

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ПУЩИНСКИЙ НАУЧНЫЙ ЦЕНТР
ИНСТИТУТ МАТЕМАТИЧЕСКИХ ПРОБЛЕМ БИОЛОГИИ

А. М. МОЛЧАНОВ

НЕЛИНЕЙНОСТИ В БИОЛОГИИ

ПУЩИНО • 1992

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ПУЩИНСКИЙ НАУЧНЫЙ ЦЕНТР
ИНСТИТУТ МАТЕМАТИЧЕСКИХ ПРОБЛЕМ БИОЛОГИИ

А. М. МОЛЧАНОВ

НЕЛИНЕЙНОСТИ В БИОЛОГИИ



ПУЩИНО • 1992

Монография посвящена методическим вопросам, возникающим при изучении нелинейных систем в естествознании. Особое внимание уделено свойствам, которые не зависят от происхождения (механика, физика, химия, биология) и уровня организации системы (размер, число частиц, временной масштаб). Прослеживаются аналогии в поведении отдельных частиц и макросистем. Книга адресована научным сотрудникам, интересующимся междисциплинарными проблемами.

Тема этой книги — математическое моделирование в биологии. В ней обсуждаются методические уроки двух десятков лет работы математика с биологами. В книге нет претензий на общность подхода и, тем более на полноту. Это личный опыт автора, который может оказаться полезным и другим.

Из опыта работы в Пушкино, а также многочисленных симпозиумов и совещаний и особенно, десяти всесоюзных школ по математическому моделированию в биологии, автор вынес одно убеждение, которое необходимо предпослать дальнейшему.

Для успешной работы с биологами математикам не хватает прежде всего математического образования. Это не оговорка и не опечатка. Именно математического. Для занятий математикой достаточно знать только свою тему, только свой инструмент. В физике уже объект диктует набор математических инструментов, но этот набор обычно невелик и хорошо известен. В биологии объект неизмеримо сложнее и никогда заранее не знаешь, что именно из математики (да и из физики и химии) потребуется для данной задачи.

«Математику на «пять» знает один господь-бог» — говаривал М. В. Остроградский своим студентам — «Я знаю ее на «четыре», а вы все на «два». Так вот в биологии математику нужно знать на «четыре с плюсом» и работать с биологом, который знает биологию не хуже. Попытка выращивания «кентавров» обычно превращает хорошего биолога в посредственного математика, а хорошего математика — в плохого биолога.

В книге обсуждаются подробно только две темы — нелинейные явления и макросистемы («термодинамика»). Общность, столь милую сердцу математика, приходится принести в жертву содержательности. Зато в приложении намечены возможные обобщения и применения.

Природа не роскошествует началами вещей. (Фр. Бэкон).

Вся первая часть может служить подтверждением этого тезиса. Принцип «все или ничего» был давно сформулирован в физиологии. «Разрывная трактовка» позволяла (до появления современной вычислительной техники) успешно работать в прикладных

вопросах нелинейной теории колебаний. Принцип «смены* лимитирующего фактора», восходящий к Либиху и развитый Игорем Андреевичем Полетаевым, стал весьма общим методом математического моделирования в биологии. Нелинейная теория колебаний выявляет общую причину таких феноменов — переходный процесс. Переходный процесс имеет тенденцию быть коротким во времени и (соответственно) узким в пространстве. Угадывается связь с принципом Ле-Шателье, однако эта волнующая аналогия еще ждет своего исследователя. Эта тема тем более актуальна, что ее экологический аспект — роль границ между фитоценозами (по С. М. Разумовскому) — имеет существенное прикладное значение.

Понимание «нормы через патологию» давно уже стало в медицине важным эвристическим приемом. Хорошие преподаватели математического анализа учат строить график функции, начиная с особых точек. Фазовые переходы играют в физике существенную роль. Экстремальные режимы в экологии становятся в наше время важным инструментом понимания и предсказания драматических событий в биосфере.

А вот в математике теория бифуркаций еще весьма юная ветвь науки. Как обычно логика, аксиоматика завершают и формализуют то, что было завоевано без них и до них. Их роль — фиксация, упрощение, «инструментализация» идей и результатов, добытых более передовыми, подвижными, интуитивными отраслями естествознания. Но только консолидация (аксиоматизация) определенного уровня («слоя», «этажа») знаний позволяет начинать движение к следующему уровню. В первой части обсуждаются также две тесно связанные идеи — *устойчивости* и *эргодичности*. Традиционное противопоставление детерминизма и стохастики постепенно все более размывается. Сейчас (несколько опережая локомотив истории) можно думать, что идеи устойчивости, детерминизма относятся к внешним характеристикам сложных стационарных режимов. Тот же самый режим изнутри выглядит как эргодический и стохастический.

С вычислительной точки зрения вероятность, стохастика вторичны. Для получения шумовых эффектов используются чисто детерминированные механизмы перемешивания (типа дробной части числа). Не исключено, что это всего лишь переоткрытие древнего биологического механизма. Так, например, правдоподобно, что единый биохимический механизм дает и Т и В лимфоциты. Стационарный режим имеет сложную форму (в пространстве

* Наглядный пример. Ока (вблизи Пушкино) является зоной смены лимитирующего фактора. На северном берегу лимитирует свет и леса, растут на высоких местах, на южном лимитирует влага и деревья «толпятся» по оврагам.

концентраций) и в одной его части возникают Т формы, в другой В формы.

Стабилизация неустойчивых состояний высокочастотными колебаниями дает более простой пример того же типа. Эта задача разобрана Н. И. Боголюбовым, и уже давным-давно иллюстрируют ее жонглеры, удерживающие быстро вращающиеся тарелки на кончиках гибких бамбуковых прутьев.

Вторая часть посвящена весьма трудному вопросу — применимости термодинамики к биологическим вопросам. Разделяемая автором точка зрения состоит в том, что эта применимость весьма ограничена.

Рассматриваются общие макросистемы, то есть системы, состоящие из большого числа одинаковых компонент. Для стационарных состояний вполне применимы обычные термодинамические утверждения, включая второе начало термодинамики. Однако эти утверждения являются не точными, а только асимптотическими, что выясняется в достаточно общей форме на основе глубоких идей А. Я. Хинчина. Если и ввести сколько-нибудь реалистические предположения о внутренней структуре компонент, то стационарные режимы становятся значительно более сложными. Серия примеров (в том числе механических систем) показывает огромное разнообразие стационарных режимов. Следует особенно подчеркнуть, что в рассматриваемой серии примеров термодинамический предельный переход (число частиц $N \rightarrow \infty$) может быть точно осуществлен. Так, например, возникает знаменитая ныне система Лоренца. Она оказывается весьма частным случаем макродинамических систем.

Выясняется, что одна температура описывает только простейшие (скалярные, энергетические, одномерные) макросистемы. Нормально же имеется несколько «температур», то есть величин типа интенсивностей, описывающих поведение системы в целом.

Термодинамические представления, прекрасно описывающие простейшие (но поэтому и важнейшие) системы типа паровой машины или двигателя внутреннего сгорания, оказываются, поэтому, *дезориентирующими* при описании более сложных (и уж тем более биологических) систем.

Указанная серия примеров, допускающих предельный переход ($N \rightarrow \infty$), подсказывает, почти вынуждает глубокую аналогию между теорией бифуркаций (для уравнений макродинамики) и теорией фазовых переходов.

Еще менее применимы стандартные термодинамические соображения для весьма сложных и глубоких эволюционных вопросов.

В приложения вынесены работы, примыкающие к двум основным темам книги, и показывающие, что биологические идеи весьма плодотворны и в традиционных областях точного естествознания.

По словам Ф. Энгельса «подлинное единство мира состоит в его материальности, которое доказывается не парой фокусничес-

ких фраз, а всем долгим и трудным развитием естествознания». Похоже, что язык дифференциальных уравнений весьма пригоден для выражения этого единства. Впрочем еще Ньютон говорил, что «математика — это единственный язык, на котором можно разговаривать с природой».

ВОЗМОЖНАЯ РОЛЬ КОЛЕБАТЕЛЬНЫХ ПРОЦЕССОВ В ЭВОЛЮЦИИ

Доорганическая эволюция. Теория Дарвина имела решающее значение для объяснения принципов эволюции в живой природе. Сейчас накоплено, по-видимому, немало фактов, позволяющих утверждать, что эти общие принципы приложимы и к доорганической эволюции.

Мысль о том, что эволюция живого является продолжением (и ускорением) эволюции неживого, не является, разумеется, ни в малейшей мере новой. Непросто, однако, сформулировать в терминах физики и химии (а в идеале — в математических терминах) такие понятия, как «естественный отбор», «индивид».

В статье делается попытка сформулировать некоторые важные понятия эволюционной теории на математическом языке, точнее, на языке теории дифференциальных уравнений.

Что может быть общего у теории эволюции с дифференциальными уравнениями? Возникает естественный вопрос, почему вообще нужно формулировать эти понятия на языке, имеющем, казалось бы, весьма малое отношение к рассматриваемой теме.

Теория эволюции имеет дело прежде всего с такими понятиями, как индивид и вид. Существуют ли на языке дифференциальных уравнений соответствующие понятия? Постараемся сначала (в максимально абстрактных терминах) сформулировать понятие индивида.

Индивид. Индивид — это, по-видимому, такой объект (совершенно не обязательно живой), который можно выделить из окружения, противопоставить среде, изучать в какой-то мере независимо от всего остального, «индивидуально». На первый взгляд кажется, что это соответствует ближе всего физическому понятию «замкнутая система». Однако это не так.

В чем же отличительное, определяющее свойство индивида? Весьма правдоподобно выглядит предположение, что таким определяющим свойством индивида является информационная замкнутость системы.

Это означает следующее.

Индивид может (и должен) обмениваться энергией и веществом с окружающей средой, но этот обмен в основном определяет

ся внутренним состоянием самого индивида, а не состоянием окружающей среды.

Состояние индивида. Индивид обладает некоторым множеством состояний, причем переход из одного состояния в другое определяется только его собственным или его предшествующим состояниями. Множество состояний, которые возможны для индивида, может быть очень велико (оно задается большим числом переменных) для сложных систем. К счастью, переменные обычно устроены иерархично. Это значит, что состояние в основном определяется небольшим числом главных переменных, а остальные описывают лишь менее существенные детали. Более того, если сосредоточить внимание на какой-нибудь детали, то она в свою очередь определяется небольшим числом переменных, а остальные — пусть и в большем количестве — снова мало существенны.

Сказанное позволяет считать, что для многих задач вполне достаточно описание индивида системой небольшого числа обыкновенных дифференциальных уравнений:

$$\frac{dx}{dt} = A(x).$$

Разберем некоторые из возможных возражений против такой постановки вопроса.

Первое возражение. Индивидом может быть автомат *, поведение которого описывается не системой дифференциальных уравнений, а дискретной схемой, алгоритмом.

Дискретные модели. Если речь идет о реальной машине, физически реализованном автомате, то совершенно ясно, что алгоритмическое дискретное описание есть не что иное, как удобная разрывная идеализация непрерывных, но весьма резких изменений, происходящих в таком приборе.

Если же алгоритм задан абстрактно, но весьма правдоподобна, хотя, по-видимому, никем не доказана, теорема о том, что, любой такой алгоритм может быть интерпретирован как алгоритм численного решения для некоторого дифференциального уравнения. Поясним на простом примере это предположение. Пусть автомат имеет n состояний S_1, \dots, S_n , которые он проходит циклически $S_k \rightarrow S_{k+1}, S_n \rightarrow S_1$. Тогда такой автомат есть часть автомата, реализующего разностную схему для уравнений гармонического осциллятора

$$\dot{x} = -y,$$

$$\dot{y} = x.$$

* Автомат по Эшби (Эшби. Введение в кибернетику. ИЛ, 1960).

Рассмотрим теперь общий автомат. Он задается конечным множеством состояний, соединенных стрелками, которые показывают переходы. Почти очевидно, что множество можно расположить в пространстве и дополнить до непрерывного векторного поля во всем пространстве, т.е. включить в непрерывную динамическую систему*.

Описательно можно сказать, что дискретные автоматы расположены среди обыкновенных дифференциальных уравнений так же, как рациональные числа среди действительных.

Память. Второе возражение — индивид может обладать памятью. Его поведение может определяться не только состоянием в данный момент, но и предшествующими моментами, т.е. его историей. Пример такого поведения моделируется интегродифференциальным уравнением

$$\frac{dx}{dt} = \int_0^t K(x) d\tau.$$

В этом случае поведение системы с памятью эквивалентно поведению системы без памяти, но более широкой. Действительно, обозначим интеграл в правой части одной буквой

$$\int_0^t K(x(\tau)) d\tau = y.$$

Тогда система записывается в виде

$$\frac{dx}{dt} = y,$$

$$\frac{dy}{dt} = K(x).$$

По-видимому, при любом разумном определении «памяти» включение памяти ** в «динамические переменные» системы порождает более широкую систему без памяти и с детерминированным поведением. С точки зрения дифференциальных уравнений это утверждение близко по духу (двойственно) к утверждению о том, что система уравнений высокого порядка может быть сведена к системе первого порядка увеличением числа переменных.

Математик мог бы предложить (возражения третье и четвертое) модель «индивида», которая описывается разностными уравнениями

* Возможность обратного включения (приближенного, конечно) доказываетс^я ежедневно опытом многих программистов, интегрирующих дифференциальные уравнения на электронных вычислительных машинах, являющихся классическим примером дискретного автомата.

** Яркий пример — «собака и автомобиль» — содержится в цитированной выше книге Эшби (с. 168).

ми или уравнениями с запаздывающим аргументом. Эти возражения есть просто варианты первых двух возражений.

Распределенные параметры. Возражение пятое состоит в следующем. Система может иметь бесконечно большое число степеней свободы. Тогда для описания ее поведения могут, например, понадобиться дифференциальные уравнения в частных производных. Об этом в сущности уже шла речь выше, когда разбирался вопрос о множестве состояний индивида.

В математике основным приемом решения задач с большим (и даже бесконечным) числом степеней свободы является аппроксимация задачи аналогичной задачей с меньшим (небольшим) числом степеней свободы. Разложение по собственным функциям, разложение в ряды, отыскание параметров в вариационных задачах, разностные схемы — все это в сущности есть аппроксимация «бесконечномерной» задачи более простой конечномерной. Успех этих методов показывает, что определяющих параметров обычно бывает немного. В тех же случаях, когда равноправных параметров в самом деле много, это нередко означает, что задача плохо поставлена, и, вводя параметры иначе, мы опять-таки выделяем небольшое число определяющих переменных. Типичный пример такой ситуации — движение большого числа одинаковых частиц, когда введение (вместо необозримого количества переменных — координат и импульсов частиц) разумных общих величин — давления и плотности — позволяет просто ответить на основные вопросы о поведении такой системы в целом*.

Равноправие моделей. Следовательно, хотя и существуют весьма различные математические модели поведения индивида (дискретные, непрерывные с памятью, с конечным и бесконечным числом степеней свободы и т. д.) все они, по-видимому, эквивалентны в том случае, когда любую из них можно аппроксимировать какой угодно другой с любой степенью точности. Поэтому за основу можно выбрать ту, которая больше по вкусу. Различия между ними могут иметь решающее значение при исследовании конкретных задач, но весьма мало существенны при методологическом анализе. Всюду в дальнейшем принята модель обыкновенных дифференциальных уравнений.

Абстрактная схема и реальный объект. Итак, пусть поведение индивида описывается системой обыкновенных** дифференциальных уравнений

* Цена, которую мы платим за такое построение, — отказ от рассмотрения поведения отдельных частиц, — может оказаться слишком высокой в некоторых важных случаях.

** Уравнения, правые части которых не содержат явно времени, носят выразительное название «автономных».

$$\frac{dx}{dt} = A(x).$$

Здесь x — вектор, имеющий несколько компонент, а функция $A(x)$ — векторная функция векторного аргумента*. Смысл компонент вектора x может быть очень разным и решающим образом зависит не только от того, какой индивид изучается, но и от того, с какой точки зрения и даже от того, на каких масштабах времени происходит изучение.

Допустим, что речь идет о живом существе, например, млекопитающем. Если наше рассмотрение ограничено небольшими временами (порядка нескольких минут), то главными являются дыхание и работа сердца. На таких интервалах времени остальные функции (в физиологическом смысле этого слова) можно, по-видимому, считать постоянными (в математическом смысле этого слова), индивид достаточно хорошо моделируется сердцем и легкими, изучаемыми изолированно. В нашу задачу не входит сейчас обсуждение вопроса о том, как дальше строится физическая, а затем и математическая модель сердца — сошлемся на классические работы Ван-дер-Поля и его последователей. Заметим только, что в данном случае выделение существенных переменных индивида имеет морфологический смысл.

Орган — модель. Сердце локализовано не только функционально, но и геометрически. Последнее, разумеется, совершенно не обязательно. Если речь идет о газообмене, то сеть капилляров является органом — и соответственно моделью индивида в данном отношении, — локализованным только функционально, но отнюдь не объемно.

Если изучать поведение животного на больших отрезках времени, например, порядка суток, то в первом приближении определяющими будут параметры, относящиеся к деятельности органов пищеварения. На таких временах лучшая модель животного — желудок. Другие функции организма могут считаться постоянными при этом рассмотрении.

Весьма любопытно сравнить прямо противоположные причины, по которым можно не учитывать различные отправления. «Долгопериодические» функции (такие, например, как размножение), имеющие масштабы времени порядка года или месяца, не успевают существенно измениться в пределах суток. Другие, «быстрые»,

* $A(x)$, в частности, может быть линейной функцией и тогда $A = \frac{\partial A(x)}{\partial x}$ есть квадратная матрица.

как-то дыхание, меняются столь быстро, что имеют значения только средние по большому числу их собственных периодов*.

Некоторые выводы. Беглый анализ, проведенный выше, приводит к следующим полезным идеям.

1. Хорошей моделью индивида часто является его собственный орган.

Слово «орган» следует понимать в его общем, так сказать, «кибернетическом» смысле. Например, мотор является одним из органов автомобиля, станок — органом цеха, арифметическое устройство — орган вычислительной машины, щель для монет — органом телефона-автомата.

2. Один и тот же реальный индивид может моделироваться совершенно разными моделями при изучении его на различных масштабах пространства и времени.

Так, в небесной механике земля — материальная точка, а при изучении сейсмических свойств — шар или даже полупространство.

Атом в элементарной кинетической теории газов — твердый шарик, а в теории взаимодействия вещества с излучением — квантовая система, описываемая уравнениями в частных производных.

3. Существуют индивиды, представляющие собой иерархию колебательных систем, «вложенных» друг в друга, по крайней мере в смысле масштабов времени. При моделировании таких систем необходимо спрашивать себя, «о каких масштабах времени идет речь?», и учитывать только те переменные, периоды которых сравнимы с изучаемым масштабом. Медленные переменные можно считать постоянными, а от быстрых останутся только их средние значения**.

Устойчивость, неустойчивость и колебания. Если считать убедительными аргументы в пользу того, что каждому индивиду соответствует моделирующая его система дифференциальных уравнений, то ниоткуда не следует, что любой системе отвечает некий индивид.

Рассмотрим, например, систему уравнений, у которой есть одно-единственное состояние равновесия и притом устойчивое.

* Пример из другой области: пусть индивид — это гейзер, рассматриваемый на временах порядка нескольких извержений. Тогда условия, в которых он работает, — сечение и длина питающего канала, температура на входе, атмосферное давление на выходе и т.д. — можно считать постоянными. Осреднение по быстрым изменениям давления и других переменных приводит к простой релаксационной модели гейзера. Ясно, что в принципе механизм работы гейзера тот же, что и у неоновой лампы, хотя совсем непросто установить соответствие между параметрами этих моделей.

** Статистические схемы возникают, как правило, при отбрасывании быстрых переменных. Поэтому всегда можно говорить прямо об «осреднении» по мере в фазовом пространстве, избегая более чем двусмысленного и дискредитированного философскими спекуляциями термина «вероятность».

Кажется, что это пример хорошей модели хорошо уравновешенного индивида. Однако если вдуматься поглубже, то окажется, что такая система не отвечает нашему интуитивному представлению об индивиде. Ведь индивид что-то делает, с ним что-то происходит, он меняет свое состояние. Если же он приходит в одно-единственное состояние, то интуитивно это воспринимается как гибель системы. Система приходит в состояние устойчивого равновесия и перестает быть системой, способной к движению. Другой противоположный случай, — когда система неустойчива, также не отвечает представлению об индивиде, ибо означает в сущности прогрессирующую несовместимость частей системы, приводящую ее к распаду.

Лучше всего представлению об индивиде, как о системе, которая с одной стороны, сохраняет свое строение, а с другой — способна к внутреннему движению, отвечают поэтому колебательные системы.

Большие промежутки времени. Итак, системы описывающие индивид, должны быть колебательными системами. Это утверждение можно понять еще следующим образом.

Пусть когда-то, давным-давно существовали разнообразные объекты. С тех пор прошло много времени, протекли миллиарды лет. Какие объекты остались с тех давних времен, кто выдержал испытание временем?

Устойчивые? Нет, так как они давно уравновесились, стали частью среды. Вспомним, что относительное постоянство является отличительной чертой именно среды.

Неустойчивые? Нет, так как они распались.

Следовательно, имеют шанс «выжить», сохраниться только колебательные системы, процессы и объекты.

Этому утверждению противоречит на первый взгляд несомненное существование неустойчивых объектов. Ярким примером активного, неустойчивого объекта является свободный кислород атмосферы. Но мы знаем*, что кислород не что иное, как продукт жизнедеятельности морских водорослей (на 90%) и вообще зеленых растений (на 10%).

По-видимому, в такой среде, как Земля, которая громадное количество времени находится в состоянии слабопеременного «проточного» (излучения Солнца!) равновесия, неустойчивые объекты могут существовать только как постоянно возобновляемый продукт деятельности нейтральных колебательных систем.

Уточнение. Переменные и параметры. В математической модели индивида его поведение всегда одинаково и ни от чего не

* «Физика и химия жизни», ИЛ, 1960, стр.34. Темп возобновления 10^{-4} за год. Это значит, что за время существования фотосинтеза (10^9 лет) водоросли 100 тыс. раз прямо возвращали в атмосферу каждую молекулу O_2 .

зависит. Но поведение реального индивида будет различным в различных средах. Можно попытаться учесть это обстоятельство, вводя зависимость правых частей от дополнительных переменных параметров

$$\frac{dx}{dt} = A(x, \alpha).$$

Параметр α подобно переменному x является вектором, число компонент которого, вообще говоря никак не связано с числом компонент вектора x . Естественно считать, что индивид не перестает быть индивидом и притом тем же самым* индивидом, если параметры немного изменились. Обойдем молчанием весьма щекотливый вопрос о том, что такое «немного», и разберем подробнее смысл и роль параметров α .

Индивид и среда. Будем помещать индивида в различные среды. До тех пор пока он вообще в состоянии работать как индивид, среды классифицируются только по результатам их воздействия на индивид. Пока изучается данный индивид, объем и форма информации, необходимой для предсказания поведения индивида, определяются только строением индивида и ничем больше. Если, в частности, среда такова, что ее воздействие на индивид не меняет параметра α то такую среду следует считать постоянной. Возьмем такой классический прибор, как термометр. В чем смысл измерения температуры? В том, по-видимому, что значение температуры позволяет предсказать какие-то явления. Так, если температура ниже нуля, то вода замерзает, а если выше нуля, то лед тает. Стоит подчеркнуть, что предсказываются только финальные состояния, а не детали переходящего процесса.

Напрашивается естественное обобщение.

Если имеется любой индивид, то стационарные значения его параметров α следует рассматривать как обобщенную «температуру» среды.

Такое словоупотребление подчеркивает два важных обстоятельства. Первое — обычная температура характеризует не только и не столько среду, сколько поведение очень простых «скалярных» систем в этой среде. Второе — чем сложнее изучаемый индивид, тем более сложной является «температура», характеризующая его равновесие со средой. Так, например, при задании химической системы нужно задавать не только температуру, но и химический потенциал.

* Говорят, правда, что в прошлом веке один французский художник изобразил Луи-Филиппа в виде перезревшей, готовой упасть груши. Представ перед судом, он с непостижимой быстротой набросал семнадцать рисунков, самый левый из которых был несомненно Луи-Филипп, а самый правый столь же несомненная груша, и попросил суд точно указать ему, где именно надлежало остановиться, дабы избежать оскорбления величества.

Необходимость разнообразных «термодинамик». Сказанное приводит к выводу, что термодинамика (на редкость неудачный термин, значительно точнее было бы — «термостатика») есть наука о равновесных средах для максимально простых индивидов, имеющих один-единственный параметр — внутреннюю энергию.

В этом смысле обычная термодинамика является, конечно, наиболее универсальной теорией, но именно по этой причине она почти бесполезна для изучения сколько-нибудь сложных систем*, она для такого изучения слишком груба.

С изложенных позиций представляется поэтому заранее обреченной на неудачу любая попытка объяснить принципами типа «минимум потока энтропии», поведение не только живых, но даже сколько-нибудь сложных неживых систем.

Сложные системы не характеризуются только энтропией или вообще одним (каким угодно) числом именно потому, что уже задание равновесных сред** требует задания всех параметров α .

Разумеется, в число этих параметров входят и температура и химический потенциал, но может быть еще и много других параметров α .

Резюмируем. Равновесие индивида со средой характеризуется значениями его внутренних параметров α . Простейшие (энергетические) системы требуют для своего описания только одного параметра — температуры. Более сложные системы содержат наряду с температурой и другие параметры, а в наиболее тонких случаях энергетические представления и, в частности, энтропия могут оказаться второстепенными.

Учет среды в математической модели. Всю «литературную» модель индивида в среде, на которую истрачено несколько страниц, можно ужать до двух строчек уравнений:

$$\frac{dx}{dt} = A(x, \alpha),$$

$$\frac{d\alpha}{dt} = \varepsilon B(x, \alpha).$$

Эта запись весьма содержательна. Здесь проведено четкое разделение переменных индивида и его параметров. Указано, что такое разделение имеет точный смысл только при $\varepsilon=0$, а при малых ε оно носит только асимптотический характер. Наличие малого параметра подчеркивает главное: влияние среды есть процесс значительно более медленный, чем внутренние изменения

* Существуют соединения, стереоизомеры которых ядовиты. Никакое вычисление энтропии не поможет уберечься от яда, так как энтропия не знает «где лево, а где право».

** Подчеркнем, речь идет о равновесной среде, т.е. о постоянстве параметров α а не о равновесии (т.е. постоянстве x) индивида, так как, с нашей точки зрения равновесие индивида равносильно его гибели.

индивида. В такой модели предположена довольно высокая степень общности: влияние среды зависит от состояния индивида, так как правая часть уравнения для α зависит и от α и от x . Это может приводить к появлению предпочтительных состояний. Более подробно этот тонкий вопрос здесь разбираться не будет. Однако один важный аспект следует разобрать уже сейчас.

В простейших случаях, например, при устойчивом периодическом режиме быстрых движений $x(t)$, можно показать, что вместо параметра α допустимо ввести новый параметр β по формуле

$$\beta = \alpha + \varepsilon P(x, \alpha, \varepsilon)$$

$$\frac{dx}{dt} = A(x, \beta),$$

$$\frac{d\beta}{dt} = \varepsilon B(\beta)$$

т.е. убрать быстрые переменные x из системы для медленных переменных. Это весьма важный факт.

Мера индивидуальности индивида. Формула для β особенно выразительно подчеркивает приближенный характер противопоставления параметров индивида его «внутренним» переменным. Довольно ясно, что это прямо относится к главному вопросу о самой возможности выделения индивида из среды. Однако сейчас мы в состоянии изъясняться уже более точно. Малый параметр ε дает количественную меру точности с которой можно говорить об индивиде. Этому параметру можно придать более наглядный смысл, заметив, что $n = \frac{1}{\varepsilon}$ дает по порядку величины число периодов собственных колебаний, за которые происходят существенные (порядка единицы) изменения параметров. Это число довольно точно соответствует понятию «добротность» в радиотехнике, где оно, к слову сказать, обычно имеет величину порядка 10^4 . Таким образом, «мера неиндивидуальности» ε есть обратная величина «добротности». Любопытно сравнить добротность таких индивидов, как человек, земля и электронная вычислительная машина. В качестве основного периода возьмем биение сердца, один год и время выполнения элементарной операции; а за большой интервал выберем среднюю продолжительность жизни, время существования Земли и время между двумя сбоями. Получается, что Земля имеет добротность 10^{10} , человек 10^9 и машина 10^6 . Самое главное, что все это очень большие числа.

Вернемся к системе уравнений. Во многих вопросах интересуются только изменением параметров индивида. В этих случаях естественно оставить только уравнение для β :

$$\frac{d\beta}{d\tau} = B(\beta),$$

введя, конечно, медленное «эволюционное*» время

$$\tau = \varepsilon t,$$

так как время t , связанное с быстрыми переменными x , не характерно для медленных изменений параметров. Произошло расщепление системы, выделение независимой системы для медленных переменных. Это расщепление, имеющее принципиальное значение, часто является источником терминологической путаницы. Так, например, говорят о стационарных, да еще устойчивых** состояниях индивида. На самом же деле речь идет о стационарных значениях параметров α , а вовсе не о стационарных значениях x , что означало бы гибель индивида.

Сложность системы и колебания. Важную роль колебательных систем можно понять и с нашей новой точки зрения. Будем помещать индивид в различные среды, т.е. менять α . Системы дифференциальных уравнений движения*** имеет обычно положение равновесия, т.е. точку x_0 для которой $A(x_0) = 0$. Характер стационарной точки зависит, конечно от параметра α . В частности, величины действительных частей собственных значений линеаризованной системы**** также будут функциями параметров α . Приравняем нулю***** одну из действительных частей

$$\rho(\alpha) = \operatorname{Re} \lambda(\alpha) = 0$$

Даже если параметр α всего один, это уравнение, вообще говоря имеет решение. Корни уравнения определяют критические значения параметра, при которых система может попадать в колебательные режимы, так как собственные значения будут

* Во избежание заметим, что здесь термин «эволюция» относится к медленным изменениям индивида (например, возрастные изменения). О связи эволюции индивида и вида см. ниже.

** Нередки снисходительные усмешки по поводу тех, кто не понимает, что «в природе реализуются только устойчивые состояния». Автор смиренно сознается, что относится к числу еретиков, полагающих, что индивиду нет решительно никакой надобности находиться в «устойчивом» состоянии. Индивиду куда больше смысла сделать устойчивыми как раз такие значения внешних (так сказать «защитных») параметров α , при которых он имел бы максимальную внутреннюю свободу. Автор склонен видеть в количестве независимых частот, которые имеют внутреннее движение, меру этой внутренней свободы, меру сложности организации индивида.

*** «Механистическая» привычка! Следовало бы говорить об уравнениях поведения.

**** Примечание для нематематиков. Эти три родительных падежа подряд соответствуют (читать надо с конца!) трем обрядам, которые совершаются при исследовании на устойчивость: 1) систему линеаризуют; 2) вычисляют собственные числа; 3) берут действительные части этих собственных чисел. После того как это проделано, ответ иногда получается, иногда не получается.

***** Это как раз тот случай, о котором в предыдущей сноске было сказано «иногда не получается».

чисто мнимыми. Если параметров α не один, а два, то можно рассчитывать на получение двухчастотных колебаний и т.д.

Итак, в пространстве параметров имеются поверхности (меньшего числа измерений), каждая точка которых определяет систему, обладающую сложным, в частности, колебательным движением. Ясно, что поведение* индивида будет тем более сложным, чем большим количеством внутренних параметров он располагает.

Возникает следующая картина. В пространстве параметров существуют точки, которым соответствует очень простое поведение индивида — например, гибель, так как соответствующая система уравнений обладает одним устойчивым состоянием.

Существует, однако, поверхность на которой возникает колебательное поведение. В точках, близких к этой границе (со стороны гибели), индивид хотя и гибнет, но это происходит по типу колебательного разряда. Чем ближе к границе, тем дольше длятся колебания, пока, наконец, на самой границе они не становятся чисто периодическими.

Механизмы усложнения. Переход системы через критическую границу имеет принципиальное значение. Становится ясным прежде всего, что разделение переменных на параметры и внутренние переменные имеют только асимптотический (предельный при $\varepsilon \rightarrow 0$) смысл. Более того, получается, что в математической модели уже заложена возможность усложнения индивида.

В самом деле, при переходе через критическую границу система приобретает лишнюю колебательную степень свободы, так как одно из переменных, имевших ранее свойство параметра (оно стремилось уравновеситься), становится теперь равноправным внутренним колебательным переменным.

Более внимательный анализ приводит к выводу, что уравнения подсказывают по крайней мере два механизма усложнения, которые могут иметь эволюционное значение.

«Биорезонанс». Первый из них состоит в том, что индивид попадает на границу устойчивости по одному из своих эволюционирующих параметров. Очень соблазнительно интерпретировать этот переход как попадание индивида в «трудные** условия».

Но если все предыдущее верно, то возникает следующая увлекательная возможность.

Сам факт попадания в трудные условия может создавать предпосылки для преодоления этих трудных условий. В самом деле если несколько индивидов попали в колебательный режим,

* Заметим, что точке в пространстве параметров α отвечает траектория (поведение) в пространстве состояний x . Очень важно не путать эти два типа переменных.

** Можно думать, что рассмотрение проточных систем приведет к утверждению, что при довольно широких предположениях увеличение «нагрузки» на проточную систему вызывает появление колебательных режимов.

то созданы условия для возникновения резонанса, причем возникающее объединение имеет, так сказать, многоклеточный характер.

Конечно, такой резонанс не обязательно означает усложнение организации. Может случиться, что из двух «резонирующих» индивидов один просто существует за счет другого: происходит, например, необратимая перекачка энергии из одного маятника в другой.

Но может случиться и так, что в результате получится более сложная организация. Очень интересна математическая задача — выяснить условия, при которых в модели может происходить это замечательное явление.

Релаксационные колебания. Есть другая интересная возможность эволюционного усложнения организации. Представим себе систему типа химической, жизнедеятельность которой подавляется конечными продуктами этой жизнедеятельности. Пусть далее есть другая система, для которой эти продукты являются исходным материалом, условием ее работы.

В этом случае явления развиваются, так сказать, в обратном порядке.

Сначала возникают резкие релаксационные колебания, так как использование «связующего» продукта носит случайный характер. Грубо говоря, второй индивид «где-то бродит» потому, например, что первый выдает еще слишком мало продукции* и «с нее сыт не будешь».

Затем система налаживается и резкие колебания сглаживаются, переходя, в идеале, к ровной стационарной работе.

Один индивид или два? Внимательный читатель уже заметил, вероятно, что совершается некий «обман» — все время речь шла только об одном индивиде, а сейчас их оказалось уже два. Однако в формальной математической схеме никакого противоречия нет. Эта схема оказывается достаточно богатой и гибкой, чтобы включить в себя на равных правах обе схемы: и последовательного, и параллельного объединения.

Математически это выглядит так: если у вас есть « n » индивидов сорта x и столько же индивидов сорта y , то мы объявляем «компаунд-индивидом» просто пару $z = (x, y)$ и уравнения для Z_k записываются вполне аналогично уравнениям для x_k .

Поэтому, когда в математической модели ожидается возникновение нового индивида из двух старых, можно считать, что все время есть один сложный индивид. Только в одних предельных условиях (например, при $t \rightarrow -\infty$) единая система распадается на две независимые, а в других (при $t \rightarrow +\infty$) может потерять смысл разбиение системы на части.

* А выдает он ее мало потому, что его вовремя от нее не освобождают.

Такой прием изучения сложных систем давно известен в физике. Хороший тому пример — метод активированного комплекса в квантовой химии.

Затягивание среды. Выход индивида на критическую границу независимо от того, как именно этот выход происходит, может иметь еще одну интерпретацию. Мы различаем внешние и внутренние переменные. При качественном скачке, когда индивид становится другим, часть бывших внешних параметров становится внутренними параметрами. Таким образом, новый индивид можно рассматривать как старый индивид, присоединивший* к себе часть среды.

Такое закрепление оптимальной среды — вещь весьма обоудострашая: жить, конечно, стало легче, но зато теперь необходимо защищать это благоприобретение. Эта необходимость может стать стимулом новой экспансии, и так — до тех пор, пока центробежные факторы не уравновесят центростремительные.

Морфология и кинетика. В заключение бегло коснемся интересного вопроса о связи строения и функционирования.

Колебательные системы, являющиеся предметом рассмотрения, как правило, как-то «устроены». У них есть обычно пространственно разделенные специализированные части. Тот же гейзер имеет подводящий канал, полость, где происходят вскипание и выход в атмосферу — явная пространственная гетерогенность, которая и является причиной колебаний.

Несомненно, что сейчас причина колебаний — гетерогенность. Но что послужило причиной появления самой гетерогенности? Ведь гейзер сам себя устроил. Вытекала горячая вода, охлаждалась, попадая на поверхность, выпадали отложения, которые и создавали структуру гейзера.

Если даже такая «монотонная» кинетика способна к созданию структур, то какова же формообразующая роль колебательных химических реакций?

Напрашивается вывод: *нынешняя структура — следствие вчерашней кинетики. Биологические структуры есть морфологическое закрепление кинетических свойств больших молекул**.* Форма этого высказывания, возможно, излишне категорична. Однако суть дела состоит в том, что в эволюционном аспекте изучение кинетики есть вместе с тем изучение морфологии (или наоборот).

* Затягивание среды не есть выдумка математика. Приведем два интересных примера. Кровь животных по солевому составу близка морской воде — миллиард лет носим мы в себе родимую стихию. Другой пример — термиты, которые добравшись до средних широт, дотащили в термитниках температуру и влажность тропического пояса.

** Большие молекулы в свою очередь представляют собой материализованную кинетику малых молекул.

Совсем непросто, конечно, в каждом конкретном* случае увидеть, как именно кинетика порождает структуру, но от этого задача не становится менее важной.

Что же сказано в статье? Выдвинуто предположение, что в теории нелинейных колебаний существует задача, очень близкая по духу к общим задачам эволюционной теории.

Это — задача о поведении на больших временах ($t \sim \frac{1}{\varepsilon}$) или даже больше решений системы уравнений вида:

$$\frac{dx}{dt} = A_0(x) + \varepsilon A_1(x, \varepsilon)$$

Основная часть статьи посвящена переводу на язык уравнений важнейших понятий теории эволюции. Смысл этой деятельности автор видит в том, что каждое переведенное, таким образом, понятие вырвано тем самым из цепких объятий «биологической специфики». За ним сразу выстраивается вереница физических, химических и даже экономических объектов, которые оказываются «гомологичными» математической модели.

Существуют прямо противоположные точки зрения на то, как далеко может быть продвинут аксиоматический метод в изучении биологических явлений. Одни считают, что можно промоделировать все, другие — что ничего. Точка зрения автора состоит в том, что «все» моделировать невозможно, да и не нужно, а вот «главное» нужно стремиться аксиоматизировать.

Если понимать под «главным» поведение систем в критических ситуациях, то обычно оказывается — если угодно в этом и состоит «символ веры», — что поведение системы бывает критическим по небольшому числу переменных, критические явления могут быть хорошо описаны довольно простыми моделями.

С заключительным замечанием согласятся, вероятно и сторонники математической «неописуемости» живых систем.

Чем дальше будет продвинута работа по моделированию, тем яснее мы будем понимать, что же есть «истинно биологического» в живых системах, такого, что не поддается уже никакому химико-физико-математическому описанию.

Обсуждение

Б. В. Вольтер. Я хочу сказать несколько слов о связи «трудных условий» и колебательных режимов. Это, возможно, будет еще одним примером к сообщению А. М. Молчанова.

Занимаясь автоматизацией самых различных химических реакторов, мы неоднократно наблюдали их в колебательных режимах. Чаще всего такие режимы появлялись вблизи границы допустимых условий работы, т.е. тогда, когда реактор оказывался в «трудных условиях».

* Например, волнующая задача о физиологическом значении конформационных движений ферментов.

Для большинства реакторов полного смещения статическая характеристика (например, зависимость температуры в реакторе от температуры входной смеси) имеет два экстремума, являющиеся точками бифуркации, которые можно назвать критическими или «трудными точками». Если входную температуру поднимать выше максимума статической характеристики, то произойдет срыв режима, что часто носит характер теплового взрыва. Если охлаждать входную смесь ниже минимума статической характеристики, то реакция заглохнет. Поэтому точки бифуркации можно считать относящимися к трудным условиям. Интересно, что колебательные режимы появляются вблизи точек бифуркации.

РЕЛАКСАЦИОННАЯ МОДЕЛЬ АДАПТАЦИИ

Рассмотрены общие свойства процесса адаптации в предположении, что это явление имеет биохимическую основу. Высказана гипотеза о существенном различии хода адаптации в изолированной системе и в комплексе систем, где может стать выгодной колебательная кинетика.

Введение

Адаптационные процессы широко распространены в биологии и встречаются на самых разных уровнях организации. Понятие приспособления очень широко, и вряд ли существует единый механизм, лежащий в основе всех адаптационных явлений. Возможно, однако, что чисто биохимические процессы составляют существо довольно широкой группы явлений физиологической адаптации (если понимать под словом «физиологические» сравнительно мягкие изменения условий). Среди них выделяется более узкая группа, поражающая воображение яркостью кинетической картины.

Нормально работающая система при небольшом* изменении условий «вдруг замирает» и долго «не подает признаков жизни». Нетерпеливый наблюдатель может счесть систему погибшей, однако по прошествии некоторого времени (иногда довольно значительного) система столь же внезапно возобновляет работу и работает «как ни в чем не бывало».

Выражения, взятые в кавычки, кажутся на первый взгляд излишне «литературными» и неопределенными. Однако именно они легче всего поддаются формализации, подсказывая мысль о релаксационной природе адаптации.

1. Быстрые и медленные переменные

Слово «вдруг» есть, в сущности, эмоциональный эквивалент утверждения, что время реакции наблюдателя значительно боль-

* Например, изменение температуры на 4-6 градусов

ше времени «срабатывания» наблюдаемой системы. Но «вдруг» система лишь изменяет состояние, а сохранять его она может так долго, что наблюдатель иногда теряет терпение. Следовательно, резкое различие масштабов времени есть внутреннее свойство системы, а не субъективная оценка ее наблюдателем. В математической модели это означает наличие малого параметра в системе

$$\varepsilon \frac{dw}{dt} = a(w, \varepsilon), \quad (1.1)$$

(в простейших случаях) возможность такого выбора переменных, которое сводит к распадению системы на быструю и медленную:

$$\varepsilon \frac{du}{dt} = f(u, v, \varepsilon), \quad (1.2)$$

$$\frac{dv}{dt} = g(u, v, \varepsilon). \quad (1.3)$$

Кинетика такой системы определяется, как известно, свойствами поверхности квазистационарных состояний, получающейся при формальном предельном переходе $\varepsilon \rightarrow 0$

$$0 = f_0(u, v), \quad (1.4)$$

$$\frac{dv}{dt} = g_0(u, v). \quad (1.5)$$

Эта система позволяет найти быстрые переменные как функции медленных переменных:

$$u = \varphi(v) \quad (1.6)$$

и, подставив найденные значения u в уравнения для v :

$$\frac{dv}{dt} = g_0(\varphi(v), v), \quad (1.7)$$

проследить эволюцию медленных переменных.

Изложенное полностью содержит «метод квазистационарных концентраций» [1]. Однако еще в классической работе Тихонова [2] было найдено условие близости реального поведения системы к полученному предельному.

Это условие устойчивости квазистационарной точки быстрой системы.

Проверка этого условия принципиально очень проста, хотя может содержать большие технические трудности. Алгоритм проверки состоит в вычислении матрицы производных в точке квазиравновесия

$$A = \left. \frac{\partial f}{\partial u} \right|_{f(u,v)=0} \quad (1.8)$$

и отыскании ее собственных значений

$$\det|A - \lambda E| = 0 \quad (1.9)$$

Если все эти числа имеют отрицательные действительные части:

$$p_i = \operatorname{Re}\lambda_i < 0, \quad (1.10)$$

то система (1.2) устойчива и движение реальной системы происходит в ε -окрестности траекторий системы (1.5), лежащих на поверхности (1.4) квазистационарных состояний.

Однако условие Тихонова (1.10) выполнено далеко не всегда. Рассмотрим на поверхности квазиравновесия точки, в которых выполнено равенство

$$\det \left| \frac{\partial f}{\partial u} \right| = 0 \quad (1.11)$$

Эти точки, называемые точками срыва, образуют границу, отделяющую область устойчивости от области неустойчивости. Исследованию поведения системы вблизи точки срыва посвящен цикл работ Л.С. Понтрягина [3] и его учеников. Существо явления, детали которого нас сейчас не интересуют, состоит в том, что траектории «срываются» с поверхности квазиравновесия, квазистационарное приближение становится грубо неверным.

Это отчетливо видно уже на простейшем примере системы двух уравнений, возникающих из уравнения Ван-дер-Поля.

Качественная картина релаксационных колебаний полностью определяется наличием области неустойчивости* на поверхности квазиравновесия и характером медленного движения (по направлению к точкам срыва!) на устойчивых участках этой поверхности.

Весьма поучительно сопоставление поведения быстрых и медленных переменных. Быстрые переменные совершают характерные «перескоки» из одного состояния (ветвь A) в другое (ветвь B), невольно вызывающие ассоциацию с внезапным «пробуждением» системы после лаг-фазы в адаптации.

Медленные переменные остаются непрерывными, но в точках, соответствующих разрывам быстрых переменных, на графиках медленных переменных образуются характерные изломы.

Релаксационные колебания взяты в качестве иллюстрации не только из-за их широкой популярности, но также и потому, что это весьма правдоподобная схема работы адаптивного механизма в некоторых крайних условиях. Однако об этом ниже.

* На плоскости точки срыва совпадают с границами участков монотонности кривой квазиравновесия.

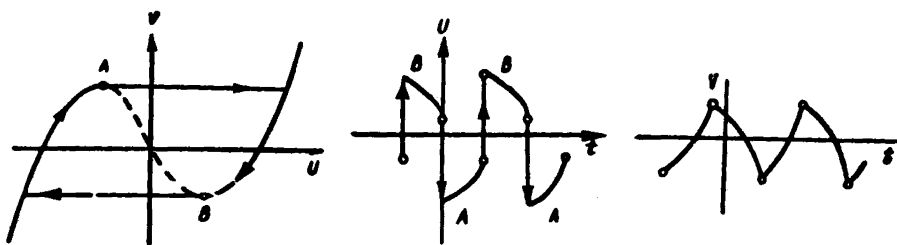


Рис. 1. «Разрывная трактовка» релаксационных колебаний. Пунктиром выделен участок неустойчивости на кривой квазиравновесия, A и B — точки срыва

Рис. 2. «Перескоки» быстрой переменной с ветви A на ветвь B и обратно

Рис. 3. Изломы в экспериментальной кривой, «выдающие» существование быстрой стадии

2. Наблюдаемые и скрытые параметры

О жизнедеятельности изучаемой системы судят обычно по немногим, ясно видимым признакам типа роста, или движения, или потребления субстрата. Глубинные процессы мало доступны и экспериментальное вмешательство (даже со скромной целью наблюдения) нередко кончается гибелью системы. Даже в опытах *in vitro* непрерывная регистрация представляет собой трудную техническую задачу [4] и наблюдаемой обычно бывает одна-единственная величина z , являющаяся довольно сложной* функцией концентраций:

$$z = F(u, v), \quad (2.1)$$

причем часто неизвестным образом зависящая от концентраций.

Если z есть наблюдаемая величина, а переменные u и v — это скрытые переменные, относящиеся к релаксационной системе, то график z обычно имеет характерные разрывы, типичные для быстрых переменных.

Разрывы будут отсутствовать лишь в том исключительном случае, когда наблюдаемая величина зависит только от медленных переменных и совсем не зависит от быстрых.

Существует, однако, широкий класс методов регистрации, связанный с накоплением продуктов реакции. В этих случаях наблюдаемая величина z связана с определяющими переменными u и v дифференциальным уравнением

* Светопоглощение, окислительно-восстановительный потенциал, выделение газов и т.д.

$$\frac{dz}{dt} = F(u, v, z) \quad (2.2)$$

которое является частным случаем уравнения для медленной переменной. Разница только в том, что переменное z в нашем случае не входит в систему (1.1) и может быть найдено после того, как основная система проинтегрирована. По существу, сказанное, есть точная математическая формулировка гипотезы: измерения ведутся настолько аккуратно, что не вносят искажений в ход процесса.

Таким образом, различие между разрывами и изломами на экспериментальной кривой характеризует вовсе не свойство системы, а метод регистрации. Для классических методов регистрации (например, наблюдения за ростом) типична меньшая точность, необходимость накопления и, следовательно, изломы на экспериментальной кривой, нередко воспринимаемые субъективно как «погрешность» эксперимента с вытекающим отсюда стремлением «сгладить» кривую. Понятно поэтому пристрастие, которое питает математик к малоинерционным методам непрерывной регистрации, выделяющим четко, скачками, наиболее интересную, определяющую сторону явления — его релаксационный характер.

3. Релаксационная модель. Состояние активности и покоя

Простую релаксационную модель адаптации можно построить с одной быстрой переменной, которая является также единственной наблюдаемой величиной z . Вторая переменная y — медленная. Любопытно, что даже эти минимальные предположения позволяют построить достаточно содержательную модель:

$$\begin{aligned} \frac{dy}{dt} &= b(y, z, \varepsilon), \\ \varepsilon \frac{dz}{dt} &= c(y, z, \varepsilon). \end{aligned} \quad (3.1)$$

Весь дальнейший анализ основан на двух основных предположениях. Первое состоит в том, что явление вообще описывается математической моделью. Второе — что система имеет рабочее состояние и состояние покоя, которые различаются наблюдательно.

Так как единственной наблюдаемой величиной является переменная z , то это означает, что кривая квазиравновесия

$$c(y, z, \varepsilon) = 0 \quad (3.2)$$

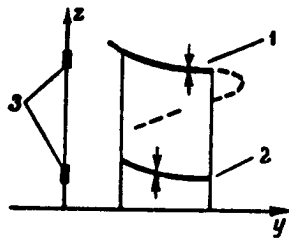


Рис. 4. Две ветви кривой квазиравновесия. z — наблюдаемая величина, y — внутреннее переменное; 1 — состояние активности, 2 — состояние покоя (шока), 3 — экспериментальные характеристики состояния покоя и активности

имеет по крайней мере два различных* корня,

$$Z_0 = f(y, \epsilon), \quad Z_1 = g(y, \epsilon), \quad (3.3)$$

один из которых соответствует покою, а другой — активному режиму.

Математически это означает, что кривая $c(y, z, \epsilon) = 0$ на плоскости (y, z) имеет в некоторой зоне изменения внутренней переменной y по крайней мере две ветви.

Сама возможность наблюдения состояний активности и покоя означает устойчивость обоих этих состояний по отношению к быстрому движению. Но отсюда уже чисто логически вытекает существование промежуточного, неустойчивого по отношению к быстрому движению, состояния.

Это промежуточное состояние трудно или даже невозможно зарегистрировать, так как даже очень малое его измерение быстро (за время порядка ϵ) перебрасывает систему либо в квазиравновесие активности, либо в квазиравновесие покоя. Однако роль этого состояния весьма существенна, ибо оно делит плоскость (y, z) на зону притяжения состояний активности и зону притяжения состояний покоя.

Следующий этап анализа — учет эволюции системы в «состоянии» покоя. Из эксперимента известно, что система «находящаяся» в покое может «спонтанно» возобновить активность. Наиболее простая интерпретация этого факта состоит в том, что суждение о состоянии системы по одной только наблюдаемой величине слишком грубо.

* Изменение шкалы измерения z

$$\dot{z} = (1 - \theta)f + \theta g,$$

$$\theta = \frac{(z - f)}{(g - f)}$$

приводит общий случай к ситуации «да-нет», когда активности соответствует значение $\theta = 1$, а покою $\theta = 0$. Качественные высказывания типа «движется», «не движется» можно поэтому считать частным случаем нормировки наблюдаемой величины.

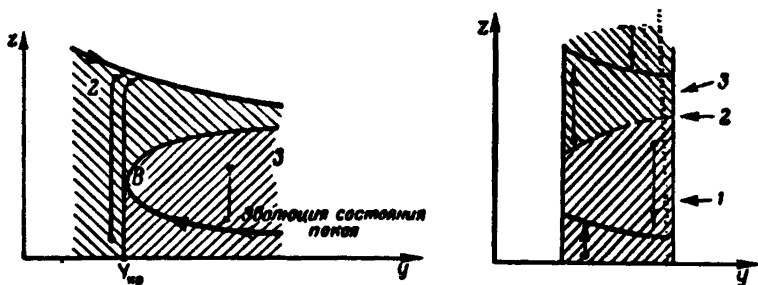


Рис. 5. Состояние неустойчивого квазиравновесия, отделяющее зону активности от зоны покоя. 1— зона покоя, 2— граница — неустойчивые квазиравновесия, 3— зона активности

Рис. 6. Спонтанное возобновление активности. 1— потеря устойчивости состояния, 2— зона активности, 3— зона покоя, 4— эволюция состояния покоя

Внутренняя, не наблюдаемая (при данных методах регистрации) переменная на самом деле медленно сдвигается вдоль линии покоя до пересечения с линией раздела.

В этот момент состояние покоя теряет устойчивость по отношению к быстрому движению и система за короткое время приходит в состояние активности. Точка y_{kp} на рис. 6 соответствует как раз такому критическому состоянию, за которым система вообще не имеет состояния покоя и может устойчиво находиться только в состоянии активности.

Скрытая переменная y испытывает в момент скачка незначительные изменения. Однако скорость ее движения

$$\frac{dy}{dt} = b(y, z, \epsilon) \quad (3.4)$$

зависящая от быстрой переменной z , меняется существенным образом. Медленные изменения y будут, конечно, продолжаться, но характер эволюции в состоянии активности обычно иной, нежели в состоянии покоя.

Математически это выражается в том, что уравнение для y в состоянии покоя,

$$\frac{dy}{dt} \approx b(y, f(y)) \quad (3.5)$$

вообще говоря, совершенно непохоже на уравнение в состоянии активности

$$\frac{dy}{dt} \approx b(y, g(y)) \quad (3.6)$$

Вполне вероятно, например, изменение направления движения y .

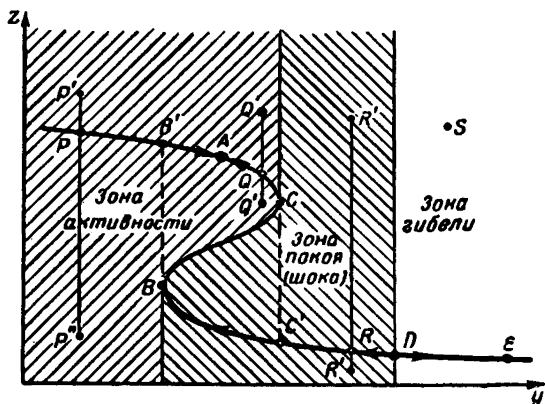


Рис. 7. Фазовый портрет адаптивной системы: А — точка устойчивой активности; D — граница обратимых изменений

Проведенный до сих пор анализ относится к механизму возвращения адаптивной системы в состояние активности. Поддержание этого состояния допускает еще более простое истолкование.

Если система уравнений имеет истинные положения равновесия, то все они расположены на кривой квазиравновесия, так как в точке истинного равновесия обращаются в нуль обе скорости:

$$\left. \begin{aligned} b(y, z, \varepsilon) &= 0 \\ c(y, z, \varepsilon) &= 0 \end{aligned} \right\} \quad (3.7)$$

как быстрого, так и медленного движения. Точек истинного равновесия может быть несколько. Они соответствуют совершенно различным режимам. Устойчивая точка на линии покоя соответствует гибели системы, устойчивая точка на линии активности — устойчивой жизнедеятельности системы. Наконец, неустойчивые точки отделяют участки квазистационарной линии с различным типом поведения.

4. Типы воздействия. Пример адаптивного поведения

Дальнейшее обсуждение невозможно без перевода на язык модели важнейшего экспериментального понятия — понятия воздействия на систему. В реальной ситуации оно может осуществляться самыми разными способами: изменением температуры, механическим повреждением, помещением в тяжелую воду, облучением и т.д. С точки зрения модели существенно только одно — в процессе воздействия описание системы уравнениями (3.1) не годится, так как в них пришлось бы включать экспериментатора.

Другое дело, когда воздействие прекращается и систему «предоставляют ее судьбе». В этот момент результат воздействия

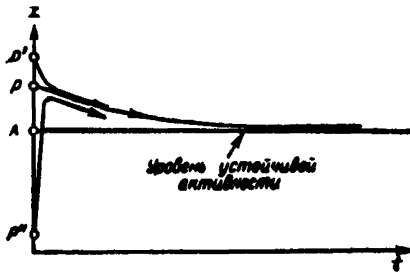


Рис. 8. Сохранение активности и выход на устойчивый режим. Воздействия P , P' , P''

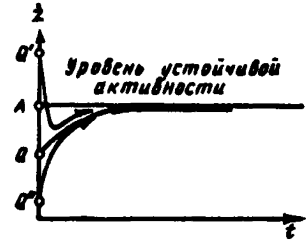


Рис. 9. Другой возможный тип выхода на устойчивый режим. Воздействия Q , Q' , Q''

сводится к тому, что система оказывается в неравновесном состоянии. Дальнейшее поведение системы, ее адаптация определяются уже только внутренними свойствами. Такая точка зрения означает, конечно, отказ от анализа воздействий (представляющего самостоятельную задачу) и концентрацию внимания на механизме адаптации.

Математически это точно соответствует задаче Коши (задача с начальными данными) для системы уравнений (3.1). Поэтому в дальнейшем под «воздействием» мы будем понимать перенесение изображающей точки системы в одну из точек ее фазовой плоскости. Более того, мы будем говорить о «воздействии P », понимая под этим, что система оказалась в точке P в результате некоторого воздействия.

Классификация воздействий оказывается, с этой точки зрения, равносильной классификации типов поведения интегральных кривых системы, что соответствует разбиению фазовой плоскости системы на области однотипного поведения.

В примере, изображенном на рис. 7, таких областей оказывается три: зоны активности, покоя и гибели системы. Приведем типичные формы поведения во времени наблюдаемой величины z при воздействиях различных видов. Во всех случаях наблюдается характерный быстрый переходный процесс, соответствующий выходу на состояние квазиравновесия. Обозначения на рис. 8-11 соответствуют обозначениям рис. 7.

5. Устойчивость и адаптивность

Проведенный анализ показывает, что реакцию системы на внешние воздействия следует характеризовать по крайней мере двумя различными показателями — устойчивостью и адаптивностью.

Устойчивость системы тем больше, чем дальше отстоит точка A от «опасной» точки срыва C . В этом случае необходимы

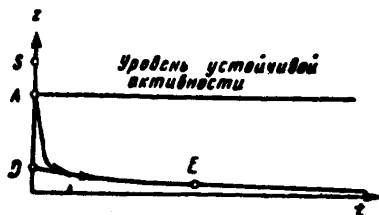
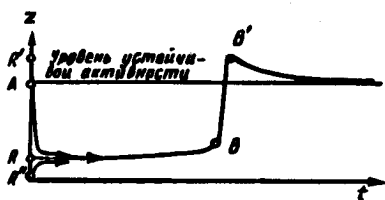


Рис. 10. Лаг-фаза BV и эволюция вдоль линии $B'A$

Рис. 11. Необратимые изменения в результате воздействия S . Гибель системы

значительные воздействия, чтобы «выбить» систему из состояния активности (рис. 12).

Иначе «переносят» сильные воздействия адаптивные системы. Сначала они «впадают в транс», но, «отлежавшись», восстанавливают состояние активности. Запас адаптивности тем больше, чем больше дуга BD (рис. 13).

Поучительно сравнение крайних ситуаций — устойчивости без адаптивности и адаптивности без устойчивости.

Устойчивая система без адаптивности (точка D слилась с точкой B) может переносить сильные воздействия, сохраняя активность. Однако любое воздействие, приводящее к шоку, равносильно гибели системы, так как при эволюции вдоль линии DE в системе развиваются необратимые изменения и она никогда уже не возвращается в состояние активности.

Наоборот, хорошо адаптивная система без устойчивости (точка A на дуге BC) сохраняет способность выходить из шокового состояния. Она, правда, теряет способность сохранять активность и периодически* впадает в шок спонтанно, без внешнего воздействия, но все же это не гибель системы.

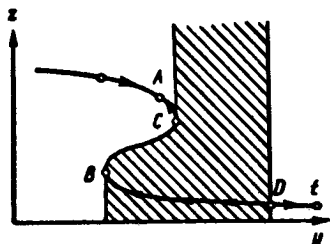
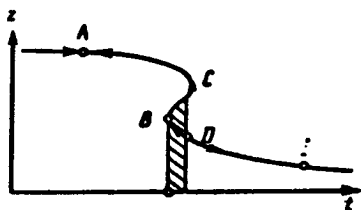


Рис. 12. Устойчивая, но малоадаптивная система

Рис. 13. Адаптивная система с малым запасом устойчивости

* Возникает колебательный режим, который ничем не отличается от релаксационных колебаний, изображенных на рис.1.

Следовательно, устойчивость и адаптивность — это принципиально разные пути стабилизации систем.

Устойчивые системы больше и продуктивнее работают. Они хорошо соответствуют благоприятным условиям, но быстро гибнут в неблагоприятных.

Адаптивные работают хуже и часто «замирают», однако они способны работать в трудных условиях.

6. Эволюционное значение адаптивных колебаний

До сих пор термин «эволюция» употреблялся в узкотехническом смысле — медленные движения в системе (3.1).

Более широкое понимание этого термина — еще более медленное изменение вида правых частей (в частности, коэффициентов) системы уравнений — лучше соответствует смыслу, который вкладывают в это слово биологи.

При таком понимании можно ставить вопрос о взаимодействии системы со средой, рассматривая систему и среду как части более широкой системы.

Правдоподобно, что в такой постановке задачи можно доказать постепенное развитие адаптивности у малоадаптивных систем, помещенных в неблагоприятную (но не гибельную) среду.

Наоборот, адаптивные системы в благоприятной среде будут увеличивать устойчивость и работоспособность, конечно, за счет снижения адаптивности.

Ожидание справедливости подобных утверждений тем более оправдано, что сходная задача в дискретной трактовке (игры автоматов) была поставлена М.Л. Цетлиным [5] и доказаны похожие утверждения.

Рассмотрим теперь еще более общую схему, учитывающую пространственную неоднородность среды.

Допустим, что в некотором направлении условия ухудшаются и существует граница, за которой условия становятся губительными для существующей «колонии» систем. В такой ситуации, в колонии постепенно возникает «градиент адаптивности». Ближе к границе системы будут становиться все более адаптивными, «расплачиваясь» за это своей устойчивостью. Новые, более адаптивные системы смогут проникнуть за старую границу. Эта экспансия будет остановлена только исчерпанием «запаса устойчивости». На новой границе смогут существовать только системы, адаптивные без устойчивости. Но мы уже знаем, что в этом случае система неминуемо становится автоколебательной.

Это обстоятельство создает совершенно новую ситуацию для колонии в целом. Для отдельной системы (клетки) автоколебания — свидетельство крайне неблагоприятных условий, а для колонии в целом они могут быть организующим фактором. В частности, ближайшие соседи, с их ничтожно малым запасом

устойчивости, будут втянуты в автоколебания. В таких условиях возможны резонансные образования [6] надклеточного уровня.

Очень соблазнительно сопоставление возникающей автоколебательности со спонтанной электрической активностью нервных клеток. Разумеется, сравнение должно быть эволюционно грамотным. Можно сравнивать либо с объектами, стоящими на грани клетки и многоклеточного организма, либо с этапом эмбриогенеза, на котором происходит закладка нервных клеток. Сравнение с эволюционно зрелыми нервными клетками, где первоначальная «бедственность» морфологически закреплена и получила смысл сигнала, вряд ли будет продуктивным.

Литература

1. Бенсон С., Основы химической кинетики, «Мир», 1964.
2. Тихонов А.Н., Матем. сб., 22, 193, 1948.
3. Понтрягин Л.С., Изв.АН СССР. Сер. матем., 22, 193, 1957.
4. Жаботинский А.М. В сб.: Колебательные процессы в биологических и химических системах, М., «Наука», 1967, с.149.
5. Цетлин М. Л., Докл. АН СССР, 149, 284, 1963.
6. Молчанов А.М. В сб.: Колебательные процессы в биологических и химических системах. М., «Наука», 1967, с.274.

УПРАВЛЕНИЕ И АДАПТАЦИЯ

(Эволюционный аспект проблемы управления)

Обсуждается гипотеза о связи управления и адаптации. Управление — это форма адаптации на уровне «коллективов», существенно связанная с распадом его на «пассивно-устойчивую» и «адаптивную» популяции и возникающая в особо критических условиях. Предложена математическая модель эволюционного возникновения управления. Указаны трудности изучения этой модели и значение предельных случаев, допускающих исследование частными методами.

Введение

Понятие управления получило широкое распространение в научной и особенно научно-популярной литературе. Затруднительно даже перечислить те многообразные оттенки вплоть до противоречивых, которыми снабжают это понятие разные авторы. Существуют три главные, весьма обширные группы работ, внутри каждой из которых термин «управление» понимается достаточно единообразно.

В теории оптимального регулирования, дисциплине в сущности чисто математической, имеется четкое определение понятия «управление». Так называют свободные параметры $u(t)$ в системе уравнений, определяющих зависимые переменные $x(t)$:

$$\frac{dx}{dt} = a(x, u).$$

Главный акцент этого определения — независимость управляющих параметров $u(t)$ от управляемых переменных $x(t)$.

Другие аспекты, например, эффективность управления, т.е. величина «усилия» Δu , необходимого для получения требуемого результата Δx , имеют второстепенное значение. Эффективность может быть, в частности, столь угодно малой и это не изменит формально безупречного деления переменных на управляющие и управляемые.

Иначе и несколько менее определенно понимают управление в биологии, в частности, в биохимии. Вместо термина «управление» часто используют термины «контроль» или «регуляция», понимая под ними процесс или результат (или и то и другое) деятельности ферментных систем самого разного типа. Акцент такого понимания управления иной. На первый план здесь выдвигается именно эффективность малых концентраций веществ-регуляторов, их специфичность. Настаивать на строгой независимости управляющих агентов от управляемых становится невозможным — и те и другие суть концентрации разных химических соединений, участвующих в общей реакции.

Наиболее абстрактно понятие управления в третьей группе работ, в кибернетике, где оно равнозначно понятиям «сообщения» или «сигнала». Это означает двойную идеализацию. Идеализация заключена, во-первых, в предположении, что сигнал вызывает требуемое изменение без всяких энергетических или материальных затрат. С другой стороны, постулируется полная независимость управления от управляемых систем. Последняя абстракция, впрочем, несколько смягчается введением понятия «обратной связи», когда допускается воздействие результатов управления на последующие сигналы. Такой подход вводит (из-за излишней жесткости дискретной схемы, столь популярной в кибернетике) принудительное запаздывание (на один «такт») совершенно необязательное не только в реальных системах, но даже и в других математических моделях.

Перечисленные различия достаточно серьезны. Они делают сомнительным существование единого понятия управления, годного для всех приложений. Как и всякое научное понятие, понятие управления описывает крайние, предельные ситуации. В естествознании такие ситуации реализуются только приближенно. Возникает поэтому принципиальный методологический вопрос о границах применимости понятия «управления».

Речь идет вовсе не о формальной строгости терминологии. Вопрос о границах применимости имеет вполне содержательную эволюционную форму. Так как свойства систем, которые мы сейчас называем «управлением» или «регуляцией», возникли не сразу, а прошли длинный путь эволюционного развития, то управление должно иметь своего эволюционного предшественника.

В статье высказывается и аргументируется гипотеза, что управление является развитием и крайней, асимптотической формой адаптации.

1. Математическая модель

В предыдущей статье [1] показано, что понятия, связанные с явлением адаптации, такие как «работоспособность», «шок», «покой», «приспособление» и т.д., могут быть (в простейшей

форме) смоделированы системой всего лишь двух уравнений. Правда эти уравнения должны быть существенно нелинейны и содержать параметры, в том числе обязательно один малый параметр ϵ ,

$$\left. \begin{aligned} \frac{dy}{dt} &= b(y, z, \epsilon, \alpha) \\ \epsilon \cdot \frac{dz}{dt} &= c(y, z, \epsilon, \alpha) \end{aligned} \right\} \quad (1.1)$$

но зато это обыкновенные дифференциальные уравнения, для которых математиками развита содержательная качественная и, частично, количественная теория.

Это обстоятельство, важное для дальнейшего, можно пояснить ссылкой на другую статью автора [2]. В ней показано, что поведение даже очень сложной системы резко упрощается в критической ситуации. В условиях, когда системе грозит потеря устойчивости (в каком бы то ни было смысле), число определяющих переменных сокращается до двух (когда система проходит через комплексный корень, т.е. через колебательный режим при потере устойчивости), а иногда даже до одного. Поэтому самые главные особенности поведения любых систем определяются не столько их внутренней структурой, сколько характером критических условий, в которые эти системы могут попадать. Гипотеза, ради которой написана эта статья, состоит в том, что управление есть одна из самых эффективных форм адаптации на уровне «коллективов».

Таковыми биологическими коллективами могут быть тканевые культуры или колонии одноклеточных, стада животных и стаи птиц, «государства» общественных насекомых.

Главный шаг в построении модели состоит в допущении, что поведение подобных коллективов может быть смоделировано системой большого числа однотипных уравнений:

$$\left. \begin{aligned} \frac{dy_i}{dt} &= b_i(y_i, z_i, \epsilon, \alpha) + \delta B_i(y_1 \dots y_n, z_1 \dots z_n, \alpha) \\ \epsilon \frac{dz_i}{dt} &= c_i(y_i, z_i, \epsilon, \alpha) + \delta C_i(y_1 \dots y_n, z_1 \dots z_n, \alpha) \end{aligned} \right\} \quad (1.2)$$

Индекс i — номер особи, правые части b_i и c_i описывают поведение отдельно взятой особи, а B_i и C_i задают взаимодействие особей друг с другом и внешней средой, представленной параметром α .

В этой системе появляется еще один малый параметр δ . В отличие от ϵ , характеризующего внутренние свойства особей, этот параметр задает величину взаимодействия их друг с другом и является в некотором смысле мерой индивидуальности каждой особи. Соответствующий большой параметр,

$$T = \frac{1}{\delta}, \quad (1.3)$$

задает по порядку величины «время жизни» отдельной особи (измеренное величиной характерного цикла жизнедеятельности).

Полученная модель характеризуется тремя масштабами времени. Средний масштаб, время порядка единицы, соответствует собственным временам жизнедеятельности отдельной особи. Назовем условно соответствующие изменения «физиологическими». Малые времена $\Delta t \sim \varepsilon$ — это времена адаптивных скачков в том смысле, в котором это явление рассмотрено в статье [1]. Наконец есть большое время, задаваемое формулой (1.3), в течение которого происходят существенные изменения с каждой отдельной особью.

2. Эволюция системы в неблагоприятных условиях

Разберем эволюцию системы (1.2) при помещении ее в неблагоприятные условия. Отложим временно кардинальный вопрос о том, какое именно отношение имеют написанные ниже квазибиологические слова к математической модели и почему эти слова что-либо описывают в биологии. Проще всего считать, что этот пункт содержит отдельную, логическую модель и оставить открытым вопрос о соответствии моделей друг другу и каждой из них «объясняемому» явлению.

Итак, пусть колония особей находится в неблагоприятной среде. Возможны два пути приспособления — один за счет увеличения устойчивости, другой за счет увеличения адаптивности. Каждый из них имеет свои достоинства и недостатки.

Увеличение устойчивости полезно отдельной особи в благоприятных условиях. Однако колония в целом будет занимать меньшую площадь, так как граница «шока» будет для нее границей гибели.

Увеличение адаптивности уменьшает активность каждой особи, так как часть времени адаптивные особи проводят в состоянии «шока», в который они впадают даже в условиях благоприятных для более устойчивых особей. Однако колония таких особей может захватить области абсолютно непригодные для более устойчивых.

Стоит заметить, что приведенные, почти очевидные, соображения основывались целиком на свойствах отдельных особей. Взаимодействие особей не играет никакой роли в подобных рассуждениях.

Следующий этап существенно опирается на взаимодействие. Предположим, что среда переменна (в пространстве или времени — это все равно, такие задачи двойственны друг другу) и особи способны взаимодействовать.

В этом случае изменение среды повлечет за собой сдвиг свойств особи либо в сторону аддитивности, либо в сторону устойчивости. Направление этого сдвига определяется типом взаимодействия.

Возникает ряд интересных вопросов. Каким должно быть взаимодействие, чтобы ухудшение условий вызвало общий сдвиг в адаптивность? Можно ли добиться тех же результатов (в сохранении жизнеспособности колонии) на пути пассивного увеличения устойчивости? Зависит ли результат от скорости ухудшения обстановки или просто от степени неблагоприятности условий?

Любопытно, что многие из подобных вопросов можно разобрать на важном частном случае модели, когда все особи тождественны друг другу. В этом случае так сказать «синхронной», точнее говоря «гомогенной», однородной культуры взаимодействие можно учесть введением еще одной (или нескольких) переменных и дело сводится к исследованию качественного поведения систем трех-четырех уравнений. Подобные задачи вполне посильны для современного анализа, опирающегося (при случае) на вычислительную технику.

Однако значительно интереснее ситуации, в которых гомогенные колонии выжить не могут. Так как гомогенные колонии являются частным случаем гетерогенных, то заведомо существуют переменные среды губительные для гомогенных систем, но в которых возможно существование систем гетерогенных, неоднородных. Простейшим типом такой системы является колония, часть особей которой сдвинуты в сторону адаптивности, а другая в сторону устойчивости. Если одна из популяций совпадает со всей колонией, то система становится гомогенной.

Если работать вблизи границы гибели гомогенных колоний, то наверняка можно найти условия, вызывающие расщепление колонии на адаптивную популяцию и устойчивую популяцию.

Именно такая ситуация является критической для интересующей нас проблемы эволюционного возникновения «управления».

3. Трудность проблемы

Довольно ясно, как следовало бы ставить задачу дальше. Почти очевидно, что адаптивная популяция — это и есть прообраз нервной сети в ткани или популяции сторожевых муравьев в муравейнике. Роль сигнала играет «впадение» адаптивных особей в шоковое состояние (или, наоборот, выход из него). Однако и без этого мы зашли слишком далеко в спекулятивных построениях. Изложенные соображения не являются теорией — их невозможно проверить. Дело обстоит даже хуже — к ним невозможно найти противоречащие факты. Трудность здесь двойного сорта. Теоретическое исследование предложенной модели находится далеко за пределами современных математических

методов. Во многих случаях это не пугает исследователя — современные вычислительные средства настолько могучи, что могут нередко заменить теоретический анализ. К сожалению не в нашем случае.

Несложная оценка показывает необозримость задачи. Для того, чтобы качественная картина хоть сколько-нибудь могла быть получена, малые параметры ε и δ следует выбрать порядка одной сотой, не больше

$$\delta = 0,01,$$

$$\varepsilon = 0,01.$$

Выберем очень скромную колонию,

$$n = 100.$$

Получается, таким образом, система двухсот уравнений, ибо каждой особи соответствует два уравнения. Даже если ограничиться парным взаимодействием и не рассматривать более сложные типы его, каждое из этих двухсот уравнений будет содержать в правой части сто существенно нелинейных членов.

Для сравнения стоит указать, что даже куда более скромную задачу интегрирования системы уравнений небесной механики — девять уравнений для девяти планет Солнечной системы — удается провести только на интервале в несколько тысяч лет.

Дальнейшему счету препятствует нарастание ошибок округления, «забывающих» результат. Заметим, что десять тысяч лет как раз соответствуют времени

$$T = \frac{1}{\delta} = 100,$$

так как в качестве «года» в нашем случае следует принять время существенного изменения самой быстрой переменной

$$\tau = \varepsilon = 0,01.$$

Однако на таких временах еще нельзя ожидать не только расщепления колонии, но даже существенных сдвигов адаптивности гомогенной колонии.

Эта грубая оценка показывает насколько далеки от количественного исследования даже простейшие эволюционные проблемы при попытке моделировать их «в лоб», минуя упрощающий теоретический анализ. Аналоговые устройства, с их чрезвычайно высоким уровнем шумов (даже лучшие из них не обеспечивают точности в три знака) для подобных задач абсолютно непригодны. Не удивительно, поэтому, что в большинстве исследований ограничиваются анализом уже существующих систем с управлением, обходя проблему их эволюционного возникновения. Но даже и это исследование обычно основано на весьма сильных упрощениях. Например, приближение рефракторной среды, соответствующей

щее кибернетической схеме «управление \equiv сигналу» и учитывающее только геометрию системы.

Вполне понятно, конечно, что во многих случаях таких предположений оказывается достаточно. Однако вопрос эволюционного возникновения разнообразных морфологических структур, связь с прохождением критического уровня в эволюции, воспроизведение этой кризисной ситуации в индивидуальном развитии имеет не только принципиальное, но и практическое значение.

В задачу статьи не входит попытка ответа на поставленные вопросы. Ее задачи намного скромнее — проанализировать методологические корни абстрактности подходов к проблеме управления в биологических системах. С точки зрения математики речь идет об асимптотических задачах с малым параметром при старшей производной, причем заранее известно, что никакие асимптотические методы не пригодны, так как необходим по крайней мере двойной и даже тройной предельный переход. Растет время ($t \rightarrow \infty$), увеличивается число уравнений ($n \rightarrow \infty$), а также стремится к нулю ($\epsilon \rightarrow 0$) малый параметр. В лучшем случае разобраны подобные задачи с одним лишь предельным переходом.

Но если все так плохо, то зачем писать об этом? Смысл данной статьи заключается в постановке задачи. Сейчас нереально исследование проблемы в общем виде. Однако нередко бывает так, что конкретные частные задачи удается проинтегрировать (или «проанализировать»), используя их частные свойства. Хороший подбор таких задач иногда заменяет (на время, конечно) общую теорию. Очень важно поэтому разбирать максимально разнообразные частные задачи — автору пока неизвестна ни одна, — позволяющие с разных сторон подойти к общей проблеме.

Литература

1. Молчанов А. М., Биофизика, 15, 352, 1970.
2. Молчанов А. М., Препринт, ИПМ, М., ТО1817.

ЭНДОГЕННЫЕ БИОХИМИЧЕСКИЕ КОЛЕБАНИЯ КАК ВОЗМОЖНАЯ ОСНОВА ФИЗИОЛОГИЧЕСКИХ РИТМОВ

Показано, что период автоколебаний может быть чувствителен к небольшим изменениям параметров биохимической системы, например, активности ключевого фермента. Выдвинуто предположение о возможности эволюционного закрепления в циркадных ритмах свойств именно таких критических эндогенных колебательных систем.

1. Введение

В литературе, ныне уже многочисленной [1], посвященной биологическим часам и их математическим моделям, обсуждается вопрос об эволюционном происхождении биологических ритмов. Нередко при этом противопоставление [2] циркадных ритмов (рассматриваемых как вынужденные) эндогенным. Считается почти само собой разумеющимся, что эти ритмы имеют разную природу. Неявно подразумевается, что огромные различия в периодах — сутки для циркадных ритмов и доли секунды для метаболических — автоматически снимают даже постановку вопроса об их общности.

Существуют тем не менее серьезные биологические и математические основания для обсуждения вопроса об эволюционной вторичности циркадных ритмов — ритмов, вызванных, в основном, суточной периодичностью светового потока. Однако фотосинтез возник, вне всякого сомнения, эволюционно значительно позже других метаболических систем (например, гликолитической). Эти системы имели собственные ритмы, никак не связанные с астрономическими явлениями. Естественно поэтому думать, что циркадные колебания возникали на основе уже имевшейся эндогенной ритмики. Но если качественно подобные соображения реальны, то количественная пропасть,

$$N = \frac{1 \text{сутки}}{1 \text{секунда}} = 86400,$$

которая разделяет метаболические и циркадные ритмы, кажется непреодолимой. Для большинства известных физических явлений или технических устройств подобное количественное различие заведомо означает разную природу колебаний.

Цель настоящей работы — показать, что биохимические системы могут охватывать огромные диапазоны частот (или амплитуд) на основе единого механизма небольшим изменением его параметров. На организменном уровне хорошей иллюстрацией могут служить органы зрения и слуха. Причина, говоря математически, состоит в существенной нелинейности биологических, в частности биохимических, систем.

2. Модель гликолитической системы

В работе [3] построена математическая модель гликолитической системы. При некоторых упрощающих предположениях система дифференциальных уравнений, описывающих кинетику гликолиза, имеет вид:

$$\left. \begin{aligned} \frac{dx}{dt} &= \beta - xy^2 \\ \frac{dy}{dt} &= xy^2 - y \end{aligned} \right\} . \quad (1)$$

Эта система была численно исследована на ЭВМ с достаточной для практики (с учетом приближенности модели) точностью. Однако теоретический анализ был проведен лишь в линейном приближении, а качественная картина нарисована ошибочно*. Между тем полный анализ системы (1) весьма поучителен. Он имеет непосредственное отношение к обсуждаемой теме о взаимосвязи эндогенных и циркадных ритмов. Самое существенное в этом анализе — изучение системы в целом. Эта особенность психологически резко противоречит привычной практике исследования физических систем, когда достаточно бывает изучить малую окрестность стационарного режима. Корень упомянутой ошибки, заметим в скобках, именно в привычке к локальному исследованию.

3. Особые точки «на бесконечности»

Фазовый портрет системы в целом нагляднее изображать не в исходных переменных, а в переменных, аналогичных переменным Пуанкаре:

* За что автор настоящей статьи (под чьим руководством была выполнена эта работа) также несет ответственность. Ошибка в основном исправлена в предлагаемой вниманию читателя статье. Предполагается более подробная публикация.

$$\begin{aligned}\xi &= \frac{x}{1+x^2+y^2} \\ \eta &= \frac{y}{1+x^2+y^2}\end{aligned}\quad (2)$$

отображающих всю плоскость (x, y) на единичный круг плоскости (ξ, η) . Для наших целей достаточно ограничиться рассмотрением полукруга, так как траектории, начинающиеся в верхней полуплоскости, расположены в ней целиком. Вниз их «не пускает» решение

$$y = 0, \quad (3)$$

которое имеет наша система.

В полукруге $\eta \geq 0$ изучаемая система имеет пять особых точек (рис. 1). Одна из них, точка F , соответствующая стационарному режиму

$$\left. \begin{aligned}\beta - xy^2 &= 0 \\ xy^2 - y &= 0\end{aligned} \right\}, \quad (4)$$

расположена внутри полукруга, остальные на границе. Достоинство переменных Пуанкаре именно в том, что становятся «ясно видимыми» эти граничные особые точки. В исходных переменных легко находится только стационарная точка F :

$$\left. \begin{aligned}x &= \frac{1}{\beta} \\ y &= \beta\end{aligned} \right\}, \quad (5)$$

а остальные являются бесконечно удаленными. Между тем кинетика системы определяется структурой как раз этих точек.

Результат исследования этих точек, детали которого опускаем, представлен на рис. 1. Бесконечно удаленные особые точки сохраняют свой тип при всех значениях параметра β . Точка же F , соответствующая стационарному режиму и устойчивая при $\beta > 1$, теряет устойчивость, когда β проходит через «линейное» критическое значение

$$\beta = 1. \quad (6)$$

По мере приближения β сверху к единице стационарный режим становится все менее устойчивым. В критической точке происходит «мягкое» рождение автоколебательного режима — именно в этом смысле речь идет о «линейном» критическом значении. Предельный цикл рождается с нулевой амплитудой и «наследует» период малых линейных колебаний — так называемое характерное время системы. По мере роста амплитуды автоколебаний растет различие между периодом предельного цикла и «характерным временем» системы (рис. 2).

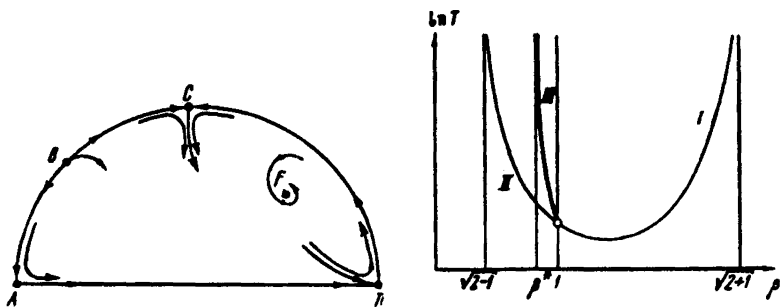


Рис. 1. Особые точки в полукруге Пуанкаре. При $\beta > 1$ стационарная точка F устойчива

Рис. 2. Зависимость в логарифмической шкале различных периодов от параметра β . I — период линеаризованных колебаний вблизи устойчивого стационарного режима; II — то же для неустойчивого режима (при $\beta < 1$); III — период автоколебаний

Понятно поэтому, что самый термин «характерное время» системы весьма неудачен в подобных ситуациях, так как оно характерно только для некоторой малой окрестности стационарного (теперь уже неустойчивого) режима.

4. Критический нелинейный режим

Численные методы позволяют исследовать поведение предельного цикла в зависимости от β сразу после рождения автоколебательного режима*. Однако прямой численный счет очень быстро становится невозможным, насколько стремительно растут (с уменьшением β) период и амплитуда автоколебаний. Возникает необходимость в теоретическом анализе ситуации. Этот анализ выявляет существование еще одного, «нелинейного», критического значения параметра

$$\beta = \beta^*, \quad (7)$$

определяемого условием слияния двух сепаратрисс. Одна из них, CN , выходит из седла C и отделяет поток, идущий из узла β от потока, проходящего в треугольнике CD . Вторая сепаратрисса MD разделяет входящие и проходящие кривые в окрестности сложной особой точки — седлоузла D .

Слияние сепаратрисс соответствует «равновесию» двух потоков. Поток из B целиком вливается в D . Поток же из неустойчиво-

* Сообщение Л. М. Когана, Л. В. Луневской, А. М. Молчанова и Е. Е. Селькова доложено на Ученом совете Института биологической физики АН СССР в Пущине в 1970 г. Авторы благодарны Г. А. Ососкову за помощь.

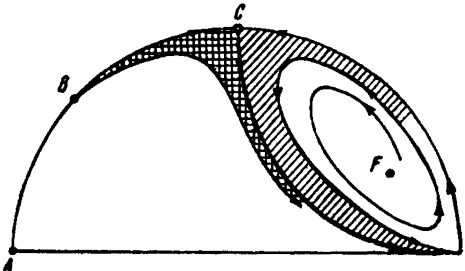
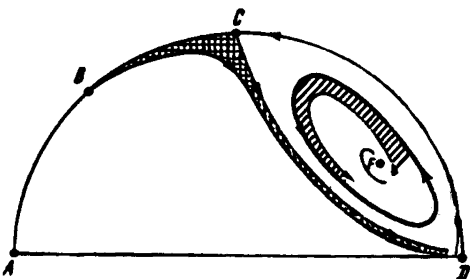
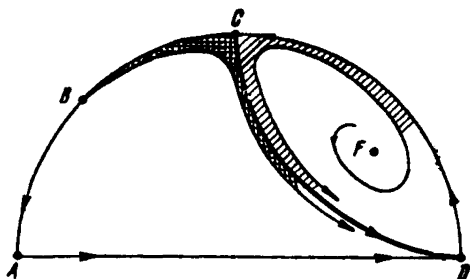


Рис. 3. Сепаратриссы образуют полуустойчивый предельный цикл — «бычий глаз» по терминологии Пуанкаре

Рис. 4. Образование предельного цикла при разрыве петли сепаратрисс

Рис. 5. Поток F вытеснил траектории потока B из двуугольника CD

го фокуса F заполняет криволинейный двуугольник CD , закручиваясь изнутри к полуустойчивому предельному циклу, составленному из двух сепаратрисс. Одна начинается в C и кончается в D , другая — целиком расположенная в бесконечности! — начинается в D и кончается в C . Эта замечательная ситуация изображена на рис. 3.

Изменение параметра β сдвигает равновесие. Небольшое увеличение этого параметра приводит к «затеканию» внутрь двуугольника CD тонкой струйки потока B . Она вытесняет поток F из окрестности точки D . Возникает предельный цикл, который тем больше, чем ближе значение β к критическому (рис. 4).

Уменьшение β приводит к обратному эффекту. Поток F «прорывается» за сепаратриссу, втекающую в D , и целиком вливается в эту особую точку. Чем ближе β к β^* , тем больше оборотов делают кривые потока F перед тем, как устремиться в D . Однако это единственное воспоминание о разрушившемся предельном цикле, который уже не существует при $\beta < \beta^*$ (рис. 5).

Численное отыскание критического значения β^* представляет собой непростую вычислительную задачу, включающую, в частности, отыскание асимптотических разложений сепаратрисс точек C и D . Кроме того, неустой-

чивость счета из «седла в седло» заставляет прибегать к специальным вычислительным приемам.

Тем не менее с принципиальной точки зрения критическое значение $\beta = \beta^*$ сходно с «линейным» критическим значением $\beta = 1$ тем, что оба они определяют границы устойчивости стационарных режимов. Первое относится к предельному циклу, второе — к стационарной точке. И то и другое является поэтому корнями уравнения

$$\text{Re} \lambda(\beta) = 0, \quad (8)$$

где $\lambda(\beta)$ — собственное значение, определяющее малые колебания вблизи стационарного режима.

Однако устойчивость стационарной точки определяется исследованием алгебраического уравнения. Это свойство локальное. Устойчивость же предельного цикла приводит к изучению интегральных уравнений. Это свойство глобальное. Ясно, что подобные задачи, типичные, по-видимому, для биохимических систем, интереснее, труднее и разнообразнее.

5. Заключение

Разобранный пример демонстрирует важную особенность биохимических систем — резкую зависимость характеристик автоколебательного режима от параметров системы. Так, например, период автоколебаний T успевает измениться от 2π до ∞ :

$$2\pi < T < +\infty,$$

при уменьшении β от первого («линейного») критического значения до второго («нелинейного»)

$$1 > \beta > \beta^*.$$

Вычисления показывают, что β^* всего лишь на десяток процентов

$$\beta^* \approx 0,9,$$

отличается от единицы — «линейного» критического значения. Конечно, не любая нелинейная система обладает столь поразительной лабильностью периода. Изученная система имеет индивидуальную особенность — она относится к классу негрубых систем. Одна из ее особых точек (а именно D) при всех значениях параметра β остается сложной особой точкой — седлоузлом — не распадаясь на седло и узел. Практически любое уточнение этой, несомненно, приближенной модели неминуемо приведет к такому распаду. Стабилизируется, конечно, и период автоколебательной системы — изменчивость его уже не будет столь резкой.

Разобранная система является частным случаем значительно более широкого класса «проточных» систем:

$$\begin{aligned}\frac{dx}{dt} &= S(x,y) - A(x,y) \\ \frac{dy}{dt} &= A(x,y) - P(x,y)\end{aligned}\tag{9}$$

Эти уравнения описывают кинетику следующих событий. Откуда-то извне (и в рамках изучаемой задачи неважно, откуда и как именно) в систему направлен поток S «субстрата реакции x ». Затем включается «механизм A переработки x в полуфабрикат». Окончательно вещество y «механизмом P » перерабатывается в некий «целевой продукт», природа и судьба которого нас не интересует. Он, в частности, может совпадать с y , и механизм P (в этом случае) есть механизм вывода продукта из системы.

В нашем конкретном случае

$$\begin{aligned}S &= \beta \\ A &= xy^2. \\ P &= y\end{aligned}\tag{10}$$

Эта система хорошо иллюстрирует общий тезис [4] о «трудных условиях» как причине возникновения автоколебаний. В данном случае автоколебания возникают тогда, когда систему «сажают на голодный паек».

В системе «заложено» конкурентное противоречие между механизмами A и P , каждому из которых необходим «полуфабрикат» y . Но механизма A вырабатывает это соединение, а поток P его из системы выносит. К тому же «квадратичный» механизм A сильнее «линейного» механизма P при больших концентрациях y , а при малых, наоборот, P сильнее A .

В результате при больших ($\beta > 1$) потоках субстрата механизм A успевает «наработать» достаточное количество y и для себя и для P . Система работает в стационарном режиме.

Иное дело малые ($\beta < 1$) потоки S . Заметного оживления A приходится ждать долго — пока в систему не «натечет» достаточное количество субстрата x . Однако увеличение производства y вызывает его быстрый выброс потоком P . Производительность A снова падает. Система «впадает» в колебательный режим.

Если еще уменьшить S , повреждение механизма A становится необратимым (при $\beta < \beta^*$). Поток P монотонно падает, происходит медленное неограниченное накопление субстрата x в системе. Работать она уже не способна.

Эти простые соображения допускают эволюционную интерпретацию. Можно предположить, что первоначально, до возникновения фотосинтеза, существовали разнообразные метаболические системы, в том числе и «влачившие жалкое существование». Их колебательный характер не только не приносил им никакой пользы, но был, так сказать, «клеямом нищеты». Среди них были самые разнообразные, и в частности с периодами, близкими

к суточным. Эти последние были совсем «слабыми» (огромный период в нашем примере означает близость к роковой границе). Однако после возникновения фотосинтеза с его вынужденным суточным ритмом выжили именно они, резонировавшие с этим внешним ритмом. Здесь не место для обсуждения деталей конкретного механизма выживания.

Суть дела в предоставлении широкого поля деятельности для главной творческой силы — естественного отбора. Именно трудные условия (в данном случае падение потока субстрата) создают широкий диапазон свойств (в данном случае неограниченный набор периодов автоколебаний) над которыми «может поработать» естественный отбор.

* * *

Автор благодарен за полезное обсуждение многим лицам, особенно В. Л. Давыдову, Г. П. Крейцеру, Г. А. Ососкову, Е. Е. Селькову, Э. Э. Шнолю, однако они не несут ответственности за рискованные, быть может, построения заключительного параграфа.

Литература.

1. *Aschoff J.* Ed. *Ciradian Clocks*, North Holland Publ. Co., Amsterdam, 1965.
2. *Brown F. A.* см. [1], стр. 231.
3. *Сельков Е. Е.*, Мол. биол., 2, 252, 1968.
4. *Молчанов А. М.* В сб.: Колебательные процессы в биологических и химических системах, стр. 274, М., «Наука»; 1967.

КОЛЕБАТЕЛЬНЫЕ РЕЖИМЫ КАК ЧАСТНЫЙ СЛУЧАЙ КРИТИЧЕСКИХ РЕЖИМОВ

Сам факт созыва Второго Всесоюзного симпозиума по колебательным процессам в биологии достаточно многозначителен. Теперь уже нет сомнения, что большое число исследователей в различных областях биологии понимают и ценят «колебательный стиль мышления». Поэтому агитация в пользу того, что «такое бывает» уже перестала быть первоочередной задачей. Колебания, с другой стороны, еще не стали модой и поветрием, и было бы очень хорошо, если бы они и не стали ими никогда. Из других примеров стремительного внедрения новых идей в многострадальную биологию очень хорошо известно, сколь велики ненужные издержки.

Главный стимул настоящего сообщения — убеждение, что своевременно и четко поставленный вопрос о границах применимости новой идеи или теории помогает сохранить их внутреннюю научную ценность и уберечь от возможной инфляции.

В общей форме вопрос о границах применимости неясно как поставить. Разумно поэтому ограничение более узким вопросом — о роли колебательных режимов в математических моделях. Пусть имеется какой-нибудь объект (не обязательно биологический), для которого построена математическая модель, описывающая его поведение. Ограничимся еще более частным случаем, когда эта модель есть система обыкновенных дифференциальных уравнений:

$$\begin{aligned}\frac{dx_1}{dt} &= f_1(x_1, \dots, x_k; \alpha_1, \dots, \alpha_l) \\ \frac{dx_2}{dt} &= f_2(x_1, \dots, x_k; \alpha_1, \dots, \alpha_l) \\ &\dots\dots\dots \\ \frac{dx_k}{dt} &= f_k(x_1, \dots, x_k; \alpha_1, \dots, \alpha_l)\end{aligned}\tag{1}$$

или сокращенно,

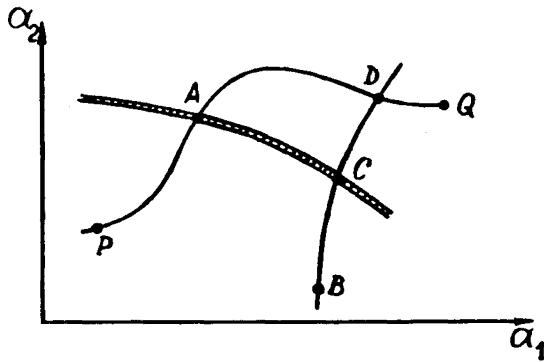


Рис. 1. Структурный портрет системы. Линии AC и BC — границы устойчивости. Линия PARQ — путь развития системы.

$$\frac{dx}{dt} = f(x, \alpha), \quad (1)$$

зависящая от параметров $(\alpha_1, \dots, \alpha_l) = \alpha$

Предположим, что в пространстве параметров имеется область, где система имеет устойчивый стационарный режим.

Теория устойчивости Ляпунова позволяет, в частности, найти границы устойчивости — линии нейтральности. Для этого нужно найти стационарную точку системы (1)

$$f(x_0, \alpha) = 0, \quad (2)$$

$$x_0 = x_0(\alpha), \quad (3)$$

линеаризовать систему в этой точке, т.е. найти матрицу ее частных производных

$$A(\alpha) = \left. \frac{\partial f}{\partial x} \right|_{x=x_0} = \left. \frac{\partial f_m}{\partial x_n} \right|_{x=x_0} \quad (4)$$

и, решив характеристическое уравнение

$$\det|A - \lambda E| = 0, \quad (5)$$

найти все k собственных чисел нашей системы, которые будут, конечно, функциями параметров: $\alpha_1, \dots, \alpha_l$

$$\lambda_m = \lambda_m(\alpha_1, \dots, \alpha_l) \quad 1 < m < k. \quad (6)$$

«Добрая половина» этих собственных чисел будут, конечно, комплексными:

$$\lambda = p + i\omega, \quad (7)$$

и часть границ устойчивости, как это вытекает из теории Ляпунова, определяется обращением в нуль действительных частей

$$p(\alpha_1, \dots, \alpha_l) = 0 \quad (8)$$

этих характеристических чисел.

Если система «в процессе развития», т.е. (в нашей модели) при движении ее параметров вдоль некоторой кривой PADQ, пересекает в точке А такую линию нейтральности, то в системе неминуемо возникает колебательный режим.

Это и есть математическая причина, по которой колебательные процессы встречаются достаточно часто, чтобы быть важным предметом исследования.

Не следует, однако, забывать и про другую менее «добрую половину» случаев, когда развитие системы происходит через точку В, лежащую на линии обращения в нуль одного из действительных корней

$$\lambda(\alpha_1, \dots, \alpha_l) = 0.$$

Явления, происходящие при переходе через такую линию бифуркации, тоже весьма интересны и моделируют, например, простейшие свойства такого важного процесса, как дифференцировка ткани.

Можно думать, что единое понимание и колебательной, и бифуркационной критичности — наиболее плодотворный подход, во всяком случае, к математическим моделям. Еще одно «усилительное» замечание. Предположим, что система в своем развитии пересекла десять критических линий. Может, конечно, случиться, что все они были чисто колебательными. Но это всего лишь один из $2^{10} = 1024$ возможных маршрутов развития. На любом другом пути развития необходимо (равноправное с колебательным) изучение бифуркационного типа усложнения системы.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ В ЭКОЛОГИИ. РОЛЬ КРИТИЧЕСКИХ РЕЖИМОВ

Введение

Экология — одна из наиболее сложных биологических дисциплин. Экосистемы едва ли не самые трудные для изучения системы из всех, с которыми приходится иметь дело в науке. Сама постановка вопроса об изучении экосистем стала возможной только в XX в. Уже это достаточно выразительно характеризует объемность проблемы.

В любую экосистему входят как биотические, так и косные (абиотические) элементы. Поэтому изучение экосистем невозможно без самого прямого участия наук геологического и географического цикла (при описании и изучении геологических, ландшафтных, гидрографических и климатических характеристик). Методы химии и физики абсолютно необходимы (но, к сожалению, совершенно недостаточны) при изучении составных элементов косной компоненты экосистемы. Суть и содержание экосистемы — живые существа, растения и микроорганизмы — предмет изучения всего комплекса наук биологического цикла.

При изучении экосистем невозможно обойтись и без социально-экономических наук, ибо основная цель экологии — использование биосферы для нужд человечества. Стоит, видимо, добавить — для человечества, поднимающегося над чисто эгоистическим «использованием» и вырастающего до понимания своей роли хранителя биосферы. Непросто взаимодействие экосистем с быстро развивающейся техникой, остро нуждающейся в разнообразных типах сырья, которое она черпает в экосистемах. С другой стороны, уже становится «противоположным общим местом» мысль об опасности необратимого загрязнения окружающей среды.

Науки математического цикла также должны найти свое место в этой важной проблеме. Можно указать, например, на статистические методы при количественных оценках структуры экосистем, на системный подход при описании и классификации многообразных внешних и внутренних связей экосистем и на более традици-

онные аспекты изучения изменения систем во времени как в эволюционном плане, так и в плане сукцессионном (возрастном).

Этот беглый перечень наглядно демонстрирует всю громадность проблемы. Возникает даже невольное сомнение: а готова ли сегодняшняя наука, состоящая из весьма разобщенных разделов, к решению подобных комплексных проблем? На это можно сказать только одно: вопрос поставлен самой жизнью, и стихийное его решение дорого обойдется и человечеству, и биосфере. Наиболее правильна, по-видимому, постановка вопроса о выделении самых острых, животрепещущих проблем.

Особенности биологических объектов

Правильное выделение главных направлений исследования и моделирования требует четкого понимания специфических особенностей биологических систем. Очень обобщенно можно указать на четыре важнейшие особенности живого: 1) сложность внутреннего строения каждой отдельной особи; 2) полифакторность внешней среды (условий жизнедеятельности); 3) незамкнутость (проточность) как в энергетическом, так и в структурном или информационном смысле; 4) существенная нелинейность — огромный диапазон внешних характеристик, при которых сохраняется жизнеспособность систем. Приведем несколько замечаний по каждому из этих разделов.

1. Сложность внутреннего строения. Этот раздел можно было бы — при желании — озаглавить иначе: «Уважение к биологическому объекту». Сколько-нибудь полное описание даже простейшего биологического объекта требует всей суммы знаний, накопленных в «предбиологическом» естествознании. Поясним простым примером эту мысль. В популярной литературе всегда приводят пример колеса, когда хотят польстить самолюбию человечества и показать, что биология «до колеса не додумалась». Но если понимать под «принципом колеса» идею замены трения скольжения трением качения, то любые шагающие, прыгающие бегающие животные реализуют этот принцип намного более гибко, чем неудобное и громоздкое техническое сооружение, именуемое колесом. Если же поставить вопрос более глубоко — о циклических процессах преобразования энергии, то биология «придумала» совершенно великолепный принцип «химического колеса» — преобразование химической энергии в механическую во всех процессах мышечного сокращения. О таком преобразователе в технике можно пока только мечтать. С чисто математической точки зрения сложность строения системы означает, что описание ее или задание ее состояния требует задания многих чисел. На математическом языке это можно сформулировать так: *фазовое пространство биологической системы многомерно.*

2. Полифакторность внешней среды. Биологические объекты сложны не только внутренне. Они функционируют в сложной,

нередко быстро меняющейся среде. Более того, есть серьезные основания думать, что сама сложность их строения носит «компенсаторный» характер. Они именно потому и сложны, чтобы в ответ на любое воздействие среды развить внешнюю защитную реакцию таким образом, чтобы сохранить максимально неизменной свою внутреннюю структуру. Особенно тщательно оберегается, конечно, генная, наследственная структура. Можно сказать, что сегодняшняя сложность биологической особи — это запасенный впрок «опыт общения со средой» ее «мамы», «бабушки», и «прапра...бабушки».

Математическое следствие: математическая модель биологической системы должна содержать много параметров (непрерывных и дискретных), задающих сложную среду, в которой функционирует изучаемая система.

3. Незамкнутость. Биологические системы никогда не бывают замкнутыми энергетически — это знают сейчас даже школьники. Существуют, однако, и более тонкие аспекты этого важного свойства. Так, например, у большинства высших растений орган фотосинтеза — это лист. Но в средней полосе деревья на зиму сбрасывают листву, которая становится частью среды. Еще более сложно взаимодействие со средой у насекомых, проходящих, например, стадии гусеницы и куколки. Эту особенность живого можно назвать «морфологической, структурной проточностью», или незамкнутостью. Почти очевидны примеры информационной проточности, скажем «химическое общение» у социальных насекомых. Но отсюда вытекает принципиальный (хотя и грустный для математика-«модельера») вывод: *необходимо совместное моделирование как биологической системы, так и среды ее функционирования.*

4. Существенная нелинейность. Диапазон изменения условий, в которых нормально работают биологические системы (например, органы зрения и слуха), значительно больше диапазона, к примеру, амплитуд и частот физических приборов или технических устройств. Во многих физических теориях вполне эффективно линейное приближение, описывающее малые отклонения от положения равновесия. Учет квадратичных поправок обычно существенно повышает точность модели. Необходимость введения нелинейности более высокого порядка нечасто встречается в «предбиологическом» естествознании.

Совсем иное положение в биологии. Там невозможно ограничиться даже сколь угодно высокими степенями. Нелинейность в биологии имеет *экспоненциальный характер*. Наиболее ясным выражением этого является физиологический «закон Вебера — Фехнера», устанавливающий логарифмическую зависимость реакции от воздействия. *Эволюционный смысл столь мощной нелинейности вполне понятен: надо услышать шорох ползущей змеи и не ослепнуть при близкой вспышке молнии.* Те биологические системы, которые не смогли охватить громадный диапазон

жизненно значимых воздействий среды, попросту вымерли, не выдержав борьбы за существование. На их могилах можно было бы написать: «Они были слишком линейны для этого мира». Но такая же судьба ожидает и математические модели, не учитывающие этой важной особенности жизни.

Сложные предбиологические системы. Границы возможностей вычислительной техники

В естествознании существуют объекты, для которых построены достаточно полные математические модели. Так, например, небесная механика предсказывает движение планет со всей точностью, на которую способны современные средства наблюдения. В этом смысле можно говорить о точной модели явления. Другой крайний случай — квантовая механика, описывающая структуру атома. В принципе нет оснований сомневаться, что уравнение Шредингера описывает поведение молекул с весьма высокой точностью. Это во всяком случае должно быть справедливо для небольших молекул, таких, например, как бензольное кольцо C_6H_6 .

На этом примере можно наглядно продемонстрировать границы возможностей вычислительной техники. Попытаемся (мысленно) просчитать на ЭВМ простейшую квантовомеханическую модель бензола — уравнение Шредингера для 24 электронов. Мы берем только по четыре валентных электрона на каждый из шести атомов углерода, откидывая внутренние (спаренные электроны и пренебрегая движением ядер. Даже при таком упрощении возникает уравнение в частных производных для ψ -функции, зависящей от 72 переменных (24×3).

Применяя стандартные вычислительные методы, используем разностную схему, взяв только лишь десять точек (а это очень мало, надо бы сто) по каждой из переменных. Всего у нас получится внушительное число точек: $N = 10^{72}$.

Чтобы оценить всю громадность этого числа, заметим, что современные ЭВМ способны производить не более миллиарда ($n = 10^9$) операций в секунду. Если предположить, что инженерам удастся повысить мощность ЭВМ еще в миллиард раз, то $n = 10^{18}$ операций в секунду. Даже и в этом (ныне совершенно фантастическом) предположении потребуется время $T = 10^{54}$ сек, невообразимо превышающее время существования Земли T_0 : $T \gg T_0 = 5 \times 10^9$ лет $= 1,5 \times 10^{17}$ сек для того только, чтобы сделать один-единственный шаг по расчету бензольного кольца прямым методом.

Точные, приближенные и элементарные модели

Технически прямой просчет точной модели какого-нибудь явления неосуществим, конечно, не только в биологии, но здесь это

случается особенно часто. Это ни в малейшей степени не означает бесполезности точных моделей. Но это предполагает необходимость тесного взаимодействия с чистой и прикладной математикой. Хорошо отработанная приближенная модель нередко радикально уменьшает объем вычислительной работы. Так, например, квантовая химия есть не что иное, как приближенный метод решения уравнения Шредингера для молекул. Квантовая химия позволяет вычислить важные (обычно энергетические) характеристики не слишком сложных, но достаточно типичных молекул и активных групп. Хотя мы и не можем найти решение точной квантовомеханической задачи, но знание точного уравнения Шредингера дает возможность оценить теоретически точность приближенного решения.

Точная, пусть даже крайне сложная математическая модель явления позволяет строить приближенные модели, отвечающие на частные вопросы. Происходит своеобразная «эмансипация» теории от эксперимента. Решающее слово остается, разумеется, за экспериментом. Но «Его Величеству Эксперименту» предоставляется возможность дать однозначный, «незамутненный» ответ. Справедливость квантовой механики, например, не нужно проверять на молекулах. Проверка теории идет на элементарном объекте (в данном случае на атоме), и притом достаточно простыми и точными экспериментами.

Мораль для математика: частные и предельные случаи теории важны не только сами по себе, но и как способ экспериментальной проверки полной, точной модели.

Возможна ли аксиоматика биологии

До последнего времени развитие математики стимулировали главным образом запросы механики, физики и техники. Возникла даже метафора «точное естествознание». Именно в рамках этого сравнительно узкого цикла дисциплин сформировался аксиоматический идеал построения науки. Суть идеала — установление (пусть даже экспериментальное) основных постулатов, но затем строго логическое построение всей остальной теории. Этот идеал буквально выстрадан всей историей становления естествознания как общественной и духовной силы. Привлекательность его очевидна. Тщательному уточнению и скрупулезной экспериментальной проверке подлежит лишь сравнительно небольшой набор постулатов. Следствия из них (модели частных явлений, как сейчас принято говорить) автоматически улучшаются по мере уточнения постулатов.

В наиболее прозрачной, классической форме этот идеал реализован в эллинской (Евклидовой) геометрии. Значительно сложнее аксиоматическая структура следующего крупного этапа научной мысли — небесной механики. В ней роль аксиом играют уравнения движения, опирающиеся на закон всемирного тяготения.

Далеко не все следствия легко обозримы. Так, например, «проблема трех тел» до сих пор не имеет удовлетворительного решения, хотя любое частное движение нетрудно найти современными вычислительными средствами.

Еще сложнее аксиоматика квантовой механики. Ее основной постулат — уравнение Шредингера — является уравнением в частных производных. Оно допускает точное решение только в простейших (хотя и важнейших) частных случаях. Чисто вычислительные подходы, как уже было сказано, чрезвычайно трудоемки и пока нереальны.

Все это, по-видимому, показывает, что аксиоматический идеал характерен, к сожалению, для «предбиологического» естествознания. Однако этот грустный вывод относится к биологии в целом. *Совершенно несомненно, что ее значительные отрасли (и притом совсем не обязательно близкие к ее нынешнему подразделению на частные дисциплины) не только могут, но и должны быть построены именно по аксиоматическому идеалу.* Ниже приводятся некоторые соображения в пользу этой точки зрения.

«Иерархия» и малый параметр

В прошлом веке была создана клеточная теория строения животных и растений. Это, вероятно, наиболее трудный и принципиальный шаг к «атомизму» в биологии, исторический сдвиг в понимании иерархичности, дискретности, «некисельности» форм существования жизни. Популяция состоит из организмов. Организм состоит из клеток. При всем глубочайшем различии между популяцией и организмом они имеют и глубокое сходство, позволяющее употребить оба раза один и тот же глагол «состоит». Клетки органа (или ткани), конечно, взаимодействуют друг с другом. Однако эта связь значительно слабее внутриклеточных связей. Относительная слабость межклеточных связей эффективно используется в эксперименте. Исследователь подбирает внешние воздействия (механические, химические или физические) настолько сильными, чтобы они могли вызвать распад ткани на отдельные клетки. Замечательно, что экспериментатор может этого добиться, сохранив полностью неповрежденными структуры и функции каждой отдельной клетки.

Отношение величины (неважно, чего именно — силы или кислотности, напряжения или температуры) воздействия «разъединяющего» к величине «разрушающего» воздействия есть, как говорят в математике, «безразмерный малый параметр» ϵ . Подобные малые параметры характеризуют количественно меру индивидуальности клетки в ткани. Обобщение этой идеи «вверх и наружу» — к биогеоценозам и биосфере, а также «вниз и внутрь» — к клеточным органеллам и макромолекулам очевидно. *Невозможно переоценить значение этой замечательной особенности живого для эффективности математических методов в биологии.*

Соседние уровни

Математика имеет дело с той же реальностью, что и все остальное естествознание. Отличаются, однако, методы, подход, точка зрения. Математика, чаще изучает связи, отношения, аналогии явлений, нежели их фактическое воплощение, реализацию. Неизбежен, конечно, проигрыш в конкретности. Но он с избытком (а только такие темы и стоят труда) компенсируется выигрышем в общности и количестве приложений. Ихтиолог, например, прогнозирующий численность некоторого вида рыб, работает обычно на уровне популяционном, не «опускаясь» до молекулярного, но и не «поднимаясь» до биосферного уровня. Эта возможность не учитывать соседние (вверх и вниз) уровни имеет общий характер и тесно связана с идеей малого параметра, точнее, двух малых параметров, причем эти параметры приобретают иную — кинетическую, временную — интерпретацию.

Среду можно считать почти постоянной: она меняется весьма медленно в масштабах времен и расстояний, характерных для изучаемого объекта. Это и есть иное — кинетическое — обличье малого параметра, меры индивидуальности. Внутреннюю среду объекта также можно считать постоянной, точнее зависящей только от существенных переменных, описывающих изучаемый объект. Однако причина в некотором смысле противоположна. Субъединицы (например, особи в популяции или клетки в особи) двигаются, изменяются, колеблются столь быстро, что имеют значения только средние значения этих быстрых переменных. Наиболее глубокое выражение эта идея находит в известной теореме А. Н. Тихонова об уравнениях с малым параметром при старшей производной. Впрочем ее частные приложения формулировались, по-видимому, и независимо: например, «метод стационарных концентраций» в химической кинетике.

Стоит ли «ломиться в открытую дверь» и «доказывать математически» возможность исследований в пределах одного уровня? Разумеется, не стоило бы, если бы не два обстоятельства. Во-первых, невозможно оставаться в рамках одного уровня, нужно знать границы применимости «одноуровневой» схемы. Нельзя, например, понять закономерности важнейших миграционных явлений только на популяционном уровне без физиологического (и даже биохимического) анализа стрессовых факторов. Математика помогает понять общую причину «прорыва» соседнего уровня — потерю устойчивости изучаемой системой. Во-вторых, математические аналогии позволяют иногда моделировать изучаемое явление на другом (более доступном эксперименту) уровне. Уже одно упоминание о «лавином» характере некоторых миграционных процессов есть указание на возможную физическую модель — аналогию.

Критические режимы, в частности колебательные

Сложность биологических явлений обычно очень велика. Даже в рамках одного уровня, даже «занулив» все дополнительные малые параметры, упростив и откинув все, что только можно упростить и откинуть, редко получают обозримую модель. В таких случаях может помочь следующий методологический прием. Изменением параметров модели (или условий эксперимента, а лучше и «условий эксперимента») сознательно и целенаправленно выводят систему на границу устойчивости. В таких критических ситуациях число существенных переменных обычно невелико, часто их всего только две. Разумеется, в других условиях «критическими» могут оказаться совсем другая пара или несколько переменных. Но если удалось добиться выхода на «линию нейтральности», то структура и свойства изучаемой системы во многом проясняются.

Можно уже чисто математически найти все возможные типы кинетики любой системы, попавшей в подобную критическую ситуацию. Это (хорошо известные из радиотехники) мягкий и жесткий режимы возбуждения автоколебаний. Возможны, кроме того, релаксационный аналог этих режимов (взрыв или мономолекулярный распад), а также комбинация этих режимов (например, дифференцировка тканей в эмбриогенезе). Все эти режимы, конечно, давно и хорошо изучены теоретически.

Существенно, однако, что в любой какой угодно сложной системе простейший тип потери устойчивости стационарного режима непременно будет протекать по одному из четырех указанных типов. *Математические модели радиоактивного распада (важнейший пример мономолекулярной реакции) или лампового генератора — это не примеры, а типичные представители, канонические формы самых сложных систем.* Очевидно, насколько возрастает с усвоением этой идеи наше уважение к подобным моделям.

Экологический аспект метода критических режимов

Экологические системы относятся к таким объектам, где следовало бы законодательно запретить применение методов, столь сочувственно описанных в предыдущем разделе. В сущности в этом и состоит одна из важнейших цепей всей программы «Человек и биосфера» — научиться не выводить экосистемы на край гибели. Именно для этого надо научиться «загонять» в кризис их модели. «Знай край, да не падай!» Хорошо известно, что немало различных экосистем поставлены как раз в критические условия (и сейчас уже несущественно, по невежеству или по другим причинам). Приходится поэтому превращать нужду в добродетель и ставить вопрос о первоочередном и тщательном,

максимально объективном изучении, сборе материалов, наблюдении и разумном вмешательстве в кризисных ситуациях. *Итак, метод моделирования критических режимов наиболее эффективен именно в таких ситуациях, когда моделирование всего более необходимо.*

Подходы системного анализа, тщательный учет всевозможных сложнейших, труднообнаруживаемых связей должны быть дополнены количественными оценками относительно вклада этих связей. Необходимо сопоставить различные модели, исходящие из альтернативных гипотез о ведущих переменных, и сравнить модельные выводы с экспериментальными, полевыми данными.

«Узкое место» и регуляция

После создания Дарвиным теории эволюции мы уже довольно ясно представляем себе процесс возникновения регуляторных механизмов. Идея малого параметра, независимо возникшая и в химии (а особенно в биохимии) под названием «узкое место в реакции», помогает понять картину эволюционного становления управляющих систем, во всяком случае некоторых из них. Первоначально это — «наиболее обнаженное», «самое уязвимое» место или стадия. Затем система либо вымирает (как вид), либо «берет в свои руки» контроль над узким местом (так, например, повышение концентрации углекислого газа возбуждает дыхательный центр). Вообще закрепление бывшего «отравителя» и постепенное превращение его в нынешнего «управителя» есть, по-видимому, одно из самых удивительных изобретений эволюции.

Может быть стоит поэтому предоставить все благодетельному воздействию времени? Так, конечно, поступить можно. Вопрос только в цене, которую заплатят биосфера, человечество и особенно цивилизация за стихийное регулирование. *К сожалению, в экологических системах мы не располагаем эволюционным временем. Мы вынуждены заменить последовательную эволюцию параллельным анализом.* Если, однако, мы хотим (а мы этого хотим!), чтобы к моменту окончания анализа было что регулировать, а самое главное, было кому регулировать,— этот анализ неминуемо должен быть модельным.

Приведенные соображения призваны осветить роль и место математического моделирования в экологических вопросах.

ЭКОЛОГИЯ И ЭРГОДИЧНОСТЬ

Предпринята попытка анализа важных экологических понятий — сукцессионного ряда и климакс ассоциации. Высказано предположение об идейном сходстве с эргодической проблематикой в математике. Подчеркнуто определяющее значение эволюционных, динамических аспектов по сравнению со структурными, морфологическими.

1. Сукцессионный ряд. Идею сукцессии можно понять на примере событий, происходящих после лесного пожара. Сначала на выгоревшей площади поселяются сорняки. Потом они вытесняются луговыми травами, место которых постепенно занимают кустарники. Кустарники создают условия, благоприятные для роста скороспелых древесных видов — осина, береза. Под их пологом развиваются основные породы. В зависимости от почвенных и климатических условий это может быть ельник, сосновый бор, дубрава.

Существенно различие длительности жизни каждого из этих пяти схематически выделенных биоценозов. Если первая стадия продолжается год-два, вторая — лет пять, а кустарники существуют обычно десять-двадцать лет, то характерный возраст березовой рощи уже сравним с продолжительностью человеческой жизни.

2. Климакс ассоциации. Ельник или сосновый бор, сменяющие березняк, живут так долго, нужны столь сильные потрясения (пожар, вырубка, наводнение) для их гибели, что психологически оправдано возникновение *понятия климакс ассоциации как заключительного, стабильного состояния.*

Присмотримся, однако, ближе к ельнику, существующему с незапамятных времен. Вот свалилась старая ель, сломала соседние деревья, они сгнили, образовалась полянка. Что на ней вырастет? А все та же трава, кустарник, осина. Основная площадь леса занята, конечно, елью, но все остальное тоже есть.

Те биоценозы, которые сменяют друг друга во времени в сукцессионном ряду, в климакс ассоциации соседствуют друг с другом в пространстве. Иными словами, климакс ассоциации есть *развернутый в пространстве временной ряд.*

Внешняя катастрофа вызывает принудительную синхронизацию фаз развития, сбрасывает на нуль возрастной индекс. Затем

случайные неоднородности все более и более разбалтывают фазы. Проходит время и возникает лес — та переливающаяся в пространстве (и во времени!) мозаика биоценозов, которую окрестные жители называют ельником (или дубравой).

3. Эргодическая гипотеза. Но почему мы за деревьями не видим леса? Потому, конечно, что основная часть площади леса занята именно деревьями. Этому очевидному замечанию можно придать содержательную и обобщающую форму.

Эргодичная гипотеза. Площади S_1, S_2, \dots, S_n , занимаемые биоценозами в климакс ассоциации пропорциональны временам T_1, T_2, \dots, T_n развития этих биоценозов в сукцессионном ряду

$$\frac{S_1}{T_1} = \frac{S_2}{T_2} = \dots = \frac{S_n}{T_n}.$$

Возможное практическое применение этой гипотезы — определение спелости древостоя по данным аэрофотосъемки. Один из теоретических выводов — значение быстрых стадий. В климакс ассоциации их роль плохо видна — мала площадь, занимаемая соответствующими компонентами. Сейчас все знают, что вытаптывание подлеска приводит к гибели всего леса. Менее известен другой выразительный и в некотором смысле, противоположный пример. Ф. Э. Фальц-Фейн, дореволюционный владелец нынешнего заповедника Аскания-Нова, огородил участок ковыльной степи. Ковыль быстро и пышно разросся и столь же быстро сгнил. Оказалось, что возобновление степи невозможно без копытных, протаптывающих узенькие полоски голой земли, на которой прорастают семена.

4. Заключение. Применение математики к биологии — одна из самых модных тем в современной науке. Тем поразительней контраст — обилие публикаций и бедность результатов. Объяснение, по-видимому, несложно. Математики плохо знают биологию, а биологи владеют, обычно, только простейшим математическим аппаратом.

Возникает грустный парадокс. Физика имеет развитый, специализированный математический аппарат. А биология (наука более трудная, включающая не только физику, но и химию) довольствуется скудной математикой прошлых веков. Надежды на мощные компьютеры неосновательны — самая могучая строительная техника бесплодна, если нет архитектора.

Главный путь синтеза биологии и математики — это правильное выделение фундаментальных (элементарных) объектов и явлений. Однако такой синтез возможен только при руководящей роли биологии и, особенно, ее наиболее глубокой и развитой ветви — теории эволюции в широком смысле слова.

КРИТИЧЕСКИЕ ТОЧКИ БИОЛОГИЧЕСКИХ СИСТЕМ (математические модели)

Последние несколько лет характеризуются бурным развитием исследований, посвященных колебательным явлениям в биохимии [1]. Однако интерес к нестандартной (в частности, колебательной) кинетике нередко принимает несколько сенсационный характер. Колебательный характер процессов воспринимается скорее как свидетельство изощренности автора очередной статьи (безразлично — экспериментатора или теоретика, а не как проявление глубинных свойств изучаемых объектов. Такое отношение к проблеме свойственно не только биохимикам. Аналогичное отношение математиков зафиксировано даже терминологически: исключительные ситуации называют случаями «вырождения». Терминологии соответствует и точка зрения: такие системы рассматриваются как досадные препятствия, мешающие применять привычные общие методы.

Между тем критические ситуации — это не только рубежи, отделяющие один тип поведения от другого, хотя одного этого было бы достаточно для серьезного их изучения. Биологическое значение критических ситуаций, по-видимому, еще глубже: это «реликты» важных этапов биологической эволюции. Полезен, вероятно, следующий эвристический принцип. *Если постепенное усиление воздействия на биохимическую систему (нормально, *in vivo*, работающую в стационарном неколебательном режиме) вызывает затягивание ее в колебательный режим, то колебательными становятся прежде всего параметры, характеризующие эволюционно наиболее позднюю связь системы.* Грубо говоря, эволюционно юная система распадается на эволюционно более древние подсистемы.

Этот «принцип слабого звена» вытекает из более общей идеи о роли кинетического резонанса в процессах усложнения биоструктур [2]. Если система эволюционно возникла в результате резонансного объединения первоначально независимых подсистем, то естественно предположить, что «обращение» эволюции усилением внешнего воздействия приведет к выделению исходных подсистем. Большинство систем утратило, разумеется, в процессе эволюции ту «рыхлость» строения, которая необходима для

затягивания в колебательный режим и расщепления на подсистемы. Тем интереснее с этой точки зрения каждый случай колебательной кинетики.

С другой стороны, биологические системы далеко не всегда стремятся избавиться от колебательных свойств. Имеются, по-видимому, случаи, когда колебательная кинетика является существенным функциональным элементом. Можно высказать рабочую гипотезу, что именно так обстоит дело в сократительных системах.

Миллиарды лет эволюции привели к отбору тонкого ферментативного механизма — «химического колеса». Этот механизм, действующий, вероятно, в циклическом режиме, реализован морфологически системой актиновых и миозиновых субъединиц [3], образующих основу сократительного аппарата. Конечно, сказанное есть не более, чем постановка вопроса.

1. Соотношение строения и кинетики.

В работе [4] найден канонический вид любой системы вблизи критической точки. Выяснено, что кинетика системы, определяемая амплитудным уравнением

$$\frac{dz}{dt} = pz + qz^2 + az^3, \quad (1)$$

не зависит от происхождения (строения, морфологии) системы и полностью определяется характером критической ситуации.

По-видимому, это частное проявление общей закономерности. *Кинетика является эволюционно первичной категорией.* Так как эволюция состоит в отборе устойчивых образований, на каждом уровне организации снова и снова возникает проблема стабилизации системы. Исходный материал меняется на каждом шагу эволюции, и соответственно его новым свойствам возникает новая комбинаторика (морфология) элементов, обеспечивающая данную кинетику. Разнообразие строения системы является фактором, компенсирующим разнообразие структурных элементов*.

Изложенные соображения уточняют роль и место математических моделей в биохимических исследованиях. В каждой проблеме выделяются два вопроса, дополняющие друг друга. Один вопрос — *какова кинетика данной системы* — допускает чисто математическое изучение, независимое от структуры изучаемой системы. Другой вопрос — *какова структура системы, обеспечивающая именно данную кинетику*, существенно опирается на знание конкретных свойств составляющих элементов. Однако даже в

* Хорошо известные механико-электрические, электрико-акустические и другие аналогии [5] можно рассматривать с этой точки зрения как различные реализации одной и той же простейшей кинетики, задаваемой линейными уравнениями.

этом случае предварительное изучение кинетики позволяет указать наиболее вероятный тип связей в системе. Более того, такое предсказание опирается во многом опять-таки на кинетические характеристики элементов, а не на детали их строения.

2. Структурный портрет системы

В теории колебаний [6] весьма полезно понятие фазового портрета системы. Так называют поведение полного семейства траекторий в фазовом пространстве. Это понятие допускает различные модификации. В некоторых случаях достаточно изобразить проекцию фазового портрета на подпространство. Иногда удобно выделить подсистему, и ее фазовый портрет называть фазовым портретом всей системы. В нашем случае, когда от времени зависят только две переменные, а остальные сохраняют постоянное значение, целесообразно поступить именно так.

Однако это понятие относится к индивидуальной системе, а в приложениях, особенно биологических, система обычно содержит параметры. Их смысл может быть весьма различен (например, количество фермента, скорость поступления субстрата, длина реактора, давление, температура и т.д.), но в математической схеме излишняя детализация противопоказана. Предположим для определенности, что система зависит от двух параметров α и β . Тогда коэффициенты канонической формы (1) будут функциями этих параметров:

$$p = p(\alpha, \beta), \quad q = q(\alpha, \beta), \quad a = a(\alpha, \beta). \quad (2)$$

В плоскости параметров α, β каждая точка изображает индивидуальную систему и каждой точке соответствует фазовый портрет системы. Но так как изобразить все эти портреты невозможно, то приходится ввести еще одно понятие — структурный портрет системы. При этом следует понимать в широком смысле слова, а именно как двухпараметрическое семейство систем (конечно, число параметров может быть и больше). Под структурным портретом целесообразно понимать изображение критических линий в пространстве (в нашем случае плоскости) параметров. В нашем конкретном случае такими линиями будут линии нейтральности

$$p = 0 \quad (3)$$

и линии кратных корней

$$4ap - q^2 = 0. \quad (4)$$

Отсюда сразу видно, что особую роль играет точка *максимальной критичности*

$$\left. \begin{aligned} p(\alpha, \beta) &= 0 \\ q(\alpha, \beta) &= 0 \end{aligned} \right\}, \quad (5)$$

лежащая на пересечении этих двух критических линий. Вполне естественна мысль о переходе от сравнительных случайных «экспериментальных» параметров α и β к структурным «каноническим» параметрам p и q .

Возникшая ситуация весьма характерна для математического моделирования. Исходные параметры α и β обычно весьма наглядны. Экспериментатор знает, что нужно делать, чтобы изменить α или β , он понимает их «физический» (или химический, или еще какой-нибудь) смысл. Обычно нелегко бывает убедить, что «формально» введенные канонические параметры p и q имеют значительно более глубокий, более адекватный задаче смысл. Полезно проследить эту коллизию на излагаемом конкретном примере.

Предположим, что функции $p(\alpha, \beta)$ и $q(\alpha, \beta)$ устроены так, что систему уравнений

$$\left. \begin{aligned} p &= p(\alpha, \beta) \\ q &= q(\alpha, \beta) \end{aligned} \right\} \quad (6)$$

можно решить относительно экспериментальных параметров α и β :

$$\alpha = \alpha(p, q), \quad \beta = \beta(p, q). \quad (7)$$

Для наших целей достаточно, чтобы это условие было выполнено в некоторой окрестности точки *максимальной критичности*. Точное условие состоит в том, что якобиан не равен нулю

$$\frac{\partial(p, q)}{\partial(\alpha, \beta)} = \begin{vmatrix} \frac{\partial p}{\partial \alpha} & \frac{\partial p}{\partial \beta} \\ \frac{\partial q}{\partial \alpha} & \frac{\partial q}{\partial \beta} \end{vmatrix} \neq 0 \quad (8)$$

именно в самой критической точке. Обращение якобиана в нуль приходится рассматривать как «вырожденный» случай, и тогда все негодование, высказанное вначале, оказывается огнем, вызванным на себя.

Выход заключается в том, что каждая задача должна ставиться на определенном уровне сложности, исключающем *любые* дополнительные условия *типа равенства*, — принцип «общего положения». Но именно эти «мешающие» равенства при переходе к более высокому уровню организации (включая, в частности, большее число параметров и поэтому большую адаптивную свободу) становятся условиями, определяющими *новую точку максимальной критичности*.

Структурный портрет нашей системы в канонических переменных определяется функцией

$$a = a(p, q), \quad (9)$$

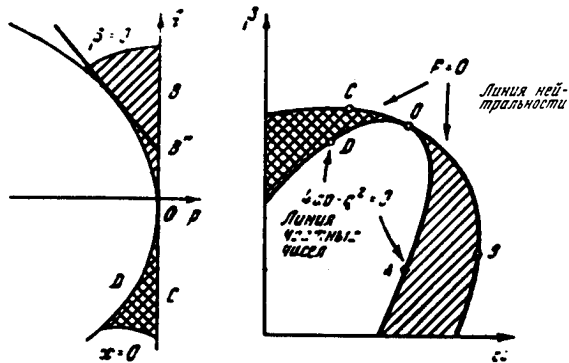


Рис. 1. Структурный портрет системы в канонических и экспериментальных параметрах. Изображены границы экспериментальной реализуемости (линии $\alpha=0$ и $\beta=0$). Случай $a < 0$

возникающей при подстановке в $a(\alpha, \beta)$ вместо α и β их выражений через p и q . Однако в окрестности критической точки имеет значение только величина a :

$$a = a(p, q) \Big|_{p=0, q=0}, \quad (10)$$

которую согласно принципу общего положения, надлежит считать отличной от нуля:

$$a \neq 0. \quad (11)$$

Поэтому структурный портрет (рис. 1) состоит из прямой $p=0$ и касающейся ее параболы $4ap - q^2 = 0$.

Эти линии являются, конечно, кривыми линиями в экспериментальных параметрах и α , и β , но факт касания сохраняется.

3. Устойчивые и неустойчивые критические точки.

Найденные линии делят плоскость параметров на четыре области, в каждой из которых система сохраняет определенный тип кинетики. Изменение типа происходит при пересечении параметрами критических линий. Основные, грубые черты системы определяются свойствами критической точки* $p=0, q=0$. Амплитудное уравнение в этой точке становится особенно простым:

$$\frac{dz}{dt} = az^3, \quad (12)$$

* Стрелки двух приборов («р-метра» и «q-метра») дрожат на красной черте, а мы смотрим, взорвется или не взорвется.

и сразу выделяются два противоположных случая соответственно по знаку числа a .

В случае положительного коэффициента

$$a > 0 \quad (13)$$

амплитуда колебаний z обращается в бесконечность, причем, как это видно из решения

$$z = \frac{z_0}{\sqrt{1 - 2az_0^2 t}}, \quad (14)$$

нарастание размаха колебаний до бесконечных размеров происходит за конечное время

$$T = \frac{1}{2az_0^2}, \quad (15)$$

определяемое величиной начального возмущения z_0 .

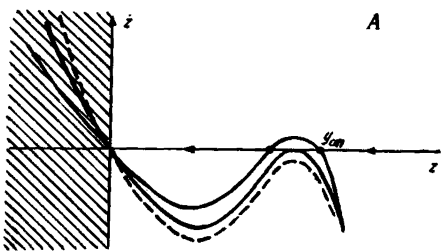
Когда формула дает ответ «бесконечность», это всегда означает, что мы пытаемся применить ее за границами применимости. В нашей задаче амплитудное уравнение было получено разложением в ряд и *отбрасыванием* членов более высокого порядка малости. Для отрицательных a колебательный режим, на который выходит решение, остается вблизи стационарной точки (в фазовом пространстве). Разложение в ряд поэтому законно.

Если же (при $a > 0$) решение выходит за пределы малой окрестности, то замена системы амплитудным уравнением незаконна. Поэтому уход решения на бесконечность не следует принимать слишком близко к сердцу: свойств реальной системы он, конечно, не отражает. Это не значит, однако, что ключевое уравнение бессмысленно. Напротив, оно совершенно правильно описывает не только факт ухода из малой окрестности, но и детали кинетики этого ухода (например, число оборотов). Но финальное движение оно описывает совершенно неверно. Это как раз то, чего следовало ожидать. Система уходит куда-то далеко, на «большой» предельный цикл или в другую стационарную точку. Для амплитудного уравнения, «рассматривающего» окрестность данной стационарной точки в микроскоп с бесконечно большим увеличением ($\varepsilon \rightarrow 0$), эти точки находятся на бесконечности, о чем уравнение и «сообщает» единственно доступным ему способом.

Случай отрицательного коэффициента

$$a < 0 \quad (16)$$

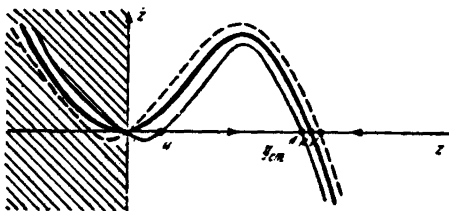
характеризуется устойчивостью стационарной точки. Решение имеет в точности тот же вид (14), что и для случая положительного a . Но смысл решения совершенно иной: начальное возмущение z_0 за время T



A

$$T = \frac{1}{2|az_0^2|} \quad (17)$$

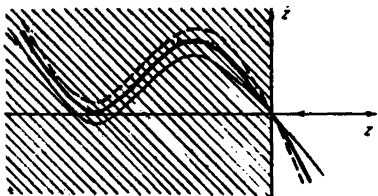
уменьшается в $\sqrt{2}$ раз. Поэтому решение при всех t остается в области применимости амплитудного уравнения, которое описывает, следовательно, асимптотически точно поведение системы.



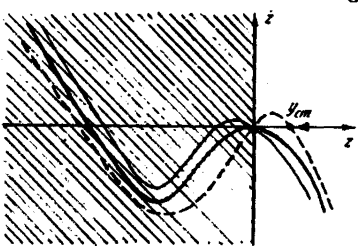
B

4. Мягкий и жесткий режимы возбуждения колебаний

Анализ, проведенный в работе [4] и приведший к выводу амплитудного уравнения (1), позволяет заключить, что характер возбуждения колебаний с кинетической точки зрения одинаков во всех системах независимо от их строения и происхождения. Можно поэтому ожидать, что понятие мягкого и жесткого режимов возбуждения колебаний, возникшее при исследовании [7] ламповых генераторов в радиотехнике, применимо к любым системам. Это ожидание оправдывается. Некоторое отличие обусловлено тем, что общие системы имеют многомерные фазовые пространства, в то время как уравнения в радиотехнике обычно второго порядка и фазовый портрет можно рисовать на плоскости. Но в малой окрестности критической ($p=0, q=0$) точки дополнительные степени свободы не проявляются и картина возбуждения колебаний может быть описана в привычных терминах мягкого и жесткого режимов.



D



C

Рис. 2. Точка А. Жесткий срыв колебания

Рис. 3. Точка В. Жесткое возбуждение колебаний

Рис. 4. Точка D. Сохранение стационарного режима

Рис. 5. Точка С. Мягкое возбуждение колебаний

Известный графический метод изучения одного уравнения [8] состоит в построении графика правой части. Пересечение кривой с осью z дает стационарные точки, а знак функции указывает направление движения. В нашем случае достаточно построить четыре графика, соответствующие типичным точкам критических линий A, B, C и D рис. 1. На приведенных графиках (рис. 2—5) область отрицательных z заштрихована, так как для исходной системы эти значения не имеют смысла, что не мешает им быть весьма полезными при изучении амплитудного уравнения. Толстые сплошные линии на этих графиках дают поведение правой части (\dot{z}) как функции z в точках на критических линиях. Тонкие сплошные линии показывают результат небольшого сдвига параметров в область между параболой и касательной к ней осью q . Пунктиром дан сдвиг внутрь параболы, и, наконец, штрих-пунктир использован при сдвиге в закритическую ($p > 0$) область. Буквой H обозначен возникающий при сдвиге неустойчивый цикл, Уст. — устойчивые циклы. Двойные стрелки показывают направление движения в силу амплитудного уравнения изображающей точки z .

5. Неизбежность гистерезиса в жестком режиме

Наиболее ярким (и экспериментально легче всего обнаруживаемым) оказывается явление гистерезиса в жестком режиме колебаний. Это явление вызывается «набуханием» и «сморщиванием» (при изменении параметров) неустойчивого предельного цикла. Сливаясь с охватывающим его устойчивым предельным циклом, он заставляет систему «свалиться» в положение равновесия. Наоборот, сжимаясь в точку, он делает неустойчивым положение равновесия, вынуждая переход системы в колебательный режим.

На графиках, изображающих зависимость \dot{z} от z (см. рис. 2—5), жесткому режиму соответствуют два положительных корня. Движение меньшего из них (он дает амплитуду неустойчивого предельного цикла) перераспределяет (рис. 6) области притяжения положения равновесия и колебательного режима.

6. Экспериментальное построение структурного портрета системы

Проведенный анализ позволяет указать план проведения серии экспериментов по построению структурного портрета системы (рис. 7). Основная цель — экспериментальное отыскание точки O максимальной критичности и построение критических линий OA , OB и OC . Предполагается, конечно, что техника эксперимента обработана настолько, чтобы достаточно уверенно фиксировать экспериментальные параметры α и β .

Элементарный анализ, на котором нет смысла останавливаться подробнее, позволяет построить серию графиков, показывающих изменение амплитуды колебаний вдоль кривых на плоскости

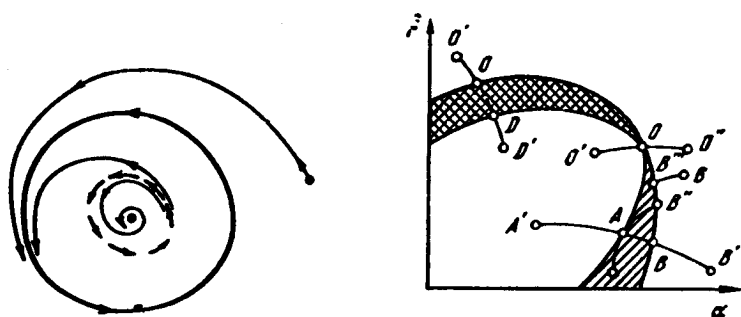


Рис. 6. Фазовый портрет системы с жестким возбуждением

Рис. 7. Различные воздействия на систему, описываемые изменением экспериментальных параметров. Случай $\alpha < 0$. OC — линия мягкого возбуждения колебаний; OB — линия жесткого возбуждения колебаний; OA — линия жесткого срыва колебаний

параметров. На рис. 8—12 приведены типичные «экспериментальные» кривые. Принадлежность точки к критической линии OC проявляется изломом в экспериментальной кривой (рис. 8). По мере приближения точки C к точке O излом увеличивается. Мягкий режим становится все более жестким. Максимального значения излом достигает при движении вдоль линии $O'O''$, проходящей через критическую точку O (рис. 9).

Гистерезис в жестком режиме позволяет построить не только линию OB , но и линию OA . Первое экспериментальное «соприкосновение» с явлением гистерезиса оставляет ощущение необратимости, если изменение параметров производить достаточно плавно (рис. 10). Лишь возвращение достаточно далеко назад, за точку A , восстанавливает состояние равновесия — тоже скачком (рис. 11).

Таким образом, точка B обнаруживает себя скачкообразным возбуждением колебаний при движении $A'ABB'$. Очень существенно, что движение начинается именно в точке A' . При обратном движении точка B ничем не выделяется, зато происходит срыв колебаний в точке A , переход через которую был незаметен при движении $A' \rightarrow B$. Общий качественный рецепт экспериментального обнаружения точки O состоит в том, чтобы мягкий режим стараться сделать более жестким, а жесткий — более мягким. В заключение полезно указать на возможность инверсных ситуаций. Нормально срыв колебаний в точке A происходит при падающей амплитуде. Можно, однако, добиться срыва колебаний при постоянной или даже растущей амплитуде.

Для этого необходимо (ясно, что экспериментально это трудная задача) менять параметры вдоль такой кривой $AB'''B$, которая в плоскости канонических параметров лежит между параболой AO и касательной AB'' , проведенной к параболе в точке A (рис. 12). Движению точно по касательной соответствуют

сохранение амплитуды и внезапный срыв колебаний в точке A (рис. 12). Однако экспериментальное обнаружение этих инверсных явлений требует поистине ювелирной техники, так как работать приходится в очень малой области $AB''A$.

7. Общие замечания.

Разобранный пример возбуждения колебаний помогает понять общее значение математических моделей. Выход системы на линию нейтральности оказывается настолько важным событием в ее истории, что строение и происхождение системы оказываются второстепенными «подробностями»: *в критических условиях логика поведения диктует логику строения*. Вдали от кризисной ситуации на первый план выступают специфика системы и особенности функционирования. Однако качественная схема сохраняется полностью и может измениться только при переходе через очередную границу устойчивости.

Качественный скачок сложности (возникновение колебательной степени свободы в нашем примере), происходящий в критической точке, невольно наводит на мысль о важнейшем эволюционном понятии — *ароморфозе* [9]. Тогда медленные количественные изменения между двумя кризисами естественно сопоставляются с *идиоадаптациями*.

Разумеется, прямое отождествление этих понятий было бы вульгаризацией. Однако противоположная крайность — отрицание какого бы то ни было сходства в этих двух рядах понятий — непростительное легкомыслие. Самое разумное поэтому при классификации разнообразнейших типов переходов через критические точки искать черты сходства и различия с теорией ароморфоза и иметь в виду построение в будущем простейшей математической модели ароморфоза.

В заключение стоит подчеркнуть, что изучение сложных систем стимулирует постановку задач, весьма непривычную для математика. Если понимать сложность (в самом простом варианте) как число параметров, от которых зависит система, то, чем сложнее система, тем «вырожденнее» ее точка максимальной критичности. Придется, например, изучать стационарные точки, где есть два нулевых корня, три корня — чисто мнимые с двумя резонансными соотношениями между ними и, кроме того, один из коэффициентов (типа a в нашем примере) канонической формы равен нулю. Математик склонен свысока относиться к таким задачам. Не исключено, однако, что классификация типов кинетики в критических точках неизбежно приведет и к таким «вырожденным» системам. Возможно, что термин «вырожденность» полезно в этих видах заменить термином «сложность».

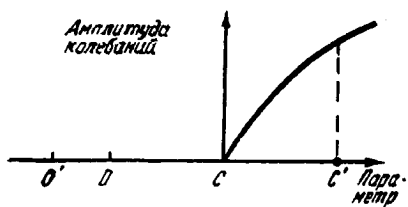


Рис. 8. Мягкий режим. Движение вдоль кривой $O' OCC'$ на рис. 7

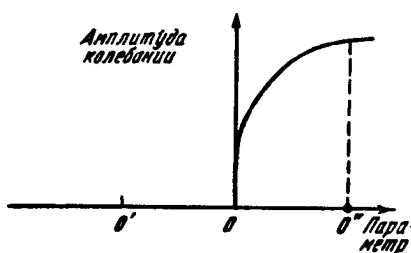


Рис. 9. Граница мягкого и жесткого режимов. В точке максимальной критичности касательная вертикальна

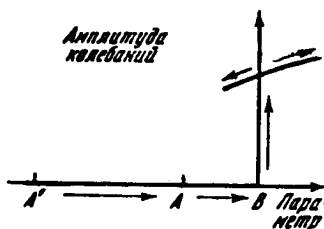


Рис. 10. Возбуждение колебаний в точке B и «невозможность вернуться к равновесию»

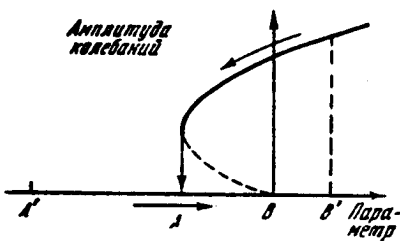


Рис. 11. Полная петля гистерезиса в жестком режиме

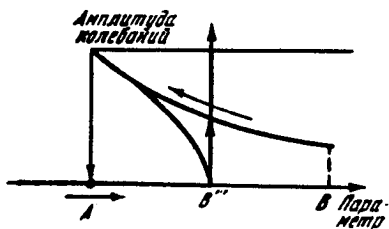


Рис. 12. Срыв колебаний при растущей амплитуде. Предельная линия соответствует критической линии OA на рис. 7

Литература

1. Колебательные процессы в биологических и химических системах. Труды симпозиума. Пушкино-на-Оке. М., «Наука», 1967.
2. Молчанов А.М.— В сб.: Колебательные процессы в биологических и химических системах. М., «Наука», 1967, с. 285.
3. Дерщевский В.И. Кинетическая модель мышечного сокращения и ее экспериментальная проверка. Автореф. канд. дисс. М., 1970.
4. Молчанов А.М. Нормальная форма нелинейной системы вблизи критической точки. Препринт № 6, ИПМ — ИБФ, 1969.
5. Ольсон Г. Ф. Динамические аналогии. М., ИЛ, 1947.
6. Боголюбов Н.Н., Митропольский Ю. А. Асимптотические методы в теории нелинейных колебаний, I. М., Гостехиздат 1955, с. 12.
7. Андронов А.А., Хайкин С.Э. Теория колебаний. М., Физматгиз, 1959.
8. Боголюбов Н.Н., Митропольский Ю.А. Асимптотические методы в теории нелинейных колебаний. II. М., Гостехиздат, 1955, с. 19.
9. Северцев А.Н. Главные направления эволюционного процесса. М.—Л., Биомедиздат, 1934.

ЭКСТРЕМАЛЬНЫЕ РЕЖИМЫ

Введение. Экстремальные состояния как поворотные точки эволюции

Биологические системы сложны. Это высказывание используется в самых различных целях. Математики используют его, чтобы объяснить весьма незначительный успех большинства попыток моделирования, а биологи — как защитное средство от вторжения непрошенных помощников и советчиков. Полезно уточнить это бесспорно верное высказывание.

Всегда ли биологические системы сложны?

Этот вопрос вызывает серьезные разногласия среди математиков и биологов. Интересно, что отрицательный ответ (равно как и положительный) дает часть биологов и часть математиков — раздел происходит по методологии, а не по профессии.

Но даже те, кто согласен, что биологические системы сложны не всегда, делятся на две резко оппозиционные группы. Одни считают, что простыми биологические системы бывают на краю гибели и что поэтому такие состояния нетипичны и бесполезны для понимания живых, функционирующих систем.

Другие согласны с первыми в том, что простыми бывают лишь *патологические* состояния, однако именно в изучении крайних, экстремальных состояний видят ключ к изучению нормы. Ясно, что симпатии автора на стороне этих «других», и небольшой, но поучительный пример поясняет такое предпочтение. Известно, что в эмбриогенезе существует стадия *морулы*, на которой все внутренние клетки *гибнут и растворяются*, служа всего лишь субстратом для внешних, поверхностных клеток, из которых возникает все тканевое разнообразие, все функциональное богатство будущего организма. Это не более чем иллюстрация, но она заставляет предполагать, что экстремальные состояния, близкие к губельным, существенны для нормального функционирования и уж во всяком случае заслуживают внимательного рассмотрения.

Вспомним теперь, что индивидуальное развитие повторяет (конечно в упрощенной и модифицированной форме) развитие вида, его эволюцию. Естественно, подобные катастрофические ситуации играли важную роль в эволюции. Возможно, что они

были точками выбора, разветвления, расхождения в разные стороны ранее близких, однотипных путей развития. Возникает желание связать математическое понятие «точки бифуркации» с биологическим, эволюционным понятием дивергенции, а также попытаться придать ясный и простой, хотя, конечно, всего лишь модельный смысл этим понятиям эволюционной теории.

При изучении биологических систем необходимо учитывать три главные особенности: 1) высокую внутреннюю подвижность; 2) их существенную нелинейность и 3) полифакторность внешней среды. Это делает малоперспективными попытки дедуктивного, аксиоматического построения теории.

В статье обсуждается альтернативная возможность — феноменологический подход к математической теории эволюции. Основой является глубокое сходство понятий нелинейной теории колебаний и функциональных идей теории эволюции.

В критических моментах эволюции структурная сложность биологических систем отступает на второй план. Ведущими оказываются сравнительно простые кинетические свойства, допускающие математическое описание.

Громоздкость многомерных систем

Развитие квантовой химии — «теории молекул» шло практически параллельно с развитием квантовой механики — «теории атома». Тем разительнее контраст между глубиной и завершенностью нерелятивистской квантовой механики и разрозненными частными результатами квантовой химии. Оказалось, что даже глубокое знание структуры атома не приводит* автоматически к пониманию свойств молекул.

Более того выяснилось, что многие важные свойства молекул вообще не зависят от деталей строения атомов. В некоторых случаях оказалось даже несущественным, какую модель атома положить в основу — грубую (но зато простую) классическую или значительно более тонкую квантовую.

Можно думать, что это одно из многочисленных и неожиданных проявлений *закона больших чисел*. Число электронов, а тем более различных парных взаимодействий оказывается довольно большим даже для небольших молекул. Это приводит к двум важным следствиям.

Во-первых, с некоторого момента становится бесполезным дальнейшее уточнение свойств малой компоненты (в данном

* На современном этапе глубокие результаты молекулярной биологии не отменяют необходимости независимого исследования клеточных структур и оргanelл.

случае атома) для понимания структуры и поведения всей системы (в данном случае молекулы).

Во-вторых, оказывается непосильной задача описания структуры и поведения системы на основе анализа точных уравнений, задающих взаимодействие всех ее компонент.

Поясним эту важную мысль на примере молекулы бензола [1]. Нет сомнения, что уравнение Шредингера, если бы его удалось решить или численно проинтегрировать, дало бы точную структуру и детали временного поведения молекулы бензола. Но все дело в этом «если бы». Молекула бензола содержит 24 валентных электрона. Поэтому уравнение Шредингера надлежит писать в 72-мерном пространстве ($72 = 24 \times 3$), даже если пренебречь движением атомов. Это приводит к разностной схеме, содержащей 10^{72} точек. При современных скоростях ($n = 10^9$ операций в секунду) потребуется $10^{72-9} = 10^{63}$ с $\approx 10^{55}$ лет для просчета одного шага по такой схеме.

Ясно, что такая степень трудности равносильна прямой невозможности.

Сверхчувствительность к изменению параметров

Большое количество независимых переменных — «проклятие размерности» — имеет еще один математический аспект. С формальной точки зрения существует теорема, гарантирующая непрерывность решения при изменении параметров системы. Однако, характер этой непрерывности оказывается весьма экзотичным при большой размерности системы.

Рассмотрим простой пример линейной системы:

$$\begin{aligned} \frac{dx_1}{dt} &= \lambda_1 x_1 + a_2 x_2, \\ \frac{dx_2}{dt} &= \lambda_2 x_2 + a_3 x_3, \\ &\dots\dots\dots \\ \frac{dx_{n-1}}{dt} &= \lambda_{n-1} x_{n-1} + a_n x_n, \\ \frac{dx_n}{dt} &= \varepsilon a_1 x_1 + \lambda_n x_n \end{aligned} \tag{1}$$

Если в этой системе $\varepsilon = 0$, то она становится треугольной, а собственные числа совпадают с числами стоящими по диагонали. Поэтому, если все λ_i отрицательны

$$\lambda_i < -\rho < 0, \tag{2}$$

то положение равновесия устойчиво. Теорема о непрерывной зависимости от параметра гарантирует сохранение этой устойчивости при достаточно малом значении параметра ε .

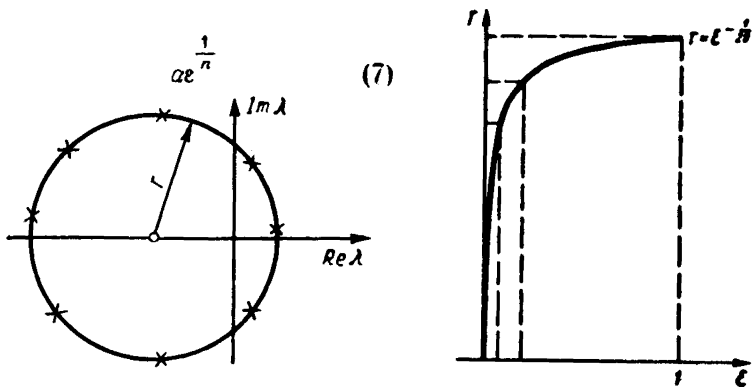


Рис. 1. Собственные числа на комплексной плоскости

Рис. 2. Вид зависимости r от ϵ при больших n

Присмотримся, однако, к реальным числам. Несложно написать в общем виде вековое уравнение для системы (1):

$$(\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n) + (-1)^{n-1} a_1 a_2 a_3 \dots a_n \epsilon = 0. \quad (3)$$

При $\epsilon \rightarrow 0$ собственные числа $\lambda_i(\epsilon)$ стремятся к λ_i , как это и полагается по общей теореме. Рассмотрим для ориентировки частный случай, когда все λ_i равны между собой

$$\lambda_i = -p \quad (4)$$

и все a_i равны между собой

$$a_i = a. \quad (5)$$

В этом случае уравнение (3) можно решить. (Достаточно сказать, чтобы невозмущенная система имела n -кратный корень):

$$\lambda = -p + r e^{\frac{2k+n-1}{n} \pi i} \quad (6)$$

Следовательно, все собственные числа расположены на окружности с центром в точке $\lambda = -p$ (рис. 1)

Радиус этой окружности зависит от ϵ :

$$r = a \epsilon^{\frac{1}{n}} \quad (7)$$

и, согласно теореме, стремится к нулю при $\epsilon \rightarrow 0$.

Однако зависимость r от ϵ чрезвычайно велика, даже если n не очень большое (рис 2.)

Так, например, уже при $n=20$, $p=1$, $a=2$ изменения ϵ всего лишь в пятом знаке достаточно, чтобы система потеряла устойчивость.

Полезно в связи с этим заметить, что во многих [2] американских работах (особенно относящихся к направлению имитационного моделирования (simulation modelling)) беззаботно толкуют о системах, содержащих многие десятки и даже сотни переменных.

Эффекты нелинейности

Существенная нелинейность биологических систем также приводит к резким изменениям свойств при небольших изменениях параметров. В работе [3] разбираются математические свойства простейшей модели [4] гликолиза:

$$\begin{cases} x = \alpha - xy^2, \\ y = xy - y. \end{cases} \quad (8)$$

Здесь x и y — концентрации некоторых соединений, а параметр α задает внешний поток субстрата в систему. Оказывается, что весьма умеренные изменения внешнего потока α приводят к бурным и драматическим следствиям в поведении системы. Опишем качественно происходящие события [4].

При достаточно обильном «питании», а именно при $\alpha \geq 1$, система работает в стационарном режиме и «нарабатывает» субстрат в продукт с постоянной скоростью α .

При $\alpha = 1$ в системе возникает устойчивый предельный цикл, субстрат по-прежнему целиком перерабатывается в продукт, однако при постоянном входе выход становится переменным, циклическим. Часть периода система накапливает субстрат, а другую часть использует для выработки продукта.

С уменьшением внешнего потока работа системы становится все более «лихорадочной», все длиннее «пустые» периоды накопления, все короче «рабочие» циклы.

Наконец, при некотором критическом значении параметра α

$$\alpha = \alpha_{кр}$$

внешний поток монотонно стремится к нулю, идет бесполезное накопление субстрата в системе.

В той или иной степени описанные явления могут происходить в любой проточной системе. Однако в приведенной системе эти свойства проявляются особенно ярко. Достаточно сказать, что все эти события разворачиваются при изменении внешнего потока α меньше, чем на двадцать процентов! (рис. 3.)

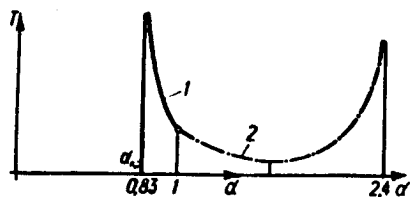


Рис. 3. Зависимость периода T от величины внешнего потока α : 1 — период предельного цикла; 2 — период малых колебаний при возвращении в положение равновесия.

Биологические и небιологические системы

Ясно, что рассмотренные свойства не являются исключительной принадлежностью именно биологических систем. Аналогичные явления могут возникать и возникают и при развитии (эволюции) иных сложных систем. Примерами могут служить: эволюция звезд и возникновение планетной системы; геологические процессы на Земле в целом или судьба данного речного бассейна; история возникновения паровозов, совершенствования и практически полного исчезновения паровозов, как транспортного средства, и даже (с очень существенными оговорками) некоторые социальные явления.

Однако есть весьма существенная разница между биологическими и небιологическими системами. Конечно, для каждого свойства можно указать небιологическую систему, обладающую заданным свойством, и даже изучать это свойство, так сказать, в чистом виде, именно на такой системе. Но биологические системы обладают обычно полным «ассортиментом» всех таких сложных и непривычных особенностей, и трудной становится противоположная задача — найти такую биологическую систему, которая *не обладала бы* заданным сложным типом поведения или структуры.

Стационарные режимы. Устойчивость по Ляпунову. Теорема Пуанкаре.

История физического естествознания оставила в наследство естествознанию биологическое понимание существенной роли стационарных режимов. Теория устойчивости Ляпунова была первоначально создана для тщательного изучения сравнительно узкой задачи — устойчивости грушевидных форм вращения жидких тел. Сейчас, однако, все более ясно, что созданный Ляпуновым алгоритм изучения вопросов устойчивости сохраняет значение для более широкого круга проблем.

Если поведение системы описывается системой обыкновенных дифференциальных уравнений

$$\frac{dx_i}{dt} = f_i(x_1, \dots, x_n; \alpha_1, \dots, \alpha_k), \quad (9)$$

то стационарное состояние определяется такой системой уравнений:

$$\left. \begin{aligned} f_1(x_1, \dots, x_n; \vec{\alpha}) &= 0, \\ f_2(x_1, \dots, x_n; \vec{\alpha}) &= 0, \\ f_n(x_1, \dots, x_n; \vec{\alpha}) &= 0 \end{aligned} \right\}. \quad (10)$$

Если

$$x_1 = a_1, \dots, x_n = a_n \quad (11)$$

— решение этой системы конечных уравнений (10), то для определения устойчивости надо построить матрицу

$$A = ((A_{ik})) = \left(\left(\frac{\partial f_i}{\partial x_k} \right) \right)_{x_i = a_i}, \quad (12)$$

и тогда вопрос об устойчивости стационарного состояния (11) решается вычислением корней характеристического уравнения

$$P(\lambda) = \text{def}((A_{ik} - \lambda \delta_{ik})) = 0 \quad (13)$$

Корни характеристического уравнения оказываются, конечно, зависящими от параметров

$$\lambda = \lambda(\alpha_1, \dots, \alpha_k) \quad (14)$$

поскольку и координаты стационарной точки

$$\left. \begin{aligned} a_1 &= a_1(\alpha_1, \dots, \alpha_k), \\ a_2 &= a_2(\alpha_1, \dots, \alpha_k), \\ a_n &= a_n(\alpha_1, \dots, \alpha_k), \end{aligned} \right\} \quad (15)$$

и коэффициенты матрицы A являются функциями параметров системы.

Кинетика системы вблизи стационарной точки полностью определяется собственными числами $\lambda_1, \dots, \lambda_n$. Теорема Пуанкаре [5] выясняет глубокий смысл этого обстоятельства. Эта теорема утверждает, что линейная диагональная система

$$\left. \begin{aligned} \frac{dy_1}{dt} &= \lambda_1 y_1, \\ \frac{dy_2}{dt} &= \lambda_2 y_2, \\ &\dots\dots\dots \\ \frac{dy_n}{dt} &= \lambda_n y_n, \end{aligned} \right\} \quad (16)$$

— есть *каноническая* форма произвольной нелинейной системы. Существует такая замена переменных:

$$\left. \begin{aligned} x_1 &= P_1(y_1, \dots, y_n), \\ x_2 &= P_2(y_1, \dots, y_n), \\ &\dots\dots\dots \\ x_n &= P_n(y_1, \dots, y_n), \end{aligned} \right\}, \quad (17)$$

которая приводит произвольную систему (9) к нормальной, канонической, линейной форме (16).

Следует подчеркнуть два обстоятельства.

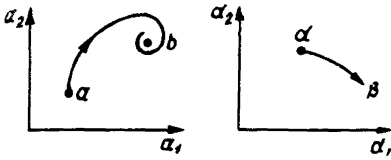


Рис.4. Переходный процесс $a \rightarrow b$, соответствующий переключению параметров $\alpha \rightarrow \beta$.

Во-первых, теорема верна только для случая, когда все собственные числа лежат целиком в полуплоскости комплексного переменного λ (проблема малых знаменателей). Это условие автоматически выполнено для устойчивых стационарных точек, когда

$$\text{Re} \lambda_i < 0 \quad (18)$$

и, следовательно, все собственные числа лежат в левой полуплоскости.

Во-вторых, замена переменных (17) определена только в некоторой достаточно малой окрестности стационарной точки (несодержащей, в частности никакой другой стационарной точки или предельного цикла).

Структурная устойчивость

Устойчивость по Ляпунову означает малое изменение траектории движения при малом изменении начальных данных. Однако во многих вопросах, особенно биологических, требуется знание изменения траектории при малом изменении внешней среды, а не внутреннего состояния системы. Математически это соответствует изменению параметров системы $\alpha_1, \dots, \alpha_k$, а не фазовых переменных x_1, \dots, x_n . Оказывается, что, эти два типа устойчивости связаны друг с другом. Опишем более точно эту связь.

Предположим, что параметры заданы и система находится в стационарной точке a , соответствующей точке α пространства параметров. Изменим значение параметров, переведя их в точку β . Равновесие нарушается, так как состоянию β соответствует равновесная точка b , отличающаяся от a . Возникает переходный процесс — движение в точку b (рис 4.).

Если в состоянии α система была устойчива по Ляпунову (более точно асимптотически устойчива), а сдвиг $\alpha \rightarrow \beta$ был выбран небольшим, то система остается устойчивой и в точке β . Это вытекает из непрерывной зависимости собственных чисел λ от параметров системы.

Структурный сдвиг $\alpha \rightarrow \beta$ мы можем интерпретировать иначе. Мы можем считать, что система уже находилась в состоянии β , но мы изменили начальные данные, поместив систему в точку a .

Эквивалентность структурной перестройки $\alpha \rightarrow \beta$ фазовому сдвигу $b \rightarrow a$ (обратный порядок!) позволяет применить к этой ситуации методы теории устойчивости Ляпунова.

Можно, в частности, оценить (при достаточно малом сдвиге $\alpha \rightarrow \beta$) время возвращения в равновесие — длительность переходного процесса.

Чисто формально это время равно, конечно, бесконечности, но как всегда в таких случаях время оценивают временем уменьшения отклонения в « l » раз.

Нетрудно показать, что это время релаксации определяется наименьшей величиной p действительных чисел $p_i = \text{Re} \lambda_i$ системы в точке β .

Итак, время релаксации может быть найдено по формуле

$$T_{\alpha \rightarrow \beta} = -\frac{1}{p(\beta)}. \quad (19)$$

Знак минус в этой формуле показывает, конечно, из-за отрицательности всех p_i в точке устойчивого равновесия. Для дальнейшего полезно выписать отношение времен релаксации для прямого и обратного процессов:

$$\frac{T_{\beta \rightarrow \alpha}}{T_{\alpha \rightarrow \beta}} = \frac{p(\alpha)}{p(\beta)}. \quad (20)$$

Динамическая асимметрия

Из формулы (20) вытекают важные следствия относительно характера переходных процессов для состояний, близких к критическому. Рассмотрим сечение пространства параметров какой-нибудь плоскостью. Это делается только для наглядности. Рассуждения не зависят от размерности пространства параметров и числа фазовых переменных. Суть дела в том, что мы рассматриваем окрестность гиперповерхности нейтральности — поверхности, пересечения которой приводят к потере устойчивости стационарного режима. На самой поверхности корень может становиться либо чисто мнимым (откуда и происходит название линия нейтральности, а слово «линия», вместо поверхность, отражает традицию ограничиваться двумерным структурным пространством), либо нулевым (рис 5.)

Линии уровня $\text{Re} \lambda = \text{const}$ на этом графике соединяют состояния с одинаковым временем релаксации к положению равновесия. Это время растет по мере приближения к линии нейтральности и обращается в бесконечность на самой линии (рис 6.).

Отсюда вытекает существенный вывод. Вблизи линии нейтральности система приобретает резко выраженную асимметрию.

Приближение к линии нейтральности вызывает длительный переходный процесс, удаление от линии нейтральности приводит к быстрому установлению стационарного режима. Это легко понять рассматривая формулу (20), показывающую, что отношение времен прямого и обратного перехода может быть весьма значительным.

Любопытно, что это обстоятельство было обнаружено экспериментально при запуске ректификационных колонн. Однако авторы сочли этот факт специфической особенностью технологическо-

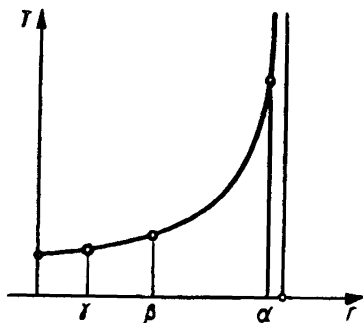
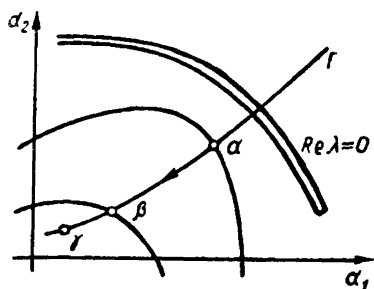


Рис. 5. Линии уровня функции $p(\alpha) = \text{Re} \lambda$. Пересечение с плоскостью α_1, α_2 . Выделен критический уровень $\text{Re} \lambda = 0$ — линия нейтральности

Рис. 6. Время релаксации к равновесию. Взят разрез вдоль кривой Γ (рис. 5.). Величина ординаты над точкой γ изображает время релаксации к положению равновесия, соответствующему этой точке пространства параметров

го процесса, поэтому общая причина этого явления ускользнула, по-видимому, от их внимания. Эффект был воспроизведен на математической модели ректификационной колонны и послужил побудительной причиной для поисков внутренней кинетической причины, не зависящей от деталей конкретного строения системы.

Этой главной причиной оказывается запас устойчивости, близость режима к экстремальным, неустойчивым состояниям.

Эволюционная роль экстремальных режимов

Затягивание системы в экстремальные состояния может быть прослежено на примере любой эволюционной системы. Интересно, что технические системы обнаруживают и в этом отношении принципиальное сходство с системами биологическими. Ограничимся одним примером.

Развитие авиации долгое время шло под знаком борьбы за скорость. Увеличение взлетной, а особенно посадочной скорости привело к тому, что упругость пневматиков шасси в принятой тогда трехколесной схеме стала причиной драматического явления, известного под названием «шимми». Этот пример особенно любопытен потому, что абстрактная «степень свободы», выходящая первой на границу устойчивости, имела здесь вполне зримые очертания переднего колеса, начинавшего «вихлять» при возрастающих посадочных скоростях.

В других случаях, (например, «фляттер») совсем не так легко было выделить ту «слабую степень свободы», которая первой подвергалась воздействию «эволюционного отбора». Поучительно, что эволюция технических устройств обнаруживает те же основные закономерности «борьбы за существование» и «выжива-

ние приспособленных», что и эволюция биологических систем не только на организменном, но и на популяционном уровне.

Мягкое и жесткое возбуждение колебаний

Исторически первая аккуратная математическая модель была построена А. А. Андроновым [6] для лампового генератора. Современное изложение, простое и наглядное, можно найти в работе [7]. Однако долгое время считалось, что схема возбуждения колебаний является специфической особенностью, если не электротехнических схем, то во всяком случае, двумерных динамических систем.

Анализ окрестности линии нейтральности, проведенный в работе автора [8], показал, что двумерные системы являются канонической формой общих систем.

Имеет место пока еще, к сожалению, недоказанный аккуратно аналог теоремы Пуанкаре для общих систем. В окрестности линии нейтральности *любая* система может быть приведена (надлежащей заменой переменных) к следующей канонической форме:

$$\left. \begin{aligned} \frac{dr}{dt} &= pr + qr^3 + ar^5 + \dots, \\ \frac{d\varphi}{dt} &= \omega + \mu r^2 + \nu r^4 + \dots, \\ \frac{dz_i}{dt} &= \lambda_i(r) z_i \end{aligned} \right\} \quad (21)$$

Здесь r и φ — амплитуда и фаза возбуждающейся степени свободы, а z_i — канонические переменные, сохраняющие стационарное состояние $z_i = 0$. Таким образом, система остается линейной, как в теореме Пуанкаре, по всем «квазистационарным» переменным z_i , становится существенно нелинейной по основным переменным r и φ , образующим возбужденную степень свободы.

Качественный характер возбуждающих колебаний совершенно такой же, как и в случае одной степени свободы. Это видно из того, что уравнение для r не содержит квазистационарных переменных z_i , хотя в уравнение для z_i это переменное r входит существенным образом.

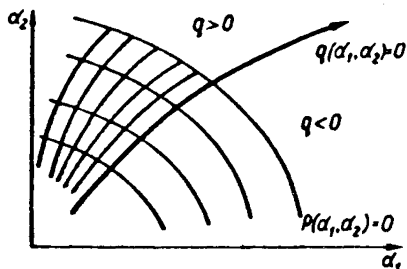


Рис. 7. Кривые $p(\alpha_1, \alpha_2) = 0$ и $q(\alpha_1, \alpha_2) = 0$ разбивают плоскость параметров α_1, α_2 , на четыре области соответственно комбинациям знаков величин p и q

Полезно изобразить «структурный портрет» возникающей системы (рис. 7). Возникающие здесь явления подробно разобраны в работе [9]. Особенно существенно явление гистерезиса, неизбежно проявляющееся при переходе линии нейтральности в жестком режиме возбуждения колебаний.

Комбинированные экстремальные режимы

В рамках настоящего сообщения невозможен сколько-нибудь подробный анализ как раз наиболее интересных и перспективных экстремальных режимов. Ограничимся поэтому в порядке постановки вопроса указанием на самые актуальные задачи.

Во-первых, необходимо проанализировать явление бифуркации стационарной точки. Выше разобрано пересечение линии нейтральности. Это явление порождается обращением в нуль действительной части комплексного корня

$$P = \operatorname{Re} \lambda = 0, \quad (22)$$

которому соответствует (в линейном приближении) инвариантная плоскость (r, φ) .

Совершенно аналогично (но более просто) можно разобрать случай обращения в нуль действительного корня

$$\mu = 0. \quad (23)$$

Есть серьезные основания предполагать, что такие важные биологические феномены, как *дифференцировка* ткани или изоляция видов имеют в основе своих математических моделей именно бифуркацию стационарной точки.

И в этом случае существуют аналоги мягкого («расхождение») и жесткого («взрыв») эволюционного пути. Эволюционирующая система в своем развитии может пересекать как линии нейтральности (22), так и линии бифуркации (23). Нет никаких априорных оснований предпочитать один тип усложнения другому.

Но если это так, то система, прошедшая через десять критических точек в своем развитии, имеет больше тысячи ($2^{10} - 10^{24}$) различных вариантов усложнения. Среди них только один соответствует чисто колебательному (на всех десяти этапах), а другой (тоже один) — чисто релаксационному типу развития.

Почти очевидно, что реальные системы реализуют все мыслимые возможности хотя, конечно, часть таких путей и может приводить к эволюционным тупикам.

Общий закон эволюции — индивидуальный путь развития повторяет путь развития вида — ставит перед математикой весьма своеобразную задачу.

Дело в том, что повторение происходит по весьма «сокращенной стенограмме» — месяцы вместо миллиардов лет. Это почти наверняка означает, что индивидуальный путь развития сливает

в одно несколько (быть может весьма много) различных этапов усложнения.

Математически это может означать путь эволюции через точку высокой кратности — высокой сложности.

Эти соображения стимулируют изучение окрестности сложных особых точек, например таких, как точка *C* на рис. 8. Точка *C* является точкой пересечения двух линий

$$\left. \begin{aligned} \operatorname{Re} \lambda = \rho(\alpha_1, \alpha_2) = 0 \\ \mu(\alpha_1, \alpha_2) = 0 \end{aligned} \right\} .(24)$$

Это означает, что в ней одновременно обращается в нуль один из действительных корней и становится чисто мнимым один из комплексных корней.

Подобные точки допускают весьма полное исследование и тесно связаны с любопытной механической задачей

$$x''' + \omega^2 x = \varepsilon f(x, x', x'') \quad (25)$$

Число возможных разнообразных типов кинетики увеличивается очень существенно и достигает полутора десятков вместо двух (или четырех) типов кинетики в обычных точках возмущения колебаний.

Заметим, что даже если не принимать во внимание доводы относительно возможной роли подобных систем в математических моделях биологических систем, изучение таких систем представляет содержательную математическую задачу.

Литература

1. Молчанов А.М. Математические модели в экологии. Роль критических режимов. — В кн.: Математическое моделирование в биологии. М., «Наука», 1975, с. 133—141.
2. Austin M. P., Cook B. G. Ecosystem stability. — J.Theor. Biol., 1974 No 45, p. 435—458.
3. Молчанов А.М. Индогенные биохимические колебания как возможная основа физиологических ритмов — Биофизика, 1971, 16, с. 878—883.
4. Сельков Е.Е. Автоколебания в гликолизе. Простая одночастотная модель. — Молекулярная биология, 1968, № 2, с. 252—266.
5. Немыцкий В.В., Степанов В.В. Качественная теория дифференциальных уравнений. Т., 1. Изд. 2. М.— Л., ГИТТЛ, 1949, с. 233—241.
6. Андронов А.А., Хайкин С. Э. Теория колебаний. М.— Л., ОНТИ, 1937. 510 с.
7. Боголюбов Н.Н., Митропольский Ю.А. Асимптотические методы в теории нелинейных колебаний. М., ГИТТИ, 1963. 406 с.
8. Молчанов А.М. Нормальная форма нелинейной системы. Препринт ИПМ № 6, М., 1969. 32 с.
9. Молчанов А.М. Критические точки биологических систем. (Математические модели). — В кн.: Математическое моделирование в биологии. М., «Наука», 1975, с. 142—153.

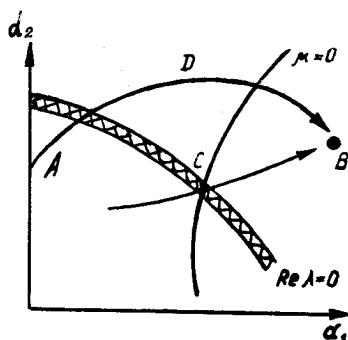


Рис. 8. Последовательный (видовой) АДВ и параллельный (индивидуальный) АСВ пути развития. Необходимость пройти через точку *C* высокой степени критичности

МЕХАНОХИМИЯ

(феноменологический аспект)

Предложена система уравнений, описывающая, по мнению автора, основные кинетические режимы, возникающие при взаимодействии механических и химических компонент. Намечены пути изучения этих сложных режимов.

История механики состоит из трех главных этапов. Начальный этап (похожий на современное состояние проблемы биологической подвижности) характеризовался разнообразием задач и пестротой методов их решения. Во времена Эйлера наступила эра вариационных принципов, которым придавался философский, мировоззренческий смысл. Однако постепенно выяснилось, что максимум функционала реализуют только очень простые (чаще всего просто положения равновесия) движения. Интересные (и сложные) движения соответствуют обычно стационарным* (а не экстремальным) значениям функционала, задающего вариационный принцип.

Жгучий (но несколько «поспешный») интерес к вариационной тематике остыл и наступил завершающий этап действительно глубокого синтеза. Был создан гамильтонов формализм и показано, что вариационный подход есть частный случай гамильтоновой схемы, после чего механика получила надежную математическую базу — теорию гамильтоновых систем и канонических преобразований.

Награда за этот тернистый и мучительный путь оказалась достойной — выяснилось, что математические методы эффективны далеко за пределами исходного физического объекта (планеты

* Уравнения Эйлера — это необходимые (но отнюдь не достаточные) условия экстремума. На плоскости, например

$$\frac{\partial f}{\partial x} = 0, \frac{\partial f}{\partial y} = 0,$$

что дает минимум ($f = x^2 + y^2$), максимум ($f = -x^2 - y^2$) и просто гиперболическую точку ($f = x^2 - y^2$) — максимум по одному направлению и минимум по другому.

и пушки — небесная механика и внешняя баллистика). Все огромное индустриальное разнообразие поддерживающих ферм вращающихся шестерней, переключающих рычагов, льющихся и перекачиваемых жидкостей, упорных грунтов и горных пород и стремительных газов — все это добрая старая механика Эйлера. Техника сегодняшнего дня — это физика вчерашнего дня.

Математический аппарат реагировал лишь несущественными уточнениями на два мощных катаклизма, потрясших основания физических представлений — релятивистский и квантовый пересмотр основ физики.

Сейчас ведущей стала биология. Возникла ли новая ситуация? Появится ли нет самостоятельная «теория биологической подвижности»? Судить об этом рано. Однако одна ветвь — механохимия — назревает. Что это такое?

Существует биоэнергетика. Она возникла на биохимической, молекулярной основе. Однако изучение митохондрий (основных энергетических комплексов клетки) показывает, что дело не сводится к чистой химии. Существенны и механические (возможно конформационные) процессы. С другой стороны, актомиозиновые комплексы — структурные единицы мышечного «кристалла» — имеют как бы «встроенные» энергокомпоненты (химической, разумеется, природы).

Главное направление сейчас — изучение конкретных систем. Однако никому не заказаны попытки «рыть туннель с другого конца». Например, нет недостатка в «термодинамических» схемах. Это изложение известного на другом языке, да и обилие таких схем вызывает подозрение. Задним числом (*a posteriori*) можно понять глубокую причину их бесплодности. Дело в том, что системы (пока они биологические) работают вдали от положения термодинамического равновесия, а также вдали от живого равновесия «спячки». По сходной причине недостаточен и чисто биохимический подход.

Существует, однако, еще одна возможность. Не претендуя на преждевременное построение «общей» теории можно попытаться угадать ее существенные элементы. Анализ разумно начать с изучения движений, близких к состоянию «живого» равновесия («спячки»). Это заведомо содержательней «термодинамики» и много проще и скромнее общей теории биологической подвижности (которая не обязана существовать).

Экстремальные режимы

Применительно к нашему случаю феноменологический* анализ выглядит следующим образом.

Предположим, что «задача решена» и математическая теория биологической подвижности создана. Разберем свойства, характерные для любой частной модели.

Прежде всего описание движения потребует несколько механических степеней свободы. Энергетика (и, вероятно, регуляторные механизмы) — это еще несколько уравнений для химических концентраций. Кроме того, надо задать внешнюю среду и условия функционирования. Поэтому в простейшем случае (пространство среды и условий) в уравнение войдет несколько параметров.

Итак

$$\frac{d\vec{x}}{dt} = \vec{a}(\vec{x}, \vec{\alpha}), \quad (1)$$

где \vec{x} — вектор размерности k ; $\vec{\alpha}$ — вектор размерности l .

Предположим, что проведено полное исследование этой системы и построен «структурный портрет», т.е. найдены границы (поверхности) в пространстве $\vec{\alpha}$, разделяющие области с различными режимами функционирования. Пересечения этих границ (многообразия размерности $l-2$ и меньше) тем интереснее (с точки зрения разнообразия поведения), чем больше поверхностей пересекается. Особый интерес представляют точки максимальной сложности** (в которых пересекаются не меньше, чем l поверхностей). Близкие аналоги этой схемы — теория ветвления решений и теория бифуркаций, но для приложений недостаточно, к сожалению, развитых до сих пор подходов.

Дискретное множество точек максимальной сложности даст полный набор этих типов кинетики («поведения»), к которым вообще способна изучаемая система.

Введем термин «точка кумуляции» для точки максимальной сложности и «экстремальный режим» для кинетики системы в этой точке.

Система в экстремальном режиме обнаруживает свойство, сходное с адаптивностью биологических систем — качественно перестраиваться при изменении внешних условий. Наоборот вдали от точек кумуляции система устойчива. Более того она является «грубой», поскольку речь идет об изменении параметров, а термин «устойчивость» использован, к сожалению, в задачах об изменении начальных данных. Любой из этих грубых режимов

* Исходящий из общих свойств, а не из конкретного механизма.

** Крайне неудачное выражение «максимальное вырождение», употребляемое математиками.

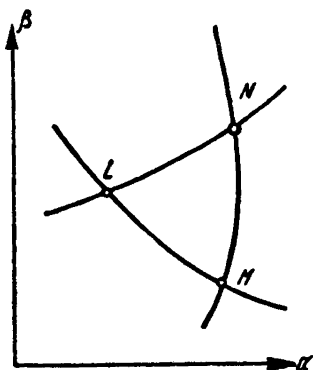


Рис. 1. Границы смены режимов: NM — линия нейтральности; LN и LM — линия нулевых корней. Точка M — одна из точек кумуляции (один нулевой и два чисто мнимых корня).

качественно схож с одним из тех, на которые распадаются экстремальные режимы при сколь угодно малом возмущении.

Различение экстремальных и грубых режимов имеет и чисто практический смысл. Оно намечает правильное «разделение труда» между ЭВМ и человеком. Математик качественно изучает экстремальный режим, «подключая» ЭВМ на количественной стадии. Затем они «вместе» анализируют (количественно и качественно) распадение экстремального режима на грубые, после чего ЭВМ «дотягивает» нужный режим количественно до нужных значений параметров.

Окрестность покоя

Разнообразие экстремальных режимов бесконечно велико, однако изучены весьма немногие, а широко известен пока один — мягкое и жесткое возбуждение колебаний (рождение предельного цикла). Цель настоящей статьи — разбор еще одного важного примера экстремального режима — механо-химического.

Рассмотрим простейшую возможность — одна механическая степень свободы (x и y), одна химическая (концентрация z) и два параметра α и β :

$$\begin{aligned} \frac{dx}{dt} &= f(x, y, z; \alpha, \beta); \\ \frac{dy}{dt} &= g(x, y, z; \alpha, \beta); \\ \frac{dz}{dt} &= h(x, y, z; \alpha, \beta). \end{aligned} \quad (2)$$

Отличие «химической» переменной от «механических» определяется характером правых частей системы.

Параметры α и β выбираем так, чтобы вблизи $\alpha=0$, $\beta=0$ система имела устойчивую стационарную точку. Линеаризуя нашу систему вблизи этой стационарной точки, мы получим три собственных числа, зависящие, конечно, от параметров α и β . Построим на плоскости α , β линии нулевых корней,

$$\mu(\alpha, \beta) = 0 \quad (3)$$

и линию нейтральности,

$$\text{Re} \lambda(\alpha, \beta) = 0, \quad (4)$$

при пересечении которой и происходит рождение предельного цикла.

Гипотеза. Точка кумуляции* основных механических режимов есть пересечение линии нейтральности с линией нулевых корней. Разберем подробнее структуру системы вблизи точки кумуляции. Запишем уравнение выделив явно главную (линейную) часть:

$$\frac{dx}{dt} = -y + a(x, y, z; \alpha, \beta); \quad \frac{dy}{dt} = x + b(x, y, z; \alpha, \beta); \quad \frac{dz}{dt} = c(x, y, z; \alpha, \beta) \quad (5)$$

Здесь величины a , b , и c являются «малыми» поправками к невозмущенной линейной системе.

В каком смысле эти поправки малы?

В классической проблеме устойчивости движения ответ ясен — это квадратичные, кубичные и т.д. члены:

$$\begin{aligned} a &= \tilde{a} = a_{11}x^2 + 2a_{12}xy + 2a_{13}xz + a_{22}y^2 + 2a_{23}yz + a_{33}z^2 + \dots; \\ b &= \tilde{b} = b_{11}x^2 + 2b_{12}xy + 2b_{13}xz + b_{22}y^2 + 2b_{23}yz + b_{33}z^2 + \dots; \\ c &= \tilde{c} = c_{11}x^2 + 2c_{12}xy + 2c_{13}xz + c_{22}y^2 + 2c_{23}yz + c_{33}z^2 + \dots; \end{aligned} \quad (6)$$

— поправки к линейной части. Эта точка зрения сохранилась и в более современной теории бифуркаций (и несколько претензионной ветви ее — теории катастроф). Однако для многих приложений (особенно биологических), когда система содержит несколько параметров, эта точка зрения недостаточна. Необходимо также учитывать сдвиг параметров из точки M . Он приводит дополнительно к малым константным и линейным поправкам, так что полное** возмущение имеет вид

$$\begin{aligned} a &= \varepsilon(a_0 + a_1x + a_2y + a_3z) + \tilde{a}; \\ b &= \varepsilon(b_0 + b_1x + b_2y + b_3z) + \tilde{b}; \\ c &= \varepsilon(c_0 + c_1x + c_2y + c_3z) + \tilde{c}. \end{aligned} \quad (7)$$

Эволюционные уравнения

Главная особенность полученных уравнений — два независимых малых параметра. Это явно написанный параметр ε (характеризующий малое отклонение от M в пространстве параметров) и подразумеваемая малая окрестность стационарной точки.

* Гармонический осциллятор $\ddot{x} + x = 0$ — это экстремальный режим механики. Весьма правдоподобно также, что $\dot{z} = 0$ есть кумуляция химических (одномерных) режимов. Естественно поэтому рассмотреть «кумуляцию» этих «кумуляций».

** При сдвиге параметров возникают и малые нелинейности, однако их можно включить в основную нелинейность $\tilde{a}, \tilde{b}, \tilde{c}$.

Выделим в уравнениях (5) эволюционную систему. Для этого следует перейти к полярным координатам

$$\begin{aligned}x &= \rho \cos \varphi; \\y &= \rho \sin \varphi; \\z &= z,\end{aligned}\tag{8}$$

выделив явно быструю фазу φ и медленные переменные z и ρ :

$$\begin{aligned}\frac{d\varphi}{dt} &= 1 + \dots; \\ \frac{d\rho}{dt} &= a \cos \varphi + b \sin \varphi; \\ \frac{dz}{dt} &= c.\end{aligned}\tag{9}$$

Метод вращающейся фазы позволяет исключить быструю фазу (надлежащей заменой переменных) из уравнений для z и ρ . Главные члены этих уравнений — просто средние по быстрой фазе от главных членов исходной системы. Дальнейшие выкладки быстро становятся весьма громоздкими, но для наших целей достаточно одного существенного обстоятельства — в системе остаются только резонансные члены:

$$\begin{aligned}\frac{d\rho}{dt} &= \rho F(\rho^2, z); \\ \frac{dz}{dt} &= G(\rho^2, z).\end{aligned}\tag{10}$$

Выпишем несколько первых членов:

$$\begin{aligned}- \frac{d\rho}{dt} &= \rho(\alpha + Az + B\rho^2); \\ \frac{dz}{dt} &= \beta + \gamma z + Dz^2 + C\rho^2.\end{aligned}\tag{11}$$

Это уравнения главного приближения, если A, B, C, D — константы. Однако эти уравнения становятся точными, если эти величины считать произвольными функциями от ρ^2 и z . В данной статье изучается только окрестность стационарного режима и для наших целей достаточно считать A, B, C и D константами. При изучении тонкой структуры экстремального режима (п.5) придется учитывать зависимость A, B, C и D только от z , да и то в линейном приближении.

Существуют, однако, задачи, в которых требуется построение в целом фазового портрета системы в точке кумуляции. В этом случае следует знать функции A, B, C, D во всем фазовом

пространстве. Существуют и промежуточные случаи. Если изучаемый режим неустойчив, то укороченная система может генерировать неограниченные решения, а в полной системе это всего лишь переходный процесс в другой стационарный режим (или на предельный цикл). В этом случае нужно выбрать функции A , B , C и D так, чтобы они правильно описывали окрестности обоих стационарных режимов — отталкивающего и притягивающего.

Выбор параметров α , β и γ соответствует удобному выбору системы координат в пространстве параметров — точка кумуляции имеет координаты $\alpha=0$, $\beta=0$, $\gamma=0$.

Экстремальный режим

Рассмотрим систему в точке кумуляции ($\alpha=0$, $\beta=0$, $\gamma=0$). Это соответствует точке зрения теории бифуркации. Система в этой точке имеет вид

$$\frac{d\rho}{dt} = A\rho z + \dots; \tag{12}$$

$$\frac{dz}{dt} = c\rho^2 + Dz^2 + \dots$$

Мы оставили только главные, квадратичные члены (не учитывалась в частности, кубичная нелинейность $B\rho^3$ в первом уравнении). Нетрудно проверить, что полученная система имеет первый интеграл,

$$I(\rho, z) = \rho^{-2k}(z^2 + \theta\rho^2) = \text{const}, \tag{13}$$

где параметры k и θ связаны с коэффициентами системы (12) соотношениями

$$k = \frac{D}{A}; \tag{14}$$

$$\theta = \frac{c}{D-A}.$$

Фазовый портрет системы, т.е. семейство кривых

$$z^2 = \theta\rho^2 = I\rho^{2k} \tag{15}$$

качественно меняется, когда параметры k и θ проходят через прямые $\theta=0$, $k=0$ и $k=1$.

Следовательно, на плоскости параметров k , θ имеются шесть областей (образованных этими тремя прямыми), внутри которых фазовый портрет сохраняет вид.

Тонкая структура экстремального режима

В предыдущем пункте разобрано влияние только главных членов — квадратичных по переменным ρ и z . Более полная картина возникает, если учесть кубические члены. При этом величины A , B , C и D нельзя считать константами, но достаточно заменить их линейными по z функциями. Зависимость от ρ^2 даст четвертый порядок и ее можно в этом приближении не учитывать, как и члены вида z^2 . Исследование можно провести различными способами, но едва ли не самый простой состоит во введении удобных переменных. Поскольку нас интересует окрестность начала координат, то разумно оставить ρ , а вместо z ввести однородную переменную $\omega, z = \omega\rho$.

Перепишем систему (11) в новых переменных (они разумны только при $\alpha=0$, $\beta=0$, $\gamma=0$, т.е. в точке кумуляции):

$$\frac{d\rho}{dt} = \rho^2 (A\omega + B\rho);$$

$$\frac{d\omega}{dt} = \rho [C + (D - A)\omega^2 - B\rho\omega]. \quad (16)$$

Исключим теперь время, разделив первое уравнение на второе:

$$\frac{d\rho}{d\omega} = \rho \frac{A\omega + B\rho}{C + (D - A)\omega^2 - B\rho\omega}. \quad (17)$$

В этом уравнении особенно прозрачна аналитическая структура системы в точке кумуляции. Окрестность начала координат z , ρ переходит в полосу вдоль оси ω ($\rho=0$) в новых переменных. Поэтому правую часть уравнения (17) можно просто разложить в ряд по степеням ρ , ограничившись двумя первыми членами

$$\frac{d\rho}{d\omega} = \rho \frac{A\omega}{C + (D - A)\omega^2} - \rho^2 \frac{\varepsilon + F\omega^2 + G\omega^4}{[C + (D - A)\omega^2]^2} + \dots \quad (18)$$

Это уравнение, содержащее новые структурные параметры ε , F и G , которые можно назвать «постоянными тонкой структуры», легко интегрируется. Это уравнение Бернулли и сводится к линейному делением на ρ^2 . Однако исследование получающегося решения — самостоятельная задача, выходящая за рамки данной статьи. Здесь ограничимся одним принципиальным замечанием. Уравнение (18) дает правильную асимптотику (при $\rho \rightarrow 0$) всюду, кроме нулей знаменателя

$$C + (D - A)\omega^2 = 0, \quad (19)$$

которым соответствуют прямолинейные решения на левых фазовых портретах ($\theta < 0$) рис. 2.

Окрестность этих двух особых точек на оси ω ($\rho=0$) необходимо исследовать прямо из уравнения (17).

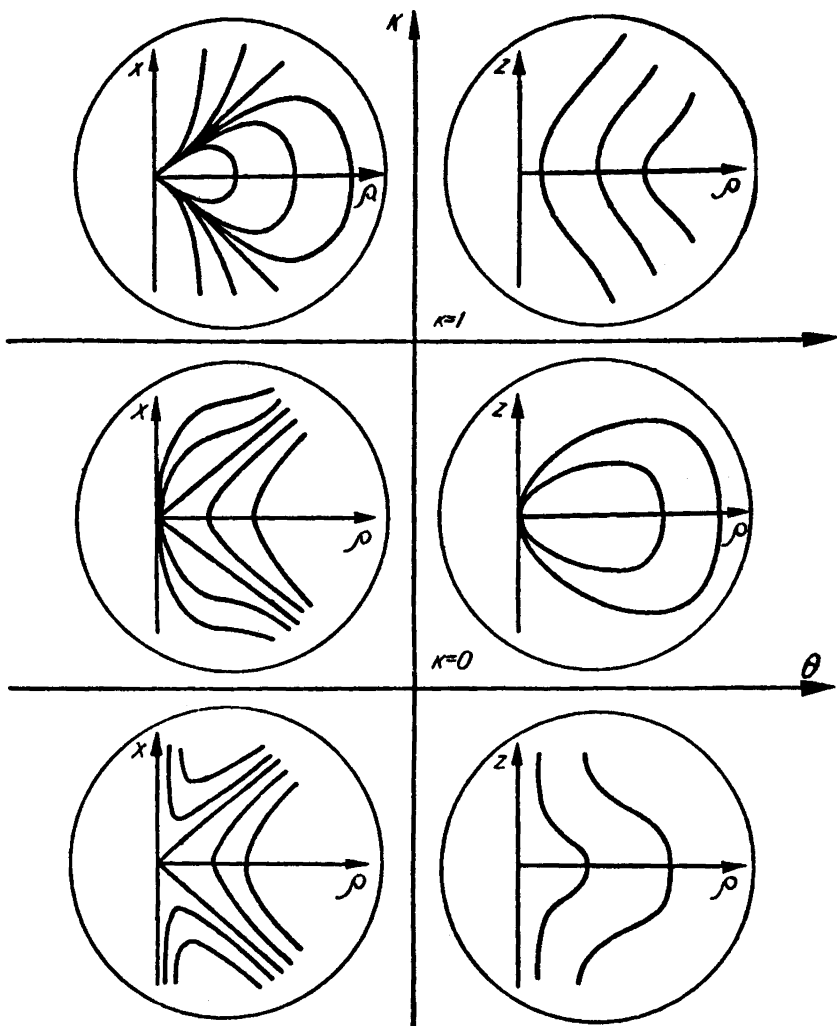


Рис. 2. Фазовые портреты в различных областях изменения параметров θ и κ

Нейтральные возмущения экстремального режима

Изучение распада экстремального режима при сдвиге параметров является трудной задачей. В данной статье ограничимся указанием одного замечательного обстоятельства, существенно облегчающего анализ ситуации. Оказывается, что из трехпараметрического (α , β и γ) семейства сдвигов двухпараметрическое семейство, задаваемое условием

$$\gamma = \frac{2D}{A} \alpha, \quad (20)$$

сохраняет первый интеграл системы. В этом смысле эти возмущения названы нейтральными. Выпишем этот интеграл. Рассмотрим систему (11), откинув в ней кубические нелинейности (в частности, $B\rho^3$, а величины A , D и C считаем постоянными), но сохранив параметры сдвига α, β и γ .

Разделим второе уравнение на первое:

$$\frac{dz}{d\rho} = \frac{Dz^2 + \gamma z + \beta + C\rho^2}{\rho(\alpha + Az)}. \quad (21)$$

Умножим обе части равенства на $(\alpha + Az)$:

$$(Az + \alpha) \frac{dz}{d\rho} = \frac{1}{\rho} (Dz^2 + \gamma z + \beta) + C\rho \quad (22)$$

Теперь становится понятным смысл условия (20). Оно означает, что уравнение (22) — линейное относительно величины

$$U = Dz^2 + \gamma z + \beta \quad (23)$$

Само уравнение (22) приобретает вид

$$\frac{A}{2D} \frac{du}{d\rho} = \frac{u}{\rho} + C\rho. \quad (24)$$

Интегрирование этого уравнения дает следующий первый интеграл:

$$\left[z^2 + \frac{\gamma z + \beta}{D(D-A)} + \frac{C}{D-A} \rho^2 \right] \rho^{-2k} = \text{const}. \quad (25)$$

Понятно, что при $\gamma=0$, $\beta=0$ этот интеграл совпадает с найденным ранее (13). В общем случае получаем фазовый портрет системы при любых β и γ :

$$z^2 + \frac{\gamma z + \beta}{D(D-A)} + \frac{C}{D-A} \rho^2 = I\rho^{2k}. \quad (26)$$

При этом подразумевается, что α связано с γ соотношением

$$\alpha = \frac{A}{D} \gamma. \quad (27)$$

Сдвиг из этой плоскости* приводит к разрушению интеграла (26).

Более подробное изучение фазового портрета (26) и явлений, происходящих при сдвиге из плоскости с линии (27), остаются за рамками данной статьи.

* Условие (27) есть уравнение плоскости в трехмерном пространстве параметров α , β , γ .

Обсуждаемый здесь вопрос идейно близок к обнаруженному Н. Н. Боголюбовым замечательному явлению высокочастотной стабилизации седловых точек.

Из системы (11) видно, что параметры β и γ суть собственные числа линеаризованной системы. Поэтому $\beta=0$, $\gamma=0$ соответствуют нейтральности системы. Сдвиг этих параметров приводит либо к устойчивости, либо к неустойчивости этой стационарной точки. Однако выполнение условия (27) сохраняет первый интеграл (26) полной системы, что является аналогом сохранения нейтральности (вообще говоря, индефинитной!) полной системы, несмотря на возникновение грубости линейного приближения. Это означает, по-видимому, что наличие быстрой фазы и вытекающее из этого осреднение приводит к таким специальным типам правых частей системы эволюционных уравнений, к которым не применены соображения «общего положения», т.е. мы попадаем в пространство коэффициентов системы на такие удивительные поверхности, на которых отказывает привычная интуиция грубых систем.

СТАЦИОНАРНЫЕ РЕЖИМЫ В БИОЛОГИИ И МАТЕМАТИКЕ

Тема моего краткого вступительного слова — современное состояние взаимодействия математики и биологии. Сейчас это состоя-



Рис. 1. Точка S — устойчивая стационарная точка. Весь интервал CS входит в область притяжения стационара S .

ние чрезвычайно поучительно, многозначительно, и, мне кажется, в дальнейшем перспективно и продуктивно. Я все прилагательные расставил, как краткие, так и длинные; теперь можно переходить к существу дела. Еще 20, ну, максимум 30 лет назад единственной (в широком масштабе) формой взаимодействия математики и биологии была ста-

стистическая обработка результатов наблюдений. Я даже помню кусочек расписания на мехмате, маленькая клеточка, в которой было такое заклинание, которое я еще тогда выучил наизусть: «Теорверниобрезнабл». Расшифровывалось это так: «Теория вероятностей и обработка результатов наблюдений». От этого заклинания сегодня мы (как мне кажется) достаточно **далеко ушли**, а связано это с важным обстоятельством, которое я попробую наглядно изобразить.

В прошлом веке и большую часть этого века единственным представлением — я повторяю, в «широких массах» — о том, что должны делать математики в биологии было представление, что математики должны изучать стационарное состояние и малые отклонения от него. И это все, что доверялось математикам. Под стационарным состоянием понималось в те времена следующее. Вот у вас есть некая ось X (биомасса, дыхание или еще что-нибудь), а стационарное состояние — точка на этой оси, к которой при малых отклонениях приближается система в процессе своей жизнедеятельности. Это есть стабильное устойчивое состояние (рис. 1). Потом появились словечки типа «метастабильность», но они все не очень далеко уходили от этого главного понимания. Если изобразить во времени — пусть « x » — это некий признак, который нас интересует, а « t » — время. Сначала с ним (s

Рис. 2. S-образная кривая — переходный процесс $C \rightarrow S$. Но в той же системе переходный процесс ($D \rightarrow S$) в тот же стационар S может быть значительно сложнее (с ускорениями и замедлениями)

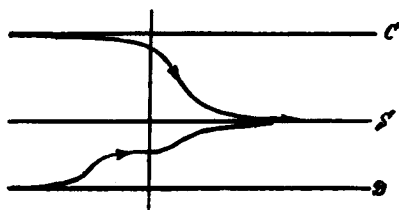
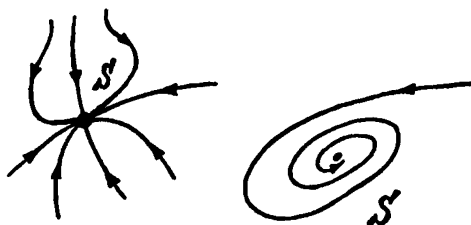


Рис. 3. Типичные устойчивые стационарные точки — узел слева и фокус справа (рис. 3).



«х»-ом) что-то такое происходит, а потом он выйдет на стационар (рис. 2, нижняя позиция).

А если он находился сверху, то он все равно выйдет на тот же стационар (рис. 2, верхняя позиция). Вот это и есть стационарное состояние; это вот и есть то, чем должна заниматься математика (рис. 3).

Лет 17 назад в Пущине был проведен первый колебательный симпозиум. К этому времени мы начали понимать, что стационарным состоянием может быть не только равновесие, но и предельный цикл, и траектория движения системы на него может снаружи и изнутри накручиваться (рис. 4). Типичный пример стационарных режимов такого типа — это система «хищник-жертва», знаменитая Вольтерровская система*.

Другие примеры — физиологические процессы, дыхание или сердцебиение. Обычно рисуют кардиограмму так: «х» (одно из «отведений») как функцию от «t», и получают весьма прихотливые кривые (рис. 5). Аналогичный график получится для любого «х» от «t»; самое главное, что процесс цикличен.

Это уже довольно серьезный сдвиг в понимании. Но еще остается, по-видимому, пройти самое главное. И у меня такое впечатление, более того, уверенность, которой я хотел бы заразить вас, что сейчас, спустя 17 лет, мы серьезно продвинулись

* В Вольтерровской системе реализуется бесконечно много циклов. Но любое реалистическое уточнение идеализированной трофической схемы приводит обычно к выявлению единственного предельного цикла.

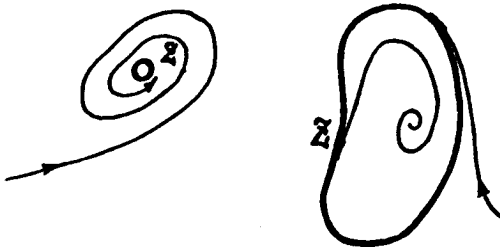


Рис. 4. Периодический стационарный режим — устойчивый предельный цикл. При малой амплитуде (справа) экспериментально неотличим от фокуса

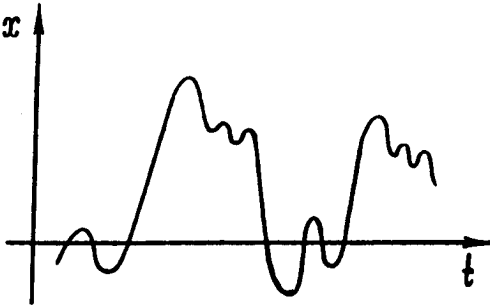


Рис. 5. Запись только одного x во времени значительно проще экспериментально, но существенно беднее по содержанию. Осциллограф хорошо, а дисплей лучше

в понимании того, что такое стационарное состояние. Старые математические представления о стационарных состояниях (более точно: те, которые использовались в биологии) были настолько бедны по сравнению с богатейшим биологическим материалом, что все время казалось, что биология — это одно, а математика — это другое. Кстати, так оно и есть; это все правда, но это не вся правда. А вот новая часть правды (как мне кажется) состоит в том, что мы сейчас осознали, что правильное понимание стационарных состояний при всей колоссальной важности описанных выше двух типов к ним не сводится. Имеется значительно большее богатство стационарных состояний. К сожалению, эта тема очень трудна, и я смогу ее только обозначить.

В математике уже давным-давно известны очень своеобразные режимы, в которых могут находиться сложные системы, как эволюционирующие, так и (особенно) стационарные. Эти режимы были впервые обнаружены в небесной механике, в сложных теоретических вопросах движения планет, очень активно разрабатывались в теории турбулентности, а в последнее время наиболее популярны стали в связи с вопросами метеорологии, в связи с неустойчивостью режимов погодных, процессов погодных. И наиболее знаменитая и популярная — слухи о ней, наверное, до всех дошли — это система Лоренца. Я не буду о ней рассказывать, не пугайтесь, но основную идейную картину нарисую.

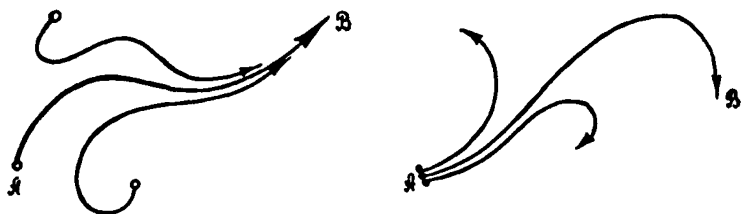


Рис. 6. Устойчивость траектории АВ в «хорошей» системе. В такой ситуации легко быть пророком.

Рис. 7. Экспоненциальная неустойчивость вблизи траектории АВ. Здесь безопаснее быть историком.

Мы привыкли вот к чему. Если у вас есть какая-то система, то ее движение в пространстве (геометрическом, фазовом или каком-либо ином) можно изобразить просто в виде линии (траектории). Обычная задача по теории устойчивости формулируется так: что будет, если мы немного отклонимся от начальной точки. Как будет вести себя новая траектория? Привычная точка зрения такова: в хорошей системе она далеко не уйдет от первой (рис. 6). Если система неустойчива, то будет сильно уходить (рис. 7). Вот наша стандартная картина, вот, что нам привычно. Вот то, что мы видим. Эта картина для плоскости (x, y) является почти неизбежной.

Но если число переменных, описывающих систему, больше двух — хотя бы три (а для биологии и сто отдать не жалко в качестве фазовых переменных даже для простейших систем), так вот там такие картинки нереалистичны уже довольно часто, и вместо таких режимов мы приходим к более сложным. Я изображу типичную картину.

Возьмем маленький объем — каплю (считайте, что это просто струйки текущей жидкости). При увеличении скорости струя становится уже, при уменьшении — шире (рис. 8). Закон Бернулли, если помните. Этот объем течет, деформируясь, и немного меняя форму. На плоскости так оно и есть.

Нечто совсем иное происходит в трехмерном пространстве — не всегда, конечно, но достаточно часто, чтобы с этим пришлось считаться. Оказывается, что картина будет совсем другой. Если мы капнем капельку окрашенной жидкости в трехмерную струю, то судьба этой окрашенной капельки будет такова: через некоторое время она превратится в нечто, похожее на листок клевера из трех *лепестков (рис. 9). Пройдет еще некоторое время и каждый из лепестков тоже разстроится, но общий объем капли сохранится прежним. Система чрезвычайно неустойчива, но локальна. Все это происходит в фиксированном небольшом объеме. Капелька не

* Или двух, или четырех, а то и больше.

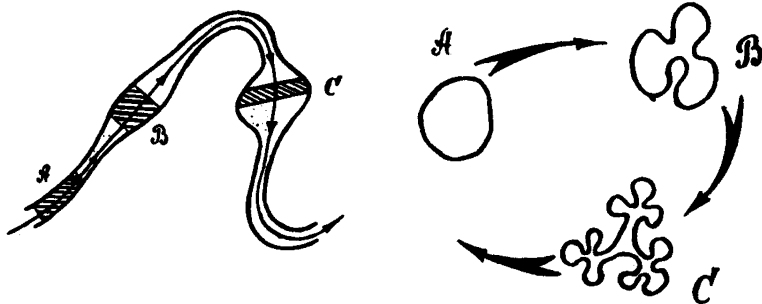


Рис. 8. Трубка тока. Один и тот же «жидкий объем» А несколько замедляется в положении В и резко тормозится в положении С

Рис. 9. Растекание капельки в локально неустойчивых системах. Фазовый объем в положениях А, В, и С один и тот же

может дробиться, она размазывается по всему небольшому объему.

Поэтому, если вы сосчитаете на ЭВМ такую систему, то картина, которую вы увидите, будет следующая: могут появиться ограниченные объемы, которые заполнены довольно прихотливыми траекториями. Качественный рисунок довольно близок к реальности. Такого типа картинок сосчитано уже довольно много. И вот эту «непонятность» и надлежит рассматривать в качестве правильного аналога стационарного состояния.

Именно то безобразие*, которое здесь происходит, — это и есть стационарный режим, типичный для трехмерных систем. Никуда далеко траектория не уходит, все остается здесь и происходит перемешивание — вот правильное слово. Вы знаете, на плоскости невозможны пересечения фазовых траекторий. Но все происходит в трехмерном пространстве, это пространственная кривая. Она нигде не самопересекается, но в проекции на любую плоскость возникает подобная типичная картина со множеством самопересечений.

Итак, кратко повторяю еще раз: старое понимание стационарных режимов — устойчивое равновесие; современное, достаточно широко распространенное понимание — предельные циклы. Поэтому то, чем сейчас серьезно следует заниматься — это новые типы стационарных режимов.

Каковы стандартные подходы к подобным системам, в которых реализуются такого рода стационарные режимы, когда мы сталки-

* Безобразие — это отсутствие образа. Точнее, отличное от привычного и приятно-понятного (в силу привычки) образа.

Рис. 10. Плотность «вероятности» нахождения системы в разных точках. Правильнее говорить о «среднем времени пребывания» изображающей точки в данном месте. Нарисованы изохроны — линии равной длительности пребывания. Не смешивать с траекториями!

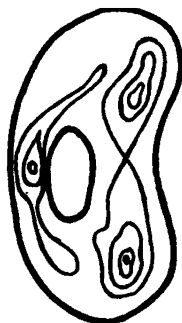
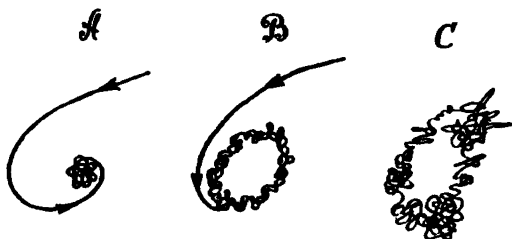


Рис. 11. Во всех трех случаях есть зоны перемешивания. Однако экспериментатор скажет, что в случае А имеется стационар, а в случае В ему удалось обнаружить предельный цикл. В случае С он либо начнет чинить прибор, либо свалит все на солнечную активность



ваемся с ними в реальном эксперименте? Конечно, статистика! Вычисляется вероятность системы в данной точке — ничего дурного в этом нет, это хорошее приближенное понимание. Но этого недостаточно. Что мы будем знать при таком подходе? Представьте себе, что основную часть времени система проводит на периферии, там она медленно движется, середину проскакивая быстро. Тогда, если вы нарисуете плотность вероятности, то у вас получится так: жирно здесь и плавно спадает к центру. Вот нормальная картина, которую вы в эксперименте получаете. Вероятностная оценка ситуации (рис. 10.) Можно однако понимать это более точно. Можно все-таки сказать, что система не просто размазана по некоторому объему, а еще есть некоторые предпочтительные вероятности переходов из одной части в другие. Но еще лучше вообще не говорить о «вероятности» переходов, лучше назвать кошку кошкой и просчитать истинную траекторию движения.

Вот разница в трех подходах. Точнее, это последовательное развитие одного и того же подхода. Потому что если область, которую заполняет траектория, достаточно мала, то мы ее воспринимаем как стационарную точку. Если эта область достаточно велика, но вытянута (типичный пример — узкое кольцо, внутри которого происходит движение) — мы говорим — линия. Если оба радиуса велики, но близки по размерам, то это воспринимается как предельный цикл. А вот если они становятся заметно различными, то мы вынуждены говорить: *система с перемешиванием*.

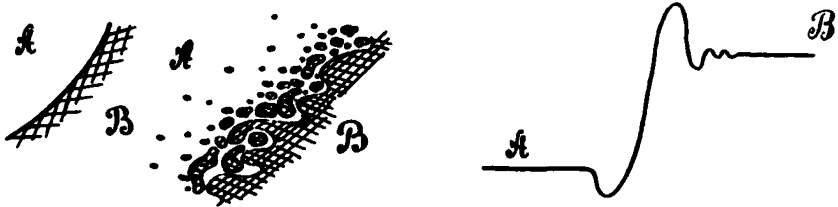


Рис. 12. Слева граница между биогеоценозами в плохих книгах, справа — в хороших

Рис. 13. Переходная зона (типа ударной волны). Сравните с рис. 2, изображающим переходный процесс (то есть во времени, а не в пространстве) и в более простом случае

Итак, в зависимости от конкретной геометрии системы мы воспринимаем стационарные режимы то как точки, то как линии, то, наконец, как объемы (рис. 11.).

Хочу подчеркнуть еще одно важное обстоятельство, тесно связанное со всем только что указанным. Мы сейчас должны большое внимание — если, конечно, хотим понять эту новую ситуацию — уделять границам. Классический пример такой. Два соседствующих биогеоценоза друг с другом **взаимодействуют**, один обычно вытесняет другой. Границу между ними часто рисуют простой линией — это один, это другой. Но если вы почитаете хорошие учебники, хорошие книжки, то вы увидите, что правильную картину рисуют иначе (рис. 12).

Упрощенно, но в целом достаточно правильно: это — состояние *A*, это — *B*, а в зоне границы прихотливо перемешаны элементы *A* и *B*. Там непременно есть еще и другие элементы, и они играют важную роль, но мы для простоты будем считать, что граница есть перемешивание двух типов, переходная зона. Возникает интересный и серьезный вопрос: как устроен этот стационар *A*? Как устроен стационар *B*? Они оба чаще всего метастабильны, т.е. стационарны только в каких-то узких условиях. И, наконец, самое любопытное: как происходит переход от одного к другому? Возникает, следовательно, интересный, самостоятельный, независимый вопрос о поведении границы.

Я ограничусь двумя замечаниями. Замечание первое: есть понятие ударной волны. Эта ситуация, когда у вас есть два стационара точечного типа и переход из состояния *A* в состояние *B* происходит довольно резко (с крутым фронтом). Тогда мы говорим, что имеется движущая ударная волна из *B* по *A* (рис. 13.). Понятие это возникло в гидродинамике, но оно достаточно хорошо применимо ко многим другим ситуациям. У этого понятия могут быть два аспекта. По оси абсцисс можно откладывать «*x*», а можно откладывать «*t*». В одном случае вы будете иметь геометрию, а в другом — кинетику. Если переходный процесс идет быстро, то пространственный фронт крутой. Обычно прост-

ранственная картина есть развертка* в пространстве** временного процесса, а временная картина есть развертка во времени пространственной ситуации. Это двойственные вещи: либо « x », либо « t ». Важно то, что уже в простых задачах не всегда реализуется S-образная кривая. Довольно часто картина бывает сложнее.

Вот если у вас есть два стационарных уровня, то картина бывает, как на рис. 13. Это достаточно широко известное явление. Но куда интереснее, когда эта система, этот переходный процесс устроен как на рис. 14. Не видно, что там было раньше, не видно, что будет потом, но очень видно, что происходит сейчас. А происходит нечто непонятное.

Извинительно поэтому, что человек, которому показывают такую картину, говорит: «Конечно, случайность». Но мне кажется, что в ситуациях, подобных этой, слишком уж комфортно и просто сказать: стохастика, вероятность, случайность. Фактически это отказ от анализа. Я расскажу свой любимый анекдот об умной девочке. Девочка была маленькая и умела считать до двух, а ей дали три ложки и попросили сосчитать. Девочка решительно отодвинула одну ложку и сказала: «Уберите эту ложку, она грязная».

Так вот мы сейчас умеем прекрасно считать до одного, просто великолепно, S-образные кривые, стационарные состояния, термодинамическое равновесие. Чуть хуже, но вполне прилично, умеем считать до двух (предельные циклы). А что касается трех, то мы решительно говорим: стохастика, случайность.

Я абсолютно убежден, что многие важные идеи и факты классической биологии могут быть поняты только вместе с математиками. Верно и обратное — серьезные части математики получают свое новое дыхание при обращении к биологическим задачам. Ищите контакты.



Рис. 14. Причиной такой «свиstopляски» конечно могут быть «наводки» от линии метро, слишком близко расположенной к физфаку МГУ. Ну, а если система и в самом деле такая?

* Запись на магнитофон.

** Пронгрывание записи.

ТЕОРЕМА ВЫРАВНИВАНИЯ. ПРИБЛИЖЕННОСТЬ ФОРМУЛЫ БОЛЬЦМАНА. «ЗАКОН» ВОЗРАСТАНИЯ ЭНТРОПИИ

Изучается математическая ситуация, типичная для массовых явлений любой природы. Показана возможность аксиоматического построения теории. Идея такого построения восходит к лекциям, прочитанным А. Я. Хинчиным в весеннем семестре 1950 г. на механико-математическом факультете МГУ.

п.1. Система и структурная функция

Аксиоматический подход основан на выделении главного понятия теории массовых явлений — понятия системы. Определение системы включает задание фазового пространства X , интегрирования по этому пространству, задаваемого неотрицательной (возможно дискретной) мерой dX , и скалярной* функции $H(X)$. Удобно обозначение

$$S = \{X, dX, H(X)\}, \quad (1)$$

явно указывающее на каждый из трех аспектов задания системы.

Важную роль в построении теории играет понятие структурной функции системы $\Omega(\epsilon)$. Эта функция задается следующим образом.

Рассмотрим пространство D бесконечно-дифференцируемых функций $f(\epsilon)$, заданных на прямой $-\infty < \epsilon < +\infty$. Формула

$$(\Omega, f) = \int_{-\infty}^{+\infty} \Omega(\epsilon) f(\epsilon) d\epsilon = \int_x f[H(X)] dX \quad (2)$$

определяет обобщенную, вообще говоря, функцию одного переменного ϵ , которая и называется структурной функцией системы.

* В частности, в статистической термодинамике $H(X)$ это энергия, точнее гамльтониан.

п.2. Объединение систем и компоненты

Любые две системы S_1 и S_2 порождают третью систему S :

$$S = (S_1, S_2) = S_1 + S_2 \quad (3)$$

объединение систем S_1 и S_2 , причем X , dX , $H(X)$ строятся по соответствующим аспектам подсистем следующим образом. Пространство X есть прямое произведение:

$$X = (X_1, X_2). \quad (4)$$

Мера dX есть прямое произведение мер dX_1 и dX_2

$$dX = dX_1 dX_2. \quad (5)$$

Функция H есть сумма функций H_1 и H_2

$$H(X) = H_1(X_1) + H_2(X_2). \quad (6)$$

Системы S_1 и S_2 называются в этом случае компонентами системы S . Обобщение на случай нескольких компонент очевидно.

Вычисление для системы S значения функционала $\Omega(\epsilon)$,

$$\begin{aligned} (\Omega, f) &= \int_X f(H) dX = \int_{X_1} \int_{X_2} f(H_1 + H_2) dX_1 dX_2 = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\epsilon_1 + \epsilon_2) \Omega_1(\epsilon_1) \Omega_2(\epsilon_2) d\epsilon_1 d\epsilon_2 \end{aligned}$$

непосредственно из определения [1] свертки обобщенных функций приводит к Теореме о свертке структурных функций компонент

$$\Omega(\epsilon) = \int_{-\infty}^{+\infty} \Omega_1(E_1) \Omega_2(\epsilon - E_1) dE_1. \quad (7)$$

п.3. Ведущая функция системы

Всюду в дальнейшем предполагается, что в пространство D входят все комплекснозначные функции, убывающие при $\epsilon \rightarrow \infty$ быстрее экспоненты. В приложениях это условие обычно выполнено, так как функция $H(X)$ чаще всего ограничена снизу и растет степенным образом при $|X| \rightarrow \infty$. В этом случае преобразование Лапласа структурной функции является настоящей (а не обобщенной) функцией и может быть записано в двух эквивалентных формах:

$$\Phi(\beta) = \int_{-\infty}^{+\infty} e^{-\beta E} \Omega(E) dE = \int_X e^{-\beta H(X)} dX \quad (8)$$

Функция $\Phi(\beta)$ называется ведущей функцией системы и обладает следующими замечательными свойствами:

1. $\Phi(\beta)$ аналитична в полуплоскости $\text{Re}\beta > 0$.
2. $\Phi(\beta)$ монотонно убывает вдоль действительной оси β .
3. $|\Phi(\beta)| \leq \Phi(\text{Re}\beta)$.
4. $\Phi(\beta)$ логарифмически выпукла на действительной оси β .

Докажем последнее свойство. Вычисление второй производной $\ln\Phi$ дает

$$\frac{d^2 \ln \Phi}{d\beta^2} = \frac{1}{\Phi^2} \left[\Phi \frac{d^2 \Phi}{d\beta^2} - \left(\frac{d\Phi}{d\beta} \right)^2 \right]$$

Первое слагаемое, стоящее в квадратных скобках, может быть записано в виде двойного интеграла, получающегося при умножении однократных:

$$\Phi \frac{d^2 \Phi}{d\beta^2} = \int_X \int_Y H^2(Y) e^{-\beta[H(X)+H(Y)]} dX dY.$$

Так как величина интеграла не зависит от обозначения переменных интегрирования, то выражение для $\Phi\Phi''$ можно записать в симметризованной форме:

$$\Phi\Phi'' = \frac{1}{2} \int_X \int_Y [H^2(X) + H^2(Y)] e^{-\beta[H(X)+H(Y)]} dX dY.$$

Подставляя аналогичное выражение для $(\Phi')^2$, получаем окончательную формулу

$$\frac{d^2 \ln \Phi}{d\beta^2} = \frac{1}{2\Phi^2} \int_X \int_Y [H(X) - H(Y)]^2 e^{-\beta[H(X)+H(Y)]} dX dY \geq 0 \quad (9)$$

из которой очевидна неотрицательность второй производной $\ln\Phi$ и, следовательно, выпуклость этой функции.

Особенно наглядно удобство для приложений использования функции Φ вместо Ω демонстрирует

Теорема о произведении ведущих функций

Ведущая функция $\Phi(\beta)$ системы равна произведению ведущих функций компонент:

$$\Phi(\beta) = \Phi_1(\beta) \Phi_2(\beta). \quad (10)$$

Доказательство непосредственно вытекает из определения компонент (3), (4), (5), (6) и определения (8) ведущей функции

$$\begin{aligned}
\Phi(\beta) &= \int_{\mathcal{X}} e^{-\beta H(X)} dX = \int_{\mathcal{X}_1} \int_{\mathcal{X}_2} e^{-\beta H_1 - \beta H_2} dX_1 dX_2 = \\
&= \int_{\mathcal{X}_1} e^{-\beta H_1(X_1)} dX_1 \int_{\mathcal{X}_2} e^{-\beta H_2(X_2)} dX_2 = \\
&= \Phi_1(\beta) \Phi_2(\beta)
\end{aligned}$$

п.4. Формула обращения для структурной функции

Как уже было сказано выше, в отличие от обобщенной функции $\Omega(\varepsilon)$ ведущая функция $\Phi(\beta)$ является настоящей и к тому же еще аналитической функцией своего аргумента. Вместе с тем задание $\Phi(\beta)$ эквивалентно заданию $\Omega(\varepsilon)$, так как функцию $\Omega(\varepsilon)$ можно восстановить [2] по $\Phi(\beta)$, используя формулу обращения преобразования Лапласа:

$$\Omega(\varepsilon) = \frac{1}{2\pi i} \int_{\Gamma} e^{\beta\varepsilon} \Phi(\beta) d\beta \quad (11)$$

В этой формуле путь интегрирования Γ может быть выбран в силу независимости интеграла от пути произвольно в правой полуплоскости, лишь бы он асимптотически выходил на прямую, параллельную мнимой оси. Наиболее удобно выбирать такую прямую, вдоль которой подинтегральное выражение имеет минимальную величину, ибо в этом случае сводятся к минимуму интерференционные явления, сильно затрудняющие (теоретическое или фактическое) вычисление интеграла. В этом преодолении интерференции и состоит, собственно, основная идея *метода перевала* [3]. Третье из свойств функции $\Phi(\beta)$ показывает, что минимум следует искать на действительной оси.

Логарифм подинтегральной функции в формуле (11) есть функция двух переменных

$$\psi(\beta, \varepsilon) = \beta\varepsilon + \ln\Phi(\beta) \quad (12)$$

растущая как при $\beta \rightarrow \infty$ (потому что тогда растет первое слагаемое, а второе убывает), так и при $\beta \rightarrow 0$ (когда слагаемые меняются ролями). Так как кроме того ψ линейным слагаемым отличается от выпуклой (в силу (9)) функции $\ln\Phi(\beta)$ значит и сама функция ψ выпуклая. Поэтому при каждом значении ε существует и притом единственная точка β , в которой достигается минимум функции ψ :

$$S(\varepsilon) = \min_{\beta} (\beta\varepsilon + \ln\Phi) = \min_{\beta} \psi(\beta, \varepsilon) \quad (13)$$

Точка β может быть найдена из уравнения

$$\frac{d \ln \Phi}{d \beta} = -\varepsilon, \quad (14)$$

получающегося приравнением нулю производной Φ по β . Именно в этой точке вертикальная прямая Γ в комплексной плоскости β дает наилучший путь интегрирования для вычисления $\Omega(\varepsilon)$. Значение S в этой точке удобно выразить через параметр β :

$$S = \ln \Phi - \beta \frac{d \ln \Phi}{d \beta} \quad (15)$$

Формулы (14) и (15) подсказывают целесообразность введения параметра β в качестве основного переменного, ибо остальные величины хорошо выражаются именно через β .

п.5. Теорема выравнивания. Интенсивность и экстенсивность

Функция $\psi(\beta, \varepsilon)$ является, подобно $H(X)$, аддитивной функцией системы в том смысле, что при объединении компонент S_1 и S_2 в новую систему S функции ψ_1 и ψ_2 складываются:

$$\psi(\beta, \varepsilon) = \psi_1(\beta, \varepsilon_1) + \psi_2(\beta, \varepsilon_2). \quad (16)$$

Это свойство вытекает из аддитивности H и мультипликативности ведущей функции $\Phi(\beta)$.

Из выпуклости функций ψ_1, ψ_2 и ψ вытекает замечательная [4].

Теорема выравнивания («закон» возрастания энтропии)

$$S \geq S_1 + S_2; \quad \beta_1 \leq \beta \leq \beta_2. \quad (17)$$

Доказательство получается немедленно из определения S_1 и S_2

$$\psi_1(\beta, \varepsilon_1) \geq \psi_1(\beta_1, \varepsilon_1) = \min_{\beta} \psi_1(\beta, \varepsilon_1) = S_1 \quad (18)$$

$$\psi_2(\beta, \varepsilon_2) \geq \psi_2(\beta_2, \varepsilon_2) = \min_{\beta} \psi_2(\beta, \varepsilon_2) = S_2 \quad (19)$$

Поэтому при всех значениях β имеет место неравенство

$$\psi(\beta, \varepsilon) \geq \psi_1(\beta_1, \varepsilon_1) + \psi_2(\beta_2, \varepsilon_2). \quad (20)$$

из которого, в частности, в точке β минимума ψ и вытекает требуемое неравенство

$$S \geq S_1 + S_2 \quad (21)$$

Кроме неравенства для «экстенсивностей»* из неравенства (20)

* *ex* — не только «бывший», но и «вне».

следует также (в силу выпуклости ψ_1, ψ_2 и ψ) двойственное неравенство $\beta_1 \leq \beta \leq \beta_2$, смысл которого состоит в том, что параметр β объединенной системы расположен между* параметрами компонент. Это обстоятельство выясняет важнейшую роль параметра β , который можно, следовательно, считать мерой *интенсивности* в противоположность экстенсивному параметру S .

Равенство (предельный случай неравенства экстенсивностей)

$$S = S_1 + S_2 \quad (22)$$

(как это опять-таки вытекает из неравенства для функций ψ) имеет место тогда и только тогда, когда

$$\beta_1 = \beta_2 = \beta_3. \quad (23)$$

Иными словами, аддитивность является предельным случаем экстенсивности, достигаемым только при равенстве интенсивностей. Отсюда вытекает, что параметр β задает также условие равновесия компонент (по отношению к объединению, смешиванию).

После сказанного не удивительно, что в статистической термодинамике β совпадает с обратной температурой (в единицах k), а S есть энтропия.

Появление аналогичных величин в теории передачи сообщений (теория «информации») означает совпадение математического аппарата, что и является, в сущности, главной темой статьи. Иногда [5] из совпадения формул заключают о совпадении физической природы объектов теорий. Такой вывод неправомерен. Плодотворность аксиоматического метода основана именно на одинаковости кинетики (или статики) объектов, имеющих совершенно разную физическую природу, что и позволяет изучать их поведение абстрактно. Так, например, поведение физического маятника и колебательного контура в радиотехнике (а также некоторых экологических систем) описывается одним и тем же уравнением, но было бы странно делать заключение об одинаковости физической сущности явлений. Другая крайность [6] приводит к бесполезному противопоставлению функции $H(X)$, взятой из химической кинетики, совершенно аналогичной функции $H(X)$, но взятой из механики. Предлагаемая статья содержит попытку найти разумную среднюю линию, основанную на аксиоматическом подходе, учитывающем самым внимательным образом свойства системы, но целиком пренебрегающем происхождением этой системы.

* *in* — «внутри»

п.6. Атомистичность

Идея атомистичности — одно из самых древних научных завоеваний. Ее первоначальной формой является представление о неделимости, неразложимости элементов, из которых состоит «все сущее». Однако развитие этой идеи все более проясняет решающую роль двух других аспектов идеи атомистичности — массовости и одинаковости (похожести) элементов, составляющих изучаемую систему. Суть дела не в том, что элемент (компонента, «атом», частица) бесструктурен, а в том, что его структура почти несущественна для свойств большой системы. Элементов очень много и строение (или функционирование) полной системы определяется не деталями структуры элементов, а основными, грубыми их свойствами. Элементарная модель твердых шариков и уточненная схема квантовой физики в равной мере приводят к закону Бойля-Мариотта для разреженных газов.

Общие утверждения, высказанные в предыдущих параграфах, справедливы для произвольных систем. Однако именно в силу всеобщности они не очень содержательны и носят скорее качественный, нежели количественный характер.

Точные утверждения, являющиеся основой многих количественных закономерностей, можно установить, сконцентрировав внимание на свойствах больших систем. Математическая идеализация состоит в предельном переходе $\frac{1}{n} \rightarrow 0$ и установлении вытекающих из него предельных соотношений.

Обычно для получения результата вводится, кроме $\frac{1}{n}$, еще какой-нибудь малый параметр. В работе [7], например, это малая плотность. В настоящей работе предпринята попытка обойтись без дополнительных гипотез, используя только один малый параметр $\delta = \frac{1}{n}$. Разложение по степеням этого малого параметра приводит к иерархии структурных свойств компонент, составляющих изучаемую систему. Нечто похожее происходит, например, в электростатике. На больших расстояниях даже очень сложная система зарядов может быть с успехом заменена точечным зарядом. При уточнении (особенно если полный заряд равен нулю) необходим уже учет структуры, но сначала лишь в форме введения эквивалентного диполя. Дальнейшие уточнения приводят к идее мультиполей.

п.7. Формула Больцмана

Из теоремы, сформулированной в конце п.3, нетрудно получить формулу для ведущей функции большой системы

$$\phi_n(\beta) = [\varphi(\beta)]^n. \quad (24)$$

Главная особенность этой формулы — превращение* аморфного индекса « n » в левой части (имеющего скорее порядковый характер) в количественный показатель степени справа, приобретающий уже алгоритмически точный смысл.

Выпишем формулу для структурной функции,

$$\Omega_n(\varepsilon) = \frac{1}{2\pi i} \int_{\Gamma} e^{z\varepsilon + n \ln \varphi(z)} dz. \quad (25)$$

где обозначение z должно подчеркнуть, что путь интегрирования лежит в комплексной области. Лапласу принадлежит идея интегрировать вдоль такой прямой Γ_β (параллельной мнимой оси)

$$z = \beta + i\xi \quad (26)$$

которая пересекает действительную ось в точке β , в которой подинтегральная функция имеет минимум (метод перевала). Ясно, что при разных ε прямую Γ придется проводить через различные точки β , определяемые уравнением

$$n \frac{d \ln \varphi}{d\beta} + \varepsilon = 0. \quad (27)$$

Существенным развитием этой мысли является предложение** именно β считать независимой переменной, а ε рассматривать как функцию β . Выше мы уже видели (15), что и экстенсивность S весьма просто выражается через β :

$$S = n \left[\ln \varphi(\beta) - \frac{d \ln \varphi}{d\beta} \beta \right]. \quad (28)$$

Наша ближайшая задача — найти выражение через n и β для структурной функции большой системы. Подинтегральная функция имеет максимум*** как раз в точке β (конечно, для ε , задаваемого формулой (27)) и максимум этот равен e^S где S — экстенсивность, вычисляемая при заданном β по формуле (28). Вынесение экспоненциального множителя за знак интеграла

$$\Omega_n(\beta) = e^{S_n \beta} J_n(\beta) \quad (29)$$

сводит задачу вычисления структурной функции к вычислению интеграла $J_n(\beta)$, под знаком которого стоит функция, модуль

* Сейчас это «очевидно» — через полтора века (1812) после первых работ Лапласа, полвека (1901) после теоремы Ляпунова и четверть века (1943) после книги Хинчина. Да и так ли уж до конца здесь все ясно?

** В книге Хинчина эта мысль четко сформулирована, но восходит она если не к Лапласу, то уж к Больцману и Гиббсу, наверное.

*** Максимум — при движении вдоль Γ и, наоборот, минимум — при движении вдоль действительной оси. В этом, конечно, суть перевальной точки.

которой всюду меньше единицы (кроме $z = \beta$, где эта функция как раз равна единице).

Так же как и $\Omega_n(\epsilon)$ новая функция $J_n(\epsilon)$ является, вообще говоря, обобщенной функцией своего аргумента β . Однако в классическом случае статистической механики обе они настоящие функции и соотношение (29) можно прологарифмировать:

$$\ln \Omega_n = S_n + \ln J_n. \quad (30)$$

Ниже будет показано, что величина J_n всегда имеет степенной порядок по n . Поэтому $\ln J_n$ — величина порядка $\ln n$ в то время как S_n пропорционально n .

Отбрасывание* $\ln J_n$ приводит к знаменитой формуле Больцмана**

$$S \approx \ln \Omega. \quad (31)$$

Естественно сохранить название *формула Больцмана* и для более общего случая. Однако распространение формулы Больцмана на случай произвольных систем возможно только в формуле (29), но не в формуле (30), так как понятие логарифма обобщенной функции неопределено.

Поэтому все дальнейшие вычисления сводятся к установлению асимптотики интеграла $J_n(\beta)$. Удобно сохранить для $J_n(\beta)$ термин «предэкспонента» (широко распространенный в физике) независимо от того «до или после» больцмановской экспоненты написан этот множитель.

Весьма любопытно «распределение обязанностей» между экспонентой и предэкспонентой. Количественной стороной величины Ω_n «заведует» экспонента, являющаяся настоящей, а не обобщенной функцией β (или ϵ , это все равно). Более того, эта функция даже аналитична в некоторой полуплоскости, лежащей правее мнимой оси в комплексной плоскости β . Предэкспонента, наоборот, почти не оказывает влияния на Ω_n в количественном отношении, но зато именно в ней сосредоточены все особенности (слагаемые типа δ — функции Дирака), определяющие принадлежность $\Omega_n(\beta)$ к определенному классу обобщенных функций.

* Это пренебрежение особенно убедительно в кинетической теории газов, когда число частиц n имеет порядок 10^{19} на кубический сантиметр.

** Критику попыток обосновать утверждение о том, что «энтропия системы в том или другом ее состоянии пропорциональна логарифму вероятности этого состояния» см. [4], стр. 105 (прим.ред)

п.8. Асимптотика предэкспоненты

Выражение для предэкспоненты имеет вид:

$$J_n(\beta) = \frac{1}{2\pi i} \int_{\Gamma} e^{n\sigma(z, \beta)} dz, \quad (32)$$

где функция $\sigma(z, \beta)$ выражается через логарифм ведущей функции одной компоненты

$$\psi(\beta) = \ln \varphi(\beta) \quad (33)$$

весьма любопытной формулой (заставляющей вспомнить преобразование Лежандра):

$$\sigma(z, \beta) = \psi(z) - \psi(\beta) - \psi'(\beta)(z - \beta). \quad (34)$$

Эта функция переменной z по самому своему происхождению устроена так, что при $z = \beta$ она обращается в нуль вместе со своей первой производной. Так как она аналитична по z , то ее можно разложить в ряд по степеням $(z - \beta)$. Этот ряд начинается, конечно, с члена второго порядка по $(z - \beta)$

$$\sigma(z, \beta) = \sum_{k=2}^{\infty} \frac{\psi^{(k)}(\beta)}{k!} (z - \beta)^k. \quad (35)$$

Самое нетривиальное соображение, определяющее весь ход дальнейших выкладок, — понимание решающей роли квадратичного члена в разложении функции $\sigma(z, \beta)$. Это и есть основная идея Лапласа — идея метода перевала.

Выявление механизма, создающего асимптотическое разложение, основано на выборе правильной переменной интегрирования вдоль пути Γ . Вместо комплексного переменного вводится новое действительное переменное τ :

$$z = \beta + \frac{i\tau}{\sqrt{n}}. \quad (36)$$

Суть дела именно в множителе $\frac{1}{\sqrt{n}}$. Он выбран как раз так, чтобы главный член в показателе (квадратичный, как мы предполагаем) не содержал n . Именно это соображение определяет масштаб переменной интегрирования. Наши ожидания оправдываются, ибо разложение $n\sigma$ приобретает вид:

$$n\sigma = -\frac{a^2\tau^2}{2} + \sum \left(\frac{i}{\sqrt{n}}\right)^k a_k \quad (37)$$

и, следовательно, все члены, кроме главного, стремятся к нулю при $n \rightarrow \infty$. Для отчетливого выявления главной идеи в формуле (37) краткими символами a_k обозначены весьма непростые функ-

ции трех переменных — индекса k , переменной интегрирования τ и параметра β :

$$a_k = -\frac{\psi^{(k+2)}(\beta)}{(k+2)!}(\tau)^{k+2} \quad (38)$$

Главное слагаемое также может быть получено отсюда при $k=0$. Однако для того, чтобы подчеркнуть его роль и особенно важный в дальнейшем факт существенной отрицательности этого члена, введено специальное обозначение:

$$\psi''(\beta) = a^2, \quad (39)$$

право на которое дает выпуклость функции $\psi(\beta)$. В результате под знаком интеграла возникает множитель $e^{-\frac{a^2\tau^2}{2}}$ быстро стремящийся к нулю при $|\tau| \rightarrow \infty$. Это означает, что основной вклад в величину интеграла дает конечный интервал изменения τ и, следовательно, ничтожно* малый (порядка $\frac{1}{\sqrt{n}}$) отрезок пути Γ в плоскости z .

Главное же сделано, выяснена идейная сторона задачи, остается только техническая стадия, впрочем довольно громоздкая.

Бесконечный ряд в $n\sigma$ порождает под знаком интеграла бесконечное произведение A .

$$A(\tau, \beta, n) = \prod_{k=1}^{\infty} e^{a_k \left(\frac{i}{\sqrt{n}}\right)^k}, \quad (40)$$

которое нужно преобразовать в бесконечную сумму

$$A = 1 + \sum_{k=1}^{\infty} \left(\frac{i}{\sqrt{n}}\right)^k A_k(\beta, \tau), \quad (41)$$

подставить эту сумму в формулу для предэкспоненты и почленно проинтегрировать. Получится искомое разложение предэкспоненты в ряд по обратным степеням \sqrt{n} :

$$J_n(\beta) = \frac{1}{2\pi\sqrt{n}} \int_{-\infty}^{+\infty} e^{-\frac{a^2\tau^2}{2}d\tau} \left[1 + \sum_{m=1}^{\infty} \left(\frac{i}{\sqrt{n}}\right)^m \frac{\int_{-\infty}^{+\infty} A_m(\beta, \tau) e^{-\frac{a^2\tau^2}{2}d\tau}}{\int_{-\infty}^{+\infty} e^{-\frac{a^2\tau^2}{2}d\tau}} \right]. \quad (42)$$

Интегралы, входящие в эту формулу, нетрудно вычислить, так как коэффициенты A_m являются многочленами от τ . Дело сводится поэтому к отысканию интегралов вида:

* В статистической термодинамике 10^{-9} .

$$\int_{-\infty}^{+\infty} \tau^{2q} e^{-\frac{a^2 \tau^2}{2}} d\tau = \frac{1}{a^{2q+1}} \frac{(2q)!}{2^q q!} \sqrt{2\pi} \quad (43)$$

Если же под знаком интеграла стоит нечетный множитель τ^{2q+1} , то интеграл просто равен нулю.

Самая громоздкая часть вычислений — это выражение коэффициентов A_m через величины a_k . Прямое разложение показательной функции к результату:

$$A_m = \sum_{l=1}^m \frac{1}{l!} \sum_{k_1+\dots+k_l=m} \dots \sum a_{k_1} \dots a_{k_l} \quad (44)$$

Подстановка выражений для a_k дает:

$$A_m(\tau, \beta) = \sum_{l=1}^m \frac{\tau^{m+2l}}{l!} (-1)^l \sum_{k_1+\dots+k_l=m} \dots \sum \frac{\psi^{(k_1+2)} \dots \psi^{(k_l+2)}}{(k_1+2)! \dots (k_l+2)!} \quad (45)$$

Из этой формулы видно, что все нечетные члены в (42) обращаются при интегрировании в нуль. Поэтому разложение $J_n(\beta)$ содержит только члены $m=2p$ и, следовательно, действительные члены, и происходит по целым (не считая общего множителя $\frac{1}{\sqrt{n}}$) степеням $\frac{1}{n}$

$$J_n(\beta) = \frac{1}{\sqrt{2\pi n} a(\beta)} \left[1 + \sum_{p=1}^{\infty} \frac{(-1)^p}{n^p} K_p(\beta) \right] \quad (46)$$

В этой асимптотической формуле коэффициенты $K_p(\beta)$ весьма громоздким образом выражаются через производные только одной функции β , а именно $\psi(\beta) = \ln \phi(\beta)$. Эти выражения имеют следующий вид:

$$K_p(\beta) = \sum_{l=1}^{\infty} \frac{(-1)^l}{a^{2p+2l+1} 2^{p+l} (p+l)!} \times \\ \times \sum_{\substack{r_1+\dots+r_l \\ +r_l}} \dots \sum_{\geq 2p+2l} \frac{(r_1+\dots+r_l)!}{r_1! \dots r_l!} \psi^{(r_1)} \dots \psi^{(r_l)}, \quad r_i \geq 2 \quad (47)$$

Обычно бывает достаточно первых двух членов разложения $J_n(\beta)$

$$J_n(\beta) = \frac{1}{\sqrt{2\pi n} a(\beta)} \left\{ 1 + \frac{1}{24n} \left[3 \frac{\psi^{IV}}{a^5} - 5 \frac{(\psi^{III})^2}{a^7} \right] + O\left(\frac{1}{n^2}\right) \right\}, \quad (48)$$

которые, впрочем, можно получить значительно проще (минуя общие выкладки) прямо из формулы (32).

Найденные разложения дают асимптотику структурной функции

$$\Omega_n(\varepsilon) = \frac{e^{n(\ln \varphi - \frac{d \ln \varphi}{d\beta} \beta)}}{\sqrt{2\pi n (\ln \varphi)^n}} \left\{ 1 + \frac{1}{24n} \left[3 \frac{\psi^{IV}}{\alpha^5} - 5 \frac{(\psi^{III})^2}{\alpha^7} \right] + O\left(\frac{1}{n^2}\right) \right\}, \quad (49)$$

которая может быть, конечно, продолжена до полного ряда. Но так как члены этого ряда получаются друг из друга сложным алгоритмом, содержащим операции дифференцирования, то, вообще говоря, нет ни малейшей надежды получить сходящийся ряд.

Однако функция Ω нигде не используется в «одиночестве». Всюду в теории она входит как весовой множитель при интегрировании какой-нибудь другой функции. Это обстоятельство позволяет применить к задаче современную точку зрения теории обобщенных функций, главная идея которой состоит в том, что подбором надлежащего пространства \mathfrak{H} можно ряду (49) придать точный смысл сходящегося ряда, если рассматривать его члены (аналитические, к слову сказать функции) как функционалы над векторами пространства \mathfrak{H} . Детали построения этого пространства \mathfrak{H} нас сейчас не интересуют [1].

п.9. Интенсивность β как независимое переменное

Внутренняя логика разбираемой схемы подсказывает, что именно β удобнее всего выбрать в качестве основного переменного.

Соберем воедино, с этой точки зрения, все основные формулы. Первый шаг состоит в вычислении ведущей функции («статистическая сумма»)

$$\varphi(\beta) = \int_X e^{-\beta H(X)} dX. \quad (50)$$

Следует отметить, что именно в этот момент происходит «обезличка» системы. Пространство X могло иметь самую разную природу. Оно могло быть фазовым пространством механической системы или химическим концентрационным пространством или пространством численностей в экологических задачах. Его геометрия могла быть довольно прихотливой, например, тор или дискретный набор точек, или бесконечномерное пространство. Функция $H(X)$ могла иметь какую угодно размерность (энергия или число букв) и любые поверхности уровня, в том числе состоящие из нескольких компонент.

При переходе к $\varphi(\beta)$ все эти индивидуальные различия полностью стираются — любая система заменяется эквивалентной одномерной системой, которая однозначно задается структурной

функцией. Было бы однако неоправданной односторонностью считать энергетические системы образцом строения.

Следующий шаг — отыскание зависимости ϵ от β . Раньше мы решали уравнение относительно β и отыскивали в функции. Теперь мы интерпретируем соотношение

$$\frac{\epsilon}{n} = \epsilon = -\frac{d \ln \varphi}{d \beta} \quad (51)$$

как задание ϵ в виде функции от α . Поучительно, что в термодинамике это соотношение есть уравнение состояния (в энергетической форме).

Наконец, основной результат — абстрактный аналог формулы Больцмана-асимптотика структурной функции также получает наиболее удобную форму в переменной β

$$\Omega_n(\epsilon) = e^{S(\beta)} \frac{1}{\sqrt{2\pi n} a(\beta)} \left\{ 1 + \frac{1}{24n} \left[3 \frac{\Psi^{IV}}{a^5} - 5 \frac{(\Psi^{III})^2}{a^7} \right] + \dots \right\}, \quad (52)$$

где величина s имеет вид

$$\frac{s}{n} = \ln \varphi - \frac{d \ln \varphi}{d \beta} \beta. \quad (53)$$

Полезной часто оказывается и «энергетическая», небольцмановская, форма асимптотики, в которой β и ϵ входят на равных правах как сопряженные величины:

$$\Omega_n(\epsilon) = \frac{e^{S^*}}{[\varphi(\beta)]^n} \cdot \frac{1}{\sqrt{2\pi n} a(\beta)} \left\{ 1 + O\left(\frac{1}{n}\right) \right\}. \quad (54)$$

В последней формуле поправочный, «предэкспоненциальный» множитель содержит функцию $a(\beta)$, которая выражается через $\varphi(\beta)$ весьма просто

$$a^2(\beta) = \frac{d^2 \ln \varphi}{d \beta^2} > 0. \quad (55)$$

В заключение полезно отметить еще одно любопытное обстоятельство. В частном случае термодинамики (разбирая для определенности пример идеального газа) можно установить следующую простую связь между интенсивностью β и температурой T :

$$\beta = \frac{1}{KT}. \quad (56)$$

Ландау [8] отмечает, что существуют ситуации, в которых имеют смысл отрицательные температуры. Он получает, после некоторых специальных рассуждений, что отрицательные температуры «лежат выше» положительных. Этот результат вытекает из формулы (56), если считать системы упорядоченными всегда в порядке убывания параметра β . Это означает, возможно, что параметр β имеет более общее и глубокое значение, нежели его частый случай — температура.

Литература

1. Гельфанд И.М., Шилев Г.Е. Обобщенные функции и действия над ними. М., Физматгиз, 1958.
2. Гельфанд И.М., Шилев Г.Е. Пространства основных обобщенных функций. М., Физматгиз, 1958.
3. Евграфов М.А. Асимптотические оценки и целые функции. М., Физматгиз, 1962.
4. Хинчин А.Я. Математические основания статистической механики. М., Гостехиздат, 1943.
5. Бриллюэн, Леон. Наука и теория информации. М., Физматгиз, 1960.
6. Гудвин Б. Временная организация клетки. (Пер. с англ.) М., «Мир», 1966.
7. Боголюбов Н.Н. О некоторых статистических методах в математической физике. Львов, 1945.
8. Ландау Л.Д., Лившиц Е.М. Статистическая физика. М., Гостехиздат, 1951, с. 237.

ОБ ОДНОЙ ТЕОРЕМЕ А.Я. ХИНЧИНА

Сформулирована и доказана теорема об аппроксимации функцией одной переменной любой сумматорной функции большого числа переменных.

Аппроксимация (в смысле интеграла от квадрата разности) становится все более точной по мере увеличения числа переменных.

Введение

А.Я.Хинчин в своих исследованиях по математическим основаниям статистической механики систематически использовал репрезентативность среднего значения сумматорных величин.

Этот факт является по существу теоремой анализа и в своих лекциях, к сожалению не опубликованных, А. Я. Хинчин особенно подчеркивал желательность освобождения от вероятностной терминологии.

Задача статьи — формулировка теоремы анализа, равносильной факту репрезентативности [1] среднего и установления связи этой теоремы с обобщенной [2] формулой Больцмана.

п. 1. Сумматорные функции. Разбиение на доминанту и флуктуанту

В книге Хинчина [1] все термины (кроме термина «сумматорная функция») взяты из термодинамики и статистической физики. Однако за четверть века, прошедших со времени выхода этой книги, стало ясно, что значение ее основных идей существенно выходит за рамки узкой задачи обоснования статистической механики. Эта самостоятельная область математического анализа — функции большого числа переменных со своими задачами и методами, но сохранившая, к сожалению чужую терминологию, неминуемо приводящую к путанице и противоречиям. Полная смена всей терминологии — это, конечно, крайность. Будет трудно понимать даже простые вещи только потому, что они сказаны непривычными словами. С другой стороны, такие слова, как «вероятность» и «энтропия», настолько перегружены истори-

ей науки, случайными*, привходящими ассоциациями, что должны быть решительно изгнаны из употребления. Новые же термины желательно вводить только в случае крайней необходимости. Автору кажется, что введение термина «доминанта» и «флуктуанта» не слишком дорогая цена за простое объяснение глубокой идеи А. Я. Хинчина.

Обычно считается (и, вообще говоря, это правильно), что сложность строения функции растет по мере увеличения числа ее аргументов. Тем более неожиданна идея Хинчина.

Для многих классов функций, важных для математического естествознания, имеет место прямо противоположная ситуация — они сводятся к функциям одной переменной E .

Суть дела, разумеется, именно в выделении правильного класса функций. Этот класс должен быть достаточно узким, чтобы можно было надеяться на существование нетривиальных закономерностей. Но вместе с тем он должен быть настолько широким, чтобы эти закономерности имели общую значимость.

Наша ближайшая задача — точная формулировка и доказательство принципа Хинчина для самого простого класса — сумматорных функций.

Определение

Функция $A(X)$

$$A(X) = a(x_1) + a(x_2) + \dots + a(x_n) \quad (1)$$

называется сумматорной, если функция $a(x)$ одна и та же для всех слагаемых, а ее аргументы x_1, x_2, \dots, x_n , задаются проекциями точки X ,

$$X = (x_1, x_2, \dots, x_n) \quad (2)$$

на все компоненты.

Теорема (Хинчин).

Любая сумматорная функция представима в виде суммы двух функций

$$A(X) = A(H) + F(X). \quad (3)$$

Первое слагаемое доминанта $A(H)$ имеет величину порядка n и сохраняет постоянное значение на каждой поверхности уровня H -функции.

* Лебег [3] пишет: «Если бы предел ρ_k был назван «тарарабумбий» круга, то вряд ли кто-нибудь позволил бы себе вывести из нее величину тарарабумбий сектора и сегмента, но делать это разрешают себе, когда вместо слова тарарабумбия употребили слово площадь! Это тягчайшее преступление против здравого смысла». Стоит добавить, что это детские игрушки по сравнению с тем, как измываются над замечательным понятием энтропии.

Второе слагаемое — флуктуанта $F(X)$ может меняться вдоль поверхности уровня, но зато относительно мала в том смысле, что ее среднее квадратичное значение имеет порядок \sqrt{n} . Флуктуанта, конечно, не всюду равномерно мала. Обязательно найдутся точки, в которых она сравнима с доминантой. Однако в тех задачах, где множества малой меры не имеют значения, функция огромного числа переменных может быть заменена функцией $A(E)$ одной существенной переменной и притом тем точнее, чем больше переменных у функции $A(X)$.

Суть идеи можно пояснить еще одним способом. Введем в пространстве X новую систему координат (E, Y) , используя функцию H , $E=H(X)$ в качестве одной из новых независимых переменных. Формула (3) в новых координатах имеет вид:

$$A(E, Y) = A(E) + F(E, Y). \quad (4)$$

Точка зрения замены переменных полезна тем, что подчеркивает главную мысль. Равноправные и симметричные переменные x_1, x_2, \dots, x_n , неудобны при рассмотрении вопросов, относящихся к системе в целом. Нужно ввести иерархию независимых переменных, выбрав $E=H(X)$ в качестве главного переменного (аналогичного радиусу в полярной системе координат). Тогда оказывается, что все сумматорные функции почти не зависят от огромного числа «угловых» переменных Y , а зависят только от «радиуса» E . Для выделения доминанты из сумматорной функции несущественно, как именно выбраны координаты Y на поверхности уровня H -функции.

$$E = H(X). \quad (5)$$

Употребление буквы E всюду в дальнейшем означает, что рассмотрение ведется в координатах E, Y . Если же выкладки удобнее вести в исходных переменных x_1, x_2, \dots, x_n , то естественно сохранить для E обозначение $H(X)$ и термин H -функция.

п.2. Усреднение как способ вычисления доминанты

Основная мысль состоит, конечно, в том, что доминанта вообще существует*. После того, как вопрос о доминанте поставлен и угадана основная формула разложения сумматорной функции, само отыскание составляющих труда не представляет.

Определяющим свойством доминанты является ее постоянство вдоль поверхностей уровня $H(X) = E$, поэтому она не меняется

* Несомненно, что замысел навеян законом больших чисел. Но когда цыпленок идеи вылутился из яйца догадки, нужно немедленно освободиться от скрупулы вероятностной терминологии, заслоняющей глаза новому существу и путающейся у него в ногах.

при осреднении по этим поверхностям. Напротив, флуктуанту можно задать как раз условием равенства нулю ее среднего значения по каждой поверхности уровня. Покажем, что это наилучший выбор флуктуанты.

Пусть есть два разложения сумматорной функции. Одно разложение

$$A(X) = A(H) + F(X) \quad (6)$$

построено так, что среднее от флуктуанты равно нулю. Второе разложение произвольно

$$A(X) = A'(H) + F(X). \quad (7)$$

Выразим среднее значение квадрата второй флуктуанты через первую:

$$\begin{aligned} \overline{F'(X)^2} &= \overline{[A(H) - A'(H) + F(X)]^2} = \\ &= \overline{[A(H) - A'(H)]^2 + 2[A(H) - A'(H)]F(X) + F^2(X)} = \\ &= [A(H) - A'(H)]^2 + 2[A(H) - A'(H)]\overline{F(X)} + \overline{[F(X)]^2}. \end{aligned}$$

Итак

$$\overline{[F'(X)]^2} = [A(H) - A'(H)]^2 + \overline{[F(X)]^2}. \quad (8)$$

В этой традиционной выкладке, идейно восходящей к тригонометрическим рядам, использованы только два свойства: постоянство любой доминанты вдоль поверхностей уровня и равенство нулю среднего от «правильной» флуктуанты. Полученное равенство доказывает экстремальное свойство правильного разложения.

Среди всех возможных разложений наименьшую квадратичную погрешность обеспечивает доминанта, получаемая осреднением сумматорной функции вдоль поверхностей уровня $H(X) = E$.

Стоит подчеркнуть два обстоятельства. Во-первых, экстремальное свойство является точным, а не асимптотическим. Во-вторых, идея осреднения вторична и вытекает из идеи доминанты.

Физики нередко [4] пишут и о «гипотезе молекулярного* хаоса» как о самостоятельной физической идее. Математик обязан уточнить, что осреднением доказывается вовсе не «доминантность» доминанты, а всего лишь единственность наилучшего разложения.

Это уточнение тем более необходимо, что доказательство сохраняет силу при произвольной функции. Для любой функции ее наилучшую аппроксимацию при помощи функции одной переменной всегда дает метод осреднения. Но это вовсе не означает, что он всегда дает хорошую аппроксимацию. Это особенно ясно

* Как будто хаос становится лучше от того, что он «молекулярный».

в случае среднего, равного нулю, когда разложение оказывается полностью бессодержательным.

В книгах [5] по физике (явно и неявно) ссылаются обычно на практический успех такой процедуры при решении некоторых конкретных задач. Однако для новых ситуаций важно знать причину успеха и предвидеть возможность неудачи. Подробный анализ привел Хинчина к выводу, что существо дела в свойствах тех конкретных функций, которые реально интересны в статистической физике — их сумматорности. Но и в этом случае необходима осторожность. Даже для сумматорных функций аппроксимация становится хорошей только асимптотически, при $n \gg 1$.

п.3. Алгоритм вычисления доминанты

Основная идея выкладки, приводящей к экстремальному свойству, допускает модификацию, удобную для дальнейшего. Умножим обе части равенства (6) на произвольную (бесконечно-дифференцируемую) функцию $f(E)$ и проинтегрируем по всему пространству X .

$$\int_X f(H) \mathcal{A}(X) dX = \int_X f(H) A(H) dX. \quad (9)$$

Второй член справа равен нулю, так как обращается в нуль его интеграл по каждой поверхности уровня $H=E$. Преобразуем полученное равенство, учитывая, что \mathcal{A} — сумматорная функция, а справа под знаком интеграла стоит функция только H .

$$\sum_X \int f(H) a(x_i) dX = \int_{-\infty}^{+\infty} f(E) A(E) \Omega(E) dE. \quad (10)$$

Свойства и асимптотика структурных функций подробно разработаны [2]. Все интегралы слева одинаковы и равны, поэтому достаточно вычислить один из них. Обозначим через Y прямое произведение всех компонент, кроме первой

$$Y = (x_2, \dots, x_n) \quad (11)$$

Тогда X можно представить в виде прямого произведения двух компонент (опуская индекс y малой компоненты)

$$X = (x, Y). \quad (12)$$

Поэтому

$$\begin{aligned} \int_X f(H) a(x) dX &= \iint_{xY} f(H' + h) a(x) dx dY = \\ &= \int_x a(x) \left[\int_{-\infty}^{+\infty} f(E' + h(x)) \Omega'(E') dE' \right] dx. \end{aligned} \quad (13)$$

Сдвинем переменную интегрирования во внутреннем интеграле и переставим порядок интегрирования. Получится следующее равенство, справедливое для любой функции $f(E)$:

$$\int_{-\infty}^{\infty} f(E) \left[\int_x^{\infty} na(x) \Omega'(E-h(x)) dx \right] dE \cong \int_{-\infty}^{+\infty} f(E) A(E) \Omega(E) dE. \quad (14)$$

Из него вытекает равенство для обобщенных функций,

$$A(E) \Omega(E) = n \int_x^{\infty} a(x) \Omega'(E-h(x)) dx, \quad (15)$$

которое служит для вычисления $A(E)$. Следует иметь в виду, что все функции, кроме $a(x)$ и $h(x)$, зависят от числа n и интерес представляет именно их асимптотическое поведение при $n \gg 1$.

Полученное равенство можно записать в виде

$$\frac{A(E)}{n} = \int_x^{\infty} a(x) K_n(E, h(x)) dx, \quad (16)$$

где ядро $K_n(E, h)$ зависит только от скалярных аргументов E, n и h и является решением уравнения

$$K_n(E, h) \Omega_n(E) = \Omega_{n-1}(E-h). \quad (17)$$

Если бы была определена операция деления на обобщенную функцию, то уравнение для ядра имело бы очевидное решение:

$$K_n(E, h) = \frac{\Omega_{n-1}(E-h)}{\Omega_n(E)}. \quad (18)$$

Однако общего определения не существует, поэтому полученная формула имеет точный смысл, если Ω_n является обычной функцией, и асимптотический смысл в случае обобщенной функции. В этом последнем случае смысл формулы (18) следующий — разложение ядра по обратным степеням получается формальным делением разложений числителя и знаменателя.

Любопытно, что уравнение (17) не только более осмысленно, но и более удобно для фактического вычисления асимптотики ядра.

Несложные, но довольно громоздкие* выкладки, основанные

* Очень интересно, что эти выкладки содержат «вычитание бесконечностей». Главные члены в Ω_n и Ω_{n-1} сокращаются, что в терминах «энтропии» (то есть логарифма структурной функции) означает вычитание членов, пропорциональных n .

на обобщенной формуле Больцмана, приводят к следующему результату

$$K_n(E, h) = \frac{e^{-\beta h}}{\varphi(\beta)} \left[1 + O\left(\frac{1}{n}\right) \right], \quad (19)$$

где величина β (интенсивность) связана с E «уравнением состояний»

$$\frac{E}{n} = -\frac{d \ln \varphi}{d \beta}. \quad (20)$$

Главное отличие этого вывода (по существу, конечно, совпадающего с выводом Хинчина в его книге) состоит в установлении прямой связи с обобщенной формулой Больцмана.

Полезно собрать вместе формулы, дающие метод вычисления доминанты. Выкладки проводятся в пересчете на одну компоненту и в главном члене асимптотической теории. Во всех формулах поправки имеют порядок $\frac{1}{n}$, так как все они являются следствием поправки к ядру, имеющей именно такую величину.

По заданной функции $h(x)$ строится функция $\varphi(\beta)$

$$\varphi(\beta) = \int_x e^{-\beta h(x)} dx. \quad (21)$$

после чего решается (относительно интенсивности β) уравнение

$$\frac{E}{n} = \frac{1}{\varphi(\beta)} \int_x h(x) e^{-\beta h(x)} dx = -\frac{d \ln \varphi}{d \beta}. \quad (22)$$

Впрочем, как уже неоднократно говорилось, удобнее считать β независимым переменным, а формулу (22) рассматривать как определение зависимости E от n и β .

Величина доминанты находится по формуле, вполне аналогичной формуле для E .

$$\frac{A}{n} = \frac{1}{\varphi(\beta)} \int_x a(x) e^{-\beta h(x)} dx + O\left(\frac{1}{n}\right) \quad (23)$$

Отличие состоит в том, что формула для E является точной, а формула для A асимптотической. Второе важное отличие заключается в необходимости интегрирования по пространству x , в то время как E можно найти (вторая часть формулы (22)) дифференцированием по переменной β .

Найденная формула для среднего значения допускает очень важное для дальнейшего обобщение. Выкладки, вполне аналогичные проведенным, дают формулу среднего для функции $a(x, y)$, зависящей от двух компонент

$$\mathcal{O}_2(\beta) = \frac{1}{\varphi^2(\beta)} \iint_{xy} a(x,y) e^{-\beta h(x) - \beta h(y)} dx dy + O\left(\frac{1}{n}\right). \quad (24)$$

Обобщение на случай большего числа компонент очевидно. Однако нужно иметь в виду, что формула

$$\mathcal{O}_k(\beta) = \frac{1}{\varphi^k(\beta)} \int_{x_1} \dots \int_{x_k} a(x_1, \dots, x_k) e^{-\beta h(x_1) - \dots - \beta h(x_k)} dx_1, \dots, dx_k + O\left(\frac{1}{n}\right) \quad (25)$$

является асимптотически правильной только в том случае, если число k фиксировано или по крайней мере растет медленнее n , так как формула для ядра имеет поправку порядка $1/n$. Впрочем, для наших ближайших целей вполне достаточно формулы (24), в которой важную роль играет поправка $O(1/n)$.

п.4. Оценка флуктуанты

Дальнейшие выкладки текстуально совпадают с аналогичными выкладками А. Я. Хинчина. Тем не менее их повторение целесообразно хотя бы потому, что книга Хинчина является сейчас библиографической редкостью.

Флуктуанта $F(X)$, так же как основная функция $\mathcal{A}(X)$, относится к классу сумматорных функций:

$$F(X) = f(x_1) + f(x_2) + \dots + f(x_n). \quad (26)$$

Однако слагаемые флуктуанты

$$f(x) = a(x) - \bar{a} \quad (27)$$

имеют среднее, равное нулю, так как вычитаемое является как раз средним значением уменьшаемого

$$\overline{f(x)} = 0. \quad (28)$$

Квадрат флуктуанты содержит n существенно положительных слагаемых и $n^2 - n$ величин, средние от которых необходимо оценить:

$$F^2(X) = \sum_{i=1}^n f^2(x_i) \sum_{i \neq k} f(x_i) f(x_k) \quad (29)$$

Так как средние не зависят от индексов, то для среднего от F^2 получаем простое выражение

$$\overline{F^2(X)} = n \overline{f^2(x)} + (n^2 - n) \overline{f(x)f(y)}. \quad (30)$$

Из формулы (24) вытекает, что в главном члене среднее от произведения равно произведению средних. Эта асимптотическая независимость разных слагаемых и приводит к тому, что главный

член среднего от квадрата флуктуанты оказывается величиной порядка всего лишь h , а не n^2 . Действительно:

$$\begin{aligned} & \frac{1}{\varphi^2(\beta)} \iint_{xy} f(x)f(y)e^{-\beta h(x)-\beta h(y)} dx dy = \\ & = \frac{1}{\varphi(\beta)} \int_x f(x)e^{-\beta h(x)} dx \frac{1}{\varphi(\beta)} \int_y f(y)e^{-\beta h(y)} dy = 0. \end{aligned} \quad (31)$$

Поэтому поправка к ядру в формуле (24) дает оценку

$$\overline{f(x)f(y)} = 0\left(\frac{1}{n}\right), \quad (32)$$

которая приводит к оценке флуктуанты,

$$\overline{F^2(X)} = 0(n). \quad (33)$$

завершающей доказательство репрезентативности доминанты.

Заключение

Наглядно говоря, теорема Пифагора является основной причиной улучшения аппроксимации сумматорной функции ее доминантой.

Конец ряда отрезков, укладываемых вдоль прямой, удаляется со скоростью n . Отрезки, используемые в качестве катетов, дают рост гипотенузы всего лишь со скоростью \sqrt{n} . Поэтому прямолинейная составляющая рано или поздно становится доминирующей.

Литература.

1. Хинчин А. Д. Математические основания статистической механики. М., Гостехиздат, 1943.
2. Молчанов А. М. Теория выравнивания и формула Больцмана. Препринт Ин-та прикл. математики АН СССР, М., 1968.
3. Лебег, Анри. Об измерении величин. М., Учпедгиз, 1938, с. 80.
4. Пригожин И. Введение в термодинамику необратимых процессов. М., ИЛ, 1960.
5. Хилл Т. Статистическая механика. М., ИЛ, 1960.

КИНЕТИКА СЛОЖНЫХ СИСТЕМ

Сложные системы могут и должны изучаться с различных точек зрения. Методический арсенал должен, конечно, включать подходы, выработанные при изучении простых систем, однако он не может ими исчерпываться. Специфика сложных систем неминуемо должна порождать и новые пути исследования.

К сожалению (или к счастью), не существует общего определения сложной системы, позволяющего однозначно отличать сложную систему от простой. Не претендуя не только на полноту, но даже и на бесспорность точки зрения, можно все же считать, что подсистемы устроены проще, нежели полная система, из них состоящая. Этот качественный подход, естественно, порождает количественный, если рассмотреть тройную иерархию вложенных друг в друга систем. Самую широкую (и самую сложную) систему будем называть «внешней средой». Основную изучаемую систему будем называть просто «системой» и предположим, что она состоит из большого числа одинаковых подсистем, которые естественно назвать «элементами». Кратко ситуацию можно, поэтому, охарактеризовать так — изучается поведение во внешней среде системы, состоящей из большого числа одинаковых элементов. Термины «среда», «система», «элемент» означают, главным образом, взаимоотношения этих трех типов систем. Так, например, «элемент» элементарен только с точки зрения «системы», но сам по себе может иметь сколь угодно сложное строение, детали которого несущественны в рамках изучаемой проблемы. «Среда» не обязана быть безбрежной пустыней, малой песчинкой которой является «система». Напротив, они могут даже совпадать. В этом случае удобно, для краткости, говорить о «замкнутой системе».

Описанная ситуация типична для широкого класса сложных систем.

В настоящее время высказываются различные точки зрения на возможность, целесообразность и своевременность попыток создания общей теории сложных систем. Наиболее простой путь состоит, по-видимому, в том, чтобы ставить и решать созревшие задачи. А там уж пусть время рассудит — образуют ли эти сто мышей одного слона. Если окажется, что они относятся не к одной стройной гармоничной общей теории сложных систем —

не беда, при условии, конечно, что они принесут пользу конкретным отраслям естествознания.

Схема — «система из большого числа одинаковых элементов, находящихся во внешней среде» — достаточно абстрактна, чтобы описывать широкий класс явлений. Противоположной опасности — бессодержательности слишком абстрактных построений — можно избежать, уточнив характер взаимодействия среды с системой и элементов друг с другом.

В рамках предлагаемой схемы удастся выяснить обстоятельство, имеющее общеметодологическое значение. Оказывается, что функционирование системы не зависит от природы составляющих ее элементов при достаточном их числе. В простейшей форме этот факт известен уже из термодинамики — переход от модели твердых шариков к громоздкому формализму квантовой механики весьма мало отражается на свойствах газа как целого, по крайней мере при высоких температурах.

Более общий подход позволяет серьезно расширить границы применимости этого общего принципа. Несущественными оказываются не только детали строения элементов, но даже характер их взаимодействия. Более того, жестко детерминированное взаимодействие и чисто стохастическая связь элементов оказываются полностью эквивалентными с точки зрения функционирования системы как целого. Эквивалентность, разумеется, является асимптотической, при большом числе элементов.

Решающим в поведении системы (в противоположность малой существованию «строительного материала») оказывается характер взаимодействия с внешней средой. Так, например, обычная термодинамика сводит взаимодействие к простейшей, скалярной форме — обмену энергией. В этом хорошо изученном случае и природа элементов и взаимодействие их оказываются полностью несущественными.

Если взаимодействие со средой носит более сложный характер, то «опрошение» элементов и взаимодействия между ними уже не будет столь «тотальным». Тем не менее и в общем случае изучаемую систему можно заменить более простой, не нарушая «равновесия» среды. Эта более простая система состоит из элементов, адекватно *отражающих только взаимодействие со средой. Взаимодействие элементов друг с другом оказывается почти всегда несущественным. Это настолько вошло в привычку, что физик, тщательно разбирающий детали взаимодействия атомов, с легкостью повествует о «случайном» столкновении молекул, переходя к статистическому рассмотрению. Нередко говорят да-

* Так, например, когда физик вычисляет дипольный момент молекулы, то он конструирует, в сущности, тот простейший элемент (диполь), которым заменяется в статистической схеме реальная молекула.

же о «гипотезе молекулярного (!) хаоса». По существу же речь идет, конечно, об эквивалентности детерминированной и стохастической схем взаимодействия элементов.

§1. Математическая модель

Простейшая математическая модель возникает при рассмотрении системы большого числа одинаковых уравнений

$$\frac{dx_i}{dt} = \sum_{k \neq i} a(x_i, x_k). \quad (1)$$

Состояние i -ого элемента описывается вектором x_i , взаимодействие любых двух элементов задается вектор-функцией двух переменных $a(x, y)$. Внешняя среда неявно может входить уже в структуру* функции $a(x, y)$, задающей парное взаимодействие. Явным образом внешняя среда определяет постановку задачи. Мы рассматриваем функцию,

$$H(x) = h(x_1) + h(x_2) + \dots + h(x_n), \quad (2)$$

которая характеризует взаимодействие со средой. Очень существенно предположение об аддитивности — суммарность функции H . Основой описания является функция $h(x)$. Аргумент x этой функции задает внутреннее состояние элемента (например, импульс и координату материальной точки), а значение функции определяет вид воздействия элемента на среду (например, совершаемую работу или, точнее, выделяемую мощность).

Производная этой функции

$$\frac{dH}{dt} = \sum_{i \neq k} \sum \frac{dh}{dx_i} a(x_i, x_k). \quad (3)$$

довольно сложным образом зависит от всех переменных x_i , причем вид этой зависимости определяется не только взаимодействием $a(x, y)$, но и свойствами внешней среды, которая в нашей задаче представлена функцией $h(x)$.

Если число переменных невелико, то уравнение (3), являющееся следствием системы (1), не прибавляет ясности проблеме. Его можно, конечно, ввести в систему вместо любого из уравнений, но система проще от этого не становится.

Иное дело системы с большим числом элементов. По мере увеличения « n » уравнение (3) приобретает все большую независимость от остальной системы.

* Диэлектрическая постоянная, например, входящая формально во взаимодействие зарядов, описывает, в сущности, свойства среды.

Это замечательное обстоятельство вытекает из теоремы Хинчина о свойствах сумматорных функций I. Правда, эта теорема непосредственно неприменима к разбираемому случаю, так как правая часть уравнения (3) является функцией бисумматорной

$$B(x) = \sum_i \sum_{i \neq k} b(x_i, x_k). \quad (4)$$

Однако весьма несложное обобщение показывает, что и для случая бисумматорной функции справедливо представление

$$B(X) = B(H) + F(X). \quad (5)$$

В этом представлении главный член, доминанта B , имеющая порядок n^2

$$B(H) = n(n-1)b \quad (6)$$

зависит только от функции H .

Флуктуанта $F(x)$, которая аккумулирует в себе все детали зависимости функции B от огромного числа остальных переменных, играет роль поправки

$$F(x) \sim n^{\frac{3}{2}}. \quad (7)$$

Эта теорема вскрывает механизм обособления одного уравнения

$$\frac{dH}{dt} = B(H) + F(x), \quad (8)$$

и выделения его из системы большого числа уравнений.

Методологически принципиальный шаг состоит в переходе к одному уравнению

$$\frac{d\varepsilon}{dt} = B(\varepsilon), \quad (9)$$

полностью отщепившемуся от всех остальных уравнений.

Исходная система состояла из большого числа равноправных переменных. Уравнение для H создает иерархию в системе, несущественную в малых системах и значительную в системах больших. Однако эти различия носят только количественный характер.

Предельный переход порождает качественно новую ситуацию. Он вынуждает введение нового понятия — системы в целом. Ее состояние описывается новым независимым переменным ε , а кинетика задается автономным* уравнением (9). Соотношение

$$H(x) = \varepsilon \quad (10)$$

* Довольно редкий случай глубокого значения чисто математической терминологии!

устанавливает соответствие исходной системы, в которой функция H зависит от всех переменных, новой системе, в которой величина сама себя опережает. Обычное противопоставление фазовых переменных термодинамическим величинам имеет аналогичный смысл. Предельный переход совершенно такой же природы, как и рассмотренный, более того, являющийся частным случаем рассмотренного, обладает, конечно, аналогичной «творческой силой». Происходит «эмансипация» системы, исчезает зависимость от составляющих ее элементов.

§2. Выделение системы в целом

Наиболее глубоким фактом в рассматриваемой задаче является, конечно, именно асимптотическое «самоопределение» системы.

Конкретный вид правой части уравнения (9) имеет меньшее значение. Тем не менее некоторые особенности строения этого уравнения оказываются интересными методологически.

Вычисление функции $B(\varepsilon)$ начинается с построения характеристической функции («статистической суммы»)

$$\varphi(\beta) = \int_x e^{-\beta h(x)} dx. \quad (11)$$

Всюду в дальнейшем параметр интенсивности* β имеет не меньшее значение, чем сама величина ε , с которой он связан соотношением

$$\varepsilon = -n \frac{d \ln \varphi}{d \beta}. \quad (12)$$

Более подробная запись этого соотношения

$$\frac{\varepsilon}{n} = \int_x h(x) \frac{e^{-\beta h(x)}}{\varphi(\beta)} dx \quad (13)$$

показывает, что β определяется условием равенства двух средних — среднего по частицам и среднего по фазовому пространству одной частицы.

Величина $B(\varepsilon)$ определяется затем осреднением взаимодействия $b(x, y)$ по фазовому пространству пары частиц:

$$B(\varepsilon) = (n-1)nb(\beta) = \frac{n(n-1)}{\varphi^2(\beta)} \int_x \int_y b(x, y) e^{-\beta h(x) - \beta h(y)} dx dy. \quad (14)$$

* В обычной термодинамике величина β соответствует обратной температуре $\beta = \frac{1}{kT}$.

С формально математической точки зрения равенства (9), (11), (12) и (14) полностью определяют систему, и задача решена.

Однако найденный алгоритм допускает иную интерпретацию, приводящую к существенным методологическим выводам.

Чтобы придти к этой интерпретации, подставим в формулу (14) интересующее нас выражение для $b(x, y)$

$$b(x, y) = \frac{dh}{dx} a(x, y) \quad (15)$$

и заменим двойной интеграл повторным, проводя сначала интегрирование по y . Введем специально обозначение для внутреннего интеграла

$$\alpha(x, \beta) = \frac{n-1}{\varphi(\beta)} \int_y a(x, y) e^{-\beta h(y)} dy. \quad (16)$$

Это обозначение позволяет записать полный интеграл в виде

$$B(\varepsilon) = \frac{n}{\varphi(\beta)} \int_x \frac{dh}{dx} \alpha(x, \beta) e^{-\beta h(x)} dx. \quad (17)$$

Полученные формулы подсказывают идею расщепления системы на невзаимодействующие компоненты. Более точно, взаимодействие в этом случае приобретает совершенно иную форму. Вместо парного взаимодействия каждой компоненты с каждой возникает взаимодействие с «эмансипированной» системой в целом, представленной параметром β .

§3. «Расщепление» компонент

Формальное построение выглядит следующим образом. Рассмотрим новую систему, состоящую из частиц, аналогичных исходным, но «движущихся в осредненном поле»:

$$\frac{d\xi_i}{dt} = \alpha(\xi_i, \beta). \quad (18)$$

Следует подчеркнуть, что термины «поле» и «движение» используются только для наглядности. Задача ставится значительно более общим образом и компонентами x_i могут быть, например, производственные мощности предприятия (в экономических моделях) или концентрация химических соединений (в химических моделях). Построенная система незамкнута, в отличие от исходной системы с парным взаимодействием, так как не задан закон изменения β . Отложив временно решение этого вопроса, вычислим сначала производную от той же сумматорной функции H

$$H(\xi_1, \dots, \xi_n) = H(\xi_1) + \dots + H(\xi_n) \quad (19)$$

но уже в силу нашей новой системы

$$\frac{dH}{dt} = \sum_i \frac{dh}{d\xi_i} \frac{d\xi_i}{dt} = \sum_i \frac{dh}{d\xi_i} \alpha(\xi_i, \beta). \quad (20)$$

Эта производная оказывается сумматорной функцией, в то время как производная в силу исходной системы была более сложной — бисумматорной — функцией.

Проведем выделение доминанты. Для этого нужно сначала построить характеристическую функцию

$$\varphi(\gamma) = \int_{\xi} e^{-\gamma h(\xi)} d\xi, \quad (21)$$

которая, конечно, только обозначениями отличается от построенной ранее функции $\varphi(\beta)$. Совершенно аналогично предыдущему происходит выделение одного уравнения

$$\frac{d\varepsilon}{dt} = C(\varepsilon) \quad (22)$$

правая часть которого вычисляется даже проще, чем правая часть уравнения (9), а именно

$$C(\varepsilon) = \frac{n}{\varphi(\gamma)} \int_{\xi} \frac{dh}{d\xi} \alpha(\xi, \beta) e^{-\gamma h(\xi)} d\xi. \quad (23)$$

Сравнение полученной формулы с формулой (17) для $B(\xi)$ показывает, что они совпадают, если $\beta = \gamma$. Это обстоятельство имеет принципиальное значение.

§4. Принцип статистической эквивалентности

Первоначальная постановка задачи состояла в отыскании переменного, уравнение для которого можно было бы, хотя бы асимптотически, интегрировать отдельно, независимо от всей огромной системы. Такой величиной оказалась сумматорная функция H . По причинам, которые, на первый взгляд, представляются формально математическими, более удобной величиной является величина β , связанная с ε соотношением (12).

$$-\frac{\varepsilon}{n} = \frac{d \ln \varphi}{d\beta}. \quad (24)$$

Поэтому уравнение для ε равносильно уравнению для β , получающемуся дифференцированием по t соотношения (24):

$$\frac{d\beta}{dt} = b(\beta). \quad (25)$$

Правая часть этого уравнения вычисляется на основании формул (9) и (17)

$$b(\beta) = - \left(\frac{d^2 \ln \varphi}{d\beta^2} \right)^{-1} \frac{1}{\varphi(\beta)} \int_x \frac{dh}{dx} \alpha(x, \beta) e^{-\beta h(x)} dx. \quad (26)$$

Совершенно аналогичное вычисление для величин ξ приводит к выводу о совпадении уравнений и, следовательно, совпадении параметров β и γ .

Полученный результат можно сформулировать следующим образом.

Асимптотическое поведение сумматорной функции

$$H(x_1, \dots, x_n) = h(x_1) + \dots + h(x_n) \quad (27)$$

в силу системы уравнений

$$\frac{dx}{dt} = \sum_{i \neq k} a(x_i, x_k) \quad (28)$$

с «парным взаимодействием» $a(x, y)$ совпадает с асимптотическим поведением той же функции

$$H(\xi_1, \dots, \xi_n) = h(\xi_1) + \dots + h(\xi_n)$$

в силу системы осредненных уравнений

$$\frac{d\xi_i}{dt} = \alpha(\xi_i, \beta) \quad (29)$$

для не взаимодействующих компонент ξ .

Уравнение, описывающее асимптотическое поведение этой функции

$$H = n\varepsilon,$$

$$\frac{d\varepsilon}{dt} = \frac{1}{\varphi(\beta)} \int_{\xi} \frac{dh}{d\xi} \alpha(\xi, \beta) e^{-\beta h(\xi)} d\xi \quad (30)$$

может быть интерпретировано в силу соотношений

$$\varphi(\beta) = \int_x e^{-\beta h(x)} dx \quad (31)$$

$$\varepsilon = - \frac{d \ln \varphi}{d\beta} \quad (32)$$

как уравнение для параметра интенсивности («температуры») осредненного «поля»

$$\alpha(x, \beta) = \frac{1}{\varphi(\beta)} \int_x a(x, y) e^{-\beta h(x)} dx. \quad (33)$$

Более кратко можно сказать так:

Система с парным взаимодействием статистически эквивалентна системе без взаимодействия, находящейся в «осредненном поле».

Это означает, что система, состоящая из большого числа одинаковых компонент, может быть заменена (с точки зрения внешней среды, воздействие на которую описывается функцией $h(x)$) значительно более простой осредненной системой.

§5. Системы второго порядка, в частности механические

Границы применимости предложенной схемы особенно ясно проявляются при попытке применить ее к механическим системам, или вообще, к системам второго порядка

$$\frac{d^2 y_i}{dt^2} = \sum_{k \neq i} a(y_i, y_k). \quad (34)$$

Запишем эту систему, вводя «импульсы»

$$\frac{dy_i}{dt} = q_i,$$

$$\frac{dq_i}{dt} = \sum_{i \neq k} a(y_i, y_k) \quad (35)$$

в виде системы первого порядка, но, конечно, с вдвое большим числом переменных. Эта система может быть записана в виде

$$\frac{dx_i}{dt} = a(x_i) + \sum_{k \neq i} a(x_i, x_k) \quad (36)$$

весьма напоминающем систему (1), с которой начато изучение. Разница только в члене с «самодействием», которого не было в системе (1).

Система (35) равносильна системе (36), если положить

$$x = \begin{pmatrix} y \\ q \end{pmatrix} \quad (37)$$

$$a(x) = \begin{pmatrix} q \\ 0 \end{pmatrix} \quad (38)$$

$$a(x', x'') = \begin{pmatrix} 0 \\ a(y', y'') \end{pmatrix}. \quad (39)$$

Вполне естественно, казалось бы, изучать систему (36) совершенно аналогично системе (1), выбрав в качестве основной функции

$$h(x) = \frac{q^2}{2} + W(y). \quad (40)$$

Однако вычисление производной от сумматорной функции вскрывает ограниченность разобранного выше подхода.

Действительно

$$\begin{aligned} \frac{dH}{dt} &= \sum_i q_i \frac{dq_i}{dt} + \sum \frac{dW}{dy_i} \frac{dy_i}{dt} = \\ &= \sum_i \sum_{k+i} q_i a(y_i, y_k) + \sum_i \frac{dW}{dy_i} q_i \end{aligned} \quad (41)$$

и без дальнейших вычислений видно, что среднее значение правой части оказывается равным нулю. Это произошло потому, что правая часть нечетна по переменным q , поэтому усреднение только по q уже дает нуль независимо от интегрирования по y .

Это обстоятельство подрывает самые основы предложенного подхода, так как главная идея состояла в том, чтобы оставить в правой части главный член, пренебрегая флуктуантой. Если же «главный» член оказывается равным нулю, то вся схема лишается смысла.

Можно было бы думать, что дело в неудачном выборе функции $h(x)$ и попытаться исправить дело хитроумным подбором этой функции. Несложное рассуждение хотя и не доказывает невозможность «хорошего» выбора, но делает эту невозможность весьма правдоподобной.

В самом деле, $h(x)$ должна быть четной функцией q , иначе разойдется интеграл,

$$\varphi(\beta) = \int_y \int_q e^{-\beta h(y,q)} dy dq \quad (42)$$

определяющий статистическую сумму. Но если функция $h(x)$ четна по q , то ее производная $\frac{dh}{dt}$ **обязательно** нечетна по q .

Действительно, эта производная имеет вид

$$\frac{dh}{dt} = \frac{\partial h}{\partial y} \frac{dy}{dt} + \frac{\partial h}{\partial q} \frac{dq}{dt} = \frac{\partial h}{\partial y} q + \frac{\partial h}{\partial q} \frac{dq}{dt}. \quad (43)$$

Первое слагаемое есть произведение четной функции от q (так как четность не меняется при дифференцировании по «чужому» аргументу) на нечетную — q . Второе слагаемое есть произведение нечетной функции (четность как раз наоборот меняется на противоположную при дифференцировании по «своему» аргументу) на четную функцию (так как $\frac{dq}{dt}$ вообще не зависит от q).

Осреднение нечетной функции с весом $e^{-\beta h}$, четным по q , дает нуль уже при внутреннем интегрировании по q .

§6. Многомерная «температура» — векторная интенсивность

Разумеется, нуль в качестве «главного» члена может получиться и для систем первого порядка. Однако такое происходит случайно и может рассматриваться, поэтому, как вырождение задачи. Нормально, если система не удовлетворяет дополнительным условиям типа равенства, это вырождение не имеет места. Поэтому разобранный выше схема дает решение поставленной задачи.

Иное дело системы второго порядка. Четность по временным производным неминуемо приводит к нечетности по скоростям и вытекающему из этой нечетности тождественному вырождению.

Возникающая трудность носит принципиальный характер. Весьма поучительно, что выход из этой трудности, который подсказывается структурой уравнения для $\frac{dH}{dt}$, имеет значение, далеко выходящее за рамки сравнительно частного вопроса о системах второго порядка.

Этот выход состоит в освобождении от предрассудка скалярной точки зрения. Основная идея состояла не в том, чтобы написать обязательно одно уравнение. Важно, чтобы из системы необозримо громадного числа уравнений выделилось небольшое число автономно меняющихся переменных. Решающее значение имеет сам этот процесс «становления» системы, ее выделения из составляющих ее компонент, приобретение ею самостоятельности.

Очень хорошо, если система оказывается настолько простой, что описывается одним уравнением. Но если одно не выделяется, а можно выделить два уравнения, то это отнюдь не является неудачей. Неудача состояла бы в том, что вообще не удастся проследить процесс «высвобождения» системы. Более того, невозможность выделить одно уравнение и возможность, тем не менее, выделить несколько автономно интегрируемых уравнений следует рассматривать как принципиальное достижение. Такая ситуация показывает, что изучаемая система является достаточно сложной и не может быть принципиально описана в рамках скалярных представлений. Очень характерен способ, которым система «объясняет» о своей сложности недооценившему ее исследователю.

Главный член, детерминированный и однозначный, «тонет» в «поправке» — флуктуанте. Очень часто такая ситуация воспринимается как свидетельство стохастичности системы. Между тем, как мы увидим ниже, это означает всего лишь бессилие исследователя, не сумевшего выделить все существенные переменные. Так, например, физический маятник будет восприниматься как вероятностная система, если наблюдатель лишен почему-либо «дара» воспринимать и измерять скорость и ограничивается измерением только положения. Одно только положение маятника ни в малейшей степени не предсказывает его положения в будущем. Напро-

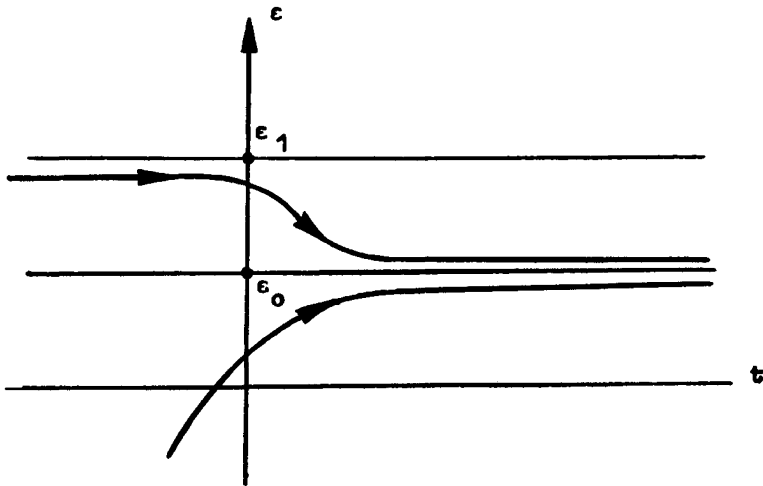


Рис. 1. Кинетика скалярного уравнения. ϵ_0 — устойчивое положение равновесия, ϵ_1 — неустойчивое

тив того, значение координаты и импульса позволяет предсказать как будущие положения, так и будущие скорости.

Возвращаясь к нашей задаче, рассмотрим не одну, а две функции

$$h(x) = \frac{q^2}{2} + W(y) \quad (44)$$

$$q(x) = \frac{dW}{dy} - q$$

и найдем производные соответствующих сумматорных функций в силу уравнений нашей системы

$$\begin{aligned} \frac{dH}{dt} &= \sum_{i \neq k} \sum q_i a(y_i, y_k) + \sum_i \frac{dW}{dy_i} q_i, \\ \frac{dG}{dt} &= \sum \sum \frac{dW}{dy_i} a(y_i, y_k) + \sum_i \frac{d^2W}{dy_i^2} q_i q_i. \end{aligned} \quad (45)$$

Совершенно аналогично скалярному случаю нужно искать средние значения правых частей. Однако в отличие от скалярного случая среднее вычисляется на поверхности, определяемой двумя равенствами

$$\begin{aligned} H &= \epsilon \\ G &= Q. \end{aligned} \quad (46)$$

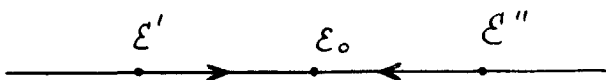


Рис. 2. Область «притяжения» устойчивого стационарного состояния ξ_0

Если бы дело происходило не в пространстве огромного числа измерений, а в обычном трехмерном пространстве, то условие

$$H = \varepsilon \quad (47)$$

определяло бы поверхность (например сферу), а два условия — линию (например, широтный круг). Довольно ясно поэтому, что происходит при переходе к большому числу переменных. Изучение системы становится более точным, так как правые части осредняются по более «тонким» множествам. Мы «не валим в кучу» все широты, «осредняя по всей Земле», а учитываем то обстоятельство, что обстановка на экваторе и вблизи полюса различается. Если же вернуться к скалярному осреднению по поверхности $H = \varepsilon$, то это соответствует замене и полюса, и экватора средними широтами.

Алгоритм вычисления содержит построение характеристической функции, зависящей уже от двух параметров интенсивности β_1 и β_2 ,

$$\varphi(\beta_1, \beta_2) = \int_x e^{-\beta_1 h(x) - \beta_2 y(x)} dx, \quad (48)$$

которую можно записать и в старой форме

$$\varphi(\beta) = \int_x e^{-\beta h(x)} dx, \quad (49)$$

если только иметь в виду векторный характер как величины h , так и сопряженной величины β .

Подробный анализ возникающих соотношений далеко выходит за рамки обсуждаемой темы. Однако одно обстоятельство важно для дальнейшего. Функция H четна по q , а функция G нечетна по q . Правые части уравнения для этих функций обладают противоположной четностью. Поэтому систему, возникающую при осреднении системы (45), можно записать в виде

$$\begin{aligned} \frac{d\varepsilon}{dt} &= Qa(\varepsilon, Q), \text{ или} \\ \frac{dQ}{dt} &= b(\varepsilon, Q), \end{aligned} \quad (50)$$

где функции $a(\varepsilon, Q)$ и $b(\varepsilon, Q)$ являются четными функциями Q .

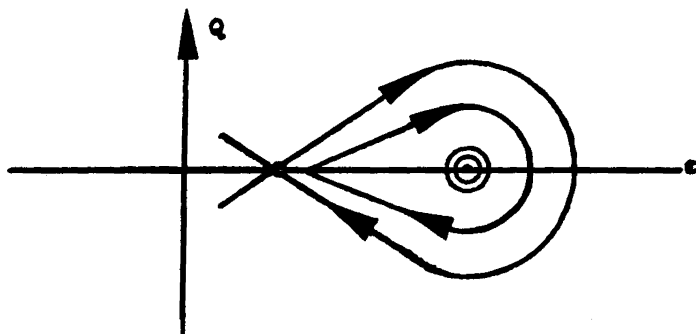


Рис. 3. Строение типичной «петли» сепаратрисы

7. Возникновение колебательной кинетики

Принципиальный вывод, который вытекает из рассмотрения систем второго порядка, состоит в том, что поведение системы в целом далеко не всегда описывается одним уравнением.

Скалярные уравнения

$$\frac{d\varepsilon}{dt} = B(\varepsilon) \quad (51)$$

определяют крайне бедную кинетику. Она целиком определяется стационарными точками этого уравнения

$$B(\varepsilon) = 0 \quad (52)$$

и состоит в асимптотическом выходе на это стационарное состояние. Вся ось разбивается неустойчивыми положениями равновесия на интервалы, внутри каждого из которых находится устойчивое положение равновесия, притягивающее все траектории из этой области.

Этот процесс довольно быстрый и большую часть времени такая система проводит вблизи положения равновесия. Неудивительно поэтому, что термодинамика, например, занимается главным образом уравновесившимися состояниями. Ей больше приличествовало бы поэтому именоваться термостатикой.

Совершенно другой характер имеет кинетика систем, несводимых к скалярным. Проиллюстрируем возникающие здесь особенности на примере системы (50). Нетрудно показать, что интегральные кривые этой системы симметричны относительно оси ε . Важную роль играют, как и в скалярном случае, стационарные точки, особенно расположенные на оси ε ,

$$\begin{aligned} Qa(\varepsilon, Q) &= 0 \\ b(\varepsilon, Q) &= 0 \end{aligned} \quad (53)$$

и взяв в качестве решения первого уравнения $Q=0$, получаем:

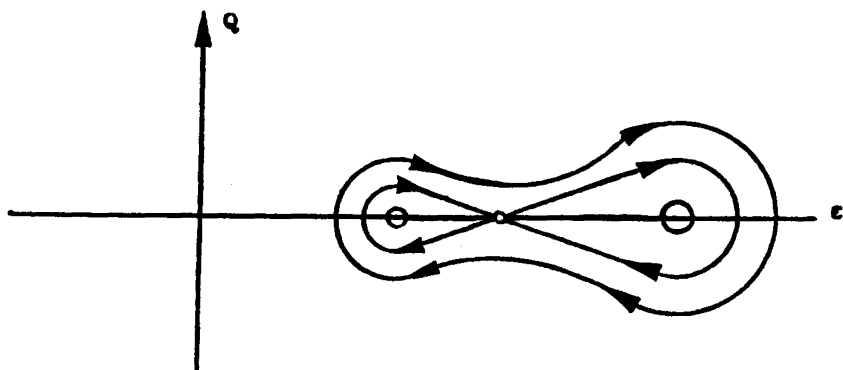


Рис. 4. Колебательный режим, охватывающий «восьмерку», составленную из сепаратрисс

$$Q = 0$$

$$b(\varepsilon, 0) = 0 \quad (54)$$

Несложный анализ показывает, что структура интегральных кривых вблизи оси $Q = 0$ определяется опять-таки неустойчивыми стационарными точками системы. Однако, в этом случае они являются седловыми, и зоной влияния (уже не притяжения) оказываются петли сепаратрисс, выходящих из этих стационарных точек. Вокруг устойчивой, лучше сказать нейтральной, стационарной точки возникает колебательный режим. Его предельная амплитуда задается, конечно, петлей сепаратриссы, окружающей центр. Более сложные колебания могут возникать при возрастании амплитуды на кривых, охватывающих два или более устойчивых центра.

Более подробный анализ, которому не место в этом методологическом очерке, показывает, что уточнение полученной кинетики выявляет медленное движение, приводящее к постепенному затуханию колебаний. Важно только подчеркнуть, что этот анализ требует расширения системы. Введение одной скалярной функции не дает никаких сведений, система воспринимается как чисто стохастическая детерминированность ее поведения «тонет» в шумах.

Уточнение средств изучения понижает уровень «шумов» и выделяет детерминированную колебательную систему, совершающую чисто периодическое движение, не содержащее эволюции.

Следующий шаг, не обсуждавшийся совсем, должен выделить, и здесь об этом можно говорить только в порядке гипотезы, эволюционное движение. Это движение может приводить к затуханию колебаний.

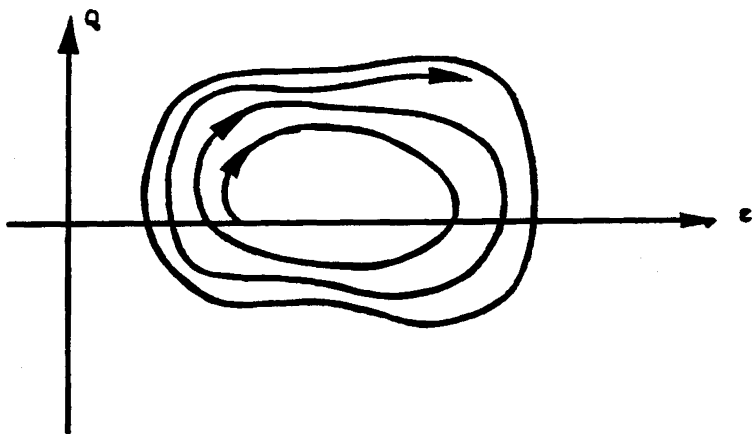


Рис. 5. Возможность возникновения предельного цикла при введении дополнительных переменных

Однако если система достаточно богата по строению и энергии, то эволюция может привести к возникновению устойчивого внутреннего самодвижения. Это наиболее интересная возможность является главным направлением исследований в рассматриваемой области.

Можно думать, что достаточно большие (и достаточно сложные системы) обладают тенденцией к возникновению «внутреннего» движения. Даже если сами компоненты просты, но взаимодействие со средой не сводится к скалярному, возникают условия для появления «самодвижения» системы.

Пример систем второго порядка вскрывает любопытный механизм принудительного появления векторного характера интенсивностей. Эта многокомпонентная, не скалярная природа параметров интенсивности является решающим условием зарождения нетривиальных типов кинетики.

Предлагаемое направление демонстрирует одну из многочисленных отличительных особенностей, которыми могут обладать (да простится мне это слово!) даже «простые» сложные системы.

ТЕРМОДИНАМИКА И ЭВОЛЮЦИЯ

Биология и математика. Развитие биологических исследований в последнее время стимулирует возрастающий интерес к колебательным химическим процессам. Все большее число исследователей приходит к мысли, что именно в периодических химических реакциях следует искать основу сложной организации процессов, происходящих в живых организмах. Как это обычно бывает в бурно развивающихся областях, нет недостатка в крайних высказываниях сторонников и противников такого взгляда. Характерная черта происходящего осмысливания богатого биологического материала — настойчивые попытки математического анализа явлений.

Попробуем разобраться в существе этих попыток. Разумно при этом поставить вопрос значительно шире и обсуждать вообще место математических методов в биологии. Известно, какую громадную, даже решающую роль играет математика в современной физике. Теоретическая физика в значительной мере — математическая наука. Однако уже в химии роль математики далеко не так велика и сходит по существу на нет по мере «приближения» к биологии. До самого последнего времени почти единственным «математическим» методом в биологии была статистическая обработка результатов наблюдений. Биолог использовал математику, так сказать, для «очистки поверхности» биологического факта, а затем «сливал» ее вместе с экспериментальной «грязью». Ему и в голову не приходило, что математика могла бы помочь и на «чистой» стадии работы. Разумеется, чистая математика.

Немного о кибернетике. Внешним образом положение дел резко изменилось в связи с возникновением кибернетики. Энтузиастам казалось, что происходит бурное вторжение физики, техники и вместе с ними математики в святая святых биологии. Крепости сдаются одна за другой, еще один решающий штурм и биологов (а заодно и композиторов) можно будет отменить и посадить всюду вместо них программистов.

Сейчас, когда страсти несколько поулеглись и можно подвести хотя бы предварительные итоги, ситуация предстает в несколько ином освещении. Произошло скорее проникновение биологии в технику, чем вторжение техники в биологию. Техника обогати-

лась идеями и в этом немалая заслуга кибернетики. Однако обратное воздействие, несомненно существующее, значительно меньше. Действие, увы, не равно противодействию. Стоит подчеркнуть, что речь идет об идейном воздействии, влиянии на теоретические представления, а не о создании новых технических средств исследования.

Дарвин и Клаузиус. Для того чтобы понять случившееся, стоит вспомнить немного историю. Грубо говоря, к концу прошлого столетия одновременно были сформулированы основные идеи классической физики и биологии. В физике это было торжество атомистических представлений и тесно связанной с ними статистической картины строения мира. В биологии утвердилась дарвиновская теория эволюции. Тепловая смерть Вселенной на одном знамени и безграничное совершенствование на другом.

Было от чего возникнуть и на долгие годы закрепиться взаимному непониманию. Рано или поздно эти точки зрения должны были столкнуться. Реально они сталкивались уже в сознании наиболее ярких представителей физического мировоззрения. Сам Больцман выдвинул идеи гигантской флуктуации. Его очень быстро «разоблачили» — значительно более «вероятной» была бы неизмеримо меньшая по размерам флуктуация. И в те времена нечего было противопоставить этому доводу. Куда более удивительно, что гипотеза Больцмана по традиции считается несостоятельной и в наше время. А ведь с тех пор мы уже узнали, что пространство — время (вакуум квантовой физики) есть равновесная форма существования материи (или поля, по модной ныне физической терминологии), что вещество (включая излучение) есть неравновесная форма материи и что выражение «громкая флуктуация» означает только, что Вселенная достаточно богата материей, чтобы обеспечить именно такие «размеры» флуктуации. Меньшей это флуктуация просто не может быть, так как слишком малые отклонения от равновесия столь же «невероятны», как и слишком большие.

Не следует думать, что только идейные разногласия мешали взаимному обогащению биологии и физики. Ведь не смущает же в наше время физиков внутренняя противоречивость их теорий, как не мешали в свое время развитию математического анализа всевозможные нелепости (тоже, кстати сказать, связанные с расходимостями), проистекавшие от неразмысляющей веры в силу новых методов. «Идите вперед, уверенность придет позже» — это обычный лозунг романтических периодов развития науки.

Нелинейности в биологии. Идейные разногласия были скорее наиболее явной формой глубокого методологического противоречия между биологией и физикой, противоречия, вытекающего из существа задач, стоявших перед этими науками.

Пропасть между ними далеко не заполнена и в настоящее время, но сейчас можно по крайней мере хотя бы оценить ее размеры.

Одним из самых характерных свойств биологических объектов является громадный диапазон воздействий, в пределах которого система нормально работает. Биологические системы обычно не просто нелинейны, они, если можно так выразиться, экспоненциально нелинейны. Это объясняется по-видимому, жесткой необходимостью. Логарифмическая шкала ответов — единственная возможность охватить все информационно значимые воздействия, сохранив сколько-нибудь приемлемые размеры органов. Системы, не сумевшие развить в себе этого свойства, просто не выдержали в борьбе за существование.

Так обстоит дело в биологии и совсем не так обстояло дело до самого последнего времени в физике — во всяком случае в математическом аппарате физики. Принцип суперпозиции, независимость колебаний, разложение по собственным функциям, теория возмущений — все это различные выражения одного и того же главного обстоятельства. В подавляющей массе в физике изучаются малые отклонения от положения равновесия, теория существенно линейна. Крайне важно добавить, что даже «сам» принцип возрастания энтропии тесно связан, как мы увидим ниже, с аддитивностью энергии. Аддитивность — идея более глубокая, чем линейность, но близкая ей по духу. Так или иначе, обе эти идеи возникли при рассмотрении очень простых систем, у которых организацией можно пренебречь.

Принципиальная недостаточность линейной трактовки стала осознаваться в физике, как всегда, под воздействием острой практической необходимости. Такая необходимость возникла прежде всего в радиотехнике и возникла, заметим, закономерно: сильная нелинейность всегда будет проявляться в большинстве устройств, «сопрягающих» технические системы с биологическими. Именно в радиотехнике всерьез начали развиваться нелинейные методы математического естествознания — достаточно напомнить имена Мандельштама, Андронова, Ван-дер-Поля.

В дальнейшем в течение довольно длительного времени существовала весьма любопытная ситуация, когда вынужденный в практической деятельности отказ от линейных представлений мирно уживался с самыми нелепыми предрассудками (типа «неотвратимости тепловой гибели Вселенной*») в общетеоретических воззрениях.

Энтропия возрастает. Что это значит. Важным этапом осмысления понятия энтропии были работы А. Я. Хинчина, в особенности его книга «Математические основания статистической механики», вышедшая в 1943 г. и ставшая ныне библиографической

* Например, Н. Винер. Кибернетика и общество. ИЛ. 1958, стр. 52.

редкостью*. В этой книге показано, что основные понятия термодинамики — такие, как температура и энтропия, — приобретают точный смысл только для предельных, идеализированных систем, обладающих следующими свойствами.

1. Система состоит из очень большого числа отдельных подсистем (частиц).

2. Система замкнута, т.е. не обменивается энергией с внешним миром.

3. Энергия системы равна сумме энергий отдельных частиц.

4. Система находится в состоянии статистического равновесия.

Отсюда, кстати, видно, что самое название «термодинамика» является дезориентирующим. Следовало бы говорить о «термостатике», так как все утверждения касаются состояний равновесия. Хинчин показывает, как следует определить, исходя из статистических представлений, понятия температуры и энтропии и доказывает следующую теорему.

Температура T равновесного состояния системы, получающейся объединением двух подсистем, заключена между температурами компонент

$$T_1 \leq T \leq T_2.$$

Энтропия S больше (или равна) сумме энтропий

$$S \geq S_1 + S_2,$$

причем знак равенства достигается только при равенстве температур $T_1 = T_2$.

Именно этой теореме нередко придают «кровожадный» характер рассуждениями приблизительно такого типа: при объединении энтропия была $S_1 + S_2$, а после наступления равновесия стала S . Так как $S \geq S_1 + S_2$, то налицо «фатальная деградация» энергии. Однако в таком рассуждении постулировано, что 1) понятие энтропии имеет смысл для неравновесных состояний (состояние в момент объединения) и 2) энтропия и в этом случае равна сумме энтропий.

Критический анализ разнообразных «усиленных» сформулированной теоремы приводит Хинчина к выводу, что все они неявно используют утверждения, равносильные доказываемому (нечто вроде многочисленных «доказательств» пятого постулата Эвклида). Следует особенно подчеркнуть, что в теореме ничего не говорится о том, как скоро достигается состояние равновесия и достигается ли оно вообще. Теорема носит условный характер:

* Настолько, что оригинал считают переводом с английского! См., например, литературу к главе 8 в книге Бай Ши-и «Турбулентные течения жидкостей и газов» (ИЛ, 1962).

«если положение равновесия достигнуто, то будут иметь место такие-то соотношения».

Статистическое равновесие и эргодичность. Немалая заслуга Хинчина состоит именно в четком разделении двух совершенно разных вопросов: о достижении положения равновесия и о вычислении характеристик равновесного состояния.

Второй вопрос значительно легче и допускает полное решение разными способами. Сам Хинчин, в частности, предложил весьма изящный аппарат — теорию сумматорных функций, — который строится в полной аналогии с математическим аппаратом теории вероятностей. Есть основания думать, что небольшое усовершенствование этого аппарата позволит изложить с единой точки зрения как частные случаи одной общей теории такие разнообразные, казалось бы, вопросы, как теория информации, классическая и обе квантовые статистики, «таландическая» термодинамика, термодинамика малых систем и термодинамика неравновесных процессов. При всей разнородности этих задач есть нечто принципиально общее, что их роднит — закономерности массовых явлений в значительной степени определяются самим фактом массовости и в меньшей мере зависят от частных свойств отдельных частиц.

Вернемся, однако, к обсуждению основного вопроса о достижении равновесия. Состояние статистического равновесия, о котором всюду идет речь, качественно отличается от привычных состояний равновесия простых механических или физических систем. Оно не означает отсутствия движения или стремления к какому-то предельному положению. Это — состояние динамического подвижного равновесия системы в целом. Настоячивые попытки разобраться в этом трудном вопросе имеют длинную историю и породили богатейшую литературу. Первый шаг был сделан Больцманом, который предположил (эргодическая гипотеза), что точка в пространстве огромного числа измерений (фазовом пространстве), изображающая мгновенное состояние системы, последовательно проходит через все точки фазового объема, совместимые с заданными значениями внешних параметров и энергии. Довольно быстро обнаружилось, что это математически невозможно, и следующее уточнение эргодической гипотезы состояло в том, что траектория проходит сколь угодно близко к любой точке — траектория всюду плотна.

Если бы это утверждение было доказано, то стало бы понятным главное свойство состояний статистического равновесия — постоянство во времени термодинамических величин. Математически термодинамическая величина есть среднее по времени вдоль траектории системы от соответствующей фазовой функции. Например, давление есть средняя величина суммы импульсов. Так вот, если траектория всюду *равномерно* плотна, то можно показать, что среднее по времени от любой фазовой функции равно среднему по пространству. Термодинамические функции

превращаются в этом случае в среднее по пространству и становятся понятным, что они не зависят от траектории. Это прекрасное и хорошо известное рассуждение. Беда только в том, что никому еще не удалось доказать эргодическую гипотезу в сколь угодно общем случае.

Роль атомистичности. Проблема была сдвинута с мертвой точки опять-таки в работах Хинчина, причем сама постановка задачи претерпела существенные изменения. Сторонники эргодической гипотезы возлагают все надежды на весьма хитрое поведение фазовой траектории. Хинчин видит главную причину в свойствах термодинамических величин — в том, что порождающие их фазовые функции сами почти постоянны. Это замечательная постановка вопроса. Отказ от привычных представлений — обычно самое трудное в науке. Нужна немалая смелость, чтобы решиться посягнуть на традиции. Вряд ли случайно, что идеи Хинчина до сих пор мало известны, объясняется это не только тиражом его книги.

Какие же свойства термодинамических величин позволяют столь радикально упростить вопрос? Этих свойств два. Первое состоит в том, что порождающие фазовые функции алгебраически устроены весьма специальным образом. Они могут сколь угодно сложно зависеть от координат одной частицы, но их зависимость от разных частиц очень проста — это суммы по всем частицам. Второе важное свойство состоит в том, что это суммы громадного числа слагаемых, так как существо задачи — изучение асимптотических, предельных свойств систем, состоящих из большого числа одинаковых объектов.

Центр тяжести задачи переносится поэтому на изучение свойств функций, и Хинчин доказывает, что сумматорные функции обладают следующим замечательным свойством.

Рассмотрим одну какую-нибудь сумматорную функцию (например, энергию) и предположим, что поверхности ее уровня ограничивают области с конечным фазовым объемом. Рассмотрим далее любую другую сумматорную функцию. Оказывается, что на каждой поверхности уровня первой функции вторая функция тоже почти постоянна. Более точно: она может — и даже обязательно будет — сильно отличаться от постоянной, но на множестве малой (порядка $\frac{1}{\sqrt{n}}$) меры, а на всем остальном множестве ее отличие от постоянной мало (тоже порядка $\frac{1}{\sqrt{n}}$).

Температур может быть несколько. Не следует, конечно, думать, что «эргодическое» направление мысли совсем не имеет отношения к делу. Имеет, и самое прямое. Ведь может случиться, что траектория лежит как раз на исключительном множестве. Это наверняка произойдет, например, в случае, когда система уравнений движения имеет, кроме энергии, еще какой-нибудь интеграл движения. Тогда траектория целиком лежит на пересечении

поверхности постоянной энергии и поверхности уровня этого нового интеграла. Пересечение образует многообразие меньшего числа измерений и является, следовательно, множеством меры нуль.

Вычисления и в этом случае основываются все на тех же соображениях теории сумматорных функций. Нужно только взять в качестве основных не одну, а две сумматорных функции. Все остальные сумматорные функции становятся в этом случае функциями двух переменных, а не одного (температуры и, например, химического потенциала, а не только температуры, как в классической термодинамике).

Может показаться, что мы вернулись к тому, с чего начали: нужно найти множество, на котором траектория всюду плотна, и по нему осреднять. Это не совсем так. Теория сумматорных функций, во-первых, показывает, что функции и до усреднения почти постоянны, а во-вторых, существенно уменьшают требования к траектории. Не обязательно, чтобы траектория была всюду плотной. Важно, чтобы она не «застряла» на исключительном множестве. Раньше нужно было доказывать, что траектория проходит через любую область и притом находится там время, пропорциональное объему, а теперь достаточно показать, что замыкание траектории не является множеством меры нуль. Но это максимально ослабленное утверждение уже действительно необходимо доказывать. Речь идет не о формально безукоризненном доказательстве (хотя и это совсем неплохо), а об уверенности, что нет дополнительных интегралов, наличие которых существенно меняет результат. Нельзя забывать поучительную историю теории теплоемкости металлов, стоившую физикам немалых мучений и путаницы. Положение вполне разъяснилось только после создания квантовых статистик (симметрической и антисимметрической), причем оказалось, что классическая статистика соответствует усреднению по всему фазовому (бесконечномерному!) пространству, а новые статистики получаются осреднением по многообразиям меры нуль (многообразия симметрических или антисимметрических функций).

«Подавляющее» большинство и «ничтожное» меньшинство. В заключение этого затянувшегося экскурса в статистические проблемы физики необходимо обсудить еще один вопрос, имеющий прямое отношение к разбираемой теме.

До сих пор обсуждались свойства и поведение системы в целом. Те же математические результаты допускают двойственное истолкование, если интересоваться поведением отдельных частиц.

Примером такого истолкования является распределение Максвелла для скоростей молекул. Чисто описательно статистика приводит к такой картине: подавляющее большинство частиц имеет характеристики, заключающиеся внутри некоторого интервала вокруг средних значений, и только ничтожное меньшинство

(порядка \sqrt{n} из общего числа n) может, и обязательно будет, существенно отклоняться от нормы. Очень важно подчеркнуть, что хотя взаимодействие между частицами в простейшей статистической схеме формально не учитывается, оно является единственным фактором, могущим обеспечить необходимое перемешивание. Более того, это взаимодействие обязательно должно носить в главном упругий характер — характер обмена механической энергией. Последнее требование, которое очень трудно сколько-нибудь точно сформулировать для общих (немеханических) систем, имеет крайне важное значение. Оно сразу показывает, что термодинамические понятия имеют смысл только на временах, малых по сравнению с временем неупругих взаимодействий.

Об этом ни в коем случае нельзя забывать, иначе наш слух может быть неприятно поражен хрустом косточек «подавляющего» меньшинства, исчезающего в прожорливой пасти «ничтожного» меньшинства. Не исключено, например, что нечто весьма похожее* происходило при эволюции газопылевого облака, из которого, как полагают, возникла наша планетная система.

Термодинамика «вымирает». Перейдем к некоторым важным выводам о месте и соотношении энтропийных и эволюционных представлений. Следует, однако, сразу же оговорить, что если до сих пор речь шла в основном о результатах, то в дальнейшем придется иметь дело с предположениями чаще, чем этого хотелось бы. Однако таково реальное положение дел в теории, и общий методологический анализ необходим для оценки уже сделанного и правильной постановки задач на будущее.

В истории науки нередко случалось так, что представления, называвшиеся взаимоисключающими и вызывавшие ожесточенную борьбу мнений, оказывались через некоторое время только крайними, полярными, предельными случаями реальной картины явления. Нильс Бор возвел даже в ранг закона природы или по крайней мере закона познания принцип дополненности. Можно сколько угодно спорить по поводу значения этого принципа, но самый факт его появления симптоматичен. Он показывает, насколько важны в науке пары полярных понятий типа волна — частица или причинность — статистичность.

Приведем соображения в пользу того, что энтропийные и эволюционные представления также являются подобной полярной парой. В настоящее время нет возможности сформулировать такое утверждение в сколько-нибудь точных терминах — в этом, собственно, и состоит одна из ближайших задач общей теории. Однако разбор достаточно разнообразных примеров серьезно облегчает правильную постановку задачи.

* Налипание частиц на небольшое число «зародышей» будущих планет и спутников.

В качестве одного из таких примеров рассмотрим упоминавшуюся выше эволюцию газопылевого облака. На ранних стадиях оно состоит из громадного числа отдельных частиц и к нему хорошо применимы статистические и термодинамические представления. Нужно только рассматривать объемы настолько большие, чтобы они содержали большое число отдельных частиц (не забудем, что $\frac{1}{\sqrt{n}}$ — мера применимости термодинамических представлений!), но вместе с тем настолько малые, чтобы все частицы находились в однородных условиях. Так как частицы взаимодействуют часто, а слипаются сравнительно редко, то налицо факты, обеспечивающие эргодичность. Однако слипание пусть медленно, но идет. Система проходит ряд состояний, каждое из которых статистически равновесно, число частиц уменьшается, неоднородности растут, организация усложняется. Заметим, что для наших целей несущественно, произошла ли конкретно наша планетная система именно таким образом. Речь идет о мысленном эксперименте. Рассмотрим заключительную стадию эволюции. Образовались большие планеты и вращаются с точностью часового механизма, более того, являются прообразом часового механизма. Куда же девались статистика и термодинамика? Несмотря на всю наивность примера, в нем содержится главное — применимость статистических представлений уменьшается в процессе эволюции, так как уменьшается число частиц.

Мы приходим к парадоксальному выводу — в этом случае термодинамика применима к ранним стадиям эволюции, в то время как стадия эволюционной зрелости описывается механикой, да к тому же еще небесной.

Область применимости термодинамики. Небольшое размышление показывает, что это не так уж удивительно. Надо вспомнить, что вопрос о взаимодействиях лежит, в сущности, вне рамок термодинамики.

Поставим задачу следующим образом. Пусть известны уравнения, точно описывающие поведение какой-то системы. Тогда в этих уравнениях главными будут члены, дающие поведение каждой частицы во внешнем поле сил. Взаимодействие мало по сравнению с главными членами, причем основную роль в нем играют члены, описывающие упругое парное взаимодействие. Есть, кроме того, сверхмалые члены, которые описывают процессы диссипации и еще более сложные процессы неупругого столкновения — например химические, которые могут приводить к исчезновению старых и образованию новых частиц.

Соответственно трем типам членов возникают три масштаба времени: малые времена, на которых отдельные частицы ведут себя как свободные; большие времена, на которых происходит несколько соударений, и сверхбольшие, эволюционные времена, когда начинают играть роль сверхмалые, неупругие члены. Для ориентации стоит упомянуть, что при обычных химических реакци-

ях в газовой фазе на $10^{10} - 10^{11}$ упругих столкновений приходится одно эффективное, «химическое» столкновение.

Как будет вести себя такая система?

Пропустим неинтересную стадию малых времен, когда ничего не происходит. Затем идут масштабы времени, на которых еще можно пренебречь неупругими членами, но решающую роль играют упругие столкновения. Вот именно эти времена, на которых устанавливается статическое равновесие, и есть подлинное царство статистики, которая принципиально исходит из того, что частицы все время сохраняют индивидуальность.

Необходимо подчеркнуть, что все «энтропийные» рассуждения неявно используют незаконную по существу экстраполяцию на эволюционные времена закономерностей, несомненно, имеющих место на временах неизмеримо меньших — временах термодинамической релаксации.

Между тем правильный вывод был бы таким: обычно эволюция — более медленный процесс, она происходит на фоне установившегося и медленно сдвигающегося термодинамического равновесия.

О математической модели. Проведенный выше методологический анализ оправдывает, по-видимому, постановку вопроса о создании более общей теории, теории, включающей и энтропийные и эволюционные принципы в качестве предельных, описывающих крайние ситуации, в которых может находиться эволюционирующая система.

Попытаемся наметить те основные черты картины развития, которые характерны для любой эволюционирующей системы — биологической, химической или физической — и которые именно в силу своей всеобщности могут служить основой для построения математической модели.

Так как это должна быть эволюционная теория, т.е. теория изменений присходящих во времени, то подходящим математическим аппаратом будет, вероятно, теория дифференциальных уравнений. При этом можно ограничиться по крайней мере на первых порах обыкновенными дифференциальными уравнениями $\frac{dx}{dt} = A(x)$ представляющими достаточное богатство возможностей.

Возможные конкурирующие схемы с дискретным временем или уравнения в частных производных можно рассматривать как предельные случаи схемы обыкновенных уравнений. Они несомненно могут быть более удобны в различных частных задачах, но

для общих соображений вполне достаточно* теории обыкновенных уравнений, значительно более развитой в алгоритмическом аспекте.

Уравнения должны быть автономными, т.е. правая часть не должна содержать явно времени t . Это означает, что состояние системы в каждый следующий момент $[x(t+\Delta)]$ определяется только ее состоянием $[x(t)]$ в данный момент и не зависит от состояния внешнего мира.

Система является, следовательно, информационно замкнутой. Во всех других отношениях (например, энергетически) система вполне может быть открытой, но она сама определяет, что и когда она будет поглощать из внешнего мира или испускать в этот мир.

Следующее важное свойство — атомистичность строения. Оно имеет два аспекта. Во-первых, число частиц (индивидуумов) очень велико. Во-вторых, частицы в основном независимы и индивидуальны.

Первый аспект означает в математической модели, что вектор x имеет громадное число измерений (например, $6n$ для газа, где $n \sim 10^{19}$, если речь идет о кубическом сантиметре).

Индивидуальность частицы выражается в математической модели специальным видом правых частей уравнений. Уравнения содержат малый параметр ε таким образом, что при $\varepsilon = 0$ система распадается на независимые уравнения, каждое из которых описывает поведение отдельной частицы. Система уравнений приобретает поэтому следующий вид.

$$\frac{dx_i}{dt} = f_i(x_i) + \varepsilon F_i(x_1, \dots, x_n).$$

Каждое переменное x_i , описывающее состояние i -й частицы, само является, конечно, вектором; возможно даже с большим числом компонент. Разумно предположить, что в систему входят частицы двух-трех видов. В частности, это означает, что среди функций f_i есть две-три существенно различных (даже зависящих от разного числа переменных), но в пределах каждого вида эти функции отличаются только номером аргумента i .

Дальнейшее уточнение касается характера взаимодействия, т.е. свойств функций F_i . Начинать нужно во всяком случае с рассмотрения только парного взаимодействия

$$F_i(x_1, \dots, x_n) = \sum_{k=1}^n F(x_i, x_k).$$

* К тому же у математиков достаточно хорошо работает «индустрия обобщений», переносящая результаты, полученные, скажем, в обыкновенных уравнениях, на уравнения в частных производных или разностные уравнения.

Есть основания думать, что все существенные явления, по крайней мере качественно, уже содержатся в модели парного взаимодействия. В результате система приобретает следующий вид:

$$\frac{dx_i}{dt} = f(x_i) + \varepsilon \sum_{k=1}^n F(x_i, x_k).$$

Можно было бы привести соображения в пользу дальнейшего упрощения модели. Однако здесь не место для столь подробной детализации.

Шаг эволюции. Как может выглядеть теория подобной модели? Заранее попросив извинения у математически настроенного читателя за головокружительные логические скачки, а у биологически настроенного читателя — за непривычную для него бедность получающейся схемы, попытаемся все же высказать ряд утверждений, составляющих основную канву возможной теории.

Эволюция подобной системы состоит, вообще говоря, из трех качественно различных этапов.

Первый этап — установление статистического равновесия, этап термодинамики в широком смысле слова.

Второй этап — медленная эволюция, проходящая ряд термодинамически квазиравновесных состояний. Эволюция завершается созданием условий для появления нового вида частиц-индивидуумов, качественно отличных от исходных, что приводит к третьему этапу. На третьем этапе происходит «взрывное размножение» новых частиц, когда их число экспоненциально быстро (разумеется, в эволюционной шкале времени) растет. Альтернативная возможность — потеря устойчивости каким-нибудь «видом» и быстрое его исчезновение. Этот этап завершается исчерпанием ресурсов для создания новых частиц. В случае «вымирания вида» не обязательно, конечно, полное исчезновение его. Вполне возможен переход на другой уровень, соответствующий новым условиям.

В результате возникает система, вполне аналогичная той, с которой было начато рассмотрение, только исчезли, частично или полностью, старые виды и созданы новые, нередко с другим масштабом собственного времени (времени «жизни» частицы-индивидуума).

Заметим, что новая система термодинамически неравновесна, так как теперь частицы в ней другие и термодинамику нужно понимать в новом смысле слова.* Ясно, что если предлагаемая схема соответствует действительности, то чисто эволюционная

* Вспомним закон парциальных давлений в обычной термодинамике, который означает, что каждый газ распространяется в пространстве, заполненное другими газами, так как если бы оно было пусто. Новый вид ведет себя, вероятно, аналогично.

и чисто термодинамическая схема являются идеализированными предельными случаями этой более общей схемы эволюционного цикла. Лучше, вероятно, говорить об «эволюционном шаге», так как в системе произошли качественные изменения, а слово «шаг» достаточно гибко и может означать не только топтание на месте, но и движение вперед. Или назад.

Расселение, эволюция, взрыв. Разберем более подробно отдельные этапы. Все три этапа вполне равноправны, и любой из них можно считать началом цикла. Начнем с первого этапа, который получил такой номер просто из уважения к истории науки.

Для предложенных схем можно построить теорию статистического равновесия по аналогии с обычной термодинамикой. Однако в самом главном пункте будет существенное отличие. Равновесное состояние характеризуется не одним «внутренним» параметром, как это будет в случае, когда нас интересует только одна энергия, а несколькими числами. Температура приобретает векторный характер. Ничего особенно нового в этом, конечно, нет. Уже в химических системах, кроме температуры, надо задавать химический потенциал, играющий роль, вполне аналогичную температуре.

Построение такой теории для двумерной температуры было намечено выше в самых общих чертах.

Следующий пункт — самый трудный во всей программе. Это — выяснение условий, при которых взаимодействие обеспечивает выход на статистическое равновесие. Слабым утешением является тот факт, что и в обычной статистической механике дело обстоит не лучше, а ведь она существует уже почти сто лет. Ясно во всяком случае одно — далеко не любой тип взаимодействия приводит к выходу на статистическое равновесие. Не следует думать, что это чисто негативный факт. Напротив, возможно, что именно таков один из механизмов «отсечения» боковых ветвей эволюционного дерева.

Дальше идет стадия медленной эволюции статистически равновесного состояния. Здесь можно было бы привести поистине огромное количество разнообразнейших примеров*. Стоит заметить, что известная теория необратимых процессов имеет дело с очень простым, идеализированным вариантом теории эволюции равновесного состояния, когда можно ограничиться линейным приближением к задаче.

* Внешние слои звезды находятся в состоянии весьма сложного термодинамического равновесия. По мере того как пояс сгорания водорода в гелий (для водородных звезд) приближается к поверхности, оболочка звезды становится, по-видимому, все менее устойчивой. Дело может кончиться взрывом, переходящим в колебания или колебаниями, раскачивающимися до взрыва — существуют обе точки зрения.

Наиболее, по-видимому, интересен для биологических приложений момент возникновения условий для образования нового вида. Ясно, что тут может быть много разных случаев. Чаще всего это будет, вероятно, переход через критическую точку, когда возникают условия для экспоненциального роста* новообразований. Сами новые частицы могли возникать на всей предыдущей стадии эволюции. Но раньше они распадались, не успев размножиться, а теперь «вид» закрепился.

О резонансах. Самый интересный вопрос — как моделировать математически появление новой частицы. Можно указать по крайней мере на один возможный путь моделирования. Будем считать новой частицей две (или более) старые частицы, образовавшие коалицию. Признаком коалиции может служить слипание частиц, когда они все дальнейшее время имеют близкие пространственные координаты. Это не очень интересный случай.

Значительно более интересен, по-видимому, другой возможный способ создания коалиции — резонанс колебательных систем. Современная физика элементарных частиц считает это настолько важным путем возникновения новых частиц, что придумано даже новое слово (не очень удачное, на наш взгляд) — «резонон» — в качестве их родового обозначения**. Можно было бы привести еще аргументы в пользу разумности предложения считать новой частицей резонансную пару старых. Но «кто доказывает слишком много, не доказывает ничего». Ограничимся поэтому указанием на ободряющий пример линейных теорий. Если уже такие простые предположения позволили систематизировать огромный материал, то от теории, существенно учитывающей нелинейность, можно ожидать не меньшего.

Дарвинизм и точные науки. Резюмируем основные идеи.

Первое соображение. Современный математический аппарат, созданный в тесном контакте и под прямым влиянием классической физики, нуждается в серьезной модернизации уже при переходе к современной физике и совсем плохо приспособлен к биологической проблематике. Поэтому взаимодействие с биологией, которое неизбежно будет состоять в попытках «прилагать» математику к биологии, должно обогащать прежде всего математику. Уроки из неминуемых неудач и полуудач должна извлекать именно математика. «Сначала руки учат голову, а затем поумневшая голова учит руки».

Второе соображение есть по существу конкретизация первого. Дарвиновская теория эволюции должна занять подобающее ей

* Типичные примеры — цепные реакции и многие инфекционные заболевания.

** Совсем замечательный пример доставляет теория сверхпроводимости, где коалиция — пары электронов с противоположными импульсами и спинами.

место в точном естествознании. Сейчас накоплено достаточно фактов, свидетельствующих, что основные принципы дарвинизма применимы к любым эволюционирующим системам от элементарных частиц до галактик.* К сожалению, «применение» носит скорее иллюстративный характер: после того, как очередная теория создана частными методами, можно с удовлетворением убедиться, что все весьма похоже на биологическую эволюцию. Между тем в дарвинизме заложена огромная творческая сила — идея организующей роли больших масштабов времени. Для выявления всеобщей значимости этой глубокой мысли необходима прежде всего математическая формализация теории, освобождение идей от их слишком частного конкретного выражения.

Сформулированная выше программа является одной из многочисленных попыток такого обобщения. Ее главной особенностью, позволяющей надеяться на продвижение, размеры которого можно, разумеется, оценить только после реализации программы, является упрощение и сужение задачи — выделение эволюционного цикла. Надежды основаны на том, что получающаяся задача уже поддается, вероятно, формализации и все еще содержит, как можно рассчитывать, важнейшие свойства эволюционного процесса.

Понятие эволюционного цикла явно формулирует идею, что любой эволюционный процесс при всей его сложности может быть разбит на значительно более простые однотипные циклы, последовательно сменяющие друг друга. Если угодно, это идея дискретности, атомистичности эволюции. При таком подходе особенно подчеркнута роль понятия «вид» в эволюции. Цикл начинается «расселением» вида, возникшего в предыдущем цикле, продолжается стадией медленной эволюции и заканчивается возникновением** нового вида или гибелью одного из старых.

О книге Гудвина. Среди многочисленных работ, появляющихся на стыке биологии, физики и математики, книга Гудвина «Временная организация клетки», несомненно, предстает собой яркое явление. Немало привлекательных особенностей характеризует эту работу. Быстрый, оперативный отклик на злободневные события в биохимии, широкая осведомленность в разнообразных областях биологии, самобытность теоретического подхода, крити-

* Например, тройные звезды, у которых все три компонента близки, «вымирают» из-за сильного взаимодействия в (эволюционно) короткий срок. Происходят столкновение и взрыв. «Выживают» системы с двумя близкими и одной удаленной звездой.

** Не нужно понимать видообразование слишком буквально. Пусть, например, внутри клетки были как геометрически закрепленные, так и «плавающие» ферменты. Тогда эволюционный цикл может заключаться в закреплении одного из плавающих ферментов вблизи своего ранее закрепленного предшественника. Метаболизм ускорится, что клетке выгодно.

ческое отношение к модным предрассудкам, подкупающая общность развиваемых положений, явное владение математическими и физическими идеями, спокойный деловой тон — все это делает книгу обаятельной для довольно многочисленного уже сейчас (и он становится все более широким) круга инженеров, физиков, химиков, математиков, интересующихся биологией и желающих принять посильное участие в работе. Идеи, развиваемые в книге, могут стать весьма популярными.

Наша цель состоит как раз в том, чтобы максимально противодействовать этой популярности.

Две главные идеи пронизывают книгу Гудвина: первая — колебательные процессы составляют суть биохимии; вторая — клетка есть набор почти независимых биохимических осцилляторов.

На наш взгляд, эти две идеи глубоко враждебны друг другу. Насколько правильна и плодотворна первая идея, настолько вредна и уводит исследование назад к старым представлениям вторая. Можно высказать догадку, основания для которой дают некоторые высказывания Гудвина об идее обратной связи, что психологически такой возврат вызван реакцией на некоторые крайности «кибернетически настроенных» исследователей. В результате в книге Гудвина происходит, как нам кажется, столкновение самых передовых идей биохимии с довольно-таки обветшалым математическим аппаратом. Перейдем к обсуждению вопроса по существу.

Центральная идея рассматриваемой книги — введение термодинамики, описывающей поведение системы независимых химических осцилляторов. Сама по себе эта тема весьма интересна, хотя подход автора слишком узок. Так как он отвлекается от характера взаимодействия между химическими осцилляторами, то нельзя утверждать ни того, что равновесие наступит, ни того, что наступившее равновесие будет описываться только одним внутренним параметром — таландической температурой. Создается впечатление, что автор просто не подозревает о возможности разных «термодинамик», соответствующих разным типам взаимодействия (при одном и том же исходном материале — химических осцилляторах).

Об этом можно было бы и не говорить (изучение всегда, разумеется, надо начинать с простейшей возможности), если бы автор правильно указал наиболее вероятную область применения своей теории.

Где, в самом деле, и когда можно ожидать хорошего согласия с реальностью теории, основанной на пренебрежении взаимодействием, на идее независимых химических осцилляторов.

По-видимому, там, где ферментов немного, а поле их деятельности велико, так что они почти не мешают друг другу и уж во всяком случае не дожидаются результатов деятельности друг друга.

Кажется очень правдоподобным, что такие условия можно встретить в «первозданном бульоне», когда только-только (скажем, первый миллиард лет биологической эволюции) начали появляться большие молекулы. Это — эпоха «первоначального накопления».

Однако Гудвин думает, что его теория относится к временной организации клетки. Это крайне неправдоподобно. Клетка — результат громадной эволюции, организация в ней находится на высших ступенях развития, все реакции идут крайне специфичным образом, последовательные стадии нередко геометрически организованы.

Не нужно понимать эти высказывания так, будто термодинамика химических осцилляторов совсем не имеет отношения к клетке. Здесь может быть уместна такая аналогия: термодинамика дает полное решение для газа, некоторые ориентировочные суждения для больших ядер (гидродинамическая модель ядра $n \sim 200$) и совсем неприменима для планетной системы, но она прекрасно объясняла многие свойства стадий эволюционного процесса, приведшего к возникновению планетной системы. Возможно «таландическая» термодинамика могла бы играть такую же роль в истории возникновения клетки. Но это уже область чистых домыслов.

В заключение необходимо, как нам кажется, с полной ясностью сказать следующее. В настоящее время нет сколько-нибудь реальной возможности создания математической модели клетки. Выбирать приходится пока между двумя крайними подходами. Один уподобляет клетку механизму, другой — газу. При всей грубости и жесткости первого подхода он неизмеримо ближе к истине чем второй, так как исходит из главного — высокой степени эволюционной зрелости такого замечательного и сложно организованного биологического объекта, как клетка.

МАКРОДИНАМИКА

«Проклятие размерности» — не часто встречается в науке такой всплеск эмоций. Он выразительно характеризует крушение экстремистских иллюзий имитационного моделирования (simulation modelling). Безоглядный оптимизм «прорыва» в экологию (имевший, впрочем, основу в стремительном росте вычислительной техники) сменился крайним пессимизмом. Тем не менее освоение ЭВМ вызвало весьма серьезные сдвиги, правда в неожиданных направлениях.

Более подробного обсуждения заслуживают два аспекта этих сдвигов: во-первых, фазовые переходы, а во-вторых, детерминизм и стохастика.

1. Фазовые переходы

Пока главным (и почти единственным) примером теории больших систем была кинетическая теория газов, простая схема фазовых переходов первого и второго рода казалась достаточной. Однако по мере изучения новых объектов (нейронных сетей в том числе) становилось ясным обилие и разнообразие стационарных режимов и переходных процессов между ними.

Параллельно шло развитие теории нелинейных колебаний. Углубленное изучение бифуркационных явлений привело к убеждению о глубокой аналогии между бифуркациями и фазовыми переходами.

Ныне гипотеза о совпадении этих понятий подкреплена обширным вычислительным экспериментом. Некоторые возможности теоретических подходов к этой интересной теме обсуждаются в приложении.

2. Детерминизм и стохастика

Еще важнее и принципиальнее вторая тема — детерминизм и стохастика. Давняя традиция противопоставляет друг другу эти понятия. Однако эволюция понятия стационарного режима заставляет взглянуть иначе на взаимоотношение детерминизма и стохастики (рис. 1).

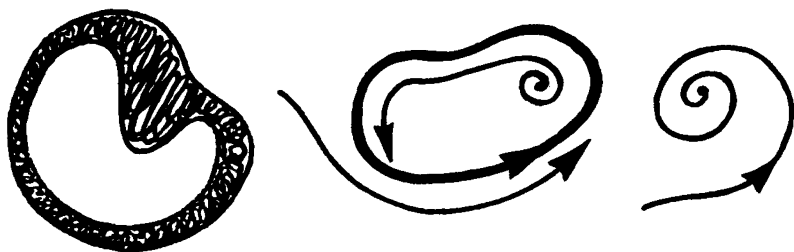


Рис. 1. Эволюция понятия стационарного режима. Положение устойчивого равновесия, предельный цикл, зона перемешивания

Зона перемешивания обладает характеристическим свойством — это множество, устойчивое снаружи (в целом) и неустойчивое изнутри (локально). Стоит заметить, что и предельный цикл обладает сходным свойством, но понято это было уже после анализа свойств зон перемешивания.

Возникающая ситуация позволяет сформулировать рабочую гипотезу: *и детерминизм, и стохастика — это два предельных способа описания сложных (и пока еще непривычных) свойств режимов перемешивания.*

При этом *детерминистический подход* выделяет те величины (типа средних значений), для которых характерен процесс установления на больших временах, вытекающий из эргодичности — устойчивости снаружи.

В противоположность этому *стохастический подход* обращает особое внимание на локальные свойства — неустойчивость изнутри.

Таким образом речь идет о новом, серьезном обобщении *теории устойчивости*. Возникает надежда на снятие противоречия между детерминистическим и стохастическим подходами, включении их в единую схему в качестве *предельных случаев*.

3. Общие замечания

Качественная теория обыкновенных дифференциальных уравнений подсказывает наглядную иллюстрацию к подобной острой ситуации — процессу смены парадигмы.

На первых порах кажется, что речь идет о непримеримом противоречии — взаимонесключающих подходах (рис. 2).

Затем накапливаются факты и в подходе А («стохастика»), и в подходе В («детерминизм»), уточняются представления, возникает «тонкая структура» первоначальных грубых представлений. Возникает догадка о новом синтезе С, с этой новой точки

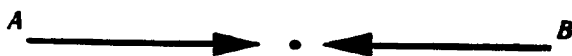


Рис. 2. Кризис парадигмы — столкновение точек зрения

зрения* переоцениваются исходные предпосылки A и B , выявляются те компоненты (α и β), которые перестают соответствовать новой точке зрения (рис. 3).

Локальная (экспоненциальная!) неустойчивость (гипотеза C) дает механизм «квазихаоса» и позволяет отказаться от формального введения плотности вероятностей («отметаемое» представление α). Та же гипотеза в детерминистском подходе B выявляет его слабое место — представление о точной предсказуемости любых деталей процесса, а не только главных, определяющих характеристик.

4. Заключение

Ныне на подъеме изучение сложных, многокомпонентных систем с сильным, нелинейным взаимодействием и велика роль теории нейронных сетей, нередко подсказывающей пути изучения других, в том числе физических систем («жидкие» кристаллы, например).

Нейронные сети — основной объект книги — характеризуются двумя существенными (для математики) свойствами.

Во-первых, они состоят из большого числа одинаковых (или похожих) элементов — нейронов. Поэтому «проклятие размерности» приобретает менее расплывчатую форму — форму необходимости предельного перехода $N \rightarrow \infty$. Во-вторых, взаимодействие этих элементов существенно нелинейно.

Считается, что это двойное осложнение. Успешно можно работать либо с нелинейностью, но тогда невысокой размерности, либо с $N \gg 1$, но тогда нужна близость к линейности.

Приложение посвящено, однако, попытке превратить «проклятие» в «благословение». Там указан класс гамилтоновых систем, для которых конструктивно осуществлен (независимо от силы взаимодействия) предельный переход $N \rightarrow \infty$. Вопрос о связи этих систем с нейронными сетями — дело будущего. Попытка носит методический характер и доказывает лишь принципиальную возможность предельного перехода.

* Философы легко узнают в этой схеме знаменитую гегелевскую триаду: тезис, антитезис \rightarrow синтез. Любопытно, однако, что «метод перевала» указывает не только на принципиальное равноправие тезиса и антитезиса, но и «цену» (необходимость отказа от α в A и от β в B), которую надо «платить» за новый синтез C . Короче — истина лежит между (а не «посередине») и сильно сбоку.

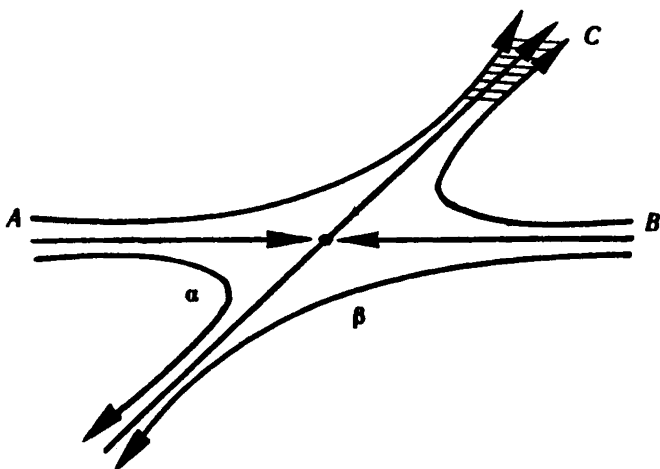


Рис. 3. Перевал. Взаимодействие тезиса А с антитезисом В приводит к синтезу С за счет отказа от устаревших представлений α и β

Однако из нее вытекает важный вывод. Термин «молекулярный хаос» связан с противопоставлением стохастичности (атомного уровня) и детерминистичности (целостной системы).

Дело, похоже, обстоит сложнее — режимы перемешивания могут возникать, как это видно из приложения, уже на макроуровне.

Методы нелинейной теории колебаний становятся, поэтому, все более актуальными при изучении Больших Систем. Это еще один аргумент в пользу глубокой внутренней аналогии (если не совпадения) будущей теории фазовых переходов и схемы бифуркаций в теории колебаний — «таблицы Менделеева» типов кинетики.

Термодинамический подход далеко не всегда достаточен при изучении многокомпонентных систем, важных для естествознания.

Главная причина состоит в том, что функционирование подобных систем (особенно биологических) происходит обычно вдали от положения термодинамического равновесия. Весьма стеснительным является также (характерное для термодинамических рассматриваний) предположение малости взаимодействия или какое-нибудь аналогичное* предположение о наличии в системе малого параметра. В результате теоретические выводы часто мало пригодны для понимания реальных ситуаций.

* Малая плотность, приближение сильной связи и т.п.

Крайне важно, поэтому, изучение таких систем (даже модельных), поведение которых можно проследить теоретически в широком диапазоне условий — «в целом».

1. Градиентные системы

Рассмотрим в пространстве векторов x

$$x = \begin{pmatrix} x_1 \\ \dots \\ x_s \end{pmatrix} \quad (1)$$

произвольную скалярную функцию h — гамильтониан

$$h = h(x) \quad (2)$$

и произвольную матрицу I

$$I = \begin{pmatrix} I_{11} & \dots & I_{1s} \\ \dots & \dots & \dots \\ I_{s1} & \dots & I_{ss} \end{pmatrix} \quad (3)$$

с постоянными коэффициентами.

Система уравнений

$$\frac{dx}{dt} = I \left(\frac{dh}{dx} \right)^* \quad (4)$$

называется *градиентной** системой или (I, h) -системой. Звездочка *, превращающая строку в столбец

$$\left(\frac{\partial h}{\partial x_1}, \dots, \frac{\partial h}{\partial x_s} \right)^* = \begin{pmatrix} \frac{\partial h}{\partial x_1} \\ \dots \\ \frac{\partial h}{\partial x_s} \end{pmatrix}, \quad (5)$$

необходима для «выравнивания тензорной размерности» левой и правой части системы уравнений.

В частном случае матрицы I , такой, что

$$I^2 = -E \quad (6)$$

(при четной размерности s фазового пространства), возникает собственно гамильтонова система.

* Впервые введены в работе А.М. Обухова (О симметризуемых нелинейных системах. ДАН СССР, 1977, т.233, с. 35-38) под названием «симметризуемых» систем. Для наших целей удобнее несколько изменить обозначения и опустить требование невырожденности матрицы I .

Пример I.

Пусть

$$x = \begin{pmatrix} p \\ q \end{pmatrix}, \quad h = h(p, q), \quad I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Тогда

$$\frac{d}{dt} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial h}{\partial p} \\ \frac{\partial h}{\partial q} \end{pmatrix} = \begin{pmatrix} -\frac{\partial h}{\partial q} \\ \frac{\partial h}{\partial p} \end{pmatrix}$$

и в координатной записи

$$\begin{cases} \frac{dp}{dt} = -\frac{\partial h}{\partial q} \\ \frac{dq}{dt} = \frac{\partial h}{\partial p} \end{cases}, \quad (7)$$

получается простейшая гамильтоновая система на плоскости (одна степень свободы). Если, кроме того положить

$$h(p, q) = \frac{p^2}{2m} + U(q) \quad (8)$$

то получаются уравнения движения материальной точки массы m в потенциальном поле $U(q)$.**Пример II.**

Рассмотрим случай, когда

$$x = \begin{pmatrix} p \\ q \\ z \end{pmatrix}, \quad h = h(p, q, z), \quad I = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Система уравнений имеет вид

$$\begin{cases} \frac{dp}{dt} = -\frac{\partial h}{\partial q} \\ \frac{dq}{dt} = \frac{\partial h}{\partial p} \\ \frac{dz}{dt} = \frac{\partial h}{\partial z} \end{cases} \quad (9)$$

После дифференцирования функции $h(p, q, z)$ в силу системы (9) получается:

$$\frac{dh}{dt} = \frac{\partial h}{\partial p} p + \frac{\partial h}{\partial q} q + \frac{\partial h}{\partial z} z = \left(\frac{\partial h}{\partial z}\right)^2 \geq 0 \quad (10)$$

Следовательно, гамильтониан h является одновременно функцией Четаева системы, порожденной этим гамильтонианом.Если, в частности, $h(p, q, z)$ не зависит от z :

$$\frac{\partial h}{\partial z} = 0$$

то из уравнений (9) получаются уравнения (7).

2. Парное взаимодействие

Пусть уравнения движения макросистемы, состоящей из большого* числа $N \gg 1$ одинаковых компонент x_i , задаются гамильтонианом \mathbf{H}

$$\mathbf{H} = \mathbf{H}(x_1, \dots, x_N) \quad (11)$$

и матрицей I

$$\frac{dx_i}{dt} = I \left(\frac{\partial \mathbf{H}}{\partial x_i} \right)^* \quad (12)$$

Обычно рассматривается *парное взаимодействие*

$$\mathbf{H} = \sum_{i=1}^N \sum_{l=1}^N H(x_i, x_l) \quad (13)$$

когда гамильтониан \mathbf{H} — это сумма N^2 гамильтонианов H

$$H = H(x', x'') \quad (14)$$

описывающих взаимодействие двух компонент x' и x'' . Слагаемые в (13) отличаются друг от друга только номерами аргументов.

3. Сумматорные функции А.Я. Хинчина

Последовательное изучение свойств сумматорных функций в статистических вопросах восходит к работе А.Я. Хинчина (Математические основания статистической механики. М., Гостехиздат, 1943) и связано с обобщением на функции Закона Больших Чисел, хотя роль симметрических функций известна еще из алгебры — формулы Ф. Виета (≈ 1590).

В нашей задаче сумматорные функции возникают сами собой, как следствие применения метода Фурье разложения функций в ряд по ортогональному базису**.

Итак, рассмотрим произвольный (пока!) базис $\{y_\alpha(x)\}$ в пространстве функций одной переменной x и разложим гамильтониан взаимодействия $H(x', x'')$ по соответствующему базису $\{y_\alpha(x') y_\beta(x'')\}$ в пространстве функций двух переменных

* Типичный пример — число Авогадро (1811) $N_A = 6,022 \cdot 10^{23} = 6022 \cdot 10^{20}$.

** Современная модификация — метод Галеркина-Ритца — не предполагает даже ортогональности.

$$H(x', x'') = \sum_{\alpha\beta} H^{\alpha\beta} y_{\alpha}(x') y_{\beta}(x''). \quad (15)$$

Подставляя полученное выражение в гамильтониан системы \mathbf{H} , получаем:

$$\mathbf{H} = \sum_{i\alpha\beta} H^{\alpha\beta} y_{\alpha}(x_i) y_{\beta}(x_i) = \sum_{\alpha\beta} H^{\alpha\beta} \left[\sum_{\alpha} y_{\alpha}(x_i) \right] \left[\sum_{\beta} y_{\beta}(x_i) \right] \quad (16)$$

Сумматорные функции

$$Y_{\alpha} = \sum_{i=1}^N y_{\alpha}(x_i) \quad (17)$$

возникают, таким образом, автоматически, при изменении* порядка суммирования. Выражая гамильтониан \mathbf{H} через новые переменные Y_{α}

$$\mathbf{H} = \sum_{\alpha\beta} H^{\alpha\beta} Y_{\alpha} Y_{\beta}, \quad (18)$$

приходим к важному выводу.

Гамильтониан есть квадратичная функция макровеличин Y_{α} , и эта квадратичность есть прямое следствие парности взаимодействия.

Отметим также, что коэффициенты квадратичной формы (18) совпадают с коэффициентами Фурье $H^{\alpha\beta}$ разложения гамильтониана взаимодействия $H(x', x'')$ по базису $\{y_{\alpha}(x') y_{\beta}(x'')\}$.

Сделаем еще одно замечание о производных сумматорной функции. Дифференцируя Y_{α} по x_i получаем:

$$\frac{\partial Y_{\alpha}}{\partial x_i} = \frac{dy_{\alpha}(x_i)}{dx_i} \quad (19)$$

Величина Y_{α} содержит N слагаемых, а производная — лишь одно, ибо остальные слагаемые не зависят от x_i и обращаются в нуль при дифференцировании. Это обстоятельство тесно связано с характером асимптотики (при $N \rightarrow \infty$) сумматорных функций. Каждая сумматорная функция порождает (в разных контекстах) три масштаба величин.

Максимальное значение имеет порядок N (по числу слагаемых).

Средняя флуктуация (корень квадратный из интеграла квадрата функции) значительно меньше и имеет порядок \sqrt{N} .

И, наконец, производные совсем малы по сравнению с N (порядка единицы) как это видно из (19).

* Отметим сходство этой процедуры с лебеговской схемой интегрирования и переходом к числам заполнения в квантовой статистике.

4. Макродинамика

Запись гамильтониана системы H в переменных Y

$$H = H(Y) \quad (20)$$

позволяет упростить уравнения движения.

Вычисляя производную $\frac{\partial H}{\partial x_i}$

$$\frac{\partial H}{\partial x_i} = \sum_{\beta} \frac{\partial H}{\partial Y_{\beta}} \frac{\partial Y_{\beta}}{\partial x_i} = \sum_{\beta} \frac{\partial H}{\partial Y_{\beta}} \frac{\partial y_{\beta}}{\partial x_i}$$

и подставляя результат в уравнения движения, получим

$$\frac{dx_i}{dt} = \sum_{\beta} \frac{\partial H}{\partial Y_{\beta}} l \left(\frac{\partial y_{\beta}}{\partial x_i} \right) * \quad (21)$$

Правые части полученных уравнений содержат как скалярные (макро) величины* L^{β}

$$L^{\beta} = \frac{\partial H}{\partial Y_{\beta}}, \quad (22)$$

так и векторные (микро) поля $a_{\beta(x)}$

$$a_{\beta}(x) = l \left(\frac{\partial y_{\beta}}{\partial x} \right) * \quad (23)$$

В этих обозначениях более четко видна структура уравнений движения

$$\frac{dx_i}{dt} = \sum_{\beta} L^{\beta} a_{\beta}(x_i). \quad (24)$$

подсказывающая идею выделения *макродинамики* (эволюции Y_{β}) из огромного множества ($N \gg 1$) μ движений x_i .

Дифференцируя по времени соотношение (17), определяющее M — величину Y_{α} ,

$$\frac{dY_{\alpha}}{dt} = \sum_i \frac{dy_{\alpha}}{dx_i} \frac{dx_i}{dt}$$

и подставляя из (24) найденные значения производных x_i , получаем:

* В дальнейшем употребляется более наглядная запись типа M -«переменное» вместо «(макро)переменное» и « μ -поле» вместо «(микро)поле».

$$\frac{dY_\alpha}{dt} = \sum_i \frac{dy_\alpha}{dx_i} \left[\sum_\beta L^\beta a_\beta(x_i) \right] = \sum_\beta L^\beta \left\{ \sum_i \frac{dy_\alpha}{dx_i} a_\beta(x_i) \right\}. \quad (25)$$

Выражение в фигурных скобках есть сумматорная функция, каждое слагаемое которой может быть разложено по базису

$$\frac{dy_\alpha}{dx} a_\beta(x) = \sum_\gamma I_{\alpha\beta}^\gamma y_\gamma(x) \quad (26)$$

Здесь постоянные $I_{\alpha\beta}^\gamma$ (коэффициенты Фурье) определяются только свойствами базиса $\{y_\alpha(x)\}$ и никак не связаны с исходными уравнениями. Наиболее отчетливо это видно, если вспомнить определение μ -поля $a_\beta(x)$ (23) и подставить его в (26):

$$\frac{dy_\alpha}{dx} I \left(\frac{dy_\beta}{dx} \right)^* = I_{\alpha\beta}^\gamma y_\gamma(x) \quad (27)$$

Суммирование по всем компонентам x_i в (25) с учетом формулы (26) приводит к полному исключению μ -переменных, и мы получаем уравнения *макродинамики*:

$$\frac{dY_\alpha}{dt} = I_{\alpha\beta}^\gamma Y_\gamma(x) L^\beta.$$

5. I-алгебры. Квадратичные многочлены

Предположение о парности взаимодействия, как видно из вывода, можно не вводить. Значительно серьезнее другое ограничение.

Проведенный анализ не претендует на формальную строгость, так как игнорировались вопросы сходимости возникающих рядов. Однако, эти выкладки становятся вполне строгими для важного случая *I-алгебр*.

Определение:

Конечномерное подпространство в пространстве функций с базисом $\{y_\gamma(x)\}$ называется I-алгеброй, если оно замкнуто относительно операции I — произведения, определенной следующим образом:

$$y_\alpha y_\beta = \frac{dy_\alpha}{dx} I \left(\frac{dy_\beta}{dx} \right)^* = I_{\alpha\beta}^\gamma y_\gamma(x) \quad (29)$$

Коэффициенты разложения $I_{\alpha\beta}^\gamma$ называются структурными константами алгебры.

Непустота множества определяемых объектов есть очевидное требование содержательности определения. В нашем случае пространство квадратичных форм образует алгебру для любой матрицы I . Это вытекает из того, что производная (градиент)

квадратичной формы есть линейная форма, а произведение двух линейных форм есть квадратичная форма.

Эту алгебру можно расширить, добавив константу. Другое полезное расширение состоит в добавлении всех линейных функций. Этот пример, демонстрируя непустоту множества I -алгебр, далеко не исчерпывает его.

Пример III.

В простейшем случае системы с одной степенью свободы конечная алгебра порождается всего двумя функциями,

$$a(x) = x, \quad b(x) = \frac{x^2}{2} \quad (30)$$

к которым мы добавим еще константу. Уже в этом случае возникает содержательная задача.

Выкладки, как это обычно бывает, проще провести заново, нежели использовать готовый результат.

Итак, пусть

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial x_i} \quad (31)$$

где

$$H = H(A, B) \quad (32)$$

и

$$A = \sum_i x_i, \quad B = \sum_i \frac{x_i^2}{2} \quad (33)$$

Уравнения для переменных x_i в новых обозначениях приобретают вид:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial A} + x_i \frac{\partial H}{\partial B} \quad (34)$$

откуда можно вывести (дифференцируя равенства (33) по t), замыкающие систему (34) уравнения макродинамики:

$$\begin{cases} \frac{dA}{dt} = N \frac{\partial H}{\partial A} + A \frac{\partial H}{\partial B} \\ \frac{dB}{dt} = A \frac{\partial H}{\partial A} + 2B \frac{\partial H}{\partial B} \end{cases} \quad (35)$$

В уравнениях (35) не использовано предположение о парности взаимодействия (и, следовательно, квадратичности гамильтониана H по M -переменным A и B). Они справедливы при значительно более общем предположении (32). Заметим, что все уравнения (34) совпадают (с точностью до обозначений) с уравнением μ -движения одной компоненты:

$$\frac{dx}{dt} = \frac{\partial H}{\partial A} + x \frac{\partial H}{\partial B} \quad (36)$$

Поэтому система из N уравнений сведена, по существу, к системе трех уравнений (35) и (36). Проинтегрировав (аналитически или численно) систему (35), мы получаем одно линейное уравнение (36) с переменными коэффициентами. Любое решение $x_i(t)$ получается подстановкой надлежащих начальных данных в решение уравнения (36).

Пример IV.

В двумерном случае,

$$x = \begin{pmatrix} p \\ q \end{pmatrix} \quad (37)$$

возникает пять M -переменных (из алгебры квадратичных многочленов):

$$\begin{aligned} P &= \sum p_i & W &= \sum \frac{p_i^2}{2} \\ R &= \sum p_i q_i \\ Q &= \sum q_i & Z &= \sum \frac{q_i^2}{2} \end{aligned} \quad (38)$$

Гамильтонова система

$$\begin{cases} \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \\ \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \end{cases} \quad (39)$$

в случае гамильтониана H , зависящего только от M -величин

$$H = H(P, Q, R, W, Z). \quad (40)$$

переписывается в виде:

$$\begin{cases} \frac{dp_i}{dt} = -\frac{\partial H}{\partial Q} - p_i \frac{\partial H}{\partial R} - q_i \frac{\partial H}{\partial Z} \\ \frac{dq_i}{dt} = \frac{\partial H}{\partial P} + p_i \frac{\partial H}{\partial W} + q_i \frac{\partial H}{\partial R}. \end{cases} \quad (41)$$

Дифференцируя равенства (38) по t и подставляя соотношения (41), получим уравнения макродинамики:

$$\left\{ \begin{array}{l} \frac{dP}{dt} = -N \frac{\partial H}{\partial Q} - P \frac{\partial H}{\partial R} - Q \frac{\partial H}{\partial Z} \\ \frac{dQ}{dt} = N \frac{\partial H}{\partial P} + P \frac{\partial H}{\partial W} + Q \frac{\partial H}{\partial R} \\ \frac{dR}{dt} = P \frac{\partial H}{\partial P} + 2W \frac{\partial H}{\partial W} - Q \frac{\partial H}{\partial Q} - 2Z \frac{\partial H}{\partial Z} \\ \frac{dW}{dt} = -P \frac{\partial H}{\partial Q} - 2W \frac{\partial H}{\partial R} - R \frac{\partial H}{\partial Z} \\ \frac{dZ}{dt} = Q \frac{\partial H}{\partial P} + R \frac{\partial H}{\partial W} + 2Z \frac{\partial H}{\partial R} \end{array} \right. \quad (42)$$

Механизм выделения макродинамических уравнений тот же самый, разумеется, что и в случае I-алгебры общего вида. Однако, результат весьма интересен. Особо отметим важный частный случай механических систем, когда гамильтониан есть сумма кинетической и потенциальной энергии:

$$H = \frac{p_i^2}{2Nm} + U\left(q_i, q_l, \frac{q_i^2}{2}, \frac{q_l^2}{2}\right). \quad (43)$$

В этом случае

$$H = \frac{1}{m}W + U(Q, Z). \quad (44)$$

и система макродинамических уравнений существенно упрощается:

$$\left\{ \begin{array}{l} \frac{dP}{dt} = -N \frac{\partial U}{\partial Q} - Q \frac{\partial U}{\partial Z} \\ \frac{dQ}{dt} = \frac{1}{m}P \\ \frac{dR}{dt} = \frac{2}{m}W - Q \frac{\partial U}{\partial Q} - 2Z \frac{\partial U}{\partial Z} \\ \frac{dW}{dt} = -P \frac{\partial U}{\partial Q} - R \frac{\partial U}{\partial Z} \\ \frac{dZ}{dt} = \frac{1}{m}R \end{array} \right. \quad (45)$$

В задачу настоящей работы не входит анализ поведения системы уравнений макродинамики, однако, следует отметить, что вычислительные эксперименты, проведенные с системой (45), обнаружили большое разнообразие возникающих в ней динамических режимов.

6. «Расщепление» компонент

Полученные результаты допускают наглядное истолкование. Изложим (конспективно) основные идеи анализа.

Пусть имеется произвольная* I -алгебра $\{y_\gamma(x)\}$

$$\frac{dy_\alpha}{dx} I \left(\frac{dy_\beta}{dx} \right)^* = I_{\alpha\beta}^\gamma y_\gamma(x) \quad (I)$$

и гамильтониан H

$$H = H(Y_1, \dots, Y_k), \quad (H)$$

зависящий только от M -переменных Y

$$Y_\alpha = \sum_i y_\alpha(x_i), \quad (Y)$$

и являющийся (в случае парного взаимодействия) квадратичным многочленом от своих аргументов.

Система (I, H) , порожденная матрицей I и гамильтонианом H

$$\frac{dx_i}{dt} = I \left(\frac{\partial H}{\partial x_i} \right)^* \quad (x)$$

и записанная в новых переменных,

$$\frac{dx_i}{dt} = \sum_\beta \frac{\partial H}{\partial Y_\beta} I \left(\frac{dy_\beta}{dx_i} \right)^*,$$

допускает «выщепление» M -движений Y :

$$\frac{dY_\alpha}{dt} = I_{\alpha\beta}^\gamma Y_\gamma \frac{\partial H}{\partial Y_\beta}. \quad (M)$$

Более того, при известных Y система μ -движений x_i распадается на *независимые* движения. Все уравнения оказываются копиями (с точностью до обозначения переменных) одного стандартного уравнения для x :

$$\frac{dx}{dt} = \frac{\partial H}{\partial Y_\alpha} I \left(\frac{dy_\alpha}{dx} \right)^*. \quad (\mu)$$

Траектория системы в многомерном M -пространстве (x_1, \dots, x_N) может быть, поэтому, *точно* заменена набором из N траекторий (отличающихся только начальными данными) в стандартном μ -пространстве x .

Полезно иллюстрировать схемой этапы анализа.

* Напомним еще раз, что I -алгебра по определению конечна и ее размерность k не зависит от количества компонент n .

В исходной системе каждая компонента взаимодействует с каждой.

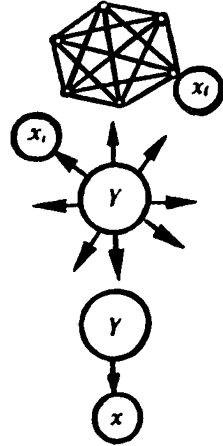
Всего N^2 связей.

Введение «лишних» M -переменных Y позволяет заменить взаимодействие действием Y на x_i .

Остается N связей.

Все компоненты одинаковы. Поэтому движение в многомерном M -пространстве равнозначно движению «тучи точек» в стандартном μ -пространстве x .

Существенна только одна связь.



Для «расцепления» компонент достаточно знания одной траектории M -системы. Вычисление такой траектории не представляет проблемы для современной вычислительной техники, так как размерность* M -системы не зависит от N , хотя само число N и может входить в коэффициент M -системы, что мы уже видели в примерах III и IV.

7. Эквивалентное поле и макродинамическое равновесие

Уравнение (μ) удобно переписать в следующем виде:

$$\frac{dx}{dt} = I \left(\frac{\partial h}{\partial x} \right)^*, \quad (x)$$

вводя «эквивалентное» поле h :

$$h(x, Y) = \sum_a \frac{\partial H}{\partial Y_a} y_a(x). \quad (h)$$

Поэтому движение x можно интерпретировать, как движение во «внешнем» поле h , параметры которого определяются M -величинами Y .

Исследование любых систем обычно начинается с изучения стационарных режимов (а нередко этим и заканчивается, особенно если для практических целей достаточно анализа именно стационарных режимов).

* Мрачный пессимизм словосочетания «проклятие размерности» не относится к M -системе. Введение «лишних» макропеременных Y может, поэтому, рассматриваться как «снятие заклятия».

«Треугольная» структура* системы уравнений для Y, x

$$\begin{cases} \frac{dY}{dt} = f(Y) \\ \frac{dx}{dt} = g(Y, x) \end{cases}$$

позволяет ввести понятие *макродинамического равновесия*,

$$f(Y) = 0, \quad Y = \text{const},$$

которое означает стационарность только M -системы. Это естественное обобщение понятия «термодинамического равновесия».

В состоянии M -равновесия эквивалентное поле h оказывается не зависящим от времени и система (x) становится автономной.

Весьма правдоподобно, что этот важный частный случай понятия эквивалентного поля соответствует понятию «самосоглазованного поля» в физике.

Однако подобный анализ этого интересного вопроса выходит за рамки настоящей публикации.

Стационарное эквивалентное поле (или просто h -поле) интересно еще и тем, что дает простую и наглядную интерпретацию предельного перехода $N \rightarrow \infty$.

Каждая траектория исходной системы порождает (как мы видели) N траекторий в пространстве одной компоненты x . Если начальные точки этих траекторий распределены достаточно равномерно, то траектории (в пределе) заполняют все пространство. Следовательно *фазовый портрет* системы может быть интерпретирован, как *предел проекции* на μ -пространство *одной* траектории исходной системы. Эта неожиданная связь обещает интересные результаты в дальнейших исследованиях и существенно расширяет область приложений качественной теории обыкновенных дифференциальных уравнений.

Заметим также, что в случае парного взаимодействия правые части уравнений макродинамики всегда квадратичны. Это указывает на глубокое родство обсуждаемых систем с общими билинейными системами (см. Молчанов А. М. Билинейные системы. Препринт. Пущино, ОНТИ НЦБИ АН СССР, 1982).

8. Стационарные режимы

Ниже приведены некоторые результаты численных экспериментов с уравнениями макродинамики, за проведение которых автор благодарен А. С. Кондрашову.

В подборе примеров нет претензии на полноту. Из них, однако, видно, что кроме положений равновесия существуют и другие,

* Состоящая в том, что поведение Y не зависит от x .

весьма разнообразные (чтобы не сказать затейливые) стационарные режимы.

Другое замечание касается традиционных методов обработки экспериментальных данных. Обычно строят зависимость от времени для каждой измеряемой величины отдельно (например, электрокардиограмма). При анализе внешнего воздействия (скажем, температуры) опять-таки строят независимые кривые.

Приводимые иллюстрации позволяют поставить на эту тему полезный мысленный эксперимент. На каждой странице (кроме шести проекций траектории макродвижения на различные плоскости) вверху расположен график $W(t)$.

Будем считать эту величину W «наблюдаемой» в отличие от «скрытых переменных» P, Q, R, Z .

Много ли можно узнать о движении пятимерного вектора P, Q, R, Z, W по поведению «наблюдаемой» величины W предоставляем судить читателю.

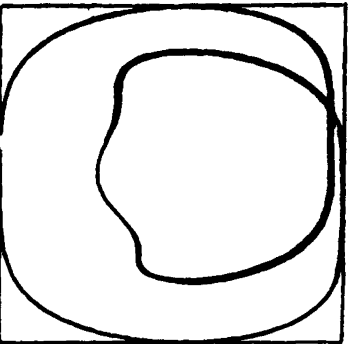
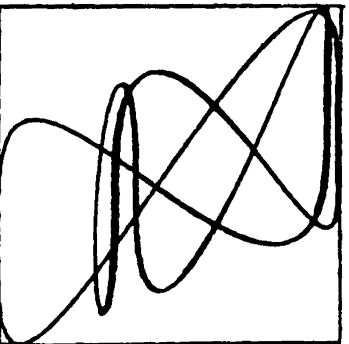
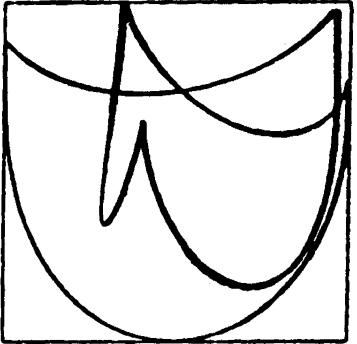
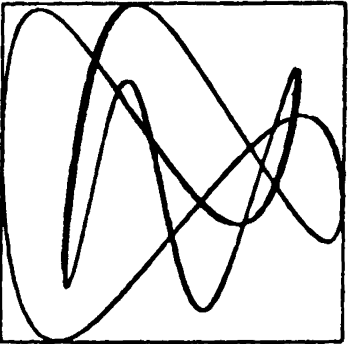
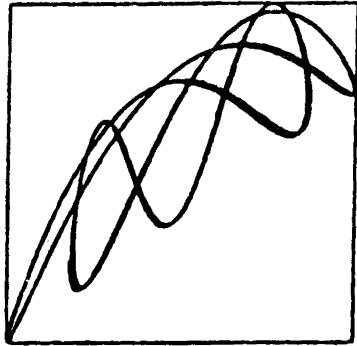
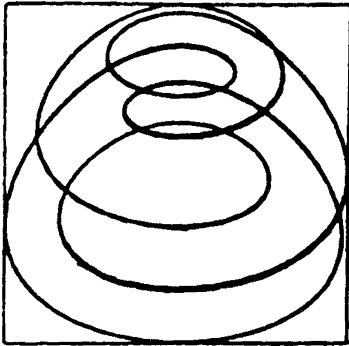
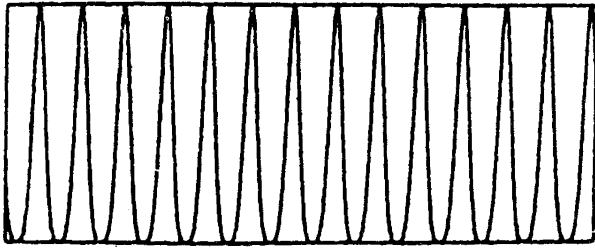


Рис. 1

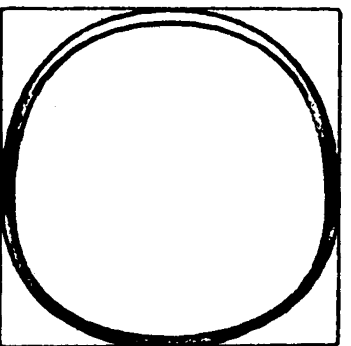
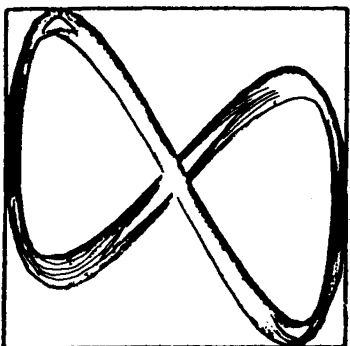
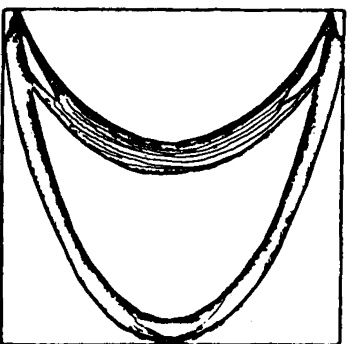
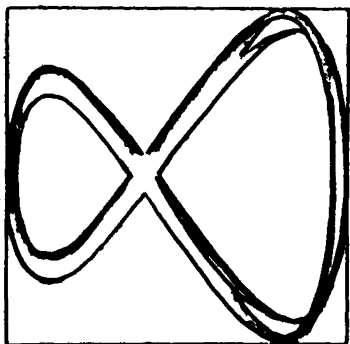
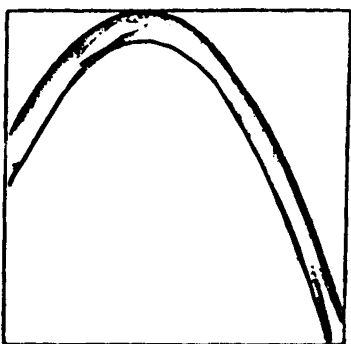
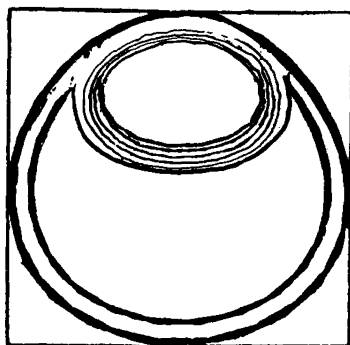
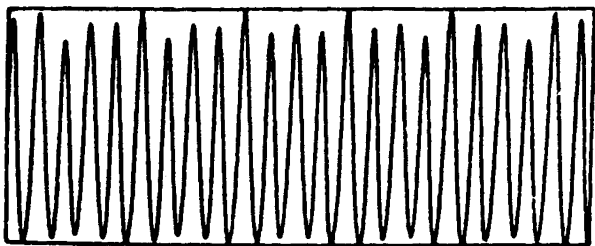


Рис. 2

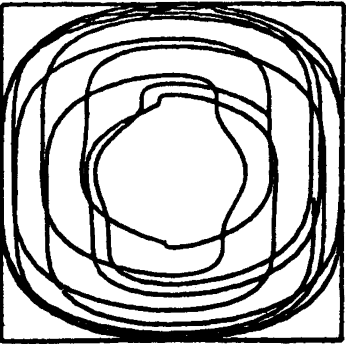
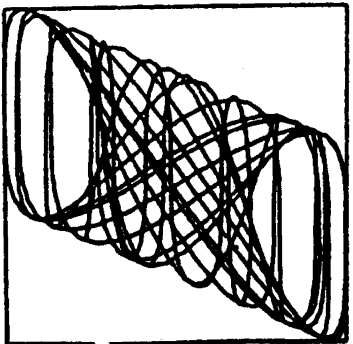
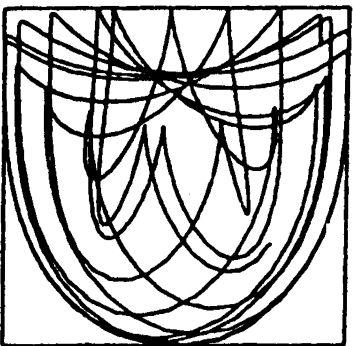
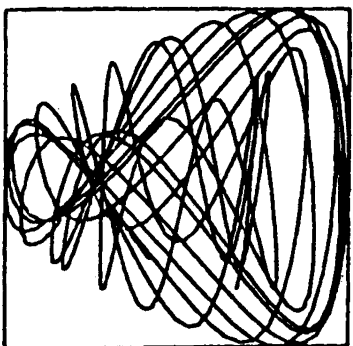
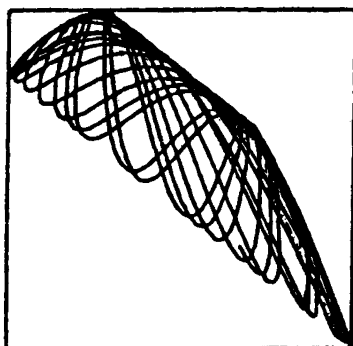
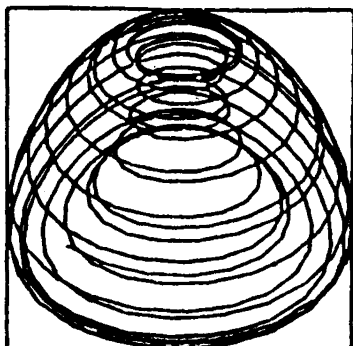
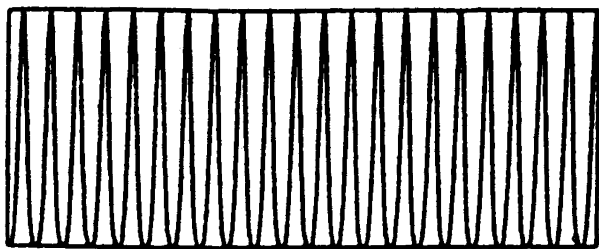


Рис. 3

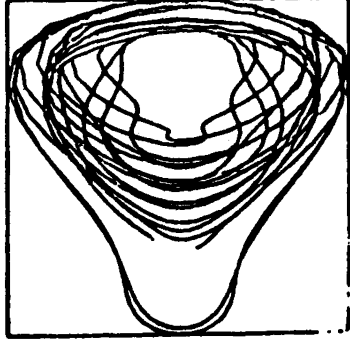
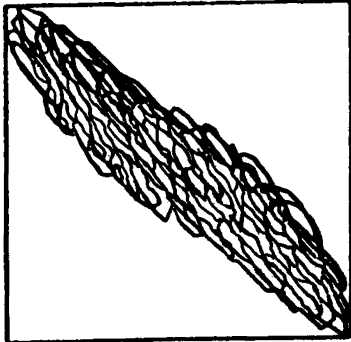
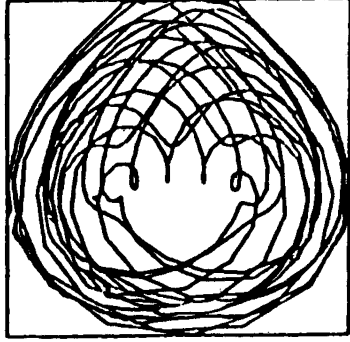
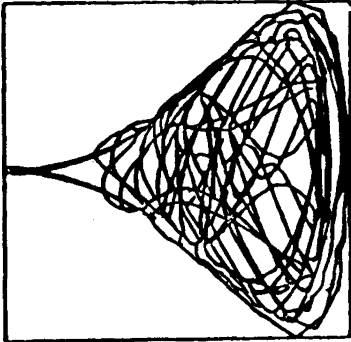
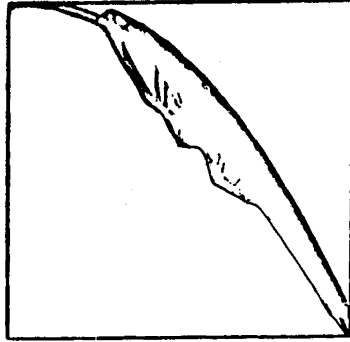
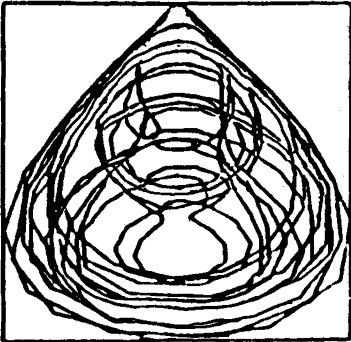
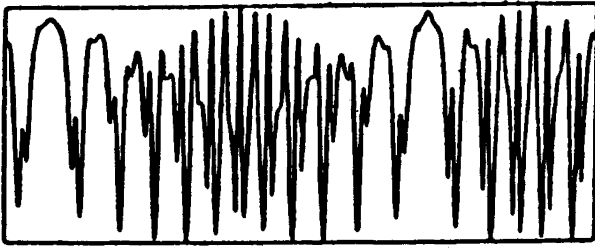


Рис. 4

ГИПОТЕЗА РЕЗОНАНСНОЙ СТРУКТУРЫ СОЛНЕЧНОЙ СИСТЕМЫ

В статье изложена история изучения резонансных движений и современное состояние вопроса. Сформулировано предположение о решающей роли «исчезающих» диссипативных факторов в формировании резонансных структур. Приведены общие соображения о значении резонансных явлений в естествознании.

Введение. История вопроса

Небесная механика стала количественной, математической наукой после формулировки Ньютоном закона всемирного тяготения (1687). Движение планет Солнечной системы определяется (в главном члене) полем тяготения Солнца, основного резко доминирующего тела системы. С точки зрения теории колебаний планетная система состоит из одночастотных колебательных подсистем, почти независимых, слабо связанных друг с другом. Необходимо подчеркнуть, что распадение на колебательные подсистемы существенно отличается от распадения на физические тела — планеты. Каждая колебательная подсистема состоит из пары физических тел — Солнца и планеты. Таким образом, в сложной колебательной системе — планетной — Солнце повторено девятикратно (по числу планет).

а. Большое вековое возмущение — резонанс Юпитера и Сатурна

Ньютон (1687) [1] объяснил кеплеровы (1609) эллипсы и разобрал кометные гиперболы. Учет следующего малого члена — взаимодействие планет — был осуществлен Лагранжем и Лапласом (1776) в форме теории возмущений [2]. Сам Ньютон считал, что взаимные возмущения планет приводят Солнечную систему в беспорядок и только божественное вмешательство сохраняет стройную картину мироздания.

Резонансные явления появляются уже на этом этапе в роли пособников дьявола. Драматическая ситуация, причиной которой явилась почти соизмеримость периодов Юпитера и Сатурна, сочувственно описана Субботиным (1937) [3]: «Большое неравенство в движении Юпитера и Сатурна, зависящее от делителя $2\omega_{\text{J}} - 5\omega_{\text{S}}$, было открыто эмпирически. Эйлер и Лагранж, после

ряда безуспешных попыток объяснить его, стали склоняться к мнению, что это неравенство указывает на действие какой-то силы, отличной от притяжения Солнца и известных планет. Истинная причина была открыта в 1784г., когда Лаплас предпринял вычисление всех неравенств первого порядка в движении Юпитера и Сатурна до третьей степени включительно».

б. Троянцы. Резонанс 1: 1 (унисон)

Существует очень немного точных решений задачи трех тел. Однако из них — треугольник Лагранжа (1772) — является резонансным движением. Наблюдение опоздало более чем на сто лет по сравнению с теорией. Только в 1904г. Вольф открыл (применив стереоскоп) первый астероид, образующий вместе с Юпитером и Солнцем равносторонний лагранжев треугольник. В настоящее время известно более дюжины [4, 5] астероидов-троянцев. Астероиды, названные именами греческих героев, предшествуют Юпитеру. В эту группу, впрочем попал (видимо, в качестве пленника) Гектор (№ 624). Побежденные троянцы замыкают это торжественное шествие во славу науки.

в. Пояс астероидов. Пробелы Кирквуда

Большую серию резонансных движений, воспринимаемых опять-таки как досадные помехи [6] в стройной теории доставляет пояс астероидов. Хорошо известны [7] пробелы Кирквуда, соответствующие резонансам 1:2, 2:5, 1:3 с обращением Юпитера. Менее заметные понижения в кривой распределения периодов обращения астероидов возникают при резонансах 1:4, 1:5, 3:5, 3:7.

Существует и противоположная ситуация — группировка орбит вблизи точек 3:4 и 2:3. В музыкальной терминологии это — «кварта» и «квинта». «Прима» также устойчива и соответствует группе троянцев.

г. Кольца Сатурна

Знаменитая «щель Кассини» (1680) имеет резонансную природу. Она занимает ту зону, в которой частички, составляющие кольца Сатурна, имели бы периоды, близкие к $1/2$ периода Мимаса, к $1/3$ периода Энцелада и $1/4$ периода Тефии. Для понимания этого явления недостаточно было обнаружить щель и открыть спутники Сатурна. С этим справился сам Кассини. Мало было даже открыты другие пробелы в кольцах Сатурна. Только в XIX веке Кирквуд (1850), сопоставив пробелы в поясе астероидов с кольцами Сатурна, осознал единый резонансный механизм образования пробелов.

д. Галилеевы спутники Юпитера

Эта единственная подсистема Солнечной системы, стимулировавшая благосклонное отношение небесных механиков к резонансу.

Резонанс был обнаружен [8] еще Лапласом (1789) в форме дополнительного первого интеграла

$$\varphi_1 - 3\varphi_2 + 2\varphi_3 \approx \pi, \quad (1)$$

где φ_1 , φ_2 , φ_3 — планетоцентрические долготы спутников Ио, Европы и Ганимеда. Этот резонанс долгое время был исключительным и в другом смысле. Только он связывал воедино три тела, а не два, как во всех остальных случаях. Форма записи этого замечательного соотношения уже предвещает правильную форму резонанса

$$\omega_1 - 3\omega_2 + 2\omega_3 \approx 0 \quad (2)$$

как целочисленного линейного уравнения для частот, а не для периодов, как обычно записывают парные резонансы.

е. Луна

Вдохновляющий пример резонансного движения демонстрирует наш единственный естественный спутник — Луна. Это тоже «прима» — унисон 1:1 — как и в случае троянцев, но уже вращение с обращением.

Джордж Дарвин (1907) [9], сын знаменитого Чарлза Дарвина, был, по-видимому, первым, кто попытался эволюционно проследить возникновение резонансной структуры. Он правильно указал главный фактор эволюции системы Луна — Земля — приливные силы. К сожалению, неверна, как показал Ляпунов (1882) [10], исходная гипотеза Пуанкаре об устойчивости грушевидной фигуры вращения жидкой массы. Поэтому теория Д. Дарвина неприменима даже к вопросу о происхождении двойных звезд, не говоря уже о системе Луна — Земля, для которой наиболее вероятно [14] происхождение из газо-пылевого облака.

ж. Солнечная активность

«Солнце менее полезно, чем Луна, ибо светит днем [11], когда и без него светло, а Луна ночью». Кроме того на Солнце есть пятна. Однако именно солнечные пятна проливают дополнительный свет на проблему возникновения резонансных структур.

Средняя периодичность [12] солнечной активности подозрительно близка к периоду обращения Юпитера. Аристарх Аполлонович Белопольский (1886) [13] попытался усмотреть в этом факте причинную связь. Мысль эта, понимаемая буквально, несомненно неверна, так как солнечная активность не является

однопериодическим процессом. Однако открытым остается вопрос, нельзя ли связать солнечную активность с приливным воздействием обеих главных планет — Юпитера и Сатурна.

Независимо от справедливости этой двухчастотной модели пятна на Солнце пополюют нашу коллекцию нетривиальных точным резонансом 1: 2 («октава») магнитных и механических явлений, ибо надежно установлено, что полярность пятен меняется на противоположную в соседних циклах.

з. Планетные расстояния или частоты?

Эмпирическое правило (Nieto, 1970) планетных расстояний Тициуса (1766) — Бодде (1772) положило начало пониманию Солнечной системы как единого целого. Это правило было пересмотрено Отто Юльевичем Шмидтом (1946) с эволюционных позиций. Однако основой рассмотрения оставались планетные расстояния.

Автор настоящей статьи попытался [15] применить к Солнечной системе идеи нелинейной теории колебаний. Такой подход вынуждал рассматривать частоты (а не большие полуоси) в качестве основных характеристик движения. Обнаружилось, что частоты планет имеют максимально возможный набор резонансных соотношений. Неосторожно высказанная заветная мысль (восходящая к Четаеву (1929) [16], как потом выяснилось*) о «квантованности» Солнечной системы, «целочисленности» ее структурного принципа вызвала дружное негодование и астрономов, и физиков: «Случайность!» — возмущаются первые [17—19]. «А что играет роль постоянной Планка?» — обличают вторые.

Гипотеза максимальной резонансности, высказанная в цитированной работе [15], может, конечно, оказаться неверной. Однако она полезна самой постановкой вопроса о судьбе любой колебательной системы, эволюционирующей под воздействием слабых диссипативных факторов. Единство подхода к резонансам вращения и обращения выгодно отличает эту гипотезу от любой модификации «правила планетных расстояний».

и. Спин-орбитальное взаимодействие

Развитие наблюдательных средств (прежде всего радиолокационной техники) позволило обнаружить новые типы резонансов. Так, у Меркурия был найден [20] резонанс 2:3 («квинта») вращения с обращением. Прошло немного времени, и наблюдения выявили удивительно сложный резонанс [21—23] вращения Венеры вокруг оси с обращением Земли и Венеры вокруг Солнца.

* Molchanov (1969) [32]

В результате этого резонанса Венера оказывается похожей на Луну — в моменты наибольшего сближения Венера почтительно смотрит на Землю. Впрочем в реальности этого резонанса сомневается ныне даже один из авторов (Goldreich, 1970). Хороший пример колебательности постижения истины!

к. Ориентация искусственных спутников Земли

Большое разнообразие задач, выполняемых спутниками, делает практически важным систематическое изучение достаточно широкого класса резонансов вращения с обращением — так называемых номинальных движений (Белецкий, 1965). Особенно важны, конечно, классические движения лунного (или меркурианского) типа для связанных, метеорологических и геодезических спутников.

Но если в «академических» вопросах дело природы позаботиться об устойчивости резонансного движения, то в инженерных вопросах «диссипативный фактор» приходится «загружать на борт спутника» в форме той или иной системы активной стабилизации.

В настоящее время имеется обширная литература (Торжевский, 1969), посвященная изучению этого круга вопросов.

Аксиоматика небесной механики

Небесная механика за неполных три века существования блестяще справилась со своей задачей — задачей объяснения видимого движения небесных тел. Субботин (1955) [24] с законной гордостью писал: «В 1950 г. удалось показать, что закон всемирного тяготения позволяет представить наблюдения (с 1780 по 1940) пяти планет Юпитеровой группы (Юпитер, Сатурн, Уран, Нептун и Плутон) со всей той точностью, которую имеют эти наблюдения». В движении планет Земной группы (Меркурий, Венера, Земля и Марс) — главная невязка — знаменитое годовое перемещение перигелия Меркурия 5,75 секунды дуги вместо теоретического 5,34 секунды дуги — была объяснена Эйнштейном (1915) на основе релятивистского уточнения закона всемирного тяготения. Другие невязки, несравненно более мелкие, устранены после открытия Де Ситгером и Спенсером Джонсом (1926) неравномерности вращения Земли [25, 26].

Небесная механика вслед за геометрией стала чисто аксиоматической наукой — таков важнейший методологический результат истории ее развития. Система аксиом небесной механики значительно проще системы аксиом геометрии и записывается тремя строчками уравнений движения:

$$\frac{dp_i}{dt} = - \frac{\partial H}{\partial q_i}$$

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}$$

$$H(p_i, q_i) = \sum_{i=1}^n \frac{p_i^2}{2m_i} + \sum_{i \neq k} \sum \gamma \frac{m_i m_k}{|q_i - q_k|} \quad (3)$$

Построение аксиоматики любого явления неминуемо ставит вопрос о границах применимости математической модели. В небесной механике этот вопрос (имеющий специфическую форму проблемы устойчивости планетной системы) испокон веку стоял драматически остро. Даже для Ньютона и Эйлера, не говоря уже о Галилее, он был прямо связан с проблемой бытия божия. Постепенно теологический накал остывал, время от времени давая знать острыми вспышками уже чисто научных споров, что под слоем пепла еще таится пламя.

Между тем шло успешное исследование математической модели — теории возмущений гамильтоновых систем. Андрей Николаевич Колмогоров (1954) предложил замечательную модификацию итерационного метода Ньютона для доказательства гипотезы об устойчивости многочастотных колебаний относительно гамильтоновых возмущений. Понимаемая буквально, эта гипотеза неверна. Максимально возможное продвижение на этом пути, по-видимому, содержит теорема Арнольда (1961). По его собственным словам она «есть метрический аналог «теоремы» Лапласа об устойчивости планетной системы». И далее: «... исключительное множество начальных данных хотя и имеет малую меру, все же всюду плотно связано и простирается в бесконечность. Движение при исключительных начальных условиях нужно еще исследовать; а ргоіг оно может оказаться осциллирующим или даже уходящим в бесконечность» [29].

Эти исследования завершают, в существенном, внутреннее изучение свойств математической модели и позволяют тем самым поставить кардинальный вопрос о том, насколько сама модель соответствует изучаемому объекту — реальной планетной системе.

Точка зрения автора настоящей статьи состоит в следующем.

Классическая небесная механика правильно моделирует поведение реальной планетной системы на малых масштабах времени (сотни тысяч — миллионы лет).

Однако во всех эволюционных вопросах (миллиарды лет) и особенно в вопросе о структуре планетной системы гамильтоново приближение недостаточно и дает искаженное представление о свойствах реальной Солнечной системы.

Консервативные и диссипативные факторы

Некритическое перенесение свойств математической модели на свойства реального объекта является методологической ошибкой. На эту ошибку, в частности (насколько известно автору), указывал Колмогоров как раз по поводу связи его гипотезы с поведением реальной планетной системы на больших временах.

Это общее соображение допускает в интересующем нас случае более точную формулировку.

Рассмотрим систему уравнений с двумя малыми параметрами ε и δ :

$$\begin{cases} \frac{dp}{dt} = -\frac{\partial \mathcal{H}}{\partial q} + \delta A \\ \frac{dq}{dt} = \frac{\partial \mathcal{H}}{\partial p} + \delta B, \end{cases} \quad (4)$$

гамильтонову в главном члене и «дважды возмущенную». Одно возмущение — чисто гамильтоново, консервативное — задается малой поправкой εh к основному гамильтониану H невозмущенной системы

$$\mathcal{H}(p, q) = H(p, q) + \varepsilon h(p, q). \quad (5)$$

Другое возмущение, значительно меньшее по абсолютной величине

$$\delta \ll \varepsilon \ll 1, \quad (6)$$

задает неконсервативные поправки δA и δB , описывающие как потерю энергии (или массы) системой, так и подкачку их в систему.

Отбросим сначала сверхмалые ($\delta \ll \varepsilon$) диссипативные факторы. В математической модели это равносильно предельному переходу $\delta \rightarrow 0$. Получится новая, полностью гамильтонова система:

$$\frac{dx}{dt} = -\frac{\partial \mathcal{H}}{\partial y} \quad \frac{dy}{dt} = \frac{\partial \mathcal{H}}{\partial x} \quad \mathcal{H}(x, y) = H(x, y) + \varepsilon h(x, y). \quad (7)$$

Предметом теории возмущений гамильтоновых систем являются именно такие системы. Ясно, что существует немало вопросов, в которых законно пренебрежение диссипативными факторами. К ним, несомненно, относятся эфемеридные задачи. К сожалению, наиболее содержательные (и трудные) результаты, касающиеся поведения гамильтоновых систем при $t \rightarrow \infty$, как раз и не могут быть перенесены на системы общего вида. Нетрудно показать, что система (7) дает хорошую аппроксимацию поведения общей системы (4) на «предрелаксационных» временах

$$t \ll \frac{t_0}{\delta} \quad (8)$$

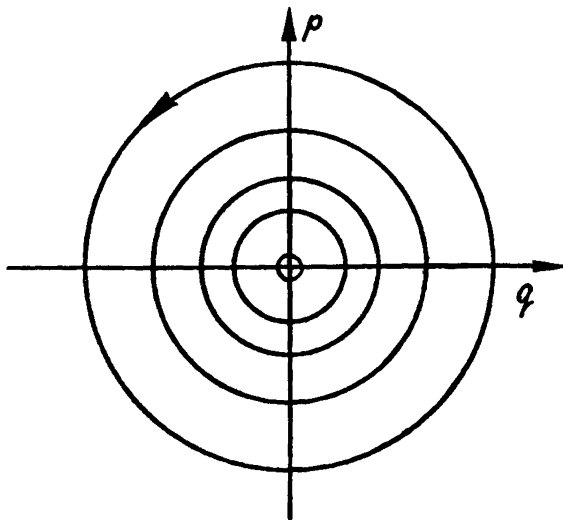


Рис. 1. Линии уровня (окружности) полной энергии $H = \frac{p^2 + q^2}{2} = E$ гармонического осциллятора

(t_0 — основной период гамильтоновой системы, например период обращения Юпитера вокруг Солнца) и совершенно непригодна именно на больших интервалах времени

$$t \gg \frac{t_0}{\delta}. \quad (9)$$

Таким образом, разумная область применимости гамильтоновой теории возмущений ограничена и сверху и снизу:

$$\frac{t_0}{\epsilon} \lesssim t \lesssim \frac{t_0}{\delta}. \quad (10)$$

На интервалах времени, много меньших, еще не имеет смысла учитывать ϵ -возмущение, а на временах, много больших, уже нельзя пренебрегать δ -возмущением.

Пример систем с одной степенью свободы

Особенно прозрачно взаимоотношение ϵ -систем и δ -систем на плоскости (одна степень свободы). Гамильтониан \mathcal{H} является первым интегралом ϵ -системы

$$\mathcal{H}(p, q) = \text{const} \quad (11)$$

в пространстве любой размерности. Однако плоскость существенно проще. Для нее значение первого интеграла полностью определяет геометрию системы — семейство интегральных кривых

$$H(p, q) + \varepsilon h(p, q) = \text{const.} \quad (12)$$

Эта формула показывает, что на плоскости гамильтоново возмущение лишь немного искажает форму регулярной траектории.

Изображающая точка подобной системы будет вечно двигаться по замкнутой траектории (12), мало отличающейся от замкнутой траектории

$$H(p, q) = \text{const} \quad (13)$$

невозмущенной системы.

Совершенно иначе ведут себя траектории δ -систем. Главной отличительной особенностью этих систем является появление избранных траекторий — предельных циклов, на которые наматываются (как снаружи, так и изнутри) все остальные траектории. Сильно модернизируя оригинальные исследования Пуанкаре (1902) и следуя идеям Ван-дерПоля (1927), напомним уравнение для гамильтониана в силу полной системы

$$\frac{d\mathcal{H}}{dt} = \delta \left(A \frac{\partial \mathcal{H}}{\partial p} + B \frac{\partial \mathcal{H}}{\partial q} \right). \quad (14)$$

Отбрасывая в правой части члены, осциллирующие при движении по невозмущенной траектории, получим новое уравнение:

$$\frac{dE}{dt} = \delta f(E), \quad (15)$$

правая часть которого есть просто среднее от правой части точного уравнения по быстрой переменной (фазе). Крылов и Боголюбов (1937) показали, что переход от уравнения для \mathcal{H} к уравнению для E может быть осуществлен при помощи замены переменных

$$\mathcal{H} = E + \delta F(E, \varphi, \delta), \quad (16)$$

исключающей быструю фазу из медленного уравнения для \mathcal{H} . «Эволюционное» уравнение для E определяет амплитуды (точнее «энергии») предельных циклов (устойчивых и неустойчивых) полной δ -системы:

$$f(E) = 0. \quad (17)$$

Они совпадают со стационарными точками эволюционного уравнения.

Экспериментальный подход

Несмотря на столь явно качественное различие в асимптотической структуре решений, обнаружить разницу в поведении δ -и гамильтоновой систем крайне трудно, и, вероятнее всего, даже невозможно. Это можно понять, если ориентировочно оценить характерные времена в реальной Солнечной системе.

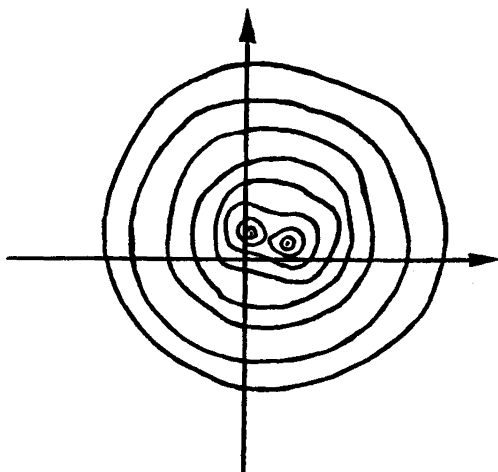


Рис. 2. Возможное искажение траекторий при гамильтоновом возмущении. Вблизи положения равновесия картина может меняться качественно, а не только количественно

Основной временной масштаб Солнечной системы задается главной колебательной компонентой — обращением Юпитера вокруг Солнца. Для наших целей достаточны оценки порядка величин. Полагаем поэтому

$$T \sim 10 \text{ лет.} \quad (18)$$

Основной малый параметр Солнечной системы, характеризующий взаимодействие планет друг с другом, имеет порядок отношения массы всех планет к массе Солнца, то есть величину порядка одной тысячной. Следовательно,

$$T_e \sim 10^4 \text{ лет.} \quad (19)$$

Диссипативные факторы оцениваются очень ненадежно. Ясно только, что речь идет о величинах (безразмерных) порядка одной миллионной или даже меньше. Поэтому

$$T_\delta \sim 10^7 \div 10^9 \text{ лет.} \quad (20)$$

Все время существования Солнечной системы оценивается (Уиппл, 1948) величиной в несколько миллиардов лет

$$T \sim 5 \cdot 10^9 \text{ лет.} \quad (21)$$

Из независимых источников (геологических, исследований метеоритов, космических) известно [34], что диссипативные факторы отнюдь не были столь малыми в ту эпоху, так от нас

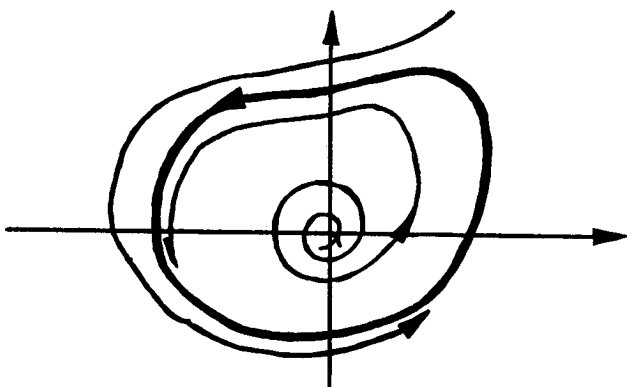


Рис. 3. Возникновение предельного цикла при неконсервативном возмущении. Роль δ -факторов.

отдаленную. Подобные соображения подтверждают разумность оценки для T_δ , во всяком случае ее верхнего предела.

Столь грандиозные различия во временах T_ε и T_δ объясняют трудность наблюдательного «уловления» диссипативных факторов. Сколько-нибудь надежные наблюдательные данные имеются (Субботин, 1937) на космогонически ничтожном отрезке времени — менее двух десятков оборотов Юпитера вокруг Солнца. Диссипативный сдвиг элементов любой орбиты лежит поэтому за пределами наблюдательной точности.

Тем не менее имеется принципиальное отличие ε - и δ -систем, позволяющее, по мнению автора, определенно предпочесть δ -модель для понимания современной структуры Солнечной системы. Главный довод — обилие резонансных явлений и процессов в реальной Солнечной системе. Резонансы непонятны и противоестественны («случайны», расположены на «множестве меры нуль») в гамильтоновых системах. В диссипативных же системах резонансы не только естественны, но и неизбежны (Molchanov, 1968).

Гипотеза Колмогорова и теорема Арнольда

В настоящее время не существует (на математическом уровне строгости) теории резонансов в нелинейных колебательных системах. Это однако, не означает, что невозможно исследование резонансных явлений «на физическом уровне строгости». При таком подходе отбрасывают [27] мешающие члены (уверяя себя

и других, что это делается из физических соображений*, и что отброшенные члены несущественны), после чего оставшиеся изучают.

На следующем этапе удается обычно обнаружить надлежащий малый параметр и построить формальное разложение, главный член которого совпадает с «физической» моделью. Разложение, как правило, оказывается расходящимся.

С этого момента задача становится чисто математической. Обоснование асимптотического разложения состоит в построении итерационного процесса, обладающего двумя свойствами:

1) сходимости при некоторых условиях (например, на множестве достаточно большой меры);

2) совпадении асимптотического ряда с формальным разложением итерации. («Достаточно большого» отрезка разложения, «достаточно далекой» итерации с «достаточно большим» отрезком асимптотического ряда).

История науки показала [28] высокую эффективность подобно-го «эстафетного» метода исследований. Нет оснований пренебрегать им и в интересующей нас задаче. Построение асимптотического разложения, не говоря уже об итерационном процессе, выходит за рамки статьи. Но даже определение главного члена и его исследование — вполне содержательная и непростая задача.

Выкладки удобнее проводить, выбрав в качестве независимых переменных первые интегралы и фазы невозмущенной системы:

$$\begin{cases} J = J(p, q) \\ \varphi = \varphi(p, q). \end{cases} \quad (22)$$

Во вполне интегрируемых гамильтоновых системах можно [29] выбрать переменные так, что в надлежащей системе координат

$$\begin{cases} J = p \\ \varphi = q. \end{cases} \quad (23)$$

В этих переменных система приобретает вид:

$$\begin{aligned} \frac{d\varphi}{dt} &= \omega(J) + \varepsilon\omega_1(J, \varphi) + \delta\omega_2(J, \varphi); \\ \frac{dJ}{dt} &= \varepsilon f_1(J, \varphi) + \delta f_2(J, \varphi), \end{aligned} \quad (24)$$

Здесь $\varphi = (\varphi_1, \dots, \varphi_l)$ — быстрые (или фазовые) переменные, а $J = (J_1, \dots, J_m)$ — медленные переменные. Правые части периодич-

* Этот «блистательный» пример физической «интуиции» обнаружен И.М.Гельфандом. Математик был бы вынужден довольствоваться «чисто формальным» соображением, что квадратное уравнение $p^2 - p = 0$ имеет только два корня $p=1, p=0$.

ны по каждой фазе с периодом 2π . Весь дальнейший анализ основан только на выписанной форме системы и не зависит от ее происхождения. В частности, при $\delta=0$ система необязательно обращается в гамильтонову. Однако различие между f_1 и f_2 следует сохранить. Оно состоит в том, что среднее по фазам от f_1 равно нулю

$$\int_0^{2\pi} \dots \int_0^{2\pi} f_1(J; \varphi_1, \dots, \varphi_l) d\varphi_1 \dots d\varphi_l = 0. \quad (25)$$

В гамильтоновых системах это условие выполнено автоматически. Рассмотрим невозмущенную ($\varepsilon=0$, $\delta=0$) систему

$$\begin{cases} \frac{d\varphi}{dt} = \omega(J) \\ \frac{dJ}{dt} = 0. \end{cases} \quad (26)$$

Медленные переменные «остановились»

$$J_k(t) = \text{const}, \quad (27)$$

а быстрые совершают условно-периодическое движение

$$\varphi_i(t) = \omega_i t + \varphi_{i0}. \quad (28)$$

Каждой точке $J = \{J_k\}$ m -мерного пространства медленных переменных J соответствует l -мерный тор пространства быстрых переменных φ , движение по которому происходит со скоростями, определенными вектором частот

$$\omega = \{\omega_i\} = \omega(J). \quad (29)$$

Гипотеза Колмогорова, доказанная Арнольдом в 1963 г., состоит в том, что гамильтоново возмущение (члены с ε) немного искажает этот тор, не меняя существенно характера движения. Однако доказательство Арнольда справедливо только для точек общего положения в пространстве J . Это, в частности, означает, что из рассмотрения исключаются все точки, лежащие на резонансных поверхностях пространства J . Каждая такая поверхность определяется целочисленным вектором резонанса (n_1, n_2, \dots, n_l) :

$$(n, \omega) \equiv n_1 \omega_1(J) + n_2 \omega_2(J) + \dots + n_l \omega_l(J) = 0. \quad (30)$$

Покажем, что вблизи резонансной поверхности возникает резонансная зона, внутри которой движение (даже при чисто гамильтоновом возмущении) существенно отличается от движения, предсказываемого гипотезой Колмогорова.

«Полубыстрые» переменные

Введем новые переменные: резонансную фазу ψ :

$$\Psi = n_1 \varphi_1 + \dots + n_l \varphi_l \quad (31)$$

и резонансную частоту ν :

$$\nu = n_1 \omega_1(J) + \dots + n_l \omega_l(J). \quad (32)$$

Эти переменные можно дополнить [32] новыми быстрыми фазами θ и новыми медленными переменными K так, чтобы возникла замена переменных, в результате которой система приобретает вид:

$$\begin{aligned} \frac{d\theta}{dt} &= \omega(\nu, k) + \varepsilon \omega_1(\theta, \psi, \nu, k) + \delta \omega_2(\theta, \psi, \nu, k); \\ \frac{d\psi}{dt} &= \nu + \varepsilon \nu_1(\theta, \psi, \nu, k) + \delta \nu_2(\theta, \psi, \nu, k); \\ \frac{d\nu}{dt} &= \varepsilon g_1(\theta, \psi, \nu, k) + \delta g_2(\theta, \psi, \nu, k); \\ \frac{dk}{dt} &= \varepsilon L_1(\theta, \psi, \nu, k) + \delta L_2(\theta, \psi, \nu, k). \end{aligned} \quad (33)$$

Изучаемая резонансная поверхность имеет в новых переменных простое уравнение

$$\nu = 0. \quad (34)$$

Поэтому как раз вблизи резонансной поверхности резонансная фаза ψ перестает быть быстрой переменной. Вид системы (33) подсказывает идею «выравнивающей» замены переменных*

$$\mu = \frac{1}{\sqrt{\varepsilon}} \nu, \quad (35)$$

которая «выщепляет» пару «полубыстрых» переменных ψ и μ . Вдали от резонанса эта пара распадается на быструю фазу ψ и медленную частоту ν .

Это преобразование в свою очередь, обнаруживает существование «полубыстрого» времени τ

$$\tau = \sqrt{\varepsilon} t, \quad (36)$$

характерного для резонансных эффектов.

В результате в системе выделяются три группы уравнений. Группа из $(l-1)$ -уравнения для быстрых фаз θ , основная группа

* Важно отметить, что *любые* варианты разложений в ряды по малому параметру не могут «выловить» резонансную зону. Это видно из того, что в формулы входит корень из ε . Следовательно (с формальной точки зрения), мы имеем дело с «ветвлением» по комплексному переменному ε .

из двух уравнений для резонансных переменных ψ и μ и наконец, группа из $(m-1)$ -уравнений для медленных переменных k :

$$\begin{aligned}\frac{d\theta}{d\tau} &= \frac{1}{\sqrt{\varepsilon}}\omega(k, \sqrt{\varepsilon}\mu) + \sqrt{\varepsilon}\omega_1 + \frac{\delta}{\sqrt{\varepsilon}}\omega_2; \\ \frac{d\psi}{dt} &= \mu + \sqrt{\varepsilon}v_1 + \frac{\delta}{\sqrt{\varepsilon}}v_2; \\ \frac{d\mu}{dt} &= g_1 + \frac{\delta}{\sqrt{\varepsilon}}g_2; \\ \frac{dk}{dt} &= \sqrt{\varepsilon}G_1 + \frac{\delta}{\sqrt{\varepsilon}}G_2\end{aligned}\tag{37}$$

Исключение быстрых переменных

Следующий шаг — главный во всей программе исследования резонансов. Он состоит в исключении быстрых переменных θ из системы и переходе к изучению эволюционной системы для медленных переменных K и резонансных ψ , μ :

$$\begin{aligned}\frac{d\psi}{d\tau} &= \mu + \sqrt{\varepsilon}a_1(\psi, k) + \frac{\delta}{\sqrt{\varepsilon}}a_2(\psi, \mu, k) + \dots \\ \frac{d\mu}{d\tau} &= b(\psi, k) + \sqrt{\varepsilon}b_1(\psi, k), \mu + \frac{\delta}{\sqrt{\varepsilon}}b_2(\psi, \mu, k) + \dots \\ \frac{dk}{d\tau} &= \sqrt{\varepsilon}F_1(\psi, k) + \frac{\delta}{\sqrt{\varepsilon}}F_2(\psi, \mu, k) + \dots\end{aligned}\tag{38}$$

Именно этот решающий переход и невозможен в настоящее время на математическом уровне строгости. В свое время автор попытался сформулировать теорему о разделении быстрых и медленных движений. К сожалению, формулировка теоремы (Молчанов, 1963) ошибочна. Автор грустно благодарен Арнольду [33], заметившему это обстоятельство. Можно предположить, что разделение движений все же имеет место, но не всюду, а на множестве асимптотически (при $\varepsilon \rightarrow 0$) полной меры. Однако в настоящее время не существует даже точной формулировки, не говоря уже о доказательстве подобной теоремы. Удовольствуемся поэтому переходом к эволюционной системе на «физическом уровне строгости». Некоторым утешением могут служить следующие соображения: во-первых, поставлена хорошая математическая задача; во-вторых, «не пропадет наш скорбный труд», так как будущее обоснование все равно будет опираться на свойства эволюционной системы; в-третьих, подобные ситуации уже случались. Так, например, Ван-дер-Поль (1927) изучал хорошие физические задачи еще необоснованным строгим методом.

Все дальнейшие выводы о свойствах общих колебательных систем имеют поэтому эвристический смысл. Строгими эти исследования станут только после строгого доказательства теоремы о замене переменных.

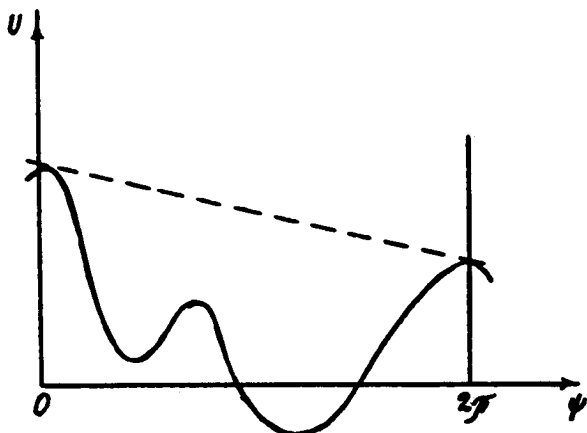


Рис. 4. Обобщенный потенциал U резонансного движения. Функция U оказывается периодичной по резонансной фазе ψ , если исходная система — гамильтонова. В общем случае U есть сумма линейной и периодической функций.

Однако в частных случаях переход к эволюционной системе может быть обоснован частными методами. Так, например, быстрые переменные могут «по своей доброй воле» не войти в эволюционную систему и тем самым избавят нас от забот. Необходимость уточнения гипотезы Колмогорова вытекает, к слову сказать, из рассмотрения как раз такого примера:

$$H(p_1, p_2, q_1, q_2) = \frac{p_1^2 + p_2^2}{2} + \varepsilon W(q_1 - q_2). \quad (39)$$

Гамильтониан этой системы с двумя степенями свободы определяется потенциалом взаимодействия W , который является произвольной функцией, четной и периодической, своего аргумента — разности q_1 и q_2 . Нетрудно проверить, что для этой системы переход к новым переменным полностью решает вопрос. Эволюционные уравнения не содержат быстрой фазы. Быстрая же фаза может быть потом «доинтегрирована» на каждой эволюционной траектории.

«О механицизме»

Рассмотрения этого пункта носят, как уже подчеркивалось, эвристический характер и не претендуют на какую-либо строгость. Полезно, однако, наметить план исследования и угадать общие контуры результата перед тем, как приступить (в других, конечно, работах) к трудным и громоздким точным доказательствам.

Итак, исследуем систему (38), положив, конечно, $\varepsilon = 0$ и тем более $\delta = 0$:

$$\begin{cases} \frac{d\psi}{d\tau} = \mu, \\ \frac{d\mu}{d\tau} = b(\psi, k), \\ \frac{dk}{d\tau} = 0. \end{cases} \quad (40)$$

Полученная система обладает весьма примечательными свойствами.

Во-первых, на рассматриваемых временах

$$\tau \sim 1, \quad t \sim \frac{1}{\sqrt{\varepsilon}} \ll \frac{1}{\varepsilon} \quad (41)$$

все медленные переменные K являются первыми интегралами:

$$K = K_0 = \text{const.} \quad (42)$$

Во-вторых, пара «полубыстрых» переменных μ и ψ удовлетворяет уравнению «механического» типа, что проявляется наиболее отчетливо, если исключить μ

$$\frac{d^2\psi}{dt^2} - b(\psi, k) = 0. \quad (43)$$

Из этого факта вытекает вывод высокой общности и значения.

Структура резонансной зоны всегда описывается ньютоновской механикой независимо от происхождения системы.

Исходная система могла быть биохимической или описывать поведение какого-нибудь экономического объекта, могла возникнуть при изучении поведения плазмы в фотосфере звезды или в задаче популяционной генетики. Все это не имеет значения. Если только в системе имеются быстрые и медленные переменные, если только в системе имеются многочастотные колебания, кинетика самых интересных, самых, если можно так выразиться, «интимных» процессов неминуемо оказывается механической по типу протекания во времени.

Более того, она оказывается ньютоновской в самом узком смысле этого слова. Уравнение (43) формально совпадает с одномерным движением материальной точки в потенциальном поле, задаваемым потенциалом $U(\psi, K)$:

$$U = - \int b(\psi, k) d\psi. \quad (44)$$

Эти рассуждения можно воспринимать как объяснение столь долгого господства «механицизма». Необходимо, однако, подчеркнуть и решающее отличие развиваемого подхода от традиционно «механицизма».

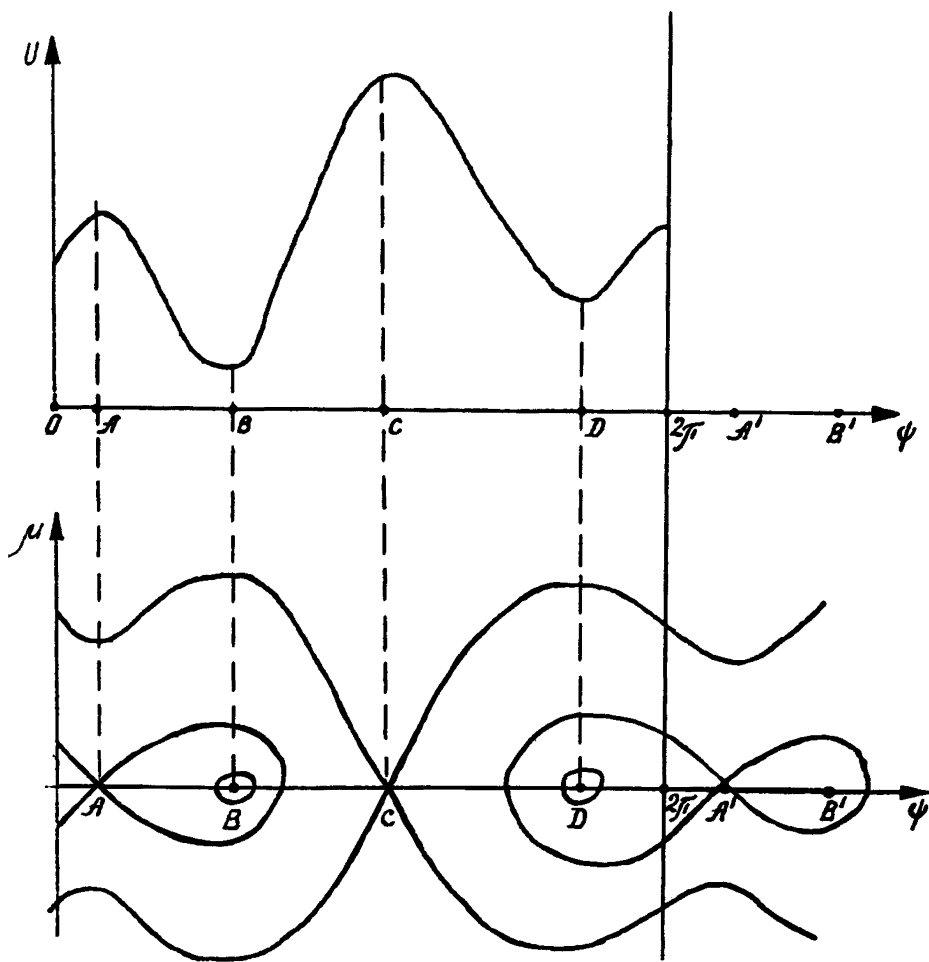


Рис. 5. Стандартное построение фазового портрета механической системы, адекватной изучаемому резонансу. Точки A', B' физически (или «биохимически») совпадают с точками A, B , соответственно, и приведены только для наглядности построения. Истинный фазовый портрет расположен на цилиндре, получающемся сворачиванием полосы шириной 2π , вырезанной из плоскости (μ, ψ)

Прежде всего, этот «механизм» «кинетический». Обычно под «механизмом» понимают «механизм» «морфологический» — утверждение, что все сущее устроено из рычагов, пружинок или, в крайнем случае, является упругой средой (эфиром) с удивительными механическими свойствами.

Здесь же утверждается нечто противоположное и, на взгляд автора, существенно более глубокое.

Независимо от исходной структуры, от морфологии механическими будут исторические, временные, эволюционные свойства любой системы — ее «поведение», кинетика, кинематика в наиболее интересных критических поворотных пунктах ее развития — в резонансных ситуациях.

Эта идея перекликается с мечтой А. Эддингтона, высказанной около полувека назад: «Задача науки — объяснить основные свойства мира не из того, что он устроен именно таким образом, а из того, что он устроен хоть как-нибудь».

Мы видим, что в одном частном случае эта любопытная программа оказывается реальной. Кроме того, приведенные соображения существенно повышают ценность и значение идей и методов, накопленных в механике и математике (теории колебаний).

Эти исследования оказываются относящимися не только к сравнительно узкому кругу чисто механических систем.

При надлежащем анализе оказывается возможным в значительно более широком классе систем (в том числе в биологических) выделить ведущие, существенные переменные. После этого замечательным образом оказывается, что соответствующие математические модели уже давно и детально изучены и пыльные архивы механических журналов обретают новую жизнь. В свою очередь, математическое моделирование в биологии оказывается основанным на прочном, надежном фундаменте математического естествознания.

Таков полезный методологический урок, который можно извлечь из сравнительно узкой темы — исследования резонансных движений.

Литература

1. *Newton Isaac. Opera quae exstant omnia*, v. 1—5. Liondini, 1779—1885. Рус. пер.: Математические начала натуральной философии. Пер. с лат. с прим. и поясн. А.Н. Крылова. В кн.: «А.Н.Крылов. Собр. трудов», т. 7. М.-Л., Изд-во АН СССР, 1936.
2. *Tisserand F. Traite de mecanique ce'leste*, v. 1. Paris, 1866.
3. *Субботин М.Ф.* Курс небесной механики, т. 2. М.-Л., 1937.
4. *Пугачик И.И.* Малые планеты. М., Гостехиздат, 1953.
5. *Самойлова-Яхонтова Н.С.* Успехи астрономических наук, 5 (1950).
6. *Ватсон Ф.* Между планетами. Пер. с англ. Б.Ю. Левина. М.-Л. Гостехиздат, 1947, с. 23.
7. *Уинпл Ф.* Земля, Луна и планеты. М.-Л. Гостехиздат, 1948, с. 35, 230.
8. *Tisserand F., Loco Cit.*, v. IV, p. 83.
9. *Darwin G.H. Scientific papers*, v. 1—5. Cambridge, 1907—1916. Рус. пер.: Приливы и родственные им явления в Солнечной системе. Пер. с англ. В.В. Серафимовича. М.-Пг., Госиздат, 1923, с. 282.
10. *Ляпунов А.М.* Общая задача об устойчивости движения. Собр. соч., т. 2. М.-Л., 1956.
11. *Козьма Прутков.* В кн.: «Поли. собр. соч.» (Б-ка поэта. Большая серия). М., Изд-во АН СССР, 1954—1965.
12. *Wolf R.* Uber die elfjahrige Periode in den Sonnenflecken und erdmagnetischen Variationen. In: «Annalen der Physik und Chemie». Lpz., Bd. 193. 1862. S. 11.

- 13 Белопольский А.А. Пятна на Солнце и их движение. Магист. дисс. М., Университетск. тип., 1886.
14. Шмидт О.Ю. Происхождение Земли и планет. М., Изд-во АН СССР, 1962.
15. Молчанов А.М. ДАН СССР, 168, 2 (1966).
16. Четаев Н.Г. В сб.: «Ученые записки Казанского ун-та», 1931.
17. Backus G.E. Critique of «The Resonant Structure the Solar System» by A.Molchanov, Icarus, 11, 88 (1969).
18. Dermott S.F., On the origin of commensurabilities in the Solar System. Mon. Not.Roy. Astr. Soc., 142, 143 (1969).
19. Henon M., Icarus, 11 (1969).
20. Colombo G., Nature, 208,575, (1965).
21. Goldreich P., Nature, 209, 1078 (1966).
22. Peal S.S., Ann. Rev.Astronomic and Astrophysic, 6, 287 (1968).
23. Peal S.S., Palo Alto. Cal. USA, 1966.
24. Субботин М.Ф. В кн.: «Большая Советская Энциклопедия», т. 2, 2-е изд. М., Госнаучиздат, 1955, с. 328.
25. Парийский Н.Н., Астрономич. ж., 22, 2 (1945).
26. Идельсон Н.И. В кн.: «Астрономический ежегодник СССР на 1941 год». М.-Л., Изд-во АН СССР, 1940.
27. Паули В. Общие принципы волновой механики. М., Гостехиздат, 1947, стр. 121.
28. Харди Г.Г. Расходящиеся ряды. Пер. с англ. Д.А.Райкова. С предисл. и обз. ст. С.Б. Стечкина. М., ИЛ., 1951.
29. Арнольд В.И., Успехи математ. наук, 18, 6, 91 (1963).
30. Борн М. Лекции по атомной механике. т. 1, Харьков-Киев, Гостехиздат УССР, 1934.
31. Биркгоф Д.Д. Динамические системы. Пер. с англ. Е.М. Левинсона. М.-Л., Гостехиздат, 1941
32. Molchanov A.M., Icarus, 8, 203 (1968).
33. Арнольд В.И., ДАН СССР, 161, 1, 9 (1965).
34. Левин Б. Ю., Изв. АН СССР, сер. физика земли, 7 (1972).

ВРЕМЯ И ЭВОЛЮЦИЯ

Введение. Кто-то (англичанин*, судя по стилю) сказал: «Когда бог создавал время, он создал его достаточно». Другой умный англичанин, Чарлз Дарвин, тоже достаточно качественно, но весьма эмоционально описывает ощущения, вызываемые у него геологическими масштабами времени: «...the mind iz stupefied in thinking over the long, absolutely necessary, lapse of years». (Цитировано по изданию Collier, New York, 1909, p.185). Не менее выразительно другое место: «...yet we must confess that it makes the head almost giddy to reflect on the years, century after century...» (loco cit., p.196).

Поучительно представить себе количественно масштаб времени, вызвавший у Дарвина столь сильные душевные движения. Оба отрывка касаются сроков формирования долины реки Санта-Крус в Патагонии. Сам Дарвин пишет, что это случилось «значительно позже образования подстилающих слоев с третичными раковинами». По современной оценке речь идет, таким образом,

* Русский вариант — «У бога дней много»

о промежутке времени в несколько десятков миллионов лет. Срок довольно скромный по сравнению с возрастом Земли, который по крайней мере во сто раз больше.

Что же так взволновало Дарвина? И почему мы, по прошествии всего лишь одного из дарвиновской чреды веков, спокойно рассуждаем о значительно больших временах? И что лучше — волнение Дарвина или наше спокойствие?

Ввиду очевидной риторичности последнего вопроса отвечать на него не надо. А вот первые два стоит обсудить.

Катастрофы или эволюция? Во всей книге Дарвина пламенеет не остывший еще накал борьбы с катастрофизмом. С теорией катастроф в нынешних популярных (да и не только популярных) книгах разделяются очень легко, попросту объявляя ее антинаучной. Однако реальному Дарвину, а особенно его другу и предшественнику Ляйеллю приходилось намного трудней. И главная трудность состояла в том, что науке того времени возраст Земли был неизвестен.

Характерно, что Дарвин, точный, почти педантичный во всем, что касается размеров, количества и даже цен (на одной только страничке 383 приведено более десятка разнообразных чисел), становится описательным, качественным и эмоциональным, как только речь заходит об оценке времени.

Едва ли не единственное количественное высказывание — «...like unto a geologist who had lived his ten thousand years...» (loco cit., p.508) — является лишь метафорой, но метафорой заставляющей призадуматься. Неужели даже для Дарвина количественным ориентиром были библейские семь тысяч лет?

Если это так, то многое становится понятным. Понятно, например, почему Ньютон, живший полутора столетиями ранее, вынужден был считать доказательством божественного творения Солнечной системы высокую симметрию ее строения. На естественное происхождение попросту не хватало времени. Именно отсюда возник ньютоновский «первотолчок», за который ему впоследствии так досталось от Энгельса.

Катастрофы Кювье проистекают, по-видимому, из того же источника. Не удавалось «уложиться» в сроки, которыми естествоиспытатели сознательно или бессознательно ограничивали естественную историю.

Логически равноправны были два выхода — или серия катастроф (которую к тому же подсказывала библейская штурмовщина шести дней творения), или радикальный пересмотр вопроса о времени. Историю Земли необходимо было, как мы теперь знаем, удлинить в миллион раз.

Деятельность Прокруста — жалкий дилетантизм по сравнению с тем, что предстояло науке. Вполне понятно поэтому, что Чарлз Ляйелль и Чарлз Дарвин — люди, больше других сделавшие для утверждения идеи эволюции — меньше всего рисковали количественными высказываниями. Собранные и осмысленные ими бога-

тейшие данные, прежде всего геологические, прямо-таки вынуждали идею огромности прошедшего времени. Однако количественные методы могли возникнуть только после и в результате того психологического перелома, который они создавали.

Большие числа. «Почему атомы такие маленькие?» — спрашивает Эрвин Шредингер в книге «Что такое жизнь с точки зрения физики?» и отвечает — «Это потому, что мы такие большие».

Сложная система (человек) неминуемо содержит громадное число простых атомов — вот главный смысл высказывания Шредингера. Однако есть и вторая сторона сопоставления «человек — атом». Нельзя безотносительно говорить «много», «мало». Всякое количественное высказывание есть высказывание об однотипном. На математическом жаргоне — параметры должны быть безразмерными. Шредингер, следовательно, подчеркивает, что всякий малый параметр есть оборотная сторона большого параметра (более точно — обратная величина). Однако психологически безразлична форма вопроса. Более того, она имеет немалое методологическое значение. Почему мы большие — более глубокий вопрос, так как сразу возникает другой. А не могли ли мы быть меньше? И насколько? И выясняется, что безобидная, казалось бы, перифраза кардинально меняет проблему. Из простого любопытства вырастает глубокая задача — каковы минимальные, критические значения объема, массы, энергии, времени и т.д., при которых может возникать жизнь.

Дарвин, в сущности, всего лишь элегичен, говоря о чреде веков. Это слегка завуалированный вздох о краткости жизни человеческой и почтительное удивление возрастом даже такой геологически обыденной вещи, как долина реки. Но при чем здесь человек с его веком? Грустно, конечно, что век такой маленький (по сравнению с чем?). Но ведь человек только стоит, смотрит и старается понять. Создавал же долину не он, а ежедневные приливы и отливы морского рукава, ставшего потом рекой Санта-Крус. Или, в нашей стороне, ежегодные «разливы рек ее, подобные морям». Человеческие масштабы здесь категория внешняя, да и астрономическое измерение времени существовавши лишь в той мере, в какой оно связано с «единичным актом творения».

Поучительно оставить наедине главные действующие лица — базальтовый массив и приливную волну. Массиву этому добрый десяток миллионов лет, и каждый день каждого года из этих миллионов лет «о скалы грозные дробятся с ревом волны и белой пеною, крутясь, бегут назад». Итого миллиард раз.

Все-таки очень легко стали мы произносить «десять в девятой степени». Писать, конечно, проще, а вот удивляться и, для пользы дела, ужасаться — разучились. Дарвин еще умел, и у него это неплохо получалось.

«В одном мгновеньи видеть вечность». Самый большой отрезок времени, сколько-нибудь поддающийся сегодня количественной

оценке — возраст нашей Галактики — составляет 10^{13} лет. Ошибка на пару порядков здесь и далее особого значения не имеет.

Самый малый «кусочек» времени, о котором имеет смысл говорить при современном уровне знаний, составляет 10^{-23} сек. Это время, за которое свет прошел бы «от одного края электрона до другого», если бы написанная фраза имела смысл. К счастью, для наших целей достаточно, чтобы была разумной оценка времени, получающаяся в результате деления классического радиуса электрона, составляющего 10^{-13} см, на скорость света — 3×10^{10} см/сек. Какое именно истолкование этот «квант времени» получит в будущей теории элементарных частиц, в данный момент несущественно.

Единственная твердая опора посреди этих двух зыбких чисел, промежутков, соединяющий эти две пучины времени — число секунд в году. Да и то надежность его весьма сомнительна — весьма и весьма вероятно, что длина суток, да и года меняется. Все же по данным на 1969 г., его можно вычислить по крайней мере с двумя знаками: $365 \cdot 24 \cdot 60 : 60 = 0,32 \cdot 10^8$, а не с точностью до порядка, как «обрамляющие» числа.

Эти трудно сопоставимые величины следует сравнивать в логарифмической шкале. И дело не только в удобстве — такая шкала, как мы увидим, имеет, по-видимому, глубокий структурный смысл.

Вся «Вечность» оказывается не очень большой — всего лишь сорок четыре порядка. Из них на долю космоса (от возраста Галактики до года) приходится тринадцать порядков, восемь порядков (от года до секунды) «отпущены» на человеческий, организменный, уровень, а остальные двадцать три порядка — больше половины — царство микромира.

Поэт W. Blake, цитата из которого (в переводе Маршака) вынесена в заголовок, призывает «to see Eternity in an hour». Но Час содержит львиную долю Вечности (28 порядков из 44), и поэтическая интуиция оказывается, следовательно, инструментом довольно грубым.

Естествоиспытатель Ч. Дарвин угадывает Вечность значительно раньше (уже в девяти порядках) и чувствует ее намного тоньше — не в простом течении, длительности, а в созидающей ритмике приливов. Если же кто-нибудь склонен видеть в приливах более разрушительное начало, нежели созидательное (созидание, впрочем, всегда происходит за счет разрушения), тому можно посоветовать прочитать дарвиновскую теорию происхождения атоллов — там уж созидательный характер микроритмики не подлежит сомнению.

Конечно, в приведенной форме «исчисление Вечности» имеет несколько юмористический характер, однако оно дает представление о современной шкале времени. Эта шкала может и будет,

конечно, меняться, особенно на «краях» — на космологическом «верху» и квантовом «низу». Но вся биологическая эволюция должна заключаться «между» указанными границами. На всю сложность жизни «отведено» около сорока порядков — хватит ли их? Или опять потребуются катастрофы? Таков совсем не юмористический смысл этих выкладок.

Забавно, что вечность Маршака, и здесь перевод весьма удачно не сохраняет «рабской верности оригиналу», близка к половине Вечности. Мгновенье, то есть миг, миганье века (глаза), продолжается несколько сотых секунды. За время жизни нашей Галактики протекло столько человеческих «мгновений», сколько в одном мгновеньи содержится элементарных «квантов» времени.

Человек — макрокосм, «мера всех вещей» — и количественно оказывается посредине между микромиром и космосом.

Математическое интермеццо. В математике, в нелинейной теории колебаний, есть задача, очень близкая по духу к разбираемой.

Предположим, что задана система уравнений

$$\frac{dx}{dt} = f(x, y), \quad (1)$$

$$\frac{dy}{dt} = \varepsilon g(x, y),$$

содержащая малый параметр ε . Каков смысл этого параметра? Его можно интерпретировать (очень существенно, что эта интерпретация не единственная) как отношение масштабов времени системы «иксов» и системы «игреков».

В самом деле, изменение x можно приближенно записать в виде

$$\Delta x \approx f \Delta t. \quad (2)$$

Из этого соотношения можно получить оценку времени

$$\Delta t \approx \frac{|\Delta x|}{|f|}, \quad (3)$$

необходимо для того, чтобы система x -в испытала изменение масштаба Δx . Следовательно, время существенных изменений, т.е. $\Delta x \sim 1$ имеет порядок единицы. Мы предполагаем, разумеется, что масштабы x , y и t выбраны разумно, т.е. так, чтобы величины $f(x, y)$ и $g(x, y)$ были порядка единицы. Аналогичная выкладка для y приводит к результату:

$$\Delta t \approx \frac{1}{\varepsilon} \frac{|g|}{|f|} \quad (4)$$

из которого вытекает, что для существенного изменения необходимо огромное время

$$T \sim \frac{1}{\varepsilon} \quad (5)$$

которое тем больше, чем меньше ε . Уже это обстоятельство наталкивает на мысль, что ε характеризует (и это еще одна интерпретация) малость взаимодействия между системами x и y .

Проведенный элементарный анализ выясняет главное. Система, описываемая вектором x , изменяется почти независимо от системы, описываемой вектором y . Обычно x называют «быстрыми», а y — «медленными» переменными.

Удобный формальный прием состоит в том, чтобы перейти к пределу, положив $\varepsilon = 0$.

Невозможно удержаться от замечания, что самая сущность математики состоит именно в изучении асимптотических, идеализированных, предельных, вырожденных, крайних ситуаций. Ситуаций достаточно простых, чтобы быть подвергнутыми логическому анализу, но достаточно сложных, чтобы сохранять главные черты явления, иначе анализ бессодержателен. Метод математики — графика, если угодно, карикатура на явление; гротеск, позволяющий заострить и выделить определяющее в явлении. Акварельные полутона противопоставлены математическому подходу, во всяком случае на стадии постановки вопроса.

Возвращаясь к нашей задаче, получаем систему:

$$\frac{dy}{dt} = 0$$

$$\frac{dx}{dt} = f(x, y), \quad (6)$$

откуда вытекает, что

$$y = y_0 = \text{const}, \quad (7)$$

и остается только уравнение для x

$$\frac{dx}{dt} = f(x, y_0) \quad (8)$$

Система для x даже формально зависит только от нынешнего «уровня» y_0 и не зависит от динамики этого уровня. Движение быстрых переменных происходит на постоянном уровне медленных. Мы не задумываясь говорим о высоте горы над уровнем моря, даже если находим на ее вершине раковины, свидетельствующие о том, что она была (и, может быть, будет) когда-то морским дном.

Вряд ли кто-нибудь из читателей заметил (это нечасто замечают даже профессионалы-математики), что все рассмотрение совершенно неравноправно по отношению к медленным и быстрым переменным.

Мы были явно пристрастны. Это находит формальное выражение в том, что время измеряется масштабом, характерным именно

для быстрых переменных. Собственное время быстрых переменных — период колебаний в случае периодических движений, время полураспада в случае движений релаксационных и т.д. — вот что выбрано за единицу измерения времени.

Попробуем восстановить справедливость. Перейдем к «медленному» времени:

$$\tau = \varepsilon t \quad (9)$$

В этом новом времени система приобретает вид

$$\left. \begin{aligned} \varepsilon \frac{dx}{dt} &= f(x, y) \\ \frac{dy}{dt} &= g(x, y) \end{aligned} \right\} \quad (10)$$

и предельный переход (полный анализ которого представляет трудную, до сих пор до конца не решенную математическую задачу) приводит к выводу:

$$\left. \begin{aligned} 0 &= f(x, y) \\ \frac{dy}{dt} &= g(x, y) \end{aligned} \right\} \quad (11)$$

Снова выделяется только одна система, на этот раз, конечно, медленных переменных. Быстрые переменные, как раньше медленные, выпадают из системы.

Угадывается любопытный общий вывод. Фиксирование определенного масштаба времени как бы «выключает», «замораживает» иные масштабы. Интересно, что «замирают» не только более медленные процессы, что вполне понятно, но и более быстрые, что кажется на первый взгляд парадоксальным.

Однако это так, и наглядно можно пояснить следующим примером. Допустим, что мы видим одновременно вращающийся винт самолета и обычные часы. Движение секундной стрелки на часах ясно видно — это основное переменное. Минутная и тем более часовая стрелки «стоят» на месте — это медленное движение. Но движение винта (быстрое переменное) мы тоже не видим. Вместо винта виден сверкающий но неподвижный круг. Математик в этом случае говорит об «осреднении по траектории быстрого движения».

И дело здесь вовсе не в оптических эффектах, не в психологии. Атом, например, «прозрачен» для быстрых α -частиц и непроницаем для других атомов. Хорошую классическую модель этого явления можно получить, хорошенько раскрутив велосипедное колесо. Палец в него совать не рекомендуется, но камешки, бросаемые раз за разом, почти все будут свободно пролетать сквозь «непрозрачное» для более медленных движений колесо.

Специфична ли квантованность для микромира? Конечно, «замораживание» быстрых процессов, при всем сходстве этого явления

с «замиранием» процессов медленных, часто имеет качественно иной характер.

Главное различие состоит вот в чем.

Медленные процессы мы можем «застать» в любой стадии их развития, а процессы быстрые всегда успевают проэволюционировать и стабилизироваться в устойчивом состоянии.

Формальное выражение это различие находит в том, что медленные переменные y_0 в системе (6) могут принимать произвольное значение, и мы субъективно воспринимаем это как непрерывность. Быстрые же переменные x должны удовлетворять первому из уравнений системы (11), т.е. выходить на квазиравновесный уровень быстрого движения. Словечко «квази» приписано для того, чтобы подчеркнуть зависимость этого уровня от медленного переменного y и особенно подчеркнуть возможность исчезновения устойчивости этого равновесия при дальнейшей эволюции y . Возникает дискретность возможных состояний быстрых переменных x .

С этой точки зрения закономерен тот исторический факт, что дискретность особенно отчетливо и принудительно была понята впервые именно в микромире в форме квантованности физических величин. Однако это всего лишь частное проявление значительно более общей закономерности*.

«Финальные» состояния (квазиравновесные, метастабильные, устойчивые), очевидно, дискретны. Именно поэтому эволюционно зрелые системы обнаруживают всегда тенденцию к дискретному, иерархическому строению. Биологические системы не представляют собой исключения из этого правила. Им «кисельность» (по терминологии Н. В. Тимофеева-Ресовского) противопоказана не в меньшей степени, чем квантовым объектам, и, в принципиальном отношении, по той же причине — по причине эволюционной зрелости.

Роль диссипации. Когда Фауст спрашивает Мефистофеля, кто он такой, тот отвечает весьма уклончиво: Ein Teil von jener Kraft, die stets das Bose will, und stets das Gute schafft.

Лукавому верить нельзя, конечно, ни в одном слове. Он и соврет, недорого возьмет, и обмануть честного христианина всегда рад. Однако на этот раз он сказал правду, хотя и не всю.

Современный научный вариант (и очередное перевоплощение Мефистофеля) — это диссипативные факторы: трение, неупругое столкновение, возрастная усталость материалов, убегание атмосферы, накопление мутаций и несть им числа. Все они заняты как будто самым дьявольским делом — повышают энтропию. И в

* Некоторые соображения по этому поводу содержатся в статье автора «The Resonant Structure of the Solar System» — «Icarus», 1968, v. 8, No 2. Развитие этих идей см. «Icarus», 1969, v. 11, No 1.

самом деле, если эти факторы достаточно сильны, они сделают свое черное дело полной деградации системы. Если бы, например, Земля испытывала при движении по орбите достаточно сильное трение, то она упала бы в конце концов на Солнце и сгорела бы там вместе со всей своей биосферой.

Однако, если эти факторы держать в узде, вместо злобного дьявола возникает легион полезных озорных чертенят, делающих действительно полезное дело.

Может показаться, что весь предыдущий импрессионизм является попыткой заменить, хотя бы на эмоциональном уровне, несколькими строчками десятки утомительных страниц чисто математического текста, — что диссипативные факторы при определенных условиях (необходимо, в частности, их «израсходование», убывание в процессе эволюции) имеют тенденцию стабилизировать сложные резонансные колебательные структуры.

Если диссипация велика, дискретных резонансных уровней мало (тем меньше, чем больше диссипация), структура оказывается бедной и неинтересной. Это, конечно, верно. Однако совсем без диссипации обойтись тоже нельзя — возникает «кисельность», стабилизирующее начало исчезает и система воспринимается как чисто стохастическая, становясь эквивалентной ей в определенном (статистическом) смысле этого слова.

Иерархия времен. Ранее бегло было разобрано взаимодействие всего лишь двух систем с резко различными масштабами времени. В реальных ситуациях цепочка «вложенных» друг в друга масштабов времени намного длиннее.

Так, например, в качестве исходного масштаба времени при изучении биологической системы (скажем млекопитающего) можно взять время единичного акта катализа. Для наиболее быстрых ферментов эта величина порядка 10^{-4} и даже 10^{-6} сек. Результатом длинной последовательности таких актов может явиться, например, элементарный этап мышечного сокращения, продолжающийся около 10^{-2} сек. За следующий масштаб примем одно биение сердца — одна секунда. Достаточно сложный акт условно-рефлекторной деятельности продолжается время, измеряемое минутами. Пищевой рефлекс характеризуется суточным ритмом. Достаточно ясно, как продолжать эту цепочку или «вставлять» в нее пропущенные звенья.

Однако проведенный выше анализ показывает, что нет надобности рассматривать (по крайней мере в первом приближении) сразу всю иерархию масштабов и соответствующих им систем. Факт «замирания», «выключения» и высших и низших масштабов времени имеет первостепенное методологическое значение. Он позволяет сводить вопрос к изучению только одной ступеньки, содержащей два соседних масштаба времени. Это весьма принципиальный шаг — выделение «кванта» эволюции, ее структурного элемента.

Задача распадается тем самым на две совершенно разные по характеру и стилю задачи.

Изучение взаимодействия двух соседних по масштабам систем неминуемо предполагает полное и последовательное кинетическое рассмотрение и классификацию возможных типов поведения. Даже при беглом взгляде ясно, что процесс на каждом уровне может быть по крайней мере трех типов — стабилизирующийся, колебательный и нарастающий. Так как уровней два, то минимальное число возможностей — девять. Реально их значительно больше, так как, например, колебательные режимы могут быть периодическими (одночастотными) и многочастотными — каждый со своими качественными особенностями.

Совсем другой, если можно так выразиться, кибернетический, системный, классификационный характер имеет задача описания всей цепочки в целом. Здесь достаточно составления «списка» типов связей соседних звеньев. Кинетические же вопросы следует считать включенными в характеристику типа связи.

Очень важно подчеркнуть, что система, приходящая в равновесие на некотором уровне, вполне может оказаться колебательной и даже неустойчивой на больших масштабах времени. Никакой связи между типами поведения на разных уровнях, вообще говоря, нет. Так, например, газ, уравновесившийся и удовлетворяющий второму началу термодинамики, вполне может быть рабочим телом в паровой машине, совершающей механическую работу. Никакого противоречия здесь нет, так как время релаксации для газа измеряется ничтожными долями миллисекунды, а механические движения разыгрываются в масштабах секунд.

Кинетика и структура. Вопрос о поведении систем во времени, изучение кинетики имеет большое методологическое значение. Это значение определяется самим существом эволюционного учения. Если сущее не было создано актом творения в готовом виде, если оно возникало постепенно, то кинетика любой системы необходимо должна предшествовать ее структуре*. Это несомненно так в плане эволюционном. Куда более интересно и важно, что воссоздание уже раз созданной структуры идет по тем же этапам, что и ее эволюционное возникновение.

В этом общий смысл известного тезиса о повторении в индивидуальном развитии развития эволюционного.

Однако это повторение не может быть только повторением. Нельзя, чтобы рождение каждого человека снова требовало полных трех миллиардов лет эволюции. Происходит головокружи-

* См. также статью автора «Возможная роль колебательных процессов в эволюции». — Труды Всесоюзного симпозиума по колебательным процессам в биологических и химических системах. М., 1967, с. 274.

тельное, в миллиарды раз, сокращение этапов, упрощение и укорочение извилистого эволюционного пути.

Одной из важнейших теоретических задач и является осознание на всех уровнях, с точки зрения любых подходов (в том числе и математического) общих принципов такого убыстрения, при котором не теряется достигнутый структурный уровень.

Можно думать, что понимание этих принципов является не только теоретической, но, вероятно, и практической, едва ли даже не технологической задачей.

Сборник, предлагаемый вниманию читателя, содержит основные материалы IV школы по математическому моделированию. Особенность данного выпуска, соответствующая особенностям самой школы, состоит в расширении предмета, выходе за рамки только экологической тематики. Можно думать, что это отражение общей тенденции углубления экологических вопросов, стремления осознать экологические проблемы с общих биологических позиций.

В статьях С.Э.Шноля, Л.Х.Гаркави и Е.Б.Квакиной, М.М.Лященко и В.Я.Бригидиной содержится «социальный» запрос к математикам. В них указаны «горячие» точки биологии в таких важных ее разделах, как теория биологической эволюции и теория неспецифической устойчивости организма.

При всем разнообразии собственного биологического содержания эти работы имеют принципиальное методологическое сходство, которое отражено в названии всей школы — «Экстремальные состояния биологических систем».

Обсуждение докладов и работа секций (в частности, секции спортивной медицины) показали плодотворность идеи «критического режима» для понимания биологических процессов. В медицине подобный стиль мышления уже давно завоевал права гражданства и даже сформулирован в афористически лаконичном высказывании — «научиться понимать норму через патологию».

Однако дело не только в распространении этого приема на другие биологические системы, в том числе экологические, особенно находящиеся под сильным антропогенным, или, лучше сказать, техногенным, воздействием. Угадывается глубокая аналогия с классическими, хорошо изученными областями математического естествознания, и прежде всего с теорией устойчивости движения А. М. Ляпунова, с ее хорошо разработанным алгоритмическим аппаратом. Более того, возникает серьезная надежда формулировки достаточно общей теории экстремальных режимов на основе естественного обобщения теории устойчивости. Суть этого обобщения, как можно предполагать, состоит в рассмотрении классов систем (или систем, зависящих от параметров) и установлении границ однотипного поведения. Иными словами, усложне-

ние объекта изучения приводит к необходимости построения не только фазовых портретов, изображающих поведение системы во времени, но и структурных портретов, характеризующих поведение систем в пространстве параметров*.

Математические модели конкретных биологических систем представлены статьями В. Я. Изакова и В. В. Дынника о мышечной подвижности. В них рассматриваются вопросы электро-механического сопряжения и энергообеспечения работы при увеличивающейся до отказа (т.е. опять-таки в экстремальных условиях) нагрузке.

Изложенное объясняет почему, кроме предметно-биологической части, в материалах школы содержатся чисто математические работы. Статьи Ю. М. Апонина и Ф. С. Березовской относятся к общей теме изучения критических режимов системы.

Особую ценность выпуску придает фундаментальная работа Александра Яковлевича Хинчина «Математические основания статистической механики».

Публикация этой работы особенно актуальна потому, что попытки биологических приложений (такие, например, как «Временная организация клетки» Гудвина или серия работ по «термодинамике» популяций) основаны на буквальном приложении готовых формул, полученных для чисто гамльтоновых систем. «Втискивание» биологических систем (или хотя бы даже моделей этих систем) в узкие рамки «механистической» теории не привело (да и не могло привести) к заметным успехам. Обращение к оригинальным глубоким идеям А. Я. Хинчина (а не к более частным результатам, содержащимся в уже опубликованных книгах) открывает, на наш взгляд, более широкие возможности. Очень важно в одном сборнике иметь не только постановку биологической проблемы, но и указания наиболее вероятного пути решения, наиболее сильного математического метода. Речь, разумеется, идет о тех проблемах, для которых представляется перспективным использование методов математического моделирования.

В целом сборник характеризует состояние дел на одном из участков взаимодействия математики с биологией. Он будет полезен не только фактами или результатами, но еще более постановкой вопросов. Вероятный круг читателей: специалисты математики, биофизики, биохимики, медики и физиологи (особенно работающие на стыках этих дисциплин), а также студенты старших курсов и аспиранты соответствующих специальностей.

А. М. Молчанов

* Один из математических аспектов этой общей схемы рассматривается в так называемой теории катастроф Р. Тома.

О работе А. Я. Хинчина

В трудное военное время вышла небольшая книга Александра Яковлевича Хинчина «Математические основания статистической механики» — ныне библиографическая редкость.

Эта книга намечала глубокий синтез вероятностных и детерминистических идей в термодинамике — одной из самых неясных областей математического естествознания. Даже сейчас наука в этом направлении продвинулась мало — так труден этот синтез. На первые послевоенные годы приходится попытка (к сожалению, незавершенная) осуществления Александром Яковлевичем намеченной им программы. Наиболее полно он изложил свои взгляды в курсе лекций, прочитанных весной 1950 г. Мне посчастливилось слушать этот замечательный курс. Перед читателем лежат мои записи этих лекций. Единственное изменение — устранение чисто личного элемента. Я считаю невозможной любую другую обработку, несмотря на очевидные недостатки студенческого восприятия, — нельзя рисковать внесением сегодняшней точки зрения в единственный, как это ни грустно, источник наших знаний об этом уникальном курсе лекций.

Две причины, делают работу А. Я. Хинчина особенно актуальной именно сейчас: 1) биология вынуждает математиков выйти за ограниченные рамки только механических систем; 2) современная вычислительная техника позволяет решать задачи, казавшиеся совершенно неприступными еще тридцать лет назад.

Уточним эти замечания.

Закону Больших Чисел повезло на общественное внимание. О нем часто говорят и пишут, хотя значительно реже понимают. Еще Анри Пуанкаре писал, что математики считают его* экспериментальным фактом, а физики думают, что все это доказано математиками. В подобной ситуации необходима ясность, и заслуга А. Я. Хинчина состоит в расширении и уточнении математической основы закона больших чисел. Александр Яковлевич

* Речь шла об одном из аспектов закона больших чисел — нормальном распределении.

показал, что закон больших чисел может быть сформулирован как теорема математического анализа (из которой, к слову сказать, несложно вывести его традиционное вероятностное истолкование). Эта формулировка настолько проста и глубока, что заслуживает широкой популяризации.

Пусть в пространстве X :

$$X = (x_1, x_2, \dots, x_n)$$

состоящем из большого числа n ($n \gg 1$) одинаковых компонент x , задана* сумматорная (сумма «одинаковых» слагаемых) функции $H(x)$:

$$H(x) = h(x_1) + h(x_2) + \dots + h(x_n).$$

Тогда любая другая сумматорная функция $\mathcal{A}(x)$:

$$\mathcal{A}(x) = a(x_1) + a(x_2) + \dots + a(x_n)$$

является (асимптотически при $n \rightarrow \infty$) функцией основной H :

$$\mathcal{A}(x) \approx A(H)$$

Теорема Хинчина утверждает, следовательно, что с увеличением числа компонент некоторые свойства (далеко не все!) системы упрощаются, а не усложняются.

Отметим важные особенности этого подхода.

Элементы. Система может состоять из любых компонент — газ из молекул, галактика из звезд, орган из клеток, лес из деревьев.

Детерминизм. Схема применима не только к статистическим (вероятностным), но и чисто детерминистическим системам. Суть дела в характере задаваемых вопросов.

Нелинейность-аддитивность. Элементы системы нелинейны. Система в целом аддитивна (сумматорность!) по элементам.

Целостность. Упрощаются (при $n \rightarrow \infty$) свойства системы в целом, при сколь угодно сложном строении (или поведении) составляющих элементов.

Однако все эти достоинства могут проявиться только при фактическом построении асимптотической зависимости $\mathcal{A} = A(H)$. Из работы Хинчина читатель узнает, что для этого необходимо вычислять кратные интегралы, к тому же зависящие от параметров (а то и от нескольких). Эта задача трудна даже для современных ЭВМ, а во время Хинчина (всего лишь тридцать лет назад) эта трудность была попросту непреодолима. Можно было рассчитывать только на те редкие (хотя и очень важные) случаи — формула Планка, распределение Максвелла, когда возможно аналитическое интегрирование.

* Функция $H(x)$ должна быть в определенном смысле положительной. Точные условия сформулированы в основном тексте.

Укажем в заключение главную причину опубликования работы Хинчина именно в данном сборнике.

Переходы через критические, экстремальные состояния биологических систем во многих отношениях напоминают фазовые переходы. Сами биологические системы обычно состоят из большого числа одинаковых, существенно нелинейных элементов. Однако математические методы, используемые при изучении этих систем, все еще находятся на уровне кинетической теории газов. Неудивительно поэтому, что достижения настолько скромны, что подрывают доверие серьезных биологов к самой идее сотрудничества с математиками. Важно поэтому обратить внимание научной общественности, особенно молодежи, на перспективный, хотя и весьма трудный, путь построения и исследования математических моделей сложных биологических объектов и явлений.

Биогеоэкологическое учение было создано В. Н. Сукачевым, продолжившим выдающиеся достижения В. В. Докучаева и В. И. Вернадского в осознании роли взаимодействия живого вещества и косной материи в развитии биосферы. Принципиальная идея В. Н. Сукачева о биогеоэкологии как естественноисторическом теле, целостной совокупности биоэкологии и других абиотических элементов биосферы на определенной территории, существенно связанных материально-энергетическими отношениями, определяет и научный метод биогеоэкологии — сопряжение естественноисторического подхода и моделирования.

Оформление биогеоэкологии как науки совпало с резким подъемом общественного интереса к проблемам взаимодействия человека со средой его обитания. Человечество, прежде не ощущавшее своего влияния на природную среду и бывшее в этом смысле «свободным», стало осознать, что среда жизни не есть нечто неизменное, не зависящее от его (человечества) деятельности. Другими словами, наступил этап «несвободного» развития человечества в среде, изменение которой вызвано самим этим развитием. Это обстоятельство, как и «внечеловеческие» пространственные и временные масштабы функционирования биогеоэкологической системы, подчеркивает необходимость применения метода математического моделирования в исследованиях биогеоэкологической структуры биосферы Земли.

Следует отметить, что у истоков применения метода моделирования в биогеоэкологии стоял Алексей Андреевич Ляпунов, один из создателей отечественной школы в кибернетике и крупный ученый-натуралист. Его статьей и открывается настоящий сборник.

Первые этапы разработки подходов к моделированию процессов в биогеоэкологии начаты были на биостанции «Миассово» в Ильменском заповеднике в конце 50-х годов. Была предпринята попытка моделирования лесного биогеоэкологии по материалам многолетних работ Института леса АН СССР в Теллермановском лесничестве Воронежской области. Итогом совместного обсужде-

ния проблемы математиками и естествоиспытателями было «открытие» огромной сложности такой работы. Не менее сложными, но и многообещающими представлялись ставившиеся задачи моделирования природных систем — популяций, биогеоценозов, океана и биосферы в целом. Это было смело и рискованно в глазах строгих математиков и слишком абстрактно в глазах большинства биологов. Сегодня мы отдаем должное проницательности и широте взглядов участников этой работы (в числе которых были А. А. Ляпунов и Н. В. Тимофеев-Ресовский), закладывавших основы теоретической биологии и биогеоценологии.

Мироощущение биологов весьма рельефно описывается принципом «все» или «ничего», наиболее отчетливо сформулированным в физиологии. Это резко противоречило идее постепенности и непрерывности, которая, как казалось, пронизывает математику. Потребовалась кипучая энергия Алексея Андреевича и вся его поистине огромная эрудиция, чтобы объяснить и заставить понять — математика готова к освоению нового поля деятельности. Более того, дискретная математика уже обладает необходимым понятийным аппаратом. И если сейчас мы понимаем, что нет противоречия между непрерывностью и дискретностью, то в этом решающая заслуга Алексея Андреевича и его школы. Пожалуй, наиболее отчетливо это понимание выражено в работах И.А. Полетаева, давнего соратника и сотрудника А. А. Ляпунова. В понятии «либиховы системы» сконцентрировано главное — скачки, разрывы (обязанные своим происхождением резкой нелинейности живых систем) сосредоточены на границах, между которыми идет плавное постепенное изменение.

Публикуемый сборник призван осветить ряд современных подходов к моделированию общих и частных явлений и процессов в биогеоценозах. Он представляет собой посильный вклад в развитие научного направления, начатого А. А. Ляпуновым и его школой.

А. Н. Тюрюканов, А. М. Молчанов, В. В. Галицкий

ОГЛАВЛЕНИЕ

Возможная роль колебательных процессов в эволюции	7
Релаксационная модель адаптации	23
Управление и адаптация	35
Эндогенные биохимические колебания как возможная основа физиологических ритмов	42
Колебательные режимы как частный случай критических режимов	50
Математические модели в экологии. Роль критических режимов	53
Экология и эргодичность	62
Критические точки биологических систем	64
Экстремальные режимы	76
Механохимия	89
Стационарные режимы в биологии и математике	100
Теорема выравнивания. Приближенность формулы Больцмана. "Закон" возрастания энтропии	108
Об одной теореме А.Я.Хинчина	123
Кинетика сложных систем	132
Термодинамика и эволюция	148
Макродинамика	165
Гипотеза резонансной структуры солнечной системы	186
Время и эволюция	205

Замеченные опечатки

Страница, строка	Напечатано	Следует читать
эмблема на обложке	GARMONIA	HARMONIA
с. 11, примечание	$\frac{\partial A(x)}{\partial x}$	$\frac{dA(x)}{dx}$
с. 52, 10 строка сверху	другую менее	другую, менее
с. 62, 16 строка снизу	Климакс ассоциации	Климакс-ассоциация
13 строка снизу	-»-	-»-
5 строка снизу	-»-	-»-
4 строка снизу	-»-	-»-
с. 88, формула (25)	$x''' + \omega^2 x =$	$x''' + \omega^2 x' =$
с. 91, примечание **	неудачное	неудачно
с. 102, подпись под рис. 4	(справа)	(слева)
с. 107, 1 строка сверху	развертка *	развертка
2 строка сверху	развертка	развертка *
с. 118, формула (38)	$(\tau)^{k+2}$	τ^{k+2}
с. 134, формула (2)	$h(x_n)$	$h(x_n)$
с. 135, примечание	редкий	нередкий
с. 174, 13 строка сверху	$= I_{\alpha\beta}^{\gamma} Y_{\gamma}(x) L_{\beta}$	$= I_{\alpha\beta}^{\gamma} Y_{\gamma} L_{\beta} \quad (28)$
с. 174, формула (29)	$\left(\frac{dy}{dx} \beta\right)^*$	$\left(\frac{dy_{\beta}}{dx}\right)^*$
с. 178, формула (H)	y_k	Y_k

Альберт Макарьевич Молчанов

НЕЛИНЕЙНОСТЬ В БИОЛОГИИ

17.05.89 г. Уч.-изд. л. 15,0. Цена 19 р. 20 к.

Тираж 335 экз. Изд. №22. Заказ 4095Р.

Отпечатано в Отделе научно-технической информации
Пушкинского научного центра РАН

19 р. 20 к.