



CONNECTIVITY,
COMPLEXITY, AND
CATASTROPHE IN
LARGE-SCALE SYSTEMS

JOHN CASTI
New York University

A Wiley — Interscience Publication
International Institute for Applied Systems Analysis

JOHN WILEY & SONS
Chichester — New York — Brisbane — Toronto
1979

Дж. Касти

БОЛЬШИЕ СИСТЕМЫ Связность, сложность и катастрофы

Перевод с английского

под редакцией д-ра физ.-мат. наук
Ю. П. ГУПАЛО

и канд. физ.-мат. наук
А. А. ПИОНТКОВСКОГО

ББК 22.17

К 28

УДК 517 : (53 + 57/59)

Касти Дж.

К28 Большие системы. Связность, сложность и катастрофы: Пер. с англ. — М.: Мир, 1982, — 216 с., ил.

В книге американского ученого излагаются принципы построения математических моделей многосвязных систем различного назначения: экономических, экологических, лингвистических и т. д. Значительное внимание уделено применению методов алгебраической топологии при выявлении в системах слабых и сильных связей и исследовании структурной устойчивости без разрушения этих связей.

Для специалистов в области сложных систем управления, а также для студентов, специализирующихся в области автоматизированных систем управления.

К 30501-131
041(01)-82 131-82, ч. 1

1502010000

ББК 22.17
517.8

Редакция литературы по новой технике

Джон Касти

**БОЛЬШИЕ СИСТЕМЫ.
СВЯЗНОСТЬ, СЛОЖНОСТЬ И КАТАСТРОФЫ**

Научный редактор Т. П. Сапожкова. Младший научный редактор Н. Н. Титова
Художник Е. В. Дубровская. Художественный редактор Л. Е. Безрученков
Технический редактор Е. Н. Подчепаева. Корректор Н. А. Гири

ИБ № 2865

Сдано в набор 07.07.81. Подписано к печати 12.03.82. Формат 60×90^{1/16}. Бумага типографская № 1. Гарнитура литературная. Печать высокая. Объем 6,75 бум. л. Усл. печ. л. 13,50, Усл. кр.-отт. 13,72. Уч.-изд. л. 11,96. Изд. № 20/1529. Тираж 10 000 экз. Зак. 1231. Цена 90 коп.

ИЗДАТЕЛЬСТВО «МИР» 129820, Москва, И-110, ГСП, 1-й Рижский пер., 2.

Ленинградская типография № 2 головное предприятие ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» ит. Евгении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли.
198052, г. Ленинград, Л-52, Измайловский проспект, 29

Copyright © 1979 International Institute for Applied Systems Analysis

© Перевод на русский язык, «Мир», 1982

ПРЕДИСЛОВИЕ РЕДАКТОРОВ ПЕРЕВОДА

Для монографий из серии, основанной МИПСА¹⁾, характерны актуальность рассматриваемых проблем и высокий научный уровень. Это неудивительно, поскольку в составлении этой своеобразной энциклопедии современного системного анализа принимают участие ведущие специалисты разных стран. К их числу принадлежит и Дж. Касти, известный советскому читателю как один из авторов книги «Методы погружения в прикладной математике»²⁾.

Основное внимание в предлагаемой читателю книге автор уделяет не столько непосредственно математическому аппарату, сколько искусству математического моделирования поведения сложных систем. Как свидетельствует опыт, накопленный математиками-прикладниками, для овладения этим искусством знаний в области математики далеко не достаточно — необходим еще и навык в переводе исходной формулировки задачи на математический язык. К сожалению, в силу сложившейся традиции этот наиболее существенный и творческий этап моделирования поведения сложных систем обычно остается «за кадром» в учебниках и монографиях, тяготеющих к аксиоматическому законченному изложению соответствующих дисциплин. Данная монография призвана в какой-то мере восполнить подобный пробел.

Первая часть книги (гл. 1, 2) целиком посвящена методологическим принципам моделирования систем. Основоположающим для автора является утверждение, согласно которому при математическом описании достаточно сложной системы неправомерно говорить о модели вообще. Существует ряд моделей, каждая из которых способна дать ответ на вполне определенный круг конкретных вопросов о поведении системы, причем каждая из них имеет свою математическую структуру. Более того, нельзя говорить о сложности системы в каком-то одном единственном значении этого понятия. Система может быть сложной в одном смысле и «простой» в

¹⁾ МИПСА — Международный институт прикладного системного анализа в Вене.

²⁾ Касти Дж., Калаба Р. Методы погружения в прикладной математике: Пер. с англ. — М.: Мир, 1976.

другом. По существу, автор формулирует своего рода «принцип дополнительности» в теории моделирования, убедительно аргументируя справедливость своей точки зрения с помощью примеров из механики, экономики, экологии.

Во второй части книги (гл. 3—5) обсуждаются довольно нетрадиционные для прикладников математические структуры, необходимые для построения моделей, отражающих те или иные стороны многообразной реальности. Эти вопросы изложены достаточно строго и в то же время с учетом интересов читателей, впервые знакомящихся с предметом.

Автор в первую очередь рассматривает проблемы, связанные с внутренней структурой сложных систем, почти не касаясь вопросов внешнего управления и оптимизации поведения таких систем. Между тем современная практика показывает, что центральной в теории сложных систем становится проблема принятия решений при наличии многих целей. По-видимому, действительно сложной следует считать систему, цели поведения которой противоречивы и с трудом поддаются формализации. Кстати, этому вопросу посвящена другая монография названной серии — *Conflicting Objectives in Decisions*, ed. by D. Bell, R. Keeney, H. Raiffa, которая хорошо дополняет книгу Дж. Касти.

Перевод книги выполнен канд. физ.-мат. наук С. П. Чеботаревым (гл. 1, 2, 5) и И. И. Чижиковым (гл. 3, 4).

*Ю. П. Гупало
А. А. Пионтковский*

*Калману, благодаря усилиям
которого теория систем перестала
быть таинством и стала математи-
ческой дисциплиной*

ПРЕДИСЛОВИЕ

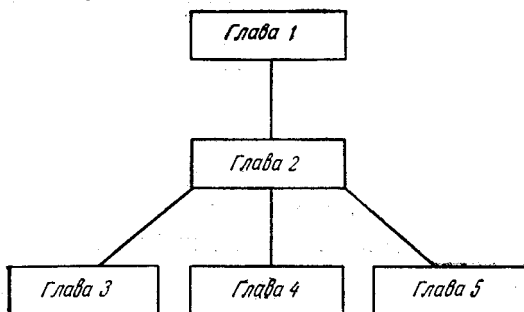
Связность, сложность, катастрофа — эти слова прочно вошли в лексикон современной теории систем. На первый взгляд эти термины удачно охватывают основные особенности процессов, пронизывающих все стороны современной жизни. Часто мы сталкиваемся с выражениями типа «сильно связанная система», «большая, сложная система», «катастрофический исход» и т. д.! Однако при более внимательном рассмотрении эти три понятия напоминают скорее улыбку Чеширского кота, чем сколько-нибудь приемлемую характеристику свойств реальных систем: прежде всего необходимо научиться описывать определенные умозрительные представления, используя четко оговоренную терминологию. До тех пор, пока словам не будет придан определенный смысл в рамках некоторой математической модели какой-то конкретной проблемы, они могут означать все, что угодно. Подобное положение совершенно неприемлемо там, где приходится принимать ответственные решения, например, связанные с формированием политики или управлением.

В этой книге мы пытаемся дать в чем-то упрощенный обзор некоторых современных работ, посвященных математическому анализу понятий связности, сложности и катастрофы. В определенном смысле книгу можно рассматривать, как развернутое общее описание, поскольку мы в основном стремились коснуться множества различных подходов к описанию и анализу этих понятий, а не к глубокому изучению отдельных результатов. На наш взгляд, такая цель не лишена смысла, поскольку сейчас полностью отсутствует единое мнение относительно того, как следует подходить к изучению этих фундаментальных вопросов. Однако многие читатели, по-видимому, захотят глубже познакомиться с затрагиваемыми в книге вопросами. Для них мы приводим списки соответствующих публикаций в конце каждой из глав.

Итак, книгу можно рассматривать как взгляд «с высоты птичьего полета» на некоторые «вершины» теории систем, при котором мелким «деталлям ландшафта» не уделяется особого внимания. Можно только надеяться, что при таком изложении основные идеи станут доступны более широкому кругу читателей. Для того чтобы наша надежда оправдалась, мы приводим множество примеров из самых различных областей, которые по крайней мере помогут читателю понять, что математические методы, доступные лишь посвященным, действительно оказываются полезными, а может быть, и убедят его, что эту книгу действительно стоило написать. Так или иначе, последнее слово за читателем.

Поскольку теория систем является бурно развивающейся областью науки, то совершенно очевидно, что в книге столь ограниченного объема просто невозможно охватить в равной мере все важнейшие результаты, полученные в области исследования связности, сложности структуры больших систем и катастроф в таких системах. На выборе тем и результатов не могли не сказаться интересы автора. Читатель, безусловно, заметит явный уклон в сторону методов, основанных на алгебраических и топологических идеях, а не на классическом анализе. Это не случайно: автор убежден, что будущее теории больших систем в методологическом плане определяется развитием тех областей алгебры и геометрии, которые изучают глобальные свойства математических объектов. Как только станет ясна общая картина, средства локального анализа уже смогут выявить «тонкие эффекты» изучаемого явления.

В общих чертах структура книги может быть представлена в виде следующей схемы:



Первые две главы посвящены описанию методологических основ моделирования систем. В частности, в гл. 1 подробно обсуждается одно из важнейших положений теории моделирования, а именно: принятое математическое описание системы предопределяет круг вопросов, которые можно исследовать в рамках данного описания. Кроме того, приводятся различные возможные варианты описания и обсуждаются их характерные свойства.

Во второй главе дается краткий обзор некоторых важных понятий теории систем, таких, как идентификация, стохастические подходы, ограничения и оптимизация. Следует заметить, что, хотя эти понятия не являются основными, они представляют собой неотъемлемые элементы анализа. Здесь же даются очень краткие вводные представления о связности, сложности и катастрофах как прелюдия к более подробному изучению этих понятий в последующих главах.

В гл. 3 рассматривается одна из основных тем книги — связность. Довольно подробное изложение элементарной теории симплициальной гомологии приводится в качестве основы для описания методологии-анализа Эткина. Эта методика анализа структурной связности далее систематически применяется для исследования ряда различных задач. Вопросы связности обсуждаются также с точки зрения алгебраической теории модулей и конечных автоматов.

Лейтмотивом четвертой главы является сложность. Начиная с обсуждения различных свойств, которыми должно обладать разумное математическое определение сложности, мы приходим к аксиоматическому определению сложности, основанному на теории подгрупп. Это определение поясняется на ряде примеров и сравнивается с возможными определениями, построенными с привлечением теории информации и топологии.

Последняя, пятая глава посвящена подробному обсуждению различных математических аспектов теории устойчивости. После краткого описания таких классических понятий, как устойчивость по Ляпунову, более подробно анализируются идеи, непосредственно связанные с понятием структурной устойчивости. В частности, рассматриваются теория

катастроф и ее связь с теорией бифуркаций. Наконец, несколько разделов посвящены «адаптируемости» динамических систем. Это понятие, введенное вначале для описания живучести экологических систем, по существу характеризует способность системы противостоять воздействию неизвестных внешних возмущений. Подробно обсуждаются различные подходы к математическому описанию этого понятия, и изложение иллюстрируется примером из экономики.

Совершенно ясно, что книгу, охватывающую столь широкий круг вопросов, невозможно было бы написать без помощи и поддержки многих лиц. Кроме того, автору еще и повезло в том смысле, что он имел возможность «обкатать» рукопись на ученых разных специальностей, пытаясь найти приемлемый компромисс между доступностью и строгостью изложения. В этой связи автор особенно признателен Э. Кастри, С. Холлингу, Р. Деннису, Д. Людвигу, К. Мьюзесу, Дж. Лейтману, М. Шекану, Д. Саалу, Дж. Десу и М. Пешелю.

Нью-Йорк
октябрь 1978

Основные понятия и методы системного анализа

Их профессия кажется нам таинственной только потому, что мы не владем ее языком.

Томас Воген

Иногда он грустно думал: «Почему?», а иногда: «По какой причине?», а иногда он думал даже так: «Какой же отсюда следует вывод?». И неудвительно, что порой он вообще переставал понимать, о чем же он, собственно, думает.

А. А. Милн, «Винни-Пух и все-все-все»¹⁾

ПОНЯТИЕ СИСТЕМЫ

В настоящее время понятие *система* широко используется почти во всех областях науки и техники. Однако до сих пор оно все еще не имеет достаточно четкого определения. Почти каждый, кто пытается уточнить, что же следует понимать под системой, прежде всего стремится создать некий умозрительный образ. Некоторым удается дать словесное описание такого образа в виде множества элементов, сложное взаимодействие которых приводит к достижению некой неосознанной цели. В ряде случаев такие индивидуальные умозрительные образы имеют так много общего, что оказывается возможным полное взаимопонимание, без которого немислимы плодотворные дискуссии, не говоря уже о сотрудничестве. Чаше, однако, в силу тех или иных причин такое взаимопонимание не достигается, несмотря на самое неподдельное стремление к общению. Именно этим обстоятельством объясняется тот факт, что одной из главных проблем, которым посвящена данная книга, является создание основы для содержательного обсуждения системно-теоретических вопросов. С этой целью на различных примерах системных задач демонстрируется возможность описания и анализа любой из них с помощью довольно ограниченного набора математических абстракций. Естественно, что при этом возникает необходимость в использовании определенного математического аппарата. Это и неудвительно, поскольку уровень знаний в области биологии, социологии, психологии, экономики (не говоря уже о таких науках, как физика и химия) настолько высок, что накопленные сведения невозможно осмыслить, не обращаясь к абстракции. К счастью, однако, для понимания большинства фундаментальных понятий системного анализа

¹⁾ Перевод Б. Заходера.

вполне достаточно иметь общее представление об обычном дифференциальном исчислении, геометрии и элементарной алгебре. В тех случаях, когда мы будем вынуждены прибегнуть к более сложному математическому аппарату (например, теории полугрупп и теории гомологий), формальное математическое изложение будет сопровождаться соответствующими системно-теоретическими рассуждениями и поясняться с помощью примеров. По нашему мнению, пояснение основных положений с помощью примеров, а не экзотических теорий поможет читателю понять существо дела, не слишком вдаваясь в подробности.

Пожалуй, лучше всего начать изложение материала с рассмотрения некоторых модельных ситуаций, или так называемых *типичных системных задач*. Анализ этих задач позволит выявить некоторые общие системные проблемы, для изучения которых, как будет показано, может быть использовано несколько математических конструкций. При этом неоднократно подчеркивается, что не существует единственной модели данной системы: существует множество моделей, каждая из которых обладает характерными математическими свойствами и каждая из которых пригодна для изучения определенного класса вопросов, связанных со структурой и функционированием системы. Поэтому особенно важно, чтобы исследователь имел в своем распоряжении как можно больше математических методов для анализа принципов построения и работы созданной им модели.

Пример. Макроэкономика

Рассмотрим экономический комплекс, состоящий из n секторов, выпускающих продукцию x_1, x_2, \dots, x_n соответственно. Предположим для определенности, что выпуск продукции измеряется в долларах в год, причем продукция, выпускаемая каждым сектором, используется как самим сектором, так и другими секторами комплекса и внешними потребителями.

Пусть a_{ij} представляет собой часть продукции, выпускаемой i -м сектором, которая необходима для производства единицы продукции j -го сектора ($i, j = 1, 2, \dots, n$). Внешнее потребление продукции, выпускаемой i -м сектором, обозначим через y_i . Тогда можно записать следующее уравнение материального баланса:

$$x_i = \sum_{j=1}^n a_{ij}x_j + y_i, \quad i = 1, 2, \dots, n.$$

Данная элементарная модель может быть использована для определения объема продукции, необходимой для удовлетворения заданного спроса при существующей технологии, которая описывается с помощью коэффициентов a_{ij} . Возможные обобщения и детализация этой модели образуют основу для так называемой модели затраты — выпуск. Матрицу технологических коэффициентов $A = [a_{ij}]$ часто называют лентьевской матрицей.

Пример. Динамика водохранилищ

Упрощенный вариант системы водохранилищ показан на рис. 1.1. Выходами системы являются сток y_1 и доля грунтовых вод y_2 в этом стоке, внешними входами — осадки r_1 и r_2 . Наполнение наземных водохранилищ в момент времени t обозначено через $x_1(t)$, $x_2(t)$ и $x_3(t)$, наполнение подземного

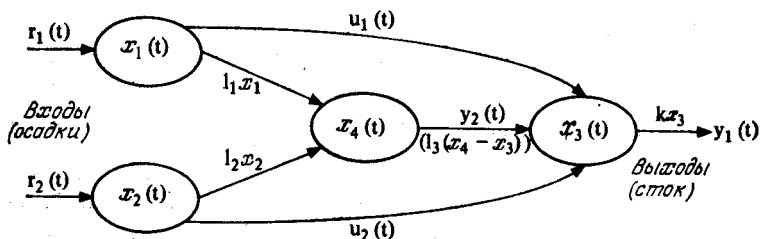


Рис. 1.1. Сеть водохранилищ.

резервуара (с учетом просачивания) — через $x_4(t)$, а попуски воды из водохранилищ — через u_1 и u_2 . Учет связи между поверхностным стоком и грунтовыми водами осуществляется с помощью выражения $l_3(x_4 - x_3)$; коэффициент k характеризует поверхностный сток, а коэффициенты l_1 и l_2 — грунтовый.

Уравнения неразрывности немедленно приводят к следующим динамическим соотношениям:

$$\begin{aligned} x_1(t+1) &= x_1(t) - l_1 x_1(t) - u_1(t) + r_1(t), \\ x_2(t+1) &= x_2(t) - l_2 x_2(t) - u_2(t) + r_2(t), \\ x_3(t+1) &= x_3(t) + l_3(x_4 - x_3) - kx_3(t) + u_1(t) + u_2(t), \\ x_4(t+1) &= x_4(t) + l_1 x_1(t) + l_2 x_2(t) - l_3(x_4 - x_3). \end{aligned}$$

Измеряемые выходы системы имеют вид

$$\begin{aligned} y_1(t) &= kx_3(t), \\ y_2(t) &= l_3(x_4 - x_3). \end{aligned}$$

Приведенное выше описание системы может оказаться полезным при изучении ряда важных вопросов, связанных с управлением паводками, оптимальной стратегией попусков (водосбросов), точным определением уровня грунтовых вод и т. д.

Пример. Система хищник — жертва

Одной из наиболее популярных проблем науки о живой природе является исследование взаимодействия сообществ хищники — жертвы в некоторой ограниченной среде обитания.

Рассмотрим для простоты экосистему с одним трофическим уровнем, в которой хищники и жертвы разделяются на два непересекающихся множества. Пусть множество хищников состоит из следующих элементов:

$Y = \{\text{люди, львы, слоны, птицы, рыбы, лошади}\}$,

а множество жертв —

$X = \{\text{антилопы, зерно, кабаны, скот, трава, листья, насекомые, рептилии}\}$.

Определение точных количественных динамических связей, существующих между хищниками и жертвами, является довольно сложной задачей. Как правило, с уверенностью можно утверждать только, что определенные хищники выбирают вполне определенные жертвы. В подобной ситуации описание системы в терминах отношения инцидентности может дать совершенно неожиданную информацию о фундаментальной структуре экосистемы.

Определим отношение λ между множествами X и Y следующим образом:

Отношение λ существует между хищником y и жертвой x тогда и только тогда, когда хищник y поедает жертву x . Отношение λ удобно описывать с помощью матрицы инцидентий Λ :

		X							
λ		Антилопы	Зерно	Кабаны	Скот	Трава	Листья	Насекомые	Рептилии
Y	Люди	1	1	1	1	0	0	0	0
	Львы	1	0	1	0	0	0	0	0
	Слоны	0	0	0	0	1	1	0	0
	Птицы	0	1	0	0	1	0	1	1
	Рыбы	0	0	0	0	0	0	1	0
	Лошади	0	1	0	0	1	0	0	0

причем, если хищник y поедает жертву x , то $\lambda = 1$, в противном случае $\lambda = 0$. Анализируя матрицу инцидентий Λ ,

можно выявить некоторые совершенно неочевидные структурные свойства системы *хищник — жертва*.

Таким образом, даже в отсутствие очевидных динамических уравнений оказывается возможным построить содержательное (и плодотворное) математическое описание изучаемой системы.

Пример. Двоичный выбор

При анализе многих системных задач, представляющих практический интерес, разумно предполагать, что система стремится минимизировать некоторую (быть может, неизвестную) потенциальную функцию. Это означает, что в отсутствие внешних возмущений система стремится к состоянию равновесия, которому соответствует минимум энергии некоторого *силового поля*, причем природа этого поля может быть различной.

Для иллюстрации этого положения рассмотрим случай, когда возможны два варианта выбора в зависимости от значений некоторой *функции полезности* $U(x, a, b)$, где x — переменная, описывающая выбор; a и b — параметры, от которых этот выбор зависит. Тогда можно определить *функцию бесполезности* как $E(x, a, b) = -U$ и построить модель, в которой эта функция минимизируется.

Допустим, что между двумя пунктами возможны маршруты A и B , стоимость которых C_A и C_B соответственно. Внешние параметры a и b являются функциями разности стоимостей $C = C_B - C_A$. Предположим, что $x < 0$ соответствует маршруту A , а $x > 0$ — маршруту B . Тогда можно построить функции $a(C)$ и $b(C)$, такие, что найдется такое число λ , что:

Если $C > 0$ и велико по модулю, то возможен выбор только маршрута A и, следовательно, $x < 0$;

Если $C < 0$ и велико по модулю, то возможен выбор только маршрута B и, следовательно, $x > 0$;

Если $0 < C < \lambda$, то наиболее вероятным является выбор маршрута A , хотя возможен выбор и маршрута B ;

Если $-\lambda < C < 0$, то наиболее вероятным является выбор маршрута B , хотя возможен выбор и маршрута A ;

Если $C = 0$, то вероятности выбора каждого маршрута одинаковы.

Для построения модели процесса выбора нам потребовалась всего лишь функция бесполезности. Другими словами, мы не испытывали необходимости в более подробном описании внутренней динамики процесса (которого для большинства социально-экономических систем просто нет). Более

того, нам не нужно даже знать точного вида функции $E(x, a, b)$. Единственно, что требуется, — это наша готовность признать сам факт существования такой функции, а все остальное следует из абстрактных математических рассуждений и имеющихся численных данных (включая и точный вид кривой, представленной на рис. 1.2, поскольку это

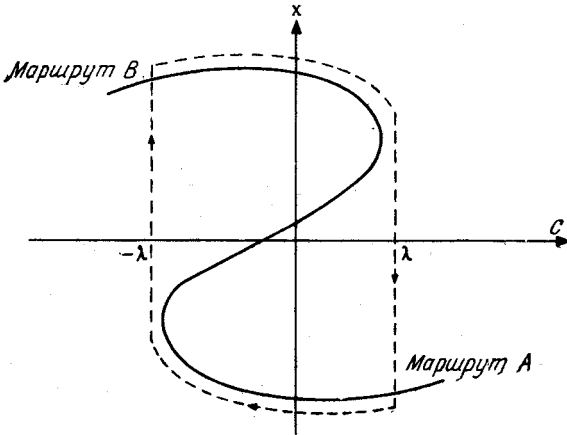


Рис. 1.2. Двойной выбор маршрутов.

необходимо для количественного моделирования данной системы). Как будет показано, для моделирования подобных ситуаций используется теория катастроф Тома.

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ

Рассмотренные примеры свидетельствуют о том, что абстрактная характеристика данной ситуации может быть получена с помощью разных типов математического описания. Однако при этом естественно возникает вопрос: а для чего вообще нужно какое-либо математическое описание? Ответ на этот вопрос в значительной степени связан с нетривиальностью современных научных результатов и необходимостью уметь выделять существенные свойства описательных моделей. Кроме того, использование именно математического описания обусловлено следующими важными соображениями.

Компактность. Словесное (или вербальное) описание системы (или процесса), как правило, представляет собой нагромождение нечетких высказываний, которые лишь затуманивают существо дела. Избавиться от таких нечетких и недо конца продуманных соображений помогает компактная

математическая символика. Математическое описание дает нам аналог знакомой картины и оказывается информативнее любого словесного описания.

Ясность. Использование математического описания позволяет каждому аспекту изучаемого процесса поставить в соответствие определенный математический символ, в результате чего становится более наглядной взаимосвязь, существующая между различными параметрами процесса. Более того, подобное сопоставление позволяет гораздо проще, чем словесное описание, установить, не были ли упущены какие-либо существенные переменные, или, напротив, не были ли внесены какие-либо дополнительные несуществующие сложности при построении описания.

Возможность численного анализа. Как только сделан выбор некоторого математического описания, последнее «начинает жить» собственной жизнью, более или менее независимой от самого исследуемого процесса. Другими словами, математическим описанием можно манипулировать в соответствии с обычными законами логики в надежде получить нетривиальное представление о самой системе. Кроме того, математическая модель дает основу для численного анализа, с помощью которого могут быть получены данные не только описательного, но и прогностического характера.

Рассмотрим кратко некоторые типы математического описания, которые мы будем использовать в этой книге.

Внутреннее описание

Со времен Ньютона динамические процессы описывали на языке дифференциальных (или разностных) уравнений, т. е. в терминах некоторых естественно выбранных переменных, таких как положение, температура, скорость и т. д. В общем виде такое описание может быть представлено как

$$\dot{x} = f[x(t), u(t), t], \quad x(0) = x_0, \quad y(t) = h[x(t), u(t), t],$$

где $x(t)$ — n -мерный вектор, компоненты которого описывают состояние системы в момент времени t , $y(t)$ — p -мерный вектор наблюдаемых выходов системы, $u(t)$ — m -мерный вектор входов системы и x_0 — начальное состояние системы. В дискретном времени динамика системы может быть описана с помощью разностных соотношений

$$\begin{aligned} x(k+1) &= F[x(k), u(k), k], & x(0) &= x_0, \\ y(k) &= H[x(k), u(k), k]. \end{aligned}$$

Наиболее важным свойством такого описания является то, что оно дает нам представление о поведении системы в

некоторой локальной окрестности текущего состояния. При этом неявно предполагается, что локальная информация может быть каким-то образом «собрана воедино», что позволит понять глобальное (во времени или пространстве) поведение системы. Такой подход оказался достаточно обоснованным для анализа многих физических и технических задач. Однако возможность его использования в случае менее изученных объектов, в особенности систем социально-экономической природы, вовсе не очевидна.

Интересно отметить, что математическое описание указанного типа начали использовать только со времен Ньютона. До этого при описании физических процессов придерживались точки зрения, высказанной Аристотелем, согласно которой *важность целого превышает важность его составляющих*. Другими словами, значимость элементов, составляющих некоторое множество, трактовалась через значимость самого множества (как целого). Взгляды Аристотеля господствовали в физике на протяжении многих столетий, пока Галилей не высказал иную точку зрения, которая впоследствии была обоснована Ньютоном: *целое объясняется свойствами его элементарных (локальных) составляющих*¹⁾.

Простые примеры локального описания можно найти в элементарной физике. Известно, например, что колебательное движение груза (маятника) единичной массы, подвешенного на нерастяжимой и невесомой нити единичной длины, описывается уравнением

$$\ddot{x} + a\dot{x} + \sin x = u(t), \quad (1.1)$$

где a — коэффициент трения, $u(t)$ — внешняя сила, действующая на груз, а $x(t)$ — отклонение груза от положения равновесия. Таким образом, уравнение (1.1) описывает мгновенное изменение положения и скорости маятника как функцию его текущего состояния (положения) и скорости, т. е. мы имеем локальное описание в координатах *положение — скорость*, что характерно для всех описаний динамических процессов на языке дифференциальных или разностных уравнений.

Внешнее описание

Тип математического описания, с которым чаще всего приходится иметь дело ученому-экспериментатору, — это связь *вход — выход*. Во многих отношениях такое описание диаметрально противоположно частному, локальному описа-

¹⁾ Сложность современной жизни, проявляющаяся в политике, экономике и социологии, стимулирует возрождение интереса к холистским теориям.

нию, поскольку оно не содержит деталей и единственным доступным источником информации является закономерность (отображение), связывающая выходы системы с ее входами. При этом ничего не известно о *внутреннем механизме* преобразования входов в выходы. По этой причине связь вход — выход часто называют *внешним описанием* системы в отличие от *внутреннего*, или локального, описания (рис. 1.3).

Внутренние и внешние описания позволяют рассматривать систему как устройство, преобразующее входы в выходы в соответствии с правилами, определенными внутренним описанием. Иными словами, система Σ является информационным процессором в некотором обобщенном смысле.

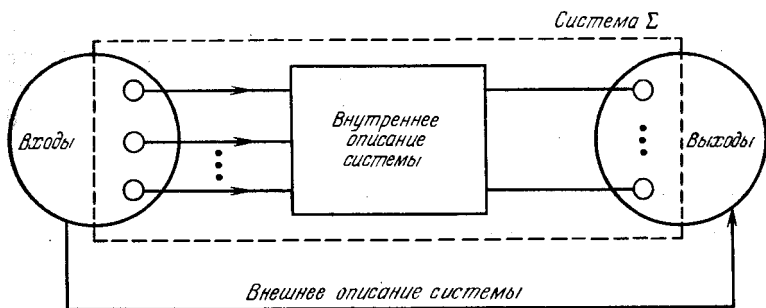


Рис. 1.3. Внутреннее и внешнее описание системы Σ .

Очевидно, что внутреннее описание говорит нам гораздо больше о способе действия системы, поскольку каждое такое описание порождает внешнее описание. Тем не менее построение модели связано с решением диаметрально противоположного вопроса: может ли внутренняя модель «объяснить» каждое внешнее описание? Ответом на этот вопрос по существу является решение так называемой *задачи реализации*, которая представляет собой один из важных аспектов математической теории систем.

Наиболее «сырая» возможная ситуация, при которой возникает необходимость в описании типа *вход — выход*, имеет место, когда мы располагаем всего лишь таблицей элементов (часто чисел), характеризующих *реакцию (выход)* системы на различные *внешние воздействия (входы)*. В этом случае внешнее описание системы эквивалентно отображению


$$f: \Omega \rightarrow \Gamma,$$

где через Ω обозначено множество возможных входов, а через Γ — множество возможных выходов системы. Как отмечалось,

во многих задачах (в частности, психологии, экономики и общественных наук) множества Γ и Ω представляют собой конечный набор элементов, связь между которыми описывается с помощью функции f .

Пример

Представим себе, что из летающего объекта вывалился загадочный ящик. Предположим, что эксперт, изучающий этот ящик, не имеет ни малейшего представления ни о его природе, ни о его содержимом. Вместе с тем эксперт может производить над ним некоторые действия (*входы*) и наблюдать их результаты (*выходы*). Предположим для определенности, что элементами множества Ω и множества Γ являются показания различных измерительных приборов. Тогда описание эксперимента типа *вход* — *выход* могло бы быть, например, таким:

Время	Вход	Выход
10.05	Эксперт не производит никаких действий	Ящик издает звуковой сигнал частотой 240 Гц
10.06	Эксперт нажал на кнопку «Х»	Частота тона возросла до 480 Гц
10.07	Эксперт случайно нажал на кнопку 	Ящик нагрелся на 20 °С и начал вибрировать

Этот довольно тривиальный пример показывает, что входы и выходы системы являются функциями времени, т. е. нельзя один и тот же эксперимент провести дважды! Единственное, что можно сделать, — это провести следующий эксперимент, который хотя и незначительно, но будет отличаться от предыдущего.

Менее тривиальный пример внешнего описания системы дает «бихевиористская» школа психологов, для которой характерным является проведение эксперимента и запись его результатов в формате *воздействие* — *реакция*. По мнению представителей этой школы, такое внешнее описание системы дает максимум информации, которую вообще можно получить о ее структуре и функционировании. В то же время

«познавательная» школа придерживается другой точки зрения, утверждая, что единственным удовлетворительным описанием системы может быть только *внутренняя модель*.

Основываясь на довольно общих результатах теории систем, можно показать, что это спор ни о чем. Обе школы, в сущности, утверждают одно и то же, и с точки зрения теории систем эти дебаты столь же содержательны, как и дискуссии относительно того, какая сторона монеты наиболее полно отражает ее стоимость.

Системы с конечным числом состояний

В тех случаях, когда предположение о *конечномерности* пространства состояний заменяется предположением о *конечности* числа его элементов, мы имеем дело с классом систем, анализ которых возможен с помощью чисто алгебраических методов. Важность такой замены трудно переоценить, поскольку совокупность систем с конечным числом состояний включает все последовательные цифровые вычислительные машины.

Математическое описание системы Σ с конечным числом состояний включает

- множество допустимых входов U ,
- множество допустимых выходов Y ,
- множество состояний Q ,
- функцию перехода $\lambda: Q \times U \rightarrow Q$,
- функцию выхода $\gamma: Q \times U \rightarrow Y$.

При этом предполагается, что множества U , Y и Q конечны. Это позволяет представить описание системы Σ в виде $\Sigma = (U, Y, Q, \lambda, \gamma)^1$.

Как отмечалось, ограничения вычислительного характера с неизбежностью вынуждают нас явно или неявно сводить каждую системную задачу к виду, указанному выше. Поэтому необходимы тщательное изучение и понимание алгебраической структуры подобных «конечных» описаний, которая основывается на теории конечных полугрупп. Хотя рассмотрение этой теории выходит за рамки нашей книги «для первого чтения», тем не менее некоторые ее простейшие понятия и методы представлены в главе, посвященной сложности.

¹⁾ В литературе такое представление часто называют *схематическим*. При этом машина характеризуется внешней функцией вход—выход $f: PU \rightarrow Y$, где $PU = \{(u_1, u_2, \dots, u_n) : n \geq 1 \text{ и } u_j \in U\}$. Здесь $f(u_1, u_2, \dots, u_n) = Y_n$ понимается как выход в момент времени n , если u_j —вход в момент времени j , $1 \leq j \leq n$.

Пример

Пусть система Σ состоит из симметрий вращения правильного треугольника. Тогда некоторые возможные конечные пространства состояний могут иметь вид

$$\begin{aligned} \begin{array}{c} a \\ \triangle \\ b \quad c \end{array} &= [a, b, c], & Q_1 &= 0, & Q_2 &= 0, & Q_3 &= 0, \\ \begin{array}{c} c \\ \triangle \\ a \quad b \end{array} &= [c, a, b], & &= 2\pi/3, & &= 1, & &= 1, \\ \begin{array}{c} b \\ \triangle \\ c \quad a \end{array} &= [b, c, a], & &= 4\pi/3, & &= 2, & &= 2. \end{aligned}$$

Для описания системы Σ достаточно любого из этих пространств состояний, однако некоторые из них, по-видимому, удобнее использовать для вычислений результата воздействия λ на состояние системы. Следовательно, пространство состояний вовсе не обязательно должно быть непосредственно привязано к реальному физическому процессу. Это чисто математическая условность, введенная для упрощения проблемы определения реакции системы на заданные внешние воздействия.

Пусть имеются два возможных отображения λ_1 и λ_2 , переводящие одно состояние системы в другое и соответствующие повороту треугольника вокруг центра тяжести на 120° и 240° соответственно. Результаты применения этих отображений к различным пространствам состояний можно представить в виде таблицы

	q	$\lambda_1(q)$	$\lambda_2(q)$
$Q_1:$	$[a, b, c]$	$[c, a, b]$	$[b, c, a]$
	$[c, a, b]$	$[b, c, a]$	$[a, b, c]$
	$[b, c, a]$	$[a, b, c]$	$[c, a, b]$
$Q_2:$	0	$2\pi/3$	$4\pi/3$
	$2\pi/3$	$4\pi/3$	0
	$4\pi/3$	0	$2\pi/3$
$Q_3:$	0	1	2
	1	2	0
	2	0	1

или

$$\begin{array}{ll}
 Q_1 \rightarrow Q_1; & \lambda_1(\alpha, \beta, \gamma) = (\gamma, \alpha, \beta) \\
 \lambda_1: Q_2 \rightarrow Q_2; & \lambda_1(q) = q + 2\pi/3 \pmod{2\pi} \\
 Q_3 \rightarrow Q_3; & \lambda_1(q) = q + 1 \pmod{3}. \\
 \\
 Q_1 \rightarrow Q_1; & \lambda_2(\alpha, \beta, \gamma) = (\beta, \gamma, \alpha) \\
 \lambda_2: Q_2 \rightarrow Q_2; & \lambda_2(q) = q + 4\pi/3 \pmod{2\pi} \\
 Q_3 \rightarrow Q_3; & \lambda_2(q) = q + 2 \pmod{3}.
 \end{array}$$

Пространство Q_1 , на первый взгляд излишне сложное, оказывается вполне оправданным для более сложных систем Σ , например в случае симметрий более общего вида, где могут присутствовать отображения типа отображений λ_3 . В то же время пространства Q_2 и Q_3 не допускают очевидных обобщений на более сложные случаи с сохранением простоты вычислений.

Способ выбора конкретного пространства состояний системы называют ее *координатизацией*. При этом важно установить, всегда ли существует координатизация, которую можно считать «хорошей» с точки зрения описания поведения системы¹⁾. Ключевым моментом проблемы координатизации является алгебраическая структура модели системы в пространстве состояний. Действительно, согласно известной теореме декомпозиции Крона — Роудза, для конечных полугрупп существует связь между произвольными преобразованиями на конечном пространстве состояний и определенными удобными способами координатизации.

Энтропия и потенциальная функция

При изучении системы с более «целеустремленной», или информационно-теоретической, точки зрения описание системы дается на языке энтропии и потенциальных функций. По аналогии с классической механикой и теорией поля можно рассматривать реакцию системы на внешнее воздействие как динамическое изменение состояния системы, в процессе которого она стремится минимизировать некоторую потенциальную функцию. В зависимости от конкретного вида системы и принятых допущений такая динамика может быть *локальной* в смысле движения системы к относительному минимуму, ближайшему к текущему состоянию, или *глобальной*

¹⁾ В дальнейшем мы уточним, что следует понимать под хорошей координатизацией, и покажем, что такая координатизация возможна для всех систем, имеющих конечное число конфигураций по отношению к переходам в пространстве состояний.

в смысле движения к абсолютному (глобальному) минимуму соответствующей потенциальной функции (рис. 1.4).

Приближенно описание динамического процесса на языке потенциальных функций включает следующие составляющие:

- пространство состояний (фазовое пространство) X ,
- набор входных функций Ω ,
- гладкое отображение $f: X \times \Omega \rightarrow R$,

где R есть пространство действительных чисел. При этом предполагается, что система ведет себя так, что при фикс-

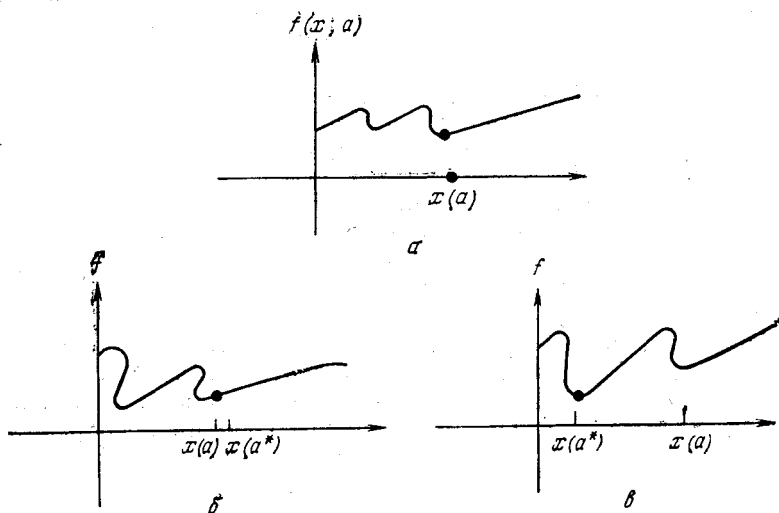


Рис. 1.4. Описание системы с помощью потенциальной функции.

a — основная идея; b — движение к локальному минимуму; v — движение к глобальному минимуму; $x(a)$ — начальное положение системы, где a — внешний параметр; $f(x, a)$ — потенциальная функция. Замена параметра a на a^* приводит к изменению положения минимума функции f .

рованном входе $\omega \in \Omega$ ее наблюдаемое состояние соответствует локальному минимуму функции f .

Использование потенциальных функций для описания хорошо изученных физических систем оказалось весьма удачной альтернативой внутренним описаниям. Успешное применение такого подхода в классической физике обусловлено существованием незыблемых вариационных принципов, таких как принципы Гамильтона, Ферма и Даламбера. В большинстве случаев внутреннее описание физического процесса на языке потенциальных функций естественным образом вытекает из описания с помощью потенциальных функций в силу уравнений Гамильтона — Якоби и Эйлера — Лагранжа.

В системах, которые являются предметом изучения общественных наук, возможность использования подобного описания не столь обоснована из-за сложности применения вариационных принципов. Однако в ряде случаев при анализе устойчивости или в теории катастроф знание точного вида потенциальной функции не является необходимым для определения важных качественных свойств системы — важен лишь сам факт ее существования (гл. 5).

С описанием системы на языке потенциальных функций тесно связана идея описания поведения систем с помощью энтропии. Как известно из классической термодинамики, энтропия является мерой беспорядка, существующего в данной физической системе. Мерой упорядоченности системы является отрицательная энтропия, или *негэнтропия*. В основе описания динамического процесса с помощью энтропии лежит предположение о преобразовании системой негэнтропии входа в информацию. Это означает, что *все замкнутые системы изменяются таким образом, что минимизируют изменение энтропии*. Таким образом, становится очевидной связь между описаниями на языке потенциальных функций и энтропии.

Чтобы показать общность описаний в терминах энтропии, перечислим основные аксиомы релятивистской теории информации, развитой Джюмэри для динамических процессов.

Аксиома 1. Система Σ является частью некоторой вселенной \mathcal{U} и развивается только постольку, поскольку она преследует некоторую цель v .

Аксиома 2. Для достижения цели v система Σ воспринимает информацию I из окружающей среды и использует эту информацию для перестройки собственной организации (внутренней структуры) a , в результате которой увеличилась бы негэнтропия n , и для оказания воздействия A на окружающую среду.

Аксиома 3. (Принцип эволюции.) Структурная энтропия E системы Σ определяется соотношением

$$dE = dI/n$$

и является неубывающей функцией эволюции Σ .

Аксиома 4. Вселенная \mathcal{U} не может наблюдать собственную эволюцию.

В силу этих аксиом уравнение состояния системы имеет вид

$$f(H_e, H_i, v) = 0,$$

где H_e — внешняя энтропия системы Σ по отношению к фиксированному наблюдателю R , H_i — внутренняя энтропия

системы Σ по отношению к наблюдателю R , v — цель системы Σ с точки зрения наблюдателя R . При таком подходе к описанию системы наблюдатель (или лицо, принимающее решения) играет особую роль, причем особый упор делается на кинематический подход, основанный на аналогах преобразования Лоренца для двух наблюдателей R и R' .

Анализируя уравнение состояния, можно заметить, что знание функции f позволяет вычислить структурную энтропию E системы с помощью соотношения

$$dI = \alpha dH_e + \beta dH_i,$$

описывающего обмен информацией, где α и β — некоторые постоянные.

Пример А. Одномерная динамика

Рассмотрим простую динамическую систему

$$\dot{x} = u(t),$$

где x и u — скалярные функции. Поскольку внешняя энтропия H_e обладает теми же свойствами, что и время t , произведем замену: $t \rightarrow H_e$. Более того, имеет смысл отождествить внутреннее состояние x с внутренней энтропией H_i . Тогда динамика системы эквивалентным образом описывается уравнением

$$dH_i - u(H_e) dH_e = 0.$$

Попытаемся теперь построить функцию состояния f в соответствии с приведенным выше ее определением. Из уравнения состояния следует, что

$$\frac{\partial f}{\partial H_e} dH_e + \frac{\partial f}{\partial H_i} dH_i + \frac{\partial f}{\partial v} dv = 0.$$

Не имея дополнительной информации о системе, можно предположить, что ее цель не меняется, и, следовательно, $dv = 0$. Интегрируя уравнение динамики, получаем

$$f(H_i, H_e, v) = H_i - \int_{H_e^0}^{H_e} u(s) ds = 0,$$

где H_e^0 — внешняя энтропия в начальный момент времени t_0 .

Проведенный анализ показывает, что система $\dot{x} = u$ не определена с точки зрения обмена информацией с окружающей средой, более того, такой обмен вообще не имеет места.

Пример Б. Стационарная динамика

Рассмотрим систему, описываемую уравнением

$$\dot{x} = \phi(x(t)),$$

которое способом, аналогичным рассмотренному в примере А, можно привести к виду

$$dH_i - \phi(H_i) dH_e = 0.$$

Чтобы получить уравнение состояния, следует записать

$$\frac{\partial f}{\partial H_i} = 1 \Rightarrow f = H_i + \chi(H_e),$$

$$\frac{\partial f}{\partial H_e} = \phi(H_i) \Rightarrow \phi(H_i) = \chi(H_e).$$

Однако эти соотношения противоречивы и уравнение динамики следует рассматривать не как уравнение состояния, а как уравнение обмена информацией

$$dI = dH_i - \phi(H_i) dH_e = 0.$$

Следовательно, система не обменивается информацией с окружающей средой и развивается с постоянной структурной энтропией, что находится в соответствии с автономным характером системы.

Остановимся теперь на релятивистском характере таких описаний. Как уже отмечалось, значения переменных H_e , H_i и v зависят от наблюдателя R . При такой интерпретации немедленно возникает вопрос о том, каковы значения этих величин относительно другого наблюдателя R' ? При заданных значениях указанных переменных для наблюдателя R из теории относительности следует кинематический подход к решению этого вопроса в римановом пространстве с геодезической

$$d\sigma^2(\Sigma | R) = c^2 dH_e^2(\Sigma | R) - dH_i^2(\Sigma | R) - dv^2(\Sigma | R),$$

где универсальная постоянная c определяется равенством

$$H_i(\mathcal{U} | \mathcal{U}) = cH_e(\mathcal{U} | \mathcal{U}).$$

Преобразования Лоренца для R и R' имеют вид

$$\begin{aligned} H_i(\Sigma | R') &= \rho [H_i(\Sigma | R) + u(R | R') H_e(\Sigma | R)], \\ v(\Sigma | R') &= v(\Sigma | R), \\ H_e(\Sigma | R') &= \rho \left[H_e(\Sigma | R) + \frac{u(R | R')}{c^2} H_i(\Sigma | R) \right], \\ \rho &= [1 - u^2(R | R')/c^2]^{-1/2}, \end{aligned}$$

где

$$u(R | R') = dH_i(R | R')/dH_e(R | R'),$$

т. е. $u(R | R')$ представляет собой *организованность* наблюдателя R с точки зрения наблюдателя R' .

В целом можно сказать, что «энтропийный» подход к анализу систем основан на трактовке системы, как некоторого единого целого. Отсюда следует, что понять сущность системы можно, лишь изучая ее взаимодействие с окружающей средой, т. е. с некой «вселенной». Взгляд на систему как на единое целое можно развить, введя понятие *связь*. Весь комплекс связей и их характеристик приводит к понятиям *структура* и *сложность* системы. Рассмотрим тип описания систем, который оказывается особенно эффективным при таких структурных исследованиях.

Множества и отношения

Принято считать, что математическими абстракциями в основном оперирует теория множеств и отношений между их элементами. Поэтому целесообразно попытаться определить понятие системы в терминах этой теории. От конструктивного определения, естественно, можно потребовать, чтобы элементы соответствующих множеств и связывающие их отношения определялись спецификой конкретной системы. Тем не менее, если мы построим даже такое «специализированное» описание системы, оно даст весьма широкие возможности для анализа не только структуры системы, но и ее поведения в динамике.

В общем случае можно предположить, что существуют два конечных множества X и Y , элементы которых как-то связаны с системой Σ . Это могут быть множества хищников и их жертв, множества типов автомобилей и дорог или множества предприятий службы быта и предлагаемых услуг. Для описания связи, существующей между двумя элементами (x, y) , $x \in X$, $y \in Y$, введем на прямом произведении X и Y бинарное отношение $\lambda \in X \times Y$.

Рассмотрим тривиальный пример, в котором X есть множество товаров, а Y — множество предприятий службы быта. Пусть для определенности

$$X = \{\text{хлеб, молоко, марки, обувь}\} \\ = \{x_1, x_2, x_3, x_4\}, \\ Y = \{\text{гастроном, универмаг, банк, почта}\} \\ = \{y_1, y_2, y_3, y_4\}.$$

Определим отношение λ на $X \times Y$ следующим образом:

Отношение λ существует между x_i и y_j тогда и только тогда, когда x_i можно купить в y_j .

В этом случае

$$\lambda = \{(x_1, y_1), (x_2, y_1), (x_3, y_4), (x_4, y_2)\}.$$

Отношение λ удобно представить матрицей инциденций

	λ	y_1	y_2	y_3	y_4
\in	x_1	1	0	0	0
\in	x_2	1	0	0	0
\in	x_3	0	0	0	1
\in	x_4	0	1	0	0

причем

$$[\Lambda]_{ij} = \begin{cases} 1, & (x_i, y_j) \in \lambda, \\ 0 & \text{в противном случае.} \end{cases}$$

С геометрической точки зрения отношение λ определяет симплициальный комплекс $K_X(Y; \lambda)$, в котором элементы множества Y рассматриваются как *вершины*, а элементы множества X являются *симплексами*. Так, элемент x_1 (хлеб) является 0-симплексом, состоящим из вершины y_1 (гастроном). Если K не содержит r -симплексов ($r \geq 3$), его можно изобразить на плоскости. Для предыдущего примера множество K имеет вид

$$y_1 \quad y_2 \quad y_4.$$

Хотя такая геометрическая структура не представляет особого интереса, тем не менее она все же показывает, что комплекс не содержит связанных компонент и что вершина y_3 (банк) не играет никакой роли в анализе $K_X(Y; \lambda)$.

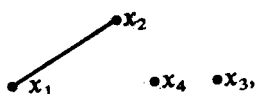
Определив подходящие множества X и Y и отношение λ , можно перейти еще к одному отношению, порождаемому λ . Это так называемое *сопряженное отношение* λ^* , которое получается, если поменять ролями множества X и Y , т. е. $\lambda^* \in Y \times X$, и строится в соответствии с правилом:

Отношение λ^* существует между y_i и x_i тогда и только тогда, когда между x_i и y_i существует отношение λ .

Матрица инцидентий для λ^* получается транспонированием матрицы инцидентий для λ , т. е.

$$\Lambda^* = \Lambda'.$$

В результате получим геометрический комплекс $K_Y(X; \lambda^*)$, в котором X — множество вершин, а Y — множество симплексов. Можно показать, что для рассмотренного примера комплекс $K_Y(X; \lambda^*)$ имеет вид



который, конечно же, гораздо более содержателен, чем полностью несвязная структура $K_X(Y; \lambda)$: вершины x_1 (хлеб) и x_2 (молоко) связаны 1-симплексом (гастроном).

Продемонстрируем общность описания систем на языке множеств и бинарных отношений еще на одном примере.

Пример. Шахматы

Чемпион мира по шахматам Эммануил Ласкер как-то заметил: «Если из 64 клеток шахматной доски вы контролируете 33, то преимущество на вашей стороне». Для нас данное высказывание представляет особый интерес, так как оно свидетельствует о том, что главное для шахматиста — это «стратегическое» содержание игры, которое мы можем выразить непосредственно в виде отношения между множеством фигур и множеством клеток шахматной доски.

Рассмотрим два отношения λ_W и λ_B , где λ_W определяет связь между белыми фигурами и клетками доски, а λ_B — то же для черных. Определим множества X и Y как

$$X = \{\text{фигуры}\}^1 \\ = \{QR, QN, QB, Q, K, KB, KN, KR, QRP, QNP, QBP, \\ QR, KP, KBP, KNP, KRP\}^2,$$

$$Y = \{\text{клетки}\}.$$

¹⁾ Используются стандартные международные обозначения для фигур: король — К, ферзь — Q, слон — В, конь — N, ладья — R, пешка — P; при этом предполагается, что клетки доски также упорядочены соответствующим образом.

²⁾ Имеются в виду следующие обозначения: буква Q в двухбуквенных символах означает ферзевый фланг, а K — королевский. Средняя буква в трехбуквенных символах (для пешек) обозначает фигуру, перел которой стоит данная пешка в начальной позиции. — *Прим. перев.*

Пусть заданы x_i и y_j ; определим отношение λ_w следующим образом:

$(x_i, y_j) \in \lambda_w$ тогда и только тогда, когда фигура x_i «атакует» клетку y_j .

Под термином «атакует» понимается одна из следующих ситуаций.

— Если ход белых и если фигура x_i не пешка и не король, то ход $x_i \rightarrow y_j$ — правильный (разрешенный) ход.

— Если фигура x_i — пешка, то клетка y_j находится под боем со стороны фигуры x_i .

— Если клетка y_j занята белой фигурой, то фигура x_i защищает эту фигуру.

— Если фигура x_i — белый король, то клетка y_j является соседней клеткой по отношению к клетке, занимаемой фигурой x_i .

— Если клетка y_j занята черной фигурой (но не королем) и если ход белых, то взятие данной черной фигуры фигурой x_i — правильный ход.

— Черный король занимает клетку y_j и находится под шахом фигуры x_i .

Аналогично определяется отношение λ_b .

Отметим, что λ_w зависит от состояния игры (от расположения фигур на доске и от того, чей ход).

Внимательный читатель, несомненно, заметил некоторое сходство между теоретико-множественным описанием системы и более привычным описанием в терминах теории графов на языке узлов и дуг (или вершин и ребер). По существу, данное выше определение совпадает с описанием на языке теории графов, если определить X , Y как множества вершин, соединенных ребрами в соответствии с отношением λ . Хотя при таком переходе в значительной степени утрачивается его гибкость, так как при этом исчезает (в лучшем случае затушевывается) универсальность отношения λ , тем не менее оно оказывается полезным во многих ситуациях.

Для теории систем наиболее существенным является описание динамики системы. Поэтому, чтобы понять, каким образом динамические переходы учитываются при теоретико-множественном описании процесса, введем понятие *образ*. Вообще говоря, образ Π есть отображение, которое каждому симплексу из комплекса ставит в соответствие определенное число, т. е.

$$\Pi: \sigma^i \rightarrow k,$$

где σ^i — симплекс из K , а k — определенная система чисел (действительных, целых и т. д.). Поскольку каждый симплекс из K обладает некоторой геометрической размерностью,

которая определяется числом его вершин, то образ Π является *ранжированным образом*¹⁾

$$\Pi = \Pi_0 \oplus \Pi_1 \oplus \dots \oplus \Pi_N,$$

где $N = \dim K$ — размерность наибольшего симплекса из K . Здесь каждое Π_i является отображением, определенным только на множестве i -мерных симплексов из K .

Поясним понятие образа на примере системы *хищник — жертва*. Напомним, что мы ввели два множества

$X =$ Множество жертв,

$Y =$ Множество хищников

с матрицей инцидентий

λ	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
y_1	1	1	1	1	0	0	0	0
y_2	1	0	1	0	0	0	0	0
y_3	0	0	0	0	1	1	0	0
y_4	0	1	0	0	1	0	1	1
y_5	0	0	0	0	0	0	1	0
y_6	0	1	0	0	1	0	0	0

Таким образом, y_1 (люди) — это 3-симплекс, y_4 (птицы) — 3-симплекс и т. д. Образ Π при этом ставит в соответствие каждому симплексу некоторое число, скажем численность популяции в данный момент. Поскольку симплексы ранжированы по их размерности, то и Π_i также ранжированы, поэтому в каждый момент

$\Pi_0: \{y_5 \text{ (рыбы)}\} \rightarrow$ численность рыб

$\{y_2 \text{ (львы)}\} \rightarrow$ численность львов

$\Pi_1: \{y_3 \text{ (слоны)}\} \rightarrow$ численность слонов

$\{y_6 \text{ (лошади)}\} \rightarrow$ численность лошадей

$\Pi_2:$ пусто

$\Pi_3: \{y_1 \text{ (люди)}\} \rightarrow$ численность народонаселения

$\{y_4 \text{ (птицы)}\} \rightarrow$ численность птиц

Полный образ для данной экосистемы имеет вид

$$\Pi = \Pi_0 \oplus \Pi_1 \oplus \Pi_2 \oplus \Pi_3.$$

Динамику системы можно теперь описать *изменениями* образа Π в каждый момент времени. Подробная интерпрета-

¹⁾ В оригинале «graded pattern». — Прим. перев.

ция этих изменений как сил, воздействующих на фиксированную геометрию комплекса, или как «свободных» изменений, допускаемых геометрией комплекса, дается в гл. 3. Отметим, что первая трактовка соответствует классическим ньютоновым силам, в то время как вторая отражает основные положения теории относительности.

ЛИТЕРАТУРА

Детальный анализ леонтьевских экономических моделей дан в работах
Gale D., *The Theory of Linear Economic Models*, McGraw-Hill, New York, 1960;

Baumol W., *Economic Theory and Operations Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1965;

Leontief W., *Mathematics in Economics*, *Bull. Am. Math. Soc.*, **60**, 215—233 (1954);

Isard W., Kaniss P., *The 1973 Nobel Prize for Economic Science*, *Science*, **182**, 568—569, 571 (1973).

Пример с водохранилищами взят из работы

Szöllösi-Nagy A., *State Space Approach to Hydrology*, in *Symposium on Mathematical Modeling in Hydrology*, University College, Galway, Ireland, April 1974.

Аналогичные примеры можно найти в работах

Bazykin A., *Elementary Model of Eutrophication*, in *Analysis and Computation of Equilibria and Regions of Stability*, H. Grümmer, ed., CP-75-8, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975; Jørgensen S., *A Eutrophication Model for a Lake*, *Ecol. Mod.*, **2**, 147—165 (1976).

Современная трактовка внутренних описаний системы в терминах переменных состояния дана в книгах

Casti J., *Dynamical Systems and Their Applications: Linear Theory*, Academic Press, New York, 1977;

Brockett R., *Finite-Dimensional Linear Systems*, Wiley, New York, 1970;

Kalman R., Falb P., Arbib M., *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.

Результаты описаний динамических процессов типа вход — выход с помощью аппарата функционального анализа содержатся в книгах
Saeks R., *Resolution Space, Operators and Systems*, Springer Verlag, Heidelberg, 1973;

Rosenbrock H., *State-Space and Multivariable Theory*, Nelson, London, 1970;

Luenberger D., *Optimization by Vector Space Methods*, Wiley, New York, 1969;

Porter W., *Modern Foundations of Systems Engineering*, Macmillan, New York, 1966;

Willems J. C., *The Analysis of Feedback Systems*, MIT Press, Cambridge, Massachusetts, 1971.

Пример с летающей тарелкой взят из классической работы

Ashby W. R., *Introduction to Cybernetics*, Chapman and Hall, London, 1965.

Более полное описание систем с конечным числом состояний дано в работе

Krohn K., Langer R., Rhodes I., *Transformations, Semigroups and Metabolism*, in *System Theory and Biology*, M. Mesarovic, ed., Springer Verlag, New York, 1968.

Вариационный подход к задачам, не связанный с потенциальными функциями, обсуждается в работе
Jumarie G. A. Relativistic Information Theory Model for General Systems: Lorentz Transformation of Organizability and Structural Entropy, *Int. J. Syst. Sci.*, 6, 865—886 (1975), в которой имеются ссылки на другие работы этого же автора, где он развивает весьма привлекательный подход к «термодинамическому» анализу больших систем.

Первой публикацией, в которой рассматривалось описание систем с помощью конечных множеств и бинарных отношений, была работа Atkin R. H., *Mathematical Structure in Human Affairs*, Heinemann, London, 1973.

Дальнейшее развитие и обобщения описания систем с помощью конечных множеств и бинарных отношений даны в книге Atkin R. H., *Multidimensional Man*, Penguin, London.

2

Основные положения и перспективы развития теории систем

Необходимо уметь мыслить абстрактно, чтобы по-новому воспринимать окружающий нас мир.

Р. Фейнман

Абстрактное мышление часто важнее, чем знания.

А. Эйнштейн

В предыдущей главе были рассмотрены некоторые основные способы описания динамики систем и их взаимодействия с человеком и окружающей средой. Теперь попытаемся показать, в чем заключается преимущество математического описания. Вообще говоря, математическое описание позволяет выделить некоторые, на наш взгляд, важные свойства данной физической системы и формально определить взаимосвязи между ее различными компонентами. Поскольку использование тех или иных математических абстракций, которыми оперирует теория систем, зависит от поставленной задачи, то мы рассмотрим различные вопросы, такие как стохастические процессы, управляемые динамические системы (процессы), идентификация и т. д. Хотя они и не являются основными в настоящей книге, их изучение по-прежнему дает важные системно-теоретические результаты и на них, в определенной степени, базируется изложение связности, сложности, устойчивости и теории катастроф больших систем.

УПРАВЛЯЕМЫЕ И НЕУПРАВЛЯЕМЫЕ ДИНАМИЧЕСКИЕ СИСТЕМЫ

Первые шаги, положившие начало долгому пути развития системного анализа, были сделаны античными астрономами. Не обладая средствами, с помощью которых можно было бы влиять на динамику изучаемых систем, они были вынуждены ограничить свой анализ лишь наблюдением, классификацией, возможно синтезом. Другими словами, их роль была пассивной: наблюдать. В аналогичном положении находятся современные исследователи, работающие, например, в области астрофизики. Они пока еще также вынуждены ограничиться только наблюдениями каких-то процессов, не имея возможности ими управлять.

Однако, насколько бы интересными и важными эти наблюдения ни были, маловероятно, что теория систем

достигла бы современного уровня развития, если бы не стремление управлять наблюдаемыми явлениями и не уверенность, что такое управление возможно. Современный исследователь призван играть активную роль в развитии наблюдаемого процесса, поскольку именно он генерирует соответствующие внешние воздействия, гарантирующие удовлетворительное поведение системы. Разумеется, при таком подходе активного вмешательства возникает множество проблем психологического и морального характера. Подобное разделение на активную и пассивную или управляемую и неуправляемую динамику позволяет наиболее наглядно выявить отличие классического и современного взглядов на системный анализ. «Кибернетический» или управленческий подход неизбежно приводит к изменению *входов* системы в зависимости от наблюдаемых *выходов*. При этом преследуется цель превратить некоторую первоначально независимую переменную в частично зависимую так, чтобы поведение системы в определенном смысле приближалось к некоторой стандартной, или желаемой, траектории. Такой процесс может оказаться более сложным, если имеется еще и *обратное преобразование*. Подобная ситуация типична для имитационного моделирования развивающихся систем. Обратное преобразование заключается в изменении и перестройке поведения системы по измеряемому выходу и является основой кибернетического регулирования и управления.

ИДЕНТИФИКАЦИЯ

Начальный этап построения математической модели данной системы состоит в идентификации существенных переменных и их взаимосвязей. В зависимости от конкретного типа выбранного математического описания идентификация может включать определение размерности пространства состояний, описание внутренней динамики системы и содержательных связей между множествами объектов, распределение вероятностей для случайных воздействий. Поскольку идентификация зависит от типа математического описания, которое в свою очередь зависит от того, насколько удачно проведена идентификация и т. д. и т. п., то процесс построения модели является итерационным: сначала выбирают математическое описание, которое затем модифицируют в зависимости от результатов идентификации, что приводит к новому описанию, и процесс повторяют.

Наиболее глубоко разработанной проблемой идентификации систем является задача построения внутреннего описания линейного отображения *вход — выход* с постоянными

коэффициентами. Для простоты изложения предположим, что данная система развивается в дискретном времени с начальным состоянием $x_0 = 0$, соответствующим начальному моменту времени $t_0 = 0$. Можно показать, что вход $u(t)$ и выход $y(t)$ системы связаны следующим соотношением:

$$y(t) = \sum_{t-\tau \geq 0} A_{t-\tau} u(\tau),$$

где все матрицы $\{A_i\}$ имеют размер $p \times m$. Тогда описание типа *вход — выход* системы Σ определяется (возможно, бесконечной) последовательностью матриц $\{A_1, A_2, \dots\}$.

Если внутреннее описание системы Σ , заданное соотношениями

$$\begin{aligned} x(t+1) &= Fx(t) + Gu(t), \\ y(t) &= Hx(t), \end{aligned}$$

согласуется с приведенным выше внешним описанием, то связь между матрицами F , G , H и $\{A_i\}$ имеет вид

$$A_t = HF^{t-1}G, \quad t = 1, 2, \dots \quad (2.1)$$

Задача реализации для линейных динамических систем состоит в отыскании $n \times n$, $n \times m$ и $p \times n$ -матриц F , G и H соответственно, удовлетворяющих соотношению (2.1) и таких, что размерность внутреннего пространства состояний n минимальна. Иными словами, задача состоит в построении по возможности более компактной модели, согласующейся с наблюдаемыми данными.

К счастью, существуют хорошие алгоритмы решения задачи реализации, если справедливо следующее предположение: *последовательность $\{A_i\}$ обладает конечномерной реализацией*. Если $\{A_i\}$ обладает реализацией размерности $n < \infty$, то первые $2n$ членов последовательности $\{A_i\}$ единственным образом определяют все остальные члены (теорема Гамильтона — Кели). Таким образом, задача состоит в определении n по данным $\{A_i\}$.

Как и следовало ожидать, подобных отработанных алгоритмов для нелинейных отображений *вход — выход* общего вида пока не существует, несмотря на попытки решения определенных классов задач с некоторой линейной или алгебраической структурой.

В отличие от наиболее общих задач идентификации (типа «от внешнего описания к внутреннему»), так называемые *задачи идентификации параметров* исследовались более интенсивно. Эти задачи обычно возникают, когда имеется твердая уверенность в правильности определения основной

внутренней структуры системы и невыясненными остаются только численные значения некоторых параметров.

Предположим, что динамика системы описывается дифференциальным (или разностным) уравнением

$$\begin{aligned}\dot{x} &= f(x, u, a), \\ y(t) &= h(x, a),\end{aligned}$$

где a — вектор неизвестных параметров, которые следует определить, основываясь на значениях наблюдаемого выхода системы $y(t)$. В некоторых случаях входная функция $u(t)$ выбирается таким образом, чтобы усилить влияние неизвестных параметров. Подчеркнем, что в данной ситуации существенным является предположение, что функции f и h , описывающие структуру системы, известны, хотя относительно их линейности никаких предположений не делается.

В качестве иллюстрации задач этого класса рассмотрим задачу о динамике численности некоторой биологической популяции, описание которой может быть получено с помощью следующих логистических уравнений:

$$\frac{dx}{dt} = rx \left(1 - \frac{x}{K}\right) - Ex, \quad x(0) = x_0.$$

Здесь $x(t)$ — численность популяции в момент времени t , r — удельная скорость ее роста в отсутствие лимитирования, K — константа, характеризующая предельные трофические возможности среды обитания (уровень насыщения численности), и E — коэффициент интенсивности изъятия особей из популяции. Предположим, что имеется возможность измерения численности популяции в каждый момент времени, т. е.

$$y(t) = x(t),$$

но численное значение параметра K неизвестно. В этом случае задача идентификации параметров состоит в определении K на основе измерения численности популяции:

$$K = \frac{-rx^2}{x + (E - r)x} \quad \text{для всех } t > 0.$$

Таким образом, для определения K достаточно знать (наблюдать) $y(t)$ на любом интервале времени. Однако в более реальных ситуациях, когда имеется лишь конечное число значений $y(t)$, приходится использовать различные приближенные методы.

Обобщения этой задачи с учетом неопределенности в измерении $x(t)$, наличия различных видов в биологической популяции и т. д. представляют собой достаточно сложные в

математическом плане задачи. Более подробно эти вопросы освещаются в работах, перечисленных в конце главы¹⁾.

Задачи идентификации систем, описываемых с применением более общего аппарата, например потенциальных функций или теоретико-множественных отношений, пока еще слабо изучены. В отличие от внутреннего или внешнего описания на языке дифференциальных уравнений описания данного типа в гораздо большей степени зависят от того, как сам исследователь представляет себе существо изучаемого процесса. Поэтому в этом случае постановка и решение задачи идентификации — больше искусство, чем наука, и состоит в основном в выделении таких множеств и отношений («энергетических» функций), которые приводят к содержательным результатам. В связи с этим определенный интерес представляет систематизация процесса разумного выбора таких множеств и отношений (гл. 3).

ОГРАНИЧЕНИЯ

Системный анализ, как и политика, — это прежде всего искусство действовать в пределах «возможного». Рассматривая математическую формулировку той или иной задачи, исследователь (или лицо, принимающее решение) должен полностью представлять себе те внутренние и внешние факторы, которые могут ограничить его выбор стратегий управления. Различные обстоятельности, связанные с объемом имеющихся ресурсов, спросом, который необходимо удовлетворить, имеющейся технологией, наличием и возможностями ЭВМ, людскими ресурсами, бюджетом времени и т. д., резко сужают круг возможностей, доступных лицу, принимающему решения.

Выделим два принципиально различных типа ограничений:

внутренние — ограничения, налагаемые структурой самой системы,

внешние — ограничения, налагаемые на поведение системы внешними факторами. Рассмотрим эти ограничения несколько подробнее.

Внутренние ограничения возникают вследствие определенной ограниченности возможностей измерять характеристики состояния системы и управлять течением процесса, т. е. они ограничивают взаимодействие системы с внешним миром. Вообще говоря, ограничения этого типа наиболее четко вид-

¹⁾ См. также: Динамическая теория биологических популяций, под ред. Р. А. Полуэктова. — М.: Наука, 1974, в частности гл. 9, где обсуждаются вопросы идентификации параметров в связи с моделированием и управлением численностью биологических популяций. — *Прим. перев.*

ны тогда, когда для внутреннего описания используют дифференциальные или разностные уравнения. Для иллюстрации понятия *внутренние ограничения* рассмотрим пример из области биомедицины.

Пример. Фармакокинетика

Предположим, что пациент, страдающий заболеванием сердца, получает дигитоксин, который в результате процесса обмена веществ превращается в дигоксин. Поскольку последний имеет способность накапливаться в организме, что в результате может привести к летальному исходу, то очень важно уметь точно определять его содержание в организме прежде, чем пациент примет очередную порцию дигитоксина.

Многокомпонентная модель, используемая для описания кинетики и превращений дигитоксина, изображена на рис. 2.1.

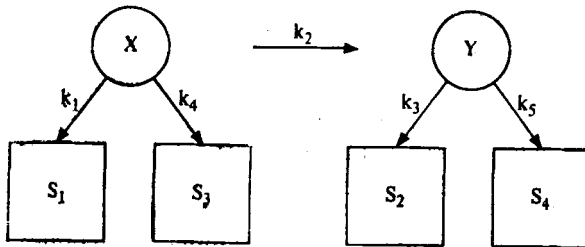


Рис. 2.1. Многокомпонентная структура процесса обмена веществ.

Здесь X — содержание дигитоксина в организме, Y — содержание дигоксина, S_1 и S_2 — мочевые выделения, S_3 и S_4 — немочевые выделения и k_i — коэффициенты диффузии, $i = 1, 2, \dots, 5$.

Обычно принято считать, что если в организм вводится некоторая доза дигитоксина, то примерно 92% этой дозы немедленно разносится по организму и около 85% от оставшихся 8% сразу превращается в дигоксин. Предполагается, что динамика концентрации лекарств X и Y может быть описана следующим образом:

$$\dot{X} = -(k_1 + k_2 + k_4)X,$$

$$\dot{Y} = k_2 X - (k_3 + k_5)Y,$$

$$\dot{S}_1 = k_1 X,$$

$$\dot{S}_2 = k_3 Y,$$

$$\dot{S}_3 = k_4 X,$$

$$\dot{S}_4 = k_5 Y.$$

Начальные условия имеют вид

$$X(0) = 0,92D, \quad Y(0) = (0,85) (0,08) D,$$

$$S_1(0) = S_2(0) = S_3(0) = S_4(0) = 0,$$

где D — введенная доза дигитоксина.

Предположим, что можно измерить содержание дигитоксина и дигоксина только в мочевых выделениях. Тогда выход системы имеет вид

$$y_1(t) = S_1(t),$$

$$y_2(t) = S_2(t).$$

Именно это весьма реалистическое с практической точки зрения ограничение и является тем, что называют *внутренним ограничением* системы. В силу этого ограничения не все внутренние переменные системы доступны для непосредственного измерения.

Имея в виду основную задачу, стоящую перед врачом, необходимо знать, достаточны ли измерения переменных y_1 и y_2 для определения неизвестной начальной дозы лекарства D . Эта задача является так называемой *задачей наблюдаемости*, обсуждение которой можно найти в литературе, приведенной в конце главы.

Внешние ограничения имеют качественно иной характер. Как отмечалось выше, они обусловлены не физическими или структурными ограничениями самого процесса, а «произволом» лица, принимающего решения, которое является «внешним» по отношению к системе. Вообще говоря, эти ограничения связаны с такими обстоятельствами, как ограниченность имеющихся ресурсов и производственных мощностей, наличие заданного спроса и т. д. Существенным моментом здесь является то, что эти ограничения налагаются извне и не имеют никакого отношения к математическим ограничениям, содержащимся в самой модели.

Типичные примеры внешних ограничений содержатся в задачах экономического управления, где требуется определить соответствующее распределение фиксированных финансовых средств для достижения определенных целей. Возьмем, к примеру, сотрудника некоторой компании, ответственного за рекламу ее продукции, причем бюджет, который компания выделила на рекламные цели, составляет M долл. Это значит, что он может истратить M долл. на размещение рекламных объявлений, скажем, в газетах, журналах, на телевидении, радио и рекламных афишах. Предположим, что вложение x_i долл. в i -й способ рекламы (i -газеты, журналы и т. д.) приводит к сбыту партии объема $f_i(x_i)$, причем

функции $f_i(\cdot)$ считаются известными. Поскольку компания заинтересована в максимизации сбыта, рекламодатель сталкивается с решением задачи максимизации

$$\sum f_i(x_i)$$

по всем распределениям $\{x_{\text{газ}}, x_{\text{журн}}, x_{\text{рад}}, x_{\text{тв}}, x_{\text{аф}}\}$
при внешнем ограничении

$$\sum x_i \leq M.$$

Следовательно, внешнее ограничение возникло из-за ограниченности бюджета, а не из образа взаимодействия системы с внешним миром.

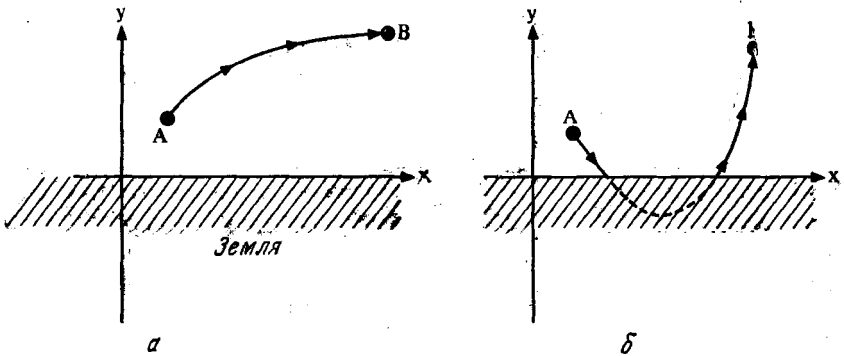


Рис. 2.2. Траектории полета самолета.

В качестве другого примера внешнего ограничения рассмотрим задачу о пилоте, которому необходимо прилететь из пункта А в пункт В за минимальное время. В зависимости от характеристик самолета и других предположений, математическое решение этой задачи может привести к оптимальной траектории, показанной на рис. 2.2, б. Очевидно, что такое решение не учитывает реальных ограничений, имеющих в этой ситуации, которые должны быть наложены извне с тем, чтобы сделать задачу осмысленной с физической точки зрения. Надлежащее внешнее ограничение ($y > 0$) привело бы тогда к оптимальной траектории, более похожей на изображенную на рис. 2.2, а.

СТОХАСТИЧЕСКИЕ СИСТЕМЫ

Несмотря на то что основная направленность настоящей книги такова, что не возникает необходимости в подробном обсуждении вопросов, связанных с неопределенностью, тем

не менее следует иметь в виду, что при анализе большинства реальных системных задач практически ничего не известно достоверно! Независимо от выбранного математического описания, неопределенности будут присутствовать в динамике, целях, ограничениях и т. п. При удачном стечении обстоятельств для неопределенных переменных будут известны с определенной достоверностью распределения вероятностей. Однако довольно часто даже распределения вероятностей заранее не известны, поэтому возникает адаптивная ситуация. В любом случае нельзя считать анализ законченным без тщательного исследования неопределенностей, присущих модели.

В дальнейшем будем придерживаться довольно смелого предположения, что всеми неопределенными эффектами можно пренебречь, т. е. будем считать, что передаточные функции, динамика состояний и т. д. известны достоверно. Такое допущение, естественно, должно быть оправдано полученными результатами, что мы и попытаемся продемонстрировать в каждом отдельном случае.

ОПТИМИЗАЦИЯ

Одна из наиболее злободневных проблем анализа систем, рассматриваемых в социально-экономических задачах, — это проблема выбора критерия, т. е. вопрос о том, каким образом следует сравнивать между собой различные реализации поведения системы. К счастью, динамические процессы, наблюдаемые в физических и биологических системах, часто протекают по вполне определенным законам, которые, как правило, являются следствием различных принципов минимума или законами сохранения. Однако перенос этих законов на объекты социальной природы в лучшем случае носит искусственный характер и, более того, часто просто невозможен.

Поскольку цель этой книги состоит в изучении структуры систем независимо от вопросов оптимизации, можно позволить себе роскошь оставить в стороне проблему выбора критерия. Тем не менее, для того чтобы продемонстрировать значимость этой проблемы, рассмотрим простой пример, иллюстрирующий ситуацию, когда выбор различных критериев приводит к качественно различным стратегиям управления.

Предположим, что динамика системы описывается одномерным линейным дифференциальным уравнением

$$\frac{dx}{dt} = u(t), \quad x(0) = c,$$

где u — вход, или функция управления. Предположим, далее, что доступные резервы управления ограничены следующим образом:

$$|u(t)| \leq 1 \quad \text{для всех } t \geq 0.$$

(Подобная ситуация возникает, например, при управлении автомобилем, и тогда функция $u(t)$ есть скорость движения.)

Одним из критериев для данного процесса может быть перевод системы из начального состояния c в некоторое заданное состояние, например $x = 0$ за минимальное время. Хорошо известно, что решение такой задачи имеет вид

$$u^*(t) = \begin{cases} +1, & c < 0, \\ -1, & c > 0, \end{cases}$$

т. е. релейное управление является оптимальным.

Предположим теперь, что мы стремимся минимизировать квадратичный функционал вида

$$J = \int_0^T (x^2 + u^2) dt.$$

Можно показать, что в этом случае оптимальный закон управления имеет вид

$$u^*(t) = \text{th}(T - t)x(t),$$

и он может быть реализован в виде обратной связи или синтеза.

Полученные результаты показывают, что изменение критерия качественно меняет характер решения. В первом случае мы имеем экстремальные управления, переключающиеся с одной границы на другую в зависимости от начального состояния. Во втором случае оптимальный закон управления строится по ходу развития самого процесса и не имеет никаких точек разрыва. Важно отметить, что, хотя динамика системы остается неизменной, выбор иного критерия приводит к качественному изменению оптимального управления.

ГЛОБАЛЬНЫЕ СВОЙСТВА СИСТЕМЫ

В настоящее время существенно увеличилось число проблем, решение которых не может быть получено редуционистскими методами, что, в свою очередь, возродило интерес к изучению и развитию холистских, или глобальных, подходов. В этой связи наша цель состоит в том, чтобы каталогизировать некоторые наиболее перспективные направления, включая вопросы связности, сложности и устойчивости.

Для иллюстрации фундаментального различия между локальным и глобальным описаниями системы рассмотрим простой пример — математический маятник (рис. 2.3).

Если отклонение маятника от вертикали обозначить через $x(t)$, то в локальной окрестности *любого* такого положения можно записать динамические уравнения движения

$$\ddot{x} + \sin x = 0, \quad x(0) = x_0, \quad \dot{x}(0) = 0,$$

в безразмерных единицах. Это уравнение описывает локальное поведение маятника в (бесконечно малой) окрестности положения $x(t)$. Редукционист попытался бы «склеить» по-

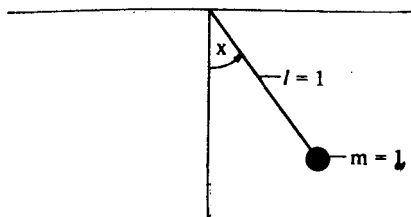


Рис. 2.3. Математический маятник.

добные локальные описания для последовательных точек в надежде достичь понимания глобального поведения. Хотя иногда такой подход оказывается успешным, непредвиденные проблемы, возникающие при его использовании, существенно снижают его эффективность.

Холист, приступая к решению этой же задачи, прежде всего заметил бы, что должны соблюдаться определенные глобальные свойства системы, и поэтому любое локальное поведение должно удовлетворять ограничениям, налагаемым глобальными свойствами. Если к тому же эти ограничения достаточно жестки, то можно ожидать, что любые локальные движения ими определяются однозначно.

В случае маятника такие глобальные ограничения определяются принципом Гамильтона — Якоби, согласно которому, глобальное движение системы соответствует минимуму полной энергии системы. Вводя гамильтониан

$$H = \text{Кинетическая энергия} + \text{Потенциальная энергия},$$

видим, что движение системы должно быть таким, что

$$H(x, \dot{x}) = (1/2) \dot{x}^2 + 1 - \cos x$$

достигает минимума. Это уравнение, очевидно, может быть сведено к уравнению движения, приведенному выше, т. е. локальные уравнения движения могут быть получены как следствие глобального принципа, а не выведены на основе

рассуждений локального характера и использования второго закона Ньютона. С концептуальной точки зрения такое различие является фундаментальным.

Для систем, рассматриваемых в социально-экономических приложениях, не существует подобных общих законов (по крайней мере сейчас), и мы вынуждены ограничиться рассмотрением ряда глобальных свойств и методов работы с ними, рассчитывая на то, что освещение различных аспектов задачи поможет понять ее структуру в целом.

В качестве примера использования глобального подхода для решения системных задач рассмотрим ситуацию с заторами на транспортной магистрали. Учитывая наличие множества факторов, влияющих на дорожную ситуацию, можно попытаться «склеить» локальные описания, полученные методом Монте-Карло или методами теории очередей и т. д. Такой подход позволяет выявить множество деталей, однако в большинстве случаев остается неясным, как можно использовать полученные результаты для анализа других дорожных ситуаций. Холист в этом случае прибегнул бы к помощи статистической физики и попытался описать подобную ситуацию одним уравнением, пренебрегая дистанцией между машинами, причинами заторов и т. д. Главным для него было бы значение параметра q — плотности потока машин (число машин в час на километр пути). Время T_A (минуты), необходимое для преодоления 1 км дороги, можно представить как сумму двух слагаемых

$$T_A = T_{A0} + k \cdot n_A,$$

где T_{A0} — время необходимое для преодоления участка дороги длиной $A = 1$ км без учета помех со стороны других машин ($q \approx 0$) ($T_{A0} = 0,5$ мин/км соответствует скорости свободного движения 120 км/ч); $k \cdot n_A$ — дополнительное время, необходимое для преодоления участка $A = 1$ км, пропорциональное числу машин n_A , находящихся на участке A в течение времени T_A (т. е. задержка в условиях заторов является линейной функцией числа торможений и ускорений, или числа n_A машин, участвующих в движении). Число n_A является произведением плотности потока машин (транспорта) q и длительности периода времени T_A :

$$n_A = \frac{q \cdot T_A}{60}.$$

Учитывая предыдущие соотношения, получаем

$$T_A = \frac{T_{A0}}{1 - k \frac{q}{60}}.$$

Функция $T_A = f(q)$ является выпуклой: каждая дополнительная машина, приводящая к росту q , не только задерживается на участке A , но и является причиной задержки других машин. При значениях $T_{A0} = 0,5$ и $k = 0,0266$ имеется хорошее согласие между кривой и экспериментальными данными (рис. 2.4). Полученное уравнение дает значения для q , лежащие гораздо ниже теоретического значения плотности

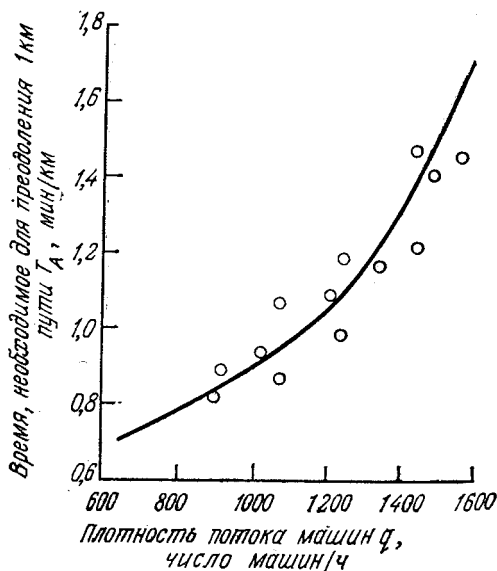


Рис. 2.4. Задержки, вызванные транспортными заторами.

Временное уравнение: $T_A = T_{A0} + k \frac{q}{60} T_A$; $T_{A0} = 0,50$ и $k = 0,0266$.

$q_{\infty} = 2,255$ машины в час, соответствующей «параличу дороги». Таким образом, глобальный (а не локальный) подход позволяет построить содержательную модель временных задержек в транспортной магистрали с заторами.

СВЯЗНОСТЬ И ГРАФЫ

Структурная связность системы является, по-видимому, наиболее существенной ее качественной характеристикой. Кажется очевидным, что с исчезновением структурной связности исчезнет и сама система, поскольку само понятие системы подразумевает наличие «чего-то», находящегося в некотором отношении (или как-то связанного) с «чем-то».

Анализ задачи построения математического описания связности может быть осуществлен с помощью различных подходов, причем наиболее удачные из них построены на использовании теории графов и алгебраической (комбинаторной) топологии. Это является вполне закономерным, поскольку вопрос о характере связности «простейших элементов» единого целого интересует алгебру в гораздо большей степени, чем любую другую математическую дисциплину.

Сущность исследования связности состоит в том, чтобы осознать и уяснить себе те математические конструкции, ко-

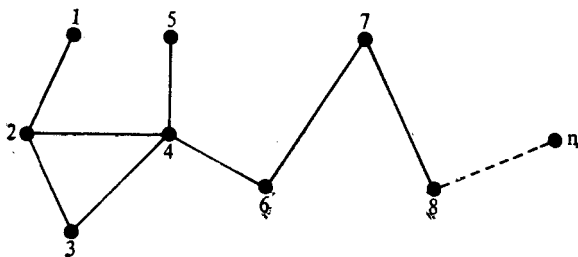


Рис. 2.5. Теоретико-графовое описание.

которые описывают характер связи между отдельными компонентами системы Σ . Если вообразить некоторую систему, в которой можно выделить n различных компонент (подсистем), то можно попытаться изобразить структуру (связную) Σ графом (рис. 2.5): n вершин изображают n подсистем системы Σ , а дуга, соединяющая подсистемы i и j , показывает, что эти две подсистемы находятся в некотором отношении или как-то связаны между собой. Например, j -я подсистема может генерировать входы для i -й подсистемы, а i -я управлять j -й и т. д. Эту схему, естественно, можно развить; так, например, можно ввести ориентацию на дугах и образовать ориентированный граф (орграф). Такое представление системы Σ позволит изучать ситуации, когда i -я подсистема влияет на j -ю, но не наоборот. Кроме того, можно учесть силу связности, сопоставив каждой направленной дуге некоторое число и т. д. Все это в конечном счете позволяет определить, какие компоненты системы Σ влияют на другие компоненты и в какой степени. По существу, теоретико-графовые модели позволяют несколько лучше понять, как можно было бы осуществить декомпозицию системы Σ на меньшие составляющие без потери тех основных свойств, в силу которых она и является системой.

Пример. Трофические структуры и экологические ниши

Рассмотрим экологическую систему, состоящую из пяти видов: птиц, насекомых, трав, антилоп и лис (рис. 2.6). Трофическая структура этого сообщества изображается оргграфом, вершины которого соответствуют видам. Дуга, проведенная от i -го вида к j -му, означает, что j -й вид является

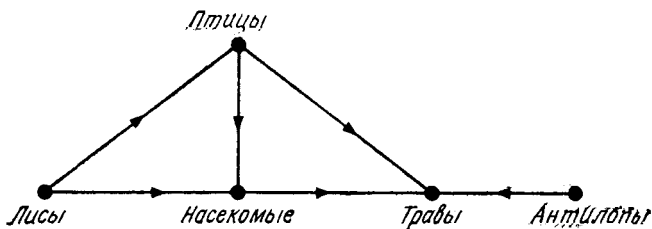


Рис. 2.6. Оргграф простой системы.

жертвой i -го вида. По данному графу можно построить матрицу смежности

$i \backslash j$	1	2	3	4	5	
Птицы	1	0	0	1	1	0
Лисы	2	1	0	1	0	0
Насекомые	3	0	0	0	1	0
Травы	4	0	0	0	0	0
Антилопы	5	0	0	0	1	0

аналогичную матрице инцидентий в теоретико-множественном описании, а также ряд других показателей, характеризующих важные аспекты системы.

Отметим, что некоторые из компонент (например, травы) кажутся более важными для системы в целом, чем другие (например, птицы), и, по-видимому, это связано с такими экологическими понятиями, как трофический уровень и борьба видов. Важно подчеркнуть, что теоретико-графовое описание позволяет непосредственно увидеть некоторые геометрические свойства матрицы смежности.

Как бы ни были важны и удобны теоретико-графовые методы для зрительного анализа связности, их использование неизбежно связано с трудностями геометрического и аналитического характера, если учитывается структура самих компонент. Из общих соображений можно ожидать, что при попытке описать многомерную структуру планарным

графом или, более общо, графом, изображенным на плоскости (это не одно и то же!), многое из геометрической структуры системы будет утеряно или в лучшем случае скрыто. По этой причине обратимся к другому возможному способу анализа связности, основанному на топологических идеях.

СВЯЗНОСТЬ И СИМПЛИЦИАЛЬНЫЕ КОМПЛЕКСЫ

Приближенно симплициальный комплекс состоит из множества вершин X и множества симплексов Y , образованных из этих вершин в соответствии с заданным бинарным отношением λ . Симплициальный комплекс $K_Y(X; \lambda)$ образован множеством симплексов Y , связанных через общие грани, т. е. через общие вершины. Например, можно положить

$$Y = X = \{\text{птицы, лисы, насекомые, травы, антилопы}\}.$$

При этом отношение λ таково: симплекс y_i состоит из всех вершин x_j , таких, что x_j является жертвой y_i . Таким образом,

$y_1 =$ „птицы“ — 1-симплекс, состоящий из вершин „насекомые“ и „травы“,

$y_2 =$ „лисы“ — 1-симплекс, состоящий из вершин „птицы“ и „насекомые“

и т. д. Отметим, что n -симплекс состоит из $n + 1$ вершин и его размер на единицу меньше числа вершин.

Вообще говоря, p -симплекс σ_p представляется выпуклым многогранником с вершинами в евклидовом пространстве E^p , а комплекс $K_Y(X; \lambda)$ — совокупностью таких многогранников в евклидовом пространстве E^a соответствующей размерности. Хотя размерность α наверняка не превышает суммы размерностей всех симплексов из $K_Y(X; \lambda)$, однако поскольку многие симплексы имеют общие грани, то размерность α на самом деле окажется меньше. В действительности можно показать, что если $\dim K_Y(X; \lambda) = n$, то $\alpha = 2n + 1$. Так, если $\dim K_Y(X; \lambda) = 1$, то наибольший порядок σ_p есть $p = 1$, поэтому можно ожидать, что трехмерного пространства E^3 будет достаточно, чтобы геометрически представить произвольный комплекс размерности 1. Это можно проиллюстрировать следующим образом: на плоскости (E^2) надо соединить непересекающимися линиями три дома H_1 , H_2 и H_3 с источником газа, воды и электроэнергии. Неразрешимость поставленной задачи иллюстрирует наше утверждение. Задача графически изображена на рис. 2.7, а ее решение в E^3 показано на рис. 2.8.

Основываясь на геометрической интуиции, можно изучать многомерную связную структуру комплекса $K_Y(X; \lambda)$ различ-

ными способами с привлечением алгебраических методов. В связи с этим рассмотрим некоторые важные понятия.

q-связность. Это понятие связано с изучением цепочек связи в $K_Y(X; \lambda)$, таких, что каждый симплекс в цепи имеет

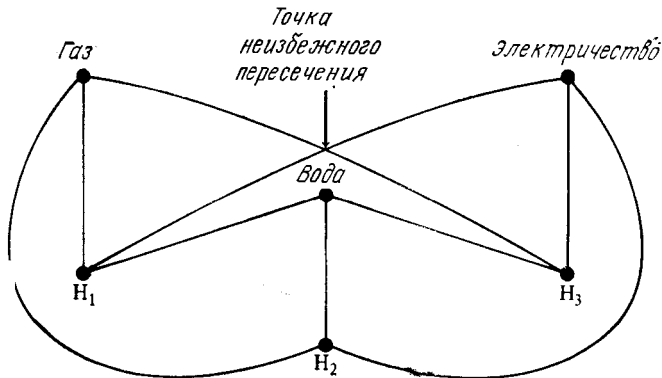


Рис. 2.7. Проблема пересечения в E^2 .

общую вершину с соседними симплексами, $q = 0, 1, 2, \dots, \dim K - 1$. Геометрически эти цепи содержат достаточно много локальной информации относительно того, каким об-

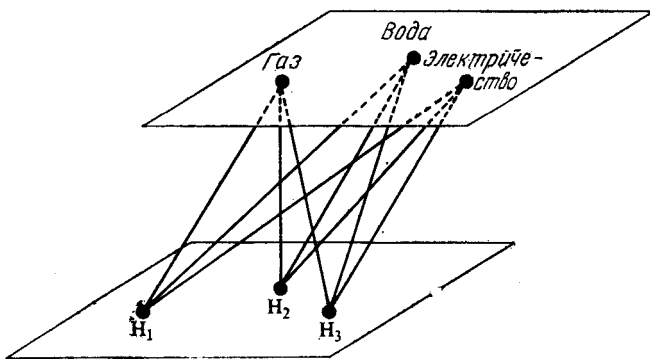


Рис. 2.8. Решение проблемы пересечения в E^3 .

разом симплексы, составляющие комплекс, связаны друг с другом. Если представить себе, что мы можем «видеть» только в пространстве размерности $\geq q$ (скажем, с помощью специальных очков), то, рассматривая комплекс $K_Y(X; \lambda)$, мы увидим, что он распадается на Q_q несвязанных элементов. Подобное геометрическое представление порождает алгебраическую теорию q -связности, позволяющую гораздо

лучше понять процессы обмена информацией внутри комплекса.

Эксцентриситет. Для того чтобы понять, каким образом отдельные симплексы «вложены» в комплекс, введем понятие *эксцентриситета*. Это понятие отражает как относительную важность данного симплекса для комплекса в целом (через его размерность), так и его значимость как связующего звена

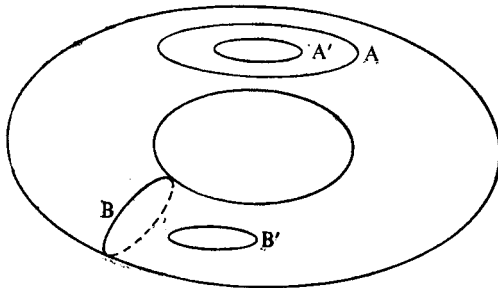


Рис. 2.9. Гомотопия на торе.

(через максимальное число его вершин, принадлежащих также любому другому симплексу). Другими словами, эксцентриситет позволяет увидеть и оценить, насколько «плотно» каждый симплекс вложен в комплекс.

Образ. Как отмечалось в гл. 1, для описания динамики системы необходимо ввести отображение каждого симплекса из $K_r(X; \lambda)$ в соответствующее числовое поле:

$$\Pi : \delta'_i \rightarrow k, \quad i = 0, 1, \dots, \dim K,$$

$$r = 1, 2, \dots, \text{card } K.$$

Образ Π отражает динамические изменения, происходящие в комплексе со временем. Поскольку каждый симплекс σ_i обладает характеристической геометрической размерностью, то же справедливо и для связанных с ним численных величин. Следует иметь в виду, что геометрическая структура налагает различные ограничения на изменение образа, т. е. на динамику системы.

Гомотопия. Вопрос о том, насколько «близким» является данный симплекс (цепь) к другому симплексу (цепи), представляет как теоретический, так и прикладной интерес. Если ввести понятие *гомотопия*, то можно получить ответ не только на этот вопрос, но и на вопрос о том, можно ли непрерывным преобразованием трансформировать одну цепь в другую, не нарушая геометрии системы. Так, например, кривые A и A' на торе (рис. 2.9) являются гомотопными, а кривые

В и **В'** нет, поскольку наличие «дырки» в центре не позволяет непрерывно деформировать **В** в **В'**. Аналогичные понятия могут быть введены и для комплекса $K(X; \lambda)$, и не исключено, что они могут оказаться полезными при анализе его структуры.

Хотя с чисто математической точки зрения изложенные геометрические понятия совершенно элементарны, они все же дают весьма подробную информацию, необходимую для понимания статической геометрии данного бинарного отношения и возможной динамики соответствующей ему связанной структуры. (Это довольно смелое утверждение будет подтверждено в следующей главе многочисленными примерами.)

СЛОЖНОСТЬ

Нет сомнения, что наиболее употребительным прилагательным в литературе по системному анализу является «сложный». Оно же является и наименее четко определяемым. Чисто интуитивно мы ощущаем, что сложная система — это такая система, статическая структура или динамическое поведение которой «непредсказуемы», «запутаны», противоречат «здравому смыслу» и т. п. Короче говоря, сложная система — это нечто весьма сложное (одна из тавтологий системного анализа). Тем не менее решение проблем, возникающих в вычислительной технике и теории алгоритмов, требовало разработки способов количественного описания этого понятия, и в результате целый ряд исследователей вынуждены были вплотную заняться вопросами сложности.

В основном сложность связана с двумя важными свойствами системы: (а) *математической структурой неприводимых компонент* (подсистем) и (б) *способом, которым эти компоненты связаны между собой*. Отсюда с очевидностью следует, что сложность присуща самой системе, а тот факт, что сложность все же связана с отношением между наблюдателем и наблюдаемым объектом, при такой трактовке затушевывается и отступает на второй план. Однако, поскольку книга носит в целом вводный характер, мы не будем затрагивать подобных релятивистских аспектов.

Первое свойство системы допускает возможность снижения видимой сложности системы путем объединения отдельных переменных в подсистемы, как это, например, имеет место в блок-схеме радиоприемника, где различные элементы системы (сопротивления, транзисторы и т. д.) сгруппированы в функциональные блоки, такие как цепь настройки или блок питания. Естественно, при такой декомпозиции преследуется

цель позволить исследователю упростить анализ системы, рассматривая ее как слабо связанную совокупность взаимодействующих подсистем. Следует, однако, отметить, что, хотя и предполагается, что взаимодействия между подсистемами будут слабыми, из этого вовсе не следует, что они действительно окажутся пренебрежимо малыми.

Второе свойство в значительной степени отражает сущность уже обсуждавшегося понятия сложности и включает такие характеристики системы, как размерность, иерархия, длина цепей связи и т. п. Кроме того, очевидно, что вопросы, касающиеся динамического поведения системы, тесно связаны как со структурой отдельных элементов, так и со способом их организации.

Одним из важных аспектов понятия сложности является ее двоякая природа. Следует различать *структурную*, или *статическую*, сложность, включающую связность и структуру подсистем, и *динамическую* сложность, связанную с поведением системы во времени. Тот факт, что эти свойства могут быть сравнительно независимыми, можно проиллюстрировать на простых примерах. Так, например, обычные часы обладают высокой степенью статической сложности, однако их динамическая сложность, по существу, равна нулю, если, конечно, часы работают как надо. Напротив, поведение нелинейного осциллятора, описываемого уравнением Ван дер Поля,

$$\ddot{x} + \lambda(x^2 - 1) + \dot{x} = 0$$

может быть весьма сложным в зависимости от значения параметра λ , и именно из-за этого «сложного» поведения он представляет теоретический и прикладной интерес. Со структурной же точки зрения осциллятор Ван дер Поля вовсе не является сложной системой.

Для иллюстрации непредсказуемого поведения, по-видимому, характерного для сложных систем, рассмотрим идеализированный линейный процесс, изображенный на рис. 2.10. (Это чисто условный пример, поэтому и его «содержательная» интерпретация также условна.)

Предположим, что гипотетическая экономическая система включает два предприятия: механическую мастерскую и электростанцию, для которых требуются рабочие двух специальностей: механики и электрики. Оба предприятия имеют фиксированное число рабочих мест и стремятся работать с полной занятостью. Смена персонала происходит достаточно быстро, так что полное число занятых рабочих равно ежегодному выпуску училищ. Всего имеется три училища:

два небольших частных училища, одно из которых готовит механиков, а другое — электриков, и одно крупное общественное училище¹⁾, готовящее равное число тех и других. Общественное училище готовит двух рабочих за один доллар. Частные училища готовят одного рабочего на одно вакантное рабочее место, но так как частные училища более требовательны к абитуриентам, производительность труда их выпускников вдвое выше, чем у выпускников общественного

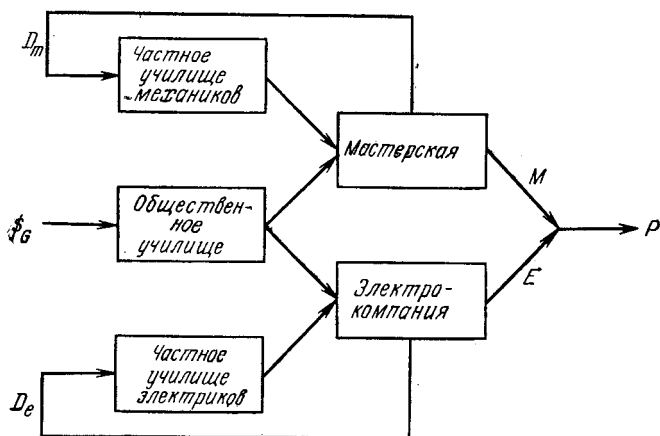


Рис. 2.10. Упрощенная блок-схема экономики развивающейся страны.

училища. Поскольку правительство субсидирует данные предприятия, они принимают на работу всех, оканчивающих общественное училище. Данная ситуация описывается следующими уравнениями:

$$M = D_m + G,$$

$$P = D_m + G + D_e,$$

$$E = G + D_e,$$

где M — число механиков, E — число электриков, P — полные производительные силы (в терминах производительности труда выпускников частных училищ), D_m — спрос на механиков, D_e — спрос на электриков, G — ежегодный выпуск общественного училища.

¹⁾ Имеется в виду училище, финансируемое из общественных средств (местного бюджета). — Прим. ред.

Отметим, что масштабирование уравнений несущественно, поскольку явления, которые мы сейчас опишем, не зависят от выбранного масштаба.

Предположим, что существует возможность управлять числом механиков и электриков и производительными силами, при этом управляющими органами являются оба предприятия и правительство. Правительство управляет переменной P , изменяя G , мастерская управляет M через D_m , и электростанция контролирует E , варьируя D_e .

В описанной ситуации возможно следующее парадоксальное поведение. Предположим, что на обоих предприятиях была полная занятость. Пусть правительство увеличивает G на единицу. Тогда предприятия в свою очередь уменьшают D_e и D_m на единицу, чтобы избежать превышения допустимой численности рабочих. Таким образом, изменение D_e и D_m приводит к уменьшению P на две единицы. Итак, увеличение G на единицу приводит к уменьшению P на единицу. Этот вывод не зависит от деталей реализации стратегий управления и определяется лишь структурой управления и целей.

Парадокс исчезает, если правительство может регулировать D_m и D_e , а не только G . Однако основная проблема возникает из-за влияния других управляющих воздействий на взаимосвязь между управляемыми (M , P , E) и управляющими (D_m , G , D_e) переменными.

Вывод, который можно сделать из анализа этого примера, состоит в том, что, казалось бы, даже в элементарных системах могут возникать совершенно неожиданные (и неприятные) явления, если сложность взаимосвязей не изучена должным образом. Другой важный вывод состоит в том, что в отличие от обычных представлений такое парадоксальное поведение вызвано вовсе не наличием нелинейности, стохастических эффектов и т. п., а порождается исключительно структурой системы, имеющимися связями и ограничениями, присущими компонентам системы.

Данный пример иллюстрирует еще один важный момент, присущий понятию сложности системы, а именно различие между сложностью неуправляемой системы и сложностью управляемой системы. Грубо говоря, сложность неуправляемой системы определяется совокупностью статической и динамической сложности в отсутствие управления, или, более общо, процессом преобразования, при котором полностью используется потенциал системы. Процесс преобразования, однако, может привести к возникновению неустойчивых конфигураций. Так, например, неустойчивые конфигурации могут возникнуть из-за разрыва между вычислительными

потребностями системы в целом и вычислительными возможностями составляющих ее подсистем.

Под сложностью управляемой системы понимается тот уровень сложности, который сопряжен с вычислениями, необходимыми для того, чтобы система была полностью управляемой. В данном случае неустойчивые конфигурации могут появиться, если быстродействие некоторых подсистем недостаточно велико, чтобы вовремя реагировать на изменения входных воздействий.

Связь между этими двумя типами сложности можно назвать *эволюционной сложностью*, и говорят, что система полностью сбалансирована, когда ее потенциальные возможности используются полностью, т. е. когда сложность управляемой и управляемой системы одинакова.

Пример. Генетическая модель Джекоба — Моно

Предположим, что функции клетки можно разделить на две группы: обмен веществ M и генетическое управление G . Механизм работы клетки можно попытаться описать следующим образом. G пытается регулировать M , воспринимая выходы M и генерируя корректирующие входы для M (обычная обратная связь в теории регулирования). Если G осуществляет свое воздействие сообразно со сложностью управляемой системы, то возникают устойчивые конфигурации и оба типа сложности совпадают. В противном же случае, т. е. когда воздействия G слишком слабы или чрезмерно велики, могут возникать различные нарушения.

Другие, более реалистические модели рассматриваются в гл. 4, где анализируются приложения теории сложности к следующим типам динамических систем:

- модели распределения ресурсов в конфликтных ситуациях,
- модели истощения природных ресурсов и загрязнения окружающей среды,
- структурные модели пространственно-временного развития.

В целом можно сказать, что сложность — многозначное понятие, включающее как статические и динамические аспекты, так и элементы, связанные с управлением. Статическая сложность, по существу, связана со сложностью подсистем, составляющих данную систему, а динамическая включает вычислительные машины или микропроцессорные элементы, что объясняется необходимостью выработки сигналов управления при наличии взаимосвязности подсистем. Наконец, сложность управляемых систем, по существу,

является мерой вычислительных возможностей, необходимых для реализации заданного поведения. В идеале математическая теория сложности должна достигнуть уровня, аналогичного уровню развития теории вероятностей. В то время как вероятность можно рассматривать как меру неопределенности в данной ситуации, сложность можно трактовать как меру понимания поведения системы.

УСТОЙЧИВОСТЬ

Наши системно-теоретические построения «покоятся» на трех китах: связность, сложность и устойчивость. Важность первых двух для понимания структуры системы была продемонстрирована нами достаточно наглядно. Что же касается устойчивости (или динамического поведения системы), то она практически еще не рассматривалась. Этот пробел можно восполнить, используя разнообразные понятия теории устойчивости.

К сожалению, термин *устойчивость* в высшей степени многозначен в литературе по системному анализу, будучи в постоянном употреблении для обозначения чего угодно, начиная с классической устойчивости по Ляпунову и кончая организационной жесткостью. Для всех возможных употреблений этого термина единственным общим моментом является интуитивное понимание того, что слово *устойчивый* обозначает, что нечто (может быть, система) способно реагировать на изменения в окружающей среде (например, возмущения, случайные помехи) и по-прежнему сохранять приблизительно то же самое поведение на протяжении определенного (возможно, бесконечного) периода времени. Совершенно ясно, что со столь нечетким и туманным «определением» устойчивости всякие попытки математического анализа устойчивости заведомо безнадежны. Тем не менее такое «определение» создает некоторую интуитивную основу для более точных определений.

Для большей ясности изложения удобно ввести две категории понятия устойчивости. Первую из них назовем *классической* и будем использовать ее для обозначения задач исследования результатов внешних воздействий на фиксированные системы, т. е. таких задач, когда изменяется только окружающая среда, но не сама система. В качестве простого примера подобной ситуации рассмотрим классический маятник (рис. 2.11).

Задача формулируется следующим образом: если сместить маятник из положения равновесия ($\theta = 0$) на некоторый угол, то может ли маятник вновь вернуться в положение

$\theta = 0$ за достаточно долгое, возможно бесконечное, время? Как из физических, так и из математических соображений очевидно, что так оно и будет для всех возмущений $\theta \neq 180^\circ$. Таким образом, $\theta = 0$ является положением устойчивого равновесия (по Ляпунову). Положение $\theta = 180^\circ$ есть положение неустойчивого равновесия, поскольку сколь угодно малое отклонение от него в конце концов приведет систему в положение устойчивого равновесия $\theta = 0$. Важно отметить, что величина начального смещения не влияет на динамику системы. Таким образом, налицо классическая ситуация, когда изменяется не структура системы, а лишь внешняя среда.

Классическая теория устойчивости в основном изучает равновесные состояния систем и динамику их поведения в малой окрестности этих состояний. Для исследования таких задач были разработаны весьма совершенные методы (гл. 5).

Подобные классические представления об устойчивости оказываются весьма плодотворными в физических и технических приключениях. Что касается их применения к анализу систем, изучаемых биологией, экономикой и общественными науками, то оно должно быть тщательно продумано и обосновано. Дело в том, что обычный режим функционирования подобных систем, как правило, далек от равновесного и, кроме того, внешние воздействия постоянно изменяют само равновесное состояние. Короче говоря, постоянные времени таких систем настолько велики, что во многих случаях ценность классического анализа устойчивости практически незаметна.

В отличие от классического равновесного подхода, центральным элементом современных взглядов на вопросы устойчивости является понятие *структурной устойчивости*. Здесь основной задачей является выявление качественных изменений в траектории движения при изменениях структуры самой системы. Таким образом, здесь изучается поведение данной системы по отношению к поведению всех «близких» к ней аналогичных систем. Если рассматриваемая система ведет себя «почти так же», как и «соседние», то говорят, что она структурно устойчива; в противном случае структурно неустойчива. Для уточнения этого понятия необходимо четко определить, что такое «близкая» система, каков класс допустимых возмущений и что значит схожесть поведения. Тем не менее основная идея остается прозрачной, достаточно

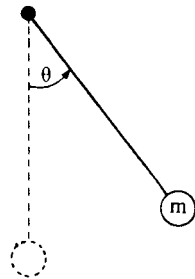


Рис. 2.11. Математический маятник

малые изменения структурно устойчивой системы должны приводить к соответственно малым изменениям в динамике ее поведения.

Пример. Простой гармонический осциллятор без трения

Динамика такой структурно-неустойчивой системы описывается уравнениями

$$\begin{aligned} \ddot{x} + c_1 \dot{x} + c_2 x &= 0, \\ x(0) &= a, \quad \dot{x}(0) = 0. \end{aligned}$$

Нас будет интересовать влияние параметров c_1 и c_2 на траекторию системы, причем из физических соображений ограничимся только случаями $c_1 \geq 0$, $c_2 > 0$.

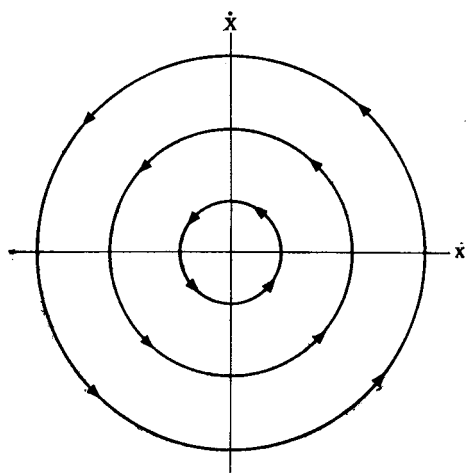


Рис. 2.12. Траектории осциллятора без трения.

Рассматривая траекторию осциллятора на фазовой плоскости (x, \dot{x}) , легко видеть, что если $c_1 = 0$, то все траектории являются концентрическими окружностями с радиусами $a\sqrt{c_2}$ и центром в начале координат (рис. 2.12). Если «ввести» в систему трение, то математически это означает, что $c_1 > 0$. Если $c_1^2 > 4c_2$, то точка равновесия $x = \dot{x} = 0$ на плоскости (x, \dot{x}) есть *узел* (рис. 2.13, а), в противном случае это *фокус* (рис. 2.13, б).

В обоих случаях начало координат является положением устойчивого равновесия по отношению к возмущениям в c_1 или c_2 . Эта ситуация резко контрастирует со случаем системы без трения ($c_1 = 0$), когда начало координат есть центр и качественная картина поведения изменяется при сколь угодно

но малых изменениях c_1 . Таким образом, при $c_1 \neq 0$ система структурно устойчива в том смысле, что качественный характер положения равновесия (узел, фокус) сохраняется при малых изменениях структуры системы.

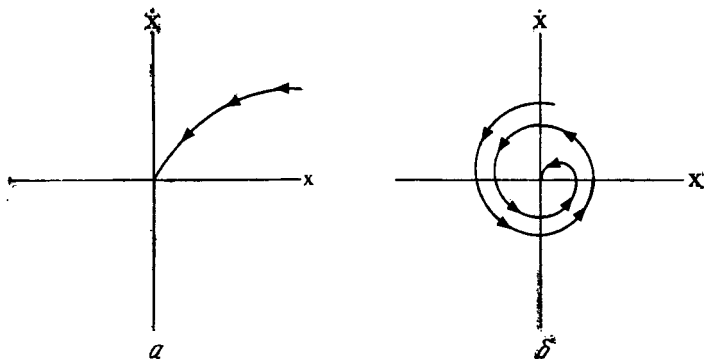


Рис. 2.13. Фазовый портрет траекторий системы на плоскости.

Поскольку идеи структурной устойчивости тесно связаны поведением траекторий системы по мере приближения к ее состоянию равновесия, представляет интерес рассмотреть те

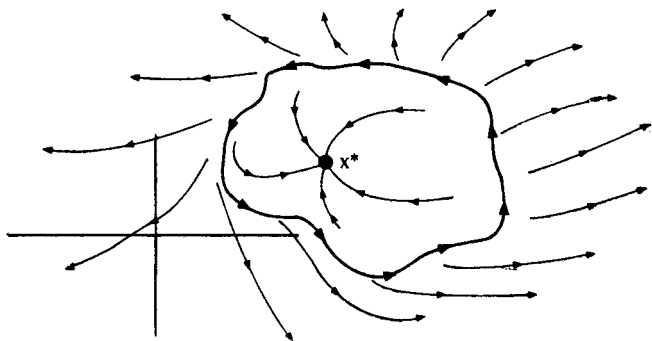


Рис. 2.14. Область притяжения фиксированной точки в R^2 .

области пространства состояний, которые соответствуют областям притяжения и отталкивания для данного состояния равновесия.

Иными словами, пусть задано равновесное состояние x^* , для простоты считающееся фиксированным. Из каких начальных состояний система в конце концов (при $t \rightarrow \infty$) придет в состояние x^* ? (Графически подобная ситуация изображена на рис. 2.14). Если допустить, что равновесные состояния

могут быть предельными циклами или периодическими траекториями, то даже в двумерном случае картина может быть довольно сложной. В случае более высокой размерности картина еще более запутана. Тем не менее вопрос об описании областей устойчивости и родственные вопросы, связанные со структурной устойчивостью, изучены достаточно глубоко.

Пример более сложной структурно неустойчивой системы.
Антисимметричная система хищник — жертва

Предположим, что m видов взаимодействуют с популяцией i -го вида, численность которой $N_i(t)$. Пусть a_i — коэффициент рождаемости i -го вида, а α_{ij} — коэффициент, характеризующий скорость уничтожения i -го вида j -м видом. Тогда динамика системы описывается уравнением Лотка — Вольтерра:

$$\frac{dN_i(t)}{dt} = N_i(t) \left[a_i - \sum_{j=1}^m \alpha_{ij} N_j(t) \right].$$

Нетривиальные равновесные популяции должны удовлетворять линейной системе алгебраических уравнений

$$\sum_{j=1}^m \alpha_{ij} N_j^* = a_i.$$

При неочевидном предположении, что матрица $A = [\alpha_{ij}]$ является антисимметрической, т. е. $\alpha_{ij} = -\alpha_{ji}$, можно показать, что при смещении системы из любого равновесного состояния ее поведение будет чисто колебательным, поскольку собственные значения любой кососимметрической матрицы чисто мнимые. (Следует отметить, что данное предположение означает, что коэффициент биохимического преобразования одного грамма жертвы j -го вида одинаков для всех хищников i -го вида, т. е. этот коэффициент не зависит от вида поедаемых особей.) Можно показать, что величина

$$Q = \sum_{i=1}^m [N_i(t) - N_i^* \log N_i(t)]$$

постоянна вдоль любой траектории системы.

Этот закон сохранения есть следствие колебательного характера поведения системы и является аналогом закона сохранения механической энергии простого гармонического осциллятора, рассмотренного выше. Однако, как только кососимметричность матрицы A нарушается, состояния равновесия системы становятся узлами или фокусами (устойчивыми

или неустойчивыми). В этом случае введение в систему сколь угодно малых изменений нарушает качественный характер траекторий, поэтому данная система структурно неустойчива. Более того, антисимметрические модели применимы только к системам с четным числом видов, поскольку из антисимметричности следует, что собственные значения матрицы A есть комплексно сопряженные числа. Если m нечетно, то действительное собственное значение матрицы A должно быть равным нулю, что приводит к вырожденности матрицы взаимодействий. Таким образом, данная система является структурно неустойчивой и в смысле вариации ее размерности.

КАТАСТРОФЫ И АДАПТИРУЕМОСТЬ

Положение равновесных состояний и соответствующих областей притяжения зависит от динамики изучаемой системы, поэтому важно знать, как они изменяются при небольшом изменении самой системы. Вопрос относительно того, приведет ли такое изменение к смещению данного состояния системы в другую область притяжения, представляет большой практический интерес, поскольку это привело бы к резким качественным изменениям в дальнейшем поведении системы. В качестве одного из инструментов исследования таких вопросов может быть использована теория катастроф.

Обычно в теории катастроф предполагается, что поведением изучаемого процесса управляет некоторая потенциальная функция, локальные минимумы которой соответствуют равновесным состояниям. Очень важно иметь в виду, что при таком подходе вовсе не обязательно точно знать, что это за функция — достаточно признать лишь сам факт ее существования. Предположим, далее, что можно измерять значения некоторых выходных переменных, генерируемых системой в ответ на входные воздействия. В «элементарной» теории катастроф предполагается, что все равновесные выходы фиксированы, т. е. фиксируют значения входных параметров, ждут (может быть, бесконечно долго!), пока не наступит равновесное состояние, затем изменяют значения входных переменных и снова ждут и т. д. Поступая таким образом, получают поверхность равновесных состояний в пространстве выходов, которую можно изобразить как функцию (многозначную) входов. В первом приближении можно сказать, что «катастрофа» происходит тогда, когда возникает скачкообразное изменение выходных параметров при непрерывном изменении входов.

Проиллюстрируем эту ситуацию на примере системы с двумя входами и одним выходом (катастрофа типа сборки с точкой возврата).

Пример. Теория «центрального места»

В экономико-географическом анализе одной из мер неравномерности распределения товаров и услуг в данном регионе является их уровень в «центральном месте». Разумеется, на этот уровень влияет множество факторов, однако наиболее существенными из них являются численность населения и расходуетый доход на душу населения. Предположим, что нас интересуют изменения равновесных уровней в

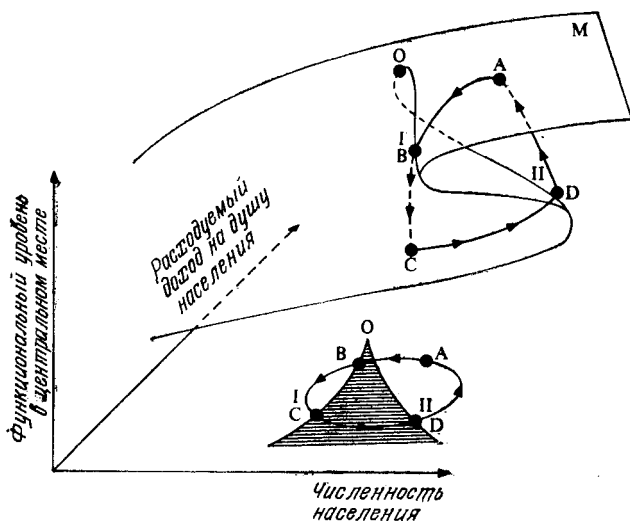


Рис. 2.15. Поверхность катастрофы в теории «центрального места».

«центральном месте» в зависимости от изменения численности населения и расходуетый доход в данном регионе. В том случае, когда истинный механизм функционирования данной экономической системы неизвестен, разумно предположить существование неизвестной нам потенциальной функции, локальным минимумом которой соответствуют равновесные уровни «центрального места» (при фиксированных входах).

Если входами являются численность населения и расходуетый доход на душу населения, а выходом уровень «центрального места» (безразмерная величина), то данную ситуа-

цию, согласно данным из гл. 5, можно изобразить как показано на рис. 2.15. Многообразие M изображает различные равновесные уровни «центрального места», как функцию двух переменных (входов). Наиболее интересным свойством поверхности M является наличие двух линий складок I и II, начинающихся в точке возврата O . На этом же рисунке изображена проекция данной картины на плоскость входов.

Рассматривая поверхность катастрофы M , легко установить, каким образом небольшие непрерывные изменения входных переменных могут привести к резким изменениям уровня «центрального места». Один из таких циклов ABCD изображен на рисунке. Уменьшение численности населения (при фиксированных доходах на душу населения) приводит к точке B , и при дальнейшем его уменьшении происходит скачкообразное уменьшение уровня (точка C). Если теперь численность населения будет увеличиваться, система непрерывно перейдет в состояние D , и далее произойдет скачок уровня к состоянию A . Следует подчеркнуть, что цикл ABCD не является динамическим движением системы! Это лишь последовательность равновесных состояний, полученных параметризацией входных параметров. Можно проследить описанные изменения, используя проекцию цикла ABCD на плоскость входных параметров. Отметим, что когда входы впервые попадают в заштрихованную область, никаких необычных изменений с уровнем не происходит. Он резко падает лишь в точке B , когда система покидает критическую область. Затем система возвращается в критическую область, и при этом с уровнем опять-таки не происходит никаких неожиданных изменений до тех пор, пока система не покидает эту область в точке D . Здесь следует обратить внимание на то, что скачки уровня (в ту или другую сторону) происходят только тогда, когда система покидает критическую область, пересекая линию складки, противоположную той, которую она пересекла при входе в область.

Для того чтобы связать эти рассуждения с нашими предыдущими замечаниями о важности анализа областей притяжения, заметим, что переход от одной области притяжения к области притяжения другого устойчивого состояния можно изобразить, как это показано на рис. 2.16. Точка x вначале принадлежит области притяжения состояния. Вследствие изменений динамики системы область притяжения P сужается с I до II, а область притяжения Q расширяется от 1 до 2. Теперь точка x притягивается к Q , а не к P . Конечно, положения P и Q сами зависят от структуры системы, поэтому объекты, изображенные на этом рисунке точками, по существу, являются областями, содержащими P и Q , но для

нас важно лишь то, что области P и Q отделены друг от друга. Следовательно, возмущения в структуре системы, приводящие к изображенной выше ситуации, порождают разрывы непрерывности в выходах, если наблюдаемые выходы оказываются равновесными состояниями.

Возвращаясь к теории катастроф, мы видим, что линии складок в точности соответствуют именно тем комбинациям входных параметров, которые приводят к скачкообразным изменениям равновесных состояний. Таким образом, теория

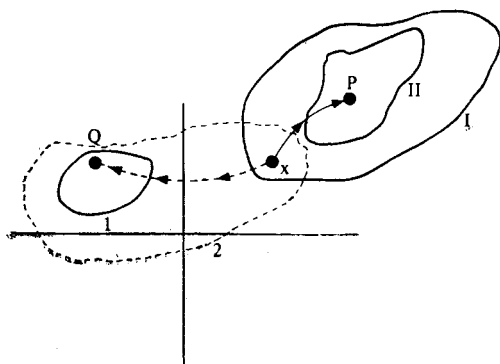


Рис. 2.16. Смещение областей притяжения.

катастроф позволяет геометрически описать изменения областей притяжения, не используя в качестве промежуточного этапа пространства состояний. Следует отметить, что теория катастроф, по существу, ничего не говорит о расположении областей притяжения в пространстве состояний. Она лишь позволяет нам определить, какие области в пространстве входных параметров могут привести к смещению областей притяжения. Этой информации часто оказывается достаточно для прикладных целей.

В заключение этой главы коснемся кратко понятия *адаптируемость*. Признано (в особенности экологами), что одним из наиболее желательных качественных свойств системы является ее способность воспринимать внешние воздействия (ожидаемые или неожиданные) без необратимых фатальных изменений в ее поведении. Иными словами, адаптируемость в некотором смысле является *мерой жизнеспособности* или *выживаемости системы*. Естественно, для формулировки этого понятия в математических терминах необходимо точно определить, какие воздействия считаются *допустимыми* и что следует понимать под *выживаемостью*. Тем не менее даже такое интуитивное описание адаптируемости показывает, что

это понятие тесно связано с понятием области притяжения и со смещением этих областей под действием естественных или искусственных возмущений. Если эти возмущения перемещают данное состояние системы в область притяжения «фатального состояния», то ясно, что система не обладает свойством адаптируемости по отношению к данному классу возмущений. В противном случае она в той или иной степени обладает этим свойством.

ЛИТЕРАТУРА

Управляемые и неуправляемые системы

Наиболее информативное и увлекательное обсуждение альтернативы *пассивное наблюдение — активное управление* дано в книге Bellman R., *Some Vistas of Modern Mathematics*, University of Kentucky Press, Lexington Kentucky, 1968.

Важность различия этих двух подходов трудно переоценить, поскольку в этом состоит основное идейное различие между современной теорией систем и классической теорией обыкновенных дифференциальных уравнений.

Идентификация

Общая проблема идентификации систем, конечно же, шире рассмотренных задач идентификации параметров и включает вопросы выделения уровней иерархии, размерностей подсистем и связей между ними и т. д. Блестящее нетехническое введение в философские основы систематического моделирования и идентификации дано в книгах

Weinberg G., *Introduction to General Systems Thinking*, Wiley, N. Y., 1975;

Vemuri V., *Modeling of Complex Systems: An Introduction*, Academic Press, New York, 1978.

Более специальное изложение современного состояния параметрической идентификации дано в работах

Mehra R. K., Lainiotis D., eds., *System Identification: Advances and Case Studies*, Academic Press, New York, 1976;

Rajbman N. S., *The Application of Identification Methods in the USSR: A Survey*, *Automatica*, 12, 73—95 (1976).

Astrom K., Eyekhoff P., *System Identification — A Survey*, *Automatica*, 7, 123—162 (1971).

Ограничения

Роль ограничений в многокомпонентной среде с различными целями обсуждается в книге

Lewin A. Y., Shakun M., *Policy Science: Methodologies and Cases*, Pergamon, N. Y., 1976.

Пример с фармакокинетикой взят из работы

Jeliffe R., et al. *A Mathematical Study of the Metabolic Conversion of Digoxin to Digoxin in Man*, USC Report EE-347, University of Southern California, Los Angeles, 1967.

С точки зрения наблюдаемости эта задача рассмотрена в книге Casti J., *Dinamical Systems and their Applications: Lin Theory*, Academic Press, New York, 1977.

Оптимизация

Из большого числа фундаментальных работ в области оптимального управления можно рекомендовать следующие:

- Kirk D., *Optimal Control Theory: An Introduction*, Prentice-Hall, Englewood Cliffs, New Jersey, 1971;
Bellman R., *Introduction to the Mathematical Theory of Control Processes*, Vols. 1, 2, Academic, New York, 1967, 1971;
Barnett S., *Introduction to Mathematical Control Theory*, Clarendon Press, Oxford, 1975;
Bryson A., Ho Y. C., *Applied Optimal Control*, Blaisdell, Waltham, Massachusetts, 1969.

Глобальные перспективы

Анализ глобальной структуры систем занимает центральное место в современной «общей» теории систем. Фундаментальные различия возможных подходов к анализу глобальной структуры рассматриваются в работах

- Koestler A., Janus, Random House, New York, 1978;
Koestler A., Smythies, eds., *Beyond Reductionism*, Macmillan, New York, 1969;
Weinberg G., *Introduction to General Systems Thinking*, Wiley, New York, 1975;
Bertalanffy L., von, *General System Theory*, Braziller, New York, 1963;
Klir G., *Trends in General System Theory*, Wiley, New York, 1972.
Ashton W. D., *The Theory of Road Traffic Flow* Methuen, London, 1966;
Dathe J., *Time-Delay Equations for Traffic Jams*, in Proc. Oper. Res. 1978, C. Schneeweiss, ed., Physica Verlag, Würzburg, 1979.

Автор признателен д-ру Дж. Десу за любезное разрешение использовать пример с транспортной магистралью.

Связность и графы

В большей части работ по связности систем для выявления компонент системы и связей между ними используются различные понятия из теории графов. Достаточно подробно эти вопросы обсуждаются в книгах Roberts F., *Discrete Mathematical Models*, Prentice-Hall, Englewood Cliffs, New Jersey, 1976;

- Harary F., Norman R., Cartwright D., *Structural Models: An Introduction to the Theory of Directed Graphs*, Wiley, New York, 1965.

Связность и симплициальные комплексы

Понятия связность, эксцентриситет и образ подробно обсуждаются в работе

- Atkin R. H., *Mathematical Structure in Human Affairs*, Heinemann, London, 1973.

Классические понятия алгебраической топологии, такие, как гомотопия, гомотопия и числа Бетти, изложены в книгах

- Giblin P., *Graphs, Surfaces and Homology*, Chapman & Hall, London, 1977;
Wallace A. D., *Introduction to Algebraic Topology*, Pergamon, London, 1957.

Сложность

Пример с экономикой взят из работы

Bristol E., The Counterintuitive Behavior of Complex Systems, IEEE Systems, Man & Cybernetics Newsletter, March 1975.

Более подробно понятия сложности управляемых и неуправляемых систем рассматриваются в статье

Gottinger H., Complexity and Information Technology in Dynamic Systems, *Kybernetes*, 4, 129—141 (1975).

Устойчивость

Из множества прекрасных книг по теории устойчивости можно порекомендовать следующие:

Bellman R., Stability Theory of Differential Equations, McGraw-Hill, New York, 1953;

Hahn W., Stability of Motion, Springer, Berlin, 1967;

Hirsh M., Smale S., Differential Equations, Dynamical Systems, and Linear Algebra, Academic Press, New York, 1974.

Ряд классических работ по устойчивости нелинейных систем содержится в сборнике

Aggrawal J. K., Vidyasagar M., eds., Nonlinear Systems: Stability Analysis, Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, 1977.

Катастрофы и адаптируемость

Современное (хотя и не установившееся) изложение теории катастроф можно найти в работах

Zeeman E., Catastrophe Theory: Selected Papers, 1972-77, Addison-Wesley, Reading, Massachusetts, 1977;

Poston T., Stewart I., Catastrophe Theory and Its Applications, Pitman, London, 1978.

Экономический пример с теорией «центрального места» взят из работы

Casti J. Swain H., Catastrophe Theory and Urban Systems, IFIP Conf. on Optimization, Nice, France, September 1975.

Общее обсуждение этого вопроса дается в работах

Holling C. S., Resilience and Stability of Ecosystems, *Annual Rev. Ecol. Syst.*, 4, 1—23 (1973);

Watters C., Foreclosure of Options in Sequential Resource Development Decisions, RR-75-12, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975.

Связность структуры больших систем

Наибольшую ценность представляет процесс познания простых форм и различий вещей, поскольку их различные комбинации создают все многообразие этого мира.

Фрэнсис Бэкон

Человек следует [законам] земли.

Земля следует [законам] неба.

Небо следует [законам] дао.

Дао следует самому себе.

Лао-Цзы. «Дао дэ цзин»

В популярной литературе по системному анализу под системой часто подразумевают совокупность взаимодействующих элементов. Хотя подобное представление едва ли можно рассматривать как исчерпывающее определение понятия *система*, тем не менее оно все же уточняет наше умозрительное представление о системе как о совокупности *связного* набора объектов. Более того, можно утверждать, что именно *связность* является сущностью понятия *большая система*, так как а fortiori система, компоненты которой не взаимодействуют, вряд ли может представлять интерес как физический процесс, не говоря уже о том, что анализ такой системы представляет собой тривиальную аналитическую задачу.

В настоящей главе обсуждаются вопросы, касающиеся описания связности структуры системы на языке математики. Поскольку в предыдущих главах были рассмотрены различные типы математических определений системы, то и вопрос о связности необходимо рассмотреть с различных точек зрения. При этом, однако, надо иметь в виду, что все эти различные концепции связности отражают единую тенденцию — выявление существенных, функционально-значимых связей системы.

Рассмотрим систему «черный ящик», которая описывается с помощью системы линейных алгебраических уравнений

$$Ax = b,$$

причем каждая компонента вектора x представляет собой некоторую подсистему, а вектор b — вход системы. В этом случае взаимодействие между подсистемами будет определяться внедиагональными элементами матрицы A и любой

анализ связности структуры процесса должен концентрироваться на изучении степени заполненности матрицы и значимости ее элементов.

Предположим теперь, что исследуемая система качественно описывается планарным графом, в котором дуга между двумя вершинами указывает на существование связи между соответствующими подсистемами. Тогда вместо вышеприведенной системы уравнений получим адекватное описание системы «черный ящик» в виде матрицы E системных взаимосвязей, (i, j) -й элемент которой равен 1, если подсистемы i и j связаны, и равен 0 в противном случае. Отметим, что внедиагональные элементы матрицы E и здесь играют важную роль, хотя для полного понимания связности структуры надо привлечь не только алгебру, но и топологию.

Таким образом, уже предварительный анализ показывает, что связность структуры системы — это сложное, многогранное понятие, требующее для своего описания применения аппарата как алгебры, так и топологии. В зависимости от способа описания системы проявляются те или иные свойства связности. Наша задача — рассмотреть наиболее интересные проблемы связности и попытаться дать ряд математических методов для их разрешения. Так как связность является, по существу, *алгебраическим* понятием, данная глава довольно насыщена математическими конструкциями и понятиями абстрактной алгебры и топологии.

КОМПЛЕКСЫ И СВЯЗИ

Из соображений геометрической наглядности начнем изучение связности структуры системы с рассмотрения комплексов. Согласно определению, данному в гл. 1, симплициальный комплекс — это естественное математическое обобщение понятия планарного графа, отражающее многомерную природу рассматриваемого бинарного отношения. Поскольку симплициальный комплекс, по существу, не что иное, как семейство симплексов, соединенных посредством общих граней, то естественной характеристикой связности могла бы служить размерность грани, общей двум симплексам. Однако нас интересует комплекс в целом, поэтому более целесообразно использовать понятие *цепь связи*, отражающее тот факт, что два симплекса могут не иметь общей грани, но могут быть связаны при помощи последовательности промежуточных симплексов. С учетом соображения размерности понятие *q-связности* может быть сформулировано следующим образом.

Определение 3.1

Данные два симплекса σ_i и σ_j комплекса K соединены цепью q -связи, если существует последовательность симплексов $\{\sigma_{a_i}\}_{i=1}^n$ в K , такая, что

σ_{a_1} — грань σ_i ,

σ_{a_n} — грань σ_j ,

σ_{a_i} и $\sigma_{a_{i+1}}$ обладают общей гранью размерности β для $i = 1, 2, \dots, n-1$,

$$q = \min \{i, \beta_1, \beta_2, \dots, \beta_n, j\}$$

(нижний индекс симплекса соответствует его геометрической размерности, например $\dim \sigma_s = s$). Можно показать, что q -связность порождает отношение эквивалентности на симплексах K .

Таким образом, задача изучения глобальной структуры связности комплекса K сводится к рассмотрению классов q -эквивалентности. Для каждого значения размерности $q = 0, 1, \dots, \dim K$ можно определить число различных классов эквивалентности Q_q . Назовем эту операцию q -анализом комплекса K , а вектор $Q = (Q_{\dim K}, \dots, Q, Q_0)$ — *первым структурным вектором комплекса*.

Информация, содержащаяся в Q , до некоторой степени отражает глобальную структуру комплекса K . Для того чтобы два симплекса A и B принадлежали одной q -связной компоненте комплекса K , необходимо наличие связывающей A и B цепи промежуточных симплексов, такой, чтобы «слабейшее» в смысле размерности звено имело размерность, большую или равную q .

Из определения 3.1 следует, что если два симплекса q -связаны, то они также $q-1, q-2, \dots, 0$ -связаны в комплексе K . Следовательно, комплекс K можно рассматривать как множество многомерных трубок симплексов, а компоненты вектора Q — как число трубок каждой размерности в K .

Иначе q -связность можно представить так. Предположим, что мы обозреваем комплекс через очки, которые дают возможность видеть его только в размерности q и выше. Тогда мы увидим комплекс K разбитым на Q_q дизъюнктивных (несвязных) кусков. Отметим, что число Q_0 идентично применяемому в топологии нульмерному числу Бетти, другие $Q_q, q \geq 1$ не совпадают с q -мерными числами Бетти. Если K обозревается на всех уровнях размерности, то Q_0 дает число несвязных компонент K .

Чтобы пояснить понятие *комплекс*, рассмотрим приведенный в гл. I пример системы, описывающей сферу обслуживания условного города. Предположим, что множество

$$X = \{\text{хлеб, молоко, марки, обувь}\}$$

представляет интересующие нас товары, а множество

$$Y = \{\text{гастроном, универмаг, банк, почта}\}$$

— предприятия сферы обслуживания. При этом возникает естественное отношение $\lambda \subset Y \times X$, связывающее эти два множества:

$(y_i, x_j) \in \lambda$ тогда и только тогда, когда товар x_j можно получить на предприятии y_i .

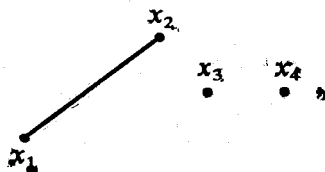
Отсюда следует, что

$$\lambda = \{(y_1, x_1), (y_1, x_2), (y_2, x_4), (y_4, x_3)\}.$$

Матрица инциденций Λ отношения λ имеет вид

λ	x_1	x_2	x_3	x_4
y_1	1	1	0	0
$\Lambda = y_2$	0	0	0	1
y_3	0	0	0	0
y_4	0	0	1	0

Геометрически комплекс может быть представлен следующим образом:



где X — множество вершин, Y — множество симплексов.

Заметим, что *пустой* симплекс y_3 не принадлежит $K_Y(X; \lambda)$, если только не принять соглашение пополнять комплекс, добавляя к K пустую вершину ϕ , представляющую (-1) -мерный симплекс.

Вышеприведенный комплекс, как это очевидно из его геометрического изображения, состоит из 1-симплекса y_1 и двух 0-симплексов y_2 и y_4 . Это означает, что данная система

обнаруживает очень низкий уровень связности. При рассмотрении комплекса K можно видеть, что $Q_1 = 1$ (симплекс y_1), в то время как $Q_0 = 3$ (дизъюнктные 0-компоненты — это симплексы y_1, y_2, y_4). Следовательно, $Q = (1 \ 3)$ есть первый структурный вектор этого комплекса.

ЭКСЦЕНТРИСИТЕТ

Несмотря на то что q -анализ оказывается довольно эффективным при изучении глобальной связности структуры, тем не менее он не дает необходимой информации о том, как каждый отдельный симплекс входит в весь комплекс. Поскольку, однако, индивидуальные свойства симплексов могут оказаться весьма существенными в рассматриваемой проблеме, важно уметь определять степень *интегрированности* каждого отдельного симплекса в структуре всего комплекса. С этой целью введем понятие *эксцентриситет*.

Определение 3.2

Эксцентриситет симплекса σ задается следующей формулой:

$$\text{ecc}(\sigma) = \frac{\hat{q} - \check{q}}{\check{q} + 1},$$

где \hat{q} — размерность симплекса σ , а \check{q} — наибольшее значение q , при котором σ становится связанным с каким-либо другим симплексом из K .

Разность $\hat{q} - \check{q}$ является *мерой необычности (нонконформности)* симплекса σ , при этом равенство $\hat{q} - \check{q} = 2$, по-видимому, информативно более значимо, когда $\check{q} = 1$, а не $\check{q} = 10$. Поэтому в качестве меры эксцентриситета будем использовать вышеприведенное отношение, а не абсолютную разность $\hat{q} - \check{q}$. Это соответствует нашему представлению о максимально эксцентричном симплексе как полностью изолированном от всех остальных.

Вернемся к комплексу, рассмотренному в предыдущем разделе. Вычисляя величины $\text{ecc}(y_1) = \infty$, $\text{ecc}(y_2) = \infty$, $\text{ecc}(y_4) = \infty$, мы видим, что каждый симплекс данного комплекса полностью изолирован от остальных.

ДЫРЫ И ПРЕПЯТСТВИЯ

Как было отмечено выше, q -анализ симплициального комплекса дает нам информацию о многомерных цепях связей симплексов, составляющих комплекс K . Особый интерес, однако, представляет вопрос о структуре, образуемой этими

цепями. Можно представить себе комплекс K в виде воображаемого многомерного швейцарского сыра с цепями q -связей, формируемыми его содержимым. В этом случае наша задача сводится к исследованию структуры дырок в таком сыре. Изучение многомерных дырок в комплексе на языке алгебраической топологии является прерогативой теории гомологий, которая оперирует такими понятиями, как *цепь*, *граница*, *группа гомологий*.

Ограничимся рассмотрением отношений между двумя конечными множествами X и Y , а именно $\lambda \subset Y \times X$ и $\lambda^* \subset X \times Y$. В этом случае оба симплициальных комплекса $K_Y(X; \lambda)$ и $K_X(Y; \lambda)$ имеют конечную размерность и конечное число симплексов.

Рассмотрим комплекс $K_Y(X; \lambda)$ с $\dim K = n$. Предположим, что на K задана ориентация, индуцированная упорядочением множества вершин X , т. е. задана нумерация вершин $(x_1, \dots, x_i, \dots, x_{k-1}, x_k)$, где $k \geq n$. Для любого целого числа p , такого, что $0 \leq p \leq n$, будем обозначать симплекс размерности p через σ_p^i , $i = 1, 2, 3, \dots, h_p$, а число всех p -симплексов в K — через h_p .

Образуем формальные линейные суммы этих p -симплексов, допуская кратность для любого σ_p . Любую такую комбинацию назовем p -цепью. Обозначим семейство всех p -цепей через C_p , а любой элемент цепи C_p — через c_p . Тогда типичная p -цепь имеет вид

$$c_p = m_1 \sigma_p^1 + m_2 \sigma_p^2 + \dots + m_{h_p} \sigma_p^{h_p},$$

где каждое $m_i \in J$, а J — произвольная абелева группа. Теперь можно рассматривать множество C_p как абелеву группу относительно операции «+» в предположении, что

$$c_p + c'_p = (m_1 + m'_1) \sigma_p^1 + \dots + (m_{h_p} + m'_{h_p}) \sigma_p^{h_p}.$$

Нулем группы является цепь 0_p , у которой каждое $m_i = 0$. Возьмем прямую сумму групп C_p , $p = 1, 2, \dots, n$, записывая получаемую при этом группу C как

$$C = C_0 \oplus C_1 \oplus C_2 \oplus \dots \oplus C_n.$$

Любой элемент из C имеет вид

$$c = c_0 + c_1 + \dots + c_n.$$

Каждой p -цепи (c_p) сопоставим определенную $(p-1)$ -цепь, которую обозначим как dc_p и назовем *границей*. Определим dc_p с помощью $d\sigma_p$ симплексов, линейная комбинация которых образует цепь c_p , если $c_p = \sum_i m_i \sigma_p^i$, то $dc_p = \sum_i m_i d\sigma_p^i$. Другими словами, потребуем, чтобы d был гомоморфизмом из C_p

в C_{p-1} . Для любого симплекса $\sigma_p = \langle x_1 x_2 \dots x_{p+1} \rangle$ можно определить $\partial\sigma_p$ следующим образом:

$$\partial\sigma_p = \partial \langle x_1 x_2 \dots x_{p+1} \rangle = \sum_i (-1)^{i+1} \langle x_1 x_2 \dots \hat{x}_i \dots x_{p+1} \rangle,$$

где \hat{x}_i означает, что вершина x_i пропущена.

Геометрическое изображение 2-симплекса $\sigma_2 = \langle x_1 x_2 x_3 \rangle$ вместе с ориентацией и индуцированными ориентациями на ребрах дается на рис. 3.1. В этом случае

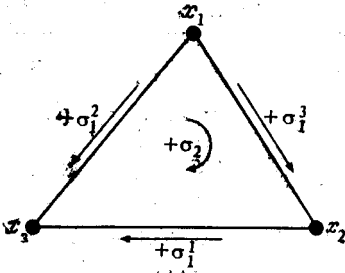


Рис. 3.1. 2-симплекс с ориентированными гранями.

$$\partial\sigma_2 = \partial \langle x_1 x_2 x_3 \rangle = (-1)^2 \langle x_2 x_3 \rangle + (-1)^3 \langle x_1 x_3 \rangle + (-1)^4 \langle x_1 x_2 \rangle,$$

т. е. $\partial\sigma_2 = \sigma_1^1 - \sigma_1^2 + \sigma_1^3$, является 1-цепью (элементом C_1).

Границу любой цепи можно рассматривать как образ этой цепи относительно оператора ∂ , который задает отображение $\partial: C_p \rightarrow C_{p-1}$ для $p = 1, \dots, n$. Оператор ∂ не

только гомоморфен (он сохраняет аддитивную структуру), но и, как легко показать, нильпотентен, так как $\partial(\partial\sigma_p) = 0$ в C_{p-2} , или, что то же самое, $\partial^2 = 0$ (тривиальное отображение).

Для случая, изображенного на рис. 3.1, имеем

$$\begin{aligned} \partial^2\sigma_2 &= \partial(\partial\sigma_2) = \partial(\sigma_1^1 - \sigma_1^2 + \sigma_1^3) = \\ &= \partial \langle x_2 x_3 \rangle - \partial \langle x_1 x_3 \rangle + \partial \langle x_1 x_2 \rangle = \\ &= \langle x_3 \rangle - \langle x_2 \rangle - (\langle x_3 \rangle - \langle x_1 \rangle) + \langle x_2 \rangle - \langle x_1 \rangle = 0. \end{aligned}$$

Так как $\partial: C_p \rightarrow C_{p-1}$ — гомоморфизм, то образ C_p относительно ∂ должен быть подгруппой C_{p-1} . Для обозначения этого образа будем использовать следующие символы: $\text{im } \partial$ или B_{p-1} , или ∂C_p . Из нильпотентности ∂ вытекает, что $\partial B_{p-1} = 0$ в C_{p-2} , или, что то же самое, $\partial(\text{im } \partial) = 0$.

Назовем p -циклами те цепи $c_p \in C_p$, у которых границы исчезают (т. е. $\partial c_p = 0$). Такие цепи образуют подгруппу группы C_p , обозначаемую символом Z_p и являющуюся ядром гомоморфизма ∂ . Очевидно, что элементы B_p , или, что то же самое, ∂C_{p+1} , являются циклами и, следовательно, $B_p \subset Z_p$. Фактически же B_p является подгруппой Z_p .

Элементы B_p называют *граничными циклами* (они являются циклами в обычном или тривиальном смысле). Те элементы Z_p , которые не являются элементами B_p , можно

отождествить с элементами фактор-группы Z_p/B_p . Любой элемент этой фактор-группы имеет вид $z_p + B_p$, и если выбрать один представитель, например z_p , из этого класса эквивалентности, то его можно обозначать $[z_p]$. Будем говорить, что два p -цикла z_p^1 и z_p^2 гомологичны (часто пишут $z_p^1 \sim z_p^2$), если z_p^1 и z_p^2 различаются только p -границей, т. е. $z_p^1 - z_p^2 \in B_p$. Легко показать, что таким образом задаваемое отношение на множестве циклов является эквивалентностью. Фактор-группа Z_p/\sim , построенная по отношению быть гомологичным, — это фактор-группа Z_p/B_p , групповая структура

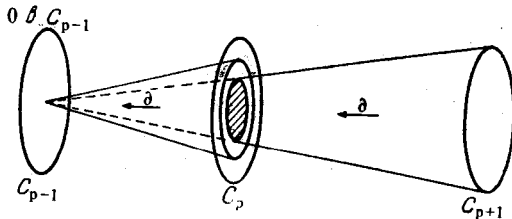


Рис. 3.2. Действие нильпотентного оператора δ на градуированной группе \hat{C} . Заштрихованная область, по форме напоминающая глаз быка, представляет группу B_p , а внутреннее кольцо, окружающее эту область, — группу циклов Z_p .

которой определена операцией «+» на элементах $z_p + B_p$. В этой групповой структуре множество B_p действует как нуль, т. е.

$$(z_p + B_p) + B_p = z_p + B_p$$

для всех z_p .

Фактор-группа Z_p/B_p называется p -группой гомологий и обозначается как

$$H_p = Z_p/B_p, \quad p = 0, 1, \dots, n.$$

Если учесть, что группа Z_p — это ядро гомоморфизма δ ($Z_p = \ker \delta$), то группу гомологии можно представить следующим образом:

$$H_p = \ker \delta / \text{im } \delta.$$

Операция δ на градуированной группе C может быть записана в виде последовательности

$$C = C_0 \oplus C_1 \oplus C_2 \oplus \dots \oplus C_p \xrightarrow{\delta} C_{p+1} \oplus \dots \oplus C_n.$$

Схематически эта операция изображена на рис. 3.2. Когда $H_p = 0$, фактор-группа Z_p/B_p имеет единственный класс эквивалентности и им является группа B_p ; каждый цикл $z_p \in B_p$, т. е. каждый цикл является граничным. Когда же

$H_p \neq 0$, в фактор-группе существует более чем один элемент и, значит, должен быть по крайней мере один цикл, который не является граничным циклом. Для примера, изображенного на рис. 3.1, имеем $H_1 = 0$, так как единственным 1-циклом является комбинация $\sigma_1^1 - \sigma_1^2 + \sigma_1^3$ (и ее кратные), которая представляется в виде $\partial\sigma_2$. Поскольку C_3 тривиальна, то и B_2 тоже тривиальна, а так как $\partial\sigma_2 \neq 0$, то и Z_2 тривиальна. В силу этих условий $H_2 = 0$. Если $H_p = 0$, то говорят, что гомология в размерности p тривиальна. Если говорят, что «гомология тривиальна» и не указывают значения p , то подразумевается, что $H_p = 0$ для всех значений p , отличных от нуля. Эта последняя группа всегда нетривиальна, за исключением случая, когда комплекс K дополнен симплексом, множество вершин которого пусто.

Можно показать, что для случая, представленного на рис. 3.1, гомология тривиальна, а $H_0 \neq 0$. Действительно, полагая границу вершины нулевым элементом группы для любой цепи c_0 вида

$$c_0 = m_1 \langle x_1 \rangle + m_2 \langle x_2 \rangle + m_3 \langle x_3 \rangle,$$

получим, что цепь c_0 должна быть 0-циклом, т. е. $c_0 \in Z_0$. Однако вершины x_1, x_2, x_3 образуют часть линейно-связной структуры в том смысле, что существуют такие 1-цепи c_1, c_1' , что

$$\langle x_2 \rangle = \langle x_1 \rangle + \partial c_1, \quad \langle x_3 \rangle = \langle x_1 \rangle + \partial c_1'$$

(действительно, достаточно положить хотя бы $c_1 = \sigma_1^3, c_1' = \sigma_1^2$). В результате имеем

$$c_0 = z_0 = (m_1 + m_2 + m_3) \langle x_1 \rangle + \partial (\text{некоторой 1-цепи}).$$

Отсюда следует, что вершина x_1 действует подобно специально выбранному 0-циклу z_0 и все возможные 0-циклы могут быть получены из формулы

$$z_0 = m z_0 + \partial (\text{некоторой 1-цепи}),$$

где цикл z_0 , состоящий из отдельной вершины, не может быть границей какой-нибудь 1-цепи. Следовательно, $z_0 \notin B_0$ и $H_0 \neq 0$. В действительности H_0 содержит единственную образующую и является аддитивной группой, изоморфной группе J , которая порождается одним своим элементом, а именно 1. Таким образом, для комплекса, представленного на рис. 3.1, имеем

$$H_0 = J,$$

или, используя символ изоморфизма, $H_0 \simeq J$.

Вышеприведенные соображения показывают, что такая структура характерна для всех линейно-связных комплексов. Это позволяет сделать следующий вывод: если комплекс K имеет k связных компонент, то 0-я группа гомологий может быть представлена в виде прямой суммы

$$H_0(K) = J \oplus J \oplus \dots \oplus J$$

с k слагаемыми. Числу k в топологии соответствует нульмерное число Бетти комплекса k , которое обозначается как β_0 .

Все рассмотренные группы C_p , Z_p , B_p являются примерами конечно порожденных свободных групп, так как между образующими любой из них не могут существовать линейные зависимости. Однако выполнение свойства «быть свободной» для фактор-группы H_p не обязательно.

В общем случае H_p может быть представлена как прямая сумма двух частей, из которых одна является свободной группой, а другая нет. Для пояснения этого соображения запишем H_p в виде

$$H_p = G_p^0 \oplus \text{Тог } H_p,$$

где G_p^0 — свободная группа, а $\text{Тог } H_p$ — подгруппа кручения группы H_p . Любой элемент h группы $\text{Тог } H_p$ обладает тем свойством, что $nh = 0$ для некоторого целого числа n . В терминах границ и циклов это означает, что элемент h может быть записан в виде $h = z_p + B_p$ (так как $h \in H_p$), при этом существует n , такое, что элемент

$$nh = nz_p + nB_p$$

должен уже содержаться в B_p (нуле фактор-группы). Отсюда следует, что хотя $z_p \notin B_p$, тем не менее для некоторого значения n должно выполняться $nz_p \in B_p$. Такое своеобразное поведение закрученных циклов характеризуется подгруппой $\text{Тог } H_p$.

Элементы свободной группы G_p^0 не могут вести себя подобным образом: если $z_p \in G_p^0$ и $z_p \notin B_p$, то и $nz_p \notin B_p$ при любом ненулевом значении n . Группа G_p^0 , очевидно, должна состоять из слагаемых вида J (число слагаемых должно равняться числу различных образующих G_p^0), в то время как подгруппа $\text{Тог } H_p$ должна состоять из слагаемых вида J_m (если J — множество целевых чисел, то J_m — группа вычетов по модулю m) для некоторых m . Это происходит потому, что группы, подобные J_m , являются аддитивными абелевыми группами со свойством $h \in J_m \Leftrightarrow mh = 0$. Каждая подгруппа

группы $\text{Тог } H_p$ должна быть изоморфной некоторой J_m для подходящего m . Число образующих (число свободных образующих H_p) называют p -м числом Бетти комплекса K и обозначают β_p .

p -ДЫРЫ

Геометрическое изображение комплекса, имеющего тривиальную гомологическую структуру, было дано на рис. 3.1. Для этого комплекса $H_1 = 0$, так как в его состав включен треугольник σ_2 . Если удалить внутренность треугольника σ_2 и оставить только его ребра, то такое изменение в строении комплекса K вызывает соответствующее изменение в группе H_1 . А именно, теперь $H_1 = J$ потому, что появилась единственная образующая группы H_1 вида

$$\sigma_1^1 - \sigma_1^2 + \sigma_1^3,$$

которая является циклом, но не является границей (σ_2 удален). Следовательно, факт появления дыры в K , ограниченной 1-симплексами, проявляется в том, что группа H_1 приобретает единственную образующую. Такую дыру в комплексе K будем называть *1-мерной дырой*. Пусть теперь комплекс K состоит из двух треугольников без внутренних частей. В этом случае группа H_1 будет изоморфна прямой сумме двух групп коэффициентов J , т. е. $H_1 \simeq J \oplus J$. Аналогично, если геометрическое изображение комплекса K имеет одну сферическую дыру (т. е. дыру, ограниченную сферой), то получим, что H_2 обладает единственной образующей $z_2 \in B_2$. Если для комплекса K существует группа $H_2 = J \oplus J$, то это означает, что K имеет в точности две 2-мерные дыры.

Нам хотелось бы особо подчеркнуть тот факт, что свободная группа G_p^0 интерпретируется как алгебраическое представление наличия p -мерных дыр в комплексе K . Точное число этих дыр задает p -е число Бетти.

Рассмотренный выше анализ q -связности предназначен для изучения топологической структуры пространства «между дырами воображаемого многомерного швейцарского сыра». Подгруппу кручения $\text{Тог } H_p$, если она нетривиальна, трудно интерпретировать в таком контексте, однако для нее можно найти другую интерпретацию.

Пример

Обозначим грани игральной кости символами $v^1, v^2, v^3, v^4, v^5, v^6$. Сопоставим их с вершинами некоторого 5-симплекса, а в качестве комплекса K будем рассматривать этот симп-

лекс и все его грани, например типичная 1-грань — это 1-симплекс $\langle v^i v^j \rangle$, где $i \neq j$. Зададим индуцированную ориентацию K с помощью естественного упорядочения вершин. Теперь проведем серию экспериментов, в которых игральная кость бросается до тех пор, пока не повторится одна и та же грань. (Последовательность $+\{v^i, v^j\}$ будем в алгебраических преобразованиях считать равной $-\{v^j, v^i\}$.) Любой результат серии последовательных бросаний можно рассматривать как элемент градуированной группы цепей

$$C_* = C_0 \oplus C_1 \oplus C_2 \oplus C_3 \oplus C_4 \oplus C_5.$$

Заметим, что границей последовательности $\langle 1\ 2\ 3 \rangle$ является 1-цепь $\langle 12 \rangle$, $\langle 23 \rangle$, $\langle 31 \rangle$.

Вначале предположим, что экспериментатор может наблюдать любую возможную последовательность и серии последовательностей. Следовательно, в этом случае в градуированной группе цепей каждый цикл является границей и поэтому

$$H_p = 0, \quad p = 1, 2, 3, 4,$$

т. е. гомология такого комплекса K тривиальна.

Теперь предположим, что экспериментатор использует помощника, который его обманывает (подделывая записи). Допустим, что в результате обмана последовательность $\langle 1\ 2\ 3 \rangle$ не встречается ни сама, ни в качестве грани какой-либо последовательности. Подобные действия помощника приводят к сильному изменению в комплексе K и в связанной с ним группе цепей. Например, последовательность $\langle 1\ 2\ 3\ 4\ 5\ 6 \rangle$ не будет встречаться, так как она содержит $\langle 1\ 2\ 3 \rangle$. Кроме того, в новом комплексе K' появится цикл

$$z_1 = \langle 12 \rangle + \langle 23 \rangle + \langle 31 \rangle,$$

который не является границей. Следовательно вмешательство помощника привело к возрастанию первого числа Бетти β_1 от 0 до 1. Именно помощник является виновником появления дыры в комплексе; группа гомологий H_1 стала изоморфна J .

Изменим порядок проведения эксперимента еще раз. Предположим, что эксперимент проводят два опытных игрока. Бросая кость, они замечают, что вероятность различных последовательностей, соответствующих типичным симплексам $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5$, равны соответственно $5/6, 5/9, 5/18, 5/54$ и $5/324$. Так как игроки намерены держать пари по результатам эксперимента, то они соглашаются приписать

симплексам некоторые веса, которые уравнивают шансы. Это вызывает введение новых (*взвешенных*) симплексов, порождающих новую группу цепей C' . Новые образующие связаны со старыми следующим образом:

$$\sigma'_1 = 54 \sigma_1; \quad \sigma'_2 = 36 \sigma_2; \quad \sigma'_3 = 18 \sigma_3; \quad \sigma'_4 = 6 \sigma_4; \quad \sigma'_5 = \sigma_5.$$

Теперь гомологии комплекса изменятся еще больше; например элемент

$$54(\langle 12 \rangle + \langle 23 \rangle + \langle 31 \rangle)$$

находится в Z'_1 , но не является элементом группы B'_1 , так как последняя состоит из элементов кратных $108 \sum_i \sigma'_i$ (108 — наименьшее общее кратное чисел 36 и 54). Отсюда получаем, что в Z'_1 существует цикл z_1 , такой, что $z_1 \notin B'_1$, а $2z_1 \in B'_1$. Это вносит в группу гомологий H_1 слагаемое, равное J_2 , а группа кручения $\text{Тог } H_1$ становится нетривиальной. В действительности

$$H_1 = J_2 \oplus J_2 \oplus \dots \oplus J_2,$$

где число слагаемых в сумме равно десяти. Остальные H_p при этом не изменились, $H_p = 0$, $p = 2, 3, 4$.

Комплекс игроков уже обладает кручением, что и отразилось в изменении H_1 . Теперь становится ясно, что кручение может быть введено в $H(K)$ разными путями, которые могут давать различные слагаемые J_m в зависимости от изменения кратности выхода эксперимента. Например, введение $\sigma'_1 = 48\sigma_1$ вызывает появление десяти слагаемых J_3 , при этом

$$H_1 = J_3 \oplus J_3 \oplus \dots \oplus J_3.$$

КОЦЕПИ И КОГРАНИЦЫ

Группе цепей C с коэффициентами из J можно сопоставить двойственное понятие, а именно группу отображений из C в J . Это позволяет ввести понятия *коцепи*, двойственное понятию цепи. Любая коцепь представляет собой гомоморфизм из C в J :

$$c^p : C_p \rightarrow J.$$

Точнее потребуем выполнения аддитивности, а именно:

$$c^p(c_p + c'_p) = c^p(c_p) + c^p(c'_p),$$

где c^p обозначает p -коцепь. Благодаря этому можно строить любую p -коцепь c^p , исходя из множества отображений p -симплексов в J . В связи с этим целесообразно ввести понятие косимплекса. Косимплексом называют отображение

$$\sigma^p : \{\sigma_p^i\} \rightarrow J.$$

Если в K входит в точности h_p p -симплексов, можно определить базис косимплексов как множество отображений $\{\sigma_p^i : i = 1, 2, \dots, h_p\}$, где

$$\sigma_p^i(\sigma_p^j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

Это позволяет любой косимплекс σ^p записать в виде суммы косимплексов σ_p^i , т. е.

$$\sigma^p = \sum_i \sigma_p^i,$$

а любую p -коцепь представить в виде линейной комбинации

$$c^p = \sum_i m_i \sigma_p^i.$$

Нулевая коцепь (для любого p) — это коцепь, у которой все $m_i = 0$. Таким образом, множество всех p -коцепей образует аддитивную группу коцепей C^p . Взяв прямую сумму

$$C^* = C^0 \oplus C^1 \oplus \dots \oplus C^n,$$

получим градуированную группу коцепей, где $n = \dim K$.

Для завершения двойственных рассмотрений определим кограничный оператор δ , сопряженный граничному оператору d . Будем обозначать $c^p(c_p)$ также с помощью внутреннего произведения (c_p, c^p) . Определим $\delta: C^p \rightarrow C^{p+1}$ при помощи равенства

$$(\partial c_{p+1}, c^p) = (c_{p+1}, \delta c^p).$$

Ясно, что оператор δ также будет нильпотентным ($\delta^2 = 0$), так как

$$\begin{aligned} 0 &= (0, c^p) = (\partial^2 c_{p+2}, c^p) = \\ &= (\partial c_{p+2}, \delta c^p) = \\ &= (c_{p+2}, \delta^2 c^p) \quad \text{для любого } c_{p+2}, \end{aligned}$$

и, следовательно, $\delta^2 c^p$ может быть только тривиальным отображением.

Теперь можно определить когомологическую группу $H^p(K; J)$, задаваемую равенством

$$H^p = Z^p / B^p = \ker \delta / \text{im } \delta.$$

СИСТЕМА ХИЩНИК—ЖЕРТВА

Исследование классической экосистемы хищник — жертва может служить прекрасной иллюстрацией применения вышеприведенных гомологических рассуждений. Рассмотрим биологическое сообщество, состоящее из 15 различных видов, и

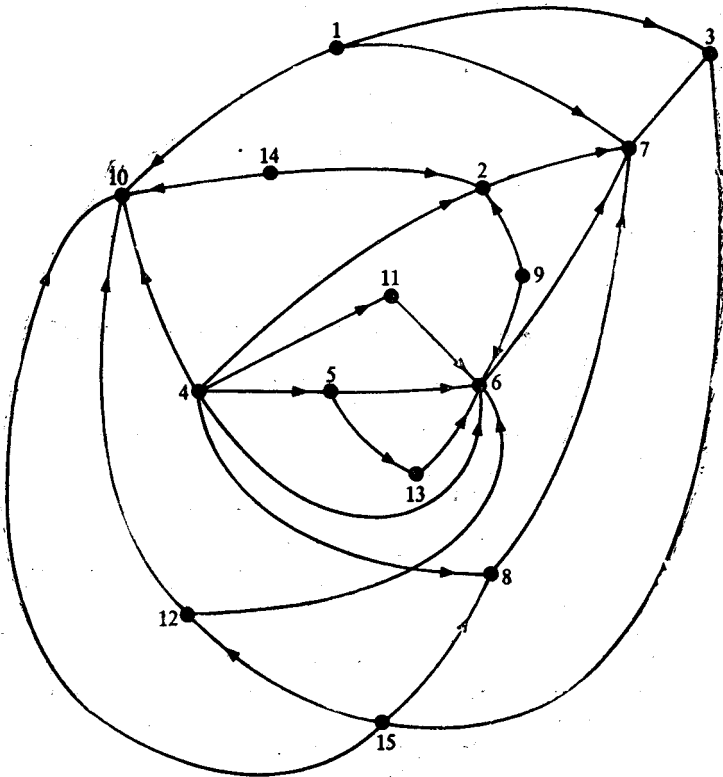


Рис. 3.3. Экосистема хищник — жертва.

1 — медведи, 2 — птицы, 3 — олени, 4 — лисы, 5 — змеи, 6 — насекомые, 7 — растения; 8 — кролики, 9 — сныть, 10 — грызуны, 11 — саламандры, 12 — скунсы, 13 — жабы, 14 — дикие кошки, 15 — волки.

представим его в виде ориентированного графа (рис. 3.3), вершины которого соответствуют видам животных и растений, а дуга от i к j означает, что вид j служит пищей для вида i .

Так как предполагается, что система замкнута, то, естественно, взять множества X и Y одинаковыми и состоящими из всех видов, входящих в сообщество. Будем использовать

символ x_i ($=y_i$) для обозначения i -го вида, анумерованного в соответствии со списком видов сообщества на рис. 3.3.

Каждый вид в сообществе может быть либо хищником, либо служит пищей для какого-нибудь вида (либо то и другое). Таким образом, для того чтобы получить полную картину взаимосвязей в биологическом сообществе, надо определить два отношения на множестве $X \times X$, а именно $\lambda_{\text{хищник}}$ и $\lambda_{\text{жертва}}$, т. е. для видов, служащих пищей. Эти отношения определяются естественным образом. Например, $(x_i, x_j) \in \lambda_{\text{хищник}}$, если x_i является хищником для вида x_j , $i, j = 1, 2, \dots, 15$. Матрица инцидентий для $\lambda_{\text{хищник}}$ легко может быть получена из рис. 3.3 и имеет вид

$\lambda_{\text{хищник}}$	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0
X_2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
X_3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
X_4	0	1	0	0	1	1	0	1	0	1	1	0	0	0	0
X_5	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
X_6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
X_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X_8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
X_9	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
X_{10}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X_{11}	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
X_{12}	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
X_{13}	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
X_{14}	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
X_{15}	0	0	1	0	0	0	0	1	0	1	0	1	0	0	0

Выполняя q -анализ, получим следующие значения связностей для хищников:

$$q=5, \quad Q_5=1, \quad \{x_4\},$$

$$q=4, \quad Q_4=1, \quad \{x_4\},$$

$$q=3, \quad Q_3=2, \quad \{x_4\}, \{x_{15}\},$$

$$q=2, \quad Q_2=3, \quad \{x_4\}, \{x_{11}\}, \{x_{15}\},$$

$$q=1, \quad Q_1=2, \quad \{x_1, x_4, x_9, x_{12}, x_{14}, x_{15}\}, \{x_5\},$$

$$q=0, \quad Q_0=1, \quad \{\text{все, исключая } x_2, x_{10}\}$$

Структурный вектор комплекса равен

$$Q = \begin{pmatrix} 5 & 0 \\ 1 & 1 & 2 & 3 & 2 & 1 \end{pmatrix}$$

Проведенное исследование показывает, что относительно хищников комплекс связан для больших и малых значений q , а для промежуточных значений q он распадается на несколько несвязных компонент.

Существование на уровне p более чем одной компоненты означает, что существуют два p -мерных вида, которые не являются p -связными. Это наводит на мысль о целесообразности введения понятия *геометрическое препятствие*. Определим вектор препятствия как $O = Q - U$, где U — вектор, все компоненты которого равны 1. Компоненты вектора O являются мерой препятствия «свободному потоку информации» в комплексе на каждом уровне размерности. В вышеприведенном примере препятствие на уровне $q = 3$ означает, что симплексы x_4 (лиса) и x_{15} (волк), хотя каждый из них и употребляет в пищу по крайней мере четыре вида животных, не связаны (прямо или косвенно) никакими четырьмя видами и, следовательно, беспрепятственный обмен пищей между лисой и волком на 3-уровне не возможен. Другими словами, вектор препятствия — это грубый индикатор возможных вариантов выбора пищи для хищников на каждом q -уровне.

Выше было отмечено, что интеграция отдельных симплексов в комплекс может быть изучена с помощью вычисления их эксцентриситетов. Следуя определению 3.2, получим следующие значения эксцентриситетов хищников:

Виды	Эксцентриситет	Виды	Эксцентриситет	Виды	Эксцентриситет
1	$1/2$	6	0	11	0
2	0	7	∞	12	0
3	0	8	0	13	0
4	2	9	0	14	0
5	1	10	∞	15	1

Таким образом, помимо видов 7 (растения) и 10 (грызуны), которые не входят в комплекс хищников, наименее однородным хищником является вид 4 (лиса). Это принципиальный вывод, так как лиса имеет достаточно пищи, кото-

рую она не делит ни с каким другим «многомерным» хищником.

Полученные результаты показывают, что эксцентриситет является мерой гибкости видов по отношению к их пище, т. е. мерой их способности к выживаемости при изменениях в биологическом сообществе. Подобная интерпретация напоминает нам понятие адаптации, но уже для отдельного вида.

Аналогичные рассуждения и анализ могут быть проведены для отношения $\lambda_{\text{жертва}}$. Чтобы сэкономить место, представим только конечные результаты q -анализа и вычислений эксцентриситета (они могут быть использованы читателями для проверки их собственных вычислений). q -анализ дает

$$\begin{aligned}
 q=5, & \quad Q_5=1, \quad \{x_6\}, \\
 q=4, & \quad Q_4=3, \quad \{x_6\}, \{x_7\}, \{x_{10}\}, \\
 q=3, & \quad Q_3=3, \quad \{x_6\}, \{x_7\}, \{x_{10}\}, \\
 q=2, & \quad Q_2=4, \quad \{x_6\}, \{x_7\}, \{x_{10}\}, \{x_2\}, \\
 q=1, & \quad Q_1=2, \quad \{x_2, x_3, x_6, x_8, x_{10}\}, \{\bar{x}_7\}, \\
 q=0, & \quad Q_0=1, \quad \{\text{все, исключая } x_1, x_4, x_9, x_{14}, x_{15}\}.
 \end{aligned}$$

следовательно, структурный вектор $Q = (1, 3, 3, 4, 2, 1)$. Для эксцентриситетов жертв получены следующие значения:

Виды	Эксцентриситет	Виды	Эксцентриситет	Виды	Эксцентриситет
1	∞	6	2	11	0
2	$1/2$	7	4	12	0
3	0	8	0	13	0
4	∞	9	∞	14	∞
5	0	10	$3/2$	15	∞

Как и в случае комплекса хищников, многомерные симплексы x_6 (насекомые), x_7 (растения) и x_{10} (грызуны) имеют большие значения эксцентриситета. Это означает, что они относительно стойки к малым изменениям в сообществе, составленном из видов, служащих пищей. Добавление или удаление хищников, по-видимому, почти не отразится на многомерных видах, служащих кому-либо пищей, — они будут оставаться жертвами для большинства видов,

Рассмотрим отношение $\lambda_{\text{хищник}}$ с гомологической точки зрения, проверив в нем наличие нетривиальных граничных циклов. Для размерности $q = 1$ имеем четыре симплекса $\sigma^5, \sigma^9, \sigma^{12}, \sigma^{14}$. Легко показать, что

$$\partial(\sigma^9 + \sigma^{12} - \sigma^{14}) = 0.$$

Это означает, что цепь $\sigma^9 + \sigma^{12} - \sigma^{14}$ является кандидатом в граничные циклы. Теперь необходимо выяснить, существует ли 2-цикл в K с границей, равной $\sigma^9 + \sigma^{12} - \sigma^{14}$. Так как единственным 2-симплексом в K является $\sigma_2 = \langle x_3 x_7 x_{10} \rangle$, а его границей является $\langle x_7 x_{10} \rangle - \langle x_3 x_{10} \rangle + \langle x_3 x_7 \rangle \neq \sigma^9 + \sigma^{12} - \sigma^{14}$, то $\sigma^9 + \sigma^{12} - \sigma^{14}$ является 1-циклом, но не является границей. Можно показать, что

$$Z = \{\sigma^9 + \sigma^{12} - \sigma^{14}\},$$

т. е. Z_1 порождается циклом $\sigma^9 + \sigma^{12} - \sigma^{14}$. Аналогично, так как $\sigma_2 = \langle x_3 x_7 x_{10} \rangle$ — единственный 2-симплекс, то $\langle x_7 x_{10} \rangle - \langle x_3 x_{10} \rangle + \langle x_3 x_7 \rangle$ — единственная граница и

$$B_1 = \{\langle x_7 x_{10} \rangle - \langle x_3 x_{10} \rangle + \langle x_3 x_7 \rangle\}.$$

Поэтому имеем

$$H_1 = Z_1/B_1, \quad H_0 \approx J,$$

а так как нет такого целого n , что $n(\sigma^9 + \sigma^{12} - \sigma^{14}) \in B_1$, то комплекс не имеет кручения и, следовательно,

$$H_1 = J.$$

Числа Бетти $\beta_1 = \beta_0 = 1$, все другие $\beta_i = 0$, $i = 2, \dots, 5$.

Вышеприведенный анализ показывает, что комплекс хищников в размерности 1 содержит дыру, ограниченную 1-симплексами.

Наглядную физическую интерпретацию этой дыры дать трудно, но ее наличие, по-видимому, означает, что на одномерном уровне существует некоторый вид циркуляции жертв между следующими хищниками: енот, скунс, дикая кошка. Аналогичный анализ для комплекса жертв показывает тривиальность гомологии на всех уровнях, так как не существует нетривиальных границ.

ИЕРАРХИЧЕСКИЕ СИСТЕМЫ И ПОКРЫТИЯ

В последние годы проблема принятия решений в иерархических системах стала темой большинства работ по математической теории систем. В результате анализа проблем коммуникации и неопределенности в больших системах было

установлено, что жесткоцентрализованное управление является недейственным и неэффективным. Следовательно, для того чтобы получить рациональное поведение системы, необходимо осуществить ее декомпозицию в локально управляемые подсистемы, поведение которых координируется органами управления на других уровнях иерархии. Чтобы убедиться в универсальности именно таких структур принятия решений, достаточно взглянуть на организационную структуру любой большой фирмы или учреждения.

Между иерархической организацией системы и способом связи ее подсистем существует очевидная зависимость. Поэтому возникает естественный вопрос, как расширить понятие *топологическая связность*, чтобы отразить в нем и иерархический аспект. Наш подход к этому вопросу предполагает использование теоретико-множественного понятия покрытия.

Определение 3.3

Семейство (конечное) множеств $A = \{A_i\}_{i=1}^n$ называется *покрытием* (конечного) множества X , если

$$A_i \subseteq 2^X$$

и

$$X = \bigcup_{i=1}^n A_i$$

(2^X — множество всех подмножеств множества X). Если, кроме того, известно, что $A_i \cap A_j = \emptyset$ ($i \neq j$), то A называется *разбиением* X .

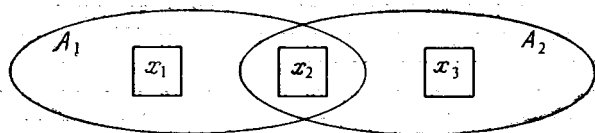


Рис. 3.4. Покрытие множества X .

Согласно вышеприведенному определению, элементы A являются подмножествами X (рис. 3.4). Следовательно, можно считать A_i как бы расположенным на $N + 1$ уровне, предполагая, что элементы X расположены на N -м уровне.

Теперь можно определить иерархию H при помощи отношения μ , задаваемого условием: $(A_i, X_j) \in \mu$ тогда и только тогда, когда $X_j \in A_i$. Такое отношение μ может быть также представлено с помощью матрицы инцидентий из нулей и единиц точно так же, как и отношения, задаваемые на

N -уровне. Эта идея может быть распространена на дополнительные уровни иерархии и связи между уровнями (рис. 3.5).

Подобный подход к изучению иерархических систем тесно связан с известной теорией типов Бертрانا Рассела, согласно которой не следует смешивать элементы множества с «множествами элементов» и «множествами множеств элементов» и т. д.

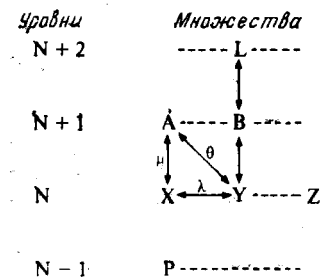


Рис. 3.5. Уровни иерархии множеств и отношений.

Введение такого принципиального логического различия сразу же избавляет нас от многих логических парадоксов, например парадокса брэдоброя.

В некотором городе парикмахер бреет только тех, кто не бреется сам. Должен ли он брить самого себя? Если X — множество мужчин в городе, то тогда в соответствии

с теорией типов парикмахера следует рассматривать не как элемент множества X , а скорее как элемент множества 2^X , состоящий из мужчин, которые бреются сами. Однако применительно к элементам 2^X нельзя решать те же вопросы, что и применительно к элементам X .

ПРИМЕНЕНИЯ ПОНЯТИЯ q -СВЯЗНОСТИ

ПРИ АНАЛИЗЕ ИГРЫ В ШАХМАТЫ

И КОМЕДИИ ШЕКСПИРА «СОН В ЛЕТНЮЮ НОЧЬ»

Рассмотрим некоторые возможности q -анализа на примерах анализа игры в шахматы и комедии Шекспира «Сон в летнюю ночь». Эти примеры отражают нетрадиционное направление использования системного анализа и свидетельствуют о том, что почти любой аспект человеческой деятельности при искусном применении математического аппарата порождает нетривиальные математические проблемы.

Шахматы

Утверждение, что игра в шахматы — это отношение (определенного типа) между игровыми фигурами и клетками доски, является почти тавтологическим. Поэтому выбор множеств X и Y , необходимых для применения q -анализа к исследованию связности рассматриваемой структуры, в этом случае совершенно очевиден, а именно:

$$X = \{\text{игровые фигуры}\}, \quad Y = \{\text{клетки доски}\}$$

Разобьем множество X на два подмножества: X_W , состоящее из белых фигур, и X_B , состоящее из черных фигур. На первый взгляд кажется многообещающим следующее отношение между X и Y : $(x_i, y_j) \in \lambda$ тогда и только тогда, когда фигура x_i занимает клетку доски y_j . Однако выводы, полученные из исследования этого отношения, не дают скольконибудь глубокого понимания структурных аспектов шахматной позиции.

Оказывается (гл. 1), что продуктивное отношение между фигурами и клетками доски задается понятием *фигура, атакующая клетку*. Поскольку множество X было разбито на белые и черные фигуры, то существует два отношения λ_W и λ_B , порождаемые правилами игры для белых и черных фигур соответственно.

Правила, приведенные в гл. 1, задают отношения λ_W и λ_B , характеризующие сложившуюся на доске позицию (клетки рассматриваются как вершины комплекса, а фигуры — как его симплексы). По-видимому, для описания позиции, полученной в некоторый момент игры, необходимы также и сопряженные отношения λ_W^* , λ_B^* .

Вышеприведенные идеи имеют важное значение для анализа позиций и автоматизации игры в шахматы. Это направление исследований очень интенсивно развивается в разных странах, и заинтересованный читатель может обратиться к литературе, приведенной в конце главы. Отметим только, что если сравнить максимальную размерность фигур, понимаемых как симплексы, порождаемые отношениями λ_W и λ_B , с классическими значениями силы фигур, то получим табл. 3.1.

Таблица 3.1

Сравнение значений шахматных фигур

Фигура	Максимальная размерность	Классическое значение
Пешка	1	1
Конь	7	3
Слон	12	3
Ладья	13	5
Королева	26	9

Из таблицы видно, что относительные разности между симплицальными значениями достаточно хорошо соответствуют относительной силе фигур, полученной эмпирически. Единственное существенное расхождение дают конь и слон.

Симплициальные значения предполагают, что слон несколько сильнее. Это объясняется тем, что все симплициальные значения вычислялись в «более выгодном для слона» предположении, т. е. когда на доске нет других препятствующих фигур, и, следовательно, мы получили максимально возможные значения. Вообще же каждому шахматисту хорошо известно, что сила шахматной фигуры зависит от конкретной позиции.

Иерархические понятия, введенные выше, соответствуют определенным шахматным понятиям, и такие известные шахматные термины, как сильное поле и слабое поле, захват центра и открытая вертикаль, могут быть интерпретированы на языке множеств, покрытий и размерности.

«Сон в летнюю ночь»

Выбор комедии Шекспира был обусловлен исключительно желанием продемонстрировать возможности применения q -анализа в такой сфере человеческой деятельности, для которой нетрадиционно использование методологии системного анализа.

Начнем с идентификации соответствующих множеств и отношений. Для этого довольно произвольно разделим пьесу на три множества:

A — пьеса, акты, сцены, мизансцены.

B — действующие лица.

C — комментарий, пьеса, сюжет, явление, реплики.

Определим элементы этих множеств и их различные иерархические уровни следующим образом:

Множество A

Уровень $N + 2$	Пьеса
$N + 1$	Акты $\{a_1, a_2, a_3, a_4, a_5\}$
N	Сцены $\{s_1, s_2, \dots, s_9\}$
$N - 1$	Мизансцены $\{sb_1, sb_2, \dots, sb_{25}\}$ (элементы и уровни получены исходя из физической композиции пьесы, т. е. появления и ухода действующих лиц)

Множество B

Все уровни N	Действующие лица $\{c_1, c_2, \dots, c_{21}\}$ (Тезей, Ипполита, Эгей и др.)
----------------	--

Множество C

Уровень $N + 3$	Комментарий (реплика, даваемая Пэком в конце пьесы. Она адресована прямо к зрителю и поэтому лежит вне непосредственного действия пьесы)
$N + 2$	Пьеса
$N + 1$	Сюжетные линии $\{P_1, P_2, P_3\}$

- P_1 : царство Тезея, занятое подготовкой к бракосочетанию Тезея и Ипполиты; репетиция и исполнение пьесы Основой и его друзьями; разрешение Тезеем затруднения Гермины и ее двух поклонников
- P_2 : царство фей и эльфов, ссора между Обероном и Титанией, трюк Оберона и его последствия для Титании и Основы
- P_3 : мир влюбленных, основной конфликт Германии, Деметрия и Лизаидра; вмешательство Оберона и Пэка и последствия такого вмешательства
- Уровень N** Явления $\{PS_1, \dots, PS_8\}$
- PS_1 : роль Тезея в жизни четырех влюбленных
- PS_2 : взаимоотношения между Тезеем и Ипполитой, церемония бракосочетания
- PS_3 : репетиция и исполнение пьесы Основой и его друзьями
- PS_4 : царство эльфов и фей в целом: их песни и волшебство
- PS_5 : взаимоотношение между Обероном и Титанией
- PS_6 : трюк Оберона, примененный к Титании, и его последствия для Основы и Титании
- PS_7 : основной конфликт четырех влюбленных
- PS_8 : вмешательство Оберона и Пэка в жизнь четырех влюбленных и последствия такого вмешательства
- Уровень $N - 1$** Реплики $\{SP_1, \dots, SP_{104}\}$ (в пьесе 104 реплики, определяемые как слова, обращенные к зрителям любым действующим лицом и развивающие сюжет, еще неизвестный зрителям)

Бинарные отношения, связывающие эти множества на различных уровнях иерархии, достаточно очевидны из определения самих множеств. Например, если Y — сюжеты, X — действующие лица, то естественно определить $\lambda \subset Y \times X$ следующим образом:

$(y_i, x_j) \in \lambda \iff$ Действующее лицо x_j участвует в сюжете y_i .

Общая иерархическая структура приведена на рис. 3.6. Составляя по этим отношениям матрицы инцидентий и проводя исследования с помощью q -анализа, можно получить некоторые интересные результаты.

Множество Y — сюжеты уровня $(N + 1)$, множество X — действующие лица уровня $(N + 1)$.

В комплексе $K_Y(X)$ все три сюжета становятся отдельными компонентами только на уровне связности $q = 8$. Это означает, что сюжетные линии могут быть различимы только зрителями, следящими за девятью действующими лицами. Аналогично, при $q = 6$ имеется всего две компоненты $\{P_1, P_2\}$, $\{P_3\}$. Следовательно, если зрители могут следовать только

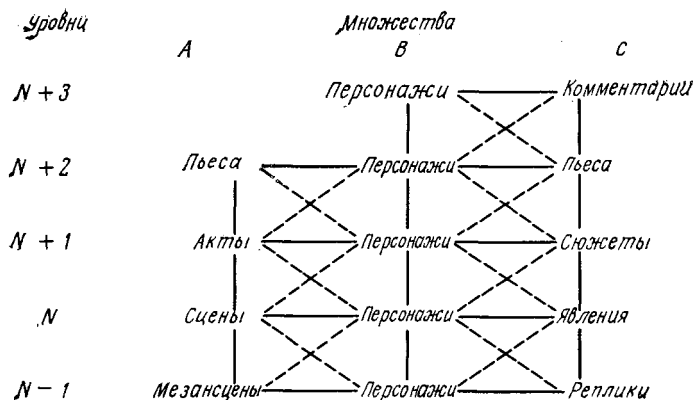


Рис. 3.6. Связность структуры комедии Шекспира «Сон в летнюю ночь».

за семью персонажами, то они видят пьесу, как бы состоящую из двух сюжетов, где P_1 и P_2 (царство Тезея и царство фей) объединены.

В сопряженном комплексе $K_X(Y)$ персонажи Гермия, Лизандр, Деметрий и Елена доминируют структуру при $q = 2$. Следовательно, зрителям, смотревшим все три сюжета отдельно, пьеса представляется как пьеса о влюбленных.

$Y =$ Сюжеты уровня $(N + 1)$, $X =$ Сцены уровня N

В комплексе $K_Y(X)$ при $q = 5$ имеется три компоненты. Следовательно, зрители, видевшие только шесть сцен, воспринимают три сюжета, не связанные друг с другом. Сюжеты P_1 и P_3 объединяются при $q = 4$, и, следовательно, зрители видят эти два сюжета как один, если следят только за пятью сценами. Все три сюжета сливаются при $q = 2$, т. е. когда зрители следят всего за тремя сценами. Этот результат отличается от результата вышеприведенного (для $q = 6$) анализа сюжетов и персонажей, где объединялись сюжеты P_1 и P_2 .

В комплексе $K_X(Y)$ хитрость Оберона, его примирение с Титанией, разрешение дилеммы влюбленных и сцена кульми-

нации церемонии бракосочетания Тезея и Ипполиты доминируют структуру при $q = 2$.

$Y =$ Явления уровня N , $X =$ Реплики уровня $N - 1$

В комплексе $K_Y(X)$ явление PS_8 доминирует структуру при $q = 35$, PS_3 — при $q = 26$ и PS_6 — при $q = 10$. Следовательно, PS_8 будет, вероятно, полностью понята зрителями, если они прослушали 36 реплик, хотя для понимания PS_3 необходимо 27 реплик, а для PS_6 — только 11 реплик. Таким образом, q -анализ дает некоторое указание о сложности явления. Критическим значением является $q = 5$, откуда следует, что все явления требуют для своего полного понимания зрителями минимум шесть реплик.

На критическом уровне $q = 0$ комплекс $K_X(Y)$ имеет тривиальную структуру. Это значит, что каждая реплика имеет отношение лишь к одному явлению, что полностью соответствует вышеприведенному ее определению.

АЛГЕБРАИЧЕСКАЯ СВЯЗНОСТЬ

Хотя описание связности системы с помощью структурного вектора, получаемого при q -анализе, представляется количественным, оно по существу носит качественный характер. Локальные системные детали при получении глобального описания «размываются», не давая какой-либо конкретной информации, относящейся к природе или структуре подсистем (симплексов). Подобный результат — естественное следствие выбранного способа описания системы, а именно языка множеств и бинарных отношений. В тех же ситуациях, когда о локальной структуре системы известно гораздо больше, вопрос связности часто может быть плодотворно изучен при использовании средств алгебры, а не топологии.

Основным определяющим требованием исследования алгебраической связности, так же как и топологической, будет условие конечности. В случае топологической связности таким подходящим условием была конечность множеств X и Y , в случае же алгебраической связности условия должны относиться к пространству состояний системы, так как мы будем иметь дело только с системами, задаваемыми внутренним описанием при помощи дифференциальных, или разностных уравнений. Так, подходящим условием конечности для систем, динамика которых линейна, является условие конечности пространства состояний.

Принципиальное преимущество, которое дает исследование алгебраической связности, — это получение систематической процедуры для декомпозиции системы в математически

неприводимые подсистемы и объединение этих подсистем в единый процесс. Необходимо отметить, что эти математические подсистемы, или «элементарные строительные блоки», в общем случае могут не иметь отношения к какой-либо естественной декомпозиции системы, получаемой исходя из ее физической структуры. Таково, к сожалению, часто неизбежное свойство процесса перевода содержательной проблемы в вычислительно-эффективную математическую модель. Как только разработана удовлетворительная модель исследуемой системы, эта модель начинает жить «своей собственной жизнью». При этом важно решить вопросы выбора координат, удобных для интерпретации результатов моделирования, и, возможно, других координат, удобных для математических вычислений на машине. В этом случае мы должны уметь переводить по желанию одно представление задачи в другое.

ЛИНЕЙНЫЕ СИСТЕМЫ

В качестве простого примера линейной системы рассмотрим динамическую систему, внутреннее описание которой может быть задано с помощью дифференциального уравнения

$$\frac{dx}{dt} = Fx(t) + Gu(t),$$

и ее выход задан равенством

$$y(t) = Hx(t),$$

где x , u и y представляют собой n -, m -, p -мерные векторные функции соответственно, а F , G , H — матрицы соответствующих порядков.

Вероятно наиболее удобный способ проиллюстрировать алгебраические аспекты линейной системы — это рассмотреть матрицу передаточной функции системы. Обозначим через \bar{x} , \bar{u} , \bar{y} преобразования Лапласа вектор-функций x , u , y соответственно. Можно показать, что отношения между преобразованными входом и выходом системы задается формулой

$$\bar{y}(z) = H(zI - F)^{-1} G\bar{u}(z) = W(z)\bar{u}(z),$$

где z — преобразование переменной. Матрица W , называемая *передаточной матрицей*, задает внешнее описание динамики системы. Предположим далее, что элементы матрицы $W(z)$ — правильные рациональные дроби от z . Можно показать, что в этом случае изучение внутренней схемы взаимосвязи входных и выходных каналов системы Σ эквивалентно изучению циклической структуры матрицы $W(z)$.

Теорема реализации

Каждая матрица передаточной функции $W(z)$ правильных рациональных дробей может быть реализована в виде прямой суммы систем

$$\Sigma_i = (F_i, G_i, H_i),$$

где G_i, H_i вычисляются по формулам, приведенным ниже, а F_i — циклическая матрица с характеристическим полиномом ψ_i , i -м инвариантом матрицы $W(z)$.

Из теоремы следует, что основные строительные блоки системы Σ получаются при вычислении инвариантов матрицы $W(z)$. Если ψ — наименьший общий знаменатель элементов матрицы $W(z)$, то ψW является полиномиальной матрицей. Применяя алгоритм вычисления инвариантов, получим представление

$$\psi W = PLQ \text{ mod } \psi,$$

где $\det P$ и $\det Q$ являются единицами кольца $K[z]/K[z]\psi$, а матрица L — диагональной матрицей, единственной с точностью до единиц того же кольца (здесь K — произвольное числовое поле). При этом оказывается, что элементы ψ связаны с элементами матрицы L следующим образом:

$$\psi_i = l_{n-i+1}, \quad i = 1, 2, \dots, q,$$

где n — степень наименьшего общего знаменателя.

Для получения элементов системы Σ_i необходимо проделать следующие операции:

1. Для каждого инварианта ψ_i матрицы W найти циклическую матрицу F_i , такую, что ее характеристический полином $\chi_{F_i} = \psi_i$, $i = 1, 2, \dots, q$. (В качестве F_i можно взять сопровождающую матрицу ψ_i).

2. Пусть $L = (l_1, l_2, \dots, l_n)$, p_i — i -й столбец P , q'_i — i -я строка Q в вышеприведенном представлении ψW . Возьмем такие полиномиальные векторы v_i и w_i , что

$$\chi_{F_i} = (zI - F_i)^{-1} = v_i(z) w'_i(z) \text{ mod } \chi_{F_i}.$$

(Эти векторы определяются однозначно с точностью до единиц $K[z]/K[z]\chi_{F_i}$.)

3. Уравнения

$$H_i v_i = (l_i / \mu_i) p_i \text{ mod } \psi_i,$$

$$w'_i G_i = q'_i \text{ mod } \psi_i$$

имеют единственное решение H_i, G_i . Здесь μ_i — наибольший общий делитель l_i и ψ . Находим эти решения и тем самым систему Σ_i .

Выводы

Размеры (размерность) блоков (подсистем), входящих во внешнее описание системы Σ , равны в точности степеням инвариантов матрицы W .

Каноническая структура линейной системы (рис. 3.7) характеризуется высокой степенью связности между ее компонентами и заметно отличается от структуры линейной системы, схема которой приводится обычно в элементарных учебниках (рис. 3.8). Простота схемы, приведенной на рис. 3.8, объясняется произвольностью ее выбора, и, несмотря на графическую привлекательность, эта схема не имеет никакого отношения к реальной системе Σ с передаточной функцией W .

Рассмотрим числовой пример.

Пусть

$$W(z) = \begin{bmatrix} \frac{1}{z+1} & 0 & 0 \\ 0 & \frac{1}{z+2} & 0 \\ 0 & 0 & \frac{1}{z+3} \end{bmatrix}.$$

Имеем $\psi(z) = (z+1)(z+2)(z+3)$, и инвариантами ψW является

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \psi \end{bmatrix}.$$

Поэтому

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$P = \begin{bmatrix} (z+2)(z+3) & 1 & 0 \\ 2(z+1)(z+3) & 2 & \\ (z+1)(z+2) & 1 & 2 \end{bmatrix},$$

$$Q = \begin{bmatrix} 1/2(z+2)(z+3) & -1/2(z+1)(z+3) & 1/2(z+1)(z+2) \\ -1 & 1 & -1 \\ 0 & -1/2 & 1 \end{bmatrix}.$$

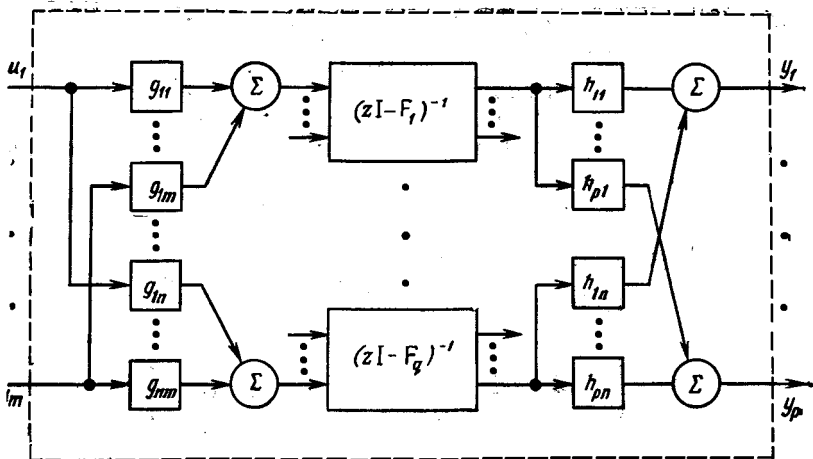


Рис. 3.7. Общая структура линейной системы.

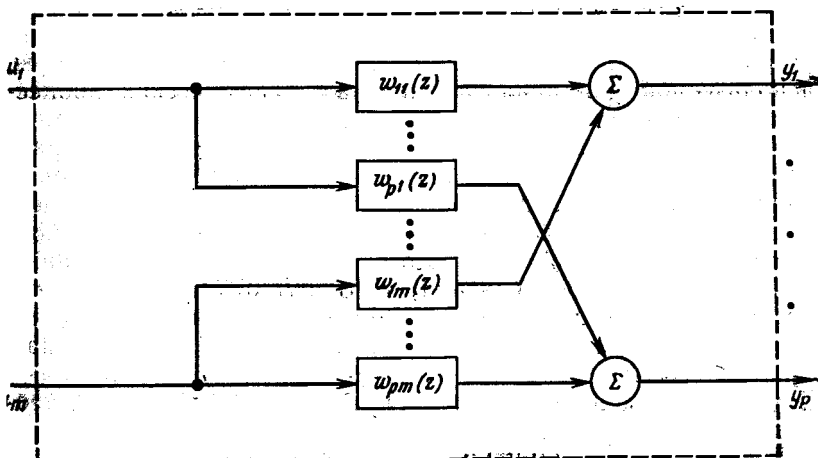


Рис. 3.8. Обычная схема линейной системы.

В качестве $v(z)$, $w(z)$ возьмем

$$v(z) = \begin{bmatrix} 1 \\ z \\ z^2 \end{bmatrix}, \quad w(z) = \begin{bmatrix} z^2 + 6z + 11 \\ z + 6 \\ 1 \end{bmatrix}.$$

Можно показать, что передаточная матрица $W(z)$ канонической внешней модели задается следующими матрицами:

$$F = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix},$$

$$G = \begin{bmatrix} 1/2 & -1/2 & 1/2 \\ -1/2 & 1 & 3/2 \\ 1/2 & -5/2 & 9/2 \end{bmatrix}, \quad H = \begin{bmatrix} 6 & 5 & 1 \\ 6 & 8 & 2 \\ 2 & 3 & 1 \end{bmatrix}.$$

Все канонические модели эквивалентны с точностью до преобразования координат $x \rightarrow Tx$ в пространстве состояний. Этот вывод легко получается из теоремы реализации.

НЕЛИНЕЙНЫЕ СИСТЕМЫ

Исследование общих нелинейных систем требует использования техники, совершенно отличной от линейной алгебры, используемой для анализа линейных динамических процессов. Это и не удивительно, так как пространство состояний нелинейной системы может и не быть векторным пространством. Для того чтобы рассмотреть достаточно широкий класс нелинейных задач, обратимся к метапринципу конечности. В этом случае разумно потребовать *конечность* пространства состояний. При этом аналогами конечномерных линейных операторов (матриц) F , G и H будут *конечные полугруппы преобразований*, возникающие из описания динамических систем с помощью автоматов. Это позволит нам сформулировать теорему декомпозиции (теорему Крона — Роудза), играющую в данном случае ту же роль, что и теорема об инвариантах в линейном случае.

ПОЛУГРУППЫ И УЗЛОВЫЕ ПРОИЗВЕДЕНИЯ

Предположим, что пространство состояний изучаемого процесса, обозначаемое Q , содержит лишь конечное число элементов. Предположим также, что \mathcal{F} — это множество преобразований пространства состояний, т. е.

$$\mathcal{F}: Q \rightarrow Q.$$

Пусть произведение двух элементов f_1 и f_2 из \mathcal{F} определяется так

$$(f_1 f_2): Q \rightarrow Q, \\ (f_1 f_2) q = f_2 [f_1(q)].$$

Это определение соответствует обычной композиции двух преобразований, которая, как известно, ассоциативна, т. е.

$$[f_1 (f_2 f_3)] q = [(f_1 f_2) f_3] q \quad \text{для всех } f_1, f_2, f_3 \in \mathcal{F}, q \in Q.$$

Можно образовать полугруппу ¹⁾ \mathcal{F}^* из групп \mathcal{F} , включив в последнюю все произведения ее элементов.

Пара объектов (Q, \mathcal{F}^*) образует так называемую *полугруппу преобразований*. Поскольку рассматривается конечное пространство Q , то и полугруппа преобразований также конечна ²⁾.

Наша цель — получить связь между произвольными преобразованиями конечного пространства состояний и определенными удобными представлениями действий этих преобразований. Для достижения этой цели введем понятие *узловое произведение координат* ³⁾. Предположим, что Q представлено при помощи декартова произведения

$$Q = X_1 \times X_2 \times \dots \times X_n,$$

т. е. каждое $q \in Q$ имеет представление $q = (x_1, x_2, \dots, x_n)$, $x_i \in X_i$, $i = 1, 2, \dots, n$. Тогда действие \mathcal{F}^* на Q *триангулировано*, если каждое $f \in \mathcal{F}^*$ может быть представлено как $f = (f_1, f_2, \dots, f_n) \in F_1 \times F_2 \times \dots \times F_n$, где $f(q) = f(x_1, x_2, \dots, x_n) = [f_1(x_1), f_2(x_1, x_2), \dots, f_n(x_1, x_2, \dots, x_n)] = q'$, т. е. $f_k: X_1 \times X_2 \times \dots \times X_k \rightarrow X_k$.

Таким образом, *триангуляция* означает, что каждая координата q' зависит только от координаты q с тем же номером и координат, ей предшествующих.

Еще одним важным понятием является понятие *k-го координатного действия*. А именно, k -е координатное действие $(Q, \mathcal{F}^*) = (X_1 \times X_2 \times \dots \times X_n; F_1 \times F_2 \times \dots \times F_n)$ — это полугруппа преобразований (X_k, G_k) , где G_k — множество всех преобразований g на X_k , таких, что

$$g: X_k \rightarrow X_k, \quad g(z) = f_k(a_1, a_2, \dots, a_{k-1}, z),$$

а $(a_1, a_2, \dots, a_{k-2}, a_{k-1})$ — произвольный фиксированный элемент $X_1 \times X_2 \times \dots \times X_{k-1}$ и $f_k \in F_k$. Другими словами, эле-

¹⁾ Полугруппой называют множество с заданной на нем ассоциативной бинарной операцией.

²⁾ Порядок (Q, \mathcal{F}^*) равен числу различных преобразований в \mathcal{F}^* .

³⁾ В алгебраической теории полугрупп общепринятым является термин *сплетение*. — Прим. переа.

менты G_k^* не только отображают X_k в себя, но и также должны быть частью триангуляции \mathcal{F}^* .

Если действие \mathcal{F}^* на Q может быть триангулировано введением координат $Q = X_1 \times X_2 \times \dots \times X_n$, то пространство состояний Q есть узловое произведение своих координатных действий, т. е.

$$(Q, \mathcal{F}^*) = (X_1, G_1^*) \omega (X_2, G_2^*) \omega \dots \omega (X_n, G_n^*).$$

Следовательно, множество $X_1 \times X_2 \times \dots \times X_n$ представляет узловое произведение координат для Q , а множество $G_1 \times G_2 \times \dots \times G_n$ — узловое произведение координат для \mathcal{F}^* .

Пример

Пусть $Q = R^2$, $X_1 = X_2 = R^1$, а $f: R^2 \rightarrow R^2$ задается верхнетреугольной матрицей

$$f = \begin{bmatrix} f_{11} & f_{12} \\ 0 & f_{22} \end{bmatrix}.$$

В качестве компонентного преобразования f_i возьмем композицию оператора, определяемого ведущим $i \times i$ блоком f , и оператора проектирования на i -ю координату.

ТЕОРЕМА ДЕКОМПОЗИЦИИ КРОНА — РОУДЗА

В теории конечных групп существует структурная теорема — так называемая теорема Жордана — Гёльдера, согласно которой любая конечная группа может быть построена из фиксированного множества простых групп (факторов Жордана — Гёльдера) и это множество групп определено однозначно (с точностью до изоморфизма). В теории конечных полугрупп аналогом такой теоремы является теорема Крона — Роудза: *любое конечное пространство состояний может быть представлено так, что множество явлений, наблюдаемых на нем, триангулировано. Кроме того, координатные действия должны быть либо (а) простыми группами перестановок, тесно связанными с (Q, \mathcal{F}^*) , либо (б) одной из трех возможных полугрупп преобразований, наибольшая из которых имеет порядок три.*

Следствием этой теоремы является тот факт, что любая система с конечным пространством состояний может быть удобно координатизирована так, что координатные действия распадаются на отдельные простые виды: группы перестановок, отмеченные в (а), должны быть таковы, что они делают первоначальную полугруппу (Q, \mathcal{F}^*) , в то время как полугруппы (б) должны быть полугруппами элементарных триг-

геров. Таким образом, независимо от сложности поведения системы, можно анализировать систему, изучая лишь сравнительно простые объекты, которые соединяются с помощью конструкции узлового произведения. Для того чтобы дать точную формулировку теоремы Крона — Роудза, необходимо ввести понятие делимости для полугрупп.

Определение 3.4

Будем говорить, что полугруппа преобразований (X, S) делит $(Y, T): (X, S) | (Y, T)$ тогда и только тогда, когда:

— существуют подмножество Y' множества Y и подполугруппа T' полугруппы T , такие, что Y' инвариантно относительно действия T' .

— существуют отображение $\theta: Y' \rightarrow X$ (\rightarrow обозначает отображение «на») и эпиморфизм $\phi: T' \rightarrow S$, такие, что $\theta(yt) = \theta(y)\phi(t)$ для всех $y \in Y'$ и $t \in T'$.

Реализация данной системы (автомата) при помощи подсистем (автоматов) осуществляется в соответствии с делением полугрупп. Теперь можем сформулировать основную задачу, имеющую непосредственное отношение к теореме Крона — Роудза.

Задача

Если X — пространство состояний, на котором действует конечная полугруппа преобразований \mathcal{F}^* , то можно ли найти декомпозицию (X, \mathcal{F}^*) в подполугруппы преобразований (X_k, G_k) , $k = 1, \dots, n$, такие, что будет получено минимальное решение $(X, \mathcal{F}^*) | (X_n, G_n) \omega \dots \omega (X_1, G_1)$, и если можно, то каковы структура компонент (X_i, G_i) и максимальное значение n ?

На языке теории автоматов соответствующая задача является факторизацией конечного автомата в наибольшее возможное число автоматов, и ее решение дается так называемой простой декомпозицией первоначального автомата.

Прежде чем перейти к формулировке теоремы Крона — Роудза, введем такие понятия, как *полугруппа триггера* и *примарная группа*.

Определение 3.5

Пусть $a \neq b$. Рассмотрим полугруппу $(\{a, b\}, \{C_a, C_b, Id\})$, где $xC_a = a$, $xC_b = b$, $xId = x$ для $x = a$ или $x = b$, которую будем записывать также в виде $(\{a, b\}, U_3)$. Полугруппа $(\{a, b\}, U_3)$ называется *полугруппой триггера порядка три*.

Определение 3.6

Конечная группа G *примарна*, если она простая¹⁾, $G \neq \{Id\}$. Если S — полугруппа, тогда $\text{Primes}(S) = \{G : G \text{ примарна и } G|S\}$.

Теорема примарной декомпозиции для конечных полугрупп, или теорема Крона — Роудза

Пусть (X, \mathcal{F}^*) заданная конечная полугруппа. Тогда существует декомпозиция её в узловое произведение $(X_1, G_1^*), \dots, \dots, (X_n, G_n^*)$, такая, что

$$(X, \mathcal{F}^*) | (X_n, G_n^*) \omega \dots \omega (X_1, G_1^*),$$

и для каждого фактора (X_j, G_j^*) либо $G_j^* \in \text{Primes}(\mathcal{F}^*)$, (X_j, G_j^*) — транзитивная группа перестановок, либо

$$(X_j, G_j^*) = (\{a, b\}, U_3).$$

Сравнивая эту теорему с теоремой Жордана — Гёльдера, можно заметить, что теория полугрупп эквивалентна теории конечных групп, дополненной «триггерной» операцией.

ДЕКОМПОЗИЦИЯ АНАЛИТИЧЕСКИХ СИСТЕМ

Если пространство состояний системы бесконечно, но конечномерно, то можно доказать справедливость теоремы, почти аналогичной теореме Крона — Роудза. Однако в этом случае потребуются более сложный математический аппарат.

Пусть система определена внутренним описанием, задаваемым дифференциальным уравнением

$$\dot{x} = f(x, u),$$

где пространство состояний M является аналитическим многообразием, $u \in \Omega \subset R^m$, а f — действительная аналитическая функция переменной x , непрерывная по u и удовлетворяющая условию Липшица по x равномерно относительно u . Обозначим через $V(M)$ алгебру Ли всех аналитических векторных полей на M . На алгебре $V(M)$ задана скобочная операция Ли $[\cdot, \cdot]$, позволяющая из двух заданных векторных полей v, w получить новое векторное поле с помощью формулы $[v, w] = (\partial v / \partial x)(w) - (\partial w / \partial x)(v)$. Для каждого $u \in \Omega$ рассмотрим векторное поле $f(\cdot, u)$ и определим алгебру Ли системы Σ как наименьшую подалгебру алгебры $V(M)$, содержащую все такие векторные поля. Сформулируем теперь конечномерный аналог теоремы Крона — Роудза.

¹⁾ Напомним, что группа проста, если у нее нет нетривиальных нормальных подгрупп. N — нормальная подгруппа G тогда и только тогда, когда $Ng = gN$ для всех $g \in G$.

Теорема Кренера

Если алгебра Ли системы Σ конечномерна, то система Σ допускает декомпозицию в параллельные каскады систем с простой алгеброй Ли с последующим каскадом одномерных систем.

Сделаем несколько замечаний относительно этой теоремы.

1. Алгебра Ли является простой, если она не абелева и не имеет нетривиальных идеалов. Таким образом, простые алгебры Ли в теореме Кренера являются аналогами простых групп из теоремы Крона — Роудза. Однако такой аналогии не существует между одномерными системами и полугруппами триггеров, так как триггеры входят в негрупповую часть автоматов. Заметим, что с точностью до изоморфизма существуют только две одномерные системы — на окружности и на прямой линии. Таким образом «простых» элементов, получаемых при декомпозиции алгебры Ли, также два.

2. В определенном смысле эта теорема дает наилучшую декомпозицию конечномерных систем. Действительно, конечномерная система неразложима тогда и только тогда, когда алгебра Ли одномерна или проста. Заметим, что эта теорема дает необходимые и достаточные условия для декомпозиции любой конечномерной аналитической системы.

В качестве примера, поясняющего эти результаты, рассмотрим билинейную систему Σ , определяемую матричными уравнениями:

$$\dot{X} = \sum_{i=1}^3 u_i B_i x, \quad X(0) = I,$$

где

$$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Пространство состояний такой системы — это группа действительных 2×2 матриц с определителем, равным 1, $M = SL(2, R)$, а входы задаются элементами $\Omega = R^3$. Проведя алгебраические выкладки, можно найти, что система Σ обладает нетривиальной параллельно-каскадной декомпозицией

$$\dot{X}_1 = u_1 B_1,$$

$$\dot{X}_2 = (X_1)^{-1} (u_2 B_2 + u_3 B_3) X_1 X_2,$$

где

$$X = X_1 X_2.$$

Этот пример показывает также, что пространство состояний нелинейной системы, как правило, не R^n .

ЛИТЕРАТУРА

Интересное использование внидиагональных элементов матрицы взаимосвязей для получения декомпозиции системы обсуждается в статье Steward D. V., Partitioning and Tearing System of Equations, *SIAM J. Num. Anal.*, 2, 345—365 (1965).

Автор предлагает версию метода «разрыва», примененного Кроном, для изучения электрических цепей. Резюме этой работы можно также найти в книге

Harr N., *The Theory of Network Diakoptics*, Academic, New York, 1970.

Подход Крона, рассматриваемый с топологической точки зрения, излагается в статье

Roth J. P., An Application of Algebraic Topology: Kron's Method of Tearing, *Quart. Appl. Math.*, 17, 1—14 (1959).

Комплекс и связи. Экцентриситет

Весь материал по q -связности и эксцентриситету взят из основополагающей работы Эткина. Более подробно эти вопросы рассматриваются в работах

Atkin R. H., *Mathematical Structure in Human Affairs*, Heinemann, London, 1974; *Combinatorial Connectivities in Social Systems*, Birkhauser, Basel, 1976; *Multidimensional Man*, Penguin, London; Atkin R. H., Casti J., *Polyhedral Dynamics and the Geometry of Systems*, RR-77-6, International Institute for Applied Systems Analysis, Laxenburg, Austria, March 1977.

Дыры и препятствия. Числа Бетти и кручение. p -дыры.**Коцепи и кограницы**

Материал по теории гомологий взят из приложения к книге

Atkin R., *Mathematical Structure in Human Affairs*, Heinemann, London, 1973.

Элементарные концепции алгебраической топологии излагаются в книгах

Александров П. С. *Комбинаторная топология*. — М. — Л.: 1947;

Giblin P., *Graphs, Surfaces and Homology*, Chapman and Hall, London, 1977;

Franz W., *Algebraic Topology*, Ungar, New York, 1968.

Более подробно и на современном уровне материал представлен в книгах Hilton P., Wylie S., *Homology Theory*, Cambridge University Press, Cambridge, 1960. [Имеется перевод: Хилтон П. Дж., Уайли С. *Теория гомологий*. — М.: Мир, 1971.]

Spanier E., *Algebraic Topology*, McGraw-Hill, New York, 1966. [Имеется перевод: Спеньер Е. *Алгебраическая топология*. — М.: Мир, 1971.]

Отношение хищник — жертва: пример гомологии

Этот пример взят из статьи

Casti J., *Connectivity, Complexity and Resilience in Complex Ecosystems*, in IFAC Symposium on Bio- and Ecosystems, Leipzig, Germany, 1977.

Приложение теории q -связности к шахматам и комедии Шекспира

Пример приложения теории q -связности к комедии Шекспира взят из книги

Atkin R., *Multidimensional Man*, Penguin, London.

Алгебраическая связность

Общая применимость «принципа конечности» к математической теории систем проиллюстрирована в работах
Eilenberg S., *Automata, Languages and Machines*, Vols. A and B, Academic, New York, 1974, 1976;
Kalman R., Falb P., Arbib M., *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969. [Имеется перевод: Калман Р., Фалб П., Арbib М. Очерки по математической теории систем. — М.: Мир, 1971.]

Линейные системы

Результаты по реализации линейных систем взяты из приведенной выше книги Калмана и др (гл. 10). См. также работу
Kalman R., *Irreducible Realizations and the Degree of a Rational Matrix*, *SIAM J. Appl. Math.*, 13, 520—544 (1965).

Полугруппы и узловые произведения. Теорема Крона — Роудза

Основные понятия и теоремы теории автоматов изложены в книгах
Ginsburg S., *Algebraic Theory of Automata*, Academic, New York, 1968;
Arbib M., *Machines, Languages and Semigroups*, Academic, New York, 1967;
Kalman R., Falb P., Arbib M., *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969. [Имеется перевод: Калман Р., Фалб П., Арbib М. Очерки по математической теории систем. — М.: Мир, 1971.]

Доказательство теоремы декомпозиции приведено в статье

Krohn K., Rhodes J., *Algebraic Theory of Machines*, *Trans. Am. Math. Soc.*, 116, 450—464 (1965).

Более полное современное изложение с большим числом примеров из биологии, психологии и физики дано в книге

Rhodes J., *Applications of Automata Theory and Algebra via the Mathematical Theory of Complexity*, Lecture Notes, Math. Dept., University of California, Berkeley, 1971.

Декомпозиция аналитических систем

Krener A., *A Decomposition Theory for Differentiable Systems*, *SIAM J. Control Optim.*, 15, 813—829 (1977).

«Ярче, звездочка, свети!
 Как найти тебя в пути?»
 «Я в двух секундах по дуге
 От места, где кажусь тебе.
 А хочешь, подсчитай точней:
 $ds^2 = - [^{1/2} (M/Y)] \cdot dr^2 -$
 $- r^2 d\theta - [^{1/2} (N/\gamma)] dt$ ».
 «Я понял», — молвил Артур ей¹⁾.

*Ральф Бартон. Сверкай, сверкай,
 маленькая звезда*

Главное — это из множества проблем выбрать наиболее простые, решение которых позволит выработать допускающие обобщения концепции.

Д. Гильберт

Прилагательное «сложный», несомненно, является наиболее часто употребляемым в системном анализе. Хотя почти в любой книге по прикладному системному анализу можно найти следующие словосочетания: *сложная система, степень сложности, сложная проблема* и т. д., тем не менее авторы, как правило, не уточняют смысла таких терминов. Если, однако, проанализировать контекст, в котором встречаются подобные сочетания слов, то становится очевидным, что понятие *сложность* включает такие факторы, как противоинтуитивное поведение системы, невозможность предсказания ее поведения без специального анализа и вычислений и т. д. Подобно понятию *времени*, каждому из нас кажется, что он понимает, что такое сложность, но это длится до тех пор, пока не возникает необходимость дать строгое определение сложности. Короче говоря, мы не можем точно определить, что же такое «сложная система», несмотря на то что способны ее распознать, если имеем дело с такой системой.

Аналогичная ситуация возникает при попытке математического описания сложности, так как обычные руководства оказываются практически непригодными. В этой связи, по-

¹⁾ Более точное значение равно 1,75 и может быть получено путем вычисления $ds^2 = g_{ij} x^i x^j$, где g_{ij} находится из $R_{ij} - \frac{1}{2} g_{ij} R = \frac{8\pi j}{c^4} T_{ij}$, где T_{ij} — тензор энергии — импульса. — *Прим. ред.*

видимому, целесообразно указать некоторые основные компоненты, которые должны присутствовать в любой математической теории сложности; дать несколько возможных мер сложности и показать на примерах, как идея сложности может быть использована при изучении конкретных проблем прикладного системного анализа.

Теория сложности систем является, по словам фон Неймана, «предпосылкой к пониманию процессов обучения и развития». Следовательно, системные инженеры должны приложить все усилия для перевода понятия сложности из области фольклера системного анализа в область развивающейся теории. В данной главе предлагается математический аппарат для такого перевода. Вначале рассматриваются составные части, которые должны входить в любую разумную математическую теорию сложности, и такие представления, как иерархическая структура, широкий диапазон шкал времени, уровни взаимодействия. Затем демонстрируется, как, используя алгебраические результаты, о которых шла речь в предыдущей главе, можно построить теорию сложности для конечных автоматов и их подгрупп преобразований. Далее предлагаются другие подходы к анализу структуры систем, описываемых с помощью множеств и отношений или задаваемых потенциальной (энтропийной) функцией. Естественно, что в пределах одной главы невозможно изложить всеохватывающую теорию сложности, однако, используя представленные здесь результаты при анализе конкретных систем и конкретных вопросов, можно извлечь много полезного и нетривиального.

Сложность — понятие многогранное, поэтому в различных проблемах проявляются разные аспекты сложности.

СТРУКТУРНАЯ СЛОЖНОСТЬ

Вероятно, первая мысль, которая приходит в голову при рассмотрении вопроса о сущности понятия *сложность*, — это считать систему сложной, если ее компоненты (подсистемы) связаны между собой запутанным, трудным для непосредственного восприятия образом. Такая ситуация представляет собой типичный пример структурной сложности. В этом случае мы имеем дело только со структурой коммуникационных каналов и схемой взаимодействия компонент подсистемы, пренебрегая при этом динамическими или вычислительными аспектами. Однако даже и здесь необходимо принять во внимание еще несколько других аспектов связности структуры,

в том числе *иерархическую структуру, схему связности, многообразие компонент, силу взаимодействия*. Рассмотрим эти вопросы более подробно.

Иерархия

Некоторые специалисты считают, что единственным определяющим фактором при решении вопроса о сложности системы является ее иерархическая организация. Вероятно, подобное утверждение связано с тем, что высокая степень сложности системы требует и высокой скорости обмена информацией между различными лицами, принимающими решения, и наоборот, необходимость иерархической структуры вытекает из требований, предъявляемых к обработке данных и контролю за выполнением решений в таких системах. Если считать эти утверждения верными, то число уровней иерархии в системе может служить приблизительной мерой ее сложности.

В качестве примера, иллюстрирующего этот «принцип иерархии», рассмотрим классическую *задачу о часовщиках*.

Два часовых дел мастера Хронос и Темпус собирают часы одинаковой конструкции, включающей 1000 деталей. Однако каждый из мастеров предпочитает свой собственный метод сборки. Темпус собирает часы последовательно, при этом, если он не собрал часы полностью и делает перерыв, то собранная конструкция распадается на части и он должен начинать все сначала. Хронос делит всю конструкцию на 10 частей и каждую такую часть еще на 10 частей. Таким образом, 10 маленьких частей при сборке дают одну большую часть, а 10 больших частей после сборки образуют часы. Поэтому, если Хронос вынужден прервать сборку часов, то он теряет только ту часть, над которой работает в данное время.

Предположим, что вероятность прерывания работы любого из них равна p . Можно показать, что вероятность того, что Темпус успешно окончит сборку часов, равна $(1 - p)^{1000}$. Для выполнения этой работы Хронос должен собрать все 110 частей, и вероятность того, что он прервется при сборке любой из них, равна $(1 - p)^{10}$. Прямые вычисления при $p = 0,01$ показывают, что в среднем Темпус должен потратить на сборку одних часов в 20 000 раз больше времени, чем Хронос.

Этот пример иллюстрирует следующее основное свойство иерархической системы: несмотря на наличие ошибок в локальных пунктах принятия решений, иерархическая система в целом может функционировать нормально. Поскольку для

любой большой системы характерны временные задержки, шум, неопределенность, неудивительно, что иерархическая структура возникла в результате необходимости управления большой системой.

Схема связности

Важным аспектом сложности является способ, которым подсистемы объединяются в единое целое. Как было показано в гл. 3, структура связности системы определяет потоки передачи информации в структуре и ограничивает воздействия, которые может оказать одна часть системы на другую. Это и есть те системные свойства, которые должны входить в любое интуитивное понятие сложности.

Используя топологические идеи, изложенные в гл. 3, рассмотрим меру сложности, которая учитывает различные q -цепи системы и их взаимодействия. Такая мера будет отражать геометрический, или размерностный, аспект сложности. Вместе с тем можно также исследовать связность и сложность с алгебраической точки зрения, взяв за исходное внешнее описание системы. Например, если имеется система, заданная с помощью линейного дифференциального уравнения

$$\dot{x} = Fx, \quad x(0) = c, \quad (4.1)$$

где F является квадратной матрицей размера $n \times n$, то заполненность матрицы F (ее структура связности) в определенной мере отражает сложность процесса. Между прочим, данный пример показывает, что большая размерность и высокая сложность системы могут быть слабо коррелированы. Размерность n системы может быть очень большой, однако если F имеет простую структуру (например, диагональная или разреженная), то уравнение (4.1) представляет систему малой сложности в том смысле, что ее поведение легко предсказать и понять. Сложность системы, описываемой уравнением (4.1), может быть охарактеризована тщательным исследованием схемы взаимодействия подсистем (схемы связности), а не ее размерностью.

Многообразие

Полуфилософский «принцип необходимости многообразия», согласно которому многообразие выходных сигналов системы может быть достигнуто только с помощью достаточного многообразия входных воздействий, также, по-видимому, имеет непосредственное отношение к сложности системы,

Будем отождествлять сложность системы с ее способностью реализовывать многие различные типы поведения оператора входа — выхода. Можно было бы назвать такую способность *сложностью управления*, так как этот аспект сложности отражает меру способности системы преобразовывать многообразие входных сигналов в многообразие выходных. Чтобы проиллюстрировать этот подход, рассмотрим проблему управления системой Σ , которая подвергается воздействию внешних помех.

Предположим, что имеются три вида помех 1, 2, 3, а лицо, принимающее решение, может осуществлять три вида управления α , β , γ . В зависимости от получаемой комбинации помехи — управление поведение системы разбивается на три категории a , b , c . Все эти возможности отражены в следующей матрице:

		Управление		
		α	β	γ
<i>Тип помехи</i>	1	b	a	c
	2	a	c	b
	3	c	b	a

В этом случае как множество управлений $\{\alpha, \beta, \gamma\}$, так и множество помех $\{1, 2, 3\}$ состоят из трех различных элементов и, как следует из таблицы, лицо, принимающее решение, может направить систему в любое желаемое состояние выхода независимо от внешних помех. Некоторый общий кибернетический принцип гласит, что

$$\text{Общее многообразие в поведении} \geq \frac{\text{Многообразие возмущений}}{\text{Многообразие управлений}}$$

Смысл этого довольно расплывчатого утверждения таков: если необходимо, чтобы система реализовала заданный вид поведения вне зависимости от внешних помех, то подавить многообразие в ее поведении можно, только увеличив множества управлений. Или как гласит *принцип необходимого многообразия* Эшби: многообразие может быть разрушено только многообразием. Такое утверждение является в некотором смысле кибернетическим аналогом второго закона термодинамики, и оно тесно связано с теорией информации по Шенону.

Уровни взаимодействия

Наконец, последний аспект понятия структурной сложности — это относительная сила взаимодействия между различными компонентами системы и уровнями иерархии.

В ряде случаев слабые взаимодействия, вообще говоря, повышают сложность системы. Однако практически этими взаимодействиями часто можно пренебречь и таким образом получить менее сложную модель системы.

Например, системе

$$\begin{aligned}\dot{x}_1 &= x_1, & x_1(0) &= 1, \\ \dot{x}_2 &= x_2, & x_2(0) &= 1, \\ \dot{x}_3 &= x_3, & x_3(0) &= 1\end{aligned}$$

логично приписать сложность 1, так как каждый жордановский блок матрицы коэффициентов имеет размер 1. Близкой к ней системе

$$\begin{aligned}\dot{x}_1 &= x_1, \\ \dot{x}_2 &= x_2, \\ \dot{x}_3 &= \varepsilon x_2 + x_3,\end{aligned}$$

(здесь ε — параметр) можно было бы приписать сложность 2, так как матрица коэффициентов

$$F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \varepsilon & 1 \end{bmatrix}$$

имеет наибольший жордановский блок размера 2 для любого $\varepsilon \neq 0$. Однако решение, полученное для второй системы,

$$x_1 = x_2 = e^t, \quad x_3 = e^t + \varepsilon^t e^t,$$

показывает, что для достаточно малых ε ее поведение сколь угодно близко к поведению первой, поэтому и ее сложность практически можно считать также равной единице. При этом, конечно, нужно учитывать ограничения на ε , вытекающие из конкретного применения данной системы.

Выводы

Вышеприведенные замечания подтверждают наше мнение, что система не может быть универсально сложной. Она может быть сложной в одних отношениях и несложной в других или может быть сложной, только если используется определенным образом. Короче говоря, структурная сложность является многогранным понятием, которое необходимо изучать с различных позиций в зависимости от целей анализа и целей самой системы.

ДИНАМИЧЕСКАЯ СЛОЖНОСТЬ

Рассмотрим некоторые аспекты сложности, которые проявляются в динамическом поведении системы.

Случайность в сравнении с детерминизмом и сложностью

Как уже отмечалось, одним из основных интуитивных показателей сложности системы является ее динамическое поведение, а именно: степень трудности наглядного объяснения и предсказания траекторий движущейся системы. В общем случае можем ожидать, что структурная сложность системы

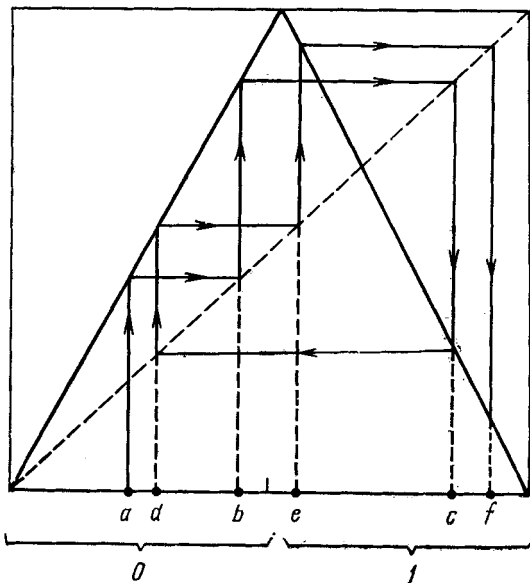


Рис. 4.1. Динамически сложный процесс.

оказывает влияние на динамическое поведение системы, а следовательно, и на ее динамическую сложность. Однако обратное не верно. Система может быть структурно простой, т. е. иметь малую структурную сложность, но ее динамическое поведение может быть чрезвычайно сложным.

Покажем, что процесс, изображенный на рис. 4.1, является структурно простым, будучи в то же время динамически сложным. Правило порождения последовательности точек a, b, c, \dots следующее: стороны вписанного треугольника и диагональ единичного квадрата используются как «отражаю-

щие барьеры». Процесс начинается с произвольной точки, расположенной в основании треугольника. Типичная последовательность абсцисс последовательности точек приведена на рис. 4.1.

Можно показать, что, приписывая каждой точке слева от середины основания треугольника число 0, а каждой точке справа — 1, получим последовательность чисел 0 и 1, порожденную этой детерминированной процедурой и математически неотличимую от последовательности, получаемой в распределении по закону Бернулли с параметром $p = 1/2$ (другие значения p могут быть получены использованием прямых, отличных от диагонали квадрата).

Этот результат имеет определенное методологическое и теоретико-системное значение. Действительно, если считать последовательность 0 и 1 выходом некоторого процесса, то не существует математического метода, позволяющего определить, является ли внутренний механизм, преобразующий вход и выход (последовательность 0 и 1), детерминированным или стохастическим. Иными словами, если не заглядывать внутрь «черного ящика», то никакие математические операции не могут помочь определить, является ли базисный механизм стохастическим или нет.

Рассмотренный пример подвергает серьезному сомнению слишком категорические утверждения о том, что глубинная природа физических процессов принципиально стохастична. Конечно, можно утверждать, что теория вероятности и статистика являются удобными инструментами для описания ситуаций, для которых характерна большая степень неопределенности. Однако нет априорных математических оснований полагать, что механизм, порождающий неопределенность, по своей природе непременно стохастичен; это может быть и некоторый детерминированный процесс, подобный вышеприведенному.

Очевидно, что если интерпретировать динамическую сложность как способность предсказывать поведение системы, то процесс, изображенный на рис. 4.1, очень сложен, так как наблюдаемый выход полностью случаен.

Для заданной последовательности выходов лучшей математической моделью для предсказания следующего выхода будет бросание монеты.

Шкалы времени

Другим важным аспектом динамической сложности является вопрос о различных шкалах времени для различных частей процесса. Часто возможны такие ситуации, когда

скорости изменения компонент одного и того же процесса различны: одни компоненты изменяются быстрее, а другие — медленнее. Типичным примером такого процесса является регулирование уровня воды в системе водохранилищ. Для управления на уровне индивидуального распределения воды требуется принимать решения ежедневно (или даже ежечасно), хотя решение об общем потоке воды через вход — выход принимается раз в месяц или раз в квартал. Очевидно, что флуктуации на локальном уровне значительно больше, чем на уровне всей системы водохранилищ.

Проблема различных шкал времени напоминает проблемы, с которыми мы сталкиваемся в численном анализе при интегрировании «жесткой» системы дифференциальных уравнений или когда имеем дело с некорректной проблемой. Простой пример некорректности представляет линейная система

$$\begin{aligned}\ddot{x} - 25x &= 0, \\ x(0) &= 1, \quad \dot{x}(0) = -5.\end{aligned}$$

Теоретически эта задача имеет решение

$$x(t) = e^{-5t}.$$

Однако при решении этой задачи численными методами в вычисления войдет дополнительный член

$$\bar{x}(t) = e^{5t}$$

с малым множителем ϵ . Таким образом, в действительности мы вычислим

$$x^*(t) = e^{-5t} + \epsilon e^{5t}.$$

Если t (или ϵ) достаточно мало, то все в порядке; однако когда ошибка округления слишком велика (большое ϵ) или когда желательно найти решение на большем интервале (большое t), то истинное решение полностью доминируется дополнительным членом $\bar{x}(t)$.

В ряде случаев трудности могут быть связаны не с вычислительными процедурами, а с самим решением системы. Для примера «жесткая» система

$$\begin{aligned}\dot{x}_1 &= x_1 + 2x_2, & x_1(0) &= 0, \\ \dot{x}_2 &= -10x_2, & x_2(0) &= 1\end{aligned}$$

имеет решение

$$\begin{aligned}x_1(t) &= -2/11 [e^{-10t} - e^t], \\ x_2(t) &= e^{-10t}.\end{aligned}$$

Таким образом, первая компонента процесса изменяется на порядок быстрее, чем вторая, и любая попытка рассчитать траекторию системы численно требует использования такого малого шага интегрирования, который позволяет аккуратно отследить «быструю» компоненту.

Это явление «жесткости» в системах (терминология инженеров-механиков), очевидно, оказывает влияние на динамическую сложность, так как точное предсказание поведения системы требует дополнительных затрат на вычисление.

Выводы

Только что приведенные примеры еще раз подтверждают, что большая размерность системы (пространства состояний или числа компонент) не обязательно означает большую сложность системы и наоборот. Сложность это слишком тонкое понятие, чтобы описывать его исключительно в понятиях размерности.

ВЫЧИСЛИТЕЛЬНАЯ СЛОЖНОСТЬ

Существуют различные теории сложности для хорошо определенных систем, и эти теории не всегда, естественно, связаны или даже сравнимы между собой. Одна из таких теорий требует использования машин Тьюринга и алгоритмов для этих машин, порождающих вычислимые функции. Структура и длина этих алгоритмов образуют основу, с помощью которой можно определить «вычислительную сложность». К сожалению, интуитивное понимание сложности, данное выше, не может быть применимо к машинам Тьюринга, так как в принципе машина Тьюринга может, вообще говоря, все (предполагается, что она имеет неограниченную память и неограниченное время для выполнения вычислений), т. е. она имеет бесконечную сложность. Ситуация становится более интересной, когда рассматривают только алгоритмы независимо от машин. При вычислении данного множества функций следует различать выполняющие это действие алгоритмы по их вычислительной сложности. Для того чтобы определить уровень сложности алгоритма, необходимо рассмотреть всевозможные вычисления и алгоритмы, а затем показать, сколько шагов данный алгоритм требует для вычисления конкретной функции.

Проблему вычислительной сложности можно исследовать с различных точек зрения (см. литературу в конце главы). Использование понятия вычислительной сложности, вообще говоря, значительно ограничивает возможности системного

анализа. Отметим только, что при любом определении сложности системы сложность всего процесса вычислений определяется сложностью компонентных вычислений. Рассмотрим аксиоматический подход к проблеме сложности.

АКСИОМЫ СИСТЕМНОЙ СЛОЖНОСТИ

Прежде чем дать математическое определение сложности системы, необходимо перечислить те основные свойства, которыми должна обладать любая ее мера, если она построена в соответствии с вышеприведенными понятиями о сложности. Как и обычно, в аксиоматических подходах можно спорить относительно начального набора аксиом, однако, как только аксиомы приняты, далее можно получить вид конкретной меры, дающей определенное понимание проблемы сложности.

Так как основные аспекты системного анализа, которые должны быть отражены в любой мере сложности, — это иерархичность, связность, динамическое поведение, то и аксиомы, которые мы будем конструировать, должны учитывать эти структурные компоненты системы. Обозначив $\theta(\Sigma)$ произвольную вещественную меру сложности, определенную для системы Σ , введем следующие аксиомы.

Аксиомы сложности

1. *Иерархия.* Если Σ_0 подсистема Σ , то

$$\theta(\Sigma_0) \leq \theta(\Sigma),$$

т. е. подсистема не может быть более сложной, чем система в целом.

2. *Параллельное соединение.* Если $\Sigma = \Sigma_1 \oplus \Sigma_2 \oplus \dots \oplus \Sigma_k$, т. е. Σ является параллельным соединением систем $\{\Sigma_i\}$, то

$$\theta(\Sigma) = \max_{1 \leq i \leq k} \theta(\Sigma_i).$$

3. *Последовательное соединение.* Если $\Sigma = \Sigma_1 \otimes \Sigma_2 \otimes \dots \otimes \Sigma_k$, т. е. Σ является последовательным соединением подсистем $\{\Sigma_i\}$, то

$$\theta(\Sigma) \leq \theta(\Sigma_1) + \theta(\Sigma_2) + \dots + \theta(\Sigma_k).$$

4. *Соединение с обратной связью.* Если присутствует операция обратной связи \odot из системы Σ_2 в систему Σ_1 , то

$$\theta(\Sigma_1 \oplus \Sigma_2) \leq \theta(\Sigma_1) + \theta(\Sigma_2) + \theta(\Sigma_2 \odot \Sigma_1).$$

(Заметим, что аксиома 3 является частным случаем аксиомы 4, если отсутствуют обратные связи.)

5. *Нормализация.* В классе систем, удовлетворяющем этим аксиомам, выделено подмножество систем \mathcal{P} , для которых

$$\theta(\Sigma) = 0 \quad \text{для всех } \Sigma \in \mathcal{P}.$$

Этих аксиом оказывается вполне достаточно для определения меры сложности систем, задаваемых различными способами; в некоторых случаях они однозначно определяют меру сложности.

Заметим, что эти более или менее стандартные декомпозиционные аксиомы учитывают изложенные выше представления о сложности. Поскольку любая система может быть представлена в виде последовательно-параллельно или каскадно (иерархически) соединенных подсистем (возможно, с обратной связью), то аксиомы 2—4 поясняют структуру связности таких декомпозиций. Иерархические аспекты отражены в аксиоме 1, в то время как динамические аспекты — в аксиоме 5. Таким образом, приведенные аксиомы сложности кажутся довольно разумными по крайней мере потому, что не упущен ни один важный аспект сложности.

При изучении процессов с конечным числом состояний, эти аксиомы являются наименьшим множеством, задающим меру сложности однозначно. Это еще раз подтверждает естественность выбранных аксиом. Наконец, аксиомы очень удобны для различных алгебраических подходов к анализу и вычислению сложности.

Теория сложности для топологических форм также представляет значительный интерес, хотя, к сожалению, до сих пор не ясно, как задавать топологию форм с помощью алгебраического аппарата. Поэтому необходимо либо найти такие процедуры, либо ввести новую меру топологической сложности. В последнем случае, для того чтобы успешно выполнить задачу, необходимо модифицировать вышеприведенные аксиомы.

СЛОЖНОСТЬ АВТОМАТОВ

С системно-теоретической точки зрения наиболее продвинутой теорией является теория сложности процессов, моделируемых конечными автоматами. Напомним, что каждый автомат M порождает конечную полугруппу преобразований (Q, \mathcal{F}^*) , где Q — пространство состояний автомата M , а \mathcal{F}^* — полугруппа всех преобразований Q . Дадим определение сложности, используя алгебраическое понятие узлового произведения, и рассмотрим связь этого определения с автоматами и аксиомами сложности.

Определение 4.1

Назовем групповой сложностью $\#_G(X, S)$ полугруппы (X, S) наименьшее целое n , такое, что

$$S \mid (Y_n, C_m) \omega (X_n, G_n) \omega \dots \omega (Y_1, C_1) \omega (X_1, G_1) \omega (Y_0, C_0),$$

где G_1, \dots, G_m — конечные простые группы, C_0, \dots, C_k — конечные комбинаторные полугруппы (триггеров); « ω » обозначает операцию узлового произведения.

Следовательно, групповая сложность полугруппы (X, S) — это минимальное число возможных комбинаций простых групп и комбинаций комбинаторных полугрупп, составляющих (X, S) . Используя теоремы декомпозиции, можно определить сложность в терминах декомпозиции фазового пространства. Например,

$$\#_G(X, S) = \min \#_G \left\{ \begin{array}{l} T: T \text{ — последовательно-параллельная} \\ \text{или каскадная декомпозиция} \end{array} \right\}.$$

Так как комбинаторные полугруппы представляют автоматы, а не вычисления, то основным сложным элементом является простая группа, в которой производятся такие элементарные арифметические действия, как сложение и умножение. В соответствии с теорией Крона — Роудза, возможна декомпозиция автомата в примитивные и неприводимые подавтоматы, причем решение задачи зависит от структуры компонент и длин вычислений. Поэтому и сложность зависит не только от длины цепи вычислений, но и от степени сложности каждой компоненты, входящей в цепь. Таким образом, сложность учитывает не только общее число вычислений в цепи (вычислительный аспект), но и внутреннюю сложность подавтоматов, отражаемую узловым произведением (структурный аспект). Эвристически вычислительная часть автомата представляется количеством «циклов» в машинной программе, которая вычисляет действие \mathcal{F}^* на Q .

Рассмотрим взаимосвязь этого определения групповой сложности и вышеприведенных аксиом сложности.

Теорема 4.1¹⁾

Пусть $\theta: M_{\mathcal{F}} \rightarrow N$ (где N — множество неотрицательных чисел, а $M_{\mathcal{F}}$ — множество всех конечных автоматов) удовлетворяет всем аксиомам сложности, причем аксиома нормализации (аксиома 5) имеет место с

$$\theta(U_3) = 0, \quad \theta(D_1) = 0,$$

¹⁾ Доказательство теоремы можно найти в работах, приведенных в конце главы.

где U_3 — триггер с тремя состояниями, а D_1 — автомат задержки на 1 шаг. Тогда

$$\theta(f) = \#_G(f^S)$$

для всех $f \in M_{\mathcal{G}}$ (f^S — это полугруппа автомата f).

Замечания. 1. Согласно теореме 4.1, функция групповой сложности является, по существу, единственной функцией, удовлетворяющей аксиомам сложности и данной нормализации. Для любой другой функции, удовлетворяющей аксиомам, должно выполняться неравенство

$$\#_{G'}(f) \leq \#_G(f) \quad \text{для всех } f \in M_{\mathcal{G}}.$$

2. Автомат задержки на 1 шаг, так же как U_3 , является комбинаторной подгруппой, действие которой задается следующим образом:

$$D_1(a_1, \dots, a_n) = a_{n-1}, \quad D_1(a_1) = \text{null},$$

где $\{a_n, a_{n-1}, \dots, a_1\}$ — любая вводимая строка.

3. $\theta(f) = 0$ тогда и только тогда, когда f представляет собой последовательно-параллельно соединенные автоматы U_3 и D_1 .

4. Для того чтобы показать, что существуют автоматы любого уровня сложности, рассмотрим автомат

$$U = \{a, b, c\},$$

$$Y = \{0, 1\},$$

$$Q = \{1, 2, \dots, n\},$$

$$\lambda(i, a) = i + 1 \pmod{n}, \quad i = 1, 2, \dots, n,$$

$$\lambda(1, b) = 2,$$

$$\lambda(2, b) = 1,$$

$$\lambda(x, b) = x, \quad x \neq 1, 2,$$

$$\lambda(1, c) = 2,$$

$$\lambda(y, c) = y, \quad y \neq 1,$$

$$\delta(q, d) = q \pmod{2}, \quad q \in Q, \quad d \in U.$$

Используя вышеприведенные результаты и некоторые положения теории полугрупп, можно показать, что сложность этого автомата

$$\theta(C) = n - 1$$

независимо от начального состояния $q \in Q$.

5. Нормализация U_3 и D_1 необходима в силу того, что они являются самыми простыми объектами теории: ни один из этих автоматов не выполняет вычислений и их действия полностью предсказуемы.

Нормализация напоминает известный в теории информации процесс выбора событий, имеющих нулевое информационное содержание. Автоматы U_3 и D_1 действуют регулярно, от них нельзя ожидать каких-либо сюрпризов, и поэтому их поведение не порождает информации. Следовательно, если обнаруживаются подсистемы, которые ведут себя подобно триггерам, то они могут быть удалены без изменения структурной сложности других подсистем, несмотря на то что вычислительная сложность может уменьшиться.

ЭВОЛЮЦИОННАЯ СЛОЖНОСТЬ И ЭВОЛЮЦИОННЫЕ СТРУКТУРЫ

В качестве дискуссионного, но многообещающего примера использования понятия сложности рассмотрим некоторые аспекты эволюции организмов с учетом следующего принципа эволюции:

Организм в процессе эволюции претерпевает такие изменения, в результате которых максимизируется контакт с окружающей средой в целях лучшего ее понимания и разумного управления последней.

«Контакт с окружающей средой» измеряется объемом памяти, необходимой для имитации поведения организма на компьютере, или, что то же самое, числом триггеров, требуемых для построения автомата, имитирующего поведение организма. Полный контакт с окружающей средой включает все формы контакта как физические, та и мысленные. Когда говорят о «понимании и разумном управлении окружающей средой», имеют в виду, что сложность взаимодействия между организмом и средой достаточно высока и численно приблизительно равна значению величины контакта. Можно показать, что увеличение контакта вызывает соответствующее возрастание сложности, а так как сложность тесно связана с пониманием (сложность — это минимальное число координат, необходимых для описания процесса), то организм, имеющий такую сложность в своих отношениях с внешней средой, должен ее понимать и быть способен на нее воздействовать.

Таким образом, если организм может увеличить одновременно контакт, управляемость, понимание и сложность, то это следует рассматривать как прогресс в эволюции. Принцип эволюции как раз и формализует эти эвристические доводы.

Перейдем к математическому описанию вышеприведенных явлений. Обозначим через \bar{E}_t конечный автомат, представляющий окружение в момент t , а через O_t — организм. Через R_t будем обозначать автомат, описывающий результат взаимодействия организма O_t со средой E_t (напомним, что результат взаимодействия между двумя автоматами R представляет по определению конечный автомат

$$R = (U, \text{range } \bar{\delta}, \text{range } \bar{\delta}, \lambda, j),$$

где $\text{range } \bar{\delta}$ — множество представлений устойчивых петель взаимодействия, а $j: \text{range } \bar{\delta} \rightarrow \text{range } \bar{\delta}$ — тождественное отображение). Далее положим $C_t = \#(R_t)$ — количество состояний R_t , а $c_t = \theta(R_t) + 1$ — функция сложности R_t , увеличенная на единицу. Таким образом, C_t является контактным числом в момент t , а c_t — числом сложности.

ПРОБЛЕМЫ ВЫБОРА И СЛОЖНОСТЬ

На уровне личных или социальных проблем сложность тесно связана с возможностью или невозможностью рационального выбора решений человеком. В связи с этим сложность отражает тип неопределенности, который не поддается обработке вероятностными методами.

Отличительной чертой сложности, присущей принятию решений, является невозможность доказать существование функции, представляющей предпочтение выборов. В этом смысле социальные проблемы выбора подобны проблемам, возникающим перед игроком в шахматы, когда для поиска «удовлетворительной стратегии» он должен решать проблему комбинаторной размерности. Однако поиск среди всех возможных вариантов находится за пределами возможностей человека или вычислительной машины. Следовательно, подход, в основе которого лежит гипотеза о нахождении оптимального поведения, не дает возможности справиться с нетривиальными проблемами выбора в сложных ситуациях. В таких случаях надо перейти к гипотезе нахождения рационального решения, которое можно было бы найти за приемлемое время.

ВНУТРЕННЯЯ СЛОЖНОСТЬ В СРАВНЕНИИ СО СЛОЖНОСТЬЮ УПРАВЛЕНИЯ

Ранее была рассмотрена функция сложности c_t , относящаяся к взаимодействию между организмом и окружающей средой, а также функция C_t , описывающая контакт между

организмом и средой. Эти два понятия тесно связаны с двумя типами сложности принятия решения, которые назовем *внутренней сложностью* и *сложностью управления*.

Внутренняя сложность определяется уровнем сложности подгруппы преобразований при полном использовании потенциальных возможностей системы. Сложность управления является числовым выражением сложности, которая получается из количества вычислений, необходимых для полного управления системой. Качественно устойчивое правило принятия решений — это такое правило, для которого эти два числа совпадают. Однако в большинстве практических ситуаций внутренняя сложность будет превышать сложность управления. В определенном смысле эти понятия, естественно, связаны с проблемами «оптимального» и «удовлетворительного» выбора соответственно. Таким образом, внутренняя сложность имеет непосредственное отношение к таким правилам принятия решения, которые в определенном смысле являются наилучшими и в общем определяются некоторыми принципами оптимальности. Однако оптимальное решение не может быть реализовано лицом, принимающим решение, в силу ограниченности вычислительных ресурсов (сложность управления). Поэтому можно определить эволюционную сложность как разность между внутренней сложностью и сложностью управления. Чем меньше эта разница, тем более устойчиво (сбалансировано) ведет себя система. (Это понятие тесно связано с адаптируемостью системы.) Для того чтобы проиллюстрировать приведенные принципы, рассмотрим типичный пример «конфликта между личными и коллективными потребностями».

На определенной ограниченной площади живут ковбои. Каждый ковбой стремится держать наибольшее возможное количество коров, что эквивалентно для него максимизации полезности. Предположим, что дополнительная полезность от приобретения еще одной коровы равна $+1$ для любого ковбоя. Однако оценка полезности этого мероприятия другими (коллективом) отрицательна в силу увеличения нагрузки на ограниченные ресурсы.

Кризисная ситуация определяется несоответствием личных выгод и коллективной необходимости или, в наших понятиях, несоответствием между внутренней сложностью и сложностью управления. В этом случае нельзя избежать конфликта, если только не добиться значительных изменений «уровня понимания» целей (за счет обучения, адаптации, управления) или изменить внутреннюю структуру потребностей объекта. Примеры такого типа часто встречаются в экологии.

ПРОГРАММА ДЛЯ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ ПОНЯТИЯ СЛОЖНОСТИ

В большинстве практических применений понятия *сложность* необходимо провести следующие этапы исследования:

— попытаться найти приемлемую аналогию между реальной системой и математической моделью системы, сформулированной на языке автоматов;

— принимая во внимание историю развития системы, вычислить различные «разумные» подгруппы преобразований системы (используя всевозможные функциональные уравнения);

— вычислить внутреннюю сложность подгрупп;

— вычислить «рациональный» уровень контактной функции «естественной» подгруппы преобразований;

— вычислить сложность управления этой «естественной» подгруппой преобразований;

— вычислить контактные функции и сложности, которые позволяют получать устойчивые конфигурации для конкретных процессов.

Достижение устойчивости конфигурации не обязательно означает, что подгруппа преобразований внутренней структуры системы идентична подгруппе преобразования, используемой при управлении, вполне достаточно, чтобы они изменялись в одном направлении для всех состояний процесса.

ПОЛИЭДРАЛЬНАЯ ДИНАМИКА И СЛОЖНОСТЬ

До сих пор наше изложение касалось изучения меры сложности систем, задаваемых при помощи теории конечных автоматов, при этом неоднократно подчеркивалось, что выбор соответствующего математического аппарата, применяемого в данной ситуации, почти полностью зависит от целей формулирования данной проблемы. Однако мера сложности, предложенная для систем, описываемых с помощью конечного пространства состояний, может и не иметь применения, если используется другое описание процесса. Рассмотрим понятие сложность в случае алгебраическо-топологического описания процесса в виде симплициального комплекса.

Так как базисным алгебраическим объектом полиэдральной динамики являются симплексы, определим меру сложности с помощью размерности симплексов и учета связей между ними. Отметим, что такое определение сложности включает только статическую структурную сложность рассматриваемой системы. Динамические же соображения, вероятно, лучше всего отображать с помощью теории конечных автоматов.

Для определения полиэдральной сложности будем использовать следующую версию аксиом сложности:

А. Система, состоящая из единственного симплекса, имеет сложность равную 1.

Б. Подсистема (подкомплекс) имеет сложность, не большую, чем весь комплекс.

В. Объединение двух комплексов образует новый комплекс, сложность которого не больше, чем сумма сложностей компонент.

Заметим, что аксиомы А—В неявно предполагают, что рассматриваемая система связана на нулевом уровне, т. е. структурный вектор Q системы имеет $Q_0 = 1$. Если это не так, то можно вычислить сложность каждой из компонент связности комплекса, а затем максимальное из этих чисел принять за сложность всей системы. Такая процедура эквивалентна рассмотрению всей системы как параллельно соединенных ее компонент связности (на 0-уровне).

В качестве меры, удовлетворяющей принятым аксиомам, возьмем следующую функцию, легко вычисляемую из структурного вектора Q :

$$\psi(K) = 2 \left[\sum_{i=0}^N (i+1) Q_i \right] / (N+1)(N+2),$$

где N — размерность комплекса K , а Q_i — i -я компонента структурного вектора Q , получаемого в процессе q -анализа. Множитель $2/(N+1)(N+2)$ введен главным образом для нормализации в соответствии с аксиомой А.

Для иллюстрации использования меры ψ , рассмотрим пример системы хищник — жертва (гл. 3). Отношение $\lambda_{\text{хищник}}$ имеет структурный вектор

$$Q_{\text{хищник}} = \begin{pmatrix} 5 & & & & & 0 \\ 1 & 1 & 2 & 3 & 2 & 1 \end{pmatrix},$$

а отношение $\lambda_{\text{жертва}}$ —

$$Q_{\text{жертва}} = \begin{pmatrix} 5 & & & & & 0 \\ 1 & 3 & 3 & 4 & 2 & 1 \end{pmatrix}.$$

Таким образом,

$$\psi(\text{хищник}) = 11/7, \quad \psi(\text{жертва}) = 50/21,$$

что указывает, что отношения жертв устроены несколько более «сложно», чем отношения хищников.

АЛГЕБРАИЧЕСКАЯ ТЕОРИЯ СИСТЕМ И СЛОЖНОСТЬ

Определение сложности, данное для конечных автоматов, вообще говоря, может быть распространено на случай конечномерных линейных систем посредством несколько искусственной алгебраической конструкции полугруппы линейной системы. Основным недостатком такого представления является то обстоятельство, что полугрупповая «реализация» линейной системы лишь случайно может оказаться минимальной (т. е. канонической). Да это и неудивительно, так как теория Крона — Роудза дает теоремы существования для последовательно-параллельных реализаций, которые, как правило, не минимальны. Математически наиболее элегантный путь — это использовать теорию модулей, которая, как было отмечено выше, всегда приводит к минимальной реализации. Однако для этого потребуются привлечение значительного алгебраического аппарата теории модулей. Подойдем к вопросу о сложности конечномерных линейных систем с позиций полиномиальной и линейной алгебры.

Для общего обоснования вида функции сложности обратимся к рис. 3.7, изображающему общую структуру линейной системы, и теореме реализации (гл. 3). Было показано, что любая линейная динамическая система Σ может быть представлена как прямая сумма подсистем Σ_i ; Σ_i определяется i -м нетривиальным фактором матрицы

$$\psi_w W(z),$$

где $W(z)$ — передаточная функция системы Σ , а ψ_w — ее характеристический многочлен. Компоненты Σ_i представляют неприводимые составные блоки системы Σ , поэтому определив меру сложности для каждого из них, получим меру сложности системы Σ .

Так как инвариантные факторы $\psi_w W(z)$ однозначно описывают структуру системы Σ , возьмем в качестве меры ее сложности

$$\xi(\Sigma) = \sum_{i=1}^q (n - \deg \psi_i + 1) \log (n - \deg \psi_i + 1),$$

где $n = \dim \Sigma$, а ψ_i — i -й нетривиальный инвариантный фактор $\psi_w W(z)$. Достаточно просто проверить, что мера ξ удовлетворяет вышеприведенным аксиомам. При этом системой, сложность которой равна 0, является циклическая система. Вполне закономерно, что мера ξ очень похожа на меру информации, содержащейся в строке символов, так как

аксиомы сложности тесно связаны с «естественными» аксиомами для меры энтропии.

Заметим, что, хотя мера ξ основана на степенях ψ_i , размерность пространства состояний также играет в ней существенную роль. Например, для двух систем с подобной циклической структурой система большей размерности будет и более сложной, что отвечает нашему интуитивному пониманию меры ξ .

Рассмотрим две системы, описываемые передаточными матрицами

$$W_2(z) = \begin{bmatrix} \frac{1}{z} + \frac{1}{z^2} & \frac{1}{z} + \frac{1}{z^2} & \frac{1}{z} \\ 0 & 0 & \frac{1}{z} \end{bmatrix}.$$

Легко показать, что матрица $\psi_w W_1(z)$ имеет единственный инвариантный фактор

$$\psi_1^1 = (z + 1)(z + 2)(z + 3).$$

Следовательно,

$$\zeta(\Sigma_1) = 0.$$

Матрица $\psi_w W_2(z)$ имеет два инвариантных фактора $\psi_1^2 = z$, $\psi_2^2 = z^2$. Таким образом, сложность e_2 системы Σ_2 равна

$$\begin{aligned} \zeta(\Sigma_2) &= \sum_{i=1}^2 (n - \deg \psi_i + 1) \log(n - \deg \psi_i + 1) \\ &= (3 - 1 + 1) \log(3 - 1 + 1) + (3 - 2 + 1) \log(3 - 2 + 1) \\ &= 3 \log 3 + 2 \log 2. \end{aligned}$$

Как и следовало ожидать, нетривиальная циклическая структура системы Σ_2 делает ее значительно более сложной, чем система Σ_1 .

НЕЛИНЕЙНЫЕ КОНЕЧНОМЕРНЫЕ ПРОЦЕССЫ

Определение сложности для системы, задаваемой нелинейным дифференциальным уравнением

$$\dot{x} = f(x, u),$$

представляет собой значительно более трудную проблему. Это объясняется тем, что для такой системы нет сжатого алгебраического представления типа $W(z)$, характеризующего структуру системы с позиций оператора входа — вы-

хода. В этом случае можно предложить следующие два подхода.

Во-первых, можно аппроксимировать конечномерное пространство состояний процесса некоторым дискретным пространством состояний подобно тому, как это делается в численных методах. Например, пространство состояний R^n может быть дискретизировано ограничением координатных направлений при помощи неравенств

$$a_i \leq x_i \leq b_i, \quad i = 1, 2, \dots, n$$

и затем введением структуры решетки в этот конечный гиперблок. После этого можно определить динамику системы с конечным числом состояний, индуцируя ее из исходной непрерывной системы. Имея теперь аппроксимацию исходной системы, можно применить к ней результаты, касающиеся процессов с конечным числом состояний. Заметим, однако, что установление обоснованности перехода к дискретной задаче является нетривиальной задачей. В любом случае необходимо установить инвариантность структуры относительно подразделения и показать, что решение дискретной задачи сходится (в некотором смысле) к истинному решению, когда решетка становится все тоньше и тоньше.

Второй подход — это использование результатов гл. 3. Если функция $f(x, u)$ аналитична по x и непрерывна по u , то теорема Кренера является аналогом теоремы Крона — Роудза. К сожалению, теория сложности для таких систем пока еще не создана, хотя, казалось бы, нет особых препятствий к распространению результатов о конечных автоматах на конечномерные аналитические системы.

СЛОЖНОСТЬ И ТЕОРИЯ ИНФОРМАЦИИ

Как уже отмечалось, между классической теорией информации по Шеннону — Винеру и понятием *сложность системы* существует тесная связь. Действительно, в некоторых ранних работах по исследованию сложности биологических систем сложность определялась как число различимых единиц, составляющих организм. Такой подход явно наводит на мысль о сравнении меры сложности с информационным содержанием строки символов. Однако теория информации не является удовлетворительной основой для определения сложности. Было показано, что система является целостным объектом, а не просто объединением своих частей. Кроме того, системные переменные не действуют по отдельности и только в совокупности с другими порождают сложные явления. Отдельные переменные могут быть не так важны, как их

комбинации, соответствующие этим явлениям, а теория информации не может идентифицировать такие комбинации. Подобно всем статическим теориям, она игнорирует тот факт, что относительное расположение элемента в структуре может оказывать сильное влияние на систему. Другими словами, численных значений частот различных элементов в системе не достаточно для объяснения ее поведения, требуется еще информация о способе, которым эти элементы связаны.

Одним из интересных подходов, удовлетворяющих этому требованию, является понятие *аналогичные явления*. При этом подходе постулируется, что исходные размеры переменными x_1, x_2, \dots, x_n можно заменить безразмерными переменными P_1, P_2, \dots, P_k , которые представляют собой комбинации x_i . Количество таких безразмерных переменных определяется согласно теореме, утверждающей, что размерное однородное уравнение

$$F(x_1, x_2, \dots, x_n) = 0$$

может быть выражено с помощью переменных P , образуемых из x_i так, что

$$f(P_1, P_2, \dots, P_{n-r}) = 0, \quad r \leq m,$$

где r — ранг размерной матрицы исходных переменных n , а m — число основных размерностей физических величин, таких, как масса, длина.

Так как фактически все физические законы размерностно инвариантны, то безразмерные комбинации P_i могут интерпретироваться как *критерии подобия*. Таким образом, появляется возможность значительно сократить число системных переменных, принимаемых во внимание. В то же самое время, обычно присущая теоретико-информационным проблемам нестационарность может быть значительно редуцирована, поскольку любые изменения, ведущие к нестационарности, вероятно, будут значительно меньше на макроскопическом уровне (P -переменные), чем на микроскопическом (x -переменные).

Необходимо поэтому концентрировать внимание на фундаментальных законах, управляющих поведением системы, а не на конкретных моделях, представляющих процесс. Другими словами, следует стремиться описать сложность в терминах инвариантных свойств структуры системы. Для того чтобы иметь возможность распознавать различные множества конфигураций системы, введем *размерностно инвариантную дискриминантную функцию*

$$Y = a_1 P_1 + \dots + a_k P_k,$$

где весовые множители (коэффициенты) определяются так, чтобы максимизировать t -статистику или F -отношение между различными группами (подсистемами).

Итак, необходимо максимизировать отношение межгрупповой вариации:

$$\phi(a_1, \dots, a_k) = \frac{n_1 n_2 (a, d)^2}{n_1 + n_2 (a, Ca)},$$

где $d' = (d_1, d_2, \dots, d_k)$ — вектор средних разностей $k (= n - r)$ размерно инвариантных функций P_i , C — внутригрупповая ковариационная матрица, а n_1 и n_2 — число наблюдений в двух группах. Пусть Y_{lk}^m значение m -й размерно инвариантной дискриминантной функции, вычисленное для k -го элемента l -й группы. Апостериорная вероятность того, что k принадлежит l -й группе, при условии, что он действительно в группе m (случай k_m), дается формулой

$$P_{lmk_m}^{m'} = \frac{P_m \exp(Y_{lmk_m}^{m'})}{\sum_{i=1}^r P_i \exp(Y_{lik_m}^{m'})},$$

где индекс m' обозначает, что берется конкретная дискриминантная функция, которая дает максимальную истинную вероятность для элементов группы, r — общее число группы, P_m — априорная вероятность того, что k лежит в группе m .

Теперь можно определить сложность системы через информационное содержание, т. е. как меру средней неопределенности в местоположении элементов. Конкретно определим сложность группы m равенством

$$H_m = \frac{1}{n_m} \sum_{k_m=1}^{n_m} G_{k_m},$$

где

$$G_{k_m} = - \sum_{i=1}^r P_{imk_m}^{m'} \log_2 P_{imk_m}^{m'},$$

а n_m — число элементов в группе m .

Можно также определить меру избыточности i -й группы

$$R_i = 1 - \frac{H_i}{\log_2 r}.$$

ЛИТЕРАТУРА

Одной из первых работ, непосредственно посвященной понятию сложности системы, является статья Simon H., The Architecture of Complexity, *Proc. Am. Soc.*, 106, 467—482 (1962), которая до сих пор остается прекрасным введением в обсуждение технических вопросов, возникающих при изучении сложности.

Morin E., Complexity, *Int. Soc. Sci. J.*, 26, 583—597 (1974);
 Weaver W., Science and Complexity, *Am. Sci.*, 36, 536—544 (1968);
 Baldwin M., Portraits of Complexity, Battelle Monographs, Battelle Institute, Columbus, Ohio, June 1975.

Структурная сложность

Обзор большинства иерархических аспектов понятия структуры системы дан в книге

Pattee H., ed., *Hierarchy Theory*, Braziller, New York, 1973.

Материал, касающийся построения надежных устройств из ненадежных компонент с точки зрения иерархических соображений обсуждается в книге

Arbib M., *Brains, Machines and Mathematics*, McGraw-Hill, New York, 1964. [Имеется перевод: Майкл М., Арбиб. Мозг, машины и математика.— М.: Наука, 1968.]

Многообразия

Классической работой по многообразиям систем и связи их с системными проблемами является книга

Ashby W. R., *Introduction to Cybernetics*, Chapman and Hall, London, 1956. [Имеется перевод: Росс Эшби У. Введение в кибернетику.— М.: ИЛ, 1959.]

Работой, заполняющей пробел между идеями Эшби и подходом к системам с точки зрения пространства состояний, является статья

Porter B., Requisite Variety in the Systems and Control Sciences, *Int. J. Gen. Syst.*, 2, 225—229 (1976).

Взаимосвязь между управлением и сложностью с позиции программирования анализируется в работе

Юдин Д. Б., Горяшко А. П. Проблемы управления и теория сложности, 1, П. — *Техническая кибернетика*, 1974, 12, с. 10—24; 1975, 13, с. 1—13.

Вопросы сложности для систем, описываемых графами и отношениями, рассматриваются в работах

Нечипоренко В. И. Структурный анализ систем.— М.: Советское радио, 1977;

Горбатов В. А. Теория частично упорядоченных систем.— М.: Советское радио, 1976;

Ashby W. R., Requisite Variety and Its Implications for the Control of Complex Systems, *Cybernetica*, 1, 83—99 (1953);

Porter B., Cybernetics and Control Engineering, *J. Dyn. Syst. Meas. Control, Trans. ASME, Ser. G.*, 95, 349 (1973).

Юдин Д. Б., Горяшко А. П. Проблемы управления и теория сложности.— *Техническая кибернетика*, 1976, 14, № 3.

Уровни взаимодействия

Большое количество примеров, свидетельствующих о необходимости рассмотрения математической теорией сложности уровней взаимодействия, приведено в работах

Gottinger H., Complexity and Dynamics: Application of Dynamic System Theory, *IEEE Trans. Syst. Man Cybern., SMC*, 6, 867—873 (1976);

Ashby W. R., Gardner M., Connectance of Large, Dynamic Cybernetic Systems: Critical Values for Stability, *Nature*, 228, 784 (1970);
Gottinger H., Towards an Algebraic Theory of Complexity in Dynamic Systems, *General Systems Yearbook*, XXII, 73—83 (1977).

Читатели, незнакомяе с техникой получения жордановой формы матрицы, могут обратиться к работам

Bellman R., Introduction to Matrix Analysis, McGraw-Hill, New York, 1960. [Имеется перевод: Белман Р. Введение в теорию матриц. — М.: Мир, 1976.]

Гантмахер Ф. П., Теория матриц. — М.: Наука, 1971.]

Вычислительная сложность

Теория вычислительной сложности алгоритмов значительно более продвинута по сравнению с аналогичной теорией для больших систем. Обзор по современному состоянию этой проблемы дается в книгах

Traub J., ed., Analytic Computational Complexity, Academic, New York, 1976;

Karp R., ed., Complexity of Computation, SIAM-AMS Proceedings, Vol. 7, American Mathematical Society, Providence, Rhode Island, 1974.

Аксиомы системной сложности

Полное изложение и обсуждение этих аксиом даны в работе Gottinger H., Complexity and Dynamics: Application of Dynamic System Theory, *IEEE Trans. Syst. Man. Cyber.*, 867—873 (1976).

Можно обратиться также к лекциям

Rhodes J., Applications of Automata Theory and Algebra, Math. Dept., University of California, Berkeley, 1971.

Процессы выбора и сложность

Процессы выбора рассматриваются в книге

Gottinger H., Leinfellner W., eds., Decision Theory and Ethics: Issues in Social Choice, Reidel, Dordrecht, 1978.

В этом отношении также интересны работы

Futia C., The Complexity of Economic Decision Rules, Bell Laboratories, Murray Hill, New Jersey, January 1975;

Gottinger H., Complexity and Information Technology in Dynamic Systems, *Kybernetes*, 4, 129—141 (1975).

Полиэдральные динамики и сложность

Результаты данного раздела взяты из работы

Casti J., Complexity, Connectivity and Resilience in Complex Ecosystems, in IFAC Symposium on Bio- and Ecosystems, Leipzig, Germany, 1977.

Сложность и теория информации

Подход к сложности с помощью теории информации изложен в работах

Sahal D., System Complexity: Its Conceptions and Measurement in the Design of Engineering Systems, *IEEE Trans. Syst. Man Cybern.*, SMC-6 (1976);

- Moshowitz A., Entropy and the Complexity of Graphs, *Bull. Math. Biophys.*, 30, 175—204; 225—240 (1968);
Колмогоров А. Н. Три подхода к качественному определению информации, Проблемы передачи информации, 1965, 1 с. 1—7;
Cornacchio J. V., Maximum Entropy Complexity Measures, *Int. J. Gen Syst.*, 3, 267—271 (1977).

Обзор большинства существующих подходов к анализу сложности с точки зрения теории информации дан в работе
Sahal D., Elements of an Emerging Theory of Complexity Per Se, *Cybernetica*, 19, 5—38 (1976).

Устойчивость, катастрофы и адаптируемость больших систем

То ли колодец был очень глубок, то ли падала она очень медленно, только времени у нее было достаточно, чтобы прийти в себя и подумать, что же будет дальше.

Льюис Кэрролл. «Алиса в Стране чудес»

Форма и гармония мироздания родились по Космической воле, затем родилась Ночь, а затем — волнующийся океан Пространства; из волнующегося океана Пространства родилось Время, ведающее порядком дней и ночей в году, — правитель всякого движения.

«Ригведа», X, 190

Замеси на опилках, чтоб клеем покрыть,
И акриды добавь по норме.
Но запомни одно: придержи свою пруть —
Сохрани симметричную форму¹⁾.

Льюис Кэрролл. «Охота на Снарка»

При исследовании больших систем основное внимание, как правило, уделялось и уделяется изучению их поведения под действием разного рода внешних возмущений. При этом в классических работах анализируются возмущения, возникающие в начальном состоянии системы или на ее внешнем входе, а в современных работах — возмущения в структуре самой системы. И в том и другом случаях основной интерес представляет вопрос, будет ли поведение системы существенно меняться в результате нежелательных, неизвестных или незапланированных изменений в режиме управления. Однако такая постановка вопроса является довольно расплывчатой, и поэтому, для того чтобы сказать что-либо определенное (или интересное) относительно хода рассматриваемого процесса, необходимо сформулировать более конкретные вопросы в рамках определенного математического аппарата.

Цель данной главы состоит в том, чтобы дать обзор основных представлений в области устойчивости, имеющих непосредственное отношение к теории систем, и проследить, как практически возникают различные понятия об устойчивости. Основное внимание будет сосредоточено на более современных понятиях, таких, как адаптируемость, катастрофы и процессы распространения возмущений. Из классических

¹⁾ Перевод А. Д. Суханова. — *Прим. ред.*

результатов рассмотрим лишь самые основополагающие, отсылая читателей к литературе для ознакомления с другими исследованиями классического характера.

Как уже подчеркивалось, проблема системного исследования может иметь много различных, неэквивалентных математических формулировок, поэтому и проблемы устойчивости должны быть сформулированы соответствующим образом. В зависимости от конкретного описания систем для их анализа будут применяться разные подходы. Оправданием такого на первый взгляд бессистемного исследования будут служить примеры, которые иллюстрируют, с одной стороны, приложения теории, а, с другой стороны, показывают важность наличия гибкости в математической постановке задачи.

ВНЕШНЕЕ ОПИСАНИЕ

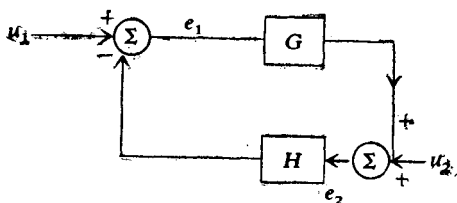
Для анализа устойчивости лучше всего рассмотреть внешнее описание системы с обратной связью

$$e_1 = u_1 - He_2, \quad (5.1)$$

$$e_2 = u_2 + Ge_1, \quad (5.2)$$

где величины e_1 , e_2 , u_1 и u_2 принадлежат некоторому расширенному функциональному пространству X , а операторы H и G отображают X в себя.

Систему (5.1), (5.2) можно представить как систему с обратной связью,



в которой оператор G является подсистемой в прямом направлении, H — подсистемой обратной связи, а величины u_1 , u_2 и e_1 , e_2 — соответственно значения на входе и отклонения. Можно считать, что значения на выходе системы представлены величинами Ge_1 и He_2 .

Исследование уравнений (5.1) и (5.2) в случае заданных u_1 , u_2 и некоторого множества $U \subseteq X$ дает нам ответы на два основных вопроса:

Имеет ли система (5.1), (5.2) единственное решение в X для e_1 и e_2 из X ?

Если система (5.1), (5.2) имеет какие-либо решения в X для e_1 и e_2 из X , то будут ли они действительно принадлежать пространству U ?

Первый вопрос по существу относится к проблеме существования и единственности, а второй — устойчивости. Вообще говоря, для изучения этих вопросов применяются различные аналитические методы, но в основном методы функционального анализа.

В качестве примера проблемы устойчивости рассмотрим случай, когда $\tilde{X} = U = L_\infty [0, \infty]$, т. е. случай существенно ограниченных функций на полупрямой. Это так называемая проблема устойчивости системы типа «черный ящик» с ограниченным входом и ограниченным выходом, представляющая очевидный интерес для практики. В дальнейшем выразим результаты исследования этой проблемы в виде зависимости от свойств операторов G и H (последние могут быть нелинейными).

ВНУТРЕННЕЕ ОПИСАНИЕ

Самым общим средством математического описания динамического процесса является дифференциальное уравнение вида

$$\dot{x} = f(x, t), \quad x(0) = c, \quad (5.3)$$

так называемое внутреннее описание. На таком описании основаны все классические результаты Ляпунова, Пуанкаре и других исследователей, причем динамика системы $f(\cdot, \cdot)$ задавалась в разных формах.

Исторически первое систематическое исследование свойств устойчивости уравнения (5.3) принадлежит Ляпунову, который рассмотрел следующую задачу: если начало координат является точкой равновесия системы (5.3), т. е. $f(0, t) = 0$ для всех t , и если система выведена «малым» возмущением из равновесия ($c \neq 0$), останутся ли траектории процесса «близкими» к началу координат для всех последующих моментов времени. В геометрической форме эта ситуация показана на рис. 5.1. Основная идея состоит в том, что если решение начинается в пределах небольших расстояний от начала координат, оно должно оставаться внутри несколько более широкого «канала», показанного на рис. 5.1 пунктиром.

Несколько более сильное определение устойчивости отвечало бы требованию $x(t) \rightarrow 0$ при $t \rightarrow \infty$, т. е. чтобы система в конечном счете возвращалась к точке равновесия. Такое определение отвечает понятию *асимптотическая устойчивость*

(согласно Ляпунову). Важно отметить, что понятия устойчивости по Ляпунову и асимптотической устойчивости являются независимыми. Легко построить примеры, для которых один из видов устойчивости имеет место, а другой нет, и наоборот. Так, система $\dot{r} = [g(\theta, t)/g(\theta, t)]r$, $\dot{\theta} = 0$, где $g(\theta, t) = \sin^2 \theta / [\sin^4 \theta + (1 - t \sin^2 \theta)^2] + 1/(1 + t^2)$, асимптотически устойчива, но ее решения становятся неограниченными, когда начальное состояние $\theta_0 = \theta(0) \rightarrow \pm \pi$ (возникает вопрос: в чем причина такого явления?).

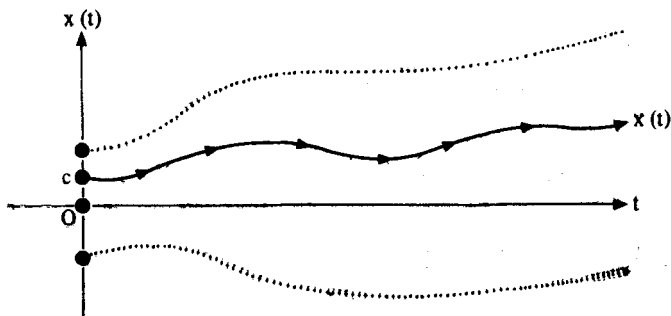


Рис. 5.1. Устойчивость по Ляпунову.

Заметим, кстати, что приведенные выше стандартные определения устойчивости подразумевают, что заранее известная точка равновесия вместе с ее ближайшей окрестностью представляет собой особенность типа центр. Как будет показано, требованию устойчивости соответствуют условия, налагаемые на свойства функции f в этой ближайшей окрестности, которые дают центр. Таким образом, с точки зрения практики, прежде чем применять какие-либо классические результаты, необходимо рассчитать все точки равновесия для f . Такой предварительный расчет может представить самостоятельную проблему в зависимости от структуры функции f . Здесь неявно предполагается, что положения равновесия f достигаются только в фиксированных точках. В общем случае они могут отвечать гораздо более сложным объектам — предельным циклам, странным аттракторам и т. п.

С каждой устойчивой точкой равновесия связана окружающая ее открытая область, называемая *областью притяжения*: устойчивая точка равновесия действует как некоторый «магнит», втягивающий любое начальное состояние внутри своей области притяжения (см. рис. 2.14).

Существенная часть современной теории устойчивости опирается на сведения о том, каким образом изменения гра-

ниц областей устойчивости и точек притяжения связаны с изменениями различных динамических параметров системы. Кроме того, в практическом отношении весьма важно иметь возможность математического описания границ данного равновесного состояния. Некоторые сведения по этим вопросам будут приведены ниже.

СТРУКТУРНАЯ УСТОЙЧИВОСТЬ

Главная особенность классических понятий устойчивости состоит в том, что они относятся к *конкретной системе* и поведению ее траекторий в окрестности точки равновесия (притяжения или отталкивания). Совершенно другого подхода требует анализ поведения *семейства траекторий*, возникающих при рассмотрении всех систем, «близких» к стандартной системе (5.3). Будем называть систему (5.3) *структурно устойчивой*, если топологический характер траекторий всех близких к ней систем такой же, как у системы (5.3). Простой пример структурной устойчивости — гармонический осциллятор с затуханием — был рассмотрен в гл. 2.

Определенные математические трудности связаны с уточнением понятия «близкая система», а также с конкретизацией смысла, который подразумевается, когда говорят о том, что траектория эквивалентна или топологически подобна другой траектории.

К концепции структурной устойчивости близка теория бифуркаций, а также ее современная популярная разновидность — теория катастроф. При анализе бифуркаций обычно предполагается, что динамика системы зависит от нескольких параметров, т. е. $\dot{f} = f(x, t, a)$, где a — вектор параметров, и исследуется характер положений равновесия при изменении параметров. Например, система

$$\begin{aligned}\dot{r} &= r(a - r^2), \\ \dot{\theta} &= 1,\end{aligned}$$

где $x_1^2 + x_2^2 = r^2$, $\theta = \text{arctg } x_2/x_1$, имеет одно положение равновесия $r = 0$ при $a < 0$. Оно соответствует устойчивому фокусу (рис. 5.2). Однако при $a > 0$ положение равновесия становится неустойчивым фокусом и возникает новое положение равновесия $r = \sqrt{a}$. Последнее представляет собой устойчивый предельный цикл, радиус которого растет как \sqrt{a} . Точка $a = 0$ является так называемой *точкой бифуркации Хопфа*. (Отметим рождение центра из устойчивого фокуса, когда параметр a проходит через критическое значение $a = 0$.)

Теория катастроф рассматривает вопрос об условиях, при которых изменение параметров системы вызывает перемещение данной точки в фазовом пространстве из области притяжения к заданному положению равновесия в область притяжения к другому положению равновесия. Простейшим является случай, когда все положения равновесия системы представляют собой фиксированные точки, которые можно

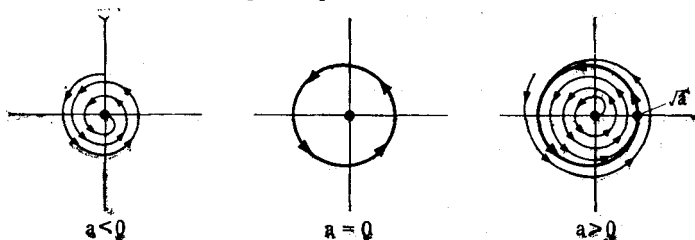


Рис. 5.2. Точка бифуркации Хопфа.

получить при помощи потенциальной функции. Это так называемая *элементарная теория*. Более сложные положения равновесия типа периодических замкнутых траекторий или аттракторов Лоренца требуют проведения исследований, выходящих за рамки данной книги. «Катастрофы» возникают при таких значениях параметра, которые приводят к сдвигу системы из области одного аттрактора в область другого аттрактора. Подробнее рассмотрим эти вопросы в одном из следующих разделов, где уточним также характер связи между теорией катастроф, анализом бифуркаций и структурной устойчивостью.

СВЯЗНАЯ УСТОЙЧИВОСТЬ И АДАПТИРУЕМОСТЬ

Ниже описывается представляющее интерес комбинированное понятие устойчивости, сочетающее классические идеи Ляпунова с комбинаторно-топологическим подходом, — понятие *связной устойчивости*. При этом рассматривается вопрос о том, останется ли состояние равновесия данной системы устойчивым (в смысле Ляпунова) вне зависимости от двойных связей между состояниями системы. Другими словами, берем в качестве исходной систему (5.3) и определяем *матрицу взаимосвязи* $E = [e_{ij}]$, такую, что

$$e_{ij} = \begin{cases} 1, & \text{если переменная } x_j \text{ оказывает влияние на пере-} \\ & \text{менную } x_i, \\ 0 & \text{в противном случае; } i, j = 1, 2, \dots, n. \end{cases}$$

Состояние равновесия $x = 0$ считается связно устойчивым, если оно устойчиво по Ляпунову для всех возможных матриц взаимосвязи E .

Понятие связной устойчивости представляет и практический интерес, поскольку при описании многих процессов наличие или отсутствие данной связи не всегда будет очевидным вследствие нарушений работы аппаратуры, неопределенности модели, случайных возмущений и т. п. Подобные ситуации особенно характерны для моделей в таких областях, как экономика, биология и энергетика.

Один из аспектов устойчивости, который привлек значительное внимание, особенно со стороны экологов, связан с понятием *адаптируемости*. Вообще говоря, адаптируемость, по-видимому, можно представить себе как определенную меру способности системы к поглощению внешних возмущений без резко выраженных последствий для ее поведения в переходном или установившемся состоянии. Внешне это представление выглядит весьма сходным с представлением о структурной устойчивости, и они действительно в значительной мере взаимно перекрываются. Однако, как и полагают, в действительности принцип адаптируемости несколько шире, поскольку надлежащим образом выбранная мера адаптируемости должна каким-то образом комбинировать возмущения собственно динамики системы с возмущениями, испытываемыми траекторией данного процесса. К сожалению, эта теория находится пока в стадии формирования, и мы располагаем лишь предварительными определениями и результатами.

ГРАФЫ

И ПРОЦЕССЫ РАСПРОСТРАНЕНИЯ ВОЗМУЩЕНИЙ

Как было показано, многие интересные системы можно успешно моделировать при помощи графов или, в более общем случае, путем использования симплициальных комплексов. Такое представление сложного процесса оказывается особенно удобным в тех случаях, когда мы не располагаем точными числовыми соотношениями между компонентами системы, необходимыми для внутреннего описания. Возникает вопрос, каким образом в эти рамки укладываются соображения об устойчивости.

Для того чтобы получить общее представление по этому вопросу, рассмотрим процесс, описываемый знаковым орграфом G . Здесь $\{u_1, u_2, \dots, u_N\}$ — вершины орграфа; предполагается, что каждой дуге G приписан знак плюс или минус, указывающий соответственно на положительную или

отрицательную связь между вершинами графа, соединяемыми данной дугой. Знаковый оргграф такого типа может быть использован для описания спроса на электроэнергию (рис. 5.3). Например, рост населения приводит к увеличению потребления энергии, поэтому дуге, соединяющей вершины, соответствующие населению и потреблению энергии, приписан знак плюс. Увеличение потребления энергии вызывает ухудшение

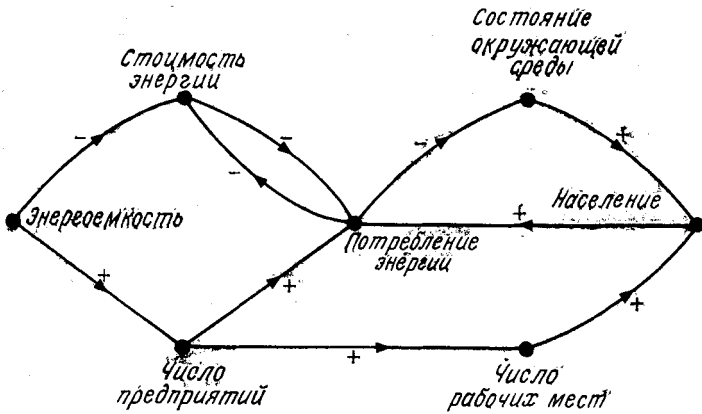
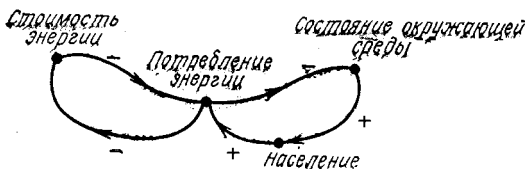


Рис. 5.3. Знаковый оргграф спроса на электроэнергию.

состояния окружающей среды, что отражено знаком минус у соответствующей дуги, и т. д. Аналогичные графы оказались полезными при исследовании различных проблем, касающихся городского транспорта, людских ресурсов во флоте, медицинского обслуживания, загрязнения атмосферы, средств обслуживания прибрежной зоны отдыха.

Заметим, что циклы знакового оргграфа соответствуют контурам обратной связи: циклы, характеризующие усиление тенденции к отклонению от данного состояния, отвечают контурам *положительной обратной связи*, а циклы, характеризующие подавление этой тенденции, отвечают контурам *отрицательной обратной связи*. Например, цикл



принадлежит ко второму из указанных типов, поскольку повышение стоимости энергии снижает ее потребление, что в

свою очередь улучшает состояние окружающей среды, приводящее в результате к росту населения, которое потребляет больше энергии, что тем самым приводит к снижению ее стоимости. В общем случае можно сформулировать следующее правило.

Цикл является контуром положительной обратной связи, если он содержит четное число дуг со знаком минус. В противном случае он является контуром отрицательной обратной связи.

Знаковый оргграф служит мощным средством исследования многих проблем. Однако он включает ряд упрощений, наиболее важное из которых состоит в пренебрежении тем обстоятельством, что одни переменные оказывают друг на друга более сильное влияние, чем другие. Иными словами, нам необходимо не только приписывать каждой дуге знак плюс или минус, но и каким-то образом указывать интенсивность связи. Таким образом, приходим к понятию *взвешенного* оргграфа как к частному случаю взвешенной связи, введенной при изучении симплициальных комплексов. Дальнейшее обобщение достигается в случае, когда каждой вершине графа приписывается численное значение и вводится интенсивность связи между двумя вершинами u_i и u_j как функция $f(u_i, u_j)$. Если теперь допустить, что численное значение, приписываемое каждой вершине, является переменным во времени, то приходим к понятию *процесса распространения возмущения* по графу G .

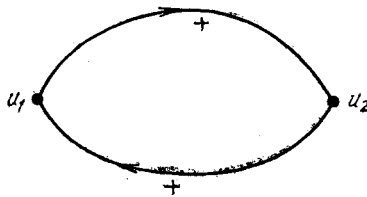
Обозначим значение в вершине u_i в момент времени t через $v_i(t)$, $i = 1, 2, \dots, N$; $t = 0, 1, \dots$. Предположим, что значение $v_i(t+1)$ зависит от $v_i(t)$ и от вершин, смежных с u_i . Таким образом, если вершина u_j смежна с u_i и если $p_j(t)$ представляет изменение в u_j в момент времени t , то следует принять, что влияние этого изменения на u_i в момент $t+1$ будет описываться функцией $\pm p_j(t)$ в зависимости от знака дуги, соединяющей u_i и u_j . В более общем случае для взвешенного оргграфа имеем следующее правило:

$$v_i(t+1) = v_i(t) + \sum_{j=1}^N f(u_j, u_i) p_j(t), \quad (5.4)$$

где через $f(u_i, u_j)$ обозначена весовая функция связи между вершинами u_i и u_j . *Процесс распространения возмущения* по оргграфу G определяется правилом (5.4) наряду с вектором начальных значений в вершинах $v(0)$ и вектором внешних возмущений $p(0)$ в каждой вершине в момент времени $t=0$. Практический интерес представляют так называемые *простые* процессы распространения возмущений, для которых $p(0)$ имеет лишь один ненулевой вход.

В связи с процессами распространения возмущений возникает много интересных вопросов. Однако в данной главе в основном исследуется устойчивость по значению и по возмущению системы по мере ее эволюции. Точнее, будем говорить, что вершина u_j устойчива по значению, если последовательность $\{|v_j(t)| : t = 0, 1, \dots\}$ ограничена. Аналогичным образом, назовем вершину u_j устойчивой по возмущению, если ограничена последовательность $\{|p_j(t)| : t = 0, 1, \dots\}$. Взвешенный орграф устойчив по возмущению (по значению), если каждая его вершина обладает этим свойством.

Нетрудно заметить, что имеет место явная аналогия между только что введенными понятиями устойчивости и представлениями об устойчивости типа «ограниченный вход — ограниченный выход», которые рассматривались для систем, заданных внешним описанием, несмотря на принципиальное и существенное различие в методах описания системы в каждом из этих случаев. В заключение данного раздела следует



отметить, что на примере простого графа видно, что устойчивость по возмущению не означает наличия устойчивости по значению, хотя обратное справедливо.

УСТОЙЧИВОСТЬ СИСТЕМЫ «ЧЕРНЫЙ ЯЩИК» С ОБРАТНОЙ СВЯЗЬЮ

Рассмотрим более подробно вопрос, поставленный выше, и проанализируем внешнее описание системы, обсуждавшееся в разделе о внешних описаниях, используя несколько математических понятий¹⁾.

Определим сначала условия, которые необходимо наложить на операторы G и H , чтобы гарантировать получение ограниченных выходов системы при ограниченных входах. Следует отметить, что для успешного решения проблем устойчивости нелинейных систем с обратной связью пока существуют только два общих подхода: теоремы о малом коэффициенте усиления и условия инертности. Чтобы пока-

¹⁾ Для читателей, не знакомых с понятиями банахова и гильбертова пространства, вполне достаточно представлять себе X как евклидово пространство R^n .

зять, каких результатов можно ожидать при использовании этих подходов, приведем следующие теоремы.

Теорема о малом коэффициенте усиления

Пусть G и H — отображения расширения X банахова пространства X над $[0, \infty]$ в себя, а $x_T(\cdot)$ обозначает сужение функции $x \in X$ на $[0, T]$. Тогда система с обратной связью

$$\begin{aligned} e_1 &= u_1 - He_2, \\ e_2 &= u_2 + Ge_1 \end{aligned} \quad (5.5)$$

устойчива, если существуют константы k_1, k_2, m_1, m_2 , такие, что

$$\begin{aligned} \|(Gx)_T\| &\leq k_1 \|x_T\| + m_1, \\ \|(Hx)_T\| &\leq k_2 \|x_T\| + m_2, \end{aligned}$$

а $k_1 k_2 < 1$ (здесь $\|\cdot\|$ обозначает норму в X).

Физическая интерпретация этой теоремы очень проста: пусть G и H соответствуют устойчивым подсистемам, тогда если G или H достаточно малы относительно предела устойчивости другой системы, то система с обратной связью (5.5) также устойчива. По существу, теорема дает точные количественные границы вместо качественного предположения «достаточной малости», и система в целом будет устойчива, если произведение коэффициентов усиления подсистем меньше единицы.

Преимущество теоремы о малом коэффициенте усиления состоит в том, что ее легко применять на практике, так как оценка коэффициентов усиления k_1 и k_2 обычно не вызывает затруднений. Кроме того, если условие $k_1 k_2 < 1$ не выполнено, то обычно определяется вид «компенсации», необходимой для его выполнения. Заметим далее, что тип устойчивости, гарантируемый теоремой о малых коэффициентах усиления, зависит от конкретного банахова пространства X . Так, если $X = L_2[0, \infty]$, то удовлетворение условий теоремы гарантирует L_2 -устойчивость. В частности, случай ограниченный вход — ограниченный выход охватывается выбором $X = L_\infty[0, \infty]$.

Самый простой пример применения теоремы о малом коэффициенте усиления соответствует случаю, когда G и H — линейные инвариантные по времени операторы, т. е.

$$\begin{aligned} (Gx)_T &= \int_0^T g(T-s)x(s) ds, \\ (Hx)_T &= \int_0^T h(T-s)x(s) ds, \end{aligned}$$

Тогда можно легко показать, что если X — пространство непрерывных функций на $[0, T]$, то условие теоремы будет выполнено, если

$$\left(\sup_{0 \leq r \leq T} |g(r)| \right) \left(\sup_{0 \leq r \leq T} |h(r)| \right) < 1.$$

Заметим, что если G — линейный инвариантный по времени оператор, а H — оператор, не обладающий памятью, то теорема о малом коэффициенте усиления сводится к критерию Попова.

Будем считать теперь X гильбертовым пространством, т. е. банаховым пространством, в котором норма $\|\cdot\|$ определяется скалярным произведением. В этом случае можно получить другой тип теоремы об устойчивости.

Теорема инертности

Пусть X — вещественное гильбертово пространство на $[0, \infty]$ со скалярным произведением $\langle \cdot, \cdot \rangle$. Тогда система (5.5) устойчива, если существуют такие константы $k, m_1, m_2, m_3, \delta, \varepsilon$, что

$$\begin{aligned} \langle (Gx)_T, (Gx)_T \rangle &\leq k \langle x_T, x_T \rangle + m_1, \\ \langle x_T, (Gx)_T \rangle &\geq \delta \langle x_T, x_T \rangle + m_2, \\ \langle x_T, (Hx)_T \rangle &\geq \varepsilon \langle (Hx)_T, (Hx)_T \rangle + m_3, \\ \delta + \varepsilon &> 0. \end{aligned}$$

В терминах теории электрических цепей физический смысл теоремы инертности таков: пусть u_1 и e_1 — функции напряжения, G — оператор проводимости, u_2, e_2 — функции токов, а H — оператор импеданса, тогда вышеприведенные неравенства означают, что если G имеет проводимость по крайней мере δ , а H имеет проводимость по крайней мере ε , то система устойчива, если суммарный уровень проводимости G и H положителен.

Заметим, что как теорема о малом коэффициенте усиления, так и теорема инертности дают только *достаточные* условия устойчивости нелинейной системы с обратной связью. Много работ посвящено также методам получения критериев неустойчивости. Однако изложение основных результатов этих работ потребовало бы слишком высокого уровня формализации по сравнению с принятым в этой книге, поэтому мы отсылаем заинтересованного читателя к ссылкам в конце главы.

ВНУТРЕННИЕ МОДЕЛИ И УСТОЙЧИВОСТЬ

В противоположность метафизическому подходу в основе математического подхода к анализу устойчивости лежит описание системы с помощью дифференциальных уравнений, причем цель анализа состоит в том, чтобы определить, будет ли данное положение равновесия системы устойчивым по отношению к возмущениям начальных условий. Различные проблемы, связанные с классической механикой и устойчивостью планетных орбит (знаменитая «задача трех тел»), поставили ряд задач, которые были в окончательном виде сформулированы на рубеже нашего столетия в работах Ляпунова, Пуанкаре и других исследователей. Основной вопрос теории устойчивости по Ляпунову состоит в том, чтобы выяснить, вернется ли система в данное положение равновесия спустя произвольно продолжительное время после возникновения начального возмущения. Мы представим ниже два наиболее важных результата, касающихся этого вопроса.

Рассмотрим сначала линейный случай, когда внутренняя модель системы описывается системой дифференциальных уравнений

$$\dot{x} = Fx, \quad x(0) = x_0 (\neq 0). \quad (5.6)$$

Здесь F — постоянная матрица размером $n \times n$, при этом предполагается, что характеристический многочлен матрицы F известен и имеет вид

$$\psi_F(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n.$$

Исследуем асимптотическую устойчивость точки равновесия $x = 0$. Решение уравнения (5.6) имеет вид

$$x(t) = e^{Ft} x_0,$$

поэтому очевидно, что в результате произвольного ненулевого начального возмущения x_0 система вернется при $t \rightarrow \infty$ в положение равновесия в том и только в том случае, если все корни характеристического уравнения матрицы F имеют отрицательные действительные части. Поскольку корни последнего совпадают с корнями многочлена $\psi_F(z)$, задача сводится к тому, чтобы с учетом известных свойств коэффициентов этого многочлена установить критерий, с помощью которого можно было бы определять, лежат ли корни многочлена в левой полуплоскости. Соответствующий критерий был получен в конце прошлого века английскими математиками Раусом и Гурвицем и был сформулирован в виде следующей теоремы.

Теорема Рауса — Гурвица

Все корни многочлена $\psi_F(z)$ имеют отрицательные действительные части в том и только в том случае, если

- 1) все $a_i > 0$, $i = 0, 1, \dots, n$,
- 2) матрица размером $n \times n$

$$A = \begin{bmatrix} a_1 & a_0 & 0 & 0 & \dots & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 & \dots & 0 & 0 \\ a_5 & a_4 & a_3 & a_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} & a_{n-2} \\ 0 & 0 & 0 & 0 & \dots & 0 & a_n \end{bmatrix}$$

имеет только положительные главные миноры.

При помощи теоремы Рауса — Гурвица сравнительно простые алгебраические выкладки дают возможность проверить устойчивость равновесия в начале координат для линейной системы, если известен характеристический многочлен матрицы F . Например, для гармонического осциллятора, описываемого системой второго порядка

$$\begin{aligned} \ddot{x} + b\dot{x} + cx &= 0, \\ x(0) &= c_1, \quad \dot{x}(0) = c_2, \end{aligned}$$

матрица, фигурирующая в условии устойчивости, имеет вид

$$A = \begin{bmatrix} b & 1 \\ 0 & c \end{bmatrix}.$$

Таким образом, применяя критерий Рауса — Гурвица, видим, что начальное возмущение будет «затухающим» в том и только в том случае, если

- 1) $b > 0$, $c > 0$,
- 2) $bc > 0$,

т. е. только тогда, когда коэффициент «демпфирования» b дает эффект положительного затухания.

К сожалению, требование, согласно которому многочлен $\psi_F(z)$ должен быть известен, существенно ограничивает практическое применение критерия Рауса — Гурвица, особенно для систем высокой размерности. В подобных случаях было бы желательно иметь в распоряжении такой критерий устойчивости, который мог бы применяться непосредственно к элементам матрицы F . Подобный критерий был разработан Ля-

пуновым на основании следующего простого физического представления о положении равновесия: равновесное состояние системы асимптотически устойчиво, если все траектории процесса, начинающиеся достаточно близко от точки равновесия, идут таким образом, что приводят к минимуму подходящим образом определенную «энергетическую» функцию, причем положение локального минимума энергии соответствует самой точке равновесия.

Рассмотрим сначала использование этого критерия применительно к общему нелинейному уравнению

$$\dot{x} = f(x), \quad x(0) = x_0, \quad (5.7)$$

а затем применительно к линейному уравнению, когда $f(x) = Fx$.

Предположим, что $f(0) = 0$ и что функция f непрерывна в окрестности начала координат.

Математическое описание энергетической функции содержится в следующем определении.

Определение 5.1

Функция $V(x)$ называется *функцией Ляпунова* (энергетической функцией) системы (5.7), если

- 1) $V(0) = 0$,
- 2) $V(x) > 0$ для всех $x \neq 0$ в окрестности начала координат,
- 3) $dV(x)/dt < 0$ вдоль траекторий системы (5.7).

Основной результат, полученный Ляпуновым, был им сформулирован в виде теоремы об устойчивости.

Теорема Ляпунова об устойчивости

Положение равновесия $x = 0$ системы (5.7) асимптотически устойчиво в том и только в том случае, если существует функция Ляпунова $V(x)$ системы.

С целью применить эту теорему к линейной системе (5.6) выберем в качестве кандидата на функцию Ляпунова следующую функцию:

$$V(x) = (x, Px),$$

где P — пока неизвестная симметрическая матрица. Для того чтобы $V(x)$ была функцией Ляпунова системы, мы должны иметь

$$\begin{aligned} \frac{dV}{dt} &= (\dot{x}, Px) + (x, P\dot{x}) = \\ &= (x, (F'P + PF)x) < 0. \end{aligned}$$

Здесь подразумевается, что уравнение

$$F'P + PF = -C$$

разрешимо при любой матрице $C > 0$.

Далее; из условий 1 и 2 следует, что матрица P должна быть положительно определенной. Следовательно, в результате мы получаем, что положение равновесия системы (5.6) в начале координат асимптотически устойчиво в том и только в том случае, если уравнение $F'P + PF = -C$ имеет решение $P > 0$ при любой матрице $C > 0$.

Следует, однако, иметь в виду, что выбранная для представления $V(x)$ квадратичная форма не является единственным кандидатом на функцию Ляпунова линейной системы (5.6). В качестве примера рассмотрим задачу из области экономики, связанную с моделированием n взаимосвязанных рынков сбыта n товаров (или услуг), поступающих из одной или нескольких связанных между собой отраслей промышленности (или систем обслуживания). Обозначив через $x(t)$ вектор цен товаров в момент времени t , получим классическую модель этой ситуации

$$\dot{x}(t) = Ax(t),$$

где $A = [a_{ij}]$ — постоянная матрица размером $n \times n$. Если все товары взаимозаменяемы, то A — матрица Метцлера, т. е. a_{ij} удовлетворяют условиям

$$a_{ij} < 0, \quad i = j,$$

$$a_{ij} \geq 0, \quad i \neq j.$$

Вопрос об устойчивости цен в такой ситуации был рассмотрен в 1945 г. Метцлером, получившим следующий классический результат:

Система Метцлера $\dot{x} = Ax$ устойчива в том и только в том случае, если главные миноры матрицы A удовлетворяют условию

$$\{-1\}^k \det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{12} & a_{22} & \dots & a_{2k} \\ \vdots & & \ddots & \\ \vdots & & & \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{bmatrix} > 0$$

для всех $k = 1, 2, \dots, n$.

Доказательство этого утверждения основано на выборе функции Ляпунова

$$V(x) = \sum_{i=1}^n d_i |x_i|,$$

где $d_i > 0$ — подлежащие определению постоянные. Соответствующий выбор d_i и использование свойств матрицы A по Метцлеру дают возможность показать, что $V(x)$ действительно является функцией Ляпунова системы. Таким образом, на основании теоремы Ляпунова об устойчивости заключаем, что начало координат является точкой равновесия системы.

Теорему Ляпунова об устойчивости особенно легко применять для одного класса нелинейных задач, в которых нелинейные члены можно считать «малыми» возмущениями главной, линейной части. Например, естественно полагать, что если динамика системы описывается уравнением

$$\dot{x} = Fx + h(x), \quad x(0) = x_0, \quad (5.8)$$

где F — устойчивая матрица (т. е. матрица, собственные значения которой лежат в левой полуплоскости), то положение равновесия $x = 0$ будет асимптотически устойчивым, если начальное возмущение x_0 и нелинейное возмущение $h(x)$ не слишком велики. Математическая формулировка этого очевидного результата может быть представлена в виде следующей теоремы.

Теорема Пуанкаре — Ляпунова

Пусть система (5.8) удовлетворяет следующим условиям:

- 1) F — устойчивая матрица,
- 2) $h(\cdot)$ — непрерывная функция переменной x , такая, что $h(0) = 0$ и $\|h(x)\|/\|x\| \rightarrow 0$ при $\|x\| \rightarrow 0$,
- 3) $\|x_0\| \ll 1$.

Тогда положение равновесия $x = 0$ асимптотически устойчиво.

Одна из трудностей, возникающих при использовании этой теоремы, связана с условием 3, которое по существу является требованием, чтобы начальное возмущение было «достаточно малым». Действительная степень малости определяется, вообще говоря, сравнительной величиной нелинейной части h и величиной действительной части корня характеристического уравнения F , ближайшего к мнимой оси.

Для того чтобы попытаться исключить условие 3 и получить достаточное условие глобальной устойчивости, необходимо наложить более строгие ограничения на динамику

системы. Соответствующий результат был получен советским математиком Н. Н. Красовским.

Теорема Н. Н. Красовского

Решение $x = 0$, соответствующее положению равновесия нелинейной системы $\dot{x} = f(x)$, асимптотически устойчиво в большом, если существует постоянная $\epsilon > 0$, такая, что матрица $J(x) + J'(x)$ имеет собственные значения, меньшие, чем $-\epsilon$, для всех x , где $J(x)$ — матрица Якоби функции f , т. е.

$$[J(x)]_{ij} = \frac{\partial f_i}{\partial x_j}.$$

Теорема Красовского является следствием теоремы Ляпунова об устойчивости, если использовать функцию $V(x) = (x, (J(x) + J'(x))x)$.

В качестве примера применения теоремы Ляпунова об устойчивости рассмотрим электрический колебательный контур с параметрическим возбуждением. Динамика такого процесса описывается уравнением

$$\ddot{x} + a\dot{x} + b(t)x = 0,$$

где $a > 0$, $b(t) = b_0(1 + f(t))$, $b_0 \geq 0$ и $f(t)$ — ограниченная функция. Здесь x — напряжение на контуре, a — сопротивление, $b(t)$ — емкость, изменяющаяся с течением времени. Приведенное выше уравнение эквивалентно системе

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = -b(t)x_1 - ax_2.$$

Исследуем устойчивость положения равновесия $x_1 = x_2 = 0$. Для этого рассмотрим энергетическую функцию

$$V(x_1, x_2) = \frac{1}{2} \left(x_2 + \frac{ax_1}{2} \right)^2 + \left(\frac{a^2}{4} + b_0 \right) \frac{x_1^2}{2}.$$

Легко показать, что

$$1) V(x_1, x_2) \geq 0,$$

$$2) V(x_1, x_2) = 0 \text{ в том и только в том случае, если } x_1 = x_2 = 0,$$

$$3) dV/dt = (-a/2)x_2^2 - (b - b_0)x_1x_2 - (ab/2)x_1^2.$$

Следовательно, положение равновесия системы в начале координат будет асимптотически устойчивым, если $dV/dt < 0$ в некоторой окрестности начала координат. Это действительно будет иметь место, если для некоторых $\alpha > 0$

$$\epsilon^2 b_0 [f(t)]^2 - a^2 (1 + \epsilon f(t)) \leq -\alpha < 0.$$

Последнее неравенство выполняется, если ϵ достаточно мало,

Итак, для достаточно малых ϵ положение равновесия системы в начале координат асимптотически устойчиво. Этот результат можно интерпретировать следующим образом. В системе действуют две противоположные силы: параметрическое возбуждение, пропорциональное ϵ , и нагрузка, демпфирующая сила $a\dot{x}$. Для того чтобы удовлетворить приведенному выше неравенству, необходимо выбрать сопротивление достаточно большим, с тем чтобы нагрузка поглотила всю энергию, выделенную за счет возбуждения. В этом случае положение равновесия в начале координат устойчиво. Если же нагрузка недостаточно велика, можно ожидать увеличения энергии, и равновесие в начале координат становится неустойчивым.

Для более подробного ознакомления с рядом интересных аспектов проблемы устойчивости рекомендуем читателю обратиться к литературе, указанной в конце данной главы. Перейдем теперь к обсуждению некоторых современных представлений об устойчивости, которые оказываются особенно полезными при изучении систем.

СВЯЗНАЯ УСТОЙЧИВОСТЬ

При решении практических задач, как правило, трудно с полной определенностью выделить внутренние связи в системе. Это обусловлено тем, что наличие или отсутствие прямых связей одной подсистемы с другой часто не может быть установлено вовсе или может быть установлено лишь с низкой точностью. Одним из возможных подходов к исследованию таких ситуаций служит предположение, что связи являются случайными переменными, описываемыми некоторыми известными функциями распределения. Для ответа на различные вопросы вероятностного характера относительно динамического поведения системы можно затем применить статистические методы. Однако ниже для исследования характеристик устойчивости системы, внутренние связи которой не установлены с большой точностью, мы собираемся использовать другой подход, заранее не опирающийся на статистическую обработку. Как было отмечено в вводном разделе, такой подход соответствует понятию *связной* устойчивости.

Рассмотрим динамический процесс, внутреннее описание которого

$$\dot{x} = A(x, t)x, \quad x(0) = x_0, \quad (5.9)$$

где x — вектор состояния системы, A — непрерывная матричная функция своих аргументов при всех $t \geq 0$ и всех $x \in R^n$.

Для изучения вопросов связности в данной задаче запишем элементы матрицы A в виде

$$a_{ij}(x, t) = -\delta_{ij}\psi_i(x, t) + e_{ij}\psi_{ij}(x, t),$$

где δ_{ij} — символ Кронекера (т. е. $\delta_{ij} = 1$, если $i = j$, и $\delta_{ij} = 0$ при $i \neq j$), ψ_i и ψ_{ij} — непрерывные функции своих аргументов. Элементы e_{ij} представляют собой компоненты матрицы взаимосвязи E системы и удовлетворяют соотношениям

$$e_{ij} = \begin{cases} 1, & \text{если переменная } x_j \text{ оказывает влияние на } \dot{x}_i, \\ 0 & \text{в противном случае.} \end{cases}$$

Понятие связной устойчивости дается следующим определением.

Определение 5.2

Состояние равновесия $x = 0$ системы (5.9) *связно асимптотически устойчиво в большом* тогда и только тогда, когда оно асимптотически устойчиво для всех матриц взаимосвязи E .

Для того чтобы получить практический метод определения наличия связной устойчивости, потребуем выполнения дополнительных условий для функций ψ_i и ψ_{ij} . Предположим, что существуют постоянные $\alpha_i > 0$, $\alpha_{ij} \geq 0$, такие, что при всех $x \in \mathbb{R}^n$ и всех $t \geq 0$

$$1) \psi_i(x, t) \leq \alpha_i,$$

$$2) |\psi_{ij}(x, t)x_j| \leq \alpha_{ij}|x_j|; \quad i, j = 1, 2, \dots, n.$$

Введем матрицу $A = [a_{ij}]$ следующим образом:

$$a_{ij} = -\delta_{ij}\alpha_i + e_{ij}\alpha_{ij}, \quad i, j = 1, 2, \dots, n.$$

Тогда основной результат может быть сформулирован следующим образом.

Теорема о связной устойчивости

Состояние равновесия системы (5.9) в точке $x = 0$ *связно асимптотически устойчиво в большом* тогда и только тогда, когда матрица A удовлетворяет условию

$$(-1)^k \det \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1k} \\ \bar{a}_{21} & \bar{a}_{22} & \dots & \bar{a}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{k1} & \bar{a}_{k2} & \dots & \bar{a}_{kk} \end{bmatrix} > 0, \quad k = 1, 2, \dots, n.$$

Замечания. 1. Условие, налагаемое на главные миноры матрицы A , в литературе по теории устойчивости принято называть условием Севастьянова — Котелянского. Экономистам матрица A известна как матрица Хикса.

2. Если ограничения на ψ_i и ψ_{ij} выполняются не для всех $x \in R^n$, а лишь в некоторой области $M \subset R^n$, то указанный выше результат имеет место в области M .

Для того чтобы исследовать размеры области связной устойчивости, определим множество чисел $\{d_i\}$ следующим образом:

$$|a_{jj}| - d_j^{-1} \sum_{i \neq j} d_i |a_{ij}| \geq \varepsilon > 0.$$

Пусть числа $\{u_i\}$ таковы, что

$$M \supset \{x \in R^n : |x_i| < u_i, \quad i = 1, 2, \dots, n\},$$

т. е. u_i определяют гиперкуб в R^n , содержащийся в M . Используя введенные выше величины, можно показать, что область

$$\left\{ x \in R^n : \sum_{i=1}^n d_i |x_i| < \min_i d_i u_i \right\}$$

является областью связной асимптотической устойчивости для системы (5.9).

Таким образом видно, что нахождение наиболее широкой области связной устойчивости, грубо говоря, эквивалентно определению чисел d_i , удовлетворяющих выписанному выше неравенству, причем наименьшее из d_i должно быть настолько большим, насколько возможно.

Для того чтобы продемонстрировать применение теоремы связной устойчивости, рассмотрим следующую модель системы хищник — жертва, включающую четыре вида. Уравнения динамики системы таковы:

$$\dot{x}_1 = a_1 x_1 + b_1 x_1 x_2 - D_1(x_1) + D_3(x_3),$$

$$\dot{x}_2 = a_2 x_2 + b_2 x_2 x_1 - D_2(x_2) + D_4(x_4),$$

$$\dot{x}_3 = a_3 x_3 + b_3 x_3 x_4 - D_3(x_3) + D_1(x_1),$$

$$\dot{x}_4 = a_4 x_4 + b_4 x_4 x_3 - D_4(x_4) + D_2(x_2),$$

где переменная $x_i(t)$ представляет популяцию i -го вида, $D_i(x_i)$ — скорость расселения популяции i -го вида, a_k и b_k — постоянные. Физически вполне оправданно принять функциональные зависимости для $D_i(\cdot)$ следующими:

$$D_i(x_i) = x_i f_i(x_i),$$

что мы и будем предполагать в дальнейшем. Наша цель состоит в определении условий, налагаемых на постоянные a_i , b_i и функции $f_i(x_i)$, при которых обеспечивается связанная устойчивость положения равновесия системы в начале координат.

После введения предположения о структуре $D_i(x_i)$ уравнения динамики системы принимают форму (5.9), в которой

$$A(x, t) =$$

$$\begin{bmatrix} a_1 + b_1 x_2 - f_1(x_1) & 0 & f_3(x_3) & 0 \\ 0 & a_2 + b_2 x_1 - f_2(x_2) & 0 & f_4(x_4) \\ f_1(x_1) & 0 & a_3 + b_3 x_4 - f_3(x_3) & 0 \\ 0 & f_2(x_2) & 0 & a_4 + b_4 x_3 - f_4(x_4) \end{bmatrix}$$

Следовательно, матрица взаимосвязи в данной задаче такова:

$$E = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

а функции ψ_i и ψ_{ij} имеют вид

$$\begin{aligned} \psi_1 &= -(a_1 + b_1 x_2 - f_1(x_1)), & \psi_2 &= -(a_2 + b_2 x_1 - f_2(x_2)), \\ \psi_3 &= -(a_3 + b_3 x_4 - f_3(x_3)), & \psi_4 &= -(a_4 + b_4 x_3 - f_4(x_4)), \\ \psi_{13} &= f_3(x_3), & \psi_{24} &= f_4(x_4), \\ \psi_{31} &= f_1(x_1), & \psi_{42} &= f_2(x_2). \end{aligned}$$

все остальные $\psi_{ij} = 0$.

Для того чтобы применить теорему, необходимо сначала найти постоянные α_i , $\alpha_{ij} \geq 0$, такие, что

$$\begin{aligned} a_1 + b_1 x_2 - f_1(x_1) &\leq -\alpha_1 < 0, \\ a_2 + b_2 x_1 - f_2(x_2) &\leq -\alpha_2 < 0, \\ a_3 + b_3 x_4 - f_3(x_3) &\leq -\alpha_3 < 0, \\ a_4 + b_4 x_3 - f_4(x_4) &\leq -\alpha_4 < 0, \\ |f_1(x_1)| &\leq \alpha_{31}, & |f_2(x_2)| &\leq \alpha_{42}, \\ |f_3(x_3)| &\leq \alpha_{13}, & |f_4(x_4)| &\leq \alpha_{24} \end{aligned}$$

Тогда условия Хикса в теореме о структурной устойчивости будут удовлетворены в том и только в том случае, если

- 1) $\alpha_1 > 0$,
- 2) $\alpha_1 \alpha_2 > 0$,
- 3) $\alpha_1 \alpha_3 > \alpha_{13} \alpha_{31}$,
- 4) $\alpha_1 \alpha_2 \alpha_3 \alpha_4 + \alpha_{13} \alpha_{31} \alpha_{24} \alpha_{42} > \alpha_1 \alpha_3 \alpha_{24} \alpha_{42} + \alpha_2 \alpha_4 \alpha_{13} \alpha_{31}$.

Таким образом, условия 1—4 определяют область в пространстве состояний (x_1, x_2, x_3, x_4) , для которой начало координат асимптотически устойчиво при любых возмущающих функциях f_i (т. е. при любых скоростях расселения) и любых a_i, b_i .

БИФУРКАЦИЯ ХОПФА

Теория связанной устойчивости дает критерии, которым должны удовлетворять параметры системы с тем, чтобы положение равновесия в начале координат было асимптотически устойчиво при всех взаимосвязях между подсистемами. Однако при изменении одного из параметров до критического значения теорема перестает быть справедливой и возникает естественный вопрос о характере преобразования исходного положения равновесия, соответствующего данному изменению параметра. В общем случае нас интересуют критические значения параметров системы, при которых точка равновесия качественно меняет свой характер (например, аттрактор переходит в центр, притяжение сменяется отталкиванием).

В случае этой простейшей разновидности проблемы «бифуркации» можно предположить, что изменяется только один параметр системы. Выше уже приводился пример проблемы такого типа, в котором кратко рассматривался принцип структурной устойчивости. Основной результат для целого класса таких задач был получен Хопфом.

Рассмотрим двумерный случай, когда динамика системы описывается уравнениями

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \mu), & x_1(0) &= x_1^0, \\ \dot{x}_2 &= f_2(x_1, x_2, \mu), & x_2(0) &= x_2^0. \end{aligned}$$

Несколько более сложным будет n -мерный случай, но основные результаты по существу остаются неизменными.

Основной результат, характеризующий изменения поведения устойчивости системы при изменении μ , заключен в бифуркационной теореме Хопфа.

Бифуркационная теорема Хопфа (в R^2)

Пусть функции f_1 и f_2 по крайней мере четыре раза дифференцируемы по каждому аргументу и $f(0, 0, \mu) =$

$= f_2(0, 0, \mu) = 0$ при всех действительных μ . Допустим далее, что матрица

$$J(f) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} (x_1, x_2) = (0, 0)$$

имеет два различных комплексно-сопряженных собственных значения $\lambda(\mu)$ и $\bar{\lambda}(\mu)$, таких, что $\operatorname{Re} \lambda(\mu) > 0$ при $\mu > 0$. Предположим также, что

$$\frac{d}{d\mu} [\operatorname{Re} \lambda(\mu)] \Big|_{\mu=0} > 0.$$

Тогда

1. Существует дважды дифференцируемая функция $\mu: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$, такая, что начальная точка $(x_1^0, 0, \mu(x_1^0))$ лежит на замкнутой траектории периода $2\pi/|\lambda(\mu)|$, радиус которой растет как $\sqrt{\mu}$ при $x_1^0 \neq 0$, $\mu(0) = 0$.

2. Существует окрестность U точки $(0, 0, 0)$ в \mathbb{R}^3 , такая, что любая замкнутая траектория, лежащая в U , — это одна из замкнутых траекторий указанного выше семейства.

3. Кроме того, если точка 0 — аттрактор при $\mu = 0$, то $\mu(x_1^0) > 0$ для всех $x_1^0 \neq 0$ и замкнутые траектории являются притягивающими.

Замечание. Замкнутой траекторией в \mathbb{R}^2 является любая точка x^* , такая, что $x(t) = x(t+T) = x^*$ для некоторого $T > 0$. В частности, точки равновесия — замкнутые траектории.

Согласно теореме Хопфа, траектории системы могут менять свое поведение при удалении μ от нуля, и характер этого изменения зависит от величины действительной части собственного значения матрицы Якоби $J(f)$ для данной системы в начале координат. Утверждение 1 в сущности означает, что равновесие в начале координат «дает бифуркацию» в замкнутую траекторию определенного радиуса, пропорционального $\sqrt{\mu}$. Если начало было притягивающей фиксированной точкой при $\mu = 0$, то, согласно утверждению 2, в достаточно малой окрестности начала (в пространстве x_1, x_2, μ) замкнутые траектории, возникающие из фиксированной точки, сами будут притягивающими.

Итак, бифуркационная теорема Хопфа касается рождения замкнутых траекторий из фиксированных точек равновесия, а также последующего динамического поведения системы,

когда параметр μ проходит через критическое значение $\mu=0$. Для рассмотренных здесь двумерных систем только замкнутые траектории могут представлять периодические решения. Для систем с большим числом измерений ситуация оказывается намного сложнее.

Пример. Уравнение Льенара

Уравнение Льенара представляет собой хорошо известную нелинейную систему дифференциальных уравнений, к которой часто приводят простые модели колебательных явлений, например в динамике популяций или при описании электрических контуров:

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1 + \mu x_2 - x_2^3.\end{aligned}$$

Исследуя поведение этой системы при изменении параметра μ от отрицательных до положительных значений, можно показать, что начало координат $x_1 = x_2 = 0$ является точкой равновесия системы при всех μ . Кроме того,

$$J(f) = \begin{bmatrix} 0 & 1 \\ 1 & \mu - 3x_2^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & \mu \end{bmatrix}$$

при $x_1 = x_2 = 0$. Собственные значения матрицы J таковы:

$$\frac{1}{2} [\mu \pm \sqrt{\mu^2 - 4}].$$

Рассмотрим значения μ , такие, что $|\mu| < 2$. В этом случае $\lambda(\mu) \neq 0$. Кроме того, $\operatorname{Re} \lambda(\mu) < 0$ при $-2 < \mu < 2$, $\operatorname{Re} \lambda(\mu) = 0$ при $\mu = 0$ и $\operatorname{Re} \lambda(\mu) > 0$ при $0 < \mu < 2$. Имеем также

$$\frac{d}{d\mu} \operatorname{Re} \lambda(\mu) \Big|_{\mu=0} = \frac{1}{2} > 0.$$

Все условия бифуркационной теоремы Хопфа выполнены, следовательно, существует однопараметрическое семейство замкнутых траекторий в окрестности начала координат.

Для того чтобы выяснить, будут ли эти замкнутые траектории устойчивыми и возникнут ли они при $\mu > 0$, необходимо использовать методы, выходящие за рамки этой книги, поскольку начало координат не является притягивающей фиксированной точкой; следовательно, мы не можем использовать вывод (3) теоремы. Оказывается, однако, что для этого уравнения периодические замкнутые траектории

действительно являются притягивающими и бифуркация имеет место при $\mu > 0$.

Заметим, что небольшое обобщение указанного примера позволяет рассмотреть общее уравнение Ван дер Поля

$$\ddot{u} + f(u, \mu) \dot{u} + g(u) = 0$$

путем замены переменных $x_1 = u$, $x_2 = \dot{u} + f(u, \mu)$. Тем самым мы приводим уравнение Ван дер Поля к общему уравнению Льенара

$$\begin{aligned} \dot{x}_1 &= x_2 - f(x, \mu), \\ \dot{x}_2 &= -g(x_1), \end{aligned}$$

которое также может быть изучено при помощи теоремы Хопфа. Отсюда следует вывод, что уравнение Ван дер Поля также приводит к устойчивым колебаниям при $\mu > 0$, возникающим в результате бифуркации от неподвижной точки в начале координат.

СТРУКТУРНО УСТОЙЧИВЫЕ ДИНАМИЧЕСКИЕ СИСТЕМЫ

Принципиальным этапом в исследовании модели динамического процесса, описываемого дифференциальным уравнением, является решение вопроса о том, могут ли малые возмущения динамики системы привести к качественно иному ее поведению. Мы уже упоминали об этом, как о проблеме *структурной устойчивости*. В данном разделе будут представлены некоторые элементарные результаты, касающиеся структурно устойчивых систем. Все математические модели, описывающие физические явления, содержат упрощения, погрешности и другие отклонения от действительности. Поэтому трудно переоценить важность роли структурной устойчивости как существенной части основы эффективного моделирования.

Рассмотрим два динамических процесса

$$\begin{aligned} \dot{x}_1 &= F_1(x_1, x_2), \\ \dot{x}_2 &= F_2(x_1, x_2), \quad \text{I} \\ \dot{x}_1 &= G_1(x_1, x_2), \\ \dot{x}_2 &= G_2(x_1, x_2), \quad \text{II} \end{aligned}$$

определенных в круге $D: x_1^2 + x_2^2 \leq 1$. Предположим далее, что векторы (F_1, F_2) и (G_1, G_2) не совпадают с касательной к границе области D и всегда направлены внутрь D . Опреде-

ление структурной устойчивости системы (I) было впервые дано Л. С. Понтрягиным и А. А. Андроновым в следующей форме.

Определение 5.3

Дифференциальное уравнение (I) называется структурно устойчивым, если существует $\delta > 0$, такое, что для любого дифференциального уравнения (II), удовлетворяющего условиям

$$|F_i - G_i| < \delta, \quad \left| \frac{\partial(F_i - G_i)}{\partial x_j} \right| < \delta, \quad i, j = 1, 2,$$

во всех точках D , существует гомеоморфизм (взаимно однозначное непрерывное отображение) $h: D \rightarrow D$, отображающий траектории системы (I) в траектории системы (II) и сохраняющий ориентацию этих траекторий.

Другими словами, система (I) структурно устойчива, если для достаточно близких систем (I) и (II) траектории (I) можно непрерывным образом деформировать в траектории (II) при сохранении направления потока.

Для получения практических условий проверки структурной устойчивости данного уравнения необходимо ввести несколько новых понятий.

Определение 5.4

Особой точкой дифференциального уравнения (I) является точка (x_1^*, x_2^*) , такая, что

$$F_1(x_1^*, x_2^*) = F_2(x_1^*, x_2^*) = 0.$$

Особенность называется *гиперболической*, если собственные значения матрицы

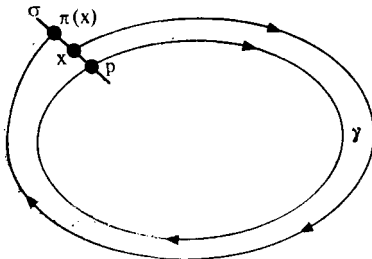
$$J = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} \end{bmatrix} (x_1^*, x_2^*)$$

имеют ненулевые действительные части. Гиперболическая особенность может быть стоком, седловой точкой или источником в зависимости от того, сколько собственных значений с отрицательными действительными частями имеет матрица J (2, 1 или 0).

В окрестности гиперболической особенности траектории системы могут иметь следующий вид:



Рассмотрим, наконец, замкнутую траекторию γ дифференциального уравнения (I). Через точку p на γ проведем малую линейную секущую σ , трансверсальную к траектории вдоль γ (рис. (5.4)). Следуя по траекториям (I), проходящим через точки x на σ , получаем отображение Пуанкаре



$$\pi: \sigma \rightarrow \sigma,$$

Рис. 5.4. Отображение Пуанкаре. определенное для всех $x \in \sigma$, достаточно близких к p .

Аналогично получаем определение гиперболической замкнутой траектории.

Определение 5.5

Замкнутая траектория γ называется *гиперболической*, если $|d\pi/dx|_{x=p} \neq 1$.

Отметим, что если $|d\pi/dx|_{x=p} < 1$, то γ — устойчивый предельный цикл и траектории имеют вид спиралей, наматывающихся на γ . Если $|d\pi/dx|_{x=p} > 1$, то γ — неустойчивый предельный цикл, а при $|d\pi/dx|_{x=p} = 1$ все траектории вблизи γ будут замкнутыми.

Используя приведенные выше определения, можем, наконец, сформулировать основной результат, принадлежащий Андронову и Понтрягину.

Теорема о структурной устойчивости (в круге)

Дифференциальное уравнение (I) структурно устойчиво в том и только в том случае, если

- 1) особенности (I) гиперболические,
- 2) замкнутые траектории (I) гиперболические,
- 3) ни одна из траекторий (I) не соединяет седловые точки.

Заметим, что, согласно условию 1, имеется лишь конечное число особенностей, а из условий 2 и 3 следует, что существует лишь конечное число замкнутых траекторий.

Фазовые портреты двух дифференциальных систем, одна из которых структурно устойчивая, а другая неустойчивая, показаны на рис. 5.5. Следует, однако, подчеркнуть, что только вид фазового портрета системы не может служить основанием для суждения о структурной устойчивости данного дифференциального уравнения. Например, пусть начало

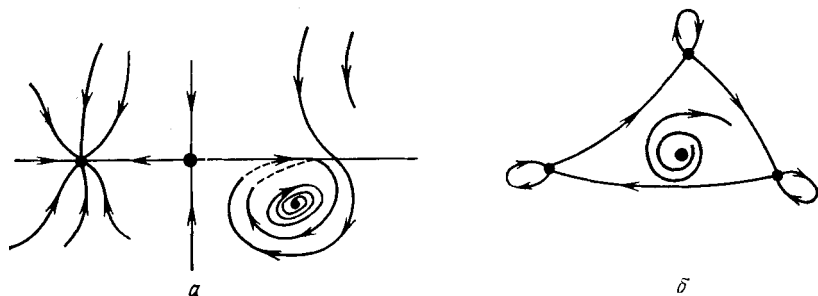


Рис. 5.5. Структурно устойчивая (а) и неустойчивая (б) системы.

координат является одной из особенностей структурно устойчивой системы (I); тогда система (II), которая получается умножением обоих уравнений системы (I) на $(x_1^2 + x_2^2)$, будет иметь точно такой же фазовый портрет, но уже не будет структурно устойчивой, поскольку в начале координат особенность системы (II) негиперболическая (оба корня характеристического уравнения нули).

Пример. Классическая система Лотки — Вольтерра

Простейшей моделью детерминированной системы двух видов хищник — жертва с непрерывным ростом популяций является система Лотки — Вольтерра

$$\frac{dH}{dt} = H(t) [a - \alpha P(t)],$$

$$\frac{dP}{dt} = P(t) [-b + \beta H(t)],$$

где $H(t)$ и $P(t)$ — популяции жертвы и хищника соответственно, а параметры a и b связывают рождаемость H и смертность P ; параметры α и β учитывают взаимодействие между видами. По очевидным и естественным соображениям ограничимся рассмотрением области $H \geq 0, P \geq 0$ и будем считать все параметры системы положительными. (Для того

чтобы удовлетворить условиям теоремы о структурной устойчивости, выберем масштабы H и P таким образом, чтобы значения H и P лежали в той части единичного круга, которая принадлежит первому квадранту.)

Точка равновесия системы, представляющая интерес в практическом отношении, определяется как

$$H^* = b/\beta, \quad P^* = a/\alpha.$$

Рассмотрев матрицу Якоби

$$J = \begin{bmatrix} a - \alpha P & -\alpha H \\ \beta P & -b + \beta H \end{bmatrix}$$

в точке (H^*, P^*) , получим, что собственные значения матрицы J в особой точке будут чисто мнимыми числами

$$\lambda = \pm i(ab)^{1/2}.$$

Таким образом, особенность не является гиперболической. Следовательно, классическая модель Лотки — Вольтерра не будет структурно устойчивой, поскольку нарушено условие 1 теоремы.

При обсуждении вопросов структурной устойчивости на интуитивном уровне обращает на себя внимание утверждение, что «система структурно устойчива, если все близкие к ней системы имеют качественно такое же поведение». Предыдущий пример показывает, что, как только мы уточняем понятия «близкие», «те же», «качественно» и «поведение», общее представление уже выглядит несколько иначе, поскольку все траектории системы являются замкнутыми, причем точка равновесия (H^*, P^*) — центр. Более того, любая система, близкая к данной и получающаяся при изменении ее параметров a , b , α , β , будет вести себя аналогичным образом.

Итак, интуитивные рассуждения могут привести к заключению о структурной устойчивости рассматриваемой системы. Однако из геометрических соображений очевидно, что траектории близких систем нельзя непрерывно отобразить друг в друга. Следовательно, согласно определению 5.3, данная система не будет структурно устойчивой, и этот вывод, разумеется, вытекает из того факта, что не все особенности системы являются гиперболическими. Таким образом, интуитивные представления и точные определения не всегда приводят к одному результату, поэтому, применяя математический аппарат, мы должны придерживаться данных определений. Указанный выше случай говорит о возможности другого определения структурной устойчивости.

Одно время предполагалось, что свойством структурной устойчивости должны обладать почти все дифференциальные системы (другими словами, ожидали, что структурно устойчивые системы образуют открытое плотное множество в множестве всех систем). И это действительно справедливо в случае одного или двух измерений, однако, как показали Смейл и Вильямс, для более высокого числа измерений справедливо обратное утверждение, т. е., вообще говоря, нет уверенности в том, что неустойчивые системы могут быть произвольно близко аппроксимированы структурно устойчивыми системами, если размерность фазового пространства $n \geq 3$. К счастью, оказывается, что в действительности существует широкий класс n -мерных систем, для которых можно установить свойство структурной устойчивости. Наиболее простыми и лучше всего изученными являются дифференциальные системы Морса — Смейла, которые содержат лишь конечное число особенностей и замкнутых траекторий. Системы другого типа, например, системы Аносова, имеют очень сложную геометрическую структуру, связанную с тем, что они содержат бесконечно много замкнутых траекторий. Такие системы будут затронуты в одном из следующих разделов.

ТЕОРИЯ КАТАСТРОФ

В последнее время большое внимание привлекает получившая широкую известность теория «катастроф», затрагивающая один из аспектов проблем, изучаемых общей теорией структурной устойчивости и теорией бифуркаций. В основных чертах теорию катастроф можно представить как метод, позволяющий получить частичный ответ на следующий вопрос: каковы обычно встречаемые типы в k -параметрическом семействе функции? Аналогичный математический метод используется также и при изучении вопроса противоположного характера: если дана функция, как будет выглядеть семейство, содержащее близкие к ней функции?

Важность этих вопросов для практического конструирования моделей обусловлена тем, что в основе применения элементарной теории катастроф лежит предположение о направленности исследований при изучении конкретной системы, хотя возможная цель и не является четко обозначенной. Вкратце задача состоит в следующем. Рассматривается система, динамика которой принадлежит к градиентному типу, и делается попытка минимизировать (локально) некоторую функцию цены. В нашем распоряжении имеется k параметров управления $\alpha_1, \alpha_2, \dots, \alpha_k$, и при этом выходные параметры системы принимают такие значения $x_1^*, x_2^*, \dots, x_n^*$

в состоянии равновесия, что достигается локальная минимизация некоторой функции

$$f(x_1, x_2, \dots, x_n; \alpha_1, \alpha_2, \dots, \alpha_k).$$

По аналогии с классической механикой функция f называется *потенциальной* (или *энергетической*) *функцией* системы. В общем случае значения x_i^* , соответствующие состоянию равновесия, зависят от выбора параметров α , поэтому

$$x_i^* = x_i^*(\alpha), \quad i = 1, 2, \dots, n.$$

Рассматривая скачкообразные изменения значений x_i^* , происходящие в результате плавного изменения параметров α , мы приходим к понятию «катастрофа».

Очевидно, существует бесконечно много систем указанного выше типа (каждой функции f соответствует своя система). Между тем многие из этих систем окажутся идентичными, если ввести преобразование координат в пространстве входных и выходных переменных α и x . Наиболее простой метод отсеять несущественные изменения системы состоит в том, чтобы выделить и рассмотреть лишь такие свойства функции f , которые имеют чисто топологический характер (в самом деле, допускаются лишь гладкие, т. е. бесконечное число раз дифференцируемые функции f и гладкие изменения координат). При этом основной результат теории катастроф, теорема Тома, дает возможность топологически классифицировать все гладкие потенциальные функции. Как мы увидим далее, наиболее замечательная особенность этой теоремы заключается в том, что такая классификация зависит только от числа k параметров управления (предполагаемого конечным).

Важность теоремы Тома для приложений обусловлена тем обстоятельством, что в общем случае функция f нам неизвестна: предполагается лишь, что она является потенциалом, описывающим динамику данной системы. При этом теорема позволяет оправдать выбор небольшого конечного числа «канонических» потенциалов в качестве моделей рассматриваемого процесса, поскольку дает возможность полагать, что «истинная» функция f , какой бы она вид ни принимала, будет отличаться от канонической модели только результатом преобразования координат. Кроме того, теорема гарантирует структурную устойчивость канонической модели. Отсюда следует, что истинная модель должна проявлять те же свойства топологического характера, что и каноническая модель. Прежде чем дать формулировку результата Тома и приступить к его обсуждению, проанализируем простой частный случай.

Рассмотрим семейство функций в R^2 , зависящее от единственного параметра $\alpha \in R$ и определяемое уравнением

$$f(x_1, x_2; \alpha) = x_1^3 - \alpha x_1 - x_2^2.$$

Критические точки функции f соответствуют $x_1^*(\alpha)$, $x_2^*(\alpha)$ и определяются уравнениями

$$\frac{\partial f}{\partial x_1} = 3x_1^2 - \alpha = 0, \quad \frac{\partial f}{\partial x_2} = -2x_2 = 0.$$

Таким образом, многообразие критических точек M_f располагается на плоскости $x_2 = 0$ вдоль кривой $3x_1^2 - \alpha = 0$, лежащей на плоскости (x, α) в R^3 . Исследуем критические точки функции $x_1^3 - \alpha x_1$ при разных значениях α .

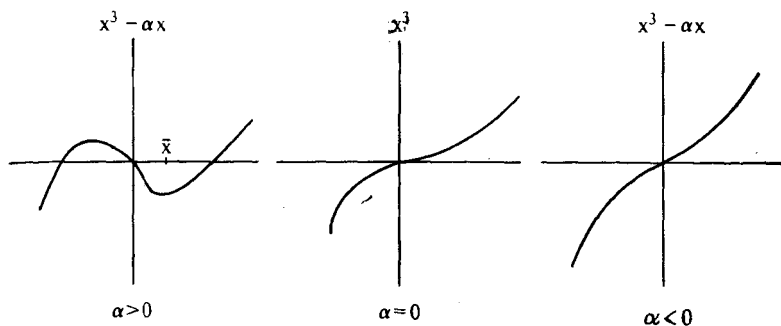


Рис. 5.6. Поведение функции $x^3 - \alpha x$ при разных значениях α .

Как показано на рис. 5.6, имеются две критические точки $x_1^3 - \alpha x_1$ при $\alpha > 0$: максимум параболического типа, где величина x_1 отрицательна, и минимум также параболического типа в некоторой точке $\bar{x} > 0$. При уменьшении α эти две критические точки сливаются в единственную точку перегиба, характерную для кубической кривой в вырожденном случае, когда $\alpha = 0$; при $\alpha < 0$ критические точки отсутствуют.

Проекция M_f в пространство α

$$\begin{aligned} \psi: M_f &\rightarrow R^k \\ (x^*(\alpha), \alpha) &\rightarrow \alpha \end{aligned}$$

называется *отображением катастроф* семейства $f(x; \alpha)$. Для большинства значений α многообразие M_f обеспечивает локальное покрытие пространства управления R^k , быть может, многолистное. В то же время, когда ψ сингулярно, число листов может внезапно меняться. Таким образом, при некото-

ром значении $\alpha = \hat{\alpha}$ происходит слияние или бифуркация критических точек функции $f(x; \alpha)$. Такая особая точка называется *точкой катастрофы* семейства $f(x; \alpha)$.

Для приведенного выше примера $f(x_1, x_2; \alpha) = x_1^3 - \alpha x_1 - x_2^2$ имеет точку катастрофы при $\hat{\alpha} = 0$. Важно отметить, что любое малое возмущение дает новое семейство

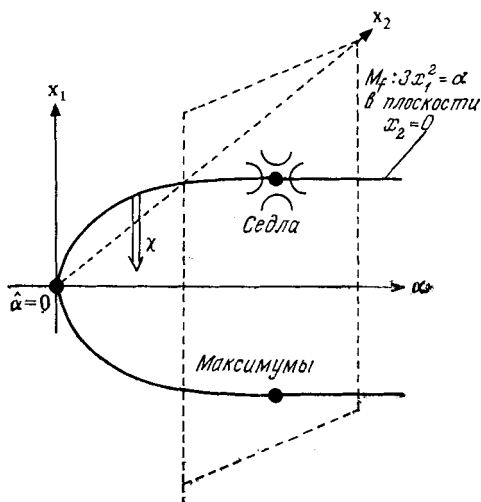


Рис. 5.7. Катастрофа типа складки.

функции $\tilde{f}(x_1, x_2; \alpha)$, которое должно обязательно иметь точку катастрофы где-то вблизи $\hat{\alpha} = 0$, причем вблизи катастрофы M_f имеет сходный с M_f вблизи $\hat{\alpha} = 0$ характер покрытия оси α . Подобная ситуация, отвечающая катастрофе типа «складки», показана на рис. 5.7.

Теперь можно перейти к обсуждению основной теоремы теории катастроф, которая впервые была сформулирована Томом, а затем дополнена Зиманом.

Теорема Тома — Зимана

Для каждого $k \leq 5$ и $n \geq 1$ существует открытое плотное множество C^∞ -потенциальных функций \mathcal{F} , таких, что

1) M_f — дифференцируемое k -многообразие, гладко вложенное в R^{n+k} .

2) Каждая особенность отображения катастроф $\psi: M_f \rightarrow R^k$ локально эквивалентна одному из конечного числа стандартных типов, называемых элементарными катастрофами. Число таких типов следующим образом зависит от k :

k	1	2	3	4	5	≥ 6
Число типов	1	2	5	7	11	∞

3) Отображение ψ структурно устойчиво в каждой точке M_f по отношению к малым возмущениям f из \mathcal{F} . Кроме того, существует каноническая форма $f(x; \alpha)$ вблизи каждой точки $(x^*, \alpha) \in M_f$ (табл. 5.1).

Таблица 5.1

Канонические формы функции $f(x; \alpha)$

k	n	Каноническая форма $f(x; \alpha)$	Название
1	1	$x_1^2 + \alpha x_1$	Складка
2	1	$x_1^4 + \alpha_1 \frac{x_1^2}{2} + \alpha_2 x_1$	Сборка
3	1	$\frac{x_1^5}{5} + \alpha_1 \frac{x_1^3}{3} + \alpha_2 \frac{x_1^2}{2} + \alpha_3 x_1$	Ласточкин хвост
4	1	$\frac{x_1^6}{6} + \alpha_4 \frac{x_1^4}{4} + \alpha_1 \frac{x_1^3}{3} + \alpha_2 \frac{x_1^2}{2} + \alpha_3 x_1$	Бабочка
3	2	$x_1^2 + x_2^3 + \alpha_3 x_1 x_2 - \alpha_1 x_1 - \alpha_2 x_2$	Гиперболическая омбилическая точка
3	2	$x_1^3 - 3x_1 x_2^2 + \alpha_3(x_1^2 + x_2^2) - \alpha_1 x_1 - \alpha_2 x_2$	Эллиптическая омбилическая точка
4	2	$x_1^2 x_2 + x_2^4 + \alpha_3 x_1^2 + \alpha_4 x_2^2 - \alpha_1 x_1 - \alpha_2 x_2$	Параболическая омбилическая точка
5	1	$x_1^7 + \alpha_1 x_1^5 + \alpha_2 x_1^4 + \alpha_3 x_1^3 + \alpha_4 x_1^2 + \alpha_5 x_1$	Вигвам
5	2	$x_1^2 x_2 - x_2^5 + \alpha_1 x_2^3 + \alpha_2 x_2^2 + \alpha_3 x_1^2 + \alpha_4 x_2 + \alpha_5 x_1$	Вторая эллиптическая омбилическая точка
5	2	$x_1^2 x_2 + x_2^5 + \alpha_1 x_2^3 + \alpha_2 x_2^2 + \alpha_3 x_1^2 + \alpha_4 x_2 + \alpha_5 x_1$	Вторая гиперболическая омбилическая точка
5	2	$1 \pm (x_1^3 + x_2^2 + \alpha_1 x_1 x_2 + \alpha_2 x_2^2 + \alpha_3 x_1 x_2 + \alpha_4 x_2 + \alpha_5 x_1)$	Символическая омбилическая точка

Примечание; для $k=4$ и $n=1$ должно быть $\alpha_3 x_1$.

Замечания. 1. C^∞ -эквивалентность двух отображений означает, что если $\psi: M \rightarrow N$, $\psi': M' \rightarrow N'$, то отображения ψ и ψ' эквивалентны при условии, что существуют

дiffeоморфизмы (взаимно однозначные отображения, C^∞) h и k , такие, что

$$k^{-1}\psi' h = \psi.$$

2. Грубо говоря, структурная устойчивость ψ в каждой точке M_f означает, что для заданной точки M_f существует некоторая окрестность J в \mathcal{F} , такая, что любая функция в этой окрестности имеет отображение катастроф, эквивалентное ψ_f .

3. При $k \geq 6$ бесконечное число отображений катастроф можно исключить при помощи более слабого определения эквивалентности. Однако на практике наиболее важное значение имеет конечная C^∞ -классификация при малых k , и именно она является наиболее приемлемой для математической постановки задачи.

ПРИМЕР ИСПОЛЬЗОВАНИЯ МЕТОДОВ ТЕОРИИ КАТАСТРОФ. ЗАГРЯЗНЕНИЕ ОЗЕР

Часто возникает необходимость объяснить (смоделировать) происходящие за короткие промежутки времени значительные колебания биомассы фитопланктона в небольших эвтрофных озерах. Поскольку некоторые особенности наблюдаемых изменений фитопланктона свидетельствуют в пользу возможности применения методов теории катастроф, ниже предлагается динамическая модель этого процесса с использованием катастрофы типа сборки.

Будем считать общую биомассу¹⁾ всех видов планктонных водорослей в озере выходной переменной x системы. В большинстве озер, особенно в случаях доминирования сине-зеленых водорослей, лимитирующим биогенным элементом является фосфор. Исходя из этого, концентрацию растворимого фосфата²⁾ следует включить в модель как один из параметров управления α_1 . В качестве эквивалента уровня освещенности и температуры в озере будем рассматривать изменения одного из представителей фитопланктона — род Апабаепа. Наблюдения показывают, что отмиранию Апабаепа предшествует ее скопление на поверхности, вызванное способностью этого вида к образованию газовых вакуолей и поднятию на поверхность в условиях небольшой освещенности. Более того, колебания концентрации растворенного фосфата оказываются противоположными колебаниям концентрации фитопланктона. Таким образом, в качестве второго параметра управления α_2 выбираем биомассу¹⁾ Апабаепа.

¹⁾ В единицах концентрации хлорофилла, мкг/мл. — *Прим. ред*

²⁾ В единицах концентрации фосфора, мкг/мл. — *Прим. ред*.

Для вывода уравнения, описывающего изменения концентрации фосфата, используется тот факт, что скорость накопления фосфора планктонными водорослями пропорциональна концентрации последних во время «цветения» воды. Когда наступает отмирание, фосфат осаждается на дно вместе с клетками и может быть «освобожден» только после окончания периода «цветения» и следующего за ним процесса отмирания. Таким образом, концентрация растворенного фосфата будет по экспоненте приближаться к равновесному уровню для данного озера.

Чтобы получить уравнение, определяющее концентрацию Апабаепа, заметим, что скорость роста Апабаепа будет зависеть только от численности ее живых клеток, в то время как снижение характерной для этого вида скорости роста будет зависеть также и от концентрации клеток конкурирующих видов, потребляющих тот же биогенный элемент. В конечном итоге уравнение для общей концентрации фитопланктона вытекает из логистической модели, которая подробно описывается в следующем разделе.

Учитывая все сказанное выше и используя каноническое уравнение сборки для выходной переменной x , получим следующую систему уравнений, моделирующую рассматриваемый процесс:

$$\dot{x} = -(c_1 x^3 - c_2 \alpha_1 x + c_3 \alpha_2),$$

$$\dot{\alpha}_1 = -c_4 x (\alpha_1 - a_0),$$

$$\dot{\alpha}_2 = c_5 \alpha_1 \alpha_2 - c_6 \alpha_2 x,$$

где c_1, \dots, c_6 — константы скорости и a_0 — равновесная концентрация фосфата.

С учетом экспериментальных данных, полученных в шт. Алабама (США) при исследовании пруда для разведения зубатки, приведенные выше уравнения были численно проинтегрированы для описания цикла эвтрофикации. При этом использовались следующие значения констант:

$$c_1 = 3,10 \text{ (мкг хлорофилла/мл)}^{-2} \text{ (время)}^{-1},$$

$$c_2 = 0,60 \text{ (мкг фосфора/мл)}^{-1} \text{ (время)}^{-1},$$

$$c_3 = 0,05 \text{ (время)}^{-1},$$

$$c_4 = 1,00 \text{ (мкг хлорофилла/мл)}^{-1} \text{ (время)}^{-1},$$

$$c_5 = 1,35 \text{ (мкг фосфора/мл)}^{-1} \text{ (время)}^{-1},$$

$$c_6 = 1,95 \text{ (мкг хлорофилла/мл)}^{-1} \text{ (время)}^{-1},$$

$$a_0 = 0.$$

Начальными значениями были $x_0 = 0,11$ мкг хлорофилла/мл, $\alpha_1^0 = 1,28$ мкг фосфора/мл, $\alpha_2^0 = 0,02$ мкг хлорофилла/мл.

Результаты эксперимента показаны на рис. 5.8—5.10. Для концентрации Апабаена и общей концентрации фитопланктона данные эксперимента достаточно хорошо согласуются с приведенной выше моделью, за исключением большого спада в пятый интервал времени. Этот спад соответствует возрастанию средней силы ветра в соответствии с измерениями и, возможно, является следствием усиленного вертикального

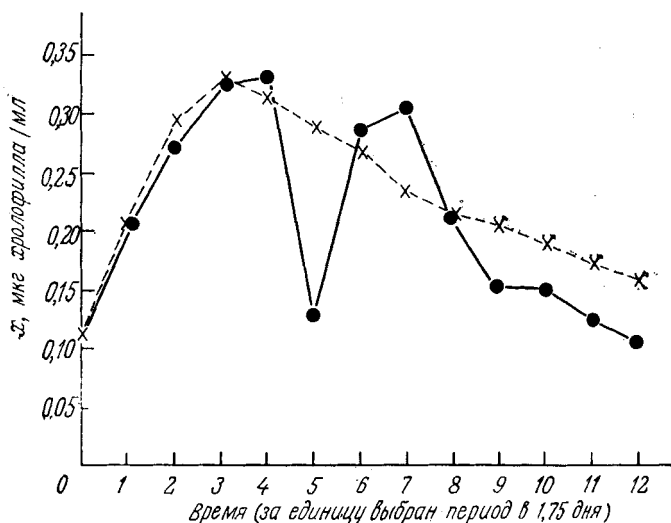


Рис. 5.8. Траектории переменных управления и переменных состояния для суммарной концентрации (x) водорослей [Parks et al., 1975].

●—● результаты наблюдений; ×—× теоретические значения.

или горизонтального перемешивания, в результате которого в месте взятия проб клетки могли переместиться из поверхностного слоя пруда.

Таким образом, моделирование дает возможность получить детерминированный диапазон, в пределах которого на цикл цветение — отмирание водорослей могут влиять случайные явления, например погода.

Совпадение теоретических и экспериментальных данных о концентрации растворенного фосфата оказывается менее точным. Подобное расхождение может быть вызвано несколькими причинами. Например, при измерении концентрации фосфата могли не учитываться растворенные удобрения из-за временных и пространственных разрывов в измерениях, которые проводились только раз в 1—3 дня и только в одной точке пруда. Кроме того, модель может быть

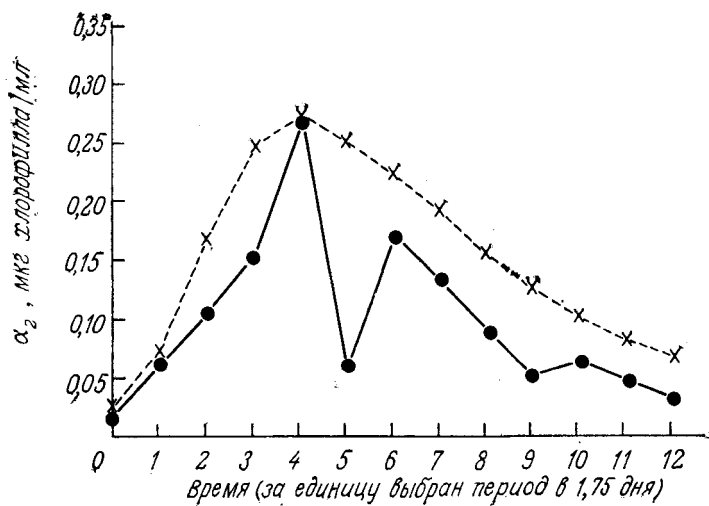


Рис. 5.9. Траектории переменных управления и переменных состояния для концентрации (α_2) Апабаена (Parks et. al., 1975).

●—● результаты наблюдений; ×—× теоретические значения.

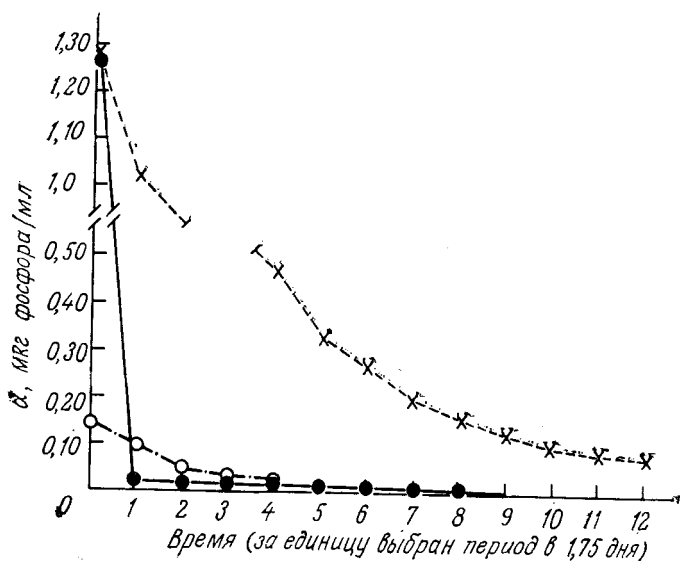


Рис. 5.10. Траектории переменных управления и переменных состояния для концентрации (α_1) фосфата в окружающей среде [Parks et. al., 1975].

●—● результаты наблюдений; ×—× теоретические значения; ○—○ экспериментальные значения.
(На рисунке по оси ординат должно быть α_1)

неправильной и нуждаться в уточнении. Поэтому только дальнейшие эксперименты, включающие более подробное изучение концентрации фосфата, могут дать материал для общей оценки пригодности модели.

КАТАСТРОФА ТИПА СБОРКИ И ЛОГИСТИЧЕСКОЕ УРАВНЕНИЕ

Возражения, которые встречает применение теории катастроф для моделирования явлений природы, сводятся к тому, что при этом не дается никакого физического обоснования для применения стандартного канонического уравнения с целью описания результирующего уравнения. Постараемся показать, что для моделирования многих процессов роста, описываемых классическим логистическим уравнением, применение канонической катастрофы типа сборки является логически обоснованным.

Логистическое уравнение, широко используемое в экологии для моделирования возрастания или уменьшения численности популяций при наличии верхнего предела емкости среды, имеет вид

$$\dot{x} = (a - dx)x,$$

где a — внутренняя скорость роста без лимитирующего влияния среды и d — вклад единицы популяции в уменьшение внутренней скорости роста вследствие влияния плотности. Очевидно, что при $ax = dx^2$ рост популяции прекращается.

Одно из допущений логистической модели заключается в том, что закон уменьшения внутренней скорости роста при добавлении к популяции каждой следующей особи является линейным относительно x . Хотя экспериментальные данные, по-видимому, подтверждают справедливость этого закона для разных видов в различных условиях среды, тем не менее могут существовать определенные сочетания видов и условий среды, для которых влияние плотности более существенно, и поэтому внутренняя скорость роста в подобных случаях может быть пропорциональна квадрату численности популяции. Например, в случае загрязнения озер поступление биогенных элементов из внешних источников минимально, что вызывает конкуренцию при использовании доступного биогенного элемента даже при высоком уровне его концентрации. Кроме того, известно, что *Апабаепа* выделяет вещество, токсичное для других видов фитопланктона, что может резко затормозить развитие фитопланктона при его большой плотности.

Уточнение логистического уравнения, учитывающее указанные эффекты, приводит к новому уравнению

$$\dot{x} = (a - dx^2)x.$$

С учетом вымирания отдельных видов имеем

$$\dot{x} = (a - dx^2)x - b,$$

или

$$\dot{x} = -(x^3 + (-a)x + b)$$

(полагаем величину d равной единице). Таким образом, логическое развитие логистической модели немедленно приводит к каноническому уравнению катастрофы типа сборки.

УСТОЙЧИВОСТЬ ПО ВОЗМУЩЕНИЮ И ПО НАЧАЛЬНОМУ ЗНАЧЕНИЮ

В разделе, посвященном графам и процессам распространения возмущений, для взвешенного орграфа было введено понятие *устойчивость по возмущению*. Теперь введем понятие алгебраического критерия устойчивости по возмущению и по начальному значению и рассмотрим некоторые вопросы относительно связи между устойчивостью графа и его топологической структурой.

Основополагающим представлением при разработке критериев устойчивости графов является представление о характеристических значениях взвешенного орграфа. Для более четкой формулировки этого представления определим матрицу взаимосвязи A для графа G следующим образом:

$$a_{ij} = f(u_i, u_j), \quad i, j = 1, 2, \dots, n,$$

где u_1, u_2, \dots, u_n — вершины графа G , $f(\cdot, \cdot)$ — весовая функция. Тогда характеристические значения G определяются как собственные значения A .

Связь между значением $v_i(t)$ в каждой вершине в момент времени t , изменением значения $p_i(t)$ и матрицей взаимосвязи графа G дается следующей теоремой.

Теорема о распространении возмущения

Для простого процесса распространения возмущения, начинающегося в вершине u_i , имеем

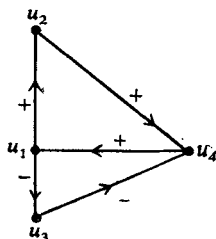
$$p_j(t) = [A^t]_{ij}$$

и

$$v_j(t) = v_j(0) + [I + A + A^2 + \dots + A^t]_{ij},$$

где A — матрица взаимосвязи для данного орграфа, $[·]_{ij}$ означает элемент соответствующей матрицы, стоящий на пересечении i -й строки и j -го столбца.

В качестве иллюстрации этой теоремы рассмотрим простой орграф



В данном случае матрица взаимосвязи такова:

$$A = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Предположим, что простой процесс распространения возмущения начинается в вершине u_1 в момент времени $t=0$, причем $v_i(0) = 0$, $i = 1, 2, 3, 4$. Простое вычисление показывает, что

$$A^2 = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

$$I + A + A^2 = \begin{bmatrix} 1 & 1 & -1 & 2 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$$

Поскольку процесс распространения возмущения начинается в вершине u_1 , получим, что $p_3(2)$ определяется элементом, стоящим на пересечении первой строки и третьего столбца матрицы A^2 , т. е. $p_3(2) = 0$.

Аналогичным образом получим $v_1(2) = v_1(0) + [I + A + A^2]_{1,1} = 0 + 1 = 1$. Очевидно, что эти результаты согласуются с теоремой о распространении возмущения.

Теперь постараемся связать представления об устойчивости по возмущению и об устойчивости по начальному значению с матрицей взаимосвязи A . По-видимому, в настоящее время неизвестны какие-либо общие необходимые и достаточные условия, однако если предположить, что оргграф G обладает различными характеристическими значениями (это условие является достаточно общим), то можно сформулировать теорему об устойчивости по возмущению.

Теорема об устойчивости по возмущению

Взвешенный оргграф G , характеристические значения которого различны, устойчив по возмущению для любого простого процесса распространения возмущения в том и только в том случае, если каждое характеристическое значение G по модулю не превосходит единицы.

Согласно данной теореме, для решения вопроса об устойчивости по возмущению необходимо лишь вычислить максимальный по модулю корень характеристического уравнения матрицы A . Если этот корень лежит вне круга единичного радиуса, то оргграф G не будет устойчивым по возмущению; в противном случае оргграф G устойчив по возмущению.

Устойчивость по начальному значению определяется на основе указанной выше теоремы об устойчивости по возмущению. Строгий критерий устойчивости дается следующей теоремой.

Теорема об устойчивости по начальному значению

Взвешенный оргграф G устойчив по начальному значению для любого простого процесса распространения возмущения в том и только в том случае, если оргграф G устойчив по возмущению для любого простого процесса распространения возмущения и единица не является характеристическим значением G .

Таким образом, как устойчивость по начальному значению, так и устойчивость по возмущению определяются путем исследования характеристических значений графа G , т. е. корней характеристического уравнения матрицы взаимосвязи A .

Пример. Борьба с насекомыми-вредителями

Рассмотрим задачу борьбы с насекомыми-вредителями культурных растений путем распыления инсектицидов. Введем следующие обозначения:

P_1 — культивируемое растение, прирост которого ограничен влиянием густоты посадки;

H_1 — растительноядное насекомое-вредитель, питающееся растением P_1 ;

W — насекомое-хищник, убивающее только насекомое H_1 ;
 G — насекомое-хищник, поедающее как насекомое H_1 , так и насекомое H_2 ;

H_2 — насекомое-вредитель, которое поедает растение P_2 ;
 I — инсектицид.

Знаковый оргграф, описывающий эту ситуацию¹⁾, показан на рис. 5.11. Отметим, что знак, приписанный дуге (x_i, x_j) ,

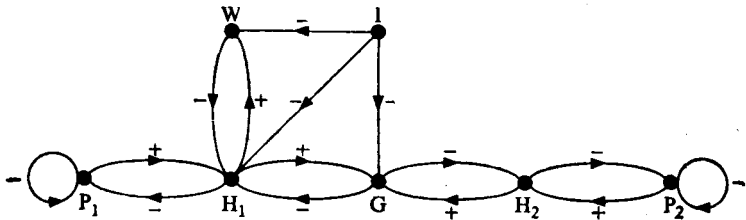


Рис. 5.11. Знаковый оргграф для задачи борьбы с насекомыми-вредителями.

указывает влияние изменения x_i на скорость изменения x_j . Матрица взаимосвязи для рассматриваемого оргграфа имеет вид

$$A = \begin{matrix} & \begin{matrix} P_1 & H_1 & W & G & H_2 & P_2 & I \end{matrix} \\ \begin{matrix} P_1 \\ H_1 \\ W \\ G \\ H_2 \\ P_2 \\ I \end{matrix} & \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

Корни характеристического уравнения матрицы A таковы:

$$\{-0,119 \pm 1,85i; -0,335 \pm 1,03i; -0,762; -0,328; 0\},$$

и, следовательно, максимальный по модулю корень имеет модуль больше единицы. Тогда, согласно теореме об устойчивости по возмущению, граф, представленный на рис. 5.11, не является устойчивым ни по возмущению, ни по значению.

¹⁾ Приведенный на рис. 5.11 знаковый оргграф не совсем точно отражает суть дела. Действительно, инсектицид должен истреблять также и насекомое H_2 (следует провести дугу со знаком минус от I к H_2), растения P_1 и P_2 должны влиять друг на друга (это приведет к дальнейшему усложнению оргграфа). — Прим. ред.

Аналогичный результат можно было предсказать и на основании анализа графа (рис. 5.11), поскольку он содержит много циклов усиления возмущения, например $G \rightarrow H_2 \rightarrow P_2 \rightarrow H_2 \rightarrow G$.

Если данный орграф не является устойчивым по возмущению, очевидный интерес представляет определение типа структурных изменений, которые оказывают стабилизирующее действие. Другими словами, желательно классифицировать устойчивые орграфы по их структурным характеристикам, что позволило бы определить способы стабилизации системы путем таких структурных изменений, которые переводят данный орграф в устойчивую структуру. К сожалению, это пока невозможно, хотя и получены полезные результаты при изучении некоторых частных классов орграфов, часто возникающих на практике.

АДАПТИРУЕМОСТЬ ДИНАМИЧЕСКИХ ПРОЦЕССОВ

В предыдущих разделах неоднократно упоминалась проблема отклика системы на возмущения ее состояния или характеризующих эту систему параметров. Грубо говоря, «адаптируемость» динамического процесса представляет собой меру возможности системы сохранить неизменность основного хода динамического процесса при воздействии возмущений. Очевидно, что такое «определение» содержит слишком много неясностей, поэтому в дальнейшем мы попытаемся сформулировать понятие *адаптируемость* на соответствующем математическом языке.

Следует сразу отметить, что практически не существует единого мнения относительно того, как именно должна определяться мера адаптируемости. Несмотря на то что между понятием адаптируемость и понятиями структурная устойчивость, бифуркация, катастрофа, связанная устойчивость и другими существует много общего, ни одно из них, по-видимому, не может дать исчерпывающего представления об адаптируемости, по крайней мере в том объеме, в котором оно используется исследователями в области экологии, где и возникло это представление, во всяком случае на терминологическом уровне. Поэтому мы вынуждены проводить наш анализ, используя собственные представления об адаптируемости, сформулированные на математическом языке, которые, однако, могут оказаться пригодными для объективного системного исследования.

Подобно понятию устойчивости, конкретная математическая формулировка понятия адаптируемости зависит от принятого математического описания системы. Все проведенные

до настоящего времени исследования в области адаптируемости основаны на описании системы с помощью обыкновенных дифференциальных уравнений. Мы также будем придерживаться такого способа описания, поскольку другие способы описания систем (при помощи графов, соотношений на входе и выходе и т. п.) пока еще изучены недостаточно. В самом деле, как будет показано, исследования, связанные с представлением об адаптируемости, все еще находятся на уровне исходных определений даже в рамках моделей, характеризующих дифференциальными уравнениями, и до сих пор нет реальных математических *результатов*. Можно только надеяться, что надлежащим образом сформулированные определения приведут к полезным для практики теоремам.

Начнем с анализа системы Σ , описываемой дифференциальными уравнениями

$$\dot{x} = f(x, a) + g(t), \quad x(0) = x_0,$$

и предположим, что в отсутствие действия внешних возмущений началом координат является точка равновесия, т. е. $f(0, a) \equiv 0$ для всех a при $g(t) \equiv 0$. Здесь a — вектор параметров системы. В свете вышешложенного центральное место в представлении об адаптируемости занимают следующие вопросы:

а) при каких условиях функция $g(t)$, характеризующая возмущения, может привести к тому, чтобы система, положение которой описывается функцией $x(t)$, покинула область притяжения D начала координат;

б) какие изменения параметров a приведут к такому искажению границы ∂D области D , чтобы положение системы оказалось сдвинутым в область аттрактора, отличного от аттрактора в начале координат.

Казалось бы, можно попытаться ответить на первый вопрос путем введения степени адаптируемости системы как меры наибольшего приближения $x(t)$ к ∂D в течение временной эволюции процесса (в предположении, что $g(t) \equiv 0$, $a = a^*$ при всех $t \geq 0$). Однако такое решение встречает возражения, которые носят скорее практический, нежели математический характер и заключаются в том, что минимальное расстояние $x(t)$ от ∂D является довольно слабым показателем того, насколько интенсивными могут быть возмущения, поглощаемые системой Σ , прежде чем она будет выведена из области D . Последнее связано с тем, что возмущение, необходимое для выведения системы Σ из области D , зависит не только от величины возмущения, но и от его направленности. Геометрически такая ситуация представлена на

рис. 5.12. Здесь область D может быть интерпретирована как потенциальная «яма» (подобно тому, как это делается при формулировке основных представлений теории Ляпунова). На рис. 5.12, *а* область D представляет собой «мелкую яму». В соответствии с этим даже в тех случаях когда $x(t)$ всегда остается вдали от ∂D , оказывается сравнительно легко сместить систему за границу ∂D . Вместе с тем если область D имеет форму типа показанной на рис. 5.12, *б*, то даже в тех случаях, когда $x(t)$ всегда лежит вблизи ∂D , необходимо существенное по величине возмущение, чтобы система Σ «перевалила через край ямы» в область притяжения другого положения равновесия.

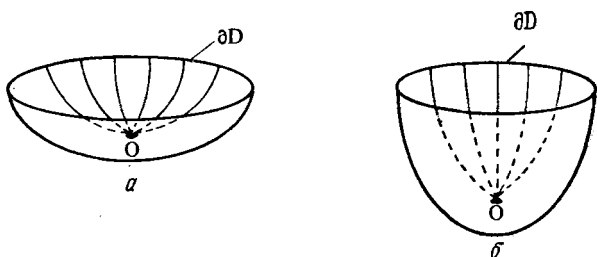


Рис. 5.12. Потенциальная яма для системы Σ .

Сказанное выше свидетельствует о том, что адаптируемость не является внутренним свойством системы Σ , а определяется характеристиками системы Σ и *классом допустимых возмущений*. В качестве простейшей иллюстрации этого факта отметим, что если минимальное расстояние $x(t)$ до ∂D равно α и $\|g(t)\| < \alpha$ при всех t , то система Σ будет «поглощать» все возможные возмущения, т. е. Σ будет обладать бесконечной адаптируемостью по отношению к возмущениям такого класса. Напротив, если *максимальное* расстояние $x(t)$ до ∂D равно β и $\|g(t)\| > \beta$ при всех t , то та же система Σ не сможет поглотить любое возмущение, т. е. Σ будет полностью неадаптируемой по отношению к данному классу возмущений. Таким образом, в отсутствие по крайней мере явного соглашения относительно класса допустимых возмущений понятие адаптируемости остается бессодержательным.

Мы уже убедились в том, что при определении меры адаптируемости необходимо принимать во внимание и величину, и направление возмущающей силы $g(t)$. Рассмотрим следующий подход к решению задачи. В каждый момент времени t построим вектор, направленный от $x(t)$ к точке на ∂D , ближайшей к $x(t)$ (рис. 5.13). Вектор $v(t)$ строится по известному вектору $x(t)$, определяемому, быть может, путем

численного интегрирования, и вектору $d(t)$, который известен, поскольку по предположению граница ∂D уже рассчитана. Таким образом, $v(t) = d(t) - x(t)$. Предположим, что функция $g(t)$ соответствует возмущению в виде «импульса» в момент $t = s$, т. е. $g(t) = \mu \delta(t - s)$, где μ — вектор, указывающий величину и направление импульса. Тогда можно попытаться путем сравнения векторов μ и $v(t)$ определить, будет ли система выведена за пределы ∂D . Как отмечалось,

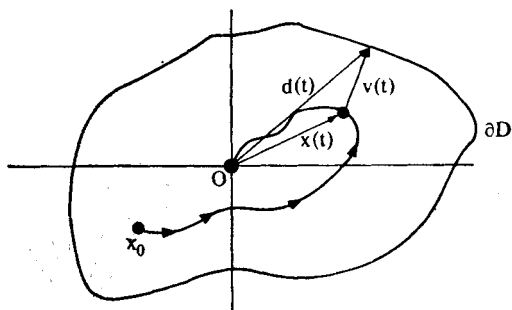


Рис. 5.13. Поведение системы Σ .

ответ на этот вопрос зависит от того, имеет ли вектор μ достаточную величину и надлежащее направление с тем, чтобы вывести $x(t)$ за границу ∂D . Отметим, что в данном случае можно не учитывать функцию $f(x)$, поскольку мы предположили, что $g(t)$ с точностью до множителя представляет собой δ -функцию.

Введем функцию m , характеризующую результат сравнения векторов μ и $v(t)$ по величине

$$m(t) = \|\mu\| - \|v(t)\|,$$

и функцию θ , характеризующую результат сравнения этих векторов по направлению,

$$\cos \theta(t) = \frac{(\mu, v(t))}{\|\mu\| \|v(t)\|},$$

где (\cdot, \cdot) означает скалярное произведение векторов, $\|\cdot\|$ — евклидову норму. При этом оказывается, что адаптируемость системы Σ в момент времени t можно характеризовать следующим полуколичественным соотношением:

низкая адаптируемость при $m(t) \geq 0$ и $\cos \theta(t) \approx 1$,
высокая адаптируемость при $m(t) < 0$ и $\cos \theta(t) < 0$.

Другими словами, система Σ адаптируема по отношению к импульсному возмущению μ в момент времени t , если вели-

чина μ очень мала или если возмущение μ стремится увести $x(t)$ от границы ∂D . Если же вектор μ по величине больше $v(t)$ и стремится подвести $x(t)$ к границе ∂D , то можно говорить о низкой адаптируемости системы Σ .

Приведенный выше анализ дает возможность заложить основу систематического математического изучения вопроса об адаптируемости системы, связанной с внешними возмущениями состояния $x(t)$. В рамках такого подхода нетрудно рассмотреть и случаи непрерывно действующих возмущений или комбинаций импульсов. Для этого достаточно применить стандартную математическую процедуру, аналогичную применяемой в теории вероятностей при переходе от дискретных к непрерывным или смешанным функциям распределения. Разумеется, при анализе таких случаев для установления степени адаптируемости системы Σ по отношению к данному классу возмущений необходимо принимать во внимание также и свободную динамику системы $f(x)$.

АДАПТИРУЕМОСТЬ И КАТАСТРОФЫ

Непосредственное внешнее воздействие на состояние равновесия является одним из факторов, благодаря которому система может быть смещена из области одного аттрактора в область другого. В предыдущем разделе были изучены

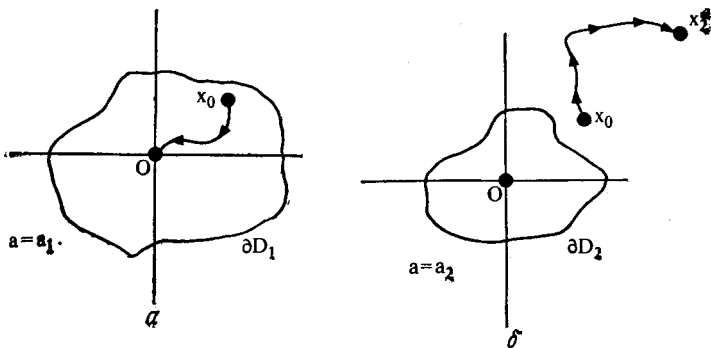


Рис. 5.14. Области притяжения начала координат.

некоторые связанные с этим вопросы. Рассмотрим теперь второй фактор, при действии которого система может быть сдвинута в область другого аттрактора. Таким фактором являются изменения в динамике системы, обусловленные изменением вектора параметров системы a .

Рассмотрим ситуацию, представленную на рис. 5.14. Когда вектор параметров системы $a = a_1$, начальное ее положение

x_0 лежит внутри области, ограниченной ∂D_1 и система стремится занять положение в начале координат. Если вектор a изменится и станет равным a_2 , то точно такое же начальное состояние x_0 будет уже лежать вне области, ограниченной ∂D_2 , и конечным итогом поведения системы будет ее переход в состояние равновесия $x = x_2^*$, расположенное далеко от начала координат. Легко представить себе ситуацию, в которой x_0 лежит очень близко к границе ∂D_1 . В этом случае даже небольшое изменение вектора a может вызвать деформацию

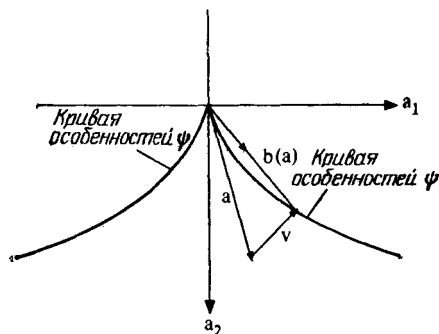


Рис. 5.15. Кривая особенностей отображения ψ катастрофы типа сборки.

∂D_1 , достаточную для того, чтобы положение системы x_0 оказалось в области притяжения к другому состоянию равновесия.

Как было показано в предыдущем разделе, только что описанная ситуация лежит в основе элементарной теории катастроф Тома—Зимана. При этом явно признается, что положения равновесия (в рассматриваемом случае x_2^* и начало координат), а также соответствующие им границы областей притяжения (∂D_1 , ∂D_2) зависят (главным образом) от вектора параметров системы a . Следовательно, существует тесная связь между отображением катастроф ψ и представлением об адаптируемости системы. Действительно, на интуитивном уровне ясно, что чем ближе к особой точке ψ расположено начальное значение вектора параметров a , тем меньше степень адаптируемости системы (по отношению к изменениям a).

Имея в виду сказанное выше, можно определить меру адаптируемости путем рассмотрения величины и направления такого изменения вектора a , которое необходимо, чтобы провести вектор a через особенность ψ (эта ситуация показана на рис. 5.15). Поскольку соответствующие рассуждения во многом повторяют изложенное в предыдущем разделе, отметим только, что представление об адаптируемости остается весьма неопределенным, пока не введено соглашение о

классе допустимых изменений вектора a . Как показано на рис. 5.15, для любого a мы строим вектор v от конца вектора a до ближайшей особенности ψ и проводим сравнение v с допустимыми изменениями a с целью оценить меру адаптируемости системы Σ для данного вектора a по отношению к допустимым возмущениям.

Пример. Модель биржевых операций

В качестве иллюстрации развитых выше представлений рассмотрим простейшую модель биржевых операций. Выходной переменной (описывающей состояние системы) является скорость изменения определенного индекса биржи (например, среднее по Доу — Джонсу), а входные переменные a_1 и a_2 характеризуют соответственно дополнительный спрос на акции со стороны основных покупателей и долю денежных средств, направленных на спекулятивные операции. Обсуждение подробностей, связанных с данной моделью, можно найти в статье, указанной в конце главы. Моделирование этой ситуации, связанное с использованием модели катастрофы типа сборки, соответствует картине, показанной на рис. 5.16.

Предположим, что единицы измерений выбраны так, чтобы рассматриваемая модель отвечала канонической катастрофе типа сборки, т. е. множество бифуркаций в пространстве исходных переменных S приводит к следующей связи между a_1 и a_2 :

$$a_2 = 5,67a_1^{2/3} \quad \text{или} \quad 27a_2^3 - 4a_1^3 = 0.$$

Эти уравнения получаются из канонического потенциала для сборки

$$f(x_1, \alpha_1, \alpha_2) = \frac{x_1^4}{4} + \frac{\alpha_1}{2} x_1^2 + \alpha_2 x_1$$

на основании условия, согласно которому вдоль линий сборки мы должны иметь

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial^2 f}{\partial x^2} = 0.$$

При помощи указанных соотношений можно исключить x и получить представленный выше результат.

Предположим, что вектор начальных параметров a описывается следующими значениями: дополнительный спрос на акции со стороны основных покупателей есть 0,1 и доля денежных средств, направленных на спекулятивные операции,

есть 0,58, т. е. $a = (0,1; 0,58)$. Ближайшая к a точка на бифуркационной кривой

$$b(a) = (0,3; 0,44).$$

Таким образом, вектор $v(a)$ имеет вид

$$v(a) = (0,2; -0,133).$$

Вектор $v(a)$ представляет собой *изменение* a , необходимое для того, чтобы пересечь бифуркационную кривую. Отсюда

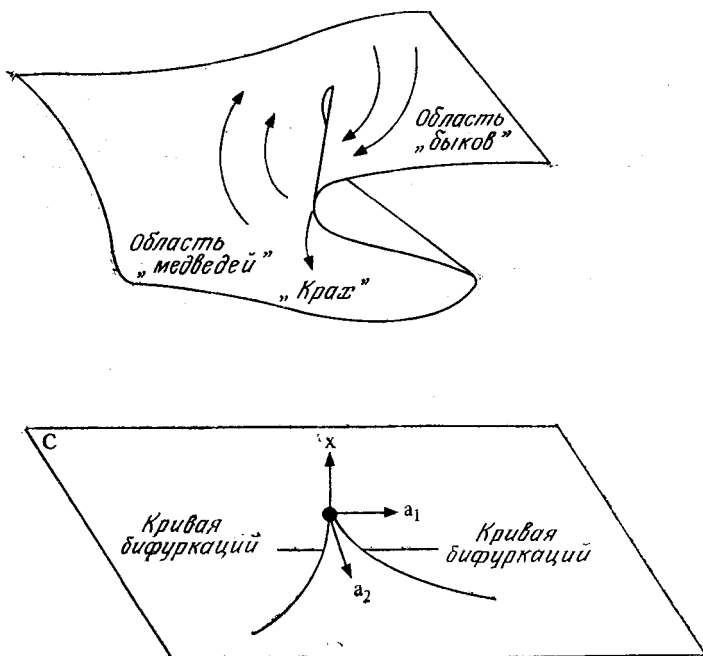


Рис. 5.16. Катастрофа типа сборки для модели биржевых операций.

следует, что модель биржевых операций, описываемая вектором параметров a , обладает почти *вдвое* более высокой адаптируемостью по отношению к изменениям a_1 (т. е. спроса на акции со стороны основных покупателей), чем к изменениям a_2 (т. е. доли средств, направленных на спекулятивные операции).

Разумеется, приведенная выше интерпретация моделей с точки зрения теории катастроф и исследование адаптируемости систем ограничены ситуациями, для которых справедливы гипотезы, положенные в основу элементарной теории катастроф. А именно:

1) динамика системы описывается гладкой функцией, т. е. является C^∞ -функцией x и a и принадлежит к градиентному типу;

2) вектор параметров системы имеет не более пяти компонентов;

3) положениями равновесия системы являются только неподвижные точки: система Σ не имеет предельных циклов, аттракторов Лоренца и других более экзотических типов начальных состояний. Другими словами, возможны лишь «элементарные» катастрофы.

Нарушение одного или даже всех этих условий сохраняет тем не менее возможность исследования адаптируемости, но уже на основании представлений о структурной устойчивости, отличных от тех, которыми оперирует теория катастроф. Рассмотрим кратко некоторые из них.

СИСТЕМЫ МОРСА — СМЕЙЛА И АДАПТИРУЕМОСТЬ

Проблема структурной устойчивости систем, размерность которых меньше или равна двум, уже обсуждалась. Из результатов, представленных в последних разделах, с очевидностью следует, что ряд основополагающих представлений об адаптируемости относится непосредственно к структурно устойчивым системам. Действительно, главным элементом в оценке степени адаптируемости является установление следующего факта: останется ли динамика системы «без существенных изменений» при наличии возмущений. Одной из возможностей получить ответ на этот вопрос является анализ устойчивости векторного поля. Такая возможность обусловлена тем, что исходную систему и систему, получающуюся в результате действия возмущений, можно рассматривать как совпадающие в качественном смысле, если они имеют сходные фазовые портреты. В данном разделе мы уточним представление о структурной устойчивости векторного поля, а также исследуем широкий класс встречающихся на практике систем, динамика которых является структурно устойчивой. Это так называемые системы Морса — Смейла. Представленные ниже результаты обеспечивают частичное обобщение теоремы о структурной устойчивости, приведенной в разделе о структурной устойчивости для диска, на случай n -мерных систем.

Прежде всего рассмотрим n -мерное C^∞ -многообразие, которое обозначим M (M представляет собой топологическое пространство, которое в U -окрестности каждой точки m

является аналогом R^n). На M можно надлежащим образом ввести операции дифференцирования (рис. 5.17), причем преобразование координат $\alpha: M \rightarrow R^n$ будет C^∞ . Простыми примерами многообразий служат R^n , сферы, торы и открытые подмножества R^n .

C^1 -векторное поле (или поле направлений для дифференциального уравнения) на M является результатом процедуры,

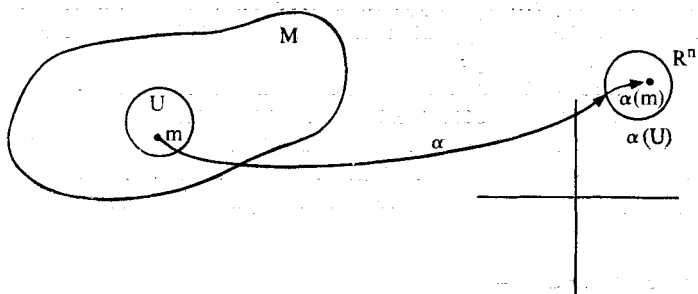


Рис. 5.17. Многообразие M и локальная окрестность точки m .

согласно которой каждой точке $x \in M$ приписывается вектор $v(x)$ так, чтобы вектор-функции $v(x)$ были гладкими (в смысле C^1).

Понятие *подобные фазовые портреты* для двух векторных полей v и w вводится следующим определением.

Определение 5.6

Два векторных поля называются топологически совместимыми, если существует взаимно-однозначное непрерывное отображение h , переводящее соответствующие полю v направленные кривые в направленные кривые, соответствующие полю w .

Таким образом, если векторные поля v и w топологически совместимы, они будут иметь одинаковое число точек равновесия, одинаковое число периодических замкнутых траекторий и сходное общее качественное поведение.

Теперь можно пояснить смысл *близких* векторных полей путем введения топологии в пространстве $V(M)$, C^1 -векторных полей на M . Будем называть поля v и w *близкими*, если они *близки* поточечно и таким же свойством обладают их первые производные. (Точные определения достаточно громоздки, поэтому за подробностями мы отсылаем читателя к ссылкам, приведенным в конце главы.) Представление о близости векторных полей дает нам возможность ввести определение *структурной устойчивости*.

Определение 5.7

Векторное поле $v \in V(M)$ является структурно устойчивым, если существует окрестность $N(v)$ в $V(M)$, такая, что каждый вектор $w \in N(v)$ топологически совместим с v .

Основная проблема структурной устойчивости модели заключается в том, чтобы установить необходимые и достаточные условия структурной устойчивости некоторого векторного поля.

Примеры. Простой гармонический осциллятор

Было показано, что элементарная система, какой является гармонический осциллятор, описывается уравнением

$$\ddot{x} = -x$$

и имеет фазовый портрет, состоящий из концентрических окружностей в фазовой плоскости (x, \dot{x}) . Это векторное поле не будет структурно устойчивым, поскольку любое векторное поле, топологически совместимое с v , имеет только периодические замкнутые траектории, причем стрелки, показывающие направление v (см. рис. 5.4), всегда можно слегка отклонить в сторону начала координат, чтобы получить близкое векторное поле с непериодическими траекториями.

Пример. Уравнение Ван дер Поля

В случае уравнения Ван дер Поля векторное поле v определяется уравнениями

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = -\varepsilon(x_1^2 - 1)x_2 - x_1, \quad \varepsilon > 0.$$

Эта система имеет одну периодическую замкнутую траекторию, причем каждая внешняя по отношению к ней траектория свертывается, приближаясь к замкнутой траектории, а каждая внутренняя разворачивается по спирали, также приближаясь к замкнутой траектории (рис. 5.19). Следовательно, это уравнение структурно устойчиво при всех $\varepsilon > 0$.

Наиболее важный класс структурно устойчивых векторных полей системы Морса — Смейла характеризуется следующими условиями:

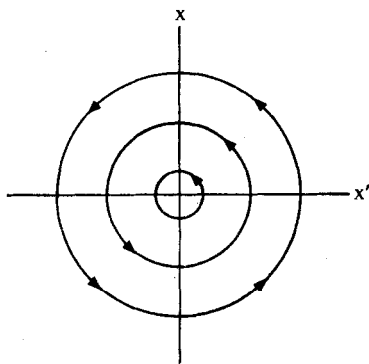


Рис. 5.18. Фазовый портрет простого гармонического осциллятора.

поле $v(x)$ имеет конечное множество точек равновесия, т. е. точек x , таких, что $v(x) = 0$, и каждая такая точка гиперболическая;

поле $v(x)$ имеет конечное множество замкнутых периодических траекторий и каждая из них является гиперболической;

устойчивые и неустойчивые многообразия точек равновесия и замкнутых периодических траекторий, пересекаясь, встречаются трансверсально (это означает недопустимость касания между устойчивыми и неустойчивыми многообразиями);

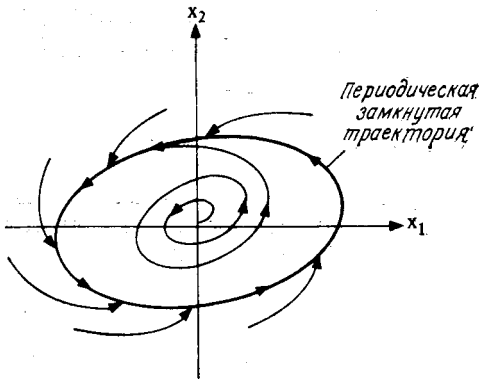


Рис. 5.19. Фазовый портрет уравнения Ван дер Поля.

неподвижными точками системы Σ являются точки равновесия вместе с точками периодических замкнутых траекторий. (Точка x считается неподвижной, если в каждой открытой окрестности U точки x при любом $T > 0$ существует $t > T$, такое, что $x(t) \in U$, если $x(0) \in U$. Другими словами, если любая кривая, описывающая решение системы Σ , начинается в U , она должна бесконечно часто возвращаться в U .)

В разделе о структурной устойчивости были даны необходимые и достаточные условия структурной устойчивости в случае, когда многообразие M представляет собой двумерный диск. В действительности оказывается справедливым следующий, более сильный результат: если $\dim M = 2$, то структурно устойчивые системы на M совпадают с системами Морса — Смейла. Если $\dim M > 2$, то могут существовать другие структурно устойчивые на M векторные поля, дополнительные к системам Морса — Смейла.

Представление об адаптируемости, рассмотренное в данной книге, очевидным образом связано с только что приве-

денными представлениями о структурной устойчивости, хотя и имеются следующие заметные расхождения.

1. Структурная устойчивость касается всего фазового портрета системы; адаптируемость обычно характеризуется асимптотическим поведением системы при $t > 0$.

2. Возмущения, рассматриваемые при анализе адаптируемости, обычно не содержат вариаций векторного поля v во всей окрестности $V(M)$. Обычно предполагается, что подмножество P от $V(M)$ задано таким образом, что $\Sigma \in P$ и допустимые вариации Σ будут также принадлежать P . Можно считать, что P описывается конечной системой параметров, содержащихся в Σ , и мы меняем Σ путем варьирования этих возмущений.

3. Представление о структурной устойчивости является слишком сильным требованием по отношению к адаптируемости, как это следует из указанных выше пп. 1 и 2. В то же время соответствующее представление об Ω -устойчивости, согласно которому налагается требование топологической совместимости только на неподвижные точки системы Σ , является слишком слабым. Последнее связано с тем, что Ω -устойчивость ничего не говорит о структурных изменениях границы области притяжения неподвижных точек.

Обсуждения, приведенные выше, дают возможность сформулировать вариант определения адаптируемой системы Σ .

Определение 5.8

Пусть непрерывная по времени система Σ описывается дифференциальным уравнением $\dot{x} = f(x)$ и P -подмножество C^1 -векторного поля на M , такое, что $f \in P$. Тогда система Σ называется адаптируемой, если

1. Существует окрестность U для f в C^1 -топологии, такая, что все системы Σ' , определенные векторными полями $f' \in \Pi \cap P$, имеют одинаковое (конечное) число аттракторов.

2. Для каждого аттрактора A_i системы Σ и каждой близкой к ней системы Σ' имеется конечное множество $a_i(\Sigma')$ аттракторов Σ' и отображения

$$\Sigma' \rightarrow Ua_i(\Sigma')$$

и

$$\Sigma' \rightarrow U\bar{B}_i(\Sigma')$$

непрерывны с C^1 -топологией на $U \cap P$ с хаусдорфовой метрикой на A_i и \bar{B}_i . Здесь \bar{B}_i — замыкание области притяжения для A_i .

Приведенное определение адаптируемости построено таким образом, что аттрактор Σ может «расщепиться» на несколько близко расположенных аттракторов Σ' и не нарушить при этом адаптируемости системы. Такой подход вполне оправдан, поскольку подобное расщепление не должно существенным образом менять асимптотического поведения системы.

Пример. Система

$$\dot{x} = \mu x - \varepsilon x^3, \quad \varepsilon > 0, \quad (5.10)$$

будет адаптируемой в указанном выше смысле по отношению к изменениям μ вблизи $\mu = 0$, даже если устойчивая фиксированная точка при $\mu < 0$ расщепится на одну неустойчивую и две устойчивые точки при $\mu = 0$. Важно отметить, что эти две притягивающие точки остаются близкими.

Согласно данному выше определению, адаптируемость является качественным свойством системы: система Σ будет либо адаптируемой, либо неадаптируемой по отношению к возмущениям, принадлежащим подмногообразию P . В зависимости от конкретной ситуации возникает несколько вариантов введения числовой характеристики, при помощи которой можно попытаться измерить степень адаптируемости Σ . Рассмотрим некоторые из них.

Минимальная адаптируемость. При таком подходе нас интересует диапазон возмущений, принадлежащих многообразию P , которые не вызывают качественных изменений в поведении системы Σ . По существу мы имеем дело с представлением об адаптируемости, обсуждавшимся в предыдущих разделах.

Одна из возможных математических формулировок минимальной адаптируемости следует из предположения, что задана метрика $d(\cdot, \cdot)$, определенная на «многообразии параметров» P , и вводится

$$S_P = \{\Sigma' \in P : \Sigma' \text{ не адаптируема в смысле определения (5.8)}\}.$$

Тогда минимальная адаптируемость может быть определена следующим образом:

$$R_{\min}(\Sigma) = d(\Sigma, S_P).$$

Итак, $R_{\min}(\Sigma)$ представляет собой расстояние от системы Σ до ближайшей (в смысле d) неадаптируемой системы Σ' в P .

Адаптируемость по скорости. В качестве меры адаптируемости, более тесно связанной с анализом чувствительности, можно рассмотреть «скорость» изменения границ областей

притяжения при возмущении системы Σ до Σ' внутри P . Очевидно, что высокая чувствительность границ области притяжения (границ «ям») не отвечает интуитивному представлению об адаптируемости, даже если система структурно устойчива, т. е. удовлетворяет условию, более сильному, чем адаптируемость.

Мера адаптируемости по скорости вводится следующим образом:

$$R_s(\Sigma) = \left\{ \lim_{h \rightarrow 0} \frac{1}{h} \sup_{\Sigma' \in P_h} \sup_i [d(A_i, A'_i), d(\bar{B}_i, \bar{B}'_i)] \right\}^{-1},$$

где P_h — шар радиуса h в P с центром в Σ , а величины A и B определены выше.

Отметим, что мера $R_s(\Sigma)$ может быть равна нулю даже в случае адаптируемой системы Σ , если положение аттракторов, или «ям», описывается недифференцируемой функцией параметров. Например, система (5.10) адаптируема, но для нее $R_s(\Sigma) = 0$.

Адаптируемость по объему. Размер области в пространстве состояний, соответствующей желательной области притяжения, также может служить в качестве меры адаптируемости, поскольку значительное уменьшение области притяжения приводит почти к таким же серьезным последствиям, как и ее полное исчезновение. Мера адаптируемости по объему вводится следующим образом:

$$R_v(\Sigma) = \lim_{h \rightarrow 0} \frac{1}{h} \sup_{\Sigma' \in P_h} |v(B) - v(B')|,$$

где B — желательная область притяжения, B' — соответствующая область притяжения для Σ' и $v(\cdot)$ — функция, измеряющая объем.

Заметим, что к определению адаптируемости по объему следует подходить с осторожностью, поскольку для моделей высокой размерности ($h > 2$) области притяжения («ям») часто имеют сложную структуру и могут иметь значительный объем, в то время как граница будет по-прежнему близкой к каждой точке области притяжения.

В заключение отметим, что по сравнению с теорией катастроф понятие адаптируемости является более широким, причем это расширение достигается, во-первых, благодаря учету аттракторов, которые имеют гораздо более сложную структуру, чем фиксированные точки, и, во-вторых, вследствие явного рассмотрения свойства областей притяжения. Перейдем теперь к анализу различных методов организации устойчивой или адаптируемой динамики системы, а также повышения степени ее устойчивости или адаптируемости.

Экологический пример адаптируемости

Приведенные выше рассуждения об адаптируемости и динамических процессах вряд ли можно сразу в полной мере оценить, поскольку они достаточно формальны. Как обычно, наиболее интересными оказываются конкретные задачи. В качестве одной из них рассмотрим важную для практики задачу моделирования экологической системы, включающей популяцию гусениц листовертки. Листовертка является насекомым-вредителем, периодически разрушающим леса на больших территориях северо-восточной части Северной Америки. Динамика всплеск численности листовертки представляет интерес с точки зрения системного подхода, поскольку наблюдаемые при этом резкие колебания и множественные временные масштабы характерны для моделей теории катастроф. Ниже описывается применение катастрофы типа сборки для характеристики некоторых аспектов адаптируемости популяции листовертки.

Динамика роста и отмирания листовертки может быть представлена следующей системой уравнений третьего порядка:

$$\frac{dB}{dt} = \alpha_1 B \left(1 - \frac{B(\alpha_3 + E^2)}{\alpha_2 S E^2} \right) - \frac{\alpha_4 B^2}{\alpha_5 S^2 + B^2},$$

$$\frac{dS}{dt} = \alpha_6 S \left(1 - \frac{\alpha_7 S}{\alpha_8 E} \right)$$

$$\frac{dE}{dt} = \alpha_9 E \left(1 - \frac{E}{\alpha_7} \right) - \alpha_{10} \frac{B E^2}{S(\alpha_3 + E^2)},$$

где $B(t)$ — плотность популяции листовертки, $S(t)$ — количество доступной листвы в лесу и $E(t)$ — переменная, характеризующая «энергетический резерв» леса, т. е. качественное состояние листьев и веток. Параметры от α_1 до α_2 включают различные константы рождаемости и смертности, скорость выедания листовертки хищниками и т. д. Искомая модель должна дать ответ на основной вопрос динамики численности: какая комбинация значений параметров вызывает резкий подъем плотности популяции листовертки с низкого равновесного уровня на высокий, или наоборот, при каких значениях параметров создается возможность снизить высокую плотность листовертки до низкого равновесного уровня.

Для ответа на поставленный вопрос рассмотрим равновесные уровни для B , S и E . Обозначим их величины через \bar{B} , \bar{S} и \bar{E} . После некоторых алгебраических преобразований

приведенных выше уравнений получаем, что величина равновесного уровня плотности листовертки \bar{B} удовлетворяет кубическому уравнению

$$- \alpha_1 (\alpha_3 + \bar{E}^2) \bar{B}^3 + \gamma \alpha_1 \alpha_2 \bar{E}^3 \bar{B}^2 - [\gamma^2 \alpha_1 \alpha_5 \bar{E}^2 (\alpha_3 + \bar{E}^2) + \gamma \alpha_2 \alpha_4 \bar{E}^3] \bar{B} + \gamma \alpha_1 \alpha_2 \alpha_5 \bar{E}^5 = 0$$

(полагаем $\gamma = \alpha_8 / \alpha_7$). Чтобы привести это уравнение к стандартному кубическому, необходимо исключить квадратный член. Это достигается путем замены зависимой переменной

$$\bar{B} \rightarrow y + \frac{\gamma \alpha_2 \bar{E}^3}{3(\alpha_3 + \bar{E}^2)}.$$

Получаем новое кубическое уравнение относительно y

$$y^3 + ay + b = 0,$$

где

$$a = \frac{[\gamma^2 \alpha_1 \alpha_5 \bar{E}^2 (\alpha_3 + \bar{E}^2) + \gamma \alpha_2 \alpha_4 \bar{E}^3 - \gamma^2 \alpha_1^2 \alpha_2^2 \bar{E}^6]}{\alpha_1 (\alpha_3 + \bar{E}^2)},$$

$$b = \frac{\left\{ \frac{-(\frac{2}{27}) \gamma^3 \alpha_1 \alpha_2^3 \bar{E}^9}{(\alpha_3 + \bar{E}^2)^2} + \gamma \alpha_2 \bar{E}^3 \frac{[\gamma^2 \alpha_1 \alpha_5 \bar{E}^2 (\alpha_3 + \bar{E}^2) + \gamma \alpha_2 \alpha_4 \bar{E}^3]}{3(\alpha_3 + \bar{E}^2)} - \gamma^3 \alpha_1 \alpha_2 \alpha_5 \bar{E}^5 \right\}}{\alpha_1 (\alpha_3 + \bar{E}^2)}$$

Поскольку равновесный уровень плотности листовертки \bar{B} связан с переменной y элементарными преобразованиями (см. выше), очевидно, что каноническая поверхность типа сборки, определяющая поведение y как функции a и b , определяет также изменение \bar{B} . Это дает нам возможность полностью представить поведение равновесного уровня \bar{B} на одной схеме (рис. 5.20). Поверхность сборки показывает, что никакие разрывы \bar{B} невозможны при $a \geq 0$, т. е. при

$$3[\gamma \alpha_1 \alpha_5 (\alpha_3 \bar{E}^2)^2 + \alpha_2 \alpha_4 \bar{E} (\alpha_3 + \bar{E}^2)] - 5 \gamma \alpha_1 \alpha_2^2 \bar{E}^4 \geq 0.$$

Например, при низком значении энергетического резерва леса ($\bar{E} \approx 0$) вспышки численности листовертки не будут иметь места, так как приведенное выше выражение всегда будет неотрицательным. Напротив, при высоких значениях \bar{E} ($\bar{E} \approx 1$) ни при каких комбинациях реальных значений параметров нельзя избежать возможных вспышек численности листовертки. Это, однако, не означает, что такая вспышка обязательно произойдет, поскольку для возможности пересечения критической ветви кривой сборки играет роль также величина переменной b .

Отметим важность приведенного выше примера, так как он позволяет нам убедиться, что при описании разрывов в развитии популяции листовертки решающее значение имеет комбинация параметров системы a и b . Представляется маловероятным, чтобы какие-либо дополнительные исследования или интуитивные представления относительно динамики численности листовертки могли подтвердить, что исчерпывающая информация обеспечивается только данной комбинацией параметров. Однако в нашем случае дело обстоит именно так. Заметим также, что, после того как параметры a и b

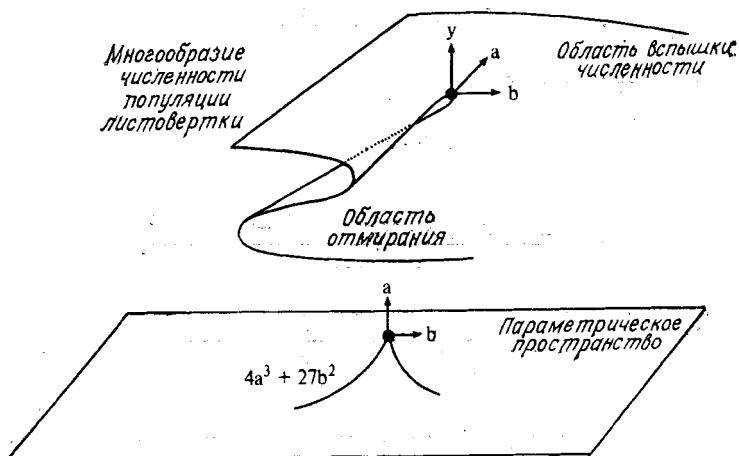


Рис. 5.20. Катастрофа типа сборки для модели развития популяции листовертки.

найлены, хорошо изученные свойства поверхности типа сборки дают возможность установить, что критические ветви в пространстве a — b , где могут возникнуть разрывы, удовлетворяют уравнению

$$4a^3 + 27b^2 = 0.$$

Таким образом, получаем конкретное алгебраическое соотношение, содержащее, хотя и в довольно сложной форме, все параметры системы. Оно дает нам полную картину изменения численности популяции листовертки в равновесном состоянии в зависимости от параметров системы.

Приведенный выше пример позволяет заключить, что параметры, имеющие физический смысл, и переменные, удобные для математического описания, как правило, совершенно различны. Причем успех исследования конкретной проблемы

часто определяется возможностью нахождения соответствующего перехода от физических параметров к математическим переменным. В рассмотренном примере успех был достигнут благодаря тому, что величина \bar{V} уже удовлетворяла кубическому уравнению, поэтому для получения удобной формы математического описания потребовался лишь тривиальный переход от \bar{V} к y . В более общих случаях могут оказаться необходимыми несколько более сложные преобразования. В этом заключается одна из трудностей на пути к успешному применению теории катастроф в системном анализе.

УСТОЙЧИВОСТЬ, УПРАВЛЕНИЕ И ОБРАТНАЯ СВЯЗЬ

Изложенные выше соображения относительно устойчивости представляют несомненный интерес, однако их рассмотрение отводит исследователю роль пассивного наблюдателя, поскольку не предусматривается возможности изменения нежелательного поведения системы путем выбора внешних условий на входе или управляющих воздействий. В принципе допущение возможности выбора внешних управляющих воздействий на модель системы означает переход исследователя с позиций пассивного наблюдения к активному вмешательству. Как отмечалось в гл. I, в философском и психологическом отношении такой переход означает более высокую ступень при нашем движении в направлении развития существенно новых подходов к анализу системы. В последующих разделах будет показано, что обеспечение возможности управления представляет собой качественный скачок в методологическом подходе к анализу системы и естественным образом приводит к одному из основных понятий современной теории систем — понятию *обратной связи*.

Для того чтобы ввести основное представление о стабилизации при помощи управления с обратной связью, рассмотрим следующее описание системы Σ :

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = c.$$

Здесь, как обычно, x означает вектор-функцию, описывающую состояние системы Σ , а управление $u(t)$ — вектор-функция, выбором которой можно распорядиться. В общем случае физические ограничения, а также ограничения ресурсов и т. п. приводят к тому, что допустимые функции управления $u(t)$ должны принадлежать некоторому множеству функций U , т. е. $u \in U$. Если предположить, что неуправляемая система (при $u(t) \equiv 0$) имеет поведение нежелательного

характера, возникает вопрос о возможности улучшения (в некотором смысле) траектории $x(t)$ путем применения управляющих воздействий из множества U . Очевидно, что математический подход к решению этого вопроса требует более точной его формулировки.

Наиболее традиционный подход к решению поставленной проблемы управления заключается в выяснении возможности стабилизации системы Σ при помощи управляющих воздействий из U . В общем случае предполагается, что положения равновесия неуправляемой системы не являются асимптотически устойчивыми в смысле Ляпунова, и предпринимается попытка стабилизировать систему путем использования управляющих функций $u(t)$. Простой пример линейной скалярной системы

$$\dot{x} = fx + u(t), \quad x(0) = c$$

при $f > 0$ показывает, что, вообще говоря, в данном случае стабилизация при помощи управляющих воздействий $u(t)$ невозможна. Действительно, из представления решения в форме

$$x(t) = ce^{ft} + \int_0^t e^{f(t-s)} u(s) ds$$

видно, что нельзя найти такую ограниченную функцию $u(t)$, которая смогла бы скомпенсировать влияние растущего экспоненциального члена ce^{ft} при всех допустимых значениях начального возмущения c .

В литературе закон управления в форме $u = u(t)$ называется законом управления типа *открытого контура*, поскольку управление не зависит от текущего состояния процесса $x(t)$, а является функцией только времени t . Одно из основных положений современной (т. е. созданной во второй половине 20-го века) теории управления состоит в том, что управление должно быть функцией состояния, т. е. структура закона управления должна иметь следующий вид:

$$u(t) = u(x(t), t).$$

Такие законы называются *законами управления с обратной связью*, поскольку при этом рассматривается состояние системы, информация о котором *поступает обратно* к исследователю, принимающему соответствующее решение на основании данных о поведении системы, характеризуемом состоянием $x(t)$. Эти два принципиально различных подхода иллюстрируются на рис. 5.21.

Для демонстрации глубоких изменений математического характера, которых можно достигнуть при помощи управления с обратной связью, вернемся к только что рассмотренному примеру скалярной линейной системы. Как было показано, никакая ограниченная функция, соответствующая закону управления типа открытого контура, не обеспечивает

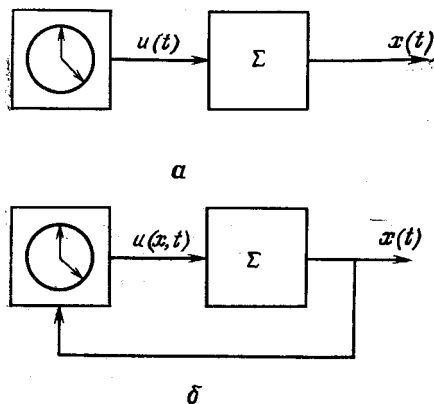


Рис. 5.21. Управление типа открытого (а) и замкнутого (б) контуров.

возможности добиться асимптотической устойчивости положения системы в начале координат, если $f > 0$. Рассмотрим теперь простой линейный закон обратной связи

$$u(x, t) = kx(t),$$

где k — постоянная, такая, что $k > f$. Используя этот закон, получим динамическое уравнение для замкнутого контура в виде

$$\dot{x} = (f - k)x(t), \quad x(0) = c,$$

из которого следует, что положение в начале координат асимптотически устойчиво для всех начальных возмущений c . Как будет показано ниже, этот вывод является частным случаем одного из фундаментальных результатов линейной теории систем — *теоремы о смещении полюсов*.

Теперь проанализируем различие между законами управления типа открытого и замкнутого контура с физической точки зрения. Отметим, что сопоставление этих двух типов законов управления можно провести следующим образом. Законы типа *открытого контура* представляют собой попытку при помощи *внешних* воздействий изменить поведение системы при сохранении соотношения между состояниями не-

управляемой системы, т. е. без изменения связей между переменными состояниями. Законы типа замкнутого контура, или обратной связи, меняют поведение системы в результате фактического изменения самой динамики системы $f(\cdot, \cdot)$. Следовательно, законы управления с обратной связью видоизменяют соотношения между переменными состояниями и в результате изменяют траекторию $x(t)$ системы путем фактического изменения ее топологии.

УСТОЙЧИВОСТЬ ПО ЛЯПУНОВУ И СМЕЩЕНИЕ ПОЛЮСОВ

Простой пример, приведенный в предыдущем разделе, показывает, что в случае скалярных линейных систем можно произвольным образом изменить характеристическое значение (и, следовательно, свойства устойчивости) системы путем использования закона управления типа линейной обратной связи. Рассмотрим теперь возможность распространения этого результата на многомерные системы.

Предположим, что система описывается линейными дифференциальными уравнениями

$$\dot{x} = Fx + Gu, \quad x(0) = c,$$

где x — n -мерный вектор состояния, u — m -мерный вектор управления, F и G — постоянные матрицы размерами $n \times n$ и $n \times m$ соответственно. Применение закона управления с обратной связью

$$u(t) = -Kx(t),$$

где K — постоянная матрица размером $m \times n$, приводит, очевидно, к системе типа замкнутого контура

$$\dot{x} = (F - GK)x, \quad x(0) = c$$

и задача сводится к поиску ответа на следующий вопрос: всегда ли при заданных F и G можно найти постоянную матрицу K , такую, чтобы собственные значения матрицы $F - GK$ были расположены в заранее выбранных точках комплексной плоскости? Весьма примечательно, что ответ на этот вопрос оказывается положительным при достаточно слабых предположениях относительно матриц F и G . Основной результат содержится в следующей теореме.

Теорема о смещении полюсов

Пусть пара матриц (F, G) полностью достижима, т. е. $(n \times m)$ -матрица

$$\mathcal{S} = [G \mid FG \mid \dots \mid F^{n-1}G]$$

имеет ранг n . Тогда для заданного произвольного множества комплексных чисел $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ всегда можно найти постоянную матрицу K , такую, что собственные значения $F - GK$ совпадают с множеством Λ .

Замечания. 1. Условие достижимости (F, G) является достаточно общим свойством, так как ему удовлетворяют «почти все» линейные системы.

2. Если множество Λ симметрично, т. е. из условия $\lambda \in \Lambda$ следует $\bar{\lambda} \in \Lambda$, то матрицу K можно выбирать действительной. В общем случае матрица K должна быть комплексной.

3. Термин *смещение полюсов* заимствован из технической литературы, где собственные значения матрицы F часто интерпретируются как «полюса» рациональной передаточной матрицы-функции системы Σ . Теорема утверждает, что если система обладает свойством достижимости, то расположение этих полюсов может быть произвольным образом изменено при помощи линейной обратной связи.

Трудно переоценить важность теоремы о смещении полюсов для практики, поскольку она позволяет обеспечить существенную гибкость в конструировании системы.

Проектировщику не надо беспокоиться о расчете определенных характеристик устойчивости системы, так как любые нестабильности в ее поведении могут быть устранены в соответствии с подходящим законом управления с обратной связью.

Пример. Равновесие в групповых взаимодействиях

Рассмотрим формальную модель социальных взаимодействий Хоманса в математической постановке Симона. Модель системы содержит четыре переменные:

интенсивность взаимодействия (или коммуникаций) между членами рассматриваемой группы, $T(t)$;

степень дружелюбия (или групповая идентификация) членов группы;

общая степень активности членов группы, $W(t)$;

степень активности воздействия внешней среды на группу (необходимая для выживания группы), $P(t)$.

Предположим, что переменные I, W, T представляют собой отклонения характеристик от некоторого желаемого идеального уровня ($I = W = T = 0$) и система, подверженная некоторым возмущениям, отклонена от этого уровня. Цель нашего исследования — показать возможность стабилизации процесса взаимодействия путем изменения внешней среды.

Переводя словесные постулаты Хоманса о взаимосвязях указанных переменных на язык дифференциальных уравнений, получим следующую математическую модель процесса взаимодействия:

$$\begin{aligned}\frac{dI}{dt} &= b(T - \beta I), \\ \frac{dW}{dt} &= c_1(I - \gamma W) + c_2(P - W), \\ T &= a_1 I + a_2 W.\end{aligned}$$

Параметры a_1 , a_2 , c_1 , c_2 , γ , β и b представляют величины различных сил взаимодействия и константы пропорциональности. Первое уравнение можно интерпретировать следующим образом: степень симпатии возрастает или убывает в зависимости от разности между активностью взаимодействий и существующим уровнем симпатии. Аналогично могут быть интерпретированы и другие уравнения.

Проводя несколько простых алгебраических преобразований, можно свести систему уравнений к стандартному виду

$$\dot{x} = Fx + Gu,$$

где

$$F = \begin{bmatrix} b(a_1 - \beta) & ba_2 \\ c_1 & -(c_1\gamma + c_2) \end{bmatrix}, \quad G = \begin{pmatrix} 0 \\ c_2 \end{pmatrix},$$

$$x = \begin{pmatrix} I \\ W \end{pmatrix}, \quad u = P.$$

Для определенности будем считать, что

$$c_1\gamma + c_2 < b(a_1 - \beta),$$

т. е. характеристические корни матрицы F неустойчивы. Теперь необходимо установить, можно ли найти линейную обратную связь, зависящую от наблюдаемых величин степени дружелюбия и групповой активности, которая сместила бы корни характеристического многочлена F в левую полуплоскость. Другими словами, необходимо определить закон

$$P(t) = -(k_1 I + k_2 W),$$

по которому стабилизируется система Σ . Очевидно, что эта задача представляет собой слабую версию задачи смещения полюсов, где множество Λ является подмножеством левой полуплоскости.

Сначала проверим реализуемость системы Σ . Легко показать, что матрицей разрешимости \mathcal{E} является матрица вида

$$\mathcal{E} = \begin{bmatrix} 0 & ba_2c_2 \\ c_2 & -c_2(c_1\gamma + c_2) \end{bmatrix},$$

имеющая ранг 2 тогда и только тогда, когда

$$ba_2c_2 \neq 0.$$

Согласно теореме о смещении полюсов, если $ba_2(c_2 \pm 0)$, то нет математических трудностей для нахождения закона обратной связи $P(t)$, по которому система стабилизируется до желаемой степени. Если система Σ устойчива, то можно использовать закон $P(t)$ для повышения степени устойчивости, смещая для этого корни F еще дальше в левую полуплоскость.

Основным недостатком теоремы о смещении полюсов является ее применимость только к линейным системам. Однако диапазон ее применимости может быть значительно расширен при помощи теоремы об устойчивости Пуанкаре — Ляпунова. Можно показать, что если система описывается дифференциальными уравнениями

$$\dot{x} = Fx + h(x), \quad x(0) = c,$$

где норма $\|c\|$ достаточно мала, а $\|h\|/\|x\| \rightarrow 0$ при $\|x\| \rightarrow 0$, то асимптотическая устойчивость системы определяется ее линейной частью, т. е. матрицей F . Таким образом, вводя в систему слагаемое, соответствующее управлению, получим систему

$$\dot{x} = Fx + h(x) + Gu, \quad x(0) = c,$$

к которой применима теорема о смещении полюсов. Следовательно, можно сместить корни F в нужную точку плоскости, гарантирующую устойчивость системы. Условие на $\|c\|$, зависящее от корня характеристического многочлена F с наибольшей вещественной частью, может быть значительно ослаблено за счет смещения этого корня еще дальше в левую полуплоскость. Связь этого результата с некоторыми рассмотренными выше понятиями адаптируемости очевидна.

УПРАВЛЕНИЕ БИФУРКАЦИЕЙ

Рассмотрим задачу нахождения такого закона управления с обратной связью, что независимо от значений вектора параметров a нелинейная система

$$\dot{x} = f(x, u, a), \quad x(0) = c$$

не имеет точек бифуркации. Другими словами, требуется найти закон управления

$$u(t) = u(x(t)),$$

такой, что динамика замкнутого контура

$$\dot{x} = f(x, u(x), a)$$

не имеет точек бифуркации ни при каком значении a .

В общем случае эта проблема не решена, поэтому сосредоточим внимание на достаточных условиях существования таких «свободных от бифуркации» законов управления. Эти условия обычно являются почти тривиальными следствиями следующего варианта общей теоремы о неявной функции. *Общая теорема о неявной функции.*

Пусть

$$f: E^n \times E^m \times E^k \rightarrow E^n$$

— непрерывно дифференцируемое отображение. Тогда существует единственное непрерывно дифференцируемое отображение

$$g: E^n \times E^m \times E^k \rightarrow E^n,$$

такое, что $g(y, u, a) = x$ для всех $x, y \in E^n, u \in E^m, a \in E^k$, удовлетворяющих $f(x, u, a) = y$, если выполняются условия:

- 1) $\det \{df/dx\} \neq 0$ для любого $(x, u, a) \in E^n \times E^m \times E^k$,
- 2) для любых $u \in E^m, a \in E^k$

$$\|f(x, u, a)\| \rightarrow \infty \quad \text{при} \quad \|x\| \rightarrow \infty.$$

По существу, теорема о неявной функции дает достаточные условия, при выполнении которых уравнение

$$f(x, u, a) = y$$

имеет единственное решение для (x, u, a, y) во всей области $E^n \times E^m + E^k \times E^n$. А именно, если матрица Якоби функции f не вырождена в интересующей нас области и f удовлетворяет условиям роста (2), то f глобально обратима.

Чтобы применить общую теорему о неявной функции к проблеме управления бифуркацией, достаточно заметить, что по предположению f имеет в начале координат точку равновесия. Кроме того, если существует закон управления $u(x)$, такой, что $[df/dx]$ не вырождена во всей области (x, u, a) и $f(x, u(x), a)$ удовлетворяет условиям роста (2) для всех a , то это означает, что у функции нет других точек равновесия и, следовательно, нет точек бифуркации. Иными словами, такой закон управления $u(x)$ гарантирует, что никакие возмущения

a не могут привести к бифуркации точки равновесия в начале координат и дать качественно другой тип поведения. Приведенный анализ, несмотря на свою простоту, может быть положен в основу успешного применения техники линеаризации к управлению проектирования системы.

Более сложная проблема — проектирование таких регуляторов, которые не дают характеристическим корням линеаризованной системы пересекать мнимую ось. Однако для канонических моделей катастроф эта проблема легко разрешима, так как поведение подобных моделей хорошо изучено, например поведение канонической модели сборки

$$\dot{x} = -(x^3 + \alpha_1 x + \alpha_2).$$

Закон обратной отрицательной связи для переменной управления α_2 имеет вид

$$\alpha_2 = -(K_1 x + K_0),$$

где K_1, K_2 — постоянные. Такой закон обратной связи приводит к замене $\alpha_1 \rightarrow \alpha_1 - K_1$. Выбирая K_1 так, чтобы иметь $\alpha_1 > 0$, можно добиться отсутствия резких «скачков» из области одного аттрактора в область другого аттрактора. Однако в случае неканонических моделей, когда используются физические переменные, а не канонические координаты, управляющие параметры часто являются нелинейными функциями физических переменных, и поэтому закон обратной связи для x вызывает изменение более чем одного параметра канонической модели. В таких ситуациях построение свободных от бифуркации законов управления является более сложной задачей, хотя в этом случае может быть использован наш общий метод.

УПРАВЛЯЕМАЯ АДАПТИРУЕМОСТЬ

В предыдущих разделах были даны различные определения адаптируемости динамических процессов. Однако все они требовали от адаптируемой системы способности сохранять свое поведение, независимо от воздействия различного вида возмущений. Поскольку наша основная цель — анализ устойчивости, то будет исследоваться вопрос глобальной асимптотической устойчивости (в смысле Ляпунова) систем, в которых допускается неопределенность в динамике, в параметрах системы и, возможно, в самих управляющих воздействиях. Мы не будем, как обычно, предполагать, что *a priori* заданы некоторые статистические свойства неопределенностей, а потребуем только, чтобы неопределенности были ограничены. Наша цель — научиться находить такие

законы управления с обратной связью, которые гарантировали бы глобальную асимптотическую устойчивость при любых ограниченных возмущениях системы. Очевидно, что если рассматривать асимптотическую устойчивость как свойство поведения системы, которое должно сохраняться вопреки внешним воздействиям, то систему, управляемую законом обратной связи подобного типа, действительно можно считать адаптируемой! Такой подход к адаптируемости — это другое выражение смысла замечаний, сделанных выше, а именно не обязательно рассчитывать устойчивость системы в начале проектирования, так как желаемые свойства устойчивости обычно можно достигнуть введением подходящих контуров обратной связи. Для успешного разрешения поставленных проблем в общем случае требуются значительные дополнительные сведения, поэтому рассмотрим только линейный случай. Основные результаты верны и для нелинейных систем, но их понимание требует значительной математической подготовки.

Итак, рассмотрим систему, описываемую линейными дифференциальными уравнениями

$$\dot{x} = Ax + Bu + Cv, \quad x(0) = c,$$

где u — m -мерный вектор управлений, v — l -мерный вектор возмущений, а A , B и C — постоянные матрицы нужных размеров. Кроме того, предположим, что матрица A устойчива, а для элементов A и B как возмущаемых известны только верхние и нижние границы их значений. Наша задача состоит в том, чтобы найти такой закон управления обратной связью $u = u(x)$, чтобы система была глобально асимптотически устойчивой при всех ограниченных возмущениях v и всех допустимых матрицах A и B .

Принимая разумные, но технически сложные предположения о неопределенностях в A и B и возмущениях v , можно показать, что i -я компонента закона стабилизирующей нелинейной обратной связи задается равенством

$$u_i(x) = \begin{cases} -\rho_i(x) \operatorname{sgn} [(b_i, Px)], & x \notin N_i, \\ \varepsilon \{y_i \in R : |y_i| \leq \rho_i(x)\}, & x \in N_i. \end{cases}$$

Здесь P является решением матричного уравнения Ляпунова

$$PA + A'P + Q = 0, \quad Q > 0 \text{ произвольно,}$$

b_i — i -й столбец матрицы B , $\rho_i(x)$ — функция, которая зависит от границ неопределенностей в A , B и возмущениях

v , (\cdot, \cdot) обозначает внутреннее произведение. Множество N_i определяется как

$$N_i = \{x \in R^n : (b_i, P_x) = 0\}.$$

Таким образом, используя закон обратной связи $u(x)$, линейную систему можно сделать асимптотически устойчивой при любых ограниченных по величине неопределенностях в матрицах A , B и ограниченных возмущениях v . Более детальное рассмотрение этого результата вместе с информацией о методе вычисления функций ρ_i можно найти в статьях, ссылки на которые даны в конце главы.

ЛИТЕРАТУРА

Исторически первые математические работы по изучению устойчивости систем были стимулированы развитием небесной механики в семнадцатом столетии. Проблема стабилизации механических систем посредством обратной связи относится ко времени исследований Максвелла об управлении паровым двигателем Уатта. Некоторые из этих основополагающих работ приведены в сборниках Bellman R., Kalaba R., eds., *Mathematical Trends in Control Theory*, Dover, New York, 1964; Aggarwal J., Vidyasagar M., eds., *Nonlinear Systems: Stability Analysis*, Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, 1977.

Внешнее описание

Основные результаты для нелинейных систем, представленных внешним описанием, появились в работах Сандберга и Зеймса, которые при изучении свойств устойчивости использовали методы функционального анализа:

Sandberg I. W., On the L_2 -Boundedness of Solutions of Nonlinear Functional Equations, *Bell Syst. Tech. J.*, **43**, 1581—1599 (1964);
Zames G., On the Input/Output Stability of Time-Varying Nonlinear Feedback Systems — I, II, *IEEE Trans. Autom. Control*, AC-11, 228—238, 465—476 (1966).

Внутреннее описание

Классическая теория устойчивости по Ляпунову подробно изложена в следующих книгах:

Hahn W., *Stability of Motion*, Springer, New York, 1967.
Lasalle J., Lefschetz S., *Stability by Lyapunov's Direct Method with Applications*, Academic, New York, 1961.
[Имеется перевод: Ла-Салль Ж., Лефшец С. Исследования устойчивости прямым методом Ляпунова. — М.: Мир, 1964.]
Casti J., *Dynamical Systems and Their Applications: Linear Theory*, Academic, New York, 1977, Chapter 7;
Арнольд В. И. Обыкновенные дифференциальные уравнения. — М.: Наука, 1975.

Структурная устойчивость

Одной из лучших вводных работ по качественной теории устойчивости является книга

Hirsch M., Smale S., *Differential Equations, Dynamical Systems and Linear Algebra*, Academic, New York, 1974.

Связная устойчивость и адаптируемость

Концепция связной устойчивости в различных ее контекстах интенсивно изучалась Д. Сильжаком. Обзор результатов этих работ можно найти в статьях

Siljak D., *Stability of Large-Scale Systems under Structural Perturbations*, *IEEE Trans. Syst. Man Cybern.*, SMC-2, 657—663 (1972);

Connective Stability of Competitive Equilibrium, *Automatica*, 11, 389—400 (1975).

Понятие связной устойчивости первоначально возникло в связи с вопросами равновесия в экономике. См. работы

Quirk J., Saposnik R., *Introduction to General Equilibrium Theory and Welfare Economics*, McGraw-Hill, New York, 1968;

Arrow K., Hahn F., *General Competitive Analysis*, Holden-Day, San Francisco, 1971;

Newman P., *Some Notes on Stability Conditions*, *Rev. Econ. Stud.*, 72, 1—9 (1959).

Графы и процессы распространения возмущений

Процессы распространения возмущений в графах и возникающие при этом представления об устойчивости изложены в книге

Roberts F., *Discrete Mathematical Models*, Prentice-Hall, Englewood Cliffs, New Jersey, 1976.

Интересное применение этого аппарата к проблемам энергетики можно найти в работах

Roberts F., Brown T., *Signed Digraphs and the Energy Crisis*, *Am. Math. Monthly*, 82, 577—594 (1975);

Building and Analyzing an Energy Demand Signed Digraph, *Environ. Planning*, 5, 199—221 (1973).

Устойчивость системы «черный ящик» с обратной связью

Теорема о малом коэффициенте усиления и теорема инертности приведены в статье

Zames G., Falb P., *Stability Conditions for Systems with Monotone and Slope-Restricted Nonlinearities*, *SIAM J. Appl. Math.*, 6, 89—108 (1968).

При некоторых условиях эти две теоремы эквивалентны. Эти условия рассматриваются в работе

Anderson B. D. O., *The Small-Gain Theorem, the Passivity Theorem and their Equivalence*, *J. Franklin Inst.*, 293, 105—115 (1972).

Внутренние модели и устойчивость

Изложение теории Ляпунова, ее многочисленные приложения и различные обобщения можно найти в классической работе

Kalman R., Bertram J., *Control System Analysis and Design via the Second Method of Lyapunov — I, II*, *J. Basic Eng. Trans. ASME*, 82, 371—400.

См. также работы, приведенные в разделе «Внутреннее описание», и книгу
Bellman R., *Stability Theory of Differential Equations*, McGraw-Hill, New York, 1953.

Связная устойчивость

См. работы Сильжака, приведенные в разделе «Связная устойчивость и адаптируемость».

Пример из экологии взят из работы

Casti J., *Connectivity and Stability in Ecological and Energy Systems*, WP-75-150, International Institute for Applied Systems Analysis, Laxenburg, Austria, November 1975.

Бифуркация Хопфа

Обзор по современному состоянию вопросов, относящихся к бифуркации Хопфа, включая перевод статьи Хопфа, приведен в книге Marsden J., McCracken M., *The Hopf Bifurcation and its Applications*, Springer, New York, 1976. [Имеется перевод: Марсден Дж., Мак-Кракен М. Бифуркация рождения цикла и ее приложения. — М.: Мир, 1980.]

Структурно устойчивые системы

Введение в проблемы, связанные со структурной устойчивостью, дано в статье

Peixoto M., *Generic Properties of Ordinary Differential Equations*, in *Studies in Ordinary Differential Equations*, J. Hale, ed., Mathematical Association of America, Washington, D. C., 1977.

Основополагающей работой является

Андронов А. А., Понтрягин Л. С., *Большие системы*. — ДАН СССР, 1937, т. 14.

Обзор современных работ дан в книге

Peixoto M., ed., *Dynamical Systems*, Academic, New York, 1973.

Теория катастроф

Классической книгой по теории катастроф является

Thom R., *Structural Stability and Morphogenesis*, Addison-Wesley, Reading, Massachusetts, 1975.

Основные работы по приложению теории катастроф представлены в книгах

Zeeman E. C., *Catastrophe Theory: Selected Papers 1972—77*, Addison-Wesley, Reading, Massachusetts, 1977;

Poston T., Stewart I., *Catastrophe Theory and Its Applications*, Pitman, London, 1978. [Имеется перевод: Постон Тн и Стюарт Иен. Теория катастроф и ее приложения. — М.: Мир, 1980.]

Современное введение в теорию катастроф и ее приложение изложено в статье

Golubitsky M., *An Introduction to Catastrophe Theory and its Applications*, *SIAM Rev.*, 20, 352—387 (1978).

Критика общего метода теории катастроф дана в статье

Sussman H., Zahler R., *Catastrophe Theory As Applied to the Social and Biological Sciences: A Critique*, *Synthese*, 37, 117—216 (1978).

Пример использования методов теории катастроф

Пример с загрязнением озера взят из статьи
 Duckstein L., Casti J., Kempf J., A Model of Phytoplankton Dieoff in
 Small Eutrophic Ponds Using Catastrophe Theory, *J. Water Resour. Res.*
 Данные к этому примеру взяты из работы
 Parks R. W., et al., Phytoplankton and Water Quality in a Fertilized Fish-
 pond, Circular 224, Agricultural Experiment Station, Auburn University,
 Auburn, Alabama, 1975.

Устойчивость по возмущению и устойчивость по начальному значению

Детальное изложение рассмотренных в этом разделе вопросов при-
 ведено в книге
 Roberts F., *Discrete Mathematical Models*, Prentice-Hall, Englewood Cliffs,
 New Jersey, 1976.

Адаптируемость динамических процессов

Качественный разбор понятия адаптируемости, возникающего в эко-
 логии, приведен в статье
 Holling C. S., Resilience and Stability of Ecological Systems, *Ann. Rev.*
Ecol. Syst., 4, 1—23 (1973).

Адаптируемость и катастрофы

Модель биржевых операций рассмотрена в статье
 Zeeman E. C., On the Unstable Behavior of Stock Exchanges, *J. Math.*
Econ., 1, 39—49 (1974).

Системы Морса—Смейла и адаптируемость

Введение в системы Морса—Смейла и градиентную динамику изло-
 жено в книгах
 Hirsch M., Smale S., *Differential Equations, Dynamical Systems and Linear*
Algebra, Academic, New York, 1974;
 Walters P., An Outline of Structural Stability Theory, in *Analysis, and*
Computation of Equilibria and Regions of Stability, H. Grüm m, ed., CP-75-8,
 International Institute for Applied Systems Analysis, Laxenburg, Austria,
 1975;
 Crüm m H., *Definitions of Resilience*, RR-76-5, International Institute for
 Applied Systems Analysis, Laxenburg, Austria, 1976.

Экологический пример адаптируемости

Модель развития популяции листовёртки взята из статьи
 Ludwig D., Jones D., Holling C., Qualitative Analysis of Insect Outbreak
 Systems: The Spruce Budworm and the Forest, *J. Anim. Ecol.*, 47, 315—332
 (1978).

Устойчивость по Ляпунову и смещение полюсов

Более подробно вопросы управления с обратной связью и устойчи-
 вости изложены в книге
 Casti J., *Dynamical Systems and Their Applications*, Academic, New York,
 1977.

Доказательство теоремы о смещении полюсов можно найти в только что приведенной книге, а также в оригинальной статье Wonham W. M., On Pole Assignment in Multi-input Controllable Linear Systems, *IEEE Trans. Autom. Control.*, **AC-12**, 660—665 (1967).

Пример группового взаимодействия взят из статьи Simon H., The Construction of Social Science Models, in *Mathematics and Psychology*, Miller G., ed., Wiley, New York, 1964.

Управление бифуркацией

Более подробно вопрос о взаимосвязи традиционной теории управления и качественной теории устойчивости рассматривается в статье Mehra R., Catastrophe Theory, Nonlinear System Identification and Bifurcation Control, Joint Automatic Control Conference, San Francisco, June 1977.

Управляемая адаптируемость

Результаты взяты из работ

Leitmann G., Guaranteed Asymptotic Stability for Some Linear Systems with Bounded Uncertainties, *J. Dynam. Syst. Meas. Control*;

Leitmann G., Wan H., Macroeconomic Stabilization Policy for an Uncertain Dynamic Economy, in *New Trends in Dynamic System Theory and Economics*, Springer, Vienna, 1977;

Gutman S., Leitmann G., Stabilizing Feedback Control for Dynamical Systems with Bounded Uncertainty, Proceedings of the IEEE Conference on Decision and Control, Florida, 1976.

Вопрос о нахождении «наилучшей» оценки состояния системы по наблюдаемым данным может быть рассмотрен с различных точек зрения; это зависит от предположений, касающихся данных. Линейные системы изучаются в следующих работах:

Luenberger D., Observing the State of a Linear System, *IEEE Trans. Mill. Electron.*, **MIL-8**, 74—80 (1964);

Luenberger D., Observers for Multivariable Systems, *IEEE Trans. Autom. Control*, **AC-11**, 190—197 (1966);

Kalman R., Bucy R., New Results in Linear Prediction and Filtering, *J. Basic Eng., Trans. ASME*, **93D**, 95—100 (1961).

Вопрос о количестве измерений при нахождении стабилизирующей обратной связи изучен в работах

Casti J., Letov A., Minimal Control Fields, *J. Math. Anal. Appl.*, **43**, 15—25 (1973); Minimal Control Fields and Pole-Shifting by Linear Feedback, *Appl. Math. Comp.*, **2**, 19—28 (1976).

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Адаптируемость 63, 135, 140, 183, 187
— динамических процессов 179
— минимальная 192
— по объему 193
— по скорости 192
— управляемая 205
Аксиомы системной сложности 117
Алгебраическая теория систем 127
q-Анализ 72
- Бифуркационная теорема Хопфа** 157
Бифуркация 168
— *Хопфа* 157
«Бихевиористская» школа 20
Ближние векторные поля 188
Большая система 70
Борьба с насекомыми-вредителями 177
- Вершина** 29
Взвешенный оргграф 143
Внешнее воздействие (вход) 19
Внутренние модели 21, 117
- Генетическая модель Джекоба** — *Моно* 57
Гиперболическая замкнутая траектория 162
— особенность 161
Глобальные свойства системы 44
Глобальный минимум 24
Гомотопия 52
Графы 47, 141
Граница 75
Граничный цикл 76
Группа гомологий 75
— отображений 82
- Двоичный выбор** 15
Декомпозиция аналитических систем 104
- Динамика водохранилищ 13
— локальная 23
— одномерная 26
— полиэдральная 125
— стационарная 27
Дыры 74, 80
- Загрязнение озер 170
Задача идентификации параметров 37
— максимизации 42
— наблюдаемости 41
— о листовертке 194
— — часовщиках 110
— реализации 37
Закон *Бернулли* 115
— управления с обратной связью 198
- Идентификация** 36
Иерархические системы 88
Иерархия 110, 118
- Классическая система Лотки** — *Вольтерра* 163
Канонические модели катастроф 205
Катастрофа типа сборки 64, 174, 184
— — складки 167
Катастрофы 63, 135, 183
Комплекс 71
— симплициальный 29, 50
Конечномерная реализация 37
Конечномерность пространства состояний 21
Координатизация 23
Кограницы 82
Коцепи 82
Критерий *Попова* 146
Критерий подобия 130
- Макроэкономика** 16
Математическое описание 16
— — внешнее 18, 136
— — внутреннее 17, 137

- Матрица взаимосвязи 140
 — инцидентий 29
 — передаточная 96
 — смежности 49
 Мера необычности 74
 Многообразие 110
 Множества 28
 Модель биржевых операций 185
- Негэнтропия 25
 Нелинейные конечные процессы 128
 Нормализация 119
- Область притяжения 61, 138
 Образ 31, 52
 Обратное преобразование 36
 Ограничения 39
 — внешние 39
 — внутренние 39
 Оптимизация 43
 Особая точка дифференциального уравнения 161
 Открытый контур 198
 Отображение 167
- Парадокс брадобрея 90
 Подобные фазовые портреты 188
 «Познавательная» школа 21
 Покрытие 89
 Полугруппа преобразований 101
 — триггера 103
 Полугруппы 100
 Потенциальная функция 23, 166
 Примарная группа 103
 Принцип Гамильтона — Якоби 45
 — необходимого многообразия Эш-би 112
 — эволюции 25
 Препятствия 74
- Равновесие в групповых взаимодействиях 201
 Разбиение множества 89
 Реакция (выход) 19
- Связность 47, 50, 70
 — алгебраическая 95
 — топологическая 89
 Связь 71
 — обратная 197
 — — отрицательная 142
 — — положительная 142
 Силовое поле 15
- Симплекс 29
 — пустой 29
 Система 11, 70
 — Аносова 165
 — Метцлера 150
 — Морса — Смейла 165
 — с конечным числом состояний 21
 — хищник — жертва 14, 62, 84, 126
 Системы 35
 — динамические неуправляемые 35
 — — структурно устойчивые 160
 — — управляемые 35
 — линейные 96
 — нелинейные 100
 — стохастические 42
 Сложность 53, 108, 123, 125, 127, 129
 — автоматов 119
 — внутренняя 123
 — вычислительная 117
 — динамическая 57, 122
 — структурная 54, 109
 — управления 123
 — эволюционная 59, 139, 160
 Смещение полюсов 200
 Соединение с обратной связью 118
 — параллельное 118
 — последовательное 118
 Сопряженное отношение 29
 Схема связности 111
- Теорема декомпозиции Крона — Рудза 102
 — инертности 146
 — Н. Н. Красовского 152
 — Кренера 105
 — Ляпунова об устойчивости 149
 — об устойчивости по возмущению 177
 — — — — начальному значению 177
 — общая о неявной функции 204
 — о малом коэффициенте усиления 145
 — — распространении возмущения 175
 — — связной устойчивости 154
 — — смещении полюсов 199, 200
 — — структурной устойчивости (в круге) 162
 — Пуанкаре — Ляпунова 151
 — реализации 97
 — Тома — Зимана 167
 Теория информации 129
 — катастроф 165
 — «центрального» места 64

- Точка фибуркации *Хопфа* 139
— катастрофы 168
Триангуляция 101
Трофические структуры 49
- Узел** 60, 61
Узловое произведение координат 101
Управление 197
— бифуркацией 203
Уравнения *Ван дер Поля* 189
— *Гамильтона — Якоби*
— логистические 38, 174
— *Льенара* 159
— *Эйлера — Лагранжа*
Уровни взаимодействия 112
Устойчивость 58, 135, 147, 197
— асимптотическая 137
— классическая 58
— по возмущению 173
— — *Ляпунову* 200
— — начальному значению 173
— структурная 59, 139, 160
- Фазовый портрет** 61
Фармакокинетика 40
Фокус 60, 61
Функция бесполезности 15
— полезности 15
- Цепь** 75
p-Цепь 75
Цепь связи 71
- Число Бетти** 72, 79, 81, 88
- Шахматы** 30, 90
Шакалы времени 113
- Эволюционные структуры** 122
Экологические ниши 49
Эксцентриситет 52, 74
Энтропия 23

ОГЛАВЛЕНИЕ

Предисловие редакторов перевода	5
Предисловие	7
1. Основные понятия и методы системного анализа	
Понятие системы	11
Математическое описание	16
Литература	33
2. Основные положения и перспективы развития теории систем	
Управляемые и неуправляемые динамические системы	35
Идентификация	36
Ограничения	39
Стохастические эффекты	42
Оптимизация	43
Глобальные свойства системы	44
Связность и графы	47
Связность и симплициальные комплексы	50
Сложность	53
Устойчивость	58
Катастрофы и адаптируемость	63
Литература	67
3. Связность структуры больших систем	
Комплексы и связи	71
Эксцентриситет	74
Дыры и препятствия	74
r -дыры	80
Коцепи и кограницы	82
Система хищник — жертва	84
Иерархические системы и покрытия	88
Применение понятия q -связности при анализе игры в шахматы и комедии Шекспира «Сон в летнюю ночь»	90
Алгебраическая связность	95
Линейные системы	96
Нелинейные системы	100
Полугруппы и узловые произведения	100
Теоремы декомпозиции Крона — Роудза	102
Декомпозиция аналитических систем	104
Литература	106
4. Сложность структуры больших систем	
Структурная сложность	109
Динамическая сложность	114

Вычислительная сложность	117
Аксиомы системной сложности	118
Сложность автоматов	119
Эволюционная сложность и эволюционные структуры	122
Проблемы выбора и сложность	123
Внутренняя сложность в сравнении со сложностью управления	123
Программа для практического применения понятия сложности	125
Полиэдральная динамика и сложность	125
Алгебраическая теория систем и сложность	127
Нелинейные конечномерные процессы	128
Сложность и теория информации	129
Литература	131
5. Устойчивость, катастрофы и адаптируемость больших систем	
Внешнее описание	136
Внутреннее описание	137
Структурная устойчивость	139
Связная устойчивость и адаптируемость	140
Графы и процессы распространения возмущений	141
Устойчивость системы «черный ящик» с обратной связью	144
Внутренние модели и устойчивость	147
Связная устойчивость	153
Бифуркация Хопфа	157
Структурно устойчивая динамика	160
Теория катастроф	165
Пример использования методов теории катастроф. Загрязнение озер	170
Катастрофа типа сборки и логистическое уравнение	174
Устойчивость по возмущению и по начальному значению	175
Адаптируемость динамических процессов	179
Адаптируемость и катастрофы	183
Системы Морса — Смейла и адаптируемость	187
Устойчивость, управление и обратная связь	197
Устойчивость по Ляпунову и смещение полюсов	200
Управление бифуркацией	203
Управляемая адаптируемость	205
Литература	207
Предметный указатель	212

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-й Рижский пер., д. 2, изд-во «Мир».