют характерные «отпечатки пальцев» и часто вывешивают «опознавательные знаки» — флаги [1], чтобы привлечь наше внимание.

Данная глава посвящена описанию восьми стандартных флагов катастроф. Первые пять флагов всегда встречаются вместе и «вывешиваются», когда физические управляющие параметры могут изменяться внутри некоторой области пространства управляющих параметров, в которой соответствующая потенциальная функция имеет более чем один минимум [1]). Остальные три флага могут встречаться даже тогда, когда потенциальная функция не имеет кратных минимумов, что чрезвычайно важно при изучении систем для которых неожиданные переходы (скачки) в другие состояния крайне не желательны. (Например, явление коллапса в строении, качественные изменения в глобальной картине погоды или чреватое взрывом состояние химической системы.)

Как только один из этих флагов зафиксирован, т. е. установлен признак, свидетельствующий о наличии катастрофы, управляющие параметры можно изменять так, чтобы стало возможным обнаружить остальные флаги, которые обязательно должны проявить себя при соответствующих условиях. Следует иметь в виду, что уже сам факт обнаружения катастрофы имеет огромное значение для описания рассматриваемой физической системы. Установление наличия и типа катастрофы в рассмотренных выше случаях возрастающей неопределенности в описании системы могут помочь определить

— упрощенную модельную потенциальную функцию, зависящую только от существенных переменных состояния и управляющих параметров, либо

— соответствующий росток потенциальной функции, который, в свою очередь, может «подсказать», какой в действительности физический процесс имеет место, или установить

---

[1]) За исключением гистерезиса, который не может иметь места, если поведение системы подчиняется принципу Максвелла.

— соответствующий тип уравнений (динамический, диффузионный и т. д.) для системы и то, каким образом потенциальная функция может входить в такие уравнения, либо

— ненужность использования уравнений вообще, если выводы, следующие из таких уравнений, зависят главным образом от канонической геометрии соответствующей катастрофы.

Хотя катастрофы обнаруживаются при качественных исследованиях уравнений, существует эффект обратной связи, который иногда позволяет получить качественные следствия даже в том случае, когда мы не знаем самих уравнений при условии, что мы в состоянии установить наличие и тип катастрофы.

1. МОДАЛЬНОСТЬ

Физическая система может иметь два или более различных физических состояния. Другими словами, описывающая систему потенциальная функция имеет более чем один локальный минимум в некоторой области изменения внешних управляющих параметров.

Пример. Катастрофа сборки становится бимодальной, если управляющие параметры лежат в пределах области сборки (рис. 9.1).

2. НЕДОСТИЖИМОСТЬ

Если система находится в состоянии равновесия, которое оказывается морсовским i-седлом ($i > 0$), то такое состояние является неустойчивым, поскольку существуют инфинитезимальные возмущения, приводящие к уменьшению значения потенциала. Всякий раз, когда потенциальная функция имеет более чем один локальный минимум, она должна иметь по крайней мере одно i-седло ($c > 0$), которое является состоянием неустойчивого равновесия.

Пример. Два слоя в области сборки (рис. 9.1), представляющие локально устойчивые минимумы, разделены срединным недостижимым слоем, представляющим неустойчивые локальные максимумы.

3. КАТАСТРОФИЧЕСКИЕ СКАЧКИ

В любом случае — используется ли принцип Максвелла или принцип максимального промедления — малые изменения в значениях управляющих параметров могут вызывать большие изменения («катастрофический скачок») в значениях переменных состояния (рис. 8.1) по мере того, как система перескакивает из одного локального минимума в другой. Если принят принцип Максвелла, этот неожиданный скачок сопровождается плавным, но не дифференцируемым изменением значений потенциала; если принят принцип максимального промедления, скачок из одного

исчезающего локального минимума в глобальный или некоторый другой локальный минимум сопровождается дискретным (разрывным) изменением значения потенциала.

Переход из окрестности одного локального минимума в другой проявляет себя в большом изменении значения переменной состояния, которое часто происходит в сверхбыстрой временной шкале.

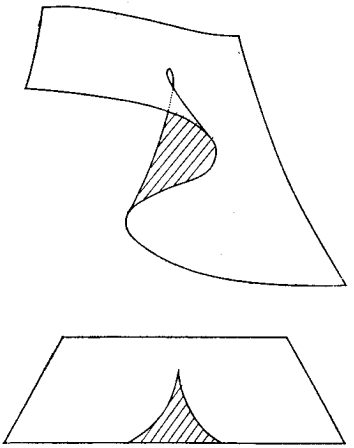


Рис. 9.1. Свойства устойчивости критических точек функции катастрофы сборки легко определяются из рассмотрения многообразия этой катастрофы.

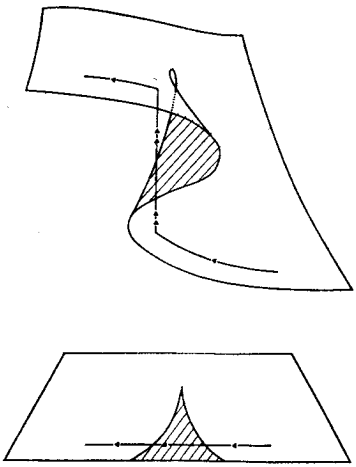


Рис. 9.2. Медленное изменение управляющих параметров может приводить к внезапному изменению переменных состояния, даже если никакие соглашения катастроф не отмечаются.

Пример. Неожиданный скачок в значении переменной состояния происходит, как только состояние системы перескакивает с одного слоя поверхности катастрофы сборки на другой (рис. 9.2). Это имеет место даже тогда, когда не используется ни один из двух экстремальных принципов.

4. РАСХОДИМОСТЬ

Конечные изменения в значении управляющих параметров приводят к конечным изменениям в значениях переменных состояния в точке равновесия. Обычно малые возмущения в исходных значениях управляющих параметров ведут лишь к небольшому изменению начальных и конечных значений переменных состояния. Однако в окрестности неморсовской критической точки малые изменения начальных значений переменных состояния могут привести к большим изменениям конечных

значений этих переменных. Неустойчивость физического процесса при возмущениях в траектории управляющих параметров называется *расходимостью.*

Пример: *Расходимость в случае катастрофы сборки.* Предположим, что некоторый физический процесс изменяет управляющие параметры (a, b) от (положительное число, ε) до (отрицательное число, ε), причем ε мало и постоянно. Тогда возмущения начальных значений управляющих параметров, которые

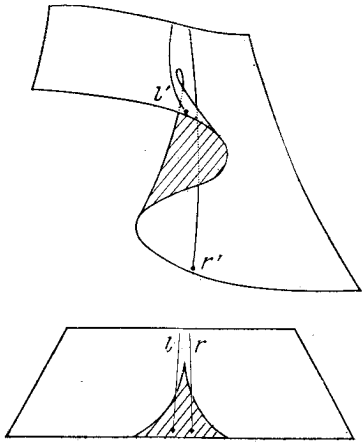


Рис. 9.3. Два близких пути в пространстве управляющих параметров могут приводить к далеко расходящимся конечным значениям переменных состояния.

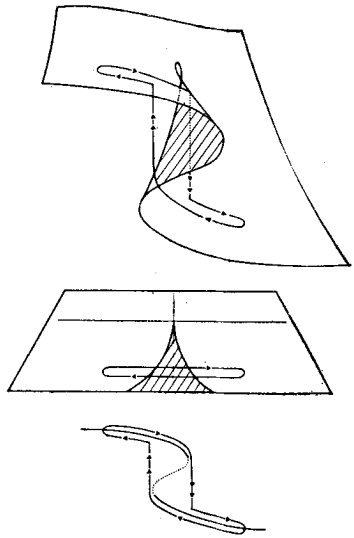


Рис. 9.4. Явление гистерезиса имеет место, когда скачок с одного листа на другой не случается при тех же значениях управляющих параметров, что и возвратный скачок.

изменяют знак b, оказывают малое воздействие на начальные значения переменных состояния, однако это возмущение оказывается весьма существенным для конечных значений (рис. 9.3).

5. ГИСТЕРЕЗИС

Гистерезис имеет место, когда физический процесс не является полностью обратимым, т. е. над той же самой точкой пространства управляющих параметров скачок из локального минимума 1 в локальный минимум 2 может и не произойти, в то время как скачок из локального минимума 2 в локальный минимум 1 имел место. Гистерезис не возникает, если принят принцип Максвелла.

Пример: *Гистерезис в случае катастрофы сборки* (рис. 9.4). Нами был принят принцип максимального промедления и сделано предположение, согласно которому управляющие параметры образуют петлю, пересекающую обе линии складки, для которой $a < 0$, а b вначале возрастает, а затем убывает. Зависящий от времени отклик переменной состояния следует за характерной петлей гистерезиса.

6. РАСХОДИМОСТЬ ЛИНЕЙНОГО ОТКЛИКА

При небольшом изменении значений управляющих параметров $(c \to c^0 + \delta c^0)$ положение точки равновесия также будет слегка изменяться $(x \to x^0 + \delta x^0)$. Связь между откликом равновесия и изменением в управлении может быть получена путем разложения $V(x; c)$ в ряд Тейлора по степеням $(x - x^0)$, $(c - c^0)$ и последующего удаления всех, за исключением линейных, членов $\nabla V = 0$ [ср. (5.2)]. В этом случае линейный отклик определяется как

$$\delta x_j^0 = -\left(V^{-1}\right)^{jk} V_{ka}\, \delta c_a^0 = \chi_{ja}\left(x^0;\ c^0\right) \delta c_a^0. \tag{9.1}$$

Линейный отклик δx_j^0 на изменение δc_a^0 определяется с помощью тензора восприимчивости χ_{ja}, который выражается через вторые производные потенциальной функции, взятые в устойчивом состоянии равновесия. Когда состояние равновесия приближается к неморсовской критической точке, $\det V_{ij} \to 0$, так что в матрице V^{-1} некоторые элементы становятся по величине слишком большими. Это означает, что функция линейного отклика χ_{ja} расходится при приближении к вырожденной критической точке.

Расходимость χ_{ja} свидетельствует не только о том, что мы находимся в окрестности вырожденной критической точки, но также и о том, что здесь присутствуют «плохие» переменные состояния. «Плохие» переменные состояния соответствуют именно тем направлениям, в которых отклик переменной состояния по отношению к малым возмущениям управления расходится. Это просто показать, если предположить, что потенциальная функция имеет морсовскую каноническую форму в точке x^0. Тогда $V_{jk} = \lambda_j \delta_j k$, так что формула (9.1) упрощается, и мы имеем

$$\delta x_j^0 = -\left(\frac{V_{ja}}{\lambda_j}\right) \delta c_a^0 \text{ (нет суммирования по } j). \tag{9.2}$$

При $\lambda_j \to 0$ отклик δx_j^0 на отличное от нуля возмущение управляющих параметров расходится.

Пример. Для катастрофы сборки $V(x;\ a,\ b) = x^4/4 + ax^2/2 + bx$ кривизна $V'' = 3x^2 + a$ в точке равновесия показана на рис. 6.2 и 6.5. Функция

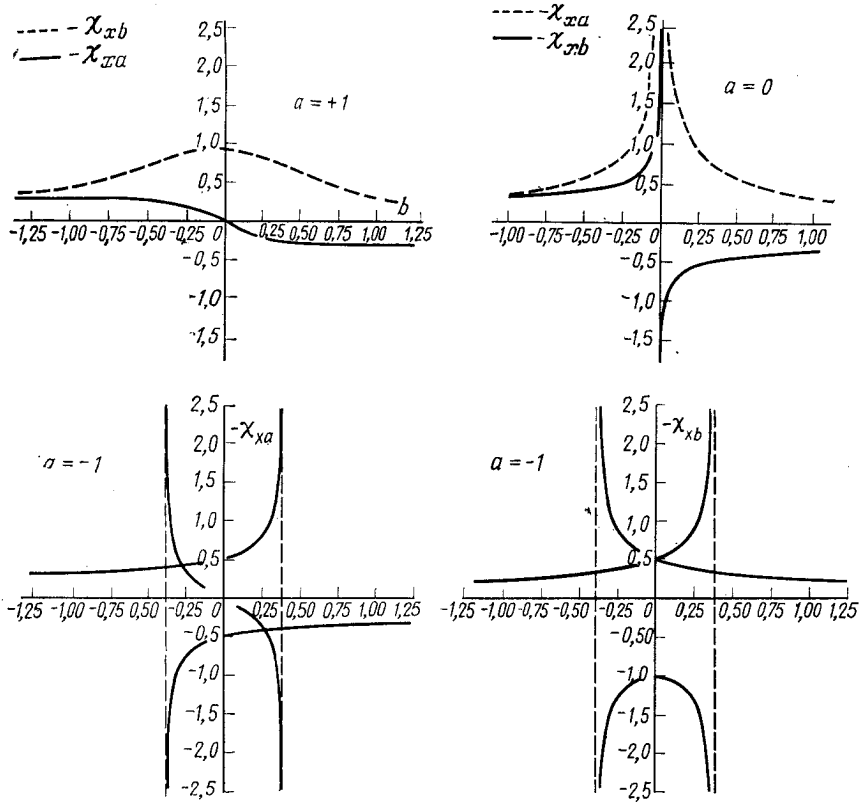


Рис. 9.5. Компоненты $-(\chi_{xa}, \chi_{xb})$ тензора статической восприимчивости катастрофы сборки как функции управляющего параметра b.

$\partial^2 V/\partial x \partial a = x_c$ может быть определена из рис. 6.2 и 6.3, в то время как $\partial^2 V/\partial x \partial b$. Для сборки тензор статической восприимчивости равен

$$(\chi_{xa}, \chi_{xb}) = \left(\frac{-x}{3x^2 + a}, \frac{-1}{3x^2 + a}\right). \tag{9.3}$$

Две компоненты этого тензора на трех путях, проходимых с возрастающим значением параметра b, изображены на рис. 9.5.

7. КРИТИЧЕСКОЕ ЗАМЕДЛЕНИЕ/СМЯГЧЕНИЕ МОДЫ

В предыдущем разделе мы пытались выяснить, каким образом (во времени) физическая система реагирует на возмущение. Однако ответ на этот вопрос неизбежно должен привести к рассмотрению временной динамики, т. е. некоторого предполагаемого уравнения движения. Если динамические явления в

системе затухают со временем, то достигается уровень аппроксимации, соответствующий шагу 7 табл. 1.1. В зависимости от характера динамического состояния физической системы она может быть либо *градиентной динамической*, либо *градиентной ньютоновской системой*, причем и в том и в другом случае она будет иметь свои собственные «катастрофические отпечатки пальцев».

7.1. Градиентная динамическая система

Уравнения, описывающие такую систему, приводились в гл. 1 [ср. (1.6)]:

$$\frac{dx_i}{dt} = -\frac{\partial V}{\partial x_i}. \tag{9.4D}$$

Если $(x^0; c^0)$ — состояние равновесия, то

$$V(x; c) = \text{Константа} + \frac{1}{2}\delta x_i \delta x_j V_{ij} + \mathcal{O}(3). \tag{9.5}$$

Уравнение движения в окрестности состояния равновесия имеет вид

$$\frac{dx_i}{dt} = -V_{ij}\delta x_j + \mathcal{O}(2). \tag{9.6D}$$

Если отбросить члены второй степени и выше, то полученные уравнения сведутся к простой системе линейных уравнений. Состояние равновесия в точке x^0 устойчиво, если все собственные значения матрицы устойчивости $V_{ij}(x^0; c^0)$ положительны. Собственные колебания имеют временную зависимость вида $e^{-\lambda_i t}$, где $\lambda_1, \ldots, \lambda_n$ — собственные значения матрицы V_{ij}. Поэтому величина $1/\lambda_i$ является характеристическим временем релаксации i-й собственной моды.

При подходе к бифуркационному множеству (предположим, что используется принцип максимального промедления) $\det V_{ij} \to 0$, так что одно или более собственных значений матрицы стремится к нулю. Следовательно, время релаксации соответствующих колебаний возрастает. Другими словами, при подходе к неморсовской критической точке для (по крайней мере) одной из мод становится все труднее релаксироваться в нулю. Это удлинение шкалы времени релаксации и называют *критическим замедлением.*

7.2. Градиентные ньютоновы системы

Подобные системы отличаются от градиентных систем (9.4D) тем, что у них первые производные по времени заменены

вторыми производными по времени (9.4N):

$$\frac{d^2x_i}{dt^2} = -\frac{\partial V}{\partial x^i}, \tag{9.4N}$$

$$\frac{d^2x_i}{dt^2} = -V_{ij}\,\delta x_j + \mathcal{O}(2). \tag{9.6N}$$

В этом случае динамический отклик в окрестности состояния равновесия (9.5) дается формулой (9.6N). Преобразование нормальных мод может быть произведено для уравнения (9.6N) так же, как для уравнения (9.6D). Если x^0 — локальный минимум, то все собственные значения матрицы V_{ij} положительны и все моды имеют периодическую временную зависимость вида $e^{i\omega_j t}$, $\omega_j^2 = \lambda_j$.

При подходе к бифуркационному множеству (используется принцип максимального промедления) $\det V_{ij} \to 0$, так что одна или более частота колебаний ω_j стремится к нулю. Это уменьшение частоты колебания для определенных мод и называют *смягчением моды.*

Пример. Для катастрофы складки матрица устойчивости равна $V'' = 3x^2 + a|_{x_c}$, где x_c — координата критической точки. Каноническая кривизна для складки была уже показана на рис. 6.2 и 6.5. Для случая градиентной динамической системы на рис. 9.6 изображены времена релаксации $1/V''$ для состояния равновесия при значениях $a = +1$, $a = 0$, $a = -1$ как функции параметра b. Для случая градиентных ньютоновых систем частота колебания для состояния равновесия как функции b при значениях $a = +1$, $a = 0$, $a = -1$ изображена на рис. 9.7. Пересчетные соотношения верны как для рис. 9.6, так и для рис. 9.7.

◇◇◇ Собственный вектор матрицы V_{ij}, соответствующий собственному значению λ_j, стремящемуся к нулю, как раз и дает «плохое» направление, связанное с рассматриваемой неморсовской критической точкой потенциала V.

◇◇◇ Градиентную динамическую систему и градиентную ньютонову систему можно рассматривать как пределы несколько более сложного классического уравнения движения

$$m\frac{d^2x_i}{dt^2} = -\frac{\partial V}{\partial x_i} - \gamma\frac{dx_i}{dt}. \tag{9.7}$$

Здесь член $-\partial V/\partial x_i$ может быть интерпретирован как консервативная сила, а $-\gamma dx_i/dt$ — как диссипативная сила трения. Обе силы действуют на частицу массы m, которая движется в пространстве $\mathbb{R}^n$ (пространство переменных состояния). В этом случае уравнение (9.7) есть не что иное, как ньютоново уравнение движения. Градиентная динамическая система получается при переходе к пределу $m/\gamma \to 0$, а градиентная ньютонова система — при переходе к пределу $\gamma/m \to 0$ [ср. (8.2), рис. 8.6].

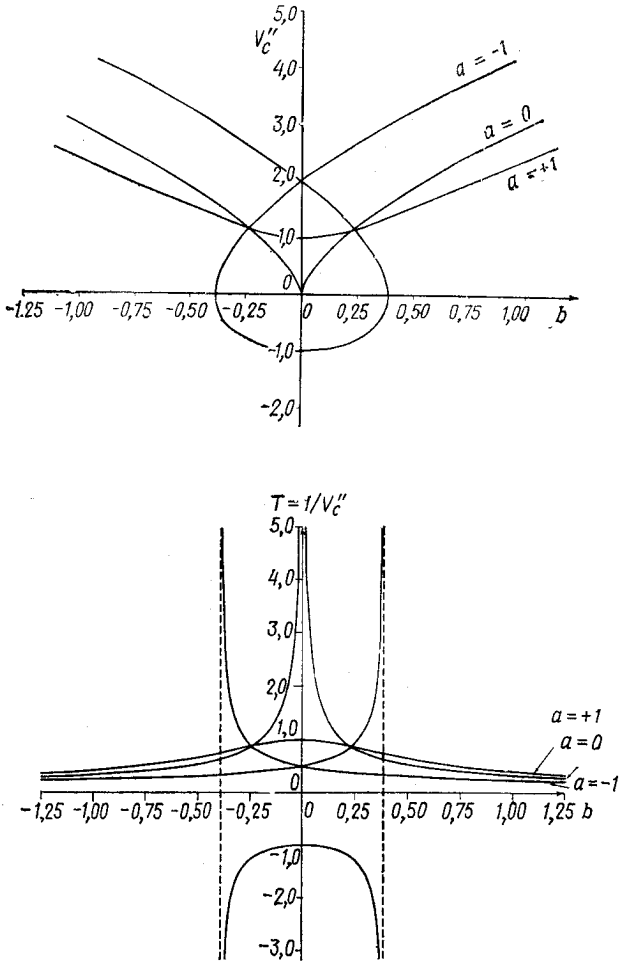


Рис. 9.6. Критическая кривизна и локальное время релаксации как функции управляющего параметра b.

Эти два предела могут быть описаны как «очень медленный» и «консервативный» пределы.

◇◇◇ Хотя градиентные динамические системы и градиентные ньютоновы системы представляют предельные случаи более общего уравнения (9.7), подобно тому как принцип Максвелла и принцип максимального промедления являются экстремумами в множестве возможностей, критические динамические свойства систем (9.7) корректно определены и за пределами диапазона $0 < \gamma/m < \infty$.

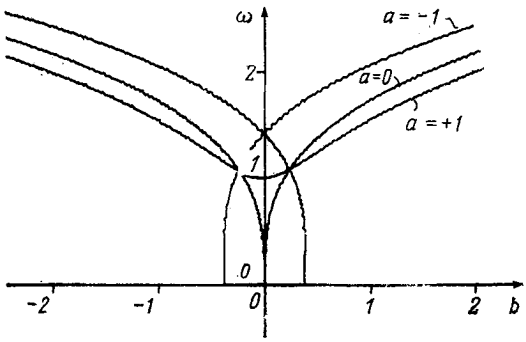


Рис. 9.7. Частоты нормальных мод показаны как функции b.

Чтобы показать это, линеаризуем потенциальную функцию в соответствии с разложением (9.5) и предположим, что существует зависимость вида $\delta x_j(t) = \delta x_j e^{i\omega t}$. В результате получим уравнение

$$[(-m\omega^2 + i\omega\gamma)\,\delta_{ij} + V_{ij}]\,\delta x_j = 0. \tag{9.8}$$

Данное уравнение имеет нетривиальное решение только в том случае, когда определитель матрицы (стоящей внутри квадратных скобок) равен нулю. При приближении к вырожденной критической точке $\det V_{ij} \to 0$, так что по крайней мере один комплексный корень ω стремится к нулю. Это и есть тот общий случай, который имеет место между критическим замедлением и смягчением моды. Как и выше, нетривиальный собственный вектор (9.8), соответствующий обращающемуся в нуль собственному значению, описывает «плохое» направление пространства переменных состояния.

◇◇◇ Если временные динамики представлены посредством оператора $f(D)$, $D = d/dt$, то уравнение движения в окрестности состояния равновесия может быть записано в виде

$$[f(D) + V_{ij}]\,\delta x_j(t) = 0. \tag{9.9}$$

Если оператор $f(D)$ линеен и предполагается существование зависимости вида $\delta x_j(t) = \delta x_j e^{\lambda t}$, то уравнение (9.9) сводится к следующему матричному уравнению:

$$[f(\lambda)\,\delta_{ij} + V_{ij}]\,\delta x_j = 0. \tag{9.10}$$

Если же f является полиномом, в котором отсутствует постоянный член, то при приближении к неморсовской критической точке $\det V_{ij} \to 0$ и по крайней мере один из корней λ урав-

нения (9.10) также стремится к нулю. (Как всегда, соответствующие собственные векторы описывают «плохие» направления.)

8. АНОМАЛЬНАЯ ДИСПЕРСИЯ

Физическая система может быть задана скорее посредством вероятности $P(x)$, определенной над пространством переменных состояния, чем посредством изолированной точки («распределения») в пространстве переменных состояния. Такая система характеризуется моментами функции распределения

$$1 = \int P(x)\,dx, \quad \langle x_i\rangle = \int x_i P(x)\,dx, \quad \langle x_i x_j\rangle = \int x_i x_j P(x)\,dx. \quad (9.11)$$

$$\vdots$$

Из этих моментов наиболее важными для нас являются среднее значение (первый момент) $\langle x_i\rangle$ и ковариация

$$\langle \Delta x_i\, \Delta x_j\rangle = \langle (x_i - \langle x_i\rangle)(x_j - \langle x_j\rangle)\rangle = \langle x_i x_j\rangle - \langle x_i\rangle\langle x_j\rangle. \quad (9.12)$$

Если теперь такая физическая система ассоциируется с потенциальной функцией $V(x, c)$, заданной над пространством $\mathbb{R}^n$ переменных функций состояния x и зависящей от управляющих параметров $c \in \mathbb{R}^k$, то не зависящая от времени вероятностная функция распределения часто связана с потенциальной функцией простым экспоненциальным путем:

$$P(x;\ c) = Ne^{-V(x;\ c)/D}, \quad (9.13)$$

где N — константа нормализации, а D — положительная константа диффузии [ср. с (8.4)].

Как правило, константа D мала. В этом случае выделяется самый глубокий минимум $V(x; c)$. Если самый глубокий минимум является морсовской критической точкой в $x = x^0$, то

$$P(x;\ c) \simeq Ne^{-V_{ij}(x-x^0)_i\,(x-x^0)_j/D}. \quad (9.14)$$

Дисперсия в этой критической точке может быть вычислена с помощью формулы (9.12). Эти вычисления несколько проще, когда потенциал V в окрестности точки x^0 приведен к каноническому виду. Это может быть сделано, поскольку операция приведения к диагональному виду $V_{ij}\Delta x_i \Delta x_j$ коммутирует с операцией взятия экспоненты. Следовательно, в случае

$$P(x;\ c) \simeq Ne^{-\Sigma \lambda_i (x_i - x_i^0)^2/D} \quad (9.15)$$

дисперсия равна

$$\langle \Delta x_i \Delta x_j \rangle = \frac{1}{2} \delta_{ij} \frac{D}{\lambda_j}. \tag{9.16}$$

Так как D мало, то дисперсия также мала, если только одно из собственных значений матрицы V_{ij} не является малым.

Это означает, что в некоторых случаях дисперсия в окрестности неморсовской критической точки может быть большой («аномальной»). В действительности при изменении управляющих параметров c_α возможны аномалии двух совершенно различных типов.

1. Минимум в точке x^0 может стать вырожденным вследствие совпадения с одной или несколькими другими критическими точками. В этом случае «плохие» переменные состояния и тип вырожденной критической точки могут быть определены следующим образом. Поскольку матрица ковариаций (9.12) является симметрической и вещественной, то ее можно привести к диагональному виду с помощью некоторого вещественного ортогонального преобразования. Это преобразование также приводит к диагональному виду и матрицу V_{ij}. Если дисперсия Δx_1 становится аномально большой после преобразования главных осей, то x_1 является «плохой» переменной состояния. В более общем виде имеем в соответствии с разложением (2.3)

$$P(x;\ c^0) \simeq N e^{-f_{NM}/D} \otimes e^{-f_M/D}. \tag{9.17}$$

Морсовская функция дает лишь члены с нормальной малой дисперсией. Неморсовские же ростки, присутствующие в разложении (9.17), должны обязательно обладать свойством нормализуемости, т. е. функция $e^{-f_{NM}/D}$ должна быть нормализуемой. Из простых ростков катастроф этим свойством обладают лишь $A_{+(2n-1)} = +x^{2n}$. Для этих ростков $\langle x \rangle = 0$ и дисперсия может быть выражена посредством гамма-функции:

$$\langle x^2 \rangle = \frac{\int\limits_{-\infty}^{+\infty} x^2 e^{-x^{2n}/D}\, dx}{\int\limits_{-\infty}^{+\infty} e^{-x^{2n}/D}\, dx} = D^{1/n} \frac{\Gamma(3/2n)}{\Gamma(1/2n)}. \tag{9.18}$$

2. В удаленном метастабильном минимуме x' значение потенциала может быть меньше, чем значение локального минимума в точке x^0. При $|V(x^0; c) - V(x'; c)| \gtrsim 10D$ вероятностное распределение является пикообразным вблизи либо точки x^0, либо x' в зависимости от того, положительна или отрицательна величина $V(x^0; c) - V(x'; c)$. В этом случае $\langle x_i \rangle \simeq x_i$ (или $\langle x_i \rangle =$

$= x_i'$) и $\langle \Delta x_i^2 \rangle \sim \frac{1}{2} D/\lambda_i^0$ (или $\langle \Delta x_i^2 \rangle \sim \frac{1}{2} D/\lambda_i'$), где λ_i^0 (или λ_i') — соответствующее собственное значение. Однако как только наименьший минимум переходит из x^0 в x', то при изменении управляющих параметров среднее значение $\langle x_i \rangle$ быстро переходит от x_i^0 к x_i'. Когда $V(x^0; c) = V(x'; c)$,

$$\langle x_i \rangle \simeq \tfrac{1}{2}(x^0 + x')_i, \quad \langle \Delta x_i^2 \rangle \simeq \left[\tfrac{1}{2}(x^0 - x')_i\right]^2. \tag{9.19}$$

В данном случае величина дисперсии может быть больше, чем в случае 1.

Пример. Предположим, что распределение вероятностей

$$\begin{aligned} P(x; a, b) &= Ne^{-V(x; a, b)/D}, \\ V(x; a, b) &= \frac{x^4}{4} + \frac{ax^2}{2} + bx \end{aligned} \tag{9.20}$$

связано с каждой точкой (a, b) плоскости управляющих параметров $\mathbb{R}^2$ катастрофы сборки (рис. 9.8).

Путь 1: $a = +1$, b возрастает. Собственное значение λ матрицы устойчивости d^2V/dx^2 как функция b для $a = +1$, $a = 0$ и $a = -1$ показано на рис. 6.5. Минимум кривизны достигается в точке $b = 0$, так что наибольшая дисперсия будет при $b = 0$.

Путь 2: $a = 0$, b возрастает. При $b \to 0$ кривизна также стремится к нулю и квадратичный член играет все более важную роль в удержании функции распределения в форме острого пика, довольно размытого в окрестности точки $x = 0$. Дисперсия в точке $(a, b) = (0, 0)$ дается формулой (9.18), из которой получаем

$$\langle \Delta x^2 \rangle = \sqrt{4D}\,\frac{\Gamma(3/4)}{\Gamma(1/4)} = 0{,}33799\sqrt{4D}. \tag{9.21}$$

Так как D по предположению мало ($D \ll 1$), то $\sqrt{4D} \gg D$ и наблюдается аномальная дисперсия в окрестности точки $(a, b) = (0, 0)$.

Путь 3: $a = -1$, b возрастает. Для значения b, лежащего вне области формы сборки, существует один минимум, и для малого D $\langle x \rangle = x_c$ является достаточно хорошей аппроксимацией. Для значений b, расположенных внутри критической области потенциала V, имеются три критических значения, но только глубочайший из двух минимумов является для нас важным при $|b| \gtrsim 10D$. Когда b проходит через значение нуль, глубочайший минимум соответствует уже отрицательному, а не положительному корню уравнения dV/dx. Следует быстрый переход $\langle x \rangle$ от положительного к отрицательному значению в соответствии с вышерассмотренным случаем 2. Этот быстрый переход тесно связан с очень большой аномалией дисперсии.

Путь 4: a убывает, $b = 0$. При $a > 0$ потенциал V имеет минимум в точке $x = 0$, и $\langle x \rangle = 0$, $\langle x^2 \rangle = D/2a$. Когда $a \to 0^+$, дисперсия возрастает и квадратичный член в усечении дисперсии становится все более важным, что и было описано выше. В точке $(a, b) = (0, 0)$ значение дисперсии дается формулой (9.21). При $a < 0$ потенциал V имеет два симметрично расположенных минимума: $x = -\sqrt{-a}$ и $x = +\sqrt{-a}$. В силу симметрии $\langle x \rangle = 0$, и приблизительное значение дисперсии дается величиной $(2\sqrt{-a}/2)^2 = |a|$ в соответствии с формулой (9.19).

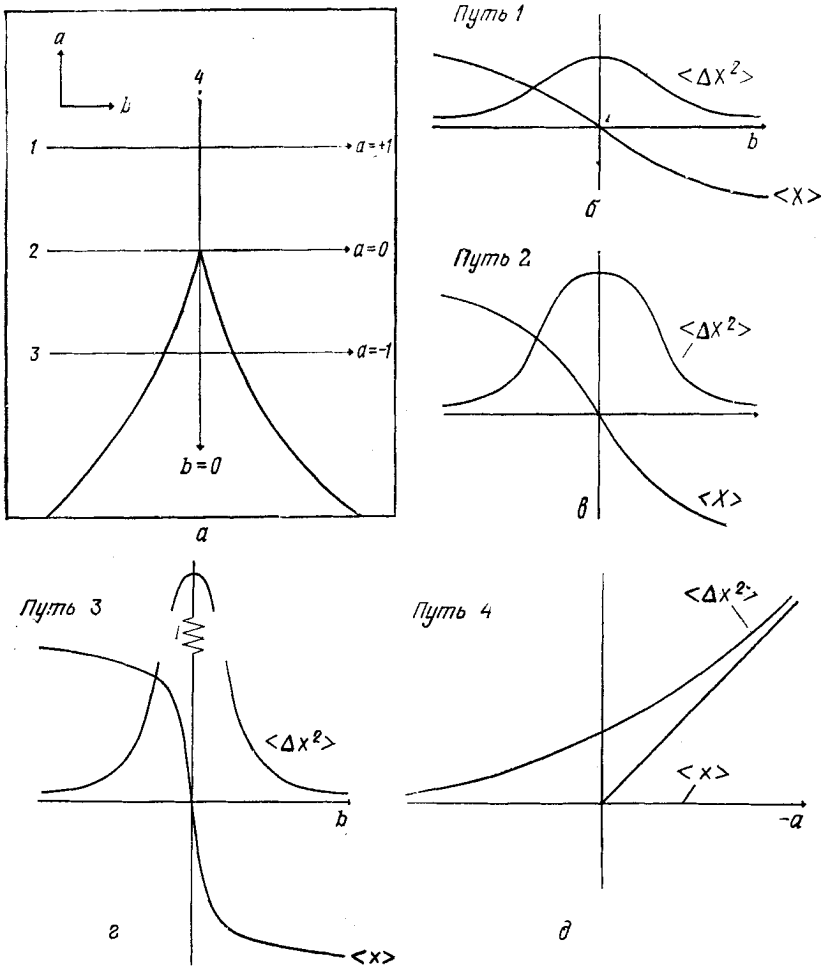


Рис. 9.8. Среднее значение $\langle x \rangle$ и дисперсия $\langle \Delta x^2 \rangle$ как функции расстояний вдоль четырех путей в плоскости $\mathbb{R}^2$ управляющих параметров катастрофы сборки. Среднее значение берется относительно не зависящего от времени распределения вероятностей (9.20).

◇◇◇ На протяжении всего этого раздела мы игнорировали вопрос о шкале времени, предположив существование асимптотической не зависящей от времени функции распределения вероятностей. Однако подобное предположение оказывается не всегда обоснованным, и поэтому необходимо соответствующим образом модифицировать описание, приведенное в случаях *1* (9.6) и *2* (9.9). Если функция распределения вероятностей

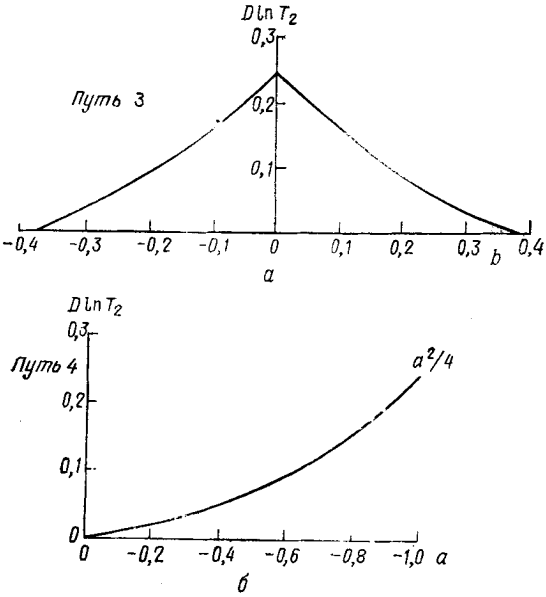


Рис. 9.9. Туннельная шкала времени T_2 для бимодальных путей *3* и *4*, изображенных на рис. 9.8.

рассматриваемой системы подчиняется уравнению диффузионного типа, то дисперсия типа (9.16) может наблюдаться во временной шкале T_1, а дисперсия типа (9.19) — во временной шкале T_2 [см. (8.31)]. Временная шкала T_1 изображена на рис. 9.6 для путей *1*, *2* и *3*. Временная шкала T_2 для состояния равновесия между двумя изолированными минимумами существует лишь в случае путей *3* и *4* (рис. 9.9). Наименьшая аномальная дисперсия (9.16) наблюдается на быстрой временной шкале T_1, наибольшая (9.19) — на более медленной временной шкале T_2. Довольно часто временная шкала, которая используется в экспериментах, лежит между T_1 и T_2. В этом случае дисперсия, ассоциированная с любым одиночным экспериментом, описывается формулой (9.16). Для того чтобы выявить большие аномалии дисперсии типа (9.19), необходимо усреднить результаты многих экспериментов. Иллюстрируем это посредством рассмотрения цикличного пути ($a = -1$, $b(t) = -b\cos\omega t$) в пространстве управляющих параметров (рис. 9.10). Если $T_1^{-1} \gg \omega \gg T_2^{-1}$, то система движется по петле гистерезиса, как это изображено на рис. 9.4 и 9.10. Среднее значение $\bar{x} = \langle x \rangle$ и дисперсия $\langle (x - \bar{x})^2 \rangle$ как функции времени вдоль одного цикла $\tau = 2\pi/\omega$ представлены на рис. 9.10. Можно

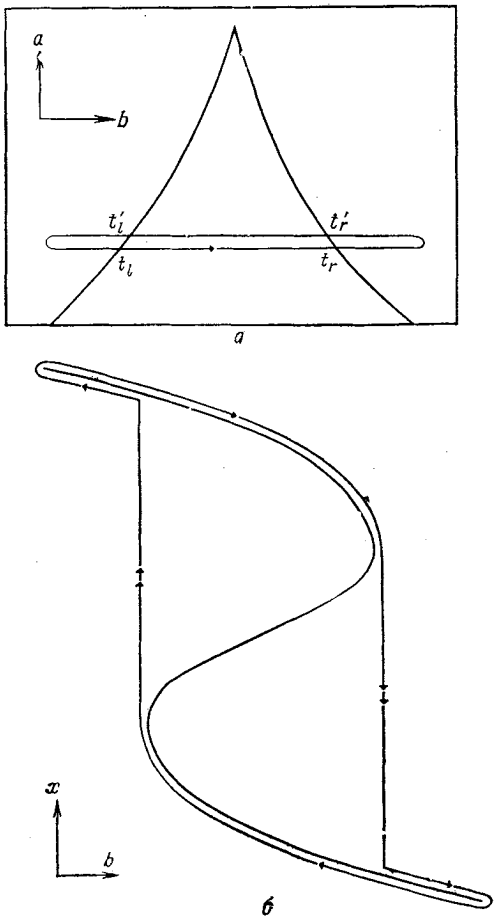
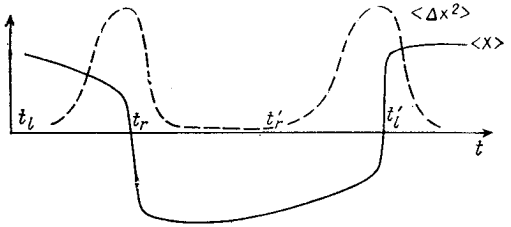
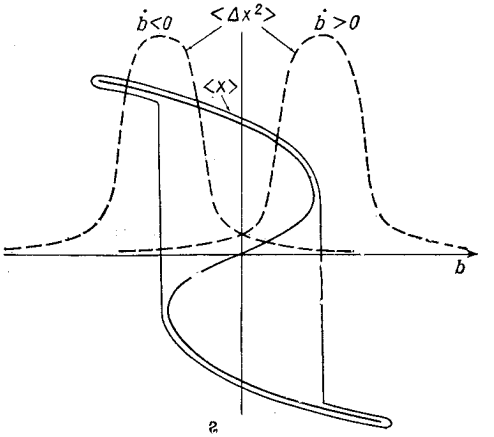


Рис. 9.10. На рисунке проиллюстрировано поведение среднего значения и дисперсии на временной шкале, в которой проводится эксперимент, и метод, используемый для обработки данных, именно в этой последовательности.

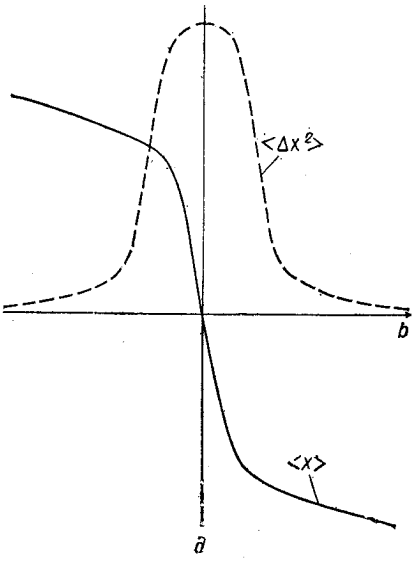
а — циклический путь, проходимый в пространстве управляющих параметров катастрофы сборки. Экспериментальная временная шкала ω^{-1} подчиняется условию $T_2 \gg \omega^{-1} \gg T_1$; *б* — система, описывающая петлю гистерезиса; *в* — изменение периодического среднего значения $\langle x \rangle$ и дисперсии $\langle \Delta x^2 \rangle$ с частотой $\omega/2\pi$; *г* — если данные отобраны в соответствии с тем, является ли *b* возрастающим или убывающим, то аномальная дисперсия имеет изображенный на рисунке вид. Для возрастающего *b* дисперсия аномально большая вблизи правосторонней линии складки. Когда же *b* убывает, дисперсия становится аномально большой вблизи левосторонней компоненты бифуркационного множества; *д* — если не делать различия между взятыми данными с $\dot{b} > 0$ и $\dot{b} < 0$, то $\langle x \rangle$ не является больше бимодальной, а дисперсия вокруг множества Максвелла становится огромной. Эти всплески чрезмерной дисперсии выравниваются аномально большой дисперсией, связанной с бифуркационным множеством.



в



г



д

изобразить их графически как функции управляющего параметра b при условии, что мы различаем пути с возрастающим b и пути с убывающим b (рис. 9.10, b). Аномальная дисперсия типа (9.18) встречается лишь на линии складки, и в этом случае она связана с временно́й шкалой T_1. Однако если мы вычислим среднее и дисперсию посредством усреднения результатов всех экспериментов, где каждый единичный эксперимент проводится в интервале времени длины π/ω с $\dot{f} \geqslant 0$ или $\dot{f} \leqslant 0$, то в этом случае мы будем иметь аномальную дисперсию, свя-

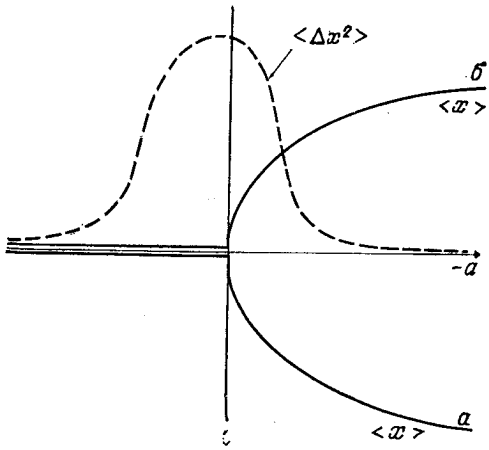


Рис. 9.11. В обычных лабораторных условиях физическая система, следуя по пути *4* (рис. 9.8), останавливается либо в левостороннем, либо в правостороннем минимуме.

занную со временной шкалой T_2, которая встречается в бифуркационном множестве в предположении принципа Максвелла. ◇◇◇ Функции распределения при применении принципа максимального промедления проявляют феномен критического замедления при приближении к бифуркационному множеству. Это следует из анализа (8.9).

◇◇◇ Когда путь 4 (рис. 9.8) наблюдается в лабораторном эксперименте, обычно происходит следующее. При $0 < a$ состояние системы определяет либо левосторонний, либо правосторонний минимум. Связанная с этим функция распределения вероятностей равна $P_l(x) \simeq Ne^{-V_l(x)/D}$ или $P_r(x) \simeq Ne^{-V_r(x)/D}$, где $V_l(x) = |a|(x + \sqrt{-a})^2$, а $V_r(x) = |a|(x - \sqrt{-a})^2$. Соответствующие среднее и дисперсия изображены на рис. 9.11. Как только один из двух минимумов выбран, туннельное время T_2 до другого минимума оказывается слишком большим по

сравнению с лабораторной шкалой времени. Если усреднить по различным экспериментам, то среднее и дисперсия ведут себя так, как это изображено на рис. 9.8. Это явление наблюдается, например, когда образец из железа охлаждается ниже температуры Кюри. Между повторяющимися экспериментами намагничивание готово идти «вверх» так же часто, как и «вниз». Дисперсия среднего «вверх» мала, что также обычно и для направления «вниз». Если результаты всех экспериментов усредняются, то среднее намагничивание равно нулю, а дисперсия громадна. Потеря точности в направлении намагничивания сама по себе может служить примером явления расходимости (разд. 4).

◇◇◇ В классической термодинамике аномалия дисперсии, связанная с ростком катастрофы складки, называется *критической опалесценцией.*

9. ВЫВОДЫ

Распознавание присутствия и типа катастрофы часто является важной предпосылкой для правильного описания физической системы. Такое исследование может быть использовано либо для упрощения работы (так как катастрофы являются каноническими) — это на одном конце спектра возможностей, либо для получения качественной или даже количественной информации, когда даже не существует математического описания (т. е. уравнений), — это на другом конце спектра. Мы обсудили некоторые характерные черты, которые могут встречаться в физических системах, наводя на мысль о присутствии катастроф.

Первые пять (модальность, недостижимость, катастрофические скачки, расходимость и гистерезис) обычно встречаются в совокупности. Они зависят от достижимости физической системой области пространства управляющих переменных, в которой потенциал имеет более чем один локальный минимум. Явление гистерезиса может быть не наблюдаемо, если поведение системы подчиняется принципу Максвелла, однако даже в этом случае иногда возможно наблюдать его (сверхохлаждение, сверхнагревание) с помощью тщательных экспериментальных методов.

Остальные три флага катастроф (расходимость линейного отклика, критическое замедление/смягчение моды и аномальная дисперсия) могут наблюдаться даже тогда, когда потенциал имеет всего лишь один локальный минимум. Последнее обстоятельство может быть использовано для установления как критических значений, так и безопасных пределов управляющих

параметров в многочисленных случаях, когда неожиданные «катастрофические скачки» могут оказаться гибельными для системы.

◇◇◇ «Отпечатки пальцев» катастроф впервые были систематизированы Зиманом [1].

Литература

1. Zeeman E. S., Catastrophe Theory, *Sci. American*, **234** (4), 65—83 (1976). Published in original from in: E. C. Zeeman, Catastrophe Theory, Selected Papers, 1972—1977, Reading: Addison — Wesley, 1977, p. 18.

Часть II.

ПРИЛОЖЕНИЯ ЭЛЕМЕНТАРНОЙ ТЕОРИИ КАТАСТРОФ В НАУКЕ И ТЕХНИКЕ

10

ТЕРМОДИНАМИКА

Классическая теория фазовых переходов (или фазовых превращений) естественным образом укладывается в рамки элементарной теории катастроф. Формулировка основных положений классической термодинамики в терминах теории катастроф позволяет рассматривать уравнение состояния вещества как n-мерное многообразие в $2n$-мерном пространстве $\mathbb{R}^n \otimes \mathbb{R}^n$ n интенсивных и n экстенсивных термодинамических переменных. Это многообразие тогда совпадает с критическим многообразием семейства потенциальных функций $\mathcal{U}$, зависящих от n управляющих параметров (экстенсивные термодинамические переменные) и n переменных состояния (интенсивные термодинамические переменные). Подобный подход приводит к очевидной геометрической интерпретации термодинамики равновесных систем: критическое многообразие является римановой поверхностью. Вычисление частных производных термодинамических величин существенно упрощается в результате анализа тензоров восприимчивости в n-мерном пространстве, касательном к критическому многообразию в $\mathbb{R}^n \otimes \mathbb{R}^n$. Кроме того, при такой формулировке напрашивается переход от термодинамики равновесных систем к термодинамике неравновесных систем по крайней мере в окрестности критического многообразия.

1. ОБЩЕЕ ОПИСАНИЕ ФАЗОВЫХ ПЕРЕХОДОВ

Для описания физической системы введем общее семейство потенциальных функций $V(x; c)$, зависящих от n переменных состояния или параметров порядка $x \in \mathbb{R}^n$ и k управляющих параметров $c \in \mathbb{R}^k$. Предположим далее, что состояние физической системы описывается значением x, минимизирующим потенциальную функцию, по крайней мере, локально. Тогда изучение такой физической системы сводится к изучению равновесия и локальной устойчивости потенциальной функции $V(x; c)$:

$$\begin{aligned} &\frac{\partial V}{\partial x_i} = 0 \text{ — равновесие,} \\ &\frac{\partial^2 V}{\partial x_i \partial x_j} > 0 \text{ — локальная устойчивость,} \end{aligned} \tag{10.1}$$

и критических значений на ветвях устойчивого равновесия.

Вообще говоря, почти для всех $c \in \mathbb{R}^k$ потенциальная функция $V(x; c)$ будет иметь только изолированные критические

точки $x^{(1)}, \ldots, x^{(p)}, \ldots$. Первое из уравнений (10.1) состоит из n уравнений состояния и может быть использовано как для определения положения изолированных критических точек, так и для выявления их зависимости от управляющих параметров c [см. (5.2)]:

$$x^{(p)} = x^{(p)}(c), \quad p = 1, \ldots . \tag{10.2}$$

Локальная устойчивость p-й критической точки определяется видом матрицы устойчивости в точке $(x^{(p)}(c); c)$. Фазовый переход происходит в тот момент, когда точка $x \in \mathbb{R}^n$, описывающая состояние физической системы, перескакивает с одной критической ветви на другую.

Фазовые переходы могут происходить при изменении значений управляющих параметров. Обычно предполагается (или это действительно имеет место), что управляющие параметры зависят всего лишь от одного параметра (например, времени), и именно этот параметр используется для описания кривой фазового перехода (кривой равновесия) в пространстве управляющих параметров:

$$\begin{gathered} c_\alpha \to c_\alpha(s), \quad s \in \mathbb{R}^1, \\ x^{(p)}(c_\alpha) \to x^{(p)}(c_\alpha(s)) \to x^{(p)}(s), \\ V^{(p)}(x^p(c); c) \to V^{(p)}\left(x_i^{(p)}(s); c_\alpha(s)\right) \to V^{(p)}(s). \end{gathered} \tag{10.3}$$

Фазовый переход имеет место, когда кривая $c_\alpha(s) \in \mathbb{R}^k$ пересекает соответствующую компоненту соответствующей сепаратрисы в $\mathbb{R}^k$. Если придерживаться принципа максимального промедления, то фазовые переходы произойдут, когда данная кривая пересечет компоненту бифуркационного множества $\mathscr{S}_B$, на которой возникают и исчезают локальные минимумы. Если же придерживаться принципа Максвелла, то можно утверждать, что фазовые переходы будут иметь место при пересечении кривой равновесия компоненты максвелловского множества $\mathscr{S}_M$, на которой не менее двух глобальных минимумов вырождены.

Часто оказывается удобным описывать фазовые переходы в соответствии с классификацией Эренфеста. Предположим, что фазовый переход происходит из-за того, что состояние системы перескакивает с p-й на q-ю критическую ветвь при $c_\alpha(s^0)$, когда s, возрастая, достигает значения s^0. В этом случае имеет место переход m-го рода, если

$$\lim_{\varepsilon \to 0} \frac{d^i}{ds^i} V^{(p)}(s)\Big|_{s^0-\varepsilon} = \lim_{\varepsilon \to 0} \frac{d^i}{ds^i} V^{(q)}(s)\Big|_{s^0+\varepsilon}, \quad i = 0, 1, \ldots, m-1, \tag{10.4}$$

и это равенство не выполняется при $i = m$. Фазовый переход называют *локальным*, или *мягким*, если

$$\lim_{\varepsilon \to 0} [x^{(p)}(s^0 - \varepsilon) - x^{(q)}(s^0 + \varepsilon)] = 0. \qquad (10.5)$$

В противном случае переход называют *нелокальным*, или *жестким*. Обычно в физических явлениях, наблюдаемых в природе, фазовые переходы имеют нулевой, первый или второй род. Если это не так, то они не укладываются в схему классификации Эренфеста.

2. ФАЗОВЫЕ ПЕРЕХОДЫ ВТОРОГО РОДА

Для описания фазовых переходов второго рода чрезвычайно полезным инструментом является модель Гинзбурга — Ландау. Суть этой модели состоит в следующем. Для описания состояния физической системы вводится параметр порядка x. Полная энергия системы представляется в виде суммы членов, описывающих кинетическую ($\sim (dx/dt)^2$) и потенциальную $V(x)$ энергию. Последняя зависит как от внешних управляющих параметров, так и от переменной состояния x. Состояние покоя системы достигается в результате минимизации суммы этих двух членов. В статическом случае ($dx/dt = 0$) состояние системы определяется величиной x, минимизирующей $V(x)$. Часто либо система подчиняется (либо для нее постулируется) некоторому принципу симметрии, сужающему возможный вид потенциальных функций до класса четных функций от x. Потенциал Гинзбурга — Ландау

$$V(x; a) = \frac{1}{4} x^4 + \frac{1}{2} ax^2 \qquad (10.6а)$$

имеет единственный минимум в точке $x = 0$ при $a \geqslant 0$ и два равных минимума в $x = \pm \sqrt{-a}$ при $a < 0$ (рис. 10.1). Фазовый переход второго рода происходит при переходе a через нуль.

Модель Гинзбурга — Ландау часто используют для термодинамического описания фазовых переходов второго рода. В этом случае управляющий параметр a в окрестности критической температуры T_c идентифицируется с $T - T_c$, а бифуркационная фазовая диаграмма представляет собой стандартный трезубец, изображенный на рис. 10.2.

С экспериментальной точки зрения в поведении фазовых переходов второго рода наблюдаются такие «особенности», которые отнюдь не следуют с очевидностью из модели (10.6а). Так, например, фазовый переход второго рода попросту исчезает при произвольно малых возмущениях, нарушающих симметрию, или

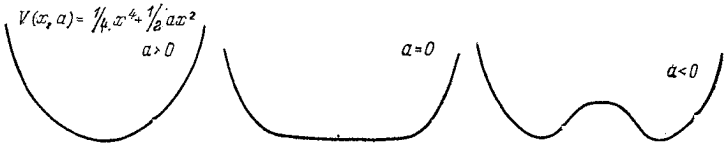


Рис. 10.1. Потенциал $V(x; a)$ имеет единственный минимум в точке $x = a$ при $a \geqslant 0$ и два равных минимума в точке $x = \pm(-a)^{1/2}$ при $a \leqslant 0$.

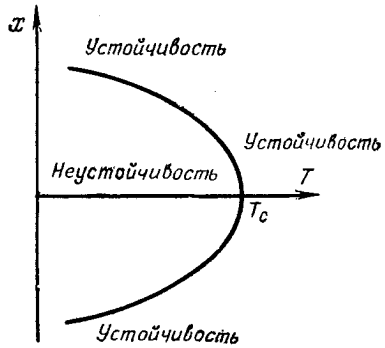


Рис. 10.2. Сечение многообразия катастрофы сборки плоскостью симметрии $b = 0$ имеет вид трезубца.
При температурах ниже критической (T_c) наблюдаются два устойчивых состояния, а выше — только одно.

может вновь возникнуть в какой-либо отдаленной точке в виде фазового перехода нулевого или первого рода. Эта особенность поведения связана со структурной неустойчивостью фазовых переходов второго рода.

Для более подробного изучения модели Гинзбурга — Ландау (10.6а), включающего анализ влияния членов, нарушающих симметрию, необходимо исследовать наиболее общий вид возмущения потенциальной функции, т. е. катастрофу сборки

$$V(x; a, b) = \frac{1}{4}x^4 + \frac{1}{2}ax^2 + bx. \qquad (10.6б)$$

Проиллюстрируем общие понятия, описанные в разд. 1, на трех примерах с катастрофой сборки.

***Пример* 0.** Будем придерживаться принципа максимального промедления и проследуем по кривой 0 (рис. 10.3). В данном случае существенной является сепаратриса, которая представляет собой бифуркационное множество $\mathcal{B}$, состоящее из двух линий складок. Фазовый переход происходит в точке $\times$, где переменная состояния системы перескакивает с исчезающего правого минимума в $x^{(r)} = \sqrt{-a/3}$ на глобальный левый минимум в $x^{(l)} = -2\sqrt{-a/3}$. Критические значения для правой и левой ветвей равны

$$V^{(r)} = a^2/12, \quad V^{(l)} = -2a^2/3. \qquad (10.7)$$

Поскольку в точке $\times$ изменение значения потенциала отлично от нуля:

$$\Delta V\,|_{\delta_B} = V^{(l)}V^{(r)} = -\frac{3}{4}\,a^2 \tag{10.8}$$

мы имеем дело с фазовым переходом нулевого рода. Величина скачка зависит только от положения точки, в которой кривая 0 пересекает линию складки, и не зависит от направления пути в этой точке.

Пример 1. На этот раз будем придерживаться принципа Максвелла и проследуем по кривой 1 (рис. 10.3). В данном случае существенной является сепаратриса, которая представляет собой максвелловское множество $\mathcal{P}_M$, состоящее из полупрямой ($a < 0$, $b = 0$). Фазовый переход происходит в точке $\times$, где переменная состояния системы перескакивает с правого минимума в

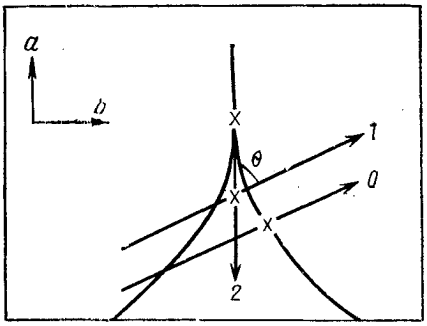


Рис. 10.3. Порядок фазового перехода зависит от принятого принципа и пути следования.

Переход нулевого рода происходит на пути 0, если справедлив принцип максимального промедления, переход первого рода происходит на пути *1* в случае справедливости принципа Максвелла, переход второго рода происходит на пути 2. Значком $\times$ обозначена точка разрыва.

$x^{(r)} = \sqrt{-a}$ на левый минимум в $x^{(l)} = -\sqrt{-a}$. Потенциальная функция в этих точках имеет равные численные значения $V^{(r)} = V^{(l)} = -a^2/4$, однако первая производная вдоль кривой 1 терпит разрыв на максвелловском множестве. Величину скачка можно определить, вычислив значение

$$\frac{d}{ds}\,V(x;\,a,\,b) = \frac{\partial V}{\partial x}\,\frac{dx}{ds} + \frac{\partial V}{\partial a}\,\frac{da}{ds} + \frac{\partial V}{\partial b}\,\frac{db}{ds}, \tag{10.9}$$

на максвелловском множестве. Производные da/ds, db/ds можно интерпретировать, как направляющие косинусы, а $db/da = (db/ds)/(da/ds) = \operatorname{tg}\theta$, как показано на рис. 10.3. Поскольку $dV/dx = 0$,

$$\Delta\,\frac{dV}{ds}\bigg|_{\delta_M} = \left(\frac{1}{2}\,x^2\cos\theta + x\sin\theta\right)^{(l)} - \left(\frac{1}{2}\,x^2\cos\theta + x\sin\theta\right)^{(r)} =$$

$$= -2\sqrt{-a}\,\sin\theta. \tag{10.10}$$

Эта функция описывает фазовый переход первого рода, так как ее первая производная претерпевает разрыв непрерывности, зависящий не только от положения точки фазового перехода, но и от направления кривой равновесия при ее пересечении максвелловского множества.

Пример 2. Рассмотрим теперь кривую *2* на рис. 10.3 (точка фазового перехода помечена знаком $\times$). Вдоль этой кривой $s = 0$ и

$$V(s) = 0, \quad s \geqslant 0; \quad V(s) = -s^2/4, \quad s \leqslant 0. \tag{10.11}$$

При $s = 0$ как сам потенциал, так и его первая производная непрерывны, однако вторая производная терпит разрыв, поскольку

$$\Delta \frac{d^2V}{ds^2} = \frac{d^2V^{(1)}}{ds^2} - \frac{d^2V^{(0)}}{ds^2} = -\frac{1}{2}, \tag{10.12}$$

где верхний индекс (1) указывает, что значение потенциала берется на любой из ветвей с ненулевым решением ($s < 0$), а индекс (0) обозначает нулевую ветвь ($s > 0$) при $s \to 0^{\pm}$. Этот фазовый переход вдоль кривой *2* является переходом второго рода.

◇◇◇ Фазовый переход второго рода является локальным, в то время как фазовые переходы нулевого и первого рода нелокальны.

◇◇◇ Фазовые переходы, которые не являются переходами ни нулевого, ни первого и ни второго рода, невозможно описать типичными путями в пространстве управляющих параметров катастрофы сборки. Их можно описать путями в пространстве управляющих параметров катастроф более высокой размерности.

◇◇◇ Не всякий фазовый переход точно укладывается в схему классификации Эренфеста.

◇◇◇ При соблюдении принципа Максвелла бифуркационное множество катастрофы сборки состоит из полупрямой ($a < 0$, $b = 0$), соответствующей фазовым переходам первого рода, и граничной точки ($a = 0$, $b = 0$), соответствующей фазовому переходу второго рода.

◇◇◇ Неочевидное поведение фазовых переходов второго рода при наличии возмущений, отмеченное в начале данного раздела, имеет совершенно естественную интерпретацию в рамках элементарной теории катастроф. Потенциал Гинзбурга — Ландау (10.6а) является наиболее общим однопараметрическим семейством четных функций, содержащих неморсовский росток. Для этого семейства могут иметь место только фазовые переходы второго рода. При нарушении симметрии возможно появление возмущений более общего вида, причем для потенциала Гинзбурга — Ландау (10.6а) таким возмущением является катастрофа сборки (10.6б). При произвольно малых возмущениях, нарушающих симметрию, путь в пространстве управляющих параметров не проходит через точку сборки $(a, b) = (0, 0)$, поэтому фазовый переход второго рода исчезает. Если, однако, принят принцип Максвелла и возмущенный путь пересекает максвелловское множество ($a < 0$, $b = 0$) в точке, вообще говоря, не слишком близкой к $a = 0$, то возникает фазовый переход первого рода. Если же принят принцип максимального промедле-

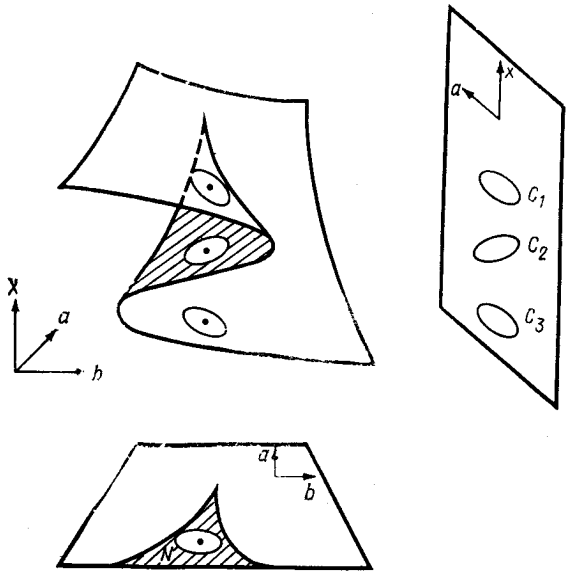


Рис. 10.4. Одной точке в пространстве $\mathbb{R}^4$ управляющих параметров могут соответствовать несколько критических точек.

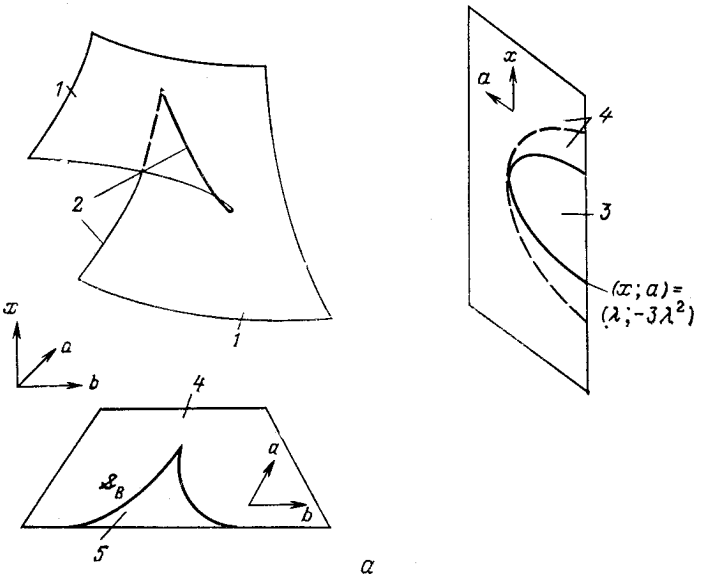
Окрестность каждой изолированной критической точки, принадлежащей k-мерному критическому многообразию, можно параметризовать точками из окрестности $N \in \mathbb{k}^k$ точки соответствующего управляющего параметра или из окрестностей C_1, C_2, C_3, ... в некотором k-мерном подпространстве из $\mathbb{R}^n \otimes \mathbb{R}^k$.

ния и путь пересекает соответствующую линию складки [опять же, вообще говоря, далекую от $(a = 0, b = 0)$], то наблюдается фазовый переход нулевого рода.

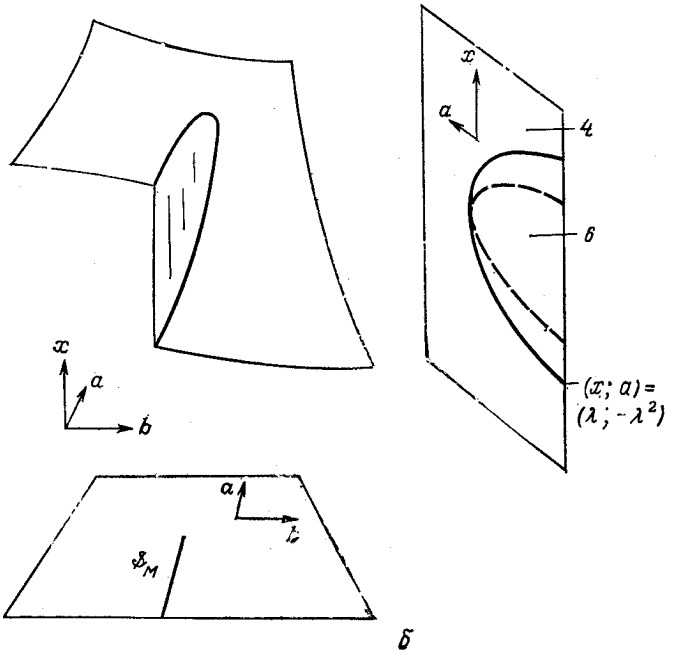
3. ТОПОЛОГИЧЕСКИЕ СООБРАЖЕНИЯ

Критическое многообразие $\nabla V = 0$ может содержать несколько критических точек $x^{(p)}$, $p = 1, 2, \ldots$ над фиксированной точкой c^0 в пространстве управляющих параметров (рис. 10.4). Поскольку критическое многообразие k-мерно, удобно, но совсем не обязательно использовать k управляющих параметров для параметризации k-мерной окрестности каждой критической точки. Такая параметризация становится несостоятельной, когда критическая точка подходит к бифуркационному множеству.

Возможны иные способы параметризации, предусматривающие использование одной или нескольких переменных состояния. Эта идея иллюстрируется на рис. 10.4. Здесь имеются три



а



б

различные критические точки над точкой (a^0, b^0), лежащей в плоскости управляющих параметров. Для точек, лежащих в окрестностях этих трех изолированных критических точек, можно использовать координаты соответствующих точек, лежащих в окрестности точки (a^0, b^0) в плоскости управляющих параметров. В некоторых случаях более удобной является параметризация через переменную состояния x и один (или два) управляющий параметр (управляющих параметров). Параметризация через пару $(x; a)$ показана на рис. 10.4. Такая параметризация удобна, поскольку имеется глобальное однозначное соответствие между точками на критическом многообразии и точками плоскости (x, y).

В тех случаях, когда состояние системы определяется минимальным значением потенциала, вид устойчивого критического многообразия зависит от того, какого принципа мы придерживаемся. На рис. 10.5, *а* показано многообразие устойчивых и метастабильных минимумов в случае реализации принципа максимального промедления. Это «урезанное многообразие», поскольку часть критического многообразия, описывающая критические точки, не являющиеся локальными минимумами, была удалена. Границей этого многообразия является кривая $(x; a, b) = (\lambda; -3\lambda^2, 2\lambda^3)$, которая проектируется на линии складок в плоскости управляющих параметров и на параболу $(x; a) = (\lambda; -3\lambda^2)$ в плоскости (x, y).

Если принят принцип Максвелла, критическое многообразие является «спаянным многообразием», поскольку его часть $\nabla V = 0$, описывающая неустойчивые критические точки и метастабильные минимумы, удалена и заменена плоским участком, интерполирующим это многообразие в промежутке между двумя минимумами. Можно дать следующую физическую интерпретацию точки, лежащей на плоском участке. Если $x^{(l)}$ и $x^{(r)}$ — положения двух равновеликих минимумов при $a < 0$, $b = 0$, то x $(x^{(l)} \leqslant x \leqslant x^{(r)})$ определяет, какая часть системы пребывает в состоянии, соответствующем минимуму в $x^{(l)}$ (скажем, в жидкой фазе), и какая часть может быть отнесена к минимуму в $x^{(r)}$, а

Рис. 10.5. Подмножество критического многообразия, отвечающее устойчивым состояниям, не является многообразием в обычном смысле ни в условиях принципа максимального промедления, ни в условиях принципа Максвелла.

а — при соблюдении принципа максимального промедления неустойчивая поверхность многообразия катастрофы сборки исчезает. Граница проецируется на кривую складок в плоскости (a, b) и на параболу $a = -3x^2$ в плоскости $(x; a)$; *б* — форма поверхности уравнения состояния при соблюдении принципа Максвелла. Плоский участок этой поверхности проецируется на максвелловское множество $\mathcal{S}_M$ в плоскости управляющих параметров (a, b), а его граница — на параболу $a = -x^2$ в плоскости $(x; a)$.

именно:

$$\begin{aligned} &\frac{x - x^{(l)}}{x^{(r)} - x^{(l)}} \text{ — часть системы, находящаяся в состоянии, отвечающем минимуму в } x^{(r)}; \\ &\frac{x^{(r)} - x}{x^{(r)} - x^{(l)}} \text{ — часть системы, находящаяся в состоянии, отвечающем минимуму в } x^{(l)}. \end{aligned} \tag{10.13}$$

Таким образом, значение переменной x описывает смесь двух чистых состояний $x^{(l)}$ и $x^{(r)}$, которые могут существовать при $a < 0$, $b = 0$; при $x^{(l)} < x < x^{(r)}$ оба этих состояния могут сосуществовать. Очевидно, что одних управляющих параметров a, b недостаточно для построения системы координат в окрестности точки, лежащей на плоском участке критического многообразия. В этом случае для однозначной параметризации критического многообразия необходимо использовать комбинацию координат и управляющих параметров (рис. 10.5, *б*).

Максвелловское множество в плоскости управляющих параметров представляет собой полупрямую ($a < 0$, $b = 0$). Вообще говоря, максвелловское множество $\mathscr{P}_M$ определяется уравнением Клаузиуса — Клапейрона (5.68) и представляет собой многообразие пространств управляющих параметров различной размерности из $\mathbb{R}^k$. В том случае, когда критическое многообразие единственным образом параметризовано через комбинацию переменных состояния и управляющих параметров (рис. 10.5, *б*), максвелловское множество в этом пространстве пусто. В подобной ситуации границами областей в параметризующем пространстве, в которых одновременно могут существовать различные фазы, являются подмногообразия, такие, как парабола, изображенная на рис. 10.5, *б*.

Линейный отклик системы на изменения значений управляющих параметров описывается уравнением (5.2). Инфинитезимальной теории линейного отклика (5.2) может быть дано топологическое толкование, если ввести понятие касательного пространства к k-мерному критическому многообразию в точке $(x^0; c^0)$ из $\mathbb{R}^n \otimes \mathbb{R}^k$. Это касательное пространство является k-мерным линейным векторным пространством. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — базисные векторы в $\mathbb{R}^n$ с центром в начале координат и $\mathbf{f}_1, \ldots, \mathbf{f}_k$ — базисные векторы в $\mathbb{R}^k$ также с центром в начале координат. Тогда базисные векторы в касательном пространстве можно определить из функции линейного отклика (5.2). Функция линейного отклика представляет собой тензор восприимчивости, поэтому (5.2) можно переписать в виде

$$\delta x^i = \chi^i_a \, \delta c^a, \quad \chi^i_a = -(V^{-1})^{ij} V_{ja}. \tag{10.14}$$

Тензор восприимчивости χ_α^i определен на критическом многообразии.

Для определения набора k базисных векторов для касательного пространства перенесем базисные векторы $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ и $(\mathbf{f}_1, \ldots, \mathbf{f}_k)$ из начала координат в $(x^0; c^0) \in \mathbb{R}^n \otimes \mathbb{R}^k$ и образуем линейную комбинацию этих базисных векторов, лежащую в касательном пространстве. Поскольку бесконечно малый вектор δv, определяемый как

$$\delta v = \delta x^i \mathbf{e}_i + \delta c^\alpha \mathbf{f}_\alpha, \tag{10.15}$$

лежит в этом пространстве, если δx^i и δc^α связаны соотношением (10.14), в качестве k линейно независимых базисных векторов можно выбрать векторы

$$|f_\alpha\rangle = \mathbf{e}_i \chi_\alpha^i + \mathbf{f}_\alpha, \quad \alpha = 1, 2, \ldots, k. \tag{10.16}$$

Такой выбор особенно удобен, если для параметризации точек, лежащих в окрестности критического многообразия, используются управляющие параметры $c \in \mathbb{R}^k$. Если же в этом участвуют и координаты (переменные состояния), то какой-либо иной выбор системы базовых векторов может оказаться более целесообразным.

Для определения полного набора базисных векторов пространства $\mathbb{R}^n \otimes \mathbb{R}^k$ в точке $(x^0; c^0)$ необходимо найти n дополнительных векторов, линейно независимых по отношению к векторам $|f_\alpha\rangle$, определенным в (10.16). Это довольно просто сделать, если ввести какую-то меру ортогональности, например скалярное произведение. Единственным действительным симметричным тензором второго рода является матрица вторых смешанных частных производных от V. Если определить скалярное произведение в точке $(x^0; c^0)$ как

$$(\mathbf{e}_i, \mathbf{e}_j) = V_{ij}, \quad (\mathbf{e}_i, \mathbf{f}_\beta) = V_{i\beta}, \quad (\mathbf{f}_\alpha, \mathbf{f}_\beta) = V_{\alpha\beta}, \tag{10.17}$$

то n векторов $|e_i\rangle = \mathbf{e}_i$, $i = 1, 2, \ldots, n$, будут ортогональны k векторам $|f_\alpha\rangle$, поскольку

$$\begin{aligned}\langle e_i | f_\alpha \rangle &= (\mathbf{e}_i, \mathbf{f}_\alpha + \mathbf{e}_j (-V^{-1})^{jm} V_{m\alpha}) = \\ &= V_{i\alpha} - V_{ij} (V^{-1})^{jm} V_{m\alpha} = 0.\end{aligned} \tag{10.18}$$

Короче, k векторов $|f_\alpha\rangle$ формируют систему «естественных» базисных векторов в касательном пространстве, а n векторов $|e_i\rangle$ образуют базис в его ортогональном дополнении в точке $(x^0; c^0)$. При таком выборе базисных векторов матрица метрики в $(x^0; c^0)$ принимает блочно-диагональный вид

$$G = \left[\begin{array}{c|c} V_{ij} & 0 \\ \hline 0 & V_{\alpha\beta} - V_{\alpha r} (V^{-1})^{rs} V_{s\beta} \end{array}\right]. \tag{10.19}$$

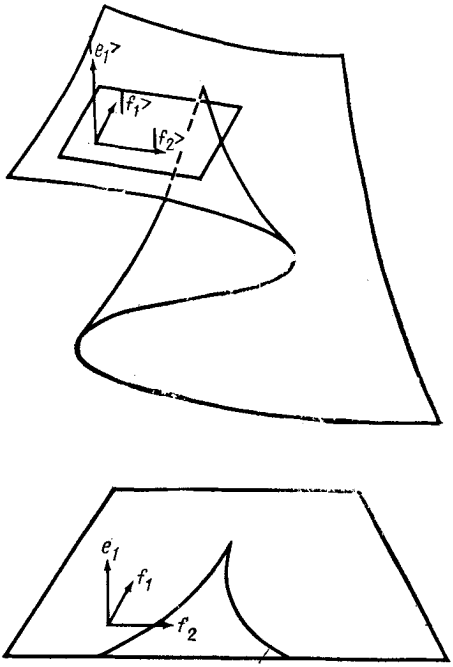


Рис. 10.6. Плоскость, касательная к многообразию катастрофы сборки, порождается базисными векторами $|f_1\rangle$ и $|f_2\rangle$, построенными из базисных векторов e_1, f_1, f_2 прямого произведения $\mathbb{R} \otimes \mathbb{R}^2$ пространств переменных состояния и управляющих параметров [см. (10.22)].

Однако $n + k$ векторов $|e_i\rangle$, $|f_\alpha\rangle$ не образуют полного базиса в $\mathbb{R}^n \otimes \mathbb{R}^k$ на бифуркационном множестве $\mathscr{S}_B$.

Пример. Для катастрофы сборки

$$V(x;\ a,\ b) = \frac{1}{4}x^4 + \frac{1}{2}ax^2 + bx,$$
$$V_{xx} = 3x^2 + a, \quad V_{xa} = x, \quad V_{xb} = 1. \tag{10.20}$$

Тензор восприимчивости в точке $(x;\ a,\ b)$, лежащей на критическом многообразии, определяется из соотношений

$$\delta x = -\frac{1}{3x^2 + a}(x,\ 1)\begin{pmatrix}\delta a\\ \delta b\end{pmatrix},$$
$$\chi_{1\alpha} = -\frac{1}{3x^2 + a}(x,\ 1), \quad \alpha = 1,\ 2. \tag{10.21}$$

Если f_1, f_2 и e_1 — базисные векторы в пространстве управляющих параметров и переменных состояния соответственно, то

$$|f_1\rangle = \mathrm{f}_1 - \frac{x}{3x^2 + a}\,\mathrm{e}_1,$$
$$|f_2\rangle = \mathrm{f}_2 - \frac{1}{3x^2 + a}\,\mathrm{e}_1 \tag{10.22}$$

образуют базис в касательной плоскости с центром в $(x;\ a,\ b)$ (рис. 10.6). В терминах смещений δa, δb метрический тензор в касательной плоскости есть

$$g_{\alpha\beta} = -\begin{pmatrix} x \\ 1 \end{pmatrix} \frac{1}{3x^2 + a}\,(x,\ 1). \tag{10.23}$$

Если для параметризации критического многообразия вместо управляющих параметров $(a,\ b)$ используется смешанная система $(x;\ a)$, то и в этом случае тензор восприимчивости определяется из преобразованного соотношения (5.2) $(V_{ij}\delta x^j + V_{i\alpha}\delta c^\alpha = 0)$:

$$(3x^2 + a)\,\delta x + x\,\delta a + \delta b = 0,$$
$$\delta b = -\,(x,\ 3x^2 + a)\begin{pmatrix} \delta a \\ \delta x \end{pmatrix}. \tag{10.24}$$

С помощью этого тензора восприимчивости можно определить базисные векторы в касательной плоскости и метрический тензор в новой системе координат.

4. КРИТИЧЕСКАЯ ТОЧКА ЖИДКОСТИ

Для многих простых веществ существуют точно определенные температуры сублимации, плавления и кипения, которые могут изменяться в зависимости от давления. Для таких веществ кривые равновесия фаз твердое тело — пар (сублимация), жидкость — твердое тело (плавление) и жидкость — пар (кипение) изображаются в виде фазовой диаграммы (рис. 10.7). Эти три кривые равновесия фаз определяются уравнением Клаузиуса — Клапейрона и являются кривыми фазовых переходов первого рода. Особый интерес на фазовой диаграмме представляют *тройная T* и *критическая C точки.* В тройной точке вещество может существовать в равновесии одновременно в твердом, жидком и газообразном состояниях; в критической точке различие между жидкой и газообразной фазами исчезает. Таким образом, точка C кривой равновесия фаз жидкость — газ соответствует фазовому переходу второго рода. (Говорят, что за точкой C вещество находится во «флуидном» состоянии.)

◇◇◇ Физическая критическая точка C соответствует математической вырожденной критической точке.

Рассмотрим критическую точку C, которая является граничной точкой на кривой фазовых переходов первого рода. Из сравнения рис. 10.7 и 10.3 с очевидностью вытекает, что 1) катастро-

фа сборки имеет место в окрестности точки C, 2) термодинамические переменные P, T связаны с математическими параметрами управления $a = a(P, T)$, $b = b(P, T)$ и 3) кривую равновесия фаз жидкость — газ можно отождествить с максвелловским множеством $\mathscr{S}_M$ катастрофы сборки ($a < 0$, $b = 0$). Если такое отождествление разумно, то получаем следующие признаки катастрофы:

— *модальность.* Вещество в текучем состоянии бимодально в окрестности кривой равновесия фаз жидкость — газ, при этом жидкая и газообразная фазы четко определены;

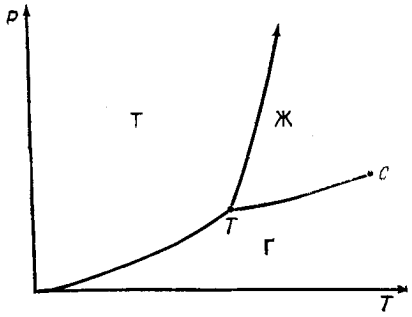


Рис. 10.7. Фазовая диаграмма простого однокомпонентного вещества. C — критическая точка, T — тройная точка.

— *недостижимость.* Этот признак катастрофы не имеет явного физического аналога;

— *внезапные скачки.* Небольшие изменения температуры и давления, в результате которых точка пространства управляющих параметров перемещается через кривую равновесия фаз, приводят к большим изменениям объема. Изменение объема происходит скачкообразно, если изменяются только переменные P и T. Если зафиксировать значения управляющих параметров для соответствующей кривой равновесия, то изменение объема, как функции полной внутренней энергии системы, будет плавным (непрерывным);

— *расходимость.* Если значения параметров P, T уменьшаются от некоторых величин, превосходящих P_c, T_c, до величин, меньших этих значений, конечным состоянием вещества может быть жидкость или газ в зависимости от слабых вариаций начальных условий. Вместе с тем, вещество можно перевести из жидкой фазы в газообразную либо скачкообразно, пересекая кривую равновесия, либо непрерывно путем обхода вокруг точки сборки C (рис. 10.8);

— *гистерезис.* Необратимый гистерезисный цикл может быть получен в результате тщательно выполненного эксперимента,

в процессе которого вначале повышают температуру, перегревая жидкость, а затем (если не возникает взрывоопасной смеси) понижают температуру, переохлаждая газ ниже точки конденсации. Если C — точка сборки, то кривая равновесия фаз будет окружена двумя линиями складок, за которыми система уже не является бимодальной. Эти линии определяют пределы перегрева и перенасыщения, они называются *спинодалями* (рис. 10.9);

— *расходимость линейного отклика.* Однокомпонентное вещество во «флуидном» состоянии имеет несколько функций линейного отклика.

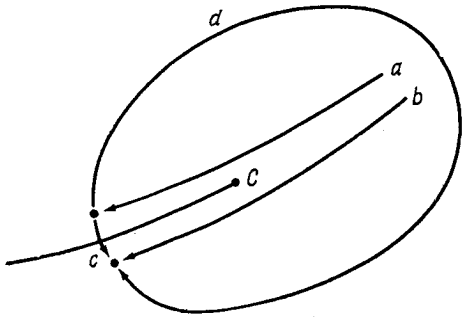


Рис. 10.8. Близкие пути из близких начальных состояний могут привести к качественно различным конечным состояниям (ср. с рис. 9.3).

Следуя по пути c, можно перевести систему из жидкого в газообразное состояние, при этом имеет место фазовый переход первого рода. На пути d фазовых переходов не происходит.

Одна из них — это изотермическая сжимаемость $k_T = -V^{-1}(\partial V/\partial P)_T$. Эта функция отклика неограниченно возрастает по мере приближения к кривой равновесия фаз (рис. 10.10);

— *критическое замедление.* Из соображений размерности скорость звука в жидкой среде зависит от $k^{-1/2}$, где k — сжимаемость (не обязательно изотермическая). С возрастанием k скорость звука стремится к нулю;

— *аномальные изменения.* При конденсации паров образуются мелкие капельки жидкости, оседающие на дне резервуара, содержащего данное вещество. Поскольку жидкость плотнее пара, при обратном переходе (кипении) небольшие пузырьки газа, образующиеся внутри жидкости, в силу тех же причин поднимаются вверх. При подходе к критической точке снизу разница в плотности двух фаз исчезает, время падения капелек и подъема пузырьков растет и размеры капелек и пузырьков возрастают до величины, при которой происходит очень сильное рассеяние света в данной среде. Из-за аномально больших ожидаемых значений радиусов пузырьков и капелек это явление называют *критической опалесценцией* (*мутностью*).

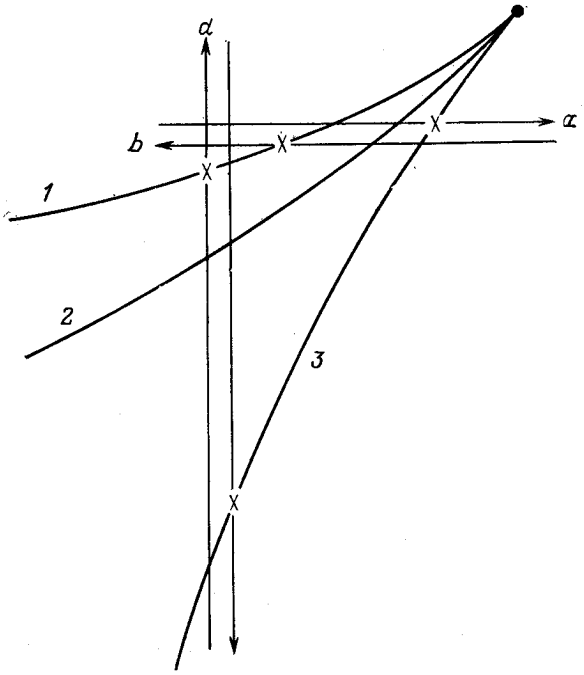


Рис. 10.9. Если критическая точка $C(\cdot)$ связана с катастрофой сборки, то она окружена кривыми складок — спинодалями.

1 — спинодаль при низкой температуре и высоком давлении; *2* — линия фазовых переходов первого рода; *3* — спинодаль при высокой температуре и низком давлении. При фиксированном давлении систему можно перегреть до температуры выше температуры кипения (определяемой $\mathscr{S}_M$, но не выше высокотемпературной спинодали (помеченной значком × на пути *a*). Вместе с тем, система не может находиться в состоянии перенасыщения ниже низкотемпературной спинодали (× на пути *b*). Спинодали $\mathscr{S}_B$ можно построить, варьируя P при фиксированном T или выбирая другие пути через кривые фазовых переходов первого рода.

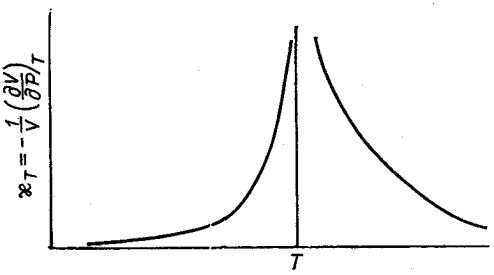


Рис. 10.10. Функции линейного отклика, такие, как изотермическая сжимаемость, терпят разрыв второго рода в точке на кривой равновесия.

5. УРАВНЕНИЕ СОСТОЯНИЯ ЖИДКОСТИ

Как было показано, уже простое сравнение фазовой диаграммы в окрестности критической точки с максвелловским множеством катастрофы сборки и последующий анализ типичных признаков катастрофы позволяют выявить существование многих физических процессов в реальных системах. И все-таки мы не можем довольствоваться только лишь качественным описанием систем. Поэтому попытаемся от качественного описания перейти к количественному описанию (уравнение состояния) системы (в частности, жидкости) в окрестности ее критической точки. Для этого необходимо прежде всего установить соответствие между математической переменной состояния x и соответствующим физическим параметром. Наиболее естественно на первый взгляд было бы идентифицировать x с объемом V. Однако это только на первый взгляд, поскольку при пересечении кривой равновесия жидкой и газообразной фаз величина V возрастает, а при пересечении в том же направлении максвелловского множества $\mathcal{S}_M$ катастрофы сборки критическое значение, доставляющее глобальный минимум, уменьшается, перескакивая с положительного значения ($b < 0$) на отрицательное ($b > 0$). Поэтому используем другую очевидную возможность и будем идентифицировать x с V^{-1}, т. е. с величиной, пропорциональной плотности ρ вещества.

Теперь нам предстоит преодолеть две технические трудности. Первая из них состоит в том, что математическая сборка имеет место при $(x; a, b) = (0; 0, 0)$, а физическая — при $(\rho_c; P_c, T_c)$. Вторая трудность связана с тем, что в отличие от физических $(\rho; P, T)$ математические $(x; a, b)$ параметры безразмерны. Обе эти трудности можно преодолеть, введя безразмерные масштабированные физические величины $(\rho_r; P_r, T_r)$, а именно:

$$\frac{\rho - \rho_c}{\rho_c} = \rho_r - 1, \quad \frac{P - P_c}{P_c} = P_r - 1, \quad \frac{T - T_c}{T_c} = T_r - 1, \\ = p - 1, \qquad = t - 1 \tag{10.25}$$

(индекс r показывает, что эти величины являются приведенными к их значениям в критической точке).

Простейшая возможная связь между математическими (a, b) и физическими $(p - 1, t - 1)$ управляющими параметрами есть линейное преобразование вида

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} p - 1 \\ t - 1 \end{bmatrix}. \tag{10.26}$$

При наличии такой связи кривая равновесия фаз должна быть прямой линией (что не так уж далеко от истины в окрестности

критической точки), которая описывается уравнением

$$0 = b = C\,(p-1) + D(t-1). \tag{10.27}$$

Приведенный коэффициент наклона m этой прямой есть

$$m = \frac{dp}{dt} = \frac{T_c}{P_c}\lim\frac{dP}{dT} = -\frac{D}{C}, \tag{10.28}$$

где предел вычисляется при приближении к критической точке вдоль кривой равновесия.

Из линейности связи (10.26) и из соотношения $x = \rho_r - 1$ вытекает следующее уравнение сборки в окрестности критической точки:

$$\begin{aligned}(\rho_r-1)^3 + (Ap + Bt - A - B)(\rho_r - 1) + \\ + (Cp + Dt - C - D) = 0.\end{aligned} \tag{10.29}$$

Группируя подобные члены, получаем

$$\begin{aligned}(\rho_r-1)^3 + p(A\rho_r + C - A) - [\rho_r(A + B) + \\ + (D + C - B - A)] = -t[B\rho_r + D - B]\end{aligned} \tag{10.30}$$

и, учитывая (10.25), окончательно имеем

$$\begin{aligned}(A - C)P_rV_r - AP_r + (D + C - B - A + 1)V_r + \\ + (B + A - 3) - \frac{1}{V_r^2} + \frac{3}{V_r} = BT_r + (D - B)V_rT_r.\end{aligned} \tag{10.31}$$

Для анализа этого запутанного уравнения сделаем ряд упрощающих, хотя и необоснованных предположений.

1. Коэффициент при V_rT_r равен нулю. Это предположение диктуется известным законом состояния идеального газа $PV = nRT$. (Поскольку идентификация с канонической катастрофой — это лишь локальное свойство, а сведение к идеальному газу (при $T \to \infty$) уже таковым не является, то, вообще говоря, трудно как-то обосновать принятое предположение.)

2. Коэффициент при V_r равен нулю.

3. Свободный член равен нулю.

Предположения 2 и 3 можно оправдать лишь тем, что при этом существенно упрощается (10.31) [см. (10.34)].

Эти предположения приводят к следующим ограничениям на коэффициенты линейного преобразования (10.26):

$$\begin{aligned}1.&\Rightarrow & D - B &= 0,\\ 2.&\Rightarrow & D + C - B - A + 1 &= 0,\\ 3.&\Rightarrow & A + B - 3 &= 0,\\ 1. + 2.&\Rightarrow & C - A + 1 &= 0.\end{aligned} \tag{10.32}$$

В результате линейное преобразование (10.26) и уравнение состояния (10.31) принимают вид

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3-D & D \\ 2-D & D \end{bmatrix} \begin{bmatrix} p-1 \\ t-1 \end{bmatrix}, \tag{10.33}$$

$$P_r V_r - (3-D) P_r + \frac{3}{V_r} - \frac{1}{V_r^2} = DT_r. \tag{10.34}$$

Сделаем еще два упрощающих предположения.

4а. Левая часть выражения (10.34) может быть представлена в виде произведения, т. е.

$$\left(P_r + \frac{\alpha_r}{V_r^2}\right)(V_r - \beta_r) = DT_r, \tag{10.35}$$

если $D = 8/3$, $\alpha_r = 3$, $\beta_r = +1/3$. Уравнение (10.35) (после подстановки $P_r = P/P_c$ и т. д.) немедленно приводит к уравнению Ван-дер-Ваальса:

$$\left(P + \frac{a}{V^2}\right)(V - \beta) = nRT, \tag{10.36}$$

где

$$a = 3P_c V_c^2, \quad \beta = {}^1/_3 V_c, \quad nR = \frac{8}{3}\frac{P_c V_c}{T_c}. \tag{10.37}$$

Здесь n — число молей, R — универсальная газовая постоянная, α, β и R имеют размерности $[n^2]$, $[n^1]$ и $[n^0]$ соответственно.

4б. Коэффициенты в (10.34) описывают приведенный коэффициент наклона m кривой равновесия фаз в критической точке [см. (10.28)], тогда $D = 2m/(m-1)$ и (10.34) уже не может быть представлено в мультипликативной форме Ван-дер-Ваальса.

Оба предположения 4а и 4б эквивалентны, если приведенный коэффициент наклона m равен 4. В этом случае возникает вопрос:

Что представляет собой приведенный коэффициент наклона $m = (T_c/P_c)\lim(dP/dT)$ *простой жидкости в окрестности ее критической точки?*

6. УРАВНЕНИЕ ВАН-ДЕР-ВААЛЬСА: КОЛИЧЕСТВЕННЫЙ АНАЛИЗ

Уравнение Ван-дер-Ваальса, предложенное впервые в 1873 г. для качественного описания жидкости вблизи ее критической точки, до сих пор не теряет своей значимости. Как из уравнения состояния из него можно получить количественные соотношения между тремя параметрами V, P, T жидкости в окрестности ее критической точки. Подобными вычислениями занимались многие

поколения исследователей. Рассмотрим некоторые из этих вопросов с позиций теории катастроф.

Уравнение Ван-дер-Ваальса описывает критическое многообразие катастрофы сборки. Зафиксировав некоторую точку на этом многообразии, можно исследовать линейный отклик системы на малые возмущения с помощью методов, описанных в разд. 3. Вместо того чтобы пользоваться тензором восприимчивости (10.21), удобнее рассматривать переменные состояния и управляющие параметры с единых позиций, записав (5.2) ($V_{ij}\delta x^j + V_{i\alpha}\delta c^\alpha = 0$) для катастрофы сборки

$$(3x^2 + a)\,\delta x + x\delta a + \delta b = 0, \tag{10.38}$$

где

$$x = \rho_r - 1 = V_c/V - 1,\quad a = \tfrac{1}{3}(P/P_c - 1) + \tfrac{8}{3}(T/T_c - 1),$$

$$\delta x = -\frac{V_c}{V^2}\delta V \quad \text{и} \quad \begin{pmatrix}\delta a\\ \delta b\end{pmatrix} = \begin{pmatrix}\frac{1}{3} & \frac{8}{3}\\ -\frac{2}{3} & \frac{8}{3}\end{pmatrix}\begin{pmatrix}\delta p\\ \delta t\end{pmatrix}. \tag{10.39}$$

Все функции линейного отклика, представляющие собой просто частные производные термодинамических величин, можно получить непосредственно из (10.38). Так, например, при постоянном объеме $\delta x = 0$ и

$$\left(\frac{V_c}{V} - 1\right)\left(\tfrac{1}{3}\delta p + \tfrac{8}{3}\delta t\right) + \left(-\tfrac{2}{3}\delta p + \tfrac{8}{3}\delta t\right) = 0, \tag{10.40}$$

поэтому

$$\left(\frac{dp}{dt}\right)_V = \frac{T_c}{P_c}\left(\frac{dP}{dT}\right)_V = \frac{8V_c}{3V - V_c}. \tag{10.41}$$

Аналогично можно определить и остальные производные термодинамических величин.

Воспользуемся длиной дуги s для параметризации пути между точкой сборки $(d, b) = (0, 0)$ и точкой (a_0, b_0) в плоскости управляющих параметров (рис. 10.11). Для достаточно малых s имеем

$$s = \frac{-x^3}{a_0 x + b_0}. \tag{10.42}$$

Если $b_0 \neq 0$, то в окрестности точки сборки

$$s = -x^3/b_0. \tag{10.43}$$

Это соотношение можно переписать в содержательных физических величинах с помощью (10.26) и связи $x = \rho_r - 1 =$

$= -v/(1+v)$, где $v = V_r - 1$. Тогда

$$s \simeq v^3/b_0. \tag{10.44}$$

Величины $P_r - 1$ и $T_r - 1$ взаимно пропорциональны вдоль любой прямолинейной траектории, проходящей через критическую

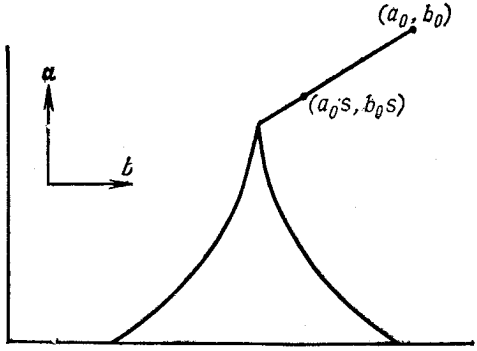


Рис. 10.11. Производные по направлению на многообразии катастрофы сборки имеют канонические значения, которые можно идентифицировать с термодинамическими свойствами уравнения состояния Ван-дер-Ваальса.

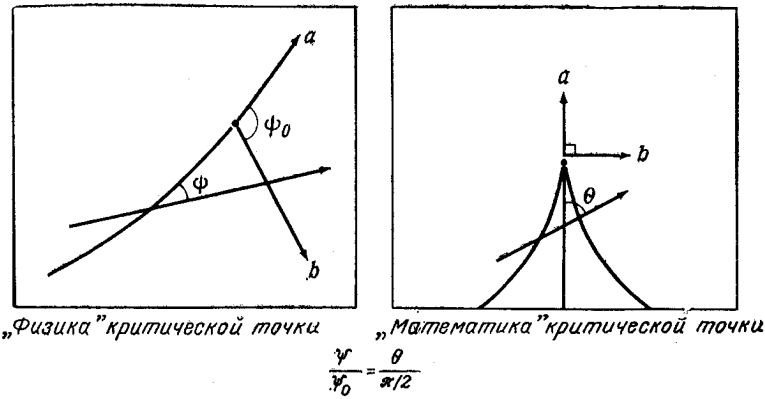


Рис. 10.12. Окрестности критической точки и росток катастрофы сборки связаны аффинным преобразованием.

точку. На любом пути, который не является «приблизительно параллельным» кривой равновесия фаз, $P_r - 1$, $T_r - 1$ приблизительно пропорциональны $(V_r - 1)^3$.

Рассмотрим путь, пересекающий кривую равновесия под углом ψ при значении приведенной температуры $T_r < 1$ (рис. 10.12). В этой точке $P_r - 1 = -(D/C)(T_r - 1)$ и $a =$

$= C^{-1}(AD - BC)(1 - T_r) \to -4(1 - T_r)$ для вандерваальсовой жидкости. Угол ψ в фазовой плоскости должен быть связан с углом θ в плоскости управляющих параметров катастрофы сборки (рис. 10.3), при этом необходимо отметить, что аффинное преобразование (10.26) не сохраняет углов. Угол между кривой равновесия (ось a) и дополнительной осью в фазовой плоскости из элементарных геометрических соображений определяется как

$$\cos\psi_0 = \frac{(1, -B/A)\cdot(1, -D/C)}{\sqrt{1+(B/A)^2}\,\sqrt{1+(D/C)^2}}. \qquad (10.45)$$

Для вандерваальсовой жидкости эта величина равна 158,84°, поэтому изменение внутренней энергии при пересечении кривой равновесия под углом ψ есть [см. 10.10)]

$$\Delta\frac{d}{ds}(U/U_c) = -2\sqrt{4(1-T_r)}\,\sin(0{,}5666\,\psi). \qquad (10.46)$$

Наконец, из геометрии канонической катастрофы сборки следует, что разность плотностей жидкости и газа вдоль кривой равновесия меняется по закону $\Delta x = \Delta\rho = 2\sqrt{-a} \sim (T_c - T)^{1/2}$.

Количественная неадекватность уравнения Ван-дер-Ваальса проявляется особенно отчетливо при попытке сравнения расчетных значений (в особенности последнего соотношения) с результатами экспериментов. На рис. 10.13 представлена зависимость [3] приведенной температуры T/T_c от ρ/ρ_c для восьми веществ в двух фазах вдоль кривой равновесия. По существу все точки ложатся на одну универсальную кривую, что наводит на мысль о существовании для этих газов некоего закона соответственных состояний. Закон соответственных состояний утверждает, что термодинамические функции состояния для каждого из этих газов изоморфны единой универсальной функции безразмерных приведенных физических величин.

Полученные экспериментальные данные позволяют сделать два важных вывода.

1. Средняя плотность жидкой и газообразной фаз вдоль кривой равновесия линейно зависит от расстояния до критической точки:

$$\frac{1}{2}\left(\frac{\rho_l}{\rho_c} + \frac{\rho_g}{\rho_c}\right) = 1 + \frac{3}{4}(1 - T_r). \qquad (10.47)$$

Этот факт известен как *закон линейных диаметров*. В связи с этим возникает вопрос: соблюдается ли закон линейных диаметров за критической точкой? Иными словами, если экстраполировать кривую раздела через критическую точку, то будет ли плотность вещества во «флуидном» состоянии удовлетворять соотношению (10.47), если его левую часть заменить на ρ_f/ρ_c

(и $T_r > 1$)? Очевидно, ответ на этот вопрос может быть получен экспериментальным путем.

2. Разность плотностей жидкости и газа вдоль кривой равновесия пропорциональна $(1 - T_r)^{\beta}$, где $\beta \simeq 1/3$, т. е.

$$\frac{\rho_l}{\rho_c} - \frac{\rho_g}{\rho_c} = \frac{7}{2}(1 - T_r)^{1/3}. \qquad (10.48)$$

Отсюда с очевидностью следует неадекватность уравнения Ван-дер-Ваальса, так как согласно этому уравнению $\Delta\rho$ зави-

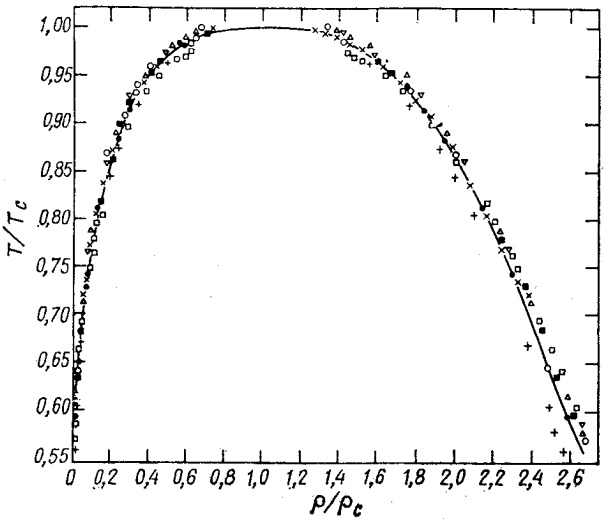


Рис. 10.13. Зависимость приведенной плотности от приведенной температуры для жидких и газообразных фаз восьми простых веществ, полученная экспериментально Гуггенхеймом [3].
+ Ne, • A, ■ Kr, × Xe, △ N_2, ▽ O_2, □ CO, ○ CH_4.

сит от $T_c - T$ как корень квадратный, в то время как эксперименты показывают, что это корень кубический. В связи с этим возникает вопрос: можно ли найти такое соответствие между математическими (x; a, b) и физическими (ρ_r; P_r, T_r) параметрами, чтобы удовлетворить экспериментальным данным? Для катастрофы сборки ответ отрицательный. Соответствие $x \sim |\rho_r - 1|^{3/2}$ не является гладким. Возможные соответствия для управляющих параметров, такие, как $a \sim [A(p-1) + B(t-1)]^{2/3}$, также неудовлетворительны, поскольку математические управляющие параметры не могут образовать систем координат в критической точке.

Итак, катастрофу сборки можно использовать для построения уравнения состояния жидкости в окрестности ее критиче-

ской точки. Однако уравнение Ван-дер-Ваальса и все подобные ему уравнения, полученные из рассуждений, основанных на усреднении или на катастрофе сборки, в количественном отношении оказываются несостоятельными, поскольку приводят к зависимости разности плотностей жидкости и газа вдоль границы равновесия (по закону квадратного корня), отличающейся от экспериментально установленной и соответствующей закону кубического корня.

7. ФАЗОВЫЕ ПЕРЕХОДЫ ПЕРВОГО РОДА

Для описания фазовых переходов первого рода [1] используется модель, аналогичная модели, описанной в разд. 2, с той лишь разницей, что в ней рассматривается потенциал шестой степени

$$V(x;\ a,\ b,\ c,\ d) = \frac{1}{6}x^6 + \frac{1}{4}ax^4 + \frac{1}{3}bx^3 + \frac{1}{2}cx^2 + dx. \tag{10.49}$$

Масштабирование для этой функции таково:

$$x \to x' = \lambda x,\quad a \to a' = \lambda^2 a,\quad b \to b' = \lambda^3 b \Rightarrow V \to V' = \lambda^6 V, \tag{10.50}$$

$$c \to c' = \lambda^4 c,\quad d \to d' = \lambda^5 d.$$

В модели Гинзбурга — Ландау присутствует принцип симметрии, ограничивающий форму потенциала A_{+5} до класса четных функций от x. [Будем обозначать такой симметризованный потенциал через $V(x;\ a,\ c)$.]

Потенциал Гинзбурга — Ландау имеет двумерное пространство управляющих параметров $\mathbb{R}^2$. Как обычно, целесообразно выяснить, каким образом бифуркационное $\mathscr{S}_B$ и максвелловское $\mathscr{S}_M$ множества разбивают это пространство на открытые подмножества. Критическое множество определяется уравнением

$$\frac{dV}{dx} = x(x^4 + ax^2 + c) = 0. \tag{10.51}$$

Точка $x = 0$ является критической, как это всегда имеет место для симметрических потенциалов. Оставшееся биквадратное уравнение может быть разрешено относительно x^2:

$$x^2 = -\frac{a}{2} \pm \sqrt{\left(\frac{a}{2}\right)^2 - c}\,. \tag{10.52}$$

Если $c < 0$, то уравнение имеет два корня разных знаков и, следовательно, два действительных решения для x;

если $0 < c < (a/2)^2$, то оба значения x^2 положительны при $a \leq 0$ и отрицательны при $a > 0$;

если $c > (a/2)^2$, то (10.52) вообще не имеет действительных корней.

В результате плоскость (a, c) управляющих параметров разбивается на открытые подмножества с параметризующими

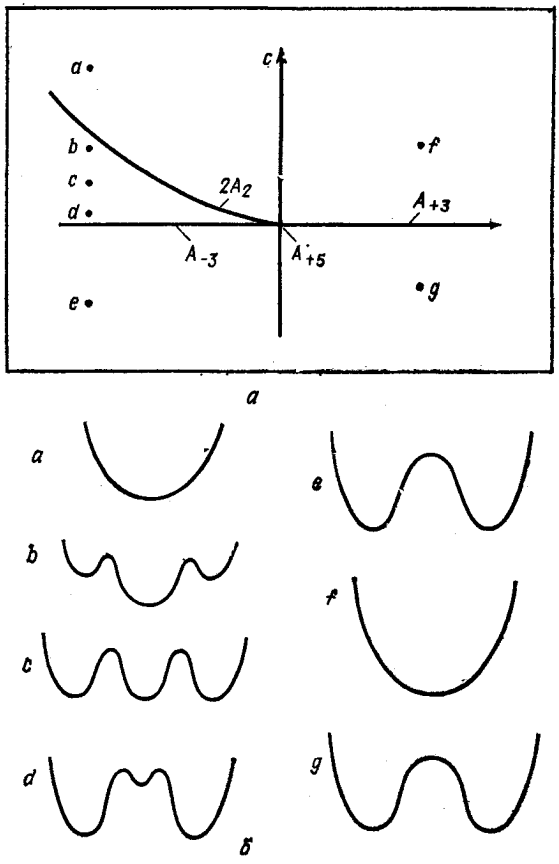


Рис. 10.14. Плоскость (a, c) управляющих параметров для симметризованной катастрофы A_{+5} разбивается на три открытые области, соответствующие функциям с одним, двумя и тремя минимумами; потенциалы соответствуют семи отдельным точкам этой плоскости.

функциями, имеющими 1, 3 или 5 критических точек (рис. 10.14). Критическое множество показано на рис. 10.15, где для простоты не показана плоскость $x = 0$ в $\mathbb{R}^1 \otimes \mathbb{R}^2$, которая на самом деле является частью критического множества.

Бифуркационное множество, как обычно, можно получить, приравняв нулю вторые производные. Геометрически оно пред-

ставляет собой «тень», отбрасываемую критическим множеством на плоскость управляющих параметров $\mathbb{R}^2$ (рис. 10.15), и включает полупрямые $a > 0$, $c = 0 (A_{+3})$, $a < 0$, $c = 0 (A_{-3})$, кривую $a < 0$, $c = (a/2)^2 (2A_2)$ и точку $a = 0$, $c = 0 (A_{+5})$. Качественный вид потенциалов для некоторых точек показан на рис. 10.14. При переходе от a к b имеются две симметричные катастрофы складки. В точке c на максвелловском множестве три локальных минимума имеют равную «глубину», поскольку критиче-

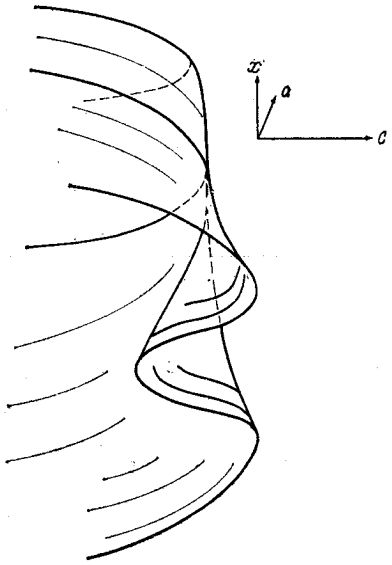


Рис. 10.15. Поверхность критических точек ($\nabla V(x; a, c) = 0$) в $\mathbb{R}^3$. (Плоскость $x = 0$ явно не показана.)

ские значения боковых минимумов приближаются к нулю сверху. На переходе от d к e имеется двойственная катастрофа A_{-3}, а на пути от f к g — стандартная катастрофа сборки.

Три минимума с равным нулю критическим значением находятся в точках $x = 0$, $x^2 = -3a/4$. Максвелловское множество определяется как

$$\mathscr{S}_M : a < 0, \quad c = \frac{3}{16} a^2. \tag{10.53}$$

Покажем, что для такого симметризованного потенциала Гинзбурга — Ландау возможны фазовые переходы нулевого, первого и второго рода.

0. Если придерживаться принципа максимального промедления, то при пересечении компонент $\mathscr{S}_B$ с $a < 0$ в подходящем

направлении имеют место фазовые переходы нулевого рода. Так, например, при пересечении кривой $c=(a/2)^2$ вдоль пути 0a (рис. 10.16) потенциал меняется на величину

$$\Delta V = V(0;\ a,\ c) - V\left(\pm\sqrt{\frac{-a}{2}};\ a,\ c\right) = \\ = -\frac{1}{6}\left(-\frac{a}{2}\right)^3. \qquad (10.54)$$

При пересечении полупрямой $a<0$, $c=0$ вдоль пути 0b изменение потенциала есть

$$\Delta V = V(\pm\sqrt{-a};\ a,\ 0) - V(0;\ a,\ 0) = -\frac{1}{12}(-a)^3. \qquad (10.55)$$

1. Если придерживаться принципа Максвелла, то фазовый переход первого рода имеет место при пересечении $\mathscr{S}_M$. На

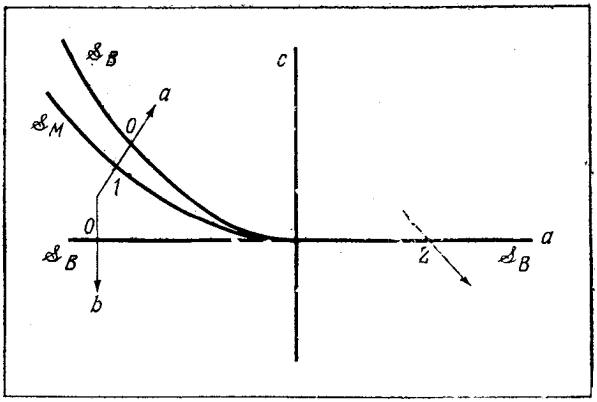


Рис. 10.16. Симметризованный потенциал Гинзбурга — Ландау претерпевает фазовые переходы нулевого, первого и второго рода почти точно так же, как и катастрофа A_{+3}.

пути 1 (рис. 10.16) первая производная потенциала изменяется на величину

$$\Delta\frac{dV}{ds} = \frac{d}{ds}V\Big|_{0;\ a,\ c} - \frac{d}{ds}V\Big|_{\pm\sqrt{-3a/4};\ a,\ c} = \\ = -\left\{\frac{1}{4}\left(-\frac{3a}{4}\right)^2\frac{da}{ds} + \frac{1}{2}\left(-\frac{3a}{4}\right)\frac{dc}{ds}\right\}, \qquad (10.56)$$

где da/ds, dc/ds — направляющие косинусы пути в пространстве $\mathbb{R}^2$ при пересечении $\mathscr{S}_M$. Для такого фазового перехода

первого рода спинодалями являются описанные выше две кривые фазовых переходов нулевого рода.

2. Фазовый переход второго рода происходит при пересечении бифуркационного множества вдоль пути 2. Величина скачка второй производной зависит от направления движения и определяется как

$$\Delta \frac{d^2V}{ds^2} = -\frac{1}{2a}\left(\frac{dc}{ds}\right)^2, \qquad (10.57)$$

где значение направляющего косинуса вычисляется на бифуркационном множестве.

8. ТРИКРИТИЧЕСКИЕ ТОЧКИ

Анализ следствий нарушения принципа симметрии, принятого в модели Гинзбурга — Ландау фазовых переходов первого рода (разд. 7), может быть выполнен теми же методами, что и анализ нарушений этого принципа в канонической модели фазовых переходов второго рода (разд. 2). В этом случае множество возможных ситуаций гораздо богаче, поскольку пространство управляющих параметров для катастрофы A_{+5} имеет более высокую размерность, чем для катастрофы A_{+3}. Первым шагом подобного анализа является введение линейного члена dx. Тогда семейство потенциалов принимает вид

$$V(x;\ a,\ c,\ d) = V(x;\ a,\ c) + dx =$$

$$= \frac{1}{6}x^6 + \frac{1}{4}ax^4 + \frac{1}{2}cx^2 + dx. \qquad (10.58)$$

Потенциал (10.58), называемый «трикритическим», представляет особый интерес для тех областей термодинамики, в которых обычно принимается принцип Максвелла. В связи с этим определим максвелловское множество $\mathcal{S}_M$ этого потенциала. В предыдущем разделе это уже было сделано в плоскости (a, c) $(d = 0)$ пространства $\mathbb{R}^3$ управляющих параметров. В указанной плоскости имеются два равновеликих минимума под положительной полупрямой $(a \geqslant 0,\ c = 0)$ и, кроме того, под кривой $(a \leqslant 0,\ c = (3/16)\,a^2)$. Эти две линии сами являются границами максвелловского множества в $\mathbb{R}^3$: полупрямая $(a > 0,\ c = 0)$ параметризует фазовые переходы второго рода, а кривая $(a < 0,\ c = 3/16\,a^2)$ — первого. На этой кривой потенциал имеет три равновеликих минимума (отсюда название «кривая тройных точек»). Граница этих кривых (т. е. $(a;\ c,\ d) = (0;\ 0,\ 0)$) называется «трикритической точкой» и является ростком катастрофы A_{+5}. Максвелловское множество потенциала $V(x;\ a,\ c,\ d)$ в $\mathbb{R}^3$ проще всего определить, выяснив, как изменяется вид потенциала при движении в пространстве $\mathbb{R}^3$ по пря-

мым линиям с постоянными (a, c) и возрастающим d. Вид потенциала в зависимости от d для четырех точек в плоскости (a, c) показан на рис. 10.17. Из этого рисунка следует, что фазовый переход первого рода происходит вдоль прямых *1* и *2*, когда глобальный минимум перескакивает с правого локального минимума на левый в момент, когда d, возрастая, проходит через нуль. Два фазовых перехода первого рода могут возникнуть на кривой *3*, когда глобальный минимум перескакивает сначала с правого локального минимума на средний при $d =$

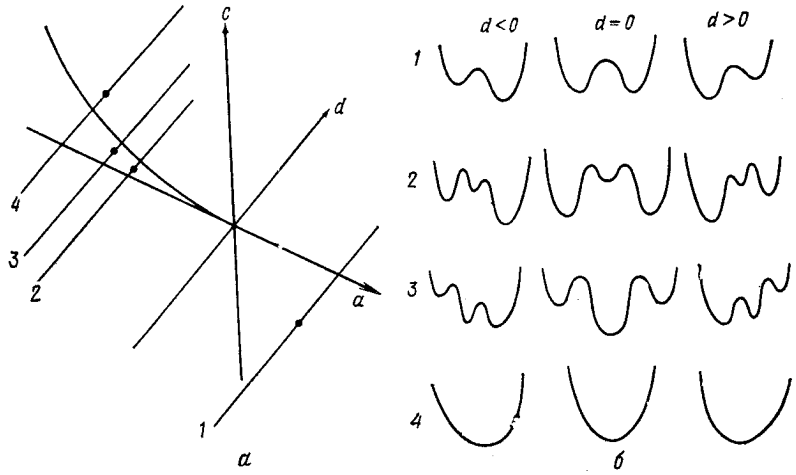


Рис. 10.17. Зависимость формы трикритического потенциала от возрастающего параметра d при фиксированных (a, c).

$= -d_0$, и затем с центрального на левый при $d = +d_0$. На прямой *4*, лежащей достаточно высоко над кривой тройных точек, при изменении d не может произойти никаких фазовых переходов.

Следствием подобных рассуждений является тот факт, что часть плоскости (a, c) ($d = 0$), лежащая ниже линий катастроф $A_{\pm 3}$, параметризует фазовые переходы первого рода, отвечающие скачкам между левыми и правыми минимумами в точках $x^{(l)}$ и $x^{(p)}$. Над кривой тройных точек симметрично относительно плоскости (a, c) расположены два «крыла» фазовых переходов первого рода, включающие боковой минимум и «средний минимум» в $x^{(m)}$. Эти крылья должны иметь границу. Кривые, ограничивающие эти крылья, должны параметризовать катастрофу A_{+3} (см. рис. 10.18). Поскольку граничная кривая является частью бифуркационного множества, легко получить ее уравнение в параметрической форме, если воспользоваться

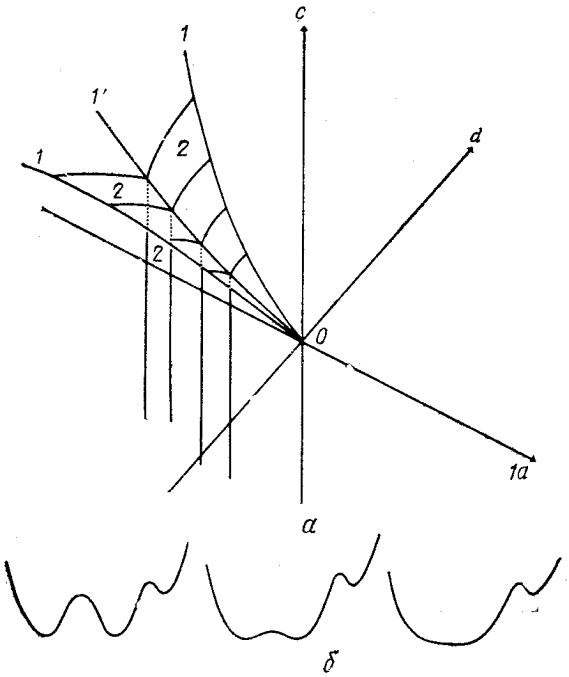


Рис. 10.18.
а — максвелловское множество трикритического потенциала, состоящее из (части) плоскости симметрии $d = 0$ и двух крыльев. Границы этих крыльев параметризуют катастрофу сборки A_{+3}; 0 — имеется нуль-мерная компонента $(a, c, d) = (0, 0, 0)$-трикритическая точка, параметризующая росток катастрофы A_{+5}; *1* — имеются три одномерные компоненты, параметризующие фазовые переходы второго рода. Это прямая ($a > 0$, $c = d = 0$) и две кривые, параметрическое описание которых дается выражениями (10.59); *1'* — имеется одна одномерная кривая, параметризующая фазовые переходы первого рода, а именно $\left(a < 0,\ c = \frac{3}{16}a^2,\ d = 0\right)$; *2* — имеются три двумерные сепаратрисы, параметризующие функции, для которых два наиболее глубоких из трех локальных минимумов имеют вырожденные критические значения. Эти поверхности пересекаются по кривой тройных точек *1'*. Границами этих трех компонент являются описанные выше кривые *1* и *1'*. Эти три сепаратрисы можно продолжить через кривую *1'*, и тогда они параметризуют функции с вырожденными метастабильными критическими значениями; *б* — изменение формы потенциалов для точек на одном из крыльев при приближении к кривой *1*.

тем же подходом, что и в случае вывода уравнений (5.9)—(5.12). В результате имеем

$$a = -\frac{10}{3}\lambda^2,\quad c = 5\lambda^4,\quad d = -\frac{8}{3}\lambda^5,\qquad \lambda \in \mathbb{R}^1,\ \lambda \neq 0, \tag{10.59}$$

где ростком сборки является $(x - \lambda)^4$.

Интегрируя уравнения Клаузиуса — Клапейрона, можно построить двумерные крылья фазовых переходов первого рода.

Проще, однако, сначала понизить размерность задачи на единицу. Для этого ограничимся рассмотрением плоскости $a = -1$. Пересечение двумерных крыльев с этой плоскостью дает одномерную кривую, уравнение которой определить сравнительно просто. Далее, используя соотношения (10.50), можно восстановить все крылья.

Топология максвелловского множества $\mathscr{P}_M$ в $\mathbb{R}^3$ для трикритического потенциала $V(x;\ a,\ c,\ d)$ показана на рис. 10.18. Сепаратрису $\mathscr{P}_B$ в $\mathbb{R}^3$ можно построить по трем сечениям $a = -1$, $a = 0$, $a = 1$, показанным на рис. 10.19, если использовать соотношения масштабирования, вытекающие из параметрического представления

$$\begin{aligned} c(x,\ a) &= -5x^4 - 3ax^2, \\ d(x,\ a) &= +4x^5 + 2ax^3. \end{aligned} \tag{10.60}$$

9. МАКСВЕЛЛОВСКОЕ МНОЖЕСТВО ДЛЯ КАТАСТРОФ ТИПА A_{+5}

Возмущением наиболее общего вида для симметризованной модели Гинзбурга — Ландау (разд. 7) фазовых переходов первого рода является катастрофа A_{+5} (10.49), имеющая четырехмерное пространство управляющих параметров. Бифуркационное множество $\mathscr{P}_B$ можно определить точно так же, как и в случае катастроф типа A_4 (гл. 5). Наибольший интерес, однако, представляет максвелловское множество $\mathscr{P}_M$, которое можно построить, проинтегрировав уравнения Клаузиуса — Клапейрона.

Особого внимания заслуживают те компоненты максвелловского множества, которые параметризуют функции с двумя или тремя равными глобальными минимумами. Грубо говоря (и тем не менее это вполне корректно), для того чтобы сделать каждое критическое значение вырожденным, достаточно одного управляющего параметра (т. е. функции от управляющих параметров).

Таким образом, имеются три трехмерные компоненты $(3 = 4 - (2 - 1))$ множества $\mathscr{P}_M$, на которых два из трех минимумов вырождены, и одна двумерная компонента $(2 = 4 - -(3 - 1))$, параметризующие функции с тремя минимумами равной глубины. Именно этой поверхности «тройных точек» и будет уделено основное внимание в данном разделе. К сожалению, уравнения Клаузиуса — Клапейрона легко интегрируются только в одномерном случае, а поверхность тройных точек двумерна. Отсюда вытекает необходимость поиска хороших методов, позволяющих понизить размерность задачи на единицу. [По-видимому, такой метод должен использовать преобразования подобия (10.50).]

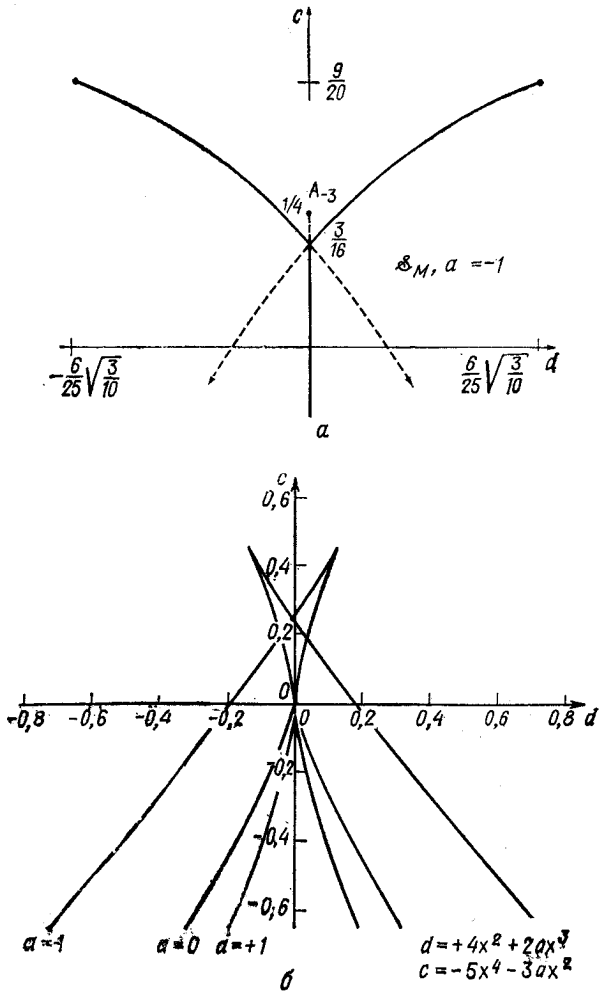


Рис. 10.19.

а — пересечение максвелловского множества $\mathcal{S}_M$ трикритического потенциала с плоскостью $a = 1$. Сплошные линии показывают значения управляющих параметров, при которых два «наиболее глубоких» минимума равны (устойчивое максвелловское множество). Штриховые линии отвечают значениям управляющих параметров, при которых эти минимумы метастабильны по отношению к третьему минимуму (метастабильное максвелловское множество). Метастабильное максвелловское множество не представляет никакого интереса с точки зрения практического применения в физических системах; *б* — пересечения бифуркационного множества $\mathcal{S}_B$ трикритического потенциала с плоскостями $a = -1, 0, +1$. Полное бифуркационное множество можно построить из этих сечений и уравнений масштаба (10.50).

Первое, что приходит в голову, — это попытаться использовать (10.50) (рис. 10.18). Однако, поскольку максвелловское множество определено для $b = 0$ [см. (10.58)], поступим иначе, а именно зафиксируем $a = \text{const}$ в (10.49), а затем воспользуемся соотношениями подобия для обратного перехода от $\mathscr{S}_M$ в $\mathbb{R}^3$ ($a = \text{const}$) к $\mathscr{S}_M$. Потенциал A_{+5} при $a > 0$ имеет не более двух минимумов, поэтому зафиксируем $a = -1$ и найдем одномерную кривую — геометрическое место критических точек в $\mathbb{R}^3$.

Для этого предположим, что $(c^1, c^2, c^3, c^4) = (-1, b, c, d)$, а $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ определяют положения трех равных локальных минимумов. Тогда уравнения Клаузиуса — Клапейрона принимают вид

$$\begin{aligned}
\delta c^2 &= (-)^{1+2} \det \begin{bmatrix} M_3^1 & M_4^1 \\ M_3^2 & M_4^2 \end{bmatrix} \delta s, \\
\delta c^3 &= (-)^{1+3} \det \begin{bmatrix} M_2^1 & M_4^1 \\ M_2^2 & M_4^2 \end{bmatrix} \delta s, \\
\delta c^4 &= (-)^{1+4} \det \begin{bmatrix} M_2^1 & M_3^1 \\ M_2^2 & M_3^2 \end{bmatrix} \delta s,
\end{aligned} \tag{10.61}$$

где

$$M_\alpha^i = V_\alpha^{(i)} - V_\alpha^{(i+1)}$$

и

$$V_\alpha^{(i)} = \frac{\partial V}{\partial c^\alpha}\bigg|_{x=x^{(i)}}, \qquad i = 1, 2, 3, \qquad \alpha = 2, 3, 4.$$

Смещения этих трех минимумов определяются с помощью (5.2′):

$$\delta x^{(i)} = -\sum_{\alpha=2}^{4} \frac{1}{V_{xx}^{(i)}} V_{x\alpha}^{(i)} \delta c^\alpha,$$

где

$$\begin{aligned}
V_{xx}^{(i)} &= \frac{\partial^2 V}{\partial x^2}\bigg|_{x=x^{(i)}}, \\
V_{x\alpha}^{(i)} &= \frac{\partial^2 V}{\partial x\, \partial c^\alpha}\bigg|_{x=x^{(i)}}.
\end{aligned} \tag{10.62}$$

Результаты численного интегрирования этих уравнений с малым шагом δs показаны на рис. 10.20 (начальные условия для интегрирования этих обыкновенных дифференциальных уравнений первого порядка получены из рис. 10.16).

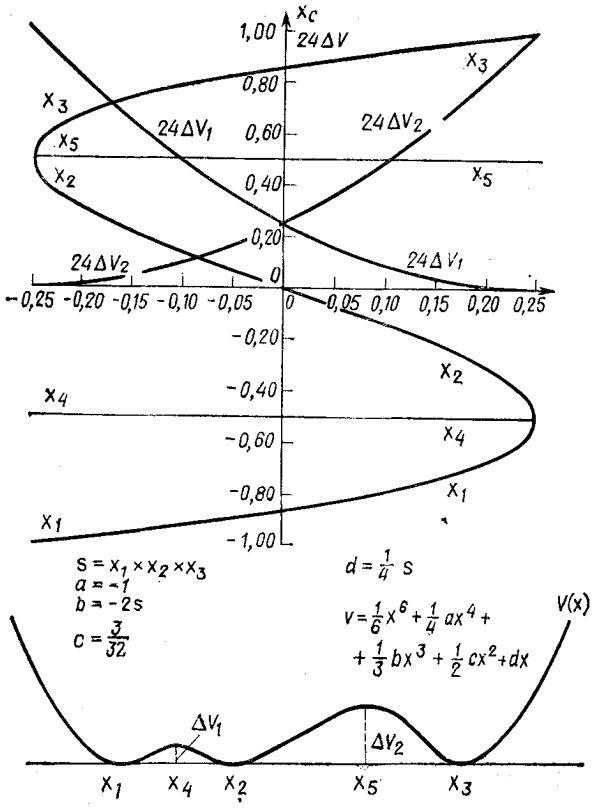


Рис. 10.20. Результаты численного интегрирования уравнений Клаузиуса — Клапейрона (10.61): параметрическое представление положений трех минимумов $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ и соответствующих значений управляющих параметров $a = -1$, b, c, d.

10. ТРОЙНАЯ ТОЧКА

Читатель, уже знакомый с «принципом лома», не мог не заметить, что мы не пользовались этим принципом при исследовании критической точки ($\sim x^4$), рис. 10.7. Однако при анализе тройной точки ($\sim x^6$) использование этого принципа напрашивается само собой. Но этот путь таит в себе ряд существенных трудностей: в то время как критическая точка, качественно характеризуемая ростком катастрофы, лежит в плоскости (P, T), трикритическая точка, качественно характеризуемая ростком x^6, не лежит в плоскости (P, T). Вместе с тем в этой плоскости лежит тройная точка, но она не является ростком катастрофы. Следовательно, варьируя P и T, можно выйти лишь в тройную

точку и на кривую равновесия, но невозможно достичь ростка катастрофы.

Для отыскания уравнения состояния вещества в окрестности тройной точки, аналогичного уравнению Ван-дер-Ваальса в случае критической точки, следует, во-первых, идентифицировать математическую переменную состояния x с физической величиной ρ посредством соотношения $x \simeq (\rho - \rho_0)$ и, во-вторых, предположить существование простейшей связи между четырьмя математическими c^α и двумя физическими P, T управляющими параметрами:

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} A_P^1 & A_T^1 & A^1 \\ A_P^2 & A_T^2 & A^2 \\ A_P^3 & A_T^3 & A^3 \\ A_P^4 & A_T^4 & A^4 \end{bmatrix} \begin{bmatrix} P \\ T \\ 1 \end{bmatrix}. \qquad (10.63)$$

$\Diamond\Diamond\Diamond$ Геометрически мы рассматриваем двумерную плоскость (P, T) физического пространства $\mathbb{R}^2$ управляющих параметров как погруженную в четырехмерное пространство $\mathbb{R}^4$ математических управляющих параметров (a, b, c, d). Плоскость (P, T) пересекает двумерную компоненту максвелловского множества $\mathscr{S}_M$ в точке (тройная точка), а три трехмерные компоненты (параметризующие потенциалы с двумя равными глобальными минимумами) — по трем одномерным кривым (кривые равновесия). Для погружения физического пространства $\mathbb{R}^2$ в математическое $\mathbb{R}^4$ требуется 12 параметров, содержащихся в 4×3-матрице (10.63).

Имея 14 свободных параметров (12 элементов матрицы, ρ_0 и масштаб полной энергии), можно, по-видимому, получить хорошее соответствие с экспериментальными данными. Однако эти параметры должны удовлетворять определенным условиям, непосредственно вытекающим из фазовой диаграммы любого вещества (рис. 10.21),

— возможность определения средней плотности в тройной точке;

— независимость двух из трех разностей плотностей в тройной точке;

— независимость двух из трех углов θ_1, θ_2, θ_3, а следовательно, и наклона любой из кривых равновесия при температуре T;

— независимость значения скрытой теплоты вдоль двух независимых путей вблизи тройной точки;

— независимость отношений P_c/P_T и T_c/T_T. Остальные $4 (= 14 - 10)$ степени свободы можно определить по скрытой теплоте или разности плотностей вдоль кривых равновесия.

Было бы очень интересно определить, насколько точно можно воспроизвести имеющиеся экспериментальные данные для хорошо изученных простых физических веществ при наилучшем вложении физического пространства $\mathbb{R}^2$ в математическое пространство $\mathbb{R}^4$ В связи с этим возникают два вопроса.

1. Каким будет уравнение состояния вещества в окрестности тройной точки?

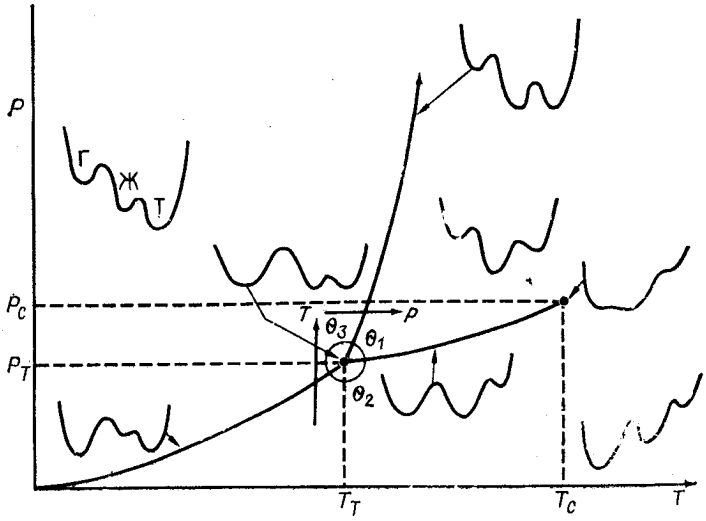


Рис. 10.21. Качественные свойства потенциала в разных областях фазовой плоскости, величина которого определяет состояние вещества в окрестности его критической точки.

2. Можно ли принцип соответствующих состояний обобщить на твердые, жидкие и газообразные фазы?

◇◇◇ При описании «флуидного» состояния вещества уравнение Ван-дер-Ваальса приводит к неправильным показателям степени по мере приближения к ростку сборки x^4. Поскольку росток x^6 не лежит в плоскости (P, T) физического пространства $\mathbb{R}^2$, можно ожидать, что классические показатели степени, вычисленные по этому аналогу уравнения Ван-дер-Ваальса, вообще говоря, дадут достаточно адекватное описание состояния вещества в окрестности его тройной точки.

11. ТОПОЛОГИЯ КАТАСТРОФ И ТЕРМОДИНАМИКА

До сих пор наше рассмотрение основных законов термодинамики носило феноменологический характер, т. е. мы пытались найти уравнения, которые могли бы воспроизвести эксперимен-

тальные данные с той или иной степенью точности. Теперь займемся другим аспектом этой проблемы и посмотрим, может ли теория катастроф быть хоть сколько-нибудь полезной при формулировании основных законов термодинамики на концептуальном уровне.

Состояние термодинамического равновесия определяется значениями соответствующих термодинамических переменных. Грубо говоря, эти переменные можно разделить на два класса: *экстенсивные*, или пропорциональные полной массе системы ($\sim M^1$), такие, как энтропия и объем, и *интенсивные*, или не зависящие от полной массы системы ($\sim M^0$), такие, как температура и давление[1]). Экстенсивные и интенсивные термодинамические переменные обычно выступают в виде пар сопряженных переменных (табл. 10.1).

Таблица 10.1. **Пары сопряженных экстенсивных и интенсивных термодинамических переменных**

Экстенсивные			Интенсивные
Энтропия	S	T	Температура
Объем	V	$-P$	Давление
Число молей	N	μ	Химический потенциал
Деформация	$S_{\mu\nu}$	$T_{\mu\nu}$	Напряжение
Намагниченность	$\mathbf{M}$	$\boldsymbol{H}$	Напряженность магнитного поля
Поляризация	$\boldsymbol{P}$	$\boldsymbol{E}$	Напряженность электрического поля
Электрический заряд	Q	Φ	Электрический потенциал
Угловой момент	$\boldsymbol{J}$	$\boldsymbol{\Omega}$	Угловая скорость
Поверхность событий черной дыры	A	k	Поверхностная гравитация черной дыры

Не все из n экстенсивных E^α и n интенсивных i_α сопряженных переменных, описывающих систему в состоянии термодинамического равновесия, являются независимыми. На самом деле только n из этих переменных могут принимать независимые значения. И как только значения n таких переменных зафиксированы, остальные n переменных определяются однозначным образом. Если, например, в качестве независимых рассматривать n экстенсивных переменных, то значения n интенсивных переменных определяются однозначно. Таким образом,

$$i_\alpha = i_\alpha(E^\beta). \tag{10.64}$$

[1]) Для описания интенсивных и экстенсивных термодинамических переменных удобно использовать соответственно ковариантные (i_α) и контравариантные (E^α) обозначения.

Соотношения такого типа называют *уравнениями состояния.* Разумеется, в качестве независимых переменных можно выбрать любые [1]) n из $2n$ термодинамических переменных. Тогда остальные n зависимых термодинамических переменных можно описать в виде функций независимых переменных, воспользовавшись уравнением состояния (10.64).

С более общих позиций уравнения состояния полезно рассматривать как описание способа погружения n-мерного многообразия равновесия в $2n$-мерное пространство $\mathbb{R}^n \otimes \mathbb{R}^n$. Координатами в этом эвклидовом пространстве являются значения n экстенсивных и n интенсивных термодинамических переменных. Задание n из $2n$ термодинамических величин единственным образом определяет состояние равновесия системы и, следовательно, соответствующую точку на многообразии уравнения состояния.

Из этих рассуждений вытекает весьма естественная связь термодинамического описания состояния системы с теорией катастроф. Критическое многообразие определяется равенством нулю градиента. В $\mathbb{R}^n \otimes \mathbb{R}^k$ пространстве n фазовых переменных и k управляющих параметров это многообразие имеет размерность k. Полезно попробовать идентифицировать многообразие уравнения состояния (термодинамика) с критическим многообразием (теория катастроф). Поскольку многообразие уравнения состояния является n-мерным многообразием, вложенным в $2n$-мерное пространство, следует выбрать n термодинамических переменных в качестве управляющих параметров (т. е. независимых переменных) и остальные n зависимых переменных — в качестве переменных состояния. Тогда многообразие уравнения состояния определяется равенством нулю градиента некоторого потенциала по переменным состояния.

Теперь остается решить проблемы выбора соответствующих переменных состояния и управляющих параметров и соответствующей функции этих переменных, критическое многообразие которой можно идентифицировать с многообразием уравнения состояния вещества, а также какого принципа следует придерживаться. [В соответствии с предыдущим анализом фазовых переходов (разд. 5 и 10) будем придерживаться принципа Максвелла.]

После того как все указанные проблемы решены, многообразие уравнения состояния оказывается спаянным. В результате интенсивные и экстенсивные переменные трактуются с не вполне единых позиций. Например, системы у которых в состоянии равновесия существует более одной фазы, описываются пло-

[1]) Это утверждение следует несколько уточнить в том случае, когда в состоянии термодинамического равновесия сосуществуют несколько фаз.

скими участками многообразия уравнения состояния (см. рис. 10.5). Точки, принадлежащие этому плоскому участку, можно единственным образом параметризовать через одни управляющие параметры при условии, что выбор этих параметров тщательно продуман. В противном случае для параметризации потребуется смесь переменных состояния и управляющих параметров.

Для иллюстрации сказанного рассмотрим простое, однокомпонентное вещество. Состояние такого вещества характеризуется двумя парами сопряженных термодинамических переменных: (S, T) и (V, P). При этом уравнение состояния есть двумерное многообразие, погруженное в пространство $\mathbb{R}^4$ с координатами $S, T; V, P$. В качестве независимых переменных можно выбрать любые две из этих четырех, и на пространство $\mathbb{R}^2$ этих управляющих параметров можно спроектировать многообразие (спаянное) уравнения состояния (рис. 10.22). В этом случае плоские участки (описывающие сосуществование различных фаз) многообразия уравнения состояния можно однозначно параметризовать через экстенсивные переменные. Этот факт остается справедливым для любых термодинамических систем. Поэтому вполне естественно рассматривать экстенсивные термодинамические переменные как управляющие параметры и, следовательно, интенсивные переменные — как переменные состояния.

По этой причине предметом нашего изучения будет критическое многообразие $\nabla_i \mathcal{U}(i_\alpha; E^\beta)$ некоторого семейства потенциальных функций $\mathcal{U}$, зависящих от n интенсивных переменных состояния i_α и n экстенсивных управляющих параметров E^β. ◇◇◇ Благодаря работам известного американского естествоиспытателя Дж. Гиббса в термодинамике произошел переход от изучения процессов к изучению состояний. В гиббсовской формулировке термодинамики состояние системы определяется некоторым термодинамическим «потенциалом», зависящим от выбора независимых переменных. Для описания однокомпонентных веществ обычно используются следующие потенциалы: $U(S, V)$ (внутренняя энергия), $G(T, P)$ (термодинамический потенциал), $A(T, V)$ (свободная энергия) и $H(S, P)$ (энтальпия)[1]). При условиях термодинамического равновесия эти функции определяются однозначно заданием их двух аргументов. Это означает, что они не являются потенциальными функциями в общепринятом смысле, поскольку состояние системы не определяется из равенства градиента нулю. Эти функции скорее можно назвать «производящими», так как их градиенты определяют

[1]) Функции, зависящие от сопряженных пар (S, T) и (V, P), не используются, поскольку их невозможно получить из четырех перечисленных функций с помощью преобразования Лежандра.

сопряженные термодинамические переменные:

$$\left(\frac{\partial U}{\partial S}\right)_V = T, \quad \left(\frac{\partial U}{\partial V}\right)_S = -P. \tag{10.65}$$

Фактически именно благодаря тому, что состояние не определяется равенством градиента нулю, эти термодинамические производящие функции связаны преобразованием Лежандра. Более того, вычисляя различные частные производные от термодинамических «потенциалов» системы, можно выявить все ее термодинамические свойства, поэтому термин «термодинамическая производящая функция» кажется более уместным.

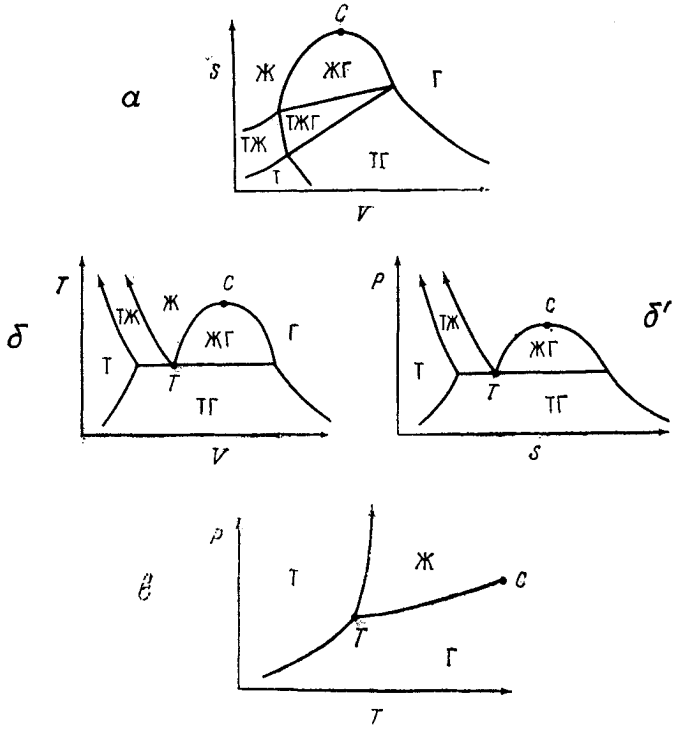


Рис. 10.22. Двумерное многообразие уравнения состояния простого, однокомпонентного вещества в $\mathbb{R}^4$ можно спроектировать на любое двумерное подпространство.

Показаны проекции на эти подпространства, порождаемые парами несопряженных термодинамических переменных: *а* — две экстенсивные переменные (S, V); *б* и *б'* — экстенсивная и не сопряженная ей интенсивная переменная, *в* — две интенсивные переменные (T, P).

Соответствие между формулировкой термодинамики равновесных систем в терминах теории катастроф и классической гиббсовской формулировкой устанавливается следующим обра-

зом. Критические точки потенциала $\mathcal{U}(i_\alpha; E^\beta)$ определяются из равенства нулю его градиента:

$$\nabla_i \mathcal{U}(i_\alpha; E^\beta) = 0, \quad 1 \leqslant \alpha, \beta \leqslant n. \tag{10.66}$$

Решая эти n уравнений состояния, можно найти интенсивные переменные i_α как функции экстенсивных переменных для каждого локального минимума (остальные критические точки не представляют интереса):

$$i_\alpha = i_\alpha^{(r)}(E^\beta), \quad r = 1, \ldots . \tag{10.67}$$

В каждом локальном минимуме значение потенциала $\mathcal{U}$ является корректно определенной функцией экстенсивных управляющих параметров:

$$U^{(r)}(E^\beta) = \mathcal{U}(i_\alpha^{(r)}(E^\beta); E^\beta). \tag{10.68i}$$

Состояние системы определяется из условия глобального минимума

$$U(E^\beta) = \min_{r=1, \ldots} U^{(r)}(E^\beta). \tag{10.68ii}$$

В силу (10.67) это условие минимума определяет значения интенсивных переменных. Гиббсовская термодинамическая производящая функция $U(E^\beta)$ равна потенциальной функции $\mathcal{U}(i_\alpha; E^\beta)$, вычисленной на критическом многообразии (определяемом уравнениями (10.66)) и выбранной в соответствии с принципом Максвелла.

В классической гиббсовской формулировке термодинамики экстенсивные и сопряженные им интенсивные термодинамические переменные связаны через градиент производящей функции:

$$i_\alpha = \frac{\partial U}{\partial E^\alpha}. \tag{10.69}$$

Окончательное соответствие можно установить, если потребовать выполнения равенства [см. (5.3, 1)]

$$i_\alpha = \frac{\partial \mathcal{U}(i; E)}{\partial E^\alpha}\bigg|_{CM}, \tag{10.70}$$

где индекс CM означает, что производные вычисляются на критическом многообразии. Два определения (10.66) и (10.70) интенсивных переменных налагают определенные ограничения на структуру потенциала $\mathcal{U}(i; E)$.

Фазовые переходы имеют место на «плоском» участке критического многообразия. Эти плоские участки проектируются на максвелловское множество в пространстве управляющих параметров. Поэтому максвелловское множество зависит от выбора

управляющих параметров, и для его определения можно воспользоваться уравнениями Клаузиуса — Клапейрона (10.68). Если в состоянии термодинамического равновесия в системе сосуществуют t различных фаз, то уравнения Клаузиуса — Клапейрона определяются из соотношений

$$\mathcal{U}_\alpha^{(1)}\,\delta E^\alpha = \mathcal{U}_\alpha^{(2)}\,\delta E^\alpha = \ldots = \mathcal{U}_\alpha^{(t)}\,\delta E^\alpha, \qquad (10.71)$$

где все производные вычисляются на критическом многообразии. Поскольку экстенсивные переменные однозначно описывают состояние системы и их вариации произвольны, то уравнения Клаузиуса — Клапейрона принимают вид

$$i_\alpha^{(1)} = i_\alpha^{(2)} = \ldots = i_\alpha^{(t)}. \qquad (10.72)$$

Иными словами, если в состоянии термодинамического равновесия системы сосуществуют t фаз, то каждая интенсивная термодинамическая переменная (T, P, ...) в каждой фазе принимает одно и то же значение. Это не что иное, как перенос условий равновесия Ньютона для механических систем на термодинамические системы. По этой причине стало обычным делом трактовать интенсивные термодинамические переменные как обобщенные силы, а экстенсивные — как обобщенные перемещения. Очевидно, что если в качестве управляющих переменных используется иной набор термодинамических переменных (p интенсивных и q экстенсивных, где $p+q=n$), то максвелловское множество оказывается иным, и уравнения Клаузиуса — Клапейрона (5.68) приводят к иным условиям сосуществования фаз.

12. МЕТРИЧЕСКАЯ ГЕОМЕТРИЯ И ТЕРМОДИНАМИКА

Изменения переменных состояния $i_\alpha \to i_\alpha + \delta i_\alpha$ и управляющих параметров ($E^\beta \to E^\beta + \delta E^\beta$) вызывают изменение величины потенциала $\mathcal{U}$:

$$\mathcal{U}(i_\alpha + \delta i_\alpha;\ E^\beta + \delta E^\beta) = \mathcal{U}^{(0)} + \delta^{(1)}\mathcal{U} + \delta^{(2)}\mathcal{U} + \ldots, \qquad (10.73)$$

причем члены первого $\delta^{(1)}\mathcal{U}$ и второго $\delta^{(2)}\mathcal{U}$ порядка даются выражениями (5.1). Если управляющие параметры изменяются достаточно медленно («квазистатически»), так что система все время остается в состоянии термодинамического равновесия, то величины δi_α и δE^β связаны соотношением (5.2). При этих условиях изменение значения потенциала $\mathcal{U}$ на критическом многообразии (равного производящей функции U) определяется как

$$U(E^\beta + \delta E^\beta) = U^{(0)} + \delta^{(1)}U + \delta^{(2)}U + \ldots, \qquad (10.74)$$

причем члены первого $\delta^{(1)}U$ и второго $\delta^{(2)}U$ порядка даются выражением (5.3). В частности, изменения потенциала $\mathcal{U}$ и производящей функции связаны следующим образом:

$$\begin{aligned} \delta^{(1)}\mathcal{U} &= \mathcal{U}^r\,\delta i_r + \mathcal{U}_\alpha\,\delta E^\alpha \\ &\downarrow \text{см} \\ \delta^{(1)}U &= \qquad\quad U_\alpha\,\delta E^\alpha \end{aligned} \tag{10.75i}$$

и

$$\begin{aligned} \delta^{(2)}\mathcal{U} &= \tfrac{1}{2}\,\mathcal{U}^{rs}\,\delta i_r\,\delta i_s + \mathcal{U}^r_\beta\,\delta i_r\,\delta E^\beta + \tfrac{1}{2}\,\mathcal{U}_{\alpha\beta}\,\delta E^\alpha\,\delta E^\beta, \\ \delta^{(2)}U &= \qquad\qquad\qquad\qquad\quad \tfrac{1}{2}\,U_{\alpha\beta}\,\delta E^\alpha\,\delta E^\beta, \end{aligned} \tag{10.75ii}$$

где

$$U_{\alpha\beta} = \mathcal{U}_{\alpha\beta} - \mathcal{U}^r_\alpha\,(\mathcal{U}^{-1})_{rs}\,\mathcal{U}^s_\beta. \tag{10.76}$$

Здесь $\mathcal{U}^r = \partial\mathcal{U}/\partial i_r$, $\mathcal{U}^r_\beta = \partial^2\mathcal{U}/\partial i_r \partial E^\beta$ и т. д.

Выражения для членов первого $\delta^{(1)}U$ и второго $\delta^{(2)}U$ порядка тесно связаны с первым и вторым началом термодинамики. С учетом (10.75i) член первого порядка можно записать в виде

$$\delta^{(1)}U = U_\alpha\,\delta E^\alpha = T\,dS - P\,dV + \mu_l\,dN_l\,\ldots\,. \tag{10.77i}$$

Это не что иное, как *первое начало термодинамики.* Второе начало термодинамики можно сформулировать в виде некоторого условия устойчивости, а именно:

$$\frac{\partial^2 U}{\partial E^\alpha\,\partial E^\beta}\,\delta E^\alpha\,\delta E^\beta = U_{\alpha\beta}\,\delta E^\alpha\,\delta E^\beta \geqslant 0, \tag{10.77ii}$$

где равенство имеет место лишь тогда, когда все приращения δE^α равны нулю. Короче говоря, *второе начало термодинамики* эквивалентно требованию положительной определенности метрического тензора $U_{\alpha\beta}$.

◇◇◇ Первые производные производящей функции U_α и потенциала $\mathcal{U}_\alpha$ на критическом многообразии равны $U_\alpha = \mathcal{U}_\alpha|_{\text{см}}$. Для вторых производных это не так. Поскольку критическое многообразие определяется из условия минимума $\mathcal{U}$, то $\mathcal{U}^{rs}$ на этом многообразии положительно определена, как и функция $(\mathcal{U}^{-1})_{rs}$. Поэтому $\mathcal{U}^r_\alpha\,(\mathcal{U}^{-1})_{rs}\mathcal{U}^s_\beta$ положительно определена, и $\mathcal{U}_{\alpha\beta}$ «более положительно определена», чем $U_{\alpha\beta}$ в том смысле, что

$$(\mathcal{U}_{\alpha\beta} - U_{\alpha\beta})\,\delta E^\alpha\,\delta E^\beta \geqslant 0, \tag{10.78}$$

причем равенство имеет место только тогда, когда все приращения δE^α равны нулю.

При изменении экстенсивных параметров $E \to E + \delta E$, производящая функция $U(E)$ становится равной

$$U = U^{(0)} + U_\alpha \delta E^\alpha + \frac{1}{2} U_{\alpha\beta} \delta E^\alpha \delta E^\beta + \ldots, \qquad (10.79)$$

где $U_\alpha = i_\alpha$. Это выражение позволяет сделать ряд важных выводов. Во-первых, классическую термодинамику можно построить, исходя из условного вариационного принципа, согласно которому ищется минимум функции Лагранжа $U - \lambda_\alpha E^\alpha$

$$\delta(U - \lambda_\alpha E^\alpha) = \delta\left(U^0 + \frac{1}{2} U_{\alpha\beta} \delta E^\alpha \delta E^\beta + \ldots\right) = 0. \qquad (10.80)$$

Сравнение с (10.69) немедленно приводит к отождествлению множителей Лагранжа λ_α с интенсивными переменными i_α:

$$\lambda_\alpha = i_\alpha \qquad (10.81)$$

в точке условного минимума.

Далее, вычисляя $\partial U/\partial E^\alpha$, находим

$$\frac{\partial U}{\partial E^\alpha} = U_\alpha + U_{\alpha\beta} \delta E^\beta = i_\alpha + \delta i_\alpha. \qquad (10.82)$$

Следовательно, положительно определенный метрический тензор $U_{\alpha\beta}$ можно интерпретировать как тензор восприимчивости, поскольку он описывает равновесный отклик δi_α интенсивной переменной i_α на изменение δE^β экстенсивной переменной E^β:

$$\delta i_\alpha = U_{\alpha\beta} \delta E^\beta. \qquad (10.83)$$

Поскольку $U_{\alpha\beta}$ положительно определена, а следовательно, не вырождена, ее можно обратить. Таким образом, отклик экстенсивных переменных на изменение интенсивных переменных имеет вид

$$\delta E^\beta = U^{\beta\alpha} \delta i_\alpha \quad (U_{\alpha\beta} U^{\beta\gamma} = \delta_\alpha^\gamma). \qquad (10.84)$$

Изменение $\delta^{(2)} U$ теперь можно записать тремя удобными способами:

$$\begin{aligned} 2\delta^{(2)} U &= U_{\alpha\beta} \delta E^\alpha \delta E^\beta, \\ &= \delta i_\alpha \delta E^\alpha, \\ &= \delta i_\alpha \delta i_\beta U^{\alpha\beta}, \end{aligned} \qquad (10.85)$$

а (10.79) можно переписать в более удобной форме:

$$U = U^{(0)} + i_\alpha \delta E^\alpha + \frac{1}{2} \delta i_\alpha \delta E^\alpha + \ldots. \qquad (10.86)$$

◇◇◇ В общем случае тензор восприимчивости, описывающий инфинитезимальный линейный отклик переменных состояния на

бесконечно малые изменения управлений, дается формулой (5.2), а в данном частном случае — формулой (10.76). Поскольку эти восприимчивости должны быть равными на критическом многообразии, должны выполняться равенства:

$$U_{\alpha\beta} = (\mathcal{U}_{\alpha\beta} - \mathcal{U}^r_\alpha (\mathcal{U}^{-1})_{rs} \mathcal{U}^s_\beta) = -(\mathcal{U}^{-1})_{\alpha s} \mathcal{U}^s_\beta. \quad (10.87)$$

Это выражение можно переписать в более простом виде

$$\mathcal{U}_{\alpha\beta} = (\mathcal{U}^r_\alpha - \delta^r_\alpha)(\mathcal{U}^{-1})_{rs} \mathcal{U}^s_\beta. \quad (10.87')$$

Уравнения (10.70) и (10.87) представляют собой единственные два ограничения, наложенные на термодинамический потенциал $\mathcal{U}$ с тем, чтобы воспроизвести термодинамику классической производящей функции U.

Пример. В случае простого, однокомпонентного вещества наиболее легко измерить следующие термодинамические функции отклика:

$$\begin{aligned} C_V &= T\left(\frac{\partial S}{\partial T}\right)_V, & C_P &= T\left(\frac{\partial S}{\partial T}\right)_P, \\ \beta_S &= -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_S, & \beta_T &= -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T, \\ \Gamma_V &= T\left(\frac{\partial S}{\partial P}\right)_V, & \alpha_P &= \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P. \end{aligned} \quad (10.88)$$

Здесь C_V и C_P — теплоемкости при постоянном объеме и давлении, β_s и β_T — адиабатическая (постоянное S) и изотермическая (постоянное T) сжимаемости, Γ_V — количество тепла, приходящееся на единицу объема при изменении давления, и α_P — коэффициент теплового расширения.

Чтобы установить связь этих термодинамических функций отклика с метрическими тензорами $U_{\alpha\beta}$ и $U^{\alpha\beta}$, определим ковариантные компоненты, записав отклик интенсивных переменных на изменение экстенсивных переменных:

$$\begin{bmatrix} di_1 \\ di_2 \end{bmatrix} = \begin{bmatrix} dT \\ -dP \end{bmatrix} = \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} dS \\ dV \end{bmatrix} = U_{\alpha\beta} \begin{bmatrix} dE^1 \\ dE^2 \end{bmatrix}. \quad (10.89)$$

При постоянном объеме ($dV = 0$) и постоянной энтропии ($dS = 0$) соответственно имеем

$$\begin{aligned} dT = A\,dS &\Rightarrow \frac{1}{A} = \left(\frac{\partial S}{\partial T}\right)_V, \\ -dP = B\,dS &\Rightarrow \frac{1}{B} = -\left(\frac{\partial S}{\partial P}\right)_V, \end{aligned} \quad (10.90)$$

$$\begin{aligned} dT = B\,dV &\Rightarrow \frac{1}{B} = \left(\frac{\partial V}{\partial T}\right)_S, \\ -dP = C\,dV &\Rightarrow \frac{1}{C} = -\left(\frac{\partial V}{\partial P}\right)_S. \end{aligned} \quad (10.91)$$

Хотя коэффициент адиабатического расширения $\alpha_S = (1/V)(\partial V/\partial T)_S$ и не входит в число функций реакции, перечисленных в (10.88), в силу симметрии

$U_{\alpha\beta} = U_{\beta\alpha}$ он тесно связан с величиной Γ_V:

$$U_{\alpha\beta} = \begin{bmatrix} \frac{T}{C_V} & -\frac{T}{\Gamma_V} \\ -\frac{T}{\Gamma_V} & \frac{1}{V\beta_S} \end{bmatrix}. \tag{10.92}$$

Контравариантный тензор восприимчивости $U^{\alpha\beta}$ вычисляется проще:

$$\begin{bmatrix} dE^1 \\ dE^2 \end{bmatrix} = \begin{bmatrix} dS \\ dV \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial S}{\partial T}\right)_P & -\left(\frac{\partial S}{\partial P}\right)_T \\ \left(\frac{\partial V}{\partial T}\right)_P & -\left(\frac{\partial V}{\partial P}\right)_T \end{bmatrix} \begin{bmatrix} dT \\ -dP \end{bmatrix} = U^{\alpha\beta} \begin{bmatrix} di_1 \\ di_2 \end{bmatrix},$$

$$U^{\alpha\beta} = \begin{bmatrix} \frac{C_P}{T} & V\alpha_P \\ V\alpha_P & V\beta_T \end{bmatrix}. \tag{10.93}$$

Поскольку ковариантный и контравариантный метрические тензоры $U_{\alpha\beta}$ и $U^{\alpha\beta}$ являются взаимно обратными матрицами, то

$$\frac{1}{\frac{C_P}{T} V\beta_T - (V\alpha_P)^2} \begin{bmatrix} V\beta_T & -V\alpha_P \\ -V\alpha_P & \frac{C_P}{T} \end{bmatrix} = \begin{bmatrix} \frac{T}{C_V} & -\frac{T}{\Gamma_V} \\ -\frac{T}{\Gamma_V} & \frac{1}{V\beta_S} \end{bmatrix}. \tag{10.94}$$

Из шести функций, перечисленных в (10.88), только три независимые [4].

Существование на критическом многообразии положительно определенного метрического тензора $U_{\alpha\beta}$ позволяет сделать ряд полезных выводов относительно общих свойств термодинамических систем; математически эти выводы могут быть представлены в виде равенств и неравенств. Хотя можно было бы проанализировать эти свойства непосредственно на многообразии уравнения состояния, удобнее перейти к n-мерному линейному векторному пространству, касательному к критическому многообразию в точке равновесия.

Пусть $\mathbf{e}^\alpha$ порождает пространство $\mathbb{R}^n$ переменных состояния и $\mathbf{f}_\beta$ — пространство $\mathbb{R}^n$ управляющих параметров. Тогда произвольная точка из $\mathbb{R}^n \otimes \mathbb{R}^n$ может быть представлена вектором $i_\alpha \mathbf{e}^\alpha + E^\beta \mathbf{f}_\beta$, и n векторов, касательных к критическому многообразию, в силу (10.16) имеют вид

$$|v_\beta\rangle = \mathbf{f}_\beta + \mathbf{e}^\alpha U_{\alpha\beta}. \tag{10.95}$$

Векторы $|v_\beta\rangle$ можно рассматривать как линейные комбинации векторов $\mathbf{e}^\alpha$, $\mathbf{f}_\beta$, определенных в начале $\mathbb{R}^n \otimes \mathbb{R}^n$ и параллельно перенесенных в точку, лежащую на критическом многообразии. Векторы $|v_\beta\rangle$ порождают касательное пространство.

Бесконечно малое перемещение δv с координатами $(\delta i_\alpha, \delta E^\beta)$ на равновесном многообразии можно идентифицировать с бесконечно малым вектором $\delta v = \delta E^\beta |v_\beta\rangle$ в касательном пространстве, поскольку

$$\delta E^\beta |v_\beta\rangle = \delta E^\beta (\mathbf{f}_\beta + U_{\beta\alpha}\mathbf{e}^\alpha) = \delta i_\alpha \mathbf{e}^\alpha + \delta E^\beta \mathbf{f}_\beta. \tag{10.96}$$

В некоторых случаях может потребоваться выразить бесконечно малый вектор (10.96) через приращения δi_α. Это можно сделать, введя в касательном

пространстве второй набор базисных векторов $|v^\alpha\rangle$ так, что

$$\overset{\delta v}{\diagup\!\!\diagup \quad \diagdown\!\!\diagdown}$$
$$\delta E^\beta |v_\beta\rangle = \delta i_\alpha |v^\alpha\rangle. \tag{10.97}$$

Связь базисных векторов $|v_\beta\rangle$ с базисными векторами $|v^\alpha\rangle$ легко определить из соотношения между компонентами:

$$\delta i_\alpha = U_{\alpha\beta}\delta E^\beta \Leftrightarrow |v^\alpha\rangle U_{\alpha\beta} = |v_\beta\rangle. \tag{10.98}$$

Скалярное произведение в касательном пространстве можно получить, идентифицируя $(\delta v, \delta v)$ с $2\delta^{(2)}U$:

$$\overset{(\delta v,\ \delta v)}{\diagup\!\!\diagup \quad \diagdown\!\!\diagdown}$$
$$\delta i_\alpha \delta i_\beta \langle v^\alpha | v^\beta\rangle \| \langle v_\alpha | v_\beta\rangle \delta E^\alpha \delta E^\beta. \tag{10.99}$$
$$\delta i_\alpha \delta E^\beta \langle v^\alpha | v_\beta\rangle$$

Сравнивая с (10.85), получаем

$$\begin{aligned} \langle v_\alpha | v_\beta\rangle &= U_{\alpha\beta}, \\ \langle v_\alpha | v^\beta\rangle &= \delta_\alpha^\beta, \\ \langle v^\alpha | v^\beta\rangle &= U^{\alpha\beta}. \end{aligned} \tag{10.100}$$

Короче говоря, базисные векторы $|v^\alpha\rangle$ и $|v_\beta\rangle$ двойственны по отношению друг к другу.

◇◇◇ Свойства метрики в касательном пространстве изучались ранее Вейнхолдом [4]. Связь между двойственным набором базисных векторов $|i_\alpha\rangle$, $|E^\beta\rangle$ Вейнхолда и векторами, введенными в этом разделе, имеет вид

$$|v_\alpha\rangle \leftrightarrow |i_\alpha\rangle, \quad |v^\beta\rangle \leftrightarrow |E^\beta\rangle. \tag{10.101}$$

Существование в векторном пространстве положительно определенного скалярного произведения приводит к неравенствам, основанным на неравенстве Шварца, и к равенствам и неравенствам, вытекающим из неравенства Бесселя.

Неравенство Шварца. Если $\mathbf{u}$, $\mathbf{v}$ — произвольные векторы линейного векторного пространства с положительно определенным скалярным произведением $(\cdot, \cdot)$, то

$$(\mathbf{u}, \mathbf{v})^2 \leqslant (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v}). \tag{10.102}$$

Конкретный выбор $\mathbf{u}$ и $\mathbf{v}$ приводит к следующим неравенствам:

$\mathbf{u}$	$\mathbf{v}$	
$\|v^\alpha\rangle$	$\|v^\beta\rangle$	$(U^{\alpha\beta})^2 \leqslant U^{\alpha\alpha}U^{\beta\beta}$,
$\|v_\alpha\rangle$	$\|v^\beta\rangle$	$\delta_\alpha^\beta \leqslant U_{\alpha\alpha}U^{\beta\beta}$,
$\|v_\alpha\rangle$	$\|v_\beta\rangle$	$(U_{\alpha\beta})^2 \leqslant U_{\alpha\alpha}U_{\beta\beta}$.

(10.103)

Для однокомпонентной жидкости метрические тензоры $U_{\alpha\beta}$, $U^{\alpha\beta}$ даются формулами (10.92) и (10.93). Из неравенства Шварца (10.102) вытекают следующие неравенства для функций отклика (10.88):

$$\alpha_P^2 \leqslant \frac{C_P \beta_T}{TV}, \quad 1 \leqslant \frac{C_P}{C_V}, \quad \frac{1}{\Gamma_V^2} \leqslant \frac{1}{TVC_V\beta_S}, \quad 1 \leqslant \frac{\beta_T}{\beta_S}. \tag{10.104}$$

Неравенство Бесселя. Если $\mathbf{u}_1, \mathbf{u}_2, \ldots$ образуют систему ортонормированных векторов в пространстве с положительно определенным скалярным произведением и $\mathbf{w}$ — произвольный вектор, то

$$(\mathbf{w}, \mathbf{w}) \geqslant \sum_i (\mathbf{u}_i, \mathbf{w})^2. \tag{10.105}$$

Это неравенство выполняется, если вектор $\mathbf{w}$ лежит в пространстве, порожденном векторами $\mathbf{u}_i$, и, в частности, если $\mathbf{u}_i$ порождают линейное векторное пространство. Поскольку $|v_\alpha\rangle$, $|v^\beta\rangle$ образуют двойственную систему, можно выбрать $\mathbf{u}_1 = |v_1\rangle/(U_{11})^{1/2}$ и $\mathbf{u}_2 = |v^2\rangle/(U^{22})^{1/2}$. Тогда $\mathbf{u}_1$ и $\mathbf{u}_2$ ортонормированы. Конкретный выбор $\mathbf{w}$ приводит к неравенствам

$$\begin{array}{lll} \mathbf{w} & & \\ |v^1\rangle & U^{11} \geqslant \dfrac{1}{U_{11}} + \dfrac{(U^{12})^2}{U^{22}}, & \\ |v_2\rangle & U_{22} \geqslant \dfrac{(U_{12})^2}{U_{11}} + \dfrac{1}{U^{22}}, & \\ |v^3\rangle & U^{33} \geqslant 0 + \dfrac{(U^{32})^2}{U^{22}}. & \end{array} \tag{10.106}$$

Для однокомпонентного вещества $n = 2$, поэтому $\mathbf{u}_1$, $\mathbf{u}_2$ порождают касательную плоскость, и неравенства обращаются в равенства, т. е.

$$\begin{aligned} \frac{1}{\beta_S} &= \frac{TVC_V}{\Gamma_V^2} + \frac{1}{\beta_T}, & \beta_T &= \beta_S + \frac{TV\alpha_P^2}{C_P}, \\ \frac{1}{C_V} &= \frac{TV\beta_S}{\Gamma_V^2} + \frac{1}{C_P}, & C_P &= C_V + \frac{TV\alpha_P^2}{\beta_T}. \end{aligned} \tag{10.107}$$

13. ЧАСТНЫЕ ПРОИЗВОДНЫЕ ТЕРМОДИНАМИЧЕСКИХ ВЕЛИЧИН

Длительное время вычисление частных производных термодинамических величин вызывало существенные затруднения и являлось камнем преткновения для многих поколений специа-

листов в различных областях науки и техники, изучавших термодинамику.

Классический способ определения этих производных состоит во введении ряда дополнительных термодинамических производящих функций, связанных с внутренней энергией U преобразованием Лежандра. Используя равенство вторых смешанных частных производных этих производящих функций, можно получить частные производные многих термодинамических величин. Однако в этом случае все зависит от удачного выбора производящих функций, который является своего рода озарением и, пожалуй, граничит с черной магией.

Процесс вычисления частных производных термодинамических величин можно существенно упростить, если учесть, что они представляют собой не что иное, как коэффициенты восприимчивости, характеризующие отклик одной термодинамической переменной на изменение другой при условии, что остальные переменные остаются фиксированными. Пространство, касательное к критическому многообразию, содержит всю информацию об отклике системы, находящейся в состоянии равновесия, на малые изменения термодинамических переменных. Точки в этом n-мерном касательном пространстве можно параметризовать любыми n из $2n$ термодинамических переменных $(i_\alpha; E^\beta)$, если, конечно, они независимые. Тогда вычисление частных производных сводится к изменению независимых координат, описывающих положение точки в касательном пространстве, что в свою очередь является довольно простой задачей алгебры линейных векторных пространств.

Проиллюстрируем эту общую схему на простом примере для $n = 2$.

Пример: $(\partial E^1/\partial E^2)_{i_2} = ?$

Решение. Тензор восприимчивости можно записать как

$$\delta E^\alpha = U^{\alpha\beta}\,\delta i_\beta. \tag{10.108i}$$

В матричной форме эта запись принимает вид

$$\begin{bmatrix} \delta E^1 \\ \delta E^2 \end{bmatrix} = \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} \delta i_1 \\ \delta i_2 \end{bmatrix}. \tag{10.108ii}$$

Данное уравнение «несимметрично» в том смысле, что экстенсивные и интенсивные переменные трактуются по-разному. Уравнение

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta E^1 \\ \delta E^2 \end{bmatrix} = \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} \delta i_1 \\ \delta i_2 \end{bmatrix} \tag{10.108iii}$$

более симметрично. Теперь надо решить, какие две из четырех термодинамических переменных следует считать независимыми. Поскольку требуется вычислить $(\partial E^1/\partial E^2)_{i_2}$, целесообразно считать приращения δE^2 и δi_2 независимыми. Перенесем независимые приращения в одну часть уравнения, оставив в другой зависимые реакции. Для этого достаточно поменять местами

соответствующие столбцы квадратных матриц (10.108iii):

$$\begin{bmatrix} 1 & -A \\ 0 & -B \end{bmatrix}\begin{bmatrix} \delta E^1 \\ \delta i_1 \end{bmatrix}=\begin{bmatrix} 0 & B \\ -1 & C \end{bmatrix}\begin{bmatrix} \delta E^2 \\ \delta i_2 \end{bmatrix}. \qquad (10.108\text{iv})$$

Выполнив операции обращения и перемножения матриц, решим это уравнение относительно зависимых приращений:

$$\begin{bmatrix} \delta E^1 \\ \delta i_1 \end{bmatrix}=\begin{bmatrix} \dfrac{A}{B} & \dfrac{B^2-AC}{B} \\ \dfrac{1}{B} & -\dfrac{C}{B} \end{bmatrix}\begin{bmatrix} \delta E^2 \\ \delta i_2 \end{bmatrix}. \qquad (10.108\text{v})$$

В результате получим $(\partial E^1/\partial E^2)_{i_2} = A/B$. Из (10.93) следует, что для однокомпонентного вещества $(\partial S/\partial V)_P = (C_P/T)/(V\alpha_P)$ и $(\partial V/\partial S)_T = \beta_T/\alpha_P$. (Эти производные невозможно получить из соотношений Максвелла.)

В табл. 10.2 приводятся различные тензоры восприимчивости, которые могут возникнуть при описании системы двумя интенсивными и двумя экстенсивными термодинамическими переменными. Независимыми переменными могут быть: две интенсивные переменные, две экстенсивные переменные, одна экстенсивная и несопряженная интенсивная переменные, пара сопряженных переменных. Тензоры восприимчивости симметричны, только если все независимые переменные интенсивные или все экстенсивные.

При вычислении частных производных термодинамических величин необходимо:

— записать тензор восприимчивости в матричной форме $\delta^\alpha_\beta \delta E^\beta = U^{\alpha\beta}\delta i_\beta$;

— выбрать независимые приращения, перенести их в правую часть уравнения, оставив зависимые переменные в левой части. Для этого необходимо переставить соответствующие столбцы и изменить их знак;

— выполнить операции обращения и перемножения матриц и, таким образом, разрешить уравнения относительно зависимых приращений;

— если необходимо, произвести преобразование базиса, введя в него линейную комбинацию независимых приращений.

Пример. Вычислить $(\partial S/\partial T)_{M,\,\sigma}$ для магнитного однокомпонентного вещества, учитывая магнитное поле только в направлении z и определяя частную производную вдоль кривой равновесия.

Решение. Данная система описывается тремя парами сопряженных переменных: (S, T), $(V, -P)$, (M, H). Если m — наклон кривой равновесия, то вдоль этой кривой

$$\frac{dP}{dT} = m \Leftrightarrow 0 = d\sigma = -dP + m\,dT. \qquad (10.109)$$

Поэтому в качестве независимых приращений следует выбрать dT, $d\sigma$, dM. Поскольку $d\sigma$ является линейной комбинацией $-dP$ и dT, то в качестве

независимых приращений выберем dT, $-dP$ и dM. Ниже приводятся шаги описанного выше алгоритма

Шаг 1:
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dE^1 \\ dE^2 \\ dE^3 \end{bmatrix} = \begin{bmatrix} A & B & C \\ B & D & E \\ C & E & F \end{bmatrix} \begin{bmatrix} di_1 \\ di_2 \\ di_3 \end{bmatrix}. \tag{10.110i}$$

Шаг 2:
$$\begin{bmatrix} 1 & 0 & -C \\ 0 & 1 & -E \\ 0 & 0 & -F \end{bmatrix} \begin{bmatrix} dE^1 \\ dE^2 \\ di_3 \end{bmatrix} = \begin{bmatrix} A & B & 0 \\ B & D & 0 \\ C & E & -1 \end{bmatrix} \begin{bmatrix} di_1 \\ di_2 \\ dE^3 \end{bmatrix}. \tag{10.110ii}$$

Шаг 3:
$$\begin{bmatrix} dS \\ dV \\ dH \end{bmatrix} = \begin{bmatrix} A - \frac{C^2}{F} & B - \frac{CE}{F} & \frac{C}{F} \\ B - \frac{CE}{F} & D - \frac{E^2}{F} & \frac{E}{F} \\ -\frac{C}{F} & -\frac{E}{F} & \frac{1}{F} \end{bmatrix} \begin{bmatrix} dT \\ -dP \\ dM \end{bmatrix}. \tag{10.110iii}$$

Шаг 4:
$$\begin{bmatrix} dT \\ d\sigma \\ dM \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ m & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dT \\ -dP \\ dM \end{bmatrix}. \tag{10.110iv}$$

Шаг 5:
$$\begin{bmatrix} dS \\ dV \\ dH \end{bmatrix} = \begin{bmatrix} A - \frac{C^2}{F} & B - \frac{CE}{F} & \frac{C}{F} \\ B - \frac{CE}{F} & D - \frac{E^2}{F} & \frac{E}{F} \\ -\frac{C}{F} & -\frac{E}{F} & \frac{1}{F} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -m & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dT \\ d\sigma \\ dM \end{bmatrix}, \tag{10.110v}$$

$$\left(\frac{\partial S}{\partial T}\right)_{\sigma, M} = \left(A - \frac{C^2}{F}\right) - m\left(B - \frac{CE}{F}\right). \tag{10.110vi}$$

Вычисления элементов матрицы A, B, C и т. д. не вызывает никаких затруднений. Например, $B = -(\partial S/\partial P)_{T, H} = (\partial V/\partial T)_{P, H}$ — это одна из функций реакции, значение которой легко определить экспериментально.

◇◇◇ Приведенный алгоритм является простым, наглядным и более широко применяется для вычисления частных производных термодинамических функций отклика, чем методы, основанные на соотношениях Максвелла. Кроме того, последние позволяют вычислить компоненты тензора восприимчивости только поочередно, в то время как предложенный алгоритм дает все элементы тензора сразу.

Метрический тензор совпадает с тензором восприимчивости только в том случае, когда все независимые приращения одного типа: интенсивные или экстенсивные. Если независимые приращения разных типов, тензор восприимчивости, вообще говоря, несимметричен (табл. 10.2), однако метрический тензор должен оставаться симметричным. В случае такой смешанной

Таблица 10.2. **Тензоры восприимчивости в случае описания равновесного состояния системы двумя интенсивными и двумя экстенсивными термодинамическими переменными**

$$\begin{bmatrix}\delta E^1\\ \delta E^2\end{bmatrix}=\begin{bmatrix}A & B\\ B & C\end{bmatrix}\begin{bmatrix}\delta i_1\\ \delta i_2\end{bmatrix}$$

$$\begin{bmatrix}\delta E^1\\ \delta i_1\end{bmatrix}=\begin{bmatrix}\dfrac{A}{B} & \dfrac{B^2-AC}{B}\\ \dfrac{1}{B} & -\dfrac{C}{B}\end{bmatrix}\begin{bmatrix}\delta E^2\\ \delta i_2\end{bmatrix}$$

$$\begin{bmatrix}\delta E^1\\ \delta i_2\end{bmatrix}=\begin{bmatrix}\dfrac{AC-B^2}{C} & \dfrac{B}{C}\\ -\dfrac{B}{C} & \dfrac{1}{C}\end{bmatrix}\begin{bmatrix}\delta i_1\\ \delta E^2\end{bmatrix}$$

$$\begin{bmatrix}\delta i_1\\ \delta i_2\end{bmatrix}=\frac{1}{AC-B^2}\begin{bmatrix}C & -B\\ -B & A\end{bmatrix}\begin{bmatrix}\delta E^1\\ \delta E^2\end{bmatrix}$$

координатной системы вычисление метрического тензора не составляет труда при условии, что

$$2\delta^{(2)}U=\delta i_\alpha\,\delta E^\alpha. \tag{10.111}$$

Это выражение необходимо переписать в терминах независимых приращений, используя тензор восприимчивости, вычисленный по описанному выше алгоритму.

Пример. Вычислить метрический тензор в системе координат с независимыми приращениями δE^1 и δi_2.

Решение. Зависимые приращения δi_1 и δE^2 можно выразить через δE^1 и δi_2, используя обращение соответствующего тензора восприимчивости, приведенного в табл. 10.2:

$$\delta i_1\,\delta E^1+\delta i_2\,\delta E^2=\left[\frac{1}{A}\,\delta E^1-\frac{B}{A}\,\delta i_2\right]\delta E^1+\delta i_2\left[\frac{B}{A}\,\delta E^1+\left(\frac{AC-B^2}{A}\right)\delta i_2\right]=$$

$$=(\delta E^1,\ \delta i_2)\begin{bmatrix}\dfrac{1}{A} & 0\\ 0 & \dfrac{AC-B^2}{A}\end{bmatrix}\begin{bmatrix}\delta E^1\\ \delta i_2\end{bmatrix}. \tag{10.112}$$

Таблица 10.3 метрических тензоров является аналогом табл. 10.2 тензоров восприимчивости.

Таблица 10.3. **Метрические тензоры в случае описания равновесного состояния системы двумя интенсивными и двумя экстенсивными переменными**

$$(\delta i_1,\ \delta i_2)\begin{bmatrix}A & B\\ B & C\end{bmatrix}\begin{bmatrix}\delta i_1\\ \delta i_2\end{bmatrix}$$

$$(\delta E^2,\ \delta i_2)\begin{bmatrix} \dfrac{A}{B^2} & \dfrac{B^2-AC}{B^2} \\ \dfrac{B^2-AC}{B^2} & C\,\dfrac{AC-B^2}{B^2} \end{bmatrix}\begin{bmatrix} \delta E^2 \\ \delta i_2 \end{bmatrix}$$

$$(\delta i_1,\ \delta E^2)\begin{bmatrix} \dfrac{AC-B^2}{C} & 0 \\ 0 & \dfrac{1}{C} \end{bmatrix}\begin{bmatrix} \delta i_1 \\ \delta E^2 \end{bmatrix}$$

$$(\delta E^1,\ \delta E^2)\begin{bmatrix} \dfrac{C}{AC-B^2} & \dfrac{-B}{AC-B^2} \\ \dfrac{-B}{AC-B^2} & \dfrac{A}{AC-B^2} \end{bmatrix}\begin{bmatrix} \delta E^1 \\ \delta E^2 \end{bmatrix}$$

14. ВОЗМОЖНЫЕ ВАРИАЦИОННЫЕ ПРЕДСТАВЛЕНИЯ

Термодинамические переменные, которые легче всего измерить экспериментально, далеко не всегда являются экстенсивными. По этой причине было бы целесообразно попытаться использовать теорию катастроф для построения такой теории термодинамики равновесных систем, в которой в качестве управляющих параметров выступали бы термодинамические переменные, наиболее легко контролируемые в условиях эксперимента. Например, для однокомпонентного вещества в качестве управляющих параметров можно было бы выбрать (T, V) или (T, P), а не (S, V). Подобную вариационную формулировку основных законов термодинамики, основанную на выборе в качестве управляющих параметров сочетания интенсивных и экстенсивных термодинамических переменных, а в качестве переменных состояния — им сопряженных, можно легко вывести из формулировки, данной в разд. 2. Для этого добавим к потенциалу $\mathcal{U}(i_\alpha;\ E^\beta)$ функцию f от интенсивных и экстенсивных переменных. Наша задача состоит в выборе функции f, которая годилась бы для вариационной формулировки основных законов термодинамики с заданным набором интенсивных и экстенсивных термодинамических переменных, выбранных в качестве управляющих параметров. Более того, мы ищем простейшую из таких функций. Для решения этой задачи разложим $\mathcal{U}+f$ в ряд до членов второго порядка включительно в окрестности некоторой точки на критическом многообразии:

$$(\mathcal{U}+f)=(\mathcal{U}+f)^0+\delta^{(1)}(\mathcal{U}+f)+\delta^{(2)}(\mathcal{U}+f)+\ldots, \qquad (10.113)$$

$$(\mathcal{U}+f)^0=\mathcal{U}(i_\alpha^{(0)};\ E_{(0)}^\beta)+f(i_\alpha^{(0)},\ E_{(0)}^\beta), \qquad (10.114.0)$$

$$\delta^{(1)}(\mathcal{U}+f)=(\mathcal{U}^r+f^r)\,\delta i_r+(\mathcal{U}_\beta+f_\beta)\,\delta E^\beta, \qquad (10.114.1)$$

$$\delta^{(2)}(\mathcal{U}+f)=\frac{1}{2}(\mathcal{U}+f)^{rs}\,\delta i_r\delta i_s+$$

$$+(\mathcal{U}+f)^{r}_{\beta}\,\delta i_r\delta E^{\beta}+\frac{1}{2}(\mathcal{U}+f)_{\alpha\beta}\,\delta E^{\alpha}\delta E^{\beta}. \qquad (10.114.2)$$

Критическое многообразие определяется равенством нулю градиента, вычисляемого по переменным состояния. Таким образом, (10.114.1) можно использовать для определения функции f, потребовав, чтобы соответствующие n из $2n$ первых производных обратились в нуль. Тогда оставшиеся n ненулевых производных определяются однозначно из соображений простоты вида функции f.

Пример 1. Проиллюстрируем вышеприведенные рассуждения на примере потенциала, зависящего от двух интенсивных и двух экстенсивных переменных, т. е. $\mathcal{U}=\mathcal{U}(i_1, i_2; E^1, E^2)$.

В качестве управляющих параметров выберем i_1 и E^2. Тогда первые производные из (10.114.1) принимают вид

$$(\mathcal{U}+f)^1,\quad (\mathcal{U}+f)^2,\quad (\mathcal{U}+f)_1,\quad (\mathcal{U}+f)_2. \qquad (10.115)$$

Далее, на критическом многообразии $\mathcal{U}^r=\partial\mathcal{U}/\partial i_r=0$. Поскольку i_1, E^2 — управляющие параметры, i_2 и E^1 — переменные состояния. Поэтому необходимо потребовать, чтобы

$$\begin{aligned}\frac{\partial(\mathcal{U}+f)}{\partial i_2}&=\frac{\partial f}{\partial i_2}=0,\\ \frac{\partial(\mathcal{U}+f)}{\partial E^1}&=\mathcal{U}_1+\frac{\partial f}{\partial E^1}=i_1+\frac{\partial f}{\partial E^1}=0.\end{aligned} \qquad (10.116)$$

В результате f не должна зависеть от i_2 и должна быть пропорциональна E^1. Простейшей функцией, удовлетворяющей этим условиям, является

$$f=-i_1E^1. \qquad (10.117)$$

Остальные две первые производные (по новым управляющим параметрам i_1 и E^2) имеют вид

$$\begin{aligned}\frac{\partial}{\partial i_1}(\mathcal{U}-i_1E^1)&=\frac{\partial\mathcal{U}}{\partial i_1}-E^1=-E^1,\\ \frac{\partial}{\partial E^2}(\mathcal{U}-i_1E^1)&=\frac{\partial\mathcal{U}}{\partial E^2}-0=i_2.\end{aligned} \qquad (10.118)$$

Итак, для перехода к вариационному описанию термодинамики равновесных систем необходимо добавить к $\mathcal{U}$ по одному члену $-i_{\alpha'}E^{\alpha'}$ для каждой интенсивной переменной $i_{\alpha'}$, выбранной в качестве управляющего параметра $\left(f=-\sum\nolimits_{r}{}'i_{\alpha'}E^{\alpha'}\right)$[1]. Градиент полученной таким образом функции по новым переменным состояния должен быть равен нулю, а первые произ-

[1]) Индекс «r» означает, что сумма «урезана», т. е. включает только управляющие параметры.

водные по управляющим параметрам равны

$$\frac{\partial}{\partial E^{\beta'}}(\mathcal{U}+f)=i_{\beta'},\quad \frac{\partial}{\partial i_{\alpha'}}(\mathcal{U}+f)=-E^{\alpha'},$$

$$(i_{\alpha'},\ E^{\beta'})-\text{управляющие параметры.} \qquad (10.119)$$

Критическое многообразие, определяемое равенством нулю градиента, такое же, как критическое многообразие, определяемое равенством $\nabla_i\mathcal{U}(i;\ E)=0$.

К какому вариационному принципу приводит описанная выше схема? Ищем ли мы минимум, седловую точку или максимум $\mathcal{U}'=\mathcal{U}+f$? Чтобы ответить на эти вопросы, необходимо исследовать $\delta^{(2)}\mathcal{U}'$:

$$\delta^{(2)}\mathcal{U}'=\delta^{(2)}\mathcal{U}-\sum_{\text{“}r^{\alpha}\text{”}}\delta i_{\alpha'}\,\delta E^{\alpha'}. \qquad (10.120)$$

Матрица вторых смешанных производных f по переменным состояния тождественно равна нулю, поскольку каждый член в сумме, представляющей f, есть произведение управляющих параметров и переменной состояния. Матрица вторых смешанных производных $\delta^{(2)}\mathcal{U}$ есть симметричная $(n\times n)$-подматрица $(2n\times 2n)$-матрицы

$$\left[\begin{array}{c|c}\mathcal{U}^{rs} & \mathcal{U}^{r}_{\beta}\\ \hline \mathcal{U}^{s}_{\alpha} & \mathcal{U}_{\alpha\beta}\end{array}\right], \qquad (10.121)$$

которая положительно определена. Поэтому $\delta^{(2)}\mathcal{U}'$ положительно определена по n новым переменным состояния. Таким образом, при фиксированных управляющих параметрах термодинамическое состояние системы определяется минимумом функции $\mathcal{U}'$ по новым переменным состояния.

Далее, каковы свойства метрического тензора, определяемого из $\delta^{(2)}U'$, где U' — значение $\mathcal{U}'$ на критическом многообразии? Эти свойства легче всего получить из соотношения

$$\delta^{(2)}U'=\delta^{(2)}U+\delta^{(2)}f=\frac{1}{2}\delta i_{\alpha}\delta E^{\alpha}-\sum_{\text{“}r^{\alpha}\text{”}}\delta i_{\alpha'}\delta E^{\alpha'}. \qquad (10.122)$$

Мы уже видели, что $\delta^{(2)}U$ положительно определена. Если бы в качестве управляющих параметров мы выбрали только интенсивные переменные, то «урезанная» сумма имела бы вид

$$\delta^{(2)}U'=\frac{1}{2}\delta i_{\alpha}\delta E^{\alpha}-\sum_{\alpha'=1}^{n}\delta i_{\alpha}{}'U,\ \delta E^{\alpha'}=-\frac{1}{2}\delta i_{\alpha}\delta E^{\alpha}. \qquad (10.123)$$

Таким образом, $\delta^{(2)}U'$ была бы отрицательно определена. От-

сюда следует, что на каждый член урезанной суммы для f приходится одна перемена знака сигнатуры метрического тензора, полученного из $\delta^{(2)}U'$. Следовательно, единственным выбором управляющих параметров,приводящим к положительно или отрицательно определенным матрицам на критическом многообразии, является выбор либо только интенсивных, либо только экстенсивных термодинамических переменных.

Пример 2. В продолжение примера 1 рассмотрим

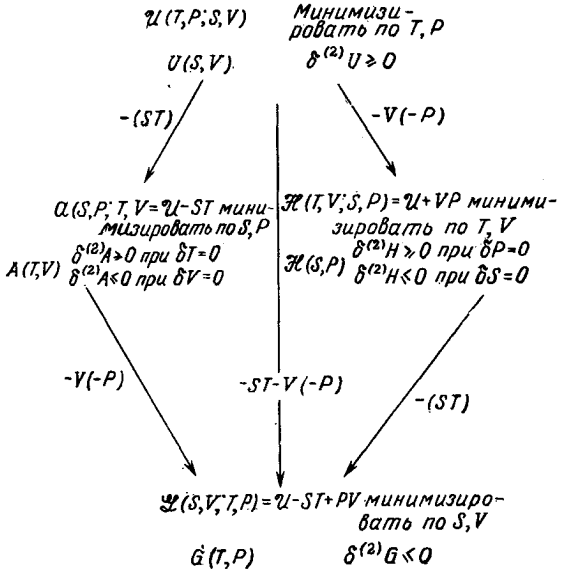
$$\mathcal{U}'(E^1, i_2; i_1, E^2) = \mathcal{U}(i_1, i_2; E^1, E^2) - i_1E^1. \qquad (10.124)$$

Критическое многообразие определяется минимизацией $\mathcal{U}'$ по E^1 и i_2. На критическом многообразии $E^1 = E^1(i_1, E^2)$, $i_2 = i_2(i_1, E^2)$ и

$$U'(i_1,E^2) = \mathcal{U}'[E^1(i_1, E^2), i_2(i_1,E^2); i_1,E^2], \qquad (10.125)$$

где выбирается глобальный минимум. Тогда $\delta^{(2)}U' \geqslant 0$ при $\delta i_1 = 0$ и $\delta^2U' \leqslant 0$ при $\delta E^2 = 0$. Коэффициенты восприимчивости можно получить из производных от U' или из табл. 10.2. Неопределенную метрику можно найти из $\delta^{(2)}U'$ в системе координат (i_1, E^2) или любой другой системе координат, если воспользоваться соотношением (10.112).

В случае простого, однокомпонентного вещества, состояние которого описывается функцией $U(T, P; S, V)$, приведенные рассуждения можно представить в виде следующей схемы



на которой также приведены свойства интересующих нас функций.

$\diamond\diamond\diamond$ Преобразование вида $(\mathcal{U} \to \mathcal{U}' = \mathcal{U} + f,\ f = -\sum_{r''} i_\alpha E^\alpha)$ называют *преобразованием Лежандра.* Оказывается, невозможно найти преобразование Лежандра, приводящее к потенциалу, для которого пара сопряженных термодинамических переменных входит в число управляющих параметров (и, как следствие, в число переменных состояния). Многие частные производные термодинамических величин могут быть вычислены из соотношений Максвелла (равенство вторых смешанных частных производных), если предварительно выбрать некоторую термодинамическую производящую функцию. Однако частные производные термодинамических величин, включающие пары сопряженных термодинамических переменных (таких, как $(\partial S/\partial P)_V$), таким способом вычислить нельзя из-за отсутствия подходящего преобразования Лежандра. Поэтому для их вычисления следует воспользоваться тензором восприимчивости (табл. 10.2).

Уравнение состояния играет фундаментальную роль при описании системы в условиях термодинамического равновесия. Мы идентифицируем это уравнение как n-мерное многообразие в пространстве $\mathbb{R}^n \otimes \mathbb{R}^n$ размерности $2n$, рассматривая соответствующие интенсивные и экстенсивные термодинамические переменные с равных позиций.

Чтобы сформулировать вариационный принцип определения многообразия уравнения состояния, была нарушена симметрия, существующая между интенсивными и экстенсивными термодинамическими переменными, так как n из них были выбраны в качестве независимых управляющих параметров, а остальные n — в качестве зависимых переменных состояния, по которым и должен минимизироваться некоторый потенциал. Если принят принцип Максвелла, критическое многообразие идентифицируется с многообразием уравнения состояния. Тогда максвелловское множество определяется как проекция критического многообразия в $\mathbb{R}^n \otimes \mathbb{R}^n$ на пространство $\mathbb{R}^n$ управляющих параметров. Наличие различных проекций, обусловленных различным выбором управляющих параметров, приводит к различным максвелловским множествам. Для любой проекции максвелловское множество определяется уравнениями Клаузиуса — Клапейрона.

С самого начала в качестве управляющих параметров были выбраны экстенсивные термодинамические переменные, поскольку они однозначно описывают систему в состоянии термодинамического равновесия. При таком выборе управляющих параметров максвелловское множество пусто, и уравнения Клаузиуса — Клапейрона приводят к условию (10.71), из кото-

рого следует, что каждая интенсивная переменная должна иметь одно и то же значение во всех сосуществующих фазах.

Иной выбор управляющих параметров может привести к непустому максвелловскому множеству в пространстве управляющих параметров. Как обычно, компоненты максвелловского множества определяются из уравнений Клаузиуса — Клапейрона, и они имеют максимальные размерности, когда управляющими параметрами являются только интенсивные термодинамические переменные.

Состояние системы, имеющей c компонент, полностью определяется $(c+1)$ парой термодинамических переменных (S, T), (N_1, μ_1), (N_2, μ_2), ..., (N_c, μ_c). Уравнение состояния представляет собой $(c+1)$-мерное многообразие в пространстве $\mathbb{R}^{2(c+1)}$. Состояние единственным образом определяется заданием экстенсивных переменных $(S, V, N_2, \ldots, N_c)$. Если в состоянии термодинамического равновесия сосуществуют ν фаз, то потенциал, описывающий данную систему, имеет ν равных минимумов. Максвелловское множество пусто в пространстве управляющих параметров, включающих одни экстенсивные переменные, и максимально в пространстве управляющих параметров, которыми являются только интенсивные переменные. В последнем случае уравнения Клаузиуса — Клапейрона для ν равных минимумов играют роль $(\nu-1)$ ограничения на интенсивные управляющие параметры. Таким образом, остается только

$$1+c-(\nu-1)=c+2-\nu \tag{10.126}$$

независимых интенсивных управляющих параметров (правило Гиббса). Кроме того, состояние системы однозначно определяется заданием значений $(\nu-1)$ соответствующей экстенсивной термодинамической переменной.

Для иллюстрации этих замечаний рассмотрим еще раз однокомпонентное вещество и выберем в качестве управляющих параметров две интенсивные переменные (T, P). Предположим, что в состоянии термодинамического равновесия сосуществуют две фазы. Тогда уравнение Клаузиуса — Клапейрона принимает вид

$$\left[\frac{\partial \mathcal{G}^{(1)}}{\partial T}-\frac{\partial \mathcal{G}^{(2)}}{\partial T}\right]\delta T+\left[\frac{\partial \mathcal{G}^{(1)}}{\partial P}-\frac{\partial \mathcal{G}^{(2)}}{\partial P}\right]\delta P=0. \tag{10.127}$$

Поскольку $\partial \mathcal{G}/\partial T=\partial G/\partial T=-S$ и $\partial \mathcal{G}/\partial(-P)=-V$ [см. (10.119)], это уравнение сводится к следующему:

$$\left[-S^{(1)}+S^{(2)}\right]\delta T+\left[V^{(1)}-V^{(2)}\right]\delta P=0, \tag{10.128}$$

или

$$\frac{dP}{dT}=\frac{S^{(1)}-S^{(2)}}{V^{(1)}-V^{(2)}}=\frac{\Delta S}{\Delta V}, \tag{10.129}$$

где ΔS — изменение энтропии системы при переходе системы из фазы 1 в фазу 2, а ΔV — соответствующее изменение объема. Состояние системы единственным образом определяется заданием значения одного независимого управляющего параметра, такого, как положение точки на кривой равновесия, и одной экстенсивной термодинамической переменной — S или V.

Состояние системы в тройной точке однозначно определяется заданием $(0 = 1 + 2 - 3)$ интенсивных параметров и двух экстенсивных переменных.

15. ФЛУКТУАЦИИ

Если g_{ij} — положительно определенная симметрическая $(n \times n)$-матрица и g^{ij} — обратная ей матрица $(g^{ij}g_{jk} = \delta^i_k)$, то

$$\langle x^i x^j \rangle = \frac{\int \dots \int x^i x^j e^{-(1/2)\, g_{rs} x^r x^s} \, dx^1 \dots dx^n}{\int \dots \int e^{-(1/2)\, g_{rs} x^r x^s} \, dx^1 \dots dx^n} = g^{ij}. \quad (10.130)$$

Этот результат можно доказать следующим образом:

1. $\langle x^i x^j \rangle$ являются компонентами контравариантного тензора второго ранга, который симметричен. То же справедливо для g^{ij}.

2. Равенство (10.130) справедливо в одной конкретной системе координат, а именно в системе, где $g_{..} = g^{..} = I_n$.

3. Поэтому оно справедливо во всех системах координат.

В окрестности любой точки $(i_\alpha;\ E^\beta)$ критического многообразия имеют место флуктуации интенсивных δi_α и экстенсивных δE^β переменных. Эти флуктуации могут быть равновесными или неравновесными, причем в первом случае приращения δi_α и δE^β связаны через тензор восприимчивости. Неравновесные флуктуации происходят вокруг точки $(i_\alpha;\ E^\beta)$ в $\mathbb{R}^n \otimes \mathbb{R}^n$; равновесные флуктуации имеют место только в n-мерном пространстве, касательном к критическому многообразию в точке $(i_\alpha,\ E^\beta)$. Неравновесные и равновесные флуктуации описываются функциями распределения вероятностей $\mathscr{P}$ и P соответственно, которые определяются как

$$\begin{aligned} \mathscr{P}(\delta i_\alpha;\ \delta E^\beta) &\simeq e^{\delta S/k} \simeq e^{-\delta^{(2)}\mathscr{U}/kT}, \\ P(\delta E^\beta) &\simeq e^{\delta S/k} \simeq e^{-\delta^{(2)}U/kT}. \end{aligned} \quad (10.131)$$

Положительно определенные выражения $\delta^{(2)}\mathscr{U}$, $\delta^{(2)}U$ представлены в (10.75ii). Эти выражения совместно с (10.130) приводят к простым соотношениям для ожидаемых значений нерав-

новесных и равновесных флуктуаций. В неравновесном случае

$$\begin{bmatrix} \langle \delta i_r \delta i_s \rangle & \langle \delta i_r \delta E^\beta \rangle \\ \langle \delta E^\alpha \delta i_s \rangle & \langle \delta E^\alpha \delta E^\beta \rangle \end{bmatrix}_{NE} = kT \begin{bmatrix} \mathcal{U}^{rs} & \mathcal{U}^r_\beta \\ \mathcal{U}^s_\alpha & \mathcal{U}_{\alpha\beta} \end{bmatrix}^{-1}. \quad (10.132)$$

Таким образом, смешанные вторые частные производные потенциала $\mathcal{U}$ могут быть выражены через неравновесные флуктуации.

В равновесном случае $\delta^{(2)}U = \frac{1}{2} U_{\alpha\beta} \delta E^\alpha \delta E^\beta$ и

$$\frac{\langle \delta E^\alpha \delta E^\beta \rangle}{kT} = U^{\alpha\beta}, \quad \frac{\langle \delta i_\alpha \delta E^\beta \rangle}{kT} = \delta^\beta_\alpha, \quad \frac{\langle \delta i_\alpha \delta i_\beta \rangle}{kT} = U_{\alpha\beta}. \quad (10.133)$$

Пример. Для однокомпонентной системы из (10.133) немедленно получаем

$$\begin{bmatrix} \langle \delta S\, \delta S \rangle & \langle \delta S\, \delta V \rangle \\ \langle \delta V\, \delta S \rangle & \langle \delta V\, \delta V \rangle \end{bmatrix} = kT \begin{bmatrix} \frac{C_P}{T} & V\alpha_P \\ V\alpha_P & V\beta_T \end{bmatrix}, \quad (10.134\text{i})$$

$$\begin{bmatrix} \langle \delta T\, \delta S \rangle & \langle \delta T\, \delta V \rangle \\ \langle -\delta P\, \delta S \rangle & \langle -\delta P\, \delta V \rangle \end{bmatrix} = kT \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (10.134\text{ii})$$

$$\begin{bmatrix} \langle \delta T\, \delta T \rangle & \langle -\delta T\, \delta P \rangle \\ \langle -\delta P\, \delta T \rangle & \langle \delta P\, \delta P \rangle \end{bmatrix} = kT \begin{bmatrix} \frac{T}{C_V} & -\frac{T}{\Gamma_V} \\ -\frac{T}{\Gamma_V} & \frac{1}{V\beta_S} \end{bmatrix}. \quad (10.134\text{iii})$$

16. ДОПОЛНИТЕЛЬНЫЕ СООБРАЖЕНИЯ

Хорошо известно, что геометрическая оптика является предельным случаем («$\lambda \to 0$») волновой оптики, описываемой уравнениями Максвелла, а классическая механика — предельным случаем («$h \to 0$») квантовой механики. Известно также, что статистическая механика является основой термостатики. Таким образом, термодинамику можно рассматривать, как предельный случай («$k \to 0$») статистической механики. (Кавычки указывают на то, что смысл предельных переходов не совсем четко определен, главным образом потому, что три величины λ, h, k не безразмерны. Эти пределы можно было бы определить точно, взяв некие безразмерные выражения от этих величин и вычислив предел, устремляя их к нулю.)

Довольно многие эффекты волновой оптики можно получить из геометрической оптики, применяя принцип Гюйгенса, так же как и многие квантовомеханические результаты следуют из классической механики, если воспользоваться подходом Фейнмана, связанным с минимизацией интеграла пути. Естественно возникает вопрос: возможно ли из термодинамической теории

«вывести» статистическую механику, используя что-то похожее на фейнмановский подход? Иными словами, какие результаты статистической механики можно получить, применяя к термодинамике некий вариационный принцип типа фейнмановской минимизации интеграла пути? Является ли $S(\mathcal{U})/k$ аналогом действия $i\Delta t\mathcal{L}(q, \dot{q})/\hbar$? Играет ли уравнение Блоха ту же роль при гипотетическом «выводе» (термодинамика → статистическая механика), что и уравнение Шредингера при переходе от классической механики к квантовой? Насколько глубока аналогия между оператором плотности ρ и матрицей рассеяния S?

В терминах симплектической геометрии в $\mathbb{R}^n \otimes \mathbb{R}^n$ классическая механика получает элегантное представление в виде гамильтоновых уравнений движения

$$\frac{dq_i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial q_i}. \tag{10.135}$$

В настоящей вариационной формулировке термодинамики аналогами уравнений (10.135) являются уравнения вида

$$i_{\beta'} = \frac{\partial \mathcal{U}'}{\partial E^{\beta'}}, \quad -E^{\alpha'} = \frac{\partial U'}{\partial i_{\alpha'}}, \tag{10.136а}$$

$$0 = \frac{\partial \mathcal{U}'}{\partial E^{\beta''}}, \quad 0 = \frac{\partial \mathcal{U}'}{\partial i_{\alpha''}}, \tag{10.136б}$$

где $i_{\alpha'}$, $E^{\beta'}$ — управляющие параметры; $i_{\alpha''}$, $E^{\beta''}$ — переменные состояния, а $\mathcal{U}'$ получается из $\mathcal{U}$, как в (10.119). Существование связи между (10.135) и (10.136а) наводит на мысль о том, существует ли некоторая естественная структура в $\mathbb{R}^n \otimes \mathbb{R}^n$, связанная с такой вариационной формулировкой термодинамики, и связана ли эта структура с симплектической или ортогональной группой $Sp(2n)$ или $SO(2n)$, или с какой-либо связанной с ними действительной формой?

Как и второе начало термодинамики, принцип Ле Шателье, обладая глубочайшим внутренним содержанием, часто формулируется довольно расплывчато. Его можно сформулировать следующим образом: *внешняя сила, возмущающая равновесное состояние некоторой системы, вызывает в системе процессы, стремящиеся ослабить действие этой силы.* Этот принцип рассматривался некоторыми авторами как постулат термодинамики равновесных систем. При такой интерпретации принцип Ле Шателье эквивалентен требованию устойчивости $\delta^{(2)}U$, которое использовалось при доказательстве положительной определенности метрического тензора $U_{\alpha\beta}$. Фактически же принцип Ле Шателье используется для вывода термодинамических неравенств (10.102—10.107). Из совокупности этих неравенств вытекает положительная определенность $U_{\alpha\beta}$.

Автор предпочитает рассматривать принцип Ле Шателье как постулат термодинамики неравновесных систем, относящийся, в частности, к процессу перехода возмущенной системы в равновесное состояние. М. Лакс в частном неопубликованном сообщении отметил, что принцип Ле Шателье можно понимать даже в более общем смысле как условие устойчивости, описывающее возврат системы к ее многообразию уравнения состояния независимо от того, является ли это многообразие равновесным или некоторым неравновесным стационарным многообразием.

Если на систему, находящуюся в состоянии термодинамического равновесия, действует «медленное» возмущение, то тензор восприимчивости есть $U_{\alpha\beta}$. Если, однако, время действия возмущения меньше времени отклика системы, система реагирует «энергичнее». Это и есть смысл принципа Ле Шателье. Кратко можно сказать, что неравновесный отклик ($\mathcal{U}_{\alpha\beta}$) системы на возмущение больше равновесного отклика ($U_{\alpha\beta}$), поскольку тензор $\mathcal{U}_{\alpha\beta} - U_{\alpha\beta}$ положительно определен [см. (10.78)]. В связи с этим возникает ряд вопросов.

Является ли принцип Ле Шателье утверждением, относящимся к положительной определенности $\mathcal{U}_{\alpha\beta} - U_{\alpha\beta}$, а не к положительной определенности $U_{\alpha\beta}$? Является ли он принципом *термодинамики* неравновесных систем ($\mathcal{U}_{\alpha\beta} - U_{\alpha\beta}$), а не *термостатики* равновесных систем? Каким образом постоянные времени, описывающие переход возмущенной системы от ее начального неравновесного отклика к конечному стационарному равновесному отклику, связаны с положительно определенными тензорами $\mathcal{U}_{\alpha\beta}$, $U_{\alpha\beta}$?

Если многообразие уравнения состояния некоторой системы получается на основе некоторого вариационного принципа (1.7)

$$\nabla V = 0, \tag{10.137.0}$$

го совершенно естественно предположить, что поведение системы во времени вне критического многообразия определяется системой градиентных уравнений типа (1.6)

$$\frac{dx_i}{dt} = -\frac{\partial V(x;\, c)}{\partial x_i}. \tag{10.137.1}$$

Тогда процесс приближения к критическому многообразию можно описать линеаризованными уравнениями вида

$$\frac{dx_i}{dt} = -V_{ij}\,\delta x_j + \mathcal{O}(2). \tag{10.137.2}$$

Вблизи критического многообразия членами более высокого порядка можно пренебречь, и временные характеристики системы оказываются обратно пропорциональны наименьшему собственному значению положительно определенной матрицы V_{ij}.

Если система смещена из состояния термодинамического равновесия, то процесс ее возвращения на критическое многообразие определяется соответствующей $(n \times n)$-подматрицей $(2n \times 2n)$-матрицы смешанных вторых частных производных $\mathcal{U}$. Если, например, интенсивные управляющие параметры фиксированы, а экстенсивные переменные возмущены, то соответствующая градиентная система уравнений имеет вид

$$\frac{\partial E^\alpha}{\partial t} = -\frac{\partial}{\partial E^\alpha}(\mathcal{U} - i_\beta E^\beta) = -\frac{\partial}{\partial E^\alpha}\left(\mathcal{U}_\beta \delta E^\beta + \right.$$

$$\left. + \frac{1}{2}\mathcal{U}_{\beta\alpha}\delta E^\beta \delta E^\gamma + \ldots - i_\beta E^\beta\right) = -\mathcal{U}_{\alpha\beta}\delta E^\beta + \mathcal{O}(2). \tag{10.138}$$

Можно помешать системе прийти в состояние термодинамического равновесия, искусственно поддерживая в ней некоторые потоки и градиенты. Если, например, два конца металлического стержня поддерживаются при разных температурах, то температурный градиент вызовет в этом стержне тепловой поток $(dS/dt \neq 0)$. Аналогично, если к концам этого стержня приложена постоянная разность потенциалов, то через него потечет ток $(\partial Q/\partial t)$. Здесь возможны перекрестные эффекты. Когда в системе поддерживаются несколько градиентов, то потоки зависят от их совокупности, т. е. вблизи критического многообразия обобщенные токи J^α связаны с обобщенными силами X_β через тензор восприимчивости L:

$$J^\alpha = L^{\alpha\beta}X_\beta, \tag{10.139}$$

где

$$J^\alpha \simeq \frac{dE^\alpha}{dt}, \quad X_\beta \simeq \nabla i_\beta. \tag{10.140}$$

В отсутствие магнитных полей матрица $L^{\alpha\beta}$ является действительной симметрической положительно определенной матрицей: $L^{\alpha\beta} = L^{\beta\alpha}$ (соотношение симметрии кинетических коэффициентов Онсагера) [5]. При сравнении этого отношения симметрии для стационарного поведения вблизи равновесия с отношением симметрии $\mathcal{U}_{\alpha\beta} = \mathcal{U}_{\beta\alpha}$, описывающим нестационарное приближение к термодинамическому равновесию, невольно напрашивается вопрос:

Какова связь $(n \times n)$-матрицы L феноменологических коэффициентов с $(2n \times 2n)$-матрицей (10.121) смешанных вторых частных производных $\mathcal{U}$, вычисленной на критическом многообразии?

17. ВЫВОДЫ

Первые содержательные приложения теории катастроф относились к области термодинамики. Из сравнения канонической модели фазовых переходов второго рода и фазовой диаграммы жидкости в окрестности ее критической точки следует неизбежный вывод о том, что в основе фазового перехода первого рода жидкость — газ лежит катастрофа сборки. Многие из признаков, характерных для этой катастрофы (см. гл. 8) можно заметить на кривой равновесия жидкость — пар, в особенности в окрестности критической точки. Отождествлению критической точки и кривой равновесия с катастрофой сборки можно придать количественный характер, введя аффинную линейную связь между переменными состояния ρ и x и между управляющими параметрами (P_r, T_r) и (a, b). Четыре довольно разумных предположения позволяют перейти от многообразия катастрофы сборки к уравнению состояния Ван-дер-Ваальса. Это классическое уравнение оказывается неверным в количественном отношении, поскольку оно дает неправильные значения для некоторых важных показателей. При этом невозможно построить правильное в этом смысле уравнение ни при каких модификациях принятого соответствия между математическими и физическими переменными состояния и управляющими параметрами.

Тройная точка представляет в чем-то больший интерес, чем критическая. Пытаясь найти уравнение состояния вещества в окрестности его тройной точки, мы рассмотрели каноническую модель фазовых переходов первого рода. Сделав всего «полшага», можно перейти от этой симметризованной модели к обсуждению трикритических точек, и еще через «полшага» — к обсуждению катастрофы A_{+5} и, в частности, ее максвелловского множества. Тогда фазовая диаграмма простого вещества есть пересечение плоскости (P, T) физического пространства $\mathbb{R}^2$ с максвелловским множеством $\mathscr{S}_M$ катастрофы A_5 в математическом пространстве управляющих параметров $\mathbb{R}^4$.

Различие между «потенциалами», характеризующими системы в состоянии термодинамического равновесия, и потенциалами и семействами потенциалов, являющимися объектом исследования в теории катастроф, и естественное отождествление уравнения состояния вещества с n-мерным многообразием в пространстве $\mathbb{R}^n$ (интенсивные переменные) $\otimes$ $\mathbb{R}^n$ (экстенсивные переменные) приводят к довольно естественной вариационной формулировке термодинамики равновесных систем в терминах теории катастроф. В этой формулировке фигурирует некоторый действительный потенциал $\mathscr{U}$, являющийся функцией n интенсивных переменных состояния и n экстенсивных управляющих

параметров. Тогда уравнение состояния идентифицируется с критическим многообразием.

При таком подходе условия равновесия и устойчивости, наложенные на семейства функций, выраженные через первые и вторые производные, эквивалентны первому и второму началам термодинамики. Более точно, первое начало термодинамики эквивалентно отсутствию бесконечно малых членов по интенсивным переменным состояния в выражении для приращения первого порядка величины потенциала в состоянии равновесия:

$$\delta^{(1)}\mathcal{U} = \mathcal{U}^r \delta i_r + \mathcal{U}_\alpha \delta E^\alpha \to 0 + U_\alpha \delta E^\alpha.$$

Второе начало термодинамики эквивалентно некоторому условию устойчивости, а именно, требованию положительной определенности членов второго порядка в разложении U на критическом многообразии:

$$\delta^{(2)}U = \frac{1}{2} U_{\alpha\beta} \delta E^\alpha \delta E^\beta \geqslant 0.$$

Положительная определенность метрического тензора $U_{\alpha\beta}$ означает, что уравнение состояния описывается римановой поверхностью. Принцип Ле Шателье, по-видимому, связан с «относительными размерами» $\mathcal{U}_{\alpha\beta}$ и $U_{\alpha\beta}$:

$$(\mathcal{U}_{\alpha\beta} - U_{\alpha\beta})\, \delta E^\alpha \delta E^\beta \geqslant 0.$$

Из положительной определенности метрического тензора $U_{\alpha\beta}$ с помощью неравенств Шварца и Бесселя выводится ряд термодинамических равенств и неравенств.

Теория линейного отклика проще всего выражается в пространстве, касательном в некоторой точке к критическому многообразию. Любая точка в касательном пространстве ($\simeq \mathbb{R}^n$) из $\mathbb{R}^n \otimes \mathbb{R}^n$ однозначно определяется заданием n из $2n$ координат. Таким образом, легко можно вычислить тензоры восприимчивости и их матричные элементы, а также частные производные термодинамических величин. В частности, метрический тензор $U_{\alpha\beta}$ оказывается тензором восприимчивости.

Если искусственно не разделять интенсивные и экстенсивные переменные, то можно дать формулировку термодинамики в терминах теории катастроф, используя смешанный набор интенсивных и экстенсивных термодинамических переменных в качестве управляющих параметров, а остальные — в качестве переменных состояния.

Пытаясь перейти границы чисто классической термодинамики, мы рассмотрели как неравновесные ($k \neq 0$), так и равновесные флуктуации, которые могут иметь место только в касательном к критическому многообразию пространстве. Корреляционные функции для флуктуаций определяются матрицами

смешанных частных производных потенциала $\mathcal{U}$ и производящей функции U.

Два обстоятельства позволяют считать, что предложенная формулировка термодинамики в терминах теории катастроф является обоснованным и полезным обобщением классической формулировки Гиббса. Первое из них связано с многообразием свойств матрицы $U_{\alpha\beta}$, которая (1) определяет $\delta^{(2)}U$, (2) является положительно определенным метрическим тензором, (3) является тензором восприимчивости, (4) описывает расположение касательного пространства в $\mathbb{R}^n \otimes \mathbb{R}^n$, (5) является корреляционной функцией флуктуаций, (6) определяет нижнюю границу неравновесной реакции $\mathcal{U}_{\alpha\beta}$.

Второе обстоятельство связано с тем, что при такой формулировке вопросов возникает больше, чем имеется ответов. В частности, такая формулировка может оказаться полезной при распространении термодинамики на случай неравновесных систем (конфигураций). В пространстве $\mathbb{R}^n \otimes \mathbb{R}^n$ «хватит места», чтобы сдвинуться с критического многообразия в некотором направлении, не лежащем в касательном пространстве. В стандартной гиббсовской формулировке термодинамики равновесных систем этого сделать нельзя, поскольку она оперирует в $\mathbb{R}^n$. Все вопросы сводятся к одному: как проложить мостик между термодинамикой равновесных и термодинамикой неравновесных систем?

Литература

1. Ландау Л. Д., Лифшиц Е. М. Статистическая физика. — М.: Наука, 1976.
2. Van der Waals J. D., Ph. D. Thesis, Leiden, 1873.
3. Guggenheim E. A., The Principle of Corresponding States, *J. Chem. Phys.*, **13**, 253—261 (1945).
4. (a) Weinhold F., Metric Geometry of Equilibrium Thermodynamics, *J. Chem. Phys.*, **63**, 2479—2483 (1975).
(b) II. Scaling, Homogeneity, and Generalized Gibbs-Duhem relations, *J. Chem. Phys.*, **63**, 24—2487 (1975).
(c) III. Elementary Formal Structure of a Vector-Algebraic Representation of Equilibrium Thermodynamics, *J. Chem. Phys.*, **63**, 2488—2495 (1975).
(d) IV. Vector-Algebraic Evaluation of Thermodynamic Derivatives, *J. Chem. Phys.*, **63**, 2496—2501 (1975).
(e) V. Aspects of Heterogeneous Equilibrium, *J. Chem. Phys.*, **65**, 559—564 (1976).
5. Onsager L., Reciprocal Relations in Irreversible Processes. I., *Phys. Rev.*, **37**, 405—426 (1931); Reciprocal Relations in Irreversible Processes. II., *Phys. Rev.*, **38**, 2265—2279 (1931).

II

МЕХАНИКА КОНСТРУКЦИЙ

Многие крупногабаритные технические конструкции могут быть описаны с помощью потенциальной функции, минимальное значение которой определяет локально устойчивое состояние конструкции. Само состояние описывается положением точки в некотором пространстве состояний конструкции. С увеличением нагрузки на конструкцию (мост, здание и т. д.) потенциальная функция изменяется. Значительная нагрузка может привести к потере устойчивости конструкции (т. е. к ее разрушению) вследствие нарушения локально устойчивого состояния, которое является для данной системы расчетным. Равновесие, устойчивость и потеря устойчивости — это основные вопросы, рассматриваемые теорией катастроф. Методы теории катастроф позволяют определить чувствительность критической, или разрушающей, нагрузки как к несовершенству конструкции, так и к динамическому воздействию. (Встречаются несколько видов жесткой чувствительности к несовершенству.) Кроме того, они оказываются эффективными при изучении составных систем, для которых возможны различные формы разрушения. Последнее обстоятельство имеет важное практическое значение, так как свидетельствует о том, что в «оптимизируемых» системах, составленных из нескольких «неприводимых» конструкционных элементов, могут проявляться неожиданные формы разрушения с жесткой чувствительностью к несовершенству, если между элементами существует сильная связь (например, разрушение опорного кронштейна).

На практике конструкции собирают обычно из большого числа отдельных элементов. Анализ процесса разрушения, как правило, проводится методами теории бифуркаций. Используемые при этом стандартные алгоритмы вычисления универсального возмущения (гл. 4) обеспечивают систематизированный подход к изучению оптимизируемых систем, спектра форм их разрушения и типа чувствительности таких систем к несовершенству.

1. СИСТЕМЫ, ОПИСЫВАЕМЫЕ ПОТЕНЦИАЛЬНОЙ ФУНКЦИЕЙ

Для описания свойств конструкции, модели или реальной физической системы необходимо прежде всего ввести координаты состояния системы $x_1, x_2, \ldots, x_n$, или так называемые параметры порядка. Полезно также ввести дополнительное множество параметров $c_1, c_2, \ldots, c_k$, которые будут представлять нагрузку (внешние силы) на систему, дефекты, возникшие при изготовлении элементов конструкции, и дефекты, появившиеся в процессе сборки системы.

Общая энергия ξ подобной системы, как правило, есть сумма кинетической и потенциальной энергии

$$\xi = KE + PE. \tag{11.1}$$

Кинетическую энергию часто можно представить в виде положительно определенной квадратичной формы от обобщенных скоростей $\dot{x}_i$, а потенциальную энергию — как функцию переменных состояния $x \in \mathbb{R}^n$ и управляющих параметров $c \in \mathbb{R}^k$:

$$KE = \frac{1}{2} M_{ij}\dot{x}_i\dot{x}_j \geqslant 0, \quad PE = V(x;\ c). \tag{11.2}$$

Рассмотрим сначала статические консервативные системы, т. е. системы, для которых $KE = 0$ и $\xi = V$.

Если предположить, что равновесные конфигурации конструкции (в пространстве параметров порядка) определяются минимумом ξ, то в этом случае условия равновесия, устойчивости и потери устойчивости могут быть записаны соответственно как

$$\nabla V = 0; \quad V_{ij} \simeq M_k^n, \quad k = 0; \quad \det V_{ij} = 0.$$

Такая форма записи наглядно показывает, почему теория катастроф может быть полезным инструментом при описании равновесия, устойчивости и характера разрушения физических конструкций.

Конструкция обычно проектируется так, чтобы были обеспечены определенные рабочие параметры. Однако подобное возможно только при условии использования совершенных материалов и при участии искусных исполнителей. Хотя мы живем в лучшем из возможных миров, тем не менее совершенные материалы и искусные исполнители — недосягаемый предел. Другими словами, созданная система не будет полностью соответствовать проектируемым параметрам.

Естественно, что заранее ничего нельзя сказать о степени нарушения такого соответствия. Однако можно предсказать чувствительность конструкции к дефектам, которые могут быть в ней обнаружены.

Поступим следующим образом. Потенциальную функцию $V(x;\ c)$, описывающую состояние идеальной системы, разложим в ряд вблизи проектируемого состояния равновесия, устойчивого при малых внешних нагрузках:

$$V(x;\ F) = \frac{1}{2} V_{ij}(F)\, x_i x_j + \text{Члены более высокой степени} \tag{11.3}$$

(постоянный член не имеет существенного значения и может быть опущен). Линейный член отсутствует в силу выполнения в точке $x = 0$ условия $\partial V/\partial x_i = 0$, и, следовательно, ряд Тей-

лора будет начинаться с квадратичных членов. Коэффициенты ряда Тейлора являются функциями управляющих параметров $c \in \mathbb{R}^k$. При отсутствии дефектов управляющими параметрами являются только внешние силы F, действующие на конструкцию. При возрастании нагрузки от нуля до тех пор, пока матрица V_{ij} остается положительно определенной, проектируемое устойчивое состояние равновесия остается локально устойчивым. Условие

$$\det V_{ij}(F) = 0 \tag{11.4}$$

определяет критическую нагрузку $F = F_p$, которую идеальная (совершенная) система уже не выдержит.

Когда росток потенциала $V(x; F)$ в точке $F = F_p$ известен, для определения вида функции $V(x; F)$, описывающей самый общий вид деформации идеальной системы, могут быть применены методы, описанные в гл. 4. Эта функция может быть использована для моделирования всех несовершенств, возникающих в системе из-за отсутствия искусных исполнителей и использования нестандартных строительных материалов.

Аналогично может быть изучена потенциальная функция $V(x; F, \varepsilon)$, описывающая несовершенную систему. Критическая нагрузка F_c, которую не выдерживает несовершенная система, определяется из соотношения

$$\det V_{ij}(F, \varepsilon) = 0. \tag{11.5}$$

Естественно ожидать, что несовершенная система имеет меньшую несущую способность ($F_c \leqslant F_p$), чем совершенная. (Иначе зачем нанимать архитектора?) Теория катастроф позволяет представить снижение несущей способности конструкции в количественном виде. Для моделей, которые рассматриваются ниже, имеем

$$F_c = F_p - k|\varepsilon|^p, \tag{11.6}$$

где k — некоторая положительная постоянная, p — положительная рациональная дробь, а ε — некоторый параметр несовершенства. Чувствительность к несовершенству при разных значениях показателя p приведена на рис. 11.1: чем меньше p, тем выше чувствительность к несовершенству.

Причиной снижения несущей способности конструкции могут быть соответствующие динамические нагрузки. Так, например, если конструкция представляет собой жилой дом или транспортное сооружение, то снижение ее несущей способности обусловливается соответственно людьми и транспортом. При этом средняя кинетическая энергия, вносимая транспортным шумом, может рассматриваться как (динамический) параметр несовершенства. Чувствительность к несовершенству конструкции,

находящейся под нагрузкой, можно определить следующим образом. Критические точки $x^{(0)}$, $x^{(1)}$, ... потенциальной функции V при любой нагрузке F определяются из уравнения $\nabla V = 0$. В каждой критической точке находятся критические значения $V^{(i)} = V(x^{(i)}; F, \varepsilon)$. Если $x^{(0)}$ — локально устойчивое состояние равновесия, а $x^{(1)}$ — наименьшее ближайшее морсовское 1-седло,

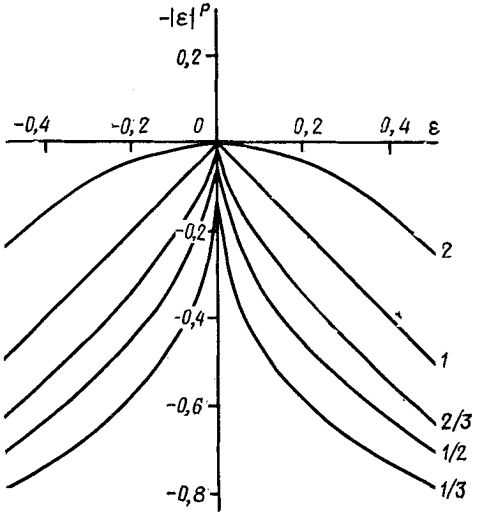


Рис. 11.1. При малом p чувствительность критической нагрузки F_c существенно зависит от параметра несовершенства ε.
Здесь $F_c = F_p - k|\varepsilon|^p$.

то динамическая чувствительность к несовершенству определяется формулой

$$\Delta E = V^{(1)} - V^{(0)}. \tag{11.7}$$

Физически это означает, что система остается в локально устойчивом состоянии $x^{(0)}$ при нулевых или малых колебаниях ($V^{(0)} + \Delta E < V^{(1)}$) до тех пор, пока кинетическая энергия, вносимая в систему извне, не станет настолько большой, что система может «перескочить» через потенциальный барьер $V^{(1)} - V^{(0)}$ в некоторую другую равновесную конфигурацию. Значения динамической чувствительности к несовершенству, получаемые из формулы (11.7), имеют при $\varepsilon \to \Delta E$ вид (11.6). Для систем, изучаемых в данной главе, динамическая чувствительность к несовершенству более существенна, чем статическая чувствительность.

◇◇◇ Две конструкции с тождественными потенциальными функциями $V(x; F, \varepsilon)$ могут различаться функциями кинетической энергии. В этом случае их поведение при статических на-

грузках будет идентичным, однако их реакции на динамическую нагрузку могут быть различными.

◇◇◇ Даже если система является консервативной, действующие на нее возмущения определенного класса могут и не быть консервативными. Таковы нагрузки, вызываемые ветром и дождем, а также некоторые виды динамических нагрузок, как, например, нагрузка на мост, вызванная движущимся поездом. Это обстоятельство следует иметь в виду, так как иначе может создаться впечатление, что анализ устойчивости и чувствительности к несовершенству полностью сводится к алгоритмам элементарной теории катастроф.

2. ВЫПУЧИВАНИЕ ЭЙЛЕРОВА СТЕРЖНЯ

В качестве конкретного примера, иллюстрирующего использование описанных выше методов, рассмотрим выпучивание эйлерова стержня под действием сжимающей нагрузки. Пред-

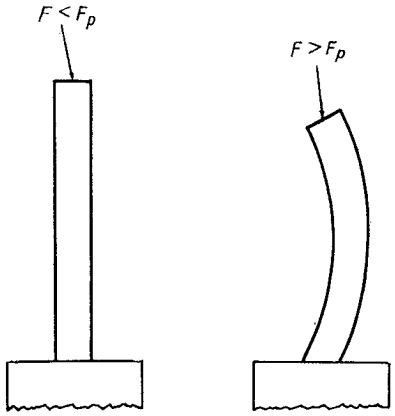


Рис. 11.2. При малых нагрузках стержень остается прямым. При больших нагрузках стержень изгибается [2].

положим, что к одному концу идеального несжимаемого стержня приложена сила F (рис. 11.2): если нагрузка (сила F), действующая на стержень, невелика, стержень остается прямым; при очень большой нагрузке F стержень сильно изгибается. Изучим подробнее процесс изгиба стержня при промежуточных значениях F.

Удобно представить функцию, описывающую поведение (форму) стержня, в виде разложения в ряд Фурье

$$y(x) = \sum_{j=1}^{\infty} a_j \sin \frac{j\pi x}{l}. \qquad (11.8)$$

Здесь $y(x)$ — горизонтальное отклонение стержня, рассматриваемое как функция расстояния x от одного из концов; l — координата другого конца стержня. Поскольку коэффициенты Фурье a_j определяют форму стержня, они играют роль переменных состояния системы; прилагаемая сила F играет роль управляющего параметра.

Фиксированная длина L несжимаемого стержня выражается через параметр l и отклонение $y(x)$ следующим образом:

$$L = \int_0^l \sqrt{1 + (dy/dx)^2}\, dx. \tag{11.9}$$

Уравнение (11.9) задает ограничение на параметр l и переменные состояния $a_1, a_2, \ldots$, которое может быть представлено также в виде

$$L = \int_0^l \left\{1 + \frac{1}{2}\left(\frac{dy}{dx}\right)^2 - \frac{1}{8}\left(\frac{dy}{dx}\right)^4 + \ldots\right\} dx =$$

$$= l + \frac{1}{2}\sum_{j=1}^{\infty} a_j^2 \left(\frac{j\pi}{l}\right)^2 \frac{l}{2} + \text{Члены более высокой степени.} \tag{11.10}$$

Из членов высшего порядка для дальнейшего рассмотрения важен лишь член четвертой степени по a_1: $-l\left(3/2^{+6}\right)(\pi/l)^4 a_1^4$.

Потенциальная энергия, накопленная в изогнутом стержне, пропорциональна

$$\frac{B}{2}\int_0^l \left(\frac{d^2y}{dx^2}\right)^2 dx = \frac{B}{2}\,\frac{l}{2}\sum_{j=1}^{\infty}\left(\frac{j\pi}{l}\right)^4 a_j^2 \tag{11.11}$$

(константа B называется *изгибной жесткостью*). Выполненная внешней силой работа равна

$$W = \int_L^l F \cdot dx = -F(L - l). \tag{11.12}$$

Потенциальная функция, описывающая состояние (форму) совершенного статического стержня, является суммой (11.11) и (11.12), т. е.

$$V_p(a_j;\ F) = \frac{Bl}{4}\sum_{j=1}^{\infty}\left(\frac{j\pi}{l}\right)^4 a_j^2 - \frac{Fl}{4}\sum_{j=1}^{\infty}\left(\frac{j\pi}{l}\right)^2 a_j^2 +$$

$$+\text{Члены более высокой степени}=\frac{l}{4}\sum_{j=1}^{\infty}\left(\frac{j\pi}{l}\right)^2\left[B\left(\frac{j\pi}{l}\right)^2-F\right]a_j^2+$$

$$+\text{Члены более высокой степени.} \qquad (11.13)$$

(Заметим, что при выводе формулы (11.13) вес стержня не учитывался.)

Состояние стержня определяется минимумом потенциальной функции $V_p(a_j;\ F)$, которая положительно определена при $F < F_1 = B(\pi/l)^2$. Как у функции возрастающей нагрузки F, первая неморсовская критическая точка потенциальной функции имеет место при $F = F_1$. В пределах неустойчивой ветви, отвечающей недеформированному стержню, дополнительные неморсовские точки существуют при F_2, F_3, ..., где $F_j = B(j\pi/l)^2$. Функция $V_p(a_j;\ F)$ будет морсовской для всех значений F, за исключением $F = F_j$ $(j = 1,\ 2,\ \ldots)$. При этих же значениях $V_p(a_j;\ F)$ является неморсовской функцией переменной состояния a_j и морсовской функцией всех остальных переменных состояния.

Для того чтобы описать состояние стержня при $F > F_1$, т. е. после выпучивания, необходимо рассмотреть члены выше второй степени по переменной состояния a_1. (Всеми остальными переменными состояния можно пренебречь.) Потенциальная функция, описывающая состояние стержня, будет иметь вид

$$V_p(a_j;\ F) \to V(a_1;\ F)=\frac{l}{4}\left(\frac{\pi}{l}\right)^2(F_1-F)a_1^2+$$

$$+\frac{3Fl}{2^6}\left(\frac{\pi}{l}\right)^4 a_1^4+\ldots\ . \qquad (11.14)$$

При $F > F_1$ величина первого коэффициента Фурье определяется формулой

$$a_1^2=\frac{8}{3F}\frac{F-F_1}{(\pi/l)^2},\quad F\geqslant F_1. \qquad (11.15)$$

Если стержень зафиксирован (ограничен) так, что он не может перейти в конфигурацию с меньшей энергией $(j = 1)$, при большем значении $F(=F_2)$ он перейдет в следующую высшую $(j = 2)$ конфигурацию, т. е. произойдет выпучивание. Вообще говоря, если первые $j-1$ мод выпучивания запрещены ограничениями, то выпучивание будет иметь место при j-й моде, форма которой определяется членом $\sin j\pi x/l$ при $F_j = B(j\pi/l)$. Значения a_j при $F > F_j$ приведены на рис. 11.3.

До сих пор мы исследовали статические свойства идеального стержня. Теперь рассмотрим влияние дефектов формы стержня на его состояние. Наиболее общий вид возмущения потенциальной

функции (11.13) включает также линейный член, поэтому

$$V_i(a_i;\ F,\ \varepsilon) = \varepsilon a_1 + \frac{l}{4}\left(\frac{\pi}{l}\right)^2 (F_1 - F)\, a_1^2 + \frac{3Fl}{2^6}\left(\frac{\pi}{l}\right)^4 a_1^4, \qquad (11.16)$$

т. е. наиболее общий вид начального дефекта стержня может быть смоделирован отличным от нуля изгибом (искривлением). Значения a_1 в состоянии равновесия определяются из условия равенства нулю градиента ($\nabla_1 V_i(a_1;\ F,\ \varepsilon) = 0$) (рис. 11.4). Как следует из рис. 11.4, приведенные кривые являются результатом

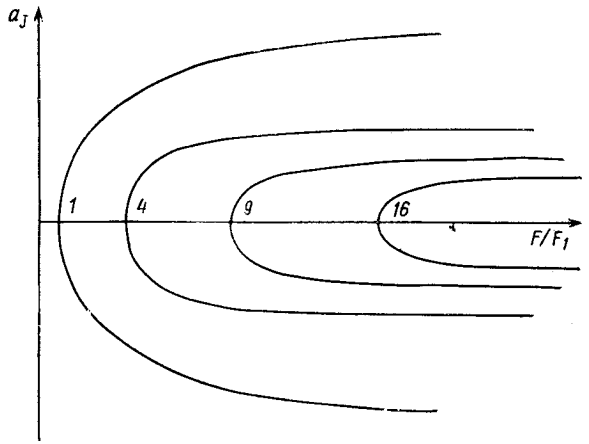


Рис. 11.3. При значительной нагрузке один из первых коэффициентов Фурье a_i может быть отличен от нуля.

сечения многообразия катастрофы сборки $x^3 + ax + b = 0$ плоскостями $b = \text{const}$. Устойчивость стержня вдоль кривых равновесия определяется свойствами устойчивости катастрофы сборки.

Если стержень находится не в статическом состоянии и его кинетическая энергия ΔE полностью определяется нижней модой, то величина параметра a_1 будет колебаться около соответствующего статическому состоянию стержня значения $a_1(F)$ (рис. 11.5).

Выпучивание нагруженного эйлерова стержня по существу представляет собой фазовый переход второго рода. Переход к выпученному состоянию является «мягким», так как состояния системы до и после выпучивания связаны непрерывным образом. Конструкции, демонстрирующие «мягкий» переход в выпученное состояние, не разрушаются при превышении предельной нагрузки — они лишь умеренно изгибаются. Поэтому можно сформулировать некоторый вид критериев безопасности для

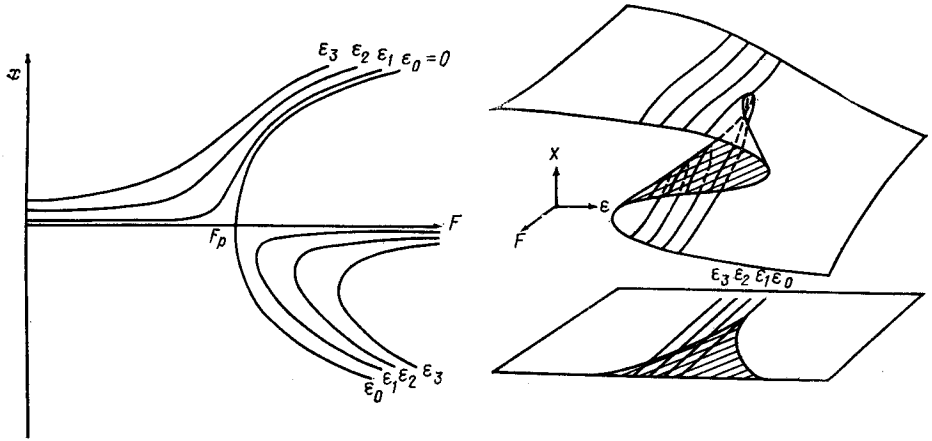


Рис. 11.4. Зависимость равновесной конфигурации несовершенного стержня от прилагаемой нагрузки и параметра несовершенства.

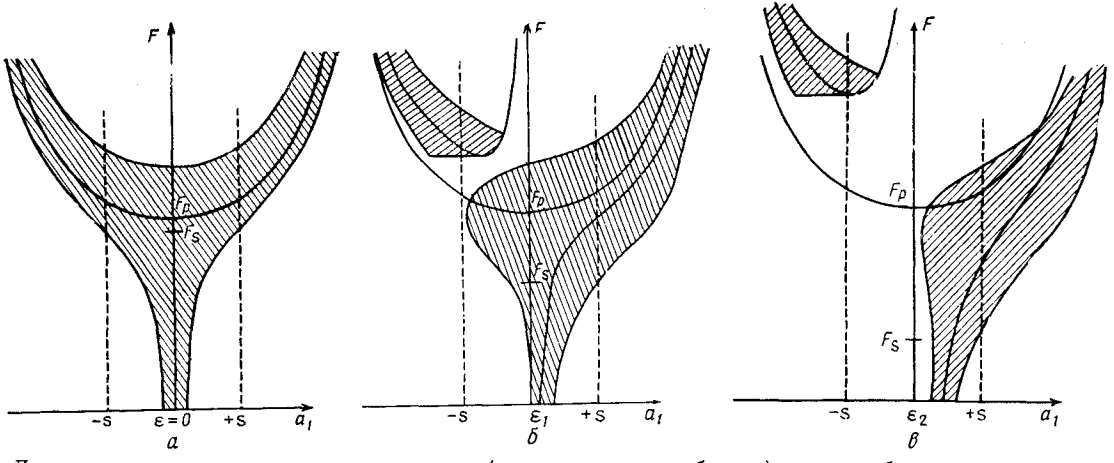


Рис. 11.5. Диапазон изменения параметров порядка (заштрихованные области) при колебаниях стержня около его равновесной конфигурации для различных значений параметра несовершенства ε.

определения пределов безопасных нагрузок. Так, для многих практических целей состояние стержня является опасным, если амплитуда начального изгиба стержня превышает предписанное безопасное значение s: $|a_1| > s > 0$. Так как амплитуда изгиба определяется из условия равенства нулю градиента, то максимальная безопасная нагрузка F_s будет определяться максимальным безопасным изгибом, т. е.

$$\nabla_1 V(a_1 = s;\ F_s,\ \varepsilon) = 0 = \varepsilon + \frac{l}{2}\left(\frac{\pi}{l}\right)^2 (F_1 - F_s)\, s + \\ + \frac{3F_s l}{2^4}\left(\frac{\pi}{l}\right)^4 s^3, \qquad (11.17)$$

$$F_s = F_c(s) - k(s)\,\varepsilon,$$

где

$$F_c(s) = \frac{F_1}{1 - \frac{3}{8}(\pi s/l)^2}, \qquad (11.18)$$

$$k(s) = \frac{(2/|s|\,l)\,(l/\pi)^2}{1 - \frac{3}{8}(\pi s/l)^2}.$$

Чувствительность безопасной нагрузки к несовершенству достаточно мягкая и зависит от параметра несовершенства в первой степени. Здесь $F_c(s)$ является безопасной нагрузкой в отсутствие дефектов. Для достаточно больших дефектов безопасных нагрузок не существует.

Аналогично может быть определена максимальная несущая способность колеблющегося стержня. Колебания имеют место вблизи состояний статического равновесия (см. рис. 11.5). Максимальная несущая способность может быть определена как значение F, при котором амплитуда изгиба достигает s. Для совершенной системы эта нагрузка находится из выражения

$$\Delta E = \frac{l}{4}\left(\frac{\pi}{l}\right)^2 (F_1 - F_s)\, s^2 + \frac{3F_s l}{2^6}\left(\frac{\pi}{l}\right)^4 s^4, \qquad (11.19)$$

что в свою очередь приводит к следующему линейному соотношению между кинетической энергией ΔE и максимальной безопасной нагрузкой:

$$F_s = \frac{F_1 - (\Delta E/l)\,(2l/\pi s)^2}{1 - 3\,(\pi s/4l)^2}. \qquad (11.20)$$

Для выпучивающихся стержней и других систем, разрушение которых происходит по схеме катастроф типа A_{+3}, чувствительность безопасной нагрузки как к несовершенству, так и к динамическим воздействиям будет весьма умеренной.

3. ПРОЩЕЛКИВАНИЕ ПОЛОГОЙ АРКИ

Работающие на сжатие балки — отнюдь не единственные элементы конструкций. Для перекрытий мостовых пролетов чрезвычайно эффективным является использование пологой арки (рис. 11.6). Если малые вертикальные нагрузки не вызывают деформации арки, то большие нагрузки приводят к ее разрушению. Попытаемся установить, что будет происходить с аркой при промежуточных нагрузках, и, в частности, определим критическую нагрузку, а также чувствительность арки как к несовершенству, так и к динамическому воздействию.

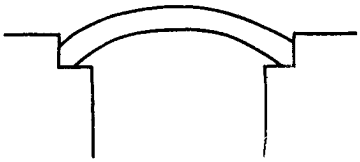


Рис. 11.6. Пологая арка — типичный элемент конструкции.

Удобно предположить, что пологая арка есть не что иное, как несжимаемая балка, к которой приложена вертикальная сила (рис. 11.2). Тогда равновесная форма такой балки при отсутствии нагружающих сил и начальных дефектов определяется как

$$y(x) = a_1^0 \sin\frac{\pi x}{l}. \tag{11.21}$$

Состояние (или форма) арки под нагрузкой может быть аналитически определено с помощью анализа ряда Фурье. Вычисления могут быть выполнены в бесконечномерном пространстве состояний, в котором переменными состояния являются коэффициенты ряда Фурье a_j. В случае прощелкивания пологой арки бесконечномерное пространство состояний может быть заменено конечномерным пространством. Для этого достаточно ограничиться двумя первыми коэффициентами Фурье:

$$y(x) = a_1 \sin\frac{\pi x}{l} + a_2 \sin\frac{2\pi x}{l}. \tag{11.21'}$$

Эти два коэффициента не являются независимыми и связаны условием постоянства длины арки:

$$L = l + \frac{l}{4}\left[\left(\frac{\pi a_1}{l}\right)^2 + \left(\frac{2\pi a_2}{l}\right)^2\right] - \\ - \frac{l}{8}\left[\frac{3}{8}\left(\frac{\pi a_1}{l}\right)^4 + \frac{3}{2}\left(\frac{\pi a_1}{l}\right)^2\left(\frac{2\pi a_2}{l}\right)^2 + \frac{3}{8}\left(\frac{2\pi a_2}{l}\right)^4\right]. \tag{11.22}$$

Соотношение между L, l и a_1^0 получается из (11.22), если $a_1 \to a_1^{(0)}$, $a_2 \to 0$. Не имеет смысла оставлять в разложении чле-

ны степени выше четвертой, так как единственный управляющий параметр F может быть использован для исключения лишь одного коэффициента ряда Тейлора. (После интегрирования ряд Фурье от x становится рядом Тейлора с коэффициентами Фурье.)

Энергия, накопленная в пологой арке, такая же, что и в изогнутом стержне, и ее величина дается формулой (11.11). Работа, совершенная внешней нагрузкой F, равна $F(a_1^0 - a_1)$. Потенциальная функция, описывающая статическую идеальную пологую арку, имеет вид

$$V_p(a_1, a_2, F) = \frac{Bl}{4}\left[\left(\frac{\pi}{l}\right)^4 a_1^2 + \left(\frac{2\pi}{l}\right)^4 a_2^2\right] - F(a_1^0 - a_1). \qquad (11.23)$$

Эта задача с ограничением и двумя переменными a_1 и a_2 может быть преобразована в задачу без ограничений с единственной переменной посредством использования условия, согласно которому переменная a_1 может быть выражена через a_2, или наоборот. Отметим, что это условие инвариантно относительно замены $a_1 \to \pm a_1$, $a_2 \to \pm a_2$. Переменная a_1 может быть представлена как функция от a_2 следующим образом:

$$a_1 = \pm f_1(a_2), \qquad (11.24)$$

где $a_1^0 = f_1(0)$.

Знаки перед функцией соответствуют двум совершенно различным физическим ситуациям: положительный знак — нормальной конфигурации арки (которая является типичным элементом конструкции); отрицательный знак — опрокинутой арке (которая практически не используется). Если вместо a_1 исключить a_2, то следует выбрать определенный знак, так как в (11.23) входит линейный член.

Положительное и отрицательное значения амплитуды второй гармоники a_2 характеризуют различные физические ситуации (рис. 11.7). При фиксированном a_1 потенциальная функция $V_p(a_1, a_2; F)$, очевидно, инвариантна относительно замены $a_2 \to -a_2$. Следовательно, полезно исключать a_1 $(a_1 > 0)$, а не a_2. Тогда потенциальная функция имеет вид

$$V_p(a_2; F) = A_0 + \frac{1}{2}A_2 a_2^2 + \frac{1}{4}A_4 a_2^4, \qquad (11.25)$$

$$A_0 = \frac{Bl}{4}\left(\frac{\pi}{l}\right)^4 (a_1^0)^2,$$

$$\frac{1}{2}A_2 = 3Bl\left(\frac{\pi}{l}\right)^4 - \frac{2F}{a_1^0}, \qquad (11.26)$$

$$\frac{1}{4}A_4 < 0$$

(постоянный член не играет существенной роли). Линейный и другие нечетные члены отсутствуют в силу соотношений инвариантности. Квадратичный член показывает, что арка сохраняет свою форму при $F < F_p = \frac{3}{2} Bla_1^0 (\pi/l)^4$. Член четвертой степени показывает, что совершенная арка разрушается ($a_1 > 0 \rightarrow$ $\rightarrow a_1 < 0$) при $F > F_p$, так как нет близкого устойчивого состояния равновесия. Потенциальная функция $V_p(a_2; F)$ проявляет катастрофу A_{-3} при $F = F_p$.

Математически разрушающаяся арка описывается с помощью катастрофы двойной сборки. Первоначально имеются два неустойчивых состояния равновесия с $a_2 \neq 0$ вблизи

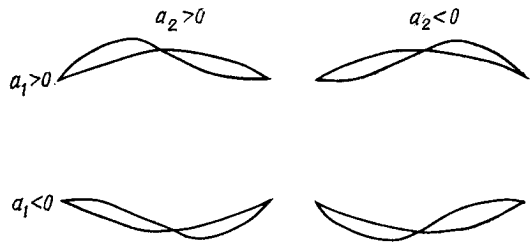


Рис. 11.7. Два разных знака перед a_1 отвечают совершенно различным физическим ситуациям. Два разных знака перед a_2 символизируют различные, но эквивалентные разрушающие моды.

устойчивого состояния равновесия $a_2 = 0$. Эти два неустойчивых состояния равновесия соответствуют положениям «прощелкивания». При возрастании нагрузки F прощелкивание арки наблюдается при меньших значениях $|a_2|$. При приближении к критической нагрузке прощелкивание может иметь место уже при $|a_2| \rightarrow 0$, и небольшое возмущение может оказаться причиной разрушения арки.

Физически прощелкивание происходит следующим образом. Вертикальная сила, приложенная к центру арки, стремится сместить ее центр тяжести вниз. Для того чтобы центр тяжести арки сместился вниз, необходимо добавить к ее форме высшие гармоники, что возможно только при увеличении энергии прогиба. Иными словами, смещение центра тяжести может произойти только тогда, когда приращение энергии за счет работы, описываемой выражением $[F(a_1^0 - a_1)]$, превысит приращение энергии деформации арки. При этом как только арка начнет двигаться вниз, ничто не может удержать ее от прощелкивания за неустойчивое положение в опрокинутое. Следовательно, можно считать, что совершенная пологая арка разрушается вследствие превышения критической нагрузки F_p.

Несовершенства пологой арки могут быть обусловлены неоднородностью строительных материалов, перемещением точки нагружения и тысячью других причин. Наиболее общий вид деформации ростка катастрофы сборки $\pm x^4$ таков: $\frac{1}{2}\varepsilon_2 x^2 + \varepsilon_1 x$. Следовательно, потенциальная функция, описывающая состояние (или форму) несовершенной пологой арки, имеет вид

$$V_i(x;\ F,\ \varepsilon_1,\ \varepsilon_2) = \varepsilon_1 x + \frac{1}{2}(F_p - F + \varepsilon_2)\, x^2 - \frac{1}{4}\, x^4. \quad (11.27)$$

[Для простоты изложения считаем $a_2 \simeq x$, и коэффициенты пропорциональности и масштаб F выбираем такими, чтобы получить наиболее простые коэффициенты в (11.27).]

При возрастании внешней нагрузки управляющие параметры следуют линейной траектории сборки в пространстве управляющих параметров. Система будет оставаться в локально устойчивом состоянии, соответствующем среднему листу многообразия катастрофы двойной сборки до тех пор, пока не достигнет бифуркационного множества. Таким образом, чувствительность к несовершенству у пологой арки будет иметь следующий вид:

$$F_c - F_p + \varepsilon_2 = -k\,|\,\varepsilon_1\,|^{2/3},$$

или

$$F_c = F_p - \varepsilon_2 - k\,|\,\varepsilon_1\,|^{2/3}. \quad (11.28)$$

Пологая арка значительно более чувствительна к несовершенству, нарушающему симметрию ($\varepsilon_1 \neq 0$), чем к несовершенству, ее сохраняющему. Поэтому можно ограничиться рассмотрением только дефектов, нарушающих симметрию. Наиболее общее нарушающее симметрию несовершенство эквивалентно перемещению прилагаемой силы на расстояние ε_1 от плоскости симметрии.

Тот факт, что в законе зависимости разрушающей нагрузки от параметра несовершенства ε надо взять степень 2/3, был показан в серии экспериментов, проведенных Рурдой [3]. В этих экспериментах параметром несовершенства было относительное смещение f/L прилагаемой силы от плоскости симметрии пологой арки. Данные экспериментов показаны на рис. 11.8.

Теперь рассмотрим уменьшение несущей способности арки вследствие динамического нагружения. Для совершенной системы, подвергаемой колебаниям в моде разрушения с кинетической энергией ΔE, критическая нагрузка определяется из выражения

$$\Delta E = \frac{1}{2}(F_p - F)\, x^2 - \frac{1}{4}\, x^4. \quad (11.29)$$

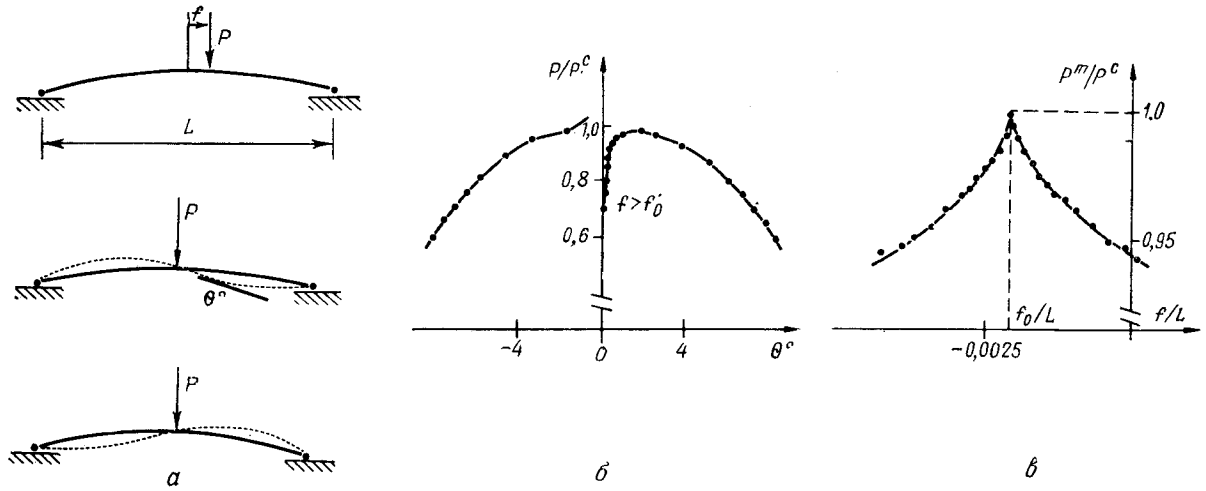


Рис. 11.8.

а — экспериментально реализуемое расположение силы (относительное смещение f/L прилагаемой силы P от центральной линии рассматривается как параметр несовершенства пологой арки); *б* — графическое представление устойчивых и неустойчивых равновесных конфигураций пологой арки в зависимости от прилагаемой нагрузки при фиксированных значениях $\varepsilon = f/l$ [3]. Устойчивые состояния равновесия имеют место лишь на центральной ветви и исчезают в точке P^M/P^C горизонтального касания. Последняя определяет разрушающую нагрузку, при которой наблюдается прощелкивание; *в* — приведенная разрушающая нагрузка как функция приведенного параметра несовершенства [7]. Сборка является откликом на несовершенства, внутренне присущие модели. Заметим, что при значении параметра несовершенства 1/4% несущая способность арки снижается на 6%. Отметим также, что модель может быть усилена посредством какого-либо несовершенства, сдвигающего сборку назад вправо.

Следовательно,

$$F_c = F_p - 2|\Delta E|^{1/2}. \tag{11.30}$$

Уменьшение несущей способности несовершенной системы может быть определено из масштабных соображений. Например,

$$\begin{gathered} x \to \lambda x, \\ |F_c - F_p| \to \lambda^2 |F_c - F_p| \Rightarrow \Delta E \to \lambda^4 \Delta E, \\ |\varepsilon_1| \to \lambda^3 |\varepsilon_1|. \end{gathered} \tag{11.31}$$

Такая поверхность может быть построена исходя из канонических свойств катастрофы сборки. Поверхность разрушения в пространстве $F - \Delta E - \varepsilon_1$ изображена на рис. 11.9. Вся по-

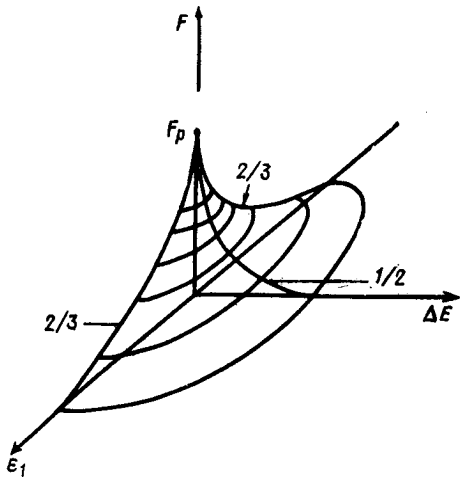


Рис. 11.9. Форма критической поверхности пологой арки в пространстве $F - \Delta E - \varepsilon_1$.

Высокая чувствительность к статическому несовершенству перекрывается еще большей чувствительностью к динамическому несовершенству.

верхность может быть восстановлена на основе масштабных соотношений (11.31) и какого-либо произвольного поперечного сечения. Ясно, что пологая арка более чувствительна к динамическим несовершенствам, чем к несовершенствам, нарушающим симметрию.

◇◇◇ Если математически различие между ростками катастроф сборки и двойной сборки выражается лишь в смене знака, то физически это различие существенно. Катастрофа сборки является глобально устойчивой, в то время как двойная сборка глобально неустойчива. У первой всегда имеется некоторое устойчивое состояние; для последней устойчивые состояния

существуют лишь в пределах области, имеющей форму сборки в плоскости управляющих параметров. В случае систем, описываемых катастрофой сборки, необходима субъективная оценка критерия безопасной нагрузки; так, при рассмотрении эйлерова стержня требуется, чтобы параметр порядка a_1 оставался по величине меньшим, чем некоторое максимальное безопасное отклонение s. Для систем, описываемых посредством двойной сборки, подобных субъективных критериев не требуется, поскольку существует объективный критерий: система разрушается при превышении предела критической нагрузки.

4. СМЕНА ТИПА УСТОЙЧИВОСТИ

В двух предыдущих примерах потенциальная функция, описывающая совершенную систему, была инвариантна относительно преобразования симметрии $x \to -x$, где x — параметр порядка системы (т. е. первый или второй коэффициент Фурье). Это, строго говоря, верно только тогда, когда мы не ограничиваемся разложениями (11.10), содержащими лишь первые два коэффициента Фурье. В данном случае мы можем сделать такое предположение, основываясь на том, что энергия прогиба быстро возрастает при переходе к последовательно более высоким модам прогиба.

Так как совершенная система была описана при помощи четной потенциальной функции $V_p(x; F)$, зависящей от единственного управляющего параметра F, то можно было бы потребовать обращения в нуль всего лишь одного коэффициента ряда Тейлора $[\sim (F_p - F)x^2]$, что неминуемо приводило бы к рассмотрению катастрофы сборки ($\sim x^4$) либо двойной сборки ($\sim -x^4$).

При отсутствии симметрии разложение в ряд Тейлора потенциальной функции $V_p(x; F)$, описывающей некоторую совершенную систему, будет иметь вид

$$V_p(x; F) = V_0 + V_1 x + \frac{1}{2} V_2 x^2 + \frac{1}{3!} V_3 x^3 + \ldots . \qquad (11.32)$$

В общем случае выберем параметр порядка x так, чтобы совершенная система имела состояние равновесия в $x = 0$. Тогда $V_1 = 0$. (Постоянный член не имеет существенного значения и может быть исключен путем переноса начала координат.) Квадратичные, кубические и члены более высокой степени в общем случае отличны от нуля. Если рассматривать изменение системы в зависимости от возрастающей прилагаемой силы F в точке $F = F_p$, то потенциальная функция может быть записана в виде

$$V_p(x; F) = \frac{1}{2}(F_p - F)x^2 + \frac{1}{3} x^3 + \ldots \qquad (11.33)$$

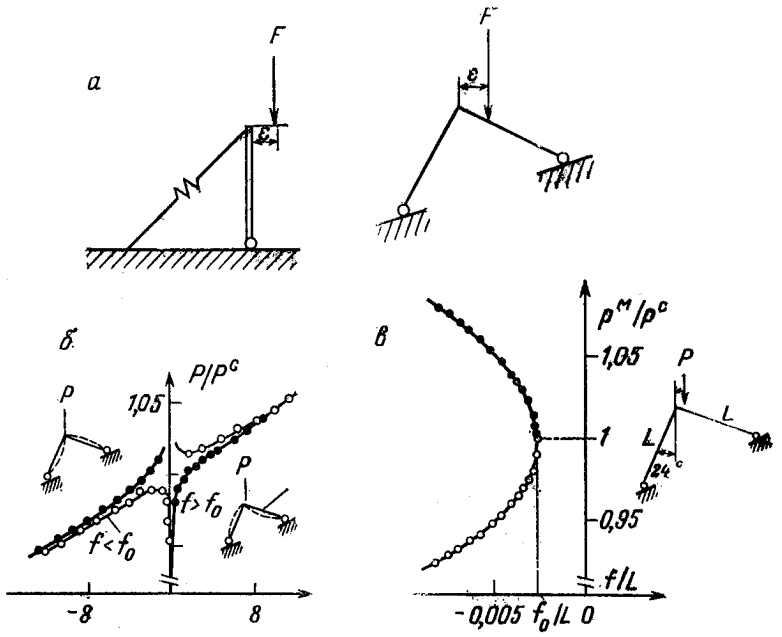


Рис. 11.10.

а — поддерживающий кронштейн и жестко сочлененная рама — две совершенные системы, в которых наблюдается смена типа устойчивости. Несовершенство моделируется с помощью смещения нагрузки; *б* — экспериментально определенные устойчивые и неустойчивые состояния равновесия совершенной жестко сочлененной рамы [3]; *в* — экспериментально найденная чувствительность к несовершенству подчиняется степенной (с показателем степени 1/2) зависимости [7]. Заметим, что при значении параметра несовершенства 1/4% величина разрушающей нагрузки снижается на 5%. Отметим также, что точка (место) разрушения не находится на касательной к оси $f/L = 0$, так как несовершенство внутренне присуще модели.

посредством замены масштабов по осям x и F. Членами четвертой и более высокой степени можно пренебречь. Две совершенные системы, описываемые потенциальной функцией, эквивалентной (11.33), изображены на рис. 11.10.

Критические точки, соответствующие V_p, определяются, как обычно, из соотношения

$$\frac{d}{dx} V_p = 0 = x\{(F_p - F) + x\}. \qquad (11.34)$$

Положение критических точек и тип их устойчивости показаны на рис. 11.11: смена типа устойчивости происходит в момент, когда две критические точки $x_1(F) = 0$, $x_2(F) = F - F_p$ проходят одна через другую.

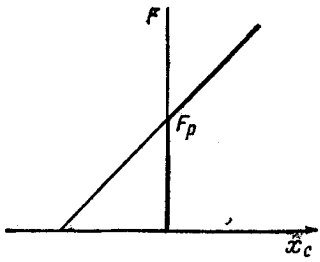


Рис. 11.11. Положения критических точек в зависимости от прилагаемой нагрузки Зависимость для устойчивой критической точки изображена жирной линией. Наблюдается смена типа устойчивости.

Потенциальная функция, описывающая соответствующую несовершенную систему, имеет вид

$$V_t(x;\ F,\ \varepsilon_1,\ \varepsilon_2,\ \ldots) = V_p(x;\ F) + p(x),$$
$$p(x) = \varepsilon_1 x + \frac{1}{2}\varepsilon_2 x^2 + \frac{1}{3}\varepsilon_3 x^3 + \ldots\ . \qquad (11.35)$$

Результаты, представленные в гл. 4, дают нам все сведения о наиболее общей деформации ростка x^3: это будет линейный член. Короче говоря, формула (11.35) может быть преобразована в каноническую форму посредством соответствующей нелинейной замены. Это полностью справедливо с математической точки зрения, но совершенно не удовлетворительно с точки зрения физики явления; такое нелинейное преобразование может привести к возникновению сложной нелинейной связи между физическим параметром (нагрузкой) F и параметрами несовершенства ε_1, ε_2, ... Поэтому, вместо того чтобы проводить нелинейное преобразование, которое приводит (11.35) к канонической форме, поступим следующим образом. Отбросим в деформационном разложении члены выше второй степени. (Квадратичные члены оставим только для того, чтобы показать, что они играют менее важную роль, чем линейные. В конечном счете их также отбросим.) Тогда потенциальная функция, описывающая несовершенную систему, примет вид

$$V_t(x;\ F,\ \varepsilon_1) = \varepsilon_1 x + \frac{1}{2}(F_p' - F)x^2 + \frac{1}{3}x^3, \qquad (11.36)$$

где $F_p' = F_p + \varepsilon_2$. Критические точки определяются соотношением

$$\frac{dV_t}{dx} = \varepsilon_1 + (F_p' - F)x + x^2 = 0. \qquad (11.37)$$

Уравнение (11.37) представляет двумерное многообразие, вложенное в пространство $\mathbb{R}^3$ с координатными осями $x - F - \varepsilon_1$ (рис. 11.12, *а*).

В этом случае состояниями равновесия несовершенной системы являются сечения этого многообразия плоскостями $\varepsilon_1 =$

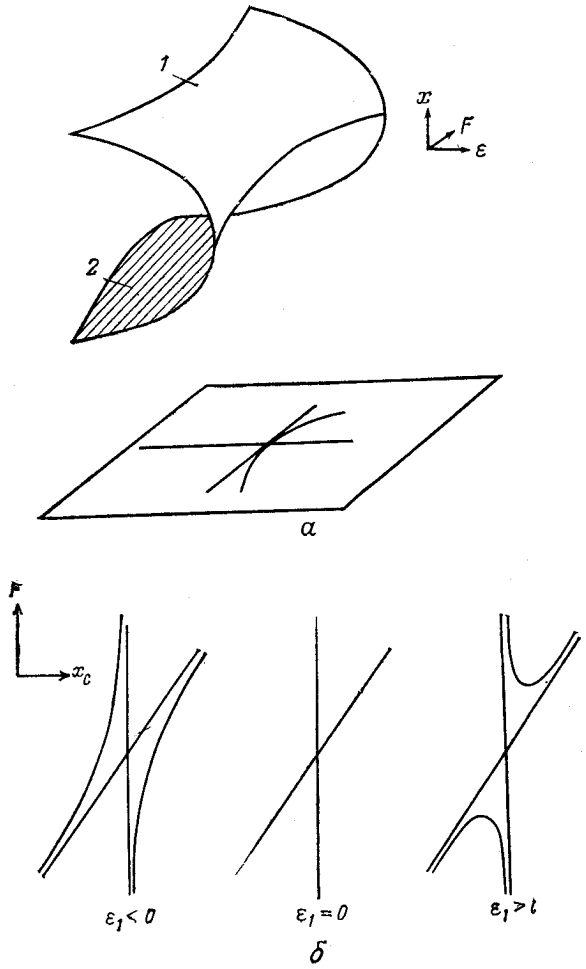


Рис. 11.12.

а — форма многообразия состояний равновесия в пространстве переменных состояния и управляющих параметров $x - F - \varepsilon_1$ для несовершенного поддерживающего кронштейна. Все точки на верхней части этой поверхности представляют локально устойчивые состояния, а точки на нижнем листе — локально неустойчивые состояния. Верхняя и нижняя ветви имеют сепаратрису, которая проектируется на параболу $\varepsilon_1 = \left(F - F_p'/2\right)^2$, лежащую в пространстве управляющих параметров. *1* — устойчивые точки, *2* — неустойчивые точки; *б* — траектории до и после выпучивания как сечения многообразия состояний равновесия плоскостью $\varepsilon_1 = \text{const}$. Свойства устойчивости вдоль этих траекторий определяются непосредственно из рассмотрения многообразия состояний равновесия.

$=$ const (рис. 11.12, *б*). Свойства устойчивости критических точек легко определяются: все точки на верхнем листе критического многообразия представляют собой локально устойчивые критические точки V_i, все точки на нижнем листе — неустойчивые критические точки.

Поведение несовершенной системы в зависимости от нагрузки F критически зависит от знака параметра несовершенства ε_1: при $\varepsilon_1 < 0$ для каждого значения F существуют две критические точки; при $\varepsilon_1 > 0$ имеется область, в которой V_i вообще не имеет критических точек. Последняя существует при

$$-\varepsilon_1 + \left(\frac{F_p' - F}{2}\right)^2 < 0. \tag{11.38}$$

Чувствительность системы к несовершенству выражается формулой

$$F_c = F_p + \varepsilon_2 - 2(\varepsilon_1)^{1/2}, \quad \varepsilon_1 > 0, \tag{11.39а}$$

$$F_c = \text{“}\infty\text{“}, \quad \varepsilon_1 < 0. \tag{11.39б}$$

При $\varepsilon_1 \leqslant 0$ локально устойчивое состояние равновесия существует при всех значениях внешней силы F (ситуация, аналогичная выпучиванию эйлерова стержня) и можно ввести субъективный критерий безопасности: *система* безопасна, если силы критической нагрузки при значении переменной состояния x превышают некоторое предписанное безопасное значение s: $|x| > s$.

При $\varepsilon_1 > 0$ необходимость в субъективном критерии отпадает. Локально устойчивое состояние равновесия перестает существовать при критической нагрузке F_c, определяемой формулой (11.39а). Очевидно, что чувствительность к несовершенству очень слабо зависит от ε_2, но сильно зависит от ε_1. Следовательно, параметр ε_1 более важен, чем ε_2, а ε_2 более важен, чем ε_3, и т. д. Именно это обстоятельство и позволяет исключить все возмущения, кроме линейного.

Степенная зависимость разрушающей нагрузки (с показателем степени 1/2) от параметра несовершенства была обоснована Рурдой для систем, которые претерпевают смену типа устойчивости в отсутствие несовершенства. Устойчивые и неустойчивые критические точки для системы, изображенной на рис. 10.10, *а*, показаны на рис. 10.10, *б*, а чувствительность к несовершенству — на рис. 10.10, *в*.

Когда потенциал (11.36) имеет две точки равновесия, неустойчивое равновесие действует как сепаратриса между областью притяжения локально устойчивого состояния и первичным хаосом. Разность энергий ΔE в локальном максимуме и

локальном минимуме равна

$$\Delta E = \frac{1}{6}[(F - F_p)^2 - 4\varepsilon_1]^{3/2}. \qquad (11.40)$$

Если система подвергается динамическому нагружению ΔE, которое можно трактовать как динамическое несовершенство, то соотношение между разрушающей нагрузкой F_c, параметром несовершенства ε_1 и фактором динамического нагружения ΔE выражается формулой

$$F_c = F_p - [4\varepsilon_1 + (6\Delta E)^{2/3}]^{1/2}. \qquad (11.41)$$

При отсутствии динамического нагружения чувствительность к разрушающей нагрузке сводится к зависимости по степенному закону с показателем степени 1/2 (11.39а). Однако при отсутствии несовершенств чувствительность к динамическому нагружению оказывается более жесткой (рис. 11.13).

Для несовершенных систем с конечной критической нагрузкой, даваемой формулой (11.39а), естественно ожидать динамические флуктуации, снижающие несущую способность. Кроме того, возможна очень жесткая чувствительность совершенной системы к динамическим факторам. Потенциальная функция, описывающая совершенную систему, является чрезвычайно плоской и глобально неустойчивой вблизи $F \simeq F_p$, так что даже очень слабые флуктуации могут вызвать переход системы через потенциальный барьер. Может показаться неожиданным, что динамические колебания могут перевести «устойчивый» случай $\varepsilon_1 < 0$ в неустойчивый. Флуктуации, превышающие разность критических значений, способны вытолкнуть систему из ее очень устойчивого состояния через потенциальный барьер в бездонные потенциальные ямы. Эти переходы из устойчивого состояния в неустойчивое в результате динамических процессов вызывают обычно беспокойство и служат источником дальнейших размышлений.

5. СОСТАВНЫЕ СИСТЕМЫ

Рассмотрим системы, сконструированные из двух компонент. Предположим, что эти две компоненты с параметрами порядка x и y являются совершенными, а также инвариантными относительно замен $x \to \pm x$, $y \to \pm y$. Предположим также, что составная система конструируется совершенно, т. е.

$$V_p(x, y; F) = V_p(\pm x, \pm y; F). \qquad (11.42)$$

В этом случае главные члены разложения в ряд Тейлора потенциальной функции, описывающей такую совершенную систему,

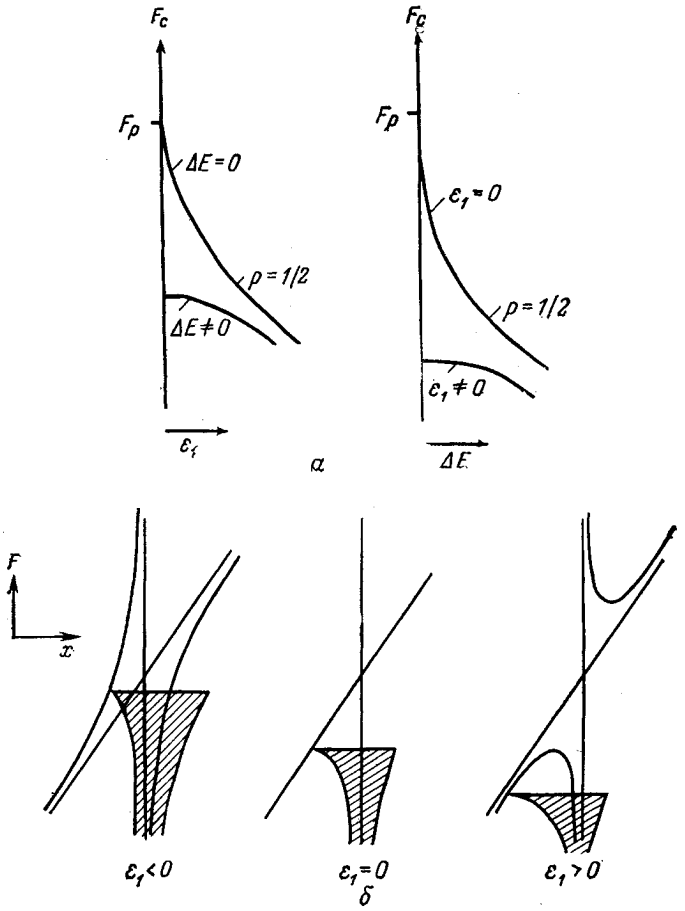


Рис. 11.13.

а — чувствительность к несовершенству систем, проявляющих смену типа устойчивости, крайне жестка. Чисто статические и чисто динамические несовершенства описываются степенными законами с показателями степени 1/2 и 1/3 соответственно; *б* — динамические несовершенства не только могут уменьшать несущую способность несовершенной системы, проявляющей смену типа устойчивости (при скачке $\varepsilon_1 \geqslant 0$), но могут также привести к скачку устойчивой системы через потенциальный барьер в хаос ($\varepsilon_1 < 0$). Заштрихованные области указывают пределы колебаний около устойчивого состояния равновесия. Система теряет устойчивость, когда колебания пересекают неустойчивый локальный максимум.

имеют вид

$$V_p(x, y; F) = \frac{1}{2}(F_1 - F)x^2 + \frac{1}{2}(F_2 - F)y^2 + \\ + \frac{1}{4}\sigma_1 x^4 + \frac{1}{4}\sigma_2 y^4 + \frac{1}{2}cx^2y^2 + \\ + \text{Члены более высокой степени;} \qquad (11.43)$$

при этом исключим из рассмотрения члены шестой и более высокой степени. Символы $\sigma_i = \pm 1$, $\sigma_2 = \pm 1$ введены для учета возможности того, что две индивидуальные компоненты разрушаются в результате катастроф типа $A_{\pm 3}$.

Прежде чем перейти к детальному анализу критических точек данной потенциальной функции, полезно сначала качественно определить ее свойства глобальной устойчивости. Это может быть сделано путем исключения квадратичных членов как несущественных по сравнению с членами четвертой степени при больших значениях $r^2 = x^2 + y^2$ (предполагается, что члены степени выше четвертой равны нулю). Например, если вклад членов четвертой степени в потенциальную функцию V_p равен

$$V_p \xrightarrow{r \text{ большое}} \frac{1}{4}x^4 + \frac{1}{4}y^4 - \frac{1}{2}\cdot 3x^2y^2,$$

$$= \frac{1}{4}(x^2 - y^2)^2 - x^2y^2, \qquad (11.44)$$

то V_p устойчива в «направлении» $x^2 - y^2$ и неустойчива в «направлении» $x^2 = y^2$ (т. е. при $x^2 - y^2 = 0$) (табл. 11.1).

Таблица 11.1. **Свойства глобальной устойчивости потенциальной функции вида** $V(x, y) = \frac{1}{4}\sigma_1 x^4 + \frac{1}{4}\sigma_2 y^4 + \frac{1}{2}cx^2y^2$

σ_1	σ_2	Условие	x^2-y^2	x^2y^2	«Морсовский тип»
$+1$	$+1$	$-1 < c$	$+$	$+$	M_0^2
		$c < -1$	$+$	$-$	M_1^2
$+1$ -1	-1 $+1$	Произвольное			M_1^2
		$+1 < c$	$-$	$+$	M_2^2
-1	-1	$c < +1$	$-$	$-$	M_2^2

Теперь свойства локальной устойчивости потенциальной функции (11.43) могут быть определены обычным способом. Условие равенства нулю градиента ведет к паре сцепленных нелинейных уравнений

$$\frac{\partial V}{\partial x} = x\{(F_1 - F) + \sigma_1 x^2 + cy^2\} = 0, \qquad (11.45а)$$

$$\frac{\partial V}{\partial y} = y\{(F_2 - F) + \sigma_2 y^2 + cx^2\} = 0. \qquad (11.45б)$$

Матрица устойчивости имеет вид

$$V_{ij}=\begin{bmatrix} F_1-F+3\sigma_1x^2+cy^2 & 2cxy \\ 2cxy & F_2-F+3\sigma_2y^2+cx^2 \end{bmatrix}. \quad (11.46)$$

Система уравнений (11.45) может иметь до девяти решений, перечисленных ниже:

1. $x=0,\ y=0,\quad V_{ij}=\begin{bmatrix} F_1-F & 0 \\ 0 & F_2-F \end{bmatrix}.$

2, 3. $F_1-F+\sigma_1x^2=0,\ y=0,\quad V_{ij}=\begin{bmatrix} -2(F_1-F) & 0 \\ 0 & F_2-F+cx^2 \end{bmatrix}.$

4, 5. $x=0,\ F_2-F+\sigma_2y^2=0,\quad V_{ij}=\begin{bmatrix} F_1-F+cy^2 & 0 \\ 0 & -2(F_2-F) \end{bmatrix}.$

6—9. $F_1-F+\sigma_1x^2+cy^2=0,\ F_2-F+\sigma_2y^2+cx=0,\quad V_{ij}=\begin{bmatrix} 2\sigma_1x^2 & 2cxy \\ 2cxy & 2\sigma_2y^2 \end{bmatrix}.$

Последнее множество решений существует, если квадратные уравнения соответствуют вещественным кривым второго порядка (эллипсам и/или гиперболам) и эти кривые пересекаются.

Множество решений $(x, y)=(0, 0)$ будем называть *нулевой ветвью*, или *стволом*. Это решение в силу симметрии существует при всех значениях параметра нагрузки. Множества решений, которые находятся в (x, F)-плоскости $y=0$ (№ 2, 3) и в (y, F)-плоскости $x=0$ (№ 4, 5), будем называть (если они существуют) *первичными ветвями*, а остальные четыре симметрических решения $(x\to\pm x,\ y\to\pm y)$ — *вторичными ветвями*. Первичные ветви ответвляются от нулевой ветви, вторичные — от первичных.

Определим свойства локальной устойчивости потенциальной функции V_p, выполняя элементарный анализ бифуркации. Предположим, что $F_2>F_1>0$. При $F=0$ решение $(x, y)=(0, 0)$ устойчиво, и оно остается утойчивым до тех пор, пока F не возрастет до $F=F_1$. В этой точке первичная ветвь $(x, y)=(\pm\sqrt{\sigma_1(F-F_1)}, 0)$ ответвляется от нулевой и существует только тогда, когда $F\geqslant F_1$ или $F\leqslant F_1$ в зависимости от типа катастрофы $A_{\pm3}$. Другая первичная ветвь ответвляется от нулевой ветви в точке $F=F_2$ и на ней $(x, y)=(0, \pm\sqrt{\delta_2(F-F_2)})$.

Вдоль нулевой ветви оси x и y являются главными, а одно собственное значение V_{ij} изменяет знак при каждой бифуркации. Вдоль первичных ветвей оси x и y также являются главными. Первая первичная ветвь $(x\neq0,\ y=0)$ представляет устойчивые морсовские критические точки M_0^2, если $\sigma_1=+1$,

или неустойчивые морсовские седла M_1^2, если $\sigma_1 = -1$. Вторая первичная ветвь ($x = 0$, $y \neq 0$) представляет неустойчивые морсовские седла M_1^2, если $\sigma_2 = +1$, или M_2^2, если $\sigma_2 = -1$.

Если вторичные ветви существуют, то они должны ответвляться от первичных ветвей, это следует из некоторых соображений, связанных с коническими сечениями. Наиболее простой способ нахождения положения этих бифуркаций состоит в изучении возможности обращения в нуль собственных значений

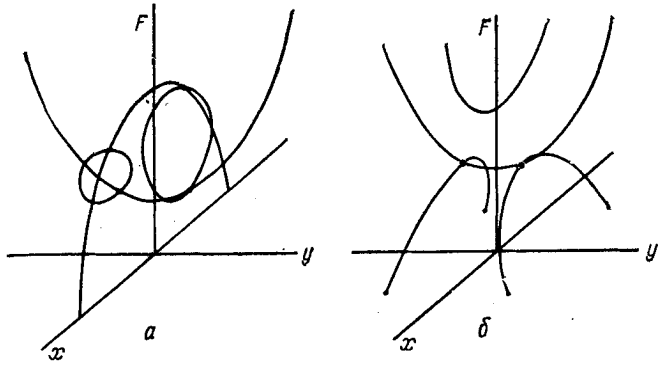


Рис. 11.14. Зависящие от коэффициентов квадратичных членов бифуркационные диаграммы для потенциальной функции вида (11.43).

а — вторичные ветви, «испускаемые» одной из первичных ветвей, «поглощаются» другой первичной ветвью; *б* — вторичные ветви являются морсовскими седлами M_1^2. Эти седла отделяют локально устойчивые состояния равновесия $(x, y) = (0, 0)$ от области глобальной неустойчивости, которая существует при больших r.

матрицы устойчивости вдоль первичных ветвей. Бифуркация на первой первичной ветви может произойти в точке

$$F = \frac{F_2 - c\sigma_1 F_1}{1 - c\sigma_1}, \qquad (11.47а)$$

если указанное значение F лежит на этой ветви. Аналогично бифуркация на второй первичной ветви может произойти в точке

$$F = \frac{F_1 - c\sigma_2 F_2}{1 - c\sigma_2}, \qquad (11.47б)$$

если это значение F лежит на ветви ($x = 0$, $y \neq 0$).

Две интересные бифуркационные диаграммы потенциальной функции (11.43) приведены на рис. 11.14. На первой из них вторичная ветвь «испускается» одной из первичных ветвей и «поглощается» другой. На второй диаграмме две первичные ветви обращены вверх, а вторичные ветви обращены вниз и неустойчивы. Нисходящие неустойчивые (M_1^2) вторичные ветви

являются потенциальными барьерами, отделяющими устойчивую (M_0^2) критическую точку $(x=0,\ y=0)$ от глобальной неустойчивости $(c<-1)$. При $F_2 \gg F_1$ или $F_1 \gg F_2$ эти седла далеко отстоят от нулевой ветви, и высота потенциального барьера значительна; при $F_2 - F_1 \simeq 0$ и малом $F_1 - F$ это уже не имеет места.

Анализ несовершенства может быть проведен для совершенных систем, описываемых потенциальной функцией (11.43); при $F_2 - F_1 \gg 0$ он сводится к анализу катастроф сборки (разд. 2 и 3). Ситуация становится более сложной, когда $F_2 \simeq F_1$. В этом случае, хотя обе бифуркации и являются «мягкими» (A_{+3}), вторичные ветви обращены вниз при $c<-1$, и, следовательно, может иметь место не только чрезвычайная, но, что хуже, непрогнозируемая чувствительность к несовершенству. Эта неожиданная чувствительность к несовершенству возникает из-за тесной взаимосвязи подобранных нагрузок разрушающих мод обеих компонент подсистемы.

◇◇◇ Тот факт, что «опасная» система может получиться в результате объединения двух «безопасных» систем, является поводом для серьезных размышлений.

6. ИНЖЕНЕРНАЯ ОПТИМИЗАЦИЯ

Многокомпонентные системы часто используются в конструкциях моста или здания. Например, несжимаемая балка и пологая арка могут быть объединены в конструкции моста (рис. 11.15). Если мост должен выдерживать максимальную нагрузку F_M, то балка не должна выпучиваться, а арка не должна прощелкивать при $F<F_M$. В действительности в практике конструирования предусматриваются дополнительные меры

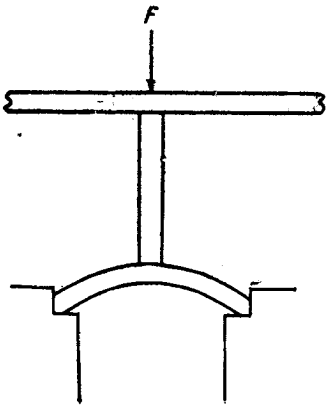


Рис. 11.15. Распространенные в технике конструкции обычно собирают из нескольких «неприводимых» структурных строительных блоков. Моды разрушения составной системы могут быть значительно более резкими, чем разрушающие моды составляющих ее элементов.

предосторожности и используются балки и арки, безопасные при нагрузках, превышающих F_M, скажем, на 50 %. Кроме того, согласно философии оптимального конструирования, целесообразно использовать структурные элементы, разрушающиеся при одной и той же нагрузке. В конце концов, как показывает практика, нет никакого смысла использовать тяжелую балку, которая может выдержать нагрузку $3F_M$ до выпучивания, если арка разрушается уже при $1{,}5F_M$. Кроме того, тяжелые балки дорого стоят, так что и с экономической точки зрения более целесообразно использовать балку, которая выпучивается также при $1{,}5F_M$.

7. НЕЖЕЛАТЕЛЬНЫЕ ПОСЛЕДСТВИЯ ИНЖЕНЕРНОЙ ОПТИМИЗАЦИИ

Инженерные оптимизационные процедуры иногда могут привести к крайне нежелательным (и даже опасным) последствиям [4]. Предположим, например, что составная система конструируется из n «неприводимых» компонент, каждая из которых разрушается вследствие катастрофы сборки $A_{\pm 3}$. Тогда потенциальная функция, описывающая совершенную систему, имеет вид

$$V(x_1, x_2, \ldots, x_n; F) = \sum_{i=1}^{n} \frac{1}{2}(F - F_i)\, x_i^2 + \sum_{i=1}^{n} \pm \frac{1}{4} x_i^4. \quad (11.48)$$

Если нижняя критическая нагрузка достигается при $F = F_1$ и этой нагрузке отвечает лишь одна разрушающая мода, то будет иметь место катастрофа типа $A_{\pm 3}$, а чувствительность к несовершенству будет такой, как описано выше.

В оптимизируемой системе $F_1 = F_2 = \ldots = F_n = (=F_p)$. При таком идеальном критическом нагружении все элементы разрушаются одновременно. Росток потенциальной функции имеет вид $\pm x_1^4 \pm x_2^4 \pm \ldots \pm x_n^4$. Универсальной деформацией этого ростка является

$$P(x) = \sum P_{i_1, i_2, \ldots, i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}, \quad (11.49)$$

где показатели i_j не выше второй степени. В этом случае несовершенная система $V_p(x;\ p) + P(x)$ может иметь до 3^n критических точек при разных значениях нагрузки F и параметрах возмущения $P_{i_1, i_2, \ldots, i_n}$. Даже если каждая отдельная компонента имеет только одну устойчивую критическую точку при $F < F_p$, то составная система может иметь множество критических точек при $F < F_p$ при условии, что отдельные компоненты сильно взаимосвязаны. Жесткая чувствительность к несовер-

шенству составной системы обусловлена нарушающими симметрию дефектами, связывающими локально устойчивую ветвь в точке $(x_1, \ldots, x_n) = (0, \ldots, 0)$ с неустойчивой ветвью, которая может существовать вблизи локально устойчивой ветви при $F < F_p$ вследствие наличия сильной связи между модами. ◇◇◇ Катастрофы с ростками $\pm x_1^4 \pm \ldots \pm x_n^4$ называют *катастрофами кратной сборки*, или *катастрофами n-кратной сборки*. Катастрофы кратной сборки не являются элементарными.

8. ПОДДЕРЖИВАЮЩИЕ КРОНШТЕЙНЫ

8.1. C_{4v}

Поддерживающий кронштейн, изображенный на рис. 11.16, обладает вышеописанным типом чувствительности к несовершенству составных систем. Если пружины вдоль осей x и y

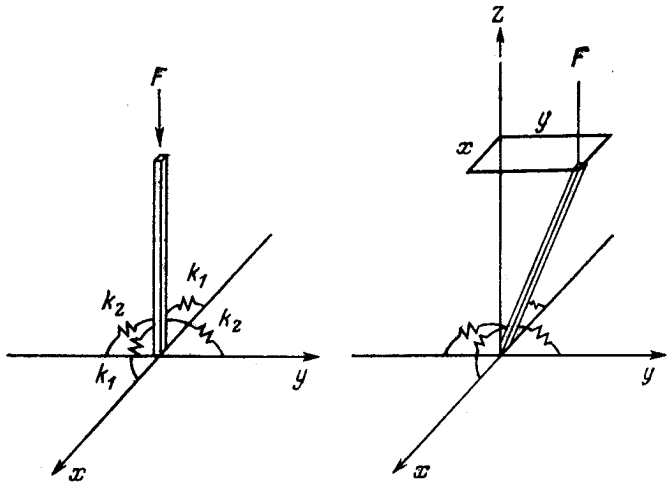


Рис. 11.16. Потенциальная функция, описывающая этот совершенный поддерживающий кронштейн, когда все пружины имеют одну и ту же константу упругости k, обладает группой симметрий C_{4v}.

имеют коэффициенты жесткости k_1 и k_2, то потенциальная функция, описывающая совершенную систему, имеет симметрию

$$V_p(\pm x, \pm y; F) = V_p(x, y; F). \qquad (11.50)$$

Разложение V_p в ряд Тейлора может содержать лишь четные степени x и y.

Если $k_1 = k_2$, то функция V_p инвариантна относительно большой группы симметрий. Такой группой является C_{4v}, т. е. группа

вращений на $\pi/2$ вокруг оси z и отражений в четырех плоскостях, содержащих ось z и прямые $x=0$, $x-y=0$, $y=0$, $x+y=0$. Результаты влияния восьми преобразований группы на координаты (x, y) любой точки из плоскости $x-y$ представлены в следующей таблице:

	E	C_4	C_4^2	C_4^3	σ_{v1}	σ_{v2}	σ_{d1}	σ_{d2}
x	x	y	$-x$	$-y$	x	$-x$	y	$-y$
y	y	$-x$	$-y$	x	$-y$	y	x	$-x$

Инвариантность функции V_p относительно группы преобразований C_{4v} означает, в частности, что коэффициенты при членах $x^p y^q$ разложения в ряд Тейлора в окрестности $(x, y)=(0, 0)$ должны равняться нулю, если p или q нечетно, а коэффициент при $x^p y^q$ должен равняться коэффициенту при $x^q y^p$.

Это утверждение может быть значительно усилено. Разложение в ряд Тейлора V_p может содержать лишь члены вида (x^2+y^2) и x^2y^2, а также произведение этих функций:

$$V_p(x, y; F)=f(x^2+y^2, x^2y^2; F). \tag{11.51}$$

Докажем справедливость подобного утверждения, используя методы теории групп.

Разложение V_p в ряд Тейлора содержит линейные члены по x и y, квадратичные члены, а также члены степени 3, 4, ... Линейное векторное пространство членов степени d имеет размерность $d+1$. В качестве базисных векторов этого пространства можно взять одночлены x^d, x^{d-1}, y, ..., y^d. Каждое преобразование группы C_{4v} отображает член степени d в член степени d в силу однородности. Например, в результате преобразования C_4 имеем

$$\begin{array}{cccc} & x^2 & xy & y^2 \\ C^4 & \downarrow & \downarrow & \downarrow \\ & y^2 & (y)(-x) & (-x)^2 \end{array} \tag{11.52}$$

или в матричной форме

$$\begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}' = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}. \tag{11.53}$$

Матрица размером 3×3, описывающая действие группы преобразований C_4 на трехмерное линейное векторное пространство, порожденное одночленами степени два, называется *матричным представлением* C_4. Аналогично могут быть вычислены матричные представления других групп преобразований в этом

пространстве и во всех других пространствах, порождаемых одночленами различных степеней.

Для рассматриваемой цели не так важны сами матрицы, как их следы (характеры). Эти характеры могут быть использованы для определения числа линейно независимых полиномиальных инвариантов, которые существуют в некотором $(d+1)$-мерном подпространстве членов степени d. Это число равно

$$\frac{1}{8}\sum_{g\in C_{4v}}\chi^d(g). \tag{11.54}$$

Здесь χ^d — след $(d+1)\times(d+1)$-матричного представления группового элемента g. Характеры для разных значений d приведены в табл. 11.2 (она повторяется очевидным способом за пределами $d=7$). Число независимых инвариантов для любой степени представлено в правой колонке таблицы. Инвариантом степени 0 является постоянный член. Инвариантом степени 2 является $x^2+y^2=I^2$. Два инварианта степени 4 суть $(x^2+y^2)=I_2^2$ и $x^2y^2=I_4$.

Таблица 11.2. **Характеры для представлений группы C_{4v} в $(d+1)$-мерном пространстве, порождаемом одночленами вида x^py^q, $p+q=d$**

d	E	C_2	$2C_4$	$2\sigma_v$	$2\sigma_d$	Число линейно независимых функций степени d, инвариантных относительно группы C_{4v}
0	1	1	1	1	1	1
1	2	−2	0	0	0	0
2	3	3	−1	1	1	1
3	4	−4	0	0	0	0
4	5	5	1	1	1	2
5	6	−6	0	0	0	0
6	7	7	−1	1	1	2
7	8	−8	0	0	0	0

Все инварианты имеют вид

$$I_2^pI_4^q,\quad p\geqslant 0,\quad q\geqslant 0, \tag{11.55}$$

где p и q — неотрицательные целые числа. Для доказательства этого утверждения достаточно показать, что число всех инвариантов степени d в точности равно числу инвариантов степени d, которые могут быть записаны в виде $I_2^pI_4^q$. Из табл. 11.2 сле-

дует, что число инвариантов степени d равно

$$\begin{cases} 0, & \text{если } d \text{ нечетно,} \\ \left[\frac{n}{4}\right]+1, & \text{если } d \text{ четно,} \end{cases} \tag{11.56}$$

где $[x]$ означает «наибольшее целое, не превышающее x». Вместе с тем число неотрицательных целых решений (p, q) для уравнения

$$2p+4q=d, \quad p\geqslant 0, \quad q\geqslant 0, \quad p, q \text{ — целые,} \tag{11.57}$$

также дается формулой (11.56). Следовательно, I_2 и I_4 являются единственными функционально независимыми инвариантами относительно группы преобразований C_{4v}, и

$$V_p(x, y; F)=\sum_{p,\ q\geqslant 0} A_{p,q}(F) I_2^p I_4^q = f(I_2, I_4; F), \tag{11.58}$$

как и утверждалось. (Предполагалось, что функции V_p, f — вещественные и аналитические.)

Точный вид потенциальной функции, описывающей совершенный поддерживающий кронштейн, который изображен на рис. 11.16, можно найти посредством прямых вычислений. Предположим, что точка (x, y) определяет положение конца кронштейна. Для простоты возьмем длину кронштейна равной единице. Тогда

$$V_p(x, y; F)=\frac{2}{2} k\theta_1^2+\frac{2}{2} k\theta_2^2+Fz. \tag{11.59}$$

(В этом выражении вес кронштейна не учитывается.) Здесь θ_1 и θ_2 — углы отклонения от направлений x и y, а $z=\sqrt{1-x^2-y^2}$. Углы θ_i могут быть выражены через x и y:

$$\theta_1=\arcsin x=x+\frac{x^3}{6}+\ldots . \tag{11.60}$$

С точностью до членов четвертого порядка разложение V_p имеет вид

$$V_p(x, y; F)=F+\left(k-\frac{1}{2}F\right)(x^2+y^2)+\left(\frac{k}{3}-\frac{F}{8}\right)(x^4+y^4)- \\ -\frac{1}{4}Fx^2y^2. \tag{11.61}$$

Нагрузка становится критической при $F=2k$. При действии критической нагрузки коэффициенты при x^4 и y^4 равны $k/12$, а коэффициент при x^2y^2 равен $-k/2$. Поскольку последний отрицателен, система глобально неустойчива.

Неустойчивые разрушающие моды встречаются в направлении $x=\pm y$. В этих направлениях потенциал сводится к сле-

дующему $(\pm x = t,\ \pm y = t)$:

$$V_p(t;\ F) = F + (2k - F)\,t^2 + \left(\frac{2k}{3} - \frac{F}{2}\right)t^4. \qquad (11.62)$$

В точке сборки коэффициент при t^4 отрицателен. Следовательно, сильное сцепление двух мод, каждая из которых является соответствующей катастрофе A_{+3} модой выпучивания и поэтому относительно безопасна, может привести к разрушающей моде, соответствующей катастрофе A_{-3}, с ее жесткой чувствительностью к несовершенству. По существу чувствительность к несовершенству в случае возмущений типа ε_1 $(x \pm y)$ имеет вид

$$F_c = F_p - k' \mid \varepsilon_1 \mid^{2/3}, \qquad (11.63)$$

в то время как чувствительность к динамическому нагружению демонстрирует каноническую степенную зависимость (с показателем степени 1/2)

$$F_c = F_p - k'' (\Delta E)^{1/2}, \qquad (11.63')$$

где $F = 2k$. Комбинированная чувствительность к статическим $(\varepsilon_1 \neq 0)$ и динамическим $(\Delta E \neq 0)$ несовершенствам описывается поверхностью, изображенной на рис. 11.9.

◇◇◇ Если $k_1 \ll k_2$, то единственная возможная разрушающая мода соответствует катастрофе A_{+3} в направлении оси x. Если $k_1 \gg k_2$, то может произойти лишь катастрофа A_{+3} в направлении оси y. В любом случае изучение разрушающих мод и параметров несовершенства сводится к одномерным задачам. Когда k_1 и k_2 становятся равными, как того требует философия инженерной оптимизации конструкций, для изучения разрушающих мод необходимо привлечь дополнительно вторую степень свободы. Вместо одной сборки получаем две; вместо трех фазовых траекторий изменения состояния после выпучивания — 3^2. Более того, при $F < F_p$ их тоже больше одной. Они представляются седлами, отделяющими локально устойчивое состояние $(x,\ y) = (0,\ 0)$ от области глобальной неустойчивости, которая существует при больших r^2.

8.2. C_{3v}

Поддерживающий кронштейн, изображенный на рис. 11.17, отличается от кронштейна, показанного на рис. 11.16, симметричным расположением пружин. Совершенная система инвариантна относительно вращений на $2\pi/3$ радиан, а также отображений в каждой из трех плоскостей, содержащих одну из пружин. Эти шесть преобразований $(E,\ C_3,\ C_3^2;\ \sigma_1,\ \sigma_2,\ \sigma_3)$ образуют группу, называемую C_{3v}. Потенциальная функция

$V_p(x, y; F)$, описывающая совершенную систему, должна быть инвариантной относительно преобразований группы C_{3v}.

Полиномиальные функции, которые являются инвариантными относительно C_{3v}, могут быть вычислены при помощи методов,

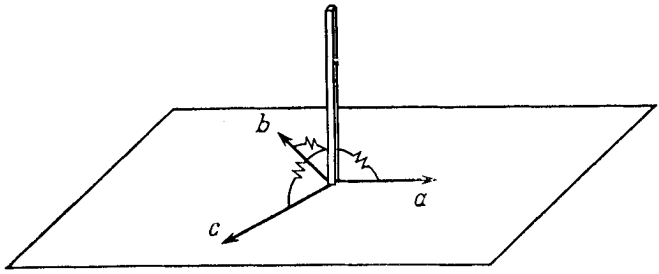


Рис. 11.17. Потенциальная функция, описывающая этот совершенный поддерживающий кронштейн, обладает группой симметрий C_{3v}.

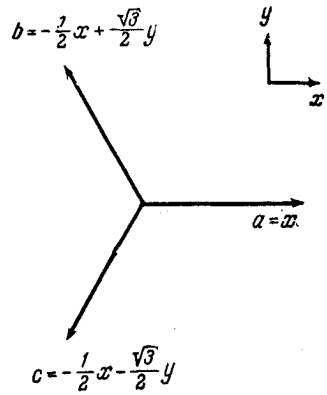


Рис. 11.18. Матрица представлений для операторов C_{3v} легко может быть построена, если распространить действие элементов группы на сверхполную систему векторов a, b, c.

использованных в предыдущем разделе. Сначала введем три вектора (рис. 11.18)

$$a = x, \quad b = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y, \quad c = -\frac{1}{2}x - \frac{\sqrt{3}}{2}y. \tag{11.64}$$

Очевидно, что

$$a + b + c = 0. \tag{11.64'}$$

Эти векторы линейно зависимы и порождают плоскость переменных состояния $\mathbb{R}^2$. Любые два из них могут быть взяты в качестве базисных векторов. Действие элементов группы C_{3v} на эти векторы описывается следующим образом:

	E	C_3	C_3^2	σ_1	σ_2	σ_3
a	a	b	c	a	c	b
b	b	c	a	c	b	a
c	c	a	b	b	a	c

Действие любого элемента группы на $d+1$ линейно независимых членов степени d может быть определено при помощи выбора a и b в качестве базисных векторов пространства $\mathbb{R}^2$. Тогда базисные векторы в пространстве членов степени d имеют вид $a^p b^q$, $p+q=d$. Например, при действии преобразования C_3 имеем

$$\begin{aligned} &C_3\colon\ a^4 \to b^4,\quad a^3b \to b^3(-a-b),\\ &a^2b^2 \to b^2(-a-b)^2,\quad ab^3 \to b(-a-b)^3,\\ &b^4 \to (-a-b)^4. \end{aligned} \tag{11.65}$$

Тогда матрица размером 5×5, представляющая C_3, определяется из

$$\begin{bmatrix} a^4 \\ a^3b \\ a^2b^2 \\ ab^3 \\ b^4 \end{bmatrix}' = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & -1 & -3 & -3 & -1 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \begin{bmatrix} a^4 \\ a^3b \\ a^2b^2 \\ ab^3 \\ b^4 \end{bmatrix}. \tag{11.66}$$

След этой представляющей матрицы равен -1.

Аналогично могут быть вычислены представляющие матрицы всех преобразований группы в любом $(d+1)$-мерном пространстве однородных многочленов степени d. Характеры этих представляющих матриц приведены в табл. 11.3, которая повторяется очевидным образом, когда d возрастает. Правый столбец таблицы содержит число линейно независимых инвариантных многочленов степени d. Это число определяется с помощью стандартного анализа характеров и равно

$$\frac{1}{6}\sum_{g\in C_{3v}} \chi^d(g). \tag{11.67}$$

Один инвариант степени d ясно виден:

$$I_d = a^d + b^d + c^d. \tag{11.68}$$

Тогда $I_0 \simeq 1$, $I_1 = 0$ (в силу 11.64′), $I_2 \simeq x^2+y^2$, $I_3 \simeq abc \simeq x^3 - 3xy^2$.

Теперь докажем, что I_2 и I_3 являются единственными функционально независимыми инвариантами.

Таблица 11.3. **Характеры для представлений группы C_{3v} в $(d+1)$-мерном пространстве, порождаемом одночленами вида $x^p y^q$, $p+q=d$ степени d**

d	E	$2C_3$	3σ	Число линейно независимых функций степени d, инвариантных относительно группы C_{3v}
0	1	1	1	1
1	2	−1	0	0
2	3	0	1	1
3	4	1	0	1
4	5	−1	1	1
5	6	0	0	1
6	7	1	1	2
7	8	−1	0	1
8	9	0	1	2
9	10	1	0	2
10	11	−1	1	2
11	12	0	0	2

Заметим, что

1. На основании анализа характеров число инвариантов степени d равно

$$\begin{cases} \left[\frac{d}{6}\right]+1, & d \bmod 6 \neq 1, \\ \left[\frac{d}{6}\right], & d \bmod 6 = 1. \end{cases} \tag{11.69}$$

2. Число неотрицательных целых решений уравнения

$$2p+3q=d, \quad p \geqslant 0, \quad q \geqslant 0, \quad p,\ q \text{ — целые},$$

равно $[d/6]+1$, если только не выполняется условие $d \pmod 6 = 1$, в противном случае оно равно $[d/6]$.

Поскольку любой инвариант степени d может быть записан в виде $I_2^p I_3^q$, то I_2 и I_3 — единственные функционально независимые инварианты. Следовательно,

$$V_p(x,\ y;\ F)=f(I_2,\ I_3;\ F). \tag{11.70}$$

Единственный управляющий параметр F может быть использован для обращения в нуль одного коэффициента ряда Тейлора, так что потенциальная функция, описывающая совершенную систему с точностью до нижнего нетривиального порядка, равна

$$V_p(x,\ y;\ F)=\frac{1}{2}(F_p-F)(x^2+y^2)+(x^3-3xy^2). \tag{11.71}$$

Наиболее общая деформация ростка x^3-3xy трехмерна и имеет базисные векторы x, y, x^2+y^2. Деформация, включающая

$x^2 + y^2$, будет давать мягкую чувствительность первого порядка к несовершенству. Следовательно, для того чтобы определить чувствительность к несовершенству поддерживающего кронштейна, вполне достаточно рассмотреть параметр несовершенства вида $\varepsilon_1 x + \varepsilon_2 y$.

Возмущенная потенциальная функция имеет вид

$$V(x, y; F, \varepsilon_1, \varepsilon_2) = x^3 - 3xy^2 + \frac{1}{2}(F_p - F)(x^2 + y^2) + \varepsilon_1 x + \varepsilon_2 y. \tag{11.72}$$

Критические точки определяются из системы уравнений

$$\frac{\partial V}{\partial x} = 3x^2 - 3y^2 + (F_p - F)x + \varepsilon_1 = 0, \tag{11.73а}$$

$$\frac{\partial V}{\partial y} = -6xy + (F_p - F)\,y + \varepsilon_2 = 0. \tag{11.73б}$$

Матрица устойчивости имеет вид

$$V_{ij} = \begin{bmatrix} 6x + (F_p - F) & -6y \\ -6y & -6x + (F_p - F) \end{bmatrix}. \tag{11.74}$$

На бифуркационном множестве $\det V_{ij} = 0$. Следовательно, бифуркационное множество определяется посредством трех уравнений (11.73а), (11.73б) и

$$(6x)^2 + (6y)^2 = (F_p - F)^2. \tag{11.73в}$$

Параметрическое представление рассматриваемого бифуркационного множества удобно дать через x и y:

$$\begin{aligned} \varepsilon_1 &= -3(x^2 - y^2) \pm 6x\sqrt{x^2 + y^2}, \\ \varepsilon_2 &= 6xy \pm 6y\sqrt{x^2 + y^2}, \\ F_p - F &= \pm 6\sqrt{x^2 + y^2}. \end{aligned} \tag{11.75}$$

Это бифуркационное множество показано на рис. 11.19, *а*. В действительности физически интересна только его наименьшая F-компонента, поскольку система разрушается на этом множестве. Такая компонента показана на рис. 11.19, *б* [5].

Вдоль любой прямой, проходящей через начало координат, имеем $y = tx$. Параметрическое представление бифуркационного множества в этом случае имеет вид

$$\begin{aligned} \varepsilon_1 &= \{-3(1 - t^2) \pm 6\sqrt{1 + t^2}\}\,x^2, \\ \varepsilon_2 &= \{6t \pm 6t\sqrt{1 + t^2}\}\,x^2, \\ F_p - F &= \pm 6\sqrt{1 + t^2}\;\; x. \end{aligned} \tag{11.76}$$

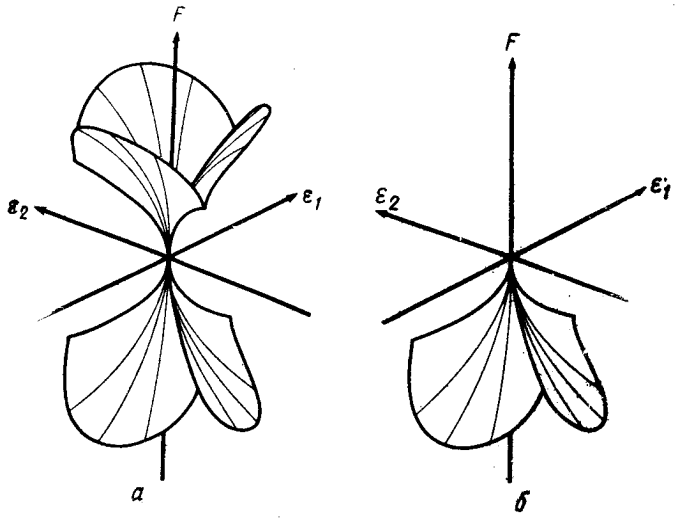


Рис. 11.19.

а — бифуркационное множество катастрофы D_{-4} в пространстве управляющих параметров $F - \varepsilon_1 - \varepsilon_2$; *б* — при возрастании F до встречи с бифуркационным множеством происходит разрушение. Поверхность разрушения является нижним листом бифуракционной поверхности [5].

Исключая определенные значения t, видим, что имеет место степенная (с показателем степени 1/2) зависимость разрушающей нагрузки от параметров несовершенства ε_1, ε_2.

◇◇◇ Степенная зависимость чувствительности к несовершенству (с показателем степени 1/2) не является такой уж неожиданной. Росток $x^3 - 3xy^2$ подобен ростку $x^3 + y^3$. Он может быть назван катастрофой «двойной складки» по аналогии с катастрофой «двойной сборки» $\pm x^4 \pm y^4$. Чувствительность к несовершенству для двойной сборки имеет ту же степенную зависимость, что и обычная сборка: показатель степени равен 2/3 для статического несовершенства и 1/2 для динамического. Аналогично чувствительность к несовершенству для двойной складки имеет ту же степенную зависимость, что и обыкновенная складка: показатель степени равен 1/2 для статического несовершенства и 1/3 для динамического несовершенства.

9. СМЯГЧЕНИЕ МОДЫ

Когда конструкция находится вблизи «точки разрушения» типа, описанного в данной главе, частота одной из разрушающих мод становится очень малой. Для того чтобы показать это, рассмотрим энергию системы, совершающей малые колебания

вблизи морсовской критической точки:

$$\xi = \frac{1}{2} M_{ij}\dot{x}_i\dot{x}_j + \frac{1}{2} V_{ij}(F)\, x_i x_j. \tag{11.77}$$

В этом случае линеаризованным уравнением движения будет

$$M_{ij}\ddot{x}_j + V_{ij}x_j = 0. \tag{11.78}$$

Соответствующее уравнение для собственных значений имеет вид

$$\det | V_{ij} - \omega^2 M_{ij} | = 0. \tag{11.79}$$

На бифуркационном множестве $\det V_{ij} = 0$, так что по крайней мере одно собственное значение $\omega = 0$ на бифуркационном мно-

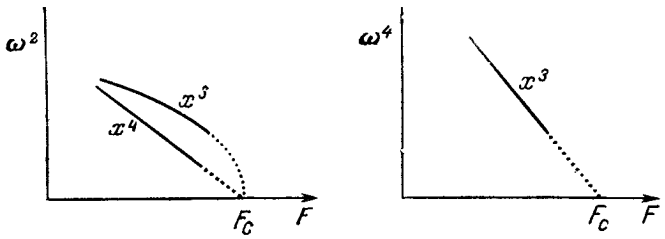


Рис. 11.20. Частота колебаний разрушающей моды убывает при приближении к критической нагрузке.

На основе зависимости ω^p от F можно не только приближенно оценить разрушающую нагрузку (при помощи экстраполяции), но определить также тип катастрофы (посредством p).

жестве, и в силу соображений непрерывности оно мало вблизи бифуркационного множества. Это уменьшение частоты колебаний нормальной моды при приближении к критической точке называют *смягчением моды.*

Смягчение моды может быть использовано для получения количественной характеристики как типа происходящей катастрофы, так и критической нагрузки. Например, если $V \sim x^3$, когда $F = F_c$, то

$$\omega^2 \sim (F_c - F)^{1/2} \tag{11.80}$$

при условии, что $F \uparrow F_c$. Вместе с тем если $V \sim x^4$, когда $F = F_c$, то $\omega^2 \sim (F_c - F)$ при условии, что $F \uparrow F_c$. На рис. 11.20 показано, каким образом может быть использована зависимость ω от F для определения критической точки системы.

◇◇◇ При приближении к бифуркационному множеству потенциальная функция становится более плоской и, кроме того, значительно отклоняется от формы функции, описывающей гармо-

нический осциллятор. Первое обусловлено смягчением моды; второе — присутствием нелинейных колебаний, которые содержат примесь более высоких гармоник. Амплитуды таких гармоник возрастают при $F \uparrow F_c$. Подобные нелинейные колебания могут быть описаны с помощью эллиптических функций и могут быть использованы в качестве еще одного способа диагностики катастроф.

10. ВЫВОДЫ

Многие конструкции могут быть описаны с помощью потенциальной функции, зависящей от нагрузки F, приложенной к конструкции. Состояние конструкции характеризуется критическими точками потенциальной функции. При этом устойчивость состояния определяется с помощью морсовской характеристики потенциала в критической точке, а критическая нагрузка, которую может выдержать конструкция, — вырождением критических точек. Короче говоря, изучение поведения статических конструкций под нагрузкой и их чувствительности к несовершенству тесно связано с теорией катастроф. Большая часть этой главы является демонстрацией этих наблюдений.

Для того чтобы найти вид потенциальной функции, описывающей идеальную, или совершенную, систему, вводятся подходящие переменные состояния. Тогда критическая нагрузка для совершенной системы определяется путем рассмотрения значений F, при которых матрица устойчивости становится сингулярной. Росток потенциала в этой неморсовской критической точке может быть использован для установления вида наиболее общих деформаций потенциальной функции, описывающей идеальную систему. Чтобы наложить существенные ограничения на вид общих деформаций, при рассмотрении принимались во внимание соображения физического характера. Это в свою очередь позволяло найти универсальную деформацию, которая может быть использована для определения чувствительности критической нагрузки к несовершенствам различных возможных типов.

Описанные методы были обобщены для изучения снижения несущей способности к нагрузке, обусловленной колебаниями конструкции. Снижение критической нагрузки определяется приравниванием критического значения разности ΔV между метастабильным локальным минимумом и нижним смежным морсовским 1-седлом величине ΔE. Установлено, что чувствительность критической нагрузки к несовершенству и колебаниям описывается степенной зависимостью $F_c = F_0 - k|\varepsilon|^p$ или $F_c = F_0 - k(\Delta E)^p$, где p принимает следующие значения:

	$\lvert\varepsilon\rvert^p$	ΔE^p	
A_{+3}	1	1	Разд. 2
A_{-3}	$^2/_3$	$^1/_2$	Разд. 3, 8
A_2	$^1/_2$	$^1/_3$	Разд. 4, 9

Для изученных систем жесткая чувствительность к несовершенству наблюдается наряду с даже более жесткой динамической чувствительностью.

Показано, что инженерная оптимизация может привести в предположении сильного сцепления элементов конструкции к неожиданным разрушающим модам с жесткими чувствительностью к несовершенству и динамической чувствительностью.

Методы теории групп могут быть использованы для привлечения теории катастроф к анализу составной системы, обладающей симметрией некоторого типа, а смягчение моды — для определения положения критической нагрузки, перед тем как произойдет разрушение.

◇◇◇ По причинам, связанным с соображениями физического характера, замены переменных, приводящие к смешиванию переменных состояния и управляющих параметров, могут оказаться нежелательными. Если же такие замены допустимы, то чувствительность к несовершенству системы, находящейся в неморсовской критической точке, может быть определена детальным анализом с использованием методов, изложенных в гл. 3—5. Подобный анализ был проведен Голубицким и Шеффером [6].

Литература

1. Thompson J. M. T., Hunt G. W., A General Theory of Elastic Stability, New York: Wiley, 1973.
2. Euler L., Methodes Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes (Appendix, De Curvis Elasticis), Lausanne and Geneva: Marcum Michaelum Bousquet, 1744.
3. Roorda J., Stability of Structures with Small Imperfections, *J. Eng. Mech. Div. Am. Soc. Civil Eng.*, **91**, 87 (1965).
4. Thompson J. M. T., Hunt G. W., Dangers of Structural Optimization, *Eng. Optimization*, **1**, 99—110 (1974).
5. Thompson J. M. T., Hunt G. W., Towards a Unified Bifurcation Theory, *J.* Appl Math. Phys., **26**, 581—603 (1975).
6. Golubitsky M., Schaeffer D., A Theory for Imperfect Bifurcation via Singularity Theory, *Commun. Pure Appl. Math.*, **32**, 21—98 (1979).

12

АЭРОДИНАМИКА

Высокоэффективный летательный аппарат чем-то напоминает породистую скаковую лошадь: оба гладкие, быстрые, стремительные и временами их поведение непредсказуемо. Эта «непредсказуемость» может быть иногда уподоблена элементарным катастрофам, и тогда непредсказуемое становится предсказуемым.

Для летательных аппаратов определенного класса нелинейные уравнения движения оказываются настолько простыми, что возможен их качественный анализ. Такой анализ позволяет провести идентификацию многообразия стационарных состояний с элементарной катастрофой [1]. Как только катастрофа идентифицирована и ее бифуркационное множество найдено, качественная динамика летательного аппарата по существу определена.

Предсказания качественного характера относительно поведения системы, основанные на идентификации многообразия уравнений состояния с некоторой катастрофой A_{+5}, можно сравнить с подробными численными решениями динамических уравнений движения. Статика теории катастроф и динамика численного интегрирования являются не только сходными, но и дополняют друг друга, причем численное интегрирование может быть использовано для нахождения устойчивых компонент многообразия стационарных состояний. При этом подробные результаты численного интегрирования могут быть интерпретированы на основе фундаментальных представлений, сформулированных в теории катастроф.

1. ОПИСАНИЕ СОСТОЯНИЯ ДИНАМИЧЕСКОЙ СИСТЕМЫ

Аэродинамика — это инженерная наука, для которой предположения о статическом поведении, пригодные для других инженерных приложений, оказываются явно неприемлемыми, так как самолет в полете не является статическим объектом. Виражи, развороты, пике, вращения и горки — все это внутренние элементы динамики, которые, как правило, взаимосвязаны между собой посредством сложных нелинейных механизмов.

Уравнения, описывающие движение летательного аппарата, сложны, нелинейны и образуют систему уравнений первого порядка. Прежде чем приступить к подробному изучению уравнений движения какого-то конкретного летательного аппарата, целесообразно сначала остановиться на общих свойствах систем нелинейных обыкновенных дифференциальных уравнений первого порядка. Семейство таких уравнений имеет вид

$$\frac{dx_i}{dt} = F_i(x; c), \quad x \in \mathbb{R}^n, \quad c \in \mathbb{R}^k. \tag{12.1}$$

Как обычно, будем рассматривать n-мерный x как систему переменных, характеризующих состояние, а k-мерный c — как систему управляющих параметров.

Одна из задач аэродинамики состоит в определении зависимости решения $x(t) \in \mathbb{R}^n$ от управляющих параметров $c \in \mathbb{R}^k$. Предметом же изучения теории катастроф является, вообще говоря, определение числа, типов и свойств устойчивости решений системы (12.1), нахождение бифуркационного множества $\mathcal{P}_B \subset \mathbb{R}^k$, в котором эти числа и типы решений изменяются, а также исследование свойств фазовых переходов, т. е. переходов системы из одного стационарного состояния в другое.

Существует несколько подходов к решению этой проблемы. Если имеется доступ к большой ЭВМ с недорогим машинным временем и эффективным алгоритмом Рунге — Кутта для интегрирования уравнений вида (12.1), то необходимо лишь запрограммировать эти уравнения и предоставить машине сделать остальное. Хотя такой подход чрезвычайно экономичен по времени (после отладки программы), может дать значительную информацию и позволяет «почувствовать», каким образом решения ведут себя как функции управляющих параметров, он не может обеспечить фундаментального изучения свойств решений. Кроме того, асимптотическое стационарное состояние (если оно существует) может зависеть от начальных условий и (или) управляющих параметров разрывным образом (рис. 12.1). К тому же численный подход позволяет определить положения сепаратрис, отвечающих разрывам, только путем детального численного отображения множеств начальных условий и управляющих параметров в произведении пространств $\mathbb{R}^n \otimes \mathbb{R}^k$, что оказывается довольно дорогим удовольствием даже для «свободного» компьютера.

Другим возможным подходом к изучению поведения катастроф динамической системы (12.1) является предварительное исследование ее стационарных решений $dx_i/dt = \dot{x}_i = 0$. (Мы особо подчеркиваем различие между стационарными и равновесными условиями. Летательный аппарат может находиться в стационарных условиях, совершая полет на высоте 13 000 м. Летательный аппарат находится в равновесии, когда он покоится на земле.) Этот подход дополняет ранее рассмотренный, но и он не лишен недостатков. Например, система (12.1) может вообще не иметь стационарных (устойчивых) решений, не говоря уже о том, что решение системы n уравнений $F_i(x; c) = 0$ может оказаться совсем не простой задачей.

В ряде ситуаций, представляющих особый интерес, обе эти трудности могут отсутствовать. В одном полезном частном случае все, кроме одной, переменные состояния $x = x_1, x_2, \ldots, x_n$ могут быть исключены из уравнений $F_i = 0$. В результате имеем

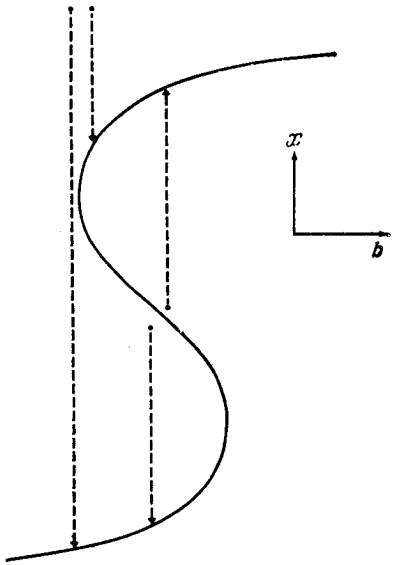


Рис. 12.1. Конечное значение переменной состояния x может разрывным образом зависеть от своего начального значения или от значения управляющего параметра b.

единственное уравнение вида

$$\Phi'(x;\ c)=0,\quad x\in\mathbb{R}^1,\quad c\in\mathbb{R}^k. \tag{12.2}$$

Тогда можно положить

$$\Phi(x;\ c)=-\int^{x}\Phi'(y;\ c)\,dy,\quad \Phi'(x;\ c)=-\frac{d}{dx}\Phi(x;\ c). \tag{12.3}$$

В этом случае стационарные состояния (12.1) могут быть идентифицированы с критическими точками функции $\Phi(x;\ c)$, называемой *функцией Ляпунова*. Эта функция для стационарных систем играет ту же роль, что и потенциальная функция для равновесных систем.

Если для динамической системы можно ввести функцию Ляпунова, то бифуркационные свойства и катастрофы динамической системы могут быть частично связаны с катастрофами $\Phi(x;\ c)$. В этом случае катастрофы, которые могут произойти (для Φ), относятся к катастрофам типа A_k. Такая идентификация стационарных состояний динамической системы с элементарными катастрофами дает информацию об адекватности полиномиальной аппроксимации $\Phi(x;\ c)$. И наоборот, если Φ есть полином степени n, то для реализации любого желаемого возмущения стационарного состояния системы требуется

в точности $n-2$ управляющих параметра. Если $k > n-2$, то число управляющих параметров окажется избыточным, а если же $k < n-2$, то недостаточным.

Значительный успех в понимании свойств динамической системы может быть достигнут, если реализовать оба вышеописанных подхода — как численный, так и аналитический: аналитический подход дает возможность определить уравнения гиперповерхности стационарных состояний динамической системы в произведении пространств $\mathbb{R}^n \otimes \mathbb{R}^n$, а метод численного интегрирования позволяет установить, каким образом динамическая система реально достигает многообразия стационарных состояний в зависимости от (1) значений управляющих параметров, (2) начальных условий и (3) времени.

2. ОПИСАНИЕ СОСТОЯНИЯ УСЕЧЕННОЙ ДИНАМИЧЕСКОЙ СИСТЕМЫ

Иногда удобно изучать динамическую систему с помощью разложения силовых функций $F_i(x; c)$ в ряд Тейлора в окрестности стационарного решения. Если предположить, что координаты переменной состояния выбраны таким образом, что $x = 0$ для $c = 0$, то разложение в ряд Тейлора правой части уравнения (12.1) относительно выбранного состояния дает

$$\dot{x}_i = F_i + F_i^j x_j + F_i^{jk} x_j x_k + \dots . \tag{12.4}$$

Здесь верхние индексы означают дифференцирование по соответствующей координате, например $F_i^{jk} = \partial^2 F_i / \partial x_j \partial x_k$. Все производные вычисляются в точке $(x; c) = (0; c)$. Для фиксированного значения управляющих параметров c матрицу F_i^j можно рассматривать как функцию линейного отклика на изменение переменных состояния, которая описывает изменение $\dot{x}_i$, соответствующее изменению x_j. В более общем случае F_i^{jk} не что иное, как компоненты тензора восприимчивости.

Коэффициенты $F_i(c)$, $F_i^j(c)$ могут быть разложены в ряды Тейлора по управляющим параметрам c:

$$\begin{aligned} F_i(c) &= F_i(0) + F_i^\alpha c_\alpha + F_i^{\alpha\beta} c_\alpha c_\beta + \dots , \\ F_i^j(c) &= F_i^j(0) + F_i^{j\alpha} c_\alpha + \dots . \end{aligned} \tag{12.5}$$

Коэффициенты этих разложений являются частными производными функций F_i, вычисленными в точке $(x; c) = (0; 0) \in \in \mathbb{R}^n \otimes \mathbb{R}^k$; коэффициенты F_i можно трактовать как функции линейного отклика на изменение управляющих параметров, коэффициенты $F_i^j(0, 0)$ и $F_i^\alpha(0, 0)$ рядов Тейлора в окрестности стационарного состояния $(0, 0)$ можно рассматривать как функ-

ции линейного отклика на изменение переменных состояния и управляющих параметров, а все другие члены разложений оказываются нелинейными по переменным состояния и(или) управляющим параметрам.

Для некоторых динамических систем можно изучать присущее им нелинейное поведение, оставляя в разложении лишь главные нелинейные члены по переменным состояния. Такие нелинейные системы описываются уравнениями (12.4), если пренебречь членами третьей и более высоких степеней. Ряд систем подобного типа обладает дополнительным свойством, а именно существует постоянный невырожденный положительно определенный симметрический метрический тензор G^{li}, такой, что

$$x_l G^{li} F_i^{jk} x_j x_k = \tilde{F}^{ijk} x_i x_j x_k = 0. \tag{12.6}$$

Для таких систем можно определить «энергетическую» функцию посредством следующих вычислений:

$$\begin{aligned} x_l G^{li} \dot{x}_i &= x_l G^{li} F_i + x_l G^{li} F_i^j x_j + x_l G^{li} F_i^{jk} x_j x_k, \\ \frac{1}{2} \frac{d}{dt} (G^{li} x_l x_i) &= \tilde{F} x_i^i + \tilde{F}^{ij} x_i x_j + \tilde{F}^{ijk} x_i x_j x_k. \end{aligned} \tag{12.7}$$

Последний член, согласно (12.6), равен нулю. Произведение $G^{il}F_l^j$, вообще говоря, не есть симметрический тензор и может быть представлено в виде суммы симметрического и антисимметрического тензоров, причем последний свертывается к нулю. Следовательно, достаточно взять $\tilde{F}^{ij}$ как симметрическую часть $G^{il}F_l^j$. Если симметрическая матрица $\tilde{F}^{ij}$ невырождена, а обратная ей матрица есть $\tilde{F}_{ij}$, то выражение (12.7) можно записать в виде

$$\frac{d}{dt} \left(\frac{1}{2} G^{ij} x_i x_j \right) = \tilde{F}^{ij} (x_i + b_i)(x_j + b_j) - \tilde{F}^{ij} b_i b_j, \tag{12.8}$$

где $b_j = \frac{1}{2} \tilde{F}^i \tilde{F}_{ij}$ есть функция только управляющих параметров. Если G^{ij} положительно определена, квадратичная форма

$$E = \frac{1}{2} G^{ij} x_i x_j \tag{12.9}$$

может интерпретироваться как энергия. В этом случае уравнение (12.8) описывает скорость изменения обобщенной энергии E динамической системы. Взятая с обратным знаком обобщенная «энергетическая» функция (12.9) будет являться функцией Ляпунова для динамических систем вышеописанного типа.

◇◇◇ Если силы F_i задаются как градиенты какой-либо функции, то

$$F_i\dot{x}_i \to -\frac{\partial V}{\partial x_i}\frac{dx_i}{dt} = -\frac{dV}{dt}. \tag{12.10}$$

Для градиентной динамической системы функция Ляпунова есть просто потенциальная функция. Для динамических систем в рассматриваемом случае функция Ляпунова $(-E)$ может быть названа *обобщенной потенциальной функцией.*

3. ЛИНЕЙНЫЙ АНАЛИЗ УСТОЙЧИВОСТИ

Как только стационарное решение (или решения), отвечающее $\dot{x}_i = 0$ в уравнении (12.1), будет найдено, необходимо исследовать его (их) свойства устойчивости. Проще всего это сделать с помощью стандартной процедуры линейного анализа устойчивости. Если $x = x^0$ — стационарное решение уравнения (12.1) при $c = c^0$, то уравнение движения в точке $x = x^0 + \delta x$, лежащей в окрестности x^0, есть

$$\delta\dot{x}_i = M_i^j\delta x_i + O(2), \tag{12.11}$$

$$M_i^j = \frac{\partial F_i}{\partial x_j}(x^0;\ c^0). \tag{12.12}$$

Члены второго порядка в выражении для δx могут быть опущены, если матрица устойчивости M размером $n \times n$ не имеет собственных значений с нулевой действительной частью (гл. 19). В пренебрежении членами второго порядка уравнение (12.11) имеет решение

$$\delta x(t) = e^{M(t-t_0)}\delta x(t_0), \tag{12.13}$$

где δx есть вектор-столбец $n \times 1$, а экспонента определяется ее степенным разложением. Свойства устойчивости вблизи x^0 можно определить, приводя M к канонической жордановой форме (гл. 14). В частности, если хотя бы одно из собственных значений матрицы M имеет положительную действительную часть, x^0 является неустойчивым стационарным состоянием. Если все собственные значения M в точке x^0 находятся в левой полуплоскости, возмущенная система будет возвращаться к устойчивому стационарному состоянию в течение интервала времени, определяемого собственным значением матрицы с наименьшей действительной частью:

$$T_1 \simeq \max_{1 \leqslant i \leqslant n} (-Re\lambda_i)^{-1}. \tag{12.14}$$

Если x^0 есть стационарное решение для $c = c^0$, то, слегка изменяя значение управляющего параметра $c \to c^0 + \delta c$, можно

получить небольшое изменение положения стационарного решения. Если все собственные значения $F_i^j(x^0; c^0)$ лежат в левой полуплоскости, то там же лежат и все собственные значения $F_i^j(x^0 + \delta x; c^0 + \delta c)$. Открытое множество в $\mathbb{R}^k$, содержащее c^0, которое локально параметризует устойчивые стационарные решения (12.1), ограничивается бифуркационным множеством $\mathscr{S}_B \subset \mathbb{R}^k$, на котором одно или более собственных значений $F_i^j[x^0(c); c]$ имеют нулевую действительную часть. Если матрица M_i^j (12.11) оказывается симметрической, как в случае градиентных систем, то все ее собственные значения будут действительными и бифуркационное множество F_i^j будет локально эквивалентным бифуркационному множеству некоторой функции катастроф. Однако матрица F_i^j не обязательно должна быть симметрической, она может иметь (и довольно часто имеет) комплексные собственные значения, и поэтому ее бифуркационное множество и катастрофы таят в себе немало неожиданного (гл. 19).

4. УРАВНЕНИЯ ДВИЖЕНИЯ ЛЕТАТЕЛЬНОГО АППАРАТА

Рассмотрим поведение летательного аппарата, для которого потеря устойчивости соответствует одной элементарной катастрофе. Для такого летательного аппарата матрица устойчивости в случае некоторых устойчивых стационарных решений имеет комплексно-сопряженные корни, однако один из ее действительных корней проходит через нуль и является ответственным за потерю устойчивости. Изучаемые ниже уравнения, описывающие движение летательного аппарата, имеют вид (12.4) (без ...). Если $x = x^0$ есть стационарное решение этой системы уравнений, то линеаризованная матрица устойчивости в окрестности этой точки имеет вид

$$M_i^j = F_i^j + 2F_i^{jk}x_k^0. \tag{12.15}$$

Прежде чем приступить к описанию возможных стационарных состояний летательного аппарата, необходимо выбрать подходящую систему переменных состояния. Эта система должна быть достаточно развернутой, чтобы наше описание не было жалкой карикатурой на поведение летательного аппарата, и вместе с тем достаточно компактной, чтобы быть удобной при использовании.

Начнем с выбора системы координат, жестко связанной с корпусом летательного аппарата. Для этого начало координат (или системы отсчета) поместим в центр масс летательного аппарата, а оси координат совместим с осями его главных моментов инерции (рис. 12.2). Ориентация осей рассматриваемого

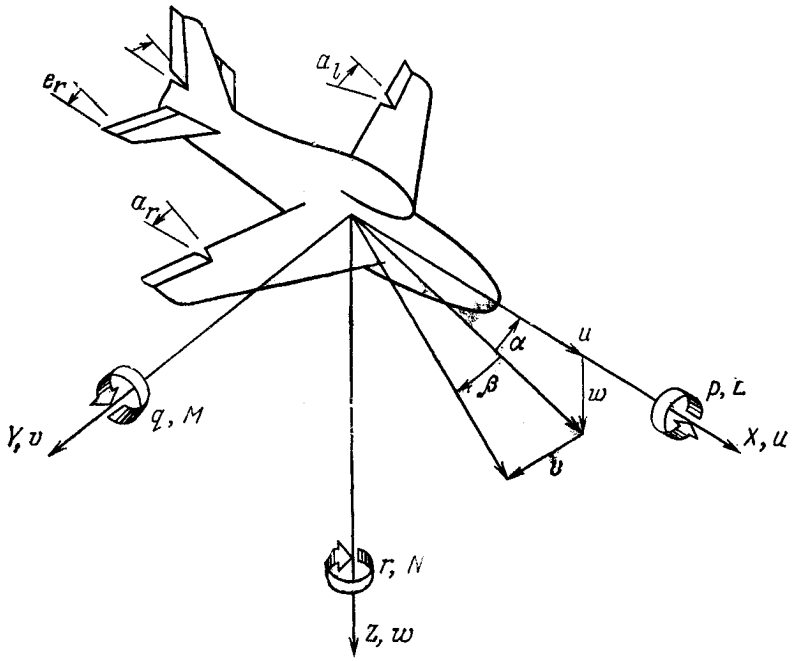


Рис. 12.2. Система координат, жестко связанная с корпусом летательного аппарата, оси которой составляют эйлеровы углы с осями инерционной системы отсчета [1, 2].

X, Y, Z — компоненты вектора силы; L, M, N — компоненты вектора вращающего момента; u, v, w — компоненты вектора скорости относительно инерционной системы отсчета; p, q, r — компоненты вектора угловой скорости по главным осям; α, β — угол атаки, угол скольжения; a_l, a_r — углы отклонения элеронов; e_l, e_r — углы отклонения рулей высоты; τ — угол отклонения руля направления.

тела относительно инерциальной системы отсчета определяется тремя углами Эйлера (ϕ, θ, ψ). Отличные от нуля компоненты тензора инерции I_{ij} по трем главным осям суть (I_{xx}, I_{yy}, I_{zz}). Компоненты вектора угловой скорости $\boldsymbol{\omega}$ по главным осям обозначим через (p, q, r), а компоненты вращающего момента $\mathbf{T}$ — через (L, M, N). Компоненты вектора скорости центра масс $\mathbf{v}$ относительно инерциальной системы отсчета обозначим (u, v, w), а компоненты силы $\mathbf{F}$ — через (X, Y, Z).

Если скорость приблизительно постоянна, то часто удобно вместо декартовой системы координат использовать полярные координаты $|\mathbf{v}|$, α, β, где α — *угол атаки*, β — *угол скольжения*. Для малых углов (в радианной мере) приближенно имеем $\alpha \simeq w/|\mathbf{v}|$, $\beta \simeq v/|\mathbf{v}|$.

Для описания устойчивости летательного аппарата во многих случаях достаточно использовать в качестве переменных состоя-

ния три компоненты угловой скорости (p, q, r), два «полярных» угла (α, β) и три эйлеровых ориентационных угла (ϕ, θ, ψ). Тогда пространство переменных состояния $\mathbb{R}^n$ является 8-мерным, и $(p, q, r, \alpha, \beta, \phi, \theta, \psi) = (x_1, \ldots, x_8) \in \mathbb{R}^8$.

Ориентация летательного аппарата определяется положениями a_l, a_r левого и правого элеронов, положениями e_l, e_r левого и правого рулей высоты и положением τ руля направления. Эти углы измеряются от положений, устанавливаемых для полета по прямой на постоянной высоте, и, как правило, они малы. Перечисленные угловые отклонения являются управляющими воздействиями $(a_l, a_r, e_l, e_r, \tau) = (c_1, c_2, c_4, c_5) \in \mathbb{R}^5$. Для последующего изложения удобно ввести следующие обозначения:

$$\begin{aligned} a &= \tfrac{1}{2}(a_l + a_r), \quad \delta a = a_l - a_r, \\ e &= \tfrac{1}{2}(e_l + e_r), \quad \delta e = e_l - e_r. \end{aligned} \tag{12.16}$$

Теперь определим структуру уравнений движения. Поскольку ньютоновские уравнения движения имеют второй порядок, аэродинамические уравнения движения также должны быть уравнениями второго порядка. Так как три компоненты угловой скорости (p, q, r) уже являются производными первого порядка по времени, то для этих трех переменных состояния получаем уравнения вида

$$\dot{x}_i = F_i(x;\ c), \tag{12.17}$$

где $1 \leqslant i \leqslant 3$.

Два «полярных» угла также выражаются через компоненты скорости w, v, которые сами являются производными первого порядка по времени. Следовательно, эти два угла также описываются уравнениями первого порядка (12.17) с индексами $4 \leqslant i \leqslant 5$. Наконец, производные по времени трех эйлеровых ориентационных углов (ϕ, θ, ψ) связаны следующими линейными преобразованиями с тремя компонентами угловой скорости (p, q, r):

$$\begin{aligned} \dot{\phi} &= p + q \operatorname{tg}\theta \sin\phi + r \operatorname{tg}\theta \cos\phi, \\ \dot{\theta} &= \quad q \cos\phi \quad - r \sin\phi. \end{aligned} \tag{12.18}$$

Это тоже соотношения типа (12.17) с индексами $6 \leqslant i \leqslant 8$, играющие роль определяющих уравнений. Короче говоря, получаем систему уравнений типа (12.1) с восемью переменными и пятью управляющими параметрами.

Уравнения движения для (p, q, r) $(1 \leqslant i \leqslant 3)$ не зависят от ориентационных углов Эйлера (ϕ, θ, ψ), а уравнения для (α, β) $(4 \leqslant i \leqslant 5)$ зависят от углов ϕ, θ из-за наличия слагаемых,

содержащих гравитационное ускорение, деленное на модуль скорости $g/|\mathbf{v}|$. Для реактивных летательных аппаратов величина $\mathbf{v}$ велика и $g/|\mathbf{v}| \to 0$, поэтому указанными слагаемыми часто можно пренебречь. В результате в случае реактивных летательных аппаратов системы первых пяти и последних трех уравнений оказываются расцепленными, и в качестве уравнений движения имеем

$$\dot{x}_i = F_i(x_j;\ c), \quad 1 \leqslant i, j \leqslant 5, \tag{12.19а}$$

$$\dot{x}_k = F_k(x_l;\ c), \quad 6 \leqslant k \leqslant 8, \ 1 \leqslant l \leqslant 8. \tag{12.19б}$$

Займемся решением первой системы уравнений (12.19а).

В случае симметричного реактивного летательного аппарата система уравнений (12.19а) имеет одно очевидное стационарное решение $x = 0 \in \mathbb{R}^5$ при $c = 0 \in \mathbb{R}^5$. Для поиска других возможных решений воспользуемся разложением системы уравнений в ряд Тейлора, как это сделано в (12.4):

$$\dot{x} = LA + LC + NL. \tag{12.20}$$

Здесь LA — линейные аэродинамические, а LC — линейные управляющие параметры

$$LA = F_i^j x_j, \qquad LC = F_i^\alpha c_\alpha. \tag{12.21}$$

Из нелинейных членов NL наиболее важными являются инерционные $IN = F_i^{jk} x_j x_k$, $1 \leqslant i \neq j \neq k \leqslant 3$, где

$$\begin{aligned} F_1^{23} &= F_1^{32} = \frac{(I_{yy} - I_{zz})}{I_{xx}}, \\ F_2^{31} &= F_2^{13} = \frac{(I_{zz} - I_{xx})}{I_{yy}}, \\ F_3^{12} &= F_3^{21} = \frac{(I_{xx} - I_{yy})}{I_{zz}}. \end{aligned} \tag{12.22}$$

Если пренебречь всеми нелинейными членами, за исключением инерционных, то система нелинейных аэродинамических уравнений движения имеет вид

$$\begin{aligned} \dot{x}_i &= \sum_\alpha F_i^\alpha c_\alpha + \sum_{j=1}^{5} F_i^j x_j + \sum_{1 \leqslant j < k}^{3} F_i^{jk} x_j x_k, \quad 1 \leqslant i \leqslant 3, \\ \dot{x}_i &= \sum_\alpha F_i^\alpha c_\alpha + \sum_{j=1}^{5} F_i^j x_j, \quad 4 \leqslant i \leqslant 5, \end{aligned} \tag{12.23}$$

а следовательно, и свойства динамической системы, изученной в разд. 2. Ее функция Ляпунова совпадает с указанной в (12.8),

а именно величина

$$E=\frac{1}{2}M\mathbf{v}\cdot\mathbf{v}+\frac{1}{2}I_{ij}\omega_i\omega_j \qquad (12.24)$$

равна функции Ляпунова, взятой с обратным знаком. Применение линейного анализа устойчивости непосредственно к системе (12.23) позволяет получить ее стационарные решения.

5. МНОГООБРАЗИЯ СТАЦИОНАРНЫХ СОСТОЯНИЙ

Стационарные решения системы (12.23) находят, полагая $\dot{x}_i = 0$. Получающиеся при этом уравнения имеют вид

$$0=F_i^a c_a + F_i^j x_j + F_i^{jk} x_j x_k, \quad 1\leqslant i\leqslant 3, \qquad (12.25а)$$

$$0=F_i^a c_a + F_i^j x_j, \quad 4\leqslant i\leqslant 5. \qquad (12.25б)$$

Два уравнения (12.25б) линейны по переменным $x_1, \ldots, x_5$. Используя их, можно представить x_4 и x_5 в виде линейных комбинаций управляющих параметров и остальных переменных состояния x_1, x_2, x_3. Подставляя эти выражения в (12.25а), получаем

$$\begin{aligned} 0&=k_1^\gamma c_\gamma + F_1'^j x_j + F_1^{23} x_2 x_3,\\ 0&=k_2^\gamma c_\gamma + F_2'^j x_j + F_2^{31} x_3 x_1,\\ 0&=k_3^\gamma c_\gamma + F_3'^j x_j + F_3^{12} x_1 x_2. \end{aligned} \qquad (12.26)$$

Здесь $k_i^\gamma c_\gamma$ — линейные комбинации управляющих параметров; $F_i'^j$ — функции, выражающиеся через первоначальные коэффициенты функции линейного отклика F_i^j; инерционные члены остаются неизменными.

Система трех совместных нелинейных уравнений (12.26) может быть решена следующим образом. Решая одновременно второе и третье уравнения (12.26) относительно x_2 и x_3, выразим последние через оставшуюся переменную состояния $p = x_1 = x$. Эти выражения представляют собой отношения квадратичных функций

$$x_2=\frac{Q_2}{Q_1}, \qquad x_3=\frac{Q_3}{Q_1}, \qquad (12.27)$$

$$Q_1=(F_2'^3+F_2^{31}x_1)(F_3'^2+F_3^{21}x_1)-F_2'^2F_3'^3,$$

$$\begin{aligned} Q_2=&-(k_3^\gamma c_\gamma + F_3'^1 x_1)(F_2'^3 + F_2^{31}x_1) +\\ &+F_3'^3(k_2^\gamma c_\gamma + F_2'^1 x_1), \end{aligned} \qquad (12.28)$$

$$Q_3=-(k_2^\gamma c_\gamma + F_2'^1 x_1)(F_3'^2 + F_3^{21}x_1) + F_2'^2(k_3^\gamma c_\gamma + F_3'^1 x_1).$$

Подстановка (12.27), (12.28) в первое из уравнений (12.26) приводит к следующему алгебраическому уравнению:

$$0 = k_1^{\nu} c_{\nu} Q_1^2 + F_1'^{1} x_1 Q_1^2 + F_1'^{2} Q_1 Q_2 + F_1'^{3} Q_1 Q_3 + F_1^{23} Q_2 Q_3. \quad (12.29)$$

Это уравнение пятой степени относительно переменной состояния $x = x_1$. Следовательно, уравнение состояния $\Phi' = 0$ получается из функции Ляпунова, которая представляет собой полином шестой степени относительно единственной переменной состояния x. С этим уравнением состояния можно связать катастрофу, требующую не более четырех управляющих параметров, т. е. катастрофу A_5. Таким образом, один из пяти управляющих параметров $c \in \mathbb{R}^5$ оказывается математически избыточным.

◇◇◇ Квадратичный член Q_1 не зависит от каких-либо управляющих параметров, в то время как квадратичные члены Q_2 и Q_3 зависят от них линейно. В результате все члены в выражении (12.29) для Φ' линейно зависят от управляющих параметров c_{ν}, за исключением членов, получающихся при развертывании $F_1^{23} Q_2 Q_3$.

6. ПРИЛОЖЕНИЕ К КОНКРЕТНОМУ ЛЕТАТЕЛЬНОМУ АППАРАТУ

Различные предположения, приводящие к уравнению состояния (12.29), могут выполняться для летательных аппаратов определенного класса, к которому относятся широко изученные реактивные летательные аппараты с короткими крыльями-стабилизаторами. Для такого летательного аппарата момент инерции I_{xx} относительно оси фюзеляжа много меньше, чем компоненты I_{yy}, I_{zz}[1]). Поскольку крылья одновременно служат и стабилизаторами, $I_{yy} \simeq I_{zz}$ и, в частности, $|I_{yy} - I_{zz}|/I_{xx} \simeq 0$, членом F_1^{23} (12.29) можно пренебречь. Вследствие этого уравнение состояния $\Phi'(x; c) = 0$ имеет пятый порядок относительно переменной состояния x и линейно относительно всех управляющих параметров. Такая модель не является безупречной, поскольку первоначальные уравнения движения (12.4) зависят лишь от двух управляющих параметров a, e и не зависят от трех остальных δa, δe, τ. Более совершенная модель должна была бы учитывать эти три дополнительные степени свободы управления.

---

[1]) Аэродинамические F_i^j и управляющие F_i^{ν} коэффициенты функции линейного отклика, а также компоненты тензора моментов инерции приведены в работе [2].

Для стандартной модели уравнение состояния имеет вид

$$\Phi'(x;\ c)=\sum_{k=0}^{5} a_k x^k=0, \tag{12.30}$$

$$a_5=-21{,}6,\quad a_4=-326{,}4a,\quad a_3=50{,}3e+358{,}9,$$
$$a_2=5412{,}6a,\quad a_1=11752{,}8e-1525{,}9,\quad a_0=-23015a. \tag{12.31}$$

(Если нет дополнительных указаний, углы a, e измеряются в радианах.) Несмотря на указанный недостаток, данная модель позволяет наглядно представить поверхность, задаваемую уравнением состояния $\Phi'=0$ в виде двумерного многообразия, погруженного в трехмерное пространство $\mathbb{R}^3$ с координатами $(x;\ a,\ c)$.

Рассматриваемое уравнение состояния, имеющее пятый порядок по x, может быть приведено к каноническому виду путем переноса начала координат $x=\tilde{x}-a_4/5a_5$. Это преобразование приводит к канонической форме катастрофы A_5, канонические коэффициенты которой являются нелинейными функциями физических управляющих параметров.

6.1. Бифуркационное множество

Теперь определим бифуркационное множество для функции Φ, поскольку бифуркационное множество для A_5 каноническое. Однако такая идентификация является нелинейной и сложной. Более простой путь определения бифуркационного множества заключается в параметрическом представлении.

Допустим, что $\Phi'(x;\ a,\ e)=0$ и $\Phi''(x;\ a,\ e)=0$ одновременно. Члены этих выражений могут быть нелинейными только относительно переменной состояния x. Таким образом, переменная состояния x может быть использована для параметрического представления управляющих параметров $(a(x),\ e(x))$ на бифуркационном множестве. Это можно сделать явным образом, если переписать $\Phi'=0$, $\Phi''=0$ в матричном виде

$$\begin{bmatrix} x^5 & x^4 & x^3 & x^2 & x & 1 \\ 5x^4 & 4x^3 & 3x^2 & 2x & 1 & 0 \end{bmatrix} \times$$
$$\times \begin{bmatrix} 0 & 0 & -21{,}6 \\ -326{,}4 & 0 & 0 \\ 0 & 50{,}3 & 358{,}9 \\ 5412{,}6 & 0 & 0 \\ 0 & 11752{,}8 & -1525{,}9 \\ -23015 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ e \\ 1 \end{bmatrix} = 0. \tag{12.32}$$

Каждому значению x соответствует пара совместных линейных уравнений для управляющих параметров (a, e). Решение этих уравнений единственно при условии, что определитель системы не равен нулю.

Параметрическое представление кривых складки $(a(x), e(x))$ и изолированных точек возврата показано на рис. 12.3. Рассматриваемое бифуркационное множество состоит из трех

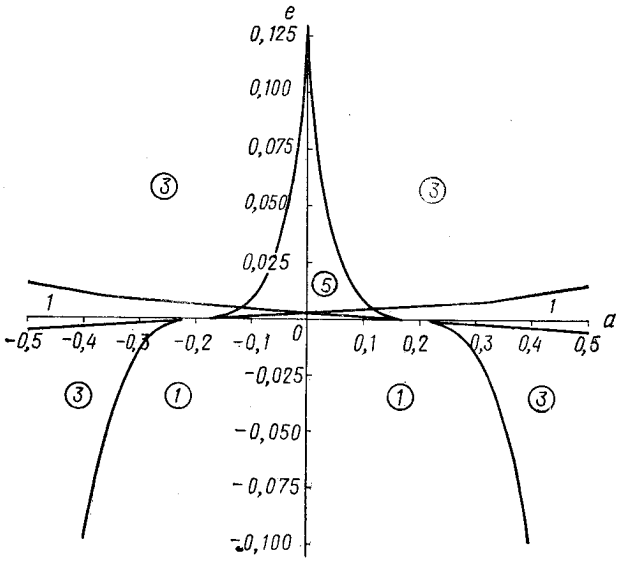


Рис. 12.3. Бифуркационное множество потенциальной функции $\Phi(x; a, e)$, описывающей состояние реактивного летательного аппарата с крыльями-стабилизаторами, изображенное на плоскости $\mathbb{R}^2$ управляющих параметров (a, e). Кривые складки ограничивают области с различным числом состояний равновесия (устойчивого и неустойчивого), указанным в кружках.

несвязных ветвей, поскольку определитель соответствующей матрицы размером 2×2 (12.32) имеет два трансверсальных нулевых сечения при $x_c = \pm 2{,}88$. Кривые складки, параметризованные значениями $-|x_c| < x < +|x_c|$, имеют по одной угловой точке, в то время как кривая складки, параметризованная значениями из внутреннего сегмента $-|x_c| < x < +|x_c|$, имеет три точки возврата.

Бифуркационное множество делит плоскость $\mathbb{R}^2$ управляющих параметров на открытые непересекающиеся области (рис. 12.3), в которых функция Φ имеет различное число (①, ③ или ⑤) критических точек, а система соответственно раз-

личное число (①, ② или ③) устойчивых стационарных состояний.

Положение вертикальных касательных $(x;\ a(x),\ e(x))$, их проекций, бифуркационное множество $\mathscr{S}_B$ в $\mathbb{R}^2$ и информация о числе критических точек являются достаточными сведениями для установления формы уравнения многообразия состояний.

6.2. Катастрофа

Сечения поверхности уравнения состояния $\Phi'(x;\ a,\ e)=0$ легко определить в параметрической форме на основании первого из матричных уравнений (12.32). Некоторые из сечений многообразия $\Phi'=0$ плоскостями $e=\text{const}$ показаны на рис. 12.4.

Вид этих сечений обычен. Для ряда значений управляющего параметра a возможны три устойчивых стационарных состояния, в то время как для других значений возможное число таких состояний равно двум или одному. Предположим, что начальные значения управляющих параметров определяются парой чисел $(a,\ e)=(0,\ 2°)$. Если e остается фиксированным, в то время как a медленно возрастает (на временной шкале это соответствует принципу максимального промедления), то вначале система x будет линейно реагировать на изменение a. При достижении $a\simeq 9°$ [1] линейный характер отклика резко нарушается и точка, описывающая состояние, скачком переходит на нижний лист. На этом листе состояние летательного аппарата совершенно не чувствительно к изменению угла отклонения элеронов $(-20<a<+20)$. Это похоже на ситуацию, при которой теряется возможность управления при $a\sim 9°$. Другими словами, если пилот попытается вернуть летательный аппарат в состояние $x=0$, уменьшая a при фиксированном e, он потерпит неудачу. Летательный аппарат не будет реагировать на изменение a до тех пор, пока оно не достигнет значения $a\simeq -20°$, при котором произойдет скачкообразный переход на верхний лист. Последующее увеличение управляющего параметра a вновь сопровождается отсутствием реакции вплоть до значения $a\simeq +20°$, при котором точка, характеризующая состояние системы, опять перескочит на нижний лист. Средний лист остается недостижимым при $e=2°$. Заставить систему вернуться на средний лист можно, лишь изменяя оба управляющих параметра — как a, так и e. Например, если a возвратить к нулю и сделать затем меньше нуля, то система скачкообразно придет к единственному устойчивому состоянию. Последующее возвращение параметра e к нулю вернет систему на центральный лист. Система может также быть возвращена на центральный лист без скачков (каким образом?).

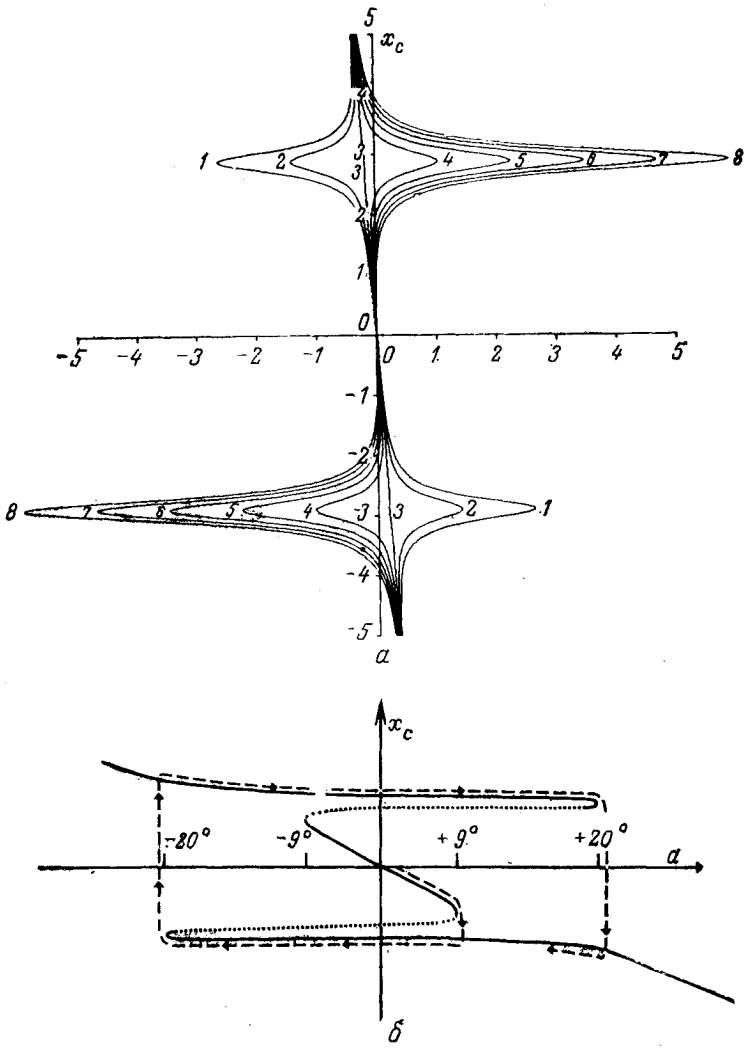


Рис. 12.4.

а — сечения многообразия равновесных состояний $\Phi'(x; a, e) = 0$ плоскостями $e = \text{const}$. *б* — при начальном состоянии $(x; a, e) = (0; 0, 2°)$ вследствие медленного изменения $a\,(0° \to 9°)$ уводит переменную состояния с центрального листа. При последующем уменьшении a переменная x остается на нижнем листе до тех пор, пока при достижении $a \to -20°$ она скачкообразно не перейдет на верхний лист. При дальнейших изменениях a центральный лист остается недостижимым до тех пор, пока параметр e остается фиксированным; ——— лист устойчивости; ····· лист неустойчивости; — — — траектория изменения состояния.

На рис. 12.5 показано сечение многообразия Φ' плоскостью $a = 0$. Для $e < 0°$ существует единственное устойчивое стационарное состояние $x = 0$. При значении $e \simeq 9°$ ветвь $x = 0$ становится неустойчивой и происходит бифуркация (A_{-3}) этой ветви на две симметричные неустойчивые ветви. Последние вращаются вокруг точки $e \simeq 0°$, знаменуя поочередный переход от неустойчивости к устойчивости. В диапазоне $e > 9°$ существуют

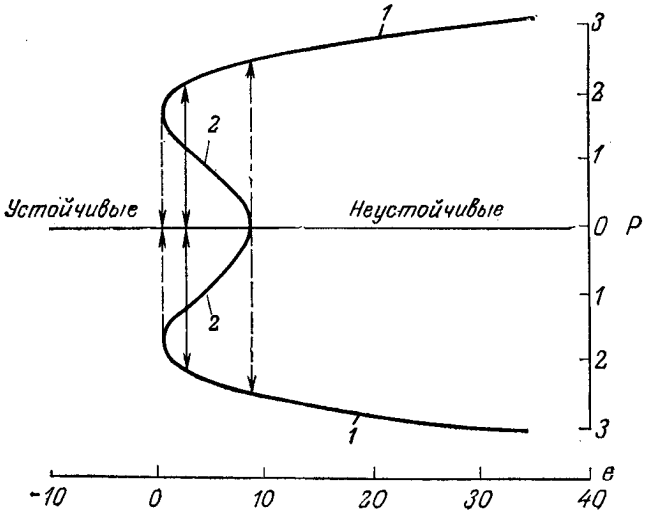


Рис. 12.5. Форма сечения многообразия стационарных состояний $\Phi' = 0$.
Свойства устойчивости различных ветвей определяются с помощью анализа устойчивости произвольной невырожденной точки многообразия стационарных состояний. Максвелловское множество показано пунктирными линиями, а бифуркационное множество — штриховыми линиями. Последние являются спинодалями. Здесь e измеряется в градусах и $p = x_1 = x$. *1* — устойчивая ветвь; 2 — неустойчивая ветвь.

два симметричных устойчивых решения, а также одно неустойчивое решение $x = 0$. Диапазону $0 < e < 9°$ отвечают три устойчивых и два неустойчивых решения.

Если выполняется принцип максимального промедления, то при возрастании e (при $a = 0$) и достижении значения $e \simeq 9°$ имеет место фазовый переход первого рода. При обратном изменении e фазовый переход первого рода наступает при $e \simeq 0°$. Налицо гистерезис. Невозможно предсказать, на какую ветвь — верхнюю или нижнюю — произойдет скачок с ветви $x = 0$ при $e \simeq 0°$. Это определяется малыми случайными возмущениями.

Если выполняется принцип Максвелла, то значение e, при котором происходит переход с ветви $x = 0$ на одну из ветвей $x \neq 0$, определяется на основе функции Ляпунова $\Phi(x;\ a = 0, e_c)$. В этом случае гистерезиса не наблюдается, но область фазового перехода первого рода очерчена спинодалями.

Эта однопараметрическая бифуркация неустойчива к возмущениям типа $a \neq 0$. На рис. 12.6 показано сечение многообразия $\Phi' = 0$ плоскостью $a = \text{const}$, *где величина* const *мала.* Недостаточная чувствительность спинодалей летательного аппарата может быть объяснена с помощью методики, описанной в гл. 11.

При фиксированных управляющих параметрах a, e перемен-

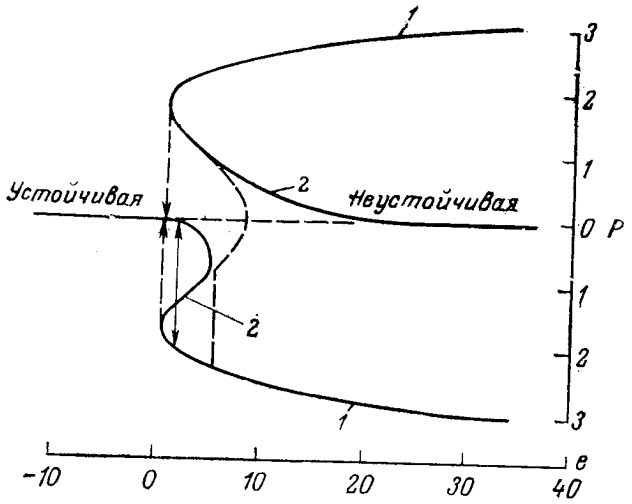


Рис. 12.6. Структурно неустойчивая бифуркация к изменениям формы $\neq 0$.

Сплошными линиями показаны сечения многообразия стационарных состояний плоскости $a \neq 0$. *1* — устойчивая ветвь; *2* — неустойчивая ветвь.

ная состояния x может принимать 1, 2 или 3 возможных устойчивых стационарных значения. Какое из них она примет, реально зависит от предыстории движения летательного аппарата. Для каждой тройки чисел $(x;\ a,\ e)$, удовлетворяющих $\Phi' = 0$, существует единственная точка $(x_1,\ x_2,\ x_3,\ x_4,\ x_5) \in \mathbb{R}^5$ пространства состояний системы (12.23). Значения x_2, x_3 однозначно определяются значениями x_1, x_2, x_3 с помощью (12.27); значения x_4, x_5 однозначно определяются значениями x_1, x_2, x_3 посредством линейных уравнений. Следовательно, если уравнение $\Phi' = 0$ является t-значным относительно (a, e), то двумерное многообразие $\mathcal{M}$ в $\mathbb{R}^5 \otimes \mathbb{R}^2$, определяемое уравнениями (12.25), также t-значно относительно $(a,\ e)$. Каждая точка многообразия $\mathcal{M}$ может быть исследована на линейную устойчивость. Линейный анализ устойчивости дает пять собственных значений. Точке $x \in \mathcal{M}$ соответствует динамически устойчивое стационарное состояние системы (12.19), если все пять собственных значений

имеют отрицательные действительные части. Потеря динамической устойчивости связана с тем, что действительная часть одного или нескольких собственных значений, возрастая, проходит через нуль. Бифуркационное множество $\mathscr{P}_B < \mathbb{R}^2$ соответствует тому, что действительное собственное значение, возрастая, проходит через нуль. Наибольшая часть устойчивых листов Φ' соответствует динамически устойчивым листам $\mathscr{M}$. Однако имеется некоторая часть устойчивых листов Φ', соответствующих точкам в $\mathscr{M}$, все действительные собственные числа которых отрицательны, но для которых имеется пара комплексно-сопряженных собственных значений, пересекших мнимую ось и сместившихся в положительную полуплоскость. Соответствующее состояние системы является неустойчивым, но вблизи него расположено устойчивое нестационарное состояние, называемое *предельным циклом*. Описание этого явления выводит за рамки данной главы.

6.3. Интегрирование уравнений движения

Теперь можно приступить к решению первоначально поставленной задачи — интегрированию уравнений движения. Исходные динамические уравнения движения определены на произведении пространств $\mathbb{R}^8 \otimes \mathbb{R}^5$ (12.19). Однако расцепление системы, достигаемое при высоких скоростях ($g/|\mathbf{v}| \to 0$), позволяет ограничиться рассмотрением усеченной динамической системы (12.23), определенной на $\mathbb{R}^5 \otimes \mathbb{R}^5$, а изучение частной модели — системы, определенной на $\mathbb{R}^5 \otimes \mathbb{R}^2$.

Для нахождения стационарных состояний динамической системы, определенной на $\mathbb{R}^5 \otimes \mathbb{R}^2$, вначале осуществляют преобразование в $\mathbb{R}^1 \otimes \mathbb{R}^2$ [(12.25)—(12.29)], а затем обратное преобразование полученных результатов из $\mathbb{R}^1 \otimes \mathbb{R}^2$ в $\mathbb{R}^5 \otimes \mathbb{R}^2$. Если положение двумерного многообразия стационарных состояний $\mathscr{M}$ в $\mathbb{R}^5 \otimes \mathbb{R}^2$ известно, то возможна разумная интерпретация интегрируемых уравнений движения.

Указанные уравнения были проинтегрированы следующим образом [1]. Предполагалось, что при $t < 0$ система находится в устойчивом стационарном состоянии, а при $t = 0$ происходит изменение значений управляющего параметра. Тогда система будет совершать асимптотический переход к новому стационарному состоянию, которое будет устойчивым в новом диапазоне значений управляющего параметра. Всегда можно преобразовать многообразие устойчивых стационарных состояний, исходя из некоторого начального стационарного состояния и производя затем последовательные и дискретные изменения управляющих параметров. Подобная процедура дает несколько больший объем информации относительно свойств устойчивости

решений, чем теория линейного нестационарного отклика. Бифуркационное множество ясно выявляется как дивергенция тензора чувствительности.

Более интересный и реалистический метод состоит в изучении (глобального) отклика системы на конечнозначное изменение управляющих параметров. Конечное состояние системы зависит от ее начального состояния $x^0 \in \mathbb{R}^5$, а также от конечных значений управляющих параметров.

Численное интегрирование уравнений было выполнено авторами работы [1]. Их результаты нормированы использованием одинаковых начальных условий $(x;\ c) = (0;\ 0) \in \mathbb{R}^8 \otimes \mathbb{R}^2$ для всех случаев интегрирования. Таким образом, от одного случая интегрирования к другому изменялись только конечные значения управляющего параметра. Некоторые части поверхности уравнения состояния оказались достижимыми из выбранного начального состояния, в то время как другие части остались в «тени» неустойчивых компонент многообразия уравнения состояния.

Изменение во времени переменных состояния (p, q, r, α, β, ϕ, θ) показано на рис. 12.7. Для указанных значений управляющих параметров существует устойчивое стационарное состояние на верхнем листе с $x \simeq -2$ и устойчивое стационарное состояние на верхнем листе с $x \simeq 1{,}6$. Устойчивые состояния на верхнем листе скрыты за потенциальным барьером, который представляет собой лист неустойчивости, так что получаемое интегрированием по времени решение должно асимптотически стремиться к устойчивому стационарному состоянию, соответствующему нижнему листу.

Кривые, приведенные на рис. 12.7, наглядно описывают эволюцию системы к стационарному состоянию. Видно, что вначале движение системы к стационарному состоянию происходит медленно. Это объясняется тем, что после изменения управляющих параметров точка, описывающая состояние, лежит вблизи среднего листа. В результате все силы F_i оказываются малыми, а движение — медленным. Не успев подойти к многообразию стационарных состояний, система тут же его покидает и переходит на складку. Пройдя складку, система удаляется от критического многообразия, продолжая движение в направлении нижнего листа. При удалении от критического многообразия скорость движения точки становится выше, вследствие возрастания сил F_i. С приближением к нижнему листу движение вновь замедляется. В конечном счете, система достигает стационарного состояния с $(x;\ a,\ e) \simeq (-2,\ 0;\ 3{,}8°,\ 2°)$. Однако если управляющие параметры имели бы значения $(a,\ e) = (3{,}7°,\ 2°)$, то конечное состояние системы лежало бы на центральном листе, причем приближение к этому листу было бы медленным.

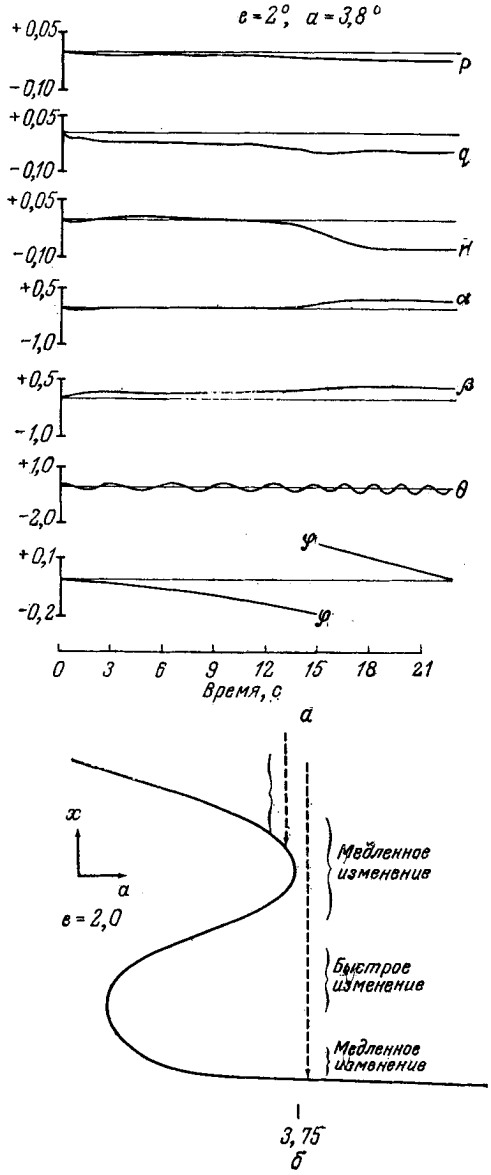


Рис. 12.7.

а — изменение во времени семи переменных состояния при достижении управляющим параметром предельных значений, переводящих систему с одной устойчивой ветви на другую [1]. *б* — критическое замедление, имеющее место при $a \to 3{,}75°$, вследствие того, что вынуждающая сила F_i $(x; c)$ достигает нуля на складке. Критическое замедление происходит в зависимости от того, будет ли стационарное состояние системы находиться на среднем или нижнем листе.

Для значений управляющих параметров $e = 2°$, $a > 3{,}8°$ изменение переменных состояния качественно подобно случаю, представленному на рис. 12.7, за исключением уменьшения масштаба времени. При $e = 2°$, $a < 3{,}7°$ конечное состояние си-

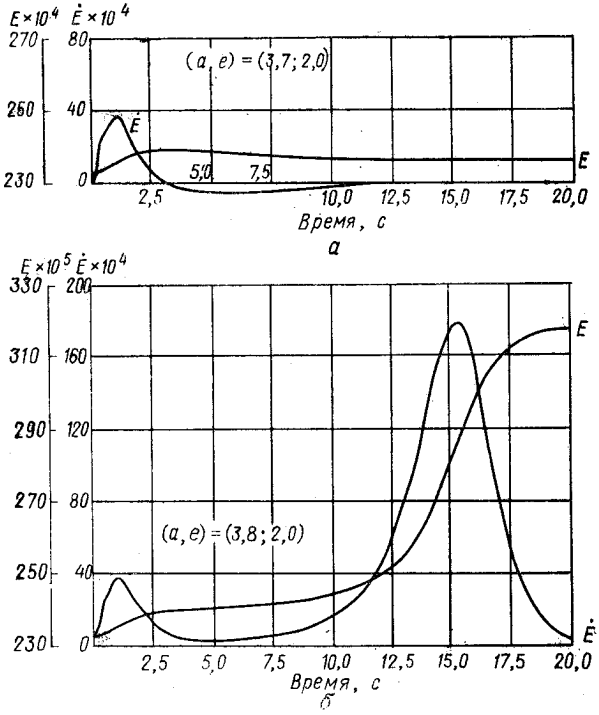


Рис. 12.8. Энергия и ее производная по времени как функции времени [1].

а — состояние системы перемещается к краю среднего листа и конечное стационарное состояние находится на среднем листе вблизи границы многообразия; функция Ляпунова $(-E)$ вначале убывает вплоть до момента $t \simeq 3$ с. В этот момент $\dot{E}$ меняет знак и затем асимптотически стремится к нулю, в то время как $-E$ возрастает до значения, соответствующего стационарному состоянию; *б* — система эволюционирует в направлении стационарного состояния, находящегося на нижнем листе.
Энергия измеряется в БТЕ; 1 БТЕ = 1050 Дж.

стемы лежало бы на центральном листе, но приближение к этому устойчивому стационарному состоянию происходило бы быстрее. Это уменьшение временного масштаба имеет место вследствие того, что силы F_i оказываются бо́льшими при удалении управляющих параметров от бифуркационного множества. Растяжение же временной шкалы при $e = 2°$, когда $a \to 3{,}75°\pm$, является примером критического замедления. Оно возникает в зависимости от того, будет ли асимптотически устойчивое состояние лежать на центральном или нижнем листе.

Система из пяти связанных нелинейных уравнений движения первого порядка может быть, вообще говоря, численно проинтегрирована при использовании нормированной комбинации начальных условий $x^{(i)} \in \mathbb{R}^5$ и конечных значений управляющих параметров $c^{(f)} \in \mathbb{R}^2$. Однако при отсутствии каких-либо сведений о структуре многообразия уравнения состояния интерпретация получаемой при таком интегрировании обширной информации в виде зависящих от времени функций, вызывает существенные затруднения. Вместе с тем, располагая сведениями о такой структуре, мы получаем возможность описывать качественное поведение траекторий $x(t) \in \mathbb{R}^5$ при произвольных $(x^{(i)}; c^{(f)}) \in \mathbb{R}^5 \otimes \mathbb{R}^2$, не интегрируя уравнения движения.

6.4. Энергия как функция Ляпунова

Кинетическая энергия летательного аппарата может быть выражена через линейную $\mathbf{v}$ и угловую $\boldsymbol{\omega}$ скорости, или через переменные состояния (p, q, r) $(=\boldsymbol{\omega})$ и (α, β) («полярные» углы вектора скорости)

$$E = \frac{1}{2} M\mathbf{v} \cdot \mathbf{v} + \frac{1}{2} I_{ij}\omega_i\omega_j, \tag{12.33}$$

$$\dot{E} = M\mathbf{v} \cdot \dot{\mathbf{v}} + I_x p\dot{p} + I_y q\dot{q} + I_z r\dot{r}. \tag{12.34}$$

Для рассматриваемой динамической системы энергетическая функция (12.33) есть частный случай общего выражения (12.9) и определена на полном многообразии уравнения состояния. На этом многообразии $\dot{E} = 0$. При приближении к стационарному состоянию $\dot{E}$ стремится асимптотически к нулю. Если $\dot{E}$ проходит через нуль (трансверсально), стационарное состояние не реализуется. Значение E может существенно измениться при скачкообразных переходах системы из одного состояния (с одного листа) в другое (на другой). Эволюция двух функций E и $\dot{E}$ показана (начальное условие $x = 0 \in \mathbb{R}^5$) на рис. 12.8.

7. ВЫВОДЫ

Цель настоящей главы состояла в том, чтобы показать, что элементарная теория катастроф может быть использована в качестве фундамента, на котором строится изучение динамических свойств некоторого летательного аппарата. Нам представляется, что именно такой путь позволяет наглядно продемонстрировать возможность использования не только методов, но и результатов элементарной теории катастроф для описания и понимания свойств динамических (не градиентных) систем с соответствующим образом записанными уравнениями движения.

Поскольку наш подход не ограничивался применением теории катастроф лишь только к летательным аппаратам, мы провели общее обсуждение используемых методов и ввели соответствующие систему и подсистему уравнений. Последняя содержит линейные, а также нелинейные члены определенного типа. Для динамических систем, описываемых уравнениями движения такого типа, можно определить обобщенную «потенциальную» функцию. В этом смысле динамические системы начинают походить на градиентные системы. Кроме того, мы предоставили читателям возможность составить представление о процедуре линейного анализа устойчивости.

В рассмотрение были также введены пространства переменных состояния и управляющих параметров, требуемых для описания простой модели летательного аппарата.

Показано, что уравнения движения могут быть существенно упрощены с помощью известных предположений, которые приближенно выполняются для реактивного летательного аппарата с крыльями-стабилизаторами. Эти предположения включают пренебрежение гравитационными членами ($g/|\mathbf{v}| \to 0$) и всеми нелинейными членами уравнений движения, за исключением инерционных. При таких предположениях получающиеся уравнения обладают свойствами ранее изученной системы уравнений. Вследствие этого многие свойства таких систем, вообще говоря, известны. В этой связи последовательная редукция произведения пространств переменных состояния управляющих параметров выглядит следующим образом:

$$
\begin{array}{c}
\mathbb{R}^n \otimes \mathbb{R}^k \to \mathbb{R}^8 \otimes \mathbb{R}^5 \\
\downarrow \quad (g/|\mathbf{v}| \to 0) \\
\mathbb{R}^5 \otimes \mathbb{R}^5 \\
\downarrow \quad (12.25б) \\
\mathbb{R}^3 \otimes \mathbb{R}^5 \\
\downarrow \quad (12.29) \\
\mathbb{R}^1 \otimes \mathbb{R}^5 \\
\downarrow \\
\mathbb{R}^1 \otimes \mathbb{R}^4
\end{array}
$$

Последняя редукция оказывается возможной, поскольку уравнение состояния $\Phi'(x, c) = 0$ является уравнением пятой степени. Наихудшей катастрофой, соответствующей такому многообразию, является A_5, которая имеет лишь четыре управляющих параметра.

Полученные результаты были использованы при описании состояния конкретного летательного аппарата. Уравнения состояния (12.30) и (12.31), включающие параметры этого летательного аппарата, позволяют легко определить бифуркацион-

ное множество. Как только последнее становится известным, сразу же можно найти связанную с ним катастрофу.

Прямое интегрирование рассматриваемой системы уравнений было реализовано Мерá, Кесселом и Керроллом [1]. Результаты этого численного интегрирования поддаются простой интерпретации, поскольку были предварительно установлены положение и структура критического многообразия. Более того, интегрирование уравнений движения позволяет обнаружить устойчивые компоненты многообразия уравнения состояния, а также бифуркационное множество (с помощью проектирования). И наоборот, знание многообразия стационарных состояний может быть использовано для получения решений уравнений движения вообще без какого-либо интегрирования.

Были использованы лишь две из четырех степеней свободы по управляющим параметрам, связанных с катастрофой A_5. Это обусловлено несовершенством модели, а не данного метода. Более реалистичная модель могла бы иметь ненулевые коэффициенты функции линейного отклика на отклонение τ руля направления, а также на разности отклонений δa, δe элеронов и рулей высоты. Однако эти дополнительные ненулевые коэффициенты внесли бы изменения в наши рассуждения лишь на стадии записи соотношения (12.31) и последующего анализа.

Литература

1. Mehra R. K., Kessel W. C., Carroll J. V., Global Stability and Control Analysis of Aircraft at High Angees of Attack, Cambridge: Scientific Systems, 1977.
2. Etkin B., Dynamics of Atmospheric Flight, New York: Wiley, 1972.

13

КАУСТИКИ И ДИФРАКЦИОННЫЕ КАРТИНЫ

Геометрическую оптику можно рассматривать как коротковолновый предел классической оптики, которая в свою очередь основана на аппроксимации уравнений Максвелла. В данной главе анализируется возможность переноса вариационного принципа из волновой оптики на ее коротковолновый предел. Проводится подробное исследование интеграла, определяющего амплитуду оптического сигнала в произвольной точке. Поскольку этот сигнал определяется ненулевыми вкладами от соседних критических точек, то можно немедленно перейти к изучению катастроф. Следует иметь в виду, что влияние критических точек Морса на амплитуду остается ограниченным, в то время как вклад неморсовских критических точек является чрезвычайно большим.

Значительное внимание уделяется изучению связи между большими воздействиями на амплитуду (каустики) и бифуркационными множествами катастроф; рассматривается случай конечной длины волны, т. е. ситуация, когда каждая каустика может быть аппроксимирована дифракционной картиной. С канонической каустикой связана каноническая дифракционная картина. Последняя может быть рассчитана, если известен интеграл Френеля от соответствующей функции катастрофы.

1. ГЕОМЕТРИЧЕСКАЯ ОПТИКА

Проблема дуализма волна — частица долгое время служила предметом словопрений философов и ученых. Дело в том, что при тщательном рассмотрении большинство объектов имело, грубо говоря, волноподобные свойства. Это означало, что эффективные размеры измерительной аппаратуры были сравнимы с исследуемой длиной волны. Однако, если длина волны много меньше, чем размер прибора, исследуемую систему практически можно рассматривать как частицу, т. е. система, подобно частице, может перемещаться вдоль определенных траекторий, даже если она фундаментально описывается волновым уравнением. Классическая механика и геометрическая оптика формально являются коротковолновыми пределами квантовой механики и электромагнитной теории.

◇◇◇ Классическая механика и геометрическая оптика с равным основанием могли бы быть названы также геометрической механикой и классической оптикой, поскольку и та и другая являются коротковолновыми асимптотиками волновых теорий. Поэтому нет ничего удивительного в том, что математические

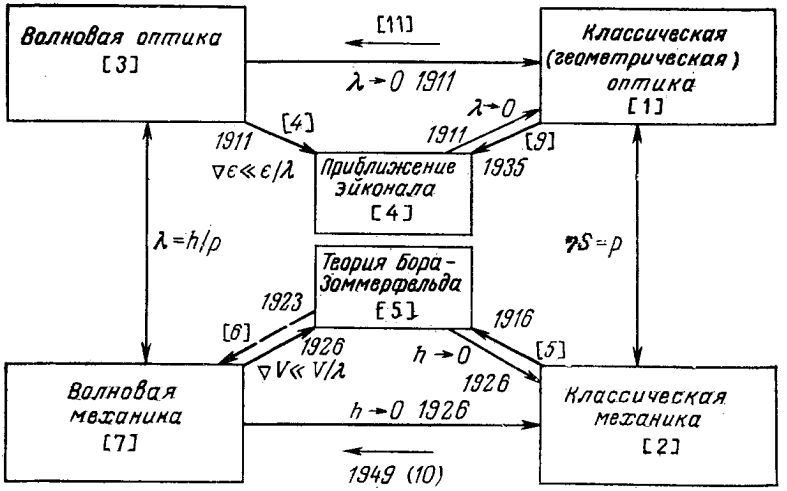


Рис. 13.1.

Существует тесная математическая связь между волновой и геометрической оптикой и механикой. Идея «заполнения пустых клеточек» схемы привела Шредингера к его волновому уравнению.

Библиография и комментарии

1. Geometrical Optics, Principle of Least Time, Oeuvres de Fermat, Vol. 2, Paris, 1891, p. 354.

2. Least Action Formulation of Classical Mechanics, W. R. Hamilton, *Trans. Roy. Irish Acad.*, 17, 1 (1833). Hamilton's Mathematical Papers, J. L. Synge, W. Conway, editors, Cambridge University Press, p. 285.

3. Wave Optics, J. C. Maxwell, A Treatise on Electricity and Magnetism: Oxford University Press, 1873.

4. The Eikonal Approximation, A. Sommerfeld and J. Runger, *Ann. Physik*, 35, 289 (1911).

5. The Quantum Patchwork, N. Bohr. *Phil. Mag.*, 26, 1 (1913). W. Wilson, *Phil. Mag.*, 29, 795 (1915). A. Sommerfeld, *Ann. Physik*, 51, 1 (1916).

6. М. Бор запрещает Гейзенбергу попытаться обобщить теорию Бора — Зоммерфельда, используя результаты матричной механики.

7. Wave Mechanics. L. deBroglie, *Nature*, 112, 540 (1923). L. deBroglie, These, Paris (1924). L. deBroglie, *Ann. Physique*, 3, 22 (1925). P. A. M. Dirac, *Sc. Am.*, 208, 45 (1963). (Шредингер выводит уравнение Клейна — Гордона и решает его для случая атома водорода. Однако свою работу он не публикует.)
E. Schrödinger, *Ann. Physik*, 79, 361 (1926). E. Sphrödinger, *Ann. Physik*, 79, 489 (1926). E. Schrödinger, *Ann. Physik*, 80, 437 (1926). E. Schrödinger, *Ann. Physik*, 81, 109 (1926).

8. The WKB Approximation. J. Liouville, *J. Math.*, 2, 16 (1837). J. Liouville, *J. Math.*, 2, 418 (1837). Lord Rayleigh, *Proc. Roy Soc. (London)*, A86, 207 (1912). H. Jeffreys, *Proc. London Math. Soc.* (2), 23, 428 (1923). G. Wentzel, *Z. Physik*, 38, 518 (1926). H. A. Kramers, *Z. Physik*, 39, 828 (1926). L. Brilloun, *Comptes Rendus*, 183, 24 (1926).

9. Вывод приближения эйконала из геометрической оптики П. Франком и Р. фон Мизесом. Die Differential- und Integralgleichungen der Mechanik und Physik, Vol. II, Braunschweig: Friederich Yiewig und Sohn (1935), Prob. 2—4.

10. Формулировка Р. Фейнманом волновой механики, основанная на интеграле действия. R. P. Feynman, *Revs. Mod. Phys.*, 20, 267 (1948). R. P. Feynman and A. R. Hibbs, Path Integrals and Quantum Mechanics, McGraw-Hill, New York, 1965.

11. Формулировка уравнений Максвелла, основанная на интеграле времени. Идея не была математически формализована. Основополагающая работа принадлежит Гюйгенсу.

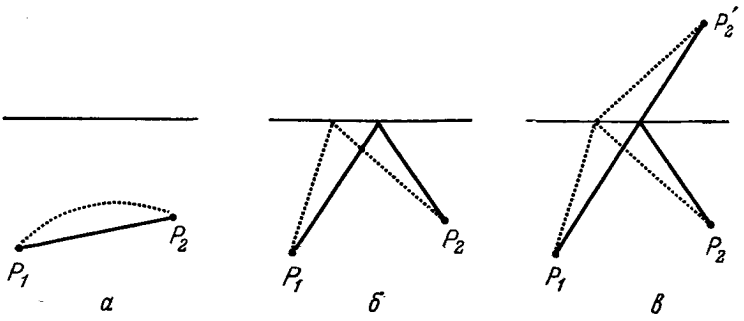


Рис. 13.2. Путь светового луча между точками P_1 и P_2 соответствует оптической длине пути.

описания основных представлений обеих дисциплин по существу идентичны.

◇◇◇ Последнее обстоятельство (т. е. идентичность гамильтоновских формулировок классической механики и геометрической оптики) и тот факт, что классическая оптика является коротковолновым пределом волновой оптики, позволили Шредингеру предположить, что если природа в некотором смысле симметрична, то классическая механика должна быть коротковолновым пределом некоторой новой дисциплины, которую по аналогии он назвал волновой механикой. Достаточно настойчивое следование этой аналогии позволило ему получить так называемое уравнение Шредингера. Связь, существующая между волновой оптикой и волновой механикой, а также между их коротковолновыми асимптотиками — классической оптикой и классической механикой — демонстрируется с помощью рис. 13.1.

Чтобы наглядно показать, почему геометрическую оптику можно считать коротковолновым пределом волновой оптики, рассмотрим два тесно связанных примера. В обоих случаях свет распространяется из точки P_1 в точку P_2 в трехмерном пространстве (рис. 13.2), но в первом случае это происходит согласно законам классической оптики, а во втором согласно законам волновой оптики. Существенно при этом, что независимо от того, как распространяется свет, в виде ли луча или в виде волны, его путь определяется на основании одного и того же принципа минимума.

Пример 1. Путь светового луча, попадающего из точки P_1 в точку P_2 (рис. 13.2, *а*), определяется, согласно принципу наименьшего времени Ферма, как

$$\delta \int_{P_1}^{P_2} dt = 0. \tag{13.1}$$

Если скорость света в некоторой эталонной среде (вакууме) равна c, а показатель преломления среды, через которую распространяется свет, есть $n(x, y, z)$, то $ds/dt = c/n$, и принцип Ферма может быть записан в виде

$$\delta \int_{P_1}^{P_2} \frac{1}{c} n\, ds = 0, \tag{13.2}$$

где ds — путь, проходимый светом за промежуток времени dt. Если показатель преломления n постоянен, то из (13.2) следует, что проходимый светом путь есть кривая минимальной длины, связывающая точки P_1 и P_2. Тогда принцип минимума может быть записан в виде

$$\delta D(P_1, P_2) = 0, \tag{13.3}$$

где вариация берется по всем путям, связывающим эти две точки.

Пример 2. Интенсивность света длиной волны λ, который наблюдается в точке P_2, есть [1]

$$I(P_1, P_2) = I_0 \left| \frac{1}{\lambda} \int \psi e^{2\pi i \Phi/\lambda}\, dx \right|^2, \tag{13.4}$$

где I_0 — интенсивность источника света в точке P_1, Φ — длина пути, связывающего точки P_1 и P_2 (посредством зеркала), ψ — мера, определенная на пространстве криволинейных интегралов, а интеграл берется по всем возможным путям, связывающим P_1 с P_2 (криволинейный интеграл). Если отраженный зеркалом M (рис. 13.2, *б*) свет полностью достигает точки P_2; то интерес представляет та область зеркала, отражение от которой вносит наиболее значительный вклад в интеграл (13.4) в коротковолновом пределе. Практически достаточно рассмотреть лишь кусочно-линейные пути с точкой излома, лежащей на зеркале. В этом случае криволинейный интеграл заменяется интегралом, взятым по поверхности зеркала:

$$\Phi = R_1(x) + R_2(x), \quad \psi = [R_1(x) + R_2(x)]^{-1}, \tag{13.5}$$

где $R_i(x)$ — кратчайшее расстояние между P_i и произвольной точкой (x_1, x_2) на поверхности зеркала. Тогда (13.4) сводится к виду

$$I(P_1, P_2) = I_0 \left| \frac{1}{\lambda} \iint \psi(x_1, x_2)\, e^{2\pi i \Phi(x_1, x_2)/\lambda}\, dx_1\, dx_2 \right|^2. \tag{13.6}$$

В предельном случае $\lambda \to 0$ вклад окрестности x^0 в быстро осциллирующий интеграл типа (13.6) является нулевым, если $\nabla\Phi \neq 0$ в x^0. Следовательно, в коротковолновом предельном случае бóльшая часть света, испускаемого в точке P_1 и достигающего точки P_2, отражается окрестностью точки $x^0 \in M \simeq \simeq \mathbb{R}^2$, такой, что

$$\delta(R_1 + R_2) = 0. \tag{13.7}$$

Этот результат эквивалентен евклидову построению, показанному на рис. 13.2, *в*. [Полезно сравнить (13.3) с (13.7).]

В результате отражения от криволинейной поверхности через некоторые точки трехмерного пространства $\mathbb{R}^3$ будут проходить два или более лучей. Вдоль огибающих (рис. 13.3) интенсивность света может быть существенно выше, чем в окрестности точек, лежащих вне огибающей. Подобное усиление интенсивности в точках огибающей может сопровождаться локальным

повышением температуры, достаточным для воспламенения бумаги или дерева. Такие огибающие называют *каустиками.*

Если нормальные интенсивности связаны с типичными кри-

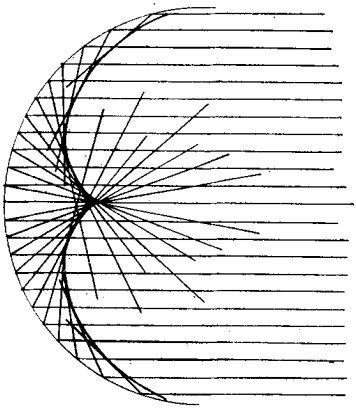


Рис. 13.3. Огибающая лучей, отраженных от криволинейной поверхности, образует каустики. Каустика, образованная отражением параллельного пучка сферической поверхностью, выглядит (предположительно) подобно катастрофе сборки.

тическими точками, в которых $\nabla\Phi = 0$ (критические точки Морса), то можно ожидать, что каустики связаны с нетипичными критическими точками.

2. МЕТОД СТАЦИОНАРНОЙ ФАЗЫ

В основе коротковолновой асимптотики лежит тот факт, что распространение световой волны описывается интегралами от быстро осциллирующих функций, имеющих резко выраженный пик в узком конусе направлений. Представим себе, что волна длиной λ распространяется от одной двумерной поверхности к другой. Предположим, что оптическая толщина среды между источником, расположенным в точке $x \in \mathbb{R}^n$ $(n = 2)$, и точкой наблюдения $\Omega \in \mathbb{R}^k$ $(k = 2)$ есть $\Phi(x; \Omega)$, а амплитуда сигнала, испущенного в точке x и принятого в точке наблюдения Ω, — $\psi(x; \Omega)$. Тогда интенсивность сигнала, принятого в Ω, определяется как

$$I(\Omega) = |A(\Omega)|^2,$$

$$A(\Omega) = \frac{1}{\lambda}\iint \psi(x; \Omega)\, e^{2\pi i \Phi(x;\,\Omega)/\lambda}\, d^2x. \tag{13.8}$$

Вектор Ω можно трактовать как точку наблюдения в физическом пространстве $\mathbb{R}^2$ или, согласно иной геометрической интерпретации, как направление наблюдения. Поскольку Ω может, вообще говоря, контролироваться наблюдателем, будем интерпретировать Ω как систему управляющих параметров (без дальнейших уточнений). Тогда естественно рассматривать x как систему переменных состояния.

Проанализируем n-мерный случай. Положим $k = 1/\lambda = 2\pi/\lambda$. Тогда соотношения (13.8) принимают вид

$$I(\Omega) = |A(\Omega)|^2,$$

$$A(\Omega) = (2\pi)^{-n/2} k^{n/2} \int \psi(x; \Omega)\, e^{ik\Phi(x;\,\Omega)}\, d^n x. \qquad (13.9)$$

Чтобы выявить, какой вклад в интеграл (13.9) дает окрестность произвольной точки x^0 области интегрирования, разложим Φ и $\ln\psi$ в ряд Тейлора в окрестности точки x^0:

$$\Phi(x; \Omega) = \Phi(x^0; \Omega) + \delta x \cdot \nabla\Phi + \ldots,$$

$$\ln\psi(x; \Omega) = \ln\psi(x^0; \Omega) + \delta x \cdot \nabla \ln\psi + \ldots. \qquad (13.10)$$

С учетом этого разложения (13.9) можно представить как

$$A_{x^0}(\Omega) = (2\pi)^{-n/2} k^{n/2} \psi(x^0; \Omega)\, e^{ik\Phi(x^0;\,\Omega)} \times$$

$$\times \int_{N(x^0)} (e^{ik\nabla\Phi + \nabla \ln\psi})^{\delta x} d^n x, \qquad (13.11)$$

где $N(x^0)$ — малая окрестность точки x^0 (в данном случае мы пренебрегаем членами более высокого порядка в разложении экспоненты). Оценка n-мерного интеграла в (13.11) сводится к последовательному нахождению n одномерных интегралов вида

$$k^{1/2} \int_{\delta x=-\varepsilon}^{\delta x=+\varepsilon} e^{(ikA+B)\,\delta x}\, dx = k^{1/2}\, \frac{e^{(ikA+B)\,\varepsilon} - e^{-(ikA+B)\,\varepsilon}}{ikA + B}. \qquad (13.12)$$

Выбирая ε так, чтобы $\varepsilon kA = 2\pi \times$ (целое число), это выражение можно свести к

$$k^{1/2}\, \frac{2 \operatorname{sh} \varepsilon B}{ikA + B} \xrightarrow{\varepsilon \to 0} \frac{2k^{1/2} B\varepsilon}{ikA + B} \xrightarrow[\to]{\lambda \to 0} 0. \qquad (13.13)$$

Последний предельный переход допустим, когда $A \neq 0$. Пока $\nabla\Phi(x^0; \Omega) \neq 0$, окрестность $N(x^0)$ можно всегда выбрать так, чтобы каждый из одномерных интегралов (13.12) обращался в нуль при предельном переходе $\lambda \to 0$.

Поскольку окрестности некритических точек дают нулевой вклад в амплитуду $A(\Omega)$, интеграл в (13.9) при $\lambda \to 0$ можно достаточно точно оценить, рассматривая лишь вклад окрестностей критических точек $\Phi(x; \Omega)$. В этих малых окрестностях функцию $\psi(x; \Omega)$ можно принять равной константе, так что

$$A(\Omega) \simeq \sum_{r=1} (2\pi)^{-n/2} k^{n/2} \psi(x^{(r)}; \Omega) \int e^{ik\Phi(x^{(r)} + \delta x;\, \Omega)}\, d^n x, \qquad (13.14)$$

$$\nabla\Phi(x^{(r)}; \Omega) = 0.$$

Описанный метод оценки значений интегралов от быстро осциллирующих функций, основанный на интегрировании лишь по окрестностям точек, в которых фаза $\Phi(x;\ \Omega)/\lambda$ меняется медленно, называют *методом стационарной фазы.*

Оценим интеграл (13.9) в окрестности критической точки Морса. Для этого, во-первых, не будем принимать во внимание вариации $\psi(x;\ \Omega)$ в окрестности этой точки (если только ψ не оказывается близким к нулю), во-вторых, разложим $\Phi(x;\ \Omega)$ в ряд Тейлора в окрестности x^0:

$$\Phi(x;\ \Omega)=\Phi(x^0;\ \Omega)+\frac{1}{2}\delta x_i\delta x_j\Phi_{ij}+\ldots \tag{13.15}$$

и, в-третьих, исключим члены более высокой степени. Затем с помощью ортогонального преобразования приведем Φ_{ij} к диагональному виду

$$(2\pi)^{-n/2}k^{n/2}\int\psi(x;\ \Omega)\,e^{ik\Phi(x;\ \Omega)}d^nx\to$$

$$\to\psi(x^0;\ \Omega)\,e^{ik\Phi(x^0;\ \Omega)}\prod_{j=1}^{n}\left[\left(\frac{k}{2\pi}\right)^{1/2}\int e^{(1/2)\,ik\lambda_j(\delta x_j)^2}dx_j\right]=$$

$$=\psi(x^0;\ \Omega)\,e^{ik\Phi(x^0;\ \Omega)}\prod_{j=1}^{n}\left(|\lambda_j|^{-1/2}e^{\pm i\pi/4}\right), \tag{13.16}$$

где знак «+» или «—» выбирается согласно знаку собственного значения $\lambda_i\neq 0$. Вклад окрестности критической точки Морса в интеграл от быстро осциллирующей функции может быть выражен в виде

$$\left(\frac{k}{2\pi}\right)^{n/2}\int\limits_{N(x^0)}\psi(x;\ \Omega)\,e^{ik\Phi(x;\ \Omega)}\,d^nx\simeq$$

$$\simeq\psi(x^0;\ \Omega)\,e^{ik\Phi(x^0;\ \Omega)}\,\frac{e^{it\pi/4}}{|\det\Phi_{ij}(x^0;\ \Omega)|^{1/2}}. \tag{13.17}$$

Здесь $t=(n-i)-i=n-2i$ — индекс инерции i-седловой точки Морса ($t=n$ в случае локального минимума).

3. СТЕПЕНЬ СИНГУЛЯРНОСТИ

До сих пор мы обсуждали вклад в значение интеграла от быстро осциллирующей функции окрестности некритической точки (13.13) и критической точки Морса (13.17), причем в первом случае он был нулевым, а во втором — конечным. Это дает основания предположить, что вклад окрестности неморсовской критической точки будет расходящимся и такие критические точки будут тесно связаны с каустиками.

Пусть x^0 — неморсовская критическая точка функции $\Phi(x; \Omega)$. Тогда с помощью гладкой замены переменных Φ можно привести к каноническому виду (2.4)

$$\Phi \doteq \mathrm{CG}(l) + M_l^{n-1}. \tag{13.18}$$

При этом вклад окрестности точки x^0 в интеграл от быстро осциллирующей функции выражается следующим образом:

$$A_{x^0}(\Omega) = \psi(x^0; \Omega)\, e^{ik\Phi(x^0;\,\Omega)} J(x^0) \int \left(\frac{k}{2\pi}\right)^{n-l/2} e^{ikM_l^{n-1}} d^{n-1}x' \times$$
$$\times \int \left(\frac{k}{2\pi}\right)^{l/2} e^{ik\,\mathrm{CG}(l)} d^l x', \tag{13.19}$$

где $J(x^0)$ — якобиан преобразования, приводящего к каноническому виду (13.18) в точке x^0. Интеграл по морсовским областям в (13.19) представлен в (13.17) так, что остается лишь вычислить вклад, вносимый за счет ростка катастрофы $\mathrm{CG}(l)$ с l переменными. Оценка и вычисление таких интегралов могут быть выполнены соответственно путем анализа размерностей и комплексного анализа.

Проиллюстрируем эти методы для одномерного случая. Предположим, что $\Phi(x; \Omega)$ и $\psi(x; \Omega)$ имеют вид

$$\Phi(x; \Omega) = \Phi(x^0; \Omega) + \Phi_r(x; \Omega),$$
$$\psi(x; \Omega) = \sum c_j (x - x^0)^j. \tag{13.20}$$

Здесь остаточный член $\Phi_r(x; \Omega)$ определяется разложением в ряд Тейлора в окрестности точки x^0. Для $|x - x^0| \ll 1$ члены $|x - x^0|^p$ с более низкими степенями меняются значительно быстрее членов $|x - x^0|^p$ с более высокими степенями $(p < q)$. Поскольку по существу мы имеем дело с быстрым изменением тригонометрических членов в быстро осциллирующих интегралах, достаточно рассмотреть лишь главный ненулевой член в $\Phi_r(x; \Omega)$:

$$\Phi_r(x; \Omega) = \gamma (x - x^0)^p, \quad \gamma \neq 0, \quad p > 0. \tag{13.21}$$

В результате имеем

$$\left(\frac{k}{2\pi}\right)^{1/2} \int_{x^0-\varepsilon}^{x^0+\varepsilon} \psi(x; \Omega)\, e^{ik\Phi(x;\,\Omega)}\, dx =$$
$$= e^{ik\Phi(x^0;\,\Omega)} \sum_{j=0} c_j \left(\frac{k}{2\pi}\right)^{1/2} \int_{x^0-\varepsilon}^{x^0+\varepsilon} (x - x^0)^j\, e^{ik\gamma(x-x^0)^p}\, dx. \tag{13.22}$$

Эти интегралы удобно вычислять, используя подстановку $y = k^{1/p} |\gamma|^{1/p} (x - x^0)$. Пределами интегрирования будут

$$\pm \varepsilon |\gamma|^{1/p} \left(\frac{2\pi}{\lambda}\right)^{1/p}. \tag{13.23}$$

Если λ — длина волны и размер окрестности интегрирования ε таков, что $\lambda^{1/p} \ll \varepsilon$, то пределы интегрирования преобразованного интеграла удобно расширить до $\pm\infty$ без заметного искажения значения интегралов:

$$\left(\frac{k}{2\pi}\right)^{1/2} \int_{x^0-\varepsilon}^{x^0+\varepsilon} (x - x^0)^j e^{ik\gamma(x-x^0)^p} dx =$$

$$= (2\pi)^{-1/2} k^{1/2-(j+1)/p} |\gamma|^{-(j+1)/p} I_{\pm}(j, p), \qquad (13.24)$$

$$I_{\pm}(j, p) = \int_{-\infty}^{+\infty} y^j e^{\pm iy^p} dy \qquad (13.25)$$

(знак плюс соответствует случаю $\gamma > 0$, а минус — случаю $\gamma < 0$). Результат (13.24) позволяет представить интеграл в ле-

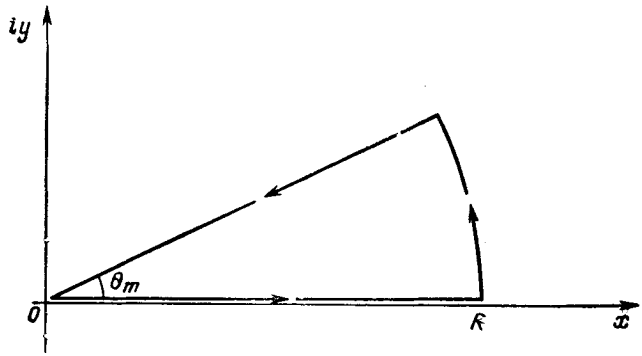


Рис. 13.4. Интеграл (13.25) можно оценить, вычисляя комплексный интеграл (13.26) вдоль трех участков показанного на рисунке секторообразного контура.

вой части в виде двух сомножителей — физического и геометрического. Физический сомножитель k^{σ} определяет зависимость амплитуды от длины волны. Геометрический сомножитель (13.25) может быть вычислен методом контурного интегрирования.

Комплексный интеграл

$$I = \oint z^j e^{izp} dz, \qquad (13.26)$$

взятый по контуру, показанному на рис. 13.4, равен нулю, поскольку в области, ограниченной контуром, нет полюсов, и мо-

жет быть записан в виде суммы интегралов по двум отрезкам прямых и дуге

$$0 = I = \int_0^R x^j e^{ix^p}\,dx + \int_0^{\theta_m} R^j e^{ij\theta} e^{iR^p[\cos p\theta + i\sin p\theta]} Re^{i\theta} i\,d\theta +$$

$$+ \int_R^0 R^j e^{ij\theta_m} e^{iR^p[\cos p\theta_m + i\sin p\theta_m]} e^{i\theta_m}\,dR. \qquad (13.27)$$

Если θ_m находится в интервале $0 \leqslant p\theta_m < \pi$, то интеграл по дуге должен дать в пределе $R \to \infty$ нулевой вклад; если $p\theta_m = \pi/2$, то третий член в (13.27) должен быть пропорционален интегралу от действительной функции:

$$\int_0^\infty x^j e^{ix^p}\,dx = e^{i(j+1/p)\pi/2} \int_0^\infty R^j e^{-R^p}\,dR. \qquad (13.28)$$

Подстановка $y = R^p$ приводит (13.28) к виду

$$\int_0^\infty x^j e^{\pm ix^p}\,dx = e^{\pm i(\pi/2)(j+1)/p} \frac{1}{p} \Gamma\left(\frac{j+1}{p}\right), \qquad (13.29)$$

где $\Gamma(n) = \int_0^\infty x^{n-1} e^{-x}\,dx$ — стандартная Г-функция. Порядок величин этих выражений — единица при $p > 0$, $j \geqslant 0$.

Выражения (13.24) и (13.25) содержат полезную информацию об относительном вкладе различных частей $\psi(x; \Omega)$ и $\Phi(x; \Omega)$ в быстро осциллирующий интеграл:

— при фиксированном p члены с бóльшими j дают меньший вклад в значение интеграла. Когда $\lambda \to 0$, основной вклад вносит главный член разложения в ряд Тейлора функции $\psi(x; \Omega)$ в окрестности x^0;

— интеграл $I_\pm$ $(j = 0, p)$ принимает значения

$$I_\pm(0, 1) = 0,$$

$$I_\pm(0, 2) = \pi^{1/2} e^{\pm i\pi/4},$$

$$I_\pm(0, p) = 2e^{\pm i\pi/2p}\Gamma\left(1 + \frac{1}{p}\right), \quad p \text{ четное}, \qquad (13.30)$$

$$= 2\cos\frac{\pi}{2p}\Gamma\left(1 + \frac{1}{p}\right), \quad p \text{ нечетное};$$

— зависимость амплитуды от длины волны на ростке катастрофы имеет вид

$$A \sim k^\sigma, \quad \sigma = \frac{1}{2} - \frac{1}{p} \quad \text{для} \quad A_{p-1}. \qquad (13.31)$$

Аналогичным образом исследуются остальные ростки катастрофы. Выбирая масштабы для аргументов $x' = k^{\alpha}x$, $y' = k^{\beta}y$, можно получить зависимость от длины волны из интеграла Френеля для ростка катастрофы $D_p(x, y) = x^2y + y^{p-1}$:

$$k^{2/2}\int_{-\infty}^{+\infty}\int e^{ik\,(x^2y+y^{p-1})}\,dx\,dy =$$

$$= k^{1-\alpha-\beta}\int_{-\infty}^{+\infty} e^{iky^{p-1}}d\,(k^{\beta}y)\int_{-\infty}^{+\infty} e^{ikx^2y}d\,(k^{\alpha}x). \qquad (13.32)$$

Из размерности аргументов следует, что $\beta(p-1) = 1$, $2\alpha = \beta = 1$, и интеграл принимает вид

$$k^{2/2}\int_{-\infty}^{+\infty}\int e^{ikD_p(x,\,y)}\,dx\,dy =$$

$$= k^{(p-2)/2\,(p-1)}\int_{-\infty}^{+\infty}\int e^{iD_p(x',\,y')}\,dx'\,dy'. \qquad (13.33)$$

Вывод. Интеграл Френеля для ростков элементарной катастрофы может быть представлен в виде произведения двух сомножителей, один из которых (k^{σ}) зависит от физических свойств через длину волны λ, а другой (интегральный) — от геометрии и безразмерных комплексных чисел

$$k^{l/2}\int e^{ik\,\mathrm{CG}\,(l)}d^lx = k^{\sigma}\int e^{i\,\mathrm{CG}\,(l)}d^lx' = k^{\sigma}I\,[\mathrm{CG}\,(l)]. \qquad (13.34)$$

Индекс σ и значение этих интегралов для ростков элементарной катастрофы сведены в табл. 13.1.

◇◇◇ Для быстро осциллирующего интеграла вида (13.9) с $\psi = 1$ вклад окрестности критической точки кратности μ может быть представлен в виде асимптотического разложения

$$A \sim \sum_{\alpha,\,\varkappa} C_{\alpha,\,\varkappa}\lambda^{n/2-\alpha}(\ln\lambda)^{\varkappa}, \qquad (13.35)$$

где $0 \leqslant \varkappa \leqslant n-1$[2].

4. КАУСТИКИ И РОСТКИ КАТАСТРОФЫ

Результаты, полученные в предыдущем разделе и обобщенные в табл. 13.1, показывают, что амплитуда, связанная с вырожденной критической точкой фазовой функции $\Phi(x; \Omega)$, расходится как $(1/\lambda)^{\sigma}$ при $\lambda \to 0$. Следовательно, бифуркационное множество $\mathcal{B}$ функции $\Phi(x; \Omega)$ определяет каустики, связанные с быстро осциллирующим интегралом. Поскольку семейство

функций $\Phi(x; \Omega)$ параметризовано с помощью n управляющих параметров $\Omega \in \mathbb{R}^n$, наихудшая катастрофа, типичная для данного случая, оказывается n-мерной. Вообще говоря, затруднительно наблюдать одновременно каустики более чем двух измерений, поскольку, как правило, исследуются дифракционные картины, полученные на двумерном экране ($\mathbb{R}^2$). Следовательно, образцы каустик, с которыми мы обычно сталкиваемся, могут быть получены как сечения $\mathbb{R}^2 \cap \mathcal{S}_B$.

Таблица 13.1. **Преобразование Френеля факторов элементарных катастроф в произведение двух сомножителей:**

$$k^{l/2} \int_{-\infty}^{+\infty} e^{ik\mathrm{CG}(l)} d^l x = k^{\sigma} I\,[\mathrm{CG}(l)]$$

Росток	Каноническая форма	σ	$I\,[\mathrm{CG}(l)]$
$A_{\pm(p-1)}$	$\pm x^p$	$\frac{p-2}{2p}$	$\left[e^{\pm i\pi/2p} + e^{\pm(-)^p i\pi/2p}\right] \Gamma\left(1+\frac{1}{p}\right)$
$D_{\pm(p+1)}$	$x^2y \pm y^p$	$\frac{p-1}{2p}$	$\left[e^{(i\pi/4)(1\pm 1/p)} + e^{(-i\pi/4)(1\pm(-)^{p+1}/p)}\right] \times$ $\times 4\Gamma\left(\frac{3}{2}\right) \Gamma\left(1+\frac{1}{2p}\right)$
$E_{\pm 6}$	$x^3 \pm y^4$	$\frac{5}{12}$	$2\cos\frac{\pi}{6}\,\Gamma\left(\frac{4}{3}\right) \times 2e^{\pm i\pi/8}\Gamma\left(\frac{5}{4}\right)$
E_7	$x^3 + xy^3$	$\frac{4}{9}$	$2\cos\frac{\pi}{6}\,\Gamma\left(\frac{4}{3}\right) \times 3\cos\frac{\pi}{9}\,\Gamma\left(\frac{11}{9}\right)$
E_8	$x^3 + y^5$	$\frac{7}{15}$	$2\cos\frac{\pi}{6}\,\Gamma\left(\frac{4}{3}\right) \times 2\cos\frac{\pi}{10}\,\Gamma\left(\frac{6}{5}\right)$

Если происходящая в окрестности Ω_0 катастрофа наивысшей размерности имеет место в самой Ω_0, то каустики, появляющиеся на экране $\mathbb{R}^2$, примыкающем к Ω_0, могут принимать лишь такие канонические формы, которые могут быть получены при рассмотрении двумерных срезов соответствующего бифуркационного множества. Например, если росток катастрофы A_4 расположен где-то в $\mathbb{R}^3$, то каустики, с которыми мы столкнемся, помещая плоский экран вблизи этого ростка в $\mathbb{R}^3$, могут принимать формы, изображенные на рис. 13.5, *а*. Аналогичное рассмотрение может быть проведено в случае ростков катастроф D_{+4} (рис. 13.5, *б*) и D_{-4} (13.5, *в*). Каноническое распределение ростков катастрофы приводит к аналогичным каноническим распределениям каустик трехмерных или большего ($n > 3$) числа измерений. Кроме того, известный канонический вид плоских сечений бифуркационных множеств и масштабные соотношения между управляющими параметрами позволяют на основе наблю-

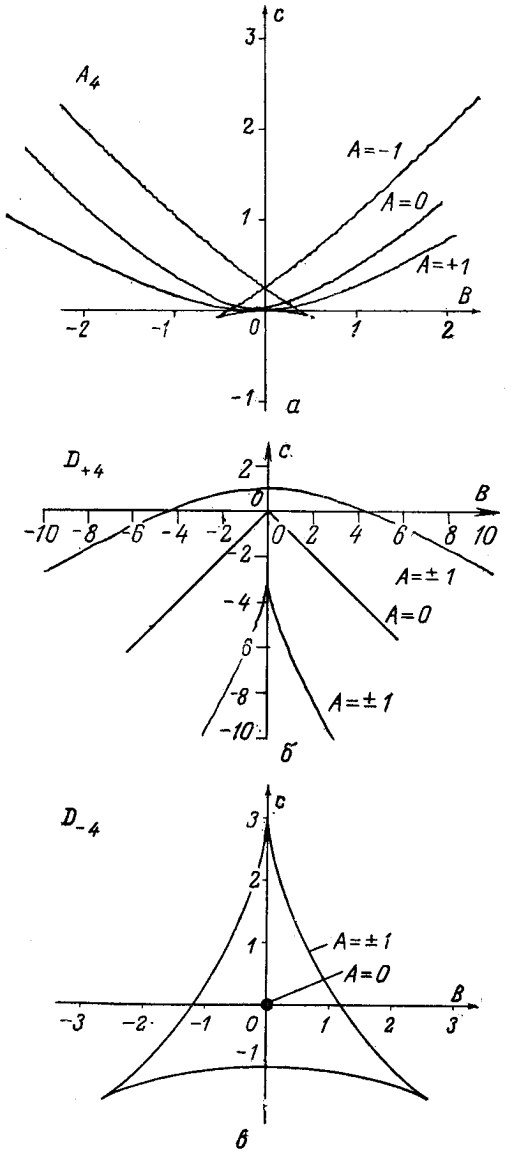


Рис. 13.5. Каустики, связанные с трехмерными катастрофами.

Катастрофы A_4, D_{+4}, D_{-4} являются сечениями соответствующего бифуркационного множества $\mathcal{S}_B$ плоскостью $\mathbb{R}^2$; изображенные на рисунках каустики получены на основании рис. 5.6, 5.11 и 5.17 соответственно.

дений каустик на двух параллельных плоскостях легко обнаружить в случае $n = 3$ плоскость, содержащую каустический росток.

◇◇◇ Обратим внимание на сходство понятий, используемых в данной главе и в гл. 10. Если фазовую $P - T$-диаграмму можно рассматривать как пересечение $\mathbb{R}^2 \cap \mathscr{S}_M$ физической плоскости $P - T$ с максвелловским множеством термодинамического потенциала $\mathscr{G}(E^\alpha; i_\beta)$, то диаграмму каустики на плоскости экрана можно представить как пересечение $\mathbb{R}^2 \cap \mathscr{S}_B$ плоскости физического экрана с бифуркационным множеством функции оптической длины пути $\Phi(x; \Omega)$. В обоих случаях, изменяя положение и ориентацию секущей плоскости $\mathbb{R}^2$, можно отобразить максвелловское множество и бифуркационное множество в пространство управляющих параметров соответствующей катастрофы. И в том и в другом случае существует дуализм между пространственными переменными и управляющими параметрами.

◇◇◇ Параллель между каустиками и термодинамическим равновесием может быть несколько расширена. Существует класс тригонометрических интегралов, зависящих от параметров Ω, которые могут быть представлены в виде

$$A(\Omega) = \int e^{i[S(x) - \langle \Omega, x \rangle]/\lambda} d^n x. \tag{13.36}$$

Здесь $\langle \Omega, x \rangle$ есть скалярное произведение двух n-мерных векторов x, Ω. Такие интегралы составляют специальный класс интегралов (13.9) с

$$\Phi(x; \Omega) = S(x) - \Omega \cdot x. \tag{13.37}$$

Метод стационарной фазы позволяет сразу же указать на подсистему точек $(x; \Omega) \subset \mathbb{R}^n \otimes \mathbb{R}^n$, которые в коротковолновом пределе дают существенный вклад в интеграл (13.36). Это многообразие определяется соотношением

$$\nabla_x \Phi(x; \Omega) = 0 \Leftrightarrow \frac{\partial S}{\partial x_j} = \Omega_j. \tag{13.38}$$

Такое n-мерное многообразие в $\mathbb{R}^n \otimes \mathbb{R}^n$ называют *лагранжевым многообразием*. Многообразия подобного типа существуют и в термодинамике, поскольку выполняется соотношение

$$\nabla_i \mathscr{U}(i_\alpha; E^\beta) = 0 \Leftrightarrow \frac{\partial G}{\partial i_\alpha} = -E^\alpha. \tag{13.39}$$

5. ДИФРАКЦИОННЫЕ КАРТИНЫ И ФУНКЦИИ КАТАСТРОФЫ

Структура каустики определяется двумерным сечением бифуркационного множества некоторой катастрофы. Амплитуда каждой компоненты каустики ведет себя как $(1/\lambda)^{\sigma}$, где показатель степени σ для простых катастроф приведен в табл. 13.1. Например, сингулярность складки для каустики подобна $(1/\lambda)^{2/6}$, а сингулярность сборки ведет себя как $(1/\lambda)^{2/4}$. Для коротковолнового предела эти результаты являются точными.

В предельном случае рассматриваемые интенсивности являются несобственными. Однако на практике интенсивности могут быть большими, но они остаются ограниченными сверху; в свою очередь длины волн могут быть малыми, но практически они ограничены снизу. При предельном переходе ($\lambda \to$ к малой величине, но $\lambda \not\to 0$) приближение стационарной фазы не является строго корректным. В окрестности каждого ростка катастрофы существуют функции с изолированными критическими точками, критические значения которых соизмеримы с λ. Следовательно, оценка амплитуды $A(\Omega)$ с помощью суммирования вкладов от изолированных критических точек (13.14) оказывается плохим приближением, и пределы интегрирования в (13.23) не могут быть расширены до $\pm\infty$.

Подобные оценки показывают, что необходимо вычислить вклады всех стационарных точек, находящихся в окрестности ростка катастрофы, для случая, когда предельное значение длины волны λ отлично от нуля. Это означает, что следует рассмотреть интеграл Френеля от стандартной деформации соответствующего ростка катастрофы. Проиллюстрируем это для ростка $A_p = x^{p+1}$. Дифракционный интеграл

$$k^{1/2}\int_{-\infty}^{+\infty} e^{ik\left[(p+1)^{-1}x^{p+1}+\sum_{j=1}^{p-1}a_j x^j\right]}\,dx = k^{\sigma} I[A_p;\ c_1,\ \ldots,\ c_{p-1}],$$

$$I[A_p;\ c_1,\ \ldots,\ c_{p-1}] = \int_{-\infty}^{+\infty} e^{i\left[(p+1)^{-1}y^{p+1}+\sum_{j=1}^{p-1}c_j y^j\right]}\,dy,$$

$$c_j = a_j k^{1-j/(p+1)}, \quad \sigma = \frac{p-1}{2(p+1)} \tag{13.40}$$

представляет собой преобразование Френеля канонической функции катастрофы. Интеграл $I[\mathrm{CG}(l)]$ есть канонический дифракционный интеграл, взятый первоначально в пространстве управляющих параметров. Для простейшей катастрофы A_2 стандартная деформация есть $(1/3)x^3 + tx$. При $t > 0$ стационарных точек не существует, а при $t < 0$ их две. Поэтому в коротковолновом пределе $\lambda \to 0$ следует ожидать, что росток катастрофы A_2 будет темным при $t > 0$ и светлым при $t \leq 0$, а амплитуда

будет уменьшаться при убывании t, поскольку кривизна в двух стационарных точках Морса будет убывать с уменьшением t. Области $t > 0$ и $t < 0$ разделяются каустикой при $t = 0$. Действительно,

$$\int_{-\infty}^{+\infty} e^{i[x^3/3+tx]}\,dx = 2\mathrm{Ai}\,(t), \tag{13.41}$$

где $\mathrm{Ai}(t)$ есть функция Эйри [3] (рис. 13.6). При положительном аргументе эта функция круто падает до нуля, а при от-

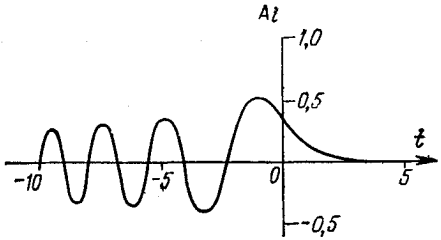


Рис. 13.6. Функция Эйри.

рицательном убывающем аргументе она быстро осциллирует, медленно убывая. На основании (13.40) имеем

$$k^{1/2}\int_{-\infty}^{+\infty} e^{ik[x^3/3+ax]}\,dx = 2k^{1/6}\,\mathrm{Ai}\,(ak^{2/3}). \tag{13.42}$$

Аналогичным образом могут быть рассмотрены другие катастрофы. Так, канонический дифракционный интеграл $I[A_3; a, b]$ представлен на рис. 13.7 [4]. Заметим, что поперечные сечения, ортогональные линиям складки, выглядят подобно функциям Эйри за вычетом некоторого фона. Происхождение функций Эйри связывают с двумя складывающимися листами, а генезис фона — с удаленным листом. Хотя со складкой можно связать существование темной области, для сборки, которая всегда имеет по меньшей мере одну стационарную точку, таких темных областей не существует.

Классические функции Эйри (1838) и Пирси (1947) непосредственно связаны с деформациями двух простейших катастроф — они являются преобразованиями Френеля от этих функций катастроф. Последнее обстоятельство дает возможность разработать методику конструирования аналогов канонических дифракционных картин более высокой размерности. Правда,

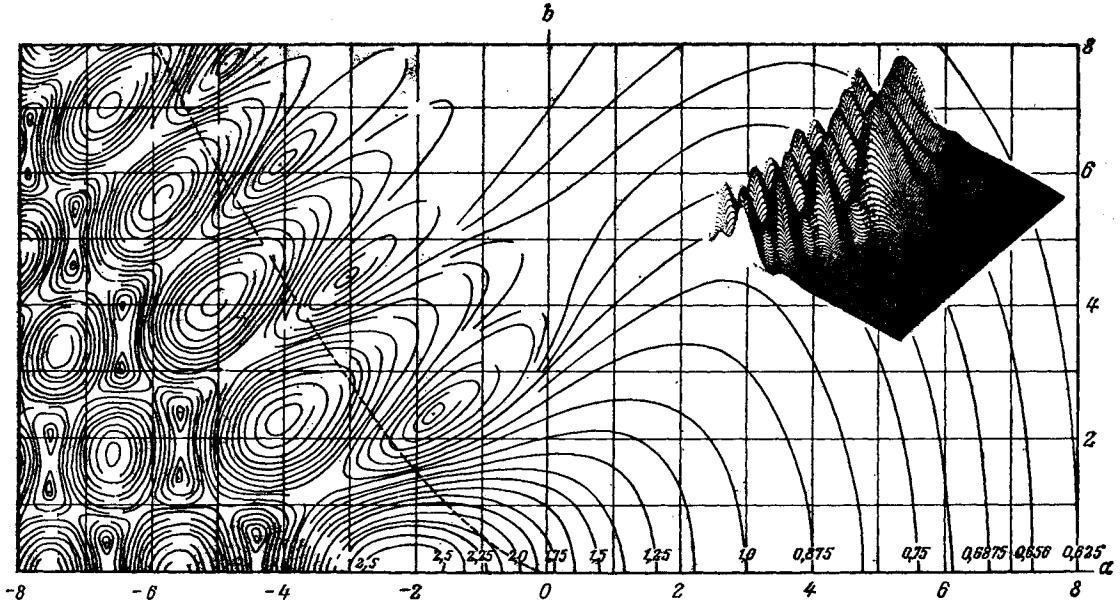


Рис. 1.37. Дифракционная картина и ее трехмерное изображение в случае катастрофы сборки [4, 5].

канонические дифракционные картины более высокой размерности не могут быть представлены графически и очень не легко найти их зрительный образ. Единственной возможностью визуализации таких картин является наблюдение на экране их двумерных сечений с последующим графическим воспроизведением после соответствующих вычислений. Например, если мы наблюдаем каустику, изображенную на рис. 13.5 для A_4 при $A=-1$, то мы знаем, что это одно из поперечных сечений бифуркационного множества A_4 плоскостью $a=\text{const}$. Следовательно, дифракционная картина на экране $\Omega_1\Omega_2$ будет иметь следующее распределение интенсивности:

$$I(\Omega_1, \Omega_2) \simeq \int_{-\infty}^{+\infty} \left| e^{i\,[x^5/5+ax^3+\Omega_1 x^2+\Omega_2 x]}\, dx \right|^2. \qquad (13.43)$$

В окрестности такой каустики дифракционная картина будет описываться функциями Эйри или Пирси. Однако если мы выберем другую плоскость наблюдения $\Omega_1\Omega_2$ так, чтобы наблюдалась каустика, изображенная на рис. 13.5 для A_4 при $A=0$, то для описания дифракционной картины, связанной с предельным переходом $\lambda \to$ малая величина $\to 0$, уже нельзя использовать классические функции, она будет описываться выражением (13.43) при $a=0$.

Если наши наблюдения относятся к плоским сечениям бифуркационного множества катастрофы A_4 при $b=\text{const}$ или $c=\text{const}$, то дифракционная картина, связанная с соответствующими каустиками, может быть описана (локально) функциями Эйри или Пирси. Однако когда плоскость наблюдения смещается так, что $b\to 0$ или $c\to 0$, то соответствующая дифракционная картина не будет больше описываться этими функциями.

Теперь становится ясной общая методика предсказания связанных с каустикой дифракционных картин в зависимости от катастрофы размерности k пространства управляющих параметров. Предположим, что плоскость наблюдения Ω_1, Ω_2 расположена на расстоянии d от начала отсчета в пространстве управляющих параметров $\mathbb{R}^k$. Тогда k управляющих параметров могут быть записаны как линейные комбинации Ω_1 и Ω_2 [ср. с (10.63)]:

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = \begin{bmatrix} M_1^1 & M_1^2 & M_1^3 \\ M_1^2 & M_2^2 & M_2^3 \\ \vdots & \vdots & \vdots \\ M_k^1 & M_k^2 & M_k^3 \end{bmatrix} \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ d \end{bmatrix}. \qquad (13.44)$$

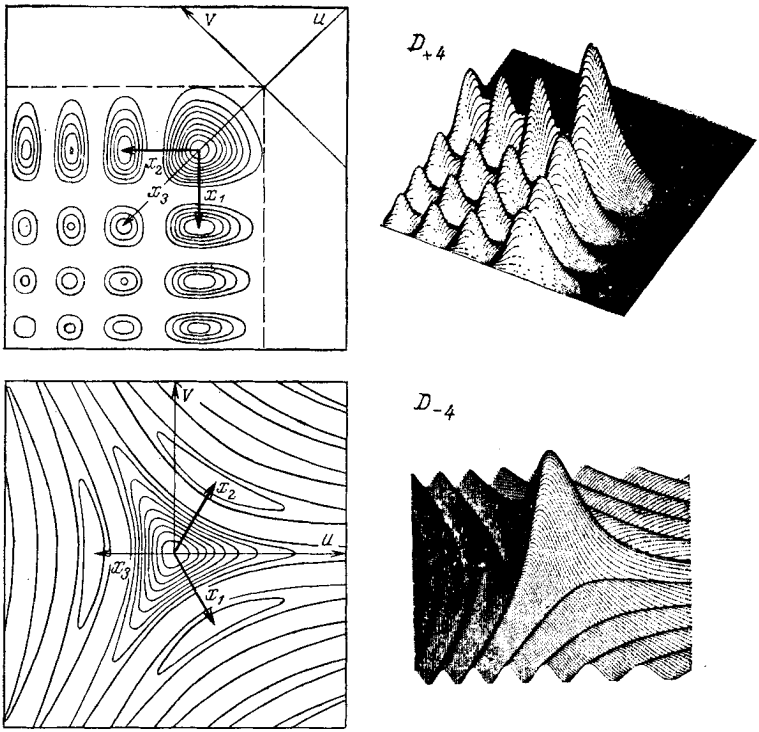


Рис. 13.8. Линии уровня и трехмерные изображения дифракционной картины в плоскости $a = 0$ для катастроф типа D_{+4} и D_{-4} [5].

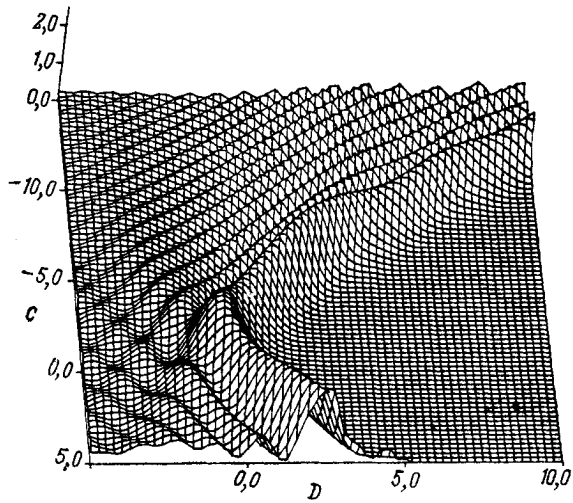


Рис. 13.9. Дифракционная картина в плоскости $b = +1$ для катастрофы $A_4 : x^5/5 + bx^3/3 + cx^2/2 + dx$. (Воспроизводится с разрешения Н. А. Поупа.)

Элементы матрицы $M_i^j (1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant 3)$ определяются структурой каустики на плоскости наблюдения. Если $\mathrm{CG}(l)$ есть росток соответствующей катастрофы и

$$P[c_1(\Omega), \ldots, c_k(\Omega)] = P(\Omega_1, \Omega_2, d) = P(\Omega) \qquad (13.45)$$

— стандартная деформация, где координаты c_i линейно выражаются через координаты точек наблюдения Ω_1, Ω_2 в плоскости и параметр расстояния d, то каноническая дифракционная картина будет иметь вид

$$I(\Omega_1, \Omega_2, d) \simeq \left| \int_{-\infty}^{+\infty} \ldots \int e^{i[\mathrm{CG}(l)+P(\Omega)]} d^l x. \right. \qquad (13.46)$$

Этот интеграл может быть численно оценен для любого погружения физического пространства $\mathbb{R}^2$ в пространство управляющих параметров $\mathbb{R}^k$.

Дифракционные картины в случае сечения катастроф D_{+4} и D_{-4} плоскостью $a = 0$ были рассчитаны Тринкхаусом и Дреппером [5] и воспроизводятся на рис. 13.8; дифракционные картины, соответствующие катастрофе A_4, были вычислены Поупом. На рис. 13.9 представлена дифракционная картина функции $F(x; b, c, d) = x^5/5 + bx^3/3 + cx^2/2 + 2x$ в сечении $b = +1$.

6. ВЫВОДЫ

Классическая (геометрическая) оптика, классическая механика и классическая термодинамика являются предельными случаями соответственно волновой оптики, волновой механики и статистической механики. В настоящей главе было показано, как может быть осуществлен предельный переход (при $\lambda \to 0$) в волновой оптике и как можно реконструировать волновую оптику из ее более старой ветви, геометрической оптики. При этом были рассмотрены связь классического вариационного принципа (Ферма) с вариационным принципом волновой оптики и метод стационарной фазы. Последний является прямым приложением разложений (2.1) — (2.3) и позволяет определить, что

— вклад окрестности некритической точки в интеграл (13.9) $\simeq 0$;

— вклад окрестности критической точки Морса в интеграл (13.9) конечен;

— вклад окрестности вырожденной критической точки в интеграл (13.9) является сингулярным типа $\simeq (1/\lambda)^{\sigma}$.

Эти результаты справедливы в предельном случае $\lambda \to 0$.

Было показано также, что каустики на экране $\mathbb{R}^2$ представляют собой просто двумерные сечения бифуркационного

множества соответствующей катастрофы. Мы «реконструировали» волновую оптику, рассмотрев возможные дифракционные картины, которые могли бы возникать в окрестности каустик при малой, но отличной от нуля длине волны. Канонические деформации катастроф позволяют нам сейчас конструировать канонические дифракционные картины, каждая из которых связана с канонической каустикой.

Приведенные математические построения без труда могут быть перенесены с оптики на механику простой заменой $\Phi(x; \Omega)/\lambda \to \mathscr{S}(x_i; x_f)/\hbar$, где $\mathscr{S}$ есть интеграл действия механической системы.

$\Diamond\Diamond\Diamond$ Полное рассмотрение связи между каустиками и теорией катастроф опиралось на работы [6—9].

Литература

1. Born M., Wolf E., Principles of Optics (1st ed.), London: Pergamon, 1959.
2. Арнольд В. И. Критические точки гладких функций и их нормальные формы. — УМН, 1975, 30 : 5, 3—65.
3. Airy G. B., On the Intensity of Light in a Neighborhood of a Caustic, *Trans. Camb. Phil. Soc.*, **6**, 379—403 (1838); см. также Abramowitz M., Stegun I. A., Handbook of Mathematical Functions, Washington, D. C.: NBS, 1964.
4. Pearcey T., The Structure of Elecromagnetic Field in the Neighborhood of a Cusp of a Caustic, *Phil. Mag.*, **37**, 311—317 (1946).
5. Trinkhaus H., Drepper F., On the Analysis of Diffraction Catastrophes, *J. Phys.*, **A10**, L11—L16 (1977).
6. Arnol'd V. I., Wave Front Evolution and Equivariant Morse Lemma, *Commun. Pure Appl. Math.*, **29**, 557—582 (1976).
7. Duistermaat J. J., Oscillatory Integrals, Lagrange Immersions and Unfolding of Singularities, *Commun. Pure Appl. Math.*, **27**, 207—281 (1974).
8. Jänich K., Caustics and Catastrophes, *Math. Ann.*, **209**, 161—180 (1974).
9. Berry M. V., Waves and Thom's Theorem, *Adv. Phys.*, **25**, 1—26 (1976).

14

КАНОНИЧЕСКИЕ ФОРМЫ ЖОРДАНА—АРНОЛЬДА

Анализ матриц или, более общо, систем линейных уравнений может быть выполнен точно так же, как и анализ функций, если учесть, что

— в качестве аналога уравнения состояния $\nabla V(x; c) = 0$ может быть использовано характеристическое уравнение $\det M(\lambda; c) = 0$;

— матрицы с невырожденными (изолированными) собственными значениями устойчивы относительно возмущения точно так же, как и функции с изолированными критическими точками;

— если функции с вырожденными критическими точками часто бывают локально эквивалентны каноническим росткам катастроф, то матрицы с вырожденными собственными значениями всегда глобально эквивалентны каноническим росткам матриц, или жордановым каноническим формам матриц. Кроме того, если росткам катастроф соответствуют канонические возмущения, то в случае жордановых канонических форм мы имеем дело с универсальной деформацией, или с так называемыми каноническими формами Жордана — Арнольда [1]).

Таким образом, исследовательская программа теории катастроф для систем линейных уравнений и функций полностью совпадает.

В данной главе анализируется связь, существующая между системами линейных уравнений, и вводится жорданова каноническая форма матрицы в тех случаях, когда имеется вырождение, рассматриваются произвольные и минимальные возмущения жордановых канонических форм, причем их значение поясняется путем обсуждения матричных аналогов элементарных катастроф типа A_2, A_3, A_4, а также определяется полный спектр самых плохих возможных вырождений, которые могут устойчиво встречаться в k-параметрическом семействе линейных операторов, и исследуется бифукационное множество, связанное с любым жордановым ростком.

1. ИССЛЕДОВАТЕЛЬСКАЯ ПРОГРАММА ТЕОРИИ КАТАСТРОФ В СЛУЧАЕ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Специалистам самого различного профиля довольно часто приходится иметь дело с системами линейных уравнений, зависящих от n переменных состояния $x \in \mathbb{C}$, которые в свою очередь удовлетворяют следующим двум условиям:

$$\mathscr{L}(\alpha x^{(1)} + \beta x^{(2)}) = \alpha \mathscr{L}(x^{(1)}) + \beta \mathscr{L}(x^{(2)}) \quad \text{(линейность)}, \quad (14.1)$$

$$\mathscr{L}(x) = 0 \quad \text{(уравнение)}.$$

[1]) Впервые исследование универсальной деформации жордановых канонических форм методами теории катастроф было выполнено В. И. Арнольдом.

Любой линейный оператор в $\mathbb{C}^n$ может быть представлен как квадратная матрица порядка n. Поэтому изучение канонических форм систем линейных уравнений сводится к изучению канонических форм матриц размером $n \times n$. При помощи преобразования подобия S (переход к новому базису) любая комплексная матрица может быть приведена к жордановой канонической форме

$$M \to SMS^{-1} = \begin{bmatrix} J_1(\lambda_1) & & & 0 \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_k(\lambda_k) \end{bmatrix}, \tag{14.2}$$

где $\lambda_1, \lambda_2, \ldots$ — собственные значения M, имеющие кратности (вырожденности) $d_1, d_2, \ldots$; J_i — жордановы матрицы, т. е. квадратные матрицы порядка d_i с λ_i по главной диагонали, 1 и 0 на диагонали, расположенной выше главной диагонали (остальные элементы — нули). Например,

$$J(\lambda) = \begin{bmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \lambda & 0 & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{bmatrix} (= \lambda^3\lambda^2). \tag{14.3}$$

(Мы будем иметь дело в основном с комплексными матрицами.)

Если семейство линейных систем $L(x; c)$ зависит от управляющих параметров $c \in \mathbb{C}^n$, то некоторые члены семейства могут быть вырожденными. В связи с этим жорданова каноническая форма, соответствующая типичному члену семейства, может быть совершенно отличной от жордановой канонической формы, соответствующей вырожденному (в большей или меньшей степени) члену семейства. Поэтому преобразование подобия, приводящее члены семейства к жордановой канонической форме, будет зависеть от управляющих параметров разрывно. Таким образом, становится очевидным, что методы, используемые при изучении функций, их канонических форм и универ-

сальных возмущений, могут быть использованы для изучения аналогичного спектра вопросов в случае систем линейных уравнений. Однако сначала рассмотрим следующие примеры.

Пример 1. Имеется линейный оператор $\mathscr{L} = I_n d/dt - M$ в пространстве $\mathbb{C}^n$, где M — квадратная матрица порядка n, $x \in \mathbb{C}^n$. Соответствующая линейная система может быть представлена как

$$\frac{d}{dt} x = M(c)\, x, \tag{14.4}$$

при этом предполагается, что матрица M зависит от управляющих параметров $c \in \mathbb{C}^n$. Если n собственных значений $\lambda_1(c)$, $\lambda_2(c)$, ..., $\lambda_n(c)$ различны при $c = 0$, то в силу соображений непрерывности они также будут различны и при малых c. Следовательно, если собственные значения $M(0)$ различны, то собственные значения произвольного возмущения

$$M(\delta c) = M(0) + \delta M, \qquad \delta M = \delta c_\alpha \frac{\partial}{\partial c_\alpha} M(c)\Big|_{c=0}$$

матрицы $M(0)$ также различны.

Пример 2. Пусть дано линейное дифференциальное уравнение

$$\dddot{x} + p\ddot{x} + q\dot{x} + rx = 0, \tag{14.5}$$

где точки обозначают производные по времени. Это уравнение может быть приведено к канонической матричной форме посредством замены

$$y_1 = x, \quad y_2 = \dot{x}, \quad y_3 = \ddot{x}. \tag{14.6}$$

Тогда

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -r & -q & -p \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}. \tag{14.7}$$

Теперь попытаемся установить, что произойдет, если немного «пошевелить» управляющие параметры $(p, q, r) \in \mathbb{C}^3$. Для этого вычислим собственные значения матрицы M, соответствующей линейному уравнению (14.5), используя характеристическое уравнение, которое получается непосредственно из (14.5) путем замены $d/dt \to \lambda$:

$$\lambda^3 + p\lambda^2 + q\lambda + r = 0. \tag{14.8}$$

Если характеристическое уравнение имеет три различных корня, то возмущение параметров ($p \to p + \delta p$ и т. д.) не оказывает влияния на каноническую структуру, соответствующую линейному уравнению (14.5) и его корням. Внимания заслуживает только случай, когда среди корней уравнения (14.5) имеются вырожденные. Не вдаваясь в подробности, можно с помощью замены $\lambda = \lambda' - (1/3)p$ перенести центр тяжести собственных значений в нуль и таким образом получить кубическое уравнение $\lambda'^3 + A\lambda' + B = 0$, не содержащее квадратичного члена. В результате становится очевидным, что интересующие нас точки в новом пространстве управляющих параметров $(A, B) \in \mathbb{C}^2$ лежат на бифуркационном множестве сборки.

Пример 3. Рассмотрим совместную систему линейных дифференциальных уравнений

$$\begin{aligned} \dddot{y} + A_1\ddot{y} + B_1\dot{y} + C_1 y &= -D_1 z, \\ \dot{z} + A_2 z &= -C_2\dot{y} \end{aligned} \tag{14.9}$$

и определим наименьшее семейство совместных линейных уравнений от y, $\dot{y}$, $\ddot{y}$, $\dddot{y}$ и z, $\dot{z}$, содержащее все возмущения вышеприведенной системы. Для

этого можно построить матрицу, соответствующую линейной системе (14.9),

$$\frac{d}{dt}\begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \\ z \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -C_1 & -B_1 & -A_1 & -D_1 \\ 0 & -C_2 & 0 & -A_2 \end{bmatrix} \begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \\ z \end{bmatrix}. \tag{14.10}$$

Этот результат наводит на мысль, что самое общее возмущение системы (14.9) может быть получено заменой

$$C_2\dot{y} \to B_2\ddot{y} + C_2\dot{y} + D_2y, \tag{14.11}$$

результатом которой является линейная система, зависящая от двух переменных состояния $(y,\ z) \in \mathbb{C}^2$ и восьми управляющих параметров $(A_1, \ldots, D_2) \in \mathbb{C}^8$. В действительности имеется меньшее семейство линейных систем, содержащее как (14.9), так и любое ее возмущение и представляющее собой 6-параметрическое семейство

$$\begin{aligned} \dddot{y} + A_1\ddot{y} + B_1\dot{y} + C_1y &= -D_1z, \\ \dot{z} + A_2z &= -D_2y \end{aligned} \tag{14.12}$$

в соответствующей системе координат.

Канонические формы линейного преобразования, существование которых было впервые доказано Жорданом, значительно упростили анализ линейных систем. Существование канонических форм семейств преобразований было доказано Арнольдом [1]. Канонические формы Жордана — Арнольда ведут к каноническому списку собственных значений и их зависимости от управляющих параметров, параметризующих члены семейства линейных уравнений. Эти канонические формы позволяют также определить, какие жордановы канонические формы присутствуют в семействе матриц и как они геометрически взаимосвязаны в пространстве управляющих параметров.

2. ВОЗМУЩЕНИЯ

Согласно гл. 3, анализ произвольных функций n переменных состояния следует начинать с поиска канонических форм таких функций в вырожденной критической точке. При исследовании систем линейных уравнений нашим «рабочим материалом» являются $n \times n$-матрицы. Поэтому роль вырожденной критической точки будут играть вырожденные собственные значения матриц, а канонических ростков в вырожденных критических точках — канонические жордановы формы (14.2) и (14.3) матриц с вырожденными собственными значениями; вместо нелинейных преобразований функций будут использоваться линейные преобразования линейных систем [1]).

[1]) Сведение к каноническим формам линейных систем при наличии вырожденности было проведено Жорданом задолго до того, как Уитни и Том решили более сложную проблему сведения к каноническим росткам функций.

Вычисления для матриц, сходные с вычислениями, которые выполняются при приведении функции к каноническому виду (гл. 3), можно найти в любой монографии [2], посвященной теории линейных векторных пространств. Поскольку, однако, в подобных работах, как правило, отсутствует информация относительно нахождения наименьшего универсального возмущения, то, по-видимому, имеет смысл рассмотреть этот вопрос более подробно.

Напомним еще раз, что, как только канонический росток определен, необходимо определить его наиболее общее возмущение. Для этого к ростку добавляют произвольное возмущение, а затем «забывают» все те члены возмущения, которые могут быть получены из ростка с помощью преобразований «подобия» (т. е. нелинейных координатных замен). По этому пути можно пойти и в случае линейных систем. Так, если имеется комплексная $n \times n$-матрица M_0, то можно добавить к ней произвольное возмущение δM, и тогда наиболее общим возмущением будет комплексная $n \times n$-матрица, все элементы которой достаточно малы. Некоторые малые матрицы δM (кстати, они не представляют для нас особого интереса, поэтому о них можно «забыть») могут быть получены посредством преобразований подобия матрицы M. Оставшиеся матрицы аналогичны универсальным возмущениям ростков катастроф, которые приведены в табл. 2.2.

Пример 1. Положим

$$M_0(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}. \tag{14.13}$$

Наиболее общее возмущение $M_0(\lambda)$ имеет вид

$$\delta M = \begin{bmatrix} \delta m_{11} & \delta m_{12} & \delta m_{13} \\ \delta m_{21} & \delta m_{22} & \delta m_{23} \\ \delta m_{31} & \delta m_{32} & \delta m_{33} \end{bmatrix}, \tag{14.14}$$

при этом любой элемент матрицы «мал» по величине. Можно получить большое семейство матриц из M_0, выполняя над ней преобразования подобия: $M_0 \to SM_0S^{-1}$. В общем преобразованные матрицы, близкие к M_0,— это лишь те матрицы, которые могут быть получены путем преобразований подобия, близких к тождественному. Следовательно,

$$SM_0S^{-1} \to (I + \delta S)\, M_0\, (I + \delta S)^{-1} = (I + \delta S)\, M_0\, (I - \delta S + \delta S^2 - \ldots) =$$
$$= M_0 + [\delta S,\ M_0] + \ldots. \tag{14.15}$$

Если мы возьмем малую матрицу δS:

$$\delta S = \begin{bmatrix} A & B & C \\ a & b & c \\ \alpha & \beta & \gamma \end{bmatrix}, \tag{14.16}$$

любой элемент которой «мал», то

$$[\delta S,\ M_0] = \begin{bmatrix} -a & A-b & B-c \\ -\alpha & a-\beta & b-\gamma \\ 0 & \alpha & \beta \end{bmatrix}. \tag{14.17}$$

Ясно, что некоторые матрицы δS_c коммутируют с матрицей M_0

$$\delta S_c = \begin{bmatrix} A & B & C \\ 0 & A & B \\ 0 & 0 & A \end{bmatrix} = \begin{bmatrix} \diagdown & \diagdown & \diagdown \\ & \diagdown & \diagdown \\ & & \diagdown \end{bmatrix}, \tag{14.18а}$$

в то время как другие матрицы δS_i не коммутируют с M_0; например,

$$\delta S_i = \begin{bmatrix} 0 & 0 & 0 \\ a & x' & z' \\ \alpha & \beta & y' \end{bmatrix}. \tag{14.18б}$$

Также очевидно, что некоторые матрицы δM_i могут быть записаны в виде коммутатора некоторой матрицы δS с M_0:

$$\delta M_i = \begin{bmatrix} \delta m_{11} & \delta m_{12} & \delta m_{13} \\ \delta m_{21} & \delta m_{22} & \delta m_{23} \\ 0 & -\delta m_{21} & -\delta m_{11} - \delta m_{22} \end{bmatrix}, \tag{14.18в}$$

в то время как другие матрицы δM_c — нет; например,

$$\delta M_c = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ x & y & z \end{bmatrix}. \tag{14.18г}$$

Можно пренебречь всеми возмущениями δM матрицы M_0 вида (14.18в), так как они «внутренне» порождаются координатными преобразованиями (т. е. заменой базиса). Следовательно, самое общее возмущение матрицы M_0 наименьшей размерности имеет вид

$$M_0 + \delta M = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ x & y & z+\lambda \end{bmatrix}. \tag{14.19}$$

Вернемся к обсуждению общей проблемы. Инфинитезимальные преобразования подобия $S \to I + \delta S$ вызывают возмущения M_0 в соответствии с

$$\begin{gathered} (I+\delta S)\, M_0\, (I+\delta S)^{-1} = M_0 + \delta M, \\ \delta M = [\delta S,\ M_0] + \mathcal{O}(2). \end{gathered} \tag{14.20}$$

Множество инфинитезимальных матриц δS образует линейное векторное пространство $\mathscr{S} \simeq \mathbb{C}^{n^2}$. Подмножество матриц в $\mathscr{S}$, коммутирующих с M_0, [ср. (14.18а)], образует линейное векторное подпространство $\mathscr{S}_c$ в пространстве $\mathscr{S}$; если $\delta S_1,\ \delta S_2 \in \mathscr{S}_c$, то

$$[\alpha\delta S_1 + \beta\delta S_2,\ M_0] = \alpha\,[\delta S_1,\ M_0] + \beta\,[\delta S_2,\ M_0] = 0. \tag{14.21}$$

Подпространство $\mathscr{P}_c$ определено однозначно. Пространство $\mathscr{P}$ может быть представлено в виде прямой суммы

$$\mathscr{P} = \mathscr{P}_c \oplus \mathscr{P}_i, \quad \mathscr{P}_c \cap \mathscr{P}_i = 0. \tag{14.22}$$

Подпространство $\mathscr{P}_i$ [ср. (14.18б)] не единственно. Матричная структура пространства $\mathscr{P}_i$ может быть выбрана подходящим образом.

Множество инфинитезимальных преобразований δM матрицы M_0 является линейным векторным пространством $\mathscr{P} \simeq \mathbb{C}^{n^2}$. Подмножество матриц $\mathscr{P}$, которые могут быть представлены в виде $[\delta S, M_0]$, образует линейное векторное подпространство $\mathscr{P}_i$ [внутреннее, ср. с (14.18в)] в пространстве $\mathscr{P}$; если $\delta M_1 = [\delta S_1, M_0]$ и $\delta M_2 = [\delta S_2, M_0]$, то

$$\alpha\delta M_1 + \beta\delta M_2 = \alpha\,[\delta S_1, M_0] + \beta\,[\delta S_2, M_0] = [\alpha\delta S_1 + \beta\delta S_2, M_0]. \tag{14.23}$$

Подпространство $\mathscr{P}_i$ определено однозначно. Пространство $\mathscr{P}$ может быть представлено как

$$\mathscr{P} = \mathscr{P}_i \oplus \mathscr{P}_c, \quad \mathscr{P}_i \cap \mathscr{P}_c = 0. \tag{14.24}$$

Подпространство $\mathscr{P}_c$ [ср. с (14.18г)] не единственно. Матричная структура пространства $\mathscr{P}_c$ может быть выбрана подходящим образом.

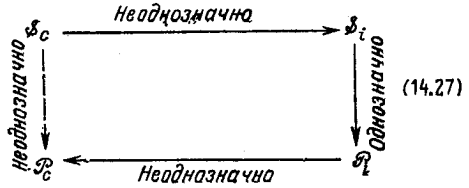
Ранее мы уже пояснили, почему между двумя пространствами $\mathscr{P}$ и $\mathscr{P}$ должен существовать некоторый тип двойственности. В частности, существует взаимно однозначное соответствие между (не единственным) пространством $\mathscr{P}_i$ и единственным пространством $\mathscr{P}_i$. Предположим, что для некоторого $\delta M \in \mathscr{P}_i$ можно найти две матрицы δS_1, $\delta S_2 \in \mathscr{P}_i$, коммутаторы которых с M_0 дают δM. Тогда $\delta S_1 - \delta S_2 \in \mathscr{P}_i$, и из

$$[\delta S_1 - \delta S_2, M_0] = \delta M - \delta M = 0 \tag{14.25}$$

следует, что $\delta S_1 - \delta S_2 \in \mathscr{P}_c$. Так как $\mathscr{P}_c \cap \mathscr{P}_i = 0$, то $\delta S_1 - \delta S_2 = 0$. Поэтому соответствие между этими пространствами таково:

$$\begin{array}{ccccc} & & \text{единственно} & & \\ \mathscr{P} = & \mathscr{P}_c & \oplus & \mathscr{P}_i & \\ & \updownarrow 1\text{–}1 & & \updownarrow 1\text{–}1 & \\ \mathscr{P} = & \mathscr{P}_c & \oplus & \mathscr{P}_i & \\ & & & \text{единственно} & \end{array} \tag{14.26}$$

Поскольку $\mathscr{P}$ и $\mathscr{P}$ изоморфны $\mathbb{C}^{n^2}$, а $\mathscr{P}_i$ изоморфно $\mathscr{P}_i$, то должно существовать взаимно однозначное соответствие между $\mathscr{P}_c$ и $\mathscr{P}_c$. Наименьшее универсальное возмущение матрицы M_0 лежит в $\mathscr{P}_c$. Таким образом, $\mathscr{P}_c$ может быть построено, как показано ниже:



(14.27)

Пример 2. Предположим, что M_0 имеет различные собственные значения. Без потери общности можно считать, что матрица M_0 диагональна и имеет собственные значения $\lambda_1 \neq \lambda_3 \neq \ldots \neq \lambda_n$. Тогда

$$\delta M = [\delta S, M_0], \tag{14.28}$$

$$\delta m_{ij} = \delta S_{ij}(\lambda_j - \lambda_i). \tag{14.29}$$

Поскольку собственные значения различны, все внедиагональные элементы возмущения δM могут быть порождены внутренним образом. Универсальное возмущение диагональной матрицы с различными собственными значениями само является диагональной матрицей. Возмущение вызывает лишь небольшое смещение значений изолированных собственных значений, не изменяя при этом знак их действительной части [1]), или свойства устойчивости M_0. В этом смысле матрица с n невырожденными (изолированными) собственными значениями аналогична морсовским функциям с невырожденными (изолированными) критическими точками.

Пример 3. Предположим, что M_0 имеет собственные значения $\lambda_1, \lambda_2, \ldots$ с вырожденностями $d_1, d_2, \ldots$. Тогда матрица M_0 может быть взята в жордановой канонической форме

$$M_0 = \left[\begin{array}{c|c} J_1(\lambda_1) & \\ \hline & J_2(\lambda_2) \end{array}\right] \qquad (14.30)$$

Инфинитезимальное преобразование подобия δS также может быть представлено в блочной форме

$$\delta S = \begin{bmatrix} S_{11} & S_{12} & S_{13} \ldots \\ S_{21} & S_{22} & \\ \vdots & & \ddots \end{bmatrix}, \qquad (14.31)$$

где $S_{ij} - d_i \times d_j$-матрица. Коммутатор дает

$$\delta M_{ij} = S_{ij} J_j(\lambda_j) - J_i(\lambda_i) S_{ij}. \qquad (14.32)$$

Полученное уравнение может быть однозначно разрешено для элементов $d_i \times d_j$-матрицы S_{ij} через элементы $d_i \times d_j$-матрицы δM_{ij}, если $i \neq j$, так как в этом случае $\lambda_i \neq \lambda_j$. Следовательно, все внеблочно-диагональные возмущения δM являются внутренними, и $\mathscr{P}_c$ может быть выбрано так, что его матричное представление будет иметь следующую структуру:

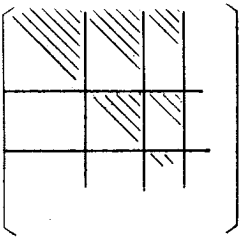
$$\mathscr{P}_c = \begin{bmatrix} (\delta M)_{11} & 0 & \cdots \\ 0 & (\delta M)_{22} & \\ \vdots & & \ddots \end{bmatrix}. \qquad (14.33)$$

Пример 4. Предположим, что M_0 имеет k-кратно вырожденное собственное значение и ее жорданова каноническая форма имеет структуру

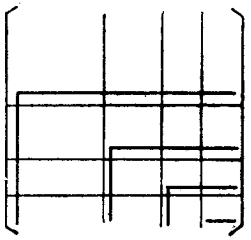
$$M_0 \to J(\lambda) = \left[\begin{array}{cc|cc} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & 1 & \\ & & \lambda & \\ \hline & & \lambda & 1 \\ & & & \ddots \\ & & & 1 \\ & & & \lambda \end{array}\right] \qquad (14.34)$$

[1]) Предполагается, что все собственные значения имеют отличные от нуля действительные части.

Подпространство $\mathscr{P}_c$ матриц может быть схематически представлено как

$\mathscr{S}_C =$  (14.35)

при этом элементы матриц на отмеченных участках диагонали равны между собой, а остальные элементы равны нулю [сравните с (14.18а)]. Подпространство матриц $\mathscr{P}_c$ может быть выбрано так, что его схематическим представлением будет [1]

$\mathscr{P}_C =$  (14.36)

при этом отмеченные элементы матриц независимы, а остальные равны нулю [сравните с (14.18г)]. Наиболее общее возмущение жордановой матрицы (14.34) $\lambda^{n_1}\lambda^{n_2}\lambda^{n_3}\ldots$ $(n_1 \geqslant n_2 \geqslant n_3 > \ldots > 1)$ [сравните с (14.3)] имеет следующую размерность:

$$D = n_1 + 3n_2 + 5n_3 + \ldots . \tag{14.37}$$

Пример 5. Наиболее общее возмущение жордановой матрицы

$$M_0 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

имеет вид

$$\delta M = \left[\begin{array}{ccc:c} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hdashline x_4 & x_3 & x_2 & x_1 \\ x_5 & 0 & 0 & x_6 \end{array}\right] \tag{14.38}$$

◇◇◇ Смотрите (14.12).

◇◇◇ Читатель, который внимательно следил за ходом наших рассуждений, не мог не увидеть аналогии, между

возмущениями, изучаемыми в гл. 4, и возмущениями, рассматриваемыми в этой главе, однако он также не мог не заметить, что мы потерпели неудачу, пытаясь довести аналогию между анализами функции и матриц до конца. Объясняется это тем, что нелинейные преобразования, описанные в гл. 4, были глобальными, в то время как линейные преобразования, рассмотренные в этой главе, локальны (инфинитезимальны). Инфинитезимальные преобразования подобия были введены для того, чтобы стало возможным работать с линейным векторным пространством $\mathscr{S}$ (алгеброй Ли), а не с группой Ли G преобразований подобия. Освобождение от этих инфинитезимальных ограничений достигается путем замен

$$\mathscr{S} \to G \text{ (группа Ли)},$$
$$\mathscr{S}_c \to C \text{ (централизатор)},$$
$$\mathscr{S}_i \to G/C \text{ (фактор-множество)}.$$

Локальное и глобальное рассмотрения дают эквивалентные результаты, хотя первое значительно проще. Это наводит на мысль о существовании «инфинитезимального аналога» методов, описанных в гл. 4.

3. ПРИЛОЖЕНИЯ

3.1. Катастрофы типа A_2

В этом и следующем разделах мы будем предполагать, что матрица M_0 уже приведена к жордановой канонической форме. Кроме того, предположим, что все собственные значения M_0 вещественны. Следовательно, можно ограничиться рассмотрением лишь вещественных возмущений δM матрицы M_0. Сделав подобные предположения, мы естественно вводим ограничение, рассматривая вместо поля комплексных чисел только поле вещественных чисел: $\mathbb{C} \to \mathbb{R}$.

Пусть M_0 — 2×2-матрица

$$M_0(\gamma) = \gamma I_2 + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{14.39}$$

Самое общее семейство матриц, содержащее вырожденную 2×2-матрицу $M_0(\gamma)$, имеет вид

$$M(\gamma;\ x,\ y) = M_0(\gamma) + \begin{bmatrix} 0 & 0 \\ y & x \end{bmatrix} = \begin{bmatrix} \gamma & 1 \\ y & \gamma + x \end{bmatrix}. \tag{14.40}$$

В общем случае ненулевое возмущение $(x,\ y) \neq (0,\ 0)$ будет снимать вырожденность собственных значений, так что жорда-

нова каноническая форма возмущенной матрицы диагональна. Собственные значения $\lambda_\pm$ матрицы $M(\gamma; x, y)$ таковы:

$$\lambda_\pm - \gamma = \frac{x}{2} \pm \sqrt{\left(\frac{x}{2}\right)^2 + y}. \tag{14.41}$$

Парабола $(x/2)^2 + y = 0$ образует сепаратрису в плоскости управления $x - y$. В открытой области $(x/2)^2 + y < 0$ собственные значения образуют комплексно-сопряженную пару, в то время как в открытой области $(x/2)^2 + y > 0$ они оба вещественны. На сепаратрисе оба значения равны $x/2 + \gamma$.

Гладкая замена переменных

$$x' = x/2, \quad t = (x/2)^2 + y \tag{14.42}$$

более просто вскрывает действие возмущений на собственные значения и связь с элементарной теорией катастроф. Вещественная часть собственных значений дается формулой

$$\operatorname{Re}(\lambda - \gamma - x') = \begin{cases} 0 & t \leqslant 0, \\ \pm\sqrt{t}, & t \geqslant 0. \end{cases} \tag{14.43}$$

Зависимость от x' тривиальна, зависимость от t каноническая. ◇◇◇ Гладкая замена переменных (14.42) может быть получена посредством преобразования подобия

$$\begin{bmatrix} 0 & 1 \\ y & x \end{bmatrix} \to \begin{bmatrix} \frac{x}{2} & 1 \\ \left(\frac{x}{2}\right)^2 + y & \frac{x}{2} \end{bmatrix} = \begin{bmatrix} x' & 1 \\ t & x' \end{bmatrix}. \tag{14.44}$$

3.2. Катастрофы типа A_3

Рассмотрим жорданову квадратную матрицу порядка три:

$$M_0(\gamma) = \begin{bmatrix} \gamma & 1 & 0 \\ 0 & \gamma & 1 \\ 0 & 0 & \gamma \end{bmatrix}. \tag{14.45}$$

Поскольку матрица (14.39) связана с катастрофой A_2, то естественно ожидать, что (14.45) будет связана с катастрофой A_3. Это каноническое соответствие форма — катастрофа выглядит достаточно просто, и его можно детально проиллюстрировать для матрицы (14.45) (однако для высших катастроф оно становится достаточно сложным).

Канонической формой Жордана — Арнольда для $M_0(\gamma)$ служит

$$M(\gamma;\ p,\ q,\ r)=\begin{bmatrix}\gamma & 1 & 0\\ 0 & \gamma & 1\\ -r & -q & \gamma-p\end{bmatrix}. \tag{14.46}$$

Как было отмечено раньше, характеристическое уравнение для этой матрицы будет кубическим, так что присутствует катастрофа сборки. Для преобразования (14.46) к более удобному виду может быть использовано преобразование подобия

$$\gamma I_3+\begin{bmatrix}-\lambda & 1 & 0\\ 0 & -\lambda & 1\\ -r & -q & -p-\lambda\end{bmatrix}\to\gamma I_3+\begin{bmatrix}c-\lambda & 1 & 0\\ 0 & c-\lambda & 1\\ -b & a & c-\lambda\end{bmatrix}, \tag{14.47}$$

$$\downarrow \text{Характеристическое уравнение} \qquad \downarrow \text{Характеристическое уравнение}$$

$$(\lambda-\gamma)^3+p(\lambda-\gamma)^2+q(\lambda-\gamma)+r=0\to x^3+ax+b=0, \tag{14.48}$$

$$x=\lambda-\gamma-c.$$

Коэффициенты p, q, r и a, b, c связаны между собой посредством формул

$$a=\tfrac{1}{3}(3q-p^2),\quad b=\tfrac{1}{(3)^3}(2p^3-9pq-27r),\quad c=-\tfrac{1}{3}p. \tag{14.49}$$

Итак, мы получили следующее соотношение между каноническими формами:

$$M(\gamma;\ p,\ q,\ r)=M(\gamma+c;\ a,\ b). \tag{14.50}$$

Один из трех параметров ($c=-p/3$) может быть поглощен собственным значением. Этого легко достигнуть посредством переноса центра тяжести собственных значений. Остальные два управляющих параметра определяют способ, которым три собственных значения расщепляются после возмущения. Характеристическое уравнение

$$\det M(\gamma+c-\lambda;\ a,\ b)=0 \tag{14.51}$$

является обычным уравнением состояния многообразия катастрофы сборки. Внутри области, образуемой сборкой $M(\gamma+c;\ a,\ b)$, имеется три вещественных различных собственных значения со средним $\gamma+c$. Вне этой области имеется одно вещественное собственное значение и комплексно-сопряженная пара. Два собственных значения вырождаются на линиях складки. Все три становятся вырожденными при $(a,\ b)=(0,\ 0)$.

3.3. Катастрофы типа A_k

Рассмотрим квадратную жорданову матрицу $M_0(\gamma)$ порядка k и ее универсальное возмущение δM:

$$M_0(\gamma+\delta M = M(\gamma;\ p_1,\ \ldots,\ p_k) = \begin{bmatrix} \gamma & 1 & & & 0 \\ & \gamma & 1 & & \\ & & \ddots & & \\ & & & 1 & \\ & & & \ddots & \\ 0 & & & & \gamma \end{bmatrix} - \begin{bmatrix} & & 0 \\ & & \\ p_k & \ldots & p_1 \end{bmatrix}. \qquad (14.52)$$

Характеристическое уравнение

$$\det M(\gamma-\lambda;\ p_1,\ \ldots,\ p_k) = \sum_{j=0}^{k} p_j(\lambda-\gamma)^{k-j} = 0,\quad p_0 = 1, \qquad (14.53)$$

является полиномиальным уравнением степени k относительно неизвестного $x = \lambda - \gamma$. Всегда можно выбрать новое начало координат на оси x ($x = x' - (1/k)\rho_1$) так, что член $(x')^{k-1}$ будет отсутствовать. Такое изменение начала координат (для функций) выполняется при помощи преобразований подобия для матриц. Полученное характеристическое уравнение будет уравнением состояния $\nabla V = 0$ катастрофы A_k. В результате этого получаем, что множество точек в пространстве управляющих параметров $\mathbb{R}^{k-1}$ для возмущения

$$M(x';\ c_1,\ \ldots,\ c_{k-1}) = \begin{bmatrix} x' & 1 & & & 0 \\ & x' & 1 & & \\ & 0 & & x' & 1 \\ -c_1 & -c_2 & \ldots & -c_{k-1} & x' \end{bmatrix}, \qquad (14.54)$$

у которого вырожденность собственных значений не полностью ликвидирована, соответствует в точности бифуркационному множеству $\mathcal{S}_B$ катастрофы A_k.

4. ВЫРОЖДЕННОСТИ В СЕМЕЙСТВАХ ЛИНЕЙНЫХ ОПЕРАТОРОВ

Семейства линейных операторов, зависящие от управляющих параметров, могут содержать отдельные матрицы с вырожденными собственными значениями. Какие же типы вырожденностей могут встречаться в k-параметрическом семействе

линейных операторов? Ответ на этот вопрос можно получить, используя результаты, изложенные в предыдущих разделах.

Если $k = 1$, характеристическое уравнение в общем случае может иметь лишь дважды вырожденный корень. Наиболее вырожденный член 1-параметрического семейства имеет следующую жорданову каноническую форму:

$$\alpha^2\beta\gamma\ldots\ . \tag{14.55}$$

Если $k = 2$, характеристическое уравнение может иметь один трижды вырожденный корень или два дважды вырожденных корня (остальные корни невырожденные) и жордановы канонические формы имеют вид

$$\alpha^3\beta\gamma\ \ldots\ \text{ и }\ \alpha^2\beta^2\gamma\ \ldots\ . \tag{14.56}$$

Эти формы соответствуют функциям с одной троекратной вырожденностью сборки или двумя двукратными вырожденностями складки.

Если $k = 3$, возможны следующие случаи:

$$\alpha^4,\ \alpha\alpha,\ \alpha^3\beta^2,\ \alpha^2\beta^2\gamma^2. \tag{14.57}$$

Очевидность первого, третьего и четвертого случаев следует из интуитивных соображений пересчета, который, кстати, может быть выполнен достаточно строго. В результате имеем один управляющий параметр для того, чтобы сделать равными два собственных значения, что аналогично ситуации, когда имеется один управляющий параметр для вырождения каждой критической точки или каждого слияния критических значений. Матрица имеет вид

$$\alpha\alpha = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}. \tag{14.58}$$

Покажем, что такие матрицы встречаются при рассмотрении случая трех управляющих параметров.

Диагональная $n \times n$-матрица λI_n обладает n-мерным универсальным возмущением. Размерность возмущения может быть понижена на единицу, если взять возмущенную матрицу, имеющую тот же след, что по существу соответствует переносу центра тяжести мультиплета в нуль (или переносу начала координат), в результате которого из потенциальной функции, описывающей катастрофы A_k (x_i-корни), исключается член $(x_1 + \ldots, \ldots + x_k)x^k$.

Более общо, $\alpha^p\alpha^q$ представляет жорданову $(p+q)\times(p+q)$-матрицу. Верхний жорданов $p \times p$-блок имеет $+1$ на диагонали, расположенной выше главной диагонали. Таков же и нижний жорданов блок. Внедиагональные блоки нулевые. Жорданова

матрица $\alpha^3\alpha^2$ имеет структуру типа (14.3). Матрица $\alpha^2\alpha$ впервые устойчиво встречается в 4-параметрическом семействе, так как $2+3\times 1-1=4$ [ср. с (14.37)]. Значения k, при которых может впервые устойчиво встречаться жорданова матрица вида $\alpha^p\alpha^q\alpha^r\ldots$, приведены в табл. 14.1. Используя эту таблицу, можно составить перечень наихудших возможных вырождений, которые могут типично встречаться в k-параметрическом семействе линейных операторов. Этот перечень зависит от k и не зависит от n размерности пространства состояний, в котором действуют линейные операторы при условии, что сумма вырожденностей не превышает n. Перечень вырожденных жордановых форм приведен в табл. 14.2. Канонические формы Жордана — Арнольда (универсальные возмущения), соответствующие каждой вырожденной матрице (ростку), могут быть легко построены, если следовать формулам (14.34) и (14.36).

Таблица 14.1. **Жордановы блоки вида $\alpha^{n_1}\alpha^{n_2}\alpha^{n_3}=\ldots=\{n_i\}$, которые могут впервые встречаться в k-параметрических семействах матриц [1]**

K	1	2	3	...	7	8	...	11	12	...	15
$\{n_i\}$	2	3	4	...	8	9	...	12	13	...	16
			1, 1	...	5, 1	6, 1	...	9, 1	10, 1	...	13, 1
					2, 2	3, 2	...	6, 2	7, 2	...	10, 2
						1, 1, 1	...	4, 1, 1	5, 1, 1	...	8, 1, 1
								3, 3	4, 3	...	7, 3
									2, 2, 1	...	5, 2, 1
											4, 4
											1, 1, 1, 1

Таблица 14.2. **Наиболее вырожденные матрицы, которые, как правило, могут встречаться в k-параметрических семействах матриц [1]**

k	Жорданова форма
1	α^2
2	α^3, $\alpha^2\beta^2$
3	α^4, $\alpha\alpha$, $\alpha^3\beta^2$; $\alpha^2\beta^2\gamma^2$
4	α^5, $\alpha^2\alpha$, $\alpha^4\beta^2$, $\alpha\alpha\beta^2$, $\alpha^3\beta^2\gamma^2$, $\alpha^3\beta^3$, $\alpha^2\beta^2\gamma^2\delta^2$

5. БИФУРКАЦИОННЫЕ МНОЖЕСТВА КАНОНИЧЕСКИХ ФОРМ ЖОРДАНА — АРНОЛЬДА

Если $M(c^0)$ — вырожденная матрица в k-параметрическом семействе линейных операторов $M(c)$, то типичное возмущение полностью ликвидирует вырожденность собственных значений. Попытаемся выяснить, как расщепление собственных значений связано с возмущением $c^0 \to c^0 + \delta c$, с учетом того факта, что подмножество возмущений меры нуль не будет ликвидировать полностью вырожденность, и какова структура этого подмножества в пространстве управляющих параметров $\mathbb{R}^k$. Поскольку конкретные жордановы «ростки» имеют канонические возмущения, то на эти вопросы можно ответить «канонически». По существу эти вопросы полностью аналогичны таким вопросам элементарной теории катастроф, как: что представляет собой критическое каноническое многообразие и каково каноническое бифуркационное множество?

Начнем с предположения, что однопараметрическое семейство $n \times n$-матриц имеет один дважды вырожденный корень. В этом случае канонической формой Жордана — Арнольда интересующей нас матрицы будет

$$\begin{bmatrix} 0 & 1 \\ t & 0 \end{bmatrix} \tag{14.59}$$

(ее собственные значения и бифуркационное множество приведены на рис. 14.1).

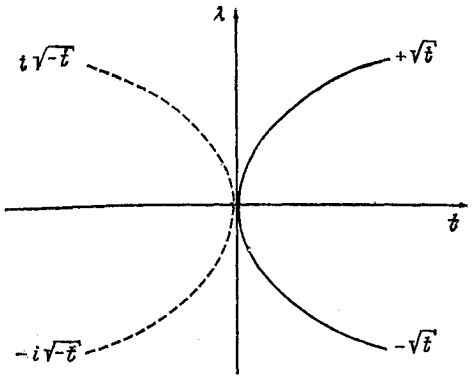


Рис. 14.1. Каноническая зависимость собственных значений возмущенной жордановой 2×2-матрицы от универсального возмущения Арнольда t [ср. с (14.59)].

Сплошные линии — вещественные собственные значения; штриховые линии — мнимые собственные значения. Бифуркационное множество расположено в $t = 0$.

Двухпараметрическое семейство интересующих нас 3×3-матриц Жордана — Арнольда имеет вид

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -b & -a & 0 \end{bmatrix}. \tag{14.60}$$

Собственные значения над плоскостью (a, b) тесно связаны с многообразием катастрофы сборки (при учете добавлений, сделанные выше). Поперечные сечения плоскостями $a = -1, 0, +1$ показаны на рис. 14.2, *б*. Бифуркационное множество в плоскости (a, b) является стандартной сборкой.

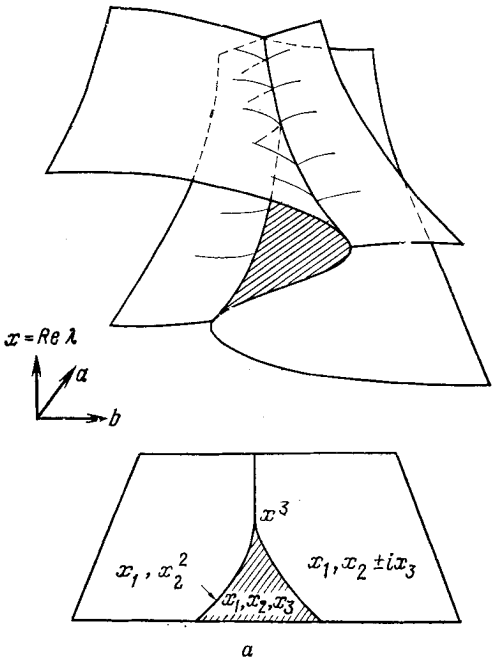
Аналогично могут быть определены бифуркационные множества жордановских ростков с $k = 3$. Жорданова каноническая форма $\alpha^4 \simeq A_4$ и ее бифуркационное множество показаны на рис. 14.3, *а*. Бифуркационное множество жордановой формы $\alpha^3\beta^2$ является прямым произведением бифуркационных множеств, соответствующих катастрофам A_3 и A_2, а поэтому оно выглядит так, как показано на рис. 14.3, *в*. Бифуркационное множество $\alpha^2\beta^2\gamma^2$ является произведением трех прямых $\mathbb{R}^1$ (рис. 14.3, *г*). Не изменяющее следа трехпараметрическое возмущение определяется как

$$\begin{bmatrix} a+z & x+y \\ x-y & a-z \end{bmatrix}. \tag{14.61}$$

Бифуркационным множеством будет конус $x^2 + z^2 - y^2 = 0$ (рис. 14.3, *б*).

◇◇◇ В общем случае бифуркационное множество, соответствующее возмущению жорданова ростка вида $J_1(\alpha)J_2(\beta)\ \ldots$, где $J_1(\alpha) = \alpha^p\alpha^q\ \ldots$, представляет собой прямое произведение бифуркационных множеств, соответствующих жордановым формам для каждого вырожденного собственного значения.

На основании соотношения, существующего между вырожденностью и возмущениями функций и матриц, можно сделать вывод о возможности существования матричных аналогов бифуркационного множества $\mathscr{S}_B$ и множества Максвелла $\mathscr{S}_M$. Как было показано, бифуркационное множество описывает остаточные вырожденности, не полностью уничтоженные возмущениями. Асимптотическая устойчивость линейной системы вида $dx/dt = Mx$ определяется собственным значением, имеющим наибольшую вещественную часть. Если теперь определить множество Максвелла канонической формы Жордана — Арнольда как множество точек пространства управляющих параметров, в которых два или более несопряженных собственных значения имеют



а

Рис. 14.2. Три собственных значения возмущенной жордановой 3×3-матрицы с канонической зависимостью от параметров универсального возмущения a, b.

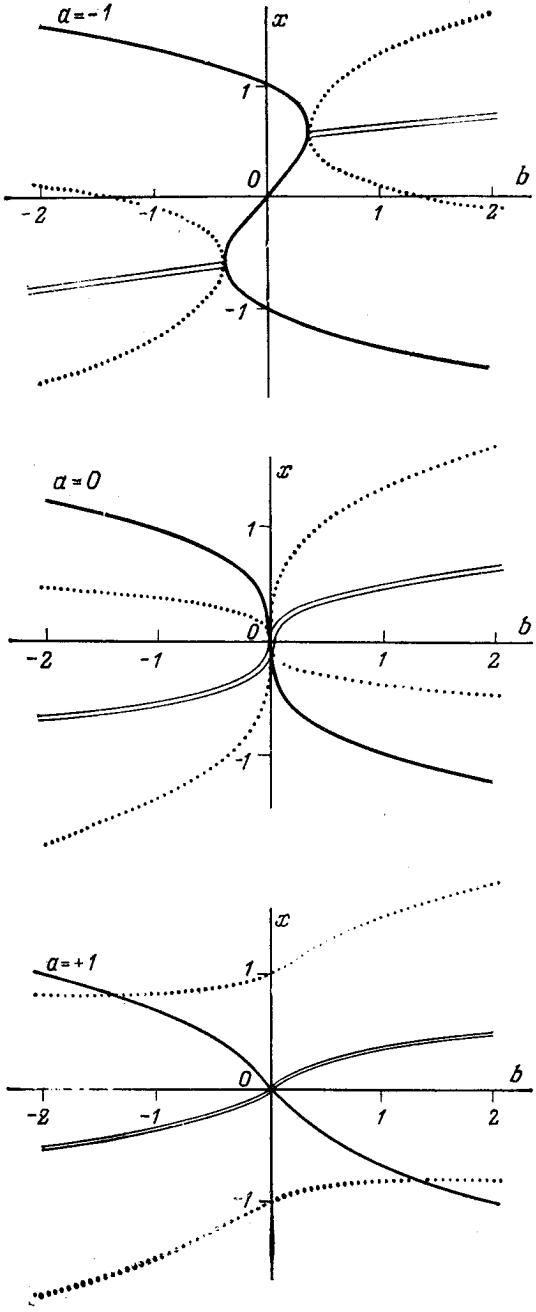
а — поверхность собственных значений над плоскостью управления выглядит подобно многообразию катастрофы сборки с присоединенными крыльями. Показаны лишь действительные части комплексных собственных значений; *б* — собственные значения как функции от b при $a = -1, 0, +1$. Двойные линии указывают равные вещественные части комплексно-сопряженных собственных значений, а расстояние между штриховыми линиями (сверху и снизу) двойными линиями дают значения мнимых частей комплексно-сопряженных собственных значений.

равные вещественные части, то динамическая асимптотическая устойчивость будет иметь место на отдельных компонентах множества Максвелла. Следовательно, множества Максвелла для матриц полностью аналогичны множествам Максвелла для функций.

6. ВЫВОДЫ

Линейные операторы могут быть представлены посредством матриц, действующих на линейных векторных пространствах соответствующей размерности. Две матрицы подобны, а соответствующие операторы эквивалентны, если их характеристические и минимальные многочлены одинаковы.

Возмущение исходного линейного оператора вызывает возмущение коэффициентов характеристического многочлена. Если



б

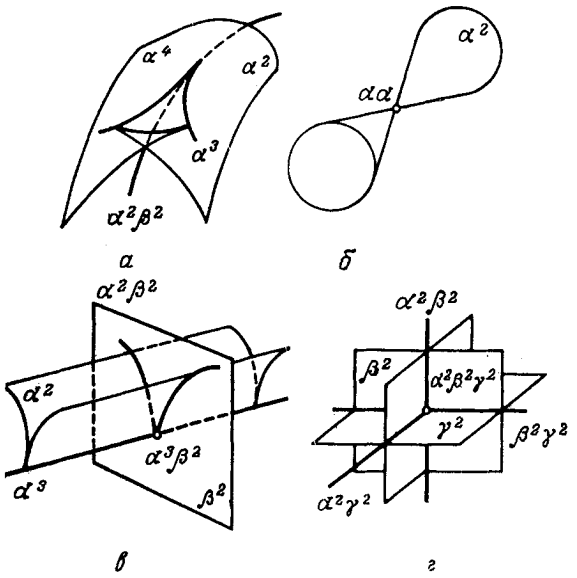


Рис. 14.3. Бифуркационное множество в пространстве управляющих параметров R^3 для канонических форм Жордана — Арнольда, зависящих от трех управляющих параметров [1].

исходная матрица имела невырожденные собственные значения, то возмущенная матрица также будет иметь невырожденные собственные значения. Когда имеет место вырожденность собственных значений, характеристический многочлен может быть записан в виде произведения сомножителей $(\lambda-\alpha)^p \times (\lambda-\beta)^q \ldots$, среди которых есть вырожденные. Возмущение этих отдельных сомножителей может быть проведено почти тем же путем, что и возмущение вырожденных корней градиента потенциала (т. е. уравнения состояния).

Наиболее общее минимальное возмущение матрицы M_0 получается после удаления «внутренних возмущений» и сдвига центра тяжести собственных значений. Наиболее общее возмущение матрицы, приведенной к жордановой форме, будет блочно-диагональная матрица (14.33). Возмущение каждого жорданова блока показано в (14.36). Канонической формой Жордана — Арнольда данной жордановой матрицы является семейство матриц минимальной размерности, которое включает все возмущения исходной жордановой матрицы.

Эти канонические формы дают каноническую информацию о произвольных возмущениях данной жордановой матрицы. Действие возмущения на спектр собственных значений канони-

чески определяется при помощи «многообразий» собственных значений (в действительности многообразий с особенностями), полностью аналогичных критическим многообразиям. Вырожденность собственных значений встречается на различных компонентах бифуркационного множества, которое обычно является прямым произведением бифуркационных множеств, связанных с жордановыми матрицами для различных собственных значений. Самая общая сепаратриса, соответствующая отдельному жорданову блоку, представляется алгебраической поверхностью

$$D = n_1 + 3n_2 + 5n_3 + \ldots - 1 \qquad (14.63)$$

в $\mathbb{C}^D$ или $\mathbb{R}^D$. Значение переменной x выбирается так, чтобы не изменять следа матрицы. Каждая компонента бифуркационного множества параметризует матрицы с вырожденными собственными значениями. Жордановы формы связаны друг с другом точно так же, как ростки катастроф и каустики.

◇◇◇ Эта глава полностью основана на работе В. И. Арнольда [1].

Литература

1. Арнольд В. И. О матрицах, зависящих от параметров. — УМН, 1975, 26 : 2, 101—114.
2. Hoffman K., Kunze R., Linear Algebra (2nd ed.), Englewood Cliffs, N. J.: Prentice-Hall, 1971.

ОГЛАВЛЕНИЕ

Предисловие редакторов перевода 5
Предисловие 7

Часть I. Элементарная теория катастроф

Глава 1. ТЕОРИЯ КАТАСТРОФ КАК ИССЛЕДОВАТЕЛЬСКАЯ ПРОГРАММА 9

Глава 2. ЛОКАЛЬНЫЙ ХАРАКТЕР ПОТЕНЦИАЛЬНЫХ ФУНКЦИЙ 13

1. Форма теоремы о неявной функции 14
2. Морсовские формы 15
3. Формы Тома 16
4. Канонические формы в окрестности критической точки . . . 17
5. Функции катастроф 18
6. Выводы 20

Глава 3. ЗАМЕНА ПЕРЕМЕННЫХ. КАНОНИЧЕСКИЕ ФОРМЫ . . . 21

1. Замена переменных 21
2. Приложения 24
3. Выводы 40

Глава 4. ЗАМЕНА ПЕРЕМЕННЫХ. ВОЗМУЩЕНИЯ 42

1. Возмущения 42
2. Приложения 43
3. Выводы 57

Глава 5. «ПРИНЦИП ЛОМА» 60

1. Потенциальная функция 60
2. Семейство потенциальных функций 64
3. Элементарные катастрофы 67
4. Бифуркационное множество и множество Максвелла 96
5. Выводы 99

Глава 6. ГЕОМЕТРИЯ СКЛАДКИ И СБОРКИ 102

1. Складка 102
2. Сборка 103
3. Двойственная сборка 109
4. Выводы 110

Глава 7. ОРГАНИЗАЦИЯ КАТАСТРОФ 111

1. $A_{k+1} \to A_k$ 112
2. $D_{k+1} \to D_k$ 113
3. Диаграммные представления 115
4. Контурные представления 129
5. Примыкания 131
6. Выводы 135

Глава 8. ПРИНЦИП МАКСИМАЛЬНОГО ПРОМЕДЛЕНИЯ И ПРИНЦИП МАКСВЕЛЛА 137

1. Необходимость принятия соглашений 137
2. Соглашения 138
3. Какие соглашения использовать 141
4. Неадекватность принципов 150
5. Выводы 151

Глава 9. ФЛАГИ КАТАСТРОФ 153

1. Модальность 154
2. Недостижимость 154
3. Катастрофические скачки 154
4. Расходимость 155
5. Гистерезис 156
6. Расходимость линейного отклика 157
7. Критическое замедление/смягчение моды 158
8. Аномальная дисперсия 163
9. Выводы 171

Часть II. Приложения элементарной теории катастроф в науке и технике

Глава 10. ТЕРМОДИНАМИКА 173

1. Общее описание фазовых переходов 173
2. Фазовые переходы второго рода 175
3. Топологические соображения 179
4. Критическая точка жидкости 185
5. Уравнение состояния жидкости 189
6. Уравнение Ван-дер-Ваальса: количественный анализ 191
7. Фазовые переходы первого рода 196
8. Трикритические точки 200
9. Максвелловское множество для катастроф типа A_{+5} 203
10. Тройная точка 206
11. Топология катастроф и термодинамика 208
12. Метрическая геометрия и термодинамика 214
13. Частные производные термодинамических величин 220
14. Возможные вариационные представления 225
15. Флуктуации 231
16. Дополнительные соображения 232
17. Выводы 236

Глава 11. МЕХАНИКА КОНСТРУКЦИЙ 239

1. Системы, описываемые потенциальной функцией 239
2. Выпучивание эйлерова стержня 243
3. Прощелкивание пологой арки 250
4. Смена типа устойчивости 256
5. Составные системы 261
6. Инженерная оптимизация 266
7. Нежелательные последствия инженерной оптимизации . . . 267
8. Поддерживающие кронштейны 268
9. Смягчение моды 277
10. Выводы 279

Глава 12. АЭРОДИНАМИКА 281

1. Описание состояния динамической системы 281
2. Описание состояния усеченной динамической системы 284
3. Линейный анализ устойчивости 286
4. Уравнения движения летательного аппарата 287
5. Многообразия стационарных состояний 291
6. Приложение к конкретному летательному аппарату 292
7. Выводы 303

Глава 13. КАУСТИКИ И ДИФРАКЦИОННЫЕ КАРТИНЫ 306

1. Геометрическая оптика 306
2. Метод стационарной фазы 310
3. Степень сингулярности 312
4. Каустики и ростки катастрофы 316
5. Дифракционные картины и функции катастрофы 319
6. Выводы 325

Глава 14. КАНОНИЧЕСКИЕ ФОРМЫ ЖОРДАНА — АРНОЛЬДА . . 327

1. Исследовательская программа теории катастроф в случае систем линейных уравнений 327
2. Возмущения 330
3. Приложения 336
4. Вырожденности в семействах линейных операторов 339
5. Бифуркационные множества канонических форм Жордана — Арнольда 342
6. Выводы 344